Scholarship

2009 Assessment Report

Statistics and Modelling

COMMENTARY

Most candidates demonstrated competent statistical skills and knowledge, with answers that were clear and accurate, and gave full reasoning when required by the question.

Some candidates ignored the instructions to start each question on a new page. Markers reported that this made some booklet pages very crowded, and it made marking difficult in some instances. Some candidates had writing that was very hard to decipher. Some candidates attempted questions out of sequential order. This effectively disadvantaged them because the 2009 examination questions were thematically linked and were also sometimes linked by content. Therefore, candidates who answered questions out of order did not benefit from this ordering.

Candidates who answered questions with concise and relevant statements showing all their working, achieved well. These candidates showed a sound understanding of all the facets of the content being assessed. Their answers also showed a degree of maturity and understanding of the context. These candidates communicated statistical ideas clearly and efficiently, and justified the steps taken to obtaining their solutions, noting any assumptions or decisions they had made. These candidates also attempted each question in its entirety by completing all parts of the question in succession. This ensured that they made several distinct points in a description, rather than either a bare minimum of points, or a repetition of the same point.

Some candidates did not show understanding of validity of predictions (e.g. the words "validity" and "reliability" were confused or were treated as if they meant the same thing). Many candidates used interpolation instead of extrapolation, leading to completely incorrect answers to Q3 (b). Some candidates displayed a lack of clarity in report writing that involved basic statistical knowledge and made vague or irrelevant comments. Some candidates did not show evidence to support their answers or justify their conclusions, and gave superficial answers.

SCHOLARSHIP WITH OUTSTANDING PERFORMANCE

Candidates who were awarded Scholarship with Outstanding Performance typically:

- showed high level critical thinking, especially in Questions 2(a), 4(b) and 6(c)
- described an appropriate method for taking a sample where proportional allocation was required, making links to the context and clearly outlining how the method would ensure randomness and equity in selection for the required sample
- when discussing both limitations of models and validity of predictions, considered the context of the investigation in addition to mathematical considerations e.g. for non-linear models
- recognised that the discussion of validity of predictions was not just based on rote-learnt responses (e.g. interpolation is good, extrapolation is bad), but examined the plots, and made sensible statements based on the fit of the models or the lack of data available in order to establish the validity of the predictions
- evaluated predictions with reference to the quality of the modelling process, rather than simply comparing their predictions to the graphed data
- calculated an index number series
- demonstrated confidence in finding optimal solutions for linear programming problems where the initial solution was not suitable (e.g. a non-whole number answer where the context required whole number values, or where the constraints changed or new constraints needed to be taken into account); using graphing skills as well as algebra skills to investigate the situation.

SCHOLARSHIP

Candidates who were awarded Scholarship but not Scholarship with Outstanding Performance commonly:

- displayed effective communication, particularly in questions 1, 3 and 4
- wrote statements supported by statistics that clearly compared two situations
- correctly used statistical terms (e.g. justify, validity, limitations)
- related their answers to the questions
- assessed the information given, rather than simply stated generic answers
- set out their answers in a clear mathematical way, showing clear evidence for their conclusions
- clearly put forward their ideas in written text
- gave answers that were in context (e.g. $, thousands)
- accurately interpreted information contained in a graph
- displayed knowledge of what a confidence interval meant
- identified and discussed key aspects of comparison between two samples, based on the interpretation of a box and whisker plot
- demonstrated that they were able to apply probability and models to solve complex problems by constructing confidence intervals and calculating probabilities
- solved probability problems successfully by using correct probability notation, with their answers clearly indicated
- used a graphics calculator successfully in order to find solutions to problems, including obtaining estimates for the constants in a fitted model
- selected appropriate probability formulae and used algebra skills to manipulate algebraic expressions required to answer Q2 (b)
- provided succinct discussions and descriptions of the graphs in Q3 (a) and Q4 (a)
- selected essential points to discuss in the graphs, and supported their statements with statistics
- obtained forecasts for time series data by selecting an appropriate model for the trend and calculating estimates for the required seasonal effect
- looked within variables in order to discuss factors relating to patterns in the data
- recognised that Q5 (a) (ii) was a "distribution of the sample means" question
- used and correctly interpreted solutions provided by using a graphics calculator
- identified and carefully worked through all the whole number possibilities for the optimal points in Q6.

OTHER CANDIDATES

Candidates who were not awarded Scholarship or Scholarship with Outstanding Performance commonly:

- did not correctly interpret and complete questions, or gave answers that were not related to the question
- did not describe features of data
- wrote vague statements or used language that lacked precision or context – e.g. for confidence interval 'the difference of the means rather than 'the difference between the mean house valuations...'
- did not draw a graph of given information
- gave vague statements relating to their graphs

- did not recognise that Q1 (a) required a comparative box plot
- did not consider the context when presenting answers e.g. correct units, including dollar signs or correct measurements (hundreds, thousands), or appropriate rounding
- did not correctly describe the sampling method in detail in Q1 (b)
- did not quote the actual statistical values to back up their comparative statements in Q1 (a)
- treated every curve as an exponential curve
- did not substitute accurately into their calculators
- made poor choices for their model equations
- contradicted themselves about the validity of their predictions
- used individual points on a scatter plot to discuss the validity of the model, rather than an aggregate view, demonstrating a lack of understanding that the model predicts an average for the response variable from a given value for the explanatory variable, not an exact value
- did not describe the validity for predictions in Q3 (b) with high correlation and data closeness
- attempted to explain why the time series occurred, rather than what they could observe from the time series
- gave answers that were not sensible in the context of the question
- did not differentiate between questions asking them to describe relationships, select appropriate models, justify the selection of models, and discuss the validity of the predictions, but repeated the same responses; or did not take into account the type of data – e.g. non-linear, linear, time series
- rounded down for optimal solutions when whole number answers were required, without considering surrounding points.

SPECIFIC COMMENTS ABOUT THE QUESTIONS:

Question One

Q1 (a) Graph drawing was generally well done. Only a few candidates presented completely incorrect graphs (line graph, bar graph). In most cases the discussion that followed was not well done. Many candidates wrote correct comparative statements but they were not supported by statistics and so their discussion became a series of vague statements with no evidence to support their claims. There were some confusing box plots where the mean was drawn as a line which made it hard to distinguish from the upper quartile. Generally it was disappointing that some candidates did not mention thousands of dollars in answering the question.

Q1 (b) Many candidates did not fully describe the sampling method. Many talked about a systematic sampling method but excluded the details e.g. beginning at a randomly selected number and then selecting every 400th number from there omitting to explain how the 400 was calculated. Some candidates did not emphasise taking samples from each of the two suburbs but lumped the 80000 homes to get a sample of 200 homes.

Q1 (c) Generally well done. In some instances the confidence interval was expressed as (180.2, 203.8) rather than (180 200, 203 800).

Q1 (d) Many comments were too vague to be awarded credit, or the generic statement was given without any context at all. Some candidates claimed that the difference would lie in the interval (a, b) where the word “mean” was not mentioned at all.

Q1 (e) Poorly done. Many of those who began correctly did not find the number of further valuations required for each suburb. $517000 was a common answer where the means were not weighted according to the size of each suburb.

Q1 (f) Very poorly done. While quite a good proportion of candidates managed to calculate that 771 should be the sample size they did not answer the question and carry on to find the number of further valuations that should have been sampled from each suburb after subtracting the 200 that had already been selected.

Question Two

Q2 (a) (i) Generally well done.

Q2 (a) (ii) While the prediction was correct in most cases, the validity comment was not well done.

Q2 (a) (iii) The limitations of the model were not well described. The limitations of using these non-linear models was interpreted by many as what happens as x approaches 0 or gets large – a lot of candidates did not go beyond that to other factors that would have an influence on the selling price.

Q2 (b) Along with Q5(a)(ii) most candidates found this the easiest outstanding question out of the six in the paper to score on. Their setting out and communication was very good. However a few did not correctly handle the algebra involved.

Question Three

Q3 (a) A similar problem to Q1 (a) with correct statements being made but not backed up with evidence. Most candidates found three points for an S. Again, many made vague comments about an outlier without mentioning exactly where the outlier is positioned.

Q3 (b) Many validity statements were incorrect or too vague. While most candidates selected the correct equation, justified the choice of this equation and correctly calculated the prediction, few discussed the validity correctly. Many answers were reversed with respect to good validity (extrapolation) for one prediction and suspect for the other (interpolation). Candidates that did not comment on the validity, only mentioned the justification of their choice.

Q3 (c) This question required description of the relationship between SP and floor area, which involved discussing the shape of the points etc with outliers and groups etc. However, many candidates described what a graph might look like for each pair of variables, or interpreted this question to be only about a "relationship" between the two variables, either age and selling price or location and selling price. The word "relationship" should have been interpreted in the broader sense as to answering such questions as how does "location" influence "selling price"? Most managed to get two or three points of discussion but very few got four or more. Many candidates, in the interpretation of the word relationship, tried to look for a linear or non linear relationship rather than identifying the effect these factors have on the selling price.

Question Four

Q4 (a) As in Q3 (a), this was not well done though most candidates managed to give a partially correct answer. Many candidates did not pick out the salient points in a time series.

Q4 (b) For validity many candidates considered only the R^2 value. Many could not calculate the seasonal effect correctly. Many candidates did not choose equation 3 and this suggests that candidates are not aware that one should use the trend fitted to the latest data to make forecasts. Once again, discussion on validity of forecast was poor.

Q4 (c) Possibly this was not taught to the majority of candidates, as very few calculated the required index number series. Many otherwise good candidates didn't know what to do. Some deflated the median sale prices into June 05 dollars using the given index numbers, and thereby got the explanation part of the question correct.

Question Five

This was largely a mechanical question which enabled many candidates to score well on all three parts.

Q5 (a) (i) Most candidates managed this question with some omitting the continuity correction.

Q5 (a) (ii) Many candidates missed the fact that this was asking about the distribution of the mean of a sample. Some managed to get full marks by correctly using probabilities relating to "totals" instead which led to the same answer.

Q5 (b) Many candidates managed this well, and most of those used the binomial formula to calculate the value of n. A very few went further and considered which of $n = 25$ or $n = 26$ was the better answer.

Question Six

Q6 (a) Most candidates managed to find one of the optimal solutions, but many did not consider <u>all</u> possible optimal integer (x, y) values near the intersection points.

Q6 (b) It was pleasing to see the number of candidates who managed to correctly translate the English into an in-equation. Unfortunately though, many candidates did not get a correct point out of this and consequently did not score in this question. There were many correct answers but with completely wrong points.

Q6 (c) Few candidates got this question completely correct. Many answers had (86, 10) as the optimal point. Those candidates who drew the graph and sketched the curve and worked through the whole number points to find the one that maximised the profit correctly, did well.