ROYAL STATISTICAL SOCIETY EXAMINATIONS, 2006

REPORTS OF EXAMINERS

General comments

Most comments made by examiners refer to specific features of questions set in this year's papers. But every year examiners draw attention to aspects of examination technique which could be improved. As we have noted in earlier reports, it is disappointing to see candidates losing marks unnecessarily. Several comments made by examiners in 2006 echo those made in recent years. We therefore repeat here the advice to candidates given at the start of last year's report, revised so as to incorporate further general comments made by examiners following the 2006 papers.

- Read any question you intend to answer slowly and carefully, and ensure you answer the question actually asked. Every year, some candidates reproduce bookwork which may have some relation to the topic, but does not answer the question itself. Examiners award marks in accordance with detailed marking schemes, which assign marks for specific answers to each part of each question. There is therefore no point in writing down what you know about a different (if similar) topic, since the marking scheme will have no marks available for this.
- Take note of the marking scheme printed on the paper. It is a waste of your time writing a detailed two-page description of some topic, if this can only be awarded 2 marks.
- When preparing for an examination, you will of course know that there will be certain details (definitions, formulae and the like) you will be expected to memorise. For any paper, candidates will of course be expected to know the definitions of all concepts relevant to the syllabus. As for formulae, it will be clear that (for example) a candidate who does not know the formula for a binomial probability function cannot fully understand the binomial distribution, so examiners may expect candidates to be able to quote that probability function when it is relevant to a syllabus. Similar examples can be given for other areas; formulae for sample variance and conditional probability (at Ordinary Certificate level) and sums of squares for appropriate analysis of variance models (at higher levels).
- Make sure you understand the difference between the instructions *explain* and *define*. An *explanation* of some concept requires one or more sentences; the concept concerned should be described in words and (if appropriate) the purpose or use should be outlined. In a mathematical examination, a *definition* is a short and precise statement, which may require the use of mathematical notation. If a definition is required, a rough description is likely to be awarded no marks.
- Ensure that you include sufficient reasoning in your answers for the examiners to be sure about the basis for any conclusions you draw. For example, writing 'the test statistic is greater than the value in tables' without stating the value, the relevant sampling distribution or the degrees of freedom will gain very few marks, if any.
- In questions requiring calculations it is understandable that errors will be made under examination conditions. When a candidate shows his or her working clearly, it is possible to give

credit for use of a correct method, even if there are errors in the numbers presented. However, when little or no working is shown it is rarely possible to assess either the method being used or the source of the error. Candidates are therefore advised to show sufficient working to make it quite clear which method is being used.

- When you complete a calculation, or finish answering a practical part of a question, try to check the plausibility of your result. For example, a variance cannot be negative, and a correlation coefficient cannot be outside the range -1 to $+1$. Similarly, a trend or regression line which does not pass through the main part of the data points is very unlikely to be correct.
- If a rough sketch diagram is required, this can be done in your answer book; there is no need to draw it accurately on graph paper. This might for example apply to a sketch of a probability density function. Of course, such sketches must always be sufficiently clear that salient features stand out properly. However, when an accurate graph or chart is required, this should always be done on graph paper; and you should make sure you include a title and label the axes. This might for example apply to histograms.

It is important to follow the instructions on the front cover of the answer book. We realise that candidates will not wish to spend time during the examination reading the front cover, so we have produced a copy you can consult on the Society's website. You are strongly encouraged to look at this before the examination, and to ensure that you follow the instructions. We draw your attention to the following instructions in particular:

1. Begin each answer on a new page.
   (You do *not* need to begin each *section* of an answer on a new page.)
2. Write the number of each question at the top of each page.
4. Graph paper should be attached opposite the answer to which it relates.
5. Enter in the space below (NOT in the side panel) the numbers of the questions attempted.
   (The question numbers should be written *in the order in which you answered the questions*. Note that the side panel is for the examiners' use only.)

It is also helpful to examiners, as well as simpler for candidates, when the answer to a question is written on consecutive pages of the answer book. We do realise that in practice candidates may sometimes need to return to a question later. If you do this, it is helpful if you indicate this clearly on the page where the earlier attempt was made.

Ordinary Certificate Paper I

General

The Ordinary Certificate syllabus covers the essential ideas of statistics in practice. Paper I on the collection and compilation of data includes the key topics of data collection in the field: what data should be collected, who from, and how data should be captured.

The overall standard on Paper I was fairly good with some excellent answers from many candidates and a few candidates obtaining high total scores. There were also some scripts of very poor quality where candidates appeared to know almost nothing about the topics on the syllabus. In general, knowledge of sampling techniques (stratified, cluster, and systematic) was poor, with many candidates appearing to be very confused as to the differences between methods.

Too many candidates did not answer the questions set. Some appeared to have noticed one or two words in the questions and to have written on the topic these suggested to them, instead of concentrating on what was asked. Some candidates did not make a proper distinction between the questions, carrying over parts of earlier questions into later ones, as mentioned in comments on individual questions below. On this paper, all questions were self-contained. It was only information displayed in boxes which related to more than one question, as was stated within the boxes themselves.

Question 1

Although part (i) asked about criteria to be considered when selecting villages, many candidates instead wrote about selection of farms or farmers (a later part of the study). In parts (ii) and (iii) some candidates concentrated more on discussing technical aspects of the sampling methods used and almost ignored the practical aspects such as choosing and interviewing farmers at a meeting (method A), and the difficulty of obtaining a systematic sample in a country where there are no reliable large-scale maps (method B). Very few mentioned that an advantage of conducting interviews at a farm is that interviewers would be able to supplement information given by farmers by observation. Many candidates implied that the views of farmers who did not grow the food crop attacked by the pest were not important, whereas the reason why some did not grow the crop could be because of attacks by the pest, and their views would therefore be important.

Question 2

Some candidates designed excellent questionnaires with good instructions for the interviewer (as required), but a very small number of candidates designed questionnaires which were clearly for self completion by the farmer. In addition a considerable number of candidates gave as answers something which was more in the nature of a record form, with tables apparently designed for an interviewer to fill in and questions for the interviewer such as ‘Does the farmer have any livestock?’. These candidates did not appear to have considered that the interviewer would have to ask the farmer questions in order to find out the answers, nor to realise the importance of standard wording in questionnaires. When a question requires a questionnaire to be designed, candidates should take note of the way in which information will be collected (by interviewer, by post, etc) and design the form accordingly.

Question 3

Judging from the answers, few candidates appeared to have had actual experience of dealing with the results from surveys. Many suggested that there should be one variable to show the type of livestock; however, this would only work if every farmer interviewed had just one type of livestock. It is possible (and, indeed, likely) that farmers will have more than one type of livestock and that different farmers will have different combinations of types. Since the numbers of each type are also to be stored in electronic form, it is necessary to have one variable for each type, and this would show the number the farmer possessed. Variables giving livestock numbers would all be discrete. Similar comments can be made about crops grown, where the variables would be continuous, and about pest control measures where the variables would be binary. Answers to open-ended questions would help determine the coding system.

Question 4

This was reasonably well done, with many candidates scoring full marks on parts (i) and (ii). Part (iii) required comments on the advantages and disadvantages of the two methods. Not all candidates gave these. Many failed to mention that there was little difference between the sample sizes required for the two methods.

Question 5

Part (i) was about selecting a sample of districts, but a large number of candidates wrote about the selection of adults, sometimes with no mention of sampling districts first. Note that a good answer to the wrong question will obtain few or no marks.

Many candidates said that they would group districts or adults into employment categories. Here they had not read the given information carefully. Employment data were not available for individual adults, and the information for districts was in terms of the total number employed in different categories. However, a stratification factor taking account of the mix of employment in districts might be devised; for example, by grouping districts according to the percentage of the population in full-time employment. Many candidates appeared to be unaware of the difference between stratified and cluster sampling.

Question 6

A few candidates wrote general accounts of non-sampling errors. This was not required. Some wrote about errors which were not a result of the selection method; for example, about the possibility of interviewers influencing the responses.

In general candidates need to write more concisely. For example, some voters on the list will no longer be at the address shown because they have moved or have died. There is no need to present these as separate items.

Question 7

This was done reasonably well by most candidates and some good points were made. A few candidates spent time commenting on the disadvantages and advantages of collecting time-use data by a questionnaire, whereas the question was about diaries. Some candidates wrote as if a questionnaire had to be used in an interview, but some of the points they made about diaries apply also to self-completion questionnaires, and some of the points they made about questionnaires do not apply to self-completion questionnaires.

Question 8

In part (i) some candidates wrote about all children in the 5-16 age group, although this part was about ‘younger’ children. A few produced a general account of what can and cannot be done in an observational study, instead of relating their answers to the situation described. Some candidates seemed to think that the time use was to be recorded in 10-minute slots for a week as in the diary described in Question 7. In general candidates were poor in considering the practicalities of the method.

In part (ii) some candidates wrote about all children in the 5-16 age group, although this part was about ‘older’ children. Some very elaborate multi-stage schemes for choosing samples of children via samples of schools were suggested in part (ii)(a), with the two schemes differing mainly in the sampling methods used at different stages. Simpler methods would have been sufficient. Some candidates lost marks as they did not then go on to suggest how the children might get access to the web questionnaire. In part (ii)(b) not all candidates related their answers to questionnaires on the web or to questionnaires that were designed for children to complete.

Ordinary Certificate Paper II

General

The overall standard of graph and chart drawing is improving, with ruled axes, labels, headings and use of broken-scale convention now the norm. Candidates should be encouraged to choose convenient scales with which to work (for example, based on units of 5 or 10) wherever possible. A handful of candidates failed to make use of the graph paper provided. Tables need to be given more consideration and planned, e.g. as to appropriate number of columns, before filling in the numbers. Short but clear headings are needed. Columns should be ruled.

Question 1

This question was the least popular on the paper. The average mark was less than half of the marks available. Some candidates, on the other hand, supplied very good succinct solutions occupying only one page. In part (b)(i), which was only worth 1 mark, most candidates failed to spot that, since 35 is midway between 30 and 40, straight linear interpolation would mean that the required answer was midway between the two quoted values. In part (b)(ii) candidates failed to read the question and did not quote the values of $1/n$ to 5 decimal places, as was required.

Question 2

Most candidates attempted this question and the average mark was quite good. In parts (i) and (iii) they lost some credit for failing to rewrite the quotations, as was required. It was surprising that in part (ii) many candidates could not evaluate the percentage correctly. In part (iii) they failed to spot that it was the increase in the number filing on-line which was required. In part (iv) some failed to spot the obvious error in the percentages.

Question 3

Candidates clearly have different methods of drawing a stem and leaf diagram. Some truncated the units figure from the value to be entered, some rounded and some entered a double figure. (The first two were considered acceptable but not the third). Some used commas to separate the values; this was also allowed but we do not recommend it as it makes alignment difficult. The labelling of the diagram including a title, labels of stem and leaf and their units was not always present. Though not a necessary part of a stem and leaf diagram, it is useful to include a column giving counts of values. The best candidates included such a column, and were thus able to check that they had 100 values in the diagram. Some candidates had missing values or wrong values. Not all values were aligned correctly, particularly when the stems had many entries.

The frequency table was generally constructed correctly, though again many answers lacked some of the necessary labelling. A total count should be added here.

Some candidates produced excellent histograms, with the frequencies correctly adjusted for the unequal class widths; some however forgot to adjust at all and a few even showed the class widths as equal when they weren't. Labelling was better than in previous years though few noted that the vertical axis represented frequency density and not frequency. A few candidates produced a frequency polygon rather than a histogram. There is no need to colour in or give different patterns to the bars of the histogram; this only wastes valuable time.

In part (iv) some candidates misinterpreted the word *diagram*, and discussed whether they preferred the frequency table to the histogram; what was required was a comparison between the stem and leaf plot and the histogram. Few candidates remarked that a properly aligned stem and leaf diagram viewed on its side makes a good histogram.

Question 4

This probability question was better answered than probability questions in previous years. However it was surprising how many candidates (even those who went on to complete the question correctly) could not give the basic definition of conditional probability. This is one of the definitions which need to be memorised. In part (ii) a very precise answer was required, not a general discussion of cases of independent events. Part (iii)(a) was generally well answered but part (iii)(b) less so. Several candidates quoted a probability which exceeded 1.

In part (iv), when device 1 is known to be working, the better candidates noted that X and Y must both be working correctly, so that the relevant conditional probabilities are 1, 0 and 1. For the case in which device 2 is known to be working, many forgot to divide by the probability that device 2 is working to give the conditional probabilities.

Note that, when a question says that there is no need to convert fractions to decimals, candidates should take the hint and not convert.

Question 5

Several candidates ignored this comparatively straightforward rank correlation question. Some could not remember the formula correctly. Only one commented on an obviously incorrect answer. Many ignored the checks: $-1 \leq r_s \leq +1$, $\Sigma d = 0$.

Comments on the value of the coefficient were satisfactory on the whole and several noted that Chelsea and Tottenham Hotspur were the two teams that contributed to the reduction in the degree of positive correlation.

There were varied suggestions as to why r_s was appropriate but only a few knew the correct reason: league position is an ordinal variable.

Question 6

This question was surprisingly poorly done. The calculations of mean and standard deviation should be routinely accomplished. It was clearly stated that the data were a random sample so the formula for standard deviation with $n - 1$ in the denominator was appropriate.

Candidates should be encouraged to write out their calculations in a tidy fashion. When calculating Σx^2, there is no need to write out each individual value of x^2; the time would be better spent squaring and summing the values a second time on the calculator to check the result. Candidates should be reminded that they cannot take the square root of a negative number and that their calculation should be redone if they find they are in this position.

Some were unfamiliar with the coefficient of variation and its purpose. It is the standard deviation expressed as a proportion (or, often, as a percentage) of the mean; when given as a percentage it is therefore appropriate to use a percentage sign when presenting its value.

The comparisons were very poorly done. Many did not bother to state the obvious fact that the mean number of heart bypass operations performed is much greater than the mean number of heart valve operations. The point of the question is that, with widely differing means for the numbers of the two operations, it was difficult to compare the variability in numbers having the two operations in the various hospitals because one would expect the standard deviation to have a larger value when the mean is much larger. The coefficient of variation allows a more appropriate comparison to be made and reveals that the variation in numbers having the two operations was not too dissimilar.

Question 7

A sizeable number of candidates did not attempt this question, and it was very disappointing that the question was so poorly answered. Being able to use and interpret percentages is a fundamental skill required for budding statisticians. That these data were taken from a popular newspaper and were incorrectly interpreted by so many candidates gives cause for concern. There was a noticeable difference in the quality of answers from different centres, so perhaps some centres should give their candidates extra practice in this topic.

Candidates need reminding that an index number is a percentage. As such, it is necessary to ask ‘Percentage of what, and when?’.

In part (i), candidates had to use the given graph to draw up a table showing chain base house price index numbers from Q2 2003 to Q4 2004. The graph shows that in the North, in Q2 2003, the percentage *change* in house price over the previous quarter was 2%. Hence the price in Q2 2003 (as a percentage of its value in the previous quarter) is 102% and so the chain base house price index for the North in Q2 2003 is 102. A chain base index number is always based on the immediately preceding time period. In every case, therefore, candidates needed to read the value from the graph, add 100 to it, and tabulate the result.

In part (ii), all these values had to be rebased on Q1 2003. The index number for the North for Q2 2003 remained the same at 102, as the previous quarter is Q1 2003. However the chain base index for Q3 2003 for the North, which was 102.5 compared to the previous quarter, Q2 2003, had to be rebased to Q1 2003. Since prices were 102.5% of what they were in Q2, they were $1.025 \times 102\%$ (or 104.6%, to 1 decimal place) of the Q1 values. That is, prices in Q3 were 4.6% higher than in Q1. Similar calculations were required for the South and for other quarters.

The answers in part (iii) were very poor. A few stated clearly that prices had risen in both regions over the whole period but the majority took the downturn in the graph to imply that prices themselves had fallen rather than the rate at which prices were rising. Given that very similar data are published regularly with regard to retail prices and earnings, it is very important that candidates should master this topic.

Question 8

This question was reasonably well done compared to previous answers on similar topics. The time charts were well drawn on the whole. Most candidates knew how to calculate and position the moving averages but some could not centre the 4Q moving averages correctly. Some candidates were obviously rushed in answering, as this was the last question on the paper; however, they should be advised that entering their results neatly in tabular form is the most efficient way of guarding against slips in calculations as ‘unusual’ values are more easily spotted. Several candidates wrote down far too many decimal places in their moving averages, which again wasted time.

Higher Certificate Paper I – Statistical Theory

General

The aim of this paper is to test the ability of candidates to understand and interpret basic statistical theory and to apply and adapt it to simple practical situations.

There were a few infringements of the rubric.

About three quarters of all candidates achieved the nominal pass mark of 50. The range of performances was wider than usual this year, and about one third of scripts were rated at distinction level. Whilst many candidates achieved impressively high standards, the work of an appreciable minority of candidates (even those who narrowly passed) shows many examples of serious mathematical weakness. Particular problems were observed in relation to misuse of sum ($\sum$) and product ($\prod$) notation and suffixes, inaccurate differentiation and integration, and confusion between discrete and continuous distributions; there was a widespread inability to deal algebraically with $|x|$.

Question 1
Most scripts achieved an acceptable standard in at least parts (i), (ii) and (iii). In part (iv), some candidates failed to realise that there are just 7 increasing and 7 decreasing sequences that are excluded. Several candidates failed to find a method for part (v). It is probably easiest to count the PINs with 3 digits equal and with 4 digits equal, and to subtract the total of these numbers from 10000.

Question 2
There were many good answers to this popular question. Part (i) was well done on the whole, although some weak candidates had incorrect answers for the probabilities of a single entry from A, B or C, due to omitting the combinatorial multiplier from their (presumed) binomial probabilities. Weaker candidates also failed to recognise that a conditional probability was required in part (ii).

Answers to part (iii) were mixed. Three essential points were expected here: first, to enumerate the 6 possibilities $(A, B, C) = (2, 0, 0)$, $(0, 2, 0)$, $(0, 0, 2)$, $(1, 1, 0)$, $(1, 0, 1)$, $(0, 1, 1)$; second, to note that events A, B and C are independent, so that in each of the 6 cases three individual binomial probabilities can be multiplied together; finally, since the 6 cases are mutually exclusive, it is simply the sum of the resulting probabilities which is required.

Question 3
The highest-scoring question, although few good graphs of the pdf were seen (the correct asymptotic behaviour at $x = 0$ and $x = 1$ was seldom seen and often no graph at all was offered). Most candidates found $k = 30$; surprisingly few used symmetry to deduce $E(X) = 0.5$. There were a few mathematically weak answers to part (ii), in which sloppy use of limits ($\pm\infty$ instead of 0 and 1, possibly reflecting over-emphasis on indefinite integrals in mathematics teaching) led to trouble, along with sloppy algebra leading to a faulty term-by-term expression of the pdf. In part (ii), some candidates

worked in decimals, which necessarily leads to limited accuracy; this approximate approach cannot provide an exact justification of the given answer 17/81.

There were a few misguided attempts to use a binomial distribution in part (iv), and several candidates failed to identify the variance of the sample mean as $\sigma^2/5$, where σ^2 is the population variance.

Question 4

Nearly all attempts scored at least half-marks on this most popular question, and part (i) was almost always correct. However, there were several errors in calculating the necessary standardised Normal (that is, $N(0,1)$ values, or z-values) in part (ii). This was due mainly to the regrettable and misleading shorthand '$2T$' for 'T_1+T_2' or '$3T$' for '$T_1+T_2+T_3$' (leading to wrong variances) and failure to take square roots of variances when forming z. Wonderful but wrong arguments were sometimes adduced in part (iii) to arrive at the given mean journey time of 16.05. Many candidates correctly applied the 'variance of the mean' formula to solve part (iv).

Question 5

This question was less popular, and gave rise to several very good but many rather poor answers. Many marks were lost for weak algebra and very sloppy use of sum and product notation. In part (i), not all candidates were aware that deriving the mean and variance, if done by means of generating functions, requires the generating functions to be derived first. Weak answers to part (ii) commonly confused the operations $\sum$ and $\prod$, simultaneously used i and x as dummy variables in the same expression, and failed to apply $\prod$ to all terms in the probability mass function when constructing the likelihood. Part (iii) required use of either the Cramér-Rao asymptotic bound for $\mathrm{Var}(\hat{\lambda})$ or the 'variance of the mean' formula, but several answers appeared to work back unconvincingly from the given result. In the final data-based part (iv), most attempts were successful in calculating the confidence interval. Calculation of the sample variance was often omitted, however, and surprisingly few candidates recognised the closeness of the sample mean and variance as support for the Poisson model.

Question 6

This was easily the least popular question on the paper, with easily the worst average score. There were very few good plots of the pdf: the modal cusp at $x = 0$ was rarely seen, and indeed the curve for $x < 0$ was often omitted. Many candidates were also quite unable to process algebra involving $|x|$. The algebraic shortcomings noted in Question 5 were also present in the ML analysis in this question.

Question 7

This was a popular question, to which many good answers were seen. Parts (i) and (ii) were well done. The mean and variance of X were correctly obtained in most cases, but the covariance calculation was more error-prone and sometimes omitted. Several candidates were under the impression that zero covariance implied independence.

The final part (v) betrayed further confusion, as a common method of solution was to 'generate' the joint distribution of U and V by obtaining the marginal distributions and then assuming independence. The safe approach to this problem is first to record the value of (U, V) for each possible (X, Y) combination and then sum the probabilities which contribute to each of the possible values of (U, V), that is, $(0, 0)$, $(1, 0)$, $(0, 1)$ and $(1, 1)$. Since it is easily found that $\mathrm{Cov}(U, V) \neq 0$, or that (say) $P(U = V = 0) \neq P(U = 0) \times P(V = 0)$, U and V are not independent.

Question 8

This question proved surprisingly unpopular, and many attempts at part (ii) were mediocre by reason of being based on the with-intercept model instead of the no-intercept model asked for. In part (i), statements of the standard assumptions and model for simple linear regression were generally satisfactory, although the equivalence of least squares and maximum likelihood under the Normal error assumption was often not stated. In part (ii)(a) the derivation of the least squares estimator of slope for the no-intercept model was seldom well done; many attempts merely quoted the with-intercept formulae, and very few good calculus-based arguments for the correct model were seen. Although the data were generally plotted well, many more marks were lost in part (ii)(b) by candidates who persisted in fitting the with-intercept model.

Higher Certificate Paper II – Statistical Methods

General

The main aim of the Statistical Methods paper is to examine the understanding of basic concepts and techniques of statistical analysis. This is primarily achieved by asking candidates to solve standard problems of estimation and hypothesis testing, with particular emphasis being placed upon assessing each candidate's ability to summarise and interpret the results obtained from statistical analyses. Additionally, candidates are asked to describe or explain, sometimes with examples, particular concepts or general methodological approaches, and provide some descriptive analyses of data.

Overall, the performance on the paper was quite good, with an average mark of over 55%. As in previous years, calculations were performed very well, although occasionally formulae (particularly for confidence intervals) were not accurately recalled. In addition, explanations of methods and concepts tended to be too brief and interpretation too limited. Candidates should also ensure that they fully address all aspects of each question. Many candidates failed to list assumptions or provide interpretation when required. Moreover, despite statistics being an applied subject, few candidates included units where appropriate (for example, thousands of farms in Question 5, and years in Question 7).

It was also noticeable that candidates tended to attempt questions in the order in which they appeared on the question paper. Even when candidates choose the 5 questions to attempt prior to answering the first, they would be well advised to start by attempting the questions they expect to find easiest. For example, although Questions 4 and 8 were those answered best, no candidate answered Question 8 first and only one candidate's first attempt was at Question 4!

The candidates' weaknesses noted in last year's report were similar to those found amongst this year's scripts. In addition to those already noted, there were several common errors in connection with tests of hypotheses:

- using a test for two independent samples when the samples were clearly related (or vice-versa);
- performing two-sample tests when only one sample was given, or vice-versa;
- confusing two-tailed and one-tailed tests, and therefore using the wrong percentage points of the relevant sampling distribution (for example, using 5% values in tables rather than 2.5% points);

- confusing the standard deviation of a random variable and the standard error of a parameter estimate.

Question 1
Not a question well answered by many candidates, with an average mark a little under 50%. Calculations were invariably carried out correctly, but statistical understanding appeared limited in many cases. One relatively common error was to perform two-sample tests when, in fact, there was a single sample to compare with the 'normal' population's parameters. In part (i), a few candidates sensibly noted that the sample variance was very close to the stated population variance (of 'normals') whereas others unaccountably concluded that the sampled population had a different variance. Candidates are reminded that they should always clearly state their null and alternative hypotheses, degrees of freedom and critical value or values of the sampling distribution of the appropriate test statistic. Additionally, they should ensure that, when a particular approach to testing is required (for example, a p-value approach), they use the required approach when writing out their solutions.

Question 2
Again, this question was not answered particularly well, although several candidates provided very good solutions.

Part (i) led to some very good answers and some very weak ones. In particular, candidates should reflect on the issues specific to the example to which they are referred rather than simply writing out general material found in textbooks. Only a few candidates noted that the non-parametric tests could be used for ordinal data whereas the parametric tests based on the Normality assumption require the data to be at least approximately on an interval scale.

In part (ii), most candidates correctly used the Wilcoxon rank sum test (or the Mann-Whitney U version of that test) and correctly computed the value of the test statistic.

Two errors occurred quite frequently. One was to view the data as paired; it is vital that candidates learn to recognise from a description of an experiment whether samples are paired or independent. The other common error was to use the Mann-Whitney U test statistic but critical values for the Wilcoxon rank sum test.

Some candidates made the unnecessary transformation of both the test statistic and critical value from the Wilcoxon rank sum test to those for the Mann-Whitney U test; candidates are strongly encouraged to recognise the relationship between these two equivalent tests; they should be able to apply either. Since, however, the critical values given in the Society's tables are for the Wilcoxon rank sum test, candidates are generally advised to use the Wilcoxon version of the test unless a question dictates the use of the Mann-Whitney U test.

Question 3
This was a relatively high-scoring question. Most candidates plotted the data adequately, although omission of the measurement units from the axis labels was a common weakness. Several candidates sensibly broke the horizontal axis after 0 and restarted it at 150. (This is of course easily done when a graph is hand-drawn, but may not be feasible with plots drawn using computer packages, as appears in the solutions to this paper on the RSS website.)

In general, the computation of the regression coefficient estimates was performed well. However, relatively few candidates computed the coefficient of determination correctly. Often, the interpretation of this coefficient was also not done well; it can be expressed as the percentage of the variability in the sea bass weights explained by the fitted model (that is, a linear relationship between the weights and lengths).

Several candidates commented on the probable curvature in the model, but there was a tendency to over-interpret the likely form of the relationship; candidates who proposed an alternative model (for example, quadratic or exponential) generally failed to spot that the shapes of the curves proposed would not address the observed form of non-linearity.

Question 4

This question was very well answered by most candidates who attempted it, although fewer than half of those sitting the examination did so. Calculations were invariably carried out correctly, with reasonable interpretation and statement of assumptions. The implication of blocking was quite well understood by most candidates, although several failed to refer directly to power in part (iii).

Question 5

The computational parts of the question were generally well answered. However, only a few candidates provided the correct units (thousands of farms) or attempted to interpret the values of the median and IQR in part (iii), although many candidates correctly identified these as the values requiring interpretation. Likewise, it was disappointing to find few candidates able to identify that either Pearson's or a rank (e.g. Spearman's) correlation coefficient would measure the *strength* of the relationship between two quantitative measures. Candidates are reminded to read questions carefully and to recognise the difference between a test and a measure of some quantity (for example, an estimate of a parameter).

Question 6

This question, whilst answered quite well by a moderate number of candidates, was poorly answered by a similar number. This resulted in a rather low average mark.

In part (i), only a few candidates identified that the *areas* (rather than the heights) of the 'blocks' of a histogram represent the frequencies. Note that there are several ways of choosing 'tick marks' on the horizontal axis. It can be argued that it's best to use convenient round numbers; here 4.85, 4.90, $4.95, \ldots$. It can also be argued that, to make the histogram present as clear as possible a picture of the data as supplied originally, one should use the class boundaries as tick marks, as was done in the solutions on the website.

Part (ii) was quite well answered by around half the candidates who attempted it. Most of those who were unable to answer the question could not use the given Normal distribution to compute the frequencies expected under the null hypothesis. Other common problems were incorrectly identifying the degrees of freedom (for the χ^2 goodness-of-fit test) and failing to note that not rejecting the null hypothesis does not necessarily mean that the data do follow a Normal distribution.

Question 7

A popular question, attempted by about three quarters of candidates. Answers were generally satisfactory.

Some candidates provided good answers to the first part, with others giving only rather basic descriptions. Some candidates wrote down the formulae for the test statistics and sampling distributions, which were not required.

A surprisingly large number of candidates made errors when computing within-pair differences in part (ii)(a), given the relatively few errors in other calculations. It is important to take care with all computations, no matter how apparently trivial. In part (ii)(b), most candidates performed the test correctly, although some did not correctly state the null hypothesis: this should be a proposal about a population parameter (or, sometimes, set of parameters) against which one tests an alternative. Also, although many performed the test correctly, a large proportion of these did not take sufficient care to state the precise rejection criterion they were applying to decide whether to reject the null hypothesis.

Rather strangely, despite the strong hint in part (a), which required candidates to calculate the mean and standard deviation of within-pair differences, several candidates tried to perform an independent-samples t test. At Higher Certificate level, within a part of a question there will often be a theme. Therefore, candidates should not expect questions to require changes in approach without being directed to apply a different technique.

Finally, when computing the 95% confidence interval for the difference in part (ii)(c), some candidates used a different sampling distribution (typically standard Normal rather than t_{n-1}) to the one used in part (ii)(b), despite the underlying data, statistics (and parameter) being the same! This is illogical as the underlying sampling distribution is dependent on the same test statistic (and degrees of freedom) and so its distribution will be the same as in part (ii)(b).

Question 8

A popular question, attempted by about two thirds of candidates and also well answered. Few weaknesses were observed. The main issue was, in common with the rest of paper, limited interpretation of the findings. For example, only a few candidates tried to interpret the confidence interval produced in part (ii), although a number of candidates included a partial interpretation in their answers to part (iii).

Also, very few candidates opted to use the recommended (Yates') continuity correction when computing the χ^2 test statistic and several candidates used the upper 2.5% point when performing a test at the 5% significance level. In relation to the latter, as any departure from the null hypothesis (of no association) is reflected in a large value of the χ^2 test statistic, only values in the upper tail of the χ^2 distribution should result in rejection of the null. Therefore, although the alternative hypothesis is two-sided, the χ^2 test is actually one-tailed!

Additionally, when, as in part (i), the test statistic exceeds the 5% point, candidates should consider more extreme percentage points of the relevant sampling distribution (under the null hypothesis) and determine the lowest at which significance is observed. In the present case, candidates could report from the Society's tables that the value exceeded the 1% point, enabling a fuller statement of the strength of the evidence against the null hypothesis to be provided.

Higher Certificate Paper III – Statistical Applications and Practice

General

The aim of the syllabus of Statistical Applications and Practice is to develop skills in data analysis, using the theoretical concepts developed in the syllabuses for the Ordinary Certificate and the earlier

Higher Certificate papers. A principal objective is the analysis of real data sets and the effective and comprehensive communication of the results. The questions in this paper require candidates to select and carry out appropriate statistical procedures, computational and graphical, and to report the findings and conclusions clearly.

Questions 1 and 6 were the most popular questions and Question 8 was least popular.

Question 1

The stating of hypotheses was not good. Frequently a two-tailed rather than one-tailed alternative was given, and the symbol $\bar{x}$ was used instead of μ to represent the *population* mean. Also in part (iii) the mean or some vague measure of location was often given, rather than the median. Variation in the value of the t statistic in part (i) was usually down to the use of $\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}$ instead of $s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)$, where s^2 is the pooled estimate of variance. The checks on the assumptions necessary for the t test were often poor or absent.

In part (iii) the rank sum calculations were usually correct, but the smaller rank sum was often used, rather than the sum for the smaller sample. There are different ways of carrying out the Wilcoxon rank sum (or Mann-Whitney U) test and these were allowed for in the marking. Use of a Normal approximation carried reduced credit, since this was not really justified in this case. A number of candidates mixed up the different approaches, using the statistic for one method and comparing with the critical value for a different method.

Question 2

As with many questions involving significance testing, some candidates stuck with a 5% level of significance and did not attempt to refine the degree of significance. (See also the comment made on Question 8 in Higher Certificate, Paper II.) Considering other significance levels may allow a stronger statement of conclusions.

In part (ii), residuals were frequently plotted against means rather than against fitted values, as instructed in the question. In part (iii) the degrees of freedom were often given as 4, rather than the value associated with s. In part (iv) there was not in general a good degree of appreciation of how to optimise the conclusion from further experimentation.

Question 3

Parts (i) and (ii) were often poor. In most of these cases, it was clear that the candidate did not have a good understanding of sampling methods.

The rest of the question tended to be answered fairly well. The majority of candidates had some notion of potential bias in the survey results in their discussion for part (d).

Question 4

The mechanical part of calculating the remaining moving average values and the expression for the regression line was generally done correctly.

However, it was surprising that some candidates who made a slip in the regression calculation resulting in a positive, rather than negative gradient, then seemingly failed to recognise that this was not sensible, even after plotting the line.

The comments and interpretation in this question were generally poor. A weighted moving average was often related to giving more weight to recent observations, and not specifically recognised as a mechanism for dealing with cyclic pattern. Representation of trend by the regression line was rarely mentioned, and the presence of a serial pattern was rarely recognised in part (iii).

Question 5

This question was not popular, and was in general poorly done. Answers were often incomplete. Interpretation of the comparison of the fitted model with the empirical distribution, plotted in part (iv), was rarely done with any appreciation of the effect on estimating survival probabilities outside the range of the data. In particular, it was rare for a candidate to point out that the plot would suggest that the answer in part (iii) may be pessimistic in its estimate of the probability of survival beyond 24 months.

Question 6

Although this was the most popular question, it was not always done well. A mean effects plot was often given rather than the interaction plot which would have been appropriate. Several answers mentioned the notion of 'parallel lines' to indicate absence of interaction; these answers were often confused or unclear in relation to the data in this question, where one of the factors had three levels. Candidates needed to make it clear that they were looking for plots that were 'similar'. Comments and interpretation were of a very variable standard.

Question 7

Most candidates gave reasonable plots and chose the most appropriate linear relationship. The justification of the choice based on the regression output was limited, often relying solely on the R^2 value. The reason given for not considering $\log x$ vs $\log y$ was very poor, often not using any reference to the shape of the existing graphs. Units (i.e. real values) for the answer to part (iv) were often not given. Part (v) was often answered well.

Question 8

This was the least popular question and attempts were generally poor. There was little attempt to look at variability within laboratories. Most candidates looked only at differences in mean results between laboratories. There was little notion of a potential relationship between strain level and cycles to fatigue. Ideas relating to cycles decreasing with increasing strain level were poorly expressed and there were no suggestions as to the possible form of the relationship or the ability to predict fatigue at strain levels other than those used in the experiment.

Graduate Diploma Paper: Statistical Theory And Methods I

General

This paper examines probability theory – Bayes' Theorem, discrete and continuous random variables, univariate and bivariate distributions, transformations of random variables, simulation, order statistics, simple stochastic processes.

The distribution of overall marks this year was bimodal – about half the candidates made good (in several cases, very good) attempts at the paper, while the remainder were able to achieve very little. This pattern was repeated in many of the individual questions. This year, even standard work on proofs was tackled with mixed success.

All candidates found 5 questions that they could start, but some of these were token attempts. The least popular questions were those that examined the multinomial distribution (Question 3) and Markov Chains (Question 8).

Candidates could generally improve their attempts to explain what they are doing and could lay out their working more clearly.

Question 1
This question examined the Law of Total Probability and Bayes' Theorem. Future candidates should be advised that questions on these topics are not necessarily easy; there is some evidence that candidates rushed into tackling this question without giving due consideration to the problem-solving they would be required to carry out in part (ii)! Half the candidates got full or nearly full marks for this question; the other half got hardly any marks.

Question 2
This sets out to test the candidates' knowledge of some basic properties of random variables, such as the relationship between the pdf and the cdf. It was reasonably well done, but several candidates seemed confused about very basic concepts.

Question 3
The question concerned the multinomial distribution. In contrast to questions on this topic in some previous years, the emphasis this time was on how this model arises (rather than on algebraic manipulations) and candidates were required to write explanations. This change of emphasis might account for the relative unpopularity of the question, though most candidates who did attempt it scored close to full marks. Perhaps surprisingly, several candidates could not write down the correct form of the multinomial distribution. Even more appeared not to have any means of calculating the factorials required to work out numerical answers.

Question 4
The topic of bivariate transformations was examined in Question 4, this year in the context of independent χ^2 random variables. This was the question with the highest mean mark; however, several candidates struggled to see how the general result in part (i) could be applied to part (ii).

Question 5
This question examined properties of the moment-generating function of the Gamma distribution. There were some excellent attempts at this question, though also some surprisingly poor attempts at what was, after all, standard bookwork. In part (i), not a single candidate mentioned that the domain of the moment-generating function $M_X(t)$ was restricted to $t < \theta$ because the required integral does not converge elsewhere.

Question 6
In this question, the topic of order statistics for a sample from a Uniform distribution was tested. This was the question where by far the lowest mean mark was scored. Candidates did not know (and could not derive) standard formulae for marginal and joint probability density functions of order statistics.

Question 7
Simulation was the focus of Question 7. Several candidates tried to solve either (a) or (b) of part (i) by inverting the probability mass function or probability density function rather than the distribution function, which appears to show a very fundamental misunderstanding of this part of the syllabus.

Question 8
Only two candidates attempted Question 8 on Markov Chains. Both were token efforts.

Graduate Diploma Paper: Statistical Theory And Methods II

General

The paper aims to test understanding of a range of statistical principles and methods, and their application in simple situations.

All but two candidates answered exactly five questions. Of the remaining two, one answered seven questions and the other answered four.

Almost all candidates showed at least some understanding of some of the material and a majority showed a good understanding. Indeed there were two outstanding performances, one of which was almost flawless. The overall level of preparedness of the candidates appeared to be satisfactory. Many candidates gathered marks efficiently, showing good examination technique. There were some strong performances from UK and overseas candidates.

Question 1
The majority of candidates attempted this question. There were many good attempts at part (i), though a few candidates failed to say that they were using the factorisation theorem to show sufficiency. Part (ii) also generated some good attempts, though some candidates forgot to justify their results by mentioning functional invariance. Part (iii) was poorly done. Few noted that T has a Poisson distribution (which makes the calculations easy) and some resorted to an asymptotic approach, which was inappropriate. There were a few reasonable attempts at part (iv); a common error was to forget that if both sides of the given strict inequality are raised to the power T, then the inequality is only strict when $T > 0$ (but $P(T > 0) > 0$ since T has a Poisson distribution).

Question 2
The majority of candidates attempted this question. Part (i) tended to be done poorly with what appeared to be guesswork, no justification being given. There were several good solutions to part (ii) but a few candidates stumbled badly over very basic properties of expectation and variance, which was disappointing. There were some excellent answers to part (iii) but many candidates appeared not to know what *consistency* means. There were some good attempts at part (iv), but a common error was to interchange the numerator and denominator in the expression for efficiency.

Question 3
The majority of candidates attempted this question. Some candidates had difficulty proving the required expectation result, but many obtained the method of moments estimator successfully. Part (ii) was mainly well done but with a few slips in the differentiation of the log-likelihood. There were some reasonable attempts at part (iii), but again some basic errors in calculus were made by some candidates. Most candidates knew the method needed for part (iv).

Question 4

The majority of candidates attempted this question. Part (i) was well answered. There were some good attempts at part (ii) but some candidates had difficulty in calculating the likelihood ratio correctly and some did not bring out the importance of the fact that the likelihood ratio is an increasing function of the given statistic. Part (iii) was essentially a piece of bookwork but was poorly done in the main. Part (iv) was also poorly done; few candidates were able to write down the likelihood ratio statistic or give its large-sample null distribution.

Question 5

Just under half the candidates attempted this question. In all three parts there were some good attempts. However, a common difficulty was that the solutions were set out poorly, thereby making it hard for candidates to keep track of the various components of the required probabilities. Thus assorted minor slips and errors tended to propagate.

Question 6

The majority of candidates attempted this question. In part (a) most candidates gave reasonable solutions. Likewise all components of part (b) were done well in the main. A few candidates clearly did not know what a Bayes estimator is, however. Also, a common error was to use a Normal approximation in part (b)(iii), even though the necessary information was given in the preamble to enable an exact Bayesian confidence interval to be obtained.

Question 7

Just under half the candidates attempted this question. In part (a) the explanations given were often disappointing and lacking in detail. For part (b)(i) attempts tended to be basically correct or totally incorrect, with little middle ground. Few candidates mentioned that the pivotal quantity has the exponential distribution with mean one. Again, for parts (b)(ii) and (b)(iii), solutions tended to be 'all or nothing'.

Question 8

Few candidates attempted this question. Those that did made good attempts to structure their answers, which was pleasing. There tended to be a reasonable list of criteria given but the criteria were often poorly described and their importance little discussed, except in the case of bias and variance. The importance of unbiasedness was greatly exaggerated by candidates.

Graduate Diploma Paper: Applied Statistics I

General

This paper is designed to test candidates' understanding of theoretical models and their practical application to realistic statistical problems. A basic grasp of theory is therefore necessary, but not sufficient, to achieve a pass mark. Candidates must be able to describe the purpose of a method, state the associated assumptions, and apply the method; they need then to interpret the results in a way that a non-statistician would understand.

As we noted in the general comments at the start of this report, marks will only be awarded for an answer to the question that is asked. Unfortunately, several candidates appear to recognise a statistical term mentioned in a question, and build their answer by repeating (often correct) theory about that term. However, most questions on this paper are set in a specific context, and the answers therefore need to relate to this context. Replication of bookwork, out of the context of the scenario in the questions, gains very little credit.

Some candidates answer only parts of a question. For example, when asked to describe associations and variability in data they referred only to associations. It is worth spending some time reading a question to try to understand what it is asking before starting to write answers. Also, some questions ask for basic definitions, and it is disappointing to find these parts omitted or answered incorrectly.

Question 1
There is usually a question on time series analysis, testing knowledge of basic definitions. This year there were many arithmetical errors in answers. A common mistake was failing to note that the expected value of a constant is its value (that is, $E(a) = a$), and there were also careless mistakes with minus signs. Working was sometimes difficult to follow. In part (ii), several candidates omitted the answer about the partial autocorrelation function or gave information that was not relevant to the question.

Question 2
This question was designed to test candidates' understanding of the basic theory of principal component analysis and their ability to interpret an analysis. In their answers, candidates tended to refer to the short names (for example, *Popn*) given to variables rather than to the actual variables; this means that the interpretations are overly mathematical and of little practical value. Often theory was presented that was not directly related to the question asked.

Several candidates misinterpreted the regression output. It is worth spending time checking that the output is what you think it is, if it is in an unfamiliar format. Poor answers to parts (i)(a) and (ii)(a) reflected a lack of understanding of basic theory.

Question 3
This was a relatively straightforward question on cluster analysis, although it tested understanding by asking for examples and interpretation. Working was often difficult to follow, and answers followed bookwork too closely rather than answering the question.

Question 4
Although this was a fairly straightforward example of a generalised linear model, few candidates attempted it. Answers to part (i)(a) were poor and methods used to calculate the 95% confidence intervals were often based on guesswork rather than appropriate theory.

Question 5
This was answered by few candidates. The mathematics involved are uncomplicated but candidates found this difficult.

Question 6
This was a question about forward selection in multiple regression. In part (i), most candidates showed a qualitative grasp of the method, but they could not always use the sums of squares presented to produce the quantitative results. Part (ii) was generally answered poorly, largely because the point of the question was misunderstood; candidates were expected to make a brief comment on the relationship between the two methods for variable selection.

Question 7

Part (i) is not difficult, but there were errors in graphs, and incomplete descriptions. Arithmetic was generally accurate, although there were errors in degrees of freedom, and several candidates did not answer the actual questions asked.

Question 8

There is usually a question on the analysis of variance (anova) requiring candidates to specify models and to do all or part of the anova by hand. It is essential that models are specified precisely, and that working is shown clearly.

Practical examples were required in part (ii); a repetition of bookwork theory was not sufficient.

Some candidates confused the two models and so got very little credit. Some candidates omitted the interaction term in one or both of the models, and others did not attempt to present the expected values of the mean squares. Errors in arithmetic lost few marks, but incorrect methods were of course penalised.

Graduate Diploma Paper: Applied Statistics II

General

This paper aims to test candidates' understanding of fundamental concepts on designed experiments and sample surveys, and their ability to apply these ideas to data.

The overall level of preparedness of the candidates sitting this paper appeared to be considerably greater than in previous years. There were some good performances, but most candidates still got less than half marks.

All but two candidates attempted 5 questions. As in previous years, many candidates lost marks by not answering the question asked or omitting sections of a question. In general, candidates seemed more comfortable when reproducing standard bookwork than when dealing with the practical aspects of surveys and experimental design.

Candidates should be encouraged to gain a deeper understanding of all topics on the syllabus, and of their application, since this paper is unlikely to be passed on standard bookwork.

Question 1

Parts (i) and (iii) of this question were done well by most candidates. In part (i), some candidates omitted the 'block' sum of squares (SS) from the analysis of variance (anova) table or forgot to calculate the interaction SS. We remind candidates that, although blocks were not explicitly mentioned in part (i) of the question, the fact that blocks were used was made very clear, and it was therefore important to account for blocking in the analysis. Unfortunately, some candidates did not read the question properly, and constructed the anova table for a single replicate of a factorial design.

Few candidates attempted part (ii), and those who did try it showed little understanding of the theory on partitioning sum of squares. In part (iii), the basic rules for good graph drawings were followed. Part (iv) was often omitted or answered poorly, with no mention of the relationship between seed rate and spacing.

Question 2

Most candidates were able to define a 'contrast' of means but some failed to derive its standard error. Most candidates correctly stated the contrasts for (a) to (d) in part (i), but hardly anyone attempted (e) on the interactions of (b), (c) and (d) with (a).

Part (ii) required candidates to test the statistical significance of each contrast. There were many mistakes. Often, the standard error of the contrast was miscalculated, so the t statistic was incorrect. Some candidates chose to calculate the sum of squares (SS) for each contrast, and compare this with the residual (error) mean square using an F test, but used the wrong formula for the SS. It should be noted that the formula for SS given by $(\sum_{i=1}^{\nu} c_i T_i)^2/(r \sum_{i=1}^{\nu} c_i^2)$ is based on totals T_i, whereas the data given in this question were mean responses. The conclusions were often poorly stated.

Question 3

Candidates could describe the essential features of a balanced incomplete block design (BIBD), but not necessarily when a BIBD would be useful; that is, when comparisons between any pair of treatments are to be made with the same precision.

Answers to parts (ii) and (iii), which were more practical, were generally very poor. The concept of treatment effect estimates seems poorly understood. Few candidates attempted to use the information provided to calculate the adjusted treatment sum of squares. In this part of the question, candidates needed to substitute the estimated treatment effects ($\hat{\tau}_i$) into the formula $(\frac{\nu\lambda}{k} \sum \hat{\tau_i}^2)$ for the adjusted treatment SS but many attempted to calculate the adjusted treatment sum of squares from the raw data given. Few obtained the correct answer.

In part (iii), tests of significance were generally carried out correctly, but the conclusions were not stated clearly.

Question 4

It was disappointing to see so few candidates attempt this question, which was a relatively straightforward application of a 3×3 factorial design for fitting a second-order response surface. There were some good attempts, but, as noted earlier, candidates did not know how to construct the polynomial contrasts required to complete the anova in part (a)(ii).

There were some good answers to part (b) in relation to mixture experiments.

Question 5

This question concerned the numbers of adults (x) and cars (y) in households sampled at random in a town. Unfortunately there was an error in the wording of the question. Ratio and regression estimators were required for the mean number of cars per household, with information about x being used to improve the accuracy of estimation. However, by mistake the question actually asked for the mean number of cars per adult. We are very sorry about this error; we have of course reviewed the scripts of all those who attempted this question to ensure that we have compensated them for any confusion arising from the mistake. The version of the question on the RSS website has been corrected, and the solution there relates to the corrected version.

Most candidates correctly calculated the ratio estimate $\hat{r} = (\Sigma y_i)/(\Sigma x_i)$ of the number of cars per adult; the ratio estimate of the mean number of cars per household is $\bar{X}\hat{r}$, where $\bar{X}$, which is obtained from the census data quoted in the question, is the mean number of adults per household in the town.

In part (b), candidates had problems remembering the formula for the variance of a ratio estimator. The formula $(1-f)(1-r^2)S_y^2/n$ for the variance of the regression estimator was given on the question paper. In this formula, r denotes the correlation coefficient between the two random variables X and Y. However, several candidates used the ratio estimate, $\hat{r} = (\Sigma y_i)/(\Sigma x_i)$, in the formula. Care needs to be taken not to confuse these two concepts.

In part (ii), several candidates did not answer the question asked, and discussed instead the advantages and disadvantages of using a telephone directory; the question required a comparison between use of a directory and use of random digit dialling.

Question 6

This question on stratified random sampling was generally done well, and most candidates scored well in parts (ii) and (iii), which were more practical. However, a few candidates did not know how to derive or calculate the variance of the estimator of the population mean from a stratified random sample.

Part (i) tested candidates' knowledge of basic theory for simple random sampling. Answers tended to be rather muddled and did not always show a full understanding of the theory.

Question 7

It was pleasing to see a good number of attempts at this question on two-stage random sampling.

Comments in part (a) were mediocre for the most part. Candidates did not really think through the potential sampling issues, many failed to mention the likely trends in quality from top to bottom, and sides to middle, of the truck, or within a crate.

Most candidates did part (b)(i) correctly. On parts (b)(ii) and (b)(iii) most candidates lost marks because they were unable to define the between-cluster variance, s_b^2, or the within-cluster variance, s_w^2, or both. Few candidates attempted part (iv).

Question 8

This question on crude and adjusted rates was generally well done by most candidates.

However, almost half the candidates calculated *age-specific* 'death' rates, rather than *age-adjusted* 'death' rates. Nevertheless, there were some good comments on the age-specific rates. Candidates who did calculate age-adjusted rates often did this for A, using B as standard, and for B, using A as standard, in order to compare A and B. It should be noted that the age-adjusted rate for A, if A is being used as standard, is the same as the crude rate for A.

Ideas for further enquiries and calculations were somewhat limited. Most candidates did not make any serious comments on the differences observed between the crude and the adjusted rates.

Graduate Diploma: Options Paper

The numbers of candidates for most components of this paper are small, and it is not possible to give detailed reports without identifying individual answers.