ROYAL STATISTICAL SOCIETY

ROYAL STATISTICAL SOCIETY EXAMINATIONS, 2013

REPORTS OF EXAMINERS

GENERAL COMMENTS

This report incorporates the comments made by examiners after marking the papers set in 2013 at all levels of the Society's examinations (Ordinary Certificate, Higher Certificate and Graduate Diploma). We would encourage all candidates intending to take the examinations in 2014 or subsequently to refer to the particular comments on the papers they expect to sit, as this is the primary means by which their examiners can communicate with them. We would also remind candidates that past papers (or specimen papers for new examinations) and reading lists are provided on the RSS website, and strongly suggest that all candidates will wish to make use of these vital resources as part of their preparation.

All levels of the Society's examinations are now fully modular. Candidates, and those advising them, should be aware of the benefits of a modular structure. Candidates do not need to sit all the modules at a particular level in the same year; indeed, we anticipate that only a small minority of candidates will do so. Candidates are most likely to be ultimately successful in passing at a particular level if they are realistic about the amount of time they have available for study and enter for an appropriate number of modules.

Most comments made by examiners refer to specific features of questions set in this year's papers. But every year examiners draw attention to general aspects of examination technique that 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 2013 echo those made in other recent years. We therefore repeat here the advice to candidates given at the start of previous years' reports, revised so as to incorporate further general comments made by examiners following the 2012 papers.

The published syllabuses for the Higher Certificate and Graduate Diploma give details of mathematical topics with which candidates at those levels are expected to be familiar before embarking on study for the Society's examinations. You must make yourself aware of the necessary mathematics background for the examinations you intend to sit, and make every effort to master it. This year again, candidates lost many marks as a result of their poor skills in algebra and calculus.

Read any question you intend to answer slowly and carefully, and ensure you answer the question actually asked. Every year, some candidates reproduce bookwork that 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.

On a related matter, be sure to carry out any specific instructions given in a question: e.g. round answers to three significant figures if that is what is asked; calculate the standard deviation, not just the variance, if that is what the question requires.

Take note of the number of marks allocated to each part of a question, as printed on the examination paper. It is a waste of your time writing a detailed two-page description of some topic, if this can only be awarded two marks.

When preparing for an examination, you will of course know that there will be certain details (definitions, formulae and the like) that you will be expected to have memorised. For any paper, candidates will 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). In recent years, examiners have regularly pointed out that candidates had quoted key formulae incorrectly and therefore gone badly wrong from the start of a question.

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 strongly advised to show sufficient working to make it quite clear which method is being used.

Be aware of the *RSS statistical tables* that are provided for candidates during all the examinations. This year, several examiners commented on the amount of unnecessary work that candidates had given themselves because they did not simply look up values in these tables. You may freely download a copy of the tables to use during your preparation for sitting the examinations, and you are strongly advised to get to know what is in these tables and how to use them efficiently.

In calculations with several steps, it is important not to round intermediate answers to too few significant figures. For example, if the final answer is to be quoted to three significant figures, then at least four significant figures will need to be retained for intermediate answers.

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 that 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 provided 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 so 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.

Year after year, examiners comment that many candidates seem a lot more comfortable with calculations and graphs than with discussion or reports. Applied statisticians need to develop excellent communication skills, so the RSS examinations assess these as well as arithmetic and mathematical skills. You should practise answering discussion questions, possibly using past papers and solutions as a guide; you will find it helpful to talk through your answers with a knowledgeable person, a tutor or a statistician you work with.

It is important to follow the instructions on the front cover of the answer book. We realise that candidates will have little time to spend on reading the front cover during the examination itself, 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.
3. Graph paper should be attached opposite the answer to which it relates.
4. 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, then it is helpful if you indicate this clearly on the page where the earlier attempt was made.

ORDINARY CERTIFICATE IN STATISTICS

Module 1 (Collection and Compilation of Data)

This paper is designed to assess a wide range of topics associated with data collection (including sampling and survey design) and data compilation. It is essential not only to understand the underlying principles but also to be able to apply them in a real-world context. It is also important to be able to explain and justify decisions and recommendations that you make. Candidates are advised to study real surveys and other projects involving data sampling. All of the questions on this paper were based on real problems known to the examiner.

The scripts were on the whole very good, with candidates engaging with the scenarios presented. Few marks were awarded for generic bookwork not tailored to the details given in the question. Marks were also lost when working was not shown (e.g. Q2) or where inadequate reasons were given (e.g. Q3 (ii)).

Several candidates found Q2 difficult, and candidates are advised to read up on the use of random numbers (in tables or computer-generated) in sampling.

There was confusion between strata and clusters, and these concepts are central to sampling theory. Candidates are strongly advised to study the definitions and also surveys that have used these methods, taking care to note the real advantages and disadvantages of the methods used.

Few candidates were able adequately to discuss the problems of missing data (i.e. genuine blank answer as opposed to no answer) in Q7. Moreover, few understood what was meant by a statistical worksheet, and how the analyst has to determine both the number of variables (often columns) required and the possible coded values when compiling such a worksheet (e.g. a spreadsheet).

While there were no arithmetical questions this year, this cannot be assumed. In any such questions (as in Q2) it is important to show all working, as some marks can be awarded for methods even when an answer is wrong.

Finally, it is important to consider how many marks there are for each part of a question, as this does indicate the likely length of answer required.

Module 2 (Analysis and Presentation of Data)

The quality of work displayed by candidates varied enormously. Some were clearly not ready to tackle a paper at this level, but the best showed a very good grasp of both technique and interpretation.

Many candidates have barely legible hand-writing. In the worst cases it was just not possible to determine whether an answer was correct or not because it could not be deciphered. It is a simple point for future candidates to bear in mind: taking the trouble to make answers legible can only gain marks.

Question 1
The unusual nature of the data values, increasing by a factor of 2 at each step, made this more difficult for candidates. It is extraordinary that many candidates were unable even to read off the mode from the table. The description called for in part (iv) is not just an attempt

to put the numbers in the table into words. The point is that the distribution is U-shaped with the peak at the left being substantially greater than the peak the right.

Question 2
Candidates were almost all able to spot the point that the information given refers to variation with age, but then quotes a single figure without reference to age. It was comparatively rare, however, to have comments on the other factors that influence daily energy requirements – height, weight, metabolic rate, activity level, etc. Almost nobody commented that, even when all such factors are taken into account, a single figure is of very limited value if there is no indication of the natural variation about that figure.

Question 3
Probability questions are generally not popular in OC2, and this one was no exception. Complete solutions were rare, and the question was simply omitted by many.

Question 4
In part (i) the appropriate graph was a simple time series with points joined by straight lines. The description should have been something like: there has been a general rise in petrol prices over this period, but with occasional dips from one year to the next. It is not appropriate to turn the sequence of numbers into words: doing that obscures rather than illuminates. The calculations in parts (ii) and (iii) were frequently done well. In part (iv), however, the simple linear interpolation defeated many, or involved very cumbersome methods.

Question 5
Parts (i) and (ii) were generally done well. In part (iii) it was quite common for candidates to use an incorrect formula for Spearman's coefficient. The comments in part (iv) were disappointing. The point was that worker P would contribute disproportionately to the product-moment coefficient. Ranking the data removes this effect and so should give a better overall indication of the association between the two sets of data.

Question 6
The appropriate graph to show relative proportions within a total is a pie chart, though a stacked bar was accepted too. Several candidates said that they knew they should be using a pie chart but didn't have a pair of compasses with them. Others drew circles by hand. Otherwise, this question was answered well.

Question 7
The main point of part (i) of this question was to distinguish between fuel consumption (litres per kilometre) and fuel economy (litres per kilometre per passenger). This caused no difficulty in the calculations or the graph, but many candidates confused the two measures when interpreting the data, often losing marks.
In part (ii) the main point was that the correlation is heavily influenced by the two outliers. When these two data points are removed the correlation is much closer to –1. Unfortunately, a good many candidates got the arithmetic wrong here and so were unable to gain the marks for interpretation.

Question 8
It was quite common to see that candidates did not know how to calculate a centred moving average in part (i).
Parts (ii) and (iii) revealed a lack of understanding of trend and seasonal variation. In particular, 'trend' was often interpreted in its rather vaguer everyday sense of how the data change from one month to the next.

Question 9
This was a popular question for candidates to tackle first. (Perhaps before they had forgotten the linear regression equation?) The calculations were often done accurately, if somewhat laboriously. The interpolation and extrapolation in part (ii) were handled competently too.

HIGHER CERTIFICATE IN STATISTICS

Module 1 (Data collection and interpretation)

Overall, most candidates displayed a pleasing grasp of relevant issues in collecting and interpreting data. However, many scripts were difficult to read, either because candidates' hand-writing is poor, or – more importantly – because the answer was poorly structured. There was some tendency among the weaker candidates to adopt the scatter gun approach of writing as much as possible without consideration of relevance in the hope of picking up marks. This was particularly evident in Question 1, perhaps because this question had the least structure to it.

Question 1
Most candidates explained simple random sampling as a method in which *each individual* in the population has equal chance of being selected. The correct answer, of course, is that *each possible sample* must have equal chance of selection. The rest of part (i) attracted a great deal of the scatter gun approach noted above.
In part (ii), quite a number of candidates strayed from the intended topic, non-sampling error, into any type of error they could think of. Part (iii), however, was generally well answered, with many candidates showing a good grasp of the difficulties associated with telephone surveys.

Question 2
The graphs envisaged in part (i) were simple line charts showing the trend in expenditure (a) from 2008-2011, and (b) across age groups. Since the purpose of the graphs was to facilitate comparisons, it was not appropriate to show males and females on separate graphs.
Part (ii) asked for 'a short report, suitable for publication in a serious newspaper'. Such a report should summarise the important features of the data without giving excessive detail. A verbal description, line by line, of the rise and fall of the numbers in the table is *not* helpful. So, for example, saying that males' expenditure is typically about 10% higher than females' expenditure is much more illuminating than giving a list of all the percentages for each age band and year.

Question 3
This was a popular question and it was frequently answered well. Most candidates were able to make a good many suggestions for improving the design of the questionnaire. However, several glaring errors in the original seemed to go unnoticed. For example, the overlap of categories in items 3, 5 and 6 was often overlooked. One suggested solution from some who did spot the overlap was to have categories £1 to £5 and £6 to £10. This, of course merely replaces one problem with another – unless the instructions say something about rounding to the nearest £.
A surprisingly large number of candidates suggested a third category in item 2: 'transgender' and 'other' were popular suggestions. Given the context, a questionnaire to be given to school children, this is not particularly realistic. It suggests that candidates were looking a little too hard for faults in the questionnaire.

Question 4
This was the most highly structured of the questions and it attracted some very good solutions. It does, however, require some careful reading. In (a)(ii) it was common to omit the estimated variance of the estimator, but whether this reflected ignorance or just a mis-read of the question was not clear. Very few made any sensible attempt at the estimated variance of the estimated proportion in (a)(iii). Part (b) was answered well by many, though a substantial minority appeared to think that a simple random sample is the same thing as a representative sample or an unbiased sample.

Module 2 (Probability models)

This paper was fairly well done overall. Almost all candidates were able to make attempts at three out of four questions. One candidate attempted all four questions, although the rubric says to attempt three; in the context of an exam, candidates are advised not to attempt more questions than they need to and to spend all their available time making their answers to those questions as full and correct as possible.

Question 1
This question examined probability, conditional probability and Bayes Theorem. It was the question with the highest mean mark. Some candidates confused themselves by setting out their working for the use of Bayes Theorem (in (i) (c)) and the Law of Total Probability (in (ii) (c)) in a haphazard manner. It is vital to label events of interest clearly and consistently in questions like these.

Question 2
This question explored various properties of the exponential distribution. On average, candidates got about half marks for it. Several struggled with the integration required in part (a), which is standard bookwork for this module. In part (b), several candidates lost marks by failing to heed the instruction to set $\lambda = 1$. Many others failed to recognise the exponential distribution as positively skewed.

Question 3
The binomial distribution and its Normal approximation were examined here. On average, candidates got about half marks for it. A lot of candidates lost most of the marks for part (i) because they made no attempt to 'derive' the expected value and variance but simply obtained them from well-known formulae for the binomial distribution. Candidates should be clear about the distinction between the instructions 'derive' and 'write down' in circumstances like this. In general, candidates demonstrated good knowledge of the Normal approximation to the binomial and the use of the continuity correction there. They became confused, especially about the value of the variance, when they had to use a Normal approximation to the mean of several binomial random variables.

Question 4
This question was about the Poisson distribution and its Normal approximation. Again, the average mark was about 10/20. Almost everyone who attempted this question did part (a) well, but note that the points should not be joined up on a graph of the probability function of a discrete distribution. Candidates seemed less familiar with the Normal approximation in this context than in the context of a binomial distribution.

Module 3 (Basic statistical methods)

This module covers the very basic statistical tests with which all who wish to use statistics in their work must be thoroughly familiar; please therefore read carefully the specific guidance that is given below on some of these tests. This year the first three questions of the paper were attempted by the majority of candidates, with the final question being markedly less popular.

Question 1
Part (i) was well done by those candidates who made the correct choice of test i.e. the paired *t*-test. The null hypothesis was actually stated in the question, although the equivalent form $\mu_A = \mu_B$ was allowed. By wording the question in terms of the *differences* for the ten universities, it was intended to point candidates towards the paired test, although some still performed a two-sample *t*-test. The only assumption required was the normality of the underlying population of differences.
Part (ii) was less well done as many candidates worked towards performing a Wilcoxon Rank Sum test rather than the Wilcoxon Signed Rank test which was actually required by the question. Marks were awarded for method in choosing the smaller of the rank sums, for stating the value looked up in tables and for outlining the comparison made to reach the final conclusion, so candidates should ensure that they always give full information in their answers.

Question 2
In part (i) the test of association was generally well done, although some candidates lost marks because they did not apply Yates correction for the 2×2 table.
In part (ii), care should always be taken when finding the standard error for a difference in proportions, as the form of this is different depending on whether a confidence interval or hypothesis test is required. The form required for the confidence interval here did not require the calculation of a pooled sample proportion, although some candidates calculated this.
The answers to 2(iii) were very diverse in nature. A few candidates obtained all three marks by saying that the hypothesis test would be directly equivalent to that performed in (i), so we would expect to get a significant result as we did in (i). Also the confidence interval obtained in (ii) does not contain the value zero and this corroborates the same conclusion.

Question 3
The confidence intervals required in parts (i) and (ii) were generally very well done although there were a few calculation errors particularly affecting the value of the sample variance. The unbiased form of s^2 with $(n-1)$ in the denominator should always be used. The test statistic in part (iii) was generally consistent with previous values. It was necessary to read the question carefully to ascertain that a one-sided alternative hypothesis was needed and so to compare the observed test statistic with the correct value from *t* tables. The salient point in (iv) was that the confidence interval in (i) was equivalent to performing a two-sided 1% test whereas the alternative hypothesis in (ii) was one-sided.

Question 4
As previously stated, this question was answered by only about one half of the candidates. However, those who did attempt it were generally successful in performing the *t* test for independent samples in part (i). About half of these also reproduced the confidence interval in part (ii) although more explanation could have been given in many solutions. Part (iii) was well done and most candidates correctly observed that since the interval contained the value 1, this indicated no significant difference in the population variances in either direction.

Module 4 (Linear models)

Question 1

(a) This is a one-way ANOVA or completely randomised design.
(b) (i) A standard one-way ANOVA is appropriate. Not all answers gave good explanations of conclusions.
(ii) Use of the actual dose allows a simple linear regression of reduction on dose level.

Question 2
In part (i) some candidates stated that *x* and *y* must be from normal distributions and were not clear about predictor and dependent variables or about the meaning of e_{ij}. This should be a starting point in all training courses.
At the end of this question the problem is of course that we do not have data in that region.

Question 3
In part (i) most scatter diagrams were well done, but the obvious curve was not always mentioned. This is relevant also in part (iv).
In parts (ii) and (iii), RSS tables can be used instead of working out the test for each statistic.

Question 4
Candidates who knew this topic mostly gave good answers (but it was the least popular question). A mention of "parsimony" in (vii) would have strengthened the recommendation of model 3, as there is so little to choose between them.

Module 5 (Further probability and inference)

Candidates are reminded that solutions often involve the use of calculus, so good skills in integration and differentiation are essential.

Question 1
This was a generally well answered question with most candidates correctly setting up the joint distribution table. In part (a) (i), recall that var(X1 - X2) = var(X1) + var(X2) - 2cov(X1, X2); a regular error was to omit the covariance term. Also, note that in part (b) (iii) it is not necessary to evaluate var(S) or var(Y).

Question 2
Although not intrinsically difficult, this was a poorly answered question. Candidates should study the solution carefully.

Question 3
This was quite a well answered question. The probability distribution is somewhat unusual, but the low numbers of marks for parts (i) and (ii) are a strong hint that candidates are not expected to carry out much mathematical manipulation. In this context, an error made by some candidates was to equate log(a + b) with log(a) + log(b). Only a few candidates realised that the second derivative of the log likelihood, evaluated at the maximum likelihood estimate, should be used in finding the confidence interval.

Question 4
In part (i), although most candidates correctly identified the method of moments estimator, many did not give a full explanation of why it was the method of moments estimator. The hints should not be ignored in parts (iii) and (iv).

Module 6 (Further applications of statistics)

The overall standard was very variable (mean 27.07, std dev 11.50), with the standard higher for those from the UK (31.11,12.22) than for those from overseas (24.08, 10.26). Several candidates wrote very little at all.

By far the clearest feature of the results is that most candidates scored most of their marks for the arithmetic and algebraic parts of the paper, rather than the discursive parts.

Question 1
Many performed the calculations adequately, but description of blocking, assumptions and potential problems were overall poor given that these are very straightforward questions. This was the most popular question, as expected.

Question 2
Definition of the normal equations was performed adequately by several candidates, but very few performed well on defining and discussing residuals. Nobody mentioned standardisation and most candidates were vague about what residual plots are actually plots of.

Question 3
Very few candidates adequately compared and contrasted polynomial regression and 1-way ANOVA, with most attempts being very poor. However, the backwards elimination was often performed well.

Question 4
Many candidates described how to construct a Shewhart chart reasonably well, but very few even tried to explain and justify the ideas and assumptions as requested. Most candidates derived the probabilities for the two methods correctly, although many did not simplify the results as much as they could. However, no-one verified the difference in rejection probabilities, and in retrospect this last part was probably not a very good question as it is hard to mark properly, in the sense of knowing how far candidates had got towards the answer.

Module 7 (Time series and index numbers)

Question 1
The first three sections were mainly numerical exercises, with a small element of real-world interpretation. These sections were generally well scored. The fourth section tested wider understanding of smoothing techniques commonly used for time series.

Question 2
This question tested candidates' knowledge of the wider characteristics of time series, and then focussed on the process of seasonal adjustment. Detail was required on how to deal with outliers in such a way as not to distort the process and conclusions of seasonal adjustment.

Question 3
Although predominantly a numerical question, it requires a familiarity with index number algebra and the relationship between price, quantity and value. Most candidates correctly spotted that the volumes in 2011 had returned to their 2007 levels, meaning that a volume index should ideally show no difference between these two years.

Question 4
This question involves the application of index number theory to a non-economic context. Some candidates did get some of the substitutions for price, quantity and value mixed up. However, those who made the correct substitutions got good marks for this question.

Module 8 (Survey sampling and estimation)

The overall standard of candidates sitting this paper seemed somewhat lower than in previous years. All candidates attempted 3 questions. There were some good scripts from about half the candidates (marks in the range 28 to 38). Other candidates performed less well. A few candidates seemed completely unprepared for a paper at this level (marks 10 to 12).

Candidates who scored well were more able to apply the basic concepts in the syllabus to practical situations, describe associated practical issues, and then comment on the implications for a survey. They provided fuller answers to all parts of a question compared to candidates with lower marks, perhaps who had not covered the whole syllabus.

As in previous years, candidates lost marks on more descriptive parts by not appreciating the practical implications of aspects of the situations described in the question. Answers need to be more targeted, and aligned with the marking scheme. For example, if there are 6 marks, there needs to be ~6 points to score 6/6.

Candidates are reminded that they are expected to memorise certain formulae as indicated in the syllabus. Marks were lost due to wrong formulae.

GRADUATE DIPLOMA IN STATISTICS

Module 1 (Probability distributions)

Performance on this paper was very variable. There were some excellent attempts but also some attempts where candidates struggled to score any marks at all on several questions. Adequate preparation for this module must include systematically working through lots of appropriate examples, such as those found in the recommended textbooks. It is not a good strategy to miss out some of the topics in the syllabus, in order to concentrate on others, since any individual question is likely to examine material listed under several different topics.

Question 1
This question primarily examined the Law of Total Probability and Bayes' Theorem, with an extension that required the use of the normal approximation to the binomial distribution. Almost all candidates attempted it and did so successfully, achieving on average well over half marks. In part (iii), some candidates got confused between two different random variables: the number of correct questions required to pass a multiple choice exam (which is a Binomial random variable in the context) and the total marks gained for the exam (which is not).

Question 2
In this question, candidates had to compare the approximate mean and variance of a Weibull distribution (obtained through a Taylor series expansion) with the true values (obtained by integrating the probability density function). About half of all candidates attempted it and they obtained just over half marks on average. A number of candidates managed parts (i) and (ii) fairly well but struggled with the integration in part (iii); they would have benefited from knowing the properties of the gamma integral, as in the hint given for this question.

Question 3
This question was about the properties of order statistics, specifically the marginal and joint distributions of the minimum and maximum of a random sample of uniform random variables. About two-thirds of candidates attempted this question, but on average they scored below half marks. Although candidates were clearly instructed to derive the results they were required to use, several of them lost a lot of marks by quoting general results and applying those.

Question 4
Candidates' knowledge of the multivariate normal distribution was tested in this question. Very few candidates attempted it, only one of them meaningfully. It is worth pointing out to future candidates that this is a very important topic in probability theory at this level, since the multivariate normal distribution is widely used in statistical applications.

Question 5
This question was based around standard bookwork – a bivariate transformation of independent gamma random variables to obtain independent gamma and beta random variables. About two-thirds of candidates attempted it and they obtained over half marks on average. There were many good attempts at this question and a few very bad ones. This is the sort of bookwork that well-prepared candidates should have seen and worked through for themselves in advance of the exam.

Question 6
This question started by testing candidates' knowledge of moment-generating functions, in the context of independent Poisson random variables, with an extension to a conditional probability mass function leading to the derivation of a multinomial distribution. More than three-quarters of candidates tried this question, but on average they scored less than half marks. A surprising number of them struggled with the standard bookwork of parts (i) and (ii). Part (iii) seemed less familiar to all candidates, but several of them made excellent attempts at the whole question.

Question 7
This question was set in the context of an extended problem, which had to be solved using various results from the course including the laws of iterated expectation and variance and the Central Limit Theorem. About half of all candidates attempted it but their average mark was very poor. Many struggled to extract necessary information from the question and structure it in a way that would have led them to a method of solution.

Question 8
About half of all candidates tried this question, on simulation. On average, they scored above half marks. A number of candidates misapplied the inverse c.d.f. method of simulation by setting $u = f(x)$ (the probability density function of X) rather than $u = F(x)$ (the cumulative distribution function of X) in part (ii). Others did not appear to know how to apply this method to the standard normal distribution using the table of its c.d.f. that is provided for the exams.

Module 2 (Statistical inference)

Likelihood is at the root of many of the methods covered by this paper, so it is vital to be able to form the likelihood function efficiently and to be able to simplify products.

Question 1
A popular, generally well answered question. All data should be used in forming the likelihood, although in parts (i) and (ii), where lambda is known, the W_i observations do not affect the maximum likelihood estimation.

Question 2
Again, popular and generally well answered. When making use of the likelihood ratio to find the critical region, be careful to get the inequality the correct way round. It is often useful to take logarithms but do not fall into the trap of equating log(a/b) with log(a)/log(b).

Question 3
Part (i) is a well-known example, but many candidates' answers were less than convincing. Part (iv) is a warning that the bootstrap method is not always appropriate.

Question 4
Although the different pressures can be treated separately in part (i), it saves time to form a single likelihood, based on all three pressures, from the outset, as it is needed in the later parts.

Question 5
It is good to see a question involving Bayesian inference being answered so well by so many candidates. However, in part (iii), several candidates assumed that the prior and likelihood were conjugate despite the posterior distribution clearly not being of the same form as the prior.

Question 6
There were relatively few attempts at this question and many candidates were apparently unaware of the formal definitions of the terms listed at the beginning of the question.

Question 7
Some good answers to this question. When evaluating the Fisher information, it is usually simpler to use the definition involving the second derivative of the log likelihood.

Question 8
A very unpopular question. Few candidates could explain how permutations are used to evaluate statistical significance. Answers to part (b) were satisfactory but often limited themselves to just one or two other desirable properties.

Module 3 (Stochastic processes and time series)

The paper consisted of an equal balance of questions on Stochastic Processes and Time Series. The quality of the best attempts was impressive but there were also candidates who were clearly not sufficiently prepared across the whole syllabus. Candidates' grasp of definitions and 'bookwork' theory was mostly very disappointing.

Question 1
Few candidates attempted to derive the equations for the transition probabilities. On the other hand in (b) the construction of the Markov chain's transition matrix was handled well. In (iii) instead of finding the expected number of steps, several candidates attempted to derive the stationary distribution presumably not realising that 3 of the 4 states were transient!

Question 2
This was the most popular stochastic processes question and there were several good answers. Most attempts opened unnecessarily with a derivation of the differential equations for the generalised birth and death process from 1st principles whereas, for 3 marks, only an explanation of the steady state equations based on balance considerations was required. Many of these also wasted time in part (ii) by repeating the same derivation for this queuing model, not realising that it is a special case of part (i). Another common error was to assume that part (ii) could be solved using results for an M/M/1 queue.

Question 3
This question on the M/G/1 queue was the least popular, perhaps because it is entirely theory. In part (i), which required an understanding of both pgf and mgf, the advice "By conditioning on *X*, ..." was not well taken or understood.

Question 4
This question on the Poisson process was popular but low scoring. Here, in contrast to Question 2, the candidates *were* expected to derive the differential equation for $p_0(t)$ from 1st principles and solve it. This would then have opened up the extension to a non-constant rate in part (ii). Many candidates simply stated general properties of the Poisson process before quoting the formula for the distribution.

Question 5
The first time series question was also popular but low scoring. Candidates seemed unfamiliar with the use of lag operator notation (eg L or B) to simplify what appeared to be an ARMA(3,2) model to the required ARMA(1,1,1) form.

Question 6
This was the most popular question and achieved the best average score. The best attempts were able to interpret the computer output provided and demonstrate insight into the methodology for selecting candidate models, fitting them and interpreting the Box-Ljung statistics. Weaknesses included confusion about the interpretation of acf and pacf and, worryingly, basic misunderstanding of P-values.

Question 7
There were a number of good attempts at this question on the spectral density function. However the sketches mostly missed the periodicity of the cosine function. Candidates also had difficulty linking the spectral density to the visual appearance of the series.

Question 8
The second most popular question was sadly one of the least successful. Most candidates were unable to give a satisfactory definition of 2^{nd} order stationarity in terms of mean, variance and autocorrelations and this greatly hampered their efforts in the rest of the question. In this question, and to some extent in Q5, some candidates made heavy weather of the calculation of variances and covariances by using the identity $Cov[X_t, X_{t-i}] = E[X_t X_{t-i}] - E[X_t]\, E[X_{t-i}]$ unnecessarily.

Module 4 (Modelling experimental data)

General

Good marks were achieved where candidates addressed all parts of a question. A common failing with questions in which the output from analysis was presented was in providing inadequate or no interpretation of the presented results in terms of the underlying science – it is essential for applied statisticians to be able to go beyond the test results and to be able to explain the results in non-technical language.

Question 1
This is a common type of design problem, requiring the statistician to balance various different constraints in constructing a suitable design. In part (i), common omissions were in failing to note the mis-match between the number of samples per run and the number of treatments for a RCBD, and the potential confounding of blocks with treatment differences in an incomplete block design. Many candidates also failed to explain the relationships in part (ii), merely re-stating what they indicated. A number of candidates found suitable designs in part (iii), but many answers suggested designs with block sizes greater than 7, or which could not be balanced, and often not considering designs using fewer than 90 experimental units. The approach to constructing the identified design requires a simple logical approach rather than complex mathematics, though cyclic permutations can be used to construct each of two halves after careful arrangement of the initial blocks.

Question 2
Another common type of design problem – relatively few answers but generally good answers. Some candidates do not appear to appreciate the difference between confounding and aliasing, and therefore do not understand the differences between the two types of design in part (i). In part (ii), a few candidates failed to identify the need for three confounding terms (and the four generalised interactions between them), and the need for these to be three- and four-factor terms to enable estimation of all main effects and two-factor interactions. There was also some uncertainty about the identification of terms in the principal block and hence in writing down the remaining blocks. Outline ANOVA tables often failed to identify the confounded terms, and often suggested allocating three-factor terms to the error as well as higher-order terms – with the available degrees of freedom from the

higher-order terms, it would be sensible to gather information on three-factor terms where available. Answers to part (iv) were fairly poor, failing to note the split-unit nature of the design and consequent impacts on estimates of interaction effects.

Question 3
Some understanding shown about the value of a Latin square design, but candidates often failed to link this to the example experiment. Some confusion about the form of model. Generally needed more interpretation of the test statistics in part (ii), not just reporting them. A remarkable number of errors in the simple calculations in part (iii) – incorrect critical values, failure to account to replication levels, failure to extract the variance estimate from the ANOVA table. Some good answers relating to contrasts, but other candidates showed no appreciation or understanding of these ideas. Interpretation often missing in part (vi), though one answer used a sketched figure to support their interpretation.

Question 4
Relatively few answers to this question, but generally poorly answered. Many answers showed no real understanding of the structure of a GLM, and incorrectly identified the presented output as relating to a GLM assuming data following a binomial distribution rather than a Poisson distribution (maybe the "Before" data should have been omitted from the question). Many errors in calculating the percentage reduction, either because of identifying the wrong form of GLM, or because of misunderstandings about the parameters being presented – many candidates failed to identify the parameters for the six treatments as indicating differences from the Untreated control. Answers to part (iii) generally showed a lack of imagination, failing to see the factorial structure of the existing treatment set, and hence to identify an extension of this structure by the addition of two further combination treatments. The concept of overdispersion is clearly not understood, but is important as most real data sets show some degree of overdispersion.

Question 5
Another poorly answered question, surprising given comments about similar questions on previous papers. Common errors were in the calculation of the correction term, in the identification of replication levels for the divisors in sums of squares, and in omitting calculations for a number of the interaction terms. Some errors in the identification of the numbers of degrees of freedom both overall and for individual terms. The calculation of the contrast value was often completely omitted, with other answers failing to include the "standardisation" of the sum of squares using the sum of squared contrast coefficients. There were a few good diagrams, though often these showed totals rather than means, or values for combinations of Variety or Weed Combination separately with Fertiliser Rate rather than for the sixteen combinations of all three factors. Interpretation (part (iv)) was often only focussed on the graph, ignoring the test results in the ANOVA table, or just reported the significance levels for the F-tests, rather than interpreting what the main effects and interactions might mean.

Question 6
The most popular and best answered question. Some good answers to part (i), though many candidates failed to connect minimum parameters and maximum variance explained. Generally good answers to part (ii), particularly the explanation of the Gauss-Markov theorem. There was considerable confusion about what the normal equations are – some good derivations, though often these failed to reach the equations for the parameter estimates. Generally a good understanding of stepwise regression approaches, though descriptions often failed to note the iterative nature of the process, and the crucial impact of the choice of thresholds. Mostly good answers about the adjusted R^2 statistic, though some confusion about the degrees of freedom in calculating this from the R^2 statistics (but there is disagreement in the literature on this!). Interpretation of the analysis output was probably the worst part of this question, with confusion about how to use the potentially contradictory

information from the presented F-tests and t-tests, and about why the different approaches should lead to different models.

Question 7
A popular question but some poor answers. Some confusion in part (i) between non-linear models and GLMs, and certainly some misunderstanding about what makes a model linear (which carried over in to part (ii)). Many answers failed to identify the differences in fitting linear and non-linear models. A common trend was to apply a log transformation to identify whether a model could be linearised. Some good interpretation of the presented analyses, using the presented graph, but too much reporting of test statistics rather than interpreting the fitted curve. Some candidates failed to interpret the form (parameterisation) of the equation shown in the output leading to incorrect calculations for the inverse prediction of the required phosphate level. Some confusion between leverage and influence, and generally poor interpretation of why the indicated points have high leverage (often just statements that the values are above some (arbitrary) threshold).

Question 8
Some good answers. Many answers to part (i) just described the questions rather than the forms of the models fitted, and therefore the differences between them. Many candidates failed to interpret the tests within the accumulated ANOVA table as providing comparisons between the successively more complex models, and failed to interpret the "intercept" parameter as being for variety A at 18°C, and the "treatxxx" parameters as indicating differences in the intercept parameter relative to this treatment. Some good interpretations, including an illustrative figure showing the parallel regression lines. Answers to part (iii) were often focussed on more experimental work rather than an extension of the analysis, possibly because of the failure to identify the "baseline" response as being for variety A at 18°C. Some good answers about pure error and lack-of-fit, with other candidates showing no appreciation of these ideas. Generally good awareness of the assumptions of normality and independence, but less so for homogeneity of variance and the appropriateness of the chosen model.

Module 5 (Topics in applied statistics)

This paper assesses 4 main topics, with a potentially wide range of topics. It is essential to understand the basic theory associated with methods but also to be able to apply them to realistic scenarios and to be able to interpret findings critically. Learning bookwork is not enough; nor is being able to interpret computer output in terms of just p-values and log-likelihoods.

It is also important for the applied statistician to be able to communicate with people who do not understand statistical terminology. When reading questions it is worthwhile to take time to establish precisely what you are being asked to do and how the information in the question is relevant.

Question 1
This was straightforward provided that candidates had committed the formulae to memory and were aware of the purpose of the test. This was not a popular question.

Question 2
Here candidates had to make decisions/recommendations. Few read the question carefully enough in order to establish which variables were relevant to each scenario. Candidates were not good at stating the distributional assumptions of methods, and answers were generally too short and without justification.

Questions 3 & 4
The bookwork parts of these questions were good, but candidates were less confident when critiquing/comparing output from packages. Candidates are reminded to provide reasons for answers, as these can gain marks even when the answers have errors/misconceptions.

Question 5
Parts (i) (ii) and (iii) were linked as this question was about confounding factors. Candidates were poor at identifying the associated issues associated with designing an observational study. Some gave a naïve study design which required unjustifiable assumptions about the variables involved. Moreover, the answer was intended to be for a doctor with little statistical expertise – a point missed by most candidates.

Question 6
Not everyone knew what a case control study was, and although several candidates had some idea of the methods to be applied very few were able to derive the table for an unmatched study from the one in the question which was for a matched study. Consequently marks were low on this question.

Question 7
Although candidates may have understood what a stratum is few were able to demonstrate this in an applied context. There were several arithmetic mistakes, but marks could be awarded where working was shown clearly. There was too much reliance on generic bookwork for the last part.

Question 8
The arithmetic in this question required careful working, but high marks could have been gained even without this. However, there was confusion over the definition of cluster sampling (as opposed to cluster analysis), and very little grasp of what the intra-cluster correlation is. These concepts are central to sampling theory. Moreover, even if an equation is unfamiliar (part (v)), the applied statistician should be able to use it when it is relevant to a topic with which she or he is familiar.