ROYAL STATISTICAL SOCIETY EXAMINATIONS, 2009

REPORTS OF EXAMINERS

General Comments

This report incorporates the comments made by examiners after marking the papers set in 2009 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 2010 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.

In 2009, the Higher Certificate became wholly modular; the Graduate Diploma will follow suit in 2010. 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 2009 echo those made in 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 2009 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 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 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 2009, several examiners noted that candidates had quoted key formulae incorrectly and therefore gone badly wrong right 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 advised to show sufficient working to make it quite clear which method is being used.

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 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 our 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 Paper I

General

Some candidates did not pay full attention to what questions were asking. In many cases, they appeared to latch on to a keyword in the question and write everything they knew on that topic (whether it was relevant to the question or not).

Question 1

This was disappointing in that a large number of candidates wrongly gave one-stage sampling methods in (i) thinking that one-stage cluster sampling and stratified sampling were two-stage, taking the division into groups as the first stage and the sampling from the groups as the second. In addition most candidates ignored the fact that it was a sample of adults that was required, and not a sample of districts or of addresses. In (ii) many candidates assumed it was an interview survey though the mode of collection was not mentioned. Most did have a reasonable idea as to the benefits and drawbacks of simple random, stratified, systematic and cluster sampling, however. In (iii) many candidates appeared to think that private households meant owner occupiers. Few wrote about adults in institutions.

Question 2

Too many candidates lost marks because they designed a self-completion questionnaire instead of one for use with an interviewer. Many candidates who correctly designed a questionnaire for use with an interviewer suggested that the interviewer asked the respondent what sex they were instead of asking the interviewer to observe this. In most cases observation and the respondent's name with title would identify the sex. Another common error was not to make the section on the demographic information into questions. Instructions to the interviewer were mainly good.

Question 3

On the whole, this question was fairly well done with candidates displaying knowledge of points for and against telephone interviewing and self completion questionnaires as ways of collecting information from victims of crime. A few candidates did not mention crime, however.

Question 4

It was clear that most candidates understood what the method of quota sampling involves, but hardly any scored highly on this question, suggesting only a few advantages or disadvantages. Some wasted time by writing in length also about the advantages and disadvantages of simple random sampling. Some were confused between non-response and the fact that a full sample size can be obtained with quota sampling.

Question 5

Parts (i) and (ii) involving calculations on a stratified sample were done correctly by nearly everyone. Part (iii) was done less well. Not every candidate made all of the points that: the SDs were out of date; this method of choosing sample sizes is not necessarily good for other variables; the sample size under optimal allocation was worryingly small for de luxe customers.

Question 6

There were some interesting responses to this question about the sources of error in an interview survey of a sample of cruise passengers. Most candidates appeared to be aware of the different sources, but in many cases the presentation was muddled and so some candidates did not score particularly well.

Question 7

Most candidates could distinguish between open and closed questions and give an example of each type. However, maybe half the candidates did not know what a closed question with multiple answers is, thinking that this is a closed question with many options rather than one where more than one option could be chosen. Some of these candidates also thought that a closed question with a single answer is a closed question with only one option, which would be a bad question.

Question 8

On the whole, this question was not done particularly well with candidates not fully distinguishing what each part was asking. Many candidates suggested looking at till receipts which would be laborious but did not suggest in (iii) that these would be sampled. Hardly anyone suggested taking a sample of newsagents or a sample of publications. Only a very few candidates suggested that an index number might be a suitable method of summarising the data, but most made sensible suggestions about summarising the money taken by such as time, publication or newsagent.

Ordinary Certificate Paper II

General

Several candidates failed to follow the rubric. Candidates are reminded that each question should be started on a new page but not necessarily a right hand page. It is not necessary to start each part of a question on a new page. Candidates are reminded that all graph paper and any second answer books used should be firmly secured to the first answer book used.

Some basic formulae, e.g. for the position of the median and for the rank correlation coefficient, were not remembered correctly. It is important that such basic formulae be committed to memory accurately.

The presentation of graphs and charts was pleasing, with almost all candidates using ruled lines, where appropriate, and clear headings, axis labels and scales.

There were some very untidy scripts. Neatness and orderliness are to be encouraged.

Candidates should be encouraged to attempt all the questions.

Question 1

Attempts at this question were somewhat disappointing. Many candidates just wrote out the formulae for mean, variance, etc. which was not required. Candidates did not seem to realise that a mean cannot be less than the minimum value in the data, they also did not appreciate that a variance cannot be negative. In the case of the correlation coefficients they did not seem to know that they range from –1 to +1 nor when they achieve their minimum values.

Question 2

On the whole this was competently answered with candidates fully prepared for such a question with protractor and compasses. The heading, sector labels and keys were generally well done. It was not necessary to use colour or shading to obtain full marks as long as the sectors were clearly identified. The fact that the data to be represented did not add up to 100% because of rounding errors caused a few problems. The easiest way to proceed was to take the given percentages and express them as a percentage of 99 before calculating the angles. In this way the degrees totalled 360. Alternative methods, such as distributing the extra 1% between the remaining sectors, were also acceptable, but simply adding the extra 1% to the 'All others' sector distorted the presentation.

Question 3

The first part of this question was generally well answered although some candidates were unable to specify the position of the median correctly and some failed to arrange the data in order of size. The range should be quoted as the actual difference between the maximum and minimum values rather than as the interval between them.

A concise paragraph was requested in the second part. Some candidates chose instead simply to give a list of points. This was marked down. The paragraphs were disappointing on the whole as they tended not to interpret the values of the statistics calculated in the context of the call centres workload, merely interpreting them in general terms.

Question 4

Many of the candidates were able to find the required histogram heights correctly. Some of the candidates did not fully appreciate that it is the class mid-points that should be used when calculating the mean (as these are the best estimates of the mean value of the members of the class), but the class right-hand end-points which should be used in plotting cumulative frequencies (as the cumulative frequency includes all those observations in the particular class). A small number of candidates did not read the question properly and ended up plotting either the histogram or the cumulative frequency curve (or both), thus wasting valuable time.

Question 5

Most candidates were able to produce the correct table. The problems after this were that many candidates could not remember the formula for rank correlation, many had an incorrect denominator, still more forgot the "1 –" at the beginning. More worryingly, some candidates did not seem to realise there was a problem when they obtained a value less than –1 or greater then +1 (see Question 1 comments). The best candidates not only commented, for example, that the correlation between high jump and hammer (0.8095) is a large positive correlation but that this meant that countries which perform well in the high jump also do well in the hammer and countries which perform badly do so in both.

Question 6

As usual, the probability question was the one ignored by more of the candidates than any other though, this year, there were quite a number of really good

answers. Alternative methods of calculation (when correct) were allowed in (ii) parts (a) and (b). In (ii)(a), a few candidates struggled with the idea of conditional probability and calculated instead the probability that the central box has a blue sum and that the first five choices were also blue, thus obtaining half the required answer.

Question 7

The time chart was generally well drawn with most candidates starting the vertical scale at zero or correctly using the broken-scale convention. The use of a key or labels to identify the different series on the chart was generally well done.

The majority of candidates did calculate and position the 3 year moving average correctly, although some were a year out.

It appears that some candidates had no experience of plotting a least-squares trend line in the format given. Others calculated the y value for every year. As it is a straight line plot, the recommended procedure is to calculate the value of y for three widely spaced years, join the points with a ruler, then rule a line right across the graph. Candidates should be aware that a trend, however calculated, should go through the midst of the data and not be consistently above or below all the data points. The final parts were often answered in general terms rather than in the context of this data set.

Question 8

Many candidates did not read the question carefully. The prices given in the table were the prices per share not the value of the shareholding.

Index numbers appear to be a difficult topic for many candidates and some are very unsure of what a price relative is. All index numbers and price relatives are percentages; the period for which they are calculated and the base period to which they refer should be clearly stated in the answer.

Some candidates did know the general formula for a weighted price relative index number but did not use the appropriate weighting, simply using the number of shares held. The appropriate weight is the current value of the shares which, although dependent on the number of shares held, reflects the importance of changes in share price to the holder's portfolio. The final result gives an indication of the percentage by which share prices have risen on average in the ten-year period. Other candidates calculated an aggregate index, which had not been requested.

Higher Certificate Module 1 (Data collection and interpretation)

General

Questions 1 and 2 were more popular than questions 3 and 4, but all questions were chosen by many candidates.

Question 1

Candidates clearly knew the difference between open and closed questions (parts (i) and (ii)) and the reasons for using the different types. Similarly, most could draft questions of each type, but a few candidates lost marks because they did not say which were which.

Question 2

In part (a) on whether a group of students appeared to have produced random numbers, parts (i) to (iii) were done fairly well, but few candidates could make a suggestion about another property that would help decide whether a sequence of numbers had been randomly generated. In (b), many candidates suggested discarding all sets of three digits greater than 350 when in fact they could also have used a second set of three digits in choosing a random sample from a population with 350 elements.

Question 3

In (a), hardly any candidates calculated any summary measures from the frequency distributions, but concentrated on the frequencies. As the age groups were of different widths, adjustment should have been made for this and in diagrams also. Many candidates drew component bar charts instead of histograms. The age distributions for men and women were in fact very similar. In (b), some of the suggestions about advantages and disadvantages of using a diary to record information about diet were a little fanciful.

Question 4

This question was open-ended. Some candidates appeared not to understand what is meant by a sampling scheme. In (ii), some candidates designed a complete questionnaire or did not ask specific questions about shopping for food; the question related to part of a questionnaire only. A sizeable number of candidates suggested other types of analysis in part (iii) rather than the kinds of tabulations they might make in a survey of shopping habits.

Higher Certificate Module 2 (Probability models)

General

Results for this module show a wide range of standard, with many very good scripts being seen but also many woefully derisory attempts. The effect of the long and very poor "tail" of both UK and overseas candidates is to depress the average score (45.1%), despite a pass rate of 50% with 21% of candidates passing at the level of distinction.

Question 1

There were 41 attempts at this question, with an average score of 8/20. While many candidates had part (i) correct, surprisingly many answers had this result multiplied by 5! (as if the order in which the hand was dealt were significant). Parts (ii) and (iii) were on average poor, although there were a few very good analyses. Many candidates attempted the 'order' method of solution, e.g.

$$1\times\frac{3}{51}\times\frac{48}{50}\times\frac{3}{49}\times\frac{2}{48}\times\frac{5!}{3!2!}$$

for (ii), but seldom found the correct final combinatorial factor to count the total number of different orders in which the hand could be dealt. Similar difficulties arose in part (iii). This question seemed particularly hard (and unpopular) for overseas candidates.

Question 2

There were 52 attempts, with an average mark of 11/20. Part (i) was usually correct. However, in part (ii) there was a common failure to realise that, because X and Y are positive,

$$\frac{X}{Y}>2.2\ \Leftrightarrow X-2.2Y>0,\ \text{ or }\frac{X}{Y}<1.8\ \Leftrightarrow X-1.8Y<0\,.$$

Candidates who overcame this hurdle often calculated Var(X-2.2Y), etc. wrongly. Many failed to realise that P(ratio differs from 2:1 by more than 10%) was simply the sum of the probabilities already found, viz.

$$P\left(X-2.2Y>0\right)+P\left(X-1.8Y<0\right).$$

The correct answer (0.0045...) was often obtained and rounded to 3 decimal places as asked, but it was not intended that the rounded figure be used in subsequent calculation. Candidates are reminded that premature approximation is inaccurate. Thus, in the last part (iii), to guarantee a final probability to 3 places as asked, it will usually be necessary to work to at least 4 significant figures throughout the calculation, only rounding at the final step. However, the major error in this part was the predominant use of the Normal approximation for B(1000, 0.0045) (or B(1000, 0.005) if using the rounded answer to (ii)), rather than the Poisson which is more accurate.

Question 3

50 attempts, average 7/20. Taken overall, this was the least satisfactory of all the questions. There were many nugatory answers to the first two parts, which showed virtually no notion of algebraic proof, suggesting that some candidates were inadequately prepared. Candidates using the moment generating function (mgf) for part (i) did not always realise that "Show that ..." would require the mgf to be derived rather than merely quoted. The final part (iii) caused some confusion. Very few candidates saw that, whilst $U = V - Z = Y + Z - Z = Y$ is a simple Poisson random variable, $T = X + Y - Z$ is the difference of two independent Poisson random variables.

Question 4

38 attempts, average 11/20. Along with many correct answers to part (i), several candidates made errors through failure to consider the limits of integration. Differentiation of the given answer for the cumulative distribution function to obtain the original given uniform pdf is not, of course, a sufficient answer. Many answers to part (ii) gave unclear explanations for the given answer for $F_Z(z)$, but these were marked with some latitude. A good number of candidates went on to obtain $E(Z)$ correctly, although poor differentiation of $F_Z(z)$ lost several candidates several marks. Similar comments apply to the parallel exercise for W in part (iii), along with several sign errors in answers due to failure to negate the derivative of $P(W > w)$ when finding the pdf. The final part was well done by those who had correct results for parts (ii) and (iii), but those whose previous errors led them to expectations which were not multiples of θ unavoidably lost marks.

Higher Certificate Module 3 (Basic statistical methods)

General

Almost all candidates answered questions according to the rubric. All four questions attracted a good number of attempts. Also, there were some excellent attempts to all four questions.

Overall, many candidates demonstrated that they could carry out the statistical methods featured in the syllabus. However, sometimes beautifully set out and numerically correct calculations were of no use because they involved application of inappropriate procedures, such as: using unpaired methods for paired data; carrying out a two-tailed test when a one-tailed test was needed; conducting a "known variance" test when the variance had to be estimated from the data. Thus, in preparing for this paper, is it certainly important to be able to apply the statistical methods featured in the syllabus, but it is also important to become familiar with **when** particular methods are appropriate.

Question 1

This involved a chi-squared goodness of fit test for a Poisson distribution. There were very few convincing explanations of why the Poisson model might apply to the error count per page. In contrast, many candidates correctly calculated the mean of the data, though a minority of candidates could not handle frequency data properly, which is very disappointing at Higher Certificate level. Many candidates calculated the relevant probabilities and expected values and were clearly aware of the chi-squared statistic. Some candidates did not combine values with low expectations and other candidates combined at the wrong point. Some candidates rounded too early and therefore obtained rather inaccurate results.

Question 2

There were many good answers to most parts of this question. Many candidates correctly calculated the t based confidence interval for the mean and the chi-squared-based confidence interval for the variance. Common errors were to use Normal rather than t values in the first and to get the degrees of freedom wrong in the second. The one-sample t test needed in part (iii) was often done well but common errors were: use of a two-tailed test when the question clearly points to a one-tailed test; incorrect reading of the t tables (0.1% rather than 1%). Very few candidates could correctly define a p-value (often the fact that it is a probability calculated assuming the null hypothesis is true was not stated), though the numerical part was done more successfully.

Question 3

In part (i) of Question 3, there were many good solutions. Most candidates used a chi-squared test, which is the natural choice, but a few used an equivalent difference of two proportions approach. In part (ii), most candidates either got this completely correct or made no progress at all. In part (iii), many candidates correctly used McNemar's test, though a few did not complete the test after calculating the value of the test statistic. Several candidates wasted a lot of time and effort in wrongly doing a standard chi-squared test, ignoring the fact that the question states that the study in question involves **matched pairs**.

Question 4

This was the least well done question. There were several excellent solutions produced. However, a large proportion of candidates did not use the fact that the data are paired and consequently used entirely inappropriate, unpaired methods in both (i) and (ii). A few candidates opted to evaluate correlation coefficients. Whilst this correctly uses the fact that the data are paired, the approach cannot answer the main issue (pointed out in the line below the table) because, for example, the Pearson correlation coefficient does not depend on the data means. (For example, I could add one million to each of Examiner B's marks – which would surely indicate that he/she tends to mark higher than Examiner A – but the correlation coefficient would remain unchanged). Part (iii) was frequently poorly done – many comments made by candidates were too vague, for example: "the paired *t* test depends on Normality" (Normality of what?); "the paired *t* test is more powerful" (under what circumstances?).

Higher Certificate Module 4 (Linear models)

General

The overall standard of candidates in this paper was reasonable: results averaged about 55% with 67% passing (32/48) of which 13 distinctions. Whilst some very good answers were seen, many candidates appeared to be well-drilled in routine analyses but weaker in showing evidence of understanding.

In many hypothesis-testing situations, candidates often failed to give the critical value of the relevant test statistic, so that it was then unclear whether they were using a one- or two-tailed test. Candidates should be aware that: (i) the exam presents an opportunity to demonstrate relevant knowledge to the examiner; (ii) partial credit may be given for a clear statement of a correct method even when some numerical details are wrong.

A worrying feature in a largely numerical paper was the tendency of several candidates to shun the summary data (given to expedite statistical analysis) and rework estimates from the raw data, losing time and often making mistakes in the process.

Question 1

Graphs were generally plotted accurately, albeit sometimes with an inappropriate *y*-scale: it would be sensible here to telescope the speed scale from 0 to 100 using the zig-zag or concertina symbol if the origin is to be retained. Most candidates correctly noted the upward trend in speeds with a possible outlier in 1985. Most calculations of the regression line were accurate, although several candidates worked from the raw data rather than use the summary statistics that were provided in order to save time. It is to be hoped that candidates' grasp of algebra is not so compromised that they lack the confidence to make correct use of the formulae provided for their benefit. Working from the raw data would imply a significant time penalty. Most candidates correctly preferred the analysis excluding the result for 1985, although the fact that this result was probably an outlier due to freak weather was seldom mentioned. Part (iii) was well done; most candidates were commendably cautious about extrapolation beyond the range of the data.

Question 2

There were many good ANOVA calculations; this analysis is clearly well-rehearsed. There were a few errors in calculating the degrees of freedom, and some candidates were under the impression that the *F* test for ANOVA is two-tailed. Conclusions were generally well reported. The basic assumptions for the analysis were fairly well stated, and many candidates were aware of the need to check the Normality of the estimated residuals. Scrutiny of residuals for (e.g.) temporal or serial correlations was rarely mentioned, however. The final part (iii) was less good, as many candidates re-estimated the variance for the A vs B comparison from the A and B data only, rather than using the (assumed common variance) error mean square from the ANOVA. As well as taking more time, this is less powerful, owing to the reduced DF available. The repeated sampling interpretation of the confidence interval (CI) was very seldom correctly given, and (although several correct tests were seen) few candidates were clear about the equivalence of accepting the null hypothesis with the CI including zero.

Question 3

This question on correlation and association was fair, slightly less satisfactory than Q1 and Q2 and rather less popular. In part (a), most candidates gave

acceptable graphs to illustrate (i) strong positive correlation and (ii) weaker negative correlation; however, several failed to make clear that a downward but nonlinear trend was required in (iii), i.e. a situation in which rank correlation is valid but not the product-moment correlation (pmcc). Part (iv) revealed that many candidates did not clearly understand that "non-monotonic" implies a trend with at least one turning point, e.g a parabola. Part (b) revealed a wide range of standards in stating the underlying assumptions for correlation (in which both variables are stochastic, with the bivariate Normal distribution as the paradigm) and for simple linear regression (in which the values of the independent variable are observed without error or conditioned on as fixed). In part (c)(i), the graph was well done but often not commented upon (and the hint of curvature was often missed). In part (ii), surprisingly, some candidates wasted time by recalculating the pmcc, or converting it to an equivalent t value, instead of merely referring the given value to the Tables. Most candidates ranked the data correctly, but many wrong calculations of the rank correlation were seen, due to misremembered formulae. When correct answers were found, candidates often failed to make the intended comparison (that the rank correlation achieved borderline significance whereas the pmcc did not).

Question 4

This question, on the interpretation of computer output of regression analysis, was the least popular and also the least well done, with an average mark of about 11/20. In part (i), dealing with the partial t tests, there was some confusion over degrees of freedom, although this number was given in the ANOVA for regression. In part (ii), the slight curvature of the plot was often missed. Most candidates were clear that x was not significant in the presence of x^2 in Model 2, and R^2 was correctly interpreted more often than not. Most answers identified a problem with non-constant error variance with this model, though the pattern of increasing scatter with x was not always clearly described. Few candidates made all possible points in favour of Model 3 in part (v). "Fitted line closer to data" and "higher R^2" were popular reasons, but higher F for regression and/or stronger partial significance of explanatory variables were usually not mentioned. The log transformation involved in Model 3 found several candidates unable to exponentiate correctly, resulting in wrong estimates for y_{10} for this model in part (vi) and hence a wrong adverse reflection on it in part (vi).

Higher Certificate Module 5 (Further probability and inference)

General

In order to be successful in this paper, candidates should note that, as well as gaining familiarity with the new statistical and probabilistic ideas, they need good skills in algebra (e.g. in evaluating products when constructing likelihood functions) and differential and integral calculus.

Question 1

This question on a bivariate continuous distribution required knowledge of definitions and skill in integration. Many candidates made a satisfactory attempt, though some struggled with the integration, e.g. several candidates thought that the integral of $1/y$ was $-1/y^2$. Common mistakes were: to integrate the joint density with respect to x instead of y when finding the marginal density of X; to assume that $E(XY) = E(X)E(Y)$, a result that only applies if X and Y are uncorrelated. Candidates should have noted that the joint density is symmetric in x and y; this saves work in (iv). Part (v) was generally not well answered.

Question 2

This question, on the evaluation and uses of the moment generating function, was popular and generally well answered.

Question 3

This question on likelihood was not so popular and attracted few very good answers. Several candidates had trouble forming the likelihood function, apparently not realising that k, in the expression for $P(X = k)$, needs to be replaced by x_i. Also, some candidates evaluated minus the reciprocal of the second derivative of the log likelihood, but failed to take its expectation.

Question 4

A fairly standard question on probability generating functions and estimators, but with a surprisingly wide spread of answers. Part (i) is essentially an exercise in differential calculus. In part (ii), candidates should be aware that the expectation of the sample mean of a random sample is equal to the expectation of individual values and be able to produce a brief proof. It was good to see that many candidates realised that a binomial distribution was involved in part (iv), though several confused themselves by introducing a random variable $\overline{W}$, even though there was only one W.

Higher Certificate Module 6 (Further applications of statistics)

Question 1

There are discussions in several of the books on the Reading List of the purpose of **blocking**. A **practical** example was required in (a) and (b). A 2×2 factorial experiment using **1**, **a**, **b** and **ab** as treatments leads to three sums of squares, that for A×B coming from the totals for (**a** and **ab**) and (**1** and **b**). Interaction is examined first, because main effects have little meaning if interaction is present. Reports need to explain this result.

Question 2

Candidates who attempted this question answered it quite well on the whole.

Question 3

Clearly some candidates knew a lot about CUSUMS, while some were skilled in medical trials, but few could deal with both.

Question 4

This was a popular question, with several good answers. Two points gave trouble. The first section revealed a weakness in knowledge of (iv) and (v), while the second section showed incomplete knowledge of how backwards elimination works and the statistical tests required. But the level of answers at some centres was very good.

Higher Certificate Module 7 (Time series and index numbers)

Question 2

Only 4 of the 20 candidates attempted this question.

Question 3

Many candidates struggled to cope with the new item whose base period price was by definition undefined. Some were also unfamiliar with the formulation of index numbers as functions of relatives and monetary values, rather than prices and quantities.

Question 4

Relatively few candidates knew the formulae for the Tornqvist and Geometric Laspeyres indices.

Higher Certificate Module 8 (Survey sampling and estimation)

General

The range of entries was very wide, but a pleasing number of candidates achieved excellent marks. There was a general trend for candidates to be happier with calculations than discussions, and some discussion based questions were not attempted by as many candidates as questions including more calculations.

Question 2

Only 7 out of 21 candidates attempted this question.

Question 3

Part (i)(b) appeared to mislead some candidates; the phrase "the interval in part (a) to be no greater in width than 4 hrs" resulted in some candidates basing solutions on an interval of ± 4 hrs instead of ± 2 hrs. Candidates must be advised to read questions carefully.

Graduate Diploma – 'traditional' version

This was the last year of the Graduate Diploma in its 'traditional' form (Statistical Theory and Methods, Applied Statistics, Options). The numbers of candidates for these papers were small, so it is not possible to give detailed reports without identifying individual answers.

Graduate Diploma Module 1 (Probability distributions)

General

This was the first paper set on this module, for which the syllabus is somewhat different from that of the traditional Statistical Theory and Methods I. Nearly all candidates made full attempts at five questions, suggesting that the time allowed is adequate.

Question 1

Several candidates who wrote down an accurate version of Bayes' Theorem were unable to apply it to the situation described, mainly because they failed to specify formally what the central event of interest, A, should be, along with the two "states of nature", E_1 and E_2.

Question 2

There were some weak attempts, but good candidates scored very well on it.

Question 3

This was a popular question: nearly all candidates offered a solution, and half of those scored at least 18 out of 20. Some candidates wasted time by not referring to their RSS tables to quote the numerical value of the binomial probability in the last part.

Question 4

Many candidates showed good preparation for this question, by applying the fact that (constant multiple apart) the integral to be evaluated was a gamma density, hence its value could be immediately written down.

Question 5

This, too, was very popular, and generally well done. The main errors were: slips in manipulation; a failure to identify the correct limits for the integration to obtain the marginal densities; omitting to specify the range on which those density expressions were non-zero. Time was wasted, and opportunities for errors introduced, by converting friendly fractions such as 8/15 to ugly four-decimal-place approximations.

Question 6

Few candidates attempted this question, and most attempts were unconvincing. The last part asks for "a detailed argument involving the Central Limit Theorem" to justify normal approximations to binomial and Poisson distributions; simply **asserting** that this step can be taken if $n > 30$ or $\lambda > 25$ does **not** constitute a detailed argument!

Question 7

The first part of this question, which was the third most popular, was generally done well, but few candidates could name the distribution of the ratio of two independent χ^2_4 distributions as $F_{4,4}$. Time was often wasted by evaluating the mean of this χ^2 distribution; the fact that the ratio of the two means is unity follows by symmetry, once it is noted that the mean is finite.

Question 8

The final question, on the expanded section of the syllabus that deals with simulation, had fewest attempts – just 5 from 26 candidates – but generated two near-perfect solutions.

Graduate Diploma Module 2 (Statistical inference)

General

This paper aims to test a range of statistical principles and methods, and their application in simple situations. The general standard of work was good.

Question 1

Not a difficult question but rather non-standard, so it is important to do exactly what is asked and not attempt short-cuts. It was generally answered well.

Question 2

A popular question, which was generally well-answered. It is important to realise that the minus sign in the probability goes with the logarithm, so that in the log likelihood you get $\log(-\log(1-p))$, not $\log(\log(1-p))$.

Question 3

A fairly standard question, but not so well answered apart from parts (i) and (iii). Few candidates gave a satisfactory definition of a confidence interval or realised that the moment generating function could be used to answer part (ii).

Question 4

This was another standard question, but with few very good answers. Many candidates failed to construct the correct likelihood function; often, this was because candidates tried to write it down directly without working through the steps. The Normal approximation was often not found correctly in part (ii) and few candidates seemed to have any idea how to answer part (iv).

Question 5

This was the first time questions had been set on jack-knife estimators and the bootstrap method, so it was not surprising that few candidates attempted this question. However, it was answered well by those who attempted it.

Question 6

A standard but unpopular question, generally answered very poorly. Many candidates did not seem to be aware that the maximum likelihood estimator (and, perhaps, the constrained mle) has to be found in constructing the test statistic.

Question 7

Quite a non-standard question, but generally very well answered. It was good to see that the lengthy hint did not deter candidates from attempting the question.

Question 8

In order to gain good marks on this type of question, it is important to answer the question asked. Many of the answers were too "generic", describing a variety of topics (e.g. likelihood, Bayesian inference, non-parametric statistics) without being clear how they related to the issues raised in the question.

Graduate Diploma Module 3 (Stochastic processes and time series)

General

This paper fell into two parts: four questions on stochastic processes and four on time series. The first three time series questions, Questions 5 to 7, turned out to be the most popular ones and the ones on which candidates achieved their best marks. In the stochastic process questions, the most straightforward question turned out to be Question 2, where candidates were able to gain marks by performing calculations on the transition matrix of a Markov chain in discrete time.

Question 1

It was disappointing that what was a rather routine question on a simple branching process was not done better. Most candidates who attempted this question succeeded only in doing a few fragments of it.

Question 2

A number of candidates successfully solved the equations to obtain the stationary distribution of the Markov chain. Very few were able to discuss adequately the concepts of irreducibility and recurrence and how the stationary distribution is related to the long-term behaviour of the chain.

Question 3

Few candidates attempted this question. It appears that the specification of the transition rates of a continuous Markov chain to model the situation described in the question was too challenging for most candidates.

Question 4

The first half of the question, which involved the derivation of the equilibrium distribution of a simple single-server queue, was done successfully by a number of candidates. The second half of the question, about the distribution of a customer's waiting time, was beyond most candidates.

Question 5

This question, which was centred around the derivation of the infinite moving average representation of a second order autoregressive process, was done reasonably well by a number of candidates.

Question 6

There were a number of excellent answers to this more applied question, where candidates had to examine sample autocorrelation functions and interpret the output from fitting a seasonal ARIMA model.

Question 7

This was the most successfully done question. A number of candidates showed that they had understood the technique of Holt-Winters forecasting and were able to put it into practice.

Question 8

Few candidates attempted this question on the theory of forecasting, but those who did made reasonable headway.

Graduate Diploma Module 4 (Modelling experimental data)

General

Marks were lost when candidates did not know basic theory, or were unable to apply it to a specific situation.

Candidates should take note of the marks allocated to each part of a question and plan their time accordingly. A section worth only 3 marks can be awarded at most 3 marks irrespective of the length of the answer.

Candidates are reminded to use specimen papers and reading lists in their preparation.

Question 1

Although this question only required standard theory from textbooks, the answers were poor. Imprecision and inaccurate statements always lose marks, but especially when the model answers are standard bookwork. Take note of the marks awarded for each part of the question. For example, since (vi) has 7 marks compared with the 2 marks for (v), candidates should expect to write more for (vi).

Question 2

Candidates should note that, in a BIB analysis, there is not a valid test for blocks as well as for treatments in the same analysis. Explanations of the BIB design were not always good or complete, and interpretations in terms of the actual treatments (amounts of Cd and Sn) were mostly absent.

Question 3

A few candidates gave good answers to parts (i), (ii) and (iii), but the practical parts, (iv) and (v), were not well done.

Question 4

Surprisingly, quite a few candidates did not know the difference between linear models containing *fixed* effects and those with *random* effects. Very few indeed knew how to make the hypothesis tests in part (ii)(b).

Question 5

In factorial analysis, do not compare main effect means when interactions are present. As in this example, such comparisons are meaningless. Reports were often poor, or totally absent, although this was a popular question.

Question 6

This question required candidates to apply theory to a specific scenario. In a question like this, it is important to answer the question – think what theory is relevant and how all information given in the question can be used. For example, some candidates failed to refer to the first graph. When you are asked for advice about a specific situation the answer should relate to that situation, and not simply repeat general theory. Always justify your answer – there can be marks for reasoning even if the examiner might not agree with your advice, but only if you have presented your reasoning.

Question 7

There were some good answers to this question. Marks were lost for imprecise reasoning or answers that did not relate to the question asked. Read the question carefully. Automatic model selection is fraught with difficulties; the applied statistician needs to be aware that choosing a good model involves far more than merely running such a method, and why.

Question 8

Exploratory analyses were not generally very convincing, and bookwork was poorly presented. Even candidates who are not familiar with a particular package are expected to be able to interpret output from it. In preparing for an exam like this, take special care to look at computer output and to check that you can interpret it.

Graduate Diploma Module 5 (Topics in applied statistics)

General

A paper like this requires candidates to be familiar with a number of applied methods: the rationale for the method; the basic theory; and how it is applied. Few candidates showed a depth of knowledge across sufficient different methods. Some failed to show much understanding of any method.

Many people can process data in a statistical package. A hallmark of a statistician is that he or she knows when such an analysis is appropriate and how to interpret it. This paper asks questions to test such skills.

Question 1

Most candidates were familiar with principal component analysis, although few could offer three reasons for its application. Many candidates lost marks in (iii) by being unable to repeat bookwork. Always justify an answer when interpreting data. If the question asks for advice then give advice; this question testing that you can interpret an analysis not just that you know in principle what such an analysis is for.

Question 2

This question was a fairly standard application of basic theory. Marks were lost for incomplete or inaccurate answers, especially in (iii)(c) and (iii)(d). Show your working carefully; marks can be awarded for methods – if the methods are clear – even if the answer is wrong.

Question 3

Answers were generally poor. Make sure that you are able to interpret output from statistical packages and that you are able to interpret coefficients in

models. The "best" model is not necessarily the one with the highest log likelihood. The applied statistician should also be aware of assumptions required for a model, how these can be checked, and the likely implications if they are not valid.

Question 4

The Weibull distribution is important in survival analysis. Not all candidates could derive the standard results in (i)(a), though most could identify some (but not all) of the different shapes of the hazard function. Parts (ii)(b) and (ii)(c) are fairly standard methods in survival analysis; learn how to do these.

Question 5

Be careful to interpret case-control studies as retrospective and not as prospective. Most candidates struggled with Mantel-Haenszel and with interpreting odds ratios. The applied statistician needs to be able to explain odds ratios to clients, and must know that the "no effect" value is 1 not 0. Adjusted odds ratios are common, and statisticians need to be able to compare adjusted and unadjusted values and to suggest reasons for the differences.

Question 6

This was a standard demography example with medical data; half the answers were quite good.

Question 7

This was standard sampling material, but few candidates made any headway with the minimisation theory. In addition, few people knew how to use the probability argument to find n when the interval width is specified. Most used $n = 227$, which is clearly wrong. With the width specified, two of the strata should have been sampled more than 100%; a few candidates gained credit by noticing this.

Question 8

Questions of planning are never easy, but simply quoting a standard list of rules for questionnaire surveys (as some candidates did) gains little credit. The first need here was clearly a sampling frame, and constructing one as none existed. The problems in this stage of planning were often totally ignored.