ROYAL STATISTICAL SOCIETY EXAMINATIONS, 2010

REPORTS OF EXAMINERS

General Comments

This report incorporates the comments made by examiners after marking the papers set in 2010 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 2011 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.

In 2010, the Graduate Diploma joined the Higher Certificate in becoming wholly modular. It has now been announced that the Ordinary Certificate will follow suit in 2011. 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 2010 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 2010 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.

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

The overall performance on this paper was satisfactory.

One of the commonest ways of losing marks was not appreciating the practical implications of aspects of situations described in the questions. It appears that training courses give lists of things to consider when dealing with different sampling problems but answers often reproduced these, in very similar words, while being theoretical to the extent of hardly referring to the actual question or suggesting things that would be impossible in practice. Too often, answers were much longer than necessary.

There were also several instances of a candidate not reading a question carefully. Many scripts showed a poor level of English, and some were barely legible.

Question 1

Quite a number of answers suggested that the description "self-service" had not been understood, and several people did not seem to know what "random" means. Sampling frames did not often get a mention; they are very important in any study.

Question 2

There were few comments on sample size in part (ii), and several poor methods suggested because of misunderstanding "self-service". Most, but not all, answers did have an element of multi-stage in them, if sometimes an unusual one.

Question 3

People in an age-group beginning at 18 are unlikely to be teaching in these colleges. Even if the standard list of age-groups does start at 18, something like 21 or 22 must surely be right for this question. Some of the covering letters and introductions were not very good; a few were excellent. It is a bad mistake to ask for people's names, especially when it is suggested that the college itself is doing the survey.

Question 4

In (i) “uniform sampling fraction” was interpreted as “the same sample size in each stratum” by a surprising number of candidates, even by some who were able to do the second part because they had the full instructions for it. Of course this error affected their answer to (iii).

Question 5

The (standard) answers to (i) were better than the answers to (ii), where possible disadvantages were often overlooked.

Question 6

Few people knew all the uses of a pilot survey. There was too much emphasis on the questions only, and relatively little on other aspects.

Question 7

“Get a better list” is hardly an answer to how the problems of not having a good list can be overcome. The emphasis on using the electoral roll was perhaps rather overdone – it does have its gaps and weaknesses, especially in the UK.

Question 8

Here many answers were irrelevant to the question as asked. It is tinned goods that are of interest. This is an observational study, and there appears to be a standard set of comments for that.

Ordinary Certificate Paper II

General

Most candidates attempted the entire paper. The general standard of presentation was good, with charts clearly drawn and annotated. The ‘comments’ parts of questions were often poorly answered and this indicates an area for improvement in the future. Candidates are reminded to read the questions carefully and to make sure that they do not omit any part of a question.

Question 1

Generally candidates were able to define the mode and the median though many did not explain exactly how to obtain the middle observation for the median. Some simply said that the mean was the average which was not sufficient. For

the choice of words some candidates did not realise that not all the words would be used (there were after all seven) and that words could be used more than once. Joshua's survey presented the most problem as candidates did not appreciate that it was categorical even though the categories were assigned numbers and thought that the mean would therefore be appropriate. For the final part many candidates misread the question and thought that they were being asked for the most appropriate measure and thus only gave one measure for each survey.

Question 2

This question was generally well answered with well presented tables. A few candidates presented the table the wrong way round and some left their entries as tally marks instead of numbers. The few who got their totals wrong would have been wise to spend a couple of minutes checking that there were 30 items in total and even counting the number of sugar of each category and the number of fat in each category. The questions that followed were generally correct except for (iv) and to a greater extent (v). Conditional probability (see also question 7) seems to be a less well understood area.

Question 3

The choice of chart, whether dual bar, stacked bar or back-to-back bar was usually appropriate. Several attempted to break the vertical scale which is not acceptable for a bar chart as it distorts the overall impression given. Some made life more difficult for themselves by choosing inappropriate scales. Multiples of 5 and 10 are much easier to work with than multiples of 3 or 4. The ratio asked for was Female:Male although several calculated Male:Female. On occasions the proportions or percentages of males and females was given rather than the ratio.

Question 4

The formula for calculating standard deviation was not always remembered correctly. It was clearly stated that a sample of beans was selected so (n-1) rather than n should have been used in the denominator of the variance. Several calculated the variance rather than the standard deviation. There was a tendency to round off too much, too early in the calculation. A few had no idea what the quartiles and the IQ range were. Students should be reminded that the coefficient of variation is a percentage. In the comments, very few were precise enough about what a higher coefficient of variation indicates. As the mean of the sample was higher than the norm, a larger absolute variation in the sample might

be expected but as there is a higher relative variation as measured by the CV, there is evidence that the merchant was justified in his views.

Question 5

This question was the least popular one. It did not take long to answer but tested whether candidates understand the basics of index numbers and percentages. Many were found wanting. Many referred to Jan 1987 as a base year rather than a base period or month. The process of rebasing was not well understood by many. The question asked for comments on the differences in price rises faced by pensioner and general households whereas many candidates discussed each series separately. Precise wording is important in the comments. At no stage did prices decrease, although in one month the rate of increase in prices declined. A results table was requested so any working should have been shown separately.

Question 6

The time charts were generally well drawn with axes ruled and clearly labelled and a key/legend included to distinguish the two series. The question asked for one benefit and one drawback of using the moving average method to estimate the trend. Many students answered this as though it were any method of estimating trend rather than specifically moving average. The whole point of a trend is to smooth out the short-term variations to find the underlying movement of a series so it is not a drawback if the method succeeds in doing this. The calculation of the moving average was generally good but the positioning of it at the centre of each seven-year period was not always correctly done. Many missed out the comments and the benefit/drawback parts of the question.

Question 7

This was answered better than probability questions in previous years. Again the definition of conditional probability was not done well. The diagram was generally correct though some candidates had trouble converting 1% and 5% to probabilities and some misread one or other of the given probabilities. Many candidates could calculate the required probability in (i) but then struggled with the conditional probabilities required in (ii) and (iii). The comments were variable with a few excellent ones and many not as good.

Question 8

Candidates seemed much more able to calculate the answers required in (i) than in previous years however some still did not appreciate the relationship between the sums of squares and products required and the correlation coefficient. Such

candidates then wasted significant time recalculating the coefficient from scratch. Some candidates who obtained an answer less than –1 or greater than +1 did not comment that this must be wrong. Generally candidates who finished part (i) managed the calculations in part (iii) also though some did think that the slope of the required line was equal to the correlation coefficient. The comments in part (ii) were generally reasonable though not many mentioned that the limits for perfect correlation were +1 and –1 and many did not explain that a negative coefficient means that one variable decreases as the other increases. The comments in part (iv) were less good with some thinking that the value was an estimate because only the sums and products had been given and not the original data.

Higher Certificate Module 1 (Data collection and interpretation)

General

Question 1 was very popular and was attempted by almost all candidates. The other questions were each attempted by about the same number of candidates. The general standard of the answers was good with few overall failures.

Question 1

Part (a) asked for comments on six questions which were thought to need amendment. Here candidates found several points needing attention. Part (b) asked for explanations and examples of why sensitive questions and hypothetical questions are likely to be troublesome. There were interesting suggestions here, and clear understanding of these types of question.

Question 2

Data were given on the heights of girls aged 6 and 10 born to six short mothers and to six tall mothers and the question asked for a report. Most candidates calculated measures of location and of spread for the six separate distributions and a few drew diagrams, but hardly anyone looked at the changes of height from ages 6 to 10. All realised that girls born to tall mothers were in general taller than those born to short mothers.

Question 3

This question had three parts, with most marks being given to the first part on distinguishing between stratified, cluster, and quota sampling. This was done well for the most part. The second part asked for a sample scheme involving cluster, stratified and quota sampling in a particular situation. The obvious combination was cluster sampling of regions, followed by stratification by urban/rural followed by quota sampling of persons to interview, but a few candidates suggested alternative and less satisfactory schemes. The last part of the question was on systematic sampling. This was the least well done part.

Question 4

The main part of the question was on non-response bias in a face-to face survey as caused by refusals and failure to contact sample members. Few candidates mentioned that those who do not take part tend to be different from those who do, but suggestions of how to reduce the bias were good. In the last part of the question, most candidates correctly calculated the three response rates.

Higher Certificate Module 2 (Probability models)

General

There was a poor overall standard of both home and overseas scripts this year. A disturbing number of scripts were derisory, showing fundamental incapacity in the algebraic and other mathematical skills that are vital for this subject. All potential candidates should note that a command of basic calculus and algebra, as outlined in the syllabus requirements, is essential. Particularly in this module, lack of prerequisite mathematical capability appears to be the principal barrier to progress, although many weaker candidates also struggle with conditional probability.

Question 1

Although generally popular, this question on the binomial distribution and relevant approximations had overall the lowest average marks. In part (i), many students failed to obtain $E(X)$ and $\mathrm{Var}(X)$, even for the B(2, p) distribution, and several wasted time attempting to derive these results even though the question said 'write down...'. Some answers were left with a general 'n' rather than '2' for the number of trials. The conditional probabilities were seldom well done: even when $P(X < 2)$ was correct, the intersection probability in the numerator was often found by multiplying $P(X = 0$ (or 1)) by the denominator, which leads to the unconditional probability and shows a fundamental lack of understanding. Part (ii)(a) was also weak, with very few clear and complete answers. The Normal approximation in (ii)(b) was better, but the continuity correction was often either wrong or omitted. A number of students failed to use the Poisson(4) distribution in (ii)(c), usually embarking on a normal approximation for which no credit was given. The final part (ii)(d) involved a Poisson approximation for the complementary random variable $200 - Y$, but only a few candidates completed it successfully. Several candidates who did (ii)(c) correctly nevertheless used a Normal approximation in (ii)(d), presumably because 'p' was not small. A few exact binomial calculations for parts (c) and (d) were seen, and given partial credit.

Question 2

This question on a Normal distribution along with a derived discrete distribution was also popular, and overall achieved the highest average score, with overseas candidates doing slightly (on average about 1 mark) better than those from the UK. Part (i) was generally well done. Part (ii)(a) involved a categorisation of underlying Normal data to build up the distribution of the discrete random variable C, but many weaker students were unable to carry this out effectively.

Most of those who obtained the distribution of C were able to find E(C), but Var(C) was more problematic. A common error in (ii)(c) was to misrepresent $C_T = C_1 + ... + C_{100000}$ as '100000C' with variance 100000^2 Var(C) rather than 100000 Var(C). A few students took Var(C_T) as Var(C)/100000, presumably due to confusion between mean and total. Although calculators normally give very accurate answers, many candidates rounded off intermediate answers to 3 or 2 significant figures, sometimes only one, causing serious inaccuracy when these results are used in later work. T he critical assessment asked for in part (ii)(d) was seldom well done. A few students appeared to have read this question as '*State* the assumptions...' and merely reiterated 'Normal and independent', although most answers recognised that independence of passengers was unlikely to hold, e.g. across members of the same family. Within-group correlation was occasionally mentioned, but without noting that the tendency to equalise baggage weights leads to a negative within-group correlation which would reduce Var(C_T). Further reasonable points made (other than those suggested in the pattern solution) were that the characteristics of passengers' luggage might vary with time, e.g. seasonally, and that it might be unreasonable to assume a single (Normal) distribution for the weights of all passengers' luggage.

Question 3

This question on the exponential distribution was rather less popular than questions 1 and 2. Although the cumulative integral of this distribution must be very often rehearsed in statistics courses, in part (i) many candidates were unable to reproduce it accurately, sign errors, incorrect limits and omission of the constant '1' being common. The graphical work was also disappointing, with many answers reproducing the pdf. Perhaps the given answer in part (ii) helped to ensure slightly better success here. In part (iii), many students obtained a correct equation for λ but failed either to cancel out the factor $\left(1-e^{-\lambda}\right)$ (leading to a linear equation in $e^{-\lambda}$) or to solve the unsimplified equation as a quadratic in $e^{-\lambda}$. A fair number of candidates obtained the correct result for $P(T > t|T > c)$, but several neglected to differentiate their result to obtain the pdf as asked. There were very few good demonstrations that $P(T–c > t|T > c) = e^{-\lambda t}$, although some answers did show awareness of the 'memoryless' property of this distribution.

Question 4

This less popular question was set to test knowledge of the Poisson distribution, along with a rider involving a Bayesian argument. In part (i), most candidates obtained the Poisson(2) distribution and many scored well, although the errors

$P(X > 2) = 1 - P(X < 2)$

and $P(X > 2) = 1 - P(X = 1 \text{ or } 2)$

were too often seen. Another common bad habit of writing algebra seen here was the inexplicable reluctance to use brackets to make the meaning clear, so that

$$1 - (P(X = 0) + P(X = 1) + P(X = 2))$$

became $1 - P(X = 0) + P(X = 1) + P(X = 2)$.

In the next part (ii), several candidates failed to notice that the boat was rigged with 100 (not 1000) metres of rope from companies A and B, but the methods used for finding P(no flaws) were usually sound. However, several candidates calculated $P(X_A = 1) \times P(X_B = 1)$ for part (ii)(b), which represents a particular way of getting 2 flaws and is of course wrong. A number of candidates also ruined their answers to parts (ii) and (iii) by working with 'average rope' in which the number of flaws was assumed to be Poisson distributed with weighted average means of $(0.2 + 0.3)/2 = 0.25$ in part (ii) and $(0.75 \times 4 + 0.25 \times 6) = 4.5$ in (iii). Many attempts at part (iii) did nevertheless succeed in establishing the Poisson(4) and Poisson(6) distributions and hence finding P(7 flaws|A) and P(7 flaws|B), and indeed several went on to find the 'inverse' probabilities P(A|7) and P(A|8) correctly. However, the necessary conditional probability argument was too much for many of the weaker candidates, a common error here being to represent the denominator probability as P(7 flaws|A) + P(7 flaws|B). Candidates with correct calculations in part (iii) might be presumed to have understood the analysis, but they did not always entirely successfully explain the contrast between P(A|7) and P(A|8).

Higher Certificate Module 3 (Basic statistical methods)

General

There were no particular problems with the paper in terms of wording and there was no evidence that candidates did not understand what was being asked of them. Most candidates attempted three questions but a few attempted two or four questions. Of the four questions, question 2 was attempted by the fewest candidates. There were some very good solutions produced by candidates to all four questions. Numerical work was by and large done well but interpretation and explanation tended to be less convincingly done.

Many candidates appeared to have no idea what a distribution is. For example, when asked for the distribution of the number of Type A manuscripts (question 1) many candidates wrote down the observed proportion of Type A manuscripts, when the answer "Binomial with n = 10 and success probability θ" was expected.

Many candidates confused significance levels with confidence levels and therefore wrote statements like "we therefore reject the null hypothesis at the 95% confidence level".

Many candidates seemed unaware that hypotheses are statements about the population, not about the sample.

Many candidates did not make it clear whether they were undertaking a two-tailed or one-tailed test or what significance level they were using. So, for example, some candidates calculated the value of a test statistic and then said only "so we reject the null hypothesis" without giving the significance level or the critical value from tables that they were using.

Some candidates did not put their results in context. For example, having done a hypothesis test, they concluded with "and so we cannot reject the null hypothesis" without saying what the null hypothesis is in terms of the particular application.

Question 1

(i) Common errors were: not calculating the pooled variance estimate correctly (it is a weighted average of variances, not a weighted average of standard deviations), not using the pooled variance estimate at all in the t statistic, not stating the conclusions clearly.
(ii) Many candidates were rather vague about what the assumptions underlying the two tests are.
(iii) The ranking calculations were mostly well done but some candidates clearly did not know how to use the statistical tables.

Question 2

(i) See general point 1 above.
(ii) See general point 3 above.
(iii) A two-tailed test is relevant here. A common error was to give the p-value for a one-tailed test.
(iv) This part gently guided candidates towards a Normal approximation to the binomial but this was not well done in the main (see general point 1 above).

(v) All the general points above applied here.

Question 3

(i) This was mostly done well. Some candidates claimed that if the 10 digits are equally likely this means that the digits generated must have a Poisson distribution and everthing went down hill from there.
(ii) Again most candidates did the required chi-squared test, though see general points 2 to 5 above.
(iii) Many candidates missed the "absolute value" statement or did not know what it means and consequently evaluated the wrong probability. Some candidates confused probability with an observed proportion.
(iv) Common errors here were the result of getting the wrong null hypothesis and not spotting that this is simply a test for a specified proportion (or equivalently a chi-squared test). Again general points 2 to 5 applied here.

Question 4

(a) This is simply bookwork on the definition of a sampling distribution and the standard results for the distributions of the sample mean and variance for a sample from a Normal population. Very few candidates scored well here.
(b)(i) This was largely done well using either a t interval or a z interval since n is large.
(ii) Those who knew the basic chi-squared result for sample variances tended to give good solutions. Others tended to guess a distribution and try to apply it (e.g. by pretending the sample variance has a t distribution).
(iii) Again those who knew the basic result about F tests tended to give good solutions. Those that did not tended to guess a distribution and try to apply it (e.g. by doing a t test on the variances).

Higher Certificate Module 4 (Linear models)

General

The general standard of both home and overseas scripts was reasonable. Despite a few derisory attempts, these results were overall far more satisfactory than those for some other papers, notably Module 2, possibly because this module is less demanding mathematically. Average marks on the four questions were broadly consistent.

Question 1

Although generally popular, this question on a completely randomised design and one-way ANOVA had lowish average marks. Many answers to the general bookwork part (i) were vague and/or poorly explained. Very few candidates mentioned the key point of comparing variations between and within treatments as the basis of analysis, and several candidates referred to blocking, which is not a feature of completely randomised designs. The model equation was sometimes omitted and occasionally stated imprecisely or in estimated form. The ANOVA calculation and test in (b)(i) was usually well done, although the null and alternative hypotheses were not always stated and a contextual conclusion was often omitted. In contrast, (b)(ii) was disappointing. Surprisingly few answers referred to the roughly linear increasing trend of treatment means with % cotton (which being on an interval scale thus naturally motivates consideration of regression), but many answers recognised the power of the regression model, if found satisfactory, to facilitate predictions of tensile strength at cotton mixes other than those tested. A few candidates wasted time by trying to fit the regression, although the question said not to.

Question 2

This question, on simple linear regression along with a linearising transformation, was fairly popular. Part (i) was generally well done, although a few candidates confused their graphs by using the same plotting symbols for points used to fix the fitted regression line as for the data. Although there were several good answers to part (ii), surprisingly many candidates failed to deal correctly with the simple log transformation required, betraying very poor algebra skills. Whilst many candidates did good calculations in part (iii), a sizeable minority used $\log_{10}$ rather than $\log_e$ (which is the intended meaning of 'log' as stated on the front of the exam paper), so condemning themselves to wrong and confusing results (including a fitted line quite separate from the linear trend of the data when plotted against $\log(x)$). Those who avoided these errors generally found the two model predictions successfully, but very few noted that the discrepancy in the first model prediction far exceeded the observable scatter of the data about the second fit.

Question 3

This question on the product-moment correlation coefficient (pmcc) r and Spearman's rank correlation r_s was the most popular of the four. The pmcc formula was reasonably well-known (albeit often in the form $\frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}}$ in which further definitions of S_{xy}, etc. were required for full marks), although some carelessness was evident in the omission of the $\sqrt{}$ and failure to square terms such as $\bar{x}$ where necessary. The formula for r_s was almost always given in the

form suitable for calculation, $1-\frac{6\sum d^2}{n\left(n^2-1\right)}$, although the question intended r_s to be given as the pmcc of the paired *x*- and *y*-ranks. Candidates also often failed to make clear that *r* measured linear association whereas r_s gave an assessment of the strength of any monotonic trend. In part (b)(i), the data were generally well plotted, and most answers identified either a slightly curvilinear increasing trend or a broadly linear trend with a possible outlier (the UK). The calculation of *r* and the subsequent test were generally good, but the relevant critical value was not always stated and the final contextual conclusion was often omitted. A few candidates wasted time by carrying out the equivalent *t* test, presumably unaware that the exact critical value is given in the RSS Tables. The calculation of r_s and the subsequent test were, however, less well done. Several candidates omitted to rank the data at all, simply taking $d = |x-y|$, while others ranked the combined *x*- and *y*-data as a single set of 20 values; both errors were of course fatal. The final part (b)(iv) was also rather weak. No candidate made the point that the data might not plausibly be a random sample from any obvious defined population as standard assumptions would require. Several answers noted that the two tests agreed in concluding positive association, but a common wrong idea was that the test giving the 'most significant' result was necessarily the better test to use. In contrast, the questionable linearity of the trend (unless the UK is omitted) was seldom mentioned as a reason for preferring the test based on r_s. A novel substitution of 'outliner' for 'outlier' was seen more than once.

Question 4

Just over half the candidates answered this less popular question on multiple regression. Part (i) on the standard assumptions for the error term was generally fairly well done. In part (ii)(a), some answers followed the hint of the question rather too slavishly by confining their comments to comparisons based on the root mean squared error without reference to linearity. The necessary tests in part (ii)(b) were usually correct, but not always carried out at the 5% level as asked. The equivalent *F* tests were sometimes done, and one-tailed tests and/or wrong degrees of freedom were sometimes used – difficulties that also attended the partial tests asked for in (ii)(c) (which were sometimes omitted altogether). The global *F* test was also not always done, and several candidates appeared to think that this test is two-tailed. 'R-sq = 96.7%' was generally well understood, in terms of attributable *y*-variation or as $\frac{\text{Regression Sum of Squares}}{\text{Total Sum of Squares}}$, but very few candidates noted from the output that it was calculated as $\frac{5220}{5400}$. The final part (ii)(d) showed room for improvement. Several candidates chose Model A because it achieves the highest value of R-sq, ignoring or unaware of the fact that the variable x_1 is not significant in the presence of x_2. In contrast, few

answers noted that x_2 *is* significant in the presence of x_1, or that Model C has both its parameters significant and achieves almost as good explanation and almost as small mean squared error as Model A.

Higher Certificate Module 5 (Further probability and inference)

General

This paper includes developments of the probability results and methods introduced in earlier modules, including bivariate distributions and generating functions, and also some of the concepts and methods of statistical inference. In order to be successful in this paper, as well as gaining familiarity with the new statistical and probabilistic ideas, candidates need good skills in algebra (e.g. in evaluating products when constructing likelihood functions) and differential and integral calculus. The standard was very varied, with 24% of candidates achieving a distinction grade but also a large number apparently not very well prepared.

Question 1

Parts (i) and (ii) were generally well answered. In part (iii), a number of candidates wrongly tried to evaluate the correlation instead of the covariance which caused them to waste time. Part (iv) was not well answered, many not realising that the joint distribution should be used to derive the distribution of $X - Y$.

Question 2

Very varied answers; several candidates obtained maximum marks while others made little progress. The joint distribution is a little intimidating, but knowledge of the distribution of X makes the solution to part (i) straightforward. Since the distribution of X has been given, parts (ii) and (iii) are standard questions.

Question 3

A standard question, generally well answered. However, a number of candidates did not multiply the densities correctly and quite a few threw away marks by not confirming that they had found the maximum of the likelihood.

Question 4

Another standard question, but surprisingly poorly answered. Answers to part (a) (definitions and explanations) were sometimes confused and inaccurate. Parts (i) and (ii) were generally satisfactorily answered, though many candidates did not say why the estimator was unsatisfactory. Few candidates made much progress with part (iv).

Higher Certificate Module 6 (Further applications of statistics)

Question 1

Analysis of variance is either well known or hardly known at all, with a basic ignorance of the assumptions underlying a linear model. There are several books on the reading lists, and examples on past papers, for candidates to work on it. In the later parts, following the instructions to calculate residuals was well done, but the "dot plot by groups" not quite so well (and people threw away marks by not plotting *S*, which they had been given).

Question 2

Regression questions have been surprisingly poorly answered in the past, and this year is no exception. People who did attempt (i) sometimes were able to derive the normal equations but not solve them; some of course did (i) and (ii) well. Regression computer programs are not as well-known as might sometimes be supposed, nor is the idea of checking assumptions.

Question 3

Degrees of freedom and the calculation of the interaction were disaster areas. This sort of thing appears in first year Psychology courses, let alone in Statistics. The result is easier to see in a graph if T is made the x-value and L the y-value, and different symbols for each level L will make the graph even clearer.

Question 4

This was the most popular question, and generally well answered except when people forgot to look at all the ways of reaching the accept/reject decision.

Higher Certificate Module 7 (Time series and index numbers)

Questions 1 and 2

The distributions of marks for these questions were poor.

Question 3

This was by far the least popular question this year – 10 candidates, fewer than half, attempted it. Several of them, however, had perfect or near-perfect answers. Note that, although the question referred to the Lowe index, it required no prior knowledge of this index.

Question 4

This was the most popular question this year – 22 candidates attempted it – and there were many excellent attempts at it. Most candidates gave the correct formulae in part (i), but some implemented them incorrectly by using *number of live births* instead of *live birth rate* (number of live births divided by number of women) in part (ii). Few, though, gave a satisfactory answer to the first part of part (iii), the description of the 'number of women index'.

Higher Certificate Module 8 (Survey sampling and estimation)

General

The standard this year was very mixed, with a distinct gap between the marks of those who failed and the lowest pass mark. The four questions all resulted in broadly similar average scores.

Graduate Diploma Module 1 (Probability distributions)

General

The overall standard of the answers offered was disappointing. Several candidates were well out of their depth, making little progress on almost all the questions they tried and making basic errors, such as misquoting the formula for a covariance and confusing Bayes' Theorem with the *definition* of conditional probability. A common fault was the failure to specify the region on which a formula for a density or distribution function would be valid. No single question proved substantially more or less popular than any other, but question 1 attracted most good answers.

Question 2

Only a small minority of those who tried this question could reproduce the bookwork for part (i), and attempts at part (iii) were poor.

Question 3

Many candidates wrote down the formula for the joint density of two independent variables, despite the question asking for the density of their sum.

Question 4

Several candidates failed to recognise that the geometric, not the binomial, was the appropriate distribution for the first part; those who attempted the last part generally made sensible remarks.

Question 5

Few candidates could successfully navigate part (iii), and attempts at part (iv) were largely perfunctory.

Question 6

This had a number of good answers but, as in parts of other questions, some candidates showed a lack of sensible examination technique: if a part of a question carries few marks, its answer cannot need two or three pages of complex manipulation.

Question 7

Some candidates recognised this question as based on the Box-Muller method, and had reasonable success with it. Work with Jacobeans was generally good.

Question 8

Simple manipulative slips marred several attempts at this question: if $U = Y - X$ and $V = Y + X$, a common slip was to deduce that $X = (U - V)/2$. Despite the clear instruction "sketch the region over which this joint density is non-zero", many candidates gave no sketch at all, while others offered a sketch of an arbitrary univariate distribution. The purpose of seeking this sketch was to guide candidates in their manipulations for the marginal distributions: inevitably, those who offered no sketch used incorrect limits in their integrations.

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 likelihood function appears regularly, so it is important that candidates can derive it, and take its log, for a random sample quickly and correctly.

Question 1

A popular question, generally well answered. In part (iii), for consistency both unbiasedness and variance tending to zero should be mentioned. In part (iv), it saves time if the gamma function is used. In part (iv), remember that although it is true that, if T_1 and T_2 are independent, then $E(T_1T_2) = E(T_1)E(T_2)$, it is not true that $\mathrm{Var}(T_1\ T_2) = \mathrm{Var}(T_1)\mathrm{Var}(T_2)$.

Question 2

It is vital in this question not to ignore the fact that $X = 0$ is not observed, i.e. a truncated distribution is obtained.

Question 3

Not a popular question, perhaps suggesting that decision theory has not been as thoroughly covered as it might be.

Question 4

In part (a) many candidates gave the Neyman-Pearson Lemma instead of the definition of a most powerful test. Most candidates correctly derived the form of the MP test, but answers to the later parts were varied.

Question 5

A standard but unpopular question, possibly because of lack of confidence in forming the likelihood function. However, there were some very good answers.

Question 6

The answers to part (a) were generally good; as a rough guide, with average size writing, the length of the answer to "essay" type parts should be about 2 to 3 lines per mark awarded, i.e. 14 to 21 lines for part (a). In part (b) (i), there is no need to use a Normal approximation. In part (b) (ii), often explanations were vague and inadequate.

Question 7

A routine question, surprisingly poorly answered in general. In part (i) basic properties of the Normal distribution can be assumed. In part (iv), explanations sometimes did not go into enough detail.

Question 8

A relatively new topic, so it is not surprising that this question was not very popular. However, there were some good answers.

Graduate Diploma Module 3 (Stochastic processes and time series)

General

This paper fell into two parts: five questions on stochastic processes and three on time series. Questions 6 and 8, on time series, turned out to be the most popular ones and the ones on which candidates achieved their best marks. Of the stochastic process questions, question 2 and question 3 were the ones attempted most frequently, with a number of good answers.

Question 1

Relatively few candidates showed that they were able to translate the Shewhart control scheme as described in the question into the transition matrix of a Markov chain, but some did and then went on to set up and solve the backward equations.

Question 2

A good number of candidates translated the quiz scenario into the transition probabilities of a Markov chain and then went on to find the corresponding stationary distribution. It was encouraging that a number of candidates successfully wrote down the six-step transition probabilities and calculated the expected winnings.

Question 3

A good number of candidates successfully attempted this question, which was a fairly standard gambler's ruin problem.

Question 4

This was the least popular question, perhaps because candidates were put off by having to deal with p.g.f.s and partial differential equations. Nevertheless, there were a few good attempts.

Question 5

A number of candidates successfully translated the queuing scenario as described in the question into the transition rates of a continuous time Markov chain and went on to find the corresponding equilibrium distribution. Candidates were asked to write down the **detailed** balance equations, but several wrote down the (global) balance equations, which made the task of finding the equilibrium distribution more difficult than it need have been.

Question 6

This turned out to be one of the most popular questions. There was a good number of more or less complete answers, which involved the solution of a second order difference equation to find the autocorrelation function of an AR(2) process.

Question 7

This was the least popular and least well done of the time series questions. It was rather disappointing how many candidates were unclear about how the sample acf and pacf of a time series may be used to suggest an appropriate model for the series.

Question 8

This was the other of the two most popular questions. Most candidates showed that they had understood Holt's method for forecasting and were able to apply it.

Graduate Diploma Module 4 (Modelling experimental data)

General

Some of the material in the Applied Statistic II paper of the traditional version of the Graduate Diploma is not now examined, so such a wide knowledge was not needed, but nevertheless the performance this year was encouraging. Candidates who scored highly were those who had read the questions carefully and answered the specific question that was being asked. In particular, good answers related the theory and analysis to the specific scenario outlined in the questions. Candidates scored good marks when they were precise in their answers and paid attention to detail.

Question 1

This was well done.

Question 2

This was not well done at all. The problem set here is a typical, and relatively easy, one of its type, where a statistician is involved at the planning stage of a programme and must give a view on what would be the best use of limited resources.

Question 3

In part (b) there is more than one possible, and useful, set of orthogonal contrasts, which would answer the questions being studied.

Question 4

The graph does suggest linearity, which could be more thoroughly tested in a later experiment.

Question 5

There was some confusion between 'linear relationship' and 'linear model'. Better answers described the models in terms of intercepts and slopes rather then just model parameters. Most candidates could provide an informal method for selecting the model; the better answers used the information provided in the table to formally compare models. Few candidates understood that R^2 can be very high even when a model is seriously inadequate, and that other diagnostics are necessary to assess goodness of fit.

Question 6

Few candidates attempted this question. Better answers distinguished between influence and leverage and gave a detailed and thorough interpretation of the regression diagnostics. It is important to identify unusual values in the table of diagnostics and then to say what is unusual about them and what this might mean. It is important to examine all the information given in the question.

Question 7

This question tested whether candidates could interpret the output from a log-linear model. This involved being able to communicate in non-technical language, identify the coding used in the model, and to interpret the output. The model selection required an understanding of the role of the scaled deviance. Better answers were those that gave precise answers to the questions asked.

Question 8

This question required candidates to specify a model, to discuss the way in which variables are coded and to interpret the output. It also asked for comments on the interpretation of confidence intervals, and the role of significance tests in pilot studies. In questions like these, few marks are obtained by repeating bookwork and the good answers were able to describe the practical interpretation of the outputs.

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. Overall, the performance this year was encouraging. Some candidates were very well prepared for some aspects of the syllabus but not for others. The topics for questions are not always predictable, so it is important to study the whole syllabus.

Candidates who scored highly were those who had read the questions carefully and answered the specific question that was being asked. In particular, good answers related the theory and analysis to the specific scenario outlined in the questions. Candidates scored good marks when they were precise in their answers and paid attention to detail.

Question 1

Good answers showed both an understanding of the mathematics underlying principal component analysis and also of the ways in which the method is used in a practical situation.

Some judgement is required in selecting and interpreting principal components and the good answers showed an understanding of the practical issues associated with the method. It is important to gain practice in interpreting output rather than just knowing how to generate it.

Question 2

The good answers addressed every part of the question, as asked. The cluster analysis is straightforward, but needs careful attention to detail. It is easy in an examination to make a simple error, so it is essential that all working is shown so that marks can be given for method.

Question 3

The first part of the question is bookwork. Answers should show each step and be accurate. Most candidates seemed familiar with Kaplan-Meier and understood some of the problems with data from time to failure (skewness and censoring). Not all were familiar with the Weibull distribution and its uses.

Question 4

Good answers showed an understanding of the Cox model, its assumptions, and how it could be used in practice. It is not sufficient just to know how a model can be fitted – it is also important to know what the underlying assumptions are,

how these are checked, and what the implications are if they are not met. This is true for any statistical model.

Question 5

The questions asked here are typical of those that may be asked of a consultant statistician. While some marks could be gained from bookwork and the application of standard definitions, good marks required the ability to interpret the computer output demonstrating a clear understanding of the meaning of the predicted values from the model. It is not uncommon for a logistic regression to have significant predictors but have poor discriminatory ability on the data.

Question 6

Demography questions, when tackled at all, are done really well, as this one was by most candidates. Some also got the idea of the effect asked for in (iii).

Question 7

Apart from missing the point about $\bar{x}$ being different from $\overline{X}$ most attempts at this were good. Either estimate is acceptable in (iii) as they give similar results.

Question 8

People were told in the footnote that they could assume results for continuous variables, but instead they assumed the discrete result for a single stratum and merely extended it to this general case. This affected two major sections of the question and so lost them several marks. Other candidates did know the dummy (0, 1) trick and used it correctly.