FHS Mathematics and Statistics Part B 2010
Examiners' Report

Part I

A. Statistics

1. **Numbers and percentages in each class**

   The figures in Table 1 are for Mathematics and Statistics. For comparison, Table 2 gives the corresponding figures for Mathematics.

	Number				Percentage			
Class	2010	2009	2008	2007	2010	2009	2008	2007
1	12	(10)	(13)	(9)	25.5	(33.3)	(32.5)	(24.3)
2(i)	22	(16)	(17)	(18)	46.8	(53.3)	(42.5)	(48.6)
2(ii)	9	(2)	(9)	(8)	19.2	(6.7)	(22.5)	(21.6)
3	4	(1)	(1)	(2)	8.5	(3.3)	(2.5)	(5.4)
Pass	0	(1)	(0)	(0)	0	(3.3)	(0)	(0)
Fail	0	(0)	(0)	(0)	0	(0)	(0)	(0)
Total	47	(30)	(40)	(37)	100	(100)	(100)	(100)

Table 1: Numbers and percentages of Mathematics and Statistics candidates in each class.

	Number				Percentage			
Class	2010	2009	2008	2007	2010	2009	2008	2007
1	55	(61)	(53)	(61)	35.7	(36.1)	(34.2)	(35.9)
2(i)	61	(76)	(74)	(79)	39.6	(45.0)	(47.7)	(46.5)
2(ii)	28	(23)	(22)	(25)	18.2	(13.6)	(14.2)	(14.7)
3	9	(5)	(5)	(4)	5.8	(3.0)	(3.2)	(2.4)
Pass	0	(3)	(1)	(1)	0	(1.8)	(0.7)	(0.6)
Fail	1	(1)	(0)	(0)	0.7	(0.6)	(0)	(0)
Total	154	(169)	(155)	(170)	100	(100)	(100)	(100)

Table 2: Numbers and percentages of Mathematics candidates in each class.

2. **Vivas**

   There are no vivas in Mathematics and Statistics.

3. **Marking of scripts**

   The following were double-marked: 3 whole-unit BE Extended Essays (in mathematics). The remaining scripts were single-marked according to detailed pre-agreed mark schemes. The extensive checking process for scripts and mark entry was the same as that used as in all parts of Mathematics/Mathematics and Statistics (see Part II, Section A); we are grateful to Helen Lowe for coordinating this process.

B. New examining methods and procedures

A number of changes were introduced in Part B 2009, the impact of which could not be fully assessed after just one year. No changes were made for 2010.

C. Suggested changes in examining methods and procedures

For Part B 2011, a new ("mathematics") whole unit option, Structured Project, and a new ("other") half unit option, Mathematics Education, are to be introduced. The assessment methods for these options involve new methodologies that are set out in the Examination Regulations for 2010–11 and the Syllabus and Synopses for Part B 2010–11.

For recommendations based on this year's examination process, see the end of Part II, Section A.

D. Notice of examination conventions to candidates

Candidates were given details of the examining conventions in the Notices to Candidates that were sent out by the Examiners. These notices also reminded candidates about the department's Examination Conventions document.

Part II

A. General comments on the examination

The internal examiners record their thanks to:

- the external examiners, Professor Bowman (Statistics external) and Professor Clarkson (Mathematics external)
- all assessors, particularly the actuarial assessor Paul King
- the administrative staff of the Department of Statistics, particularly Jan Boylan, and the Mathematical Institute administrative staff
- Waldemar Schlackow for managing the exam database
- the graduate students who acted as checkers for the scripts and mark entry.

Setting and checking of papers and marks processing

Course lecturers were appointed as assessors and provided draft questions, which were then checked by nominated checkers before being considered by the examiners. With one exception, lecturers marked the questions on their course.

The external examiners scrutinised the draft questions and provided helpful input which was incorporated into the final versions. (The Statistics external looked at the questions on statistics courses, the Mathematics externals looked at the questions on mathematics courses.) The actuarial assessor scrutinised and commented on papers BS3 and BS4.

Mark processing and checking was carried out according to the usual procedures: scripts were checked to ensure that all work had been marked and that all marks had been correctly totalled and recorded. In this way a small number of errors were corrected and each change was signed-off by one of the examiners.

Determination of University Standardised Marks

The Examiners followed the established procedures, outlined below, to determine the University standardised marks (USMs). When doing this the Examiners took note of:

- examiners' reports from 2009
- guidance from the MPLS Division passed on by the Statistics Academic Committee (and similarly from the Mathematics Teaching Committee) on the percentage of candidates that might be expected in each class
- reports solicited from assessors on the work they had marked, including assessors' estimates of where they considered class boundaries might fall for the scripts they had marked.

The scaling process described below was not used for a small number of units that were marked directly to the USM scale (this year, just 3 extended essays). For all other units, raw marks were converted to USMs by scaling; each half-unit was scaled separately, and if a candidate took both halves of a unit their USM for the unit was the average of the USMs for each half. (Different half-units have different scalings to allow for differences in difficulty.) For the purpose of determining the scaling, Mathematics, and Mathematics and Statistics candidates are not distinguished, they are treated as a single cohort. The algorithm used to provide a starting point for the Examiners' deliberations on scaling is as follows. It uses the average USM from Part A.

The initial map from raw mark r to USM has four segments. Suppose that for some particular half-unit, n_{70} (n_{60}) of the candidates had a USM of 70 (60) or above in Part A. Then a USM of 70 (60) is set at the raw mark, r_{70} (r_{60}) say, of the n_{70}'th (n_{60}'th) ranked candidate on that half-unit. Take the straight line through the two points $(r_{60}, 60)$ and $(r_{70}, 70)$ and truncate it at points P_1 and P_2 where it meets the lines USM $= 72$ and USM $= 57$ respectively. A line segment joins P_1 to $(50, 100)$, and another joins P_2 to $(0, 20)$. But the latter segment is broken, at P_3, where it intersects the line USM $= 37$, and the final segment of the map is from P_3 to $(0, 0)$. So the initial map from raw mark r to USM is the four segments joining $(0, 0)$, P_3, P_2, P_1, $(50, 100)$. (Note that the raw mark for a half-unit is out of a maximum of 50, while the USM is scaled to a maximum of 100.)

For a well-set paper taken by a large number of candidates, this algorithm yields a piecewise linear map which is close to linear, usually with a somewhat steeper first and last segment. If the paper is too easy or too difficult, or is taken by only a few candidates, then the algorithm can yield anomalous results. For all papers, changes to the initial maps are considered: usually these changes are made by adjusting the position of the points P_1, P_2, P_3 by hand so as to produce a fair map for each half-unit. As in past years, a preliminary meeting of examiners was held (three days before day 1 of the Final Examiners' Meeting) to assess the results produced by the algorithm and to make changes if necessary, so that the starting point for day 1 of the final meeting was a set of USM maps yielding a tentative class list with class percentages roughly in line with historic data.

Day 1 of the Final Examiners' Meeting began with an overview of the methodology and of this year's data, followed by a review of the mathematics and statistics papers taken by a large number of candidates. For each half of these papers, the data and initial scaling were scrutinised in turn, and the USM maps were adjusted as deemed appropriate. The Statistics external examiner was present for this session with the Mathematics examiners (who included the Chairman of Mathematics and Statistics) in order for him to get an overview of this year's exams, to observe the discussion of mathematics papers and thereby achieve consistency when participating in the discussion of the statistics papers (in doing this we followed the procedure setup in 2009).

The meeting then adjourned to allow the external examiners to look at scripts before reconvening to complete the task of adjusting USM maps for all papers. The Mathematics and Statistics examiners who were not Maths examiners joined the Maths examiners at the reconvened meeting, observing the adjusting and finalising of the USM maps for mathematics papers in order to achieve consistency when participating in the adjusting and finalising of the USM maps for statistics papers which followed (again, in doing this we followed the procedure setup in 2009). During this process the positions of borderlines and the operation of the Strong Paper Rule were considered.

At their final meeting the following afternoon, the Mathematics and Statistics examiners checked the positions of borderlines for their cohort, taking account of the operation of the Strong Paper Rule. Candidates near borderlines were considered in detail. The Examiners made a small number of final changes in order to arrive at a class list which, in their academic judgement, was in line with the candidates' performance.

The final positions of P_1, P_2, P_3 for the units/half-units taken by at least two Mathematics and Statistics candidates are given in Table 3. For the corresponding information for the remaining mathematics units/half-units, see the Examiners' Report on Mathematics Part B.

Paper	P_3	P_2	P_1
BS1	(24.2, 37)	(52.7, 57)	(72, 72)
BS2a	(9, 37)	(17, 57)	(41, 72)
BS3a	(9, 37)	(17, 57)	(41.6, 72)
BS3b	(9, 37)	(18, 57)	(34.2, 72)
BS4a	(10, 37)	(19.2, 57)	(39, 72)
BS4b	(9, 37)	(17.9, 57)	(41, 72)
B5a	(10.9, 37)	(23.8, 57)	(42, 72)
B5b	(13.2, 37)	(28.7, 57)	(42, 72)
B8a	(9.6, 37)	(24, 57)	(36, 72)
B8b	(9, 37)	(17, 57)	(38.2, 72)
B10a	(10, 37)	(20, 57)	(40, 72)
B10b	(9.8, 37)	(21.4, 57)	(39.4, 72)
B11a	(9, 37)	(16.1, 57)	(35.6, 72)
B21a	(9.1, 37)	(19.8, 57)	(40.8, 72)

Table 3: The points P_1, P_2, P_3 used for scaling. Note that the raw mark for each paper is out of 50, except BS1 which is out of 100.

Table 4 gives the rank of candidates by weighted average USM at the end of Part B.

Av USM	Rank	Candidates with this USM or higher	%
83	1	1	2.1
81	2	3	6.4
80	4	5	10.6
75	6	6	12.8
73	7	7	14.9
72	8	9	19.2
71	10	10	21.3
70	11	12	25.5
69	13	13	27.7
68	14	14	29.8
67	15	17	36.2
66	18	20	42.6
65	21	22	46.8
64	23	24	51.1
63	25	28	59.6
62	29	30	63.8
61	31	31	66.0
60	32	34	72.3
59	35	35	74.5
58	36	36	76.6
57	37	37	78.7
56	38	39	83.0
55	40	41	87.2
51	42	43	91.5
48	44	45	95.7
42	46	46	97.9
40	47	47	100

Table 4: Number and percentage of candidates scoring a given USM or higher at the end of Part B.

Comments on results

Considering the combined Mathematics/Mathematics and Statistics cohort (see the data in Tables 1 and 2), the overall percentage of Firsts this year (67 out of 201 candidates, 33.3%) is in line with divisional guidance, however the percentage of results of 2(ii) and below is higher than usual and higher than might be expected a priori.

Turning to the Mathematics and Statistics cohort, from Tables 1 and 2 the percentage of results of 2(i) and above is approximately the same as for Mathematics (last year it was slightly higher than for Mathematics). The percentage of Firsts (25.5%) is lower than for Mathematics and hence we gave considerable attention to this in our deliberations: we paid particular care to the scaling of BS1 (see below) and to candidates near the 1/2(i) borderline when making final changes. We note from the detailed data in Section C below that, this year, Mathematics candidates tended to score slightly higher than Mathematics and Statistics candidates on papers taken by a reasonable number of both types of candidate (this appears to be a change from last year).

At 2(ii) level and below, we were not able to reconcile the divisional guidance with the performance of Mathematics and Statistics candidates at Part A last year plus Part B this year: there was a higher percentage of results of 2(ii) and below than might have been expected – this was also true for Mathematics. Regarding this we note that the final scalings were often to scale up by more than might be desirable at the 2(i)/2(ii) level, e.g. candidates often receiving USMs of 60 and above despite scoring less than 50% in terms of raw marks.

We had concerns at the 2(i)/2(ii) borderline because of the low raw marks achieved by candidates at this level – as above, some of the raw marks at this level were scaled up quite a bit. This may indicate that the reduction of choice to two questions out of three on each section (introduced in 2009) may have had the effect of reducing average raw marks for weaker candidates while stronger candidates have continued to perform well. Low raw marks make it hard to assess candidates' knowledge and understanding, and to discriminate properly between them at a borderline, and we recommend below that higher raw marks are aimed for in future.

When scaling BS1 we viewed the initial scaling cautiously as it partly depends on Part A results and BS1 is arguably the Part B paper that is furthest from Part A: some different skills are required for applied statistics and, in particular, BS1 has a 34% contribution from assessed practical work. We found that BS1 was different from other papers and we straightened the original raw mark to USM mapping considerably. We judged that the appropriate scaling was no scaling for higher marks (e.g. a raw mark of 72 became a USM of 72), with some upward scaling (less than for other papers) at low marks. It seems likely that examiners will want to consider scaling BS1 in a similar way in future, as a special case, and we flag this for the future below.

The Strong Paper Rule did flag some candidates close to borderlines, though in the end no candidate dropped down a class because of the rule. Two candidates near the 1/2(i) borderline and one candidate near the 2(ii)/3 borderline were flagged by the rule and, following our final changes, all three of were awarded the higher classification. There was one further candidate below the 1/2(i) borderline who was close to having a high enough weighted average USM for a First, but who would then have been flagged by the rule.

Recommendations for the future

Currently setters are asked to set questions so that "a 2(i)/2(ii) borderline student should be able to gain at least half the marks". In view of the comments above, in particular that the raw marks of such candidates on whole units were frequently below 50 out of 100, we suggest that a target of 13 marks (out of 25) on a question is too low.

- We recommend that the target raw mark on a question for a 2(i)/2(ii) borderline student be 15 (or 16) out of 25.

This would involve setting up mark-schemes so that 15 (or 16) marks could be gained on material which has been seen before (in lectures or on problem sheets) or which requires straightforward adaptation of such material. It would also mean that the loading of marks was shifted towards the earlier and easier parts of questions, so enabling the papers to discriminate better between weaker candidates. A consequence would be that fewer raw marks would be available for harder parts of questions, but we do not consider this a problem: only a small number of scripts attracted nearly full marks this year, there were few USMs close to 100, and we did not see evidence of USMs close to 100 leading to unjustifiably high overall averages.

We are aware that the Mathematics Part B report recommends a target of 16 out of 25 for a 2(i)/2(ii) borderline student. An argument in favour of 15 out of 25 is that this corresponds exactly to the 2(i)/2(ii) borderline of 60 out of 100 without scaling. An argument in favour of 16 out of 25 is that students at the 2(i)/2(ii) borderline often have at least one weak question per paper and a target of 16 out of 25 makes some allowance for this.

As mentioned above, we also flag the issue of scaling BS1 for future examiners.

- We recommend that BS1 be considered a special case when scaling because of its different nature from other papers in Parts A and B, in particular because of its assessed practical work.

From our experience this year we anticipate that little (or no) scaling for higher marks (say 70 and above), and some scaling up for lower marks (but maybe less than for other papers), would be reasonable for BS1 in most years.

B. Equal opportunities issues and breakdown of the results by gender

Table 5 shows the numbers and percentages of male and female candidates in the various classes. There is no evidence for departure from a model in which gender and class are independent.

	Female		Male	
Class	Number	%	Number	%
1	4	16	8	36.4
2(i)	14	56	8	36.4
2(ii)	4	16	5	22.7
3	3	12	1	4.6
Pass	0	0	0	0
Fail	0	0	0	0
Total	25	100	22	100

Table 5: Breakdown of results by gender.

C. Detailed numbers on candidates' performance in each part of the examination

The performance of candidates is summarised below. The detailed question data is for Mathematics and Statistics candidates only.

Paper BS1 Applied Statistics

Number of candidates: 47 Average raw mark: 65.55 Average USM: 67.81
(Number of Maths candidates: 1)

Question	Average mark		StDev	Number of attempts	
	All	Used		Used	Unused
Q1	12.08	13.47	6.22	21	4
Q2	12.19	12.56	4.23	39	2
Q3	13.13	13.32	3.93	34	2
Q4	16.15	16.15	4.33	44	0
Q5	16	16		1	0
PR	24	24	3.71	47	0

Note that the individual question marks are out of 22 and the practical component (PR) is out of 34.

Paper BS2a Foundations of Statistical Inference

Number of candidates: 18 Average raw mark: 30.78 Average USM: 66.83
(Number of Maths candidates: 2 Average raw mark: 47.5 Average USM: 92.5)

Question	Average mark		StDev	Number of attempts	
	All	Used		Used	Unused
Q1	18.55	18.55	5.78	18	0
Q2	12.11	13.13	5.80	15	2
Q3	6.4	7.66	2.70	3	2

Paper BS3 Stochastic Modelling

Number of candidates: 44
BS3a: Av raw 24, Av USM 59.39, BS3b: Av raw 22.82, Av USM 58.02
(Number of Maths candidates: 9
BS3a: Av raw 33.22, Av USM 68.33, BS3b: Av raw 25.89, Av USM 63.44)

Question	Average mark		StDev	Number of attempts	
	All	Used		Used	Unused
Q1	12.56	12.56	6.30	16	0
Q2	13.40	13.95	5.59	42	2
Q3	8.80	8.96	6.10	30	1
Q4	14.59	14.59	6.55	42	0
Q5	8.52	8.75	4.39	33	1
Q6	6.64	8.5	5.70	12	5

Paper BS3a Applied Probability

Number of candidates: 3 Average raw mark: 22.67 Average USM: 60.33)
(Number of Maths candidates: 58 Average raw mark: 32.95 Average USM: 68.09)

Question	Average mark		StDev	Number of attempts	
	All	Used		Used	Unused
Q1	15	15	4.24	2	0
Q2	10.33	10.33	6.50	3	0
Q3	7	7		1	0

Paper BS4 Actuarial Science

Number of candidates: 42
BS4a: Av raw 27.4, Av USM 62.67, BS4b: Av raw 25.66, Av USM 59.39
(Maths: Number of candidates: 61
BS4a: Av raw 29.21, Av USM 64.8, BS4b: Av raw 27.62, Av USM 62.85)

Question	Average mark		StDev	Number of attempts	
	All	Used		Used	Unused
Q1	13.21	13.51	5.53	37	1
Q2	12.34	12.43	4.54	23	3
Q3	13.64	15.86	7.21	23	5
Q4	9.8	11.33	7.14	21	4
Q5	15.03	15.03	6.24	31	0
Q6	11.54	12	5.95	29	2

Paper B5a Techniques of Applied Mathematics

Number of candidates: 2 Average raw mark: 19.5 Average USM: 50
(Maths: Number of candidates: 6 Average raw mark: 33 Average USM: 66.67)

Paper B8a Mathematical Ecology and Biology

Number of candidates: 18 Average raw mark: 26.94 Average USM: 60.78
(Maths: Number of candidates: 47 Average raw mark: 26.6 Average USM: 60.36)

Paper B10 Martingales Through Measure Theory and Mathematical Models of Financial Derivatives

Number of candidates: 2
B10a: Av raw 36, Av USM 69, B10b: Av raw 31, Av USM 65
(Maths: Number of candidates: 13
B10a: Av raw 34.23, Av USM 69.54, B10b: Av raw 32.17, Av USM 67)

Paper B10b Mathematical Models of Financial Derivatives

Number of candidates: 33 Average raw mark: 29.76 Average USM: 64.36
(Maths: Number of candidates: 63 Average raw mark: 30.06 Average USM: 64.6)

Paper B11a Communication Theory

Number of candidates: 6 Average raw mark: 25.33 Average USM: 62.67
(Maths: Number of candidates: 36 Average raw mark: 27.08 Average USM: 64.44)

Paper B21a Numerical Solution of Differential Equations I

Number of candidates: 2 Average raw mark: 30.5 Average USM: 67
(Maths: Number of candidates: 18 Average raw mark: 30.11 Average USM: 65.5)

D. Comments on papers and individual questions

The following comments were submitted by assessors. (Statements suggesting where possible borderlines might lie have been removed, but the Examiners took note of this guidance when determining the USM maps.) We include comments on the statistics papers BS1–BS4; for comments on mathematics papers, see the Examiners' Report on Mathematics Part B. The comments relate to all candidates taking these papers, not just Mathematics and Statistics candidates.

BS1a Applied Statistics I

The questions seemed to be of approximately equal difficulty (marks correlate well across questions, taking the marks from questions that go forward, rather than the marks from questions attempted).

Q1 Unpopular GLM question. Several did this as a 3rd question which will not be counted toward their final mark. This was generally a bad idea, as their other two questions were in want of further refinement. Hardly any students attempted to 'Interpret the model' in 1(b)(ii), but a fair number correctly judged that the direction of ploughing made Easting the natural explanatory variable in 1(b)(iv). Students had difficulty defining deviance residuals (this is really just testing their memory) but were able to interpret them.

Q2 Many students did not know that the models in an ANOVA are nested, and treated the p-values as evidence for dropping single variables from the full model. Some threw out all the variables with large p-values. This happens to be correct; these students may have been lucky. Few students mentioned correlation at 2(a)(iv) and claimed without reflection that age was not explanatory for foot width! While many students saw that the test in 2(a)(v) is a one sided t-test, not many realised that they could get the p-value by halving the p-value for the F-test. Students who began by writing down the correct residual sum of squares in 2(b) generally finished the question correctly.

Q3 This is a standard format question from previous years and was popular. Students did fairly well on the now familiar 3(a). In 3(b), many students did not count the parameters in `Age~material+Layer` correctly. They either treated the variables as dimensionful, or didn't allow for the intercept as baseline. 3(b)(iii) is a standard F-test, and was fairly well done, subject to getting the degrees of freedom correct. Question 3(c) was completed by some 10 students. It was in fact an optional problem sheet question this year, so a well prepared student would have seen it.

BS1b Applied Statistics II

Of all students that answered a question in part b of the exam, one chose question 5 and the remainder chose question 4.

Question 5 was relatively well answered by this 1 student with the a few problems in the kernel regression part as in previous years.

Question 4 was very well answered in the bookwork parts of the questions a) and b), with many students achieving full points. The later parts c) and d) were also relatively well answered by the majority of students while a few students encountered difficulties here (as expected), often confusing the way the distribution of the Wilcoxon test statistics under the null are calculated for two-sample and single-sample tests.

BS1 Practicals: 1st practical

The results of the first assessed practical were more or less as expected. The majority scored 12–18 (out of 25), about a quarter scored more than 19, while there were a handful of cases with very low marks mainly because they did not attempt all the questions.

In the first question all attempted and most successfully did the exploratory analysis part. The biggest problem here was that many did not do the pairs plots and those that did were not able to interpret correctly. Almost everyone spotted the high correlation between the two densities. That fact led them to discard DensityA from their model, many of them even before fitting any model. Also almost everyone added the binary variable to indicate the different type of wood for boards I and J. Since most discarded DensityA a priori they ended up with the model with the pre-manufacturing variables. None tried the alternative model with the post-manufacturing variables. Another issue was that many students did not use the ANOVA command and thus the F-test in their model selection process. Instead they based their selection on AIC or the t-test for individual variables. We were expecting to see at least an F-test comparing their final model with the full model. That was also the case in questions 2 and 3. Diagnostics were in general fine.

In the second question the majority decided to drop variable DensityA before they even beginning their analysis because of the high correlation with DensityT. Therefore they ended up with a model with Pressure and Shelling as explanatory variables and were not able to find any physical interpretation for that model. Also in the initial model they did not use the additional binary variable for wood type which was used in question1.

In the third question, a good number of students fitted models with both DensityA and DensityT as response. In the first part of they question they got full marks even if they only used DensityT as response. The problem was in the second part of the question where only a handful of students were able to answer. The most common method they used to prove it was a t-test, where they only used only the data with (Pressure=30, Shelling=32) and (Pressure=35, Shelling=42). Thus they had two populations with 3 and 2 observations respectively to compare. Furthermore, even when they were doing that many used the values of DensityT instead of DensityA for the t-test. Another general comment is that many students had problems interpreting the p-values. They seemed confused about when a p-value favours or gives evidence to reject the null hypothesis. Sometimes they did not even appear certain which hypotheses they were testing.

BS1 Practicals: 2nd practical

The results of the second assessed practical showed that students do not have any problems in fitting and using simple cases of generalised linear models in R. However they have some real difficulties when things get a bit more complicated.

The problem that everybody encountered and no one properly addressed was that there was not a straightforward method to reduce the explanatory variables. Here straightforward means either the deviance (ANOVA) test or the AIC criterion. Thus most of the students started removing variables from the model without any justification. Nobody noticed the problem with speed level 5; had they noticed, model selection would have been much simpler. Instead they were removing individual points they identified as outliers, mostly by using the residual values and/or leverage and Cook's distance values.

A large number of students tried to use a transformation of the response. So instead of using a Poisson distribution for the accident counts they used the integer part of the log(counts) or sqrt(counts) and then fitted a Poisson GLM using these "rounded" observations. The justification they provided for doing so was that the raw data/counts had very large variance. Unfortunately many students followed this unorthodox and statistically wrong approach.

Furthermore many appeared to have problems interpreting the final model especially when this involved interactions. Also, not many used any kind of odds-ratio and misinterpreted the probabilities of accident for a given set of explanatory variables.

BS1 Practicals: 3rd practical

The results for the third assessed practical were significantly better than the previous two practicals. The overall performance was very good. Students found most parts quite straightforward.

The most problematic part in Question 1 was the Kolmogorov-Smirnov test for goodness-of-fit of the studentised residuals to a standard normal distribution. A large number of students have not simulated new data sets but instead they simulated standard normal samples for residuals.

In Question 2 not many students attempted part (d) or the question about finding out why IRLS is not converging (in part (c)).

BS1 Practicals: 4th practical

Students did extremely good work on this practical. There was not any specific part of the practical where they appear to have serious problems. The marks range from 15 to 25 (with one exception, who attempted less than a quarter of the practical).

They mostly lost marks in question 2, where they had to fit OLS, robust and resistant regression models. They seem to have some difficulties in commenting on the effect of the extreme outliers and in general giving an interpretation of the models.

Surprisingly many had problems with question 3, the bootstrap assessment of variability. Although they used this same method in lectures and classes and have almost identical R code to what they needed in the practical, some still made mistakes in their bootstrap code. Also many, instead of plotting the slopes against each other, made scatterplots of each slope separately, which made it difficult to compare afterwards.

BS2a Foundations of Statistical Inference

Q1. This question on the exponential family was attempted by most candidates and was reasonably done. Many, but not all, candidates attempted (e) on finding a maximum entropy prior, which distinguished good candidates.

Q2. This was a standard Bayesian Statistics Question where p in the Binomial has a prior Beta distribution. Most candidates attempted this question, with the standard parts (a)–(c) done reasonably and the harder parts (d) on a minimax estimator, (e),(f) on Bayesian hypothesis testing poorly done.

Q3. Very few candidates attempted this question and their marks were low. There were k normal groups of data with normal priors for the mean and hyperpriors for the prior means and variance.

BS3a Applied Probability

Q1: (a)–(c) Straightforward bookwork, generally well-attempted. (d) Computing size-biasing. Satisfactory overall. (e) Attempts began correctly, some confusion arising.

Q2: (a)–(c) Very straightforward. (d) A standard jump-chain calculation, well attempted. (e) Convergence to equilibrium: calculations most students did, but justifications were quite often incomplete or absent.

Q3: (a) Trivial bookwork. (a) A simple split into uses. Some confusion. (b) 'Thinning': About half got it, some by computation. (c)(i) Fairly well attempted. (c)(ii) Some saw geometric structure and computed fine. (c)(iii) Only a few students cracked all details.

In summary Q1 and Q2 were fairly straightforward. In particular, all students attempted question 2, which included a very straightforward (a)–(c), totalling 9 points (the vast majority got at least 7). Q3 did include some more understanding, particularly for the latter parts of part (c).

BS3b Statistical Lifetime Models

There was a huge gap between performance in question 4 – which was generally good – and performance in questions 5 and 6. This exam – and particularly questions 5 and 6 – differs from some of the earlier exams in BS3b in being more problem-solving oriented. The questions were not particularly mathematical, did not require much algebra, and did not require students to recognise a trick or be ingenious, but they did need to integrate information that was presented in a somewhat nonstandard way. Thus, the latter part of question 5 asked them to do perfectly standard things with proportional hazards models and matrix models – things that they have done and could do (at least, those who came to classes) on problem sheets – but without those keywords to trigger them. Question 6 seemed to be done in a rush, sometimes after submitting work for question 5 as well. The students are clearly struggling (with a few exceptions) to generalise their knowledge. It probably would be good to make more points available simply for appropriately defining major concepts and reprising some important derivations from the lectures.

Question 1: (a) About half the students got this completely right, either by the model method, or by the essentially correct alternative $-2\log(1-\frac{1}{2}q_x)$. The others mostly got confused about the time period, and had essentially half the correct rate. Some computed only ${}_1q_x = 1-(1-\frac{1}{2}q_x)^2$, but didn't convert it to actual rates. (b) Most did this essentially right, except for having the wrong rates from part (a). (c) Mostly completely right or nothing right. Some had an essentially correct formula, but substituted or calculated incorrectly. (d) Most got full credit. A few got some credit for saying ${}_sq_x = s_1q_x$, but not knowing how to use that. (e) Most got part (i). Part (ii) had many parts: They had to say what the mean and variance are, mention binomial or hypergeometric distribution, mention independence, uncorrelated, or martingale, and the Central Limit Theorem (or at least say something sensible about asymptotic normality).

Question 2: This problem turned out to be very hard. Depressingly, because the only thing that was hard about it was that there were relatively few keyword cues to which tools from the course they needed to apply. Once they knew what the question was about, the methods were relatively straightforward, but few got that far. Few recognised the link to proportional hazards models, which lost a number of marks right there.

(a)(i) Some (incorrectly) treated mortality as constant. (a)(iii) Many suggested increasing the sample size, which would not necessarily make it "more accurate". (b) If they mentioned both right censoring and left truncation, they got some marks, even if they didn't provide adequate explanation. Some had sensible explanations but didn't directly relate to the transition to Stable, others had the right terms but the wrong explanation. Some used a wrong term (for instance, left censoring instead of left truncation), but had an explanation that showed they actually knew which thing they were talking about. (c)(i) Some vaguely said that we could combine data from the two groups, without explaining how that might work. (c)(ii) Partial credit for an elaborate solution (mostly but not completely correct) that depended on the simplifying assumption of constant mortality rates. One student had a not-quite correct solution that worked from the right principles. (c)(iii) Some neglected to convert survival probabilities into mortality rates. Some didn't use the matrix model at all, but did what would have been a sensible computation for a single-decrement model (showing that they know the relevant techniques).

Question 3: Very few students chose this question, and no one did it very well. Most seemed to do it in a great hurry. (a) This was very similar to an exam question from a couple of years ago, one that we went over in revision classes, and yet most students simply worked under the (incorrect) assumption that the times T_i were iid, rather than their differences. (b) Some seemed to have no understanding of what unbiased means. Or they made vague statements like "since this is not using all the possible information" it must be biased. (No credit for this answer.) No credit for those who claimed (mistakenly) that the sum of the times is gamma distributed, and therefore the estimator is unbiased, since it still would be biased in that case. Very few actually understood that they needed to compute an expectation. (c)(i) Surprisingly little success here. Again, this is something they've done many times, but the context seems to have thrown them off. A few people at least showed that they understood what event they were supposed to be computing the probability of, though they then erred in computing it. (c)(ii) Most did pretty well. Some calculation errors. (c)(iii) Many reprised the erroneous approach they applied to part (a), and got substantial partial credit if they followed that through consistently. (c)(iv) Full credit to those who did correct calculations of variance, and applied them correctly, based on erroneous likelihoods from previous sections. Partial credit to people (a few) who showed they understood that they needed to use a normal approximation and add the variances of the two estimators, but who computed the variance incorrectly.

BS4a Actuarial Science I

Question 1: This question was the most popular question attracting both many very good answers and quite a few very poor answers. All candidates knew what a yield is but many lost a mark or two in (a) for not writing it down carefully. (bi) was mostly well-done, except that many candidates did not realise that the discounted value of the perpetuity is $1/i$ only when $i > 0$. For (bii), most candidates did not remember the concept of a running yield, which would have only cost them two marks, but some wrote pages and pages of guesswork losing time as well as marks, instead of focussing on what they could do. Almost everybody could do (biii); some only calculated the net yield, not the gross yield (before tax). In (biv), some did not see that the question was getting at time value of money, but most of those who did scored some cheap marks here.

Question 2: This question was attempted by many, but not as well-done as I had expected. Many could not formulate the two methods in (a), bookwork, although most could apply them in (b), and only a handful could produce reasonable proofs; students with special cases (avoiding time-dependent forces of interest) and intuitive arguments were generously rewarded. (bi) was very well-done by most students; a few marks were lost where students confused nominal and effective rates or annual and monthly amounts – while 3% was given unambiguously as a nominal rate, I accepted 5% as either nominal or effective as the wording was not quite clear. There was a wide spread of marks for (bii), quite a few full marks, but many who did not discuss the (optimal) strategy of repaying 10% of the loan outstanding at year 4 and the remainder at year 5; a few students assumed that the monthly repayment amounts changed with overpayments, which made their answers harder than assuming that payments stayed the same (effectively reducing the term of the mortgage). Those attempting (c) generally scored two or three marks for the set-up, but only one person multiplied by p to get annual *rates* from actual p^{th}ly interest components to attract the final mark.

Question 3: This question was a bit less popular than the other two, but those who seriously attempted it generally scored good marks with a very good number of excellent answers. (a), and also (b), were done well, but a non-negligible proportion of students displayed that they had not learned material about time-varying forces of interest and accumulation factors. (ci) was fine, just some students forgot to divide the coupon rate by p, for the coupon payments. Some marks were lost in (cii) for confusing nominal and annual rates; more so, time was lost by unnecessarily calculating annual rates in the first place, whether correctly or wrongly applied afterwards. (ciii) had fewer attempts than the other parts. There was a slight typo in (ciii) – a missing index k in the mean of the stochastic rates – this was announced near the beginning of the exam and only one student mentioned this in their script; all attempts I have seen seemed unaffected by the typo. Nobody found the correct standard deviation; marks were awarded generously for attempts, but false independence claims led to a loss of a mark.

BS4b Actuarial Science II

None of the questions seemed particularly more or less popular than the others.

Q4: For (a), saying what a fund being immunised meant caused surprising problems – some candidates gave conditions under which a fund is immunised instead – and plenty of candidates did not get full marks on (a). In (b) many candidates could find the two equations satisfied by A and C, but some then had trouble determining the values of A and C. Part (c) was more challenging – most candidates found this part difficult. Most who tried (d) managed to pick up some marks, though not many scored full marks.

Q5: Some candidates seemed to find (a) and (b) quite hard for the first two parts of a question. Many candidates could make reasonable attempts and earn some of the marks available for (a)–(e) and (f)(i), but there were only a few convincing attempts at (f)(ii).

Q6: Many candidates lost marks in (a)(i) by only attempting the "if and only if" proof in one direction – some candidates were not able to prove it fully in the other direction, but other candidates seemed not to notice that they had missed out part of the required proof. Part (a)(ii) was done quite well, but many candidates lost marks in (a)(iii) by not producing a precise description as asked. In general attempts at (b) were quite good: when unable to do this fully, many candidates were able to do something sensible and earned some marks.

E. Comments on the performance of identifiable individuals and other material which would usually be treated as reserved business

[*Removed in web version*]

F. Names of members of the Board of Examiners

Examiners: Prof AW Bowman (External), Prof PA Clarkson (External), Dr E Gaffney, Dr CN Laws (Chairman), Prof CJH McDiarmid, Prof NMJ Woodhouse.

Assessors for Papers BS1–BS4: Mr D Clarke, Prof RC Griffiths, Dr A Hammond, Mr PR King, Prof SL Lauritzen, Dr NF Meinshausen, Dr GK Nicholls, Dr DR Steinsaltz, Dr M Winkel.

29-09-2010