APPLICATION OF A MODEL FOR BORDERLINE MARKS

Dr. M. F. AL-BAYATTI

Liverpool John Moores University

School of Computing & Mathematical

Sciences, James Parsons Building,

Byrom Street,

Liverpool L3 3AF

U.K.

e-mail: M.F.ALBAYATTI@livjm.ac.uk

Abstract

The purpose of this study is to provide a rational approach for the adjustment of student marks and for the measurement of module effects. A statistical model is defined that incorporates the effects of different modules and different student ability. The model is applied to two sets of data and compared with the usual approach. Highly significant differences between marks of different modules are found and it is argued that this model is preferable.

1. Introduction

We consider the problem of borderline marks in the School of Computing and Mathematical Sciences (CMS) at Liverpool John Moores University (JMU). There are five main programmes : Computer Studies (CS), Mathematics Statistics and Computing (MSC), Applied Statistics and Computing (ASC), Software Engineering (SE) and Higher National Diploma in Computing (HND). The first four programmes of study are covered by the Integrated Credit Scheme (ICS) regulations: each module has a different number of credits and each student has to obtain a certain number of credits at each level. The fifth programme is a non-ICS programme in which all students have to do the same modules which are equally weighted.

Assessment is managed by the Module Assessment Board (MAB) and the Programmes Assessment Board (PAB). In the PAB meeting, for ICS modules, a failure may be compensated if (a) 80% of the credit required at that level has been gained, (b) the level mark is at least 40% and (c) a mark of at least 30% has been achieved in failed modules. For non-ICS programmes, if a student scores between 30% and 39% in a module, then compensation may be permitted. In deciding upon compensation, consideration is given to the student's overall performance in that module and elsewhere on the course. Compensation is allowed in only one module in each year for non-ICS programmes (HND Computing - Student Handbook (1994)). However a mark/marks less than 30% (say 29%) may be discussed in the MAB meeting to allow compensation for ICS or non-ICS modules before the meeting of the PAB. The borderline marks are those marks that fall within the compensatable range for ICS and non-ICS programmes.

A statistical model is defined in the next section that provides a rational approach for the adjustment of student marks and for the measurement of module effects. The model is applied to marks obtained at JMU and compared with the usual approach. It is argued that this model is preferable. More importantly, the model can be applied easily at other institutions: it is not just appropriate for JMU.

2.The model

We aim to adjust examination marks to correct for the fact that different students take different subjects (or modules), which are assessed with lower or higher marks. The main factors which contribute to determining an examination mark Y_{ij} for student S_i in module P_j are:

1. The difficulty of the module
2. Examination difficulty for that module
3. Severity of marking
4. Lecturer/Examiner
5. Student ability
6. Interactions between the student and factors 1 to 4
7. Random variations

For practical purposes, the factors 1 to 4 are combined in a single measurement, β_j, for module j and the interactions, 6, are absorbed into the error terms, 7, denoted by the symbol e_{ij}.
The ability of student i (factor 5) is represented by α_i. Error terms are assumed to be independent with zero mean and the variance of e_{ij} is assumed to be independent of i or j.

Model 1 for the mark for student i in module j is

$$Y_{ij} = m_j \alpha_i + e_{ij} \qquad \ldots(1)$$

where m_j is the credit for module j and α_i represents the student's mark per credit. Model (1) is the usual implicit model, where the average of the marks Y_{ij} (or $m_j Y_{ij}$), is used to indicate student i's overall performance.

Neither the automatic nor the non-automatic compensation awarded to students has a scientific basis, because no account has been taken of the different modules involved in examinations. As it may be easier for students to obtain higher marks in some modules than in others, some students suffer handicap because they take modules in which relatively low marks have been given. They could then have a smaller chance of passing. In other words credits should not be awarded for either ICS or non-ICS modules by compensation, as if the factors 1 to 4 do not exist. Account of the different modules involved should be taken into consideration for students' results to be regarded as comparable.

The following model recommended by Green, Baldock and AL-Bayatti (1981) for the adjustment of examination marks is applied.

$$Y_{ij} = m_j(\alpha_i + \beta_j) + e_{ij} \qquad \ldots(2)$$

where m_j represents the number of credits for module j and β_j is the effect of the jth module. Let q_j be the number of students taking module j. We assume that the e_{ij} are independent and have distribution $N(0, m_j\sigma^2)$, and that

$$\sum K_j \beta_j = 0, \qquad \ldots(3)$$

where $K_j = m_j q_j$.

Using the method of least squares, we minimise, subject to equation (3), the weighted sum of squares:

$$S = \sum\sum \{Y_{ij} - m_j(\alpha_i + \beta_j)\}^2 / m_j + \lambda \sum K_j \beta_j, \qquad \ldots(4)$$

where λ is the Lagrangian multiplier, and $\sum\sum$ represents summation of all cases where i and j occur together. It is easily found that $\lambda = 0$, Haeussler and Paul (1996), and

$$\sum_{i(j)} \{Y_{ij} - m_j(\alpha_i + \beta_j)\} = 0 \qquad \ldots(5)$$

$$\sum_{j(i)} \{Y_{ij} - m_j(\alpha_i + \beta_j)\} = 0 \qquad \ldots(6)$$

Mark adjustment based on (5) and (6) has been practised by a number of examining authorities. The adjustment principle has been to arrange for the mean mark of each paper to be shifted so that it becomes equal to the mean estimated ability of the students taking that module, $\{\alpha_i\}$ from (6) regarded as a formula for the estimated ability of students $\{i\}$ and (5) expresses the adjustment principle.

This type of problem has been discussed by Hasofer (1977) and Backhouse (1978) who used a model similar to model (2) with equally weighted marks. A reference to the analysis of an unbalanced blocks design is given by Scheffe' (1959, PP. 112-119) with $m_j = 1$ and unequal number of students taking paper j. The papers of Biggins et al. (1986) and Biggins and Yue (1993) consider the problem of combining examination marks when not every student takes every paper and obtain an overall measure of each student's performance in the examination. Daley and Seneta (1986) use a multiplicative model for the adjustment of examination marks.

Assuming, the usual approximate normality of the data (Central limit Theorem) we can perform an F test to compare Models 1 and 2, using normal linear theory. Suppose that the sums of squared residuals about the fitted models are Q_1 and Q_2 with degrees of freedom V_1 and V_2 respectively, where

$$V_1 = \sum\sum m_j - \sum{}_i S_i$$
$$V_2 = \sum\sum m_j - \sum{}_i S_i - \sum{}_j P_j$$

The test criterion is

$$F = \frac{(Q_1 - Q_2)/(V_1 - V_2)}{Q_2 / V_2} \qquad \ldots(7)$$

distributed as F with $V_1 - V_2$, V_2 degrees of freedom.

3.Model's application

Models (1) and (2) are compared using two sets of examination marks in CMS at JMU, one from the ICS and the other from non-ICS programmes. These two sets are level 1 examination marks for ASC and HND programmes of study in one of the recent years. To find the adjusted marks we need to apply the models to all examination marks for each programme of study.

3.1 ASC data

There were 12 students and 14 modules in this programme. Table 1 shows the original marks and Table 2 the original and adjusted student averages with the effect of the modules according to the linear least squares model.

TABLE 1

ORIGINAL MARKS (PER CREDIT) OF ASC

Module	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Credit Studt.(i)	1	2	2	3	2	3	2	3	5	2	2	2	3	3
1	38		27	44	29	45	42	29	37		3	42		26
2	64	43	34	66	55	76	69	58	75		72	74		64
3	48	53	47	69	39	51	70	72	24		40	19		15
4	77		56	73		75	89	86	79	70	84	85		90
5	49		44	50	23	59	61	60	42		53	46		45
6	63	60	81	73	58	59	88	63	74		59	78		72
7			49	48	35	56	66	42	37		55	34		40
8	32		62	75	54	63	72	68	71		29	59	33	63
9	49	51	27	46	57	45	66	74	30		63	37		40
10	58	51	80	68	62	75	81	81	79		71	84		89
11	80	53	44	70	71	80	88	65	74		87	64		72
12	59		70	65	74	50	68	59	45	45	41	65		48

TABLE 2

ORIGINAL AND ADJUSTED STUDENT AVERAGES
AND THE EFFECTS OF MODULES (ASC)

	Student Averages (α)			
i	Model 1	Model 2	j	β For Model 2
1	33.61	33.12	1	- 3
2	64.17	64.39	2	- 10
3	44.17	44.39	3	- 6
4	79.00	78.84	4	4
5	48.39	47.91	5	- 5
6	69.40	69.62	6	3
7	45.22	44.61	7	14
8	59.52	61.87	8	5
9	47.20	47.42	9	- 2
10	75.00	75.22	10	- 10
11	70.83	71.06	11	- 3
12	55.87	56.07	12	- 1
			13	- 29
			14	- 3

The borderline marks were obtained by student numbers 2, 5 and 8. Table 3 shows the adjusted marks for those students.

TABLE 3
ADJUSTED MARKS

Stud. No.	M o d u l e s												
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
2	61	54	58	69	59	68	78	70	62		61	64		62
5	45		42	52	43	51	62	53	46		45	47		45
8	59		56	66	57	65	76	67	60		59	61	33	59

The suggested model has changed the borderline marks, for student 2 in module 3, student 5 in module 5 and student 8 in module 1 and 11, to pass marks. If the mark in module 11 for student 8 had been changed in the MAB meeting from 29 to 30, as happens sometimes, he/she would be compensated under the ICS regulations and awarded the credits for modules 1, 11 and 13. Model (2) has not changed the marks for module 13 because this student is the only one in this programme of study that has taken this module. Probably module 10 would cause only a small change in any estimation but it does provide a little information. In other words there were not enough students to give a good indication of the difficulty of this module.

An improvement that can be made, in such a case, is to apply the model on the school level (i.e. marks from a number of programmes of study). Different groups of students take different combinations of modules, and model (2) is applicable for unbalanced data.

To compare Models (1) and (2), the residual mean squares are calculated and shown in Table 4

TABLE 4
MEAN SQUARES ABOUT THE FITTED MODELS

	d. f.	Sum of squares	Mean squares
Model 1	340	585242.600	1721.300
Model 2	326	39382.440	120.805

A simple F test is performed as follows:

$$F_{14,326} = \frac{545860.16/14}{39382.44/326} = 322.752$$

The critical value is $F_{14,326}(0.001) = 2.64$. Thus 322.752 is significant at the 0.1% level, giving strong evidence of definite differences between marks obtained on different modules. It follows that Model (2) is superior to Model (1) in representing the data.

3.2 HND data

Some modules in this programme of study have two elements, Coursework and Examination. A mark less than 40% in either element is considered a fail. Therefore, we treat each element as a module. It is inconvenient to include the original marks in a table like Table 1, there being 54 students and 8 elements. Here we apply Model (2) with $m_j = 1$ for all j, since all the elements are equally weighted. Table 5 shows the original and adjusted students averages with the effect of the modules according to the linear least squares model.

TABLE 5
ORIGINAL AND ADJUSTED STUDENT AVERAGES
AND THE EFFECTS OF MODULES (HND)

	Student Averages				Student Averages	
i or j	Model 1	Model 2	β	i	Model 1	Model 2
1	79.50	79.55	9	28	48.88	48.92
2	77.75	77.80	-9	29	52.13	52.17
3	76.75	76.80	6	30	51.88	51.92
4	73.38	73.42	-3	31	47.88	47.92
5	69.88	69.92	4	32	46.25	46.30
6	69.13	69.17	3	33	45.75	45.80
7	66.13	66.17	-9	34	45.25	45.30
8	65.88	65.92	2	35	45.50	49.40
9	68.50	68.55		36	46.13	46.17
10	62.63	62.67		37	46.00	46.05
11	63.25	63.30		38	42.63	42.67
12	63.88	63.92		39	44.50	44.55
13	60.63	60.67		40	41.50	41.55
14	60.88	60.92		41	40.50	40.55
15	60.00	60.05		42	41.38	41.42
16	63.00	63.05		43	39.00	39.05
17	58.50	58.55		44	37.63	37.67
18	58.63	58.67		45	34.88	34.92
19	59.38	59.42		46	34.75	34.80
20	57.50	57.55		47	28.50	28.55
21	56.13	56.17		48	19.75	19.80
22	57.75	57.80		49	17.00	17.05
23	53.13	53.17		50	18.50	18.55
24	56.38	56.42		51	15.75	15.80
25	55.63	55.67		52	15.25	15.30
26	53.25	53.30		53	11.25	11.30
27	50.13	50.17		54	11.50	11.55

The values of β_j represent, in order, the effect of first and second elements of module 1, first and second elements of module 2, first element of module 3, first and second elements of module 4 and first element of module 5. These values suggested that all Coursework elements have positive effects while all the written Examination elements have negative effects.

To compare Models (1) and (2) the residual mean squares were calculated and are shown in Table 6.

An F test is performed, using (7), as follows:

$$F_{8,370} = \frac{16698.840/8}{43110.540/370} = 17.915$$

As the critical value for $F_{8,370}(0.001)$ is 3.27, 17.915 is significant at the 0.1% level, giving strong evidence of differences between marks obtained on different modules. The superiority of Model (2) over Model (1) in representing the HND data is clear.

TABLE 6

MEAN SQUARES ABOUT THE FITTED MODELS

	d. f.	Sum of squares	Mean squares
Model 1	378	59809.380	158.225
Model 2	370	43110.540	116.514

4. Conclusion

The examination assessment of students' performance without taking into account other students' performances, is unsatisfactory. The ASC and HND sets of data provide strong evidence of differences between the marks for different modules which are equally or unequally weighted. The model considered here has retained the same overall average mark as in the original data.

Model (2) helps an examination board assess student performance on different modules. It accommodates the fact that higher marks might be recorded on some modules than on other modules. Model (2) can be easily applied at other institutions.

References

BACKHOUSE, J, K. (1978). Comparability of Grading Standards in Science Subjects at GCE A-level. (School Council Examinations Bulletin 39.) London: Evans/Methuen Educational.

Biggins, J.D., Loynes, R.M. and Walker, A.N. (1986). Combining examination marks. British Journal of Mathematical and Statistical Psychology, 36, 150-167.

Biggins, J.D. and Yue, K.K. (1993). Scaling and combining examination marks. British Journal of Mathematical and Statistical Psychology, 46, 153-179.

Daley, D.J. and Seneta, E. (1986). Modelling examination marks. Australian Journal of Statistics, 28(2), 143-153.

GREEN, J. R., BALDOCK, G. R. AND AL-BAYATTI, M. F. (1981). Adjustment of Examination Marks. Br. J. educ. Psychol., 51, 32- 38.

HAEUSSLER, Jr, E. F., PAUL, R. S. (1996). Introductory Mathematical Analysis For Business, Economics, And The Life and Social Sciences. Eighth Edition. Prentice Hall International Inc. Upper Saddle River, New Jersey 07458.

HASOFER, A. M. (1977). Some mathematical problems in mark scaling. Aust. Math. Soc. Gaz., 4, 55-60.

LIVERPOOL JOHN MOORES UNIVERSITY (1994). HND Computing - Student Handbook.

SCHEFFE', H. (1959). The Analysis of Variance. London: Wiley.