

Mathematical Modelling and Statistical Analysis of School-Based Student Performance Data

Jessica Y. C. Tan

*Thesis submitted for the degree of
Master of Philosophy
in
Applied Mathematics and Statistics
at
The University of Adelaide*

School of Mathematical Sciences



THE UNIVERSITY
of ADELAIDE

January 2013

Contents

	Page
List of Tables	ix
List of Figures	xiii
Abstract	xvii
Signed Statement	xix
Acknowledgements	xxi
1 Introduction	1
1.1 The Problem	1
1.2 Background and Literature Review	2
1.2.1 National Assessment Program - Literacy and Numeracy	2
1.2.2 <i>My School</i>	4
1.2.3 Value-added Measurement	6
Theory	6
1.2.4 Rasch Scaling	9
1.2.5 The Estimation of School Effects	12
1.2.6 Hierarchical and Longitudinal Modelling	13
1.3 Underlying Research Question	15
1.4 Outline	16
2 Data Analysis	17
2.1 The Basic Skills Test Data	17
2.2 The Variables: Background Information	18
2.3 Descriptive Statistics	21
2.3.1 Number of Participants in Schools	21
2.3.2 Univariate Statistics	21
Drop-off in participants in 2004	22

	Test scores	25
	Cohort Analysis	28
2.3.3	Number of Tests: Longitudinal View	43
2.3.4	Test Aspects	44
2.3.5	Missing Test Aspects	44
	Missing Literacy Writing Results	46
	Goodness of Fit - Binomial Model	47
2.4	Cleaning the Data: Forensic Statistics	49
2.4.1	Consistency of School Data	49
2.4.2	Consistency of Student Data	50
2.4.3	Score Check	50
2.4.4	Inference of Data	61
2.5	Bivariate Statistics	61
2.5.1	Categorical Variables	61
2.5.2	Quantitative Variables	71
2.5.3	Principal Component Analysis (PCA)	74
	Background Theory	74
	Results	74
2.6	Final Data Sets	77
3	Initial Model Selection	81
3.1	Manual Reduction of Data and Predictor Variables	81
	Categorical Variables	82
	Quantitative Variables	82
	Missing Data	83
3.2	Statistical Reduction of Predictor Variables	83
3.2.1	Theory	84
	Linear Regression	84
	Model Selection	84
3.2.2	Significant Variables in Linear Regression	85
	School-Number Model	86
	School-Covariates Model	90
3.2.3	Simplest Main Effects Model using stepAIC	91

	School-Number Model	94
	School-Covariates Model	99
3.2.4	Investigation of <code>procyear</code>	106
3.2.5	Comparison of School-Number and School-Covariates Models	107
	Statistical Theory	111
	Transformed Data	113
	Significant Schools	114
3.3	Discussion	117
4	Hierarchical Modelling: Mixed Effects Model	119
4.1	Theory of Linear Multilevel Mixed Effects Models	120
4.2	Hierarchical Model Formulation	122
4.3	Hierarchical Model Selection	124
4.3.1	Model Selection by Markov Chain Monte Carlo Sampling	126
4.3.2	Model Selection by Likelihood Ratio Test	128
4.3.3	Model Selection by <code>glmulti</code>	130
4.4	Analysis of Results	133
5	Bayesian Hierarchical Modelling	135
5.1	Theory	135
5.1.1	Bayesian Statistics	135
5.1.2	Markov Chain Monte Carlo Simulation and Sampling	136
	Gibbs Sampling	137
5.2	Hierarchical Modelling Using BUGS	138
5.2.1	The Hierarchical Model	138
5.2.2	The Program	139
5.2.3	Directed Acyclic Graphs	140
5.2.4	Analysis of BUGS output	141
	Validity of the Model: Diagnosis of Convergence	144
	Visualisation of Results	146
5.3	Hierarchical Modelling Using Stan	153
5.3.1	Theory	153
	Hamiltonian Monte Carlo	153
	No-U-Turn Sampler	154

5.3.2	Results	155
	Validity of the Model: Diagnosis of Convergence	155
	Interpretation of Regression Coefficients	158
6	Model Validation	159
6.1	Student-level Prediction	159
6.1.1	Prediction Intervals for <code>lmer</code>	160
6.1.2	Prediction Intervals from Stan	161
6.1.3	Comparison of <code>lmer</code> and Stan	162
6.2	Analysis of School Effect	165
6.3	Heteroscedasticity and School Size	169
6.4	Conclusion and Impact	171
7	Initial Longitudinal Analysis	173
7.1	Summary Statistics of Data from Sequential Tests	173
7.2	Grade 3 and Grade 5 Tests	174
7.2.1	Individual Scores	175
7.2.2	Difference in Scores	175
	Simple Linear Regression	178
	Hierarchical Modelling using Linear Multilevel Mixed Effects Models	182
7.3	Grade 3, 5 and 7 Tests	186
7.3.1	Longitudinal Modelling	191
8	Conclusion	195
8.1	Summary and Conclusions	195
8.2	Practical Implications and Future Work	196
A	Coding of Variables	201
A.1	Test Variables	201
A.2	School Variables	201
A.3	Student Variables	203
B	Plots	207
B.1	Boxplots of LL Rasch and NN Rasch for Categorical Variables	207

B.2	Grade 3 and Grade 5 Tests	228
C	Output	233
C.1	Chapter 3: Initial Model Selection	233
C.1.1	Full model with Main Effects	233
	School-Number Model	233
	School-Covariates Model	244
C.1.2	Simplest Main Effects Model by stepAIC	245
	School-Number Model	245
	School-Covariates Model	256
C.2	Chapter 7. Initial Longitudinal Analysis	258
C.2.1	Grade 3 and Grade 5 Tests	258
	School-Number Model	258
References	267

List of Tables

	Page
2.2.1 Test variables	18
2.2.2 School variables	19
2.2.3 Student variables	20
2.3.1 Number and percentage of participants in each calendar year . . .	22
2.3.2 Number and percentage of participants in each grade	22
2.3.3 Number of participants for each calendar year and grade	23
2.3.4 Number of participants and schools each year	24
2.3.5 Descriptive statistics for raw and Rasch scores for Literacy and Nu- meracy aspects	25
2.3.6 Coding for cohort numbers	28
2.3.7 Number of participants in each cohort	29
2.3.8 Mean scores for each cohort in each grade	42
2.3.9 Number of participants by year, grade and number of tests to date .	43
2.3.10 Number of participants with test aspects in each year and grade . .	45
2.3.11 Number and proportion of participants with missing aspects	46
2.3.12 Number of participants with missing LW by year	46
2.3.13 Number of participants with missing scores for aspects	48
2.3.14 Observed and expected frequencies for the goodness-of-fit χ^2 test of the Binomial model	48
2.4.1 Anomalies in student data	51
2.4.2 Number of schools and participants with anomalies in the sum of scores	52
2.4.3 Number of participants in 2003 and 2004 with Literacy Flag 1 . . .	52
2.4.4 Number of participants in 2001 and 2002 with Literacy Flag 3 . . .	53

2.4.5	Linear regression output from LL vs (LR+LS+LW) for Literacy Flag 1 tests	58
2.4.6	Linear regression output from LL vs LR, LS and LW individually for Literacy Flag 1 tests	58
2.5.1	ANOVA for LL Rasch against <code>p_g_nesb</code>	62
2.5.2	ANOVA for NN Rasch against <code>p_g_nesb</code>	62
2.5.3	ANOVA for NN Rasch against <code>visa_sub_c</code>	62
2.5.4	<i>P</i> -value and largest difference in group means for categorical variables against LL Rasch and NN Rasch	65
2.5.5	<i>P</i> -value and R^2 value for quantitative variables against LL Rasch and NN Rasch	73
2.5.6	The standard deviation, proportion of variance explained and the cumulative proportion for each of the principal components	76
3.2.1	Regression output for <code>gpokm</code> and school size from the school-covariates model	91
3.2.2	Regression output for <code>gpokm</code> and school size from the simplest-stepAIC school-covariates model	102
3.2.3	Comparison of models with and without <code>procyear</code> - simplest-stepAIC school-number model	102
3.2.4	Comparison of models with and without <code>procyear</code> - simplest-stepAIC school-covariates model	104
3.2.5	Comparison of models with and without <code>p_g_nesb</code> - simplest-stepAIC school-number model	104
3.2.6	Comparison of models with and without <code>p_g_nesb</code> - simplest-stepAIC school-covariates model	104
3.2.7	Type II Anova test for the full school-number model	105
3.2.8	Type II Anova test for the full school-covariates model	105
3.2.9	Summary linear regression output of <code>procyear</code> in the school-number model	106
3.2.10	Summary linear regression output of <code>procyear</code> in the school-covariates model	107
3.2.11	Subset of data to illustrate raw and fitted scores	108
3.2.12	Summary output of $\text{Raw} \sim \text{Model}$ linear regression	108

3.2.13	Summary output of linear regression of transformed data	114
3.2.14	Comparison of the number of statistically significant over- and under- performing schools	115
3.2.15	χ^2 test for association between the significance groups of schools and school covariates	117
4.3.1	Output of fixed and random effects from linear mixed effects model	125
4.3.2	Estimates and P -values of fixed and random effects estimated by Markov Chain Monte Carlo sampling	127
4.3.3	Estimates and P -values of fixed and random effects estimated by maximum likelihood	129
5.2.1	Summary output of hierarchical model without <code>procyear</code> using OpenBUGS	143
5.3.1	Summary output of hierarchical model with <code>procyear</code> as fixed effect using Stan	157
6.1.1	Counts and proportions of students in each performance category from the <code>lmer</code> model	161
6.1.2	Counts and proportions of students in each performance category from the Stan model	162
6.1.3	3×3 table of the counts of students for performance categories under <code>lmer</code> and Stan models	162
7.1.1	Counts of students who sat two appropriate sequential tests	174
7.2.1	ANOVA between Grade 3 and 5 scores and <code>schoolno</code>	177
7.2.2	Linear regression output of school-covariates model	181
7.2.3	Output of fixed and random effects from linear mixed effects model	183
7.2.4	Estimates and P -values of fixed and random effects estimated by Markov Chain Monte Carlo sampling	184
7.2.5	Estimates and P -values of fixed and random effects estimated by maximum likelihood	185
7.3.1	Counts of data for all the categorical predictors in the data of Grade 3, 5 and 7 scores	187
7.3.2	Counts of students in the categories of <code>atsi</code> and <code>school_edu</code>	193

C.1.1	Linear regression output of school-number model	233
C.1.2	Linear regression output of school-covariates model	244
C.1.3	Linear regression output of simplest-stepAIC school-number model .	245
C.1.4	Linear regression output of simplest-stepAIC school-covariates model	256
C.2.1	Linear regression output of school-number model	258

List of Figures

	Page
1.2.1 National Assessment Scale for NAPLAN	3
1.2.2 Exemplar of <i>My School</i>	5
2.3.1 Number of participants in each school.	21
2.3.2 Ratio of participants in 2004 compared to 2003 for all schools in all grades.	24
2.3.3 Histograms of Literacy raw scores over all years and grades.	26
2.3.4 Histograms of Literacy Rasch scores over all years and grades.	27
2.3.5 Boxplots of Literacy scores for the cohorts in Grade 3	30
2.3.6 Boxplots of Numeracy scores for the cohorts in Grade 3	31
2.3.7 Boxplots of Literacy scores for the cohorts in Grade 5	32
2.3.8 Boxplots of Numeracy scores for the cohorts in Grade 5	33
2.3.9 Boxplots of Literacy scores for the cohorts in Grade 7	34
2.3.10 Boxplots of Numeracy scores for the cohorts in Grade 7	35
2.3.11 Histograms of raw and Rasch LL and LW Grade 3 scores in Cohort 8	37
2.3.12 Mean aspect score in each multi-year cohort and across grades.	39
2.3.13 Cohort mean scores along grades in each Literacy aspect	40
2.3.14 Cohort mean scores along grades in each Numeracy aspect	41
2.3.15 Proportion of participants in each school missing LW.	47
2.4.1 Proportion of participants in a school which have Literacy Flag 1 in 2003	54
2.4.2 Proportion of participants in a school which have Literacy Flag 1 in 2004	55
2.4.3 Proportion of participants in a school which have Literacy Flag 3 in 2001	56

2.4.4	Proportion of participants in a school which have Literacy Flag 3 in 2002	57
2.4.5	Plot of LR+LS+LW versus LL for Literacy Flag 1 tests	59
2.4.6	Plot of LR+LS+LW versus LL for Literacy Flag 1 tests	60
2.5.1	Heatmap of the P -values from the χ^2 test for independence between the categorical explanatory variables.	63
2.5.2	Boxplot of LL Rasch against disability	67
2.5.3	Boxplot of LL Rasch against procyear	68
2.5.4	Boxplot of LL Rasch against gradedyear	68
2.5.5	Boxplots of LL Rasch and NN Rasch for gender	69
2.5.6	Boxplot of LL Rasch against procyear split into grades	70
2.5.7	Heatmap of the correlation between the quantitative explanatory variables.	72
2.5.8	Barplot of the variances explained by the principal components. . .	75
2.5.9	3D scatter plot of the scores of the first three principal components.	75
2.6.1	Flowchart of data sets.	79
3.2.1	Coefficient plot for the school-number model	88
3.2.2	Subset of coefficient plot for the school-number model	89
3.2.3	Coefficient plot for the school-covariates model	92
3.2.4	Subset of coefficient plot for the school-covariates model	93
3.2.5	Coefficient plot for the simplest-stepAIC school-number model . . .	96
3.2.6	Subset of coefficient plot for the simplest-stepAIC school-number model	97
3.2.7	Coefficients for the simplest-stepAIC school-number model plotted against the regression coefficients for the school-number model . . .	98
3.2.8	Coefficient plot for the simplest-stepAIC school-covariates model . .	100
3.2.9	Subset of coefficient plot for the simplest-stepAIC school-covariates model	101
3.2.10	Coefficients for the simplest-stepAIC school-covariates model plot- ted against the regression coefficients for the school-covariates model	103
3.2.11	Linear regression of original raw mean school effects against the fitted original model mean school effects	109

3.2.12	Diagnostic plots to assess validity of assumptions in the linear regression of $\text{Raw} \sim \text{Model}$.	110
3.2.13	Plot of residuals versus school size.	111
3.2.14	Plot of residuals versus school size for linear regression on transformed data	114
3.2.15	Prediction intervals of the original and transformed data	116
4.0.1	Hierarchy of the school education system.	120
4.3.1	AIC value for the best 100 models in the exhaustive search of the main effects model	132
4.3.2	The relative weights of model terms	132
5.2.1	Directed acyclic graph for hierarchical model in equation (5.2.2).	141
5.2.2	Histogram of parameters' posterior distribution.	146
5.2.3	Density plot of parameters' posterior distribution	147
5.2.4	Traceplots of parameters	148
5.2.5	Running means of parameters in each chain.	148
5.2.6	Autocorrelation plots of parameters.	149
5.2.7	Crosscorrelations of parameters.	150
5.2.8	Highest posterior density plot for the school and student parameters.	152
5.3.1	Comparison of NUTS with Metropolis and Gibbs sampling	155
5.3.2	Autocorrelation plot of parameters from Stan.	156
6.1.1	Plot of the Stan fitted values against the <code>lmer</code> fitted values	164
6.2.1	Ranked proportion of students who are over-performing in a school plotted with the 95% confidence interval	165
6.2.2	Ranked proportion of students who are under-performing in a school plotted with the 95% confidence interval	166
6.2.3	Normal Q-Q plot of the random effects estimates.	168
6.3.1	Plot of the variance of the residuals for each school against school size	170
6.4.1	Student scores for school 115	171
7.2.1	Individual Rasch scores for Grades 3 and 5	176
7.2.2	Difference versus Average in Grade 3 and Grade 5 scores.	177
7.2.3	Mean difference in Grade 3 and Grade 5 values in each school	178

7.3.1	Descriptives for the Grade 3, 5 and 7 data	188
7.3.2	Individual Rasch scores for Grades 3, 5 and 7	189
7.3.3	Possible trends between Grades 3, 5 and 7	190
7.3.4	Longitudinal and hierarchical structure of the education system. . .	191
B.1.1	Boxplot of LL Rasch against <code>procyear</code> split into grades.	227
B.2.1	Descriptives for the Grade 3 and 5 data - page 1	229
B.2.2	Descriptives for the Grade 3 and 5 data - page 2	230
B.2.3	Descriptives for the Grade 3 and 5 data - page 3	231

Abstract

In order to improve the education system so that students are not only reaching the minimum standards of literacy and numeracy but are also given the opportunity to excel, accurate measures of school performance are vital. These measures need to focus on student progress so that schools and teachers can focus on improving all students - particularly those most in need.

A current trend in primary and secondary education in Australia is national testing, with the introduction of the National Assessment Program - Literacy and Numeracy (NAPLAN). This, along with similar tests like the Basic Skills Tests, measures student performance and progress over sequential years. The issue now is how to interpret these results for individuals, schools and educational authorities.

The education system can be modelled with a hierarchical model due to the nesting of variables at the education system, school, classroom and student levels. Intertwined with the hierarchical nature of the model is also the longitudinal aspect of the data as students sit biennial NAPLAN assessments. We mainly investigate hierarchical modelling techniques, such as linear multilevel mixed effects and a Bayesian approach, to fit a statistical model which incorporates student and school predictor variables. In addition to hierarchical modelling methods, we also consider longitudinal techniques.

The objective of this thesis is to investigate and determine what conclusions about student progress and school performance can be reliably drawn from regular standardised system-wide assessment, such as NAPLAN. Various models are investigated and the end result is a model which comprehensively and accurately models the results of students, from which we can predict students' scores and assess the effect of schools in that way.

Declaration

I, Jessica Tan, certify that this work contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I give consent to this copy of my thesis, when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library catalogue and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

Signature: *Date:*

Acknowledgements

Firstly, I would like to thank my supervisors, Professor Nigel Bean and Dr Jono Tuke, for their invaluable supervision and support during these two years of my Masters - at the end of this journey, you are more like friends than supervisors. Both of you have given generously of your time and mathematical expertise and your thoughts are insightful and explanations clear, reflecting the depth and breadth of your knowledge. I accredit you and thank you for teaching me research skills and adding tools to my virtual toolbox. I consider it a privilege that you both agreed to supervise me.

To my family, you get a big thank you. Thank you, Mum and Dad, for allowing me to spend two years completing this Masters and for your continual support and encouragement throughout my entire education. Thank you to my brother Darren for helping me take a break from work on occasion, and to my sisters, Sarah and Hannah.

Thanks is also due to Dr Murray Thompson, a great friend and mentor from my school days, who provided a wealth of information about the education theory aspect of my project. I would also like to acknowledge Professor John Keeves for his help and Dr Darmawan for providing the data set.

My “home away from home” has been the School of Mathematical Sciences at the University of Adelaide and I would like to express my appreciation for the wonderful sense of community among the office staff, lecturers and students. A great part of my enjoyment of university has been because of all my friends and fellow postgrads - there are too many to name individually - but I would like to particularly thank Pricey, Michael, Hayden, Sophie, Kale, Ben, Mingmei, Kate M, Kate R, Kyle, Lydia

A and Max for all the fun and laughs we have had.

“Mathematics is the queen of the sciences.”

(Carl Friedrich Gauss)

*“The essence of mathematics is not to make simple things complicated,
but to make complicated things simple.”*

(Stan Gudder)

*“What lies behind us, and what lies before us are tiny matters,
compared to what lies within us.”*

(Ralph Emerson)