The use of statistics in the workplace: A survey of research graduates in diverse disciplines

John A Harraway
Department of Mathematics and Statistics
University of Otago
PO Box 56, Dunedin, New Zealand
jharraway@maths.otago.ac.nz

Richard J Barker
Department of Mathematics and Statistics
University of Otago
PO Box 56, Dunedin, New Zealand
rbarker@maths.otago.ac.nz

1. Introduction

The best approaches to the teaching of undergraduate statistics at university and the content of such programmes are sources of extensive debate. These discussions apply both to courses for the statistics professional and to service courses for graduates in many disciplines which use statistics.

The MEANS project based at the Royal Statistical Society Centre for Statistical Education in the University of Nottingham (Holmes, 1998) set out to (i) identify the statistical skills and knowledge needed by specialist and non-specialist statisticians; (ii) identify examples of good statistical training and assessment practice; and (iii) promote a closer correspondence between training in Higher Education and employment needs. But the MEANS project concentrated on specialist statisticians with much of the effort spent surveying opinions of universities and employers with less emphasis on graduates in the workplace. There was limited feedback on service courses and an internet search found little material which linked statistics teaching with the future needs of employment. Holmes (1998) suggests further studies are needed to assess the effectiveness of service teaching in statistics. Our study bridges this gap by identifying the attitudes of graduates now in the workforce to the content of statistics courses which they studied at university.

Higgins (1999) argues the case for a data specialist major to meet the demands of employers. This major would give greater prominence to nonmathematical statistics and could be an excellent supporting subject for a major in another discipline which uses statistics. We have set out to establish opinion on the quality of statistics teaching at university in light of employment demands. We have surveyed honours and research graduates, now employed, to identify statistical concepts and methodology used in the workplace and whether these topics were taught at university. Our purpose is to adapt courses in statistics to better prepare graduates for future employment.

Lo and Lam (1998) conducted a survey similar to our study but which was restricted to the teaching of quantitative methods in a Department of Management Sciences. The main objectives of their survey were to (i) record graduates' perceptions of teaching quality and learning experience while studying quantitative analysis modules; (ii) gather graduates' views on the relevance of course curriculum and syllabus content to the current needs of the business world; and (iii) find out the extent of use of quantitative techniques by their graduates in the workplace. Our survey, however, targets a wider range of subjects and is directed more at the content of statistics courses as well as the quality of the teaching and relevance of the topics with a view to altering what we teach.

2. The Survey

The target population for our survey comprised post graduate and research students in marketing, finance, economics, the biological sciences, nutrition, food science, psychology and

statistics, although in the case of statistics the undergraduate degree was included. Five of the eight New Zealand Universities agreed to provide us with the addresses of their graduates in the selected subjects. In two cases the questionnaires were mailed on our behalf by the university in order to preserve confidential information, but all responses were returned to our address in post-paid reply envelopes. Graduates from the period 1995 to 2000 were included to ensure most surveyed would be in employment and able to recall what they had been taught at university because of their relatively recent graduation. This aspect of our survey was suggested by Brennan *et.al.* (1994).

The questionnaire was developed by reviewing the Graduate Opinion and Employers' Survey of the University of Otago and choosing 46 major subject topics in statistics after discussion with statisticians and other university staff. The eight page questionnaire was then trialled on a group of 20 local graduates from the target population. Some minor changes were made as a consequence and the average response time of 15 minutes noted. This information was included in a covering letter sent out with the survey document in order to encourage response.

The questionnaire comprised six sections. Sections 1 and 2 collected details of university degrees, major subject and main field of employment. Section 3 aimed to establish the frequency of use in the workplace of 46 statistical techniques and eight computer packages and whether these techniques had been taught at university. Section 4 investigated the type of statistical activity in the workplace, the extent to which statistical help was available in the place of employment, and where and when training in statistics and mathematics had been received at university. Section 5 surveyed course experiences in relation to statistical methods taught at university while section 6 asked respondents for any comments they would like to make about the provision of short courses and workshops, the relevance of statistical training for their employment and recommendations for improving university statistics teaching.

The survey was distributed by post in April-May 2002. There were 2758 questionnaires sent to graduates in our target group and 721 completed questionnaires returned over the next two months. Non-respondents were followed up by a second mailing in November 2002 and this resulted in a further 216 completed questionnaires. There were 338 undelivered questionnaires due to wrong address. The effective response rate was therefore 38.7%. All the recommendations made by Edwards (2002) for increasing response rates to postal questionnaires were used except for survey length and incentives. An initial comparison of opinion in the initial and followup responses showed no important differences in the calculated proportions and for this reason all 937 complete questionnaires are used to produce the results in the next section. All 937 were manually checked for incompleteness and inconsistency. A sample of records was checked and no mistakes were identified.

3. Results

This paper reports on three findings from the survey: (i) gaps in the statistical training of researchers; (ii) the types of statistical activities carried out in the workplace by graduates in our target population; and (iii) desirable courses and workshops in statistics to provide appropriate training to fill the gaps.

Table 1 lists the statistical techniques most widely used in the workplace by 337 biological sciences research graduates. The table also records, for each technique, the percentage of these graduates who were not taught that technique at university. The numbers indicate that introductory methods, including descriptive statistics, basic tests, simple linear regression, the analysis of variance and multiple comparisons, are taught to most of the students but there are gaps in the teaching of modelling and multivariate techniques. In addition there should be greater exposure to survey design, power analysis, mark-recapture models, simulation and bioinformatics.

Similar analyses for all the other groups in our study indicate that the introductory methods are taught adequately. In the case of the 277 psychology research graduates there is evidence of a need for more coverage of modelling, survey methods and multivariate methods with an emphasis on exploratory and confirmatory factor analysis in particular. Meta analysis should also be taught but simulation, forecasting, mark/recapture models, data mining and bioinformatics are not widely

used by psychology graduates in the workplace. The responses for 78 economics and finance graduates show need for more coverage of multiple, non-linear and logistic regression, with, in addition, forecasting, moving averages and autoregressions. After the introductory methods, the 81 marketing graduates indicate only a need for more instruction in survey design, forecasting and multivariate techniques of principal components, factor analysis and cluster analysis. Nutrition and food sciences have few respondents to date and results are not reported.

Table 1. Statistics use in the workplace by 337 Biological Sciences Graduates

Statistical Technique	Number using Technique	Percentage not taught Technique
Descriptive statistics and graphical procedure	242	7.9
Basic tests e.g. *t*-test, estimation, chi-square	198	6.6
Simple linear regression and correlation	178	10.1
Analysis of variance, F-test	177	11.9
Contrasts and multiple comparisons	107	17.8
Survey design including stratification, clusters	93	22.6
Factorial design	80	18.8
Multiple regression	79	37.2
Repeated measures, split-plot designs	75	29.3
Multivariate analysis of variance	74	55.4
Principal component analysis	72	62.5
Blocking	70	28.6
Non-linear regression	64	53.1
Cluster Analysis	63	65.1
Power analysis	62	48.4
Mark/Recapture models	54	53.7
Logistic regression	53	52.8
Mathematical statistics e.g. advanced inference	53	30.2
Nonparametric regression	48	50.0
Random effects and mixed models	43	51.2
Multidimensional scaling and ordination	42	69.0
Simulation including bootstrapping	42	85.7
Discriminant analysis	40	75.0
Estimation theory e.g. maximum likelihood	38	71.1
Log-linear modelling and contingency tables	37	59.5
Canonical correlation analysis	35	82.9
Correspondence analysis	32	68.8
Randomisation testing	31	67.7
Cross over designs	22	27.3
Bioinformatics	20	85.0

The 118 statistics respondents include students who have passed a first degree with a major in statistics or a research degree in statistics. Most of our 46 listed statistical techniques are covered adequately but not non-linear and logistic regression, survival analysis, loglinear modelling, Bayesian statistics, simulation, power analysis and data mining.

A classification of statistical procedures and activities carried out in the workplace by the 937 respondents is presented in Table 2. The most common activity for 663 (71%) of respondents involves the carrying out of data analyses. Next in importance is understanding analyses in published research (542 or 58%) while only 10% of respondents have no need of statistics. There was strong evidence of an association between type of activity and area of study (c^2_{48} = 491.5, p < 0.001). A correspondence analysis with two dimensions identifies an association between financial analyses and the economics/finance graduates and an association between market research and the marketing graduates. The biological scientists and psychologists are involved with designing studies, understanding published research and carrying out their own analyses while the statistics graduates appear to cover the full range of activities.

Table 2. Statistical activities in the workplace

Activity	Statistics	Econ/ Finance	Marketing	Bio Science	Food Science	Nutrition	Psych	Total
Designing Studies	24	21	27	181	6	16	141	416
Reading published research	38	39	34	209	6	21	195	542
Own analyses	79	60	46	259	18	15	186	663
Understanding a consultant	13	7	6	79	2	12	31	150
Promotion and reports	49	34	45	158	11	13	115	425
Market research	19	10	53	12	4	1	22	121
Financial analysis	38	47	26	27	1	2	28	169
Quality control	13	4	7	29	11	5	14	83
None needed	18	4	5	42	0	1	28	98

A further analysis shows no assistance is available in the workplace for 30% of all graduates in the survey, 40% have access to an internal consultant and 34% have handbooks and manuals available. Overall, 48% (449) of the respondents say they would benefit from the provision of short courses in statistics tailored to their employment needs.

In the final part of the survey, where respondents were invited to identify courses which would be helpful for meeting their employment needs, 396 nominated topics with 103 requesting a course on multivariate methods, 92 courses on advanced regression and generalised linear models, 87 a course on population estimation and survey design, 76 a course on data management and statistical packages, and 39 a course on stochastic processes and time series.

4. Conclusions

The survey results reported here (i) show there is inadequate statistics training for research graduates in the biological sciences in particular; and (ii) identify, for graduates in all the selected areas, several courses which could be provided at specialist workshops or summer schools.

5. Acknowledgements

This work was supported by a Research into University Teaching Grant from the University of Otago and a Grant from the New Zealand Statistics Association.

REFERENCES

Brennan, J., Lyon, E.S., Schomburg, H. and Teichler, U. (1994). The experiences and views of graduates messages from recent surveys. *Higher Education Management*, 6(3), 275-304.

Edwards, P., Roberts, I., Clarke, M., DiGuiseppi, C., Pratap, S., Wentz, R. and Kwan, I. (2002). Increasing response rates to postal questionnaires: systematic review. *British Medical Journal*, 324, 1-9.

Higgins, J.J. (1999). Nonmathematical statistics: a new direction for the undergraduate discipline. *The American Statistician*, 53(1), 1-6.

Lo, H.P. and Lam, J.K.P. (1998). Employees' perceptions on teaching, learning and use of quantitative methods: a survey report. *Proceedings of Fifth International Conference on Teaching Statistics*, Singapore, 1, 319-325.

Holmes, P. (1998) Matching education, assessment and employment needs in statistics. *Report on MEANS project.* http://science.ntu.ac.uk/rsscse/means/means.html

RÉSUMÉ

Les recherches pour établir le recours à la statistique dans le lieu de travail s'accomplissent par une enquête auprè des chercheurs diplômés dans cinq universités néo-zélandaises entre les ans 1995 et 2000. Des diplômés dans la statistique et plusieurs autres matières qui comprennent l'application de la statistique forment la population visée. On signale des lacunes dans l'enseignement universitaire de la statistique avec le but de modifier le matériel du cours universitaire pour mieux préparer des chercheurs et des experts-conseil en statistique pour la vie active.