STATISTICAL EDUCATION AT SOUTH AFRICAN UNIVERSITIES WITH SPECIAL REFERENCE TO THE NEEDS OF CONSULTING STATISTICIANS.

PIERRE SPAFFORD DU TOIT

A Dissertation submitted to the Faculty of Education, University of the Witwatersrand, in fulfilment of the requirements for the Degree of Master of Education.

Johannesburg, 1987

Declaration

I, Pierre Spafford du Toit, hereby declare that this thesis is my own work and that it has not been presented to any other University for the purpose of obtaining a Degree.

P S du Toit

P S DU TOIT

November, 1987

ABSTRACT

This study concerns statistical education at university level in South Africa.

An analysis of the data from two questionnaires identified the statistical methods commonly used by groups of statisticians employed in different sectors of the economy, and showed that there were significant differences between the usages by some of the employee groups.

Estimates of the teaching emphasis given to these methods by universities established that certain methods get greater emphasis than their use warrants, while others get less.

Comments from sixty five statisticians on the effectiveness of their training indicated that the teaching of statistics is too theoretical.

Continuing education courses, as a means of upgrading the statistical abilities of users of statistics, are discussed, and the experience gained in this field at the University of the Witwatersrand is presented.

Suggestions are made concerning the use of educational theory pertaining to curriculum development to improve the effectiveness of statistical education.

The value of this study is that it is the first detailed survey of the opinions and needs of consulting statisticians in the work place in South Africa. It brings to the attention of statistical teachers useful educational theory which has possibly been neglected in the past.

ACKNOWLEDGEMENTS

I wish to acknowledge my grateful thanks to my supervisor, Professor D.D. Russell of the University of the Witwatersrand, for the invaluable advice and assistance he gave me, a novice in the theory of education.

I would like to thank Professor J.M. Juritz of the University of Cape Town, for allowing me to join her team investigating the statistical needs of industry, and for letting me use the data from the survey we carried out. Similarly my thanks go to Professors L.G. Underhill and A.G.W. Steyn for giving me permission to use the results of their surveys.

This thesis would probably never have been completed without the enthusiastic encouragement and excellent secretarial service received from Mrs. J. Landey and the editorial help given so willingly by Mrs. M. Manuel.

Finally I would like to express my sincere appreciation to my wife, Gay, for her patient support over my many years of post-graduate study.

TABLE OF CONTENTS

		Page No.
Chapter 1.	INTRODUCTION AND LITERATURE REVIEW	1.
1.1.	Historical background	1.
1.2.	The importance of this study	4.
1.3.	Aims of the study	7.
1.4.	Frame of reference	8.
1.4.1.	Definition of a statistician	9.
1.4.2.	Private sector	10.
1.4.3.	State and semi-state bodies	11.
1.4.4.	Universities	12.
1.5.	The purpose of university education	12.
1.6.	The curriculum	14.
1.6.1.	Curriculum development	14.
1.6.2.	The aims and objectives of statistical training	15.
1.6.3.	The selection of learning experiences	23.
1.6.4.	The selection of course content	27.
1.6.5.	Evaluation of the curriculum in terms of efficiency and effectiveness	30.
1.7.	Reports on requirements of statistical training	36.
1.7.1.	Report of the International Statistical Institute	37.
1.7.2.	Report of the American Statistical Association: "Preparing statisticians for careers in industry"	38.
1.7.3.	ASA report: "Preparing statisticians for careers in the federal government"	40.
1.7.4.	Training requirements suggested by other authors	40.
1.7.5.	Discussion	41.
1.8.	Comparison of the organisational functions of universities and business corporations	42.
1.9.	The critical components in industrial statistics	43.
Chapter 2.	RESEARCH DESIGN	46.
2.1.	Overall design	46.
2.2.	Questionnaire No.1	51.
2.3.	Pilot survey for Questionnaire No.1	58.
2.4.	Non-response	60.
2.5.	Questionnaire No.2	63.
2.6.	Qualitative data	66.

Page No.

Chapter 3. STATISTICAL EDUCATION AT SOUTHERN AFRICAN UNIVERSITIES 70.

3.1. Departments of Statistics at Southern African universities 70.

3.2. Staff numbers 70.

3.3. Preponderance of staff in the professional ranks 72.

3.4. Statistical methods taught to major students 73.

3.5. Service courses in Statistics 75.

3.6. Summary 77.

Chapter 4. SURVEY OF STATISTICAL METHODS USED IN COMMERCE 78.

4.1. Aims of the survey 78.

4.2. Questionnaire No.1 80.

4.3. The sample 81.

4.4. Profile of the respondents 81.

4.5. Findings from replies to survey 85.
- 4.5.1. Reasons for firms not using statistical methods 85.
- 4.5.2. Use of Statistics, Operations research and Quality control 86.
- 4.5.3. Firms using statistical methods 87.
- 4.5.4. Computer packages used by business 89.
- 4.5.5. Statistical methods used by business 91.

Chapter 5. SURVEY OF THE STATISTICAL METHODS USED BY MEMBERS OF THE S.A. STATISTICAL ASSOCIATION AND THE OPERATIONS RESEARCH SOCIETY OF S.A. 96.

5.1. Questionnaire No.2 96.

5.2. Statistical methods used by SASA and OR Society 96.

Chapter 6. COMPARISONS BETWEEN THE VARIOUS EMPLOYEE GROUPS 103.

6.1. Comparison of academic qualifications 103.
- 6.1.1. Source of information 103.
- 6.1.2. Method of statistical analysis 105.
- 6.1.3. Comparison of qualification profiles for different employee groups 106.
- 6.1.4. Conclusions 110.

6.2. Frequencies of usage of statistical methods 110.

6.3. Method of statistical analysis 113.

Page No.

6.4. Comparison of statistical method usage for different pairs of employee groups 114.
- 6.4.1. Stock Exchange versus UCT Business School (from Questionnaire No.1) 114.
- 6.4.2. Subgroups of the S.A. Statistical Association (from Questionnaire No.2) 115.
- 6.4.3. Operations Research Society versus the subgroups of the S.A. Statistical Association (from Questionnaire No.2) 117.
- 6.4.4. Business (i.e. Stock Exchange + UCT Business School from Questionnaire No.1) versus the subgroups of the S.A. Statistical Association (from Questionnaire No.2) 121.
- 6.4.5. Business (i.e. Stock Exchange + UCT Business School from Questionnaire No.1) versus the Operations Research Society (from Questionnaire No.2) 125.

6.5. Overall conclusions on usages of statistical methods 125.

Chapter 7. CORRESPONDENCE BETWEEN UNIVERSITY TEACHING AND THE USAGES OF STATISTICAL METHODS BY VARIOUS EMPLOYEE GROUPS 127.

7.1. Questionnaire on university teaching 127.

7.2. Analysis of ranks for routine usage 131.

7.3. Analysis of ranks for routine and occasional use 133.

7.4. Discussion 136.

Chapter 8. OPINIONS AND SUGGESTIONS FROM PRACTISING STATISTICIANS 140.

8.1. Introduction 140.

8.2. Training considered adequate 141.

8.3. Statistics courses supply a basic foundation 142.

8.4. Training considered too theoretical 145.

8.5. Calls for more practical training with suggestions for achieving this 147.
- 8.5.1. Improved balance between theory and practice 147.
- 8.5.2. Genuine data sets 153.
- 8.5.3. Practical work with outside organisations 155.
- 8.5.4. Practical projects 156.
- 8.5.5. The problems of statistical practice 161.
- 8.5.6. Exposure to statistical consulting. 163.

8.6. A wide variety of statistical techniques is needed 166.

8.7. Computer program packages 169.

Page No.

8.8. Continuing education 170.

8.9. Communication skills 170.

8.10. Other suggestions from the literature 172.
- 8.10.1. Model syllabus 172.
- 8.10.2. Co-operation between universities and industry 173.
- 8.10.3. Postgraduate work 174.
- 8.10.4. Teaching resources and journals 175.

Chapter 9. CONTINUING EDUCATION COURSES FOR BUSINESS AND INDUSTRY 176.

9.1. Course at University of the Witwatersrand 176.
- 9.1.1. Description of courses 176.
- 9.1.2. Course evaluation 181.

9.2. In-house courses 185.
- 9.2.1. Course at AECI Limited: November, 1985 185.
- 9.2.2. Course at Adcock-Ingram Laboratories: July, 1985 187.
- 9.2.3. Course at Middelburg Steel and Alloy: July, 1986 188.
- 9.2.4. Course at Anglo American Gold Division Services at Welkom: December, 1986 188.

9.3. Summary of findings 192.

Chapter 10. SUMMARY OF FINDINGS AND DISCUSSION 194.

10.1. Statistical methods "in common use" 194.
- 10.1.1. Methods "in common use" in all five employee groups 194.
- 10.1.2. Methods "in common use" in four employee groups 197.
- 10.1.3. Methods "in common use" in three employee groups 197.
- 10.1.4. Methods not "in common use" 198.

10.2. Differences between employee groups 198.
- 10.2.1. Difference in qualification profiles 198.
- 10.2.2. Differences in usage of statistical methods 200.

10.3. Teaching emphasis in universities 203.

10.4. Comments by statisticians on adequacy of university training 204.

10.5. Effectiveness of continuing education in statistics 207.

10.6. Answers to the research questions 207.

Chapter 11. CONCLUSIONS AND RECOMMENDATIONS 208.

11.1. Conclusions 208.
- 11.1.1. The aims of this study 208.
- 11.1.2. Usage of statistical methods 209.
- 11.1.3. Differences in usage of statistical methods 210.

		Page No.
11.1.4.	The effectiveness of statistical training at SA universities	211.
11.1.5.	The adequacy of statistical curricula	212.
11.1.6.	The need for continuing education in statistics	213.
11.2.	Suggestions for improving statistics curricula	214.
11.2.1.	Aims and objectives of a curriculum	215.
11.2.2.	Choice of course content	217.
11.2.3.	Choice of learning experience	219.
11.2.4.	Course evaluation	222.
11.3.	Overall conclusions	223.
BIBLIOGRAPHY		225.
APPENDIX A	: Questionnaire No 1 sent to Stock Exchange and UCT Business School samples	230.
APPENDIX B1	: Questionnaire No 2 (Version sent to the OR Society)	233.
APPENDIX B2	: Questionnaire No 2 (Version sent to the S.A. Statistical Association	234.
APPENDIX C	: Questionnaire to Universities on teaching of service courses	235.
APPENDIX D1	: Evaluation questionnaire for statistics courses at University of the Witwatersrand	238.
APPENDIX D2	: Evaluation questionnaire for statistics courses at AECI	240.
APPENDIX E	: List of comments from Questionnaire No 2	241.
APPENDIX F	: Syllabus suggested by Education Committee of S.A. Statistical Association	254.

TABLE		Page No.
1.1	Broad syllabus for applied statisticians (Underhill, 1984)	28.
1.2	Some suggested topics within each specialisation of Category 3 (Underhill, 1984)	29.
1.3	Evaluation of a teaching project (Juritz, 1986)	33.
3.1	Statistical staff at Southern African universities	71.
3.2	Number of universities teaching each method in a given year	74.
3.3	Extent of service courses at 18 universities in 1983 together with the student enrolment	76.
3.4	Statistics courses offered by departments other than statistics departments	76.
4.1	Activity of the Company (some reported more than one activity)	82.
4.2	Status of company (one of the sixty five respondents did not report)	83.
4.3	Number of employees per company (one of the sixty five respondents did not report)	84.
4.4	Reasons for not using statistical methods	85.
4.5	The use of Statistics, Operations research, and Quality control	86.
4.6	Effectiveness of statistical techniques	87.
4.7	Numbers of users of statistics (Four companies gave no information)	87.
4.8	Qualification of users of statistics	88.
4.9	Computer packages used by respondents	90.
4.10	Methods used Routinely in two samples : Frequency and percent rounded to nearest 10%	94.
4.11	Methods used overall (ie Routinely + Occasionally) in two samples : Frequency and percent rounded to nearest 10%	95.
5.1	Methods used Routinely : Percentages of employees using each method	97.
5.2	Methods used Overall (ie Routinely + Occasionally) : Percentages of employees using each method	98.
5.3	Methods used Routinely by 20% or more of respondents with percentage usage rounded to the nearest 10%	99.

TABLE		Page No.
5.4	Methods used Routinely and Occasionally by 30% or more of respondents with percentage usage rounded to the nearest 10%	101.
6.1	Academic qualifications extracted from SASA Directory (as percentages with actual frequencies in brackets)	104.
6.2	Academic qualifications from Questionnaire No 2 (as percentages with actual frequencies in brackets)	104.
6.3	Example of a contingency table from qualification profiles (actual frequencies in each cell)	105.
6.4	Ratios of Bachelors : Honours : Masters : Doctors (expressed as percentage of total frequency of all 4 degrees)	106.
6.5	Ratios of Bachelors : Honours : Masters : Doctors (expressed as percentage of total frequency of all 4 degrees)	108.
6.6	Methods used Routinely : Frequency of employees using each method	111.
6.7	Methods used Overall (ie Routinely + Occasionally) : Frequency of employees using each method	112.
6.8	Methods for which usage differs significantly between OR Society and University	118.
6.9	Methods for which usages differ significantly between OR Society and State and semi-state bodies	119.
6.10	Methods for which usages differ significantly between the Business group and University	121.
6.11	Methods for which usages differ significantly between the Business group and State and semi-state bodies	122.
6.12	Methods for which the usages differ significantly between the Business group and Private sector	123.
7.1	Weighted totals of frequencies of universities teaching a given method in each year, in ranked order	129.
7.2	Routine usage - Ranks for teaching emphasis and method usage for each statistical method	131.
7.3	Values of Spearman's correlation coefficient (Routine usage)	132.
7.4	Routine and Occasional usage : Ranks for teaching emphasis and method usage for each statistical method	134.
7.5	Values of Spearman's correlation coefficient (Routine and Occasional usage)	135.

TABLE | Page No.

5.4 Methods used Routinely and Occasionally by 30% or more of respondents with percentage usage rounded to the nearest 10% 101.

6.1 Academic qualifications extracted from SASA Directory (as percentages with actual frequencies in brackets) 104.

6.2 Academic qualifications from Questionnaire No 2 (as percentages with actual frequencies in brackets) 104.

6.3 Example of a contingency table from qualification profiles (actual frequencies in each cell) 105.

6.4 Ratios of Bachelors : Honours : Masters : Doctors (expressed as percentage of total frequency of all 4 degrees) 106.

6.5 Ratios of Bachelors : Honours : Masters : Doctors (expressed as percentage of total frequency of all 4 degrees) 108.

6.6 Methods used Routinely : Frequency of employees using each method 111.

6.7 Methods used Overall (ie Routinely + Occasionally) : Frequency of employees using each method 112.

6.8 Methods for which usage differs significantly between OR Society and University 118.

6.9 Methods for which usages differ significantly between OR Society and State and semi-state bodies 119.

6.10 Methods for which usages differ significantly between the Business group and University 121.

6.11 Methods for which usages differ significantly between the Business group and State and semi-state bodies 122.

6.12 Methods for which the usages differ significantly between the Business group and Private sector 123.

7.1 Weighted totals of frequencies of universities teaching a given method in each year, in ranked order 129.

7.2 Routine usage - Ranks for teaching emphasis and method usage for each statistical method 131.

7.3 Values of Spearman's correlation coefficient (Routine usage) 132.

7.4 Routine and Occasional usage : Ranks for teaching emphasis and method usage for each statistical method 134.

7.5 Values of Spearman's correlation coefficient (Routine and Occasional usage) 135.

TABLE Page No.

7.6 Statistical methods with noteworthy differences in rank (> or = 16) 137.

7.7 Statistical methods for which the usage does not match the teaching emphasis 138.

10.1 Example of calculation of index of usage (for Basic statistics) 195.

10.2 Statistical methods "in common use" in all five employee groups (in order of merit in terms of the index of usage) 196.

10.3 Statistical methods "in common use" in only four employee groups (in order of merit) 197.

10.4 Statistical methods "in common use" in only three employment groups (in order of merit) 198.

10.5 Methods grouped under the user category which used them most frequently. (Frequency of significance of methods in brackets) 202.

FIGURE		Page No.
1.1	The sources of educational objectives (UTMU, 1976:115)	17.
1.2	Traditional teaching (Steyn, 1984)	24.
1.3	What is a performance? (Romney et al, 1978)	32.
1.4	Comparison of the organizational functions of Universities and Corporations (Snee et al, 1980)	43.
1.5	The critical components in industrial statistical training (Snee et al, 1980)	44.
2.1	The research design diagram	47.

APPENDICES

		Page No.
A	Questionnaire No 1 sent to Stock Exchange and UCT Business School samples.	230.
B1	Questionnaire No 2 (Version sent to the OR Society)	233.
B2	Questionnaire No 2 (Version sent to the S.A. Statistical Association	234.
C	Questionnaire to Universities on teaching of service courses	235.
D1	Evaluation questionnaire for statistics course at University of the Witwatersrand	238.
D2	Evaluation questionnaire for statistics course at AECI	240.
E	List of comments from Questionnaire No 2	241.
F	Syllabus suggested by Education Committee of S.A. Statistical Association	254.

CHAPTER 1. INTRODUCTION AND LITERATURE REVIEW

1.1. Historical background

In the medical profession there is a need for the general practitioner as well as for the specialist. Similarly in the field of statistics one finds applied practitioners and pure mathematical statisticians. The former work on real-life problems in commerce and industry while the latter immerse themselves in the pursuit of mathematical research. Both are necessary, for without the contact with the real world the theoretician would be developing ever expanding bodies of purposeless abstractions. Without the theoretical backup, the applied statistician would lack adequate tools with which to work.

As the demands of modern living require industry and commerce to become more efficient, so is there an increasing need for the problem-solving techniques of statistics. The technicians in research and development, the marketing teams in business, the planners plotting the future courses of their companies and the production crews in the factories no longer achieve their goals on their own. Sound decisions must be based on reliable information which comes from the processing of data. These days data tends to be accumulated in ever increasing masses in data banks and on storage disks and to be disgorged onto endless sheets of digital printout. Usually the information contained in the data can be distilled and extracted only with the sophisticated methods developed by the theoretical specialists. These methods are interpreted and used by applied statisticians, who act as the vital interface between the mathematical purists and the decision makers in industry and commerce.

A healthy balance is required between the specialists and the general practitioners and University syllabuses should be designed to produce both types of statistician. However there seems to be a world wide realisation that all is not well. Could it be that the teaching of statistics is being monopolized by the specialists and directed solely to their needs? Do the most brilliant students get enticed to stay on at University to become the teachers of the next generation, often without experiencing the needs of the worlds below the ivory towers? Are the students of average ability ignored and left to flounder with a feeling of failure. Perhaps many wander off to be accepted with open arms by the computer industry.

Research is being undertaken overseas into the adequacy of university statistical training for preparing statisticians for careers in industry and in government. In South Africa it has been recognised that a detailed investigation is required into the status of statistics and the adequacy of the training of statisticians. To this end the S.A. Statistical Association (SASA) formed an Education Committee in 1982 with the following objectives:

(a) To cultivate contacts between the statistics departments at the universities and interested organisations outside the universities.
(b) To encourage contact between universities.

(c) To press for the inclusion of statistics in school syllabuses.

(d) To propagate closer contact between the SASA and practising statisticians.

(e) To promote the statistical profession as well as the SASA.

A sub-committee on the training of university students, under the chairmanship of Professor J.M. Juritz was given the following tasks:

(a) To evaluate existing university syllabuses in the light of the needs of employers and, if necessary, to recommend changes. Also to encourage discussion about syllabuses.

(b) To propagate interaction between university lecturers and personnel in outside organisations.

(c) To investigate the feasibility of compulsory vacation work for statistical students.

A second sub-committee on practical statistics, under the chairmanship of Professor F.E. Steffens was asked to prepare a data book of case studies of typical South African problem situations. They were also to arrange seminars and workshops.

A third sub-committee under Professor D.J. Stoker was to negotiate the introduction of statistics at school level. A fourth sub-committee under Mr. T.C. Gilfillan was given the responsibility of publicity. In particular they were to organise an annual competition for student projects.

The purpose of this study is to ascertain what statistical methods are used by statisticians in South Africa, whether South African universities are providing adequate training for the personnel who use these methods and to make recommendations for improving the training. This study therefore falls naturally under the terms of reference of Professor Juritz's sub-committee. The author was never formally a member of her sub-committee, but co-operated with her and two of her colleagues in a sample survey of statisticians in commerce and industry (See Questionnaire No 1 in Section 2.2). The author, with Professor Juritz's permission, conducted a second survey of members of SASA and of the Operations Research Society of SA (See Questionnaire No 2 in Section 2.5). This dissertation is essentially an analysis of the results of these two questionnaires, together with a chapter on continuing education. The conclusions and recommendations from this study will be reported to the Education Committee of SASA and to the OR Society.

1.2. The importance of this study

In 1981 the President of the S A Statistical Society, Professor F Lombard, entitled his presidential address "The neglect of applied statistics in South Africa" (Lombard, 1981). This illuminating discourse is so pertinent to this study that the following extracts and paraphrases will be used to highlight the purpose and scope of this study.

The challenge faced by the pioneers in the statistical field in SA was to show that statistics was a respectable academic discipline, that it had a broad theoretical foundation and that the problems confronted by it were sufiiciently different from those of pure mathematics to warrant the establishment of a separate discipline of statistics. So successful were these efforts that today there is hardly a university in SA which does not have an autonomous department of statistics. The successful establishment of theoretical statistics did have one consequence which was probably not foreseen. Undergraduate teaching of statistics became geared to the needs of those few students who showed the potential for proceeding to graduate work at the M.Sc. and Ph.D. level and 'statistics' became 'mathematical statistics' with the emphasis mainly on 'mathematical'. Only a small fraction of the undergraduate curriculum was concerned with realistic practical or applied work, and the needs of the greater majority of students who would, upon completion of a Bachelor's or Honours degree, be absorbed into commerce or industry, were largely ignored. The teaching of statistical theory became an end in itself, with little or no regard for what was in reality the ultimate goal, namely the training of people who would be sufficiently skilled in the analysis of data to be able to meet the demands of commerce, industry and government.

Fellingham and Kotze (1981) confirmed Lombard's concern when they stated that universities in S.A., which should be producing the "general practitioner" type of applied statistician, have become the major consumers of statistical manpower, and that their courses are designed more to equip their top students for university posts than to fit the average students (who form the majority) for applied statistical consulting.

Lombard (1981) made the following suggestions for improving statistical education:

> Firstly, we must recognise that there is a gulf separating the practitioners and the theoretician and that the Statistical Association has a duty, in the interest of the profession as a whole, to see that it is narrowed. Secondly, all those involved in teaching should see to it that undergraduate statistics curricula place more emphasis on the applied aspects of our subject, in order that those students who go into commerce and industry after attaining a Bachelor's degree are better prepared for the tasks awaiting them. At some of our universities attention has already been given to this matter. and this is indeed a hopeful sign for the future. Thirdly, we must obtain a clear picture regarding the present status of statistics and of statisticians in the private sector. Finally, we should actively seek contact with practising statisticians and their employers, learning about the gaps they perceive in the training of statisticians and of the problems of communication between the statistician and his employer.

The third suggestion of Lombard is really basic to the others, for without a clear picture of the status of statistics in the private sector in South Africa, it has not really been established that there is a gulf that needs bridging. It could well be that there are no gaps in the training of practising statisticians that need to be filled. The questionnaires on which this study is based were designed to find out whether there are indeed fundamental differences between theoretical and applied statisticians and whether their training should be differentiated to suit their peculiar needs. This is set out more fully in the next section.

1.3. Aims of the study

The aims of this study are to ascertain what statistical methods are used by statisticians employed in different sectors of the economy in South Africa, to evaluate whether South African Universities are providing adequate training for the statisticians who use these methods, and, to recommend possible improvements to statistical education at university level.

The aims can be achieved by answering the following research questions, each of which requires a separate research effort.

Research Question 1 What statistical methods are used by statisticians employed by universities, by state and semi-state organisations and by industrial and commercial firms in the private sector?

Research Question 2 Does the usage of these statistical methods differ between the different groups of statisticians mentioned in Research Question 1?

Research Question 3 What teaching emphasis is being given in university statistics courses to the methods used by the statisticians mentioned in Research Question 1?

Research Question 4 Do the statistical curricula offered by the universities provide an adequate training for consulting statisticians?

Research Question 5 Is there a need for continuing education in statistics, and, if so, what form should this take?

Research Question 6 What improvements to present statistical curricula can be recommended for consideration by university teachers?

1.4. Frame of reference

In discussing the aims of this study in the previous section, mention is made of statisticians and their areas of employment, namely the private sector, state and semi-state bodies and universities. These terms are defined below.

1.4.1. Definition of a statistician

A statistician is normally defined as a university graduate who has majored in statistics. Operation Research practitioners are not usually thought of as statisticians, but they will usually have had some statistics courses in their curricula and they use many of the methods like forecasting, time series, simulation, inventory control, network models and others that are commonly considered to be statistical methods or applications of statistical techniques. Therefore for the purposes of this study an OR practitioner will be defined as a statistician.

The definition makes no mention of personnel with statistical training from Technikons or elsewhere. It is possible that such people could have responded to Questionnaire No.1 of this study (See Section 2.2), which was sent to business firms, but this is unlikely to have caused any serious bias to the results.

There are approximately 430 members of the S.A. Statistical Association, approximately 410 members of the Operations Research Society and an unknown number of statisticians who are not members of the Association and OR Society.

A distinction is sometimes made between a theoretical statistician and a practising statistician. For example Duncan & Durban (1980) imply that the former is to be found in universities and the latter in state and private sector employ. In South Africa this distinction is very blurred.

because many university statisticians do consulting work, either privately or as part of their job, while many statisticians in state and semi-state employ are encouraged to do research work of a purely theoretical nature. This is a very healthy development, which emphasises the balance required in the training of statisticians: the theoretical statisticians require training in the arts of consulting, and the practising statisticians require adequate theory to give them the tools they need for effective consulting.

The term "practising" is not a good one in that statisticians doing theoretical research are also practising their profession. In this dissertation it is preferred to use the term consulting statistician for both practising statisticians and theoretical statisticions who are temporarily engaged in consulting. For the purposes of curriculum planning, consideration has to be given to both the theoretical and the consulting aspects of statistical education, without stipulating what the subsequent major-time occupation of the statistician will be.

1.4.2. Private sector

Information on the private sector originated from two questionnaires. The sampling frames used for the two questionnaires therefore define that part of the private sector on which this study was based.

Questionnaire No. 1 was sent to 75 companies listed on the Johannesburg Stock Exchange, of which 41 responded. A further 9 unlisted companies were sent the questionnaire of which only Escom, Old Mutual and United Building Society replied.

Questionnaire No. 2 was sent to members of the S.A. Statistical Association and the Operations Research Society. Of the SASA members there were 15 respondents who worked for firms in the private sector. Most of the 19 respondents from the OR Society worked in the private sector.

A complete list of the companies sampled and those that responded is given in Sections 2.2. and 2.5. The validity of recommendations based on information from these companies depends on the degree to which they are representative of the private sector in South Africa. This is discussed in Section 2.4.

1.4.3. <u>State and semi-state bodies</u>

Of the respondents to Questionnaire No. 2 there were 25 members of SASA who worked for State or semi-state institutions. A list of these institutions is given in Section 2.5.

1.4.4. Universities

Of the respondents to Questionnaire No. 2 there were 25 members of SASA who lecture at university. A list of the universities to which they are attached is given in Section 2.5. Of the 21 possible universities, there were replies from 14. There is unlikely to be any serious bias in drawing conclusions from this sample of university staff.

1.5. The purpose of university education

This study revolves around the need for improving university statistical education. It is therefore appropriate that we consider what is the purpose of university education in general and then proceed to discuss curricula.

Education is defined by Wheeler (1967:11) as the deliberate attempt to bring about changes in the behavioural patterns (ie. knowledge, skills, habits, sensitivities, attitudes and values) of people by presenting to them certain learning experiences. He states that "education is an important social institution built up around society's interest in the inculturation or socialization of the young. Indeed it is a major process in the maintenance of a society". Education can be divided into two broad categories, namely general education which is concerned with maintaining the culture of the society, and vocational education which is directed towards the interests of a special class or group.

One of the best known statements of the aims of general education comes from Newman (1947:157) in his "Idea of a University". It begins as follows:

> A university training aims at raising the intellectual tone of society, at cultivating the public mind, at purifying the national taste, at supplying true principles to popular enthusiasm and fixed aims to popular aspiration, at giving enlargement and sobriety to the ideas of the age, at facilitating the exercise of popular power, and refining the intercourse of private life.

Attempts to satisfy such lofty ideals were probably more common in preceding centuries when education was tho preserve of the privileged elite. In these modern times an education aimed only at enlarging the mind would impose obvious limitations on the career prospects of the less affluent.

Beard (1972:12) does not favour the opposite extreme where teachers concentrate on imparting to their students "a body of knowledge", and "see no need to provoke them into thinking about the subject". Knowledge is expanding so rapidly that she recommends a greater emphasis on the ability to independently obtain information and apply it to new situations. What is required is "ingenuity in dealing with the unexpected and flexibility in facing new circumstances".

1.6. The curriculum

As stated in the previous section, this study is concerned with improving university education and hence, ipso facto, involves curricula. In particular, Research Question No 4 of Section 1.3 asks

> Do the statistical curricula offered by the universities provide an adequate training for consulting statisticians.

Now the word curriculum is often used very loosely by the layman, but in terms of educational theory it has a very specific meaning. Much has been written in the educational literature concerning curriculum development. In this section we will look closely at this theory,

1.6.1. Curriculum development

Heywood (1984:11) states that "a curriculum is more than a syllabus or a collection of subjects". According to Wheeler (1967:11) "it is with the deliberate, systematic planned attempts to change behaviour that curriculum is concerned. By 'curriculum' is meant the planned experiences offered to the learner under the guidance of the teaching institution".

Wheeler (1967:30) sets out the process of curriculum development in a way that has stood the test of time. It consists of the following phases:

1. The selection of aims and objectives.
2. The selection of learning experiences calculated to help in the attainment of these aims and objectives.

3. The selection of content (subject matter) through which certain types of experience may be offered.
4. The organization and integration of learning experiences and content with respect to the teaching/learning process.
5. Evaluation of the effectiveness of all aspects of phases 2, 3 and 4 in attaining the aims and objectives of phase 1.

These phases are related and interdependent and combine to form a cyclical process whereby phase 5 affects phase 1 and restarts the next cycle. These phases will be discussed in greater detail in Section 1.6.2. to 1.6.5.

1.6.2. <u>The aims and objectives of statistical training</u>

(i) <u>The distinction between aims and objectives</u>

Gerrans (1986) distinguishes between aims and objectives as follows:

> The former may be seen as the general global targets or goals which are, hopefully, to be attained, and the latter, often called behavioural objectives, as the more precise outcomes of particular courses of action with respect to student achievement.

Beard a.. Hartley (1984:24-25) differentiate between aims and goals. "Aims are broad general statements. Goals are more specific statements. Goals explain how aims are to be achieved"

The University of London Teaching Methods Unit (UTMU) publication (1976:112) has a useful chapter on educational objectives. A clear distinction is drawn between objectives and aims as follows:

> Educational objectives are more or less precise statements of what a student should be able to do as a result of completing (part of) a course. Normally they describe what the students should be able to do at the end of their course that they would not do (at least not as well) at the beginning. The emphasis thus is on changes in behaviour which can be attributed to the students' educational experiences. This behavioural emphasis distinguishes objectives from course aims, with which they are often confused. Aims are usually taken to be very broad statements of what a teacher intends to do and hopes his course will achieve; objectives are statements of what a student should be able to do.

The UTMU publication (1976:114) then goes on to argue that there are three primary sources of objectives for a course:

(a) The subject matter and the interest and perceptions of the staff who teach it.

(b) The requirements of the profession for which the student is being prepared.

(c) The desires and abilities of the students.

These three sources accord with three views of education:

(a) As part of a process of passing our culture from one generation to the next.

(b) As a functional device for producing qualified manpower.

(c) As a social service providing educational opportunities for citizens.

All courses will have objectives from all three sources and it is only when objectives from different sources conflict that a decision has to be made as to which view of education prevails. The successful attainment of these objectives is influenced by a number of "facilitating" factors such as library facilities, teaching methods and methods of evaluating achievement. Their diagram is presented below as Figure 1.1.

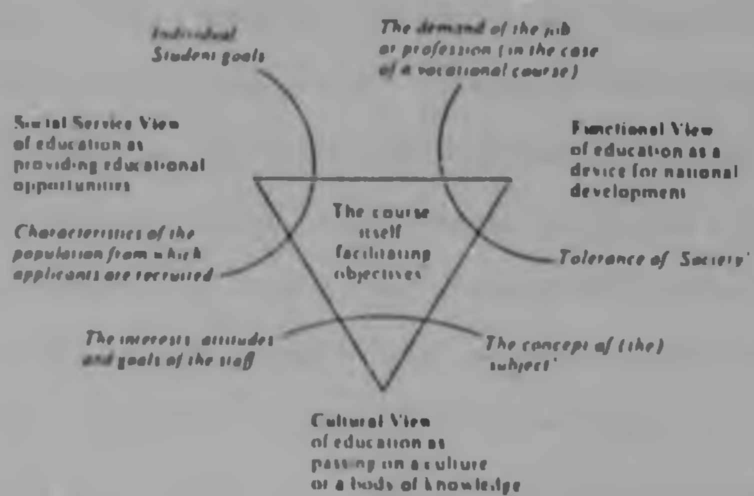

FIGURE 1.1. <u>The sources of educational objectives.</u>

<u>(UTMU,1976:115)</u>

The top right hand influence results in the development of skills, the bottom influence is the acquisition of knowledge, while the top left hand influence represents the growth of attitudes. The degree to which any one of the three sources of objectives pushes into the triangle has enormous influence on educational purposes and it is unlikely that uniformity can ever be expected from one group of people to another.

A good example of the interplay of the three sources of educational objectives shown in Figure 1.1. is supplied by a typical first year statistics course. Statistics represents a body of knowledge that from the cultural view must be passed on to each new generation. It contains many useful skills such as the estimation of statistical parameters like the standard deviation and correlation coefficient, which, from a functional view, are tools for decision-making in research and business. Its value from the social service view can be seen in changes in thinking patterns and attitudes. One example of this is the realisation that measurements vary one from another, so that quoted figures are never exact, but merely estimates of some unknown true value. Another example of a change in thinking patterns is the improvement in logic and problem-solving ability which results from being taught statistical probability. The educational benefits derived from these exercises far exceed their immediate value to the statistics profession as such.

(ii) <u>Suggested aims and objectives for the BSc in Statistics</u>

The Faculty of Science of the University of the Witwatersrand has proposed aims and objectives for the BSc which are based on those of Beard and Hartley (1984:36). These have been adapted here specifically for the Statistics major.

The Faculty of Science proposes the following as the primary aim for the BSc:

The education of students to become scientists.

A negative objective was that it should not be the aim of the Faculty to train students to become technicians.

Beard's objectives can be rewritten as follows for the Statistics degree.

(a) To know the basic facts and theories of Statistics and, to a lesser extent, of cognate disciplines;

(b) To possess intellectual skills such as the ability to think independently and creatively, to communicate effectively, to understand, to be able to solve problems, to enquire, and to be capable of self-education;

(c) To possess the necessary practical skills of a Statistician.

In order to achieve "a balance between knowledge and skills", the following subsidiary objectives are proposed:

Knowledge

At the end of the BSc degree students should:

Know the basic terminology of Statistics.

Know the principles (basic laws and concepts of Statistics).

Understand some of the uses to which Statistics is put.

Skills

University teaching in general should enable the students:

To write coherently.

To be verbally articulate.

To make their own independent judgements.

To obtain information efficiently.

To think creatively, imaginatively and in abstract terms.

To cooperate with colleagues and other professionals in their future careers.

To develop adaptability, i.e. to cope with changing patterns of knowledge (both general technological advances and new ideas in Statistics).

To develop powers of observation.

To become familiar with specialised computer hardware and software.

Attitudes

The teaching of Statistics at university level should foster in the students:

Enthusiasm for learning.

Questioning and enquiring approaches to knowledge.

Scholarly concern for accuracy.

Awareness of moral, social, economic, political and scientific problems of society.

(iii) Suggested aims and objectives for a specific learning exercise

There are innumerable aims in statistical training, but one example is given below for two reasons: firstly, it serves as a useful illustration of aims and objectives as defined in the UTMU publication, and, secondly, it highlights one of the most important aims in the statistical training of statisticians who will be consulted by technical and research personnel.

The example has been adapted from Juritz (1986), who identified eight stages in the conduct of an observational study, i.e. a study in which phenomena are observed, but not derived from a designed experiment.

Aim: Enable the graduate statistician to be able to conduct an effective statistical consultation with a client who proposes to undertake an observational study.

Objective 1: Be able to establish beforehand, in conjunction with the client, the motivation for the study and to formulate the problem.

Objective 2: Be able to recommend a sampling scheme.

(This will involve defining the population, deciding on the method of sampling, the size of the sample and estimating the cost of the survey, which may have to be tailored to fit a predetermined budget).

Objective 3: Be able to advise on the collection of information.

(This will involve the choice of the variables to be measured, the design of the questionnaire or other means of collecting the data, and consideration of whether or not the variables measured reflect the underlying phenomena).

Objective 4: Be able to advise on or carry out the computerisation of the data in a form suitable for statistical analysis.

(This may involve entering the data into a data-file on a mainframe computer for later analysis by a statistical package, or entering the data into a program especially written for the particular job on a micro-computer, or some other procedure for capturing the data).

Objective 5: Be able to advise on or carry out a check of the computerized data.

(This will involve the identification of outliers and the correction of mistakes).

Objective 6: Be able to advise on or carry out a statistical analysis of the data.

(This may involve a choice of the most efficient method from various possible alternatives, or a decision on whether the quality of the data or the purpose of the study justifies a complex (and possibly expensive) analysis rather than a simpler treatment of the data).

Objective 7: Be able to interpret the results to the client.

Objective 8: Be able to write a report on the analysis.

(In carrying out objectives 7 and 8 it is important to avoid terminology peculiar to statistics and to use a style of language appropriate to the recipients of the report).

1.6.3. The selection of learning experiences

(i) Lecturing: the traditional learning experience

Lecturing as a method of teaching is usually the commonest learning experience to which university students are exposed. Beard and Hartley (1984:153) note that lecturing

> has so frequently been attacked by educational psychologists and by students that some justification is needed to retain it. Critics believe that it results in passive methods of learning which tend to be less effective than those which fully engage the learner.

But although students

> fairly often comment on poor lecturing technique, they praise lectures which are clear, orderly synopses in which basic principles are emphasized, but they dislike too numerous digressions or lectures which consist in part of the contents of a textbook..

They feel that important advantages of lecturing are its efficiency in reaching large audiences, its economy measured in teacher time per student, and its flexibility in adapting quickly to changes in knowledge or teaching conditions.

(ii) <u>A deficiency in traditional statistical teaching</u>

Steyn (1984) points out that traditional statistical teaching consists of the four-part cycle shown in Figure 1.2., although it is not unknown for the process to stop at Stage 1.

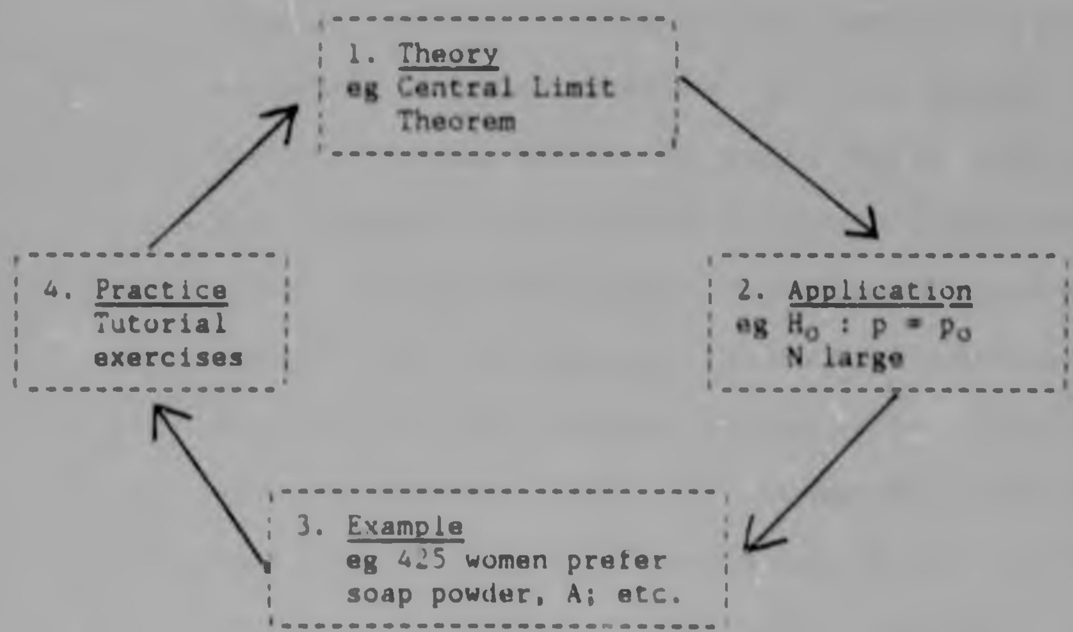

FIGURE 1.2. Traditional teaching (Steyn, 1984)

After applying the teaching process set out in Figure 1.2. to numerous topics the lecturer may feel that the student is equipped to stand alone in the real world of applied statistics. However, Steyn points out that there is a serious deficiency in this process of learning, namely, that it concentrates only on the analysis aspect of the applied research process and deals with the statistical tools in isolation.

(iii) Teach less, learn more

The learning experiances to which students are exposed must be based on the aims and objectives of the courses making up the degree. If these objectives are to be pursued, some changes in the traditional form of University teaching are going to be necessary. Gerrans (1986) states:

> It is probably the case that most change which has occurred in university teaching has taken place within the all-pervading "lecture-format", which has remained largely intact. Thus it has been an essantially egocentric process based on what the teacher does and what the teacher chooses to teach. That such an approach has achieved only a modicum of success is not surprising since, in the final analysis, education is concerned with learning not with teaching. It does not matter in itself how or what the teacher chooses to teach, but how and what the students learn. The overemphasis on the role of the teacher to the exclusion of the student is particularly marked in the often impersonal milieu of university education.

From the foregoing it would appear that one strategy for improving effectiveness in education is to transfer the present emphasis from "how and what is taught" to "how and what is learnt". Given the realities of the university situation it would be ridiculous to propose abandoning lectures completely. Clearly the answer lies somewhere between this extreme and the present situation. But where?

Aarons (1983), in an article entitled "Achieving Wide Scientific Literacy", comments succinctly on the problem:

> The notion that the understanding of science can be achieved by purely verbal inculcation seems to me to be a principal source of failure. Experience makes it increasingly clear that verbal presentations - lecturing to large groups of intellectually passive students and having them read text material - leave virtually nothing in the student's mind that is permanent and significant. Much less do they help the student attain what I consider the marks of a scientific literate person.
>
> What is the alternative? I think it is essential to back off, to slow up, cover less, and give students a chance to follow and absorb the development of a small numbe of major scientific ideas, at a volume and pace that make their knowledge operative rather than declarative.

This proposal is often equated with the "lowering of standards" but it is undeniable, though not often accepted, that such strategies would be setting far higher intellectual standards and demanding far higher performance than we now attain, with all our vaunted "coverage". This would not be a "watering-down"; rather it would be a "raising-up". I know at first hand that students do respond to such a demand.

1.6.4. The selection of course content

Course content must be carefully planned to satisfy one or more of the three educational views discussed in Section 1.6.2. A considered decision must be taken on the appropriate content to either develop skills, or to pass on knowledge or to foster the growth of attitudes, or some combination of all three objectives.

Rao (1982) has suggested general topics, some of fundamental importance and some which relate to particular broad areas of application. Underhill (1984) has adapted Rao's suggestions as shown in Table 1.1.

TABLE 1.1. Broad syllabus for applied statisticians (Underhill, 1984)

Category 1 Basic Theory	Category 2 Basic Applications	Category 3 Specialisations
Probability theory Linear models and analysis of variance Categorical data and Chi-square tests Statistical inference (a) hypothesis testing (b) estimation (c) decision theory	Design of Experiments Survey sampling Simulation Exploratory data analysis (including detection of recording errors, outliers and faked data, modelling and graphical display methods like biplots and correspondence analysis)	Econometrics Technometrics Biometrics Statistical research

Underhill (1984) then adapts Table 1.2. from Rao (1982) to show desirable statistical methods needed in each specialisation.

TABLE 1.2. Some suggested topics within each specialisation of Category 3 (Underhill, 1984)

Technometrics (for statisticians in industry)	
Quality Control	Life testing, reliability
Acceptance sampling	Response surfaces
Operations research	Stochastic processes
Econometrics (for statisticians in business and government)	
Econometric models	Portfolio theory
Time series	Economic and social indicators
Biometrics (for statisticians in medical and biological sciences)	
Bioassay	Applied multivariate analysis
Genetics	Epidemiology
Life tables	
Statistical research	
Advanced topics in probability theory, inference, multivariate analysis and the above topics	

It is extremely unlikely that any South African university with the standard staff-to-student ratio could offer four alternative specialisations. However, the idea of having two general directions for research and for applied statisticians is implicit in these ideas of Rao.

1.6.5. Evaluation of the curriculum in terms of efficiency and effectiveness

(i) Evaluation methods

The decision of whether or not the training was successful will be based on somo measurement of change in behaviour. Certain aims and objectives have been set and corresponding behaviours expected. Did these intended behaviours materialise? Did the student acquire the "knowledge, skills, habits, sensitivities, attitudes and values" mentioned by Wheeler (1979:11)? Were the learning experiences suitable for achieving the aims and objectives or could others have been more effective?

Much has been written on evaluation methods for educational programmes. Berk (1981:4) defines an evaluation as "the process of applying 'scientific procedures' to collect 'reliable and valid information' to make 'decisions' about an 'education programme'. Fink and Kosecoff (1978:1) differentiate between two kinds of evaluations:

> One is to improve a program and the other is to determine the effectiveness of a program. Improvement and effectiveness evaluations are always distinguished from one another by how information is used rather than by the kinds of information collected or the stage at which it is gathered. In an improvement context, evaluation information is used to modify and improve a program; in an effectiveness context, information is used to establish the program's quality and outcomes.

Other authors such as Morris and Fitz-Gibbon (1978:9) term these two types (viz improvement and effectiveness evaluations) formative and summative evaluations, respectively.

Patton (1981:187) categorises various types of evaluation including:

Effectiveness evaluation. To what extent is the program effective in attaining its goals.

Efficiency evaluation. Can inputs be reduced and still obtain the same level of output or can greater output be obtained with no increase in inputs.

According to Piper (1978:8) "both effectiveness and efficiency imply some notion of quality" but as pointed out by Goldschmid (1978:53) "the two are interdependent: instruction which is effective, but very costly, is inefficient; inexpensive methods which are ineffective are also inefficient". He suggests that "what we should look for are cost-effective approaches, ie educational practices which at less cost provide the same or better instructional quality, or else strategies which at the same cost are more effective".

How can effectiveness be defined? Romney et al (1978) in defining the performance of a programme give indirect definitions to its four constituents, viz commitment, utilization efficiency and effectiveness, illustrated in Figure 1.3.

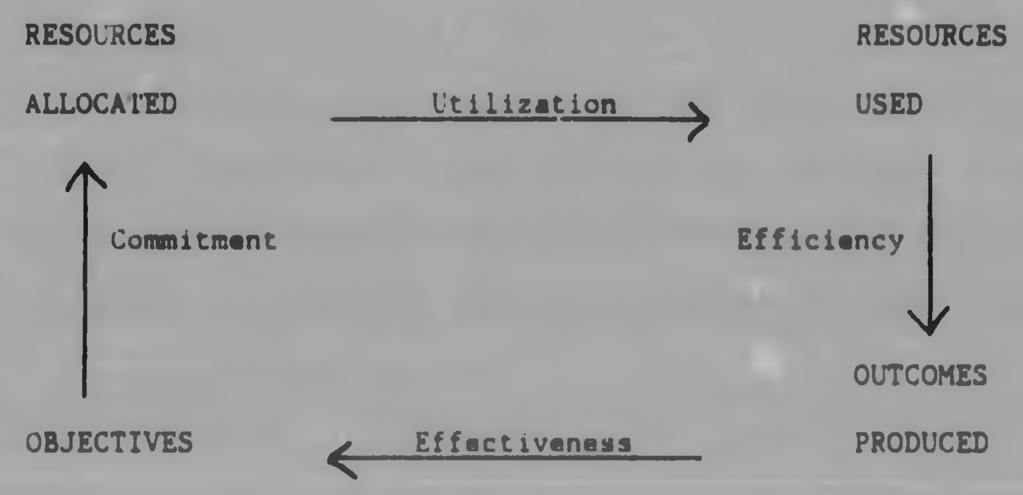

FIGURE 1.3. What is a performance? (Romney et al, 1978)

Commitment is the match between the objectives and the resources allocated to various activities. Utilization is the relationship between the resources allocated and the resources actually used. Efficiency is the relationship between the resources used and the outcomes produced. Effectiveness is the match between the outcomes produced and the objectives which led to those outcomes.

(ii) Examples of actual evaluations at SA universities

A number of examples are quoted below which show how helpful questionnaires can be in improving teaching.

Example 1

The survey teaching project at UCT discussed in Sections 1.6.2. and 8.5.1. was followed by an eight question evaluation. Juritz (1986) reports the results of forty seven replies as shown in Table 1.3., giving the median class as ( ) and the modal class as [ ]:

TABLE 1.3. Evaluation of a teaching project (Juritz, 1986)

I found the practicals	Disagree			Agree		Mean
1. Interesting.	1	2	3	[(4)]	5	3.5
2. Relevant.	1	2	[(3)]	4	5	3.4
3. Useful.	1	2	[(3)]	4	5	3.1
4. Helped me understand theory.	1	(2)	[3]	4	5	2.3
5. Made me feel I could analyse a questionnaire.	1	2	[(3)]	4	5	3.3
6. Difficulties with computer system killed my interest.	[1]	2	(3)	4	5	2.6
7. Computing took too much time.	[1]	(2)	3	4	5	2.0
8. The BMDP language is difficult.	1	[(2)]	3	4	5	2.0

Example 2

Steyn (1984) reports that an evaluation questionnaire was used following the project work at the University of Pretoria. Of the students, 80% regarded the project work as extremely worthwhile (uiters sinvol) and interesting. There were no negative comments and typical remarks (translated from Afrikaans) were as follows:

It was fun (baie pret/lekker).

Very good experience. Greatly increased my insight. Opened up a new world for me. A lot of work, but worth the effort. Why was it not prescribed in the first year?

Example 3

In 1986 the Academic Staff Development Centre of the University of the Witwatersrand conducted a course evaluation of the Mathematics and Statistics a service course for Commerce students. Much of value was learned from this evaluation which enabled the lecturing team to effect meaningful improvements for the current year.

The questionnaire consisted of twenty four multiple choice questions with the following five-scale alternative answers: Strongly disagree/Disagree/Neither agree nor disagree/Agree/Strongly agree. The questions were as follows:

1. Tutorial assignments are relevant to what is presented in class.
2. Tutorials are interesting and stimulating.
3. Tutorials make students think.
4. Tutorials form a valuable part of the course.
5. The course was extensive in its coverage of subject matter.
6. The time allocated to this course was reasonable.
7. The course is well structured with clearly defined objectives
8. The textbook makes a valuable contribution to the course.
9. The textbook is easy to read and understand.
10. Lectures and tutorials were well integrated.
11. The department sets sufficient assignments for me to assess my progress.
12. The tutorial assignments helped me assess my understanding of each topic.
13. Effective use was made of worksheets.
14. Effective use was made of notes and other handouts.
15. Effective use was made of examples and exercises during the course.
16. I learned a good deal of factual material in this course.
17. I gained a good understanding of concepts and principles in the field.
18. Audience participation was encouraged in class.
19. Alternative explanations of difficult points were made in class.
20. The pace of the lectures was adapted to the needs of the class.

21. Interesting digressions from the course were made during the class.

22. Good discipline was maintained during lectures.

23. The Drop-In tutorials should be replaced by a better tutorial system.

24. The solutions to the tutorial assignments aided my understanding of each topic.

These multiple choice questions were followed by three open-ended questions, as follows:

(a) Which aspect of the course did you like BEST?

(b) Which aspect of the course did you like LEAST?

(c) Please make any other comments you may have about the course in the space provided below.

1.7. Evidence of the need to improve statistical education

Having defined the term curriculum and having discussed in some detail the techniques of curriculum development, it is appropriate to consider how they can be used to improve statistical education. The first obvious question is whether or not this education needs improvement. The answer to this, in the South African context, will be decided by an analysis of the results of the two questionnaires used in this study. However, there is considerable information available in the literature concerning the need for improvement in statistical education overseas.

Three reports on university statistical training have appeared recently, namely that of the International Statistical Institute (ISI) on the Integration of Statistics (Duncan & Durbin, 1980), and the Reports of the American Statistical Association (ASA) on Preparing Statisticians for Careers in Industry (Snee et al, 1980) and on Preparing Statisticians for Careers in Federal Government (Eldridge et al, 1982).

1.7.1. Report of the International Statistical Institute

A committee of the ISI was constituted under the joint chairmanship of J.W. Duncan and J. Durbin (called the Committee in what follows). A summary of the Committee's report has appeared in the SASA Newsletter (Steyn, 1981) and some of the points highlighted by Steyn are as follows:

(i) Statistical education at university level was found to be too theoretical in nature. It failed to equip students for employment as practising statisticians by the State, industry and other organisations.

(ii) With reference to statistical literature the Committee noted with great concern that it was completely dominated by theoretical articles of a mathematical nature. The fact that this literature was usually incomprehensible to the practising statistician led to a further widening of the gap between theoretical and practising statisticians.

(iii) With reference to membership of Statistical Associations, the Committee noted that, with few exceptions, members in most countries were largely theoretical statisticians, who were also often associated with universities. This situation was a consequence of the poor communication between theoretical and practising statisticians.

(iv) With reference to the attitude of theoretical statisticians, the Committee noted that they tended to look down on their practising colleagues who "never got further than an experiment".

1.7.2. <u>Report of the American Statistical Association: "Preparing statisticians for careers in industry"</u>

This report by Snee et al (1980) focuses particularly on the training of statisticians working in the physical, biological, chemical and engineering sciences (also referred to as industrial statisticians). The report was circulated to heads of statistics groups in a wide variety of companies for review and comment. Sixteen replies were received from organisations representing the motor, chemical, communications, food, petroleum, pharmaceutical and tyre industries. All agreed that the report accurately reflected the broad needs of industry.

The report discusses the profile of an effective industrial statistician and summarises the desirable qualities as follows:

(i) Is well trained in the theory and practice of statistics.

(ii) Is an effective problem solver.

(iii) Has good oral and written communication skills.

(iv) Can work within the constraints of the real world.

(v) Knows how to use computers to solve problems.

(vi) Is familiar with the statistical literature.

(vii) Understands the realities of statistical practice.

(viii) Has a pleasing personality and is able to work with others.

(ix) Gets highly involved in the solution of company problems.

(x) Is able to extend and develop statistical methodology.

(xi) Can adapt quickly to new problems and challenges.

(xii) Produces high-quality work in a timely fashion.

With regard to statistical training they state that:

The course work should take into account that the industrial statistician is a general practitioner and should contain a balance between theory and practice. Most statisticians in industry are necessarily generalists because they are presented with problems from a wide variety of subject matter areas and must employ many different statistical tools in solving them. It is impossible for all statisticians to be exposed to all useful statistical techniques during their training. What is needed, however, is for this aspect of statistical practice to be recognized and for the educational process to be designed accordingly. Briefly stated, the student should be exposed to:

(i) Statistical theory and methodology useful in the solution of a wide range of real problems, and

(ii) A philosophy to use as a guide in applying these methodologies and in developing others as needed in solving problems.

This approach not only gets the students off to a good start but also provides them with a proper framework for future growth and development.

1.7.3. ASA report: "Preparing statisticians for careers in the federal government"

This report (Eldridge et al, 1982) recommends that federal government needs statisticians with a solid understanding of data systems concepts and applied statistical methods. While there were isolated needs for pure theorists, statistical applications remain the primary tool of the government statistician. For this reason tools such as computer programming, applied sampling theory, data system design, incomplete data procedures, technical writing and technical consulting are of primary importance.

1.7.4. Training requirements suggested by other authors

Daniel (1969) suggests that the common characteristic of industrial statisticians in many diverse fields is their strong interest in working on practical problems and the satisfaction they get from helping others. According to Marquardt (1979) a good consulting statistician will become totally involved in a project. The natural result is an invitation to be co-author of the resulting research report.

Hunter (1981) describes Wisconsin's Masters degree examination which was designed to help students become practising statisticians. He identifies three consulting roles, namely, helper, leader and colleague, and suggests statisticians have a primary responsibility in providing guidance about the scientific method itself.

Snee et al (1980) quote Ott as wanting sufficient problem background so that the student can see how statistical methods are often modified in practice to develop an adequate but not necessarily optimum solution.

Daniel (1969) recommends that those interested in statistics as a profession should obtain solid foundations in science or engineering and mathematics. Knowledge of basic science is valuable in communicating with scientists. (Author's note: This is certainly also true in other fields like the social sciences, medicine, education, etc.)

1.7.5. Discussion

There is no question but that the Statistical Association in America and Great Britain are worried about the health of the profession. Statistical education is considered to be too theoretical and so is the literature. The qualities of a good statistician are defined and course work to achieve these qualities is suggested.

It is implicit in the discussion that university education is not achieving the desired objectives. The two questionnaires used in the present study will be used to confirm whether the position in South Africa is similar to that overseas. But undoubtedly the techniques for curriculum development discussed in the previous section will have a use in improving statistical education.

1.8. Comparison of the organisational functions of universities and business corporations

It is important to realise that in any educational programme, the teaching aspect cannot stand on its own. There must be an interplay and cross-stimulation between research, consulting and teaching. There is possibly a legitimate criticism of university staff in South Africa, that they are not, in general, in touch with the needs of industry and commerce. This familiarity can only come with consulting experience.

The ASA report on training statisticians for industry (Snee et al, 1980) has a diagram comparing the interactions between functions. (See Figure 1.4.) Just as in a business organisation Research, Development and Manufacturing determine the success of the Marketing performance, so in a university programme is the end product, Consulting, determined by the quality of Teaching and the Theoretical and Applied Research undertaken in the department.

Similarly, just as Marketing must provide feed-back to the other three business departments, so also must the consulting function report back and influence the research and teaching programmes.

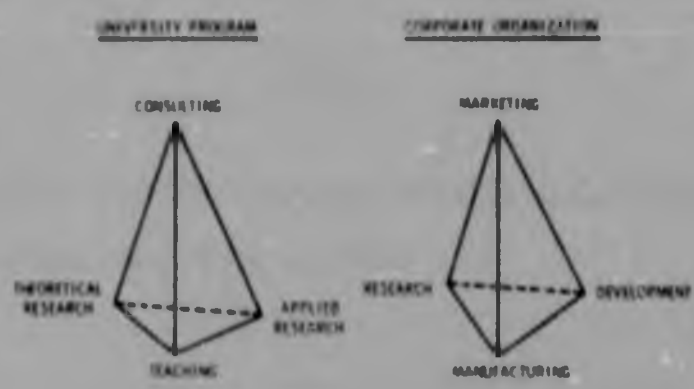

FIGURE 1.4. Comparison of the organizational functions of Universities and Corporations. (Snee et al, 1980)

1.9. The critical components in industrial statistics

A diagram highlighting the critical components in industrial statistics is presented in the ASA report (Snee et al, 1980). This is shown in Figure 1.5.

FIGURE 1.5. The critical components in industrial statistical training. (Snee et al, 1980)

The authors do not expand on these critical components, but they can clearly be related to the three views of education discussed in Section 1.6.2. The course itself will be a mixture of the functional aspect (competence with computers and skill in consulting will be needed by the industrial statistician), the cultural aspect (the body of statistical knowledge must be passed down to the next generation largely through course work) and the social service aspect (the graduates should develop problem solving abilities as well as communication skills). The balance of these three aspects will depend on the faculty philosophy. There will be inputs from both the teachers, dependent on faculty experience, individual teaching experience and research ability, and from the students, depending on their personalities.

As was stressed in the last section, teaching cannot stand in isolation. There are many influences and inputs which together combine to produce a finished product of the required standard. It will be the aim of this study to investigate to what extent the South African situation differs from that reported in the international literature. The manner of going about this investigation will be discussed in the next chapter.

CHAPTER 2. RESEARCH DESIGN

2.1. Overall design

This study arose as a consequence of the unease felt by the statistical profession both here in South Africa, as well as overseas, that all is not well with statistical education. For some years the International Statistical Institute, the American Statistical Association and the South African Statistical Association have been active in investigations into statistical training. Some information was therefore available before this study was initiated. In particular, Snee et al (1980) had decided on a list of important statistical methods in common use among statisticians in the U.S.A..

The research design for this study was not formulated at the start as a integrated whole. The surveys and data used here arose from a number of sources each responding to different investigational needs, mainly originating from the work of the Education Committee of the SASA. The overall design can be summarized in Figure 2.1.

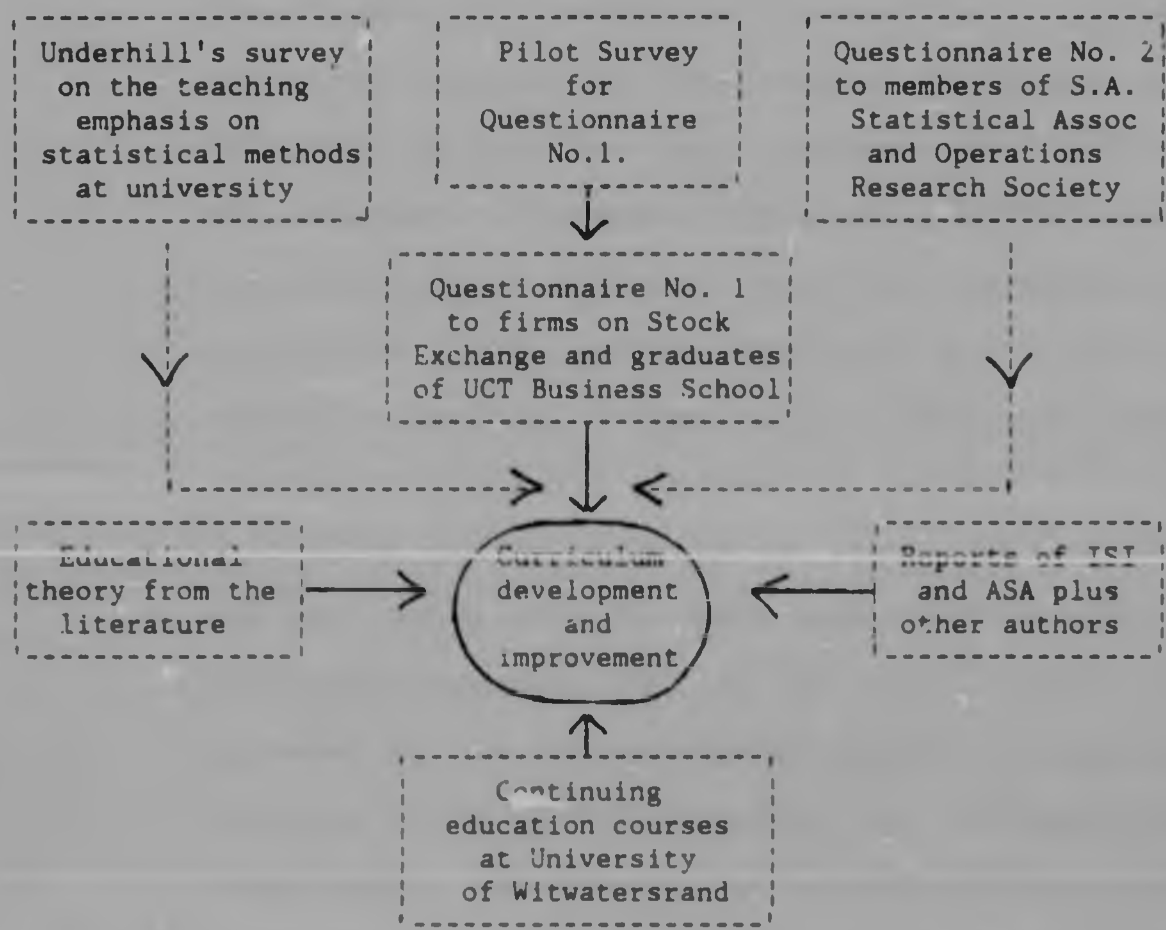

FIGURE 2.1. <u>The research design diagram</u>

(a) <u>Underhill's survey</u>

Underhill (1982) used the list of statistical methods from Snee et al (1980) to carry out a survey of thirteen major South African universities to establish which of the methods was being taught in what year of study. He has given permission for his data to be used in this study. The results of his survey are presented in Section 3.4.

(b) Questionnaire No.1

Questionnaire No.1, which is discussed fully in Section 2.2., was designed by three members of staff at the University of Cape Town, namely Professor J.M. Juritz of the Department of Mathematical Statistics, and Professors A.H. Money and J. Affleck-Graves of the Graduate School of Business. It consisted basically of the list of statistical methods of Snee et al (1980) with some amendments. Additional questions were included to gather information on the users of the methods, the manner of use and the computer facilities available to the users. In November/December, 1983 it was sent to a sample of graduates of the UCT Business School and to a sample of companies quoted on the Johannesburg Stock Exchange. The former sample was prepared and administered by the UCT academics, and the latter sample by the author at the University of the Witwatersrand. The data from the survey were processed by computer at UCT and a scientific paper was prepared for publication in the S.A. Statistical Journal by J.M. Juritz with A.H. Money, J. Affleck-Graves and P.S. du Toit as co-authors. The UCT authors have given permission for the data to be used in this study. The results are discussed in Chapter 4.

A copy of Questionnaire No.1 is attached as Appendix A.

(c) <u>Pilot survey for Questionnaire No.1</u>

A pilot survey, discussed in detail in Section 2.3., was carried out by the author to test Questionnaire No.1. Eleven copies of the draft questionnaire were posted to people in business. As a result of the pilot survey certain suggestions were made to amend the questionnaire.

(d) <u>Questionnaire No.2</u>

Questionnaire No.1 was designed to establish the usage patterns of statistical methods in industry and commerce. Questionnaire No.2, discussed in detail in Section 2.5., was an abbreviated form of Questionnaire No. 1 designed and administered by the author. It was sent to all members of the S.A. Statistical Association and Operations Research Society. The respondents could be grouped into a number of employee categories, namely, statisticians employed by universities, by State and semi-state bodies and by the private sector. This enabled comparisons to be made between the method usages of the various employee groups. It was also possible to compare the teaching emphasis, given by the universities to the various statistical methods, with the usages of the methods by the different employee groups.

There are two slightly different forms of Questionnaire No.2, which are attached as Appendices B1 and B2.

(e) <u>Reports of the International Statistical Institute and the American Statistical Association</u>

The requirements for training statisticians in the UK and USA have been highlighted in these reports. They were discussed in Section 1.7. Their recommendations will be used in Chapter 8, where they will be matched and compared with recommendations originating from South African statisticians.

(f) <u>Education theory</u>

Teaching and learning are extremely complex processes and a vast literature on educational theory has built up over the years. Traditionally university teachers receive no training in education prior to appointment and there is little incentive to acquire any later. Yet, as Gerrans (1986) points out, "to be effective any teaching strategy should have a sound theoretical basis." He argues that

> educational theory is to a teacher what mathematics is to a physicist or chemistry to a biologist. Yet in university teaching the findings of educational research are either never considered or even actively resisted.

A review of the literature of education theory must be part of the design plan. A case will be made in Chapter 11 for using the results of educational research to improve statistical teaching at university level.

(g) Continuing education courses

Continuing education, as a method for upgrading the knowledge of users of statistical methods, is of interest in this study. The contact with people from commerce and industry is very useful. A number of part-time courses in statistics have been given at the University of the Witwatersrand as well as some in-house courses at various companies. The results of the evaluation questionnaires following some of these courses are presented and recommendations for this type of course are made in Chapter 9 and Section 11.1.6.

2.2. Questionnaire No.1

In this and the following sections the two questionnaires and the pilot survey are discussed in detail. Consideration is also given to the question of non-response and to the nature of qualitative data, since these two aspects of the study are vital to the validity of any conclusions that are drawn.

The survey was conducted in two parts. Firstly, a random sample of 160 names was selected from the register of past students of the Graduate School of Business of the University of Cape Town. This part of the survey was undertaken by the three staff members of the University of Cape Town mentioned in Section 2.1. (b).

Secondly, the author undertook a survey of companies listed on the Johannesburg Stock Exchange. The sample was non-random with stratification based on the stock exchange classification, after exclusion of Mining Holdings, Property, Property Trusts, Kruger Rands and Cash Assets. From the remaining 500 companies seventy five (15%) were selected according to the following criteria:

(i) The share price should be greater than 1000c and the dividend yield greater than 10%.

(ii) Where (i) could not be met, the share should be among the leaders in its category.

(iii) When a choice had to be made among companies satisfying (i) and (ii), companies with well-known household names were preferred.

(iv) All independent Mining Houses were chosen.

(v) A selection of the larger mines were chosen from a cross section of Mining Houses.

It was decided not to choose the companies within the strata at random because the major purpose of the survey was not to draw inferences about the population of all Stock Exchange firms, but rather to gain access to statisticians in industry in a representative cross-section of business. The effect of the above criteria was to select the larger and more progressive firms, which were likely to be using statistical methods and employing statisticians or consulting them. There was no point in including small or struggling businesses in the sample, as they would be unlikely to give any information on the use of statistics in industry. The sampling method ensured representativeness by the stratification. There was an indirect random choice of statisticians rather than a deliberate random choice of firms.

The following seventy five companies were chosen for the sample:

Mining:

Coal: Amcoal (Anglo), Wit Colls (Barlow Rand)

Diamonds: Anamint (Anglo), De Beers (Anglo)

Gold: Rand Randfontein (JCI)

Klerksdorp Buffels (Gen Mining), Harties (Anglovaal)

West Wits Kloof (GFSA)

Metals and Minerals:

Copper: Palamin

Platinum: Implats (Un. Corp)

Financial:

Mining Houses: Anglo American, AngloVaal, Fed Mynbou, Gencor, GFSA, Johnnies, TC Lands.

Banks & Fin. Serv: Barclay, Nedbank, Stanbic, Volkskas.

Insurance: Guardian, SA Eagle, Mutual and Federal.

Investment Trust: Fugit.

Industrial:

Ind. Holdings: Barlow, C G Smith, Fed Volks, M&R, Picbel, Plate Glass, Rennies.

Beverages & Hotels: S.A. Breweries, Southern Sun.

Building & Constr.: Grinaker, Masonite, Toncoro.

Chemicals & Oil: AECI, Sasol, Sentrachem, Triomf.

Cloth, Foot, Textiles: Natal Con, Romatex, Seardel.

Electronics: Aberdare, Altech.

Engineering: Afrox, Dorbyl, Haggie, Std. Bradd, Stew & Lloyds.

Fishing: SWA Fish.

Food: Kanhym, Prem. Gp., Tiger Oats.

Furn & Household: Ellerine, Tedelex.

Motor: Dunlop, Toyota.

Paper & Packaging: Consol, Metal Box, Sappi.

Pharm & Medical: Adcock, S.A. Drug.

Printing & Publish: SAAN.

Steel & Allied: Cullinan, Hiveld.

Stores: Edgars, OK, Pick 'n Pay.

Sugar: Tongaat.

Tobacco & Match: Lion Match, Remgro.

Transportation: Putco, Safmarine.

The percentages of companies chosen in broad groups of similar categories was as follows:

Mining and Mining Houses:	17 out of 101 = 17%
Financial (other than Mining Houses):	8 out of 47 = 17%
Industrial : Building and Construction, Electrical, Engineering, Steel & Allied	12 out of 77 = 16%
Industrial : Others	38 out of 275 = 14%
Total:	75 out of 500 = 15%

In addition to the above companies, questionnaires were sent to the following nine non-quoted companies:

Building Societies: UBS, S.A. Perm.

Insurance: S.A. Mutual, Sanlam.

Semi-state: Armscor, Escom, Industrial Development Corp., Iscor, S.A. Transport Services.

From these 75+9 = 84 questionnaires there were twenty four responses (29% response). The questionnaire included a suggestion to respondents that they remain anonymous should they so desire.

Of the twenty four respondents, fourteen used statistics and three of these chose to remain anonymous. There were seven of the twenty four respondents who did not use statistics, of which three remained anonymous. The remaining three of the twenty four respondents stated that they used statistics, but they failed to mark any statistical methods listed on the questionnaire as being used.

The list of the twenty four Stock Exchange respondents is as follows:

Companies using statistics (14)	Not using statistics (7)
Anglovaal	Federale Volksbeleggings
AECI	Metal Box Flex Packaging
Barclays	Premier Group
Dorbyl	Seardel
Escom	Three anonymous
Haggie Rand	
Rennies	**Stated that they used statistics but failed to mark any methods (3)**
Romatex	
Tongaat-Hulett Sugar	
UBS	Randfontein Estates
S.A. Mutual	Stanbic
Three anonymous	Standard Brass

There were twenty companies from the UCT sample that returned their questionnaires to Wits University. As a number of these were, by chance, quoted on the Stock Exchange, and as some others were of a similar type to certain quoted companies, it was decided to pool these with the twenty four companies from the Stock Exchange to carry out a preliminary analysis. Of these twenty companies from the UCT sample, there were six which stated that they did not use statistics.

The list of the twenty UCT respondents who returned their questionnaires to Wits is as follows:

Companies using statistics (14)	Not using statistics(6)
AC Nielsen (Market Research)	Kimberlet Engineering Works
Arthur Young (Chartered Accountants)	Maneng (Manufacturing)
Cleminson-Plaskitt (Manufacture)	Richard Ellis (Marketing)
Dunlopillo (Manufacture)	Stand. Charter Mer. Bank
Epol (Manufacture)	SA Foundation
Everite (Building)	Suncrush
Ford Motor Company	
Hoek and Wiehahn (Auditors)	
Hulett Aluminium	
ICI (SA)	
Lindsay Saker	
PE Corporate Services	
Southern Life	
Whiteleys (Manufacturer)	

In Section 6.4. comparisons are made of method usages between employee groups. One group in Section 6.4 is called the Stock Exchange group, and it consists of the twenty eight of the above companies which responded to the question on method usage. These are the fourteen from the Stock Exchange sample that use statistics plus the fourteen from the UCT sample that use statistics.

The group called the Business School group in Section 6.4. is made up of the balance of the UCT respondants who sent their questionnaires directly to UCT. There were fourteen of them who used statistics and marked the methods they used and seven who did not use statistics, as follows:

Companies using statistics (14)	Not use statistics (7)
Adcock-Ingram Laboratories	Caltex Refineries
Nampak Paper	C I Caravans
National Panasonic	Grinaker Construction
Metal Closures Group	J R Kohn
Old Mutual Planning and Development	Strip Form Packaging
Old Mutual Properties	TA Computers
Rennie Grinaker Zimbabwe	One anonymous company
Samancor	
Zinchem (Gold Fields)	
Five anonymous companies	

From the combined Stock Exchange and UCT Business School samples there were 65 responses out of 244 questionnaires sent out, giving a response rate of 27%. Of these 65 responders, 42 used statistics and 23 did not.

2.3. Pilot survey for Questionnaire No.1

Questionnaire No.1 which was designed by the three UCT staff members mentioned in Section 2.1.(b), was tested by the author by posting eleven copies of the questionnaire to various acquaintances in business firms, as set out below:

Technical Manager of Triomf Crop Protection
Technical Manager of NPI (Pty) Ltd (Explosives Manufacturer)
Director of FMS Management Services
Senior Statistician at AC Nielsen (Market Research)
Chief Engineer in Directorate of Water Affairs
Research Statistician at AECI Ltd, Modderfontein
Head, Management Services Group, AECI Ltd, Sasolburg
Statistician in Management Services Division of Standard Bank
Managing Director of Vrede Textiles
Technical Manager of Jabula Food
Statistician at Brooke Bond-Oxo

Two of the above people did not respond, one because he did not receive the questionnaire and the other due to pressure of work. The latter stated that the questionnaire sat in his in-tray and always got sidelined because he was daunted by its length. This was one of the reasons for reducing the size and contents of Questionnaire No.2.

As a result of the feedback from the nine respondents, the author made a number of suggestions to the designers of the questionnaire in Cape Town, many of which were incorporated in an improved version of Questionnaire No.1.

One aspect which was not clarified entirely revolved around Question 9.1. of the questionnaire: "How many people in your company use these techniques?"
and Question 9.2.: "Who applies the Statistical/Operations Research methods?"
In a big company there are people who carry out or use the statistical techniques (eg statisticians, etc.) and people who interpret and use or apply the results of the techniques (eg Technical Managers, Directors, etc.). When is one "using" a method and when is one being "applied"?

A second problem that arises is that some companies interpreted Question 9.2. as "at least one user has the ticked qualification". Thus large companies showed many "users" in Question 9.1., but only ticked a few blocks in Question 9.2. Other smaller companies showed the qualifications of each user, so that the number of ticks in Question 9.2. equalled the number of users shown in Question 9.1.

The result of the above problem is that the exact number of users of different qualification cannot be determined. The results of these two questions are reported in Section 4.5.3.

2.4. Non-response

The failure to measure some of the sample units in the selected sample is known as non-response. Cochran (1977:292) classifies non-response into four categories, namely:

(i) Non-coverage. The failure to locate or visit some sample units.

(ii) Not-at-homes. The respondent is temporarily absent when the interviewer calls.

(iii) Unable to answer. The respondent does not have the required information or may be unwilling to give it. Skillful wording and pre-testing of the questionnaire can reduce this type of non-response.

(iv) The "hard-core". Respondents adamantly refuse to respond, are incapacitated or permanently absent during the period in which the field work is undertaken. This is a source of bias that persists no matter how many call-backs are made to complete all returns.

Cochran considers that

> it is convenient to think of the population as divided into two 'strata', the first consisting of all units for which measurements would be obtained if the units happened to fall into the sample, the second of the units for which no measurements would be obtained.

Non-response will not introduce a bias if the characteristics of both strata are the same, but it has often been found that this is not so. In this postal survey there was a 73% (179 out of 244) non-response from the combined Stock Exchange and UCT Business School samples, an 85% (365 out of 430) non-response from the SASA and a 95% (391 out of 410) non-response from the OR Society. The question arises as to whether the stratum of responders differs from that of the non-responders.

When considering the eighty four companies (seventy five on Stock Exchange and nine others) that were sampled, it is important to realise that it is not the firms themselves that are of interest, but rather the statisticians that work for them. As discussed in Section 2.2., the sample of companies was merely a means of finding and sampling statisticians. The method of sampling companies was designed to select those which were more likely to employ statisticians or consult them. The effect of non-response of companies is therefore not very serious.

When considering the 160 graduates of the UCT Business School and the members of the SASA and OR Society, it must be remembered that the questionnaire was mainly concerned with statistical methods in use. Therefore it is possible that all those statisticians who were in managerial positions and therefore no longer active users of the methods, would probably not have felt the questionnaire applied to them.

It is very likely that the respondents were interested in education. In this respect it is probable that the interests of those in the response and non-response strata were different. Hopefully, their use of statistical methods was independent of their interest in education. There is really no way of testing this without undertaking a much larger study than was possible in this survey. Any inferential conclusions that are made in this dissertation will only be as reliable as the sample is representative of the population of all statisticians.

To get an idea of the representativeness of the sample of businesses, it is informative to pool all the respondents who use statistical methods and classify them according to type of employment. There are 28 (Stock Ex.) + 14 (UCT) + 25 (State) + 25 (Univ) + 15 (Private) + 19 (ORSoc) + 11 (Pilot Survey) less 6 unspecified returns =131 respondents.

University/Technikon/School	27	(21%)
State and semi-state	27	(21%)
Escom, Iscor, Armscor	7	(5%)
<u>Private sector</u>:		
Mining and mining houses	16	(12%)
Financial (Insurance, Banks, Building Societies, etc.)	13	(10%)
Industrial A: Bldg, Elec, Enging, Metal	8	(6%)
Industrial B: Other	33	(25%)
Total	131	(100%)

The cross-section of companies which responded appears to be reasonably representative of the South African economy. In particular the ratio of all the companies on the Stock Exchange in the four categories

Mining : Financial : Industrial A : Industrial B

is 20:10:15:55 (adding to 100).

For the Private sector respondents above, the ratio is 23:19:11:47 (adding to 100).

In view of this representative spread of statisticians across the whole spectrum of the economy, it is hoped that the results obtained in this study will not cause any biased conclusions to be drawn.

2.5. Questionnaire No.2

The main purpose of this questionnaire was to obtain information on the usage of the thirty three statistical methods listed in Questionnaire No.1, from statisticians of the SASA in university, State and semi-state employ, as well as those in the private sector. It was also sent to members of the OR Society. As it was being sent to individuals, most of whom were not working for private companies, it was decided to dispense with the questions on size and type of company. The opportunity was taken of designing the questionnaire in a much more concise form than Questionnaire No.1 had been, and, in fact, it was possible to fit it onto a single A4 sheet. The order of the list of thirty three statistical methods was changed slightly on the form sent to the OR Society members, following a request from the Society's committee for what they considered to be a better grouping of similar methods.

Questions 9.8 and 9.9 from Questionnaire No.1 on computer package usage and type of micro-computer used, were repeated in Questionnaire No.2.

The following open-ended question was asked:

> Please comment on whether you felt your statistical training gave you the skills necessary for the adequate performance of your job(s).
>
> Suggestions for improved training would be appreciated.

The replies to these questions will be discussed in Chapter 8.

There were sixty five returns from the members of the SASA, of which twenty five were employed by universities, twenty five by State or semi-state bodies, and fifteen in the private sector. The nineteen members of the OR Society who replied, consisted of fourteen from the private sector, three from semi-state bodies, one from a university and one from a high school. All nineteen members of the OR Society have been kept together as a fourth category of employees.

The institutes and companies for which the respondents worked are as follows:

State and semi-state (25)

Cent. Econ. Advis. Serv.	(1)
CSIR, Nat Bldg Res Inst	(1)
CSIR, Nat I Trans & Rds	(1)
CSIR, Nat R I Math Sci	(2)
Dept of Agriculture	(2)
Human Sciences Res C	(5)
I Maritime Technology	(2)
Kaapse Technikon	(1)
Medical Res Council	(2)
Med. R.C., I Biostat.	(7)
PLAN Assoc.	(1)

OR Society (19)

AECI	(1)
Allied Bldg Soc.	(1)
Anglo-American	(1)
Anglovaal	(1)
Armscor	(1)
Colonial Mutual	(1)
Commercial Union	(1)
Escom	(1)
Gencor	(1)
Gold Fields	(1)
High School, Swaziland	(1)
JCI	(3)
Medical Res Council	(1)
Nat Bldg Res Inst	(1)
Rand Mines	(1)
Tongaat-Hulett Sugar	(1)
Wits Actuarial Dept	(1)

Universities (25)

Medunsa	(1)
Natal (Pmb)	(3)
Oranje-Vrystaat	(3)
Port Elizabeth	(1)
Potchefstroom	(1)
Pretoria	(2)
RAU	(1)
Rhodes	(1)
Stellenbosch	(1)
UCT	(2)
Univ of North	(1)
UNISA	(4)
Western Cape	(2)
Witwatersrand	(2)

Private sector (15)

Barlows Manuf.	(1)
De Beers, Kimberley	(1)
DM Hawkins Conslt	(1)
Escom	(1)
Goldfields	(1)
Iscor	(2)
Market Res Africa	(1)
Multicon Pta	(1)
Sankorp, Jhb	(1)
Trust Bank	(1)
UBS	(3)
Western Deep Levels	(1)

2.6. Qualitative data

Statisticians work almost exclusively with quantitative data, which are amenable to numerical analysis and easy to summarize and tabulate. Qualitative data is so seldom dealt with by statisticians that some discussion concerning it is appropriate. Patton (1980:22) devotes a chapter to the nature of qualitative data from which the quotations which follow have been taken.

> Qualitative data consist of detailed descriptions of situations, events, people, interactions and observed behaviours; direct quotations from people about their experiences, attitudes, beliefs and thoughts. The data are collected as open-ended narrative without attempting to fit program activities or peoples' experiences into predetermined, standardized categories, such as the response choices that comprise typical questionnaires or tests.

Qualitative data can be forced into a quantitative form by attaching numerical values to response categories, for example

Strongly agree/Agree/Disagree/Strongly disagree

However, much of the "depth and detail" of the information is lost by this process. Open-ended questions enable the evaluator to "record and understand what peoples' lives, experiences, and interactions mean to them in their own terms and in their natural settings."

The researcher can "understand and capture the points of view of other people without predetermining those points of view through prior selection of questionnaire categories".

Patton (1980:23) gives a very illuminating case study in which quantitative responses to questionnaires showed that 88 - 97% of the respondents (school teachers) reacted negatively to various questions concerning a new accountability system introduced into their schools. School authorities and school board members initially dismissed these results as biased, inaccurate and caused by a teacher-union campaign to discredit them. However, their tune changed drastically on reading the open-ended comments, which dramatically highlighted the disastrous collapse in teacher morale resulting from the atmosphere of fear, intimidation and distrust brought about by the accountability system.

Qualitative data are voluminous, variable in content and difficult to analyse. Patton (1980:299) states that "there is no right way to go about organising, analysing and interpreting qualitative data. Each qualitative analyst must find his or her own process". Patton suggests various methods of analysis including Content Analysis, in which "every paragraph of every comment is organised into topics and files", thereby enabling the content to be indexed and classified. A card is punched for each idea in each paragraph and a computer retrieval system permits a print-out of all passages on any subject.

The qualitative data in this study arise from open-ended questions used in evaluation questionnaires following continuing education courses and also from the comments from statisticians in Questionnaire No 2 on the adequacy of their training. The method of analysis has bean to develop classification systems by convergence of topics as suggested by Guba (1978:53) and reported by Patton (1980:311-313). The "recurring regularities' in the data are noted. They represent patterns which can be sorted into categories. The categories are then judged by two criteria: "internal homogeneity" and "external heterogeneity".

The first criterion concerns the extent to which the data that belong in a certain category hold together or 'dovetail' in a meaningful way. The second criterion concerns the extent to which differences among the categories are bold and clear.

This effort at uncovering patterns, themes and categories is a creative process that requires making carefully considered judgements about what is really significant and meaningful in the data. Since qualitative analysts do not have statistical tests to tell them when an observation or pattern is significant, they must rely on their own intelligence, experience and judgement. Sometimes this leads to the making of the qualitative analyst's equivalent of the type I and type II errors defined for quantitative statistics.

In analysing the qualitative data arising from the open-ended questions in this study, a simplified form of Content Analysis is used. Comments are grouped into topics as shown in Chapt.r 8.

CHAPTER 3. STATISTICAL EDUCATION AT SOUTHERN AFRICAN UNIVERSITIES

In a study of this type, which is centres on South African university education, it is very necessary to list the universities together with their staffing resources and the nature of the statistical courses they offer. This is attempted in this chapter.

3.1. Departments of Statistics at Southern African universities

There are 21 universities in the Republic of South Africa and its former "homelands". Of these all but two have statistics departments. There are four universities in adjacent territories who are members of the Commonwealth Universities and three of them have statistics departments.

3.2. Staff numbers

Staff numbers shown in Table 3.1. were obtained from the various university calendars for 1985 and 1986.

TABLE 3.1 Statistical staff at Southern African universities

University	Professor	Assoc Prof.	Sen. Lect.	Lecturer	Jun. Lect.	Total
Bophuthatswana	-	-	2	1	2	5
Durban-Westville	-	-	-	2	-	2
Fort Hare	1	-	-	3	-	4
Natal(Durban)	1	1	1	3	-	6
Natal(Pmb)	1	-	2	2	-	5
Oranje-Vrystaat	2	1	3	3	-	9
Port Elizabeth	3	-	1	1	1	6
Potchefstroom	5	1	1	3	3	13
Pretoria	3	1	3	7	-	14
RAU	2	-	2	2	2	8
Rhodes	2	-	-	2	1	5
Stellenbosch	3	-	2	3	1	9
Transkei	-	1	-	2	1	4
U C T	3	3	2	3	-	11
Univ. of North	1	-	-	4	1	6
UNISA	5	1	2	5	2	15
Venda	-	-	-	-	-	0
Vista	-	-	-	-	-	0
Western Cape	1	-	2	2	-	5
Witwatersrand	3	1	3	4	1	12
Zululand	-	-	1	1	-	2
TOTALS	36	10	27	53	15	141
Botswana	1	-	1	6	1	9
Lesotho	1	-	1	2	5	9
Swaziland	-	-	1	3	2	6
Zimbabwe	-	-	-	-	-	0
TOTALS	2	0	3	11	8	24

Note 1. This information was obtained from University Calendars for 1985 and 1986.

2. The University of Zimbabwe offers statistics courses given by members of the Mathematics Department.

3.3. Preponderance of staff in the professorial ranks

Underhill (1982) sent out a questionnaire survey to South African universities and received replies from thirteen or them. The results of the survey, as reported in the Newsletter of the S.A. Statistical Association of March 1983, show the following position with regards to statistical staff:

Rank	Filled	Vacant
Professor & Associate Professor	26	3
Senior Lecturer	21	15
Lecturer	31	7
Junior/Temporary Lecturer	17	0

It is clear that in 1983 there was a critical shortage of senior lecturers. This situation seems to have deteriorated by 1985 as the ratio of Prof and Ass.Prof : S.Lecturer : Lecturer : J.Lecturer is seen in Table 3.1 to be 46 : 27 : 53 : 15.

The staff complement in the 13 universities who replied in 1982 was 29 : 36 : 38 : 17.

The proportion of professors and senior lecturers to lecturers and junior lecturers in both cases is similar, for the 1985 incumbents and the 1982 complement, which are, respectively:

$$\frac{46+27}{53+15} = 1,1 \quad \text{and} \quad \frac{29+36}{38+17} = 1,2$$

The proportion of professors to senior lecturers is very different for the 1985 incumbents and the 1982 complement, which are, respectively:

$$\frac{46}{27} = 1,7 \quad \text{and} \quad \frac{29}{36} = 0,8$$

Thus there seems to be a tendency for universities to promote senior lecturers to professorial rank faster than intended by the original staffing complement. This is probably due to the difficulty experienced by universities in retaining the services of good experienced statisticians in the face of competition from other universities, semi-state bodies and industry.

3.4. Statistical methods taught to major students

The questionnaire of Underhill (1982) included a similar list of statistical methods as that used in Questionnaires No 1 and No 2. The results of this survey are shown in Table 3.2. The relationship between these methods taught at university and the methods used in industry is examined in Chapter 7.

TABLE 3.2 Number of universities teaching each method in a given year

Statistical methods used	1st year	2nd year	3rd year	Hons
1. Basic Statistics	13	0	0	0
2a. Graphical display, data summary	13	0	0	0
2b. Stem & leaf, Box & whisker	4	0	0	0
3. Regression analysis	13	7	13	8
4. Correlation analysis	12	5	10	7
5. Analysis of variance	5	4	13	9
6. Design of experiments	2	1	1	9
7. Variance component estimation	1	1	7	4
8. Bioassay	0	1	1	1
9. Contingency tables Chi-sqared	12	5	5	3
10. Quality control & Accept. sampling	5	4	5	4
11. Non-linear estimation	0	0	1	6
12. Non-parametric methods	9	5	6	9
13. Ranking & Paired data analysis	7	5	8	5
14. Multivariate analysis	0	0	10	11
15. Probability modelling	4	3	5	7
16. Linear programming	4	4	9	9
17. Non-linear programming	0	0	4	8
18. Dynamic programming	0	0	1	9
19. Inventory control	1	2	4	9
20. Simulation	0	3	3	7
21. Reliability & Life testing analysis	0	0	2	5
22. Time series analysis	9	2	3	11
23. Forecasting	4	1	1	10
24. Survey sampling	4	4	5	10
25. Categorical data analysis	3	2	4	5
26. Queueing theory	1	2	8	10
27. Factor analysis	0	0	1	9
28. Decision theory	0	1	7	7
29. Network models (PERT/CPM)	1	3	5	6
30. Multidimensional scaling	0	0	0	4
31. Conjoint analysis	0	0	0	0
32. Indices	8	1	1	0
33. Econometrics	0	0	0	2

Note: Table 3.2. is taken from the Report of the Education Subcommittee by L.G. Underhill and J.M. Juritz in the Newsletter of SASA, March, 1983. It is reprinted with the authors' permission.

3.5. Service courses in Statistics

In late 1983 Steyn (1984) made a survey of the following 18 universities in South Africa:

Zululand	Cape Town
Durban - Westville	Orange Free State
North	Port Elizabeth
Pretoria	Natal (Durban)
Stellenbosch	Natal (Pietermaritzburg)
Witwatersrand	UNISA
Fort Hare	Rhodes
Rand Afrikaans	PU for CHE (Potchefstroom)
Western Cape	PU for CHE (Vanderbijlpark)

The questionnaire was designed to yield information on the nature and extent of service courses at these universities. A copy of the questionnaire is attached as Appendix C.

Table 3.3. shows the extent of service courses in the four main areas together with the approximate length of the courses. The total number of service courses offered by all universities was 54, with a median number per university of 2 and a maximum of 11 different service courses at one of the universities.

TABLE 3.3. Extent of service courses at 18 universities in 1983 together with the student enrolment

Study area	Semesters	Students
1. Commerce, Administration and Architecture	2	7430 (69%)
2. Biological sciences (Agriculture, Medical, Dental, Pharmaceutical, Botany, Zoology, Domestic science, Occupational therapy)	1 or less	2070 (19%)
3. Engineering	1	690 (6%)
4. Social sciences (Psychology, Communications, Sociology, Geography, Speech therapy)	1 or less	630 (6%)
TOTAL		10820 (100%)

Note: Table 3.3. is taken from Steyn (1984) with the author's permission.

The numbers of students shown in Table 3.3. are those lectured to by statistics lecturers. Many departments offer courses themselves which have a high or total statistical content. Table 3.4. shows the number of courses which are lectured to by staff in their own departments. Unfortunately the student numbers are not known.

TABLE 3.4. Statistics courses offered by departments other than statistics departments

Subject	Number of Universities
Mathematics	1
Botany/Zoology	2
Psychology	7
Sociology	1
Engineering	1
Business (MBA, MBL)	5
Education	3
Geography	6

Note: Table 3.4. is taken from Steyn (1984) with the author's permission

3.6. Summary

Some background information on statistics teaching at South African universities has been included in this chapter. This involves staffing, the statistical methods taught to major students and the service courses taught to non-major students.

There are some 140 teachers of statistics at the 21 universities in the Republic and National States. Of these, 46 (or 33%) are professors or associate professors, which appears to be a disproportionately large number.

In Chapter 7 the information on statistical methods taught to major students will be compared with the usage of these methods by statisticians employed in various sectors of the economy.

With reference to service courses there were a total of 10820 students being taught elementary statistics in 1983. When one considers that this averages 77 students per academic, it is obvious that service teaching places a high burden on the statistics staff of our universities. Some departments prefer to teach their own statistics courses. Departments that do this most frequently are Psychology, Geography and the Business Schools.

CHAPTER 4. SURVEY OF STATISTICAL METHODS USED IN COMMERCE

This chapter is a presentation of the results from the responses to Questionnaire No 1. The results of Questionnaire No 2 will be presented in Chapter 5.

4.1. Aims of the survey

It is an unfortunate fact that many University lecturers of Statistics have little or no personal experience of what statistical methods are used by practitioners in commerce and industry. Very often the staff of statistics departments are made up of people who were the top students in their year and were encouraged to stay on in the department to do postgraduate work and eventually to make a career as a teacher and researcher. Fortunately, as stated earlier in Section 1.4.1., there is an increasing tendency in South Africa for university statisticians to do consulting work, either privately or as part of their job. This is a healthy development, but does not alter the fact that few teachers of statistics know what techniques are used over the whole spectrum of the business world.

Very little information is available in the literature on what statistical methods are actually in use, except for the "informal survey" conducted by Snee et al (1980) of only twenty American industrial statistics groups.

In order to gather information about the usage of statistical methods in South Africa, the Education Committee of the SASA decided to conduct a survey of business firms by means of Questionnaire No 1, which has been discussed fully in Section 2.1(b) and 2.2. As explained in Section 2.1(b), a paper was subsequently prepared for publication in 1985 by J.M. Juritz, with the author and two other co-workers as co-authors. Certain paragraphs of this chapter follow this paper fairly closely. The author has Professor Juritz's permission to use the paper in this dissertation and will show these passages with quotation marks, although they are not always exact quotes.

"The main aim of this survey was to establish which parts of the subject material are frequently used, which are considered of little practical value and which parts, although perhaps potentially useful, are not used at all. Perhaps there are serious gaps in the statistical training being offered at universities. It could be that there are methods that are used in commerce and industry which are not dealt with in academic courses."

Another aim was to establish what computer packages were being used. "These packages make a wide variety of sophisticated statistical techniques available to users who may have little or no formal statistical training." It could be that more emphasis should be given to educating the users in the correct application of these techniques. This would not only optimise the use of packages but minimise the danger of using techniques on data to which their application is inappropriate or invalid.

4.2. Questionnaire No.1

As set out in Section 2.1(b), the questionnaire was designed by three staff members of the University of Cape Town, viz. Professors Juritz, Money and Affleck-Graves. A small pilot survey of 11 questionnaires was carried out by the author and some minor amendments were made to the questionnaire as a consequence of replies received.

"The information sought in the questionnaire fell into four categories:

(i) the type of business or industry;

(ii) the number of people using statistical methods and their qualifications;

(iii) the methods used and whether their usage was routinely or occasional;

(iv) the computer facilities, both hardware and software.

Although interest centred principally on statistical methods, some questions on operations research techniques were also asked since some of them, such as simulation and inventory control, are closely linked to statistical methods. A separate question was also asked about the use of quality control. This subject seems to be developing in its own right, with very little acknowledgement being given, at least in the minds of practitioners, to the fact that it is essentially a special statistical method. The check list of topics is an extended version of those presented by Snee et al (1980) in their report to the American Statistical Association on the training of statisticians for industry.

Computer packages play an important role in the application of statistics, and much can be deduced about the type and quality of an application from the knowledge of the computer package used. So the questionnaire included a check list of well-known computer packages. There is no doubt that the micro-computer will also play an increasingly important role in business in the future, so a question on its use was also included."

4.3. The sample

As explained in Section 2.2 the sampling frame consisted of most of the companies on the Johannesburg Stock Exchange, the register of past students of the Graduate School of Business of the University of Cape Town and a further nine semi-state companies, building societies and insurance companies.

Sixty five responses were received from the 244 sent out, giving a response rate of 27%. None of the non-respondents were contacted. Confidentiality was guaranteed and, in fact, some respondents did not give information on company name and address. The companies which responded are listed in Section 2.2.

4.4. Profile of the respondents

In the discussion that follows, the data from the two parts of the survey are pooled, except for the data on methods used (Tables 4.10 and 4.11 in Section 4.5.5.). (Note that the discussion of the results in Sections 4.4 and 4.5 follows very closely the draft of the joint publication prepared by Professor J.M. Juritz on behalf of the 4 co-authors, and is used here with her permission. Because of some reluctance on the part of the Journal of the SASA to publish articles of this type, this paper has not yet been accepted for publication in that journal.)

A profile of the respondents is given in Table 4.1, 4.2 and 4.3 below. Note that some companies reported more than one activity.

TABLE 4.1. <u>Activity of the Company (some reported more than one activity)</u>

Activity of Company	User of Statistics	Non-User
Banking	3	1
Building and Construction	3	1
Chemical Industry	3	0
Manufacturing	23	7
Marketing	11	2
Mining	4	0
Other	15	8
Total	62	19

Note. Table 4.1. is used here with the authors' permission.

"Companies that reported their activities under 'other' in the user category included three insurance firms, a computer bureau and two firms engaged in Economic and Market Research. A property investment firm also reported using statistical methods.

Activities of non-users listed under 'other' included a firm of Marine Architects, two retail firms, a computer bureau and an industrial holding company. It is surprising that one banking firm and two marketing firms reported that they were non-users. It is possible in these cases the questionnaire did not reach a person competent to answer it.

Twenty-seven percent (27%) of the responses came from companies that operated on a local level, 50% on a national and 23% were international companies. In the South African context this would imply that they were subsidiaries of an international company whose headquarters were overseas."

TABLE 4.2. Status of company (One of the sixty five respondents did not report)

Status of Company	% of Sample	Users of Statistics	Non-users	% Users
Local	27%	10	7	58%
National	50%	25	7	78%
International	23%	11	4	73%
Total	100%	46	18	72%

Note: Table 4.2. is used here with the authors' permission

"Although it would appear that fewer local companies use statistical methods, there were no significant differences in the percentages of users in the three types of companies."

TABLE 4.3. Number of employees per company (One of the sixty five respondents did not report)

Number of Employees	Users of Statistics	Non-users	% Users
100 or fewer	2	7	22%
101 - 500	10	3	77%
501 - 1000	3	4	43%
More than 1000	31	4	88%
Total	46	18	X^2 = 18,88 P < 0,001

Note: Table 4.3 is used here with the authors' permission

"55% of the responses came from very large companies employing more than 1000 people. There were significant differences in the size of the company between user and non-user groups. (P < 0,001 using the Chi-squared test). This could be attributed to the fact that the smallest companies, those of 100 or fewer employees, tended not to use statistical methods, (seven non-users compared with two users) but of the thirty five very large companies only four reported that they did not use statistical methods. Further inv tigation of these cases showed that one of these was a holding company, but two of them were large retail firms. It was suspected that the latter do use some statistical methods but the person filling in the questionnaire was not aware of them. There might be some justification for deleting these cases from the sample, but it was decided to retain them."

4.5. Findings from the replies to the survey

4.5.1. Reasons for firms not using statistical methods

(Note that this section follows Juritz et al (1985) very closely)

Reasons for not using statistical methods are given in Table 4.4.

TABLE 4.4. Reasons for not using statistical methods

1	They are not applicable to our business	7
2	There are no staff with the necessary expertise	3
3	We have done so in the past and found them useless	0
4	We know nothing about them	3
5	Others (please comment)	5

Note: Table 4.4 is used here with the authors' permission.

"Among the reasons given for non-users were : 'Company only in operation a short time'; 'don't think they are needed' - this from a large construction company who used operations research methods. It was a relief to note that no firm reported that it did not use statistics because they had proved useless."

4.5.2. Use of Statistics, Operations research and Quality control

(Note that this section follows Juritz et al (1985) very closely).

Questions 4, 5 and 6 of the questionnaire asked separately whether the company used Statistics (Yes/No), OR (Yes/No) or QC (Yes/No). Only sixty two respondents filled in complete information on these three questions. The eight respondents who stated they were users of QC but not statistics appeared to be unaware that QC is essentially an applied statistical technique. Although not all OR techniques can be claimed to be statistics, in this survey the four OR users, who did not also use QC, stated that they were users of both OR and statistics. For comparative purposes all QC users will also be classed as users of statistics. The data are shown in Table 4.5.

TABLE 4.5. The use of Statistics, Operations research and Quality control

Methods as stated	No. of Companies	Methods taken to be	No. of Companies
None	8	None	8
Statistics	9	Statistics	9
Operations Research	0		
Statistics + OR	4	Statistics + OR	4
Quality Control	8	Statistics + QC	19
Statistics + QC	11		
OR + QC	1	Statistics + OR + QC	22
Statistics + OR + QC	21		
Total	62	Total	62

Note: Table 4.5 is adapted from Juritz et al (1985), with the authors' permission.

4.5.3. Firms using statistical methods

(Note that this section follows Juritz et al (1985) very closely).

Of the users of statistical methods, forty one commented on the effectiveness of the techniques, as shown in Table 4.6.

TABLE 4.6. Effectiveness of statistical techniques

Useless	1	2	3	4	5	6	7	Very effective
Percentage ex 41			5%	12%	37%	24%	22%	

Note: Table 4.6 is used here with the authors' permission.

"Usually within a company more than one employee was using statistics, and in large companies, many were. A frequency distribution of the number of users is given in Table 4.7."

TABLE 4.7. Numbers of users of statistics (Four companies gave no information)

Number of users	Number of companies	%
1 - 5	25	58%
6 - 10	6	14%
11 - 15	2	5%
16 - 20	6	14%
21 - 60	2	5%
Over 60	2	5%
Total	43	101%

Note: Table 4.7 is used here with the authors' permission

Median number of users per company is 5. Semi-interquartile range is 6.

The formal statistical qualifications of the users ranged from none (i.e. Matriculation only) to a doctorate in the subject.

The qualifications of the users are given in Table 4.8.

TABLE 4.8. <u>Qualification of users of statistics</u>

Qualification	Statistics	Other
Matriculation	-	14
Bachelors degree	34	9
Honours degree	11	11
Masters degree	5	16
Doctorate	3	2
Total	53	52
Other	10	

Note: Table 4.8 is used here with the authors' permission.

"Of the Bachelors graduates sixteen had a statistics or operations research major and eighteen had included some statistics courses in their degree. Among the 'other' qualifications reported were technical diplomas in quality control, marketing and building management and chartered accountancy. So it would appear that at least six of the ten 'other' cases had had some statistical training though not from a university. Most of the matriculants were employed by large companies and were possibly engaged in routine statistical tasks.

Although the individual qualifications of all the users are not known it would appear from Table 4.8. that about half of them had some formal statistical training. However, the overall level is low compared to that recommended by Snee et all (1980), who say that a training program for statisticians in industry should be aimed at the Masters level or higher.

Admittedly South African business and industry is not as sophisticated as the American, but it is clear that the industrial expansion that is so vital for the full development of the country will place heavy demands on statisticians. The general level of statistical qualifications in commerce and industry needs to be raised if it is to meet this challenge."

4.5.4. Computer packages used by business

(Note that this section follows Juritz et al (1985) very closely).

"Apart from the Box-Jenkins forecasting package and perhaps SAS, there appears to be relatively little use made of statistical packages. Many reported using programs developed by their company."

TABLE 4.9. Computer packages used by respondents

Statistical Package	Number of Users Ex 42	
	Occasional	Routine
BMDP	1	2
COMET	0	1
GLIM	0	0
SPSS	1	0
SAS	4	2
FMPS	2	1
MDSX	3	1
SIBYLRUNNER	0	1
BOX-JENKINS	0	2

"However, thirty eight respondents out of forty four reported using a micro-computer, with IBM, Apple and Hewlett-Packard being the most common.

These two findings highlight the need to include some training in statistical computing in our courses. Even if the practising statistician does not actually write the program himself he could reasonably be expected to advise on algorithms or efficient numerical techniques for a given task."

4.5.5. Statistical methods used by business

Juritz et al (1985) analysed the pooled data taken over both the Stock Exchange and the UCT Business School samples. In this section the data on usage of statistical methods are kept separate. Later tests of significance show that it is justifiable to pool the two samples.

Twenty eight of the Stock Exchange sample and fourteen of the Graduate School of Business sample gave information on the methods they used. These methods have been ranked in order of their total frequency of use and are given at the end of this chapter in Table 4.10. (Routinely) and Table 4.11. (Routinely + Occasionally).

When considering curricula it is necessary to decide where to draw the line between methods that are used commonly and those that are used less often. For the purposes of statistical training, a method that is used by 30% or more of the respondents, either frequently or occasionally, is obviously a method which should be included in the curriculum. Similarly a method used routinely by 20% or more of the respondents must be considered important from an educational point of view.

The use of the above two criteria on the total usage over both Stock Exchange and UCT Business School samples in Tables 4.10 and 4.11, results in the following sixteen methods being judged as important. (Note the order of merit is very similar from both Tables and this is an approximately average order):

Graphical display and Data summary

Basic statistical methods

Forecasting

Regression analysis

Time series analysis

Analysis of variance

Quality control and acceptance sampling

Linear programming

Correlation analysis

Inventory control

Simulation

Network models (PERT/CPM)

Survey sampling

Indices

Contingency table Chi-squared tests

Multivariate analysis

If the two criteria are applied to both samples separately then an additional two methods join the list, namely, Non-linear estimation and Design of experiments.

This list of commonly used methods achieves the main aim of the survey, which was to see which of the statistical methods are more frequently used. Individual teachers of statistics can use this list to judge whether their teaching emphases agree with the needs of industry. Later in this dissertation an analysis of statistics teaching at university level is presented.

An additional aim, which was to investigate the usage of computer packages, revealed that little use is made of them. On the other hand, 86% of respondents used micro-computers. This suggests the need for a greater emphasis in university courses on computer programming and on the use of packages.

The overall profile of statistics usage in industry suggested that the smaller, local companies were often ignorant of the potential benefits of statistical techniques. Only about half the users of statistics had some formal training in the subject. Juritz et al (1985) concluded that the general level of statistical training needs to be upgraded if South African businesses and industries are to meet the challenges of the future.

TABLE 4.10. <u>Methods used Routinely in two samples: Frequency and percent rounded to nearest 10%</u>

Methods used Routinely	Stock Exchange		UCT Bus School		Total
	Freq.	% ex 28	Freq.	% ex 14	% ex 42
Graphical display & Data summary	24	90	8	60	80
Basic statistical methods	22	80	9	60	70
Forecasting	17	60	9	60	60
Analysis of variance	11	40	7	50	40
Time series analysis	9	30	8	60	40
Regression analysis	11	40	6	40	40
Quality control & acc sampling	11	40	5	40	40
Inventory control	11	40	3	20	30
Correlation analysis	7	30	4	30	30
Indices	11	40	0	0	30
Linear programming	6	20	4	30	20
Simulation	8	30	2	10	20
Survey sampling	6	20	4	30	20
Network models (PERT/CPM)	8	30	0	0	20
Multivariate analysis	4	10	1	10	10
Reliability and Life testing	3	10	1	10	10
Contingency table Chi-sq tests	2	10	1	10	10
Ranking and paired comparisons	3	10	0	0	10
Decision theory	2	10	1	10	10
Design of experiments	2	10	0	0	5
Variance component analysis	2	10	0	0	5
Non-linear estimation	1	4	1	10	5
Logistic Regression	1	4	1	10	5
Non-linear programming	1	4	1	10	5
Categorical data analysis	0	0	2	10	5
Queueing theory	2	10	0	0	5
Discriminant analysis	2	10	0	0	5
Bioassay	1	0	0	0	2
Dynamic programming	1	0	0	0	2
Multidimensional scaling	1	0	0	0	2
Conjoint analysis	0	0	1	10	2
Non-parametric methods	0	0	0	0	0
Factor analysis	0	0	0	0	0
Total	190		79		

TABLE 4.11. Methods used overall (i.e. Routinely + Occasionally) in two samples: Frequency and percent rounded to nearest 10%

Methods used Routinely + Occasionally	Stock Exchange		UCT Bus School		Total
	Freq.	% ex 28	Freq.	% ex 14	% ex 42
Graphical display & Data summary	25	90	12	90	90
Basic statistical methods	24	90	12	90	90
Forecasting	23	80	11	80	80
Regression analysis	19	70	10	70	70
Linear programming	16	60	9	60	60
Analysis of variance	18	60	7	50	60
Time series analysis	14	50	10	70	60
Quality control & acc sampling	16	60	8	60	60
Simulation	15	50	8	60	50
Correlation analysis	15	50	7	50	50
Inventory control	14	50	5	40	50
Network models (PERT/CPM)	14	50	5	40	50
Survey sampling	11	40	4	30	40
Indices	12	40	1	10	30
Multivariate analysis	9	30	3	20	30
Contingency table Chi-sq tests	10	40	1	10	30
Non-linear estimation	8	30	2	10	20
Reliability and Life testing	6	20	2	10	20
Ranking and paired comparisons	7	20	1	10	20
Design of experiments	8	30	0	0	20
Decision theory	5	20	2	10	20
Non-parametric methods	4	10	1	10	10
Queueing theory	4	10	1	10	10
Multidimensional scaling	4	10	1	10	10
Logistic Regression	2	10	2	10	10
Non-linear programming	2	10	2	10	10
Dynamic programming	4	10	0	10	10
Factor analysis	4	10	0	0	10
Discriminant analysis	4	10	0	0	10
Variance component estimation	3	10	0	0	10
Categorical data analysis	1	0	2	10	10
Bioassay	1	0	1	10	5
Conjoint analysis	1	0	1	10	5
Total	323		131		

CHAPTER 5. SURVEY OF THE STATISTICAL METHODS USED BY MEMBERS OF THE S.A. STATISTICAL ASSOCIATION AND THE OPERATIONS RESEARCH SOCIETY OF S.A.

5.1. Questionnaire No. 2

This chapter is a presentation of the results of Questionnaire No 2.

This questionnaire was used to investigate the usage of statistical methods by members of the SASA and the OR Society. It was a reduced form of Questionnaire No.1, as discussed in Section 2.5. The basic information that was obtained from it pertained to the qualifications of the respondents, the statistical methods used and the computer packages and microcomputers used by the respondents. Academic qualifications for the various employee groups are compared in Section 6.1. The method usage is discussed in Sections 5.2 and in Chapter 6. The software and hardware used by respondents is not really pertinent to this dissertation, beyond confirming the conclusions of Section 4.5.4., namely, that statistical curricula should contain some training in computing and use of packages. The information or microcomputers and packages used has therefore not been included in this thesis.

5.2. Statistical methods used by SASA and OR Society

The percent usages of methods by the nineteen members of the O.R. Society and the sixty five members of the SASA (divided up into twenty five State or semi-state employees, twenty five University employees and fifteen in the Private sector) are set out below. Table 5.1. shows the percent usage of the methods used routinely while Table 5.2. shows the percent usage of the methods used both routinely and occasionally.

TABLE 5.1. Methods used Routinely: Percentages of employees using each method

Statistical methods used Routinely	SA Statistical Association				Oper. Res. Soc.	SASA + OR Soc.
	Univ. % ex 25	State % ex 25	Priv. % ex 15	Total % ex 65	% ex 19	% ex 84
1 Basic statistics	88	100	93	94	68	88
2 Graphical displ	76	92	93	86	74	83
3 Regression anal	92	76	67	80	53	75
4 Correlation anal	76	68	60	69	47	64
5 Anova	84	80	40	72	37	64
6 Design of expts	52	40	33	43	21	38
7 Var compt estmn	16	8	7	11	11	11
8 Bioassay	8	8	0	6	0	5
9 Chi squared	88	68	40	69	16	57
10 Quality control	28	8	40	23	16	21
11 Non-linear estmn	28	32	0	23	0	18
12 Non-parametrics	60	44	7	42	0	32
13 Ranking/Pair comp	36	44	7	32	5	26
14 Multivar anal	52	64	47	55	21	48
15 Logistic regn	32	40	0	28	0	21
16 Linear program	24	20	13	20	16	19
17 Non-linear prog	8	0	7	5	11	6
18 Dynamic prog	12	4	0	6	5	6
19 Inventory cntr	16	0	13	9	11	10
20 Simulation	44	32	20	34	37	35
21 Reliab Life test	12	16	13	14	5	12
22 Time series anal	52	16	47	37	32	36
23 Forecasting	44	24	47	37	37	37
24 Survey sampling	44	52	33	45	11	37
25 Categorical data	40	56	13	40	5	32
26 Queueing theory	32	0	0	12	11	12
27 Factor analysis	32	36	27	32	0	25
28 Decision theory	24	8	0	12	0	10
29 Network models	16	4	13	11	26	14
30 Mulidim scaling	20	24	7	18	0	14
31 Conjoint anal	4	4	7	5	0	4
32 Indices	20	12	20	17	11	15
33 Discrim analysis	40	44	13	35	5	29

TABLE 5.2. Methods used Overall (i.e. Routinely + Occasionally): Percentages of employees using each method

Statistical methods used Routinely + Occasionally	SA Statistical Association				Oper. Res. Soc. &	SASA + OR Soc.
	Univ. % ex 25	State % ex 25	Priv. % ex 15	Total % ex 65	ex 19	% ex 84
1 Basic statistics	96	100	100	98	95	98
2 Graphical displ	96	96	100	97	89	95
3 Regression anal	96	92	93	94	89	93
4 Correlation anal	96	88	87	91	74	66
5 Anova	96	92	87	92	74	88
6 Design of expts	88	72	60	75	58	71
7 Var compt estmn	56	36	20	40	21	36
8 Bioassay	32	20	7	22	5	18
9 Chi squared	92	84	67	83	63	79
10 Quality control	48	20	47	37	32	36
11 Non-linear estm	80	64	53	68	21	57
12 Non-parametrics	88	84	53	78	47	71
13 Ranking/Pair comp	60	60	33	54	37	50
14 Multivar anal	92	76	80	83	63	79
15 Logistic regn	72	76	13	60	5	48
16 Linear program	56	56	53	55	74	60
17 Non-linear prog	44	28	13	31	32	31
18 Dynamic prog	24	4	13	14	47	21
19 Inventory cntr	16	4	13	11	32	15
20 Simulation	84	64	60	71	79	73
21 Reliab Life test	48	24	33	28	21	26
22 Time series anal	68	64	73	68	63	67
23 Forecasting	52	56	73	58	95	67
24 Survey sampling	56	68	53	60	26	52
25 Categorical data	56	64	40	55	5	44
26 Queueing theory	32	8	40	25	21	24
27 Factor analysis	76	52	67	65	37	58
28 Decision theory	32	12	13	20	21	20
29 Network models	24	12	27	20	63	30
30 Multidim scaling	48	48	40	46	11	38
31 Conjoint anal	16	4	33	15	0	12
32 Indices	28	20	33	26	21	25
33 Discrim analysis	80	72	60	72	37	64

The application of the two criteria discussed in Section 4.5.5. can be used here to differentiate between the methods used commonly from those used less often. The criteria are either that 30% or more of the respondents use a method routinely or occasionally, or 20% or more use a method routinely. Applying these criteria gives the listings shown in Tables 5.3. and 5.4., respectively.

TABLE 5.3. Methods used Routinely by 20% or more of respondents with percentage usage rounded to the nearest 10%

University	% ex 25	State & semi-state	% ex 25	Private sector	% ex 15	Op. Res. Society	% ex 19
Regression	90	Basic stats	100	Basic stats	90	Basic stats	70
Basic stats	90	Graphics	90	Graphics	90	Graphics	70
Chi-squared	90	Anova	80	Regression	70	Regression	50
Anova	80	Regression	80	Correlation	60	Correlation	50
Graphics	80	Correlation	70	Multivariate	50	Anova	40
Correlation	80	Chi-squared	70	Time series	50	Simulation	40
Non-paramet	60	Multivariate	60	Forecasting	50	Forecasting	40
Expt design	50	Categ data	60	Anova	40	Time series	30
Multivariate	50	Sampling	50	Chi-squared	40	Network	30
Time series	50	Non-paramet	40	QC	40	Expt design	20
Simulation	40	Ranked comp	40	Expt design	30	Multivariate	20
Forecasting	40	Discrim anal	40	Sampling	30		
Sampling	40	Expt design	40	Factor anal	30		
Categ data	40	Logist anal	40	Simulation	20		
Discrim anal	40	Factor anal	40	Indices	20		
Ranked comp	40	Non-lin estm	30				
Queueing	30	Simulation	30				
Factor anal	30	Forecasting	20				
Logist anal	30	Mult scaling	20				
QC	30	Linear prog	20				
Non-lin estm	30						
Linear prog	20						
Dec theory	20						
Mult scaling	20						
Indices	20						

In Table 5.3. there are nine methods that are common to all four employee groups, namely, Basic statistics, Graphical display, Regression, Correlation, Anova, Experimental design, Multivariate analysis, Simulation and Forecasting.

It can be seen in Table 5.3. that eleven statistical methods are used by 20% or more of the OR Society respondents. Of these, Network models are not represented in the other employee groups, while Time series does not appear in the State and semi-state list. These two methods together with Simulation and Forecasting are used more by OR practitioners than by other statisticians.

In the list for the Private sector there are five new methods, namely Chi-squared, Quality control, Survey sampling, Factor analysis and Indices. Quality control and Indices are not used by State employees.

In the State and semi-state list eight new methods appear, namely, Categorical data analysis, Non-parametrics, Ranking/paired comparisons, Discriminant analysis, Logistic analysis, Non-linear estimation, Multidimensional scaling and Linear programming.

In the University list there are five methods in excess of those of State and semi-state employees, namely, Time series, Queueing, Quality control, Decision theory and Indices. At least one of the University respondents stated in his questionnaire that as a teacher of statistical methods he "used" the methods he taught. It can well be that the increased number of methods in the University list over the State and semi-state list is due to other academics considering "teaching" to be "using".

TABLE 5.4. Methods used Routinely and Occasionally by 30% or more of respondents with percentage usage rounded to the nearest 10%

University	% ex 25	State & semi-state	% ex 25	Private sector	% ex 15	Op. Res. Society	% ex 19
Basic stats	100	Basic stats	100	Basic stats	100	Basic stats	100
Graphics	100	Graphics	100	Graphics	100	Forecasting	100
Regression	100	Regression	90	Regression	90	Graphics	90
Correlation	100	Anova	90	Correlation	90	Regression	90
Anova	100	Correlation	90	Anova	90	Simulation	80
Chi-squared	90	Chi-squared	80	Multivariate	80	Correlation	70
Multivariate	90	Non-paramet	80	Time series	70	Anova	70
Expt design	90	Multivariate	80	Forecasting	70	Linear prog	70
Non-paramet	90	Logist regn	80	Chi-squared	70	Chi-squared	60
Simulation	80	Expt design	70	Factor anal	70	Multivariate	60
Non-lin estm	80	Discrim anal	70	Expt design	60	Time series	60
Discrim anal	80	Sampling	70	Simulation	60	Network	60
Factor anal	80	Non-lin estm	60	Discrim anal	60	Expt design	60
Logist regn	70	Simulation	60	Non-lin estm	50	Non-paramet	50
Time series	70	Time series	60	Non-paramet	50	Dynam prog	50
Ranked comp	60	Categ data	60	Linear prog	50	Ranked comp	40
Var compt	60	Ranked comp	60	Sampling	50	Factor anal	40
Linear prog	60	Linear prog	60	QC	50	Discrim anal	40
Sampling	60	Forecasting	60	Categ data	40	QC	30
Categ data	50	Factor anal	50	Queueing	40	Non-lin prog	30
Forecasting	50	Mult scaling	50	Mult scaling	40	Invent cntr	30
QC	50	Var compt	40	Ranked comp	30		
Reliab/Life	50			Reliab/Life	30		
Mult scaling	50			Conj anal	30		
Non-lin prog	40			Indices	30		
Bioassay	30						
Queueing	30						
Dec theory	30						

In Table 5.4. there are sixteen methods that are common to all four employee groups, namely, Basic statistics, Graphical display, Regression, Correlation, Anova, Design of experiments, Chi-squared, Non-parametrics, Ranking/Paired data comparisons, Multivariate analysis, Linear programming, Simulation, Time series, Forecasting, Factor analysis and Discriminant analysis.

The OR group have three methods not included in the Private sector group, namely, Network models, Dynamic programming, Non-linear programming and Inventory control. It is interesting that Forecasting is used by all the OR respondents but only by 50% to 70% of the members of the other employee groups.

In the Private sector list there are eight methods not used 30% or more by the OR group, namely, Non-linear estimation, Survey sampling, Categorical data analysis, Queueing theory, Multidimensional scaling, Reliability/Life testing, Conjoint analysis and Indices. There are four methods not used 30% or more by the State and semi-state group, namely Queueing theory, Reliability/Life testing, Conjoint analysis and Indices.

All the methods used by the State and semi-state employees are also used by University statisticians, but the latter use five additional methods over and above the State group, namely, Reliability/Life testing, Non-linear programming, Bioassay, Queueing and Decision theory

CHAPTER 6. COMPARISONS BETWEEN THE VARIOUS EMPLOYEE GROUPS

This chapter consists of statistical analyses of the results of Questionnaires No 1 and 2, which were presented in Chapters 4 and 5, together with additional information on academic qualifications. Comparisons are made between the academic qualifications of the different employee groups as well as between usages of statistical methods.

6.1. Comparison of academic qualifications

6.1.1. Source of information

In 1984 the S A Statistical Association published a Directory of members of SASA, based on information obtained from questionnaires sent out in October, 1981, and June, 1982. The Directory contains data on the academic qualifications of members and their place of employment.

Questionnaire No. 2 sent out in January, 1985, as part of this study, to members of SASA and the OR Society (See Appendix B) forms a second source of information on the qualifications and place of employment of statisticians. The two sources are not independent of each other because statisticians appear on both lists. Because of the time difference some members had obtained higher degrees in the interval between samples.

As explained in Section 2.3, there was some ambiguity about Questions 9.1 and 9.2 of Questionnaire No 1, which resulted in the data on the numbers of staff with various qualifications to be inexact. Therefore the information and qualifications from Questionnaire No 1 cannot be used in the analyses carried out in this chapter.

Tables 6.1. and 6.2. show the profile of academic qualifications for different employee groups, extracted from the above two sources of information.

TABLE 6.1. Academic qualifications extracted from SASA Directory (as percentages with actual frequencies in brackets).

Employee groups	Bachelor	Honours	Masters	Doctor	Unspecified	Total
University	3 (5)	10 (15)	27 (39)	54 (79)	6 (9)	100 (147)
State/semi-state	2 (1)	18 (9)	32 (16)	40 (20)	8 (4)	100 (50)
Private sector	0 (0)	39 (9)	43 (10)	18 (4)	0 (0)	100 (23)
Sub Total	3 (6)	15 (33)	29 (65)	47 (103)	6 (13)	100 (220)
Not specified	11 (4)	11 (4)	23 (8)	20 (7)	34 (12)	99 (35)
TOTAL	4 (10)	14 (37)	29 (73)	43 (110)	10 (25)	100 (255)

TABLE 6.2. Academic qualifications from Questionnaire No. 2 (as percentages with actual frequencies in brackets).

Employee groups	Bachelor	Honours	Masters	Doctor	Unspecified	Total
SASA - Univ.	8 (2)	8 (2)	12 (3)	72 (18)	0 (0)	100 (25)
SASA - State	21 (5)	16 (4)	21 (5)	42 (10)	0 (0)	100 (24)
SASA - Private	20 (3)	6 (1)	40 (6)	27 (4)	7 (1)	100 (15)
SASA -Total	16 (10)	11 (7)	22 (14)	50 (32)	1 (1)	100 (64)
OR Society	33 (6)	11 (2)	39 (7)	17 (3)	0 (0)	100 (18)
TOTAL SASA + OR	19 (16)	11 (9)	26 (21)	43 (35)	1 (1)	100 (82)

6.1.2. Method of statistical analysis

The frequencies for the four degrees from Tables 6.1. and 6.2. can be tabulated in a series of 2 x 4 contingency tables for each pair of groups as shown in Table 6.3. The frequencies for the unspecified class are not included in case they bias the analysis.

TABLE 6.3. Example of a contingency table for qualification profiles (actual frequencies in each cell).

Employee group	Bach.	Hons.	Mast.	Doct.	Group Totals
Directory: University	5	15	39	79	138
Directory: State	1	9	16	20	46
Degree totals	6	24	55	99	184

The chi-squared test can be used to test whether the ratio of Bachelor : Honours : Masters : Doctors is similar in the two groups. This is a test of the independence of the two categories in the contingency table (viz. groups and academic qualifications). Because the samples from the Directory and Questionnaire No.2 contain common SASA members, it is not valid to use the Chi-squared test to compare the two University subgroups or the two State or two Private subgroups.

Care has to be taken with contributions to X^2 which are based on a small expected frequency. Opinions differ on how small the expected frequency can be allowed to be, and, as stated by Rayner (1969:295), these opinions do not seem to be very solidly backed by actual research. A widely quoted rule is that the expected frequency should not be less than 5, but Cochran (1936) has given alternatives. For contingency tables Cochran recommends that a minimum expected cell frequency of 1 can be permitted, provided the number of such cells is not too great (e.g. not more than 1 in every 5 cells). (See Rayner (1969:313)). This rule will be used here, with the proviso that large contributions from cells with small expected frequencies will be treated with caution.

6.1.3. Comparison of qualification profiles for different employee groups

Note all chi-squared values in this section have 3 degrees of freedom. The X^2 values required for significance at the 5%, 1% and 0,1% significance levels with 3 degrees of freedom are 7,82; 11,35 and 16,27, respectively.

(a) SASA versus OR Society

The ratios of the four degrees, expressed as a percentage of the total frequency of all four degrees, are shown in Table 6.4.; but the X^2 values are calculated from actual frequencies, and not percentages.

TABLE 6.4. Ratios of Bachelors : Honours : Masters : Doctors (expressed as percentage of total frequency of all 4 degrees).

	Bach	Hons	Mast	Doc
SASA Directory (207 members)	3	16	31	50
SASA from Quest. No. 2 (63 members)	15	11	22	51
OR Soc from Quest. No. 2 (18 members)	33	11	39	17

The ratios for the 2 SASA groups, viz. from Directory and Questionnaire No. 2 are significantly different at the 1% level (X^2 = 12,04). The main cause can be seen in Table 6.4., as the high proportion of Bachelors degrees in the Questionnaire sample. It is not possible to say whether this difference is caused by non-response to the questionnaire by the higher qualified personnel of SASA, or whether the profile of SASA has changed between 1982 and 1985.

The ratio for the OR Society is significantly different from SASA Directory at the 1% level (X^2 = 25,82) but just fails to show up significantly different from SASA Questionnaire (X^2 = 7,43). The causes for the difference can be seen in Table 6.4. as the low proportion of PhD's and the high proportion of Bachelors degrees in the OR Society.

(b) <u>University versus State and semi-state versus Private sector</u>

The three subdivisions of the SASA can be compared with each other. The ratios of the four degrees, expressed as percentages of the total frequency of all four degrees are shown in Table 6.5.; but the X^2 values are calculated from actual frequencies, not percentages.

TABLE 6.5. Ratios of Bachelors : Honours : Masters : Doctors (expressed as percentage of total frequency of all 4 degrees

	Bach	Hons	Mast	Doc
SASA Directory :				
University (138 members)	4	11	28	57
State & semi-state (46 members)	2	20	35	43
Private sector (23 members)	0	39	44	17
SASA from Questionnaire No. 2 :				
University (25 members)	8	8	12	72
State & semi state (24 members)	21	16	21	42
Private sector (14 members)	21	7	43	29
OR Soc from Quest. No.2 (18 members) :	33	11	39	17

The ratio in the University group does not differ significantly from the ratio in the State and semi-state group for either the data from the SASA Directory ($X^2 = 3{,}93$) or Questionnaire No. 2 ($X^2 = 4{,}72$). The State and semi-state group does not differ significantly from the Private sector for both the SASA Directory ($X^2 = 6{,}06$) and Questionnaire No. 2 ($X^2 = 2{,}50$).

However, the University ratio does differ significantly from the Private sector for both the SASA Directory, at the 0,1% level (X^2 = 18,97) and Questionnaire No. 2, at the 5% level (X^2 = 7,97). The reason for this can be seen in Table 6.5. as the high proportion of Ph.D's in the University group. Conversely there are higher proportions of Honour's and Master's degrees in the Private sector for Directory, and higher proportions of Bachelor's and Master's degrees in the Questionnaire sample.

The OR Society is significantly different at the 5% level from the Private sector of the SASA Directory (X^2=10,68), but is not significantly different from the Private sector of SASA from Questionnaire No. 2 (X^2=1,07). As can be seen in Table 6.5., the significance is caused by the difference in Bachelor degrees, rather than a difference in Ph.D's.

The OR Society is significantly different at the 1% level from both University subgroups (X^2 = 25,53 and 13,53) and from the Directory State subgroup (X^2 = 14,67), but is not significantly different from the Questionnaire No. 2 State subgroup. The reasons for the significant differences are the PhD discrepancies. The State subgroups do not both show up as significant because of the difference in Bachelor degrees.

6.1.4. Conclusions

Within the SASA there were no significant differences between the University subgroups and the State and semi-state subgroups. Similarly the State and Private sector subgroups were not significantly different. There were significant differences between the University and Private sector subgroups, caused by the much higher proportion of PhD's in the University subgroup than in the Private sector subgroup (Average over both Directory and Questionnaire No. 2 of 60% PhD's (97 out of 163) for University to 22% for PhD's (8 out of 37), for Private sector).

It is clear that the OR Society has a different profile of qualifications than the SASA. As can be seen in Table 6.4., the former have only 17% of members (3 out of 18) with PhD's compared to the 50% of the SASA Directory (103 out of 207) and 51% of SASA from Questionnaire No. 2 (32 out of 63).

The OR Society is similar in profile to the Private sector with respect to PhD's but not Bachelor degrees. It is the high proportion of PhD's in the University and State subgroups that causes the overall significant difference between the OR Society and the SASA.

6.2. Frequencies of usage of statistical methods

Frequencies of usage in each employee group are shown in Tables 6.6. and 6.7.

TABLE 6.6. Methods used Routinely : Frequency of employees using each method

Statistical Methods Used Routinely	Questionnaire No.1			Questionnaire No.2				
				SA Statistical Association				Oper. Res. Soc. ex 19
	Stock Exch. ex 28	UCT Bus. Sch. ex 14	Busin. Group ex 42	Univ. ex 25	State ex 25	Priv. ex 15	Total ex 65	
Basic statistics	22	9	31	22	25	14	61	13
Graphical display/Data summary	24	8	32	19	23	14	56	14
Regression	11	6	17	23	19	10	52	10
Correlation	7	4	11	19	17	9	45	9
Anova	11	7	18	21	20	6	47	7
Design of experiments	2	0	2	13	10	5	28	4
Variance component estimation	2	0	2	4	2	1	7	2
Bioassay	1	0	1	2	2	0	4	0
Chi-squared tests	2	1	3	22	17	6	45	3
Quality control	11	5	16	7	2	6	15	3
Non-linear estimation	1	1	2	7	8	0	15	0
Non-parametrics	0	0	0	15	11	1	27	0
Ranking/Paired comparisons	3	0	3	9	11	1	21	1
Multivariate analysis	4	1	5	13	16	7	36	4
Logistic regression	1	1	2	8	10	0	18	0
Linear programming	6	4	10	6	5	2	13	3
Non-linear programming	1	1	2	2	0	1	3	2
Dynamic program	1	0	1	3	1	0	4	1
Inventory control	11	3	14	4	0	2	6	2
Simulation	8	2	10	11	8	3	22	7
Reliability/Life testing	3	1	4	3	4	2	9	1
Time series	9	8	17	13	4	7	24	6
Forecasting	17	9	26	11	6	7	24	7
Survey sampling	6	4	10	11	13	5	29	2
Categorical data analysis	0	2	2	10	14	2	26	1
Queueing theory	2	0	2	8	0	0	8	2
Factor analysis	0	0	0	8	9	4	21	0
Decision theory	2	1	3	6	2	0	8	0
Network models	8	0	8	4	1	2	7	5
Multidimensional scaling	1	0	1	5	6	1	12	0
Conjoint analysis	0	1	1	1	1	1	3	0
Indices	11	0	11	5	3	3	11	2
Discriminant analysis	2	0	2	10	11	2	23	1
Total Frequency	190	79	269	325	281	124	730	112

TABLE 6.7. Methods used Overall (i.e. Routinely + Occasionally): Frequency of employees using each method

Statistical Methods Used Routinely	Questionnaire No.1			Questionnaire No.2				
				SA Statistical Association				Oper. Res. Soc. ex 19
	Stock Exch. ex 28	UCT Bus. Sch. ex 14	Busin. Group ex 42	Univ. ex 25	State ex 25	Priv. ex 15	Total ex 65	
Basic Statistics	24	12	36	24	25	15	64	18
Graphical display/Data summary	25	12	37	24	24	15	63	17
Regression	19	10	29	24	23	14	61	17
Correlation	15	7	22	24	22	13	59	14
Anova	18	7	25	24	23	13	60	14
Design of experiments	8	0	8	22	18	9	49	11
Variance component estimation	3	0	3	14	9	3	26	4
Bioassay	1	1	2	8	5	1	14	1
Chi-squared tests	10	1	11	23	21	10	54	12
Quality control	16	8	24	12	5	7	24	6
Non-linear estimation	8	2	10	20	16	8	44	4
Non-parametrics	4	1	5	22	21	8	51	9
Ranking/Paired comparisons	7	1	8	15	15	5	35	7
Multivariance analysis	9	3	12	23	19	12	54	12
Logistic regression	2	2	4	18	19	2	39	1
Linear programming	16	9	25	14	14	8	36	14
Non-linear programming	2	2	4	11	7	2	20	6
Dynamic program	4	0	4	6	1	2	9	9
Inventory Control	14	5	19	4	1	2	7	6
Simulation	15	8	23	21	16	9	46	15
Reliability/Life testing	6	2	8	7	6	5	18	4
Time Series	14	10	24	17	16	11	44	12
Forecasting	23	11	34	13	14	11	38	18
Survey sampling	11	4	15	14	17	8	39	5
Categorical data analysis	1	2	3	14	16	6	36	1
Queueing theory	4	1	5	8	2	6	16	4
Factor analysis	4	0	4	19	13	10	42	7
Decision theory	5	2	7	8	3	2	13	4
Network models	14	5	19	6	3	4	13	12
Multidimensional scaling	4	1	5	12	12	6	30	2
Conjoint anal.	1	1	2	4	1	5	10	0
Indices	12	1	13	7	5	5	17	4
Discriminant analysis	4	0	4	20	18	9	47	7
Total Frequency	323	131	454	502	430	246	1178	277

6.3. Method of statistical analysis

The frequencies for the thirty three statistical methods from Tables 6.6 and 6.7 for any pair of employee groups can be set out in a series of 2x2 contingency tables of the following type:

	Group 1	Group 2
Use Method		
Do not use method		

There are thirty three of these contingency tables for each of the eleven pairs of employee groups which are of interest, ie 363 tables in all. On each table a chi-squared test can be carried out to test whether the usages of the method by the two employee groups differ significantly. The chi-squared test has one degree of freedom and is carried out by means of the following formula, which includes Yates's correction for continuity:

X^2 = (Difference of cross-products taken positively minus half the total)^2 divided by the product of the marginal totals.

It was thought unwise to use the 5% level of significance with so many consecutive tests, because, on average, 1 in 20 tests would have been wrongly declared significant. The tests were therefore made at the 1% and 0,1% significance levels. The chi-squared values for significance at the 1% and 0,1% levels for one degree of freedom are 6,64 and 10,83, respectively.

As stated in Section 6.1.2., care must be taken when expected cell frequencies fall below 5. When this happens with these 2x2 contingency tables Fisher's exact probability test is carried out.

6.4. Comparisons of statistical method usage for different pairs of employee groups

6.4.1. Stock Exchange versus UCT Business School (from Questionnaire No 1)

Routine usage

There are no significant differences at the 1% level between the usages by the Stock Exchange and the UCT Business School groups of any of the thirty three statistical methods.

Routine + Occasional usage

Again there are no significant differences at the 1% level between usages of the thirty three methods by the two groups.

Conclusions

The two groups of respondents derived from the sample of Stock Exchange firms and the sample from UCT Business School appear to be homogeneous with respect to their usages of the thirty three statistical methods. It is therefore perfectly justifiable to pool the two groups to form an employee group which will be called the Business group.

6.4.2. Subgroups of the SASA (from Questionnaire No 2)

(a) University versus State and semi-state institutions

Routine usage

The difference in usage of Queueing theory between the University and State groups is significant at the 1% level (Fisher's exact probability = 0,002). The usage ratio is 32:0% for University:State.

Routine + Occasional usage

There are no significant differences at the 1% level between the usages of the thirty three methods.

(b) University versus Private sector

Routine usage

Two methods are significantly different in usage at the 1% level, namely Chi-squared contingency tests (Fisher's exact probability = 0,002. Usage ratio is 88:40% for University:Private sector) and Non-parametric tests (Chi-squared = 9,00. Usage ratio is 60:7% for University:Private sector).

Routine + Occasional usage

The difference in usage of Logistic regression is significant at the 1% level (Chi-squared = 10,67. Usage ratio is 72:13% for University:Private sector).

(c) State and semi-state institutions versus Private sector

Routine usage

None of the methods were significant at the 1% level.

Routine + Occasional usage

The difference in usage for Logistic regression is significant at the 0,1% level. (Chi-squared = 12,36. Usage ratio is 76:13% for State:Private).

Conclusions:

(a) There is a significant difference between the usages of Queueing theory between the University and State members of the SASA. The usage ratios are 32% and 0%, respectively.

(b) Significant differences in method usage show up between members of the SASA from the Universities and the Private sector, for Chi-squared contingency tests, Non-parametric tests and Logistic regression. In all three cases the University usage is greater.

(c) There is a significant difference between State and semi-state bodies and the Private sector in the usage of Logistic regression, in the ratio 76% to 13%, respectively.

In view of the above significant differences in usage, the subgroups cannot be considered as completely homogenous. Each subgroup will be compared separately with the Operations Research Society and with Business, and it will appear that the University and State subgroups react similarly, but that both tend to react differently from the Private sector.

6.4.3. Operations Research Society versus the subgroups of the SASA (from Questionnaire No 2)

(a) OR Society versus University

There are a number of methods showing significant differences in usage between these two groups. These are shown in Table 6.8.

TABLE 6.8. Methods for which usage differs significantly between OR Society and University

Statistical methods significant at 1% (**) or 0,1% (***) levels	Routine usage		Routine + Occasional	
	X^2 or Fisher's P	Ratio of OR:Univ	X^2 or Fisher's P	Ratio of OR:Univ
Regression analysis	P=,004***	53:92%		
Analysis of variance	8,44 **	37:84%		
Chi-squared tests	20,09***	16:88%		
Non-linear estimation			12,85***	21:80%
Non-parametrics	14,73***	0:60%	6,72 **	47:88%
Logistic regression			16,97***	5:72%
Categorical data analysis			10,21***	5:56%
Discriminant analysis			6,76 **	37:80%
Forecasting			7,53 **	95:52%

(b) OR Society vs State + semi-state bodies

There are a number of methods showing significant differences in usage between these two groups. These are shown in Table 6.9.

TABLE 6.9. Methods for which usages differ significantly between OR Society and State and semi-state bodies

Statistical methods significant at 1% (**) or 0,1% (***) levels	Routine usage		Routine + Occasional	
	X^2 or Fisher's P	Ratio of OR:State	X^2 or Fisher's P	Ratio of OR:State
Analysis of variance	6,76 **	37:80%		
Chi-squared tests	9,86 **	16:68%		
Logistic regression	P=,001***	0:40%	19,03 ***	5:76%
Categorigal data analysis	10,21 **	5:56%	13,33 ***	5:64%
Dynamic programming			P=,001***	47:4%
Network models			10,40 **	63:12%

(c) OR Society versus Private sector

Routine analysis

There were no significant differences at the 1% level between usages by these two groups.

Routine + Occasional usage

The usages of the methods were not significantly different.

Conclusions

(a) The Operations Research Society shows significant differences in usage from the University members of the SASA. Methods that are used more frequently by University members are Regression analysis, Analysis of variance, Chi-squared tests, Non-linear estimation, Non-parametric methods, Logistic regression, Categorical data analysis and Discriminant analysis.
Forecasting is used more frequently by the OR Society,

(b) The OR Society shows significant differences from the State and semi-state members of the SASA. Methods that are used more frequently by the State and semi-state members are Analysis of variance, Chi-squared tests, Logistic regression and Categorical analysis. Methods used more frequently by the OR Society are Network models and Dynamic programming.

(c) The OR Society does not differ significantly from the Private sector.

6.4.4. Business (i.e. Stock Exchange + UCT Business School from Questionnaire No 1) versus the subgroups of the SASA (from Questionnaire No 2)

(a) Business versus University

There are a large number of methods for which the usages between these two groups differ significantly. These are shown in Table 6.10.

TABLE 6.10. Methods for which usages differ significantly between the Business group and University

Statistical methods significant at 1% (**) or 0,1% (***) levels	Routine usage		Routine + Occasional	
	X^2 or Fisher's P	Ratio of Bus:Univ	X^2 or Fisher's P	Ratio of Bus:Univ
Regression analysis	15,22 ***	40:92%		
Correlation analysis	13,77 ***	26:76%	11,90***	52:96%
Analysis of variance	9,28 **	43:84%	8,84 **	60:96%
Experimental design	17,50 ***	5:52%	27,41***	19:88%
Variance compt estmn			17,26***	7:56%
Bioassay			P=,004***	5:32%
Chi-squared tests	40,42 ***	7:88%	24,59***	26:92%
Non-linear estimation			17,80***	24:80%
Non-parametrics	29,11 ***	0:60%	34,62***	12:88%
Rank/Paired comparisons	P=,004***	7:36%	9,91 **	19:60%
Multivariate analysis	10,86 ***	12:52%	22,79***	29:92%
Logistic regression	P=,004***	5:32%	24,98***	10:72%
Non-linear programming			8,83 **	10:44%
Categorical data analysis	P=,001***	5:40%	17,26***	7:56%
Queueing theory	P=,004***	5:32%		
Factor analysis	P=,009***	0:32%	27,84***	10:76%
Multidimensional scaling			8,96 **	12:48%
Discriminant analysis	P=,001***	5:40%	30,86***	10:80%

(b) Business vs State + semi-state bodies

There are a large number of methods for which the usages between these two groups differ significantly. These are shown in Table 6.11.

TABLE 6.11. Methods for which usages differ significantly between the Business group and State + semi-state bodies

Statistical methods significant at 1% (**) or 0,1% (***) levels	Routine usage		Routine + Occasional	
	X^2 or Fisher's P	Ratio of Bus:State	X^2 or Fisher's P	Ratio of Bus:State
Correlation analysis	9,61 **	26:68%	7,31 **	52:88%
Analysis of variance	7,36 **	43:80%		
Experimental design	P=,001***	5:40%	16,34 ***	19:72%
Variance compt estmn			P=,004***	7:36%
Chi-squared tests	24,89***	7:68%	18,74 ***	26:84%
Non-linear estimation	P=,004***	5:32%	9,04 **	24:64%
Non-parametrics	P=,000***	0:44%	31,33 ***	12:84%
Rank/Paired comparisons	10,75***	7:44%	9,91 **	19:60%
Multivariate analysis	17,42***	12:64%	12,34 ***	29:76%
Logistic regression	P=,001***	5:40%	27,84 ***	10:76%
Categorical data analysis	19,90***	5:56%	22,22 ***	7:64%
Factor analysis	P=,000***	0:36%	12,77 ***	10:52%
Multidimensional scaling			8,96 **	12:48%
Discriminant analysis	13,02***	5:44%	24,98 ***	10:72%
Quality control			7,36 **	57:20%
Inventory control	8,61 **	33:0%	10,83 ***	45:4 %
Forecasting	7,57 **	62:24%		

(c) Business versus Private sector

There are a number of methods for which the usages between these two groups differ significantly. These are shown in Table 6.12.

TABLE 6.12. Methods for which the usages differ significantly between the Business group and Private sector

Statistical methods significant at 1% (**) or 0,1% (***) levels	Routine usage		Routine + Occasional	
	X^2 or Fisher's P	Ratio of Bus:Priv	X^2 or Fisher's P	Ratio of Bus:Priv
Experimental design			P=,005***	19:60%
Chi-squared tests	P=,006***	7:40%		
Non-parametrics			P=,002***	12:53%
Multivariate analysis			9,98 **	29:80%
Categorical data analysis			P=,007***	7:40%
Factor analysis	P=,003***	0:27%	P=,000***	10:67%
Discriminant analysis			P=,000***	10:60%

Conclusions

(a) The Business sample (ie. Stock Exchange and UCT Business School from Questionnaire No 1) was significantly different from all three subgroups of the SASA (Questionnaire No 2). With the University subgroup the methods with significantly different usages are Regression analysis, Correlation analysis, Analysis of variance, Experimental design, Variance component estimation. Bioassay, Chi-squared tests, Non-linear estimation, Non-parametric methods, Ranking/Paired comparisons, Multivariate analysis, Logistic regression, Non-linear programming, Categorical data analysis, Queueing theory, Factor analysis, Multidimensional scaling and Discriminant analysis, which were used more frequently by the University members.

(b) With the State + semi-state subgroup the significant methods were Correlation analysis, Analysis of variance, Experimental design, Variance component estimation, Chi-squared tests, Non-linear estimation, Non-parametric methods, Ranking/Paired comparisons, Multivariate analysis, Logistic regression, Categorical data analysis, Factor analysis, Multidimensional scaling and Discriminant analysis, which were used more frequently by the State subgroup. Quality control and Inventory control were used more by Business.

(c) With the Private sector subgroup the significant methods were Experimental design, Chi-squared tests, Non-parametric methods, Multivariate analysis, Categorical data analysis, Factor analysis and Discriminant analysis, which were used more frequently by the Private sector.

6.4.5. <u>Business (i.e. Stock Exchange + UCT Business School from Questionnaire No 1) versus the Operations Research Society (from Questionnaire No 2)</u>

<u>Routine usage</u>

None of the methods showed a significant difference in usage between these two groups.

<u>Routine + Occasional usage</u>

Three methods showed significant differences in usage at the 1% level, namely, <u>Experimental design</u> (X^2 = 7,48 1% level. Usage ratio, Business:OR Society 19:58%), <u>Non-parametric methods</u> (Fisher's P=0,004. Ratio 12:47%), and <u>Dynamic programming</u> (Fisher's P=0,002. Ratio 10:47%).

<u>Conclusion</u>

The three methods which showed significant differences were Experimental design, Non-parametric methods and Dynamic programming, all of which were used more frequently by the OR Society

6.5. <u>Overall conclusions on usages of statistical methods</u>

The usages of statistical methods by the Stock Exchange firms and the UCT Business graduates are similar. Hence pooling of the two samples to form a Business group is justified.

Of the three subgroups of the SASA, the University showed a greater usage of Queueing theory than the State + semi-state group, and a greater usage of Chi-squared tests, Non-parametric tests and Logistic regression than the Private sector. The State group showed a greater usage of Logistic regression than the Private sector. (Note: All these and subsequent differences are significant at the 1% level.)

There were no significant differences between the OR Society and the Private sector. The OR Society used Forecasting more than the University group, but used the following methods less: Regression, Anova, Chi-squared tests, Non-linear estimation, Non-parametric tests, Logistic regression, Categorical data analysis and Discriminant analysis. The OR Society used Dynamic programming and Network models more than the State group, but used the following methods less: Anova, Chi-squared tests, Logistic regression and Catagorical data analysis.

The Business groups used Experimental design, Non-parametric tests and Dynamic programming less than the OR Society. The Business group used Quality control, Inventory control and Forecasting more than the State group. There were eighteen methods which the three SASA groups used significantly more than the Business group. These are shown in Tables 6.10 (No new methods occur in Tables 6.11 and 6.12).

It will be suggested in Section 10.2.2 that these methods, which show significant differences, classify themselves and the employee groups into a number of clear-cut categories.

CHAPTER 7. CORRESPONDENCE BETWEEN UNIVERSITY TEACHING AND THE USAGES OF STATISTICAL METHODS BY VARIOUS EMPLOYEE GROUPS

This chapter consists of an analysis of the correlation between the teaching emphasis on statistical methods and their use in consulting work.

7.1. Questionnaire on university teaching

Underhill (1982) conducted a postal survey of SA universities to find out which of thirty three statistical methods were taught in the three undergraduate years and in the Honours year. Thirteen responses to the questionnaire were received. The data are tabulated in Table 3.2. in Section 3.4. Most methods are taught in more than one year and a quantitative measure of the teaching emphasis placed by educators on each method has to take account of this.

The methods taught have to be ranked in terms of the relative importance placed on each by the teacning profession. To do this a weighted total has been calculated for each method of the frequencies of universities teaching the method in each year. The weights used are 1;2;3;4, so that, for example, the weighted total for the method of Survey sampling is:

$$(1\times4) + (2\times4) + (3\times5) + (4\times10) = 67,$$

where, from Table 3.2, four universities teach this method in the first year, four in the second year, five in the third year and ten in the Honours year. As only thirteen universities are involved in this study, it is obvious that most are teaching Sampling at both elementary and advanced levels.

The weights were chosen to reflect the teaching emphasis, so that the quantity and complexity of a course in the third year is considered three times as great as that for a first year course. As far as the ranking of the waighted totals are concerned, it makes very little difference whether the weights 1;2;3;4 or 1;2;2;4 are used. A second weighted total of the first three years was calculated with weights 1;2;3;0 to reflect the Bachelor's degree. Both weighted totals were than ranked in ascending order of magnitude.

The weighted totals in ranked order are shown in Table 7.1.

TABLE 7.1. Weighted totals of frequencies of universities teaching a given method in each year, in ranked order

Bachelors (Weights 1;2;3;0)		Honours (Weights 1;2;3;4)	
Statistical Method	Total	Statistical Method	Total
Regression analysis	66	Regression analysis	98
Correlation analysis	52	Analysis of variance	88
Analysis of variance	52	Correlation analysis	80
Ranking & Pair data anal	41	Linear programming	75
Linear programming	39	Multivariance analysis	74
Chi-sq cont tables	37	Non-parametric methods	73
Non-parametric methods	37	Queueing theory	69
Multivariate analysis	30	Survey sampling	67
Queueing theory	29	Time series analysis	66
Quality cont & accept sampl	28	Ranking & pair data anal	61
Survey sampling	27	Probability modelling	53
Probability modelling	25	Inventory control	53
Variance comp estim	24	Decision theory	51
Decision theory	23	Chi-sq cont tables	49
Time series analysis	22	Forecasting	49
Network models	22	Network models	46
Categorical data analysis	19	Quality cont & accept sampl	44
Inventory Control	17	Non-linear programming	44
Simulation	15	Design of experiments	43
Basic statistics	13	Simulation	43
Graphical display	13	Var comp estim	40
Indices	13	Dynamic programming	39
Non-linear programming	12	Categorical data analysis	39
Forecasting	9	Factor analysis	39
Design of experiments	7	Non-linear estimation	27
Reliab & Life Testing	6	Reliab & Life Testing	26
Bioassay	5	Multidimensional scaling	16
Factor analysis	3	Indices	13
Non-linear estimation	3	Basic statistics	13
Dynamic programming	3	Graphical display	13
Multidimensional scaling	0	Bioassay	9
Conjoint analysis	0	Econometrics	8
Econometrics	0	Conjoint analysis	0

Spearman's correlation coefficient can be used to quantify the relationship between the rankings in Table 7.1. with the rankings of the method usages of the different employee groups which can be obtained from Tables 6.6 and 6.7. A large correlation coefficient will indicate that the teaching emphasis matches the usage by the particular employee group. For the purpose of this comparison Basic statistics and Graphical display have been omitted as they are the most used, but require little effort in teaching. Hence their respective ranks will differ enormously. Econometrics and Discriminant analysis have been omitted because they are not common to Underhill's questionnaire and Questionnaires No 1 and 2. For this reason Probability modelling and Logistic regression have also been left out. This leaves twenty nine methods for comparing teaching effort with usage of the methods.

When two or more of the statistical methods have the same teaching index or the same method usag , the rank allocated to them is the mean of the two or more ranks. Thus, for example, in Table 7.1 the weighted totals for the Bachelor years are 52 for both Correlation analysis and Analysis of variance. Therefore the rank given to them in Table 7.2 is (2+3)/2 = 2,5. Similarly, for example in Table 7.2. usages in the Private sector for Linear programming, Inventory control, Reliability/Life testing, Categorical data analysis and Network models were all equal to 2 out of 15 respondents. The rank positions of these methods (in the order of writing down the ranked methods after leaving out four of the methods mentioned above) were 14th, 15th, 16th, 17th and 18th, so all were allocated the mean rank of (14 + 15 + 16 + 17 + 18)/4 = 16.

7.2. Analysis of ranks for routine usage

The ranks for routine usage are obtained from Table 6.6. and are tabulated in Table 7.2. The corresponding values of Spearman's correlation coefficient are shown in Table 7.3.

TABLE 7.2. Routine usage : Ranks for teaching emphasis and method usage for each statistical method

Statistical Method	Rank of Teaching index		Rank of routine usage				
			S.A. Stat. Assoc.			O.R.	
	Bach.	Hons.	Univ.	State	Priv.	Soc.	Bus.
Regression	1	1	1	2	1	1	3,5
Correlation	2,5	3	4	3,5	2	2	7,5
Anova	2,5	2	3	1	7	4	2
Design of expts	22	18	7	11	9,5	8,5	20,5
Var compt estmn	12	20	23	21,5	21,5	15,5	20,5
Bioassay	24	28	27,5	21,5	27	26	25,5
Chi-squared	6,5	13	2	3,5	7	11	16
Quality control	10	16	16,5	21,5	7	11	5
Non-linear estm	26	24	16,5	13,5	27	26	20,5
Non-parametrics	6,5	6	5	9,5	21,5	26	28,5
Rank/Pair comp	4	10	13	9,5	21,5	20,5	16
Multivar anal	8	5	7	5	4	8,5	13
Linear prog	5	4	18,5	17	16	11	10
Non-linear prog	20	17	27,5	28	21,5	15,5	20,5
Dynamic prog	26	21	25,5	25	27	20,5	25.5
Inventory cntr	17	11	23	28	16	15,5	6
Simulation	18	19	10	13,5	12,5	4	[illegible]
Rel/Life testing	23	25	25,5	18,5	16	20,5	14
Time series	14,5	9	7	18,5	4	6	3,5
Forecasting	21	14	10	15,5	4	4	1
Survey sampling	11	8	10	7	9,5	15,5	10
Cat data anal	16	22	12	6	16	20,5	20,5
Queueing theory	9	7	14,5	28	27	15,5	20,5
Factor anal	26	23	14,5	12	11	26	28,5
Decision theory	13	12	18,5	21,5	27	26	16
Network models	14,5	15	23	25	16	7	12
Multidim scal	28,5	26	20,5	15,5	21,5	26	25,5
Conjoint anal	28,5	29	29	25	21,5	26	25,5
Indices	19	27	20,5	20	12,5	15,5	7,5

TABLE 7.3. Values of Spearman's correlation coefficient. (Routine usage)

Employee Group	Bachelors degree teaching	Honours degree teaching
SASA : University Group	0,635	0,670
SASA : State & semi-state	0,502	0,438
SASA : Private sector	0,427	0,466
Operations Research Society	0,542	0,596
Business (Stock Exchange and UCT Business School)	0,496	0,527

All the correlation coefficients are significant at the 5% level (r = 0,364 and 0,478 are required for significance at the 5% and 1% significance levels respectively). It is more important to consider the relative magnitude of the correlation coefficients. The best agreement between teaching emphasis and method usage is with the University group. None of the differences in rank equal or exceed 16. (Differences in rank of 16 or more occur about 5% of the time, so that 16 is taken as a cut-off value for highlighting statistical methods which have noteworthy differences in rank.)

The Operations Research group shows the next best agreement. Methods with noteworthy differences in rank (shown in brackets) are Non-parametric methods (20) for Honours and Non-parametrics (19.5), Forecasting (17) and Ranking/paired comparisons (16.5) for Bachelors.

The Business group is third in agreement with the following noteworthy differences in rank : Non-parametrics (22.5) and Indices (19.5) for Honours and Non-parametrics (22) and Forecasting (20) for Bachelors.

The State & semi-state group is fourth in agreement with the following noteworthy differences in rank : Queueing (21), Inventory control (17) and Categorical analysis (16) for Honours and Queueing (19) for Bachelors.

The correlation is weakest for the Private sector, with the following noteworthy differences in rank: Queueing (20) for Honours and Queueing (18), Ranking/Paired data (17.5) and Forecasting (17) for Bachelors.

7.3. Analysis of ranks for routine and occasional usage

The ranks for routine + occasional usage are obtained from Table 6.7. and are tabulated in Table 7.4. The corresponding values of Spearman's correlation coefficient are shown in Table 7.5.

TABLE 7.4. Routine and Occasional usage : Ranks for teaching emphasis and method usage for each statistical method

Statistical Method	Rank of teaching index		Rank of routine + occas. usage				
			S.A. Stat. Assoc.			O.R.	
	Bach.	Hons.	Univ.	State	Priv.	Soc.	Bus.
Regression	1	1	2	1,5	1	2	2
Correlation	2,5	3	2	3	2,5	5	8
Anova	2,5	2	2	1,5	2,5	5	3,5
Design of expts	22	18	6,5	7	9,5	11	17
Var comp estmn	12	20	14.5	18	24	22.5	26.5
Bioassay	24	28	23	22	29	27,5	28,5
Chi-squared	6,5	13	4,5	4,5	7,5	8,5	14
Quality control	10	16	19	22	15	17	5,5
Non-linear estm	26	24	9	10,5	12,5	22,5	15
Non-parametrics	6,5	6	6,5	4,5	12,5	12,5	21
Rank/Pair comp	4	10	12	13	20,5	14,5	17
Multivar anal	8	5	4,5	6	4	8,5	13
Linear prog	5	4	14,5	14,5	12,5	5	3,5
Non-linear prog	20	17	21	19	26,5	17	24
Dynamic prog	26	21	26,5	28	26,5	12,5	24
Inventory cntrol	17	11	28,5	28	26,5	17	9,5
Simulation	18	19	8	10,5	9,5	3	7
Rel/Life testing	23	25	19	20	20,5	22,5	17
Time series	14,5	9	11	10,5	5,5	8,5	5,5
Forecasting	21	14	17	14,5	5,5	1	1
Survey sampling	11	8	14,5	8	12,5	19	11
Cat data anal	16	22	14,5	10,5	17	27,5	26,5
Queueing theory	9	7	23	26	17	22,5	21
Factor analysis	26	23	10	16	7,5	14,5	24
Decision theory	13	12	23	24,5	26,5	22,5	19
Network models	14,5	15	26,5	24,5	23	8,5	9,5
Multidimen scal	28,5	26	19	17	17	26	21
Conjoint anal	28,5	29	28.5	28	20,5	29	28,5
Indices	19	27	25	22	20,5	22,5	12

TABLE 7.5. Values of Spearman's correlation coefficient (Routine and Occasional usage)

Employee Group	Bachelors degree teaching	Honours degree teaching
SASA : University group	0,531	0,516
SASA : State & semi-state	0,536	0,529
SASA : Private sector	0,431	0,535
Operations Research Society	0,512	0,669
Business (Stock Exchange & UCT Business School)	0,500	0,604

All the correlation coefficients are significant at the 5% level. Again it is the relative magnitudes of the correlation coefficients that are of greater importance than their significance.

Greatest agreement between teaching and method usage is achieved with the Operations Research group. Methods with noteworthy differences in rank are Simulation (16) for Honours and Forecasting (20) for Bachelors.

The next three groups have similar correlations. For the Business group the only noteworthy method is Forecasting (20) for Bachelors. For the State & semi-state group the noteworthy methods are Queueing (19) and Inventory control (17) for Honours and Queueing (17) for Bachelors.

For the University group the noteworthy methods are Queueing (16) and Inventory control (17.5) for Honours and Non-linear estimation (17) and Factor analysis (16) for Bachelors.

The Private sector group shows least agreement between teaching and usage. For Honours there are no methods with rank difference 16 or more, while for Bachelors noteworthy methods are Factor analysis (18.5) and Ranking/Paired data analysis (16.5).

7.4. Discussion

The Spearman's correlation coefficients in Tables 7.3. and 7.5. vary from 0,427 to 0,669, so that the correlation between teaching emphasis and method usage is only moderate in strength. It must be stressed that the weighted total of frequencies used to quantify teaching emphasis is not a sensitive measure. On the basis of the information from the correlation coefficients it must be concluded that the methods that are used are given reasonable emphasis in the teaching programme, although some, which are mentioned below, seem to be somewhat out of line.

Certain methods have differences of rank which equal or exceed the cut-off value of 16, some more often than others. These methods are listed in Table 7.6. together with the mean of the differences in ranks and the mean of the actual ranks, for those methods whose rank difference equalled or exceeded 16.

TABLE 7.6. Statistical methods with noteworthy differences in rank (> or = 16)

Statistical method with rank difference > or = 16	No. of times > or = 16	Mean rank diff.	Mean usage rank	Teaching rank Bach.	Teaching rank Hons.
Queueing theory	7	18,6	26,4	9	7
Non-parametrics	4	21,0	27,3	6.5	6
Ranking/Paired data	3	16,8	20,8	4	10
Inventory control	3	17,2	28,2	17	11
Forecasting	5	18,8	2,2	21	14
Factor analysis	2	17,3	8,8	26	23
Simulation	1	16	3	18	19
Non-linear estmn	1	17	9	26	24
Categorical data anal	1	16	6	16	22
Indices	1	19,5	7,5	19	27

It can be seen in Table 7.6 that there are four methods which are given greater emphasis in teaching than seems to be warranted by their ranking in usage. These are Queueing theory, Non-parametric tests, Ranking/Paired data analysis and Inventory control. There are 6 methods whose usage appears to warrant greater teaching emphasis. These are Forecasting, Factor analysis, Simulation, Non-linear estimation, Categorical analysis and Indices.

In Table 7.7. the match between usage and teaching emphasis is examined for the individual employee groups separately.

TABLE 7.7. Statistical methods for which the usage does not match the teaching emphasis

Statistical method	Employee group for which the teaching emphasis appears to be :		
	too great	about right	too small
Queueing theory	All groups		
Non-parametrics	Priv, OR Soc, Business	University, State	
Rank/Paired data	Priv, OR Soc, Business	University, State	
Inventory control	Univ, State, Private	OR Society	Business
Forecasting		University, State	Priv, OR Soc, Business
Factor analysis		OR Society, Business	Univ, State, Private
Simulation			All groups
Non-linear estmn		OR Society	Univ, State, Priv, Bus.
Categ data anal	OR Society, Business	Private	University, State
Indices		University, State	Priv, OR Soc, Business

As can be seen in Table 7.7. the teaching emphasis on Non-parametrics and Ranking is about right for the University and State groups, while that for Inventory control is about right for the O.R. Society, but not enough for the Business group. For the remaining groups the emphasis on Queueing, Non-parametrics, Ranking and Inventory control is too great.

The teaching emphasis on Forecasting and Indices is about right for the University and State groups but too small for the other three groups. Factor analysis and Categorical analysis have too light a teaching emphasis for the University and State groups but are either too heavily emphasised or are about right for the other groups. Simulation has too little teaching emphasis for all groups, as does Non-linear estimation with the exception of the OR Society.

CHAPTER 8. OPINIONS AND SUGGESTIONS FROM PRACTISING STATISTICIANS

8.1. Introduction

One aspect of this study, as set out in Section 2.5. was to obtain opinions of statisticians in both public and private sectors on whether they considered that their training had equipped them adequately for their careers, and how they felt university training could be improved. Note that here we are dealing with the functional aspect of education, ie the development of skills for use in the profession.

To obtain information on this aspect the following question was included in Questionnaire No. 2, which was sent to members of the OR Society and the SASA:

> "Please comment on whether you felt your statistical training gave you the skills necessary for the adequate performance of your job(s). Suggestions for improved training would be appreciated."

There were nineteen replies to this question from those members of SASA employed in State or semi-state Institutions, twenty from the University group, eleven from the Private sector and fifteen from the OR Society.

As discussed in Section 2.6. it is extremely difficult to summarise the views of sixty five people, but many comments fall into certain broad categories. For completeness all of the comments are listed in Appendix E in the original form in which they were written on the questionnaires. The code number of the respondent, together with the employee group, is given in brackets. Occasionally the comments of one respondent expressed two ideas which had to be split and included in the text under two categories. In the text a note is made when only part of a comment is used or if the comment is translated from Afrikaans. Where necessary certain words are included in brackets so that the quote reads intelligibly.

8.2. Training considered adequate

There were only four of the sixty five respondents who gave unqualified approval to their training. However, three of them are employed in specialised fields, so the approval is more by omission than of a positive nature. The remaining sixty one all had reservations of some sort or suggestions for improvement.

The comments are as follows:

Comment 1: (My training) "gave me the necessary background for the performance of my job."

Comment 2 (Part only): "I work as an epidemiologist, not a data analyst. My training in statistics was more than adequate."

Comment 3: "As I am not employed as a Statistician I can't really comment. However, I feel that my training was adequate."

Comment 4: (My training was adequate) "as a lecturer in Statistics/Biometry."

8.3. Statistics courses supply a basic foundation

All the eleven respondents quoted below feel that university training supplies the necessary basic knowledge. Most think it is impossible to broaden the training due to the time constraint, and that the specialist skills [illegible] only be developed after graduation, usually by self-[illegible]. In other words they felt the training was as good as could be expected under the constraint of limited time. [illegible] it was not in itself adequate and is only completed after a period of training in the work situation.

Note that the first two respondents quoted below are very experienced statisticians of considerable standing in the profession in S.A.

Comment 5 (Translated): "It is impossible to train a statistician in four years. The University courses must give the necessary foundation - the practical training can never be completely and satisfactorily undertaken at university. A thorough two-year training after completion of a four-year university degree is necessary for that.

Comment 6: "I do not think any training can provide adequate skills over a short period of (at most) 5 years. You can only develop a feeling for what the subject is all about in formal training. The rest is experience and cooperation with colleagues and self-training: READ,READ,READ.....".

Comment 7: "The training supplied me with a very good basic 'tool-kit'. It has been up to me to learn the skills in the use of the 'tools' supplied and to acquire more specialised 'tools' for solving certain problems. Statistics is a VERY broad field and I do not see how an undergraduate course can improve on supplying a sound, basic 'tool-kit'".

Comment 8: "My training gave me the basic tools and the ability to assimilate more specialised techniques at a later stage".

Comment 9: "My statistics training gave me a useful background but my practical experience and application were self taught".

Comment 10: "In some areas, the statistical training has helped in my work. In others, the statistics required would not be adequately provided at the level of formal statistical education undertaken to date as many aspects are treated in depth only at post-graduate statistical levels. This applies especially to the 'specialist' techniques. Inclusion of 'advanced training' at undergraduate level would not really be appropriate in view of the broad spectrum of topics which need to be covered".

Comment 11: "The year of honours (i.e. fourth year) is essential as not enough practical problems are handled in the undergraduate course".

Comment 12: (The necessary skills) "only come with experience. Good theoretical background is necessary. (There is) too much emphasis in exams on proofs of theorems - students should be taught to think!"

Comment 13 (Translated): "Training initially inadequate but with continuing studies quite adequate".

Comment 14: "Necessary skills can only develop over time".

Comment 15: "I don't believe that even an Honours degree in Statistics is adequate training. One must gain practical experience before one can obtain a true understanding of the theory".

Comment by author:

These respondents tend to take a rather pessimistic view of possible improvements to university training, largely because of the time constraint. They suggest that the practical application of the theory taught over four or five years can only be learned during in-service training - a sort of housemanship period. It is hoped that this is not the true position, because undoubtedly many statisticians are dissatisfied with their training. In this dissertation an attempt will be made to recommend improvements.

8.4. Training considered too theoretical

(i) Comments from the literature

One comment from the ISI report mentioned in section 1.7.1 was

> Statistical education at university level was found to be too theoretical in nature. It failed to equip students for employment as practising statisticians by the State, industry and other organisations.

As mentioned in section 1.7.2, the Snee report suggested that students be exposed to

(a) Statistical theory and methodology useful in the solution of a wide range of real problems and

(b) A philosophy to use as a guide in applying these methodologies and in developing others as needed in solving problems.

(ii) Comments from South African statisticians

The most commonly voiced complaint about statistics courses, as taught at university, is that they are too theoretical in nature. At least fifty two of the comments (80%) referred to this, or to the need for a more practical approach to teaching or made suggestions for improving the practical training. The following comments specifically mentioned, or implied, that courses were too theoretical, but they did not suggest ways of improving this. Some of the other comments, which did contain suggestions for improvement, are set out in Section 8.5.

Comment 17: "My courses were too theoretical. They were techniques looking for problems to which they might possibly be applied".

Comment 18 (Part only): "My statistical training was mostly academic".

Comment 19 (Translated. Part only): (Training should have) "more practical applications and less theory".

Comment 20 (Translated): "The training is very theoretical. One only learns the practical applications in the work situatic ".

Comment 21 (Translated. Part only): "A theoretical academic training does not prepare one adequately for the work of a lecturer. Practical experience is most essential".

Comment 22 (Part only): "Undergraduate stuff" (ie training) "was too theoretical: M.Sc. was more practically orientated and therefore more use".

Comment 23: "My statistics training was geared to OR and most of it was self study. Hence YES it did provide relevant skills. But I think that courses along the lines of Stu Hunter's design of expts. and Gene Woolsey's Quick and Dirty Operations Research are good for getting a basic grasp quickly. I would call this training. I would call formal course work using Wilks, Feller, Scheffe, Draper & Smith, etc., education".

Comment 33 (Translated. Part only): "The university training is very theoretical and far removed from practice. The experience that I have gained in the one and a half years in my present position is of great value and has brought theory and practice together - which is a deficiency at university".

Comment 55 (Part only): "Training not practical enough".

8.5. Calls for more practical training with suggestions for achieving this

In this section there are thirty nine comments which basically call for more practical training, but which include suggestions on how improvements can be made. The comments have been subdivided into groups with suggestions of roughly the same type. Some of these suggestions are similar to certain of the recommendations made in the two major reports on the training of statisticians, namely, the ISI report (Duncan and Durbin, 1980) and the ASA report (Snee et al, 1980). Where this happens the subsection is introduced with the appropriate quote from these two reports.

One such comment from the ISI report is appropriate here, namely,

> In the light of the growing realisation of the role that statistics should play in the modern world and the demand for statisticians in the public and industrial sectors, the training of practically-oriented students was of the greatest importance.

8.5.1. Improved balance between theory and practice

(i) Comments from the literature

The first recommendation of the Snee committee was as follows:

Programmes for training industrial statisticians should contain a balance of theory and practice. The theory that is taught should be that which is needed to understand the statistical methodology useful in practice. Relevant and useful theory is also important because industrial statisticians often have to adapt existing methodology to fit the problem and sometimes develop new methodology.

(ii) Comments from South African statisticians

A most interesting and important comment came from a highly respected South African statistician, who has had experience as the head of a university statistics department, has worked for a semi-state body and who was one of the most active and sought after consultants in this country. His comment is as follows:

Comment 24: "If I were designing courses today I'd replace about one-third of the current curriculum with more nuts and bolts courses on data analysis including enforced usage of at least 2 program packages. Our theoretical coverage is of first class international standard, with actual applied statistics the area with the most room for improvement".

Other comments, which are in agreement with the call for balanced training made by the Snee report, are as follows:

Comment 25 (Part only): "An appropriate balance between theory and practice is required".

Comment 28 (Part only): "To improve training, MANY more practical examples and applications of statistics should be included with the theory".

Comment 29 (Translated): "Greater concentration on practical applications (required). More specific goal-orientated courses. Present training gives techniques, but not enough applied practical training".

Comment 31 (Translated. Part only): "For my work as statistical consultant there were too few applied courses up to Masters level".

Comment 36 (Translated. Part only): "Practical work can be given much greater emphasis, especially with an eye to a career in Statistics".

Comment 38 (Part only): "Only after a few years of working has the theory and the practical been linked together".

Comment 43: "I feel that more emphasis could have been placed on practical situations in order to make the theory more relevant and to equip one better for the working situation".

Comment 44 (Translated): "Greater emphasis should be placed on the practical application of statistical techniques at University".

Comment 45 (Translated): "My training definitely gave me the ability to do my work satisfactorily. A possible improvement is the greater emphasis that could possibly be placed on practical aspects for people who want to go into practice (in die praktyk wil gaan)".

Comment 47 (Translated): "Statistical training was necessary but more practical training can be given".

Comment 48 (Translated): "The Mathematical Statistics courses can definitely be directed more towards the practical (op die praktyk toespits)".

Comment 50: "Felt that not enough practical application of theory was given, but as this was during 1963-1966 and training has changed since then - hope this has been remedied".

Comment 51: "Need more practical applications during studies".

Comment 52: "Please more practical examples".

Comment 53: (My training) "basically did" (give me the necessary skills), "but the practical experience is far more important".

(iii) Suggested improvement: Balanced teaching from the University of Cape Town

In Section 1.6.2. mention was made of the aims and objectives of a teaching procedure used at the University of Cape Town and documented by Juritz (1986). This set out the eight stages of an observational study as follows:

Stage 1: Motivation of the study and formulation of the problem.

Stage 2: Sampling method, size, cost, etc.

Stage 3: Design of questionnaire and collection of data.

Stage 4: Computerisation of data.

Stage 5: Checking of data (including outliers).

Stage 6: Statistical analysis.

Stage 7: Interpretation of results.

Stage 8: Report on the analysis.

Juritz states that "statistical courses tend to place heaviest emphasis on Stage 6, the analysis, and to a slightly less extent to Stage 2, the sampling. Stage 4, the computerisation, is learned almost incidentally and practically no attention is given to Stages 3, 5, 7 and 8." Statisticians, when being consulted, have to give advice on Stages 2 and 3 and if they do not give advice on Stage 4 they "may have a time-consuming task getting the data read by the computer." It is essential, therefore, to expose students to Stages 2, 3 and 4. Juritz continues

> Stage 7, the interpretation of the results of our analysis, is another issue often avoided in our teaching and our literature. Yet, if we do not participate in this we may end up as mere technicians. Report writing, Stage 8, is an art in itself, where we must learn to be brief and to the point. If we do not

write a report, we are in danger of presenting the client with what to him may be a meaningless jumble of numbers. Worse still, without a concise statement of results, we may see over-extravagant claims being drawn from very slender evidence.

At the University of Cape Town they have, over the past ten years, evolved a method of exposing students to all stages of the statistical study. This takes the form of a weekly practical class in which a questionnaire is designed and analysed using the BMDP statistical package. Over a ten week period the students are taken through all Stages 2 to 8.

The similarities between the UCT teaching method and the project system of Pretoria University, discussed in section 8.5.3 (iii), are very obvious. Other universities could well consider whether room could be found in their curricula for a similar teaching experience for their students.

8.5.2. Genuine data sets

(i) Comments from the literature

The committee of the ISI recommended that education programmes should include the use of genuine data sets in examples. Similarly, the Snee committee stated that

Real problems and associated data are essential in illustrating the statistical methodology being taught. Sufficient background to the problem should be given so that the student can see how the statistical analysis contributed to the solution of the problem. This can also be accomplished, with varying degrees of success, by having students conduct their own experiments (Hunter, 1977) or criticising and analysing data reported in the subject matter literature. It is also appropriate to have a one-or-two-semester course on "Analysis and Interpretation of Data" which is totally based on real problems.

(ii) Comments from South African Statisticians

It is obvious from the following comments that statisticians in South Africa are also aware of this need for real data.

Comment 19 (Translated. Part only): "It does not help to learn about ideal conditions which you never meet in practice".

Comment [illegible] emphasis on analysis of real (not so well-behaved) data, even at undergraduate level, is required".

Comment 38 (Part only): "I think the training should contain more practical work, not practical work out of a textbook, but in a real situation".

Comment 41 (Part only): "If only we had done more practical work. All distributions are not normal in practice".

Comment 49 (Part only): "We had no practical training other than computing statistics for data sets for which the area and technique were specified".

Comment by a respondent to Questionnaire No.1: "In practice, life is never simple and we seem to spend more time 'correcting' data for its problems, eg colinearity, non-stationarity, small sample size, non-Gaussian, etc., than we do applying the analysis techniques! I feel that more emphasis in learning programmes should be put on getting the data into the 'right' form".

Comment by Juritz (1986):

Juritz is of the opinion that text-book examples, used in tutorials do not provide an adequate training from a practical view point of view, because with these

(a) There is usually insufficient information as to how or why the data were collected, and scant motivation is given to the problem.

(b) There is no need to consider how well the measured variables reflect the underlying phenomenon.

(c) Data presented are usually scrupulously clean with all doubtful values removed.

(d) The choice of analysis is obvious from the context.

(e) There is no need to explain or discuss the implications of the analysis.

(f) The results or purpose of the analysis may be far from the student's interest or understanding

8.5.3. Practical work with outside organisations

(i) Comments from the literature

The ISI committee recommended the obtaining of practical experience in data analysis by making it compulsory for students to work for outside organisations during their training. (Also recommended by Jones and Kanji (1980).) From the following comments it is clear that local statisticians feel the same.

(ii) Comments from South African statisticians

Comment 21 (Translated. Part only): "Practical experience is most essential for lecturers and can be achieved by one or more of the following:

(a) Practical orientated (academic) training at student level.

(b) Practical experience at an outside organisation before undertaking lecturing".

(Note: This respondent is a university lecturer)

Comment 31 (Translated. Part only): "Project work in the third year and honours year should be further augmented by compulsory vacation work under an experienced statistician".

Comment 33 (Translated. Part only): "It would be beneficial for a statistician to spend some part of his training in the work situation (in die praktyk)".

Comment 34 (Translated. Part only): "Practical experience is essential and can be obtained from outside organisations".

Comment 54 "Perhaps experts could give 'guest lectures' to students on specific subjects".

8.5.4. Practical projects

(i) Comments from the literature

The third recommendation of the ISI committee was the undertaking of practical projects by individual students whereby valuable experience can be gained in the planning of experiments, surveys, analysis of data collected and the writing of a report. (Also recommended by Kanji (1979) and Griffiths and Evans (1976).)

(ii) Comments from South African statisticians

The following comments show that local statisticians also hold this viewpoint.

Comment 27 (Part only): "More emphasis on project work, even at undergraduate level, is required".

Comment 28 (Part only): "Perhaps small projects for industry should be encouraged".

Comment 31 (Translated. Part only): "Project work in the 3rd year and Honours year should be (undertaken)".

Comment 32 "More practical-orientated material should be included in statistics courses, as well as practical projects".

Comment 35 (Translated. Part only): I recommend that a case study approach (such as I experienced with my MBL training) is a very practical and useful didactic teaching method that can also be used in statistics".

(iii) Suggested improvement: Project work at the University of Pretoria

Steyn (1984) stresses the importance of exposing the student to the whole perspective of the research process, which includes planning, data collection, data editing and coding, analysis of data and finally the drawing of conclusions. The cycle then completes itself with better planning for any subsequent investigation.

(ii) Comments from South African statisticians

The following comments show that local statisticians also hold this viewpoint.

Comment 27 (Part only): "More emphasis on project work, even at undergraduate level, is required".

Comment 28 (Part only): "Perhaps small projects for industry should be encouraged".

Comment 31 (Translated. Part only): "Project work in the 3rd year and Honours year should be (undertaken)".

Comment 32 "More practical-orientated material should be included in statistics courses, as well as practical projects".

Comment 35 (Translated. Part only): I recommend that a case study approach (such as I experienced with my MBL training) is a very practical and useful didactic teaching method that can also be used in statistics".

(iii) Suggested improvement: Project work at the University of Pretoria

Steyn (1984) stresses the importance of exposing the student to the whole perspective of the research process, which includes planning, data collection, data editing and coding, analysis of data and finally the drawing of conclusions. The cycle then completes itself with better planning for any subsequent investigation.

At the University of Pretoria, where Steyn was a staff member, this is achieved by including a compulsory sample survey in the third year Mathematical Statistics course. Controversial campus topics make popular subject matter for the surveys, such as Censorship in South Africa, Personality differences between smokers and non-smokers, Rag - is it worth the effort, and many other topics. The student has to prepare a memorandum setting out the background and purpose of the survey with a provisional plan for executing it. After discussions with the lecturer and approval of the plan, data is collected, edited and if necessary computerised. A report of 20-30 pages setting out the complete project from the planning stage to the conclusions is then submitted. The project is rounded off by a presentation of 15-20 minutes to classmates and lecturing staff.

The objectives that are achieved with these projects include, learning the basis of scientific planning, data collection, sensible use of the computer, deciding on the optimum statistical procedures for processing the available information and communication in writing as well as verbally.

There were no negative comments. Steyn suggests that the project mark must make a significant contribution to the overall mark. Class prizes and intervarsity competitions are a further encouragement. He makes the further point that although the lecturer has to make a substantial contribution in time and effort, there is the satisfaction of exploring the essence of the subject Statistics on an informal level with the students.

(iv) Suggested improvement: Case studies at the University of the Witwatersrand

At the University of the Witwatersrand there was for some years a compulsory third year topic for major students in statistics called "Methods of Applied Statistics". Fatti (1984), who had been the major force in designing and implementing the course, describes it as consisting of "twelve to thirteen case studies or projects, each of them based on a different statistical technique. (The topic was worth 3 credit points out of a total of 13 required for Mathematical Statistics III.)"

Each case study was handled in the following manner:

The lecturer would present a brief description of the relevant technique, giving the relevant formulae and motivations but without any formal proofs. (These techniques were usually new to the students; some of them were dealt with, at a more formal level, in the Honours Course.) Thereafter the computational aspects or available computer programs or packages were discussed, followed by suggested case studies, problems and/or data to which the technique may be applied. Students were encouraged to come up with their own problems or to gather their own data, but in practice most of them used the case studies and data given by the lecturer. Great stress was placed on the written presentation of the reports.

Fatti makes the following comments:

On the positive side:

(a) The students were introduced to a wide range of applied statistical techniques - beyond the normal range of topics formally dealt with at an undergraduate level.

(b) It gave them practice in tackling reasonably substantial applied projects in a limited amount of time.

(c) It gave them experience in data gathering and analysis.

(d) It gave them practice in writing reports.

On the negative side:

(a) The problems were usually too well-structured, giving the students little experience in problem formulation.

(b) The problems were too technique-orientated, thereby giving the students little experience in the selection of the appropriate techniques.

(c) The cases were not always very interesting or appropriate, and pointed to the need for a good source of case studies.

8.5.5. The problems of statistical practice

(i) Comments from the literature

Another recommendation of the Snee committee was as follows:

> Discussion of the realities of statistical practice is needed, such as problem formulation and analysis, importance of graphical displays, problems of historical data and the use of the statistical literature. Students must understand how statistical methods are used in an industrial environment if they are to make significant contributions to the organisations that employ them.

(ii) Comments from South African statisticians

There is no doubt that the fresh graduate finds it very difficult to make the change from the abstract academic environment to the rough and tumble of the work situation. The following comments express a degree of the bewilderment that the new statistician can feel.

Comment 18 (Part only): "I still encounter some few problems when applying the theory which I learned to practical problems".

Comment 39 (Translated): "More practical real-life training - ie to regard the subject not just as elegant manipulations but as an essential applied discipline".

Comment 40 (Part only): "The application of the theory in the work situation proves extremely difficult".

Comment 42: "Statistics training provided adequate theoretical skills, ie. the 'mechanics' - insufficient training given to the topic 'recognition' and solution of real-life statistical problems".

Comment 46: "I felt my training at undergraduate level provided me with adequate mathematical skills to progress in a research capacity however lacked dismally to prepare me to handle any practical problem. I feel our undergraduate courses could do more to equip students with the necessary tools to tackle applications they might encounter in industry, agriculture and biometrics".

Comment 48 (Translated): "The Mathematical Statistics courses can definitely be more practically orientated. The theoretical approach helps little when it is necessary to think practically".

8.5.6. Exposure to statistical consulting

(i) Comments from the literature

A further recommendation of the Snee committee is as follows:

> Students need hands-on experience with real problems. This is effectively accomplished by consulting intern programmes. Working in consulting laboratories, taking consulting courses, and participating in joint seminars with other disciplines are also useful. Experience with real problems can also be included in course work although face-to-face contact with the scientist is not possible in this situation.

(ii) Comments from South African statisticians

The need for training in statistical consulting does not appear to feature highly on the list of needs in this country and only two comments mentioned consulting. The author feels particularly strongly on the need for consulting experience, and has made some comments below.

Comment 25 (Part only): "Students in 3rd year and above should be exposed to Statistical Consulting (e.g. Students in other faculties requiring statistical help in their experimental work)".

Comment 34 (Translated. Part only): "Practical experience is very necessary and can be obtained during in-service training by consultation with staff".

Comment by author:

During a period as consultant to a large chemical and industrial firm, the author has been asked to assist with the in-service training of a series of newly recruited statisticians, appointed fresh from university. These new appointees, each in turn, worked alone without the benefit of reporting to an experienced senior statistician. The author noted that they each suffered very serious difficulties with their first few consultations. It was as if they were paralysed with fright on being presented with real data in a face-to-face situation and on being expected to come up with an immediate answer. Such was the extent of the "stage-fright" that one did not recognise data requiring a one-way analysis of variance, while another failed to recognise that a two-sample t-test was required.

The author feels that this initial "shock" with consulting work can be overcome by giving students exposure at university to what Carter et al (1986) term "deciphering the researcher's requests, choosing appropriate analyses and performing these analyses, and also reporting results in writing and verbally to the investigator."

(iv) Suggested improvement: Teaching statistical consulting at the University of the Witwatersrand

Greenacre (1984) mentions that Zahn & Isenberg (1983) consider the consulting session to consist of four parts:

(i) Identification of relevant aspects of the problem situation.
(ii) Definition of the client's goals.
(iii) Determination of the actions to be taken.
(iv) Discussion of the various aspects of the consulting relationship and who will do what when.

They conclude by stating that "in consulting there is no substitute for practice, high-quality experience, observation of experts and competent supervision."

It is an unfortunate fact that often a new graduate is employed by a South African company that is too small to have a statistics or operations research section. The inexperienced graduate has then to face his or her first few consultations without the practice, experience or competent supervision that Zahn & Isenberg consider to be so essential.

For this reason the University of the Witwatersrand has started an organised consulting service, for both internal research workers and external clients, in which Honours students are allocated certain periods during the week in which they will attend consulting sessions together with a staff member. Initially the student is merely an observer, but with experience becomes confident enough to take over the major role in the consultation

The author no[illegible] that both the students he was associated with showed considerable nervousness prior to their first involvement in consulting. Once the "ice had been broken" they gained confidence and in later consultations were able to make a valuable contribution. This experience confirmed the author's conviction that we do our students a grave disservice if we send them out into industry with no experience in consulting, especially if they start their careers in one-person statistics departments.

8.6. A wide variety of statistical techniques is needed

(i) Comments from the literature

One of the recommendations of the Snee committee was as follows:

> It is important to include a wide variety of statistical techniques in the course work. Industrial statisticians are general practitioners and use many statistical tools because of the wide range of problems and subject matter areas they encounter in their work.

(ii) Comments from South African statisticians

The thirteen respondents listed below all suggested the incorporation of certain topics into the undergraduate syllabuses. These obviously deal with subjects that are useful in their particular fields of application. It is unlikely that all the topics could be included in the usual statistics programme due to the lack of time that most lecturers complain of.

Comment 16: "I feel that my statistics training gave me some skills. MODELLING, DYNAMIC PROGRAMMING, LINEAR PROGRAMMING & SIMULATION were self-taught".

Comment 37 (Translated. Part on [illegible] "Statistical training does not concentrate enough on pro[illegible]m-solving techniques in business and economic directions (eg market research)".

Comment 55: "Training not practical enough - very little training on Time Series, Forecasting, Regression - especially effects on Regression if classical assumptions do not hold. Training on Multivariate too Mathematical. People we appoint to do these jobs have to relearn it first".

Comment 56: "Personally, there should be a modular course on Directional Data and Time-Series Analysis at undergraduate level. These methods often have to be used in industry".

Comment 57 (Part only): "In South Africa we need Statistics with Time Series, Forecasting and Quality Control training".

Comment 58 (Part only): "The school of thought of Dr E W Deming which advocates serious consideration of 'Statistical Process control' was not highlighted to the right and appropriate perspective".

Comment 59 (Translated. Part only): "Training in statistical process control is critically necessary at Universities and Technikons".

Comment 60 (Part only): (My training) "covered almost nothing on Experimental design".

Comment 61: "I think that simulation and model development (i.e. more OR) should be incorporated in statistics courses".

Comment 62 (Part only):"I feel a greater emphasis should be placed on data inspection and multidimensional methods".

Comment 63: "Academic training gave very little attention to aspects of market research".

Comment 64: "Perhaps universities could do more work on spatially-defined data".

Comment 65: "Maybe the geostatistical theory could also have formed part of an under-graduate course".

8.7. Computer program packages

(i) Comments from the literature

Another recommendation of the Snee committee is as follows:

> Students should be familiar with the well-known statistical computer program packages and at least one programming language. An industrial statistician will probably need a computer early in his or her career.

In this modern age of computers this is such an obvious recommendation that one is almost surprised that it is necessary to make it. It is perhaps a reflection on university courses that the following comments were made.

(ii) Comments from South African statisticians

Comment 24 (Part only): "If I were designing courses today I'd replace about one-third of the current curriculum with more nuts and bolts courses on data analysis including enforced usage of at least 2 program packages".

Comment 36 (Translated. Part only): "Students should be thoroughly schooled in the use of program packages, for example SAS. It would be sensible if they could get a broad overview of what is available and could learn to apply them themselves".

Comment 60 (Part only): (My training) "had almost no practicals on the use of statistical packages".

Comment by Juritz et al (1985): In Section 4.5.4. the suggestion is made "to include some training in statistical computing in our courses".

8.8. Continuing education

Continuing education is a means for revising and upgrading the statistical skills of both graduate statisticians and other personnel who find they require statistical techniques in their work. This is highlighted in the following two comments:

Comment 26: (I had) "no training other than Summer School. Regular courses in basic techniques would be valuable. Occasional in others necessary".

Comment 30: (Part only): "Suggestion on advanced training for statisticians: I think there is still a lot of scope for training through the medium of workshops on more specialized subjects. The HSRC workshop on SAMPLING and the UCT workshop on REPEATED MEASURES were particularly worthwhile for me".

8.9. Communication skills

The Snee committee had the following to say on this matter:

> Enhancement of communication skills is a valuable addition to the programme. The ability to communicate orally and in writing with non-statisticians at all technical and managerial levels is the key to the success of the industrial statistician. The preparation of written reports and associated critiques can be an integral part of the course work. Oral communication can be improved as part of the consulting experience by oral presentations of statistical analyses.

Beard (1972:13) makes a very valuable comment on communication skills. She could almost have had the statistical profession in mind in highlighting "the need to prepare students to cooperate with specialists in other fields, to work in teams, to communicate information to committees consisting of specialists who make different assumptions, employ different vocabularies and concepts and have somewhat different modes of thinking". She therefore recommends "that communication skills should be fostered in the majority of courses in higher education and it would undoubtedly be beneficial for students and their teachers to have some knowledge of group dynamics".

Comment by author:

None of the respondents commented on this aspect of the statistician's work. The author's experience as an industrial statistician has convinced him that most statisticians do not communicate well with non-statisticians. Written reports tend to be too theoretical in nature. Engineers, in particular, are usually interested only in the "bottom line" and do not relish wading through pages of preliminaries concerning null hypotheses, one and two tail alternatives, transformations of data, etc.

8.10. Other suggestions from the literature

8.10.1. Model syllabus

The ISI Committee recommended the design of a model syllabus, updated from time to time, that could serve as a guide for the training of statisticians.

One of the tasks given to Professor Juritz' subcommittee of the Education Committee of the SASA was to evaluate existing university syllabuses in the light of the needs of employers and, if necessary, to recommend changes, and also to encourage discussion about syllabuses.

The subcommittee produced a suggest draft syllabus. The first year's work was similar to that at UCT which is based on Exploratory Data Analysis. Topics were suggested for the subsequent years of study, but these are in such a general form that the syllabus is not much help in the context of this study. The statistical methods investigated here might or might not be emphasised in the implementation of the syllabus. The syllabus is attached as Appendix F.

The syllabus was circulated to the members of the SASA and discussion was called for. It is not known what comments were received and this matter does not seem to have been taken any further.

8.10.2. Co-operation between universities and industry

The Snee report lays great stress on the benefits of a fruitful interaction between university and industry as shown by the following comments from this report:

> The lecturing staff's interest in real problems is an important factor, as is their attitude toward statistics in industry. It is worthwhile to introduce students to industrial statisticians whenever possible and expose them to their activities and the results of their work.

Universities and industry should work together to develop appropriate training programmes. Improved communication is needed so that each community can understand the other's needs and problems. It is important for university staff to take an active role in both statistical and non-statistical professional societies.

Comment by author:

With reference to the recommendation on introducing students to industrial statisticians it is interesting to note that visits by third year students to industrial statisticians was the practice at Rand Afrikaans University. At the Johannesburg Technikon quality control practitioners are invited to address industrial engineering students.

8.10.3. Postgraduate work

The Snee committee made the following additional recommendations, which were not matched by comments from the South African statisticians:

Academic programmes should be aimed at the M.Sc. level and higher, as it is generally agreed that the M.Sc. degree is a minimum requirement for the professional statistician.

Departments should offer opportunities for student post-graduate research that are stimulated by real problems.

Baskerville (1981) reports on a graduate course at the University of Western Ontario called "Statistical Consulting, Computing, and Data Analysis", which uses a case study approach along with detailed discussion of a number of current projects.

Kanji (1979) suggests that research is less important at M.Sc. level than a project related to a real problem. Andrews (1979) gets his students to search the subject matter literature for experiments with real data that is reported with little or no statistical analysis. They must then comment on experimental design and reported analysis, and then perform the appropriate analysis and write up the results.

8.10.4. <u>Teaching resources and journals</u>

The ISI Committee pleaded for financial support for the writing of books and reference works that could be used in courses.

Tanur (1982) wants the American Statistical Association to do more in producing and making available teaching resources and in fostering good teaching techniques, and especially in producing a journal devoted exclusively to teaching along the lines of the British "Teaching Statistics" and the "College Mathematics: Suggestions on How to Teach it" of the American Mathematical Association.

Comment by author:

It appears that Tanur's suggestion has been taken up by The American Statistician, which now includes articles on teaching. The Journal of the SA Statistical Association unfortunately does not yet seem to have adopted this policy. The author feels that the Journal of the SA Statistical Assocation should carry regular articles on educational theory. This would not only be beneficial to lecturers in statistics but would also go some way to countering one of the criticisms in the ISI report, namely, that "the statistical literature was usually dominated by theoretical articles of a mathematical nature. The fact that this literature was usually incomprehensible to the practising statistician led to further widening of the gap between theoretical and practising statisticians."

CHAPTER 9. CONTINUING EDUCATION COURSES FOR BUSINESS AND INDUSTRY

9.1. Courses at University of the Witwatersrand

9.1.1. Description of courses

One of the practical recommendations suggested by participants in the survey of members of SASA and OR Society (viz Questionnaire No 2) was the running of courses in statistics for personnel in business and industry (See Section 8.8). Questionnaire No 1 had shown that many companies were using personnel with very little statistical training to carry out the statistical work they required. It seemed logical that firms would wish to upgrade the statistical training of such personnel. It was decided therefore to design such courses and evaluate their effectiveness by getting the participants to complete an evaluation questionnaire at the end of each course.

Permission was obtained from the appropriate authorities to run part-time courses under the auspices of the Centre for Continuing Education of the University of the Witwatersrand, Johannesburg. After much deliberation among members of the Department of Statistics at the University of the Witwatersrand it was decided to run two courses entitled Introductory Statistics and Experimental Design, and Introductory Statistics and Analytical Assurance. The courses would each be the equivalent of a one-week full-time course, but would be spread over 9 evenings in June 1985 between 4 and 7 p.m. The idea here was that companies would be more amenable to paying for staff to attend courses which caused a minimum of disruption to their routine duties.

The reason for two courses was because two separate target populations had been identified. Course 1 was designed for engineers and scientists in industry and commerce who were involved in technical work in general and in research and development projects in particular. Course 2 was aimed at chemists and laboratory managers involved in maintaining analytical standards or undertaking problem-solving projects.

The objectives of the courses were as follows:

Course 1: Introductory Statistics and Experimental Design.
An introductory section was aimed at upgrading the basic statistical knowledge of those whose training in this field was inadequate and supplying a refresher course for those whose training required revision and revitalisation. In the latter part of the course the participants were trained in a number of statistical techniques which have proved themselves invaluable to sound technical work, such as multiple regression, analysis of variance and experimental design. The problem-solving techniques of quality assurance were also dealt with.

Course 2: Introductory Statistics and Analytical Assurance. The introductory section was identical to that of Course 1. In the latter part of the course the participants were trained in modern analytical assurance and in the problem-solving techniques of quality assurance. Regression and analysis of variance were also dealt with but in less detail than in Course 1.

Because the first part of both courses was the same, the participants all attended the same lectures on the first 5 days. On the last four days of each course the participants separated for specialised lectures held on different days to the other course.

The course leader was the author and the other lecturer was Professor H S Sichel, a Visiting Professor in the Department of Statistics and former Head of the Wits Business School. Both lecturers had had considerable experience in industry and commerce, particularly Professor Sichel who had been consultant to a number of large S A corporations and mining houses.

The course contents are shown below, with the dates in June, 1985:

Course 1

Mon 10 Descriptive statistics and probability.

Tues 11 Binomial, Hypergeometric, Poisson and Negative Exponential distributions.

Wed 12 Normal and Log-Normal distributions.

Thur 13 Tests of significance using Normal and t-distributions. Chi-squared test for contingency tables, and goodness of fit tests.

Mon 17 Correlation and regression, Multiple regression.

Tues 18 Analysis of variance: one- and two-way.

Thur 20 Factorial experiments.

Mon 24 Experimental design.

Wed 26 Quality Assurance, Cusums.

Course 2

Mon 10 Descriptive statistics and probability.

Tues 11 Binomial, Hypergeometric, Poisson and Negative Exponential distributions.

Wed 12 Normal and Log-Normal distributions.

Thur 13 Tests of significance using Normal and t-distributions. Chi-squared test for contingency tables and goodness of fit tests.

Mon 17 Correlation and regression, Multiple regression.

Wed 19 Analysis of variance: One way and nested.

Fri 21 Quality Assurance, Cusums.

Tues 25 Acceptance sampling and methods of sampling.

Thur 27 Analytical assurance.

The course attracted some thirty participants, most of whom were of a high academic calibre, including a number with PhD's. The overall reaction to the course was very favourable (with one exception, namely, a participant who was very disappointed in the lack of practical exercises).

From the experience gained from these courses and similar courses at AECI Ltd, it was decided that it was unnecessary to have two separate courses emphasising, respectively, Experimental Design and Quality Assurance. The two were therefore combined into a single course which was run in June, 1986 over a thirteen evening period. The contents of the course was essentially the same as that shown in Section 9.2.4., although the order of presenting the topics was as follows:

Descriptive Statistics, Probability.
Discrete distributions: Binomial, Hypergeometric, Poisson
Continuous distributions: Negative exponential, Normal
Sampling distributions, Hypothesis tests using Normal.
One and two sample t-tests
Tests of proportions, Confidence intervals
Chi-squared tests and goodness of fit (to a log-normal example)
One-way, two-way and nested Anova.
Randomized blocks and factorial designs.
Correlation and regression, Multiple regression.
2^n factorials and fractional replicates of 2^n, EVOP.
Quality Assurance, Control charts, Acceptance sampling, Cusums.

The results of the course evaluations are presented in Section 9.1.2.

9.1.2. Course evaluation

The evaluation questionnaire used at the end of the 1985 courses is shown in Appendix D. The questionnaire used in 1986 was shorter and simpler in that Questions 7 to 12 were not applicable and were omitted. The results of the evaluation questionnaires are a combination from the Wits courses for both years and the AECI and Middelburg courses.

The answers to the more important questions are summarised below.

Q.1. : Will this course be helpful to you in your current (or anticipated future) position?

	No	Slightly	Yes	Very much
Number	0	11	39	18
Percentage	0	16%	57%	27%

Q.2. : Did you find it interesting personally?

	No	Slightly	Yes	Very much
Number	0	6	44	18
Percentage	0	9%	65%	26%

Q.3. : Was the general level about right?

	No, too elementary	Yes, about right	No, too advanced
Number	4	55	7
Percentage	6%	83%	11%

Q.4. : Rate the following sections of the course

		Would have liked less	About right	Would have liked more
Descriptive statistics	Number Percent	11 21%	35 69%	5 10%
Distributions	Number Percent	10 20%	34 65%	8 15%
Tests of significance	Number Percent	2 4%	29 57%	20 39%
Chi-squared tests	Number Percent	2 4%	29 58%	19 38%
Analysis of variance	Number Percent	1 2%	28 55%	22 43%
Experimental design	Number Percent	1 2%	21 53%	18 45%
Correlation & regression	Number Percent	1 2%	26 52%	23 46%
Quality Assurance	Number Percent	5 11%	30 65%	11 24%
Acceptance sampling	Number Percent	6 13%	29 62%	12 25%
Cusums	Number Percent	2 4%	31 66%	14 30%
Analytical Assurance	Number Percent	2 6%	23 72%	7 22%

Comments from the open-ended questions

a) Favourable overall opinion

1. A very useful overview of applications of statistics. Lectures were well presented and interesting.

2. Very good course for persons who require to update or improve their statistical knowledge.

3. This is the first course where practical applications made statistics more understandable to me.

4. Overall course is very important before going on to separate in-depth courses.

5. Very well presented with sound theoretical and practical advice.

6. Liked best: The many very interesting points that came out of this course.

7. Liked best: Refreshing of long forgotten principles. Furnishing me with new ideas and applications.

b) Unfavourable overall opinion

1. Brochures stated there were workshops - none were held. I would not have attended had I known the format of the course would be so poor. Very disappointed in general.

c) <u>Comments favouring practical applications and worked examples</u>

1. As far as possible the lecturers should emphasise the practical aspects, and cover in less depth the descriptive stats and distributions found in any stats textbook.
   Liked best the practical advice in carrying out tests - the sort of thing you can't find in textbooks.

2. Liked best: The applications of theory, good practical examples.
   Concentrate on applications. Real life (not synthesised) examples are most interesting.

3. Liked least: Having the full time taken by a lecture. I would have preferred to have part lecture/part workshop.
   I enjoyed the real examples which both lecturers were able to provide.

4. Suggested improvements: More worked examples, for quick reference in future. Suggest questions with answers written out in full, enabling one to go back at a later date to see how to practically use a method correctly.

5. Liked least: Far too much time spent developing background theory, especially in "Distributions" - e.g. moments of dirtributions - would have been better to move from basic concepts to applications with formulae being accepted rather than derived from first principles in each case.

6. The primary aim of the course should be to impart skills not readily available in textbooks.

7. Combine the two courses into one introductory course. Follow up with in-depth courses is very important.

8. A list of current books/literature that are useful should be given to cover the topics considered.

9.2. In-house courses

9.2.1. Course at AECI Limited: November 1985

Subsequent to the 1985 Wits courses, an in-house course was requested by AECI Ltd at Modderfontein, Johannesburg. They preferred an all-day week-long course running from 8 a.m. to 4 p.m.. The course content covered was basically the same as the Wits courses. The first three days were common to both courses and the fourth day consisted of specialist lectures for Course 1 and the fifth day those for Course 2. The time available (viz seven hours x four days = 28 hours) was similar to the Wits courses (viz three hours x nine evenings = 27 hours). From the comments it will be obvious that it is undesirable to lecture all day for four days. The concentration of the participant flags under the deluge of statistical concepts. For a course run on consecutive days a period of three hours per day is more desirable. For all day courses, two days per week is more conducive to attentive learning.

Course evaluation

The evaluation questionnaire consisted of Questions 1 to 6 and Question 14 of the Wits questionnaire (See Appendix D). The results are included with those of Section 9.1.2., except for the following three comments, which refer to the timing of the course.

Comments from the course evaluation

1. Liked least: Whole day lectures - too much to absorb. Would have preferred morning lectures only for about two weeks - would have absorbed more.

2. Liked least: Too much to digest all at once. It would be better to spread it over a few weeks instead of condensed into four days.

3. Liked least: Too many facts in too short a period. I would have liked to have had some more time to get more familiar with some terms before moving on to new items.

9.2.2. Course at Adcock-Ingram Laboratories: July 1985

This course consisted of 8 sessions, each of 2 hours. Each session was held on a Friday from 1 to 3 p.m. The arrangement was not ideal because the weekly interval between sessions was too long. As the staff finished work at 3 p.m. on a Friday, their attention tended to wander as the end of the week approached.

The course content was as follows:

Session 1: Descriptive statistics.

Session 2: Probability, Binomial Distribution.

Session 3: Hypergeometric, Poisson and Binomial distributions with special reference to Acceptance Sampling.
Normal and Log-normal distributions.

Session 4: Tests of Significance and Confidence Limits using Normal and t-distributions.

Session 5: Quality Assurance with emphasis on problem solving techniques.

Session 6: Control Charts.

Session 7: Acceptance Sampling, Operating Characteristic curves related to BS 6001.

Session 8: Sampling Plans related to BS 6001. Cusums.

9.2.3. Course at Middelburg Steel and Alloy: July 1986

Having learned from the AECI course that four days in a row is too long, and having found from the Adcock-Ingram course that a few hours once a week was too little and too far apart, it was decided to run the Middelburg course on each Tuesday and Wednesday over three consecutive weeks. This arrangement was preferable to the previous courses as far as student fatigue was concerned. A problem that arose was that the course was given by one lecturer (travel and hotel costs put this constraint on out-of-town courses). The order of presenting the topics must be organised so as to intersperse tutorial exercises between short lectures of about half an hour in duration. If this is not done the students find a one-person presentation monotonous and it places a strain on the lecturer. The course contents were basically the same as those for the Anglo-American course in Section 9.2.4., but with slightly greater emphasis on Quality Assurance.

9.2.4. Course at Anglo-American Gold Division Services at Welkom: December 1986

This course was similar to that at Middelburg Steel and Alloy, but with less emphasis on Quality Assurance. It was run on each Tuesday and Wednesday over three consecutive weeks. The order of presenting the topics was changed to mix those that were conducive to tutorial exercises with those that were more of a lecture-only nature. The contents and order of presentation were as follows:

Day 1: Descriptive statistics / Probability / Binomial distribution.

Day 2: Poisson / Hypergeometric / Negative exponential / Normal distribution / Normal tables / Quality Assurance in general.

Day 3: Sampling distributions / Test mean using Normal / One and two sample t-tests / One-way and Two-way Anova.

Day 4: Nested Anova / Testing proportions / Confidence Intervals / Control Charts / Randomised blocks design / Factorials.

Day 5: Correlation / Regression / Multiple Regression / Non-linear regression / Acceptance sampling.

Day 6: Chi-squared Goodness of fit and Contingency tables / Case study on how not to carry out an R & D project / 2^n factorial / 2^n fractional replications / EVOP / Cusum.

Course evaluation

The evaluation was done using a standard Anglo-American questionnaire. The answers to the questions it contained are as follows:

1. Do you think that less time should have been spent on any aspect of the course? (Please specify)

   10 answered No; 1 suggested less time on QA.

2. Do you think that more time should have been spent on anything? (Please specify)

   7 answered No; Others suggested Forecasting (twice), Practical exercises and any aspect (by a beginner!).

3. Do you think that any subject should be completely eliminated from the course? (Please specify)

   11 answered No.

4. Do you think that any other subjects should be added to the course? (Please specify)

   7 answered No; Others suggested Non-parametric statistics, Forecasting (twice), More on Sample size.

5. Did you personally think that the course as a whole was long enough?

Increase its length by			Just right as it is	Reduce its length by		
100%	50%	20-30%		20%	30[illegible]	[illegible]
	1	2	8			

6. Did you think that there was a satisfactory balance between lectures and practical work (case studies, worked examples, role plays, etc.)

Too much lecturing	Adequate	Too much practical
3	8	

7. What did you personally think about the level of the course?

Much to basic	Very Simply	Easy	Just Right	Diffi-cult	Very heavy going	Much to Advan-ced
		1	9	1		

8. What is your overall impression of the course?

Dreadful	Poor	In-adequate	Adequate	Good	Very Good	Exce-llent
				6	3	2

Comments

1. A well presented course with good notes. My only concern is over the absorption by students and I think revision in 12 months time should be considered.
2. Very useful and usually enjoyable course. Lecturer coped very well with students of mixed statistical background. Good to see relevant examples to our industry, and class participation encouraged. The actual meaning of statistical expressions (formulas) were well explained rather than note learning. Occasionally too much time was given to "lecture-hoggers".
3. Pre-course material could have helped.
   Not enough time allowed to answer exercises.
4. Very well presented and instructive course. Especially the design of experiments.
5. The discussion on QA was possibly out of place since we are not heavily involved in mass QA as opposed to specific QA problems. However, it was nevertheless interesting to see how it is practised.

9.3. Summary of findings

It appears that the course developed by the author was favourably received and fitted a real need. Of the sixty eight participants on the Wits, AECI and Middelburg courses, 84% felt the course was helpful or very helpful in their current positions; 91% found it interesting or very interesting; 83% found the level about right, 6% found the level too elementary and 11% too advanced.

A good indication of the degree to which a topic is considered important and needed in practice is the percentage of the participants who would have liked greater emphasis placed on the topic. The ranking is as follows:

Topic	Percentage of participants who would have liked a greater emphasis
Correltation & regression	46%
Experimental design	45%
Analysis of variance	43%
Hypothesis testing	39%
Chi-squared tests	38%
Cusums	30%
Acceptance sampling	25%
Quality assurance	24%
Analytical assurance	22%
Distributions	15%
Descriptive statistics	10%

The open-ended questions showed quite clearly that people using statistics in the business world are not interested in the theoretical aspects of the subject. They appreciate a presentation that shows the usefulness of the techniques and is based on real-life examples. Lecturers with experience in industry have an advantage as far as credibility is concerned.

As far as the timing of the course is concerned, all-day lecturing for more than two consecutive days is too concentrated a presentation to be digested and absorbed.

On the other hand, too long a time interval between sessions is not desirable. An acceptable arrangement was three evening sessions per week, each of three hours per evening. For in-house courses, where travelling was involved, a reasonable compromise was two consecutive days every week for three weeks.

The topic which appeared to be missed the most was Forecasting. That this is a much needed technique is borne out by the excellent response of over sixty participants enrolled for a two-day course which was given at Wits University in July, 1987.

A problem that always occurs with this type of course is the heterogeneous nature of the participants with respect to statistical background. They vary from people with no previous knowledge of the subject to statistics majors who desire to upgrade their understanding of one or other topic like Experimental design or Quality control. The lecturer has to pitch the level of the course somewhere in the middle. Course evaluations show that it is unavoidable that for certain individuals the pace is too slow, while for others the level is too high and the pace too fast. The lecturer cannot help boring someone at some stage and causing others to flounder.

CHAPTER 10. SUMMARY OF FINDINGS AND DISCUSSION

10.1. Statistical methods "in common use"

10.1.1. Methods "in common use" in all five employee groups

In Sections 4.5.5. and 5.2. two criteria were used to differentiate between methods used commonly and those used less often. These were:

(i) 30% or more routine and occasional usage of a method,

(ii) 20% or more routine usage of a method.

These criteria were applied to the various employee groups and lists of methods which satisfied these criteria were presented in order of merit in Sections 4.5.5. and 5.2.

In order to compare methods over all employee groups simultaneously, an "index of usage" is proposed.

The index of usage can be calculated by taking the mean of all percentage usages in each of the five employee groups for both routine and routine + occasional use. This index is a weighted mean in that the frequencies for routine usage appear twice in both the routine and the routine + occasional percentage, ie. routine use is weighted more heavily than occasional use.

The percentages from each employee group are weighted equally, since no account is taken of the unequal numbers of respondents in each group from which the percentages are computed. Equal weights have been chosen because the frequencies of respondents do not represent the true unknown frequencies in the population of statisticians. For example, the calculation of the index for Basic statistics is shown in Table 10.1.

TABLE 10.1. Example of calculation of index of usage (for Basic statistics)

Group	Business		University		State		Private		OR Soc	
Type of usage	Rout.	Rout. + Occ.	Rout.	Rout. + Occ.	Rout.	Rout. + Occ.	Rout.	Rout. + Occ.	Rout.	Rout. + Occ.
% for Basic Statistics	74	86	88	96	100	100	93	100	68	95

The index of usage for Basic statistics is the mean of the ten percentages = 90%.

(The first two percentages come from Tables 4.10. and 4.11., viz. (22+9)/42=74% and (24-12)/42=86%. The remaining percentages come from Tables 5.1. and 5.2.)

There are eight methods which satisfy the above criteria for "common use" in all employee groups for both routine and routine + occasional use. They are presented in the upper section of Table 10.2. There are a further three "commonly used" methods which occur in all groups, but sometimes only for routine + occasional use. They are shown in the lower section of Table 10.2. The index of usage is used to place the methods in order of merit.

TABLE 10.2. Statistical methods "in common use" in all five employee groups (in order of merit in terms of the index of usage)

Statistical method	Index of usage
Basic statistical methods	90%
Graphical display & data summary	88%
Regression analysis	77%
Analysis of variance	69%
Correlation analysis	67%
Forecasting	57%
Multivariate analysis	54%
Simulation	50%
Contingency table Chi-squared tests	55%
Time series analysis	51%
Linear programming	40%

10.1.2. Methods "in common use" in four employee groups

In Table 10.2 the methods satisfied the criteria for "common use" in all five employee groups. The next most important set of statistical methods are those which appear as "commonly used" in only four of the five employee groups. They are shown in Table 10.3. The employee group in which the method does not occur as a "commonly used" method is shown in the last column.

TABLE 10.3. Statistical methods "in common use" in only four employee groups (in order of merit)

Statistical method	Index of usage	Group in which method is not "in common use"
Design of experiments	45%	Business
Non-parametric methods	40%	Business
Survey sampling	40%	OR Society
Discriminant analysis	37%	Business
Factor analysis	34%	Business
Quality control	33%	State
Non-linear estimation	31%	OR Society
Ranking/Paired data analysis	31%	Business

10.1.3. Methods "in common use" in three employee groups

The next most important set of statistical methods are those that appear as "commonly used" in three of the five employee groups. They are shown in Table 10.4. The two employee groups in which they do not occur as "commonly used" methods are shown in the last column.

TABLE 10.4. Statistical methods "in common use" in only three employment groups (in order of merit)

Statistical method	Index of usage	Group in which method is not "in common use"
Categorical data analysis	29%	Business, OR Society
Indices	22%	Business, OR Society
Multidimensional scaling	21%	Business, OR Society

10.1.4. Methods not "in common use"

The remaining eleven statistical methods are not "used commonly". They are Variance component estimation, Bioassay, Logistic regression, Non-linear programming, Dynamic programming, Inventory control, Reliability and Life testing data analysis, Queueing theory, Decision theory, Network models (PERT/CPM) and Conjoint analysis.

10.2. Differences between employee groups

10.2.1. Differences in qualification profiles

In Section 6.1. it was shown that the Operations Research Society differed significantly from the S.A. Statistical Society mainly in respect of proportion of PhD's, in that 17% (3 out of 18) of OR Society members have PhD's against 50% (135 out of 270) of SASA members (103 out of 207 for Directory and 32 out of 63 for Questionnaire No.2).

Information on the qualifications of SASA members came from two sources, namely, the SASA Directory and from Questionnaire No. 2. The qualification profiles from these two sources were significantly different mainly because of a difference in proportion of Bachelors degrees (3% and 16%, respectively), but they had almost exactly the same proportion of PhD's (50% and 51%, respectively).

Within the three subgroups of both sources of SASA information, University members did not differ significantly from the State and semi-state subgroup, but did differ significantly from the Private sector. This was due to differences in PhD proportions of 57% (Directory) and 72% (Questionnaire No.2) for University against 17% and 29% for Private sector.

The State and Private subgroups did not differ significantly although there was a disproportion in PhD's of 43% and 42% for State and 17% and 29% for Private sector.

To summarize : there is a rough grouping in qualification profiles. The University and State and semi-state subgroups form one grouping with a high proportion of PhD's (55% on average). The Private sector of SASA and the OR Society form a second grouping with a lower proportion of PhD's (20% on average)

Note that the information on the qualifications of individual users of statistical methods from Questionnaire No 1, which was sent to the Stock Exchange and UCT Business School samples, was not suitable for these comparisons. The reasons for this were discussed in Section 2.3.

10.2.2. Differences in usage of statistical methods

In Section 10.1. an index of usage was calculated for the various statistical methods. The following eight methods, which showed a significant difference in usage between two employee groups in Section 6.4., had indices of usage of 40% or higher:

> Regression analysis, Correlation analysis, Analysis of variance, Design of experiments, Chi-squared tests, Non-parametric tests, Multivariate analysis and Forecasting.

In attempting to highlight methods whose use differs between employee groups, it is not helpful to include the above commonly used methods. For example, the University:OR Society usage ratio for Regression analysis is 92:53% and for Analysis of variance is 84:37%. The fact that these differences in usage are statistically significant is more a reflection of the greater use of statistical testing by University statisticians than by OR practitioners, rather than an indication of a preference of the University group for there two methods.

The above eight commonly used methods were therefore removed from the methods that showed significant differences in Section 6.4. Those methods which remain classify themselves into two clear-cut groupings. The first grouping in Table 10.5. contains eleven methods which are more theoretical or academic in nature when compared with the four methods of the second grouping that are of a more applied type.

These two groupings of methods also classify the users into two categories, namely, the University and State & semi-state subgroups of the SA Statistical Association on the one hand (where the only difference is a greater use of Queueing theory by University over the State group) and the OR Society and Business (i.e. the Stock Exchange and UCT Business School samples) on the other hand (where the only difference is a greater use of Dynamic programming by the OR Society over the Business group). In between these two categories lies the Private sector of SASA, which appears to overlap into both the University/State and Business/OR Society categories.

In Table 10.5 the methods grouped under University and State & semi-state are those with a higher frequency of use by either University or State as compared with either the OR Society or Business. Those grouped under OR Society and Business have greater frequency of usage by either the OR Society or Business over either University or State. The Private sector of SASA lies in between these two categories, showing up as more applied than the University/State category, but less applied than the OR Society/Business category.

The frequency with which the methods showed up as significant is shown in brackets in Table 10.5. These frequencies are taken over both the Routine and Routine + Occasional analyses in Section 6.4.

TABLE 10.5. Methods grouped under the user category which used them most frequently. (Frequency of significance of methods in brackets)

User Categories		
"Academic" or "theoretical" methods used more frequently by University or State & semi-state subgroups of S.A. Statistical Association	"Applied" methods used more frequently by OR Society or Business (i.e. Stock Ex + UCT Business School)	Methods used by Private sector of SASA with usages overlapping in the "theoretical" category
Logistic regression (4)	Quality control (1)	Private > Business for
Categor data analysis (4)	Dynamic prog (1)	Cat data anal (1)
Non-linear estimation (3)	Inventory cont (1)	Factor analysis (1)
Discriminant analysis (3)	Network models (1)	Discrim anal (1)
Var compt estimation (2)		
Ranking/Pair comparison (2)		State & Univ > Private
Factor analysis (2)		Logistic regn (2)
Multidimensional scaling (2)		
Bioassay (1)		
Non-linear programming (1)		
Queueing theory (1)		

10.3. Teaching Emphasis in Universities

In Chapter 7 Spearman's rank correlation was used to test whether the teaching emphasis on the various statistical methods matched the usage of those methods. Teaching emphasis was estimated by a weighted index which unfortunately was not a very sensitive measure. The correlations were all significant but only of medium strength, implying that the teaching of the methods was reasonably satisfactory. However, there were four methods for which greater emphasis in teaching appeared to be given than seemed to be warranted by their usage, and six methods whose usage appeared to warrant greater teaching emphasis. These are set out in Table 7.7. and summarized below.

As can be seen in Table 7.7, in all five employee groups there is too much emphasis in teaching on Queueing, whereas Simulation gets too little emphasis. For the University and State groups the teaching emphasis on Non-parametric methods, Ranking/Paired data analysis, Forecasting and Indices is about right, but there is too much emphasis on Inventory control and too little on Factor analysis, Categorical analysis and Non-linear estimation.

For the more business orientated groups, ie for Business, OR Society and sometimes the Private sector, there is too great an emphasis on Non-parametrics, Rank/Paired data and Categorical analysis and too little on Forecasting and Indices. For the OR Society the emphasis on Inventory control, Factor analysis and Non-linear estimation is about right.

The grouping of these methods again confirms the similarity in usage of the University and State groups as against the OR Society and the Business group. As before the Private sector group overlaps the other two groupings.

These methods, for which teaching is out of line with usage, suggest the possible merit of splitting the statistics curriculum to allow for a "theoretical stream" aimed at supplying graduates for University and State and semi-state employ and an "applied" stream for the business orientated students.

10.4. Comments by statisticians on adequacy of university training

Those respondents who expressed a measure of satisfaction with their training were largely those who accepted that statistics is too broad a field to be covered in a three or even a four year curriculum. They felt that they had received a solid basic knowledge and the ability to develop for themselves the necessary skills required for specialist applications. There were fifteen out of 69 respondents who seemed sufficiently satisfied not to make a specific request for improvement.

The main complaint about statistics courses is that they are too theoretical in nature and that not enough practical examples are discussed. There were 52 respondents (80%) who either referred to this or made suggestions for improving the practical training.

It is obvious that the student often gets the feeling that the data used in examples are artificial, too well-behaved and not from real life. When real data is met in practice the newly graduated statistician tends to flounder.

Other suggestions included calls for practical work with outside organisations, for practical projects as part of the curriculum, for exposure to consulting and for training with computer packages. Where suggestions have included the addition of topics to a curriculum, they unfortunately failed to suggest what existing material should be left out of the programme, which is no doubt already packed full to capacity. That this is a common problem is shown by Beard (1972:27) who states:

> "The solution most commonly favoured when subject matter increases is simply to lengthen the corresponding courses. As shown by student complaints, a danger which too often materialises as courses are extended, is that they become overloaded, leaving the students too little time to think and burdening them with excessive detail".

Many of the opinions of practising statisticians documented in Chapter 8 can be related to Figure 1.1. in Section 1.6.2. This shows three sources of objectives which correspond to three views of education, namely, the Functional View, the Cultural View and the Social Service View. Many of the comments imply that the Cultural View (ie the more theoretical or academic influence) dominates, and the statisticians who made those comments either find that acceptable or can see no alternative.

One of the respondents (See Comment 23 in Section 8.4.) raised the interesting question of what constitutes education and what training. The functional aspect of education (developing skills) is clearly training. The cultural aspect (acquisition of knowledge) and the social service aspect (growth of attitudes) are catered for by the more theoretical or abstract parts of the syllabuses of the statistical courses making up the degree.

When a respondent comments that the training is too theoretical, he or she is implying that the cultural and social service aims are dominating the functional aspect. Some complain about the dominance of the theoretical approach without suggesting how this can be prevented. Others make suggestions for increasing the influence of the vocational aspect, but often seem to ignore the necessity for the other aspects to diminish as one increases.

A further interesting point is raised by another respondent (See Comment 22 in Appendix E), namely, "Difficult to remember enough details (about my training) as to what was good or bad, or what could have been better." What is being implied is that given time the graduate forgets the details of both the functional and the cultural parts of any course. It follows that it is not necessary to cover everything in the under-graduate degree. It is the social service aspect which develops in the graduate the attitudes and skills to make good any difficiencies in the educational system. As long as a sound balance between the three educational aspects is achieved, the graduate can overcome any barriers in his or her future career.

10.5. Effectiveness of continuing education in statistics

The attendance at the Continuing Education courses at Wits University in 1985 and 1986, and the request for in-house courses from four companies show that there is a demand for this type of course. The evaluation questionnaires show that the course must not be of too theoretical a nature, and that participants appreciate real-life examples.

The content and manner of presentation of the type of course proposed here have evolved over some years and will continue to be improved upon. However, the course as given to Anglo American in December, 1986 appears to satisfy their needs admirably. To this extent the course can be considered to be effective in filling the need for statistical training of this type. One proviso must be made, namely, that there is a demand for topics not included in these courses. In particular, Time series and Forecasting are often requested.

10.6. Answers to the research questions

The findings which have been summarized in this chapter go a long way to answering the research questions which were posed at the start of this study. These answers will be discussed in more detail in the next chapter in the form of conclusions.

CHAPTER 11. CONCLUSIONS AND RECOMMENDATIONS

11.1. Conclusions

11.1.1. The aims of this study

As stated in Section 1.3, the aims of this study were to ascertain what statistical methods are used by statisticians employed in different sectors of the South African economy, to evaluate whether our universities are providing adequate training for the statisticians who use these methods, and to recommend possible improvements to statistical education at university level.

It was planned to achieve these aims by answering six research questions, which are also set out in Section 1.3. Information was collected by means of three questionnaires, namely, Questionnaire No 1 to statisticians in business, Questionnaire No 2 to members of the SA Statistical Association and the Operations Research Society and Underhill's questionnaire to the universities. Valuable information is also available in the literature. Each of these research questions will now be discussed in turn to ascertain to what degree the aims of the study have been achieved.

11.1.2. Usage of statistical methods

Research Question 1 asked what statistical methods were used by statisticians in South Africa. Information from a list of thirty three statistical methods was obtained by means of Questionnaires No 1 and 2. Clear usage patterns emerged which showed that it was justifiable to divide statisticians into five distinct groups, namely, members of SASA employed by Universities, by State and semi-state bodies and by the Private sector, members of the OR Society and, finally, non-members of the SASA and OR Society employed in commerce and industry.

This subdivision of statisticians into subgroups is confirmed by an examination of their qualifications (See Sections 6.1 and 10.2.1). This shows up particularly in the percentages of PhD's in each group (60% for University staff, 43% for State and semi-state personnel, 22% for SASA members in the Private sector and 17% in the OR Society).

The information from Questionnaires No 1 and 2 established that the following eight methods were commonly used by all employee groups (See Section 10.1.1):

Basic statistics, Graphical display and data summary, Regression analysis, Analysis of variance, Correlation analysis, Forecasting, Multivariate analysis and Simulation.

Eight methods used commonly by only four of the employee groups were as follows (See Section 10.1.2):

Experimental design, Non-parametric tests, Survey sampling, Discriminant analysis, Factor analysis, Quality control, Non-linear estimation, Ranking and paired comparison data analysis.

The following three methods were used commonly by only three of the employee groups (See Section 10.1.3):

Categorical data analysis, Indices and Multidimensional scaling.

The remaining fourteen methods were only used commonly by two or less of the employee groups.

11.1.3. Differences in usage of statistical methods

Research Question 2 asked whether the method usages differed between employee groups. This question has been partially answered in the affirmative in the previous section. This answer is confirmed in Section 10.2.2, where it is shown that the usages of the statistical methods differed for the various categories of statisticians. The relative usages in fact classify statisticians into broad categories, namely, those employed by the Universities and by State and semi-state bodies on the one hand and those in industry and commerce and members of the Operations Research Society on the other hand. In between these two groupings, and over-lapping them to an extent, is a group consisting of members of the SASA who work in the Private sector. The methods in these groupings are set out in Table 10.5 in Section 10.2.2.

11.1.4. The effectiveness of statistical training at SA universities

Research Question 3 asks what teaching emphasis is given at university to the methods used by statisticians. Information was available from a questionnaire of Underhill (1982) on the teaching emphasis given to the various statistical methods by universities in South Africa. Rank correlation analysis in Chapter 7 showed that there was reasonable agreement between usages of the methods by the various employee groups and teaching emphasis. However, there are some anomalies which have been discussed in Section 10.3.

In particular, there was agreement across all five employee groups that Queueing theory was given too much emphasis in university teaching, while Simulation was given too little emphasis.

In the University and State and semi-state grouping it appeared that there was not enough teaching emphasis on Factor analysis, Categorical data analysis and Non-linear estimation, but that Inventory control received too much emphasis.

In the more business orientated groups there was too much on Non-parametric methods, Ranking and paired comparison data analysis and Categorical analysis.

11.1.5. <u>The adequacy of statistical curricula</u>

Research Question 4 enquired whether the statistical curricula offered by the universities provide an adequate training for consulting statisticians. The core question that requires answering is whether consulting statisticians considered their training had equipped them adequately for their careers. The question was asked of all members of the SASA and the OR Society, and 69 comments were received. Of these 68% were reasonably satisfied with their statistical training, while 25% specifically expressed discontent. The biggest complaint is that the training is too theoretical in nature. Even some of those who felt they had received a good basic knowledge, made requests for a greater practical orientation. Overall there were 80% of respondents who referred to the excessive emphasis on theory, or the need for a more practical approach, or who made suggestions for improving the practical training.

If one out of every four consulting statisticians is dissatisfied [illegible] or her training, then one is forced to conclude tha[illegible] serious deficiencies in our educational programme[illegible] Fortunately there is much that can be done to rectify the situation. A number of academics have already initiated schemes which will hopefully improve matters. These will be summarized in Section 11.2.3, and it is hoped that those that prove successful will be introduced at other universities.

11.1.6. The need for continuing education in statistics

Research Question 5 asks if there is a need for continuing education in statistics and, if so, what form this should take.

In Section 4.5.3 Juritz et al (1985) estimated that only half of the users of statistics in commerce and industry had some formal statistical training (See Table 4.8). Many of these had not majored in statistics. On the basis of this estimate it must be concluded that continuing education courses are essential for the 50% of users with no or very little statistical training. Even statistics majors have participated in the Wits courses and expressed satisfaction at having their knowledge refreshed and experiencing new insights into topics they felt they had known reasonably well.

The fact that approximately 100 people have attended the Wits courses during 1985 and 1986, and that 4 companies have requested in-house courses confirms the supposition that they are fulfilling a real need.

The form that the course takes depends on the particular field of activity and the interests of the participants. The format of the course given to Anglo American Gold Division (See Section 9.2.4) has proved very successful for technical and research staff of graduate or diploma level. The course includes an introductory section followed by techniques such as multiple regression, experimental design and quality assurance.

There is another advantage to the running of these courses. As stated in Section 1.8, teaching cannot stand in isolation. The contact with people from commerce and industry has all the benefits of industrial consulting. Designing and running a course for the actual users of statistics is a very helpful way of keeping an academic's feet firmly on the ground.

11.2. Suggestions for improving statistics curricula

The final and most important research question that needs to be answered concerns the improvements in curricula which can be suggested as a result of this study. The answer lies in the effective use of educational theory. As mentioned earlier, Gerrans (1986) has stated that "educational theory is to a teacher what mathematics is to physicist, or chemistry to a biologist." The phases of curriculum development as stated by Wheeler (1967:30) and set out in Section 1.6.1. will be used in Section 11.2.1 to 11.2.4 as a framework for the recommendations that arise from this study.

11.2.1. Aims and objectives of a curriculum

It is doubtful whether many statistics lecturers give much thought to the aims and objectives of their courses. When an academic is given the task of lecturing on a particular topic, the chances are that he or she will start by considering how much theory can be fitted into the given time period. Current textbooks on the particular topic will be consulted, and decisions made as to what aspects of the topic are needed for a systematic coverage of the subject matter. Time constraints may demand that certain aspects be omitted or dealt with more superficially than others. Requests may even be made for additional lectures to deal with the subject matter more fully.

Textbooks at the elementary level are usually a formal development of lecture notes used by some other lecturer for a course whose aims and target population may have been quite different from those for which they are now to be used. Textbooks on advanced topics tend to be expositions of the author's high-level research, which may have little practical application in the student's subsequent work situation.

In designing any course, the lecturer should start by giving some thought to the three views of education discussed in Section 1.6.2., namely:

(a) the process of passing our culture, in the form of a body of knowledge, from one generation to the next,

(b) the functional aspect of preparing statisticians for their profession, and

(c) the social service of exposing the young to the general benefits of education.

Most statistics courses will involve some emphasis on all three aspects, but will probably be dominated by one of them.

The interaction of these three aspects of education should not be the result of the haphazard, unplanned consequences of the lecturer's personal whims and preferences. There should be a formal aim to each course, and the aims of each course should be part of an overall, integrated plan for the university degree as a whole. It could well be that certain courses in the curriculum are deliberately biased towards one aspect of education and other courses to a different aspect, with balance being achieved overall.

This study has shown that there are two major subdivisions among statisticians. The statistical methods used by those employed in the Universities and by State and semi-state bodies are different from those employed in the business world. Any curriculum which is optimal for the development of one type of statistician will not necessarily suit the other type. It follows that on deciding on the aims of a curriculum, the major decision is whether to bias the training towards one or other group, or whether to have a general all-purpose curriculum which is not ideal for either. In a large department of statistics with many students it might be possible to have alternative options or topics that could separately satisfy most of the specialist needs of both types of statistician.

Even if a department of statistics is unable to offer an ideal programme, the staff should have a clearly defined policy in terms of written aims and objectives. Any deficiencies are then acknowledged and can perhaps be corrected at a later stage when the necessary resources become available.

11.2.2. Choice of course content

(i) General curriculum

The surveys carried out as part of this study have supplied considerable information on the usage of statistical methods in the work place. The thirty three methods have been ranked in terms of degree of use in Section 10.1. The eleven methods in Table 10.2. form a basic core which must obviously be included in the curriculum as they are in common use by all sections of the profession. They are as follows:

Basic statistical methods, Graphical display and data summary, Regression and Correlation analysis, Analysis of variance, Contingency table Chi-squared tests, Multivariate analysis, Linear programming, Simulation, Time series analysis and Forecasting.

Table 10.3. lists eight methods which, except for Quality control, are used commonly by the three groups of the SASA, but not by one of either the Business group or the OR Society. If, as discussed in Section 11.2.1., the aim of the curriculum is to offer a general training without specialisation, then these methods should also be included in the curriculum. The methods are as follows:

Design of experiments, Quality control, Non-linear estimation, Ranking and paired comparisons, Non-parametric methods, Survey sampling, Factor analysis and Discriminant analysis.

(ii) Specialized "theoretical" curriculum

If a department of statistics decides to offer specialised topics for students intending to make careers in the Universities or State and semi-state bodies, then the three methods in Table 10.4. could be added to the nineteen methods of the general curriculum in (i) above. These three methods are:

Categorical data analysis, Indices and Multidimensional scaling.

(iii) Specialized "applied" curriculum

The curriculum for the business orientated student should include the nineteen methods of the general curriculum in (i) above and perhaps some other applied methods such as:

Inventory control, Dynamic programming and Network models.

(iv) Additional methods

Table 10.5. lists methods which show marked differences in usage between employee groups. If additional topics are required then they can be chosen from Table 10.5. Thus the "theoretical" curriculum could have some of the following methods added:

Logistic regression, Variance component estimation, Ranking and paired comparison data analysis, Multidimensional scaling, Bioassay, Non-linear programming and Queueing theory.

(v) Other considerations in choice of method

It must be borne in mind that the methods suggested in (i) to (iii) above are chosen solely on the basis of frequent usage in consulting work. This means that the choice is made from a vocational viewpoint. As suggested in Section 11.2.1, a balance between the three aspects of education must always be striven for. Thus the cultural and social service aspects could well dictate the choice of one or more methods which are seldom used in practice.

11.2.3. Choice of learning experience

Gerrans (1986) has called for "a change in the role of the teacher from a giver of information to a manager of learning". As Aarons (1983) has pointed out, "experience makes it increasingly clear that verbal presentations - lecturing to large groups of intellectually passive students and having them read text material - leave virtually nothing in the student's mind that is permanent and significant".

What are the alternatives? Obviously a sudden and abrupt departure from the traditional lecture/tutorial format is not possible. The respondents to Questionnaire No 2 made a number of useful suggestions, which are presented here as suggestions for improving our teaching:

(i) An improved balance between theory and practice (Section 8.5.1).

Sixteen comments called for such improvement, which could be brought about by better planning through well thought out aims and objectives. Juritz (1986) has drawn attention to the advantages of setting out objectives in Section 8.5.1 (iii). If this is done then deficiencies in the teaching programme are immediately obvious.

(ii) Genuine data sets.

Six respondents commented on this (See Section 8.5.2). The Education Committee of the SASA has taken steps to rectify the excessive use of textbook examples by producing a data book with real-life examples from local applications. (See the Steffens Sub-committee in Section 1.1).

(iii) Practical work with outside organisations.

There were five comments. Again the Education Committee is aware of the problem and one of its sub-committees has been tasked with investigating the feasibility of compulsory vocation work for statistical students (See Section 1.1).

(iv) Practical projects.

Five respondents called for projects. This method of teaching has been tried with considerable success at Pretoria University and at the University of the Witwatersrand (See Section 8.5.4 (iii) and (iv)). The Education Committee has encouraged projects by holding an annual competition for undergraduates (See the Gilfillan Sub-committee in Section 1.1).

(v) Problems of statistical practice.

Six respondents commented on the problem they had adapting to the realities of statistical practice. One called for "more practical real-life training", while another complained that "my training lacked dismally to prepare me to handle any practical problem. I feel our undergraduate courses could do more to equip students with the necessary tools to tackle applications they might encounter in industry, agriculture and biometrics". The Snee committee recommended discussion of these difficulties and suggested how students could be helped to overcome them (See Section 8.5.5).

(vi) Exposure to statistical consulting.

Two respondents mentioned this need, which also appears high on the list of recommendations of the Snee committee. The author has been made acutely aware of this deficiency in our training (See Section 8.5.6 (iii)). Wits University have introduced a scheme for introducing Honours students into the consulting programme. (See Section 8.5.6 (vi)).

11.2.4. Course evaluation

Cerrans (1986) notes that when students' performances are assessed, usually by examinations and tests, very little information is obtained which helps teachers to teach. Examinations have a so-called "back-wash effect" in that they affect the student's study and learning habits. If examinations predominantly assess memory then students identify education with the gathering of information rather than recognising the need to develop higher cognitive skills.

An important way of getting feedback from the class on the effectiveness of the teaching process is by way of questionnaire evaluations during or at the end of the course. University staff as a whole are probably not very enthusiastic about exposing themselves to potential criticism from their students. However, there is much to be learned from such questionnaires, which have become traditional in continuing education courses.

Three examples of student evaluations of different learning experiences at the Universities of Cape Town, Pretoria and the Witwatersrand have been discussed earlier in Section 1.6.5. In all three cases the academics concerned felt that considerable benefit had accrued.

11.3. Overall conclusions

The first part of the overall aim of this study, namely, to ascertain what statistical methods are used by statisticians employed in different sectors of the South African economy, has been achieved. A considerably quantity of data has been collected which should enable the profession to plan for the future, without having to base projections on mere guesswork.

The second part of the aim was to evaluate whether South African universities are providing adequate training for the statisticians who use these methods. Here the answer probably has to be a qualified no. Statistical education in this country is acknowledged to be second to none from the theoretical or cultural viewpoint. Lombard (1981) considers the establishment of theoretical statistics to have been very successful. One of South Africa's most prominent and respected statisticians commented in the questionnaire that "Our theoretical coverage is of first class international standard, with actual applied statistics the area for the most room for improvement".

This study has shown that there is no need for despondency in our ability to correct any deficiencies that may exist in our teaching programmes. Areas of weakness have been pinpointed by comments from statisticians in the work place. There is a vast reservoir of knowledge in the literature of educational research, which has probably been sadly neglected by the teachers of statistics. There are undoubtedly great benefits to be obtained from its application

Above all we must never forget, as has been stressed by Gerrans (1986), that

> education does not stop at graduation; hence the ability to assume responsibility for one's own education throughout subsequent life is one of the most important skills that students should acquire. Of equal importance is the development of the ability to solve problems, to be able to confront new and unexpected situations with confidence, and to be able to ask questions rather than to expect to be given answers. If these attitudes and skills can be developed in our students, the details of the syllabus become of secondary importance.

BIBLIOGRAPHY

(N.B. Those references referred to in this dissertation are marked with an asterisk).

* Aarons,A.B.(1983), "Achieving Wider Scientific Literacy", Daedalus, 112, 91.

* Baskerville,J.C.(1981), "A Systematic Study of the Consulting Literature as an Integral Part of Applied Training in Statistics", American Statistician, 35, 121-123.

* Beard,R.M.(1972), Teaching and Learning in Higher Education, Penguin Books.

* Beard,R.M. and Hartley,J.(1984), Teaching and Learning in Higher Education, London: Harper and Row Publishers.

* Berk,R.A.(1981), Educational Evaluation Methodology:the State of the Art, London: John Hopkins University Press.

Bliss,C.I.(1969), "Communication Between Biologists and Statisticians, A Case Study", The American Statistician, 23, 15-20.

Boen,J.R.(1972), "The Teaching of Personal Interaction in Statistical Consulting", The American Statistician, 26, 30-31.

Boen,J.R., and Fryd,D.(1978), "Six-State Transactional Analysis in Statistical Consultation", The American Statistician, 32, 58-60.

Bowker,A.H.(1965), "Quality and Quantity in Higher Education", Journal of the American Statistical Association, 60, 1-15.

Box,G.E.P, Hunter,W.G., and Hunter,J.S. (1978), Statistics for Experiments: An Introduction to Design, Data Analysis, and Model Building, New York: John Wiley.

Bross,I.D.J.(1974), "The Role of the Statistician:Scientist or Shoe Clerk", The American Statistician, 28, 126-127.

* Calvert,A.J.R.(1979), "Indications of Performance", Proceedings of 15th Annual Conference of Society for Research into Higher Education, Dorset: Direct Printers.

Cameron,J.M.(1969), "The Statistical Consultant in a Scientific Laboratory", Technometrics, 11, 247-254.

Carter,A.M.(1966), An Assessment of Quality in Graduate Education, Washington, D.C.: American Council of Education.

Carter,R.L., Scheaffer,R.L. and Marks,R.G. (1986), "The Role of Consulting Units in Statistics Departments", The American Statistician, 40, 260-264.

* Cochran,W.G.(1977) Sampling Techniques, New York: John Wiley.

Cox,C.P.(1968), "Some Observations on the Teaching of Statistical Consulting", Biometrics, 24, 789-802.

* Daniel,C.(1969), "Some General Remarks on Consultancy in Statistics", Technometrics, 11, 241-246.

Deming,W.E.(1965), "Principles of Professional Statistical Practice", Annals of Mathematical Statistics, 36, 1883-1900.

* Duncan,J.W. and Durbin,J.(1980), "Report of the International Statistical Institute on the Integration of Statistics", International Statistical Review, 48, 139-168.

* Eldridge,M.D., Wallman,K.C., Wulfsberg,R.M., Bailer,B.A., Bishop,Y.M., Kibler,W.E., Orleans,B.S., Rice,D.P., Schaible,W., Selig,S.M., and Sirken,M.G.(1982), "Preparing Statisticians for Careers in the Federal Government", The American Statistician, 36, 69-81.

Entwhistle,N. and Hounsell,D.(1977), eds., How students learn, University of Lancaster.

* Fatti,L.P.(1984), "Case Studies", in Symposium on Statistical Education in South Africa, 16-17.

Fellegi,I.(1978), "Some Thoughts on Training Needed for Government Statistical Work", in Report of the American Statistical Association Conference on the Transfer of Methodology Between Academic and Government Statisticians, Washington: American Statistical Association, 43-52.

* Fellingham,S.A. and Kotze,T.J.V.W.(1981), "Manpower survey of members of the S.A. Statistical Association", Newsletter of S.A. Statistical Association, Dec 1981.

* Fienburg,S.E.(1978), "The Analysis of Cross-Classified Categorical Data", Cambridge, Massachussets: The MIT Press.

* Fink,A. and Kosecoff,J.(1978), An Evaluation Primer, London: Sage Publications.

Gehan,E.A.(1980), "The Training of Statisticians for Cooperative Clinical Trials: A Working Statisticians Viewpoint", Biometrics, 36, 699-706.

* Gerrans,G.C.(1986), Memorandum on Some Aspects of University Teaching and Learning, Johannesburg: University of Witwatersrand.

* Goldschmid,M.L.(1978), "The Efficiency and Effectiveness of Teaching in Higher Education", University of London Teaching Methods Unit, Leicester: Cavendish Press.

* Greenfield,A.A.(1979), "Statisticians in Industrial Research: The Role and Training of the Industrial Consultant", The Statistician, 28, 71-82.

* Griffiths,J.D. and Evans B.E.(1976), "Project Work in Statistics", The Statistician, 25, 117-123.

Guba,E.G.(1978), Toward a Methodology of Naturalistic Inquiry in Educational Evaluation, C.S.E. Monograph Series in Evaluation, No 8, Los Angeles: Centre for the Study of Evaluation, University of California Los Angeles.

Hammond,D.(1980), "The Training of Clinical Trials Statisticians: A Clinican's View", Biometrics, 36, 679-685.

* Heywood,J.(1984), Considering the Curriculum during Student Teaching, London: Kogan Page.

Hooke,R.(1980), "Getting People to Use Statistics Properly", The American Statistician, 34, 39-42.

* Hunter,W.G.(1977), "Some Ideas About Teaching Design of Experiments, with 2^5 Examples of Experiments Conducted by Students", The American Statistician, 31, 12-17.

Hyams,L.(1971), "The Practical Psychology of Biostatistical Consultation", Biometrics, 27, 201-212.

* Jones,J.D. and Kanji,G.K.(1980), "The Role of Professional Experience in Statistical Education", The Statistician, 29, 196-203.

* Juritz,J.M.(1986), "Teaching Applied Statistics with BMDP using a class questionnaire", Paper presented to Second International Conference on Teaching Statistics, University of Victoria, B.C., Canada.

* Juritz,J.M., Affleck-Graves,J., Du Toit,P.S. and Money,A.M. (1985), "Survey of the Statistical Methods used in Business and Industry in South Africa", Unpublished manuscript.

* Kanji,G.K.(1979), "The Role of Projects in Statistical Education", The Statistician, 28, 19-27.

Kimball,A.W.(1957), "Errors of the Third Kind in Statistical Consulting", Journal of the American Statistical Association, 57, 133-142.

* Lombard,F.(1981), "The neglect of applied statistics in South Africa", Presidential Address to the S.A. Statistical Association, Newsletter of S.A. Statistical Association, Dec., 1981.

Malinvaud,E.(1978), "University Teaching for Future Official Statisticians", in Report of the American Statistical Association Conference on the Transfer of Methodology Between Academic and Government Statisticians, Washington: American Statistical Association, 53-67.

* Marquardt,D.W.(1979), "Statistical Consulting in Industry", The American Statistician, 33, 102-107.

Meads,R.(1976), "Statistical Consulting in a University", The Statistician, 25, 213-218.

Milton,P.D., and Freund,R.J.(1977), "Organization for the Conduct of Statistical Activities in Colleges and Universities", The American Statistician, 3, 113-117.

Modular Instruction in Statistics, (1976), Report of the American Statistical Association Study in Modular Instruction, American Statistical Association, Washington, D.C.

* Morris,L.L. and Fitz-Gibbon,C.T.(1978), How to Deal with Goals and Objectives, London: Sage Publications.

Murthy,M.N.(1978), "Statistical Education, Training, Research and Their Utilization," in Report of the American Statistician Association Conference on the Transfer of Methodology Between Academic and Government Statisticians, Washington: American Statistical Association, 68-93.

* Newman,J.H.(1947), The Idea of a University, London: Longmans Green.

* Patton,M.Q.(1980), Qualitative Evaluation Methods, London: Sage Publications.

* Patton,M.Q.(1981), Creative Evaluation, London: Sage Publications.

* Piper,D.W.(1978), ed. The Efficiency and Effectiveness of Teaching in Higher Education, University of London Teaching Methods Unit, Leicester: Cavendish Press.

* Rao,C.R.(1982), "Optimum balance between statistical theory and applications in teaching", in Proceedings of the First International Conference on Teaching Statistics, 1, 34-49.

* Rayner,A.A.(1969), A First Course in Biometry for Agriculture Students, Pietermaritzburg: University of Natal Press.

Roberts,H.V.(1978), "Statisticians can matter", The American Statistician, 32, 45-51.

* Romney,L.C., Gray,R.G. and Weldon,H.K.(1978), "Departmental Productivity: A Conceptual Framework", National Centre for Higher Education Management Systems, Denver, Colorado.

Schucany,W.R.(1972), "Some Remarks on Educating Problem Solvers", in Proceedings of the Computer Science and Statistics Sixth Annual Symposium on the Interface, Berkeley, California: University of California Press, 28-32.

Snedecor,G.W.(1950), "The Statistical Part of the Scientific Method", Annals of the New York Academy of Sciences, 52, 792-799.

* Snee,R.D., Boardman,T.J., Hahn,G.J., Hill,W.J., Hocking,R.R., Hunter,W.G., Lawton,W.H., Ott,R.L., and Strawderman,W.E.(1980), "Preparing Statisticians for Careers in Industry", The American Statistician, 34, 65-75.

Sprent,P.(1970), "Some Problems of Statistical Consultancy", Journal of the Royal Statistical Society, Ser.A, 133, 139-165.

* Steyn,A.G.W.(1981), "Samewerking Tussen Statistici", Newsletter of S.A. Statistical Association, Sept., 1981.

* Steyn,A.G.W.(1984), "Projekwerk", in Symposium on Statistical Education in South Africa, 20-22.

* Tanur,J.M.(1982), "Quality of Statistical Education: Should ASA Assist or Assess?", The American Statistician, 36, 90-93.

Tarter,M., and Berger,B.(1972), "On the Training and Practice of Computer Science and Statistical Consultants", in Proceedings of the Computer Science and Statistics Sixth Annual Symposium on the Interface, Berkeley: University of California Press, 16-23.

* Underhill,L.G.(1982), "University statistical education. Survey of South African University Statistics Departments", Report to Education Committee of S.A. Statistical Association.

* Underhill,L.G.(1984), "Choice of Tutorial Matter", in Symposium on Statistical Education in South Africa, 7-12.

* University of London Teaching Methods Unit(1976), "Improved Teaching in Higher Education", Leicester: Cavendish Press.

Watts,D.G.(1970), "A Program for Training Statistical Consultants", Technometrics, 4, 737-740.

* Wegner,T.(1983), "A Survey of Quantitative Methods in South African Management", South African Business Management, 14, 89-95.

* Wheeler,D.K.(1967), Curriculum Process, London: Hodder and Stoughton.

* Whittle,P.A.(1974), "Teacher-Pupil Interaction" in The Art of the Science Teacher, eds. Sutton,C.R. and Haysom,J.T., McGraw Hill.

Woodward,W.A., and Schucany,W.R.(1977), "Bibliography for Statistical Consulting", Biometrics, 33, 564-565.

Workshop on Statistical Consulting(1979), Moderator,M.B.Wilk, in Proceeding of the Computer Science and Statistics 12th Annual Symposium on the Interface, ed. J.F. Gentleman, Ontario: University of Waterloo, 69-95.

* Zahn,D.A. and Isenberg,D.G.(1983), "Non-statistical Aspects of statistical consulting", The American Statistician, 37, 297-302.

APPENDIX A

UNIVERSITY OF CAPE TOWN

(WITH WHICH IS INCORPORATED THE SOUTH AFRICAN COLLEGE)

Telephone 698531

DEPARTMENT OF
MATHEMATICAL STATISTICS
RONDEBOSCH
7700

Dear Sir,

THE STATISTICAL EDUCATION PROJECT

Members of the Departments of Mathematical Statistics at the University of Cape Town, the University of the Witwatersrand amd the Graduate School of Business, U.C.T., are engaged on a project on statistical education. Our aim is to improve the preparation of the students who intend making their careers in commerce and industry.

Information of the statistical methods used by your company would be of great assistance to us in designing the courses we give. We enclose a short questionnaire and ask if you could complete it and return it to us. We can assure you that all information you give us will be confidential. You need not identify your company when you reply if you do not wish to do so.

Thank you for your co-operation, we greatly appreciate it. If you would like any more information please contact one of us at the address given at the end of the questionnaire.

Yours truly,

J. M. Juritz.

Prof. J.M. Juritz
Prof. A.H. Money
Prof. J. Affleck-Graves
Mr. Pierre du Toit

OFFICE USE

NAME OF COMPANY (Can be left blank) ....

ADDRESS ....

....

....

TELEPHONE NUMBER ....

NAME OF PERSON RESPONSIBLE FOR COMPLETING THIS QUESTIONNAIRE

(CAN BE LEFT BLANK) ....

POSITION IN COMPANY ....

1. Type of activity the company is involved in.

1	Banking	
2	Building & Construction	
3	Chemical	
4	Manufacturing	
5	Marketing	
6	Mining	
7	Other. Please specify	

2. Is your company a

1	Local Company	
2	National Company	
3	International Company	

3. Total number of employees in the organisation

1	100 or Fewer	
2	101 to 500	
3	501 to 1000	
4	More than 1000	

4. Are statistical methods used by the company?

1	Yes	
2	No	

5. Are Operations Research methods used by the company?

1	Yes	
2	No	

6. Does your company use quality control?

1	Yes	
2	No	

7. (a) Does the company employ a Statistician and/or Operations Researcher?

1	Yes	
2	No	

(b) Does your company use a Statistical Consultant?

1	Yes	
2	No	

8. My company does not use statistical methods because

1	They are not applicable to our business	
2	There are no staff with the necessary expertise	
3	We have done so in the past and found them useless	
4	We know nothing about them	
5	Other (please comment)	

If your company does NOT use Statistical/Operations Research methods please answer question 1[illegible] next and return the questionnaire to the address given on page 7.

9. USER OF STATISTICAL AND/OR OPERATIONS RESEARCH TECHNIQUES.

9.1 How many people in your company use these techniques? [ ]

9.2 Who applies the Statistical/Operations Research methods?

1	Matriculant		
2	Bachelor's Graduate	with Statistics/Operations Research Major	
3	" "	with some Statistics/Operations Research courses	
4	" "	with no Statistics/Operations Research courses	
5	Honour's Graduate	in Statistics/Operations Research	
6	" "	in another discipline	
7	Master's Graduate	in Statistics/Operations Research	
8	" "	in another discipline	
9	Doctor's Graduate	in Statistics/Operations Research	
10	" "	in another discipline	
11	Other	please specify	

9.3 How often are the Statistics/Operations Research techniques applied?

1	Daily	
2	Weekly	
3	Monthly	
4	Quarterly	
5	Annually	
6	Less than once a year	

9.4 For how long have these methods been used by your company

1	Less than 1 year	
2	Between 1 and 5 years	
3	More than 5 years	

9.5 In which department are the Statistics/Operations Research persons situated?

1	Data Processing	
2	Management Accounting	
3	Planning	
4	Statistics/Operations Research	
5	Marketing	
6	Other - Specify	

9.6 Please rate the effectiveness of these techniques on the scale below.

Useless 1 2 3 4 5 6 7 Very Effective

Comment:

9.7 Given below is a list of Statistical/Operations Research techniques. Please tick the ones your company uses and indicate how frequently they are used.

		Occasionally	Routinely
01	Basic Statistical Methods (means;standard deviations)		
02	Graphical display and data summary		
03	Regression Analysis		
04	Correlation Analysis		
05	Analysis of Variance		
06	Design of Experiments		
07	Variance Component Estimation		
08	Bioassay		
09	Contingency Tables Chi-Square tests		
10	Quality Control and Acceptance Sampling		
11	Non-linear Estimation		
12	Non-Parametric Methods		
13	Ranking and Paired Comparison Data Analysis		
14	Multivariate Analysis		
15	Logistic Regression		
16	Linear Programming		
17	Non-linear Programming		
18	Dynamic Programming		
19	Inventory Control		
20	Simulation		
21	Reliability and Life Testing Data Analysis		
22	Time Series Analysis		
23	Forecasting		
24	Survey Sampling		
25	Categorical Data Analysis		
26	Queueing Theory		
27	Factor Analysis		
28	Decision Theory		
29	Network Models (PERT/CPM)		
30	Multidimensional Scaling		
31	Conjoint Analysis		
32	Indices		
33	Discriminant Analysis		
34	OTHER. Please specify.		
35			
36			

9.8 Given below is a list of Statistical/Operations Research computer packages. Please indicate the ones used and indicate how frequently.

		Never	Occasionally	Routinely.
1.	BMDP			
2.	COMET			
3.	GLIM			
4.	SPSS			
5.	SAS			
6.	STATJOB			
7.	FMPS			
8.	MDSX			
9.	SIBYL RUNNER			
10.	BOX-JENKINS FORECASTING			
	NAME OTHER (IF USED)			

Does your company use a micro, or mini computer?

1	YES
2	NO

If yes: what type?

10. Would you like to receive the results of this survey?

1	Yes	
2	No	

11. Would your company be interested in sending personnel to seminars on the applications of Statistical/Operations Research techniques?

1	Yes	
2	No	
3	Maybe	

Thank you for your co-operation.

Please return the questionnaire in the enclosed envelope.

If you have any queries please contact:

The Statistical Education Project,

Prof. June M. Juritz,
Dept. of Mathematical Statistics,
University of Cape Town,
RONDEBOSCH.
7700

Ph. (021) 69.8531 Ext. 414

Prof. J. Affleck-Graves,
Graduate School of Business,
University of Cape Town,
RONDEBOSCH.
7700

Ph. (021) 69.5382 Ext. 31.

Mr. Pierre du Toit,
Dept. of Statistics,
University of the Witwatersrand,
1, [illegible]muts Avenue,
JOHANNESBURG.
2000

Ph. (011) 716.3310

Prof. A.H. Money,
Graduate School of Business,
University of Cape Town,
RONDEBOSCH.
7700

Ph. (021) 69.5382 Ext. 12.

APPENDIX B1

STATISTIESE ONDERWYSPROJEK

S.A. STATISTIESE VERENIGING

Die Subkomittee van die Onderwyskomittee van die SASV insake opleiding van Universiteitstudente is tans besig om inligting in te samel oor watter statistiese tegnieke deur werkgewers van statistici gebruik word. Die doel van hierdie studie is om die statistiese opleiding op universiteitsvlak te verbeter. 'n Vraelys is verlede jaar aan 'n steekproef van ons belangrike maatskappye uitgestuur. Om 'n gebalanseerde beeld van die gebruik van statistiek in ander buite-instansies te verkry word die volgende vraelys aan alle lede van die SASV en die Operasionele Navorsings-vereniging van S.A. gestuur. Ons sal dit hoog op prys stel as u hierdie vraelys sal beantwoord en aan Mnr P S du Toit, Department van Statistiek, Universiteit van die Witwatersrand, 1 Jan Smuts Laan, Johannesburg, 2001 terugpos.

STATISTICAL EDUCATION PROJECT

S.A. STATISTICAL ASSOCIATION

The Education Sub-committee of the SASA on University Training is presently gathering information on the statistical techniques used by employers of statisticians. The aim of the exercise is to improve the teaching of statisticians at university level. Last year a questionnaire was sent to a sample of our major companies. In order to obtain a balanced picture of the use of statistics in other institutions the following questionnaire is being sent to all members of the SASA and the Operations Research Society of S.A. We would greatly appreciate your co-operation in answering the questionnaire and returning it to Mr P S du Toit, Department of Statistics, University of the Witwatersrand, 1 Jan Smuts Avenue, Johannesburg, 2001.

QUESTIONNAIRE ON THE USE OF STATISTICAL TECHNIQUES

(Om ruimte te bespaar is hierdie vraelys in Engels opgestel. Antwoord asseblief in Afrikaans as u dit verkies.)

1. Name .... 2. Member of SASA [ ] or O R Society [ ] or both [ ]

3. Institution/Company.......................... 4. Your position in Institution/Company ....

Address ....

5. Tick your academic qualifications:

Matric (or equivalent)		University degree(s) (Specify)	
Technical Certificate or Diploma		Other (Specify)	

6. What were your major subjects .... 7. If Statistics was not your major, state what Statistics courses you have attended...

8. Does your Institution/Company use Statistical methods [ ], Operations Research methods [ ], Quality Control [ ].

9. Given below is a list of techniques. Please tick the ones used in your present or previous jobs, showing frequency of use.

		Occasionally	Routinely			Occasionally	Routinely
01	Basic Statistics (means & SD's)			18	Dynamic Programming		
02	Graphical display & data summary			19	Inventory Control		
03	Regression Analysis			20	Simulation		
04	Correlation Analysis			21	Reliability & Life Testing Data Analysis		
05	Analysis of Variance			22	Time Series Analysis		
06	Design of Experiments			23	Forecasting		
07	Variance Component Estimation			24	Survey Sampling		
08	Bioassay			25	Categorical Data Analysis		
09	Contingency Tables Chi-square tests			26	Queueing Theory		
10	Quality Control & Acceptance Sampling			27	Logistic Regression		
11	Non-linear Estimation			28	Linear Programming		
12	Non-parametric Methods			29	Non-linear Programming		
13	Ranking & Paired Comparison Data Analysis			30	Decision Theory		
14	Multivariate Analysis			31	Network Models (PERT/CPM)		
15	Discriminant Analysis			32	Multidimensional Scaling		
16	Principal Component & Factor Analysis			33	Indices		
17	Conjoint Analysis			34	Other (Specify)		

10. Given below is a list of computer packages. Please tick the ones used, showing frequency of use by entering O for Occasionally and R for Routinely in the appropriate blocks.

1. BMDP [ ] 2. COMET [ ] 3. GLIM [ ] 4. SPSS [ ] 5. SAS [ ] 6. STATJOB [ ] 7. FMPS [ ]
8. MDSX [ ] 9. SIBYL RUNNER [ ] 10. Packages with Box-Jenkins (Specify) [ ] 11. Others (Specify)______

11. Does your Institution/Company use a micro or mini-computer (Yes or No)?......

If yes what type(s)? ....

Thank you very much for assisting is with this study. We will publish the results in the respective newsletters.

May we ask one last question? Please comment on whether you felt your statistical training gave you the skills necessary for the adequate performance of your job(s). Suggestions for improved training would be appreciated.

....

....

....

....

....

....

APPENDIX B2

STATISTIESE ONDERWYSPROJEK

S.A. STATISTIESE VERENIGING

Die Subkomitee van die Onderwyskomitee van die SASV insake Opleiding van Universiteitstudente is tans besig om inligting in te samel oor watter statistiese tegnieke deur werkgewers van statistici gebruik word. Die doel van hierdie studie is om die statistiese opleiding op universiteitsvlak te verbeter. 'n Vraelys is verlede jaar aan 'n steekproef van belangrike maatskappye uitgestuur. Om 'n gebalanseerde beeld van die gebruik van statistiek in ander instansies te verkry word die volgende vraelys aan alle lede van die SASV en die Operasionele Navorsingsvereniging van S.A. gestuur. Ons sal dit hoog op prys stel as u hierdie vraelys sal beantwoord en aan Mnr P S du Toit, Departement van Statistiek, Universiteit van die Witwatersrand, Jan Smutsrylaan 1, Johannesburg, 2001, terugpos.

STATISTICAL EDUCATION PROJECT

S.A. STATISTICAL ASSOCIATION

The Education Sub-committee of the SASA on University Training is presently gathering information on the statistical techniques used by employers of statisticians. The aim of the exercise is to improve the teaching of statisticians at university level. Last year a questionnaire was sent to a sample of our major companies. In order to obtain a balanced picture of the use of statistics in other institutions the following questionnaire is being sent to all members of the SASA and the Operations Research Society of S.A. We would greatly appreciate your co-operation in answering the questionnaire and returning it to Mr P S du Toit, Department of Statistics, University of the Witwatersrand, 1 Jan Smuts Avenue, Johannesburg, 2001.

(sien keersy)
(see reverse side)

QUESTIONNAIRE ON THE USE OF STATISTICAL TECHNIQUES

(Om ruimte te bespaar is hierdie vraelys in Engels opgestel. Antwoord asseblief in Afrikaans as u dit verkies.)

1. Name R.F.K[illegible]
2. Member of SASA [X] or O R Society [ ] or both [ ]
3. Institution/Company Medunsa
4. Your position in Institution/Company Medical Natural Scientist

Address ....

5. Tick your academic qualifications : Matric (or equivalent) [X] Technical Certificate or Diploma [ ] University degree(s) Specify B.Sc [X] Other (Specify) [ ]

6. What were your major subjects Mathematics ~~Stat~~ Mathematical Statistics ....

7. If Statistics was not your major, state what Statistics courses you have attended ....

8. Does your Institution/Company use Statistical methods [ ] Operations Research methods [ ] Quality Control [X].

9. Given below is a list of techniques. Please tick the ones used in your present or previous jobs, showing frequency of use. O indicates occasional usage and R routine usage.

01	Basic Statistics (means & SD's)	✓	R	18	Dynamic Programming	O	R
02	Graphical display & data summary	O	R	19	Inventory Control	O	R
03	Regression Analysis	O	R	20	Simulation	O	R
04	Correlation Analysis	O	R	21	Reliability & Life Testing Data Analysis	O	R
05	Analysis of Variance	O	R	22	Time Series Analysis	O	R
06	Design of Experiments	O	R	23	Forecasting	O	R
07	Variance Component Estimation	O	R	24	Survey Sampling	O	R
08	Bioassay	O	R	25	Categorical Data Analysis	O	R
09	Contingency Tables Chi-square tests	O	R	26	Queueing Theory	O	R
10	Quality Control & Acceptance Sampling	O	R	27	Factor Analysis	O	R
11	Non-linear Estimation	O	R	28	Decision Theory	O	R
12	Non-parametric Methods	O	R	29	Network Models (PERT/CPM)	O	R
13	Ranking & Paired Comparison Data Analysis	O	R	30	Multidimensional Scaling	O	R
14	Multivariate Analysis	O	R	31	Conjoint Analysis	O	R
15	Logistic Regression	O	R	32	Indices	O	R
16	Linear Programming	O	R	33	Discriminant Analysis	O	R
17	Non-linear Programming	O	R	34	Other (Specify)	O	R

10. Given below is a list of computer packages. Please tick the ones used, showing frequency of use by entering O for Occasionally and R for Routinely in the appropriate blocks.

1. BMDP ___ 2. COMET ___ 3. GLIM ___ 4. SPSS ___ 5. SAS ___ 6. STATJOB ___ 7. FMPS ___ 8. MDSX ___
9. SIBYL RUNNER ___ 10. BOX-JENKINS FORECASTING ___ 11. Others (Specify) NWASIATPACK....... ___

11. Does your Institution/Company use a micro or mini-computer (Yes or No)? .. Yes
If Yes, what type(s)? HP 150 , HP 9845B, Olivetti

Thank you very much for assisting us with this study. We will publish the results in the respective newsletters. May we ask one last question? Please comment on whether you felt you statistical training gave you the skills necessary for the adequate performance of your job(s). Suggestions for improved training would be appreciated.

I don't use my Statistical training ....
....
....
....
....
....

(sien keerey)
(see reverse side)

APPENDIX C

BESONDERHEDE IN VERBAND MET DIENSKURSUSSE IN STATISTIEK WAT GEDURENDE 1983 AAN U UNIVERSITEIT DEUR U DEPARTEMENT AANGEBIED IS.

'n Dienskursus in statistiek word beskou as 'n verpligte inleidende kursus wat deur studente gevolg word wat primêr in ander vakdissiplines studeer. Hierdie kursusse is gewoonlik 'n vereiste vir die verwerwing van 'n graad of diploma, maar word ook soms op 'n ad hoc basis gereël sonder dat dit in die jaarboek gespesifiseer is. Die verpligte eerstejaar statistiekkursus vir B.Com studente moet as 'n dienskursus beskou word alhoewel sommige van die studente dit as 'n hoofvak mag neem.

NAAM VAN UNIVERSITEIT: DURBAN - WESTVILLE

FAKULTEIT WAARBINNE U DEPARTEMENT VAL VIR BEGROTINGSDOELEINDES:
HANDEL EN ADMINISTRASIE

Voltooi vir elke dienskursus:

1. Kursusnaam: STATISTIEK 112

Semesterkursus	Jaarkursus	Ander (spesifiseer)
X		

- Vir watter graad/studierigting is hierdie kursus verpligtend? B. Com.
- Aantal lesings per week: 5
- Aantal studente in 1983: 250 ±
- Naam van handboek wat in 1983 gebruik is: FREUND + WILLIAMS : BUSINESS STATISTICS - THE MODERN APPROACH + SCHAUM OUTLINE SERIES - THEORY OF FINANCE
- Leerplan (baie kortliks): WAARSKYNLIKHEID, ANALISE VAN TYDREEKSE INDEKSSYFERS, LINIÊRE PROGRAMMERING, SAAMGESTELDE RENTE EN JAARGELDE
- Verdere opmerkings oor hierdie kursus:

2/...

2. . Kursusnaam: ______________________

Semesterkursus	Jaarkursus	Ander (spesifiseer)

. Vir watter graad/studierigting is hierdie kursus verpligtend? ______________

. Aantal lesings per week: ______________

. Aantal studente in 1983: ______________

. Naam van handboek wat in 1983 gebruik is: ______________

. Leerplan (baie kortliks): ______________

. Verdere opmerkings oor hierdie kursus: ______________

3. . Kursusnaam: ______________________

Semesterkursus	Jaarkursus	Ander (spesifiseer)

. Vir watter graad/studierigting is hierdie kursus verpligtend? ______________

. Aantal lesings per week: ______________

. Aantal studente in 1983: ______________

. Naam van handboek wat in 1983 gebruik is: ______________

. Leerplan (baie kortliks): ______________

. Verdere opmerkings oor hierdie kursus: ______________

(Kopieer hierdie bladsy indien meer as drie dienskursusse aangebied word)

3/...

By baie universiteite word dienskursusse in statistiek deur ander departemente aangebied of vorm dit 'n integrale deel van die leerplanne van ander kursusse.

Meld (indien u wil) die name van sodanige kursusse

SIELKUNDE III

STATISTICS + EXPERIMENTAL DESIGN (INGINIEURSWESE)

MATHEMATICS OF FINANCE + STATISTICS (M.B.A.)

APPENDIX D1

Centre for Continuing Education, University of the Witwatersrand, Johannesburg

Statistics Courses: June 1985

Evaluation Questionnaire : Do not sign

Mark your answer with a cross in the appropriate block

1. Will this course be helpful to you in your current (or anticipated future) position?

No	Slightly	Yes	Very much

2. Did you find it interesting personally?

No	Slightly	Yes	Very much

3. Was the general level about right?

No, too elementary	Yes, about right	No, too advanced

4. Rate the following sections of the course:

	Would have liked more	About right	Would have liked less	Presentation			Notes and Handouts		
				Poor	Av.	Good	Poor	Av.	Good
Descriptive Stats.									
Distributions									
Significance tests									
χ^2 tests									
Analysis of Variance									
Experimental Design									
Correlation & Regn.									
Quality Assurance									
Accept. Sampling									
Cusums									
Analytical Assur.									

5. What did you like best about the course?

____

6. What did you like least about the course?

____

7. Would you have preferred the present course in a different format, bearing in mind the willingness of your company to (a) release you for longer periods, and (b) pay for additional modules?

Suggest OVERALL course in following format			
As presented viz 9×3=27hrs	Daily over fortnight eg 14×2=28 hrs	Alternate days eg 14×2=28 hrs	Other (please specify)

OR

Suggest course in separate modules eg Introductory Stats, Correlation & Regression, Experimental Design, Quality Assurance, Analytical Assurance.		
Each Module 5×3=15 hrs	Each Module 10×2=20 hrs	State preferred format for each module

8. Did the time 4-7 p.m. suit you? | Yes | No |

If No, please state your preferred time, bearing in mind the willingness of your company to release you.

____

9. Is there a preferred time of year? | Yes | No | If Yes, please state

____

10. What other improvements would you suggest for the present course?

____

11. Which of the following separate in-depth courses would you or your colleagues be interested in:

Probability Distributions	Experimental Design	Multiple Regression	Quality Assurance	Analytical Assurance
Industrial Sampling	Time Series & Forecasting	Use of Computer Package SAS	Other (please state)	

12. Should such specialist courses be at

9 a.m.-5 p.m.	2 p.m. - 5 p.m.	4 p.m. - 7 p.m.	Daily
State other preferred times			Alternate days

13. Did you find out about this course by receiving a brochure from the professional society to which you belong? | Yes | No |

If Yes, state which society ______________________

If No, please state how you found out ______________________

____

14. Any other comments?

____

____

____

____

____

____

____

APPENDIX D2

STATISTICS COURSE : 18-21 NOVEMBER, 1985
FINAL EVALUATION QUESTIONNAIRE
DO NOT SIGN

1. Will this course he helpful to you in your current (or anticipated future) position?

2. Did you find it interesting?

NO	SLIGHTLY	YES	VERY MUCH

3. Was the general level about right?

No, too elementary
Yes, about right
No, too advanced

4. Rate the following sections of the course content:

		Mean SD Histogram Probab.	Binomial Poisson NDX^2	Test of Signif.	Regression	Anal. of Variance	Exptal Design
I would have liked more time spent on this							
I would have liked less time spent on this							
Presentation (Reasons below if you want to)	Good						
	Av.						
	Poor						

5. What did you like best about the course? ....
....

6. What did you like least about the course? ....
....

7. Other comments you may like to make: ....
....
....
....

APPENDIX E

The 65 comments received in answer to the question in Questionnaire No 2 on the adequacy of training (See Section 8.1.) are listed below. They have all been used in the text, either in full or as part-quotes.

In parenthesis below each comment is shown its code number and the employee group to which the respondent belongs.

> "It gave me the necessary background for the performance of my job".
>
> (1 SASA: State)

> "I work as an epidemiologist, not a data analyst. My training in statistics was more than adequate (Note: I trained in Canada in Epidemiology and Statistics)".
>
> (2 SASA: State)

> "As I am not employed as a Statistician I can't really comment. However, I feel that my training was adequate".
>
> (3 OR Society)

> "Yes as a lecturer in Statistics/Biometry".
>
> (4 SASA: Univ)

"Dit is onmoontlik om 'n statistikus in vier jaar op te lei. Die Universiteitskursusse moet die nodige onderbou gee - die praktiese opleiding sal nooit bevredigend en voldoende op universiteit kan geskied. 'n Deeglike twee jaar opleiding na voltooiing van 'n vier-jarige universiteitsgraad is daarvoor nodig".

(5 SASA: State but ex Univ)

"No - I do not think any training can provide adequate skills over a short period of (at most) 5 years. You can only develop a feeling for what the subject is all about in formal training. The rest is experience and cooperation with colleagues and self-training: READ,READ,READ.....".

(6 SASA: Univ)

"The training supplied me with a very good basic 'tool-kit'. It has been up to me to learn the skills in the use of the 'tools' supplied and to acquire more specialised 'tools' for solving certain problems. Statistics is a VERY broad field and I do not see how an undergraduate course can improve on supplying a sound, basic 'tool-kit'".

(7 SASA: State)

"My training gave me the basic tools and the ability to assimilate more specialised techniques at a later stage".

(8 OR Society)

"My statistics training gave me a useful background but my practical experience and application were self taught".

(9 SASA: Private)

"In some areas, the statistical training has helped in my work. In others, the statistics required would not be adequately provided at the level of formal statistical education undertaken to date as many aspects are treated in depth only at post-graduate statistical levels. This applies especially to the 'specialist' techniques. Inclusion of 'advanced training' at undergraduate level would not really be appropriate in view of the broad spectrum of topics which need to be covered".

(10 SASA: Private)

"Reasonable, however, the year of honours (i.e. fourth year) is essential as not enough practical problems are handled in the undergraduate course".

(11 OR Society)

"Only comes with experience. Good theoretical background necessary. Too much emphasis in exams on proofs of theorems - students should be taught to think!"

(12 SASA: Univ)

Opleiding aanvanklik onvoldoende maar met voortgesette studie wel voldoende".

(13 SASA: Univ)

"No. Necessary skills can only develop over time".

(14 SASA: State)

"I don't believe that even an Honours degree in Statistics is adequate training. One must gain practical experience before one can obtain a true understanding of the theory".

(15 SASA: State)

"I feel that my statistics training gave me <u>some</u> skills. MODELLING, DYNAMIC PROGRAMMING, LINEAR PROGRAMMING & SIMULATION were self-taught".

(16 OR Society)

"My courses were too theoretical. They were techniques looking for problems to which they might possibly be applied".

(17 OR Society)

"My statistical training was mostly academic. I still encounter some few problems when applying the theory which I learned to practical problems".

(18 SASA: Univ)

"Nee: Meer praktiese toepassings en minder teorie. Dit help nie om ideale toestande te leer en jy kry nooit met dit te doen in die praktyk nie. Kan dit miskien net doen vir agtergrond".

(19 SASA: State)

"Nee, die opleiding is baie teoretiese. Die praktiese toepassing leer mens eers as jy in die praktyk staan".

(20 SASA: Univ)

'n Theoretiese akademiese opleiding pas 'n persoon <u>nie</u> goed genoeg toe vir die werk van dosent nie. Praktiese ervaring is baie noodsaaklike en kan verkry work deur een of meer van die volgende:

(a) Praktiesgerigte (akademiese) opleiding as student

(b) Praktiese ervaring by buite-instansie voordat doseerwerk begin word".

(21 SASA: Univ)

"I think it was an essential and valuable part. Difficult to remember enough details as to what was good or bad, or what could have been better. Undergraduate stuff was too theoretical: M.Sc. was more practically orientated and therefore more use".

(22 OR Society)

"My statistics training was geared to OR and most of it was self study. Hence YES it did provide relevant skills. But I think that courses <u>along the lines</u> of Stu Hunters design of expts. and Gene Woolsey's Quick and Dirty Operations Research are good for getting a basic grasp quickly. I would call this training. I would call formal course work using Wilks; Feller; Scheffe; Draper & Smith etc. education".

(23 SASA/ORSoc: Univ)

"If I were designing courses today I'd replace about one-third of the current curriculum with more nuts & bolts courses on data analysis including enforced usage of at least 2 program packages. Our theoretical coverage is of first class international standard, with actual applied stats the area with most room for improvement".

(24 SASA/ORSoc: Private but ex Univ)

"An appropriate balance between theory and practice is required. Students in 3rd year and above should be exposed to Statistical Consulting (eg. Students in other faculties requiring statistical help in their experimental work)".

(25 SASA/ORSoc: State but ex Univ)

"No training other than Summer School. Regular courses in basic techniques would be valuable. Occasional in others necessary".

(26 OR Society)

"More emphasis on project work and analysis of real (not so well-behaved) data, even at undergraduate level, is required".

(27 SASA: Univ)

"Yes. To improve training, MANY more practical examples and applications of statistics should be included with the theory. Perhaps small projects for industry should be encouraged".

(28 OR Society)

"Groter konsentrasie op praktiese toepassings. Meer spesifiek doelgerigte kursusse. Huidige opleiding gee tegnieke, maar nie genoeg toepaslike praktiese toepassingsopleiding nie".

(29 SASA: State)

"In most areas, yes. Suggestion on advanced training for statisticians: I think there is still a lot of scope for training through the medium of workshops on more specialized subjects. The HSRC workshop on SAMPLING and the UCT workshop on REPEATED MEASURES were particularly worthwhile for me".

(30 SASA: Univ)

"Vir my werk as statistiese konsultant was daar te min praktykgerigte kursusse tot op M-vlak. Projekwerk in die 3de jaar, en honneursjaar behoort verder aangevul te word met verpligte vakansie werk onder 'n ervare statistikus".

(31 SASA: Univ)

"No. More practical-orientated material should be included in statistics courses as well as practical projects".

(32 SASA: State)

"Die universiteitsopleiding is baie toerities en ver verwyderd van die praktyk. Die kennis wat ek in die een en 'n half jaar opgedoen het by my huidige werk is van groot waarde en kom teorie en praktyd bymekaaruit - wat 'n leemte is op universiteit. Dit sal goed wees as 'n statistikus as deel van sy opleiding 'n sekere tyd in die praktyk moet wees".

(33 SASA: State)

"Praktiese ondervinding is baie noodsaaklik en kan verkry word deur indiensopleiding deur konsultasie van p[illegible] of aan buite-instansies".

(34 SASA: Univ)

"My statistiese opleiding het my die nodige voordighede geleer (U.P.)
Ek beveel egter aan dat 'n gevallestudie benadering (soos wat ek ervaar het met my MBL opleiding) 'n baie praktiese en nuttige didaktiese onderrig wyse is wat ook in Statistiek gebruik kan word".

(35 SASA: Private)

"Praktiese werk kan, veral met die oog op 'n beroep in die Statistiek, baie meer beklemtoon word. Dit kan gedoen word deur die praktiese sy van die kursus groter gewig (in terme van eksamenpunte) te laat dra sodat die studente dit self ook met groter erns sal bejeen.
Studente behoort deeglik geskool te word in die gebruik van programmatuur-pakette, byvoorbeeld SAS. Dit sou sinvol wees as hulle 'n breë oorsig kon kry oor dit wat beskikbaar is en dit wat self kon leer toepas".

(36 SASA: State)

"Nee, meer praktiese ondervinding in probleemoplossing, hantering van groot databasisse en veral interpretasie van resultate word benodig. Statistiese opleiding konsentreer nie genoeg op probleem-oplossingstegnieke in handels en ekonomiese rigtings (bv. marknavorsing) nie".

(37 SASA: Private)

"No, only after a few years of working has the theory and the practice been linked together. I think the training should contain more practical work, not practical work out of a textbook, but in a real situation".

(38 SASA: State)

"1. Ja.

2. Meer praktiese - real life - opleiding sonder om die teorie te verminder - dws om die vak nie net as mooi bewerkings te beskou nie maar as 'n noodsaaklike toepasbare dissipline".

(39 SASA: State)

"My training was sufficient. However the application of the theory in the work situation proves extremely difficult. Management is not prepared to spend time on analysis of data which inevitably will have to be scrapped".

(40 OR Society)

"Yes. My training gave me a theoretical background. If only we had done more practical work. All distributions are not normal in practice".

(41 SASA: Private)

"Statistics training provided adequate theoretical skills i.e. the 'mechanics' - insufficient training given in the topic 'recognition' and solution of real-life statistical problems".

(42 SASA: Univ)

"I feel that more emphasis could have been placed on practical situations in order to make the theory more relevant and to equip one better for the working situation".

(43 SASA: State)

"Meer klem moet gele word op die praktiese toepassing van statistiese tegnieke aan die Universiteit".

(44 SASA: Private)

"My opleiding het defnitief die kunsdigheid gegee om my werk bevredigend te kan doen. 'n Moontlike verbetering in opleiding is die groter klem wat daar moontlik vir persone wat in die praktyk wil gaan, op praktiese aspekte geplaas kan work".

(45 SASA: Univ)

"I felt my training at undergraduate level provided me with adequate mathematical skills to progress in a research capacity however lacked dismally to prepare me to handle any practical problem. I feel our undergraduate courses could do more to equip students with the necessary tools to tackle applications they might encounter in industry, agriculture and biometrics".

(46 SASA: Univ)

"Statistiese opleiding was noodsaaklik maar meer praktiese opleiding kan verleen word".

(47 SASA: Univ)

"Die Wiskundige Statistiek kursusse kan defnitief meer op die praktyk toespits. Die teoretiese benadering help weinig wanneer prakties gedink moet word".

(48 SASA: State)

"No, but it was a while back - courses seem to have improved. We had no practical training other than computing statistics for data sets for which the area and technique were specified"

(49 SASA: State)

"Felt that not enough practical application of theory was given, but as this was during 1963-1966 and training has changed since then - hope this has been remedied".

(50 SASA: State)

"Need more practical applications during studies".

(51 OR Society)

"Please more practical examples".

(52 OR Society)

"Yes, it basically did, but the practical experience is far more important".

(53 OR Society)

"Perhaps experts could give 'guest lectures' to students on specific subjects".

(54 SASA: Univ)

"Training not practical enough - very little training on Time Series, Forecasting, Regression - especially effects on Regression if classical assumptions do not hold. Training on Multivariate too Mathematical. People we appoint to do these jobs have to relearn it first".

(55 OR Society)

"Personally, there should be a modular course on Directional Data and Time-Series Analysis at undergraduate level. These methods often have to be used in industry".

(56 SASA: Private)

"My statistical training was adequate for my job. In South Africa we need Statistics with Time Series, Forecasting and Quality Control training".

(57 SASA: Univ)

"Yes! Basically what is covered by UNISA syllabus is more than enough for the type of work I am involved in. Nevertheless, the school of thought of Dr E W Deming which advocates serious consideration of 'Statistical Process control' was not highlighted to the right and appropriate perspective".

(58 SASA: Private)

"Aanvanklike statistiese opleiding vorm sekerlik die basis. Die ontvikkeling is egter so snel dat dit beslis gou verander. Opleiding in statistiese prosesbeheer (X, R, ens) is krities noodsaaklike aan Universiteite en Technikons".

(59 SASA: Private)

"By and large my statistical training was not adequate for the performance of my job. It covered almost nothing on Experimental design and had almost no practicals on the use of statistical packages. But also, in retrospect, I was probably unqualified for the job. It needed someone with a strong research background e.g. Ph.D.."

(60 SASA: Univ but was Private)

"No, I think that simulation and model development (i.e. more OR) should be incorporated in statistics course".

(61 SASA: State)

"My training gave me a good background which enables me to grasp new techniques easier. I do however feel a greater emphasis should be placed on data inspection and multidimensional methods".

(62 OR Society)

"Academic training gave very little attention to aspects of market research".

(63 SASA: Univ)

"Perhaps universities could do more work on spatially-defined data".

(64 OR Society)

"Maybe the geostatistical theory could also have formed part of an under-graduate course".

(65 SASA: Private)

APPENDIX F

2½ YEAR SYLLABUS - FOR DISCUSSION

½ Course First Year

Mathematics Requirement:

A Little calculus desirable but not essential.

Σ notation

$\binom{n}{r}$

Students: All B.Sc/B Comm/B.Eng.

1. Exploratory Data Analysis (à la Tukey)

   (a) Single Variable.

   5 number summaries
   Smoothing FIT + RESIDUAL
   Significant Figures
   Means, Std. dev.

   (b) Relationship between variables.

   Y and X
   Fitting a line
   Two way tables

2. Probability distributions

   Independent Events
   AXIOMS Addition/Multiplication
   Frequency distributions

3. Confirmatory data analysis.

   What is the probability of obtaining a result like this by chance?

   Testing locations: Z and t test

   χ^2 test.

4. Data Collection.

   Sampling

2/....

YEAR 2 AND 3

4 Half Courses or 2 Full Courses. (240 lectures)

Mathematics Requirement

Mathematics I completed / Maths II concurrently recommended

/ Computer experience

STUDENTS: Potential Statistics Majors
B.Sc Mathematical Sciences
Actuarial

1. (a) Frequency Distributions as Models for Data - Properties of frequency distribution.

Independent Bernoulli trials.

Binomial, Geometric, neg. Binomial, Hypergeometric

Counted data - Poisson Process.

Continuous data : Uniform, Normal, Gamma Beta.

(b) Further properties of Frequency distributions:

Moment Generating Functions /Means Variances
Limit Theorems
Tchebyshev Inequality

2. Bivariate Distributions.

Conditional distributions
Conditional Expectation
Correlation

3. Independent Random Variables.

Samples
Sums of Random Variables/Mean and Variance of a sum
Central Limit Theorem

4. Functions of Random Variables.

Change of Variable
Max, Min, intro. to order Statistics.

3/....

5. Fitting Models to Data.

Data Collection and Summarisation
Fitting of distributions
Estimation of Parameters
The Likelihood Function
Properties of estimators
Methods of Estimation:
- Minimum Variance Bound
- Minimum Variance Unbiased Estimator
- Maximum Likelihood
- Method of Moments
- Least Squares
- Bayesian Estimation

Practical Statistics

Use of Packages

Introduction to Simple Statistical programming

6. Confidence Intervals.

General Methods

7. Goodness of Fit tests.

Probability Integral Transformation
Probability Paper
Normal Plots/Exponential Plots
χ^2-Goodness of Fit

Test for Indepedence:

Fitting model P(AB) = P(A)P(B)

Practical Work

8. Testing Hypotheses about Parameters.

Tests, Size, Power, Critical Regions
Neyman-Pearson Lemma

Best Tests.
Uniformly most powerful tests
Likelihood Ratio tests

9. Samples from a Normal Distribution.

The Likelihood-Ratio test on the mean
Students t distribution
Confidence Intervals
Testing

The Sample Variance:

The χ^2 distribution of S^2
Additive properties of χ^2

4/....

Comparing two normal samples.

2 Sample t test
F test
Confidence Intervals

10. Samples from k Normal distributions.

Likelihood ratio test of equality of k means
The Analysis of Variance

One Way layout Practical Work

Multiple Comparisons

Power of F test

Sample sizes

11. Design of Experiments.

Completely Randomised design Practical Work
Randomised Blocks
Latin Squares
Factorial Experiments

12. Relationships between Random Variables.

Correlation Pearson/Rank
Least Squares
The Linear Model $Y = X\beta + e$
Multiple Regression
Testing Hypotheses on β
Analysis of Residuals/Outlier Tests
Choice of Model - How many β_i's?

13. Generalised Linear Model. (Non normal Populations)

Some function of parameter follows a linear model.

The link function . Practical Work

$\text{Log}\ \theta = a + bX$ Poisson errors

Logit models

14. Unbalanced Data.

Anova/Regression

5/....

15. Relationships Evolving in Time.

Introduction to Time Series
Multiplicative Model
Estimation of Seasonal/Trend/Cycles
Smoothing
Fitting non linear models/ Gompertz
Prediction

Practical Work

Iterative methods

16. Multivariate Distributions.

Correlation Analysis
Hotellings T^2
Multivariate Normal

17. View of Statistical Packages.

Discussion of Programmes
Comparison of Packages

18. Sequential Testi[illegible]

19. More Sampli[illegible]

20. Special Topics in Statistics.

Survival Analysis
Quality Control
Life Testing
Contingency Tables
Experimental Designs
Circular Distributions
Selection and Ranking of Populations
Census

21. More Probability Theory.

Birth and Death Processes
Queues
Limit Theorems
Assymptotic Distributions

22. Non parametric Statistical Tests.

23. Violation of Assumptions.