

DATA PROCESSING AND STATISTICS: SOME INTERACTIONS

by Murray Jorgensen

Modern data management technology impacts on the statistician in two ways: firstly it opens new possibilities and greater flexibility for a statistician to manage his or her data, making feasible the analysis of much larger data sets. Secondly this new capability gives the statistician an interest in the design of information systems that are not in the first instance statistical, in order that these systems may be constructed in such a way as to permit the extraction of data appropriate for statistical analysis.

A Database Management System (DBMS) is a computer software product under the control of which data is stored, organised into related files, and retrieved. To use a modern DBMS it is not necessary to know the physical details of how the data is actually stored, but only to have a conceptual picture of the logical relationships among the data items. Commonly databases are organised hierarchically. As hierarchies are common in Statistics as well it is useful to give some examples.

- (i) A national crop variety evaluation might have data available at the following levels: region, site, trial, block, plot, sub-plot.

At the region level weather information might be available while soil-type, farmer's name, etc. might be available at the site level.

- (ii) A survey might collect data at the levels of:
city, mesh block, household, person
- (iii) In fisheries data collection at least 3 levels can be identified. Firstly we have data about the sampling "station" such as latitude, longitude and depth. Secondly data pertaining to the occasion of sampling such as date, time, gear used, name of responsible person. At a third level we have information on the composition of the catch. If this is rather complex the catch level may itself be decomposed into more than one level.

In a modern database, at each level, a different type of record is used to store the information pertaining to that level. The fields that make up a record are divided into **key** fields and **data** fields. The key fields contain sufficient information to uniquely identify a record. For instance in example (ii) above the key of a record of the "household" record type will contain sufficient information to uniquely identify the household referred to.

A DBMS will provide a means of storing information about the fields of a record type. This facility may be called a "schema" or a "data dictionary". In it we store such information as the data type (real, integer, character) of each field or variable, variable names, value labels, variable ranges, and so on.

But statistical packages in general expect their data input files to be in the form of a simple rectangular matrix of observations by variables. When a database is interfaced with a statistical package a retrieval program is written in the language of the DBMS to build such a "flat" file. Information must be communicated to the statistical package at a particular "level" of the hierarchy. If this level is the lowest, such as the sub-plot level in the first example, information must be "brought down" from higher levels by chasing up the tree. This information will then be written out many times at the lowest individual level. If we wish to move information to a statistical package at a higher level then the lower-level information must be summarised in some way: by summing, taking means, medians, counts, variances, or percentage points of the lower-level data. The richness and flexibility of the retrieval language will determine exactly what can be done. It is at this point that many commercially-oriented DBMSs fall down: they are often limited to such simple summaries as means and totals. Facilities for handling missing values are also absent from many DBMSs.

It may seem almost frivolous to talk about the transfer of labels and variable names to statistical packages but experience has shown that this facility is very important. P-STAT solves the problem by being a statistical package as well as a DBMS. SIR (Scientific Information Retrieval) has an interface which permits the creation of SPSS, SAS and BMDP system files. If a DBMS language is sufficiently expressive it is possible to develop procedures which create files containing the data to be transferred and expressions in the language of the statistical package which control the reading in and labelling of the data. The Ministry of Agriculture and Fisheries uses SIR procedures to build input files for GENSTAT and MINITAB in this way.

The other alternative is to do things the hard way and write out a pure data file from the DBMS and write a statistical package program to input, label and analyse the data. Errors and inconsistencies can often creep in when this is done.

Whatever approach is adopted the use of statistical packages in partnership with DBMSs provides a very powerful tool for studying, analysing and reporting on large sets of data. In Hext (1983) seven data management systems are described and evaluated for their usefulness in statistical applications.

A knowledge of what can be done with statistical packages leads the statistician on to wonder what could or should be done statistically with the very large amounts of data held in commercial dp systems. Most

people have only a vague idea of the difference between scientific and commercial computing along the lines of "they use COBOL, we use FORTRAN". Increasingly language is coming to seem less important, each package having its own internal language owing something to a variety of predecessors. I believe that a better distinction between the two types of computing is to say that scientific computing involves intensive processing of small, static, datasets. (It may even involve no data at all and be totally theoretical in nature). In contrast commercial computing involves the processing of large and constantly changing datasets. The processing itself is usually of a fairly simple nature.

Many statisticians see themselves as Data Analysts and the modern statistician has a range of powerful computer tools at his or her disposal to undertake this role. The question I wish to raise is: can we bring our analytical tools to bear on data within commercial information systems? Is there any need to do this? I would say certainly! Many businesses have captured within their computer systems a reflected picture of the firm and its changing performance in the different sectors of its market. Statistical analysis of data extracted from corporate information systems could provide valuable information to managers about the performance of their institution. The interfacing of databases with statistical packages discussed above provides one means of doing such extraction of data.

There are, however, problems connected with the dynamic nature of commercial information systems. There are limitations on the amount of historical information that can be held in disk storage and extraction retrievals will usually produce a "snapshot" of the data when what is required is a time series. Yet a statistician is normally only interested in a small fraction of the information in such a system. Furthermore advances in the theory and practice of the analysis of categorical data means that the data storage requirements can be compressed substantially. What a statistician might best be able to work with is an automatically generated "state description" which tracks the evolution of the system through time. This would usually be available only if a statistician were involved in the design of the information system.

I have indicated in the first part of my article how modern data management technology can assist the statistician in taking on large sets of data for analysis. I also feel that the statistician has a role in the field of Management Information. However I feel that the statistician will never be fully effective in this area unless he or she has an input into the design of management information systems. Both the statistician and the computer scientist recognise an area called "Data Analysis". At the moment each means something different by the term. But it is possible that both areas are part of a larger whole and could be profitably knit together. I hope that this happens, not only in an academic sense, but in the working together of all who are responsible for the collecting, processing and interpreting of data.

REFERENCE

Hext, G.R. (1983) *Database systems for statistical applications*. (Civil Service College Handbook 24). HMSO, London.

THE SEVENTH AUSTRALIAN STATISTICAL CONFERENCE

by Ken Russell

The 7th (biennial) Australian Statistical Conference was held at the University of Queensland from 27th to 31st August 1984. The list of participants records full-time attendance by 250 people, of whom ten were from New Zealand. Another 50 people attended for one day only.

The featured sessions at the Conference dealt with Medical Statistics, Analysis of Sample Survey Data, Geophysical Signal Processing, Estimation using Capture-Recapture and Related Methods, Modelling of Biological Systems, Inference, Multivariate Analysis, Discrete Multivariate Analysis, Design and Analysis of Experiments, and The Role of Statistics in Public Policy. The invited speakers were George Seber (Auckland; "Recent Developments in Population Estimation"), Richard Smith (Imperial College; "Statistical Problems of Extreme Values"), T.W. Anderson (Stanford; "Components of Variance in MANOVA"), Robin Sibson (Bath; "Smooth Variation Across the Plane: Analysis and Presentation"), Shelby Haberman (Hebrew University of Jerusalem; "Canonical Analysis and Maximum Likelihood"), and Richard Cormack (St Andrews; "Loglinear Models for Capture-Recapture").

The Conference began with the Presidential Address of John Darroch (Flinders). This was a most entertaining and illuminative lecture entitled "Probability and Criminal Trials—Some Comments on Bayes' Theorem Prompted by the Splatt Trial and Royal Commission". The accused in this celebrated South Australian case had been convicted of murder on the basis of circumstantial evidence and some "probabilistic" arguments by the prosecution. This had been followed by petitions for his release, a Royal Commission, and suggestions that a policeman may have planted some evidence. (I am sure that the New Zealanders in the audience shared my feeling of *deja vu* as parts of this case were detailed.)

The Presidential Address set the stage for what was a most stimulating conference. Scheduled talks were held from 9 a.m. to 5 p.m. each day except for Wednesday and Friday, when they finished at noon. In general, there were three or four parallel sessions at any one time, except when the Invited Addresses were being presented. I felt that the general presentation of papers was better than at earlier Conferences. There were fewer speakers who couldn't be heard, and distinctly fewer transparencies which couldn't be read—but the exceptions still remain. When will some speakers learn that photocopies of typed material are almost never legible, even from the front row?

Pleasant memories remain of the Brisbane Conference. The University campus was attractive, especially with its riverside location, and the Lunch Cruise on the Wednesday afternoon was a definite highlight. The Conference organisation was good, the morning and afternoon teas lavish, the Welcoming Party and Conference Dinner very pleasant—and the weather was superb! As well, some of us were fortunate enough to attend a barbecue held at the home of Tony Swain, one of the Conference organisers.

Less memorable were some of the bathrooms at Union College and the disappointing organisation of the poster display: there were only three posters, and these

were on show in an out-of-the-way room where few people saw them.

Overall, it was a pleasant Conference. The 8th ASC will be held in Adelaide in 1986, and I imagine that the wine buffs are already anticipating one excursion at that meeting.

DEADLINE FOR NEXT ISSUE

The deadline for submitted material for the June, 1985 issue of this newsletter is May 30. Please send all notices of seminars, news items, letters-to-the-editor, etc. to . . .

John Reynolds,
Newsletter Editor,
AMD/DSIR
P.O. Box 1335
Wellington.

The deadline for "News and Announcements" for the April, 1985 issue of *The New Zealand Statistician* is March 29.

P and PoW SCIENCE AWARD TO A MEMBER OF THIS ASSOCIATION!

Dr Peter J. Thomson, a member of the Institute of Statistics and Operations Research at Victoria University of Wellington and an enthusiastic member of this association, has been awarded a Prince and Princess of Wales Science Award by the Royal Society of New Zealand. The award has enabled Peter to visit the Time Series research group led by Dr Akaike at the Institute of Statistical Mathematics, Tokyo. A brief report on Peter's travels so far appears below.

TALES OF A NONSTATIONARY TIME SERIES ANALYST

by Peter Thomson

On July 3rd, as many of you know, I commenced my first twelve months sabbatical leave. Now, just over six months later, I am struggling to recall details of events that have happened over the intervening period. The 6 months have gone by very quickly but, paradoxically, Wellington seems to be part of a dim and distant past. However, to business. In response to a request by John Reynolds I have written a brief report of our travels to date.

The first four months of my leave were spent as a Visiting Fellow at the Department of Statistics, Institute of Advanced Studies, Australian National University, Canberra, Australia. Having visited Canberra many times before, it was a relatively easy matter to settle in to life in the Department and in Canberra itself. The ANU treat visitors very well and we were no exception, being provided with a University flat at a reasonable rental and a well-appointed office in the Department itself together with secretarial and computational facilities. Over the four months at the ANU I completed work with Professor Ted Hannan on a time series delay estimation problem. The relaxed atmosphere and lack of timetabled commitments meant that there was always time available to discuss problems of mutual interest with Ted and other colleagues. This I enjoyed very much.

In addition I also had time to do further work on the interpretation of the space-time spectrum of some DSIR seismograms, a project which I and Professor David

Vere-Jones had commenced prior to my departure from Wellington. Aspects of this work, including graphical displays of the space-time spectrum, formed part of a paper I presented to a session at the 7th Australian Statistical Conference held in Brisbane. This particular session was devoted to geophysical time series with the express aim of providing a forum where statisticians and geophysicists might exchange real problems for relevant statistical methodology. Representatives of mining and oil companies were there and contributed to the session. However my own feeling at the end of the session was that real progress in this area would best be made by an interdisciplinary research team of statisticians and geophysicists tackling specific problems. Overall the conference was well attended by statisticians from all over Australia and New Zealand together with a number of well known statisticians from other countries. All in all it was a successful conference both academically and otherwise.

We concluded the Australian part of our leave with a short, but very pleasant trip to the Division of Mathematics and Statistics, CSIRO, Sydney. There Dr Murray Cameron and I had time to discuss joint research and map our future joint work. It is always a pleasure to visit DMS in Sydney. They are a lively and talented group of people involved in many interesting research projects and are very capably led by the ever amiable Dr Geoff Eagleson.

From what information I could glean, the Australian statistical scene looks to be reasonably healthy overall. However I sense that New Zealand universities and government scientific statistical organisations are somewhat more closely tied to applied statistical problems and practice than their Australian counterparts. I believe that such an emphasis more often proves to be a strength than a weakness.

The next two months of my leave were spent as a Research Fellow at the Institute of Statistical Mathematics, Tokyo, Japan. There I worked mainly with members of the Fifth Division, Dr Akaike's time series group. This was our first visit to Japan and, because of language and cultural differences, it could well have been a difficult two months. However this was not the case. The entire family experienced the most stimulating and rewarding two months that we have ever had. The hospitality of our Japanese friends and colleagues was almost overwhelming. In addition the atmosphere of the Institute, not to mention the excellent facilities, made it a very pleasant place to work indeed. We were very fortunate in having access to reasonably priced accommodation; a conventional Japanese style small flat. This, together with the Prince and Princess of Wales Science Award that I was fortunate to be offered, helped reduce the high cost of living in Japan.

Our accommodation was organised through a friend and colleague, Dr Ogata, who also coordinated a lecture tour around universities in the southern half of Japan. On this tour we visited Hiroshima University which has one of the largest groups of university statisticians, Kagoshima University where I gave a paper at a special Time Series conference and met many of Japan's leading time series analysts, Kyoto University and Kanazawa University. I also gave seminars at Keio University, Tokyo Institute of Technology, Hitotsubashi University and of course the Institute itself.

In terms of time series research there appeared to be a marked difference between the University scientists and those at the Institute. The former were concerned more with mathematical and theoretical developments

whereas the latter were concerned more with generating innovative practical methods for the analysis of large and complex time series data sets. The more I learnt of the novel approaches advocated by Dr Akaike and his group the more impressed I became. I believe that, using information criteria such as AIC, they are exploring new and important dimensions in time series model building. Many of these essentially involve the fitting of conceptually simple likelihood models to large data sets where the parameters of these models are allowed to evolve by regarding them as slowly varying stochastic functions over time. Using simple conditioning, and, in effect, a model for the random parameters, the more complex likelihoods needed to fit extensive data sets are simply derived. In many cases these likelihoods can be constructed relatively easily using the Kalman Filter.

In terms of their respective strengths in applied statistics, the contrast between the Institute and the University groups was quite stark. In this regard the Japanese University statisticians appear to have similar problems to those confronted by many western university statistical groups in the past. When statistics, an essentially practical subject, is locked in a theoretical department it loses much of its relevance. The level of funding of Japanese universities and their structure would appear to indicate that any satisfactory resolutions to these problems are unlikely to come soon.

All in all our period in Japan and the Institute was a thoroughly stimulating, rewarding and enjoyable experience, both for me professionally and the family as a whole. We are now in London where I will be working for six months in the Department of Economics, London School of Economics and Political Science, with Professor Peter Robinson. Here we have been very fortunate to obtain accommodation in London. Again, time appears to be passing by at a phenomenal rate and it will not be long before we are back in Godzone and (gulp!) I have to front up to teaching again.

NEED YOUR SPIRITS LIFTED?

"For better or worse, statistical inference has provided an entirely new style of reasoning. The quiet statisticians have changed our world—not by discovering new facts or technical developments but by changing the ways we reason, experiment, and form our opinions about it."

Hacking, Ian (1984) *Trial by number*.
Science 84, Vol. 5, No. 9, 69-70.

THE XIIth INTERNATIONAL BIOMETRIC CONFERENCE, TOKYO, SEPTEMBER 2-8, 1984

by Bryan Manly

For me, the International Biometric Conference offered a rare opportunity for combining business and pleasure. On the one hand, the conference attracted some 400 participants, including many if not most of the best known biometricians. On the other hand, there was plenty of time for sightseeing in one of the most interesting cities in the world.

The conference programme was too full for me to give anything like a full account of it. There were 14 invited paper sessions involving about 40 individual papers, 37 contributed paper sessions, involving about 170 papers, and two poster sessions with 25 contributions. In this brief report I will merely mention what were the highlights for me.

The conference opened with the usual welcoming speeches and a lengthy address by the President, Pierre Dagnelie. I remember particularly his remarks

concerning biometrics in the developing countries. Apparently membership of the society in these countries as a proportion of the total membership has declined substantially over the past 25 years. There are at least 30 countries of at least 5 million inhabitants in which the Biometric Society does not have a single member. Unfortunately, this is probably a true reflection of the development of biometrics in these countries. The Council of the society will be examining what can be done to improve this situation in the near future.

There were a large number of papers on proportional hazards models. Recent developments in this area were the subject of one of the invited paper sessions organised by John Kalbfleisch, and also one of the contributed paper sessions. However these models also appeared in papers in many of the other sessions.

Quite by chance I attended a session on nearest neighbour analysis in field and variety trials and discovered that this is currently an area in which there is some controversy. It is claimed that substantial gains are to be had by designing and analysing experiments taking into account nearest neighbour effects. However, some of these claims are disputed and some speakers made a plea for a conservative approach to using these methods because not enough is known to justify using them in place of classical experimental designs, at least on a routine basis. It will be interesting to see the outcome of this debate.

The analysis of DNA data was the subject of two sessions on the final day of the conference. This is a topic that will surely develop substantially in the next few years. At the present time there is still a good deal of ignorance about what is going on with DNA sequences so there is some interest in simply detecting elements of non-randomness. I attended the sessions with the hope of learning something about the topic, which I did. However I came away with the feeling that everyone was working in the dark and that there are some crucial facts that are still not known. Again, it will be interesting to see future developments in this area.

Apart from the three particular topics that I have mentioned there were many other papers of interest to me. Only once or twice during the week did I feel that none of the sessions were attractive. In short, this was a first rate conference, just as I expected it would be when I arranged to go.

PAGING DR PAPADAKIS

The paper by J. S. Papadakis, *Methode statistique pour des experiences sur champ* (Bull. Inst. Amel. Plantes a Salonique, 1938, No. 23), appears regularly in the references of papers and books on the analysis of field experiments, but Ken Russell has met only one person who has ever seen the paper itself and that was Graeme Wilkinson (of Nearest Neighbour fame). Does anyone in New Zealand have a copy of this paper? If so, please contact Ken at ISOR, Victoria University of Wellington, Private Bag, Wellington.

NEWS FROM THE INSTITUTE OF STATISTICS AND OPERATIONS RESEARCH (ISOR) AT VICTORIA UNIVERSITY OF WELLINGTON

by Ken Russell

The Annual Meeting of the Institute was held early in December. Tony Vignaux (Professor of Operations Research) was elected to a three-year term as Chairman, and Dr John Reynolds (AMD) and Dr Colin Cryer (Wellington Clinical School) were elected as Honorary Members. The Institute expressed its regret that it would be losing the assistance and camaraderie

of Tim Ball (AMD) when he moved to Auckland to join the Deming Institute.

Considerable time was spent during 1984 in negotiations to retain the Assistant Statistician's post in ISOR. The successful outcome was the creation of the position of Research Statistician (oddly enough, *not* an academic post) and the appointment to it of Stephen Haslett. Megan Clark, Shirley Pledger and Ken Russell were all promoted to the position of Senior Lecturer on 1 January, as was Philip Morrison (an ISOR member who spreads the statistical gospel from within the Geography Department).

At the time of writing, Tony Vignaux is somewhere in California learning about the latest techniques in simulation, and our peripatetic Statistics Professor, David Vere-Jones, is actually at V.U.W., after three months in Canberra and a week at the Summer Research Institute in Auckland. We are unaware of any plans by David for further travel in the near future.

ISOR recently acquired a Christmas present in the form of a DEC Micro VAX computer. This was purchased with the financial assistance of the Applied Mathematics Division, DSIR; it will be used for a wide variety of tasks, including development of statistical software.

Among the special features will be a UNIX operating system and the use of the statistical computer language S. Sounds impressive, doesn't it? The only problem is that we were sent the wrong operating system . . . Watch this space for further developments.

The Foundation meeting of MULES (N.Z.) was held at a local restaurant at the end of October. MULES was formed in Sydney in November 1973, in the belief that statisticians should get together socially now and again, and its meetings took the form of biannual dinners at ethnic restaurants. These were initially intended for Statistics staff and postgraduate students at The University of New South Wales, but were later widened to include other statisticians in Sydney. Franchises are available at minimal cost for areas in New Zealand other than Wellington; please contact Ken Russell for details. (If you're wondering why it's called MULES, see page v of Volume 2 of Kendall & Stuart.)

MAF BIOMETRICS SECTION NEWS

David Harte will be joining the Biometrics group at the head office of the Ministry of Agriculture and Fisheries this month. He will be working mainly in the area of Quality Assurance. David is currently working in the Traffic Research Branch of the Ministry of Transport. He completed an MSc at Victoria University of Wellington in 1982 with a thesis entitled, "Self similar stochastic processes".

PACIFIC STATISTICAL CONGRESS

Plans are progressing well for the Pacific Statistical Congress—1985, which will take place from 20-24 May, 1985, at the University of Auckland. So far there are about 130 registrations and 100 papers. About one third of the papers will be presented by New Zealanders, another third by Australians, and the remainder will be by people from further afield. Support from the Australian and New Zealand governments and the Biometric Society is being sought to enable statisticians from the South Pacific island nations to attend.

Keynote speakers include:

Professor Peter Armitage, University of Oxford, UK (Some Statistical Aspects of Clinical Trials)
Professor Gene H. Golub, Stanford University, USA

Professor Dennis V. Lindley, University College London, UK (On re-reading Harold Jeffreys)
Professor Alastair J. Scott, University of Auckland, NZ (Fitting logistic models to survey data)
Dr Terry P. Speed, CSIRO, Australia (Applications of cumulants and their generalisations).

Invited speakers include:

Dr David A. Dickey, North Carolina State University, USA (Preliminary transformations in time series)
Professor M. Ross Leadbetter, University of North Carolina, USA (Extreme Value Theory and Dependence)
Professor J. N. K. Rao, Carleton University, Canada
Professor Shayle R. Searle, Cornell University, USA
Dr Masaharu Tanemura, Institute of Statistical Mathematics, Japan
Dr John R. van Ryzin, Columbia University, USA
Dr George P. H. Styan, McGill University, Canada
Dr Geoff A. Watterson, Monash University, Australia (Estimating Species Divergence Times).

The list of sessions being arranged includes:

Population Genetics, Statistical Computing, Analysis of DNA Sequence Data, Bootstrapping and Jackknifing, Medical Statistics, Statistics in National Development, Extreme Value Theory, Stochastic Processes, Applied Time Series, Sample Surveys, Statistical Consulting, Experimental Design, Applied Probability, Bayesian Statistics, Statistics in Ecology, and Microcomputers in the South Pacific. In parallel with the main programme, the Auckland Mathematics Association will run a one-day course on "Statistics on microcomputers in Secondary Schools" on Friday, 24th May.

Although the deadline for abstracts has passed, there may still be room for some late papers in the programme. Please send abstracts, as soon as possible, to:

Bryan Manly,
Department of Mathematics,
University of Otago,
P.O. Box 56,
Dunedin.

The Annual General Meeting of the New Zealand Statistical Association will be held at the same venue as the congress on Thursday, 23rd May at 12.15 p.m.

NZIAS CONFERENCE

The NZ Institute of Agricultural Science is holding its annual conference at Lincoln College from September 2-6, 1985. The theme of the conference is 'Water'. Any member of the statistical association who wishes to contribute a paper is asked to contact Chris Dyson, Biometrics Section, MAF, P.O. Box 24, Lincoln for further details.

ANOTHER CLIPPING FROM THE DEPO FILE

"The report says that the collection of current epidemiological data has been 'too haphazard and uncoordinated' to provide a way to assess the drug's potential long-term hazards. Two studies currently under way may eventually supply the needed information. David Thomas of the University of Washington is conducting a large study funded by WHO to evaluate Depo-Provera's potential cancer risk among contraceptive users in several countries, and Upjohn is also sponsoring an epidemiological study of New Zealand women who use Depo-Provera. Panel members say that both studies are well designed."

Sun, M. (1984) Panel Says Depo-Provera Not Proved Safe. *Science*, Vol. 226, 950-951.

**UNIVERSITY OF OTAGO
LECTURER IN STATISTICS**

or

**ASSISTANT LECTURER IN
STATISTICS**

Applications are invited for appointment to a position of Lecturer or Assistant Lecturer in Statistics in the Department of Mathematics. Applicants should be prepared to take up the position on or around 1 February 1986. The University will provide travel and moving expenses.

Salary:

Lecturer: \$NZ21,660-\$NZ25,684 per annum

Assistant Lecturer:

\$NZ16,123-\$NZ18,732 per annum

plus an annual cost of living allowance of \$417.

Further particulars are available from the undersigned, P.O. Box 56, Dunedin, New Zealand.

Applications quoting reference number A84/67 close on 15 April, 1985.

D. W. Girvan
Registrar

GENSTAT RULES, OK?

"My comments in this paper have, of course, included what I have intended to be constructive and considered criticisms of a few details of current Genstat output. These criticisms are as nothing compared with what I might have said about various other statistical programmes and packages which are, I understand, widely available."

Preece, D. A. (1984) Genstat Analysis of Variance and the Distant Client. *Genstat Newsletter No. 13*, 35-38.

WELCOME TO THE REAL WORLD

"I suspect that too often, when lecturers talk about practical exercises, what they mean is practice at routine design and analysis in relation to some rounded problem. But it is the ragged edges that statistical consultancy is all about. Discovering these and weighing them in the balance is why applied problems are seldom routine, and are never dull."

Gore, S. M. (1984) Teaching Experimental Design: Prescribed by a Medical Statistician. *The Statistician*, 33, 243-247

SAS Applied Time Series Analysis and Forecasting

**A course for those who want to use SAS/ETS tools
for analysing and forecasting time series data**

The SAS Applied Time Series Analysis and Forecasting course covers the use, limitations, and interpretation of SAS/ETS procedures designed to analyse time series data. This course addresses the application and interpretation of the output produced by the following SAS procedures:

PROC FORECAST for generating automatic forecasts on many univariate time series

PROC AUTOREG for estimating parameters in regression models where the error structure is an autoregressive process

PROC ARIMA for analysing and forecasting univariate time series and transfer function data

PROC STATESPACE for analysing and forecasting multivariate time series

PROC SPECTRA for estimating the spectral and cross-spectral densities of multivariate time series

PROC X11 for seasonally adjusting monthly or quarterly time series.

Mathematical aspects and practical applications of each procedure, in addition to annotated output, will be covered. Given prior notice, special interests of course participants can be emphasised.

Course participants should have some familiarity with reading and creating SAS data sets and some experience with fitting regression models, but a brief review of these topics is provided.

Course instructors will be:

Dr John Brocklebank, Manager of Statistical Training, SAS Institute Inc., Cary, North Carolina, USA, and,

Dr David Dickey, Associate Professor, Department of Statistics, North Carolina State University, USA.

The three-day course will be held on May 14-16, 1985 at the James Cook Hotel, The Terrace, Wellington. The fee for the course is \$345 (includes Teas, Lunches, Course Notes and Materials).

For further information and registration material contact,

Dr Kevin Hall,

SAS Institute (NZ) Ltd.,

P.O. Box 10-109,

Wellington. Phone: (04) 727-595

THE W. EDWARDS DEMING INSTITUTE OF NEW ZEALAND

The Deming Institute has commenced operations in Auckland. The institute was established to promote improved quality and productivity in New Zealand's manufacturing and service industries and in government, and to provide the means whereby this improvement can be achieved. The means are, of course, the philosophy and methods of Dr W. Edwards Deming, in particular Deming's "Fourteen Points of Management" (for a listing of these, see, for example, the *National Business Review*, May 21, 1984) and the principles and procedures of statistics.

Membership of the institute is limited to organisations and companies. The member organisations pay an annual subscription of \$5,000 and each organisation is represented in the institute by a/the chief executive. This is in keeping with Deming's strategy of spreading the quality assurance gospel, in the first instance, to top management. The institute will accordingly be conducting seminars and workshops for chief executives, acting as a clearing-house for information on Quality and Productivity programmes, publishing a newsletter for its members, and, establishing a library of relevant books, professional articles and videotapes. On a fee basis, the institute will also undertake consulting and educational services for its members and run Deming Seminars.

At this point the reader is probably asking, "What has all this got to do with statisticians and other lower forms of life?" Well, it seems that the effectiveness of statisticians in industry or for that matter any organisation is related to the managerial climate in which they work. Statisticians working in organisations belonging to the institute can probably expect their managers and directors to have their consciousnesses raised with respect to the utility of statistical methods. Statisticians employed by member companies can probably reasonably expect to attend a Deming Seminar or QA workshop run by the institute. Member organisations without the services of a statistician may well create positions for suitably trained statistics graduates.

If your employing organisation does not yet belong and you require further information you could contact:

The Director of Programmes,
The W. Edwards Deming Institute of New Zealand,
P.O. Box 68-173
Auckland. Telephone (09) 389-169

WELLINGTON REGIONAL MEETING

The New Zealand Statistical Association will be holding a one-day Wellington Regional Meeting on Wednesday, 28 August, 1985. The venue for the meeting will be the Easterfield Building Conference Room, Victoria University of Wellington. Intending speakers should contact Garry Dickinson, Department of Statistics, Private Bag, Wellington.

WORKSHOP ON STATISTICAL INFERENCE FOR STOCHASTIC PROCESSES

by M. Kelly Mara

The 25th Summer Research Institute of the Australian Mathematical Society was held at the University of Auckland from January 14 to February 1, 1985. The research institute was divided into one-week workshops according to field of interest. I attended the Stochastic Processes/Time Series session which was

held in the first week. The mornings were allocated to semi-formal talks given by invited speakers with the afternoons being devoted to workshop and problem discussion sessions.

The workshop benefitted considerably from the input of the invited speakers: David Brillinger (University of California, Berkeley), Peter Lewis (Naval Postgraduate School, Monterey) and David Vere-Jones (ISOR Victoria University of Wellington). In fact we were witness to David Brillinger's attempt to become a real "New Zilnder" by lecturing in dirty white shorts. (Actually, the shorts were clean and blue, but being statisticians we were asked to forget reality and use our imaginations.)

David Brillinger presented a unified view of methods for handling dependent point, spatial and spatial/temporal processes. An interesting practical problem he addressed concerned the response of a neuron to the changing image in an eye—the input is a spatial/temporal process and the output is a point process. Peter Lewis spoke about nonlinear models for time series with particular reference to gamma autoregressive processes. He made most antipodean attendees envious of his access to computer programs and pre-release packages. Rumour has it that he has an IBM370 in the corner of his office. David Vere-Jones and Daryl Daley (ANU, Canberra) spoke about aspects of point processes.

Personally, I learned a great deal from the week—not only the statistical aspects but also the types of very interesting problems in fields such as physics, zoology, geophysics and neurophysiology for which spatial process statistics can yield some insights.

My preference would have been for a slightly more structured session but in general the relaxed nature of the institute provided a welcome change to very formal conferences.

A CONSULTANT'S CHECKLIST

The following is excerpted for your edification from Brian Joiner's notes for a Statistical Consulting course he used to give in the Statistical Laboratory at the University of Wisconsin-Madison. The editor advises you to paste a copy of this checklist inside the cover of your battered copy of Snedecor and Cochran (or Steel and Torrie).

- Listen a lot.
- Never (well, almost never) interrupt a client.
- Always (well, almost always) allow the client to interrupt you.
- Ask a lot of questions that begin with phrases like, "Let me see if I understand this, . . .".
- Don't smell: shower daily, wear clean neat clothes, brush teeth, get decayed teeth fixed.
- Take good notes, from the beginning.
- Convey a helpful resourceful attitude.
- Helping the client clarify his/her own goals is often your most important function.
- Keep things simple: don't dazzle your client—or yourself; use as simple words, simple designs and simple analyses as will suffice.
- Put things in writing, and give copies to the client.
- At the end of each meeting spend a few minutes clarifying what has been decided and who does what next. Put these in writing.
- Know a wide variety of statistical methods (so you'll be less inclined to try to fit the problem to the wrong solution).
- Interact frequently with the client—never go off and do a lot of work at the client's expense or on the client's behalf without discussing approaches and

intermediate results.

- Make realistic cost estimates and discuss them (cost may be dollars and/or time).
- Find out how important this project is to the client, and how much she/he wants to commit to your efforts.
- Be timely: an approximate answer in a few days is almost always preferable to an "exact" answer some months later.
- Make sure you and the client understand the goals of the project, and whether or not these goals are attainable by the planned course of action. Get the goals agreed to, in writing.

HAVING PROBLEMS WITH THE PUNTERS?

"Client expectations of the money cost of statistics covers the waterfront. I've had some people come to me with fifty observations in a t-test type of problem ready to spend \$200 and others come in with \$1,000 worth of work and be shocked that the bill will exceed \$100. Some first-time clients think there is no charge at all, that they are doing me a favor to let me do something with their data because they assume statisticians are sexually aroused by fondling numbers."

Boen, J. R. and Zahn, D. A. (1982) *The Human Side of Statistical Consulting*. Lifetime Learning Publications, Belmont, California.

VACANCIES

APPLIED MATHEMATICS DIVISION, DSIR

Applied Statistician Operational Research Analyst

The Applied Mathematics Division of DSIR has vacancies in its Industrial Statistics Section and its Operational Research Section. The Statistics position is a temporary one for one or two years but it is likely to become permanent. It is based in Wellington. The Operational Research position is a permanent one. It is based in Auckland but it may be appropriate for an appointee to have the first year in Wellington.

For both vacancies the work will involve taking part in AMD's service for Industry and Government. This includes carrying out projects, providing advice and conducting seminars for clients. Currently the Industrial Statistics Section is particularly interested in quality assurance and industrial experiments. The Operational Research Section has interests in transport/location problems, inventory control, production scheduling, and simulation.

The appointees will also be encouraged to carry out research in areas of interest to the Division.

Requirements:

- Good honours degree, MSc, Diploma, or PhD.
- Ability to use computer packages.
- Ability to communicate.
- Potential to solve practical problems.
- Enthusiasm, common sense, initiative, etc.

Apply to:

The Director,
Applied Mathematics Division, DSIR,
P.O. Box 1335, Wellington.

OVERSEAS CONFERENCES

American Statistical Association/Biometric Society/ IMS Joint Meeting

The 1985 joint meeting of the American Statistical Association, Biometric Society and the Institute of Mathematical Statistics is to be held from August 5-8 in Las Vegas, Nevada. For further information write to ASA, 806 15th Street, N.W., Washington DC, 20005, USA.

International Statistical Institute—45th Biennial Session

This conference, which marks the centenary of ISI, is to be held in Amsterdam, Netherlands, August 12-22, 1985. For further information write to ISI, 428 Prinses Beatrixlaan, P.O. Box 950, 2270 AZ Voorburg, Netherlands.

XIIIth International Biometric Conference

To be held in Seattle, Washington, USA, July 27 to August 1, 1986. For further information write to Dr Gerald van Belle, Dept of Biostatistics, University of Washington, Seattle, Washington 98135, USA.

The Second International Conference on Teaching Statistics (ICOTS II)

This conference will be held in Victoria, British Columbia, Canada, August 11-16, 1986. Sessions will take place on teaching statistics at all levels, from school to university and in government, business and industry. The New Zealand coordinator for ICOTS II is . . .

John C. Turner,
Department of Mathematics,
University of Waikato,
Private Bag,
Hamilton, New Zealand.

American Statistical Association/Biometric Society

The 1986 joint meeting is to be held from August 18-21 in Chicago, Illinois. For further information write to ASA, 806 15th Street, N.W., Washington DC, 20005, USA.