SIMS ENVIRONMETRICS PROJECT

During the period from 1984 to 1987, a study team centered in the Department of Statistics participated in a $1-million study on acid rain, an environmental problem plaguing North America, Europe and other parts of the world. Work on the project began in the spring of 1984 when Bill Caselton (Civil Engineering), Al Marshall, John Petkau, and Jim Zidek joined with groups at the University of Washington, Stanford University and Rand Corporation in making an application, co-ordinated by the Societal Institute for the Mathematical Sciences (SIMS), to the Environmental Protection Agency for a Co-operative Research Agreement to study statistical problems associated with acid rain.

Acid rain is caused when substances in emissions from industry and other sources undergo chemical reactions in the atmosphere, are converted to acids and fall to the earth. Although acid rain may well have a substantial effect on the environment, its sources are unclear; in particular, how much of it is the result of industry and other human activities, and how much of it is natural is uncertain. Volcanoes, for example, can emit thousands of tons of material into the atmosphere which can be converted to acid rain. The precise measurement of the impact of acid rain, necessary for regulation and control, is exceedingly difficult.

Realistic approaches to these and related issues involve mathematical and statistical problems of great variety and depth. Objectives of the study were to develop models for the various processes involved, techniques for designing monitoring networks and setting guidelines, and methods for analyzing spatial-temporal data. These models, techniques and methods were to be used to detect the impact of acid rain, to measure trends, to map acid rain deposition levels, and to determine how those levels vary from region to region.

During the three year period, the study group expanded to include additional faculty (Mohan Delampady, Harry Joe, Jian Liu and Ruben Zamar). In addition, graduate students and visitors participated to varying extents. The initial objectives were met, but the many new projects identified along the way led to the submission of an application for a continuation of the Co-operative Research Agreement involving a much broader scope of investigation. A $1.5-million study to investigate statistical problems in the areas of human exposure, quality assurance and global climate change, in addition to acid deposition, was approved for the period from 1988 to 1991.

One major focus of the study was the development of methods to obtain estimates of trend, as well as seasonal and spatial effects, from data obtained over a period of several years from a monitoring network. The overall objective was a map of North America with areas of increasing, decreasing or no local temporal trends identified. Shiying Wu, a student working with Jim Zidek, developed a simple resistant method based on repeated medians and Jian Liu investigated an alternate approach based on parametric time series models for the data at individual stations, with overall trends and geographical effects estimated via a simple variance components model. Application of these methods to the chemical components available in the 1980-86 data from the NADP monitoring network yielded results of direct relevance to the National Acidic Precipitation Assessment Program in the United States.

When dealing with concentrations of pollutants, typically extremes rather than averages are compared with air quality standards; extreme value theory is therefore very important for the analysis of environmental data. In the case of a single pollutant, extreme quantiles can be estimated by threshold methods based on fitting the generalized Pareto model to observations exceeding a high threshold or to a small number of extreme order statistics. But concentrations of different pollutants may be highly dependent, even in the extremes, so the health effects of several pollutants simultaneously having high levels should be considered and air quality standards modified accordingly. Together with Richard Smith and Ishay Weissman, Harry Joe developed a bivariate threshold method, based on models for the joint distribution of a bivariate observation subject to some function of the components exceeding a high threshold. Applications include estimating the probability that thresholds are simultaneously exceeded by two variables, inferences on the dependence of two variables in the extremes, and estimation of a bivariate upper quantile curve. Related work includes investigation of parametric families of multivariate extreme value distributions and of an approach to modeling of peak exposures over different microenvironments based on stationary Markov models.

Kriging is the most common approach to interpolating network data. Jim Zidek has developed an alternative based on regarding the process in question as locally expandable and dealing with the coefficients for different sample paths in an empirical Bayes analysis, leading to the posterior distribution of a future value. The resulting theory is very flexible and generally applicable; inferences about nonmonitored sites with mean and quantile contours and estimates of reliability are automatically generated. Related work includes the development, by Mohan Delampady and Jean-Francois Angers, of a hierarchical Bayesian approach to smoothing a temporal or spatial process and work on the robust estimation of scale by Ruben Zamar and Doug Martin, the eventual objective of which is robust variogram estimation and the development of resistant interpolation techniques for spatial data.

A new approach to environmental monitoring network design, based on the

principle that designs should maximally reduce uncertainty (as measured by entropy), was developed by Jim Zidek, together with Bill Caselton and Lisa Kan. This methodology was employed in a systematic assessment of the NADP/NTN network, to determine how the number of stations might be reduced to cut the costs of operation and so gain resources which might better be deployed to judiciously augment other existing networks.

The possibility of relationships between air pollution and morbidity or mortality is of concern to environmental agencies responsible for health-based standards for ambient air pollution. A dose-response relationship is essential for quantitative assessment of the potential health and economic impacts of different standards. A two-year record of daily counts of hospital admissions and emergency room visits for respiratory illnesses from a community with several local sources of pollution were analyzed in detail by John Petkau, together with Keith Knight, Brian Leroux and John Millar. The hospital admissions data yielded inconclusive results, but the emergency room visits data provided a clear indication of relationships at lags up to two days. Related work includes a follow-up validation study based on the subsequent two-year period and a major new study in a different community based on daily lung function and respiratory symptom measurements in groups of asthmatic and control children.

Work on the project was organized around biweekly working sessions where ongoing research was discussed. These provided a convenient mechanism for involving graduate students in research at an early stage of their studies. In addition, a seminar series which met irregularly facilitated the exchange of knowledge on many areas of environmetrics among researchers in government, industry and universities. Rapid dissemination of completed research was accomplished via the SIMS technical report series; over the period of the study, a total of 49 technical reports were issued at UBC. Most of this work has now appeared in the peer-reviewed literature.

Much of the research envisioned at the outset was completed by the end of the second period, so another continuation was not sought. This team approach to methodological research targeted to specific subject-area problems of importance was a very stimulating experience, so the Department anticipates similar future endeavours, perhaps on problems in biostatistics.

BIOSTATISTICS RESEARCH GROUP

The Biostatistics Research Group (BRG) provides:

- a base and a focal point for faculty and graduate students who develop and disseminate theory and methods for data analysis in the health and life sciences; and
- a resource for researchers who seek such support.

Faculty affiliated with the Department of Statistics have been much involved in biostatistical activities for many years. Activities include methodological research, collaborative research, consultation, graduate student instruction and supervision, and support and service to the broader biostatistical community.

Topics of recent and current methodological research include designs for clinical trials, stochastic modeling of processes in cancer, methods for the analysis of epidemiological data, models for dose response experiments, modeling to assess synergism between drugs, measures of association and agreement, diagnostic accuracy, drug screening experiments, modeling of longitudinal categorical data, intra-laboratory comparisons, and Markov and hidden Markov models for multi-state phenomena.

Application areas of collaborative research are also diverse, encompassing cancer and its precursors, health promotion in gerontology and dentistry, Alzheimer's disease, epilepsy, multiple sclerosis, and Parkinson's disease in neurology, low-tension glaucoma in ophthalmology, infectious diseases, paediatrics, and studies of the relationships between air pollution and human health.

A substantial proportion of both our M.Sc. and Ph.D. students have written their theses on biostatistical topics. Many of these former students are now employed in health sciences research environments, either in Vancouver or elsewhere. Others occupy faculty positions in colleges and universities in various countries.

A series of meetings was held to plan how these current and future activities in the discipline of biostatistics might be better coordinated.

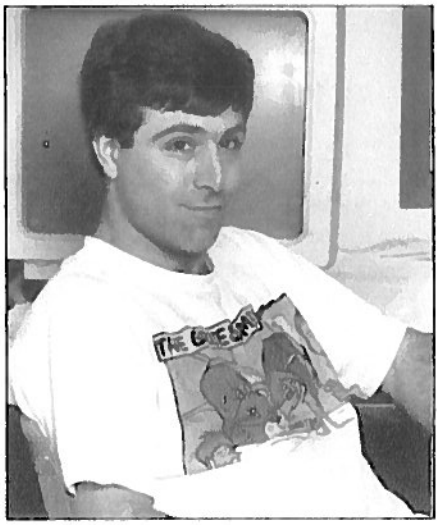
The Biostatistics Research Group was establishedin the spring of 1991 as an informal research group. The objective was to create a mechanism to facilitate and promote interactions among researchers at UBC and in the greater Vancouver area who are actively involved in development of biostatistical methodology and application of such methodology to scientific problems in the health and life sciences. Individuals who volunteered to serve as the initial executive of the BRG included Andrew Coldman, Ned Glick, Nancy Heckman, Harry Joe, John Petkau, Martin Puterman, Michael Schulzer, and James Zidek. The executive oversees the activities of the BRG and leads this informal research group in developing new initiatives and undertaking new activities. Beyond this, there is no formal membership in the BRG; rather, members identify themselves by active involvement in BRG activities.

To promote research and to communicate developments in biostatistical theory, methods, and computing, the Biostatistics Research Group sponsors:

- Biostatistics Research Reports, a series of articles intended for publication in statistical or other methodologic journals;
- Biostatistics Research Seminars, featuring local or invited speakers;
- BRG Journal Club discussions of biostatistics methods in the recent literature; and
- BRG Working Sessions for interactive discussions of participants' research in progress.

The Biostatistics Research Group encourages its participants to collaborate with researchers, health professionals, and industries involved in biostatistics projects. Such collaborative projects are often conducted in conjunction with the Statistical Consulting and Research Laboratory (SCARL). Sustained collaborative projects typically involve interdisciplinary funding initiatives, joint publications, and thesis work by graduate students.

In all its activities, the Biostatistics Research Group seeks to advance the development of biostatistics and to promote good biostatistics practice, in co-operation with other academic units and related agencies, including: UBC faculties or schools such as Medicine, Dentistry, Nursing, and Pharmaceutical Sciences; hospitals and their research institutes; clinical trials and health promotion centres; the B.C. Ministry of Health; the B.C. Cancer Agency and other clinical research organizations; the Biostatistics Section of the Statistical Society of Canada; and the Canadian Society for Epidemiology and Biostatistics.



Bertrand Clarke

ASSISTANT PROFESSOR

B.Sc. (1984), University of Toronto

Ph.D. (1989), University of Illinois

Assistant Professor, Purdue University (1989-92)

The University of British Columbia (1992-present)

Research Interests

A major attraction of research in statistics is that one is free to explore diverse areas of study.Thus, although my primary research interest is in Bayesian asymptotics, I am interested in mathematical modelling in biology.

The Bayesian model for an experiment includes a quantity called a prior density. Often, the prior is intended to reflect the pre-experimental beliefs of an investigator. Then, once data has been collected, the beliefs of the experimenter are updated and summarised in a quantity called the posterior density, which we can use to obtain inferences. Within this formulation there are many natural and interesting questions. Here are a few: How do we choose a prior in the first place? What happens to the posterior as we accumulate ever more data of various forms?

Interestingly, the Bayesian model lends itself to the statistical theory of standardised tests. Suppose we wish to assess some trait in each individual in a population. Quantify this trait by a number, representing, for instance, mathematics ability. Roughly, the proportion of people at each ability level corresponds to a prior. If we administer a test to an individual we obtain data from which we can estimate that person's ability level. Thus, in a sense, we have a posterior for each examinee. Our task is to combine the results over many examinees so as to make better inferences.

Another circumstance in which a Bayesian model arises naturally is in information theory. Suppose we want to transmit a message to someone far away and we know that the message will be corrupted. That is, the message we send, a sequence of zeroes and ones, is not always the message received. In this case, the prior represents the probabilistic behaviour of the agent supplying the messages we wish to transmit, and the most important random component in the problem is the way messages are corrupted. This model can be used in many coding and transmission problems.

Unrelated to the foregoing, probabilistic methods are useful for many applications in biology. Biologists in many specialties have been collecting enormous amounts of data over the past decade and while there are some overarching theories - evolution for instance - little modelling of general structure has been done. Working in collaboration with various biologists, I am interested in developing general theories for the organisation of organisms. For instance, is there a characteristic pattern of connectivity between genes? Is there a dynamic model for a cell, such as a bacterium, which can relate the major functions of a cell to each other? Seeking real-world applications of this line of inquiry is premature; however, without a deep conceptual understanding of the organisation of organisms few applications will be possible.

Selected Publications

Clarke, B. and Barron, A.R. (1990). "Information-theoretic asymptotics of Bayes methods." *IEEE Transactions on Information Theory*, **36**, 453-471.

Clarke, B. and Ghosh, J.K. (1991). "Posterior normality given the mean, with applications." Technical Report 91-68, Department of Statistics, Purdue University.

Clarke, B. and Mittenthal, J.E. (1992). "Modularity and reliability in the organisation of organisms." *Bulletin of Mathematical Biology*, **54**, 1-20.

Clarke, B. and Wasserman, L.A. (1992). "Noninformative priors and nuisance parameters." *Journal of the American Statistical Association*. To appear.

Nancy E. Heckman

ASSOCIATE PROFESSOR

B.S. (1977), Tufts University

M.A. (1982), University of Michigan, Ann Arbor

Ph.D. (1982), University of Michigan, Ann Arbor

Assistant Professor, State University of New York, Stony Brook (1982-1985)

The University of British Columbia (1984-present)

Research Interests

I am interested in non-parametric, or smoothing, regression. In the general regression problem, each y observation is assumed to depend upon an observed variable, x. In classical (parametric) regression, the form of this dependence is quite specific. For instance, one might assume that Y is linear in x, with random error. The least squares regression estimate of the linear relationship is easy to find, often without a calculator. And a linear dependency of Y upon x is certainly easy to picture.

The widespread use of computers has reduced the need for easily calculated estimates of the relationship between variables. And the development of computer graphics allows the statistician to literally look at forms of dependencies which are far more complex than the classically assumed linear relationships. Instead of determining the "best-fitting" line, one can calculate and graph a "best-fitting smooth curve", through such computer-intensive techniques as spline and kernel smoothing.

My research interests in non-parametric regression have focused in three areas: semi-parametric models, regression function shape, and non-normal likelihoods.

In the semi-parametric model, one assumes that the expected value of Y is α's + $g(t)$, where α is an unknown parameter vector, g is an unknown but smooth function and s and t are known covariates. This compromise between parametric and smoothing regression allows one to estimate a as accurately as one could if one knew the exact form of g.

Smoothing methods are often used to help a researcher study the shape of a regression function, or to determine if the regression function shows certain features, such as a bump. Often the determination of such shape characteristics are of greater interest than the actual estimation of the regression function. I have studied the problem of "bump-hunting" and of examining the convexity of a regression function.

Many standard smoothing methods have been developed to handle the case of normally distributed responses. I am interested in extending these methods to other models, including generalized linear models (e.g., Poisson data) and backcalculation models, which arise in the estimation of the rate of HIV infection in the population.

In addition to my research in nonparametric regression, my research has involved the study and development of designs for collecting data. Currently I am studying transect designs. A transect is a path along which data are collected and this type of data collection is often used for censusing animal populations. The transect design issues are quite different from those in the standard design problem in which one chooses a few locations (instead of a path) to collect data.

Selected Publications

Heckman, N. (1986). "Spline smoothing in a partly linear model." *Journal of the Royal Statistical Society*, **B48**, 244-248.

Heckman, N. and Woodroofe, M. (1991). "Minimax Bayes estimation in nonparametric regression." *Annals of Statistics*, **19**, 2003-2014.

Heckman, N. (1992). "Bump hunting in regression analysis." *Statistics and Probability Letters*, **14**, 141-152.

Fan, J., Heckman, N. and Wand, M. (1993) "Local polynomial smoothing in generalized linear models." Technical Report 128, Department of Statistics, UBC.

Heckman, N. and Rice, J (1993). "Line transects of two-dimensional random fields: estimation and design." Technical Report 132, Department of Statistics, UBC.

Harry Joe

PROFESSOR

B.Sc. (1978), University of Victoria

M.Sc. (1979), The University of British Columbia

Ph.D. (1982), Florida State University

The University of British Columbia (1982-present)

Research Interests

My research has branched to many areas in statistics. I take a problem-solving point of view for my research in that the research is not tied to any area of statistics and I learn new things as necessary to solve a new problem. Over recent years, I have worked in areas such as majorization and entropy in statistics, multivariate dependence, paired comparisons and ranking data, lotto games and distributions of k-tuples, environmetrics and extreme values. A common theme that underlies my diverse research is the use of multivariate dependence concepts for statistical models. On the mathematical side, I have discovered that there are many bivariate concepts that do not seem to generalize to the multivariate situation and one of my research goals is to better understand where generalizations are possible.

One specific aim is to study the general class of multivariate distributions in order that multivariate models with specified properties may be constructed; there is little research in this area which has been dominated by the use of the multivariate normal distribution. This requires studying of methods for constructing parametric families and searching for new stochastic constructions and representations for models.

Specific situations where I have made use of multivariate concepts are in: (a) extreme value inference for multivariate maxima and for dependent sequences, (b) measures of multivariate and conditional dependence that are valid for all type of data (continuous, categorical, mixed, etc.), (c) models for ranking data, and (d) models for repeated measures or multivariate binary or ordinal response.

Some of my multivariate problems go back to K. Pearson, but now there is computational power available to try new approaches. Much of my research makes use of the computer as a valuable tool; this includes the use of the computer for symbolic manipulation, quasi-Newton optimization, numerical integration, numerical linear algebra and Monte Carlo methods.

Selected Publications

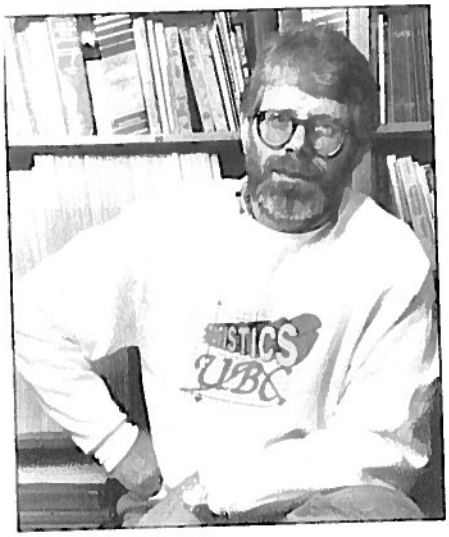
Joe, H. (1985). "An ordering of dependence for contingency tables." *Linear Algebra and its Applications: Special Statistics Issue*, **70**, 89-103.

Joe, H. (1988). "Majorization, entropy and paired comparisons." *Annals of Statistics*, **16**, 915-925.

Joe, H. (1989). "Relative entropy measures of multivariate dependence." *Journal of the American Statististical Association*, **84**, 157-164.

Joe, H. (1993). "Generalized majorization orderings and applications." In: Stochastic Inequalities, M. Shaked and Y. Tong (eds.). *IMS Lecture Notes-Monograph Series*, **22**, 145-158.

Joe, H. (1993). "Parametric families of multivariate distributions with given margins." *Journal of Multivariate Analysis*. To appear.



Bent Jørgensen

ASSOCIATE PROFESSOR

B.Sc. and M.Sc. (1979), Aarhus University

Ph.D. (1987), Odense University

Assistant Professor, Aarhus University (1979-80)

Assistant Professor (1980-84) and Associate Professor (1984-87), Odense University

Visiting Scholar (British Council Scholarship), Imperial College (1981-82)

Visiting Professor, Imperial College (1983)

Visiting Professor, University of Waterloo (1984)

Associate Professor, Instituto de Matemática Pura e Aplicada, Rio de Janeiro (1987-92)

The University of British Columbia (1992-present)

Research Interests

My primary research interest is generalized linear models, which are regression models based on non-normal error distributions, applicable to a wide range of different types of data. This is currently a very active field of research and many diverse questions are being investigated.

The original definition of generalized linear models due to Nelder and Wedderburn included as special cases the classical multiple regression models, models for positive data, and discrete data models such as logistic regression and log-linear models. I am particularly interested in studying new error distributions for use in generalized linear models. An example of a distribution that is far from being normal, but still fits in the generalized linear models framework is the von Mises distribution for angles or spherical data. Another example is compositional data (continuous proportions), for which a class of new models was proposed in recent joint work with O.E. Barndorff-Nielsen, allowing compositional data to be analyzed by generalized linear models.

A further example is a model for positive continuous data with a positive probability component at zero. This model has applications in the analysis of car insurance data, for example, where the zero component corresponds to the case of no claims for a given insurance policy and the positive part of the distribution is the claim-size distribution. Together with M. C. Paes de Souza, I have used this model to analyze a set of Brazilian car insurance data.

All these models are special cases of the class of dispersion models, which has many properties in common with the normal distribution, while being a very flexible class of distributions, as the above examples illustrate. Inference for regression models based on distributions of the dispersion model type may be based on generalizations of the familiar F and t-tests from analysis of variance and multiple regression analysis. This hence provides a general and fairly simple method of inference that can be applied to any generalized linear model with error distribution in the class of dispersion models. Many dispersion models are partly of exponential family form, and the exponential family plays an important role in the analysis of dispersion models.

The study of generalized linear models has had many spin-offs, both in terms of new statistical methods, but also results of a purely mathematical or probabilistic interest, related to topics such as infinite divisibility, stable distributions and central limit theorems.

A lot of work has been done recently on generalized linear models for correlated data, with applications to diverse areas such as growth curves, time-series analysis and variance components. Many of the methods that have been proposed are based on semi-parametric models, for example, Wedderburn's method of quasi-likelihood. However, it is possible to find parametric models for such applications, based on exponential dispersion models.

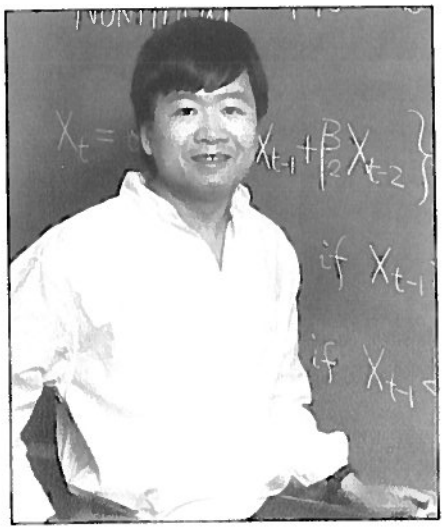
Selected Publications

Jørgensen, B. (1987). "Exponential dispersion models (with discussion)." *Journal of the Royal Statistical Society*, **B49**, 127-162.

Jørgensen, B., Seshadri, V. and Whitmore, G.A. (1991). "On the mixture of the inverse Gaussian distribution with its complementary reciprocal." *Scandinavian Journal of Statistics*, **18**, 77-89.

Barndorff-Nielsen, O.E. and Jørgensen, B. (1991). "Some parametric models on the simplex." *Journal of Multivariate Analysis*, **38**, 106-116.

Jørgensen, B. (1992). "Exponential dispersion models and extensions: A review." *International Statistical Review*, **60**, 5-20.



Jian Liu

ASSOCIATE PROFESSOR

B.S. (1982), Zhongshan University

M.S. (1984), Jinan University

Ph.D. (1987), Colorado State University

The University of British Columbia (1987-present)

Research Interests

My current research interests are in general nonlinear time series analysis and linear time series models with non-Gaussian innovations. Other interests include multivariate techniques, multivariate segmented regression, change-point problems and applications in finance, econometrics, environmental study and engineering.

In recent years, nonlinear time series analysis has attracted much attention both within and outside of the statistical community. This is partially due to its close connection with the growing interest and development in dynamical systems. Some of the merits of nonlinear models include the possibilities of modeling phenomena such as sharp spikes, limit cycles and change of experimental or physical conditions, which could never be dealt with effectively in the conventional way. These types of models provide us with wider applications in economics, finance and other business subjects, industry, geological and atmospheric sciences.

Many financial data exhibit sharp spikes and heavy-tailed distributions. So linear time series models with innovations of this type are applicable. Modeling techniques have as their objective optimal prediction. Another common situation in this context is that necessary information comes from many different channels, with the objective being to pool this information together to form some index for, say, portfolio management. Ideas from conventional multivariate analysis, multiple time series analysis and nonlinear time series analysis may be integrated for this purpose.

Selected Publications

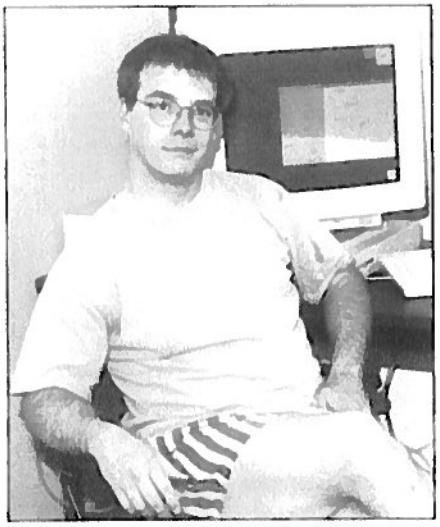
Liu, J. and Brockwell, P.J. (1988). "The general bilinear time series models." Journal of Applied Probability, 25, 553-564.

Brockwell, P.J., Liu, J. and Tweedie, R.L. (1992). "On existence of stationary threshold ARMA processes." Journal of Time Series Analysis, 13, 95-107.

Liu, J. and Susko, E. (1992). "On strict stationarity and ergodicity of nonlinear ARMA models." Journal of Applied Probability, 29, 363-373.

Espen, E. and Liu, J. (1993). "Temporary components of stock prices: New univariate results." Journal of Financial and Quantitative Analysis.

Liu, J. (1993) "Global trend of acid wet deposition - a parametric modelling of sulfate concentrations." Environmetrics. To appear.



Jean Meloche

ASSISTANT PROFESSOR

B.Sc. (1982), Université de Montréal

M.Sc. (1984), Université de Montréal

Ph.D. (1989), Université de Montréal

Visiting Scholar, Stanford University (1988-89)

The University of British Columbia (1989-present)

Research Interests

My primary research interest is in the field of image processing. A typical problem in image processing is the following. A radar camera is installed on the side of a highway, taking pictures of speeding vehicles. For all kinds of reasons, including the fast motion of the vehicle, the atmospheric conditions and the presence of water or mud on the pavement, it may be difficult to read the vehicle's license plate number from the picture. The problem is then to "clean up the dirty picture", that is to construct an estimate of the true image from a noisy picture of that image.

Depending on the amount of knowledge we have about the object in the picture, the problem can be considered as nonparametric (the image belongs to an infinite dimensional space) or as parametric (there are finitely many possible license plate numbers). One of the most interesting aspects of such problems is the wide variety of goals and circumstances in which they arise. In situations where little is known about the true image, only the local characteristics (texture) of the true image are modelled. In others, like the radar camera example, substantial benefit will result from a global model.

Most of my research focuses on the particular case where the true image is black and white. As in all statistical methods, removal of noise is based on the existence of some structure in the systematic part of the underlying model. In fact, the more structured the true image, the easier it is to estimate from a noisy picture of that image. One aspect of the structure in an image is its texture. In Meloche and Zamar (1992), we identified a family of optimal estimates parametrized by the texture of the image. What is remarkable is that the texture parameter can be estimated from a noisy picture of that image, therefore making it possible to adapt the estimate to the texture of the true image. My research is now focusing on the extent to which the estimation of this texture affects the estimation of the true image itself.

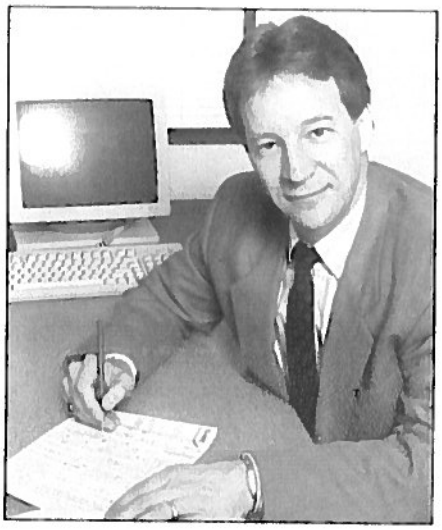
Selected Publications

Meloche, J. (1990). "Asymptotic behaviour of MISE of kernel density estimates for dependent observations." *The Canadian Journal of Statistics*, **18**, 205-211.

Meloche, J. (1991). "Estimation of a symmetric density." *The Canadian Journal of Statistics*, **19**, 151-164.

Meloche, J. and Fraiman, R. (1991). "Counting bumps." Technical Report 112, Department of Statistics, UBC.

Meloche, J. and Zamar, R. (1992). "Black and White Image Restoration." Technical Report 115, Department of Statistics, UBC.



John Petkau

PROFESSOR

B.Sc. (1971), University of Manitoba

Ph.D. (1975), Stanford University

Instructor, Massachusetts Institute of Technology (1974-76)

Visiting Associate Professor, Harvard School of Public Health (1981-82)

SERC Visiting Research Fellow, University of Sussex (1986-87)

The University of British Columbia (1976-present)

Research Interests

My research interests include the design of sequential experiments for potential clinical applications, and the development and application of new statistical methodology in subject-area research.

Due to ethical considerations, clinical experiments are inherently sequential and proposed designs should reflect this reality. Traditional designs oriented towards testing hypotheses may not always be the most relevant in clinical settings. My research has focused on developing analytical and computational methods for an alternate approach which combines Bayesian and decision-theoretic aspects and attempts to more explicitly incorporate considerations of patient benefit. The results have implications for how such clinical experiments should be carried out and are intended to stimulate consideration of relevant criteria for designing clinical experiments. My recent and current research in this area involves the investigation of group sequential designs, the effect of truncation in a specific model for comparing an experimental to a standard treatment, sequential designs for drug screening, and the applicability of this general approach to problems involving dichotomous responses.

Additional research activities involve collaboration with subject-area investigators. Such activity not only allows statisticians to contribute directly to advancement of general scientific knowledge, but also often reveals needs for methodological development and thereby creates new areas of statistical research. One major involvement, with a multiple sclerosis research group at UBC, concerns appropriate methodology for the analysis of clinical trials in multiple sclerosis. Current activity is focused on a recently-completed 3-year multi-centre beta interferon clinical trial in which physical status (ordinal) scores as well as (binary, count, and continuous) responses resulting from magnetic resonance imaging scans were obtained repeatedly over the period of the trial.

A second major involvement is an outgrowth of the SIMS Environmetrics Project. A series of studies have been completed of relationships between ambient levels of air pollution and hospital admissions and emergency room visits for respiratory problems over a four year period in Prince George, BC. Current activity, in a study headed jointly with Dr. Sverre Vedal of the Respiratory Division of the Department of Medicine at UBC, is focused on peak flow and respiratory symptom measurements obtained on a daily basis over a two year period from cohorts of asthmatic and control children in Port Alberni, BC. The findings of the study will be directly relevant to the support or modification of the 24-hour standard for PM10 (ambient inhalable particles smaller than 10 micrometers in diameter) recently established by the U.S. Environmental Protection Agency.

Selected Publications

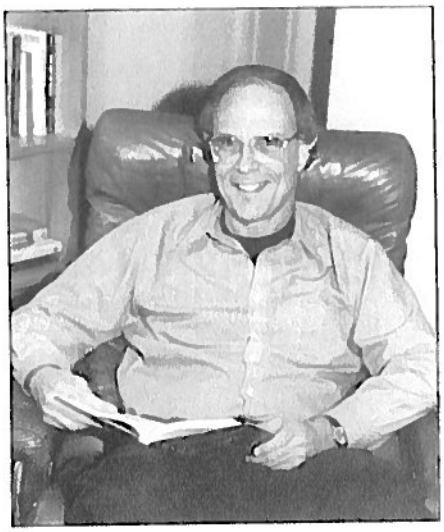
Chernoff, H. and Petkau, A.J. (1986). "Numerical solutions for Bayes sequential decision problems." *SIAM Journal on Scientific and Statistical Computing*, **7**, 46-59.

Le, D.N. and Petkau, A.J. (1988). "The variability of rainfall acidity revisited." *Canadian Journal of Statistics*, **16**, 15-38.

Bather, J.A., Chernoff, H. and Petkau, A.J. (1989). "The effect of truncation on a sequential test for the drift of Brownian motion." *Sequential Analysis*, **8**, 169-190.

Petkau, A.J. and Sitter, R. R. (1989). "Models for quantal response experiments over time." *Biometrics*, **45**, 1299-1308.

Kastrukoff, L.F., Oger, J.J., Hashimoto, S.A., Sacks, S.L., Li, D.K., Palmer, M.R., Koopmans, R.A., Petkau, A.J., Berkowitz, J. and Paty, D.W. (1990). "Systemic lymphoblastoid interferon therapy in chronic progressive multiple sclerosis I. Clinical and MRI evaluation." *Neurology*, **40**, 479-486.



Michael Schulzer

PROFESSOR

B.A. (1958), The University of British Columbia

M.A. (1959), The University of British Columbia

M.D. (1963), The University of British Columbia

Ph.D. (1967), University of Washington

The University of British Columbia (1968-present)

Research Interest

Since I have both a medical degree and a Ph.D. in Statistics, I act regularly as both a researcher in and consultant to medical research groups as well as teaching regularly in the Statistics Department. I provide specialized short courses of instruction to such groups at the Vancouver General Hospital and at the University of British Columbia Hospital. These courses deal with the design of medical experiments and methodology for analyzing results.

I am currently a coinvestigator in a Low Tension Glaucoma Study, funded by the Foundation for Glaucoma Research, headquartered in San Francisco. This longitudinal study, which began in 1986, involves hundreds of patients from more than thirty-three centres in North America, Europe, and Japan. The purpose of this collaborative randomized study is to determine the effect of treatment on the progression of this disease.

I am also involved in the medical and statistical aspects of a Medical Research Council-funded study in neurodegenerative diseases, in which several groups are currently investigating clinical, pharmacological, physiological, chemical and epidemiological features of Parkinson's disease, Amyotrophic Lateral Sclerosis, Huntington's disease and Alzheimer's disease.

Selected Publications

Schulzer, M., Drance, S.M., Carter, C.J., Brooks, D.E., Douglas, G.R. and Lau, W. (1990). "Biostatistical evidence for two distinct chronic open angle glaucoma populations." *British Journal of Ophthalmoogy*, **74**, 196-200.

Schulzer, M. Anderson, D.R. and Drance, S.M. (1991). "Sensitivity and specificity of a diagnostic test determined by repeated observations in the absence of an external standard." *Journal of Clinical Epidemiology*, **44**, 1167-1179.

Schulzer, M., Fitzgerald, J.M., Enarson, D.A. and Grzybowski, S. (1992). "An estimate of the future size of the tuberculosis problem in sub-Saharan Africa resulting from HIV infection." *Tubercle: International Journal of Tuberculosis*, **73**, 52-28.

Schulzer, M. and the Normal Tension Glaucoma Study Group (1992). "Intraocular pressure reduction in normal tension glaucoma patients." *Ophthalmology*, **99**, 1468-1470.

Schulzer, M., Mak, E. and Calne, D.B. (1992). "The anti-Parkinson efficacy of Deprenyl derives from transient improvement that is likely to be symptomatic." *Annals of Neurology*, **32**, 795-798.

Ruben H. Zamar

ASSOCIATE PROFESSOR

Degree in Business (1973), University of Cordoba

M.Sc. (1977), CIENES, Chile

M.Sc. (1982), Federal University of Pernambuco, Brazil

Ph.D. (1985), University of Washgton

Instructor, Federal University of Pernambuco, Brazil (1974-76)

The University of British Columbia (1986-present)

Research Interests

My research interests centre on the theory of robustness, the development of robust procedures, and the detection and testing of outliers. Classical statistics, with either the frequentist or the Bayesian approach, derives optimal procedures under perfectly specified parametric models. Robust statistics, on the other hand, is concerned with the behaviour of statistical procedures under deviations from these models.

The theory of robustness attempts to formalize and measure the basic notion of stability in the context of a statistical model. This theory, then, deals with the formal definition of stability for statistical procedures and with the tools for measuring and optimizing it.

Another aspect of my research interest is the development of robust and resistant procedures for the estimation of multivariate statistical models, including their error scale. I am currently working on robust regression (parametric and non-parametric), robust orthogonal regression and robust principal components.

The combination of robust methods with classical diagnostic tools can improve the power and reliability of both. In particular, outlier tests can benefit from the use of robust techniques to deal with the so-called "masking problem" of multiple outliers. On the other hand, robust procedures followed by diagnostics and testing could lead to more robust and efficient estimates.

Selected Publications

Yohai, V.J. and Zamar, R.H. (1988). "High breakdown point and high efficiency estimates of regression by means of the minimization of an efficient scale." *Journal of the American Statistical Association*, **83**, 406-413.

Martin, R.D., Yohai, V.J. and Zamar, R.H. (1989). "Minimax bias robust regression." *Annals of Statistics*, **17**, 1608-1630.

Zamar, R.H. (1990). "Robustness against unexpected dependence in the location model." *Statistics and Probability Letters*, **9**, 367-374.

Zamar, R.H. (1992). "Bias robust estimation in orthogonal regression." *Annals of Statistics*, **20**, 1875-1888.

Le, N. and Zamar, R.H. (1992). "A new test for effects in factorial designs." *Journal of Statistical Computation and Simulation*, **41**, 41-54.

James V. Zidek

PROFESSOR

B.Sc. (1961), University of Alberta

M.Sc. (1963), University of Alberta

Ph.D. (1967), Stanford

Sessional Lecturer, University of Alberta (1962-63)

Honorary Research Fellow, University College London (1971-72)

Senior Research Scientist, Commonwealth Scientific and Industrial Research Organization, (1975-76)

The University of British Columbia (1967-present)

Research Interests:

My research has two major disparate themes. The first concerns the issue of how best to use statistical information in decision making. Producing an estimate would be an example of such a decision. But my work considers the problem in the abstract. The work takes account of the inevitable uncertainty in the information generated provided by the data. It incorporates the knowledge gained from previous data or from prior experience. And it takes account of the potential consequences (or value) of any specific decision that might be taken.

A specific problem under investigation, is how a team or committee of experts should select a mutual decision when the members of the team (of say medical specialists) have different value scales, different information, and different levels of uncertainty about their information. Some surprising conclusions are that under some circumstances it is better for the team to avoid rather than go for additional information. And sometimes there will be no basis for a consensual choice of a decision, necessitating a random choice say by tossing a biased coin.

Environmental studies provide the second source of my current research topics. Again abstract theory forms the core of the work, with excursions into the implementation of the resulting theory. My initial interest was prompted by the need to develop a statistical design for assessing the negative environmental impact, if any, of wildcat (exploratory) drilling in the Beaufort Sea. That was following by a search for trends in the acidity of acid rainfall. This led to the development of methods of designing networks for monitoring spatial pollution fields, for interpolating between existing monitoring sites to find out what goes on in between,and finally to methods of smoothing the noise out of the data obtained from such network data. Now in turn, the use of this methodology in applications is being explored. In particular, a current study is exploring air pollution, and the effect of human exposure to such pollution, both long and short term. Meanwhile, the development of theory continues.

Selected Publications

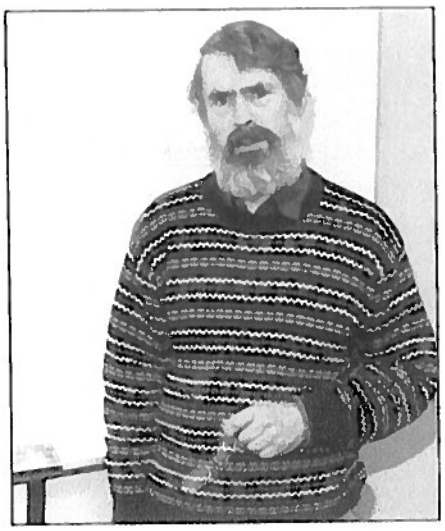
Zidek, J.V. (1988). "Group decision analysis and its application to combining opinions." *Journal of Statistical Planning and Inference*, **20**, 307-325.

Caselton, W.F., Kan, L. and Zidek, J.V. (1992). "Quality data networks that minimize entropy." *In: Statistics in the Environmental and Earth Sciences*. A.T. Walden and P. Guttorp (eds.), 10-38.

Delampady, M., Yee, I. and Zidek, J.V. (1992). "Hierarchical Bayesian analysis of a discrete time series of Poisson counts." *Statistics and Computing*, **3**, 7-15.

Le, N.D. and Zidek, J.V. (1992). "Interpolation with uncertain spatial covariances: A Bayesian alternative to Kriging." *Journal of Multivariate Analysis*, **43**, 351-374.

Schumacher, P. and Zidek, J.V. (1993). "Using prior information in designing intervention detection experiments." *Annals of Statistics*, **21**, 447-463.



Albert W. Marshall

PROFESSOR EMERITUS

B.S. (1951), University of Oregon

Ph.D. (1958), University of Washington

Acting Assistant Professor, Stanford University (1958-60)

Staff member, Institute for Defense Analyses (1960-61)

Staff Member, Boeing Scientific Research Laboratories (1961-70)

Visiting professor, University of Washington (1970-71)

Professor, University of Rochester (1971-75)

The University of British Columbia (1975-present)

Research Interests:

I have a variety of interests, most of which can be classified as reliability theory, multivariate distribution theory, or inequalities.

Reliability theory became my main research interest during my years with the Boeing Corporation. There, colleagues and I were concerned with characterizations of "wear-out" as it might be reflected in the corresponding life distributions. To this end, the classes of life distributions with monotone hazard rate, monotone hazard rate average, and the so-call new better than used distributions were introduced and studied with various coauthors. At that time, I also began to consider problems associated with component dependence in system reliability. We were interested particularly in models to explain sources of dependence, and from these models to develop classes of parametric life distributions that could be sued in the analyses of systems. More recently, I became interested in the notion of a "system", particularly because the various multistate coherent systems in the literature do not adequately address the applicability problems sometimes encountered with the original two-state coherent system.

Multivariate distributions have held a fascination for me since my first course in probability theory. Even in two dimensions, interesting and easily posed questions abound. Because marginals do not determine a joint distribution, it is important to learn how to chose between alternatives for data analysis. Such questions are my major current interest.

Nearly all of my research has in one way or another involved inequalities, and I long had an interest in inequality theory (if there is such a thing). More explicitly, I am interested in the methods which can be used to derive inequalities, and I am interested in unifying concepts which help show how various inequalities are related. A prime example is the ordering of "majorization", a concept which occupied me for some time, and which is still finding new and diverse applications.

Selected Publications

Esary, J.D., Marshall, A.W. and Proschan, F. (1973). "Shock models and wear processess." *Annals of Probability*, **4**, 627-649.

Marshall, A.W. and Olkin, I. (1979). *Inequalities: Theory of Majorization and Applications*. Academic Press, Inc., New York.

Marshall, A.W. and Shaked, M. (1986). "NBU processes with general state space." *Mathematics of Operations Research*, **11**, 95-109.

Marshall, A.W. (1991). "Multivariate stochastic orderings and generating cones of functions." In: *Stochastic Orders and Decision under Risk*. K. Mosler and M. Scarsini (eds.). IMS Lecture Notes - Monograph Series, **19**, 231-247.

Marshall, A.W. and Olkin, I. (1993). "Multivariate exponential and geometric distributions with limited memory." *Journal of Multivariate Analysis*. To appear.