UBC

STATISTICS

DEPARTMENT

THE
UNIVERSITY OF
BRITISH
COLUMBIA

UBC
STATISTICS
D E P A R T M E N T

T A B L E   O F   C O N T E N T S

The University

AERIAL SHOT OF THE UBC CAMPUS

The University of British Columbia has an enrollment of approximately 26,600 students, with about 4,600 of these in the Faculty of Graduate Studies. The campus, which overlooks the mountains and sea, occupies approximately 4 square kilometres at the western tip of the peninsula on which Vancouver is located. It is approximately 12 kilometres from the downtown area of Vancouver, the bus and railroad stations, and approximately 13 kilometres from the airport. Vancouver, with a population of one and a half million in the metropolitan area, offers its residents a variety of activities, from attending symphony performances and theatrical productions to outdoor recreation, such as hiking and sailing. The climate is mild, with summer temperatures seldom above 25°C, and winter temperatures only occasionally below freezing. While it rarely snows in the city of Vancouver, the nearby mountains are snow-capped for half of the year, providing excellent opportunities for skiers, and beautiful vistas for non-skiers.

The Department

The Department of Statistics, in the Faculty of Science, was formed in 1983. At present the Department consists of ten regular faculty members (including one joint appointment in the Faculty of Medicine), a Director of our Statistical Consulting and Research Laboratory, two Professors Emeriti, four Associate Members from Mathematics and Commerce, two Adjunct Professors, two Honorary Research Associates, and over twenty graduate students. Typically, the Department also hosts a number of long-term visitors and several post-doctoral fellows each year.

While the orientation of the Department is theoretical, most members also have a strong interest in applications of statistics, including consultation and collaboration with subject-area researchers. This manifests itself not only through links with other UBC departments and off-campus organizations, but also through the importance placed upon the Department's Statistical Consulting and Research Laboratory (SCARL).

In addition to the Department's weekly seminar, there are regular Biostatistics Research Group meetings and weekly discussion sessions attended by all of those involved in SCARL.

DEPARTMENT MEMBERS

Research Interests of Department Members

Clarke, B.
Asymptotics, Bayesian statistics, information theory, mathematical biology.

Coldman, A.
Biostatistics, quantitative epidemiology, biological modelling.

de Jong, P.
Time series, linear models, forecasting, forensic statistics, actuarial statistics.

Gill, P.S.
Correlated observations, optimal experimental designs, linear models.

Greenwood, P.E.
Optional stoppping, weak convergence of processes, random fields, asymptotic efficiency, likelihood theory for stochastic processes.

Greig, M.
Applications of statistics, contingency tables, computational algorithms, statistical software, experimental design, combinatorics.

Heckman, N.E.
Smoothing methods in regression.

Joe, H.
Multivariate models and dependence, extreme value inference.

Jørgensen, B.
Generalized linear models, exponential families, conditional inference.

Le, N.
Spatial interpolation and design, biostatistics, robust time series.

Liu, J.
Time series analysis, stochastic modelling, multivariate analysis, applications in finance and economics.

Marshall, A.W.
Multivariate distribution theory, inequalities, reliability theory, system reliability.

McCabe, B.
Change point problems, time series econometrics, statistical theory.

Meloche, J.
Smoothing, density estimation, image processing.

Nash, S.W.
Multivariate analysis, correspondence analysis, design of experiments.

Petkau, A.J.
Sequential decision problems, design of clinical trials, biostatistics, environmetrics.

Puterman, M.L.
Mixture models, generalized linear models, Markov decision processes.

Schulzer, M.
Medical statistics, statistical applications to ophthalmology, disease interactions, neurodegenerative diseases.

van Eeden, C.
Truncated parameter spaces, density estimation, nonparametrics, subset selection.

Zamar, R.H.
Robus t estimation and testing, outlier tests and other diagnostic tools.

Zidek, J.V.
Smoothing, Bayesian decision analysis, monitoring network design, spatial interpolation.

OFFICE STAFF from left to right Christine Graham, Donna Pavlov, and Shirley McFee

Undergraduate Studies
The Programs

Statistics, like computer science, derives from the information age; indeed it has been defined as the science of extracting information from data. Consequently statistics is one of the most pervasive of all disciplines. Everyone encounters statistical information daily. And statisticians find employment in medical (including pharmaceutical) research, the production of government statistics, environmental monitoring and assessment, the insurance industry as actuaries, industrial quality management, and many other diverse areas.

Nowadays students seeking a career in statistics find entry increasingly through an M.Sc. or Ph.D. The evolving undergraduate and graduate programs in statistics at UBC recognize this fact. So the undergraduate programs now emphasize the role of statistics within the general framework of knowledge acquisition. And they emphasize the development of the computational and mathematical skills needed for understanding that role and for graduate training. The undergraduate programs provide a broad education in statistics. Such an education enables a better understanding of the role of chance (uncertainty) in all aspects of life today. These programs aim at transferring statistical knowledge and skills, including the fundamentally important capacity for critical statistical reasoning.

Students aspiring to graduate programs or careers in statistics, acquire the needed prerequisite skills and knowledge at the undergraduate level, the Honours program being ideal in this regard. The Combined Honors Program in Mathematics and Statistics gives added emphasis to mathematical training. The Majors program with a smaller statistics requirement enables students to tailor their undergraduate studies to meet diverse educational objectives. Students in increasing numbers are enrolling in the Mathematical Sciences program which balances computer science, mathematics, and statistics. This program provides a very broad range of complementary technical skills and knowledge and it emphasizes the close historical linkages among these three disciplines.

Specific Objectives

The programs aim to help students develop an understanding of the concepts of observation, hypothesis, evidence and validation. The development of creative and critical reasoning skills is emphasized, along with oral and written communication skills. Specifically, the undergraduate programs in statistics foster an understanding of:

- the historical evolution of statistics along with its current directions;
- the scientific method and why statistics is essential in its implementation;
- the distinction between inductive and deductive inference along with an understanding of the role of probabilistic reasoning within statistics;
- the distinction between designed and observational studies in establishing cause and effect relationships;
- ways, including the use of randomization, of dealing with uncontrolled variation;
- the importance of design;
- the importance of good data quality and the determinants of data quality, including sampling and nonsampling errors;
- how observation is converted into information;
- the fundamental importance of variability along with its relationship to uncertainty and its impact on decision making;
- the meanings of probability;
- the various purposes of models including description, inference, and knowledge representation;
- a working knowledge of the processes involved in modelling, model fitting, model validation, and model improvements;
- the processes involved in getting to statistics from substantive questions;
- the role statistics plays in broad areas of application including actuarial science, environmental science, medical science, in industry, and in shaping public policy.

Undergraduate Courses

In terms of their overall goals, the courses at the core of the undergraduate programs may be blocked as follows:

- basic ideas in describing and coping with random variation: Stat 200, Math/Stat 302, Stat 304
- generating/assessing models and methods: Math 303, Stat 305, 406.

4

- implementing models and methods: Stat 306, 346, 404, 441, and 442
- generating and assessing data: Stat 344, 405, and 445.
- applications courses: Case Studies and Statistics in Quality Assurance are examples of topics in this area.

Graduate Studies
M.Sc. Program

The M.Sc. program normally takes two years for completion, although students with strong backgrounds in both statistics and mathematics may be able to complete it in one year.

The principle objective of the M.Sc. program is to provide students with a solid background in basic statistical theory and its applications. A thesis is required; for most students this will consist of a report of applied statistical work. This work may be undertaken as a project for the Statistical Consulting and Research Laboratory or for an agency outside the University. Students are expected to give a presentation of their thesis work in an informal seminar. There are no examinations other than course examinations.

Students seeking an M.Sc. in Statistics should have an undergraduate background in both mathematics and statistics, and possibly a few courses in computer science.

Ph.D. Program

Statistics students will normally be accepted as Ph.D. candidates only after they have received an M.Sc. in Statistics. The Ph.D. program typically involves 3-4 years of study beyond the M.Sc., but a student with a strong undergraduate background in mathematics and statistics (including courses in linear algebra, real analysis, linear models and mathematical statistics) may be able to complete both the Master's and Ph.D. program in 4 years. No formal course requirements are specified beyond the M.Sc. level, but students should take a sufficient number of courses to insure a broad exposure to both graduate level statistics and mathematics.

Students are expected to take a qualifying exam on fourth year B.Sc./first year M.Sc. material within their first year. After passing this examination, the student should decide on an area for thesis research and find a supervisor. Within the first two and a half years, the student is expected to prepare a thesis proposal and to take an oral comprehensive examination. In this exam, the student is questioned on the research proposal, two chosen areas of statistics (which may be related to the proposal), and mathematical statistics and probability. At the conclusion of his/her research, the student must defend it at an oral examination given by the Faculty of Graduate Studies.

Graduate Courses

The following courses are offered on a regular basis by the Statistics Department, although not all courses are offered every year. Courses in probability are offered by the Mathematics Department.

519	Theoretical Statistics
520	Statistical Decision Theory
521	Foundations of Multivariate Analysis
522	Asymptotic Theory and Conditional Inference
530	Bayesian Inference and Decsion
531	Reliability Theory
532	Sequential Statistical Procedures
533	Survival Analysis
534	Statistics in Quality Control
541	Applied Multivariate Analysis
542	Analysis of Categorical Data
543	Time Series Analysis
544	Theory of Sampling
545	Data Analysis
546	Nonparametric Statistical Methods
547	Topics in Statistics (recent offerings have included Density Estimation, Collection and Analysis of Official Statistics, Generalized Lienar Models, Spatial Statistics, Estimation Based on Ranks)
550	Techniques of Statistical Consulting
551	Statistical Consulting

Admissions Procedure

Students wishing to enter the program in September should apply by mid-March. Students wishing to be considered for teaching or research assistant support are advised to apply by January. Application forms and further information are available from:

Graduate Committee
Department of Statistics
The University of British Columbia
Vancouver, B.C. V6T 1Z2
Canada

Financial Aid

All applicants are automatically considered for teaching and research assistantships. A teaching assistant may be assigned to grade papers, assist in computing labs, or conduct problem sessions. Research assistantships are typically funded by SCARL or by research grants held by individual faculty members.

Natural Sciences and Engineering Research Council graduate fellowships are available to Canadian citizens and landed immigrants. Applications for these awards, funded by the Canadian government, must be made through the Department.

University Graduate Fellowships are available to both Canadian and foreign students. Applications, due in January, are made through the Department.

Statistics Fund for Excellence

As part of the campus phase of the World of Opportunity Campaign, the Department of Statistics proposed the creation of an endowment fund, to be called the Statistics Fund for Excellence. The purpose of this fund was to facilitate our continued efforts to enhance our academic programs, particularly our efforts to attract increased numbers of excellent students at both the graduate and undergraduate levels.

Income from the fund would be used to provide: prizes to recognize and encourage outstanding undergraduate and graduate students; assistance with travel expenses for incoming graduate students from foreign countries for whom the cost of air travel is a fundamental obstacle to entry into graduate studies at UBC, and for current graduate students to travel to research conferences; supplementary support, in the form of Statistics Graduate Fellowships, for outstanding incoming or current graduate students; and, journal subscriptions to upgrade and sustain the departmental Reading Room, thereby enhancing student access to the current research literature. These proposed expenditures are intended to increase the visibility of the discipline, to recognize outstanding students and encourage them to identify with the discipline, to allow us to be more competitive in recruiting graduate students, and to enhance graduate studies in the discipline at UBC.

This proposal was approved and at the end of 1992, after the first full year of fund-raising, over one-third of the initial target had already been raised. To date, all contributions have been made by the department's faculty and staff. All such donations to the fund during the first five years of fund-raising (until the end of 1996) will be matched by provincial government or other university sources.

Others may also donate to this fund. Individuals wishing to do so may mail cheques made payable to UBC/Statistics Fund for Excellence to the Department of Statistics. Receipts for income tax purposes will automatically be issued to all donors. All income derived from the fund will be used for the purposes described above.

STATISTICAL CONSULTING AND RESEARCH LABORATORY

The Statistical Consulting and Research Laboratory (SCARL) is operated by the Department of Statistics. SCARL's services are available to all faculties as well as to off-campus clients. The services SCARL provides cover the spectrum of statistical design and analysis, as well as ancillary areas. SCARL offers assistance in the formulation of research questions, in the design of experiments and surveys, in the choice and explanation of statistical methodology, in statistical computing and graphics, in data analysis, and in the interpretation of findings. Ancillary areas include organizing short courses and presentations, and budget estimation for grant proposals. The level of assistance SCARL provides is matched to clients' needs, and ranges from brief discussion and advice to a total service, and also to in-depth collaborative research.

Through such collaborations, SCARL assists in the professional development of faculty and in the training of graduate students in the Department of Statistics. SCARL works in conjunction with graduate courses in statistical consulting, and facilitates the training of students in this skill. In addition, graduate students are involved as research assistants at all levels of specific projects, from meeting with clients to analyzing their data and producing final reports. Some of these projects develop into thesis research topics for graduate students in Statistics. Most faculty members have some involvement with SCARL, and both methodological and applied

publications have resulted from this involvement. Faculty and graduate students meet weekly to review and discuss current and future work.

SCARL plays an active role in continuing education on and off campus, giving seminars and workshops to various departments, research groups, and teaching hospitals, all involved in quantitative research. These seminars on topics of interest to particular groups provide an effective means of introducing them to the scope of services and expertise provided by SCARL, as well as enhancing their awareness of some of the principles of good experimental design and of recent advances in statistical methodology and software.

In addition to the involvement of faculty and graduate students, SCARL has some permanent staff, and is headed by the Director, Dr. Malcolm Greig. The permanent staff gives SCARL the flexibility to handle inquiries at all times of the year, and to be able to take on small-to-medium projects fairly easily. SCARL's access to the department's state-of-the-art computing hardware and software permits avenues of research which would be prohibitively expensive on the University's central facilities; it also encourages students to experiment with statistical packages and novel techniques, sometimes of their own devising.

The SCARL office is always receptive to enquiries, and interested parties are encouraged to enquire about the work and the services of SCARL.
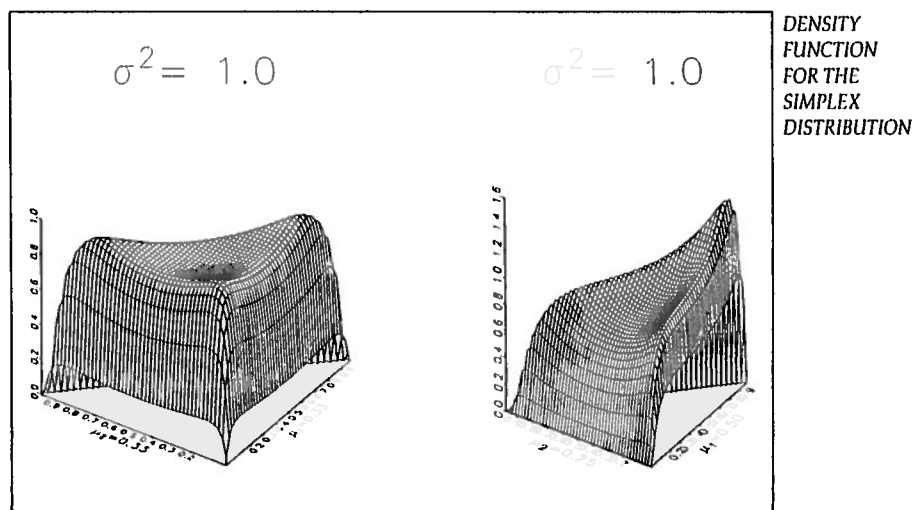
Computing Facilities

The main computing facility is a cluster of machines and terminals linked by Ethernet. This facility is constantly expanding and improving, and at the moment, it includes 15 SUN SPARCstations, 10 NCD X-window terminals, one PC 386 and one Macintosh II. In addition, members of the department have access to the many other computing facilities available on the UBC network.

The departmental computing facility is used by faculty members, graduate students, research staff and visitors. SCARL, assisted by faculty members, is responsible for the maintenance and development of this facility for the benefit of all department members.

There is a separate computing laboratory for undergraduate instruction, shared with the Department of Mathematics. In the past few years, this laboratory has consisted of 25 PC's (286 and 386) on a Novell network. A new under-graduate instructional laboratory, based on a second cluster composed of

a SPARC station server and 25 NCD X-window terminals, will be added in the near future.

The operating systems for most of the computing equipment is SunOS 4.1 or Solaris 2.1. On our main cluster, there are many software development tools including C and Fortran compilers, the Slatec mathematical library, the Linpack library, and SUN's Guide user interface editor, to name a few. In addition, there are a number of specialized pieces of software such as Splus, Maple, Xgobi, Xlispstat and CART. Splus provides extensive graphics capability as well as a programming environment where new statistical methods can easily be implemented. The computing facilities available on the UBC network provides access to yet more software including Mathematica, BMDP, GLIM, MINITAB, SAS, SHAZAM, and SPSS.
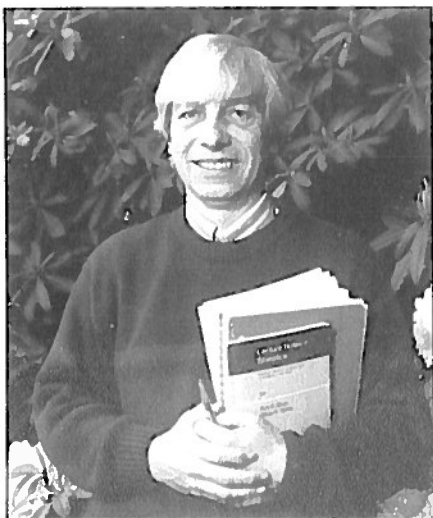
Malcolm Greig

DIRECTOR, STATISTICAL
CONSULTING AND RESEARCH
LABORATORY

B.Sc. (1964), University College London

Ph.D. (1972), University of California at Berkeley

Post-Doctoral Fellow, Simon Fraser University (1970-72)

The University of British Columbia (1972-present)

Research Interests

My main statistical research interests are in applied statistics, contingency tables, computational algorithms and statistical software. Applied statistics is a broad field, but so is my interest, which includes both the statistics themselves and their applications.

Before I joined the Department of Statistics, I was with University Computing Services at UBC. Taken jointly, for over twenty years some part of my responsibilities has been statistical consulting. As a consultant, one encounters a great variety of problems. There is considerable variety in the fields of application (I have had clients from all twelve faculties at UBC), in the statistical techniques required, and in the statistical sophistication needed. Of course, many problems are routine from a statistical point of view, but many are not; for example, some clients only turn to consultants when the problem is not a textbook case.

The primary aim of a consultant is to help the client, and to answer the client's questions (or at least, the questions the client ought to be asking). Sometimes there are additional constraints; as well as the obvious time and cost constraints, one might prefer to use techniques the client can understand, explain, and defend and publish. In the consulting process, the consultant must understand the client's problem, translate this into an appropriate statistical model, decide how to analyze that model and how to implement the analysis, and then interpret the results for the client. The role of the consultant in solving these problems is an intellectually stimulating one.

Aside from the challenge cited above, in the course of the year a consultant will typically meet several problems that could be developed into statistical research projects and publications. When one has to answer a client's question, sometimes the inadequacies of the available statistical techniques are brought home quite forcefully. Of course, finding research topics is only part of it; one must pick out those one can do something with, and do it. All the selected publications I have listed arose out of clients' questions, either directly or indirectly.

Another major area of my research interest is in combinatorics. My initial interest arose from some requests for balanced incomplete block designs (BIBD's). I have concentrated on solving, by construction, existence problems in block designs. Although BIBD's have some well-known statistical applications, the size of designs I have looked at is too large for statistical applications, although there are applications in non-statistical areas. In addition to BIBD's, I have worked on construction problems in resolvable and almost resolvable BIBD's, group divisible designs, transversal designs, pairwise balanced designs, tournaments and coding theory.

Selected Publications

de Jong, P., Greig, M. and Madan, D. (1983). "Testing for random pairing." *Journal of the American Statistical Association*, 78, 332-336.

Greig, M. (1990). "Some balanced incomplete block design constructions." *Congressus Numerantium*, 77, 121-134.

Greig, M. (1993). "Reversal of Hayter's inequality in multiple comparison procedures." Submitted for publication.

Greig, M. (1993). "Balanced incomplete block designs with a block size of 7." Submitted for publication.

Greig, M. (1993). "Designs from configurations in projective planes." Submitted for publication.

Ping H. Ma

SCARL CONSULTANT

B.Sc. (1984), University of British Columbia

M.Sc. (1987), University of British Columbia

The University of British Columbia (1987-present)

Research Interests

I am interested in methodological issues arising from particular consulting projects. Most of my statistical consulting is collaborative work with researchers from various departments at UBC. For example, comparison of the bond strength of various dental cements revealed the breaking mechanism involved more than one mode. This led to modeling of mixture data and, more importantly, provided a better understanding of the breaking mechanism. This led to a future study of the basic fracture mechanism of dental cements. Since fracture toughness is a physical property of material, independent of other extraneous factors, we believe that it will provide a better way to characterize performance — as compared with the usual parameters, such as tensile and shear bond strengths — and hence a better predictor of in-vitro performance of dental cements.

I am also interested in examining very skewed data arising from ecologic studies or assessment of soil contamination. Instead of transforming the data, I have been trying to use mixture models or other more flexible models, such as the generalized Poisson distribution, to describe the underlying processes; the results usually provide the researchers with a better understanding of the underlying processes. In assessment of soil contamination, heavy censoring, of up to 20% or 40% (for example, due to levels below detection limit), is not uncommon and makes the analysis more challenging.

In addition to consulting projects which have arisen locally, I also worked with marine engineers in Newfoundland. The main objective of a series of projects was to model the performance of ship hulls in relation to a set of design parameters of the ship architecture, as well as another set of parameters describing the sea condition in which the ship travelled. Questions of interest include prediction of ship performance, optimal design of hull forms (within practical constraints of ship architecture) involving response surface methodology and design of better experiment.

I am currently working on a project regarding factors which may affect the quality of salmon. In this study, fish quality is assessed by human panelists for 28 different attributes, each of which is recorded on a visual analog scale. Since, for instance, neither too fishy nor not fishy at all is desirable, it is not obvious how to combine the scores (of various attributes) in order to evaluate and to compare quality. Data-reduction techniques, such as principal components, provide various kinds of combined scores, but often it is difficult to interpret the results derived from such combined scores.

Selected Publications

Duleba, A., Rowe, T., Ma, P.H. and Colins, J.A. (1992). "Prognostic factors in assessment and management of male infertility." *Human Reproduction,* **7,** 1388-1393.

Mojon, P., Hawbolt, E.B., MacEntee, M.I. and Ma, P.H. (1992). "Early bond strength of luting cements to a precious alloy." *Journal of Dental Research,* **71,** 1633-1639.

Zilsel, J., Ma, P.H. and Beatty, J.T. (1992). "Derivation of a mathematical expression useful for the construction of complete genomic libraries." *Gene,* **120,** 89-92.

Rick White

INSTRUCTIONAL LAB MANAGER AND
SCARL CONSULTANT

*B.Sc. (1990), The University of British
Columbia*

*M.Sc. (1992), The University of British
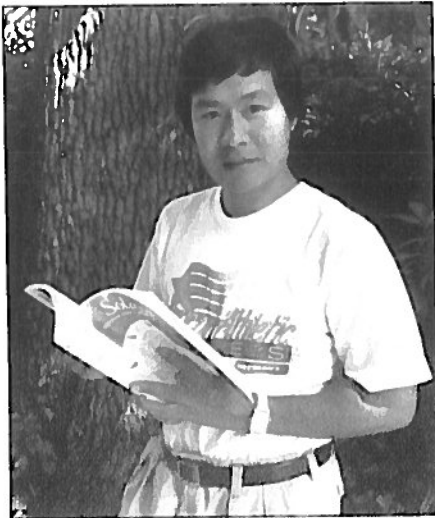Columbia*

Research Interests

My research interests are in biostatistics and multivariate statistics. I am currently collaborating on several biostatistical projects with investigators from the Department of Medicine at UBC.

My main current project is a study, directed by John Petkau and Sverre Vedal (Respiratory Division), of relationships between ambient levels of air pollution and peak flow and respiratory symptoms in children. These respiratory responses were obtained on a daily basis over a two year period from cohorts of asthmatic and control children in Port Alberni, B.C. The primary focus is on the respiratory effects of ambient inhalable particles smaller than 10 micrometers in diameter (PM10).

I am also involved in analyzing the data collected in a clinical trial for herpes carried out by an investigator from the Division of Infectious Diseases. Work is just beginning on the final longitudinal analyses of the data collected in a multi-centre clinical trial for multiple sclerosis, in which a research group centered in the Division of Neurology was one of the participating teams, but this project will require substantial effort in the near future. Much of my work on these projects requires extensive and intensive computing, so I also have strong interests in both statistical computing and exploratory data analysis.

Selected Publications

White, R.A. (1992). "The Detection and Testing of Multivariate Outliers", M.Sc. thesis.

Hongbin Zhang

STATISTICAL COMPUTING ANALYST

B.Sc. (1985), Peking University

M.Eng. (1987), Academia Sinica

M.Sc. (1993), The University of British Columbia

The University of British Columbia (1993-present)

Research Interests:

As an undergraduate, I studied mathematics. As a graduate student, I studied computer science and statistics, with concentrations in computer graphics and biostatistics respectively. In my M.Sc. project, under the direction of Dr. Harry Joe, I studied problems related to the 30 day operative status of coronary artery bypass surgery patients. We were able to identify risk factors and set up risk stratification algorithms. My interest in applied statistics continues and, through my work at SCARL, my interests have broadened to include statistical computing.

My job as Statistical Computing Analyst involves many different activities. This includes anything from keeping the Department's network up and running to advising users about computing and statistical problems. In addition, I am also responsible for doing statistical data analyses and computing, as well as providing software development for statistical methodological research.

Selected Publications

Zhang, H. (1988). "Implementation of 2c level of GKS on Sun Workstation." *Software Industry,* **12,** 20-24.

Zhang, H. (1993). "Risk prediction models for binary response variable for the coronary bypass operation", M.Sc. thesis.