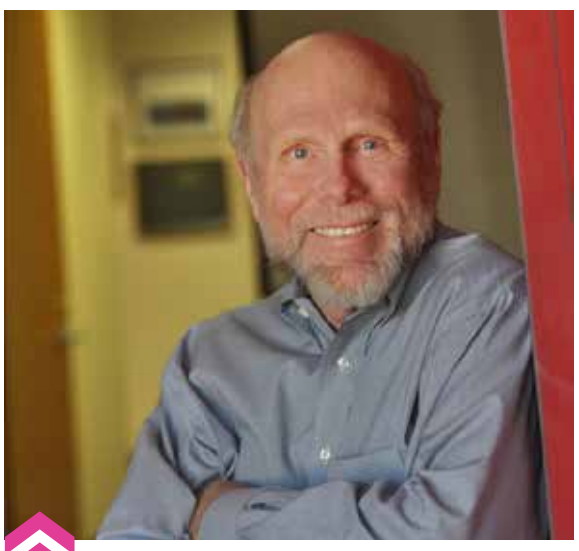
DAVID SPROTT DISTINGUISHED LECTURE BY

RAYMOND J. CARROLL



RAYMOND J. CARROLL

Raymond J. Carroll (www.stat.tamu.edu/~carroll) is the Jill and Stuart A. Harlin '83 Distinguished Professor of Statistics at Texas A&M University, and the Director of the Texas A&M Institute for Applied Mathematics and Computational Science (http://iamcs.tamu.edu). He is also Distinguished Professor of Statistics at the University of Technology, Sydney. He received an honorary doctorate from the Institut de Statistique, Université Catholique de Louvain, in 2012.

David A. Sprott (1930-2013)

Professor David Sprott was the first Chair (1967-1975) of the Department of Statistics and Actuarial Science at the University of Waterloo and first Dean of the Faculty of Mathematics (1967-1972). The David Sprott Distinguished Lecture Series was created in recognition of his tremendous leadership at a formative time of our department, as well as his highly influential research in statistical science.

Constrained maximum likelihood estimation for model calibration using summary-level information from external big data sources.

Thursday, September 24, 2015 | 4 p.m.
MC 4021, University of Waterloo
Reception will follow in the M3 Bruce White Atrium

Information from various public and private data sources of extremely large sample sizes are now increasingly available for research purposes. Statistical methods are needed for utilizing information from such big data sources while analyzing data from individual studies that may collect more detailed information required for addressing specific hypotheses of interest. We consider the problem of building regression models based on individual-level data from an "internal" study while utilizing summary-level information, such as information on parameters for reduced models, from an "external" big-data source. We identify a set of constraints that link internal and external models. These constraints are used to develop a framework for semiparametric maximum likelihood inference that allows the distribution of the covariates to be estimated using either the internal sample or an external reference sample. We develop extensions for handling complex stratified sampling designs, such as case-control sampling, for the internal study. Asymptotic theory and variance estimators are developed for each case. We use simulation studies and a real data application to assess the performance of the proposed methods.

UNIVERSITY OF WATERLOO
FACULTY OF MATHEMATICS

C006080

C008818