DISTINGUISHED LECTURE SERIES IN STATISTICAL SCIENCE

Raymond J. Carroll

Department of Statistics, Texas A&M University

Raymond J. Carroll is Distinguished Professor of Statistics and Nutrition at Texas A&M University, and Distinguished Professor at the University of Technology Sydney, Australia. He was the first statistician to receive a U.S. National Cancer Institute MERIT Award. He is the Director of the Texas A&M Institute for Applied Mathematics and Computational Science. He served as editor of Biometrics and the Journal of the American Statistical Association. He has won many honors in the profession, including the COPSS Presidents' Award and the Fisher Award and Lecture. Prof. Carroll's research has long been diverse, and has been focused on complex data scenarios, recently including methodology for multivariate functional data of different types, data integration across platforms/outcomes, uncertainties of predictor measurement (measurement error), gene-environment interactions, general semiparametrics, molecular biology of nutrition, radiation epidemiology and the risk of thyroid cancer, and physical activity. He is deeply involved in nutritional surveillance, nutritional epidemiology and dietary patterns, where the data complexity is immense due to the multivariate nature of dietary intakes and the uncertainties of dietary ascertainment.

October 4 and 5, 2017 The Fields Institute, Room 230

General Lecture: Oct 4, 2-3 pm

A Personal Tour: The Complex Nature of Nutrient and Food Intakes, and the Effects this Complexity has on Understanding Dietary Distributions and Their Effects on Mortality and Chronic Disease

I have spent many years developing statistical methods to understand how to measure dietary intakes in a population and how to relate such measures to mortality and chronic diseases. In animal experiments, clinical trials of different dietary patterns have been performed by my colleagues at Texas A&M and show, in stunning fashion, that for example a fish oil enhanced diet is protective against colon cancer, DNA damage, deleterious gene expression, etc., when compared to a corn oil enhanced diet (think potato chips).

In humans, the statistical questions are much more difficult, because it is impossible, in current practice, to measure an individual's long-term average dietary intakes across multiple foods and nutrients. This impossibility has many facets, which I will briefly review. This statistical issue, along with the media focus on dietary bullets (kale anyone?), has resulted in massive confusion, and sometimes silly conclusions. For example, what % of U.S. children eat an alarmingly poor diet: a simplistic analysis says 30%, but a more nuanced one that accounts for the measurement properties of dietary instruments says 8%, a massive difference. My aim in part is to show that focusing on dietary patterns, instead of magic bullets, leads to far more robust statistical conclusions than focusing on one potential bullet at a time.

Technical Lecture: Oct 5, 2-3 pm

Semiparametric Analysis of Complex Polygenic Gene-Environment Interactions in Case-Control Studies

Many methods have been proposed recently for efficient analysis of case-control studies of gene-environment interactions using a retrospective likelihood framework that exploits the natural assumption of gene-environment independence in the underlying population. We will review some of this literature and discuss some of the fairly astonishing gains in efficiency that are possible. However, for polygenic modeling of gene-environment interactions, a topic of increasing scientific interest, applications of retrospective methods have been limited due to a requirement in the literature for parametric modeling of the distribution of the genetic factors, which is difficult because of the complex nature of polygenic data.

We propose a fully general, computationally simple, efficient semiparametric method for analysis of case-control studies that allows exploitation of the assumption of gene-environment independence without any further parametric modeling assumptions about the marginal distributions of any of the two sets of factors. The method relies on the key observation that an underlying efficient profile likelihood depends on the distribution of genetic factors only through certain expectation terms that can be evaluated empirically. We develop asymptotic inferential theory for the estimator and evaluate numerical performance using simulation studies. An application of the method is illustrated using a case-control study of breast cancer.

For more information, please visit:
www.fields.utoronto.ca/activities/17-18/distinguished-lecture-series-statistical-sciences

Ontario

CANSSI INCASS

NSERC CRSNG