****

**Short Bio of Professor Lei Sun**

Lei Sun is a Full Professor of Statistics and Biostatistics at the University of Toronto. She studied mathematics at Fudan University and obtained her PhD in statistics from the University of Chicago in 2001. Prof. Sun’s research area is in statistical genetics and genomics, developing statistical methods and computational tools to solve problems arising from genetic studies of complex human traits. Her research is funded by both the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Canadian Institutes of Health Research (CIHR). Prof. Sun is the recipient of the [2018 NSERC Discovery Accelerator Supplements Program Award](http://www.nserc-crsng.gc.ca/Professors-Professeurs/Grants-Subs/DGAS-SGSA_eng.asp), and she is also the recipient of the [2017 CRM-SSC prize in Statistics](https://ssc.ca/en/awards/2017/lei-sun), award by the Centre de Recherches Mathématiques (CRM) and the Statistical Society of Canada (SSC) “in recognition of a statistical scientist’s professional accomplishments in research during the first fifteen years after having received a doctorate”.

**Title:** The X factor: robust approaches to X-chromosome-inclusive whole-genome association studies

**Speaker:**

Lei Sun

Professor

Department of Statistical Sciences, Faculty of Arts and Science

Division of Biostatistics, Dalla Lana School of Public Health

University of Toronto

<http://www.utstat.toronto.edu/sun/>

**Abstract:**

`Whole-genome’ association scans often exclude X-chromosome variants, due to a number of analytical considerations. Some are relatively straightforward. For example, can the method handle both binary and continuous outcomes? Is the analysis robust to sex-specific allele frequencies? Other issues are subtler, some of which had not been given much thoughts before. For example, we know that for association analysis of an autosome SNP, the choice of the reference allele only affects the direction of the effect estimate but not the strength of the association. But does this remain true for X-chromosome variants? Further, what is the added value of considering sex-genotype interaction? And finally, how do we handle the random, skewed, or no X-inactivation uncertainty?

In this talk, I first focus on the traditional location/mean test and present a unifying, regression-based association test that simultaneously resolves these X-chromosome complexities (Chen et al. *arXiv*:1811.00964). I will provide theoretical justifications and empirical evidence for its robustness in the presence of various aforementioned model uncertainties, as well as for its improved power under certain alternatives as compared with existing approaches. I will then discuss extension of the recent scale-test (Soave and Sun, 2017 *Biometrics*) to the X-chromosome, where testing for phenotypic variance heterogeneity is used to inform presence of interactions that are difficult to model directly due to missing information, multiple testing, or other considerations (Deng et al. 2019 *Genetic Epidemiology*). Application to the UK Biobank data points to a possible X-chromosome polygenic inheritance model and suggests that height could be potentially enriched for gene-gene and gene-environment interactions.

For both location and scale association tests of the X-chromosome, I emphasize the importance of recognizing sex as an inherent confounder. The examples here also demonstrate that `simple’ linear regression remains an important analytical framework in the era of big and complex data.