UTHealth
The University of Texas Health Science Center at Houston
School of Biomedical Informatics

# DSICCR Tuesday Seminar Series

## Tuesday at Noon, Click Here to Join

## StocSum: stochastic summary statistics for whole genome sequencing studies


Han Chen, Ph.D.
Associate Professor
Department of Epidemiology, Human Genetics, and Environmental Sciences
School of Public Health
University of Texas Health Science Center at Houston



Genomic summary statistics, usually defined as single-variant test results from genome-wide association studies, have been widely used to advance the genetics field in a wide range of applications. However, many applications that involve multiple genetic variants also require their correlations or linkage disequilibrium (LD) information, often obtained from an external reference panel. While these approaches have been successful for common genetic variants from European-ancestry individuals, it is usually difficult to find external reference panels that represent the LD structure for rare genetic variants, such as those from whole genome sequencing (WGS) studies, or underrepresented and admixed populations, limiting the scope of applications for genomic summary statistics. Here we introduce StocSum, a novel statistical framework for generating, managing, and analyzing stochastic summary statistics using random vectors. Regardless of the complex sample correlation structure, StocSum always scales linearly with both the sample size and the number of genetic variants in computing stochastic summary statistics from individual-level genotype data. We develop various downstream applications using StocSum including single-variant, conditional association, gene-environment interaction, variant set tests, as well as meta-analysis and LD score regression tools, all of which do not depend on the study sample size and are therefore applicable to large samples. We also demonstrate the accuracy and computational efficiency of StocSum without any external reference panels, using two cohorts from the Trans-Omics for Precision Medicine WGS studies, with longitudinal phenotype measures and/or study participants from underrepresented and admixed populations.


Tuesday, September 6th, 2022. 12p – 1p. Webcast


Contact: Xiaohong.Bi@uth.tmc.edu


#SBMIseminar