Scalable protein-DNA binding changer test for insertion and deletion of bases in the genome

Sunyoung Shin
Assistant Professor
Department of Mathematical Sciences
University of Texas @ Dallas
Friday, April 1, 2022
3:30 p.m.
Marrs McLean Science Building 301

Abstract: Mutations in the noncoding DNA, which represents approximately 99\% of the human genome, have been crucial to understand disease mechanisms through dysregulation of disease-associated genes. One key element in gene regulation that noncoding mutations mediate is the binding of proteins to DNA sequences. Insertion and deletion of bases (InDels) are the second most common type of mutations, following single nucleotide polymorphisms, that may impact protein-DNA binding. However, no existing methods can estimate and test the effects of InDels on the process of protein-DNA binding. We develop a novel statistical test, named binding changer test (BC test), using a Markov model to evaluate the impact of InDels and identify InDels altering protein-DNA binding. The test predicts binding changer InDels of regulatory significance with an efficient importance sampling algorithm generating background sequences in favor of large binding affinity changes. Simulation studies demonstrate its excellent performance. The application to human leukemia data uncovers candidate pathologic InDels on modulating MYC binding in leukemic patients. We develop R package atIndel, which is available on GitHub.

Bio: Dr. Sunyoung Shin is an Assistant Professor of Statistics in Department of Mathematical Sciences at University of Texas at Dallas. She did her postdoctoral training at University of Wisconsin Madison for 3 years on statistical and computational genomics. Her PhD in Statistics is from University of North Carolina at Chapel Hill in 2014. Her research interests include development of statistical learning and testing methods to analyze and integrate big data.

Please join us for refreshments in the department foyer at 3:00.