DBDS Workshop in Biostatistics
Remote Access Only:
Contact kkanagaw@stanford.edu for Zoom dial-in details.

DATE:	October 8, 2020
TIME:	2:30-3:50pm
TITLE:	Statistical analysis of single cell CRISPR screens
SPEAKER:	**Eugene Katsevich** Assistant Professor of Statistics Wharton University of Pennsylvania

Abstract:
Mapping gene-enhancer regulatory relationships is key to unraveling molecular disease mechanisms based on GWAS associations in non-coding regions. This problem is notoriously challenging: there is a many-to-many mapping between genes and enhancers, and enhancers can be located far from their target genes. Recently developed CRISPR regulatory screens (CRSs) based on single cell RNA-seq (scRNA-seq) are a promising high-throughput experimental approach to this problem. They operate by infecting a population of cells with thousands of CRISPR guide RNAs (gRNAs), each targeting an enhancer. Each cell receives a random combination of CRISPR gRNAs, which suppress the action of their corresponding enhancers. The gRNAs and whole transcriptome in each cell are then recovered through scRNA-seq. CRSs provide more direct evidence of regulation than existing methods based on epigenetic data or even chromatin conformation. However, the analysis of these screens presents significant statistical challenges, some inherited from scRNA-seq analysis (modeling single cell gene expression) and some unique to CRISPR perturbation screens (the confounding effect of sequencing depth). In this talk, I will first give some background on single cell CRISPR screen technology. I will then present the first genome-wide single cell CRS dataset (Gasperini et al. 2019) and discuss challenges that arose in its initial analysis. Finally, I will present a novel methodology for the analysis of this data based on the conditional randomization test. The key idea is to base inference on the randomness in the assortment of gRNAs among cells rather than on the randomness in single cell gene expression, since the former is easier to model than the latter.

Suggested Readings:
- "Towards a comprehensive catalogue of validated and target-linked human enhancers" (Nature Reviews Genetics 2020).
- "A Genome-wide Framework for Mapping Gene Regulation via Cellular Genetic Screens" (Cell 2019).
- "Panning for gold: `model-X' knockoffs for high dimensional controlled variable selection" (Journal of the Royal Statistical Society, Series B 2018) and "Fast and Powerful Conditional Randomization Testing via Distillation" (arXiv 2020).
- "Conditional resampling improves sensitivity and specificity of single cell CRISPR regulatory screens" (bioRxiv 2020).