Statistical problems of DNA methylation data analysis

Shuying Sun, Texas State University

Friday, November 19
12:00 - 1:00 pm

Abstract: DNA methylation is an important type of epigenetic event. It regulates the expression of certain genes that are closely related to tumor growth and normal cell development. DNA methylation involves the addition of a methyl group (CH_3) to a cytosine site (C) that pairs with a guanine (G) in a human genome. This pair is often called a CG site, and there are many CG sites within each gene, which is usually at least several hundred bases long. There are about 29,000,000 CG sites in a human genome. On these CG sites, methylation may occur. There are different types of methylation patterns that are associated with cancers. The identification of cancer methylation patterns is an important research topic, which may pave the way for discovering new DNA methylation biomarkers that are useful for better diagnosis and treatment. In this talk, I will discuss the statistic problems and projects related to the identification of DNA methylation patterns. For example, what should we do if the normality assumption is violated? How should we deal with the heterogenous data in cancer samples? How to find and interpret the key hub CG sites or genes that are highly correlated with others? How to deal with the multiple-test correction problems?

Bio: Dr. Shuying Sun received her Ph.D. in statistics from the University of Toronto in Canada and is currently an associate professor in the Department of Mathematics, Texas State University. Dr. Sun has interests in statistical genetics and bioinformatics and has published 30 peer-reviewed research articles in high-impact journals. Dr. Sun's research focuses on addressing challenging genetic and epigenetic questions using statistical and computational methods. She has been developing statistical methodologies and software packages for genomic and epigenomic problems using Bayesian methods, hidden Markov models, Markov Chain Monte Carlo algorithms, and linear models.