

Statistics Seminar  
Department of Mathematical Sciences

<b>DATE:</b>	Thursday, Month 31, 2017
<b>TIME:</b>	1:15pm - 2:15pm
<b>LOCATION:</b>	WH 100E
<b>SPEAKER:</b>	Yuan Fang, Binghamton University
<b>TITLE:</b>	Temporal probabilistic modeling of bacterial compositions derived from 16S rRNA sequencing

**Abstract**

The number of microbial and metagenomic studies has increased drastically due to advancements in next-generation sequencing-based measurement techniques. Statistical analysis and the validity of conclusions drawn from (time series) 16S rRNA and other metagenomic sequencing data is hampered by the presence of significant amount of noise and missing data (sampling zeros). Accounting uncertainty in microbiome data is often challenging due to the difficulty of obtaining biological replicates. Additionally, the compositional nature of current amplicon and metagenomic data differs from many other biological data types adding another challenge to the data analysis. To address these challenges in human microbiome research, we introduce a novel probabilistic approach to explicitly model overdispersion and sampling zeros by considering the temporal correlation between nearby time points using Gaussian Processes. The proposed Temporal Gaussian Process Model for Compositional Data Analysis (TGP-CODA) shows superior modeling performance compared to commonly used Dirichlet-multinomial, multinomial and non-parametric regression models on real and synthetic data. We demonstrate that the nonreplicative nature of human gut microbiota studies can be partially overcome by our method with proper experimental design of dense temporal sampling. We also show that different modeling approaches have a strong impact on ecological interpretation of the data, such as stationarity, persistence and environmental noise models.

From:

<http://www2.math.binghamton.edu/> - **Binghamton University Department of Mathematical Sciences**

Permanent link:

<http://www2.math.binghamton.edu/p/seminars/stat/191031>

Last update: **2019/10/30 17:12**

