

Data Science Seminar Series

Wednesday, February 15, 11:30-12:30 pm, OM1241

TITLE

Inference with Missing Data

SPEAKER

Dr. Mateen Shaikh

ABSTRACT

Missing data often occurs with data analysis. The missing data problem is a very old and well-known issue with a variety of historical solutions in statistics. This seminar will explore two methods of inference: multiple imputation and the expectation maximization algorithm. The first generates many pseudo-datasets by stochastically replacing the missing values through some combination of a prior or resampling. The second approach replaces not the missing values, but the corresponding sufficient statistics with their expected values. Again, the use of some prior on the missing data is exploited. Results from previous and current students conducting projects/theses on these problems are discussed.

BIOGRAPHY

Dr. Mateen Shaikh is an Assistant Professor of Statistics and Data Science at Thompson Rivers university. His major areas of research include model optimization in statistical machine learning, incomplete data, and models for mining unusual or problematic data.

https://www.tru.ca/science/masters-degrees/mscads/Data_Science_Seminar_Series.html