Colloquium Series
Department of Mathematics and Statistics

Catalina Medina

UC Irvine

Real-Time Inference of Infectious Disease Dynamics Using Viral Genetic Sequences

Abstract: The COVID-19 pandemic demonstrated that fast and accurate analysis of continually collected infectious disease surveillance data is crucial for situational awareness and policy making. Coalescent-based phylodynamic analysis can use genetic sequences of a pathogen to estimate changes in its effective population size, a measure of genetic diversity. These changes in effective population size can be connected to the changes in the number of infections in the population of interest under certain conditions. Phylodynamics is an important set of tools because its methods are often resilient to the ascertainment biases present in traditional surveillance data (e.g., preferentially testing symptomatic individuals). Unfortunately, it takes weeks or months to sequence and deposit the sampled pathogen genetic sequences into a database, making them available for such analyses. These reporting delays severely decrease precision of phylodynamic methods closer to present time, and for some models can lead to extreme biases. Here we present a method that affords reliable estimation of the effective population size trajectory closer to the time of data collection, allowing for policy decisions to be based on more recent data. Our work uses readily available historic times between sampling and reporting of sequenced samples for a population of interest, and incorporates this information into the sampling model to mitigate the effects of reporting delay in real-time analyses. We illustrate our methodology on simulated data and on SARS-CoV-2 sequences collected in the state of Washington in 2021.

Monday, Feb. 3, 1:05 – 1:50 pm in 1-109