JOEY FINGOLD, University of Guelph

A statistical framework for modelling underreported infectious diseases

Count time series models are popular for modelling surveillance data of infectious diseases, such as COVID-19 and influenza, due to their flexibility and interpretability. However, observed data on daily infections is subject to underreporting, as not every infected individual seeks medical care. Such a model is challenging to fit since it involves high-dimensional and serially correlated integer-valued unknowns. One solution is to consider an analogous continuous-valued surrogate model whose values are post-hoc mapped to integers. This procedure produces biased estimates as inference is performed using samples from such a surrogate model and not the thinned count autoregression itself. We propose a sequential importance sampling procedure to correct this misspecified model. We demonstrate its validity in a simulation study and its applicability for epidemic curve reconstruction using rotavirus data from Germany.