Notes on observational processes in epidemic models

Sang Woo Park and Benjamin M Bolker

July 15, 2019

1 Introduction

Mechanistic analyses of an epidemic time series allow us to infer the underlying transmission mechanism, estimate biologically relevant parameters, and predict the course of an epidemic. In order to make a precise and accurate inference, disease modelers often focus on developing more realistic process models. For example, England and Wales measles time series from the prevaccination era have been analyzed using many different models – these model account for time-varying transmission rates (Fine and Clarkson, 1982), realistic age structure (Schenzle, 1984), metapopulation structure (Xia et al., 2004), continuous-time infection process (Cauchemez and Ferguson, 2008), and extra-demographic variability (He et al., 2009).

Despite the amount of effort put into developing better process models, disease modelers often neglect details of the observation process associated with disease case reports. Most disease models assume that new cases are reported instantaneously when an individual is infected or shows symptom; one exception is the model by (He et al., 2009), which assumes that infections are counted upon recovery (because diagnosed cases are put to bed rest and are effectively no longer infectious).

Here, we use a simple SIR model to study how observation process affects statistical inference of underlying parameters. We show that making wrong assumptions about the observation process can lead to biases in parameter estimates and narrow confidence intervals.

2 Methods

The Susceptible-Infected-Removed (SIR) model describes how disease spreads in a homogeneous population:

$$\begin{aligned} \frac{dS}{dt} &= -\beta S \frac{I}{N} \\ \frac{dI}{dt} &= \beta S \frac{I}{N} - \gamma I \\ \frac{dR}{dt} &= \gamma I, \end{aligned} \tag{1}$$

where β is the contact rate, γ is the recovery rate, and $N = S + I + R$ is the total population size. Incidence measured upon infection ($i_1(t)$) can be defined by the integral:

$$i_1(t) = \int_t^{t+\Delta t} \beta S \frac{I}{N} dt, \tag{2}$$

where Δt is the reporting time step. Alternatively, we can keep track of cumulative incidence, C, by adding a new state variable $dC/dt = \beta SI/N$ and take the difference between two consecutive reporting periods: $i_1(t) = C(t + \delta t) - C(t)$. Likewise, incidence measured upon recovery ($i_2(t)$) can be defined by the integral:

$$i_2(t) = \int_t^{t+\Delta t} \gamma I dt, \tag{3}$$

or by the difference in cumulative number of recovered cases: $i_2(t) = R(t + \Delta t) - R(t)$. Finally, we assume that the reported incidence follows a negative binomial distribution with mean $\rho i_1(t)$ or $\rho i_2(t)$, where ρ is the reporting rate, and a dispersion parameter θ.

3 Results

In order to introduce the problem, we first compare the dynamics of $i_1(t)$ and $i_2(t)$ (Figure ?). Lags in reporting time delay the "observed" epidemic peak timing and reduce the size of the peak. However, these differences in the reporting time have small effect to the overall shape of the epidemic curve. In the presence of observation and process error, we may not expect to be able to distinguish between the two reporting processes. Note that incidence measured at these two periods are different from prevalence, which is represented by the state variable $I(t)$; the observed dynamics of incidence depend on the reporting time step, whereas those of prevalence do not (hence we expect incidence and prevalence to be similar only when the reporting time step is equal to the disease generation time). While it is uncommon, some models do not make a clear distinction between the two.

Although

References

Cauchemez, S. and N. M. Ferguson (2008). Likelihood-based estimation of continuous-time epidemic models from time-series data: application to measles transmission in London. *Journal of the Royal Society Interface 5*(25), 885–897.

Fine, P. E. and J. A. Clarkson (1982). Measles in England and Wales—I: an analysis of factors underlying seasonal patterns. *International journal of epidemiology 11*(1), 5–14.

He, D., E. L. Ionides, and A. A. King (2009). Plug-and-play inference for disease dynamics: measles in large and small populations as a case study. *Journal of the Royal Society Interface 7*(43), 271–283.

Schenzle, D. (1984). An age-structured model of pre-and post-vaccination measles transmission. *Mathematical Medicine and Biology: A Journal of the IMA 1*(2), 169–191.

Xia, Y., O. N. Bjørnstad, and B. T. Grenfell (2004). Measles metapopulation dynamics: a gravity model for epidemiological coupling and dynamics. *The American Naturalist 164*(2), 267–281.