Estimating COVID-19 Reproduction Numbers

Kevin Durant

June 2020

1 Introduction

This is an *almost* fully Bayesian approach to inferring the effective reproduction number of a COVID-19 epidemic based on reported infection counts in a given region. This number refers to the expected number of new infections caused by a single infected individual at a given stage of the epidemic.

The method discussed here is an extension of one described elsewhere by Kevin Systrom [2], which is itself based on work by Bettencourt and Ribeiro [1]. The key point made by the latter authors is that under the assumptions of a standard epidemic susceptible-infected (SIR) model, the effective reproduction number of a virus at time t can be estimated from the number of new cases recorded between times $t-1$ and t, and t and $t+1$.

More specifically, let R_t be the effective reproduction number at time t, and assume that this number remains constant over the interval $(t-1, t]$ (most likely a single day)[1]. Likewise, let λ_t be the average number of new infections that occur during this interval—i.e., the current rate of infection. One then has the following approximation [1]:

$$\lambda_t \approx \lambda_{t-1} \exp(\gamma(R_t - 1)),$$

in which γ is the reciprocal of the serial interval, or infectious period, of the virus. Equivalently,

$$R_t \approx \frac{1}{\gamma} \log\left(\frac{\lambda_t}{\lambda_{t-1}}\right) + 1.$$

The approach outlined here involves modelling the number of observed infections k_t as a stochastic process, and assuming that k_t depends on an underlying rate of infection λ_t via a negative binomial distribution (a Poisson

[1]Note that our notation differs slightly from that used by Bettencourt and Ribeiro—our R_t and λ_t correspond to their R_{t-1} and $\Delta T(t)$ respectively.

distribution can also be used). One can then infer the infection rates λ_t analytically, and use them to estimate effective reproduction numbers by applying equation (1).

2 The stochastic process

Let $\mathbf{k} = k_1, \ldots, k_{t-1}$ be a sequence of observed infection counts, and $\boldsymbol{\lambda}$ the corresponding sequence of unknown infection rates. The primary assumption is that each k and λ are related via a negative binomial distribution:

$$\mathrm{P}(k_t \mid \lambda_t, \mathbf{k}) = \mathrm{P}(k_t \mid \lambda_t) = NB\left(k_t \mid r, \frac{\lambda_t}{r + \lambda_t} = p_t\right), \tag{1}$$

where r is an unknown dispersion parameter and p_t is a reparameterisation of λ_t as a 'success' probability. Parameterised in this way, the negative binomial distribution converges to a Poisson distribution of rate λ_t as $r \to \infty$, and by adjusting r we can control the level of variance inherent to the distribution (smaller values of r result in higher variance).

Inference of the rate sequence $\boldsymbol{\lambda}$ is performed iteratively, by repeated application of Bayes' rule:

$$\mathrm{P}(\lambda_t \mid k_t, \mathbf{k}) \propto \mathrm{P}(k_t \mid \lambda_t)\, \mathrm{P}(\lambda_t \mid \mathbf{k}).$$

The first term on the right-hand side—the likelihood function—is simply the negative binomial distribution given above. The second term is a predictive prior on λ_t given only the *past* infection counts $\mathbf{k}$.

The conjugate prior for the negative binomial likelihood function (with known dispersion) is the beta distribution, and the change-of-variable formula for probability density functions implies that the corresponding prior on λ_t is the beta prime distribution, i.e.:

$$\mathrm{P}(p_t) = B(p_t \mid \alpha, \beta) \Rightarrow \mathrm{P}(\lambda_t) = BP(\lambda_t \mid \alpha, \beta)$$

In the case of a Gaussian stochastic process, one would derive such a predictive prior under the simple assumption of additive Gaussian noise on the previous latent variable, resulting in a distribution that is similar to the previous posterior but with higher variance. This involves solving an integral of the form

$$\mathrm{P}(\lambda_t \mid \mathbf{k}) = \int \mathrm{P}(\lambda_t \mid \lambda_{t-1})\, \mathrm{P}(\lambda_{t-1} \mid \mathbf{k})\, d\lambda_{t-1},$$

which is tractable in the Gaussian case.

Although the situation is not quite as straightforward here, we can achieve a similar outcome by assuming the relationship to the previous posterior directly: specifically, if

$$\mathrm{P}(\lambda_{t-1} \mid \mathbf{k}) \sim \Gamma(\lambda_{t-1} \mid \alpha_{t-1}, \beta_{t-1}),$$

we might assume a predictive prior of the form

$$\begin{aligned} \mathrm{P}(\lambda_t \mid \mathbf{k}) &\sim \Gamma(\lambda_t \mid \alpha_{t-1}/\tau, \beta_{t-1}/\tau) && (2) \\ &= \Gamma(\lambda_t \mid a_t, b_t). && (3) \end{aligned}$$

This is a straightfoward prior that has the same mean as the predictive posterior on λ_{t-1}, but a variance that is larger by a factor τ (the mean and variance of a gamma distribution are α/β and α/β^2 respectively).

One could just as easily make use of a predictive prior that introduces additive noise, unlike the multiplicative noise described above. The main reason for choosing multiplicative noise here is that changes in scale made to λ_t result in additive changes to R_t, so in this way one is effectively introducing additive noise into the derived reproduction number process.

The remaining details of the stochastic process now follow from assumptions (1) and (2). Firstly–and most importantly—the posterior on λ_t is given by

$$\begin{aligned} \mathrm{P}(\lambda_t \mid k_t, \mathbf{k}) &\propto \mathrm{P}(k_t \mid \lambda_t)\, \mathrm{P}(\lambda_t \mid \mathbf{k}) \\ &= \Gamma(k_t \mid \phi, \lambda_t)\, \Gamma(\lambda_t \mid a_t, b_t) \\ &\sim \lambda_t^{\phi} \exp(-k_t \lambda_t) \cdot \lambda_t^{a_t - 1} \exp(-b_t \lambda_t) \\ &\Rightarrow \Gamma(\lambda_t \mid a_t + \phi, b_t + k_t) \\ &= \Gamma(\lambda_t \mid \alpha_t, \beta_t). \end{aligned}$$

Secondly, we can derive the marginal likelihood of observation k_t given the previous observations:

$$\begin{aligned} \mathrm{P}(k_t \mid \mathbf{k}) &= \int \mathrm{P}(k_t, \lambda_t \mid \mathbf{k})\, d\lambda_t \\ &= \int \Gamma(k_t \mid \phi, \lambda_t)\, \Gamma(\lambda_t \mid a_t, b_t)\, d\lambda_t \\ &= \frac{k_t^{\phi-1} b_t^{a_t}}{\Gamma(\phi)\Gamma(a_t)} \int \lambda_t^{a_t + \phi - 1} \exp(-(b_t + k_t)\lambda_t)\, d\lambda_t \\ &= \frac{b_t^{a_t} k_t^{\phi-1}}{(b_t + k_t)^{a_t + \phi}} \frac{1}{\mathrm{B}(a_t, \phi)}, \end{aligned}$$

where $\mathrm{B}(x, y)$ denotes the beta function. This allows us to compute the overall marginal likelihood iteratively, since

$$\mathrm{P}(\mathbf{k}) = \prod_{i=1}^{t-1} \mathrm{P}(k_i \mid k_1, \ldots, k_{i-1}).$$

The marginal likelihood will allow us to compare the relative appropriateness of values of ϕ and τ, as well as of the alteration to the predictive prior suggested in the next section.

A regression-based prior

References

[1] L. M. A. Bettencourt and Ru. M. Ribeiro. "Real Time Bayesian Estimation of the Epidemic Potential of Emerging Infectious Diseases". In: *PLOS ONE* 3.5 (May 2008), pp. 1–9.

[2] K. Systrom. *The Metric We Need to Manage COVID-19*. Apr. 12, 2020. URL: `http://systrom.com/blog/the-metric-we-need-to-manage-covid-19`.