PHIGHT notes

April 19 2021

Let I_t denote *new* infections in day t and Y_t death in day t. Let

$$\begin{aligned} I_t &= I_{t-1}e^{c_t} + \delta_t \\ Y_t &= \sum_{s=1}^{t} f(s,t) I_s + \xi_t \end{aligned} \tag{1}$$

where δ_t and ξ_t are mean 0 random variables (independent of the other variables), $f(s,t)$ denotes the probability that someone infected at time s dies at time t, and $c_t > 0$ represents the evolution of the epidemic at time t. Taking $c_t = c$ a constant of time assumes stationarity (everything remains the same - not realistic). *Note that c_t is what I called B_t, i.e. the derivative of log death. I will show that below.*

We write

$$f(s,t) = d(s) f_0(s,t) \tag{2}$$

where $d(s)$ is the probability that someone infected at time s will eventually die and $f(s,t)$ is the probability that someone infected at time s and who will eventually die, will die at time t. We take $f_0(s,t)$, on the scale of days, to be a Gamma with mean 23.9 days and coefficient of variation 0.40.

At this point, we might use (1) as our model. But the I_t's are not observed. Instead, we recurse the first equation in (1) to find

$$\mathbb{E}[Y_t] = \sum_{s=1}^{t} I_1 f(s,t) e^{\sum_{r=1}^{s} c_r}. \tag{3}$$

Simplified Model We take $f_0(s,t)$ in (2) to be a point mass at $\delta = 24$ days (the mean time from infection to death). Then we get

$$\mathbb{E}[Y_t] = I_1 d(t-\delta) e^{\sum_{r=1}^{t-\delta} c_r}.$$

If we approximate $\log \mathbb{E}[Y_t]$ with $\mathbb{E}[\log(Y_t)]$ we further obtain

$$\mathbb{E}[L_t] = \log(d(t-\delta)) + \log I_1 + \sum_{r=1}^{t-\delta} c_r \equiv \nu(t), \tag{4}$$

where $L_t = \log(Y_t + 1)$, or equivalently,

$$\mathbb{E}[L_{t+\delta}] = \log(d(t)) + \log I_1 + \sum_{r=1}^{t} c_r \equiv \nu(t+\delta). \tag{5}$$

Note that if the exponential growth rate is constant, $c_r = c$, then $\nu(t+\delta) = \log(d(t)) + \log I_1 + c.t$, with derivative

$$\partial \mathbb{E}[L_{t+\delta}]/\partial t = d'(t)/d(t) + c.$$

Hence if c is not constant,

$$\partial \mathbb{E}[L_{t+\delta}]/\partial t = d'(t)/d(t) + c_t,$$

where $c_t = c(t)$ is the exponential growth of the epidemic at time t.

Note that the probability of dying $d(t)$ is allowed to change smoothly over time, which it likely did as hospitals were better prepared during the second wave. But assuming that the probability of dying is constant (OK assumption if we look at a short window of time; a few months maybe), then

$$\partial \mathbb{E}[L_{t+\delta}]/\partial t = c(t).$$

This is what we called $B(t)$, although it corresponds to the derivative of log deaths at $t+\delta$, not at t. (I can estimate the "state" of the disease at time t by looking at deaths δ days later.)

The cumulative death count at time $t+\delta$ is

$$\mathbb{E}(C_{t+\delta}) = \sum_{s=1}^{t+\delta} \mathbb{E}(Y_s) = I_1 \sum_{s=1}^{t+\delta} d(s-\delta) e^{\sum_{r=1}^{s-\delta} c(r)},$$

and if the death probability is constant, then

$$\mathbb{E}(C_{t+\delta}) = I_1 \ d \ \sum_{s=1}^{t+\delta} e^{\sum_{r=1}^{s-\delta} B(r)}.$$

That is to say, the relationship between B and cumulative death is definitely not linear or easy to imagine.

So why look at B? It is the instantaneous exponential growth rate, which should be able to capture changes in the disease dynamics much better than observed cumulative death incidence.

References

Unwin, H. J. T., Mishra, S., Bradley, V. C., Gandy, A., Mellan, T. A., Coupland, H., Ish-Horowicz, J., Vollmer, M. A., Whittaker, C., Filippi, S. L. et al. (2020). State-level tracking of COVID-19 in the United States. *Nature communications* **11** 1–9.