

# Case reproductive numbers across a simple epidemic

The variance conservation conjecture is true.

The conjecture asserts that the “excess” (i.e., non-Poisson) variation of individual case reproductive numbers in the standard SIR model outbreak is 1.

## Proof

Define  $x$  as the proportion of the population susceptible,  $y$  as the proportion infectious, and incidence  $i = Bxy$ .

Let  $f$  be the distribution of residence times in the infectious compartment. This is exponential in the standard SIR and we will make that assumption later. Define  $F$  as the survival distribution function  $F(t) = 1 - \int_{\tau < t} d\tau f(\tau) = \int_{\tau > t} d\tau f(\tau)$  (these are equivalent because the full integral of the distribution  $f$  is one).

Individuals are characterized by their infection time  $\tau$  and recovery time  $\rho$ . The size of each such class is  $w(\tau, \rho) = i(\tau)f(\rho - \tau)$ . The expected case reproductive number (which we will use to calculate the mean and the excess variance) is  $C(\tau, \rho) = B \int_{\tau < t < \rho} x(t) dt$ .

Define the raw moments of  $C$  as  $C_k = \iint_{\tau < \rho} d\tau d\rho w(\tau, \rho) (C(\tau, \rho))^k$ . We expect  $C_0 = Z$ , where the final size  $Z = \int dt i(t)$ . From there, we will calculate  $\mu_C = C_1/C_0$ , and the squared CV  $\kappa_c = C_0 C_2 / C_1^2 - 1$ . We expect  $\mu_C = 1$ . Thus the conjecture is equivalent to  $\kappa_c = 1$ .

We have:

$$C_0 = \iint_{\tau < \rho} d\tau d\rho w(\tau, \rho) \quad (1)$$

$$= \int d\tau i(\tau) \int_{\rho > \tau} d\rho f(\rho - \tau) \quad (2)$$

$$= \int d\tau i(\tau), \quad (3)$$

$$= Z, \quad (4)$$

as expected.

Next:

$$C_1 = \iint_{\tau < \rho} d\tau d\rho w(\tau, \rho) C(\tau, \rho) \quad (5)$$

$$= B \iiint_{\tau < t < \rho} d\tau dt d\rho i(\tau) f(\rho - \tau) x(t) \quad (6)$$

$$= B \iint_{\tau < t} d\tau dt i(\tau) x(t) \int_{\rho > t} d\rho f(\rho - \tau) \quad (7)$$

$$= B \iint_{\tau < t} d\tau dt i(\tau) x(t) F(t - \tau) \quad (8)$$

$$= B \int dt x(t) \int_{\tau < t} d\tau i(\tau) F(t - \tau). \quad (9)$$

Note that the inner integral in (9) counts the number of individuals who entered the infectious class before time  $t$  and remained until then –  $y(t)$ . Thus:

$$C_1 = B \int dt x(t) y(t) \quad (10)$$

$$= \int dt i(t) \quad (11)$$

$$= Z, \quad (12)$$

and  $\mu_C = 1$ , as expected.

In an attempt to pull similar tricks, we expand  $C^2$  in our expression for  $C_2$  as an integral over a square, which we write as twice the integral over one of the two symmetric triangles:

$$C_2 = \iint_{\tau < \rho} d\tau d\rho w(\tau, \rho) (C(\tau, \rho))^2 \quad (13)$$

$$= 2B^2 \iiint \int_{\tau < s < t < \rho} d\tau ds dt d\rho i(\tau) f(\rho - \tau) x(s) x(t) \quad (14)$$

$$= 2B^2 \iiint \int_{\tau < s < t} d\tau ds dt i(\tau) x(s) x(t) \int_{\rho > t} d\rho f(\rho - \tau) \quad (15)$$

$$= 2B^2 \iiint \int_{\tau < s < t} d\tau ds dt i(\tau) x(s) x(t) F(t - \tau) \quad (16)$$

$$= 2B^2 \iint \int_{s < t} ds dt x(s) x(t) \int_{\tau < s} d\tau i(\tau) F(t - \tau) \quad (17)$$

Now the inner integral counts the number of individuals who entered the infectious class before one time ( $s$ ) and remained until another ( $t$ ), which is much less pretty. There is no obvious way to move forward from here without using the Markovian property of our chosen infectious distribution, i.e.,  $F(a + b) = F(a)F(b)$ :

$$C_2 = 2B^2 \iint \int_{s < t} ds dt x(s) x(t) \int_{\tau < s} d\tau i(\tau) F(t - s) F(s - \tau) \quad (18)$$

$$= 2B^2 \iint \int_{s < t} ds dt x(s) x(t) F(t - s) \int_{\tau < s} d\tau i(\tau) F(s - \tau) \quad (19)$$

$$= 2B^2 \iint \int_{s < t} ds dt x(s) x(t) F(t - s) y(s) \quad (20)$$

$$= 2B \iint \int_{s < t} ds dt x(t) F(t - s) i(s) \quad (21)$$

$$= 2B \int dt x(t) \int_{s < t} ds F(t - s) i(s) \quad (22)$$

$$= 2B \int dt x(t) y(t) \quad (23)$$

$$= 2 \int dt i(t) \quad (24)$$

$$= 2Z. \quad (25)$$

This gives  $\kappa_C = 2 - 1 = 1$ , as desired.

## Comments

We calculate  $y$  from  $i$  and  $f$ , and close the loop by calculating  $i$  using  $x$  and  $y$ . But we never calculate the population susceptibility  $x$ . This is a tiny bit troubling, since it seems like there must be some assumptions about  $x$  necessary for our result. What it seems to mean is that the variance conservation conjecture is true more broadly than conjectured – population susceptibility can change, or susceptibles can be replenished, and the excess variation in  $R_c$  is still 1 as long as the epidemic is bounded. Further, a limit argument suggests that it should also be true on average for a persistent periodic system.

It would be interesting to think about if there are any ways forward from (17) for other infectious-time distributions.