Estimating the ratio between the contagiousness of two viral strains

G. Della Croce[1], G. Mastrantonio[1], F. Cerutti[2], V. Ghisetti[2], M.Gasparini[1], E. Bibbona[1],

[1]DISMA, Politecnico di Torino, Italy, [2]Amedeo di Savoia Hospital, ASL Città di Torino, Italy

Abstract

- We propose a new method for the estimation of the contagiousness ratio between a new emerging variant and the one currently dominating within a population.
- Our method takes in input incidence data and epidemic curves. It is based on a discrete-time SIR with two strains in its deterministic and stochastic version.
- The method is applied to the ISS quick surveys data on virus SARS-CoV-2, for the Piedmont Italian region from December '21 to January '22 (Omicron and Delta variants)
- The estimated contagiousness ratio is $[3.7, 5.48]$ with the deterministic model and $[3.21, 3.98]$ with the stochastic one.

The Data

Let us consider a given situation in which there is a disease with two strains, that compete for the infection of the population. Strain 1 (Delta, δ) is currently dominating, while strain 2 (Omicron, o) starts emerging at the beginning of the analysis. Every day t, the available data are:

- $\mathbf{I_t^{tot}}$ **Total active cases**, $I_t^{tot} = I_t^1 + I_t^2$;
- $\mathbf{Y_t^{tot}}$ **Total new cases**, $Y_t^{tot} = Y_t^1 + Y_t^2$;
- $\mathbf{n_t}$ random samples from Y_t^{tot} sent to sequencing facilities; $\mathbf{z_t}$ of them turn out to belong to strain 2

$\longrightarrow z_t/n_t$ provides an estimate of strain 2 **relative incidence**, and it is natural to state that:

$$z_t \sim \text{HyperGeom}\left(Y_t^{tot}, Y_t^2, n_t\right)$$

The Model

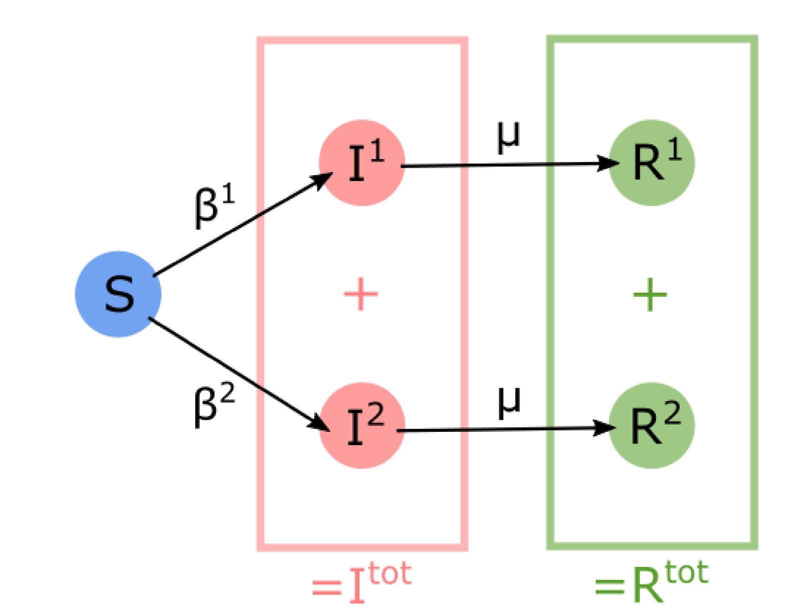

Figure 1: Two-strained discrete-time SIR model

Since our data are daily observations, discrete-time models will be adopted. In particular, one deterministic and one stochastic, both based on the compartmentalisation depicted in Figure 1.

Goal of our analysis is the **estimation** of parameter **k**, defined as:

$$k = \frac{\beta^2}{\beta^1} = \frac{\mathcal{R}_0^2}{\mathcal{R}_0^1}$$

By defining the new daily recoveries as $\rho_t^{tot} = R_t^{tot} - R_{t-1}^{tot}$ (split as $\rho_t^{tot} = \rho_t^1 + \rho_t^2$), we can easily state that, $\forall t$:

$$\mathbf{I_t^2 = I_{t-1}^2 + Y_t^2} - \rho_t^2 \tag{1}$$

This provides a recursive relation for the I_t^o terms $\longrightarrow \mathbf{I_0^2}$ additional parameter of our model.

Deterministic Approach

The terms Y_t^2, ρ_t^2 (hence I_t^2), $\forall t$, are assumed to follow a deterministic law given by:

$$\mathbf{Y_t^2} = \frac{\mathbf{kI_{t-1}^2}}{\mathbf{kI_{t-1}^2 + I_{t-1}^1}}\mathbf{Y_t^{tot}}, \qquad \rho_t^2 = \frac{\mathbf{I_{t-1}^2}}{\mathbf{I_{t-1}^2 + I_{t-1}^1}}\rho_t^{tot}$$

- Each term I_t^2 can be computed recursively (Eq. (1)), and it is ultimately a function of k and I_0^2.
- The log-likelihood of the z_t's can be explicitly computed and numerically maximised.

Par.	Est.	As. 95% C.I.	P.B. 95% C.I.
k	4.4	$[3.7, 5.48]$	$[3.76, 5.57]$
I_0^o	21.3	$[5.75, 64.5]$	$[4.92, 56.3]$

Table 1: ML estimates and 95% C.I. of the parameters

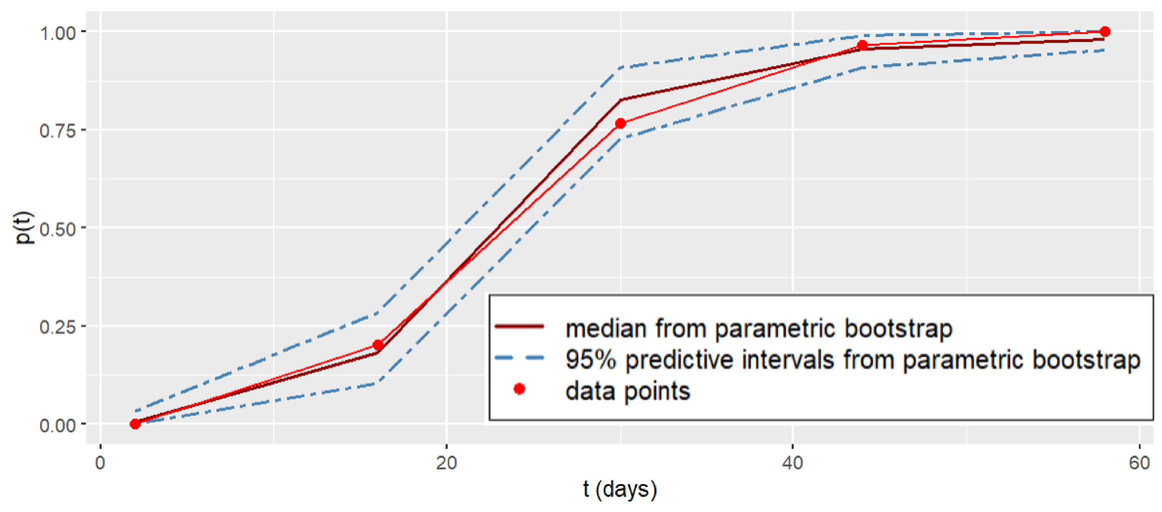

Figure 2: Omicron relative incidence curve $p(t)$

Stochastic Approach

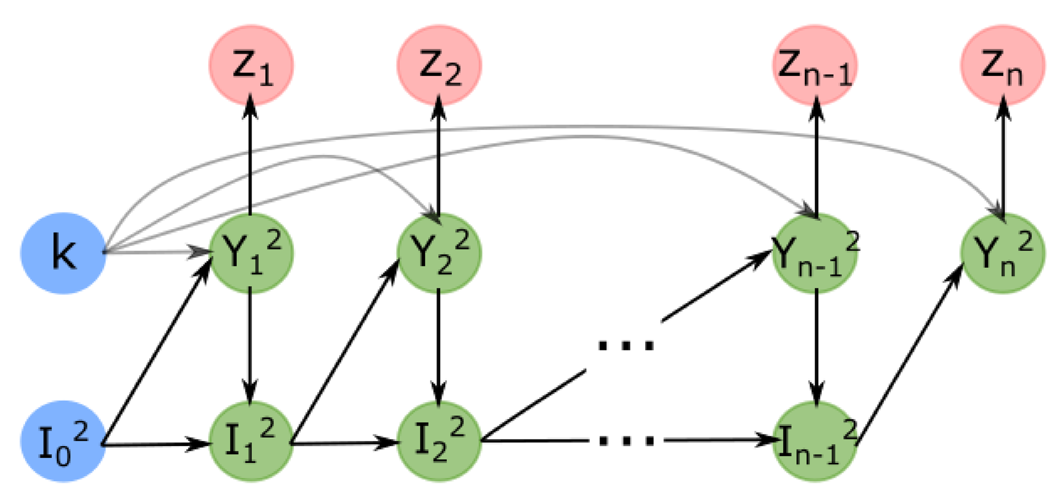

Figure 3: Stochastic model

The stochastic model results in a **Hidden Markov Model**, in which the hidden process is represented by the couple (Y_t^2, I_t^2), while the observable process is constituted by the z_t terms:

$$z_t | Y_t^2 \sim \text{HyperGeom}\left(Y_t^{tot}, Y_t^2, n_t\right)$$

$$Y_t^2 | (k, I_{t-1}^2) \sim \text{Bin}\left(Y_t^{tot}, \frac{kI_{t-1}^2}{I_{t-1}^{tot} + (k-1)I_{t-1}^2}\right)$$

$$I_t^2 | (I_{t-1}^2, Y_t^2) \sim \text{ShiftedHyperGeom}\left(I_{t-1}^2 + Y_t^2, I_{t-1}^{tot}, I_{t-1}^2, \rho_t^{tot}\right)$$

The latter because of Eq. (1) and the fact that

$$\rho_t^2 | I_{t-1}^2 \sim \text{HyperGeom}\left(I_{t-1}^{tot}, I_{t-1}^2, \rho_t^{tot}\right)$$

Our inferential procedure is based on a **MCMC algorithm**, where the prior distributions assigned to the parameters are $k \sim \text{Unif}(0.05, 20)$ and $I_0^2 \sim$ DUnif(1,1000).
When proposing new values in our MCMC, some admissibility constraints need to be fulfilled.

Constraints to be imposed

- $I_{t-1}^* \geq \rho_t^*$, for $* = 1, 2$
- $\rho_t^2 \geq 0$ and $\rho_t^2 \leq \rho_t^{tot}$
- Bound constraints for each variable

The traceplots obtained for the parameters are depicted in Figure 4 and 5. The posterior medians and 95% credible intervals are $\mathbf{3.34}$ and $[\mathbf{3.1}, \mathbf{3.6}]$ for $\mathbf{k}$, $\mathbf{126}$ and $[\mathbf{88}, \mathbf{169}]$ for $\mathbf{I_0^2}$.
The estimated Omicron relative incidence curve $p(t)$ is depicted in Figure 6.

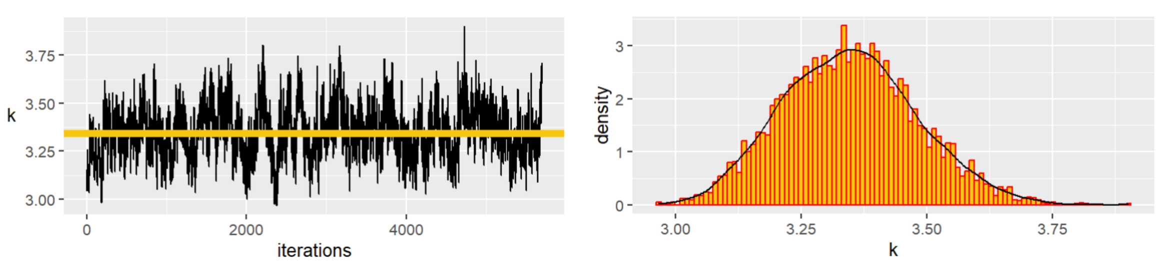

Figure 4: Traceplot and posterior density of parameter k

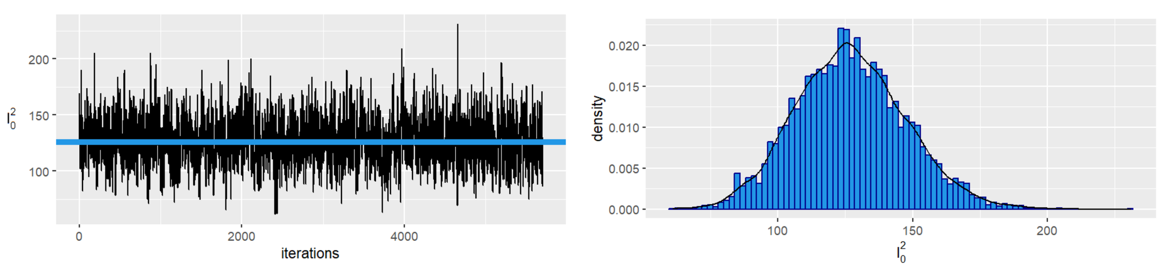

Figure 5: Traceplot and posterior density of parameter I_0^2

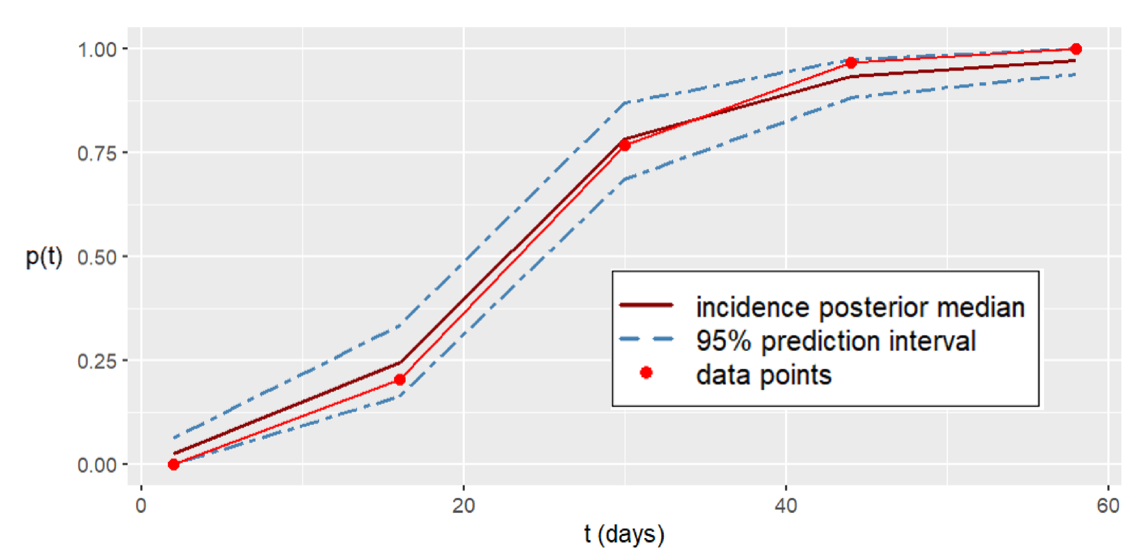

Figure 6: Omicron posterior incidence curve

In **conclusion**, what we observe:

- Both methods fit well the data, Omicron appears to be **3** or **4** times **more contagious** than Delta
- The stochastic method is certainly more robust, but the deterministic one performs equally well and its computational time is drastically smaller

Contact Information

- Email: giulia.dellacroce@polito.it
- Email: enrico.bibbona@polito.it