Some remarks on the uncertainty analysis of R_0 in the SIR model

Luiz Max F. de Carvalho[1], Daniel A. M. Villela[1], Flavio Coelho[2], and Leonardo S. Bastos[1]

[1]*Program for Scientific Computing (PROCC), Oswaldo Cruz Foundation, Brazil,* lmax.procc@gmail.com
[2]*School of Applied Mathematics, Getulio Vargas Foundation (FGV), Brazil,* fccoelho@fgv.br

Keywords: Basic reproductive number; uncertainty; logarithmic pooling; Gamma ratio distribution.

The basic reproductive number, R_0, is a key quantity in epidemic modelling and defines the threshold of disease-free equilibrium for many deterministic models of disease. It is usually a quantity of interest when constructing mathematical models of infectious diseases and developing and evaluating mitigation strategies. While most well studied mathematical models are deterministic in nature, acknowledging uncertainty on parameter values is important because predictions can be sensitive to changes in parameter values. In addition, most interesting models are nonlinear which makes correlations between parameters difficult to study analytically.

A seminal paper by Poole & Raftery [2] discusses the issue of propagating uncertainty through a deterministic model. The authors use an operator called the logarithmic pooling (LP) operator to combine prior distributions induced by the structure of the model. The operator takes a set of probability distributions $F(\boldsymbol{\theta})$ and a set of weights $\boldsymbol{\alpha}$ and combines them into a single probability distribution $\pi(\boldsymbol{\theta})$.

Here we are concerned with the following setting: suppose K experts express their opinions on a parameter $\boldsymbol{\theta} \in \Theta \subseteq \mathbb{R}^p$ by means of a set of (proper) probability distributions $F(\boldsymbol{\theta}) = \{f_1(\boldsymbol{\theta}), f_2(\boldsymbol{\theta}), \ldots, f_K(\boldsymbol{\theta})\}$. Supposed further that we are also interested in a quantity $\mathbf{y} \in \mathcal{Y} \subseteq \mathbb{R}^q$ that relates to $\boldsymbol{\theta}$ through a deterministic model $M(\boldsymbol{\theta}) = \mathbf{y}$. We would like to obtain one distribution on $\mathbf{y}$ to represent our uncertainty about this quantity *induced* by the uncertainty we have on $\boldsymbol{\theta}$. This, however, begs the question of in which order the pooling and propagation (or inducing) operations should be performed. One could either (a) compute the induced distributions for each expert and then combine these using LP or (b) combine the distributions first and then apply $M(\cdot)$ to obtain a distribution on $\mathbf{y}$.

While it is relatively straightforward to show that procedures (a) and (b) will yield the same distribution if $M(\cdot)$ is invertible, the question remains open for the case of non-invertible models, which correspond to the vast majority of epidemic models.

In this paper we study the uncertainty analysis for R_0 in the constant population Susceptible-infectious-Removed (SIR) model. We show that if uncertainty about the transmission rate β and recovery rate γ can be expressed using Gamma distributions, the induced distribution on R_0 has a closed-form solution which is a variant of the generalised Beta prime distribution. This distribution, henceforth called Gamma ratio (GR) distribution has been derived in a more general setting by Coelho & Mexia [1] and is sub-exponential, being log-concave only on an interval that depends on its four parameters.

We explore the case when one has a set of distributions on both β and γ and desires to obtain a distribution on R_0. This model is non-invertible, since, for a population of size N, $R_0 = \beta N/\gamma$. Thus, the distributions on R_0 obtained using procedures (a) and (b) need not coincide. We show that they indeed yield different distributions. While procedure (a) induces a reasonably well-behaved GR distribution, procedure (b) leads to a distribution with heavier tails and for which

closed-form solutions are available only under some restrictions on the original distributions on the parameters. A in-depth analysis of tail behaviour and some general guarantees are provided for both distributions, focusing on the comparison the two. We consider this comparison to be key in understanding how to best deal with the order of pooling and inducing operations in non-invertible models.

Finally, an application to the uncertainty analysis of R_0 estimates for Ebola in West Africa is considered. We combine distributions for the epidemic parameters estimated from data from different countries in order to obtain a pooled distribution on R_0 that reflects uncertainty about this quantity for Ebola in general.

Acknowledgements

[PUT YOURS IN, BOYZ]

Bibliography

[1] Carlos A Coelho and Joao T Mexia. On the distribution of the product and ratio of independent generalized gamma-ratio random variables. *Sankhyā: The Indian Journal of Statistics*, pages 221–255, 2007.

[2] David Poole and Adrian E Raftery. Inference for deterministic simulation models: the bayesian melding approach. *Journal of the American Statistical Association*, 95(452):1244–1255, 2000.