Uncertainty analysis of key epidemiological quantities

Luiz Max F. de Carvalho
School of Applied Mathematics, Getúlio Vargas Foundation.

June 4, 2020

Abstract

In these notes we give a few key epidemiological quantities which can be computed in closed-form for a few important ODE-based epidemic models, such as SIR and SEIR. We also show one can use these to carry out prior predictive checks. Key-words: Bayesian inference; mathematical epidemiology; prior predictive checks; final epidemic size; peak size.

SIR model

The SIR model we are interested in this note is given by the system:

$$\begin{aligned}
\frac{dS}{dt} &= -\beta S I, \\
\frac{dI}{dt} &= \beta S I - \gamma I, \\
\frac{dR}{dt} &= \gamma I,
\end{aligned}$$

where $S(t) + I(t) + R(t) = 1 \, \forall t$, β is the transmission (infection) rate and γ is the recovery rate. The basic reproductive number is

$$R_0 = \frac{\beta}{\gamma}. \tag{1}$$

Final epidemic size

Now, we would like to know what the final epidemic size would be. This is $\lim_{t\to\infty} R(t) := R(\infty)$, which leads to $S(\infty) = N - R(\infty)$. To compute $S(\infty)$, first write

$$\frac{dI}{dS} = -1 + \frac{N}{R_0 S}, \tag{2}$$

which gives

$$I(t) = -S(t) + \frac{N}{R_0} \log S(t) + C, \tag{3}$$

where C can be determined from the initial conditions (Miller, 2012) and thus:

$$S(\infty) = I(0) + S(0) + \frac{N}{R_0} \log\left(\frac{S(\infty)}{S(0)}\right) \tag{4}$$

$$R(\infty) = N - S(\infty) \tag{5}$$

Letting $a = R_0/N$ and $b = N - \log S(0)$, we arrive at the following expression for $S(\infty)$:

$$S(\infty) = -\frac{1}{a} W\left(-ae^{-b}\right), \tag{6}$$

where W is the Lambert product log function.

Peak size

To find the maximum value of $I(t)$, i.e., the peak size, $I_{\max}$, we need to solve $\frac{dI}{dt} = 0$:

$$I(\beta S - \gamma) = 0 \implies \bar{S} = \frac{1}{R_0}. \tag{7}$$

Plugging $\bar{S}$ into equation (3) gives

$$I_{\max} = S(0) + I(0) - \frac{1}{R_0}\log S(0) - \frac{1}{R_0} + \frac{1}{R_0}\log\frac{1}{R_0}, \tag{8}$$

$$= S(0) + I(0) - \frac{1}{R_0}\left[1 + \ln(S(0)R_0)\right]. \tag{9}$$

Making the approximation $S(0) + I(0) \approx S(0) \approx N$, we get

$$I_{\max} = N - \frac{\log R_0 + 1}{R_0}, \tag{10}$$

for the number of individuals that are infectious at the peak.

SEIR model

For the SEIR model the system is

$$\begin{aligned}
\frac{dS}{dt} &= -\beta S(I + \epsilon E),\\
\frac{dE}{dt} &= \beta S(I + \epsilon E) - \kappa E,\\
\frac{dI}{dt} &= \kappa E - \alpha I,\\
\frac{dR}{dt} &= \alpha I,
\end{aligned}$$

with $S(0) = S_0$, $E(0) = E_0$, $I(0) = R(0) = 0$ and $S(t) + E(t) + I(t) + R(t) = N$.

Under this model $S(\infty)$ can be calculated using the expression in (6) by writing making

$$b = R_0 - \log S(0) - \frac{\epsilon\beta}{N}(N - S(0)), \tag{11}$$

$$R_0 = \frac{\beta N}{\gamma} + \frac{\beta N \epsilon}{\kappa} = \beta N\left(\frac{\kappa + \gamma\epsilon}{\gamma\kappa}\right). \tag{12}$$

The peak size for the SEIR model can be computed if we consider infectious and exposed individuals jointly (Feng, 2007) by writing $Y(t) = E(t) + I(t)$. We can then proceed analogously to the case of the SIR model and arrive at

$$Y_{\max} = S(0) + Y(0) - \frac{1}{R_0}\left[1 + \ln(S(0)R_0)\right]. \tag{13}$$

Uncertainty analysis

In the context of a (Bayesian) statistical analysis of deterministic epidemic models, some or all of the parameters, θ, are unknown and we want to estimate them from data. Assuming we can represent uncertainty about the parameters as a joint probability distribution $\pi(\theta)$, we can then ask what the distribution on the quantity of interest (q.o.i.) such as R_0, $R(\infty)$ or $I_{\max}$ induced by $\pi(\theta)$. Let $\varphi(\theta)$ be the q.o.i. The predictive check procedure can be summarised as follows. For a number $M \in \mathbb{N}$ of iterations, do for $i = 1, \ldots, M$:

1. Draw $\theta^{(i)} \sim \pi(\cdot)$;
2. Compute $\varphi^{(i)} = \varphi(\theta^{(i)})$.

The distribution $\pi(\varphi(\theta))|J|$ can then be approximated from the samples $\boldsymbol{\varphi}$.

As a first illustration of this procedure, we place a prior directly on R_0, i.e., we interpret the basic reproductive number as a parameter rather than a derived quantity from the parameters. We take COVID-19 as a basis and elicit Gamma and log-normal distributions with mean 2.5 and standard deviation 0.65 (shown in Figure 1b). The resulting distribution on $I_{\max}$ is shown in Figure 1c.

As a second example, we consider a similar setting, by choose to place priors on β and γ. As with the previous analysis, we elicit Gamma and log-normal priors on the parameters. We follow Grinsztajn et al[1] and set $E[\beta] = 2$ and $\text{Var}(\beta) = 1$ and $E[\gamma] = 0.4$ and $\text{Var}(\gamma) = 0.25$. We show the resulting densities for β and γ in Figure 2a and b, respectively. The resulting distribution on I_{max} is shown in Figure 2c, and indicates that these priors, which seem reasonable, lead to a very wide prior on I_{max} – also on R_0, not shown here.

Acknowledgements

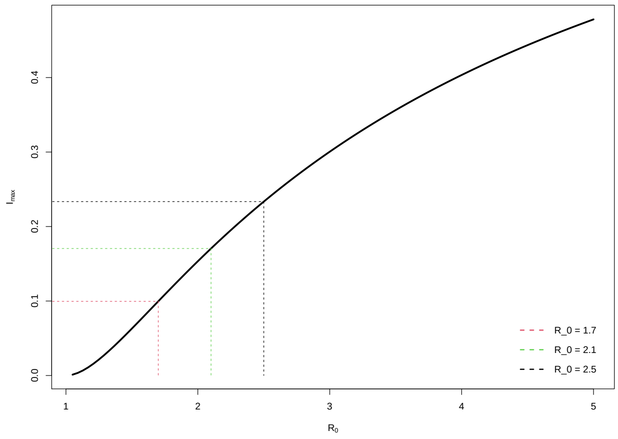
LMC would like to thank Charles Margossian (Columbia) for stimulating discussions that prompted him to write these notes in hopes these analyses become common place in statistical applications of epidemic models.

References

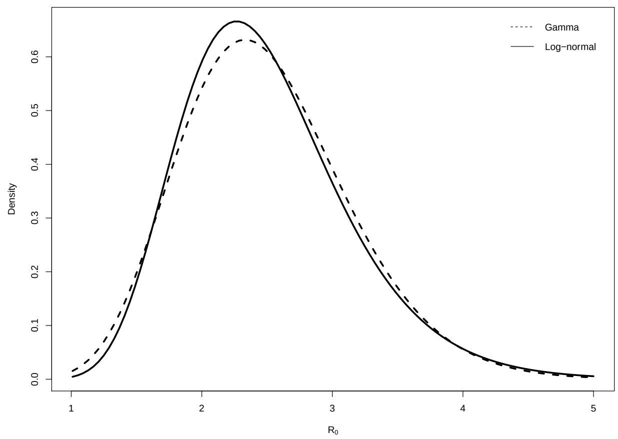
Feng, Z. (2007). Final and peak epidemic sizes for seir models with quarantine and isolation. *Mathematical Biosciences & Engineering*, 4(4):675.

Miller, J. C. (2012). A note on the derivation of epidemic final sizes. *Bulletin of mathematical biology*, 74(9):2125–2141.

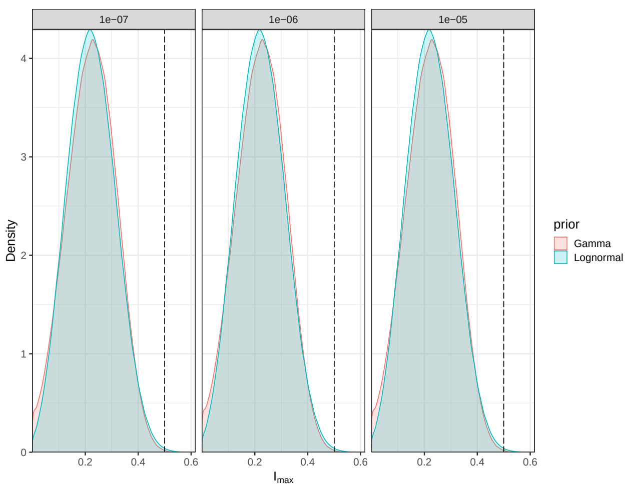
[1] Case study available at https://mc-stan.org/users/documentation/case-studies/boarding_school_case_study.html.



(a) $I_{\max}$ as a function of R_0.

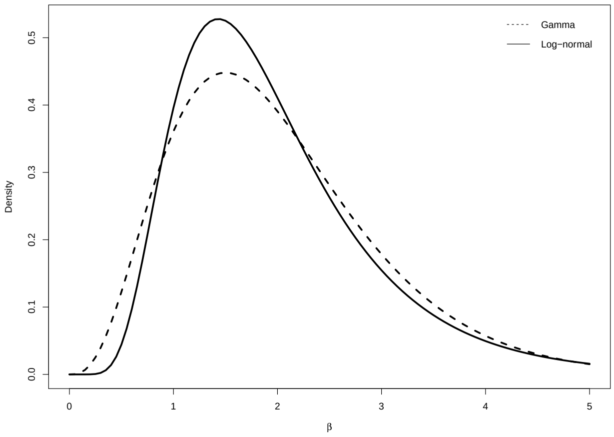


(b) Priors for R_0.

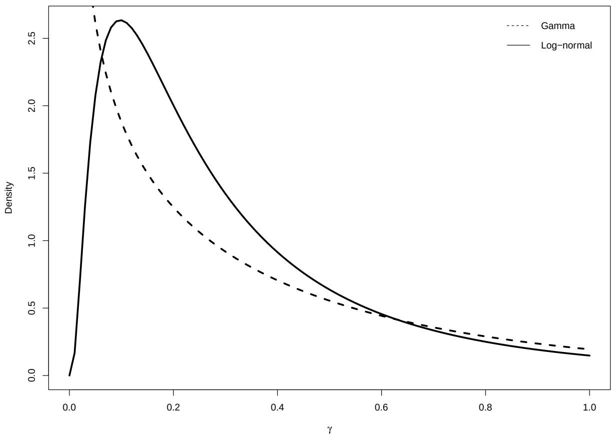


(c) Induced distributions for $I_{\max}$.

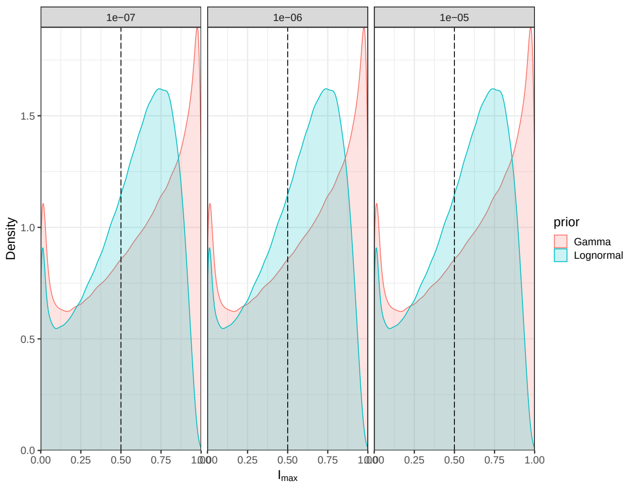
Figure 1: **Uncertainty analysis of the peak size for the SIR model, priors on** R_0. In panel A we show the peak size $I_{\max}$ as function of the reproductive number R_0. In panel B we show two priors for R_0 and panel C shows the induced distributions on $I_{\max}$ for three values of $I(0)$: 10^{-7}, 10^{-6}, 10^{-5}. Vertical dashed line shows $I_{\max} = 1/2$. In this example, $N = 1$, i.e. the system is normalised.



(a) $\pi(\beta)$



(b) $\pi(\gamma)$.



(c) Induced distributions for $I_{\max}$.

Figure 2: **Uncertainty analysis of the peak size for the SIR model, priors on β and γ.** Panels A and B show prior distributions on β and γ, respectively. In panel C we show the induced distributions on $I_{\max}$ for three values of $I(0)$: 10^{-7}, 10^{-6}, 10^{-5}. Vertical dashed line shows $I_{\max} = 1/2$. In this example, $N = 1$, i.e. the system is normalised.