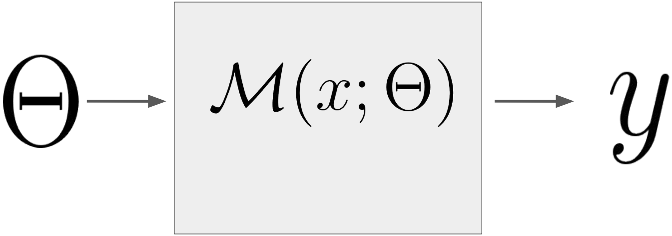
Bayesian inference for deterministic epidemic models

Luiz Max F. de Carvalho [lmax.fgv@gmail.com]

School of Applied Mathematics, Getulio Vargas Foundation (FGV), Rio de Janeiro.
Presented at "XLII Congresso Nacional de Matemática Aplicada e Computacional" (CNMAC)

Motivation: Bayesian inference of deterministic models

- Pick a prior measure $\pi(\theta)$;
- Study its **pushforward** through $\mathcal{M}(x;)$;



- Compute $p(\Theta \mid \boldsymbol{y})$.

Overview

Authors

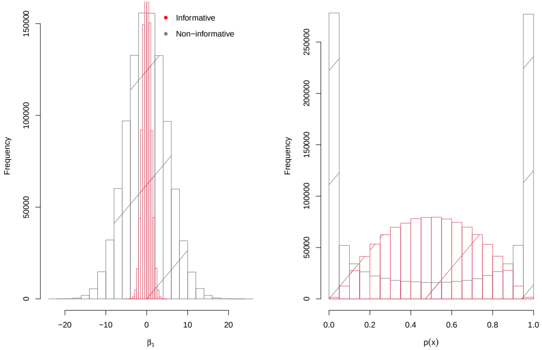
This is joint work with Márcio Bastos, Daniel Villela, Leo Bastos and Flávio Coelho.

- Complex models need careful prior elicitation;
  Induced prior
  - On the basic reproductive number;
  - On the peak height and final epidemic size.
- A few mathematical tricks we can use to make models more robust or easier to fit and/or analyse;
  - Equations in log space;
  - Approximate parametrisation in terms of R_0.

There is no such thing as an 'uniformative' prior!

Consider

$$p_\beta(x) = \frac{1}{1 + \exp(-\beta_0 + \beta_1 x)}$$



For more details, see, e.g. Seaman III et al. (2012).

The basic reproductive number, R_0

Consider a Susceptible-Infectious-Removed (SIR) model:

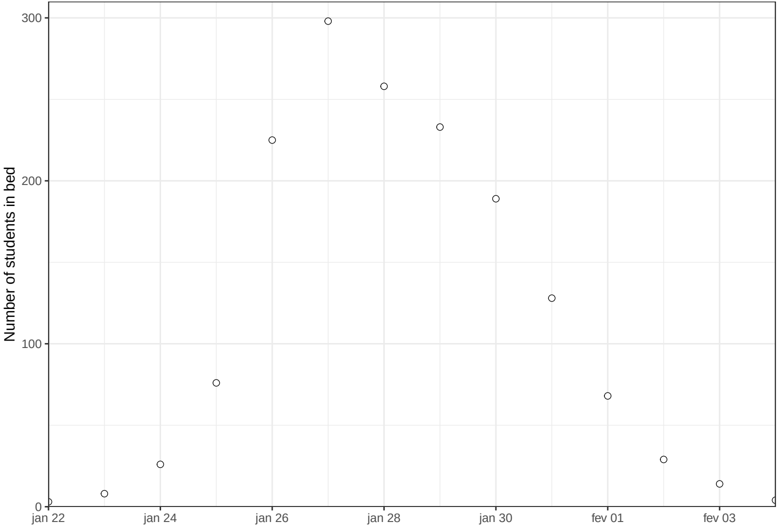
$$\begin{aligned}
\frac{dS}{dt} &= -\beta \frac{SI}{N}, \\
\frac{dI}{dt} &= \beta \frac{SI}{N} - \gamma I, \\
\frac{dR}{dt} &= \gamma I,
\end{aligned}$$

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

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

Motivating application: Influenza in a boarding school

In 1978, 512 out of 763 lads got came down with the flu.



Gamma priors on the rates

Suppose the *a priori* uncertainty about parameters can be represented by Gamma (shape/scale) distributions: $\gamma \sim \text{Gamma}(k_\gamma, \theta_\gamma)$ and $\beta \sim \text{Gamma}(k_\beta, \theta_\beta)$. The pdf of R_0 is given by (Clancy et al. (2008)):

$$f_{R_0}(r \mid k_\beta, \theta_\beta, k_\gamma, \theta_\gamma, N) = \frac{(N\theta_\beta\theta_\gamma)^{k1+k2}}{\mathcal{B}(k_\beta, k_\gamma)(N\theta_\beta)^{k_\beta}\theta_\gamma^{k_\gamma}} r^{k_\beta - 1}(\theta_\gamma r + N\theta_\beta)^{-(k_\beta+k_\gamma)}, \quad (2)$$

where $\mathcal{B}(a, b) = \Gamma(a+b)/\Gamma(a)\Gamma(b)$ is the Beta function. The expectation of the Gamma ratio distribution is then

$$E[R_0] = \frac{N\theta_\beta}{\theta_\gamma}\frac{k_\beta}{(k_\gamma - 1)},$$

which is defined only for $k_\gamma > 1$. The variance can be computed as

$$\text{Var}(R_0) = \left(\frac{N\theta_\beta}{\theta_\gamma}\right)^2 \frac{(k_\beta + k_\gamma - 1)k_\beta}{(k_\gamma - 2)(k_\gamma - 1)^2},$$

and only exists for $k_\gamma > 2$. The mode is

$$\frac{N\theta_\beta}{\theta_\gamma}\frac{k_\beta - 1}{(k_\gamma + 1)}. \quad (3)$$

Log-normal priors on the rates

Now, take $\gamma \sim \text{Log-normal}(\mu_\gamma, \sigma_\gamma)$ and $\beta \sim \text{Log-normal}(\mu_\beta, \sigma_\beta)$. It is straightforward to show that the induced distribution on R_0 is a log-normal distribution with parameters $\mu_{R_0} = \ln N + \mu_\beta - \mu_\gamma$ and $\sigma_{R_0} = \sigma_\beta^2 + \sigma_\gamma^2$. Under the justification of employing a non-informative prior, researchers might be tempted to choose $\mu_\beta = \mu_\gamma = 0$ and $\sigma_\beta = \sigma_\gamma = 100$, say[1].

This apparently non-informative choice of hyperparameters leads to a prior on R_0 for which $E[R_0] = N + \exp(10^4)$ and $\Pr(R_0 > 100) = 0.49$, which are not reasonable. In general, under log-normal priors for the rates, we have

$$E[R_0] = \exp\left(\ln N + \mu_\beta - \mu_\gamma + \frac{\sigma_\beta^2 + \sigma_\gamma^2}{2}\right),$$

$$\operatorname{Var}(R_0) = \left[\exp\left(\sigma_\beta^2 + \sigma_\gamma^2\right) - 1\right] \exp\left(2\{\ln N + \mu_\beta - \mu_\gamma\} + \frac{\sigma_\beta^2 + \sigma_\gamma^2}{2}\right).$$

[1] See e.g. Ho et al. (2018), section 5.1.

Half-normal priors on the rates

A further choice of priors for positive quantities is the half-normal (truncated at zero). Let

$$\begin{aligned} \beta &\sim \text{Normal}^{+}(\mu_\beta, \sigma_\beta), \\ \gamma &\sim \text{Normal}^{+}(\mu_\gamma, \sigma_\gamma), \end{aligned} \tag{4}$$

and $R_0 = \beta/\gamma$ i.e., taking $N = 1$ for simplicity. This gives

$$f_{R_0}(r) = \exp\left(-\frac{(\mu_\beta/r - \mu_\gamma)^2}{2(\sigma_\beta^2/r^2 + \sigma_\gamma^2)}\right) \frac{\sqrt{2}\Gamma\left(1, \frac{m(r)^2}{2v(r)}\right) v(r) + \left(2\sqrt{\pi} - \Gamma\left(\frac{1}{2}, \frac{m(r)^2}{2v(r)}\right)\right) m(r)\sqrt{v(r)}}{2\pi\sigma_\beta\sigma_\gamma\sqrt{2}[1 - F_\beta(0)][1 - F_\gamma(0)]},$$

with

$$\begin{aligned} m(r) &= \frac{\mu_\beta\sigma_\gamma^2 r + \mu_\gamma\sigma_\beta^2}{\sigma_\gamma^2 r^2 + \sigma_\beta^2}, \\ v(r) &= \frac{\sigma_\beta^2\sigma_\gamma^2}{\sigma_\gamma^2 r^2 + \sigma_\beta^2}. \end{aligned}$$

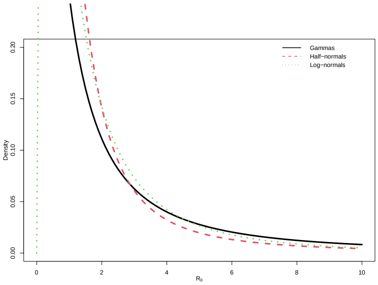
Nae moments!

$E\left[R_0^t\right] = \infty$ for all $t \geq 1$

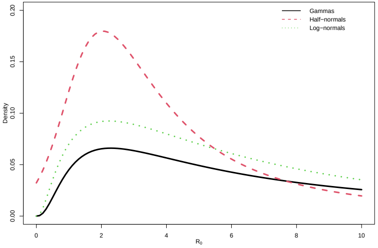
Induced prior on R_0

Prior on R_0

- 'Non-informative': means and variances equal to 1;
- 'Informative': $E[\beta] = 2$, $\text{Var}(\beta) = 1$, $E[\gamma] = 0.4$, $\text{Var}(\gamma) = 0.5^2$,

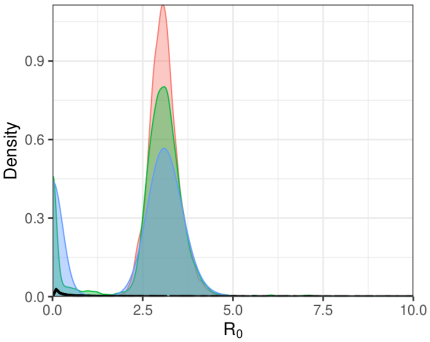


'Noninformative'

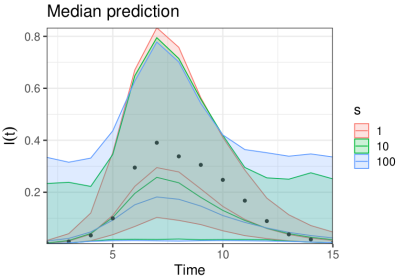


'Informative'

Shenanigans everywhere!



R_0



$I(t)$

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

$$\frac{dI}{dS} = -1 + \frac{N}{R_0 S}, \tag{5}$$

which gives

$$I(t) = -S(t) + \frac{N}{R_0} \log S(t) + C, \tag{6}$$

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

$$S(\infty) = I(0) + S(0) + \frac{N}{R_0} \log\left(\frac{S(\infty)}{S(0)}\right) \tag{7}$$

$$R(\infty) = N - S(\infty) \tag{8}$$

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

$$S(\infty) = -\frac{1}{a} W\left(-ae^{-b}\right), \tag{9}$$

where W is the Lambert product log function.

Peak height

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

$$I(\beta S - \gamma) = 0 \implies \bar{S} = \frac{1}{R_0}. \tag{10}$$

Plugging $\bar{S}$ into equation (6) gives

$$\begin{aligned} I_{\text{max}} &= S(0) + I(0) - \frac{1}{R_0} \log S(0) - \frac{1}{R_0} + \frac{1}{R_0} \log \frac{1}{R_0}, \qquad (11) \\ &= S(0) + I(0) - \frac{1}{R_0} \left[1 + \ln(S(0)R_0)\right]. \qquad (12) \end{aligned}$$

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

$$I_{\text{max}} = N - \frac{\log R_0 + 1}{R_0}, \tag{13}$$

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 (9) by writing

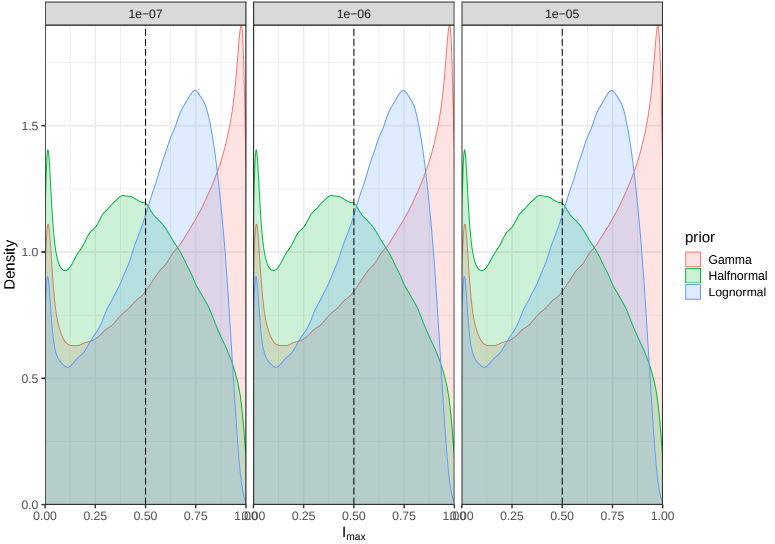
$$
b = R_0 - \log S(0) - \frac{\epsilon\beta}{N}(N - S(0)),
$$

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

Writing $Y(t) = E(t) + I(t)$ (Feng, 2007):

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

Example: I_{max} under different priors



Prior sensitivity analysis

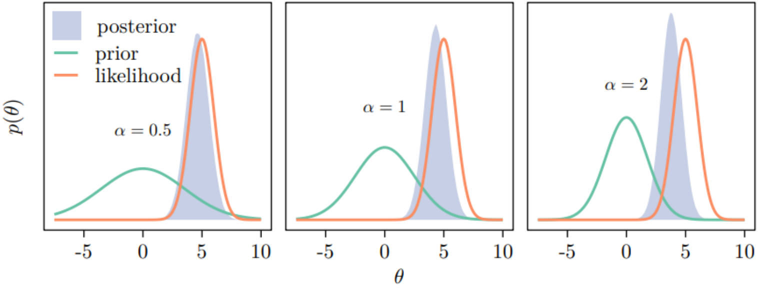
The posterior is

$$p(\theta \mid \boldsymbol{y}) \propto L(\boldsymbol{y} \mid \theta)\pi(\theta)$$

What happens if we look at [2]

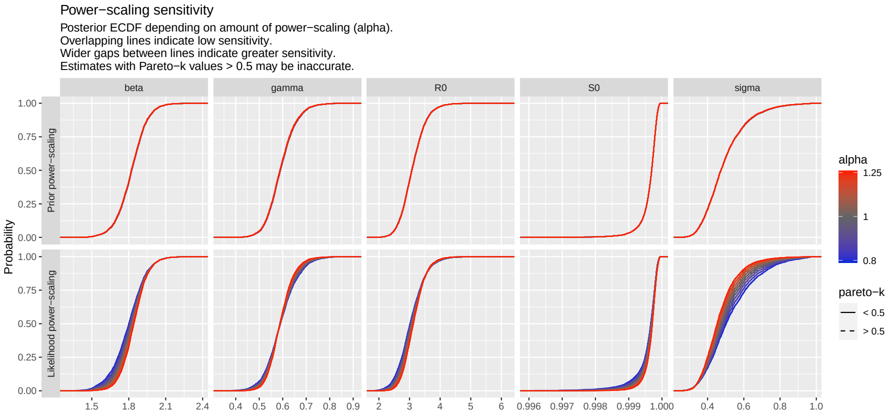
$$p_{\boldsymbol{\alpha}}(\theta \mid \boldsymbol{y}) \propto L(\boldsymbol{y} \mid \theta)^{\alpha_1}\pi(\theta)^{\alpha_2}$$

and let α_i vary on $[0, \infty)$.



[2] See Kallionen et al. 2023 for more details.

Prior sensitivity analysis: results



Detour I: reparametrising the ODEs

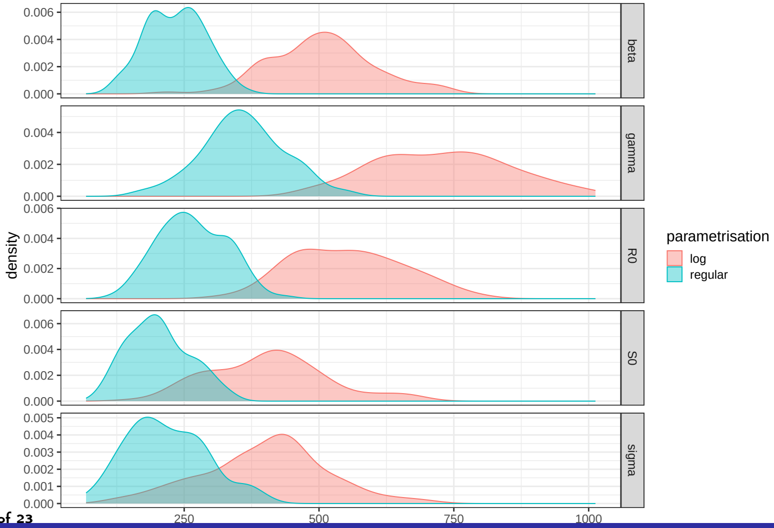
High school calculus gives

$$
\begin{aligned}
\frac{d \log S}{dt} &= -\beta I, \\
\frac{d \log I}{dt} &= \beta S - \gamma, \\
\frac{d \log R}{dt} &= \frac{\gamma I}{R},
\end{aligned}
$$

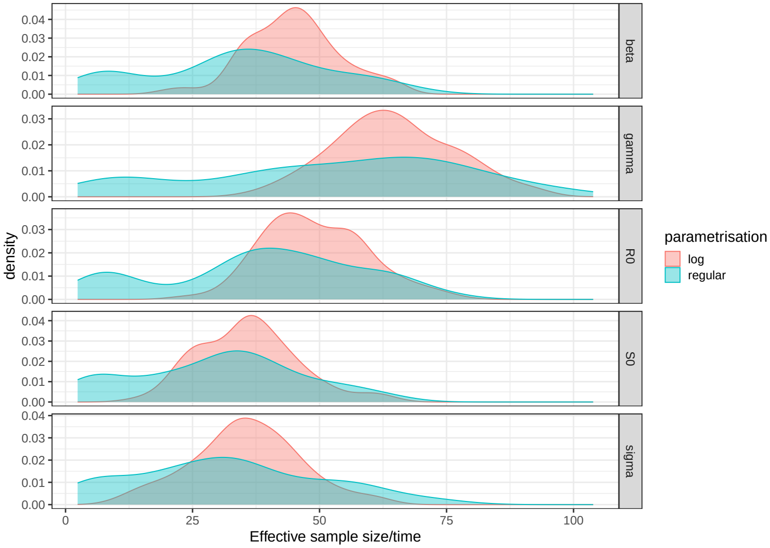
which is useful if you want to keep $I(t) > 0$ but keep a simple likelihood, i.e., use a log-normal likelihood. This is more numerically stable and plays nicer with the ODE solver.

It works!

Ran 100 chains under the 'regular' and 'log' parametrisations. Computed the effective sample size



Or does it?



0.04
0.03
0.02
0.01
0.00
beta
gamma
R0
S0
sigma
density
0
25
50
75
100
Effective sample size/time
parametrisation
log
regular

If we take the ratio:

$$\frac{dS}{dR} = \frac{-\beta SI}{\gamma I} = -\mathcal{R}_0 S \tag{14}$$

It can be integrated to $S(t) = S_0 e^{-\mathcal{R}_0 R}$. We can then substitute into the standard $\frac{dR}{dt}$ to get

$$\frac{dR}{dt} = \gamma \left(N - R - S_0 e^{-\mathcal{R}_0 R} \right) \tag{15}$$

An approximate solution can be obtained by assuming $R_0 R$ remains small for then $e^{-R_0 R} \approxeq 1 - (\mathcal{R}_0 R) + (\mathcal{R}_0 R)^2$ and (15) reduces to the first order quadratic ODE

$$\frac{dR}{dt} \approxeq \gamma \left(N - S_0 + [S_0 \mathcal{R}_0 - 1]R - (S_0 \mathcal{R}_0^2 / 2) R^2 \right) \tag{16}$$

which takes the standard solution

$$R(t) = \frac{1}{\mathcal{R}_0^2 S_0} \{ (S_0 \mathcal{R}_0) - 1 + \alpha \tanh[(\alpha \gamma t/2) - \phi] \} \tag{17}$$

where the amplitude $\alpha = \sqrt{[S_0 \mathcal{R}_0 - 1]^2 + 2 S_0 I_0 \mathcal{R}_0^2}$, and the phase $\phi = \tanh^{-1}\{(S_0 \mathcal{R}_0 - 1)/\alpha\}$.

Thank you very much for your attention!

- LMC would like to thank FGV EMAp for financial support and Charles Margossian (Flatiron), Américo Cunha (Princeton/UERJ), Erik Volz (Imperial College London) and Philip O'Neill (Nottingham) for stimulating discussions;

- If this piqued your insterested, keep an eye on https://github.com/maxbiostat/R0_uncertainty;

References

Clancy, D., O'Neill, P. D., et al. (2008). Bayesian estimation of the basic reproduction number in stochastic epidemic models. *Bayesian Analysis*, 3(4):737–757.

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

Ho, L. S. T., Crawford, F. W., Suchard, M. A., et al. (2018). Direct likelihood-based inference for discretely observed stochastic compartmental models of infectious disease. *The Annals of Applied Statistics*, 12(3):1993–2021.

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

Seaman III, J. W., Seaman Jr, J. W., and Stamey, J. D. (2012). Hidden dangers of specifying noninformative priors. *The American Statistician*, 66(2):77–84.