Estimation and Dynamics in Household Disease Models

Casper Woroszylo and Grzegorz Rempala

Mathematical Bioscience Institute, the Ohio State University

Overview

Using cross-sectional data, we fit a dynamical model of infection. More specifically:

1. Determine household structure for a disease from 64 households within a 1 month period.
2. Use this likelihood and MLE estimates for parameters to generate data for households.
3. Derive functional equations for imputing susceptible and infective compartments.
4. Use generated averages of households to fit a dynamical model in steady state using a MCMC procedure. This provides point estimates as well as 95% credible intervals for the parameters and imputed states.

Household Structure

At the household level we have the number of Infants, Juveniles and Adults:

1. Healthy I, J, A
2. Diseased D_I, D_J, D_A
3. Contamination status (V)

I	J	A	DI	DJ	DA	V
0	0	2	0	0	0	0
0	1	2	0	0	1	0
0	1	7	0	0	0	0
0	4	6	0	0	0	1
1	3	5	0	1	3	0
0	5	3	0	0	0	1

Consider the following likelihood for a given household $(x_I, x_J, x_A, x_{DI}, x_{DJ}, x_{DA}, x_V)$:

$$\prod_{k\in\{I,J,A\}} P[D_k = x_k | N_k = x_k + x_{Dk}, V = x_V] \times P[N_k = x_k + x_{Dk}] P[V = x_V]$$

where

$$D_k | (N_k, V) \sim \text{Binomial}(N_k, p_k(V))$$

$$N_k \sim \text{Poisson}(\lambda_k).$$

We use this to generate **average** households.

Dynamical System

We observe no diseased infants in the data. Hence we look at Juveniles (1) and Adults (2).

$$\frac{d}{dt}S_1 = -\beta_{12}S_1I_2 - \beta_{11}S_1I_1 - \phi_1 S_1 V - \alpha_1 S_1 + \delta_1 D_1 + (\gamma_1 - \nu_1)I_1$$

$$\frac{d}{dt}S_2 = -\beta_{21}S_2I_1 - \beta_{22}S_2I_2 - \phi_2 S_2 V - \alpha_2 S_2 + \delta_2 D_2 + (\gamma_2 - \nu_2)I_2$$

$$\frac{d}{dt}I_1 = \beta_{21}S_2I_1 + \beta_{11}S_1I_1 + \phi_1 S_1 V - \gamma_1 I_1$$

$$\frac{d}{dt}I_2 = \beta_{12}S_1I_2 + \beta_{22}S_2I_2 + \phi_2 S_2 V - \gamma_2 I_2$$

$$\frac{d}{dt}D_1 = \alpha_1 S_1 + \nu_1 I_1 - \delta_1 D_1$$

$$\frac{d}{dt}D_2 = \alpha_2 S_2 + \nu_2 I_2 - \delta_2 D_2$$

The system represents the **average** behavior of the system over time.

Since the epidemic is not taking off, we can assume that the data come from steady state period, meaning the **average state of a household is not changing with respect to time.**

MCMC and Synthetic Likelihood

The Susceptible and Infectives are not observed. We can impute them using functional forms of non-diseased individuals:

$$\tilde{D}_1 = S_1 + I_1 = \left(\frac{\gamma_1}{\beta_{11}I_1 + \beta_{12}I_2 + \phi_1} + 1\right) I_1 = f_1^{\boldsymbol{\theta}_1}(I_1, I_2)$$

$$\tilde{D}_2 = S_2 + I_2 = \left(\frac{\gamma_2}{\beta_{22}I_2 + \beta_{21}I_1 + \phi_2} + 1\right) I_2 = f_2^{\boldsymbol{\theta}_2}(I_1, I_2)$$

$$\tilde{D}_1 = \frac{(\alpha_1 - \nu_1)I_1 + \delta_1 D_1}{\alpha_1} = f_3^{\boldsymbol{\theta}_3}(I_1)$$

$$\tilde{D}_2 = \frac{(\alpha_2 - \nu_2)I_2 + \delta_2 D_2}{\alpha_2} = f_4^{\boldsymbol{\theta}_4}(I_2)$$

Outline of MCMC procedure (V=0,1 separately):

1. Simulate average households (e.g. 50 households) B times.
2. Initialize Gamma priors on parameters and I_1, I_2
3. Draw from posterior of parameters via M-H
4. Draw from posterior of I_1, I_2 via M-H using

$$l_1 \propto \exp\left(\sum_{i=1}^{B}(\tilde{d}_i^1 - (f_1^{\boldsymbol{\theta}_1} + f_3^{\boldsymbol{\theta}_3})/2)^2/2\sigma_1^2\right),$$

$$l_2 \propto \exp\left(\sum_{i=1}^{B}(\tilde{d}_i^2 - (f_2^{\boldsymbol{\theta}_2} + f_4^{\boldsymbol{\theta}_4})/2)^2/2\sigma_2^2\right).$$

5. Repeat 3 and 4 until chain converges.

Results

The following are results are for the posterior means, standard deviations, and associated credible inter vals in the settings $V = 1$ and $V = 0$. We generated 100 average households of size 50.

	MC_mean	MC_sd	2.50%	97.50%
beta11	0.5350	0.4056	0.0366	1.5541
beta12	0.5128	0.4041	0.0417	1.5410
phi1	0.5254	0.3991	0.0540	1.5526
gamma1	0.6371	0.4382	0.0798	1.6457
beta22	0.4801	0.3796	0.0364	1.5018
beta21	0.5266	0.4094	0.0388	1.5610
phi2	0.5537	0.4279	0.0376	1.7793
gamma2	0.7937	0.6352	0.0761	2.3740
alpha1	0.7401	0.4461	0.1541	1.8767
nu1	0.2514	0.1844	0.0333	0.7074
delta1	0.6994	0.4830	0.0621	1.9595
alpha2	0.8236	0.4810	0.1446	2.0287
nu2	0.2233	0.1777	0.0138	0.6557
delta2	0.6217	0.4502	0.0561	1.8455
I1	1.6693	0.3838	0.8531	2.3165
I2	4.2566	0.6144	2.8644	5.1508

Figure 1: MCMC results: n = 50, V = 1

	MC_mean	MC_sd	2.50%	97.50%
beta11	0.5316	0.4191	0.0454	1.6545
beta12	0.4824	0.4220	0.0313	1.5507
gamma1	0.7024	0.4806	0.0733	1.9213
beta11	0.4742	0.4037	0.0429	1.6455
beta21	0.4882	0.3670	0.0323	1.3771
gamma2	0.7316	0.5022	0.0824	1.9509
alpha1	0.8006	0.4851	0.1282	1.9716
nu1	0.2656	0.1995	0.0165	0.7505
delta1	0.6647	0.4867	0.0573	1.9252
alpha2	0.7496	0.4683	0.1345	1.9672
nu2	0.2228	0.1773	0.0199	0.6745
delta2	0.6649	0.4925	0.0585	1.8735
I1	1.7448	0.4658	0.6885	2.5085
I2	3.7009	0.8404	1.6611	4.9509

Figure 2: MCMC results: n = 50, V = 0

Acknowledgements

This is joint work with Boseung Choi, Rebecca Garabed, and Jiyoung Lee.

This research has been supported in part by the Mathematical Biosciences Institute and the National Science Foundation under grant DMS 1440386.

Bibliography

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. 2013. *Bayesian Data Analysis*, 3rd Edition. Chapman & Hall/CRC Press, London.
Casella, George, and Roger L. Berger. Statistical Inference. Pacific Grove, Calif: Brooks/Cole Pub. Co, 1990.
end citation formats

Contact

Email: woroszylo.1@osu.edu
Site: casperw.github.io