

# Estimation of the effect of climate on infectious diseases

## Table of contents

<b>1</b>	<b>Motivating Question</b>	<b>1</b>
<b>2</b>	<b>Data &amp; model</b>	<b>2</b>
2.1	State Process (Evolution) Equation . . . . .	2
2.2	Latent Driver Process . . . . .	2
2.3	Observation Equation . . . . .	3
2.4	Priors . . . . .	3
2.5	Full Conditional Derivations . . . . .	3
2.5.1	1. Full Conditional for $Y_t$ . . . . .	3
2.5.2	2. Full Conditional for $U_t$ . . . . .	4
2.5.3	3. Full Conditional for $\alpha$ . . . . .	5
2.5.4	4. Full Conditional for $\beta_\mu$ . . . . .	6
2.5.5	5. Full Conditional for $\beta_\sigma$ . . . . .	6
2.5.6	6. Full Conditional for $\gamma$ . . . . .	6
2.5.7	7. Full Conditional for $\phi$ . . . . .	7
2.5.8	8. Full Conditional for $\eta$ . . . . .	7
2.5.9	9. Full Conditional for $\sigma_U^2$ . . . . .	7
2.5.10	10. Full Conditional for $\sigma_o^2$ . . . . .	7

## 1 Motivating Question

If climate change (through some sort of variable such as temperature or precipitation) is affecting the number of cases of infectious disease, it is an outstanding question how strong this effect must be to identify it from some background autocorrelated value.

How strong does the signal of climate change need to be to detect it?

## 2 Data & model

Assuming we have some data on observed cases of a given infectious disease. The relationship between those observed cases and actual cases is a state process with some observation error,  $\epsilon_o$ . Cases themselves are now given as a state space model where the number of cases at time  $t + 1$  are driven by the effect of both temperature variance (consistent through time) and mean temperature (increasing through time), as well as an unobserved driver that is correlated through time with the mean temperature.

Let: -  $Y_t$  represent the true number of cases at time  $t$ . -  $Y_t^{\text{obs}}$  represent the observed number of cases at time  $t$ , which includes observational error  $\epsilon_o$ . -  $X_\mu$  represent the mean temperature at time  $t$ , and let  $X_\sigma$  represent the temperature variance, assumed constant over time. -  $U_t$  represent the unobserved driver correlated with mean temperature  $X_\mu$ .

### 2.1 State Process (Evolution) Equation

For the true cases  $Y_t$ , we specify that the cases at time  $t + 1$  are driven by the temperature effects and an unobserved driver  $U_t$  as follows:

$$Y_{t+1} \mid Y_t, X_\mu, U_t \sim \text{Poisson} \left( Y_t \exp \left( \alpha + \beta_\mu X_\mu + \beta_\sigma X_\sigma + \gamma U_t \right) \right) \quad (1)$$

where: -  $\alpha$  is an intercept term. -  $\beta_\mu$  is the effect of mean temperature  $X_\mu$  on the cases. -  $\beta_\sigma$  is the effect of temperature variance  $X_\sigma$ . -  $\gamma$  captures the effect of the unobserved driver  $U_t$  on the cases.

### 2.2 Latent Driver Process

We assume that the unobserved driver  $U_t$  has temporal correlation and is influenced by the mean temperature:

$$U_{t+1} \mid U_t, X_\mu \sim \mathcal{N}(\phi U_t + \eta X_\mu, \sigma_U^2) \quad (2)$$

where: -  $\phi$  is an autoregressive parameter governing the temporal correlation of  $U_t$ . -  $\eta$  is the strength of the correlation between  $U_t$  and  $X_\mu$ . -  $\sigma_U^2$  is the variance of  $U_t$ .

## 2.3 Observation Equation

The observed cases  $Y_t^{\text{obs}}$  are related to the true cases  $Y_t$  with observation error  $\epsilon_o$ :

$$Y_t^{\text{obs}} \mid Y_t \sim \text{Poisson}(Y_t e^{\epsilon_o}) \quad (3)$$

with  $\epsilon_o \sim \mathcal{N}(0, \sigma_o^2)$ , where  $\sigma_o^2$  is the variance of the observation error.

## 2.4 Priors

The priors for the parameters could be specified as follows:

$$\alpha \sim \mathcal{N}(0, 10) \quad (4a)$$

$$\beta_\mu \sim \mathcal{N}(0, 1) \quad (4b)$$

$$\beta_\sigma \sim \mathcal{N}(0, 1) \quad (4c)$$

$$\gamma \sim \mathcal{N}(0, 1) \quad (4d)$$

$$\phi \sim \text{Uniform}(-1, 1) \quad (5a)$$

$$\eta \sim \mathcal{N}(0, 1) \quad (5b)$$

$$\sigma_U^2 \sim \text{Inverse-Gamma}(2, 1) \quad (5c)$$

$$\sigma_o^2 \sim \text{Inverse-Gamma}(2, 1) \quad (5d)$$

## 2.5 Full Conditional Derivations

### 2.5.1 1. Full Conditional for $Y_t$

The likelihood for  $Y_t$  comes from the Poisson model and the observation model:

$$Y_{t+1} \mid Y_t, X_\mu, U_t \sim \text{Poisson}(Y_t \exp(\alpha + \beta_\mu X_\mu + \beta_\sigma X_\sigma + \gamma U_t)) \quad (6)$$

$$Y_t^{\text{obs}} \mid Y_t \sim \text{Poisson}(Y_t e^{\epsilon_o}) \quad (7)$$

The joint likelihood is:

$$p(Y_t | Y_{t+1}, Y_t^{\text{obs}}, \alpha, \beta_\mu, \beta_\sigma, \gamma, U_t, \epsilon_o) \propto \text{Poisson}(Y_{t+1} | Y_t \exp(\alpha + \beta_\mu X_\mu + \beta_\sigma X_\sigma + \gamma U_t)) \times \text{Poisson}(Y_t^{\text{obs}} | Y_t e^{\epsilon_o}) \quad (8)$$

Simplifying this:

$$p(Y_t | Y_{t+1}, Y_t^{\text{obs}}, \alpha, \beta_\mu, \beta_\sigma, \gamma, U_t, \epsilon_o) \propto \exp(-Y_t \exp(\alpha + \beta_\mu X_\mu + \beta_\sigma X_\sigma + \gamma U_t)) \times (Y_t \exp(\alpha + \beta_\mu X_\mu + \beta_\sigma X_\sigma + \gamma U_t))^{Y_{t+1}} \times \exp(-Y_t e^{\epsilon_o}) (Y_t e^{\epsilon_o})^{Y_t^{\text{obs}}} \quad (9)$$

This simplifies to:

$$p(Y_t | Y_{t+1}, Y_t^{\text{obs}}, \alpha, \beta_\mu, \beta_\sigma, \gamma, U_t, \epsilon_o) \propto Y_t^{Y_{t+1} + Y_t^{\text{obs}}} \times \exp(-Y_t (\exp(\alpha + \beta_\mu X_\mu + \beta_\sigma X_\sigma + \gamma U_t) + e^{\epsilon_o})) \quad (10)$$

This is proportional to a Gamma distribution with shape  $Y_{t+1} + Y_t^{\text{obs}}$  and rate  $\exp(\alpha + \beta_\mu X_\mu + \beta_\sigma X_\sigma + \gamma U_t) + e^{\epsilon_o}$ . Therefore, the full conditional for  $Y_t$  is Gamma-distributed, and there is a closed-form solution.

### 2.5.2 2. Full Conditional for $U_t$

The likelihood for  $U_t$  comes from the Normal distribution and the Poisson likelihood for  $Y_{t+1}$ :

$$U_{t+1} | U_t, X_\mu \sim \mathcal{N}(\phi U_t + \eta X_\mu, \sigma_U^2) \quad (11)$$

$$Y_{t+1} | Y_t, U_t \sim \text{Poisson}(Y_t \exp(\alpha + \beta_\mu X_\mu + \beta_\sigma X_\sigma + \gamma U_t)) \quad (12)$$

The joint likelihood is:

$$p(U_t | U_{t+1}, Y_{t+1}, Y_t, \alpha, \beta_\mu, \beta_\sigma, \gamma, X_\mu, \sigma_U^2) \propto \mathcal{N}(U_{t+1} | \phi U_t + \eta X_\mu, \sigma_U^2) \times \text{Poisson}(Y_{t+1} | Y_t \exp(\alpha + \beta_\mu X_\mu + \beta_\sigma X_\sigma + \gamma U_t)) \quad (13)$$

Substituting the likelihoods:

$$\begin{aligned}
p(U_t \mid U_{t+1}, Y_{t+1}, Y_t, \alpha, \beta_\mu, \beta_\sigma, \gamma, X_\mu, \sigma_U^2) \propto \\
\exp \left( -\frac{(U_{t+1} - (\phi U_t + \eta X_\mu))^2}{2\sigma_U^2} \right) \times \\
\exp(-Y_t \exp(\alpha + \beta_\mu X_\mu + \beta_\sigma X_\sigma + \gamma U_t))
\end{aligned} \tag{14}$$

This is a mixture of Gaussian and Poisson terms, which does not have a closed-form solution. Numerical methods such as Metropolis-Hastings would be used for approximation.

### 2.5.3 3. Full Conditional for $\alpha$

The prior for  $\alpha$  is  $\mathcal{N}(0, 10)$ , and the full conditional is:

$$p(\alpha \mid Y, X_\mu, X_\sigma, U) \propto \mathcal{N}(0, 10) \times \prod_{t=1}^T \text{Poisson}(Y_{t+1} \mid Y_t \exp(\alpha + \beta_\mu X_\mu + \beta_\sigma X_\sigma + \gamma U_t)) \tag{15}$$

Substituting the Poisson likelihood:

$$p(\alpha \mid Y, X_\mu, X_\sigma, U) \propto \exp \left( -\frac{\alpha^2}{2 \cdot 10^2} \right) \times \prod_{t=1}^T \exp(-Y_t \exp(\alpha + \beta_\mu X_\mu + \beta_\sigma X_\sigma + \gamma U_t)) \tag{16}$$

This simplifies to:

$$p(\alpha \mid Y, X_\mu, X_\sigma, U) \propto \exp \left( -\frac{\alpha^2}{2 \cdot 10^2} - \sum_{t=1}^T Y_t \exp(\alpha + \beta_\mu X_\mu + \beta_\sigma X_\sigma + \gamma U_t) \right) \tag{17}$$

This is a Gaussian-exponential mixture, and there is no closed-form solution. Approximation methods (e.g., numerical optimization) would be needed.

#### 2.5.4 4. Full Conditional for $\beta_\mu$

The prior for  $\beta_\mu$  is  $\mathcal{N}(0, 1)$ , and the full conditional is:

$$p(\beta_\mu \mid Y, X_\mu, X_\sigma, U) \propto \mathcal{N}(0, 1) \times \prod_{t=1}^T \text{Poisson}(Y_{t+1} \mid Y_t \exp(\alpha + \beta_\mu X_\mu + \beta_\sigma X_\sigma + \gamma U_t)) \quad (18)$$

Substituting the Poisson likelihood:

$$p(\beta_\mu \mid Y, X_\mu, X_\sigma, U) \propto \exp\left(-\frac{\beta_\mu^2}{2}\right) \times \prod_{t=1}^T \exp(-Y_t \exp(\alpha + \beta_\mu X_\mu + \beta_\sigma X_\sigma + \gamma U_t)) \quad (19)$$

This leads to a Gaussian-exponential mixture, which does not have a closed-form solution.

#### 2.5.5 5. Full Conditional for $\beta_\sigma$

The full conditional for  $\beta_\sigma$  follows the same structure as for  $\beta_\mu$ :

$$p(\beta_\sigma \mid Y, X_\mu, X_\sigma, U) \propto \mathcal{N}(0, 1) \times \prod_{t=1}^T \text{Poisson}(Y_{t+1} \mid Y_t \exp(\alpha + \beta_\mu X_\mu + \beta_\sigma X_\sigma + \gamma U_t)) \quad (20)$$

This does not have a closed-form solution either.

#### 2.5.6 6. Full Conditional for $\gamma$

The full conditional for  $\gamma$  is:

$$p(\gamma \mid Y, X_\mu, X_\sigma, U) \propto \mathcal{N}(0, 1) \times \prod_{t=1}^T \text{Poisson}(Y_{t+1} \mid Y_t \exp(\alpha + \beta_\mu X_\mu + \beta_\sigma X_\sigma + \gamma U_t)) \quad (21)$$

This does not have a closed-form solution.

### 2.5.7 7. Full Conditional for $\phi$

The prior for  $\phi$  is  $\text{Uniform}(-1, 1)$ , and the full conditional is:

$$p(\phi \mid U, X_\mu) \propto \text{Uniform}(-1, 1) \times \prod_{t=1}^T \mathcal{N}(U_{t+1} \mid \phi U_t + \eta X_\mu, \sigma_U^2) \quad (22)$$

This is a product of Gaussian distributions. The full conditional does not have a closed-form solution but can be approximated numerically.

### 2.5.8 8. Full Conditional for $\eta$

Similarly, for  $\eta$ :

$$p(\eta \mid U, X_\mu) \propto \mathcal{N}(0, 1) \times \prod_{t=1}^T \mathcal{N}(U_{t+1} \mid \phi U_t + \eta X_\mu, \sigma_U^2) \quad (23)$$

This is a Gaussian-exponential mixture, and no closed-form solution exists.

### 2.5.9 9. Full Conditional for $\sigma_U^2$

For  $\sigma_U^2$ , the prior is  $\text{Inverse-Gamma}(2, 1)$ , and the full conditional is:

$$p(\sigma_U^2 \mid U, X_\mu, \phi, \eta) \propto \text{Inverse-Gamma}(2, 1) \times \prod_{t=1}^T \mathcal{N}(U_{t+1} \mid \phi U_t + \eta X_\mu, \sigma_U^2) \quad (24)$$

The full conditional is an **Inverse-Gamma distribution** with updated shape and rate parameters. This has a **closed-form solution**.

### 2.5.10 10. Full Conditional for $\sigma_o^2$

For  $\sigma_o^2$ , the prior is  $\text{Inverse-Gamma}(2, 1)$ , and the full conditional is:

$$p(\sigma_o^2 \mid Y^{\text{obs}}, Y) \propto \text{Inverse-Gamma}(2, 1) \times \prod_{t=1}^T \mathcal{N}(\log(Y_t^{\text{obs}}/Y_t) \mid 0, \sigma_o^2) \quad (25)$$

This conditional is also **Inverse-Gamma** with parameters updated based on the data. Hence, it has a **closed-form solution**.