Advanced Bayesian Inference with Case Studies
Project: Salm: Extra-Poisson Variation in Dose-Response Study

Samar ElAyed, Nassim Karimi

2024 - 2025

1 Introduction

This project analyzes the effect of quinoline dose on TA98 Salmonella revertant colonies using a **Bayesian approach**. Counts of revertant colonies were observed across doses with three plates per dose level. The aim is to model the dose-response relationship while addressing **over-dispersion**.

Objective

1. Model the dose-response relationship using a **Poisson random effects model**.
2. Estimate model parameters (α, β, γ, τ) via **Bayesian inference**.
3. Compare results from **from-scratch MCMC implementations** (Metropolis-Hastings, Gibbs Sampling) with those using **PyMC**.

2 Model

Data

Observed counts $\{y_{ij}\}$ follow:

$$\begin{aligned} y_{ij} &\sim \text{Poisson}(\mu_{ij}), \\ \log(\mu_{ij}) &= \alpha + \beta \log(x_i + 10) + \gamma x_i + \lambda_{ij}, \\ \lambda_{ij} &\sim \mathcal{N}(0, \tau). \end{aligned}$$

Prior choices

We use weakly informative priors for all parameters in order to reflect minimal prior knowledge while keeping the inference stable.

- For the regression parameters α, β, and γ, we use *weakly informative prior*:

  $$\alpha, \beta, \gamma \sim \mathcal{N}(0,\ 10^6)$$

  This is a normal distribution centered at 0 with a very large variance. It allows the parameters to vary freely over a wide range, while still regularizing extreme values. This helps prevent numerical instability that can arise with completely flat (improper) priors.

- For the precision parameter τ , we use:

  $$\tau \sim \text{Gamma}(0.001,\ 0.001)$$

  This is a standard choice in Bayesian hierarchical models when little prior information is available. It places most of its mass near zero but with heavy tails, allowing the posterior to adapt depending on the data. However, such vague priors can sometimes cause poor mixing or overly heavy tails in practice.

For our Bayesian hierarchical Poisson regression with random effects, we evaluated three MCMC approaches:

3 Bayesian Model Selection for Poisson Regression

3.1 Pure Gibbs Sampling

Pure Gibbs Sampling is infeasible for this model due to the lack of conjugacy in the full conditionals for α, β, and γ. This arises from the Poisson likelihood with a log-link and the non-standard distributions for $\lambda_{ij}|\text{rest}$. The only exception is τ, which has a tractable Gamma conditional.

3.2 Full Metropolis-Hastings

The Full Metropolis-Hastings approach is also impractical. Challenges include the high dimensionality of the model (particularly with λ_{ij}), difficulties in tuning the proposal distribution, and low acceptance rates, making it unsuitable for our needs.

3.3 MH within Gibbs (Selected Approach)

The chosen method is a hybrid strategy combining Gibbs steps and Metropolis-Hastings (MH). Gibbs steps are used for parameters with tractable conditionals, such as τ with its Gamma distribution. MH steps are applied for α, β, γ (using random walk proposals) and λ_{ij} (with adaptive sampling). This approach ensures efficient mixing, naturally addresses non-conjugacy, and allows for flexible proposal tuning.

4 Full Conditional Distributions

In this section, we derive the full conditional distributions targeted in the Metropolis-within-Gibbs algorithm, based strictly on our hierarchical model and the actual implementation in code. The full conditionals are not necessarily conjugate, and are therefore sampled using Metropolis–Hastings (MH) steps where required.

(a) Precision τ (MH Step on log-scale)

The precision τ governs the prior variance of the latent variables λ_{ij} through:

$$\lambda_{ij} \sim \mathcal{N}(0,\ \tau^{-1}) \quad \Rightarrow \quad p(\lambda_{ij} \mid \tau) \propto \tau^{1/2} \exp\left(-\frac{\tau}{2}\lambda_{ij}^2\right)$$

The conditional posterior is thus proportional to:

$$p(\tau \mid \lambda) \propto \tau^{a+n/2-1} \exp\left(-\tau\left(b + \frac{1}{2}\sum_{i,j}\lambda_{ij}^2\right)\right) \tag{1}$$

where $n = 6 \times 3 = 18$ is the total number of random effects.

Although this corresponds to a Gamma distribution, our implementation does not use a direct Gibbs update. Instead, we define $u = \log(\tau)$ and perform a Metropolis–Hastings step with:

$$u^* \sim \mathcal{N}(u, \sigma_\tau^2), \quad \tau^* = \exp(u^*)$$

The acceptance ratio is computed using the log-posterior difference and includes $(u^* - u)$.

(b) Random Effects λ_{ij} (MH Step)

Each latent effect λ_{ij} appears both in the Poisson likelihood and in its Gaussian prior. The log-mean is defined as:

$$\log(\mu_{ij}) = \alpha + \beta\log(x_i + 10) + \gamma x_i + \lambda_{ij} = \eta_{ij} + \lambda_{ij}$$

The likelihood contribution is:

$$p(y_{ij} \mid \lambda_{ij}) \propto \exp\left(y_{ij}(\eta_{ij} + \lambda_{ij}) - \exp(\eta_{ij} + \lambda_{ij})\right)$$

The prior contribution is:

$$p(\lambda_{ij} \mid \tau) \propto \exp\left(-\frac{\tau}{2}\lambda_{ij}^2\right)$$

Thus, the full conditional posterior is proportional to:

$$p(\lambda_{ij} \mid \cdot) \propto \exp\left(y_{ij}(\eta_{ij} + \lambda_{ij}) - \exp(\eta_{ij} + \lambda_{ij}) - \frac{\tau}{2}\lambda_{ij}^2\right) \tag{2}$$

Since this form is not standard, we perform a Metropolis–Hastings update:

$$\lambda_{ij}^{\text{new}} \sim \mathcal{N}(\lambda_{ij}^{\text{current}}, \sigma_\lambda^2) \tag{3}$$

(c) Regression Coefficients θ (MH Step)

Each regression coefficient $\theta \in \{\alpha, \beta, \gamma\}$ appears in :

$$\log(\mu_{ij}) = \alpha + \beta \log(x_i + 10) + \gamma x_i + \lambda_{ij}$$

The full conditional log-posterior for θ is:

$$\log p(\theta \mid \cdot) \propto -\frac{\theta^2}{2 \cdot 10^6} + \sum_{i,j} (y_{ij} \cdot \eta_{ij} - \exp(\eta_{ij} + \lambda_{ij})) \tag{4}$$

where η_{ij} depends linearly on θ.

This posterior is not conjugate, and is thus sampled using a Metropolis–Hastings step:

$$\theta^{\text{new}} \sim \mathcal{N}(\theta^{\text{current}}, \sigma_\theta^2) \tag{5}$$

Algorithm Pseudocode

1. Initialize parameters: α, β, γ, τ, and λ_{ij}
2. For each MCMC iteration:
   (a) Propose new α using a random walk Metropolis–Hastings step
   (b) Propose new β using a random walk Metropolis–Hastings step
   (c) Propose new γ using a random walk Metropolis–Hastings step
   (d) For each λ_{ij}:
      - Propose a new value using a random walk
      - Accept or reject using the change in log-posterior
   (e) Propose new τ by random walk on $\log(\tau)$:
      - Propose $u^* \sim \mathcal{N}(\log(\tau), \sigma_\tau^2)$
      - Set $\tau^* = \exp(u^*)$
      - Accept or reject using log-posterior difference and $J = (u^* - u)$.

5 Results

Parameter	Breslow's Reported Values	MCMC Results	Agreement?
α (Intercept)	2.203 ± 0.363 (Var = 0.132)	1.891 ± 0.342 (Var = 0.117)	**Good**
β (Log-dose coefficient)	0.311 ± 0.099 (Var = 0.010)	0.392 ± 0.093 (Var = 0.009)	**Good**
γ (Dose coefficient)	$-9.74 \times 10^{-4} \pm 4.37 \times 10^{-4}$	$-0.00127 \pm 4.06 \times 10^{-4}$	**Good**
σ (Random effects SD)	0.268 (no std. dev. reported)	0.236 ± 0.069 (Var = 0.005)	**Good**

Table 1: Comparison between Breslow's reported values and MCMC results from our implementation.

The MCMC results from our Metropolis-Hastings within Gibbs sampler show a good match with Breslow's published values for most parameters. The intercept (α) was estimated at **1.89 ± 0.34**, which is close to Breslow's reported value of **2.203 ± 0.363**, and lies within one standard deviation. The random effects standard deviation (σ) was about **0.236 ± 0.069**, which is also close to Breslow's estimate of **0.268**. The log-dose coefficient (β) was estimated at **0.392 ± 0.093**, slightly higher than Breslow's

0.311 $\pm$ **0.099**, but the values are still close and the intervals overlap. This means the difference is not important. The dose coefficient (γ) was estimated at **-0.00127** $\pm$ **0.00041**, which is very close to zero. This suggests that this term might not be necessary in the model. The precision parameter (τ) had a large mean value (**24.65**) with high variability ($\pm$**29.66**), showing that this parameter mixes poorly or is strongly affected by the prior. However, the standard deviation $\sigma = 1/\sqrt{\tau}$ was more stable, and gave results that are easier to interpret.

6 PyMC

We applied the PyMC algorithm to validate our results. PyMC is a Python library for Bayesian modeling that automates sampling, builds probabilistic models, and validates results with tools like Posterior Predictive Checks (PPC), streamlining complex analyses.

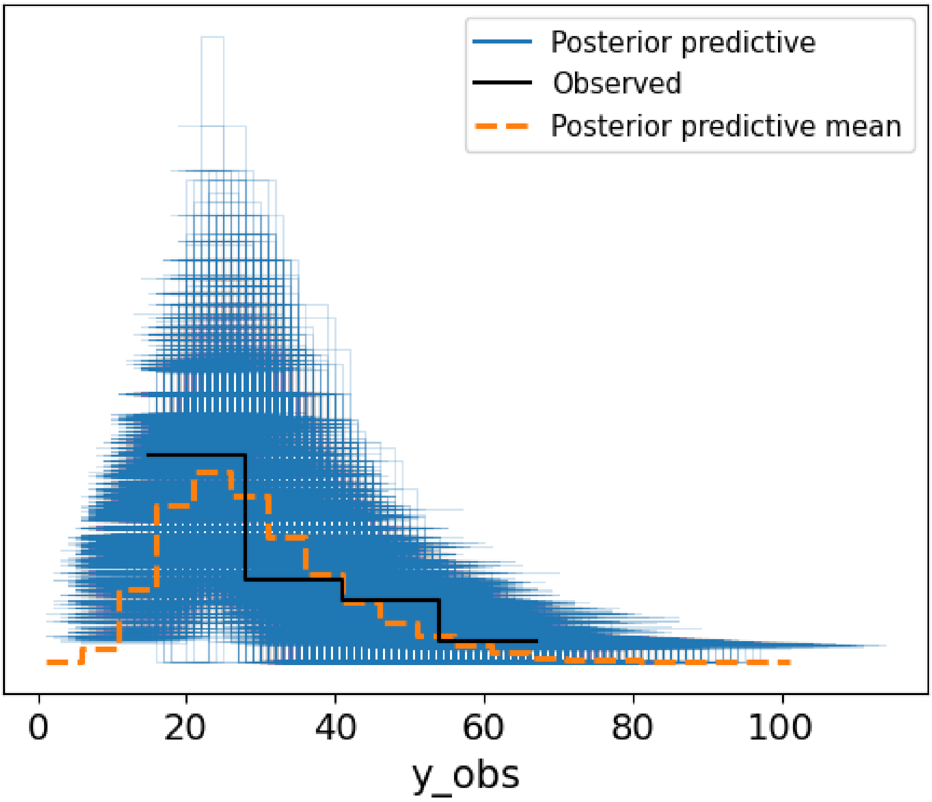

Figure 1: PyMC PPC

We conducted a Posterior Predictive Check (PPC) to compare the observed data with the model's posterior predictive distribution. The PPC evaluates how well the Bayesian model captures the data by analyzing key elements: the blue region, representing the posterior predictive distribution, should align closely with the black line, which shows the observed data. A well-fitting model displays a good general fit, as seen when the black line lies within the blue region, indicating the model captures the data's essence. The orange dashed line, the posterior predictive mean, further confirms the model's accuracy if it mirrors the observed trend. However, the distribution's right-skewed nature suggests the model sometimes predicts higher counts, hinting at over-dispersion or potential limitations in the Poisson assumption. Despite this, the predictive mean aligns well with the data, highlighting the model's overall effectiveness in capturing the dose-response relationship.

Appendix A: Trace Plots and Posterior Densities from MCMC simulation

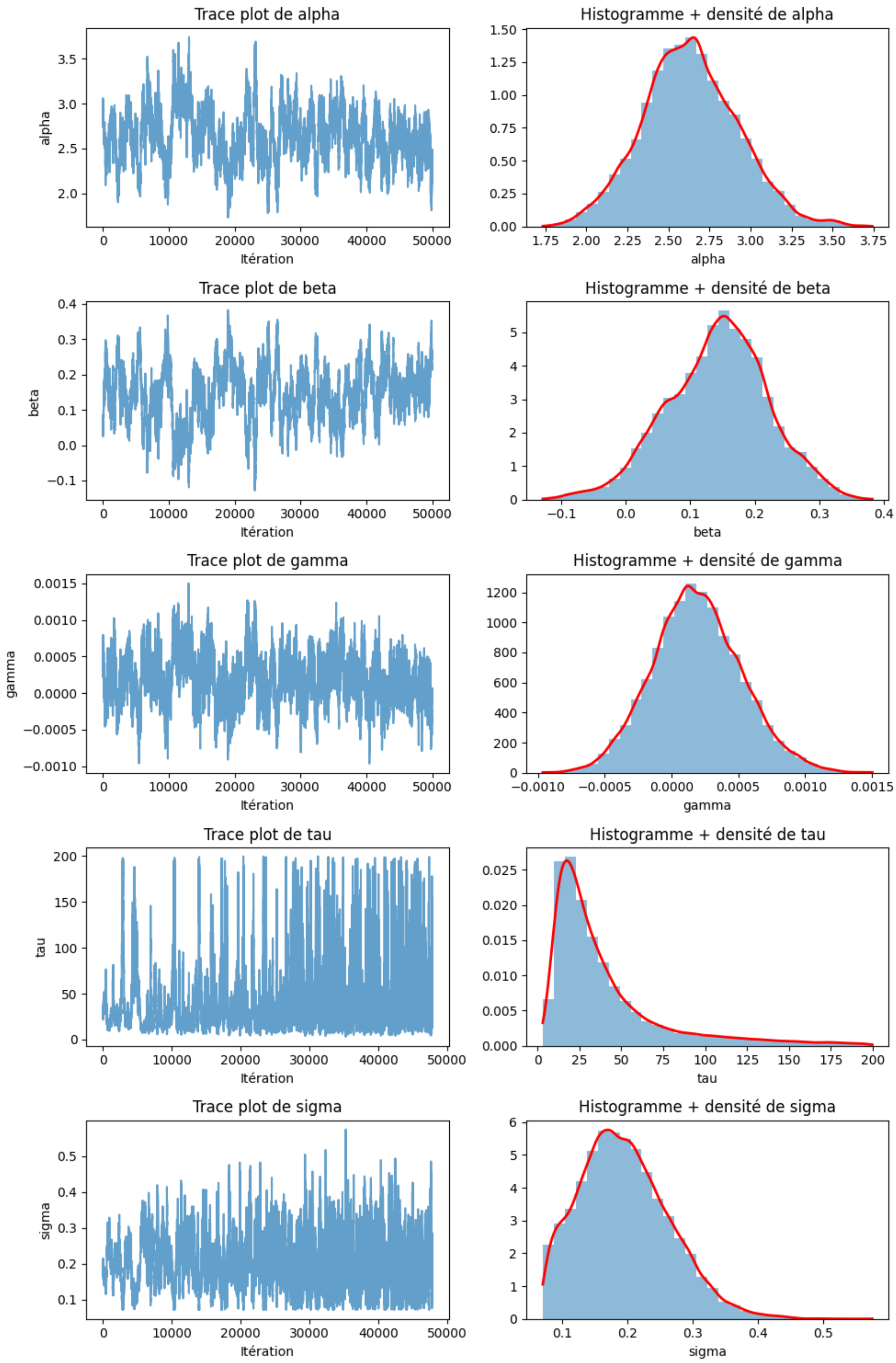

Figure 2: Trace plots and posterior densities for the parameters α, β, γ, τ, and σ.

Figure 2 shows the trace plots and posterior densities for all model parameters. We observe good mixing and convergence for most parameters, especially α, β, and γ. The parameter τ shows higher variability and occasional large jumps, which is expected due to the vague Gamma prior. The derived standard deviation σ is more stable and easier to interpret. Overall, the chains appear to explore the posterior space well.

The full implementation of the Metropolis-within-Gibbs sampler, along with plots and analysis, is available in the notebook on GitHub:

`https://github.com/samarelayed/Advanced_Bayesian_Inference_Salm_Project`