

Biostatistics 682: Applied Bayesian Inference

Lecture 6: Simple linear regression and Introduction to Gibbs Sampler

Jian Kang

Department of Biostatistics
University of Michigan, Ann Arbor

Motivating Example: Clinical Trial

Setup:

- n patients enrolled, each receives drug dosage x_i (mg).
- Outcome y_i : observed reduction in systolic blood pressure (mmHg).

Statistical model:

$$y_i = \alpha + \beta x_i + \epsilon_i, \quad \epsilon_i \sim N(0, \sigma_\epsilon^2).$$

- α : average baseline reduction at zero dose.
- β : slope parameter = effect of each additional mg of dosage.
- σ_ϵ^2 : residual variance = patient-to-patient variability.

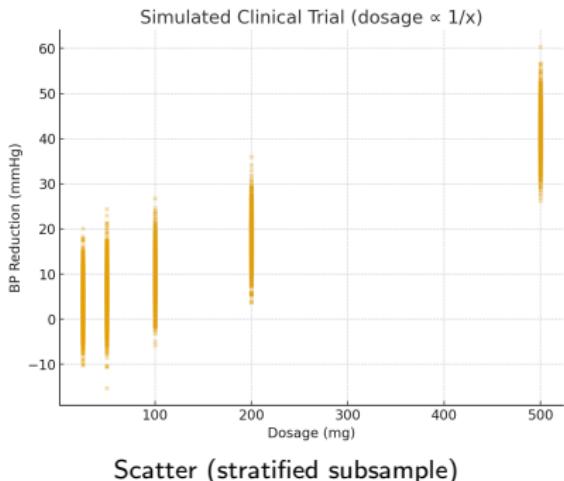
This simple linear regression connects *clinical questions* to *posterior inference*.

Data Collected

Design: $n = 100,000$; dosages in $\{25, 50, 100, 200, 500\}$ mg. BP reduction recorded to 1 decimal point (mmHg)

Summary by dosage group

Dose	Mean	SD	n
25	4.00	4.98	51,815
50	6.02	5.01	26,105
100	10.01	5.01	13,027
200	18.04	4.99	6,413
500	41.93	4.85	2,640



Takeaway: Mean BP reduction increases with dose; variability is roughly constant across groups.

Clinical Questions in a Drug Trial

Clinicians often ask:

- How likely that the treatment effect is positive?

$$P(\beta > 0 \mid y, x)$$

- How large is the expected blood pressure reduction for a new patient at a given dose?

$$E(y_* \mid x_* = 50, y, x)$$

- How much uncertainty remains about variability in responses?

$$E(\sigma_\epsilon^2 \mid y, x)$$

- How can we formally incorporate prior evidence from earlier phase trials into this analysis?

Bayesian inference directly answers these questions using posterior distributions of parameters and predictive distributions for new patients.

Bayesian Simple Linear Regression

Model

$$y_i = \alpha + \beta x_i + \epsilon_i, \quad \epsilon_i \sim N(0, \sigma_\epsilon^2).$$

Equivalent Model Representation

$$y_i | \alpha, \beta, \sigma_\epsilon^2 \sim N(\alpha + \beta x_i, \sigma_\epsilon^2).$$

Natural Conjugate Priors

$$\alpha \sim N(0, \sigma_\alpha^2), \quad \beta \sim N(0, \sigma_\beta^2), \quad \sigma_\epsilon^2 \sim G^{-1}(a_\epsilon, b_\epsilon).$$

Equivalent Precision Prior

$$\tau_\epsilon^2 = 1/\sigma_\epsilon^2 \sim G(a_\epsilon, b_\epsilon).$$

Similarly, we can define $\tau_\alpha^2 = 1/\sigma_\alpha^2$ and $\tau_\beta^2 = 1/\sigma_\beta^2$.

Simple Linear Regression

$$y_i \mid \alpha, \beta, \tau_\epsilon^2 \sim N(\alpha + \beta x_i, \tau_\epsilon^{-2})$$

$$\alpha \sim N(0, \tau_\alpha^{-2})$$

$$\beta \sim N(0, \tau_\beta^{-2})$$

$$\tau_\epsilon^2 \sim G(a_\epsilon, b_\epsilon)$$

where τ_α^2 , τ_β^2 , a_ϵ and b_ϵ are the pre-specified hyperparameters.

Posterior distributions

- Joint posterior distribution

$$\pi(\alpha, \beta, \tau^2 | y, x) \propto \pi(y | x, \alpha, \beta, \tau^2) \pi(\alpha) \pi(\beta) \pi(\tau^2)$$

- The closed form of the joint posterior distribution can be complicated.
- The conditional posterior distribution of each parameter (full conditional) has the closed form.
- Full conditional distributions

$$\pi(\alpha | \beta, \tau_\epsilon^2, y, x) \propto \pi(y | x, \alpha, \beta, \tau_\epsilon^2) \pi(\alpha)$$

$$\pi(\beta | \alpha, \tau_\epsilon^2, y, x) \propto \pi(y | x, \alpha, \beta, \tau_\epsilon^2) \pi(\beta)$$

$$\pi(\tau_\epsilon^2 | \alpha, \beta, y, x) \propto \pi(y | x, \alpha, \beta, \tau_\epsilon^2) \pi(\tau_\epsilon^2)$$

Likelihood w.r.t. α, β, τ^2

Likelihood (Normal with precision $\tau^2 = 1/\sigma^2$)

$$\pi(y | x, \alpha, \beta, \tau^2) = (2\pi)^{-n/2} (\tau^2)^{n/2} \exp\left\{-\frac{\tau^2}{2} \text{SSE}(\alpha, \beta)\right\}.$$

Up to constants in (α, β, τ^2)

$$\pi(y | x, \alpha, \beta, \tau^2) \propto L(\alpha, \beta, \tau^2; y, x) = (\tau^2)^{n/2} \exp\left\{-\frac{\tau^2}{2} \text{SSE}(\alpha, \beta)\right\}.$$

SSE and summary statistics

$$\text{SSE}(\alpha, \beta) = \sum_{i=1}^n (y_i - \alpha - \beta x_i)^2.$$

Define

$$S_x = \sum_{i=1}^n x_i, \quad S_{xx} = \sum_{i=1}^n x_i^2, \quad S_y = \sum_{i=1}^n y_i, \quad S_{xy} = \sum_{i=1}^n x_i y_i. \quad S_{yy} = \sum_{i=1}^n y_i^2$$

Expanded in (α, β) :

$$\begin{aligned} \text{SSE}(\alpha, \beta) &= n\alpha^2 - 2\alpha \underbrace{(S_y - \beta S_x)}_{\text{const wrt } \alpha} + \underbrace{\beta^2 S_{xx} - 2\beta S_{xy} + S_{yy}}_{\text{const wrt } \alpha} \\ &= \beta^2 S_{xx} - 2\beta \underbrace{(S_{xy} - \alpha S_x)}_{\text{const wrt } \beta} + \underbrace{n\alpha^2 - 2\alpha S_y + S_{yy}}_{\text{const wrt } \beta} \end{aligned}$$

Full Conditional: Intercept

Simplify the full conditional of α :

$$\begin{aligned}\pi(\alpha | \beta, \tau_\epsilon^2, y, x) &\propto \exp \left\{ -\frac{1}{2} \tau_\epsilon^2 (n\alpha^2 - 2\alpha(S_y - \beta S_x)) \right\} \exp \left(-\frac{1}{2} \tau_\alpha^2 \alpha^2 \right) \\ &\propto \exp \left\{ -\frac{1}{2} [(n\tau_\epsilon^2 + \tau_\alpha^2)\alpha^2 - 2\alpha\tau_\epsilon^2(S_y - \beta S_x)] \right\}\end{aligned}$$

This implies that

$$\alpha | \beta, \tau^2, y, x \sim N(\tilde{\mu}_\alpha, \tilde{\tau}_\alpha^{-2})$$

with

$$\tilde{\tau}_\alpha^2 = n\tau_\epsilon^2 + \tau_\alpha^2, \quad \tilde{\mu}_\alpha = \frac{\tau_\epsilon^2(S_y - \beta S_x)}{\tilde{\tau}_\alpha^2} = \frac{\tau_\epsilon^2(S_y - \beta S_x)}{n\tau_\epsilon^2 + \tau_\alpha^2},$$

Full Conditional: Intercept

Simplify the full conditional of α :

$$\begin{aligned}\pi(\alpha | \beta, \tau_\epsilon^2, y, x) &\propto \exp \left\{ -\frac{1}{2} \tau_\epsilon^2 (n\alpha^2 - 2\alpha(S_y - \beta S_x)) \right\} \exp \left(-\frac{1}{2} \tau_\alpha^2 \alpha^2 \right) \\ &\propto \exp \left\{ -\frac{1}{2} [(n\tau_\epsilon^2 + \tau_\alpha^2)\alpha^2 - 2\alpha\tau_\epsilon^2(S_y - \beta S_x)] \right\}\end{aligned}$$

This implies that

$$\alpha | \beta, \tau^2, y, x \sim N(\tilde{\mu}_\alpha, \tilde{\tau}_\alpha^{-2})$$

with

$$\tilde{\tau}_\alpha^2 = n\tau_\epsilon^2 + \tau_\alpha^2, \quad \tilde{\mu}_\alpha = \frac{\tau_\epsilon^2(S_y - \beta S_x)}{\tilde{\tau}_\alpha^2} = \frac{\tau_\epsilon^2(S_y - \beta S_x)}{n\tau_\epsilon^2 + \tau_\alpha^2},$$

Useful Fact

$$\pi_\xi(x) \propto \exp \left(-\frac{1}{2}(ax^2 - 2bx) \right) \Leftrightarrow \xi \sim N(b/a, 1/a)$$

Full Conditional: Slope

Simplify the full conditional of β :

$$\begin{aligned}\pi(\beta | \alpha, \tau_\epsilon^2, y, x) &\propto \exp\left\{-\frac{1}{2}\tau_\epsilon^2 (\beta^2 S_{xx} - 2\beta(S_{xy} - \alpha S_x))\right\} \exp\left(-\frac{1}{2}\tau_\beta^2 \beta^2\right) \\ &\propto \exp\left\{-\frac{1}{2} [(\tau_\epsilon^2 S_{xx} + \tau_\beta^2)\beta^2 - 2\beta\tau_\epsilon^2(S_{xy} - \alpha S_x)]\right\}.\end{aligned}$$

This implies that

$$\beta | \alpha, \tau_\epsilon^2, y, x \sim N(\tilde{\mu}_\beta, 1/\tilde{\tau}_\beta^2)$$

with

$$\tilde{\tau}_\beta^2 = \tau_\epsilon^2 S_{xx} + \tau_\beta^2, \quad \tilde{\mu}_\beta = \frac{\tau_\epsilon^2 (S_{xy} - \alpha S_x)}{\tilde{\tau}_\beta^2} = \frac{\tau_\epsilon^2 (S_{xy} - \alpha S_x)}{\tau_\epsilon^2 S_{xx} + \tau_\beta^2}.$$

Full Conditional: Precision τ_ϵ^2

Using the likelihood and Gamma prior:

$$\begin{aligned}\pi(\tau_\epsilon^2 \mid \alpha, \beta, y, x) &\propto (\tau_\epsilon^2)^{n/2} \exp\left\{-\frac{\tau_\epsilon^2}{2} \text{SSE}(\alpha, \beta)\right\} (\tau_\epsilon^2)^{a_\epsilon-1} \exp(-b_\epsilon \tau_\epsilon^2) \\ &\propto (\tau_\epsilon^2)^{a_\epsilon + \frac{n}{2}-1} \exp\left\{-\left(b_\epsilon + \frac{1}{2} \text{SSE}(\alpha, \beta)\right) \tau_\epsilon^2\right\}.\end{aligned}$$

Therefore,

$$\tau_\epsilon^2 \mid \alpha, \beta, y, x \sim G\left(a_\epsilon + \frac{n}{2}, b_\epsilon + \frac{1}{2} \text{SSE}(\alpha, \beta)\right),$$

where $\text{SSE}(\alpha, \beta) = \sum_{i=1}^n (y_i - \alpha - \beta x_i)^2$.

Full Conditional: Precision τ_ϵ^2

Using the likelihood and Gamma prior:

$$\begin{aligned}\pi(\tau_\epsilon^2 \mid \alpha, \beta, y, x) &\propto (\tau_\epsilon^2)^{n/2} \exp\left\{-\frac{\tau_\epsilon^2}{2} \text{SSE}(\alpha, \beta)\right\} (\tau_\epsilon^2)^{a_\epsilon-1} \exp(-b_\epsilon \tau_\epsilon^2) \\ &\propto (\tau_\epsilon^2)^{a_\epsilon + \frac{n}{2}-1} \exp\left\{-\left(b_\epsilon + \frac{1}{2} \text{SSE}(\alpha, \beta)\right) \tau_\epsilon^2\right\}.\end{aligned}$$

Therefore,

$$\tau_\epsilon^2 \mid \alpha, \beta, y, x \sim G\left(a_\epsilon + \frac{n}{2}, b_\epsilon + \frac{1}{2} \text{SSE}(\alpha, \beta)\right),$$

where $\text{SSE}(\alpha, \beta) = \sum_{i=1}^n (y_i - \alpha - \beta x_i)^2$.

Useful Fact (Gamma)

If $\pi_T(t) \propto t^{a-1} \exp(-bt)$ for $t > 0$, then $T \sim G(a, b)$.

Gibbs Sampler: Overview

Goal: sample from the posterior $\pi(\alpha, \beta, \tau_\epsilon^2 | y, x)$ via iterating full conditionals.

Full conditionals (from previous slides):

$$\alpha | \beta, \tau_\epsilon^2, y, x \sim N(\tilde{\mu}_\alpha, \tilde{\tau}_\alpha^{-2}), \quad \tilde{\tau}_\alpha^2 = n\tau_\epsilon^2 + \tau_\alpha^2, \quad \tilde{\mu}_\alpha = \frac{\tau_\epsilon^2(S_y - \beta S_x)}{\tilde{\tau}_\alpha^2}.$$

$$\beta | \alpha, \tau_\epsilon^2, y, x \sim N(\tilde{\mu}_\beta, \tilde{\tau}_\beta^{-2}), \quad \tilde{\tau}_\beta^2 = \tau_\epsilon^2 S_{xx} + \tau_\beta^2, \quad \tilde{\mu}_\beta = \frac{\tau_\epsilon^2(S_{xy} - \alpha S_x)}{\tilde{\tau}_\beta^2}.$$

$$\tau_\epsilon^2 | \alpha, \beta, y, x \sim G\left(a_\epsilon + \frac{n}{2}, b_\epsilon + \frac{1}{2} \sum_{i=1}^n (y_i - \alpha - \beta x_i)^2\right).$$

Gibbs Sampler: Overview

Goal: sample from the posterior $\pi(\alpha, \beta, \tau_\epsilon^2 | y, x)$ via iterating full conditionals.

Full conditionals (from previous slides):

$$\alpha | \beta, \tau_\epsilon^2, y, x \sim N(\tilde{\mu}_\alpha, \tilde{\tau}_\alpha^{-2}), \quad \tilde{\tau}_\alpha^2 = n\tau_\epsilon^2 + \tau_\alpha^2, \quad \tilde{\mu}_\alpha = \frac{\tau_\epsilon^2(S_y - \beta S_x)}{\tilde{\tau}_\alpha^2}.$$

$$\beta | \alpha, \tau_\epsilon^2, y, x \sim N(\tilde{\mu}_\beta, \tilde{\tau}_\beta^{-2}), \quad \tilde{\tau}_\beta^2 = \tau_\epsilon^2 S_{xx} + \tau_\beta^2, \quad \tilde{\mu}_\beta = \frac{\tau_\epsilon^2(S_{xy} - \alpha S_x)}{\tilde{\tau}_\beta^2}.$$

$$\tau_\epsilon^2 | \alpha, \beta, y, x \sim G\left(a_\epsilon + \frac{n}{2}, b_\epsilon + \frac{1}{2} \sum_{i=1}^n (y_i - \alpha - \beta x_i)^2\right).$$

Inputs: data $\{(x_i, y_i)\}_{i=1}^n$ and hyperparameters $\tau_\alpha^2, \tau_\beta^2, a_\epsilon, b_\epsilon$.

Outputs: posterior draws $\{(\alpha^{(t)}, \beta^{(t)}, \tau_\epsilon^{2(t)})\}_{t=1}^T$.

Theoretical Foundation of Gibbs Sampling

Key idea: Instead of sampling from the joint posterior directly, sample sequentially from each *full conditional* distribution.

- Suppose parameters are $\theta = (\theta_1, \dots, \theta_p)$ with posterior $\pi(\theta | y)$.
- If we can draw from $\pi(\theta_j | \theta_{-j}, y)$ for each j , then the Gibbs sampler iteratively updates components one at a time.
- The resulting Markov chain has stationary distribution

$$\pi(\theta | y).$$

- **Convergence:** Under mild conditions (irreducibility, aperiodicity), the chain converges to the joint posterior, regardless of starting values.
- **Ergodic averages:** For any function $h(\theta)$,

$$\frac{1}{T} \sum_{t=1}^T h(\theta^{(t)}) \rightarrow \mathbb{E}[h(\theta) | y], \quad \text{as } T \rightarrow \infty.$$

Theoretical Foundation of Gibbs Sampling

Key idea: Instead of sampling from the joint posterior directly, sample sequentially from each *full conditional* distribution.

- Suppose parameters are $\theta = (\theta_1, \dots, \theta_p)$ with posterior $\pi(\theta | y)$.
- If we can draw from $\pi(\theta_j | \theta_{-j}, y)$ for each j , then the Gibbs sampler iteratively updates components one at a time.
- The resulting Markov chain has stationary distribution

$$\pi(\theta | y).$$

- **Convergence:** Under mild conditions (irreducibility, aperiodicity), the chain converges to the joint posterior, regardless of starting values.
- **Ergodic averages:** For any function $h(\theta)$,

$$\frac{1}{T} \sum_{t=1}^T h(\theta^{(t)}) \rightarrow \mathbb{E}[h(\theta) | y], \quad \text{as } T \rightarrow \infty.$$

Takeaway: Gibbs sampling is exact in the limit — we only need closed-form full conditionals.

Gibbs Sampler: Algorithm

Initialization: Choose $(\alpha^{(0)}, \beta^{(0)}, \tau_\epsilon^{2(0)})$. Pre-compute $S_x = \sum x_i$, $S_{xx} = \sum x_i^2$, $S_y = \sum y_i$, $S_{xy} = \sum x_i y_i$.

For $t = 1, 2, \dots, T$ **do:**

- ① **Sample** α : $\alpha^{(t)} \sim N(\tilde{\mu}_\alpha^{(t)}, 1/\tilde{\tau}_\alpha^{2(t)})$.

$$\tilde{\tau}_\alpha^{2(t)} = n \tau_\epsilon^{2(t-1)} + \tau_\alpha^2, \quad \tilde{\mu}_\alpha^{(t)} = \frac{\tau_\epsilon^{2(t-1)} (S_y - \beta^{(t-1)} S_x)}{\tilde{\tau}_\alpha^{2(t)}},$$

- ② **Sample** β : $\beta^{(t)} \sim N(\tilde{\mu}_\beta^{(t)}, 1/\tilde{\tau}_\beta^{2(t)})$.

$$\tilde{\tau}_\beta^{2(t)} = \tau_\epsilon^{2(t-1)} S_{xx} + \tau_\beta^2, \quad \tilde{\mu}_\beta^{(t)} = \frac{\tau_\epsilon^{2(t-1)} (S_{xy} - \alpha^{(t)} S_x)}{\tilde{\tau}_\beta^{2(t)}},$$

- ③ **Sample** τ_ϵ^2 : $\tau_\epsilon^{2(t)} \sim G(a_\epsilon + \frac{n}{2}, b_\epsilon + \frac{1}{2} \text{SSE}^{(t)})$.

$$\text{SSE}^{(t)} = \sum_{i=1}^n (y_i - \alpha^{(t)} - \beta^{(t)} x_i)^2,$$

Return: Discard first B iterations (burn-in) and use the rest for posterior summaries.

SSE via Summary Statistics

Recall

$$\text{SSE}(\alpha, \beta) = \sum_{i=1}^n (y_i - \alpha - \beta x_i)^2.$$

With the precomputed sums

$$S_x = \sum x_i, \quad S_{xx} = \sum x_i^2, \quad S_y = \sum y_i, \quad S_{xy} = \sum x_i y_i, \quad S_{yy} = \sum y_i^2,$$

we can compute, for any (α, β) ,

$$\boxed{\text{SSE}(\alpha, \beta) = S_{yy} - 2\alpha S_y - 2\beta S_{xy} + 2\alpha\beta S_x + n\alpha^2 + \beta^2 S_{xx}}.$$

This avoids an $O(n)$ pass each iteration.

Gibbs Sampler: Algorithm (with fast SSE)

Initialization: Choose $(\alpha^{(0)}, \beta^{(0)}, \tau_\epsilon^{2(0)})$. Pre-compute $S_x, S_{xx}, S_y, S_{xy}, S_{yy}$.

For $t = 1, 2, \dots, T$ **do:**

- ① **Sample α as before.**
- ② **Sample β as before.**
- ③ **Sample τ_ϵ^2 :**

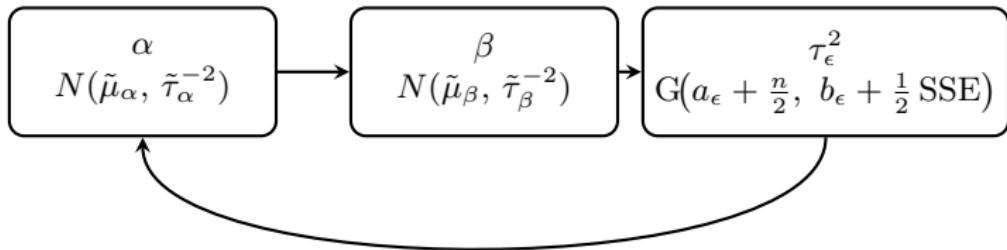
$$\text{SSE}^{(t)} = S_{yy} - 2\alpha^{(t)}S_y - 2\beta^{(t)}S_{xy} + 2\alpha^{(t)}\beta^{(t)}S_x + n(\alpha^{(t)})^2 + (\beta^{(t)})^2S_{xx},$$

$$\tau_\epsilon^{2(t)} \sim G\left(a_\epsilon + \frac{n}{2}, b_\epsilon + \frac{1}{2}\text{SSE}^{(t)}\right).$$

Return: Discard burn-in B and summarize the rest.

Gibbs Sampler: Practical View

In simple linear regression: each step corresponds to a familiar distribution.



Cycle:

$$\alpha^{(t)} \rightarrow \beta^{(t)} \rightarrow \tau_\epsilon^{2(t)} \rightarrow \alpha^{(t+1)} \rightarrow \dots$$

Note: All steps are *direct draws* from known forms (Normal or Gamma)

Posterior Summaries and Prediction

Point/interval estimates: For any parameter $\theta \in \{\alpha, \beta, \tau_\epsilon^2\}$,

$$\hat{\theta}_{\text{post-mean}} = \frac{1}{T-B} \sum_{t=B+1}^T \theta^{(t)}, \quad (1-\gamma) \text{ CI from quantiles of } \{\theta^{(t)}\}.$$

Posterior predictive for a new x_* :

$$y_* \mid x_*, \alpha, \beta, \tau_\epsilon^2 \sim N(\alpha + \beta x_*, 1/\tau_\epsilon^2).$$

Generate $y_*^{(t)} \sim N(\alpha^{(t)} + \beta^{(t)} x_*, 1/\tau_\epsilon^{2(t)})$ and summarize $\{y_*^{(t)}\}$.

Monte Carlo Estimators (from Gibbs samples)

Let $\{(\alpha^{(t)}, \beta^{(t)}, \tau_\epsilon^{2(t)})\}_{t=B+1}^T$ be the retained Gibbs draws (after burn-in B). Define $\sigma_\epsilon^{2(t)} = 1/\tau_\epsilon^{2(t)}$.

Posterior means (sample averages):

$$\hat{\alpha} = \frac{1}{T-B} \sum_{t=B+1}^T \alpha^{(t)}, \quad \hat{\beta} = \frac{1}{T-B} \sum_{t=B+1}^T \beta^{(t)}, \quad \hat{\sigma}_\epsilon^2 = \frac{1}{T-B} \sum_{t=B+1}^T \sigma_\epsilon^{2(t)}.$$

Posterior credible intervals (componentwise): For $\theta \in \{\alpha, \beta, \sigma_\epsilon^2\}$,

$$100(1 - \gamma)\% \text{ CI for } \theta = [\text{Quantile}_{\gamma/2}\{\theta^{(t)}\}, \text{Quantile}_{1-\gamma/2}\{\theta^{(t)}\}].$$

Posterior probability of positive slope:

$$P(\beta > 0 \mid y, x) \approx \frac{1}{T-B} \sum_{t=B+1}^T \mathbf{1}\{\beta^{(t)} > 0\}.$$

Predictive Inference via Monte Carlo

For a new patient with dose x_* :

$$y_*^{(t)} \sim N(\alpha^{(t)} + \beta^{(t)}x_*, \sigma_{\epsilon}^{2(t)}), \quad t = B+1, \dots, T.$$

Posterior predictive mean and interval:

$$\widehat{\mathbb{E}[y_* | x_*, y, x]} = \frac{1}{T-B} \sum_{t=B+1}^T y_*^{(t)}, \quad 100(1-\gamma)\% \text{ PI} = [\mathbf{Q}_{\gamma/2}\{y_*^{(t)}\}, \mathbf{Q}_{1-\gamma/2}\{y_*^{(t)}\}].$$

Probability of clinically relevant reduction (e.g., ≥ 5 mmHg at $x_* = 50$ mg):

$$P(y_* \geq 5 | x_* = 50, y, x) \approx \frac{1}{T-B} \sum_{t=B+1}^T \mathbf{1}\{y_*^{(t)} \geq 5\}.$$

Notes: Report Monte Carlo SEs if desired; always check trace plots and effective sample sizes.

One-Slide Summary of Posterior Inference

- **Slope (dose effect):** $\widehat{\beta} \approx 0.080$, 95% CI = [0.080, 0.0803], $P(\beta > 0 | y, x) = 1.00$.
- **Residual variance:** $\widehat{\sigma}_\epsilon^2 \approx 24.9$, 95% CI = [24.7, 25.1].
- **Prediction at 50 mg:** $\widehat{E[y_* | x_*=50]} \approx 6.07$, 95% PI = [-3.73, 16.0],
 $P(y_* \geq 5) = 0.586$.

Computation rule: All estimates are based on Monte Carlo averages and quantiles of Gibbs samples.