Bayesian Computation

Gwangsu Kim

JBNU

Second semester of 2024

How to use Posteriors I

- In Bayesian analysis, posteriors are distribution. It means that we have the probability of posterior.
- But it cannot be easy to obtain the credible interval of posterior means althogh we know the density.

How to use Posteriors II

- Example:

$$\pi(\beta|y,x) \propto \left(\frac{\exp(\beta x)}{1+\exp(\beta x)}\right)^y \phi(\beta; 0, \sigma_0^2).$$

- This requires the general methods to obtain the random samples of postriors.

Numerical Integration I

- Simulation methods

$$E\Big[h(\theta)\Big|y\Big] = \int h(\theta)\pi(\theta|y)d\theta \approx \frac{1}{S}\sum_{s=1}^{S} h(\theta^s),$$

where θ^ss are from random samples follwong the law of $\pi(\theta|y)$.

- It requires a large sample szie to improve the accuracy.

Numerical Integration II

- Deterministic methods

$$E\Big[h(\theta)\Big|y\Big] = \int h(\theta)\pi(\theta|y)d\theta \approx \frac{1}{S}\sum_{s=1}^{S} w_s h(\theta^s)p(\theta^s|y),$$

where w_s and θ^ss are weights (volume) and points, respectively.

- As dimesions become large, complexity of calculations grows very fast.

Rejection Sampling and Inportance Sampling I

- Rejection Sampling: in the previous lecture slide.

- Importance Sampling

  1. Note

$$E[h(\theta|y)] = \frac{\int h(\theta)q(\theta|y)d\theta}{\int q(\theta|y)d\theta} = \frac{\int \left[h(\theta)q(\theta|y)/g(\theta)\right] g(\theta)d\theta}{\int \left[q(\theta|y)/g(\theta)\right] g(\theta)d\theta}.$$

Rejection Sampling and Inportance Sampling II

2 Then use the

$$\frac{\frac{1}{S}\sum_{s=1}^{S} h(\theta^s)w(\theta^s)}{\frac{1}{S}\sum_{s=1}^{S} w(\theta^s)},$$

where $w(\theta^s) = \frac{q(\theta^s|y)}{g(\theta^s)}$ and θ^ss are from random sample following the law of $g(\theta)$.

Rejection Sampling and Inportance Sampling III

3 The efficiency of this sampling is measured by

$$\frac{1}{\sum_{s=1}^{S}(\tilde{w}(\theta^s))^2},$$

where $\tilde{w}(\theta^s) = w(\theta^s)S/\sum_{s'=1}^{S} w(\theta^{s'})$.

- It means that large variation of $\tilde{w}(\theta^s)$ gives the low efficiency.