

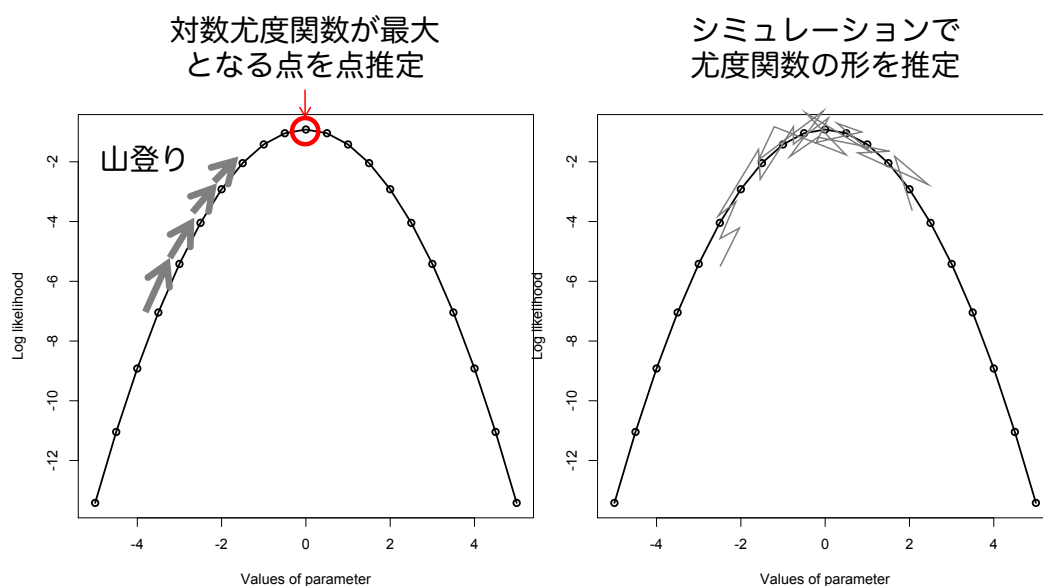
ベイズ統計

古谷知之

講義概要

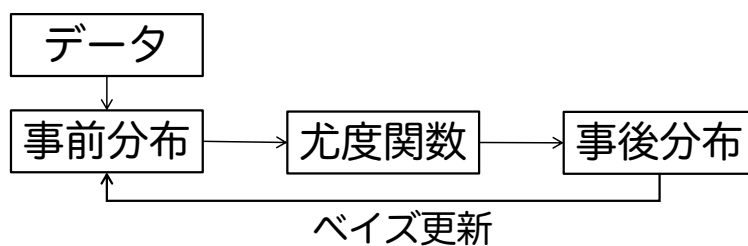
- シミュレーションによる正規分布の推定
- マルコフ連鎖モンテカルロ法
- メトロポリス法による正規分布のシミュレーション
- ギブス・サンプリング
- 二変量正規分布のギブス・サンプリング

最尤法とベイズ法の違い



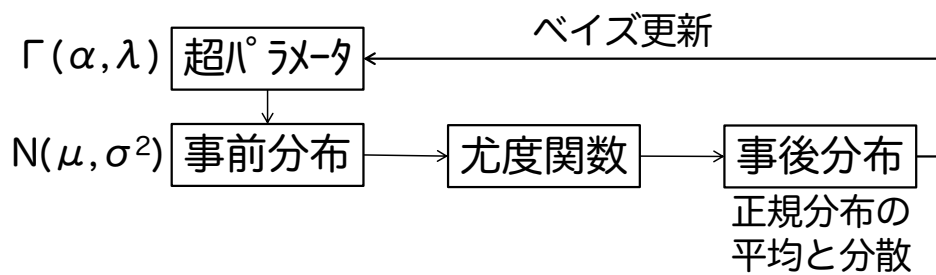
経験ベイズ推定

- 事前分布をデータから与える方法を経験ベイズ(empirical Bayes)推定という
- 事前分布のパラメータを超パラメータ(hyper parameter)という
 - 事前分布の分散(精度)など

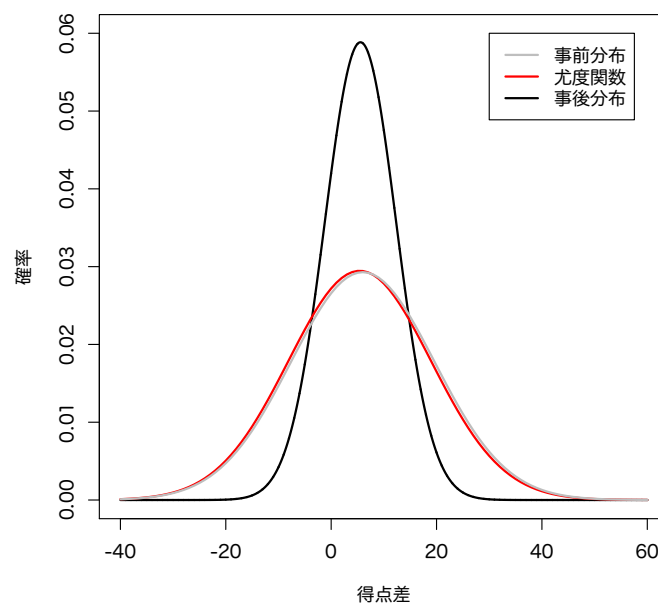


階層ベイズ推定

- 事前分布をベイズ推定する方法を階層ベイズ (hierarchical Bayes) 推定という
- 事前分布の超パラメータをベイズ推定する
 - 事前正規分布の分散(逆ガンマ分布に従う)の超パラメータのベイズ推定



正規分布の推定



ベイズ更新の繰り返し計算

- 何回繰り返せばいいのか？
- 計算結果の信頼性は？
- 効率的な計算方法は？ → 今日はココ

マルコフ連鎖モンテカルロ法

- Markov-chain Monte Carlo (MCMC)
- マルコフ連鎖
 - 一つ前の状態から次の状態を決める方法



- モンテカルロ法
 - 乱数を生成(サンプリング)する方法
- ベイズ統計ではシミュレーションにより事後分布をサンプリングする際に用いる

マルコフ連鎖モンテカルロ法

Metropolis–Hasting法

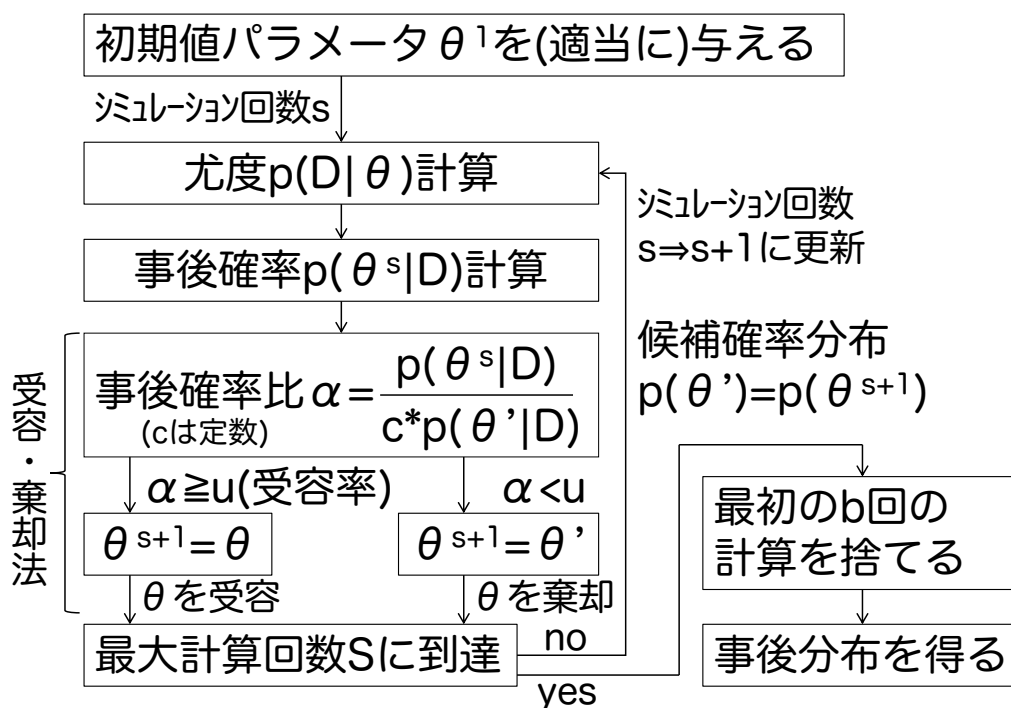
ギブス・サンプリング (Gibbs sampler)

- ・ 完全な条件付き確率の集合から順々に標本を取り、緩い条件下で安定した事後分布を持つマルコフ連鎖を得る
- ・ 未知パラメータ2個以上

Metropolis法 (酔歩法)

- ・ 完全な条件付き確率から順々に標本を得るのが難しい時に用いる
- ・ 事後分布の候補分布と提案分布の比を計算し受容するか棄却するかを判断

MCMCによるベイズ推定のイメージ

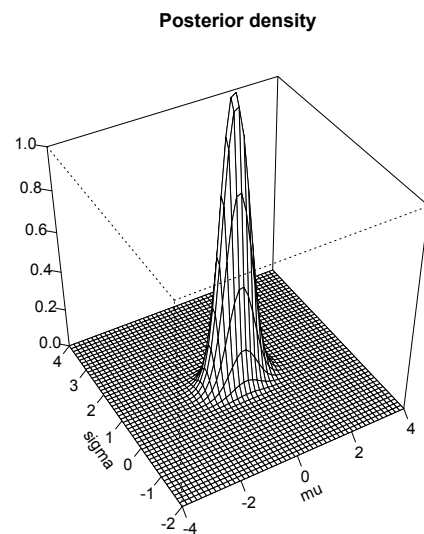
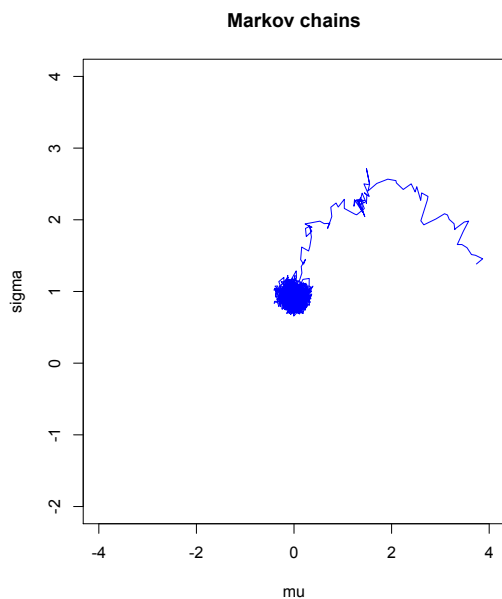


MCMCの業界用語

- シミュレーションする = 標本を得る
- サンプルング = (標本を)生成する = シミュレーションや事後分布の結果（の一部）を取り出す
- burn-in(バーン・イン、稼働検査期間) = シミュレーションを捨てる回数
- 繰り返し計算回数のことを、stepとかdrawなどという
- 同時実行されるMCMC数をchain数という

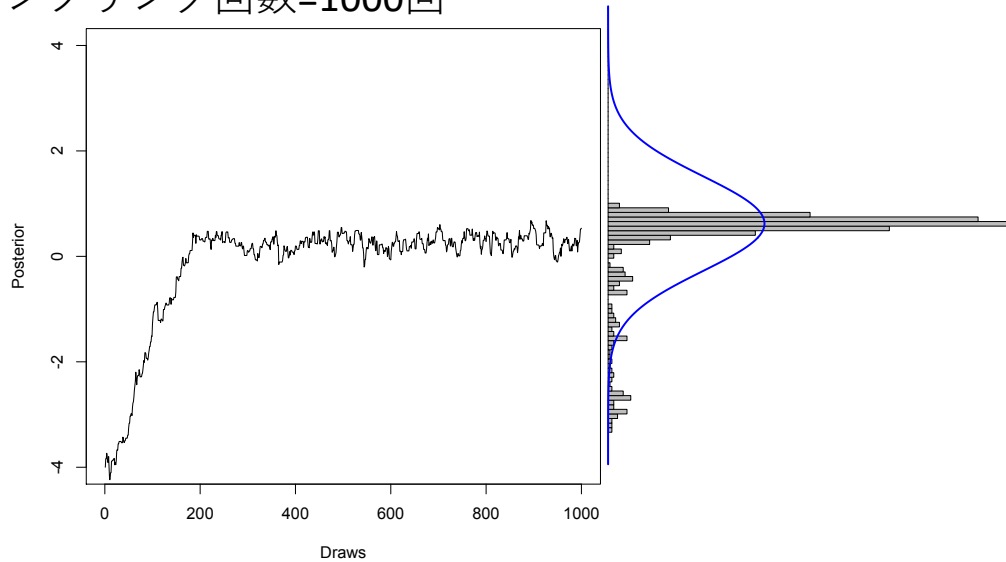
メトロポリス法のイメージ

- ムービー



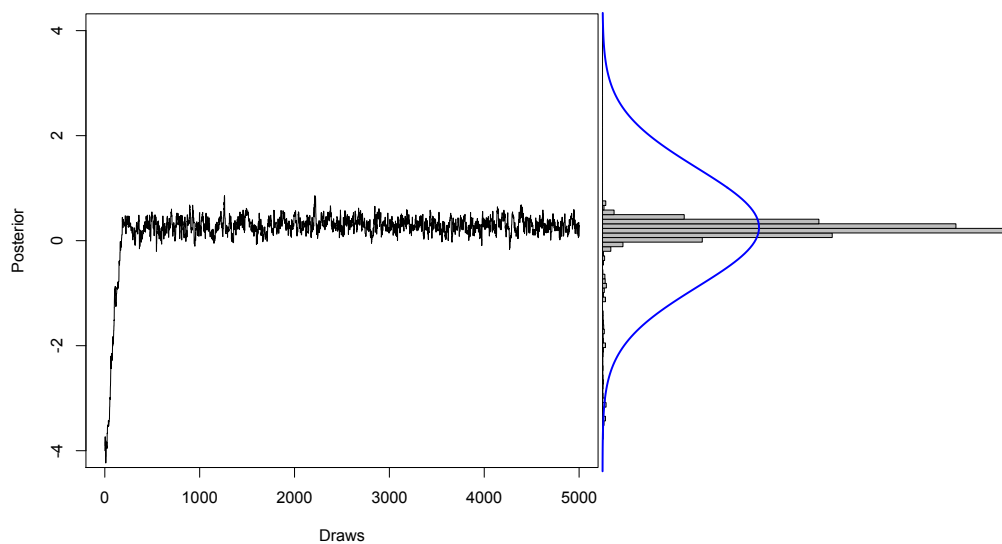
メトロポリス法によるサンプリング(1)

- サンプリング回数=1000回



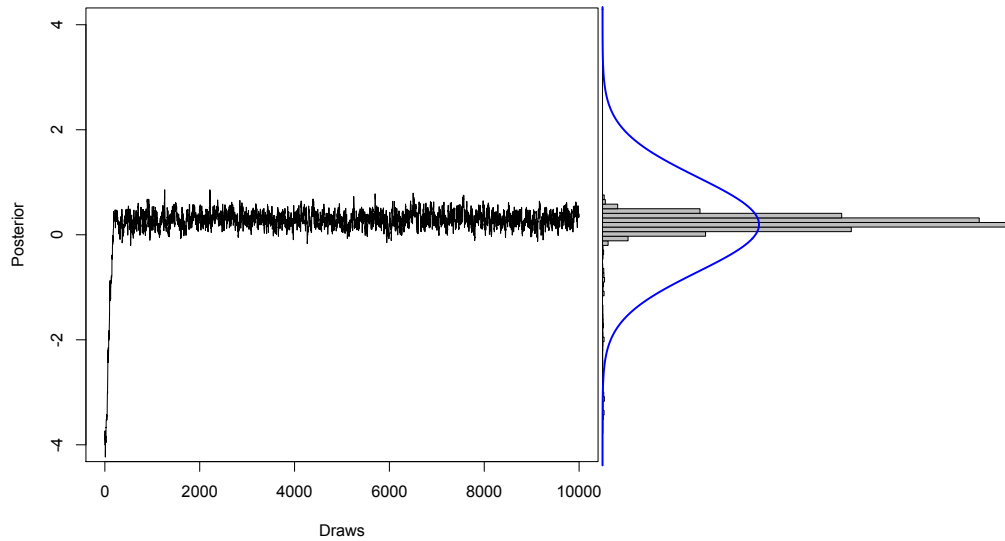
メトロポリス法によるサンプリング(2)

- サンプリング回数=5000回

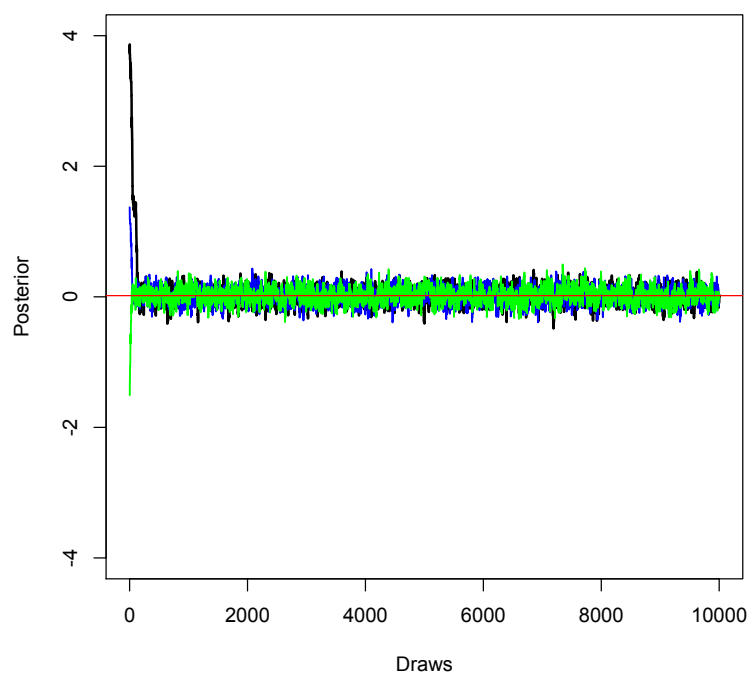


メトロポリス法によるサンプリング(3)

- サンプリング回数=10000回



メトロポリス法によるサンプリング(4)



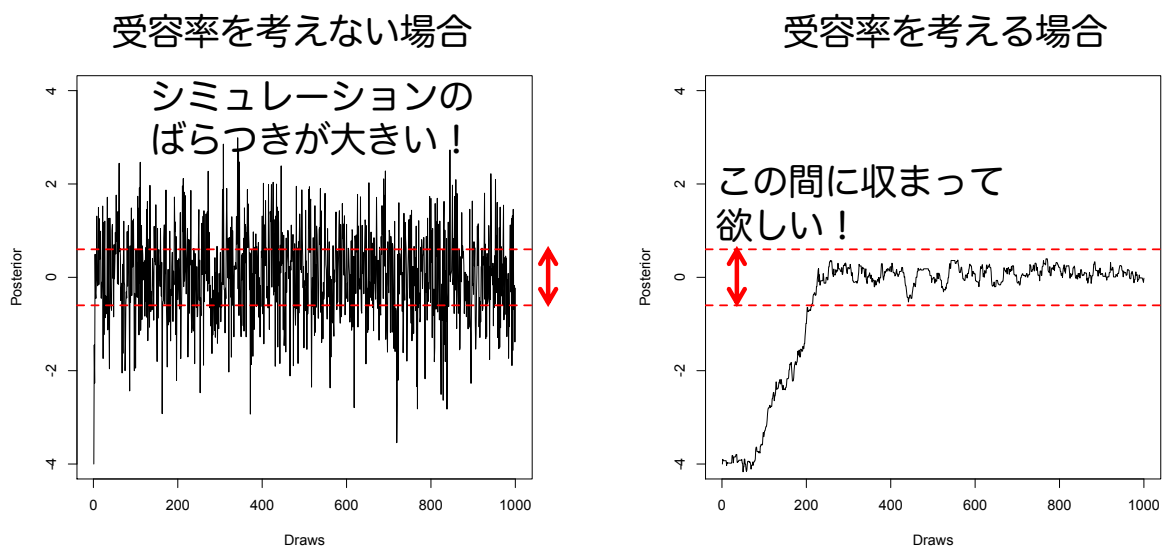
棄却サンプリングと受容率

- $p(\theta^s|D)$ から乱数発生が難しいとき
- 正規分布など乱数発生が容易な候補確率 $p(\theta')$ から乱数を発生する
- 区間 $[0,1]$ 上で受容率 u の一様乱数を生成
- 受容率を下回る事後確率は棄却サンプリングとして捨てる

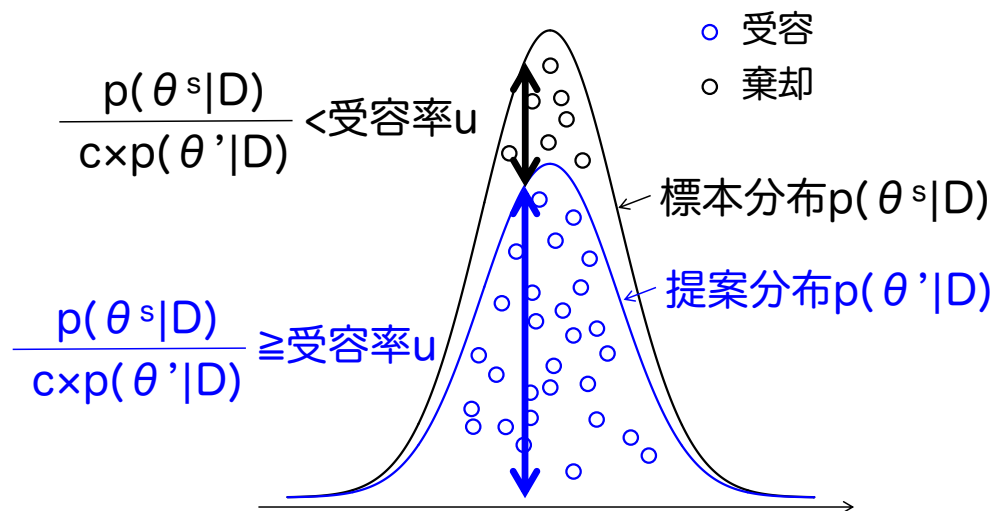
$$\text{事後確率比 } \alpha = \frac{p(\theta^s|D)}{c \times p(\theta')} \geq \text{受容率 } u$$

(c は定数)

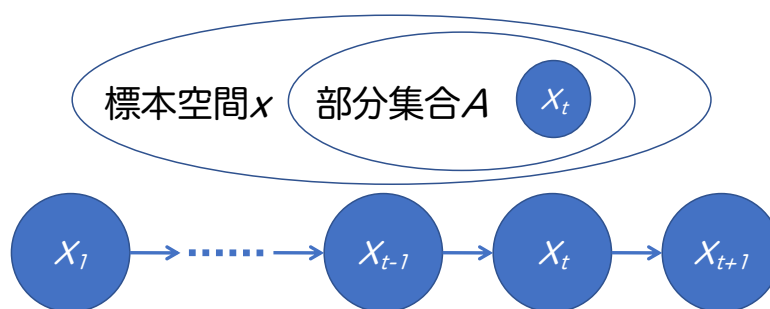
棄却サンプリングと受容率



受容-棄却法 (acceptance-rejection)



マルコフ連鎖によるシミュレーション

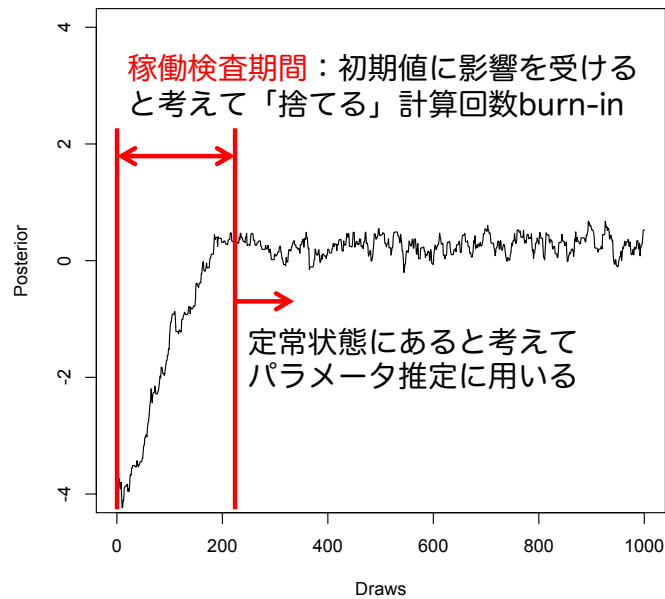


推移確率

$$\frac{P(X_{t+1} \in A | X_1, \dots, X_t)}{\text{事後確率}} = \frac{P(X_{t+1} | X_t) \cdot P(X_t | X_{t-1}) \cdots P(X_2 | X_1) \cdot P(X_1)}{\text{推移核} \quad \text{初期値}}$$

推移核が初期値(事前確率)の影響を受けないとき、
マルコフ連鎖は**定常状態**にあるという

マルコフ連鎖によるシミュレーション



推移核

- 推移核：状態 x から状態 y への遷移を以下のように表す

$$X_t = x \rightarrow X_{t+1} = y$$

$$K(x, y) = P(X_{t+1} = y | X_t = x)$$

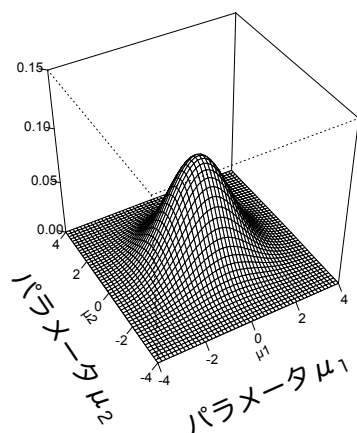
- メトロポリス・ヘイスティングス(M-H)法もギブス・サンプリングも、この推移核をどう考えるかの違いに過ぎない

ギブス・サンプリング

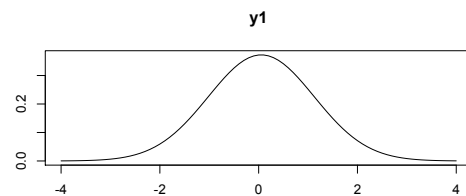
- 目標分布：最終的に求めたい事後分布(の標本分布)
- 目標分布を2つ以上に分割できると考える。例えば、 $y = (y_1, y_2)$
- このとき求める事後分布は、 $p(y_1|y_2)$ と $p(y_2|y_1)$

ギブス・サンプリング

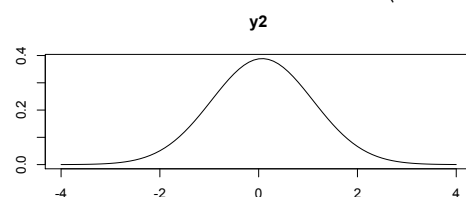
- 例えば、二変量正規分布を考える
2つの未知パラメータに対する
目標分布（求めたい分布）



事後分布 $p(y_1|y_2) \sim N(\mu_1, \sigma_1^2)$



事後分布 $p(y_2|y_1) \sim N(\mu_2, \sigma_2^2)$



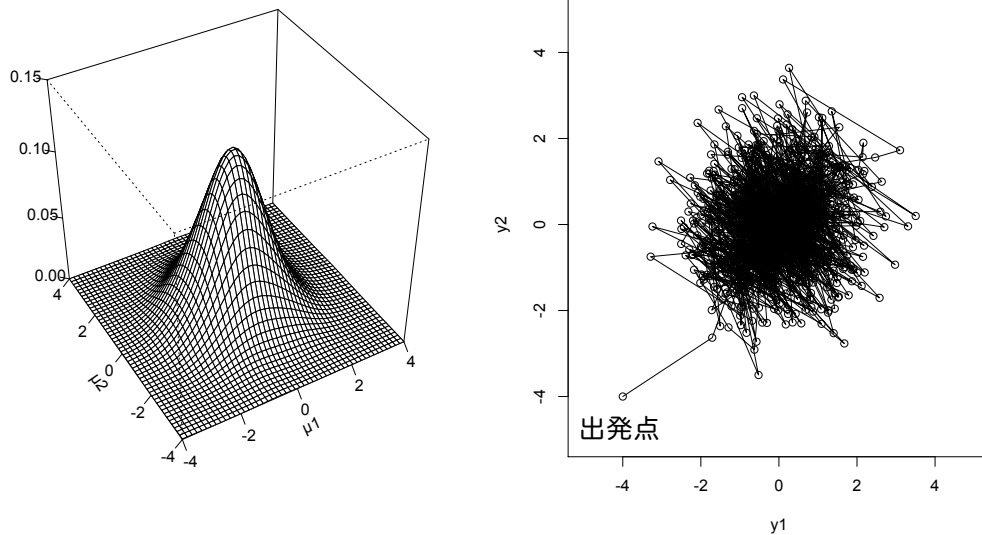
ギブス・サンプリング

- 1) 計算回数 $s = (0, \dots, S)$ を決める
- 2) $s = 0$ に対し初期値 $x^s = (x_1^s, \dots, x_n^s)$ を決める
- 3) x^s から x^{s+1} を生成する
 - 1) $x_1^{s+1} \sim p(x_1^s | x_2^s, \dots, x_n^s)$ から x_1^{s+1} を生成
 - 2) $x_2^{s+1} \sim p(x_2^s | x_1^{s+1}, \dots, x_n^s)$ から x_2^{s+1} を生成
 - $\vdots$
 - n) $x_n^{s+1} \sim p(x_n^s | x_1^{s+1}, \dots, x_{n-1}^{s+1})$ から x_n^{s+1} を生成
- 4) $s = S$ のとき計算終了

二変量正規分布のギブス・サンプリング(1)

- 2つの正規分布 $y_1 \in Y_1$ と正規分布 $y_2 \in Y_2$
- 事前分布として初期値を適当に与える
 - Y_1 の平均 μ_1 、標準偏差 ρ : $y_1 \sim N(\mu_1, \rho)$
 - Y_2 の平均 μ_2 、標準偏差 ρ : $y_2 \sim N(\mu_2, \rho)$
 - (y_1, y_2) の初期値 (y_1^0, y_2^0) 、事後標準偏差の初期値 (σ_1, σ_2)
- (y_1^s, y_2^s) から (y_1^{s+1}, y_2^{s+1}) を生成する
$$y_2^s \sim N\left(\mu_2 + \rho \frac{\sigma_2}{\sigma_1} (y_1^s - \mu_1), \sigma_2^2 (1 - \rho^2)\right)$$
$$y_1^{s+1} \sim N\left(\mu_1 + \rho \frac{\sigma_1}{\sigma_2} (y_2^{s+1} - \mu_2), \sigma_1^2 (1 - \rho^2)\right)$$

二変量正規分布のギブス・サンプリング(2)



正規母集団同時事後分布のギブス・サンプリング(1)

- ある正規母集団 Y から観測値 $y \in Y$ (標本数 n) を得た場合に、平均 μ と分散 σ^2 の事後分布をギブス・サンプリングから得たい
- 事前分布
 - 平均: $\mu \sim N(\mu_0, \sigma_0^2)$
 - 分散(精度): $\tau \sim 1/\sigma_0^2 \sim Ga(v_0/2, S_0/2)$
⇒ 正規分布の分散は逆ガンマ分布、精度はガンマ分布に従う
- 尤度関数

$$p(y|\mu, \sigma^2) \sim N(\mu, \sigma^2)$$

正規分布の平均のベイズ推定

- 平均 μ の事後分布

$$\begin{aligned}f(\mu|Y) &\propto \prod_{i=1}^n f(y_i|\mu, \sigma^2) \cdot \pi(\mu) \\&\propto \exp\left(-\frac{\sum_{i=1}^n (y_i - \mu)^2}{2\sigma^2}\right) \cdot \exp\left(-\frac{(\mu - \mu_0)^2}{2\sigma_0^2}\right) \\&= \exp\left(-\frac{(\mu - \mu_1)^2}{2\sigma_1^2}\right) \\ \mu_1 &= \frac{\mu_0/\sigma_0^2 + n\bar{y}/\sigma^2}{1/\sigma_0^2 + n/\sigma^2}, \frac{1}{\sigma_1^2} = \frac{1}{\sigma_0^2} + \frac{n}{\sigma^2}\end{aligned}$$

正規分布の分散のベイズ推定

- 正規分布のベイズ推定における尤度関数

$$\begin{aligned}\prod_{i=1}^n f(y_i|\mu, \sigma^2) &\propto \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_i - \mu)^2}{2\sigma^2}\right) \\&\propto \left(\frac{1}{\sigma^2}\right)^{\frac{n}{2}} \exp\left(-\frac{\sum_{i=1}^n (y_i - \mu)^2}{2\sigma^2}\right)\end{aligned}$$

- ここで $\alpha - 1 = \frac{n}{2}$, $\lambda = \sum_{i=1}^n (y_i - \mu)^2$, $\tau = \frac{1}{\sigma^2}$ と置き換えると、次式が得られる

$$\prod_{i=1}^n f(y_i|\mu, \sigma^2) \propto (\tau)^{\alpha-1} \exp(-\lambda\tau)$$

正規分布の分散のベイズ推定

- 正規分布に従う尤度関数

$$\prod_{i=1}^n f(y_i|\mu, \sigma^2) \propto (\tau)^{\alpha-1} \exp(-\lambda\tau)$$

- これはガンマ関数 $Ga(\alpha, \lambda)$ に比例する

$$Ga(\alpha, \lambda) \propto (\tau)^{\alpha-1} \exp(-\lambda\tau)$$

- 事前分布の精度 τ に対して、自然共役事前分布としてガンマ分布を用いることと同じ意味

正規分布の分散のベイズ推定

- 精度の事前分布

$$\begin{aligned} Ga(\alpha, \lambda) &\propto (\tau)^{\alpha-1} \exp(-\lambda\tau) \\ &\propto Ga\left(n/2 - 1, \sum_{i=1}^n (y_i - \mu)^2/2\right) \\ &\propto Ga\left((n-2)/2, (n-1) \cdot \sum_{i=1}^n (y_i - \mu)^2/(n-1)/2\right) \end{aligned}$$

分散

$$\begin{aligned} &\propto Ga((v_0 + n)/2, v_0 \cdot S_0/2) \\ &\propto Ga((v_0 + n)/2, S_0/2) \end{aligned}$$

この辺は教科書
によって異なる
場合があります

正規母集団同時事後分布のギブス・サンプリング(2)

- 条件付き事後分布
- 事後平均

$$\mu_1 = \frac{\mu_0/\sigma_0^2 + n\bar{y}/\sigma^2}{1/\sigma_0^2 + n/\sigma^2}, \frac{1}{\sigma_1^2} = \frac{1}{\sigma_0^2} + \frac{n}{\sigma^2}$$

- 事後分散

$$\sigma^2 | \mu, y \sim IG(v_1/2, S_1/2)$$

$$v_1 = v_0 + n$$

$$S_1 = S_0 + \sum_{i=1}^n (y_i - \mu)^2$$

正規母集団同時事後分布のギブス・サンプリング(3)

- 条件付き事後分布
- 事後平均

$$\mu_1 = \frac{\sigma_0^2}{\sigma_0^2 + \sigma^2/n} (\bar{y} - n) + \mu_0$$

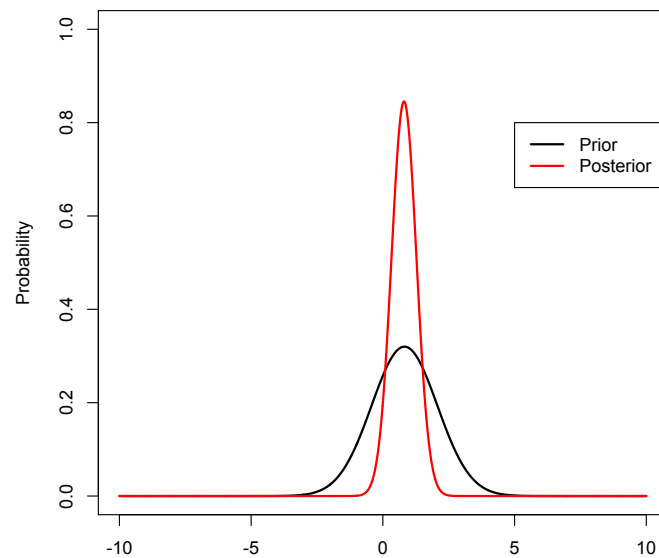
$$\sigma_1^2 = \frac{\sigma_0^2}{\sigma_0^2 + \sigma^2/n} \cdot \frac{\sigma^2}{n}$$

- 事後分散

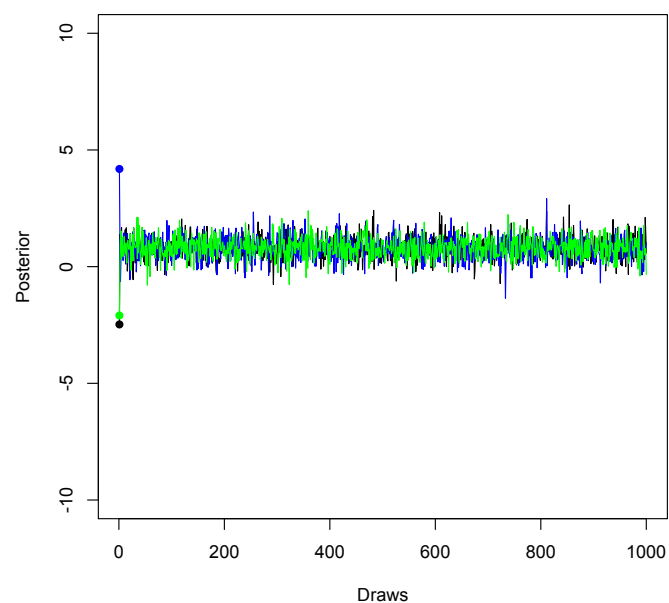
$$\sigma^2 | \mu, y \sim IG(v_1/2, S_1/2)$$

$$v_1 = v_0 + n, S_1 = S_0 + \sum_{i=1}^n (y_i - \mu)^2$$

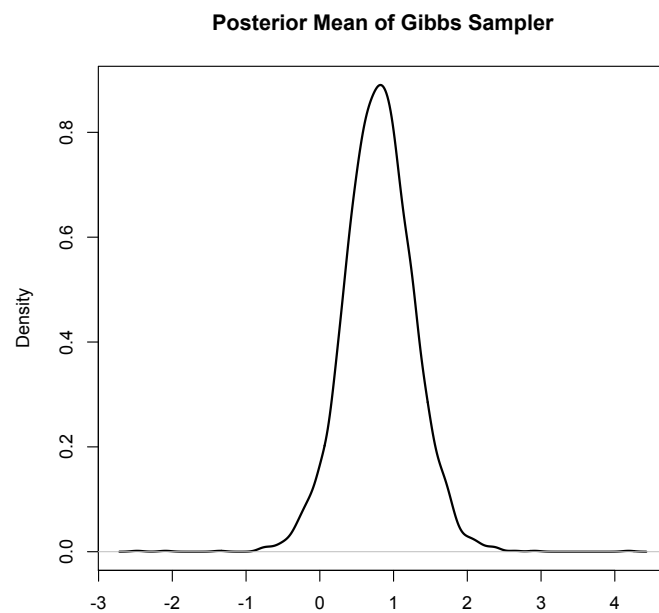
正規母集団同時事後分布のギブス・サンプリング(4)



正規母集団同時事後分布のギブス・サンプリング(5)



正規母集団同時事後分布のギブス・サンプリング(6)



正規母集団同時事後分布のギブス・サンプリング(7)

