ベイズ統計

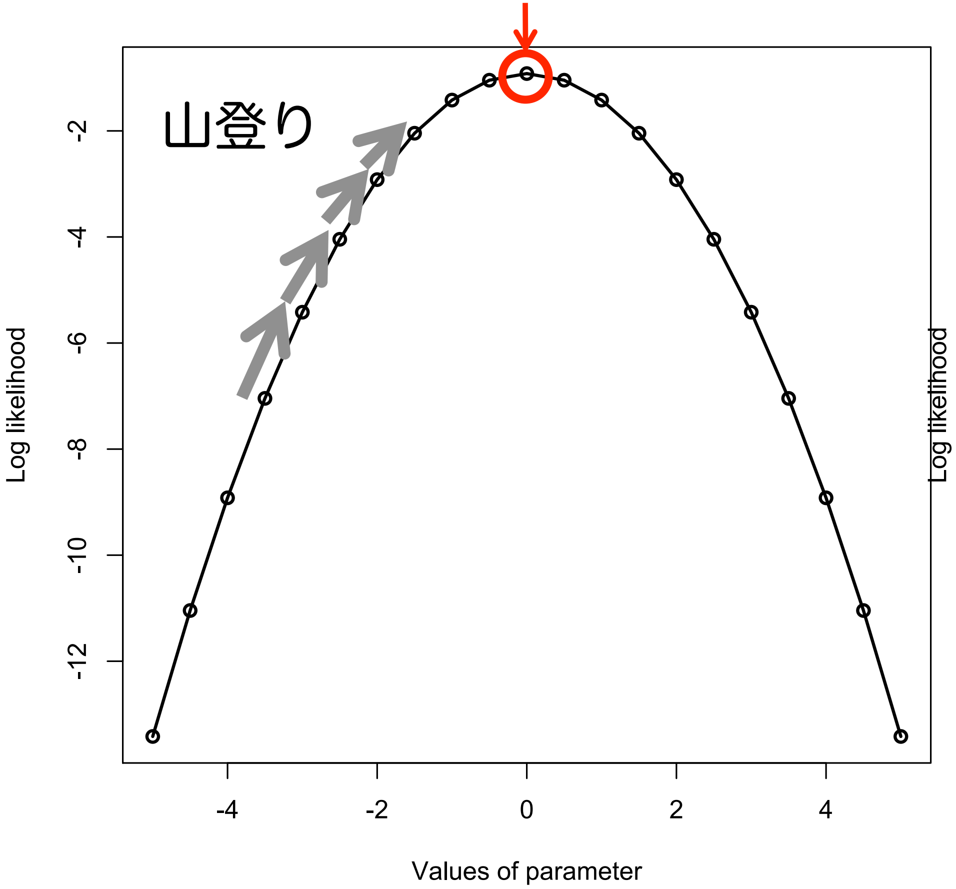
古谷知之

講義概要

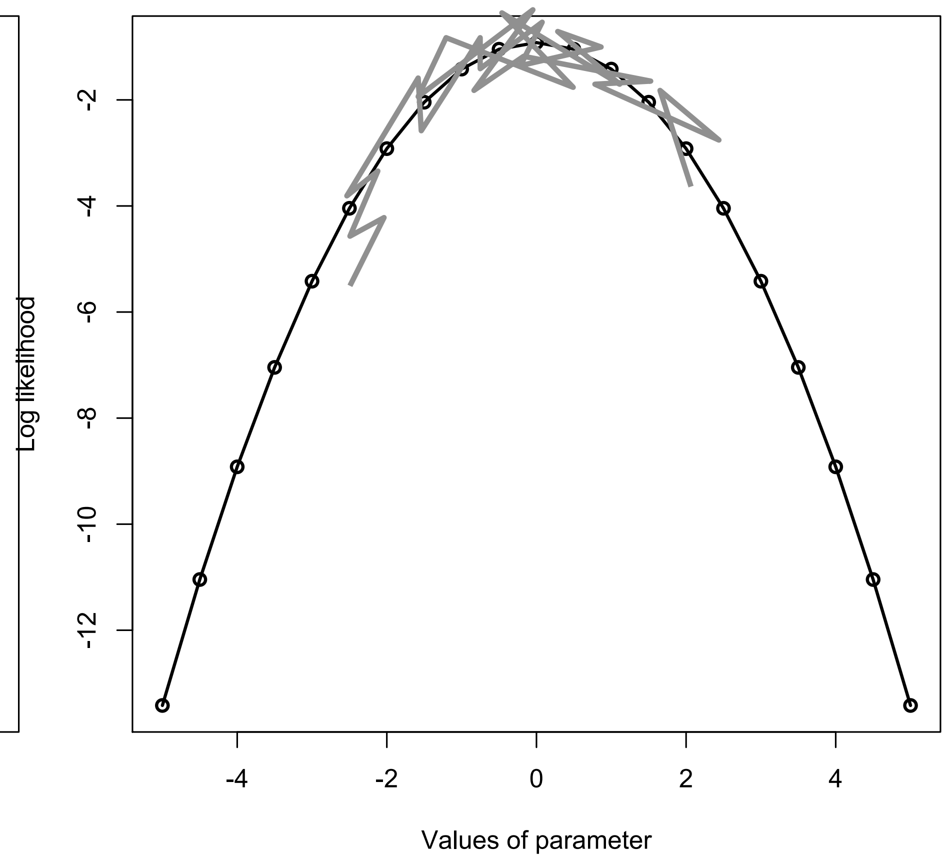
- シミュレーションによる正規分布の推定
- マルコフ連鎖モンテカルロ法
- メトロポリス法による正規分布のシミュレーション
- ギブス・サンプリング
- 二変量正規分布のギブス・サンプリング

最尤法とベイズ法の違い

対数尤度関数が最大
となる点を点推定

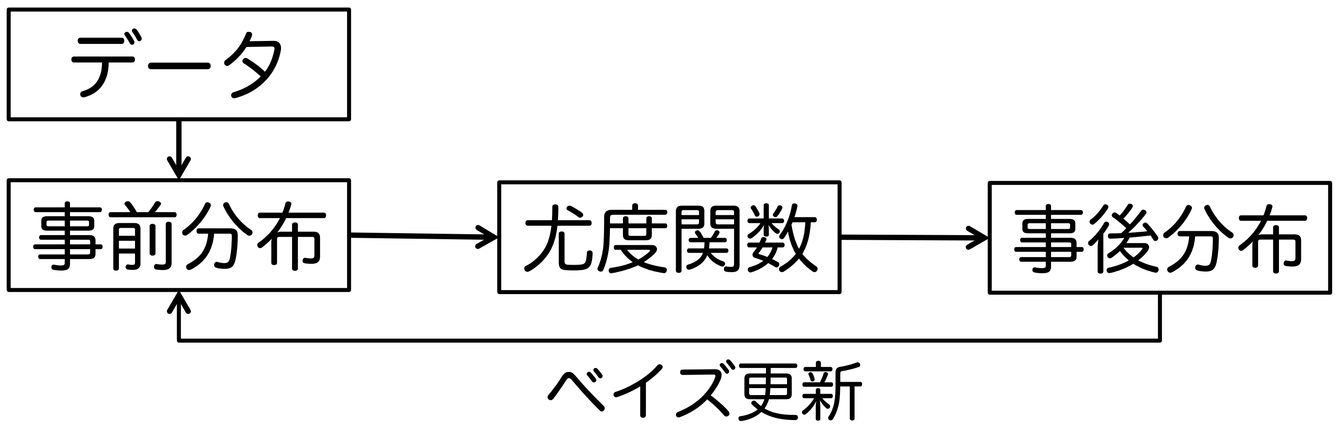


シミュレーションで
尤度関数の形を推定



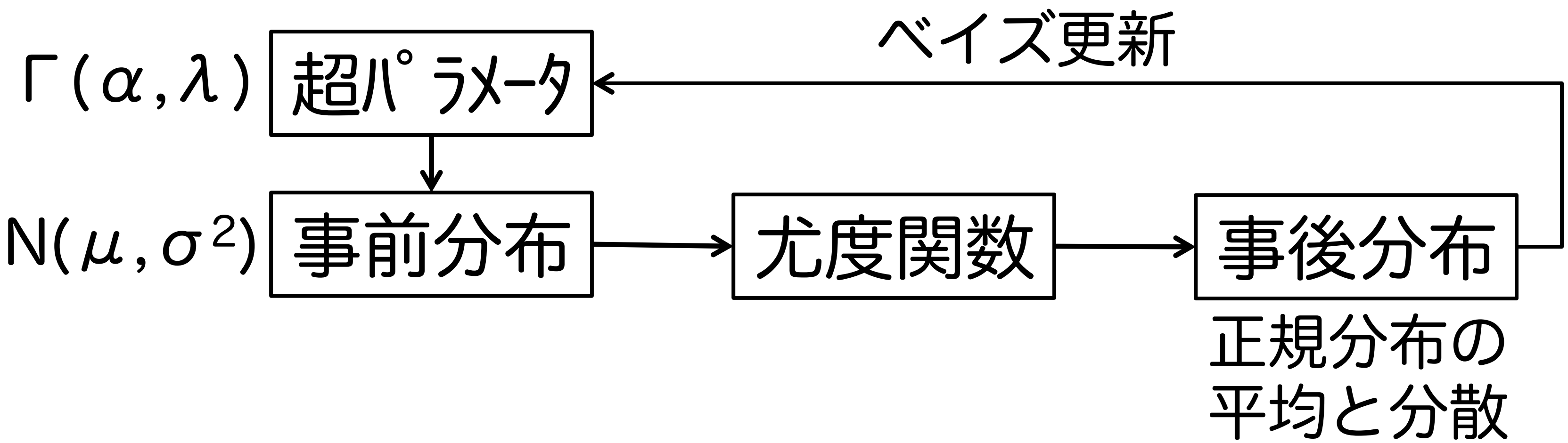
経験ベイズ推定

- 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

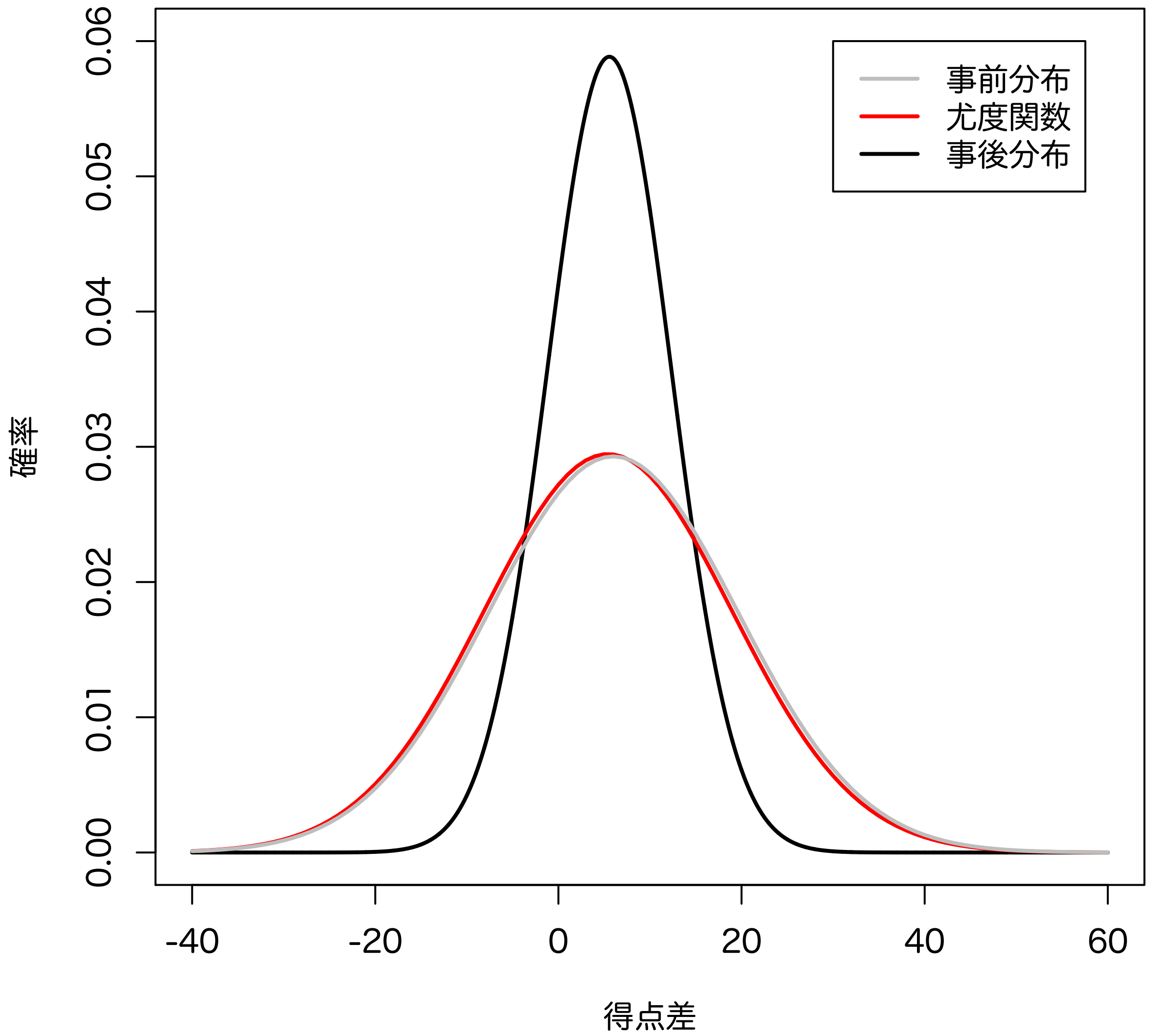


階層ベイズ推定

- 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

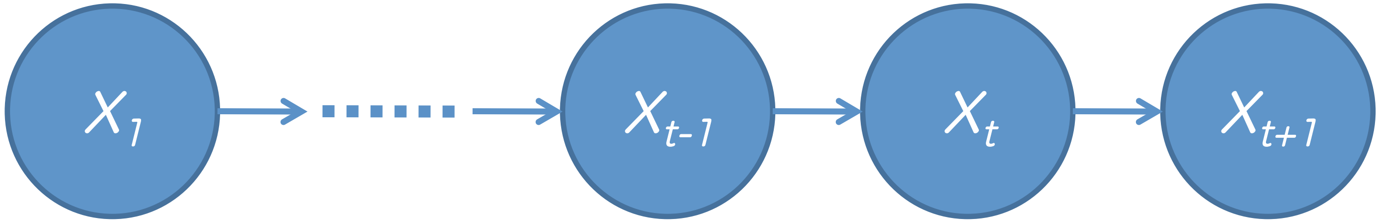


ベイズ更新の繰り返し計算

- 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

・未知ﾊﾟﾗﾒｰﾀ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

初期値パラメータθ^1を(適当に)与える

ｼﾐｭﾚｰｼｮﾝ回数s

尤度$p(D|\theta)$計算

事後確率$p(\theta^s|D)$計算

受容・棄却法

事後確率比$\alpha = \dfrac{p(\theta^s|D)}{c*p(\theta'|D)}$

(cは定数)

$\alpha \geqq u$(受容率)

$\theta^{s+1} = \theta$

θを受容

$\alpha < u$

$\theta^{s+1} = \theta'$

θを棄却

最大計算回数Sに到達

no

yes

ｼﾐｭﾚｰｼｮﾝ回数

s⇒s+1に更新

候補確率分布

$p(\theta') = p(\theta^{s+1})$

最初のb回の

計算を捨てる

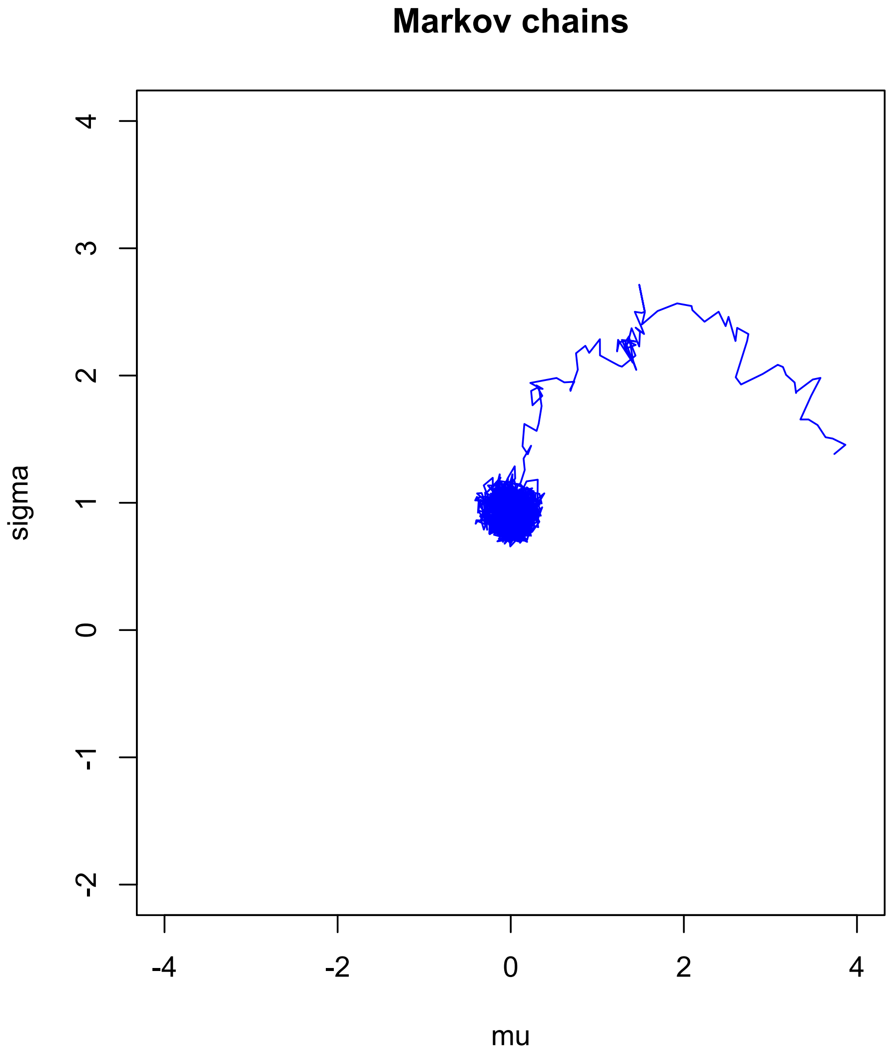
事後分布を得る

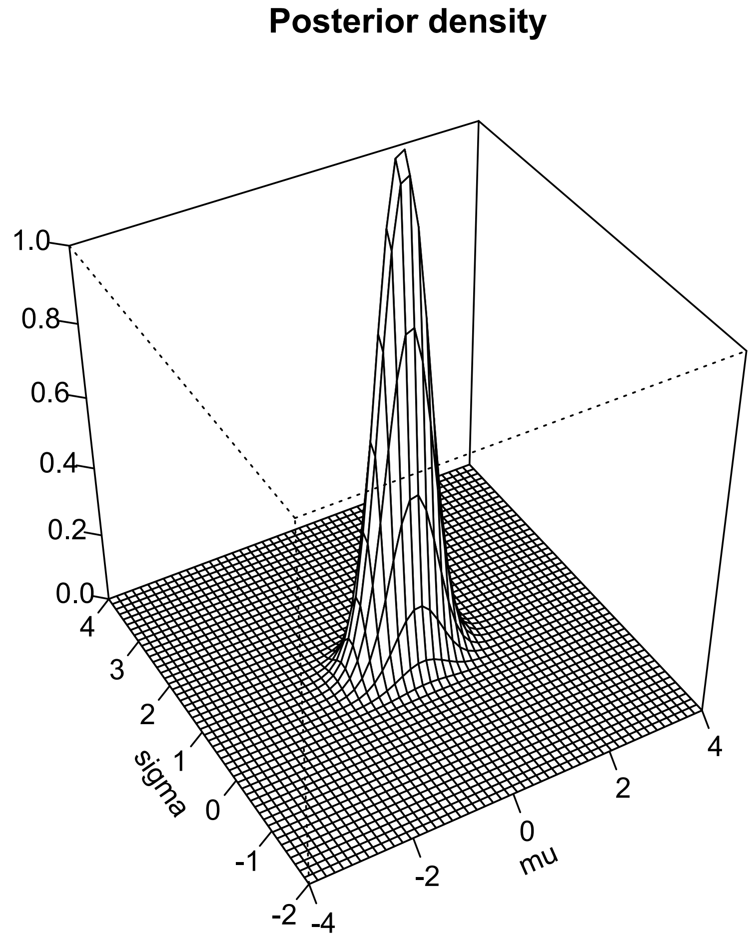
MCMCの業界用語

- シミュレーションする＝標本を得る
- サンプリング＝(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

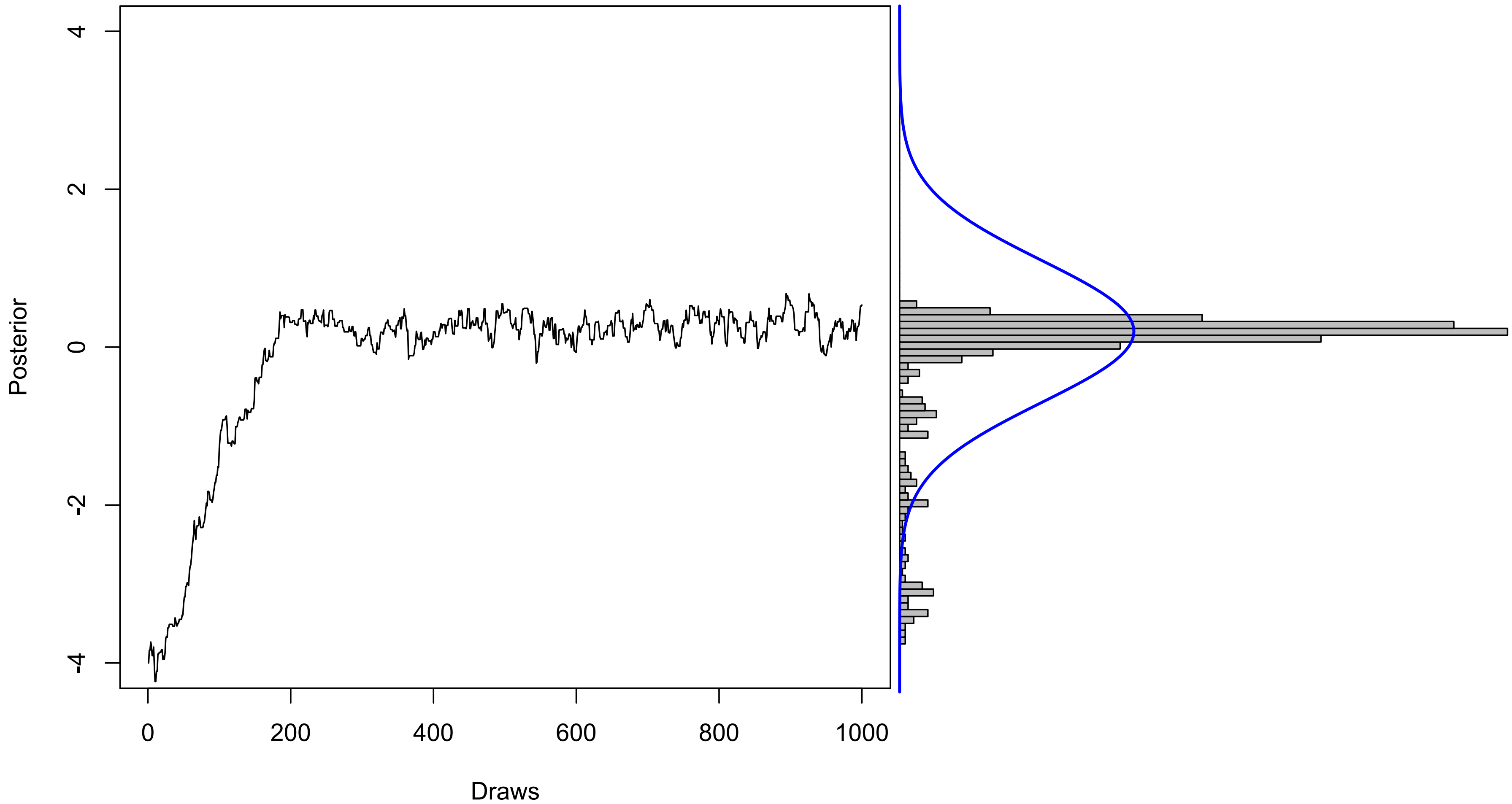
- ムービー





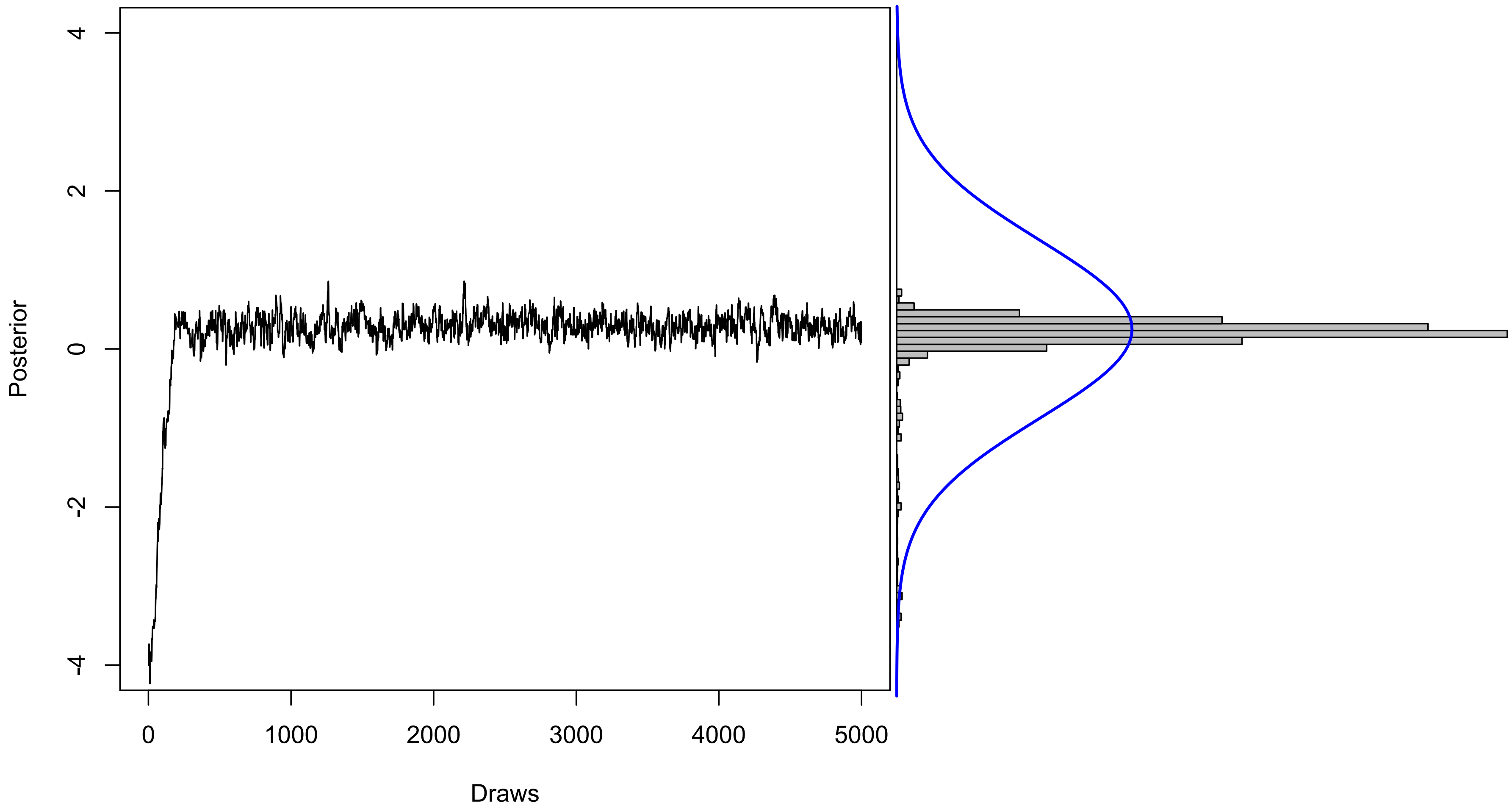
メトロポリス法によるサンプリング(1)

- サンプリング回数=1000回



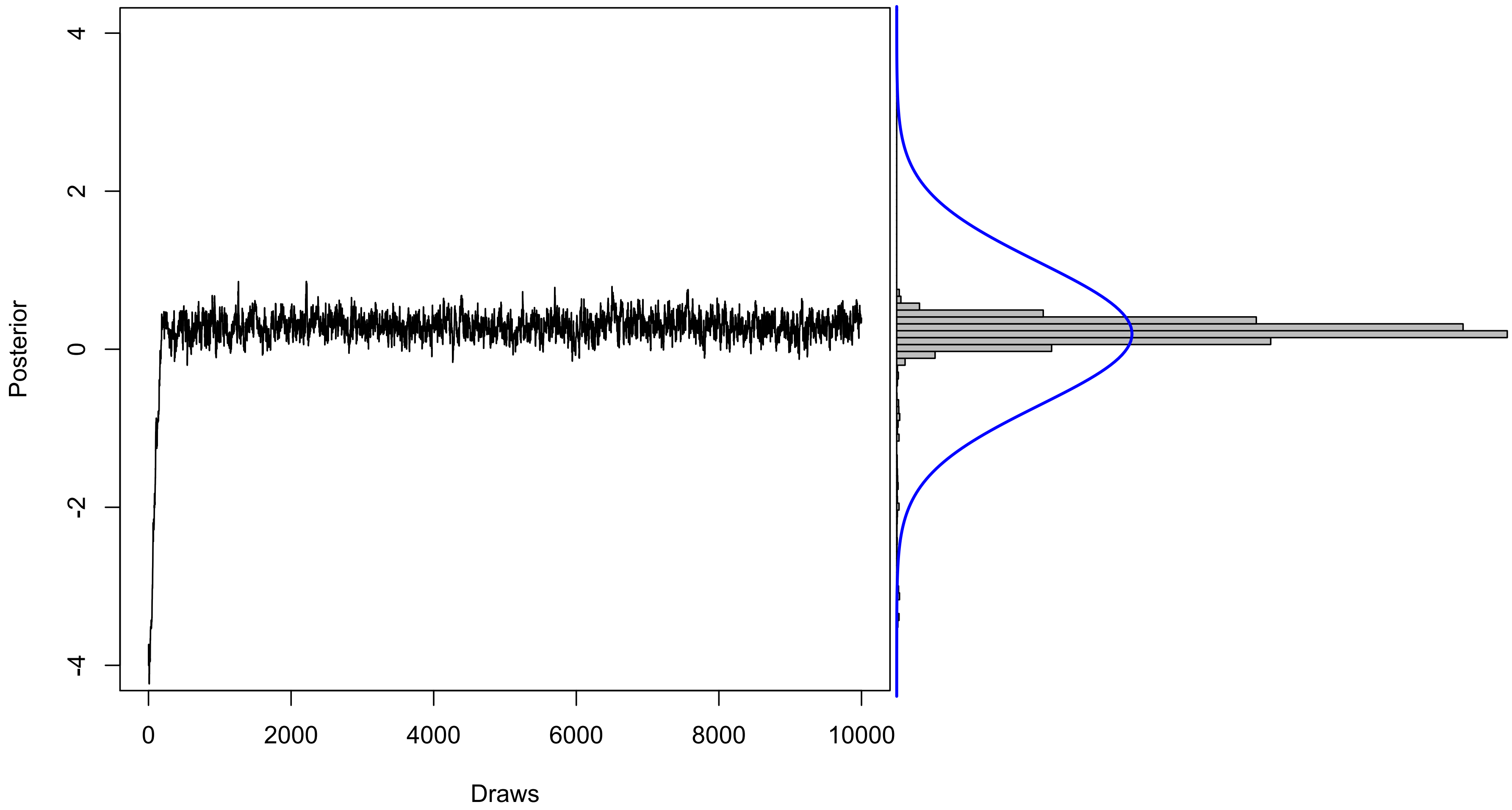
メトロポリス法によるサンプリング(2)

- サンプリング回数=5000回

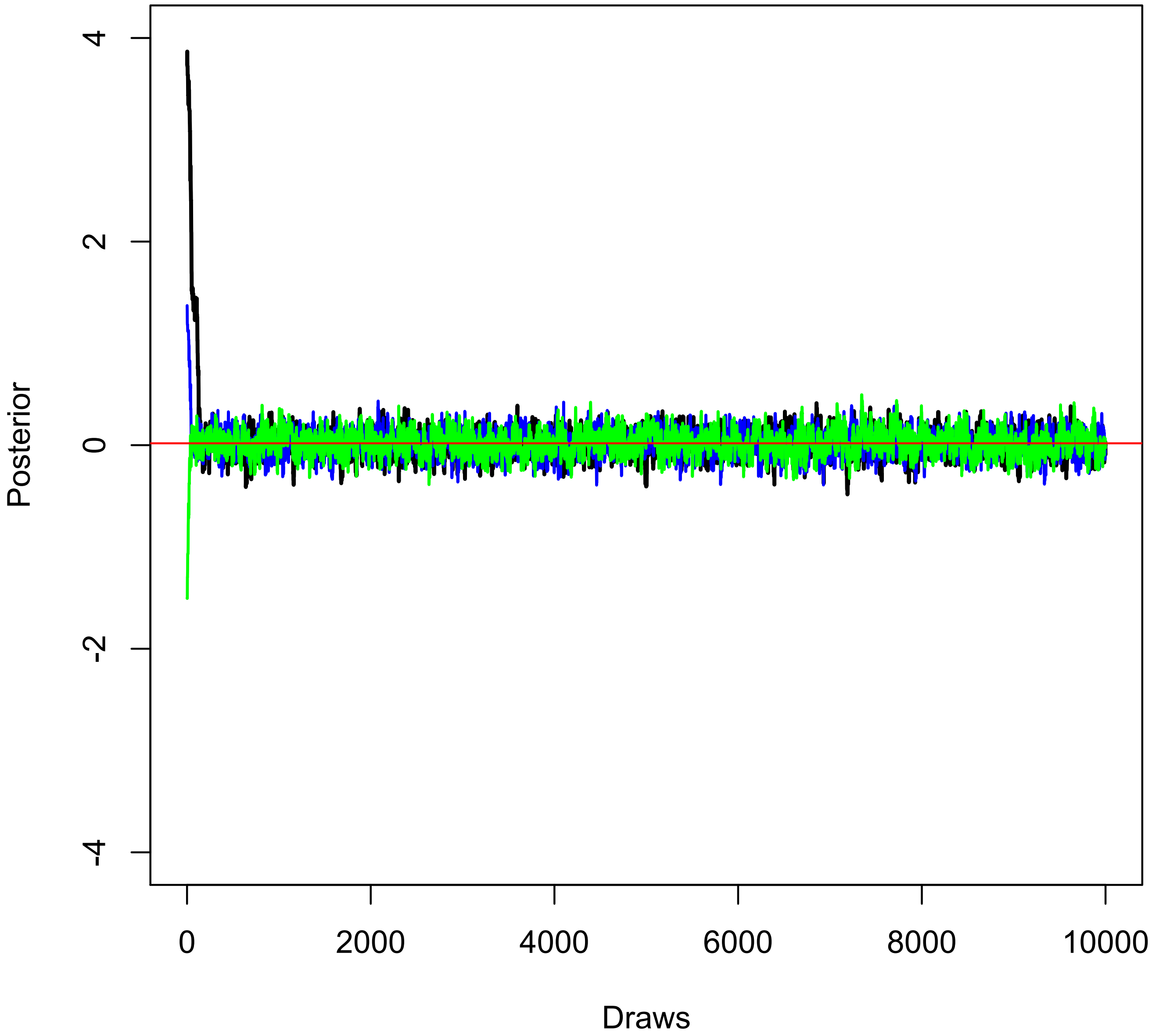


メトロポリス法によるサンプリング(3)

- 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)



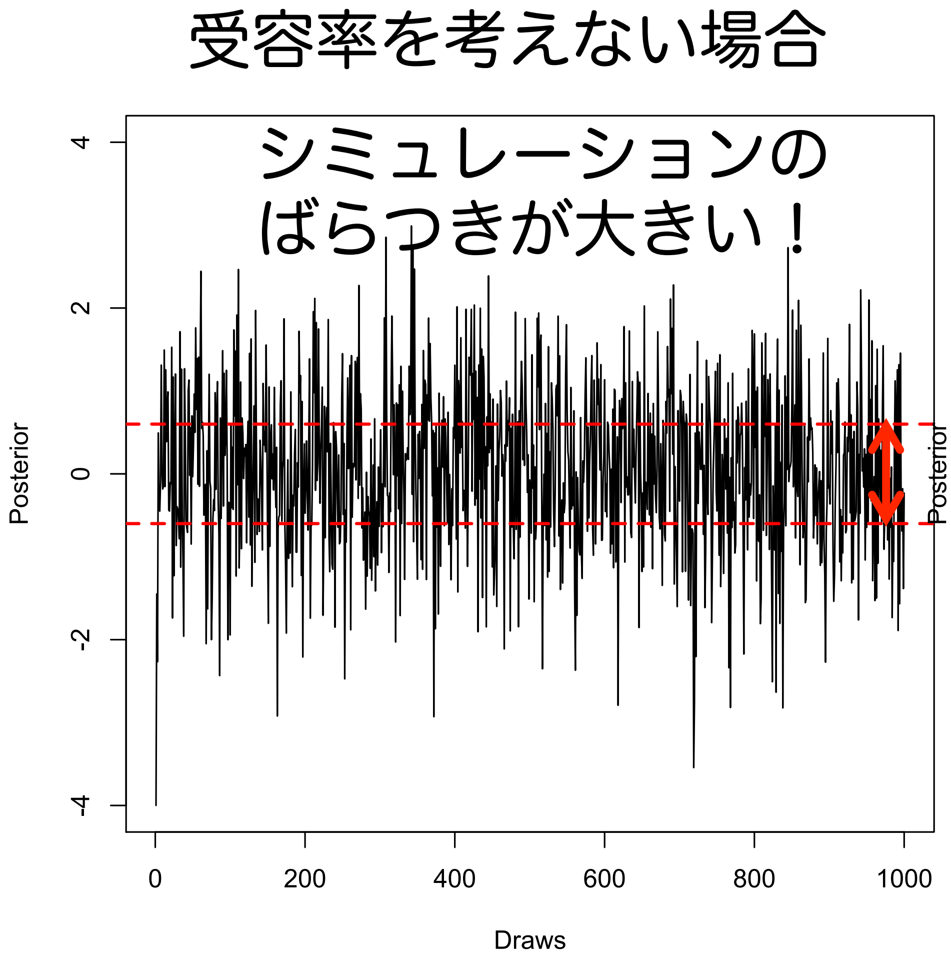
棄却サンプリングと受容率

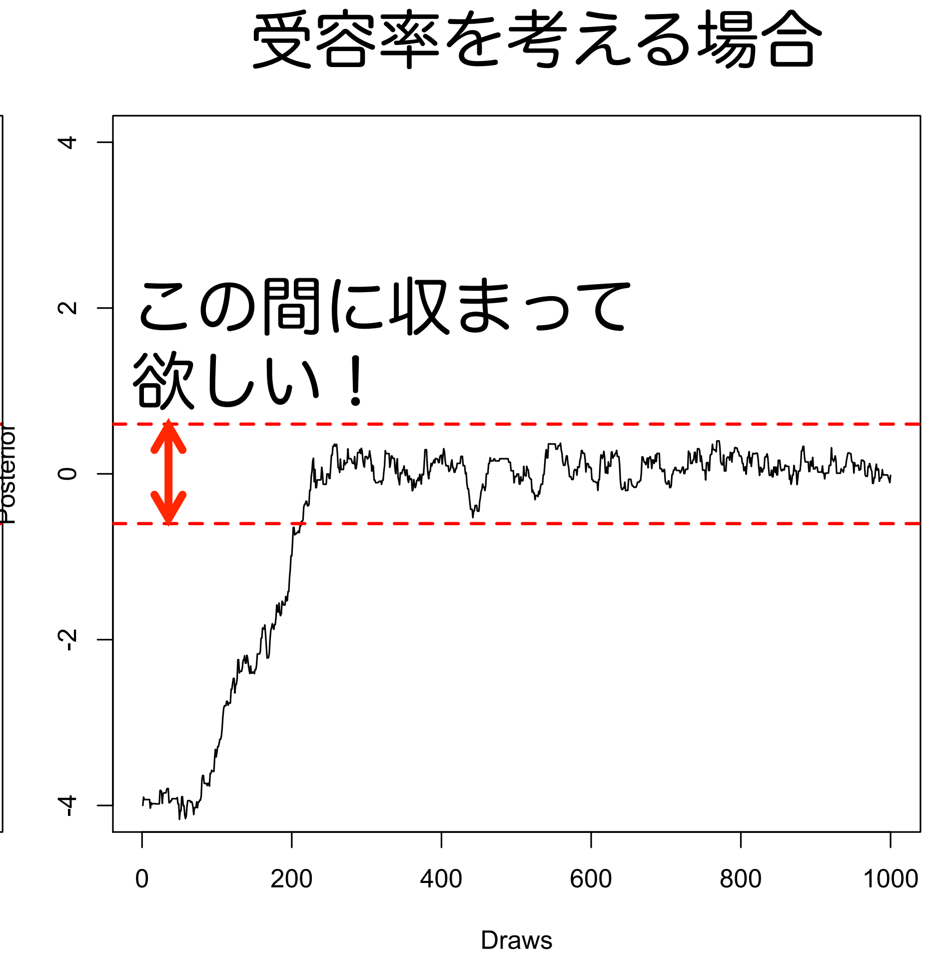
- $p(\theta^s|D)$から乱数発生が難しいとき
- 正規分布など乱数発生が容易な候補確率$p(\theta')$から乱数を発生する
- 区間[0,1]上で受容率uの一様乱数を生成
- 受容率を下回る事後確率は棄却サンプリングとして捨てる
- 計算時間を要することもある

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

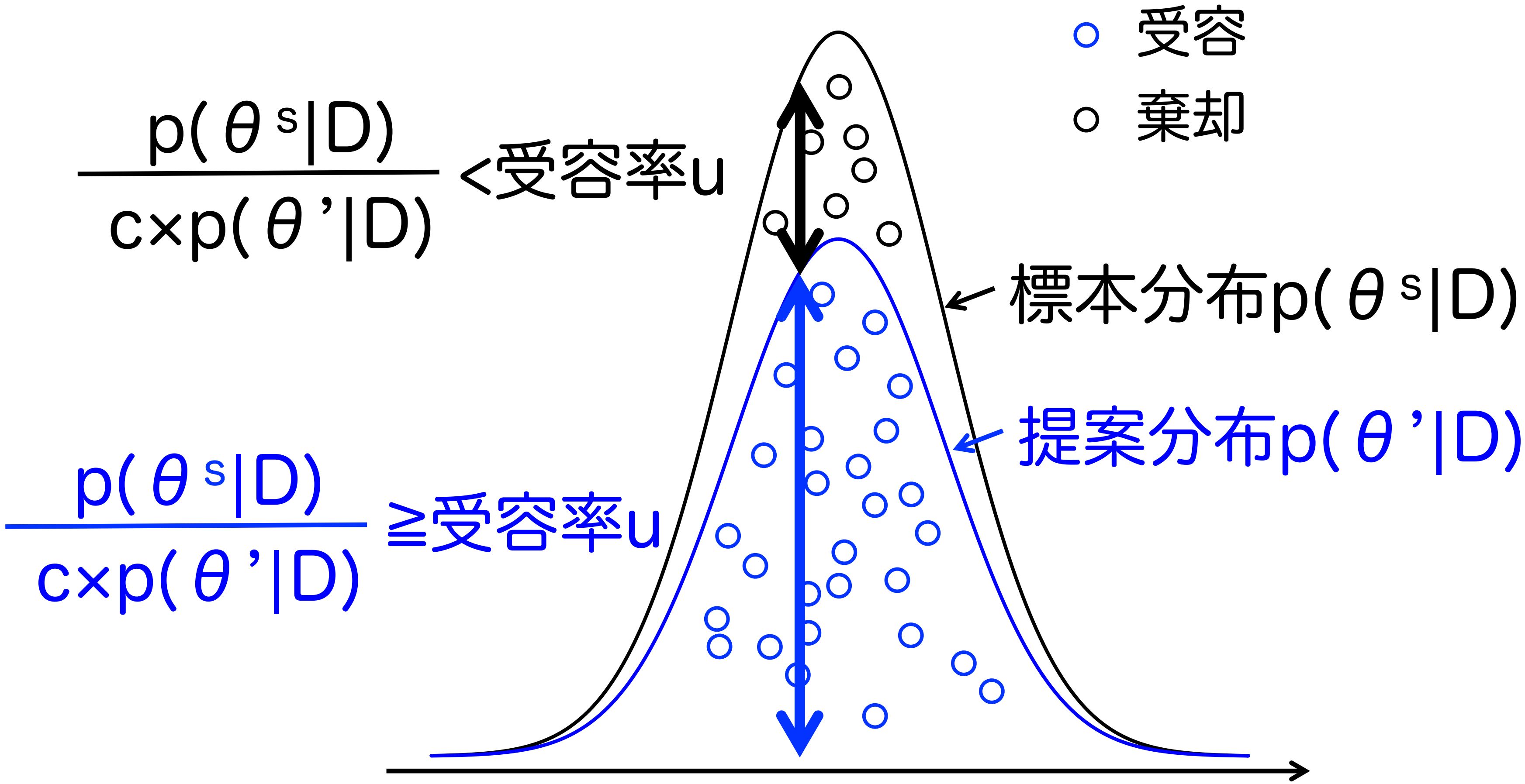
(cは定数)

棄却サンプリングと受容率

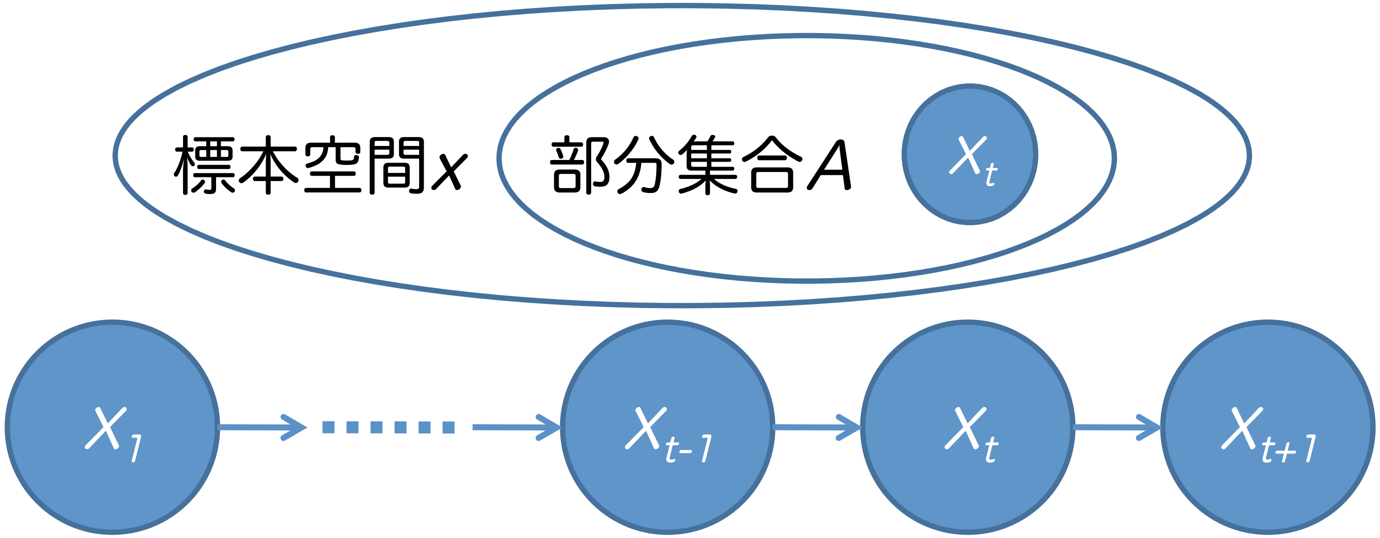




受容-棄却法
(acceptance-rejection)



マルコフ連鎖によるシミュレーション



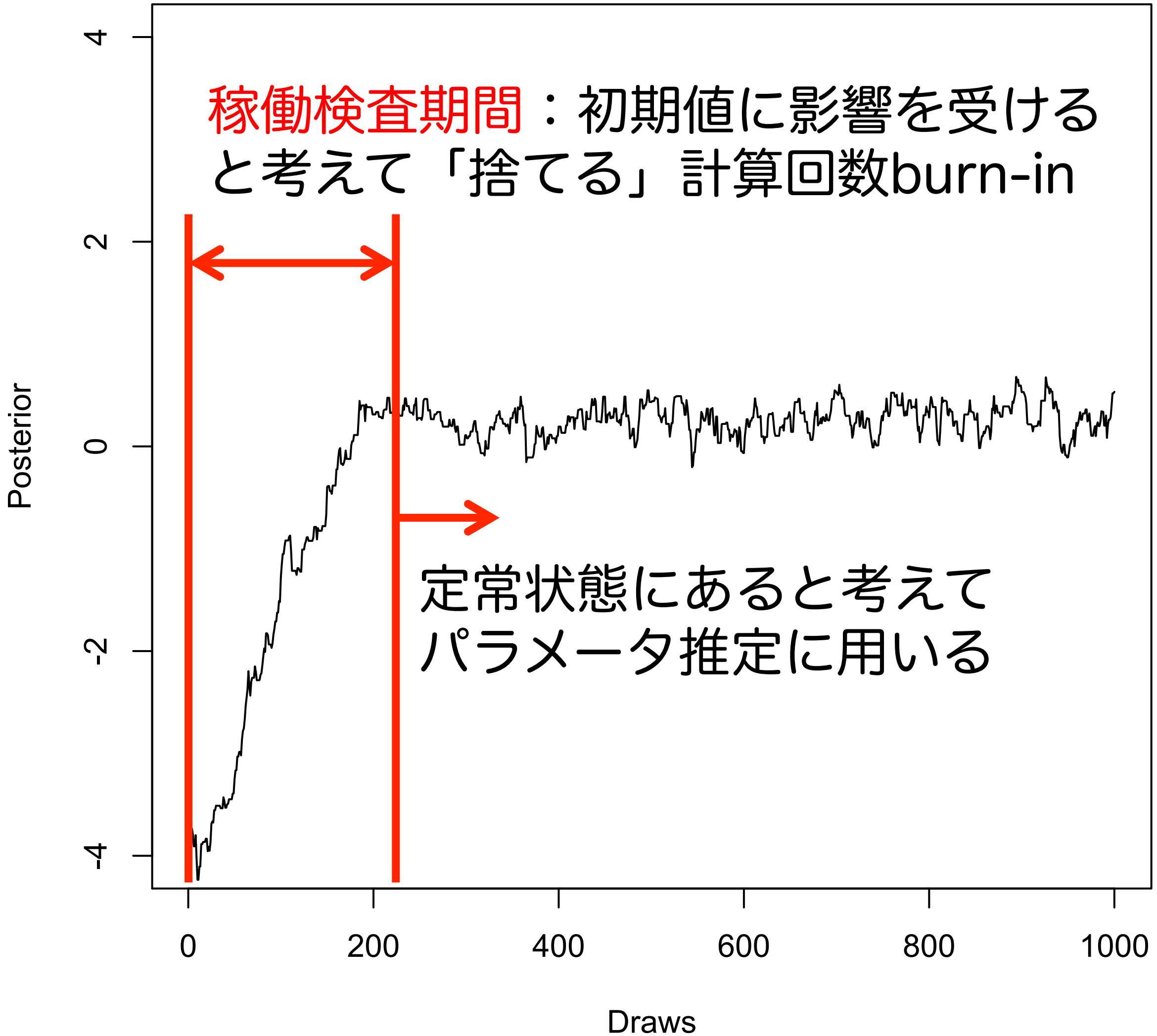
推移確率

$$\underbrace{P\left(X_{t+1}\in A \mid X_1,\cdots,X_t\right)}_{\text{事後確率}} = \underbrace{P\left(X_{t+1}\mid X_t\right)}_{\text{推移核}}\cdot P\left(X_t \mid X_{t-1}\right)\cdot$$

$$\cdots\cdot P\left(X_2\mid X_1\right)\cdot \underbrace{P\left(X_1\right)}_{\text{初期値}}$$

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

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

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

- メトロポリス・ヘイスティング(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

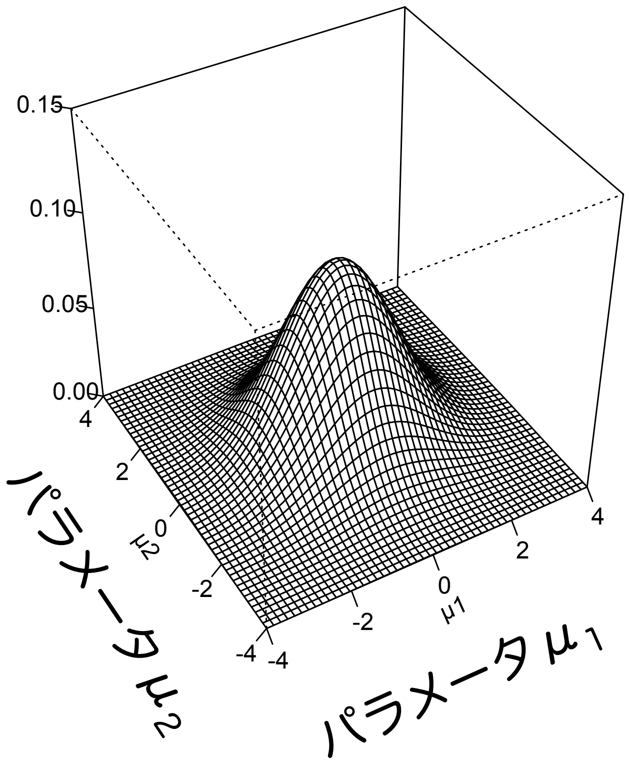
- 目標分布：最終的に求めたい事後分布(の標本分布)
- 目標分布を２つ以上に分割できると考える。例えば、$y=(y_1,y_2)$
- このとき求める事後分布は、

$p(y_1|y_2)$と$p(y_2|y_1)$

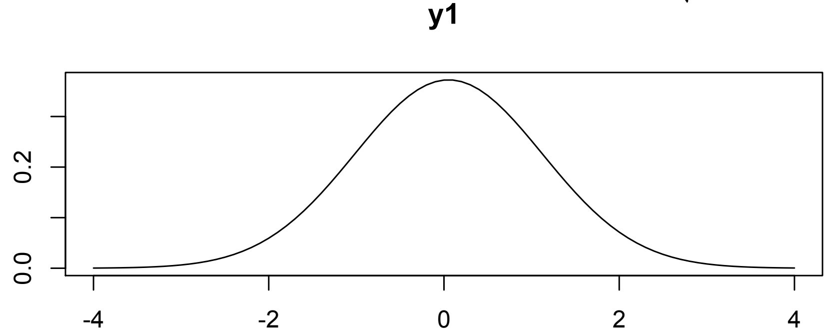
ギブス・サンプリング

- 例えば、二変量正規分布を考える

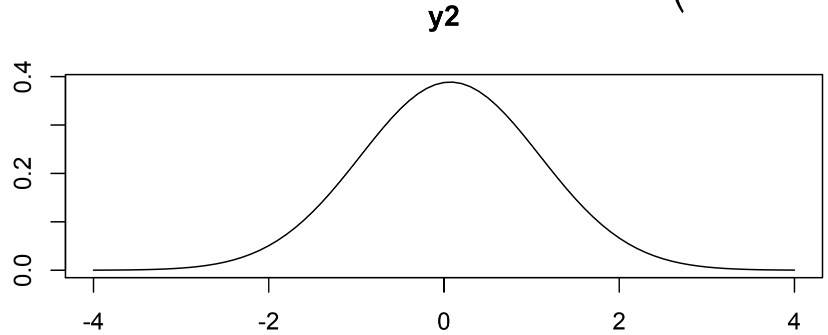
２つの未知パラメータに対する
目標分布（求めたい分布）



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



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



ギブス・サンプリング

1) 計算回数s=(0, …, S)を決める

2) s=0に対し初期値 $x^s = \left(x_1^s, \ldots, x_n^s\right)$ を決める

3) x^sからx^{s+1}を生成する

1) $x_1^{s+1} \sim p\left(x_1^s \mid x_2^s, \ldots, x_n^s\right)$ から x_1^{s+1} を生成

2) $x_2^{s+1} \sim p\left(x_2^s \mid x_1^{s+1}, \ldots, x_n^s\right)$ から x_2^{s+1} を生成

:

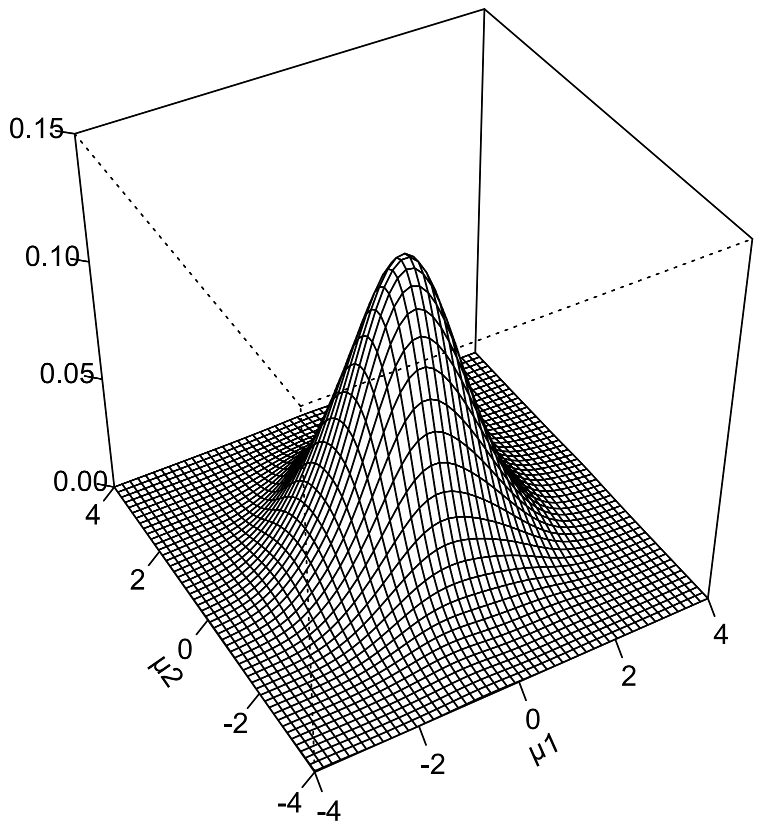
n) $x_n^{s+1} \sim p\left(x_n^s \mid x_1^{s+1}, \ldots, x_n^{s+1}\right)$ から x_n^{s+1} を生成

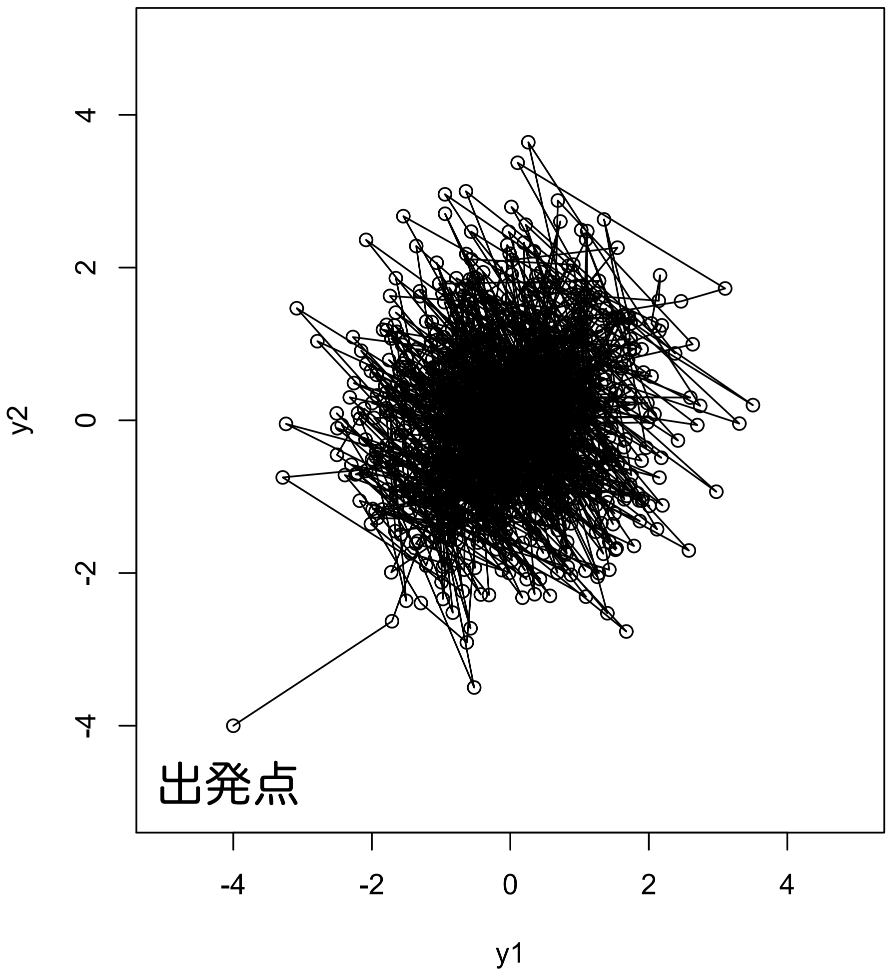
4) s=Sのとき計算終了

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： $y_1 \sim N\left(\mu_1, \rho^2\right)$
  - Y_2の平均μ_1、標準偏差ρ： $y_2 \sim N\left(\mu_2, \rho^2\right)$
  - yの初期値y_0、事後標準偏差の初期値σ_1，σ_2
- y^s からy^{s+1} を生成する 重要なのはココ！
  - $y_2^{s+1} \sim N\left(\mu_2 + \rho \frac{\sigma_2}{\sigma_1}\left(y_1^s - \mu_2\right), \sigma_1^2\left(1 - \rho^2\right)\right)$
  - $y_1^{s+1} \sim N\left(\mu_1 + \rho \frac{\sigma_1}{\sigma_2}\left(y_2^{s+1} - \mu_1\right), \sigma_1^2\left(1 - \rho^2\right)\right)$

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)を得た場合の平均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 $\mu \sim N\left(\mu_0, \sigma_0^2\right)$
  - 分散(精度) $\tau = 1/\sigma_0^2 \sim G\left(\nu_0/2, S_0/2\right)$
    ⇒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 $p\left(y \mid \mu, \sigma^2\right) \sim N\left(\mu, \sigma^2\right)$

正規分布の平均のベイズ推定

- 平均μの事後分布

$$f\left(\mu \mid Y\right) \propto \prod_{i=1}^{n} f\left(y_i \mid \mu, \sigma^2\right) \cdot \pi\left(\mu\right)$$

$$\propto \exp\left(-\frac{\sum_{i=1}^{n}\left(y_i - \mu\right)^2}{2\sigma^2}\right) \cdot \exp\left(-\frac{\left(\mu - \mu_0\right)^2}{2\sigma_0^2}\right)$$

$$= \exp\left(-\frac{\left(\mu - \mu_1\right)^2}{2\sigma_1^2}\right)$$

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

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

$$\prod_{i=1}^{n} f\left(y_i \mid \mu, \sigma^2\right) \propto \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\left(y_i - \mu\right)^2}{2\sigma^2}\right)$$

$$\propto \left(\frac{1}{\sigma^2}\right)^{\frac{n}{2}} \exp\left(-\frac{\sum_{i=1}^{n}\left(y_i - \mu\right)^2}{2\sigma^2}\right)$$

- ここで $\alpha = n/2 - 1,\ \lambda = \sum_{i=1}^{n}\left(y_i - \mu\right)^2 \big/ 2$ と置き換えると、

$$\prod_{i=1}^{n} f\left(y_i \mid \mu, \sigma^2\right) \propto \left(\frac{1}{\sigma^2}\right)^{\alpha-1} \exp\left(-\lambda \cdot \frac{1}{\sigma^2}\right)$$

正規分布の分散のベイズ推定

- さらに、$1/\sigma^2$を精度τで置き換えると

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

- これはガンマ関数Ga(α, λ)に比例する

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

- 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

$$
\begin{aligned}
Ga(\alpha,\lambda) &\propto (\tau)^{\alpha-1}\exp(-\lambda\cdot\tau) \\
&\propto Ga\left(n/2-1, \sum_{i=1}^{n}(y_i-\mu)^2\Big/2\right) \\
&\propto Ga\left((n-2)/2, (n-1)\cdot \overbrace{\boxed{\sum_{i=1}^{n}(y_i-\mu)^2\Big/(n-1)}}^{\text{分散}}\Big/2\right) \\
&\propto Ga\left((\nu_0+n)/2, \nu_0\cdot S_0/2\right) \\
&\propto Ga\left((\nu_0+n)/2, S_0/2\right)
\end{aligned}
$$

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 $\mu \mid \sigma^2, y \sim N\left(\mu_1, \sigma_1^2\right),$

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

  - 事後分散 $\sigma^2 \mid \mu, y \sim IG\left(\nu_1/2, S_1/2\right),$

$$\nu_1 = \nu_0 + n,$$

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

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 $\mu \mid \sigma^2, y \sim N\left(\mu_1, \sigma_1^2\right)$

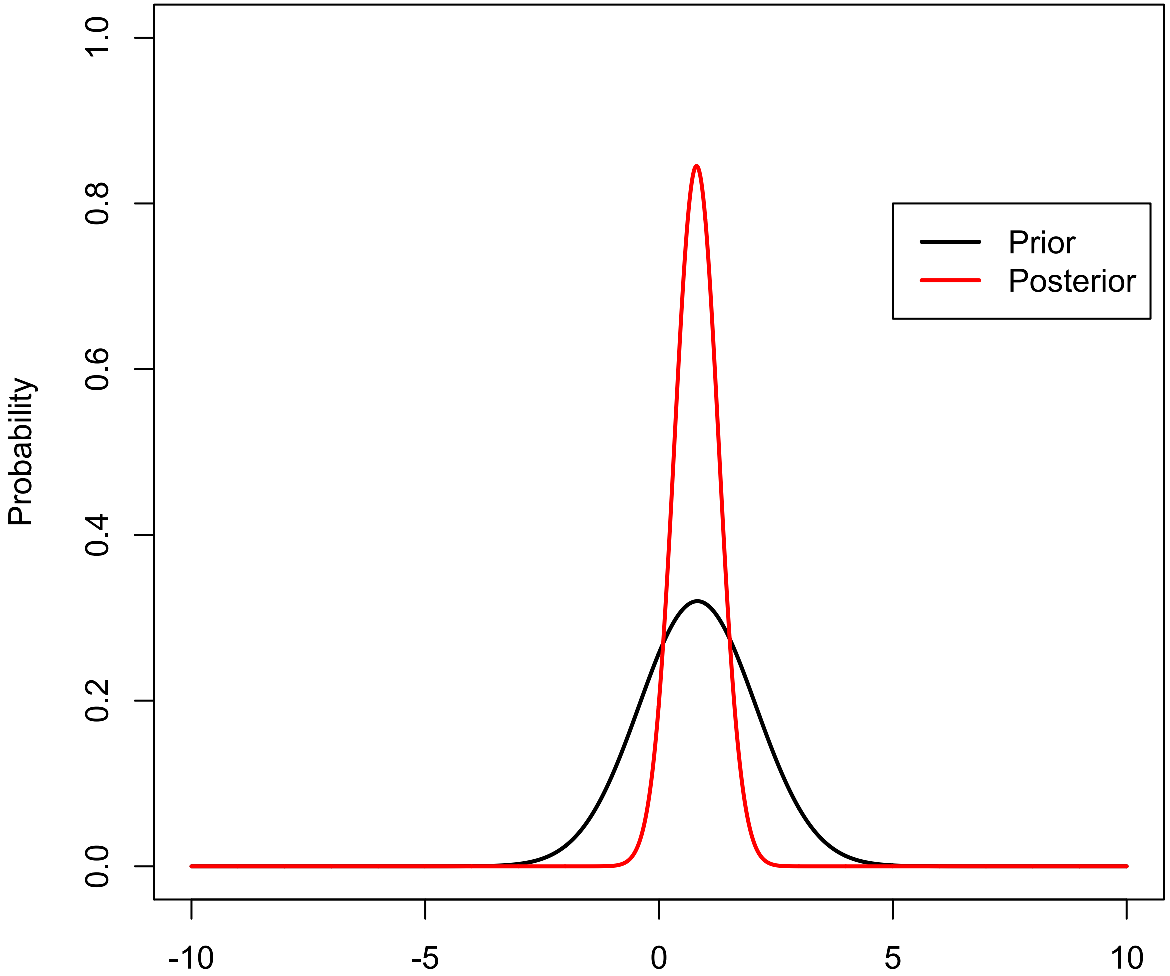
式変形
バージョン

$$\boxed{\begin{aligned} \mu_1 &= \frac{\sigma_0^2}{\sigma_0^2 + \sigma^2/n}\left(\bar{y} - \mu_0\right) + \mu_0, \\ \sigma_1^2 &= \frac{\sigma_0^2}{\sigma_0^2 + \sigma^2/n} \cdot \frac{\sigma^2}{n} \end{aligned}}$$

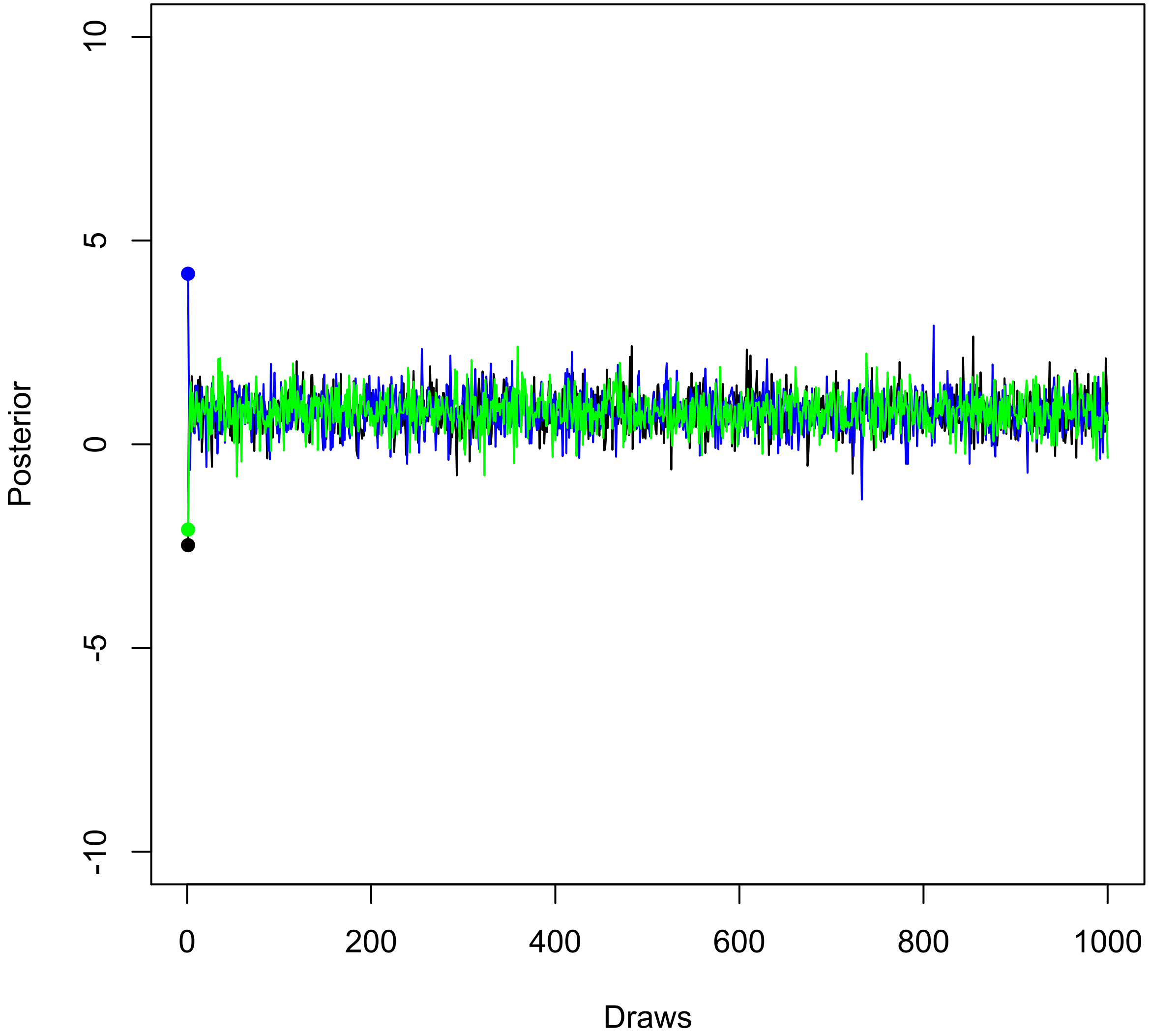
  - 事後分散 $\sigma^2 \mid \mu, y \sim IG\left(\nu_1/2, S_1/2\right)$

$$\nu_1 = \nu_0 + n,\ S_1 = S_0 + \sum_{i=1}^{n}\left(y_i - \mu\right)^2$$

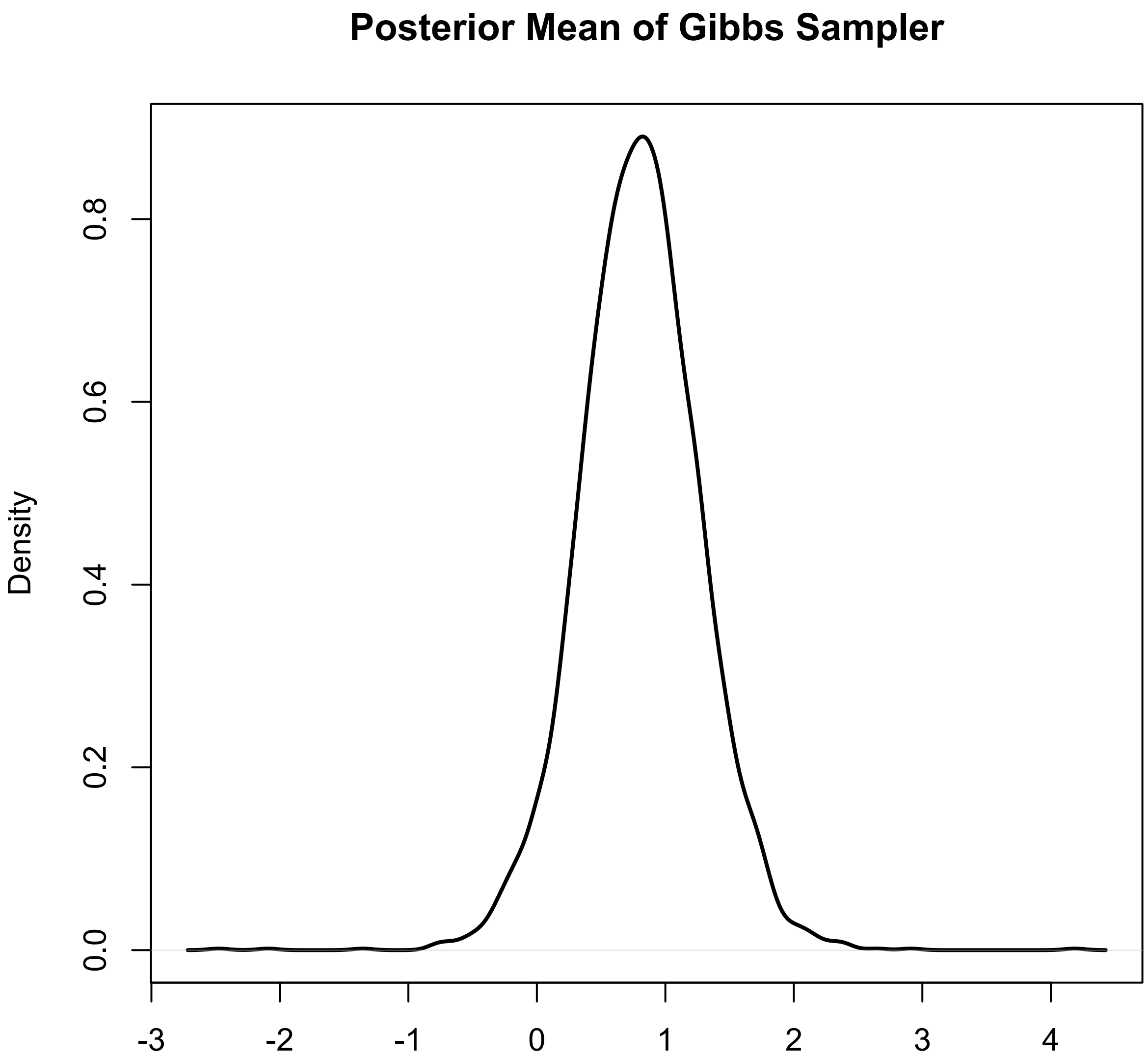
正規母集団同時事後分布のギブス・サンプリング(4)



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)