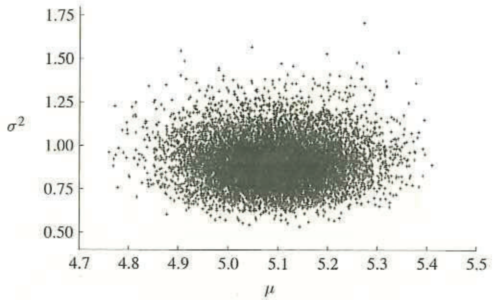

図 3 ギブス・サンプラー．標本の散布図

れた確率標本の時系列プロット（標本径路）と，推定された周辺事後確率密度関数である．ギブス・サンプラーの最初の 1000 回は初期値に依存する期間（稼動検査期間，**burn-in period**）であるとして捨てて，それ以降の 10000 回を用いている．標本径路を見ると状態空間をまんべんなく安定的にサンプリングして不変分布である事後分布に収束しており，またこのことは，図 3 の確率標本の散布図からも確認することができる．さらに表 2 は，事後分布の平均，標準偏差およびパラメータの 95% 信用区間の推定結果であり，95% 信用区間が真の値を含んでいることがわかる．

表 2 事後分布の平均，標準偏差，95% 信用区間

パラメータ	事後平均	事後標準偏差	95% 信用区間
μ	5.091	0.095	(4.907, 5.280)
σ^2	0.908	0.133	(0.680, 1.206)

ここで，さまざまなモデルに対してギブス・サンプラーを用いた推論の例を示そう．

例 5（階層ベイズモデル）

Y_{ij} を第 i グループの第 j 観測値として

$$Y_{ij} = \mu_i + \epsilon_{ij}, \quad i = 1, \cdots, K, \quad j = 1, \cdots, n_i,$$

$$\epsilon_{ij} \sim \text{i.i.d. } N(0, \sigma^2)$$

$\sigma^2 = IG(n_0/2, S_0/2)$

$\mu_i = N(\mu, \tau^2)$

$\mu = N(\mu_0, \sigma_0^2)$

$\tau = IG(k_0/2, R_0/2)$

とする．ただし K はグループ数，n_i は第 i グループからの標本の大きさである．事前分布として平均 μ_i は $\mu_1, \mu_2, \cdots, \mu_K \sim \text{i.i.d. } N(\mu, \tau^2)$ であり，さらに $\mu \sim N(\mu_0, \sigma_0^2)$, $\sigma^2 \sim IG\left(\frac{n_0}{2}, \frac{S_0}{2}\right)$, $\tau^2 \sim IG\left(\frac{k_0}{2}, \frac{R_0}{2}\right)$ とする*14．このとき，$\boldsymbol{y} = (y_{11}, y_{12}, \cdots, y_{Kn_K})'$, $\tilde{\mu} = (\mu_1, \cdots, \mu_K)$ として尤度関数は

$$f(\boldsymbol{y}|\tilde{\mu}, \sigma^2) = \prod_{i=1}^{K}\prod_{j=1}^{n_i} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(y_{ij}-\mu_i)^2}{2\sigma^2}\right\}$$

階層ベイズ法

であり，事前確率密度関数は

$$\begin{aligned}
&\pi(\tilde{\mu}, \mu, \sigma^2, \tau^2) \\
&= \pi(\tilde{\mu}|\mu)\pi(\mu)\pi(\sigma^2)\pi(\tau^2) \\
&= \prod_{i=1}^{K} \pi(\mu_i|\mu) \times \pi(\mu)\pi(\sigma^2)\pi(\tau^2) \\
&\propto \prod_{i=1}^{K} (\tau^2)^{-\frac{1}{2}} \exp\left\{-\frac{(\mu_i-\mu)^2}{2\tau^2}\right\} \times \exp\left\{-\frac{(\mu-\mu_0)^2}{2\sigma_0^2}\right\} \\
&\quad \times (\sigma^2)^{-\left(\frac{n_0}{2}+1\right)} \exp\left\{-\frac{S_0}{2\sigma^2}\right\} \times (\tau^2)^{-\left(\frac{k_0}{2}+1\right)} \exp\left\{-\frac{R_0}{2\tau^2}\right\}
\end{aligned}$$

である．したがって同時事後確率密度関数は

$$\begin{aligned}
&\pi(\tilde{\mu}, \mu, \sigma^2, \tau^2|\boldsymbol{y}) \\
&\propto f(\boldsymbol{y}|\tilde{\mu}, \sigma^2)\pi(\tilde{\mu}, \mu, \sigma^2, \tau^2) \\
&\propto (\sigma^2)^{-\left(\frac{n_0+n}{2}+1\right)} (\tau^2)^{-\left(\frac{k_0+K}{2}+1\right)} \\
&\quad \times \exp\left\{-\frac{\sum_{i=1}^{K}\sum_{j=1}^{n_i}(y_{ij}-\mu_i)^2 + S_0}{2\sigma^2} - \frac{\sum_{i=1}^{K}(\mu_i-\mu)^2 + R_0}{2\tau^2}\right\} \\
&\quad \times \exp\left\{-\frac{(\mu-\mu_0)^2}{2\sigma_0^2}\right\}, \quad n = \sum_{i=1}^{K} n_i
\end{aligned}$$

See 例 3 のアナロジィ

となる．このことから，まず μ_i $(i = 1, \cdots, K)$ の条件付事後確率密度関数は

$$\pi(\mu_i|\tilde{\mu}_{-i}, \mu, \sigma^2, \tau^2, \boldsymbol{y}) \propto \exp\left\{-\frac{(\mu_i - m_i)^2}{2s_i^2}\right\}$$

$f(y|\tilde{\mu}, \sigma^2) \cdot \pi(\mu_i)$

*14 この例のように第 1 段階として μ_i に事前分布を設定し，さらに第 2 段階としてその事前分布のパラメータ μ に事前分布を階層的に設定するようなモデルを，階層ベイズモデル（hierarchical Bayes model）という．

ただし $\tilde{\mu}_{-i} = (\mu_1, \cdots, \mu_{i-1}, \mu_{i+1}, \cdots, \mu_K)$,

$$s_i^{-2} = \tau^{-2} + n_i\sigma^{-2}, \quad m_i = s_i^2\left(\tau^{-2}\mu + \sigma^{-2}n_i\overline{y}_i\right),$$

$$\overline{y}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} y_{ij}$$

である．次に μ, σ^2, τ^2 の条件付事後確率密度関数は

$$\pi(\mu|\tilde{\mu}, \sigma^2, \tau^2, \boldsymbol{y}) \propto \exp\left\{-\frac{(\mu - m)^2}{2s^2}\right\}$$

$$\pi(\sigma^2|\tilde{\mu}, \mu, \tau^2, \boldsymbol{y}) \propto (\sigma^2)^{-\left(\frac{n_1}{2}+1\right)} \exp\left\{-\frac{S_1}{2\sigma^2}\right\}$$

$$\pi(\tau^2|\tilde{\mu}, \mu, \sigma^2, \boldsymbol{y}) \propto (\tau^2)^{-\left(\frac{k_1}{2}+1\right)} \exp\left\{-\frac{R_1}{2\tau^2}\right\}$$

ただし

$$s^{-2} = \sigma_0^{-2} + K\tau^{-2}, \quad m = s^2\left(\sigma_0^{-2}\mu_0 + \tau^{-2}K\overline{\mu}\right), \quad \overline{\mu} = \frac{1}{K}\sum_{i=1}^{K}\mu_i$$

$$n_1 = n_0 + n, \quad S_1 = S_0 + \sum_{i=1}^{K}\sum_{j=1}^{n_i}(y_{ij} - \mu_i)^2$$

$$k_1 = k_0 + K, \quad R_1 = R_0 + \sum_{i=1}^{K}(\mu_i - \mu)^2$$

である．以上をまとめると

$$\mu_i|\tilde{\mu}_{-i}, \mu, \sigma^2, \tau^2, \boldsymbol{y} \sim N(m_i, s_i^2), \quad i = 1, \cdots, K,$$

$$\mu|\tilde{\mu}, \sigma^2, \tau^2, \boldsymbol{y} \sim N(m, s^2),$$

$$\sigma^2|\tilde{\mu}, \mu, \tau^2, \boldsymbol{y} \sim IG\left(\frac{n_1}{2}. \frac{S_1}{2}\right)$$

$$\tau^2|\tilde{\mu}, \mu, \sigma^2, \boldsymbol{y} \sim IG\left(\frac{k_1}{2}. \frac{R_1}{2}\right)$$

となるので，これをもとにギブス・サンプラーを行えばよい．

例 6（単純回帰モデル）

次のような単純回帰モデルを考える．y_i は第 i 個体の被説明変数で確率変数，x_i は説明変数で定数，ϵ_i は誤差項で確率変数，$\beta_0, \beta_1, \sigma^2$ はパラメータとする．つまり

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \quad \epsilon_i \sim \text{ i.i.d. } N(0, \sigma^2), \quad i = 1, 2, \cdots, n$$

とおく．このとき尤度関数は

$$f(\boldsymbol{y}|\beta_0, \beta_1, \sigma^2) = \prod_{i=1}^{n}\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{(y_i - \beta_0 - \beta_1 x_i)^2}{2\sigma^2}\right\}$$

$$\propto (\sigma^2)^{-\frac{n}{2}}\exp\left\{-\frac{\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2}{2\sigma^2}\right\}$$

である．ここで $(\beta_0, \beta_1, \sigma^2)$ の事前分布を，それぞれ $\beta_0 \sim N(b_0, B_0), \beta_1 \sim N(c_0, C_0), \sigma^2 \sim IG(n_0/2, S_0/2)$ とおくと

$$\pi(\beta_0, \beta_1, \sigma^2)$$
$$= \pi(\beta_0)\pi(\beta_1)\pi(\sigma^2)$$
$$\propto \exp\left\{-\frac{(\beta_0 - b_0)^2}{2B_0}\right\} \times \exp\left\{-\frac{(\beta_1 - c_0)^2}{2C_0}\right\} \times (\sigma^2)^{-\left(\frac{n_0}{2}+1\right)}\exp\left\{-\frac{S_0}{2\sigma^2}\right\}$$

となり，同時事後確率密度関数は

$$\pi(\beta_0, \beta_1, \sigma^2|\boldsymbol{y}) \propto (\sigma^2)^{-\left(\frac{n_0+n}{2}+1\right)}\exp\left\{-\frac{S_0 + \sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2}{2\sigma^2}\right\}$$

$$\times \exp\left\{-\frac{(\beta_0 - b_0)^2}{2B_0} - \frac{(\beta_1 - c_0)^2}{2C_0}\right\}$$

となる．これより条件付確率密度関数は

$f(y|\beta_0\beta_1\sigma^2) \cdot \pi(\beta_0) / f(y)$

$$\pi(\beta_0|\beta_1, \sigma^2, \boldsymbol{y}) \propto \exp\left\{-\frac{(\beta_0 - b_1)^2}{2B_1}\right\}, = f(y|\beta_0\beta_1\sigma^2) \cdot \pi(\beta_0)$$

$$\pi(\beta_1|\beta_0, \sigma^2, \boldsymbol{y}) \propto \exp\left\{-\frac{(\beta_1 - c_1)^2}{2C_1}\right\}, = f(y|\beta_0\beta_1\sigma^2)\pi(\beta_1)$$

$$\pi(\sigma^2|\beta_0, \beta_1, \boldsymbol{y}) \propto (\sigma^2)^{-\left(\frac{n_1}{2}+1\right)}\exp\left\{-\frac{S_1}{2\sigma^2}\right\} = f(y|\beta_0\beta_1\sigma^2)\pi(\sigma^2)$$

ただし

$$b_1 = B_1\left\{B_0^{-1}b_0 + \sigma^{-2}\sum_{i=1}^{n}(y_i - \beta_1 x_i)\right\}, \quad B_1^{-1} = B_0^{-1} + n\sigma^{-2},$$

$$c_1 = C_1\left\{C_0^{-1}c_0 + \sigma^{-2}\sum_{i=1}^{n}(y_i - \beta_0)x_i\right\}, \quad C_1^{-1} = C_0^{-1} + \sigma^{-2}\sum_{i=1}^{n}x_i^2,$$

$$n_1 = n_0 + n, \quad S_1 = S_0 + \sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2$$

である．したがって事後分布は

$$\beta_0|\beta_1, \sigma^2, \boldsymbol{y} \sim N(b_1, B_1),$$
$$\beta_1|\beta_0, \sigma^2, \boldsymbol{y} \sim N(c_1, C_1),$$
$$\sigma^2|\beta_0, \beta_1, \boldsymbol{y} \sim IG\left(\frac{n_1}{2}, \frac{S_1}{2}\right)$$

であるので，これをもとにギブス・サンプラーを行えばよい．

例 6 では，回帰係数 β_0, β_1 を別々にサンプリングしているが，これを以下のように同時にサンプリングしてもよい．

例 6（単純回帰モデル．続き）

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \quad \epsilon_i \sim \text{ i.i.d. } N(0, \sigma^2), \quad i = 1, 2, \cdots, n$$

は

$$\boldsymbol{y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon}, \quad \boldsymbol{\epsilon} \sim N(0, \sigma^2 I_n)$$

と書くことができる．ただし $\boldsymbol{y} = (y_1, \cdots, y_n)'$, $\boldsymbol{\beta} = (\beta_0, \beta_1)'$, $\boldsymbol{\epsilon} = (\epsilon_1, \cdots, \epsilon_n)'$, I_n は n 次元単位行列で

$$X = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}$$

である．このとき尤度関数は

$$f(\boldsymbol{y}|\boldsymbol{\beta}, \sigma^2) = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2\sigma^2}(\boldsymbol{y} - X\boldsymbol{\beta})'(\boldsymbol{y} - X\boldsymbol{\beta})\right\}$$

となる．$\boldsymbol{\beta}, \sigma^2$ の事前分布を，それぞれ $\boldsymbol{\beta} \sim N(\boldsymbol{b}_0, B_0), \sigma^2 \sim IG(n_0/2, S_0/2)$ とおくと（$\boldsymbol{b}$ は 2×1 ベクトル，B_0 は 2×2 行列），事前確率密度関数は

$$\begin{aligned}
&\pi(\boldsymbol{\beta}, \sigma^2) \\
&= \pi(\boldsymbol{\beta})\pi(\sigma^2) \\
&\propto \exp\left\{-\frac{1}{2}(\boldsymbol{\beta} - \boldsymbol{b}_0)'B_0^{-1}(\boldsymbol{\beta} - \boldsymbol{b}_0)\right\} \times (\sigma^2)^{-(\frac{n_0}{2}+1)} \exp\left\{-\frac{S_0}{2\sigma^2}\right\}
\end{aligned}$$

である．したがって

$$\begin{aligned}
\pi(\boldsymbol{\beta}, \sigma^2|\boldsymbol{y}) \propto (\sigma^2)^{-(\frac{\overset{n_0+n}{n_1}}{2}+1)} \\
&\times \exp\left\{-\frac{S_0 + (\boldsymbol{y} - X\boldsymbol{\beta})'(\boldsymbol{y} - X\boldsymbol{\beta})}{2\sigma^2}\right\} \\
&\times \exp\left\{-\frac{1}{2}(\boldsymbol{\beta} - \boldsymbol{b}_0)'B_0^{-1}(\boldsymbol{\beta} - \boldsymbol{b}_0)\right\}
\end{aligned}$$

であり，条件付確率密度関数は

$$\pi(\boldsymbol{\beta}|\sigma^2, \boldsymbol{y}) \propto \exp\left\{-\frac{1}{2}(\boldsymbol{\beta} - \boldsymbol{b}_1)'B_1^{-1}(\boldsymbol{\beta} - \boldsymbol{b}_1)\right\},$$

$$\pi(\sigma^2|\boldsymbol{\beta}, \boldsymbol{y}) \propto (\sigma^2)^{-(\frac{n_1}{2}+1)} \exp\left\{-\frac{S_1}{2\sigma^2}\right\}$$

ただし，

$$\boldsymbol{b}_1 = B_1\left(B_0^{-1}\boldsymbol{b}_0 + \sigma^{-2}X'\boldsymbol{y}\right), \quad B_1^{-1} = B_0^{-1} + \sigma^{-2}X'X,$$

$$n_1 = n_0 + n, \quad S_1 = S_0 + (\boldsymbol{y} - X\boldsymbol{\beta})'(\boldsymbol{y} - X\boldsymbol{\beta})$$

である．事後分布は

$$\boldsymbol{\beta}|\sigma^2, \boldsymbol{y} \sim N(\boldsymbol{b}_1, B_1), \quad \sigma^2|\boldsymbol{\beta}, \boldsymbol{y} \sim IG\left(\frac{n_1}{2}, \frac{S_1}{2}\right)$$

となるので，これを用いてギブス・サンプラーを行う．

3.2　メトロポリス–ヘイスティングスアルゴリズム

ギブス・サンプラーのように条件付事後分布 π_i^* からのサンプリングが簡単にできない場合には，メトロポリス–ヘイスティングスアルゴリズムを用いる．メトロポリス–ヘイスティングスアルゴリズムの特徴は，マル