



MCMCの基礎と統計科学への応用

東京大学大学院経済学研究科

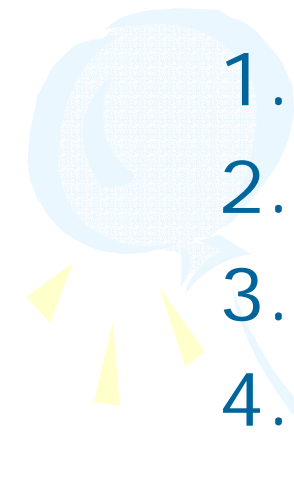

大森裕浩

2004年2月23日



全体の流れ(1)

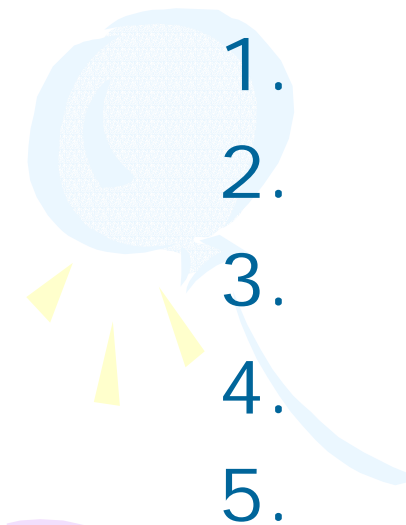
I. MCMCの基礎

- 
1. はじめに
 2. ベイズの定理
 3. ギブス・サンプラー
 4. MHアルゴリズム
 5. 収束の判定方法
 6. MCMCのためのソフトウェア
- 



全体の流れ(2)

II. 統計科学への応用

- 
1. 回帰モデル(誤差項が正規分布やt分布)
 2. トービットモデル(打ち切りのある回帰モデル)
 3. プロビットモデル
 4. ロジットモデル
 5. 見かけ上無関係な回帰モデル



I. MCMCの基礎

1. はじめに
2. ベイズの定理
3. ギブス・サンプラー
4. MHアルゴリズム
5. 収束の判定方法



1. はじめに

ベイズ的アプローチとMCMCを用いると…

ポートフォリオ選択。

運用者の相場観や投資家心理など主観的な情報を取り込んだ上で(不確実性も考慮)期待効用を最大化するような解を与えることが可能。

株式収益率のモデル。

よく知られる確率的ボラティリティ変動モデル。多くの潜在変数がモデルに存在するためにこれまでは通常的最尤法では解くことが困難だったが、MCMCにより大型計算機を使うことなく推定が可能に。



2. ベイズの定理

事前確率・主観確率

関心のある事象A(例:ある病気である)に関する事前情報を確率 $\Pr(A)$ として表現する。

例: $\Pr(A)=0.02$

データの観測

事象Aに関連する事象B(例:陽性の検査結果)を観測する。ただし、Aが与えられたときのBの条件付確率 $\Pr(B|A)$ 、 $\Pr(B|A^c)$ は既知とする。

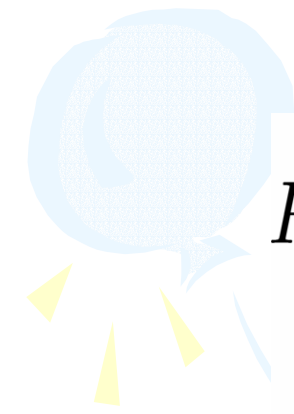
例: $\Pr(B|A)=0.8$, $\Pr(B|A^c)=0.1$



2. ベイズの定理 (続き)

事後確率

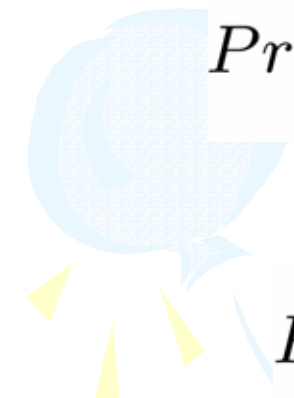
事象Bが観測されたときのAの事後確率は、


$$\begin{aligned} Pr(A|B) &= \frac{Pr(A \cap B)}{Pr(B)} \\ &= \frac{Pr(B \cap A)}{Pr(B \cap A) + Pr(B \cap A^c)} \end{aligned}$$

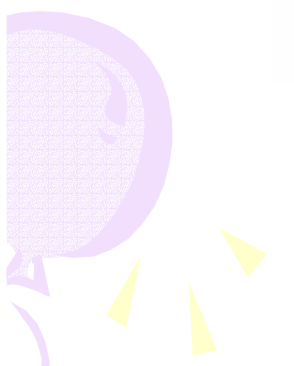


2. ベイズの定理 (続き)

したがって


$$Pr(A|B) = \frac{Pr(B|A)Pr(A)}{Pr(B|A)Pr(A) + Pr(B|A^c)Pr(A^c)}$$

例


$$\begin{aligned} Pr(A|B) &= \frac{0.8 \times 0.02}{0.8 \times 0.02 + 0.1 \times (1 - 0.02)} \\ &= \frac{0.016}{0.114} = 0.1404 \end{aligned}$$



2. ベイズの定理 (連続変数)

- 関心の対象: A

例: 株式の真の収益率, 未知のパラメータ

- 観測する情報: $B \quad X$

例: 株式の観測される収益率, データ

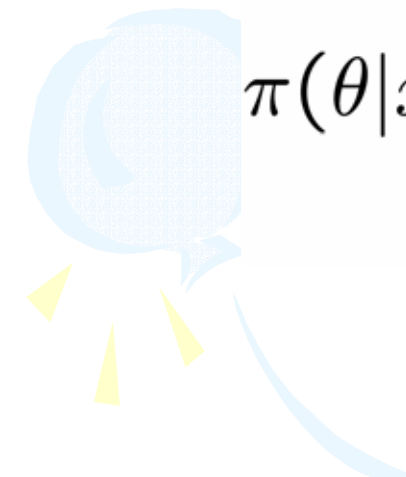
- A の確率密度関数は (\quad)

- X の確率密度関数は $f(x | \quad)$

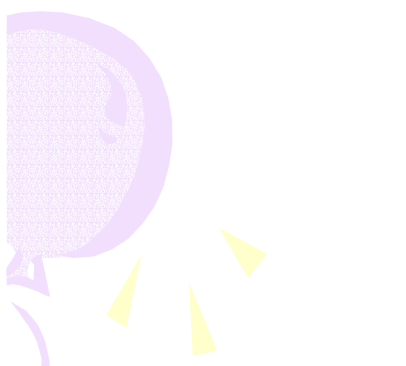


2. ベイズの定理 (続き)

- の事後確率密度関数は $(\quad | x)$


$$\pi(\theta|x) = \frac{f(x|\theta)\pi(\theta)}{\pi(x)} = \frac{f(x|\theta)\pi(\theta)}{\int f(x|\theta)\pi(\theta)d\theta}$$

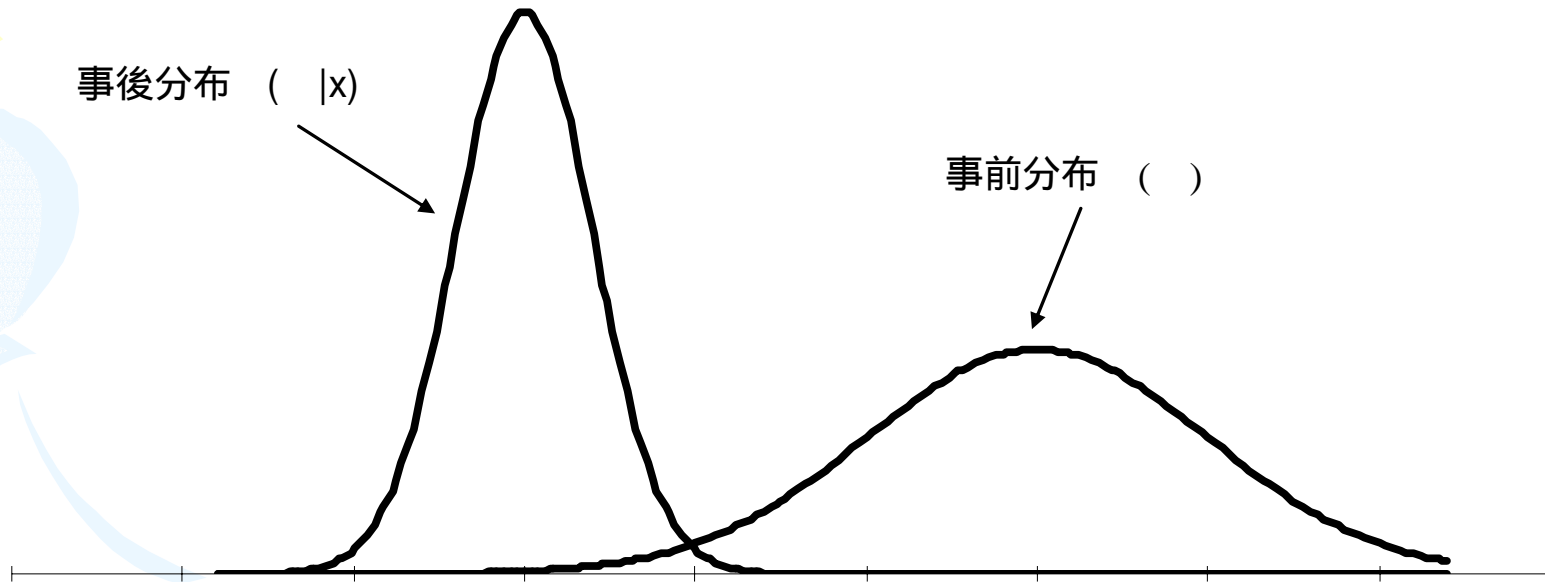
あるいは単に


$$\pi(\theta|x) \propto f(x|\theta)\pi(\theta)$$

ベイズの定理(続き)

事後分布 ($\pi(\theta|x)$)

事前分布 ($\pi(\theta)$)





2. ベイズの定理 (続き)

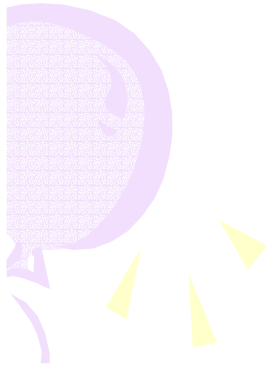
例1.

$X_1, X_2, \dots, X_n \sim N(\mu, \sigma^2)$. (σ^2 :known)



μ の事前分布は $N(\mu_0, \sigma_0^2)$ として (μ_0, σ_0^2) は既知であるとする

$$\pi(\mu) = \frac{1}{\sqrt{2\pi}\sigma_0} e^{-\frac{(\mu-\mu_0)^2}{2\sigma_0^2}}$$


$$f(x|\mu) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}$$

2. ベイズの定理 (続き)

事後分布は, 以下のような正規分布

$$\begin{aligned}\pi(\mu|x) &\propto f(x|\mu)\pi(\mu) \\ &\propto \exp\left[-\frac{\sum_{i=1}^n (x_i - \mu)^2}{2\sigma^2} - \frac{(\mu - \mu_0)^2}{2\sigma_0^2}\right] \\ &\propto \exp\left[-\frac{(\mu - \mu_1)^2}{2\sigma_1^2}\right]\end{aligned}$$

ただし

$$\begin{aligned}\mu_1 &= \frac{n\sigma^{-2}\bar{X} + \sigma_0^{-2}\mu_0}{n\sigma^{-2} + \sigma_0^{-2}} \\ \sigma_1^{-2} &= n\sigma^{-2} + \sigma_0^{-2},\end{aligned}$$

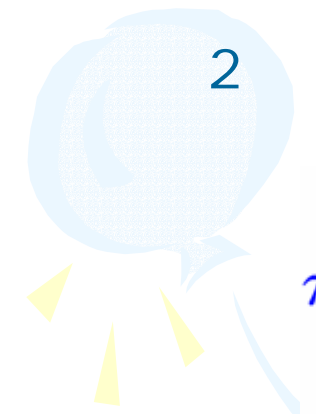


2. ベイズの定理 (続き)

例2.

$X_1, X_2, \dots, X_n \sim N(\mu, \sigma^2)$. μ : known

σ^2 の事前分布はIG ($n_0/2, n_0S_0/2$)


$$\pi(\sigma^2) = \frac{\left(\frac{n_0 S_0}{2}\right)^{\frac{n_0}{2}}}{\Gamma\left(\frac{n_0}{2}\right)} (\sigma^2)^{-\left(\frac{n_0}{2}+1\right)} e^{-\frac{n_0 S_0}{2\sigma^2}}$$

$$\propto (\sigma^2)^{-\left(\frac{n_0}{2}+1\right)} e^{-\frac{n_0 S_0}{2\sigma^2}}$$



2. ベイズの定理 (続き)

事後分布は,

$$\begin{aligned}\pi(\sigma^2|x) &\propto f(x|\sigma^2)\pi(\sigma^2) \\ &\propto (\sigma^2)^{-\left(\frac{n_1}{2}+1\right)} e^{-\frac{n_1 S_1}{2\sigma^2}}\end{aligned}$$



ただし

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



つまり, σ^2 の事後分布は IG ($n_1/2, n_1 S_1/2$)



2. ベイズの定理 (続き)

例3.

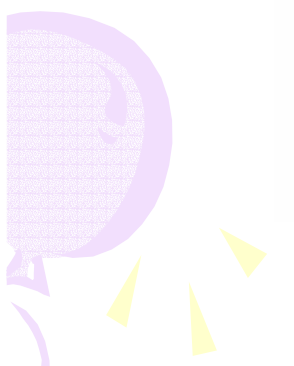
$$X_1, X_2, \dots, X_n \sim N(\mu, \sigma^2).$$



μ | σ^2 の事前分布は $N(\mu_0, \sigma^2/m_0)$.

σ^2 の事前分布は $IG(n_0/2, n_0 S_0/2)$

$$\pi(\mu|\sigma^2) = \frac{\sqrt{m_0}}{\sqrt{2\pi\sigma}} e^{-\frac{m_0(\mu-\mu_0)^2}{2\sigma^2}}$$


$$\pi(\sigma^2) = \frac{\left(\frac{n_0 S_0}{2}\right)^{\frac{n_0}{2}}}{\Gamma\left(\frac{n_0}{2}\right)} (\sigma^2)^{-\left(\frac{n_0}{2}+1\right)} e^{-\frac{n_0 S_0}{2\sigma^2}}$$



2. ベイズの定理 (続き)

事後分布は,

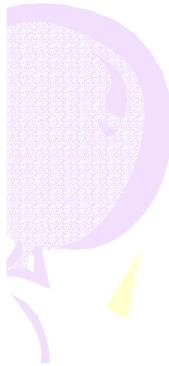
$$\begin{aligned}\pi(\mu, \sigma^2 | x) &\propto f(x | \mu, \sigma^2) \pi(\mu | \sigma^2) \pi(\sigma^2) \\ &\propto \sigma^{-1} (\sigma^2)^{-\left(\frac{n_1}{2} + 1\right)} e^{-\frac{n_1 S_1 + m_1 (\mu - \mu_1)^2}{2\sigma^2}}\end{aligned}$$



ただし

$$m_1 = m_0 + n, \quad n_1 = n_0 + n,$$

$$\mu_1 = \frac{m_0 \mu_0 + n \bar{x}}{m_0 + n},$$


$$n_1 S_1 = n_0 S_0 + \sum_{i=1}^n (x_i - \bar{x})^2 + \frac{m_0 n}{m_0 + n} (\bar{x} - \mu_0)^2$$



2. ベイズの定理 (続き)

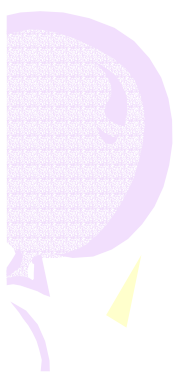
周辺事後分布は、同時分布を積分することにより

$$\sigma^2|x \sim \text{IG} \left(\frac{n_1}{2}, \frac{n_1 S_1}{2} \right)$$



および

$$\pi(\mu|x) \propto \left\{ 1 + \frac{m_1(\mu - \mu_1)^2}{n_1 S_1} \right\}^{-\frac{n_1+1}{2}}$$



他の例では同時事後分布が複雑で積分できないことが多い。これに対して、条件付分布は簡単に計算できることも多い。



2. ベイズの定理 (続き)

条件付事後分布は、

$$\sigma^2 | \mu, x \sim \text{IG} \left(\frac{n_1 + 1}{2}, \frac{n_1 S_1 + m_1 (\mu - \mu_1)^2}{2} \right)$$



および

$$\mu | \sigma^2, x \sim N \left(\mu_1, \frac{\sigma^2}{m_1} \right)$$



ギブス・サンプラーは条件付事後分布を利用して、同時事後分布からの確率標本を発生させる。

3. ギブス・サンプラー

例3 (続き). 同時事後分布 $(\mu, \sigma^2 | x)$ からの確率標本(サンプル)を以下のようなモンテカルロ法(ギブス・サンプラー)により得ることができる。

1. 初期値 $\mu^{(0)}, \sigma^{2(0)}$ を決める。
2. $\mu^{(1)}, \sigma^{2(1)}$ を次のようにサンプリングする。

$$\mu^{(1)} | \sigma^{2(0)}, x \sim N \left(\mu_1, \frac{\sigma^{2(0)}}{m_1} \right)$$

$$\sigma^{2(1)} | \mu^{(1)}, x \sim \text{IG} \left(\frac{n_1 + 1}{2}, \frac{n_1 S_1 + m_1 (\mu^{(1)} - \mu_1)^2}{2} \right)$$

3. ギブス・サンプラー(続き)

3. 同様に $i=1, 2, \dots$ について $\mu^{(i+1)}, \sigma^{2(i+1)}$ を次のようにサンプリングする.

$$\mu^{(i+1)} | \sigma^{2(i)}, x \sim N \left(\mu_1, \frac{\sigma^{2(i)}}{m_1} \right)$$

$$\sigma^{2(i+1)} | \mu^{(i+1)}, x \sim \text{IG} \left(\frac{n_1 + 1}{2}, \frac{n_1 S_1 + m_1 (\mu^{(i+1)} - \mu_1)^2}{2} \right)$$

4. すると $\mu^{(i)}, \sigma^{2(i)}$ ($i=N, N+1, \dots$ で N は十分大きな数) は同時事後分布 $(\mu, \sigma^2 | x)$ からのサンプルとみなすことができる. これを用いて推論を行う(事後分布の平均や密度関数など).

3. ギブス・サンプラー (続き)

例: ソフトウェア WinBUGS を用いると...

- モデルの指定

```
model{
  #  $f(x | \mu, \tau^2) \propto \tau^{-2} \exp(-\frac{x-\mu}{\tau})$  尤度関数
  for (i in 1:N) {
    x[i] ~ dnorm(mu,tau) #  $x \sim N(\mu, \tau^2)$ 
  }
  #  $(\mu, \tau^2)$  事前分布
  tau ~ dgamma(1.0E-3,1.0E-3) #  $\tau^2 \sim IG(0.001,0.001)$ 
  mu ~ dnorm(0.0,1.0E-3) #  $\mu \sim N(0,1000)$ 
  var <- 1/tau
}
```

<http://www.mrc-bsu.cam.ac.uk/bugs/> からダウンロード可能。フリーウェア。

3. ギブス・サンプラー (続き)

- データの指定 (この例では $X_1, \dots, X_{30} \sim N(10, 25)$ のシミュレーションデータ)

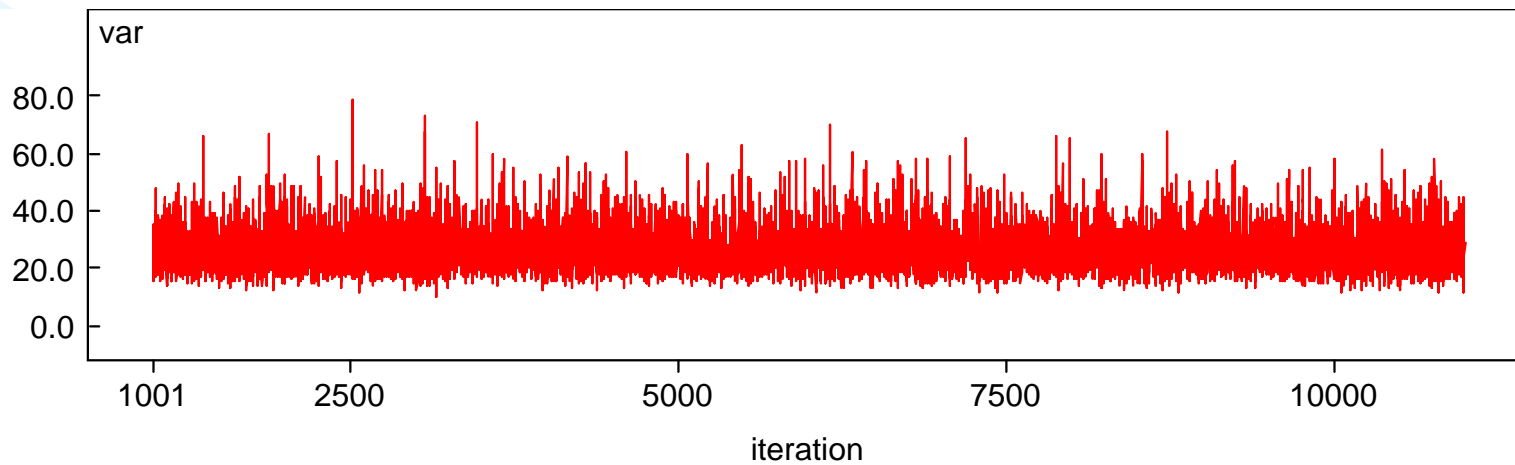
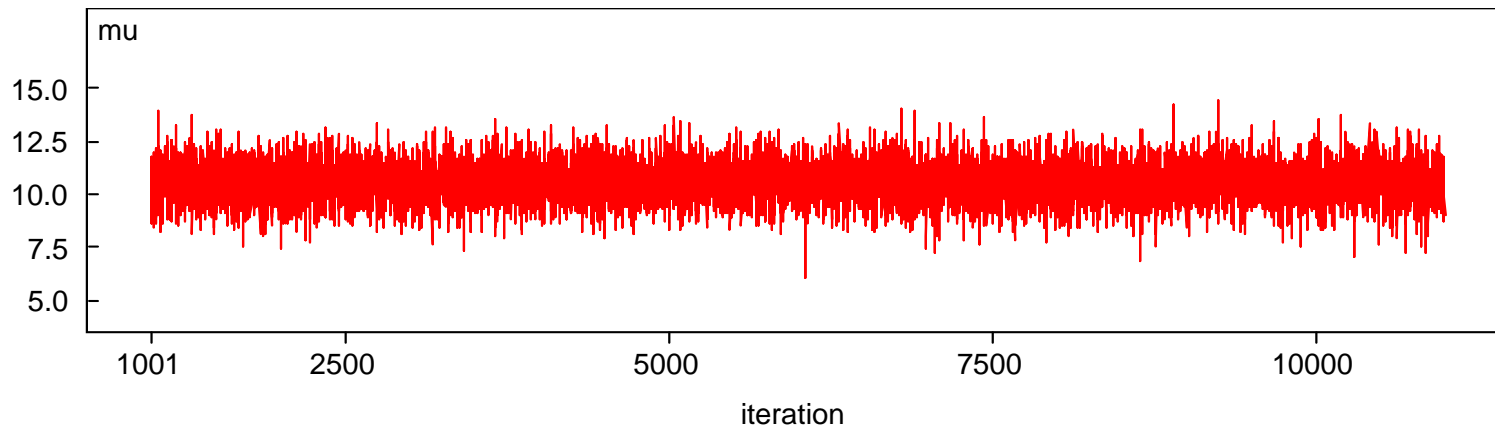
```
list(X=c(13.08, 2.44, 13.82, 14.14, 15.63, 18.99, 4.37, 11.41,  
13.53, 9.43, 12.69, 5.30, 17.73, 4.31, 11.42, -2.72, 14.15,  
9.56, 18.29, 12.09, 11.16, 8.93, 6.14, 4.72, 7.99, 7.80, 11.35,  
11.76, 13.62, 14.47), N=30)
```

- 初期値の指定 (初期値は適当に決めるが、あまりデータと整合的でないと、同時事後分布からのサンプルを得るまでに時間がかかる)

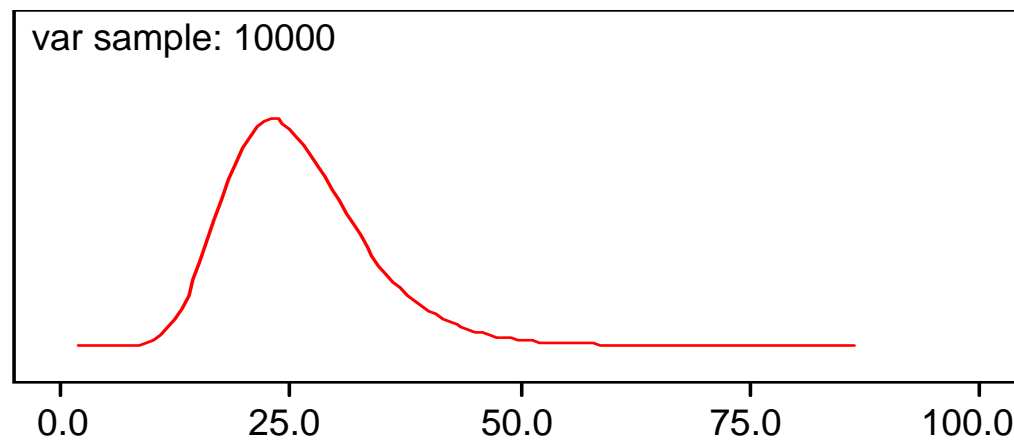
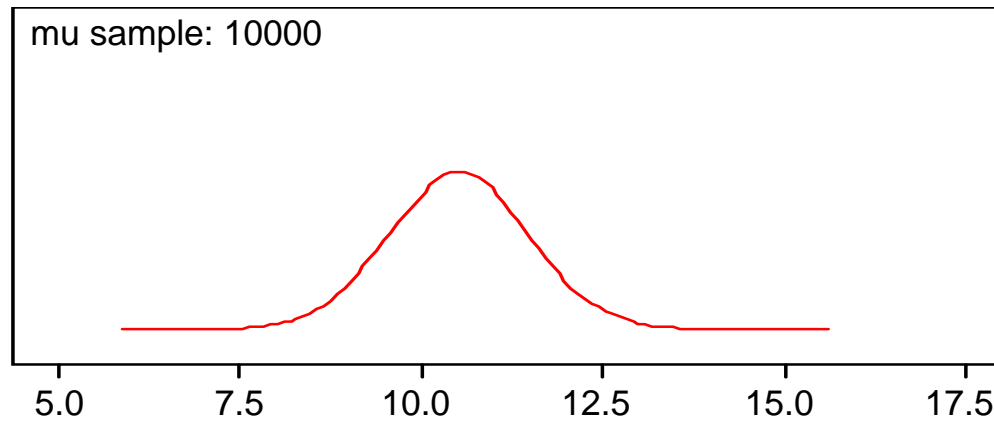
```
list(mu=10 tau=1/25)
```

- 最初の1000回を初期値に依存する期間 (burn-in period) として捨て、残りの10000回を用いる。

3. ギブス・サンプラー (続き)



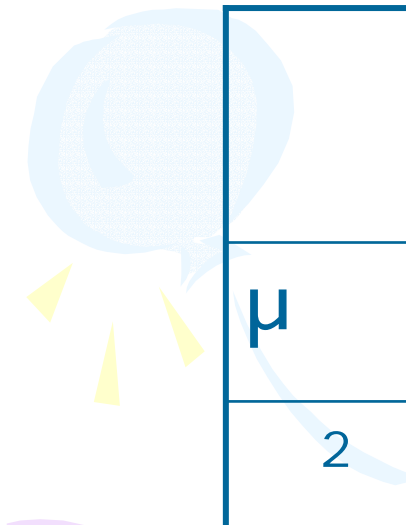
3. ギブス・サンプラー (続き)





3. ギブス・サンプラー (続き)

- 計算結果のまとめ

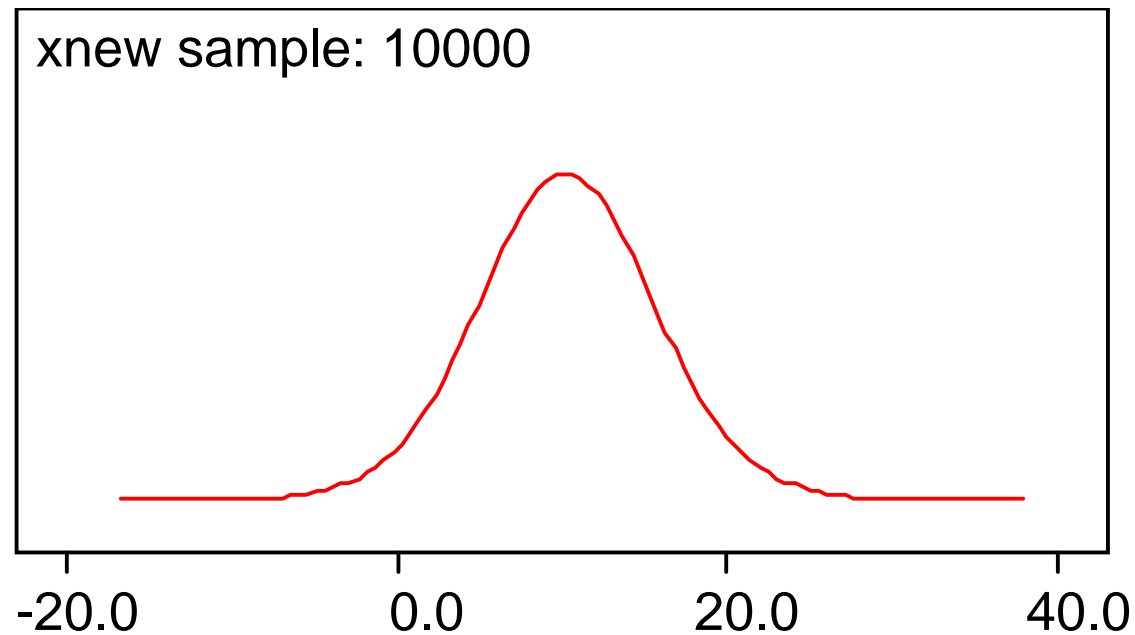


	事後 平均	標準 偏差	MC error	2.5 %	中央 値	97.5 %
μ	10.56	0.918	0.0093	8.76	10.56	12.38
2	26.40	7.415	0.0681	15.62	25.22	42.26

3. ギブス・サンプラー(続き)

- 新しい X の予測分布

$X_{\text{new}}^{(i)} \sim N(\mu^{(i)}, \sigma^{2(i)})$ をプログラムの最後につければよい。



3. ギブス・サンプラー(まとめ)

事後分布 $(x_1, \dots, x_m | X)$ からのサンプリング.

1. まず初期値 $(x^{(0)}) = (x_1^{(0)}, \dots, x_m^{(0)})$ を選ぶ.
2. $i=0, 1, 2, \dots$ について以下を繰り返す.
 - a. $x_1^{(i+1)} \sim (x_1 | x_2^{(i)}, \dots, x_m^{(i)} | X)$
 - b. $x_2^{(i+1)} \sim (x_2 | x_1^{(i+1)}, x_3^{(i)}, \dots, x_m^{(i)} | X)$
 - c. ...
 - d. $x_m^{(i+1)} \sim (x_m | x_2^{(i+1)}, \dots, x_{m-1}^{(i+1)} | X)$
3. $i > N$ (十分大きな数)について $x^{(i)}$ を保存.

$(x^{(i)} (i=1, 2, \dots))$ はマルコフ連鎖になっている(「マルコフ連鎖」モンテカルロ法).

ここでは条件付事後分布がよく知られた分布であり、かつ、そこからの乱数発生が容易であることが必要となる。そうでなければMH(メトロポリス・ヘイスティングス)アルゴリズムを用いることになる。



4. MHアルゴリズム

MH(メトロポリス・ヘイスティングス)アルゴリズム

- 一般的に**目標分布** (target distribution) である事後分布に適用できるサンプリングのためのアルゴリズム
- **提案分布** (proposal distribution) を用いたサンプリング
- 得られた**サンプル候補は棄却し捨てることもある**
- ギブス・サンプラーもMHアルゴリズムの特別なケース(サンプルを捨てない)

4. MHアルゴリズム(続き)

事後分布 $(\theta_1, \dots, \theta_m | x)$ からのサンプリング.

1. 初期値 $\theta^{(0)} = (\theta_1^{(0)}, \dots, \theta_m^{(0)})$ を選ぶ.
2. $i=0, 1, 2, \dots$ について以下を繰り返す.
 1. $\theta^{(i)} = (\theta_1^{(i)}, \dots, \theta_m^{(i)})$ が与えられたときに提案分布 Q を用いて候補 θ' を発生させる. Q の密度関数(提案密度)を $q(\theta' | \theta^{(i)}, x)$ とおく($q(\theta' | \theta^{(i)}, x)$ と書くこともある).
 2. 提案された θ' は目標分布 $(\theta | x)$ からのサンプルではないので, 目標分布からのズレを修正するために, 確率

$$\alpha(\theta^{(i)}, \theta' | x) = \min \left[\frac{\pi(\theta' | x) q(\theta^{(i)}, \theta' | x)}{\pi(\theta^{(i)} | x) q(\theta', \theta^{(i)} | x)}, 1 \right]$$

で採択し, $\theta^{(i+1)} = \theta'$ とする. 棄却したら $\theta^{(i+1)} = \theta^{(i)}$ とする.

3. $i > N$ (十分大きな数) について $\theta^{(i)}$ を保存.

4. MHアルゴリズム(続き)

MHアルゴリズムの特徴

- 事後密度の高い候補点が提案されれば必ず移動し、そうでなければ密度の比の確率で移動する.
- 事後密度の低い候補点についても小さい確率ではあるけれども、移動する確率がある.
- 詳細釣り合方程式(detailed balance equation)を満たす.

$$\pi(\theta^{(i)}|x)p(\theta^{(i)}, \theta'|x) = \pi(\theta'|x)p(\theta', \theta^{(i)}|x)$$

- 近似のよい提案分布を構成できなければ、なかなか目標分布に収束しない.

どうしたらよい提案分布を見つけることができるのか？

4. MHアルゴリズム(続き)

どうやって提案分布を選ぶのか？

1. 酔歩過程MHアルゴリズム.

酔歩過程(random walk)を用いて $x^{(i+1)} = x^{(i)} + z$, $z \sim f(z)$ とする. このとき, 提案密度は $q(x^{(i)}, x^{(i+1)} | x) = f(x^{(i+1)} - x^{(i)})$.

- z の分布には平均0の多変量正規分布や多変量t分布がよく用いられるが, 分散をどうとるのが難しい.
- 分散を小さくとると状態空間の一部の周辺でしか候補点を提案しないので, 状態空間全体からのサンプリングが進まない.
- 分散を大きくとると目標分布である事後分布の密度が低い点が提案されやすく, 受容率が下がってしまう.
- いずれにしても $x^{(i)}$ ($i=1, 2, \dots$) の自己相関が高くなり, 非効率的なサンプリングである.

4. MHアルゴリズム(続き)

2. 独立(independence) MHアルゴリズム. $q(\theta^{(i)} | x)$ の形。

- q は $\theta^{(i)}$ に依存しない。
- 候補点 θ' の受容確率は

$$\alpha(\theta^{(i)}, \theta' | x) = \min \left\{ \frac{w(\theta' | x)}{w(\theta^{(i)} | x)}, 1 \right\},$$

$$w(\theta | x) = \frac{\pi(\theta | x)}{q(\theta | x)}$$

- w は, 目標密度 π と提案密度 q の比. 提案密度に比べて相対的に目標密度の方が高い点を選ばれやすい。

4. MHアルゴリズム(続き)

- 提案分布としては, 例えば目標分布のモードないしその近似値を θ^* として正規分布 $N(m, V)$ で近似

$$m = \theta^* + V \left. \frac{\partial \log \pi(\theta|y)}{\partial \theta} \right|_{\theta=\theta^*}, \quad V = \left\{ - \left. \frac{\partial^2 \log \pi(\theta|y)}{\partial \theta \partial \theta'} \right|_{\theta=\theta^*} \right\}^{-1}$$

- この近似分布は, 対数事後密度 $\log \pi(\theta|y)$ を θ^* のまわりで2次の項まで展開することによって得られる.
- しかし, 多峰型の事後分布の近似にはよくない(混合形にする方法もある).
- V は

$$V = v_0 \left\{ - \left. \frac{\partial^2 \log \pi(\theta|y)}{\partial \theta \partial \theta'} \right|_{\theta=\theta^*} \right\}^{-1}$$

とすることによって, 受容率が高くなるように試行錯誤で v_0 の値を変えることもできる.



4. MHアルゴリズム(例)

例: 1階の自己回帰過程

$$y_t = \phi y_{t-1} + \epsilon_t, \quad \epsilon_t \sim \text{i.i.d. } N(0, \sigma^2),$$

$$t = 1, 2, \dots, n,$$

$$y_0 \sim N(0, \sigma^2 / (1 - \phi^2)), \quad |\phi| < 1$$



事前分布

– $\phi \sim U(-1, 1)$

– $\sigma^2 \sim \text{IG}(n_0/2, n_0 S_0/2)$

4. MHアルゴリズム (例続き)

- 事後分布

$$\begin{aligned}\pi(\phi, \sigma^2 | y) &\propto f(y | \phi, \sigma^2) \pi(\phi) \pi(\sigma^2) \\ &\propto \frac{\sqrt{1 - \phi^2}}{\sqrt{2\pi}\sigma} e^{-\frac{(1 - \phi^2)y_0^2}{2\sigma^2}} \times \sigma^{-n} e^{-\frac{1}{2\sigma^2} \sum_{t=1}^n (y_t - \phi y_{t-1})^2} \\ &\quad \times (\sigma^2)^{-\left(\frac{n_0}{2} + 1\right)} e^{-\frac{n_0 S_0}{2\sigma^2}} \times I_\phi(-1, 1)\end{aligned}$$

- 条件付事後分布

$$\begin{aligned}- \sigma^2 | y &\sim \text{IG}(n_1/2, n_1 S_1/2) \\ n_1 &= n_0 + n + 1,\end{aligned}$$

$$n_1 S_1 = n_0 S_0 + (1 - \phi^2) y_0^2 + \sum_{t=1}^n (y_t - \phi y_{t-1})^2$$

4. MHアルゴリズム (例続き)

– の場合には既知の分布ではなく

$$\begin{aligned}\pi(\phi|\sigma^2, y) &\propto \sqrt{1-\phi^2} e^{-\frac{(1-\phi^2)}{2\sigma^2}y_0^2 - \frac{1}{2\sigma^2}\sum_{t=1}^n (y_t - \phi y_{t-1})^2} I_\phi(-1, 1) \\ &\propto \sqrt{1-\phi^2} e^{-\frac{1}{2\sigma_\phi^2}(\phi - \mu_\phi)^2} I_\phi(-1, 1) \\ &\propto \sqrt{1-\phi^2} \times N(\mu_\phi, \sigma_\phi^2) \times I_\phi(-1, 1)\end{aligned}$$

ただし

$$\mu_\phi = \frac{\sum_{t=1}^n y_t y_{t-1}}{\sum_{t=1}^{n-1} y_t^2}, \quad \sigma_\phi^2 = \frac{\sigma^2}{\sum_{t=1}^{n-1} y_t^2}$$

提案分布は切断正規分布 $TN_{(-1,1)}(\mu, \sigma^2)$!

4. MHアルゴリズム (例続き)

- 提案分布に従って $\phi' \sim \text{TN}_{(-1,1)}(\mu, \sigma^2)$ を発生させると、

$$\begin{aligned} \frac{w(\phi'|\sigma^2, y)}{w(\phi^{(i)}|\sigma^2, y)} &= \frac{\pi(\phi|\sigma^2, y)/q(\phi'|\sigma^2, y)}{\pi(\phi^{(i)}|\sigma^2, y)/q(\phi^{(i)}|\sigma^2, y)} \\ &= \frac{\sqrt{1 - \phi'^2}}{\sqrt{1 - \phi^{(i)2}}} \end{aligned}$$

となるので、確率

$$\alpha(\phi^{(i)}, \phi'|\sigma^2, y) = \min \left\{ \frac{\sqrt{1 - \phi'^2}}{\sqrt{1 - \phi^{(i)2}}}, 1 \right\}$$

で ϕ' を採択すればよい。

4. MHアルゴリズム (例続き)

サンプリングの手順

1. 初期値 ($\phi^{(0)}$, $\sigma^2(0)$) を決める.

2. $i=0, 1, 2, \dots$ について以下を繰り返す.

– $\sigma^2(i+1) | \phi^{(i)}, y \sim \text{IG}(n_1/2, n_1 S_1/2)$

– $\phi' | \sigma^2(i+1), y \sim \text{TN}_{(-1,1)}(\mu, \sigma^2)$ を発生させ, 確率

$$\alpha(\phi^{(i)}, \phi' | \sigma^2, y) = \min \left\{ \frac{\sqrt{1 - \phi'^2}}{\sqrt{1 - \phi^{(i)2}}}, 1 \right\}$$

で受容し $\phi^{(i+1)} = \phi'$ とし, 棄却されたら $\phi^{(i+1)} = \phi^{(i)}$ とする.

3. $i > N$ (十分大きな数) について ($\phi^{(i)}$, $\sigma^2(i)$) を保存.

4. MHアルゴリズム (例続き)

計算例 (Ox言語 による) 教育用には無料で配布されている.

(詳しくは<http://www.e.u-tokyo.ac.jp/~omori>を参照).

```
for(i=0;i<10000; ++i){
  // 2のサンプリング
  a=0.5*(0.001+nobs+1);
  b=0.5*(0.001+(1-phi^2)*y[0]+sumsqrc(y[1:nobs-1]-phi*y[0:nobs-2]));
  var_e=1/rangamma(1,1,a,b);
  // のサンプリング
  mu_phi=(y[0:nobs-2]*y[1:nobs-1])/sumsqrc(y[1:nobs-2]);
  var_phi=var_e/sumsqrc(y[1:nobs-2]);
  // ' | 2(i+1), y ~ TN_{(-1,1)}(\mu, \sigma^2)
  do{phi_n=mu_phi+rann(1,1)*sqrt(var_phi);} while(fabs(phi_n)>=1);
  // 確率 で受容する.
  if(ranu(1,1)<= (0.5*sqrt(1-phi_n^2))/(0.5*sqrt(1-phi^2))){phi=phi_n;}
}
```


4. MHアルゴリズム (例続き)

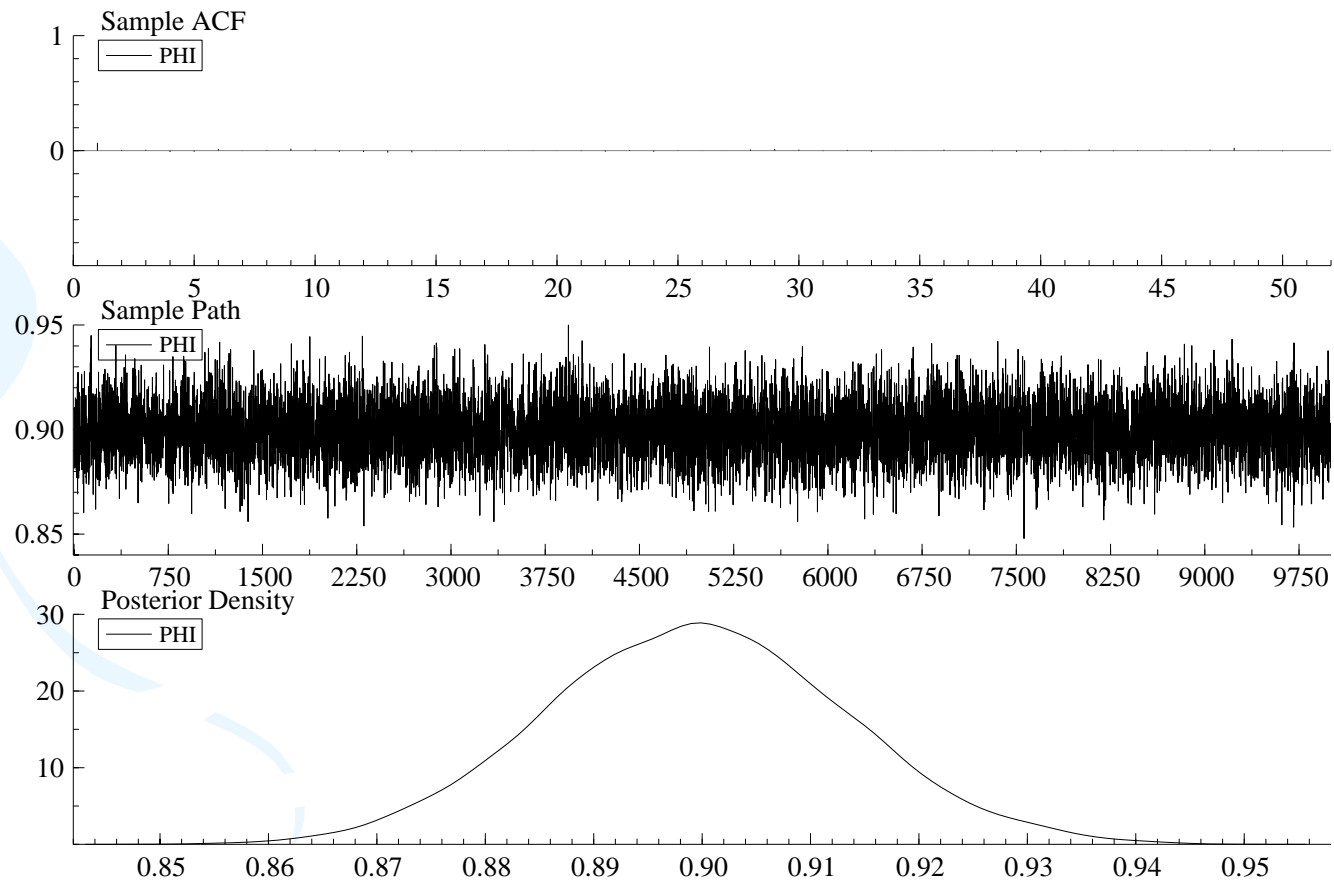
計算結果

(データは真の値を $\theta = 0.9$, $\sigma^2 = 1.0$ として
 y_0, \dots, y_{1000} を発生させた)

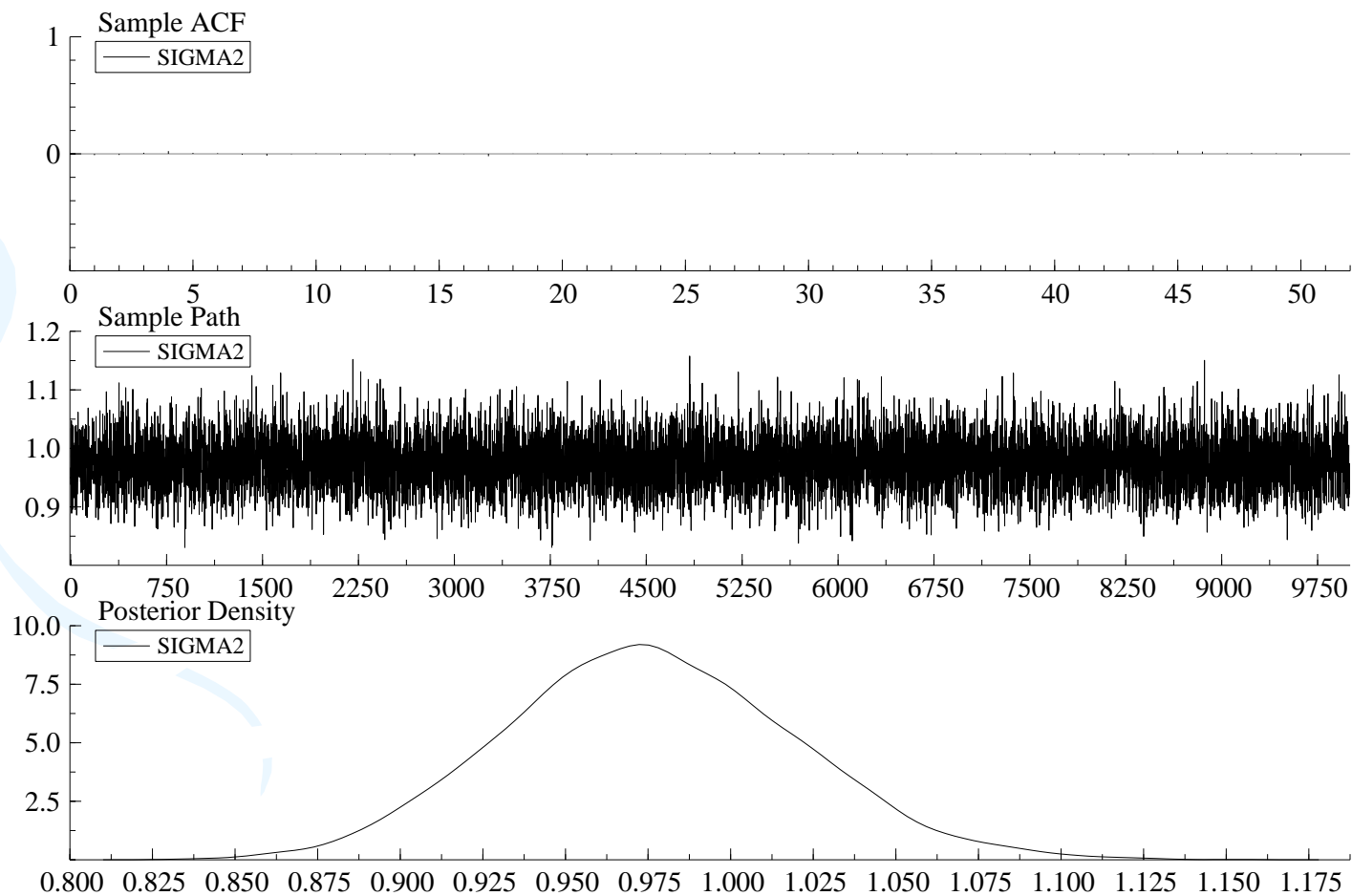
提案された θ の採択率は96.3%と非常に高い。

	事後平均	標準偏差	2.5%	97.5%
	0.899	0.014	0.873	0.927
2	0.976	0.044	0.892	1.065

4. MHアルゴリズム (例続き)



4. MHアルゴリズム (例続き)



4. MHアルゴリズム (例続き)

例: ソフトウェアWinBUGSを用いると(ただしWinBUGSはこの例をSlice samplerという方法でサンプリングしている。)

- モデルの指定

```
# Model 観測値:  $y_0, \dots, y_{999}$  の1000個
model{
  tau0 <- tau*(1-phi*phi) #  $f(y | \tau^2) = \tau^{-2}$  尤度関数
  y[1] ~ dnorm(0,tau0) #  $y_0 \sim N(0, \tau^2/(1-\tau^2))$ 
  for( i in 2 : N ) {
    ymu[i] <- phi*y[i-1] #  $y_t \sim N(y_{t-1}, \tau^2)$ 
    y[i] ~ dnorm(ymu[i],tau)}
  phi ~ dunif(-1,1) #  $\phi \sim U(-1,1)$ 
  tau ~ dgamma(1.0E-3,1.0E-3) #  $\tau^2 \sim IG(0.001,0.001)$ 
  var <- 1/tau
}
```

4. MHアルゴリズム(例続き)

- データの指定(この例では $\mu=0.9$, $\sigma^2=1.0$ のシミュレーションデータ)

```
⇒ list(N=1000,y=c(
  -1.4958246547,
  .....
  -0.2742255020)) ←
```

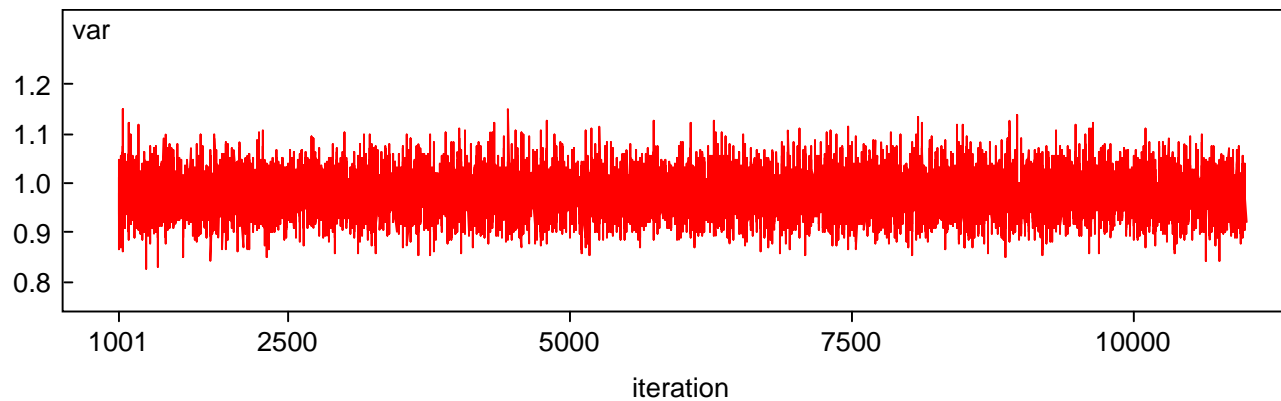
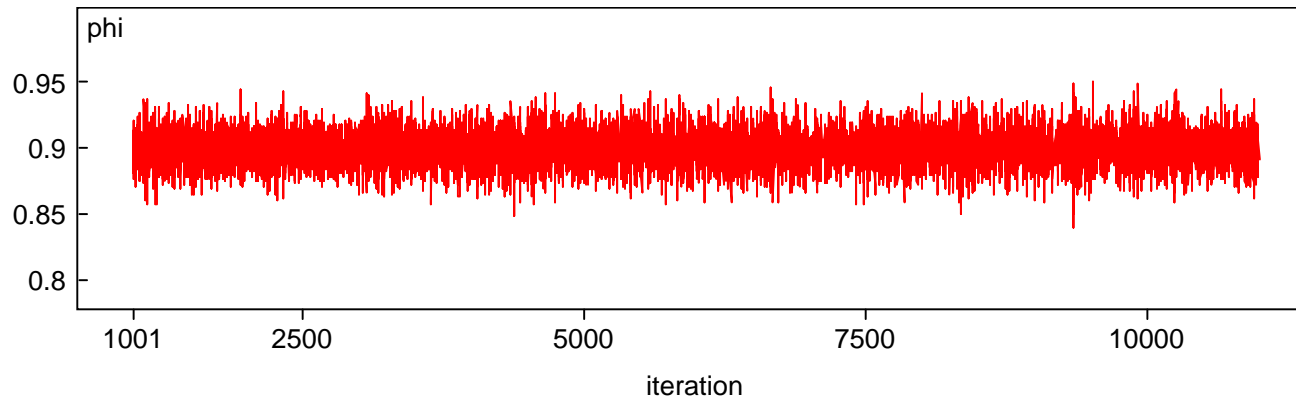
観測値の個数が多いのでFoldという矢印で囲む形式を使う。

- 初期値の指定(初期値は適当に決めるが、あまりデータと整合的でないと、同時事後分布からのサンプルを得るまでに時間がかかる)

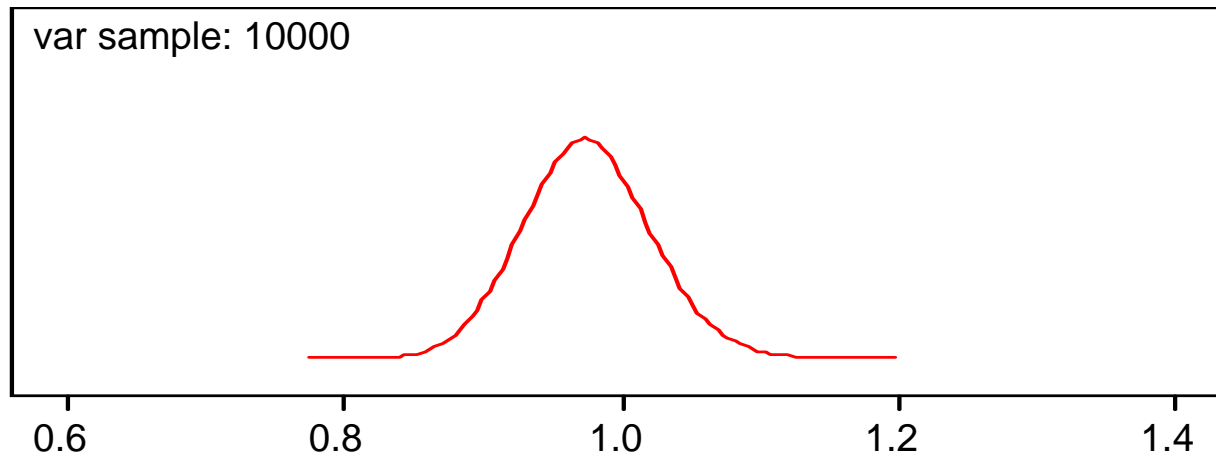
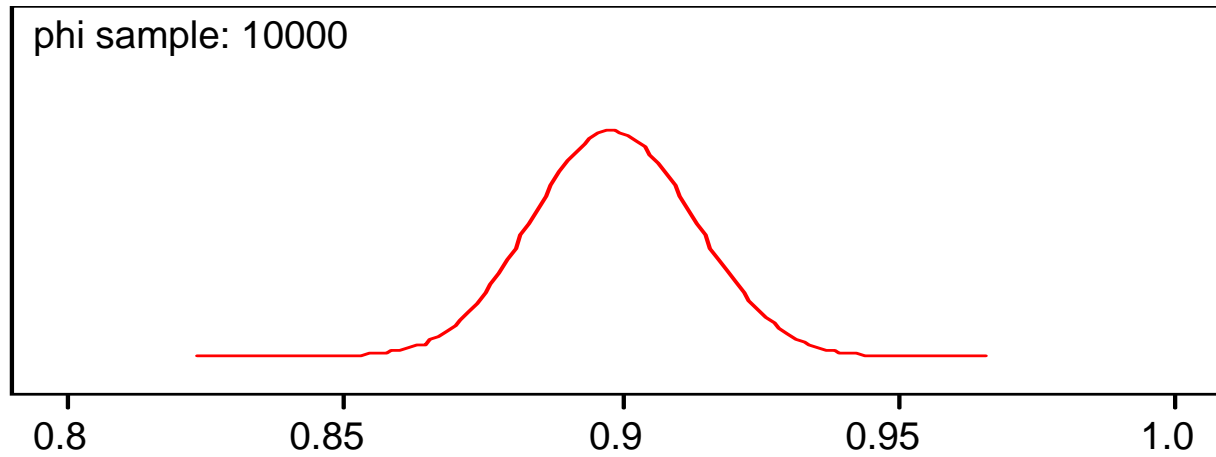
```
list(phi=0 tau=1)
```

- 最初の1000回を初期値に依存する期間(burn-in period)として捨て、残りの10000回を用いる。

4. MHアルゴリズム (例続き)



4. MHアルゴリズム (例続き)



4. MHアルゴリズム (例続き)

- 計算結果のまとめ

	事後平均	標準偏差	MC error	2.5 %	中央値	97.5 %
	0.899	0.014	0.0001	0.873	0.899	0.927
2	0.978	0.044	0.0004	0.895	0.977	1.068



5. 収束判定の方法

- 図による方法

- 標本系列や標本平均の時系列プロット

得られた標本の変動が初期値に依存せず安定的であるかどうかを見て、収束したかどうかを判定

- 標本自己相関関数のプロット

自己相関が高いときには、状態空間の一部からあまり移動できていないことを意味する

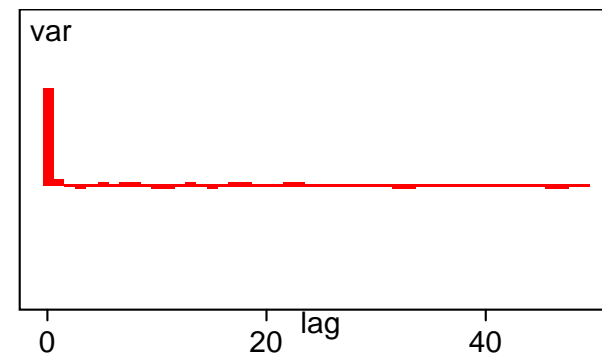
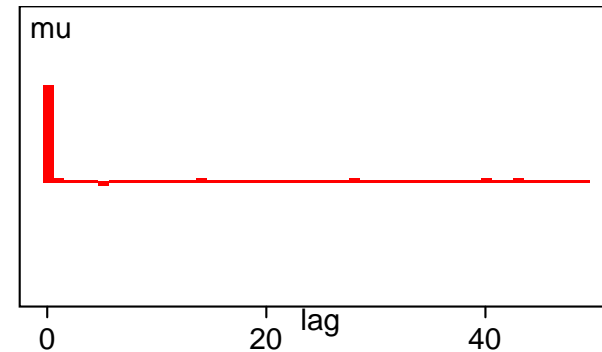
- 標本平均・百分位数の時系列プロット

標本平均・百分位数が安定的になってきているかを観測

5. 収束判定の方法(続き)

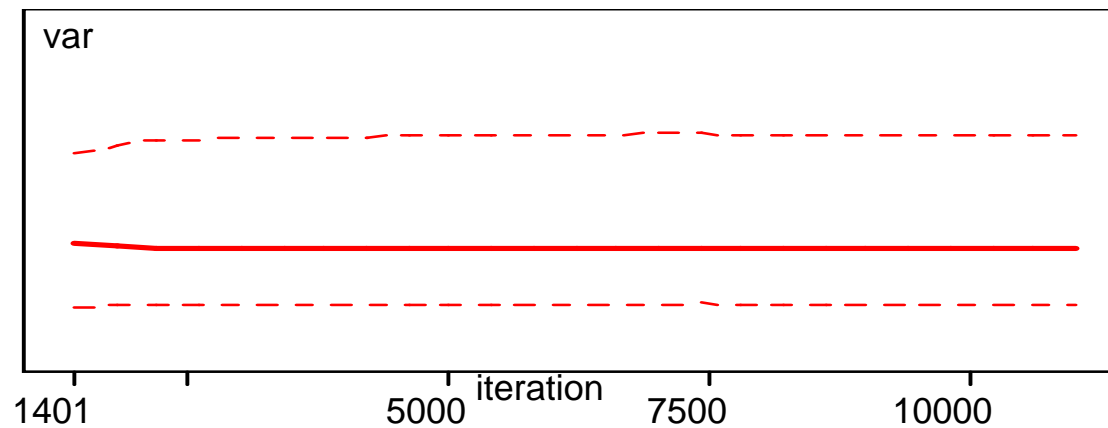
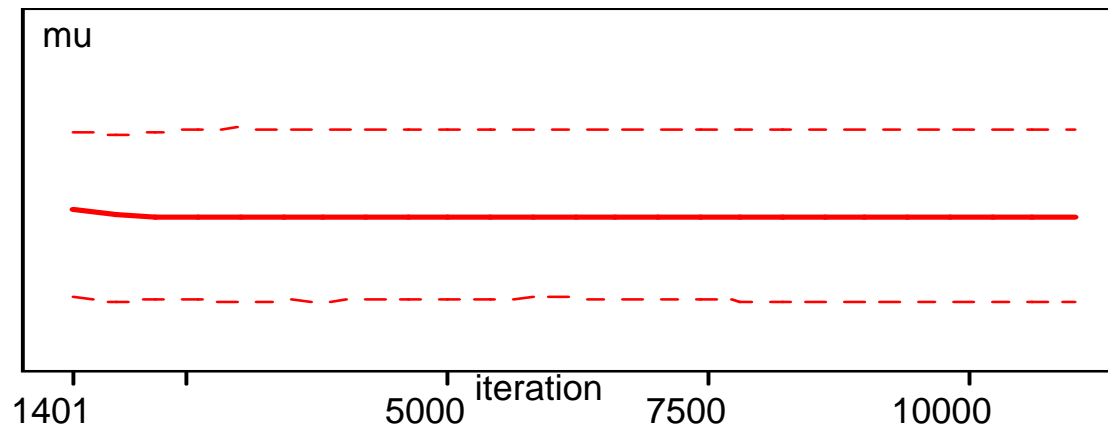
「3. ギブス・サンプラー」における例で、時系列プロットはすでに紹介した。標本自己相関関数を見ると、ラグ1以降はほとんど相関がない。

目標分布への収束が早い。



5. 収束判定の方法 (続き)

標本平均の時系列プロット





5. 収束判定の方法 (続き)

- 数値による方法

- Geweke (1992)

- 得られた標本系列の前半(10%)と後半(50%)でパラメータ θ の関数 $g(\theta)$ の期待値が同じかどうかを仮説検定.

- Gelfand and Smith (1990), Gelman and Rubin (1992), Liu and Liu (1993)

- 異なる初期値からいくつかの標本系列を別々にサンプリングして, 標本系列間の安定性に基づいて収束判定を行う

- Raftery and Lewis (1992)

- 標本百分位数の収束を用いる

- Heidelberger and Welch (1983)

- マルコフ連鎖の一様エルゴード性の検定

- Mykland et al. (1995)

- 確率過程の再生時間の概念を用いるもの



5. 収束判定の方法 (続き)

Cowles and Carlin (1996) は13種類の方法を実験に基づいて比較. 他にRobert and Casella (1999)やMengersen, Robert and Guihenneuc-Jouyaux(1999)がある.



これらの方法の多くはS-PlusやRの関数群で行うことができる. いずれもインターネットのサイトからダウンロードすることができる.

- CODA (Convergence Diagnostic and Output Analysis Software)
- BOA (Bayesian Output Analysis)
- Geweke, Gelman and Rubin, Raftery and Lewis, Heidelberger and Welchの診断方法
- 標本自己相関関数, 標本系列のプロット, 事後密度関数, 事後平均, 標準誤差, 95%信用区間を出力



5. 収束判定の方法 (続き)

- BOAによるGeweke(1992)の結果例
- 帰無仮説: 期待値が等しい(収束)

GEWEKE CONVERGENCE DIAGNOSTIC:

=====

Fraction in first window = 0.1

Fraction in last window = 0.5

Chain: sample1

	mu	var
Z-Score	0.9682253	-0.9140632
p-value	0.3329318	0.3606836

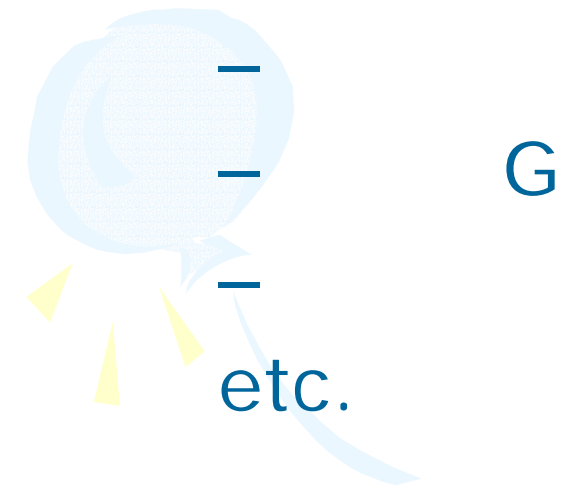
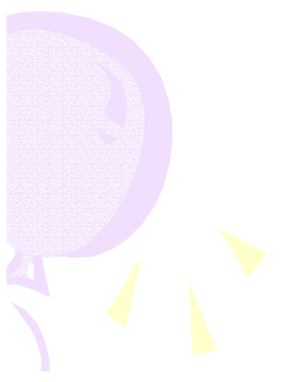
帰無仮説を受容

収束



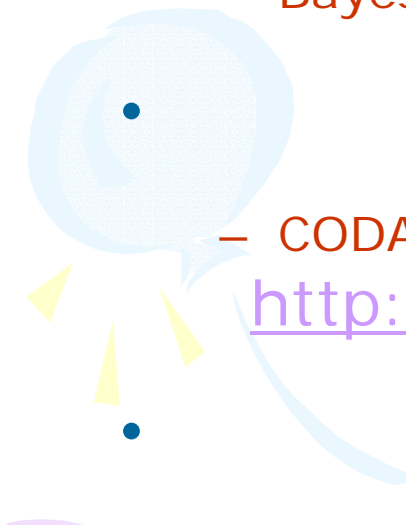
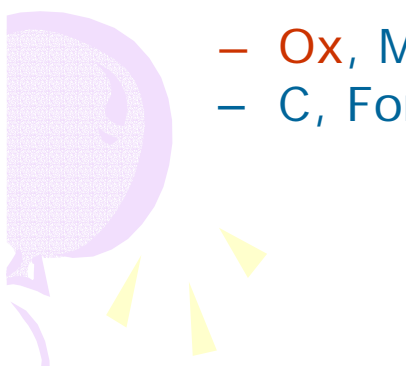
5. 収束判定の方法 (続き)

収束を早めるには…

- ブロック・サンプリング
 - 変数の変換
 - 一般化Gibbs Sampler
 - パラレル・テンパリング
- etc.
- 
- 



6. MCMCのためのソフトウェア

- プログラミングが容易だが、カスタマイズしにくいもの
 - WinBUGS (最もよく用いられている.フリーウェア)
 - BayesX (TSPやSTATAなどに近い.フリーウェア)
 - 収束判定のための関数
 - CODA, BOA (R言語およびSplus言語のための関数.フリー)
(<http://www.public-health.uiowa.edu/boa/>)
 - プログラミングがやや難しいが、カスタマイズしやすいもの
 - Ox, Matlab, GAUSS(行列言語, Oxは教育目的にはフリー)
 - C, Fortran
- 
- 



II. 統計科学への応用

1. 回帰モデル(誤差項が正規分布やt分布)
2. トービットモデル(打ち切りのある回帰モデル)
3. プロビットモデル
4. ロジットモデル
5. 見かけ上無関係な回帰モデル



1. 回帰モデル

線形回帰モデル

$$y_i = \beta_1 x_{i1} + \cdots + \beta_K x_{iK} + \epsilon_i,$$

- y_i : 確率変数で被説明変数, $i=1, \dots, n$
- $x_i = (x_{i1}, \dots, x_{iK})'$: 既知の定数で説明変数
- $\beta = (\beta_1, \dots, \beta_K)'$: 未知のパラメータ
- $\epsilon_i \sim N(0, \sigma^2)$ は確率変数で誤差項.



1. 回帰モデル(続き)

- 回帰モデルは

$$y|\beta, \sigma^2 \sim N(X\beta, \sigma^2 I)$$

$$y = (y_1, \dots, y_n), \quad X = (x_1, \dots, x_n)'$$

と表現でき、尤度関数は

$$f(y|\beta, \sigma^2) \propto \sigma^{-n} e^{-\frac{1}{2\sigma^2}(y-X\beta)'(y-X\beta)}$$

- 事前分布

$$\beta \sim N(b_0, B_0), \quad \sigma^2 \sim IG\left(\frac{n_0}{2}, \frac{n_0 S_0}{2}\right)$$



1. 回帰モデル(続き)

- 事後分布

$$\beta|\sigma^2, y \sim N(b_1, B_1),$$

$$\sigma^2|\beta, y \sim IG\left(\frac{n_1}{2}, \frac{n_1 S_1}{2}\right)$$



ただし

$$n_1 = n_0 + n,$$

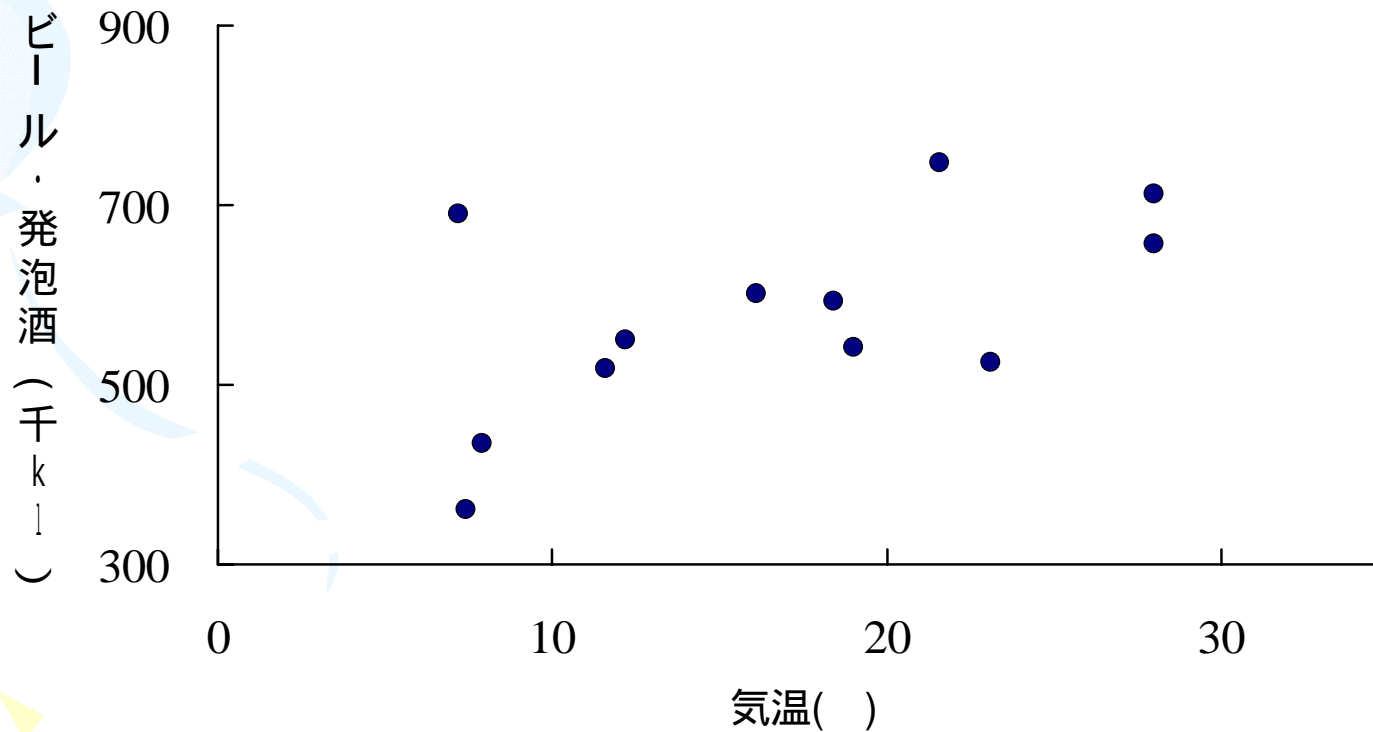
$$n_1 S_1 = n_0 S_0 + (y - X\beta)'(y - X\beta),$$

$$b_1 = B_1(B_0^{-1}b_0 + \sigma^{-2}X'y),$$

$$B_1^{-1} = B_0^{-1} + \sigma^{-2}X'X$$

1. 回帰モデル(続き)

(例) 気温(Y)とビール・発泡酒の課税数量(千キロリットル)(X)の関係



1. 回帰モデル(続き)

例: WinBUGSのコード

- モデルの指定

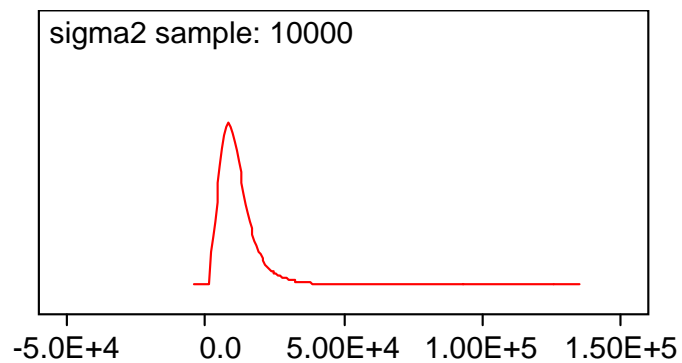
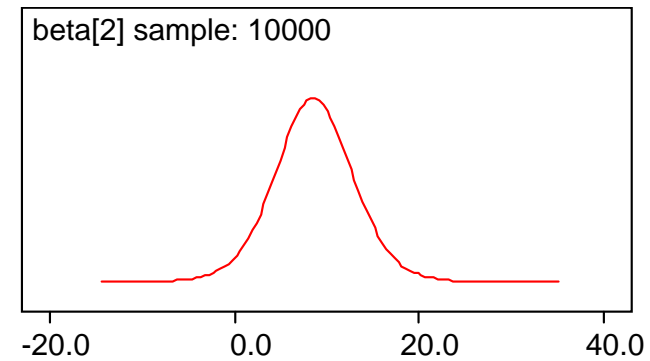
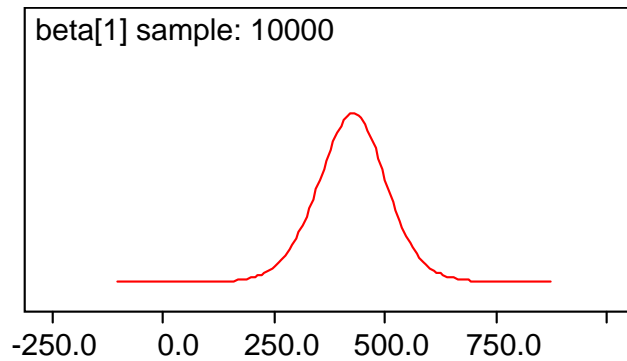
```
model{
  #  $f(x | \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$  尤度関数
  for(i in 1:n){ #nは観測値の個数
    mu[i] <- beta[1]+beta[2]*x[i] #  $\mu_i = \beta_1 + \beta_2 x_i$ 
    y[i] ~ dnorm(mu[i],tau) #  $y_i \sim N(\mu_i, \sigma^2)$ 
  }
  #  $(\beta, \sigma^2)$  事前分布
  for(i in 1:K){ #Kはベータの次元=2
    beta[i] ~ dnorm(0,1.0E-6) #  $\beta_i \sim N(0, 10^6)$ 
    tau ~ dgamma(1.0E-6,1.0E-6)
    sigma2 <- 1/tau #  $\sigma^2 \sim IG(10^{-6}, 10^{-6})$ 
  }
}
```

1. 回帰モデル(続き)

- 計算結果のまとめ

	事後平均	標準偏差	MC error	2.5 %	中央値	97.5 %
1	429.2	78.34	0.833	271.3	429.9	584.7
2	8.813	4.305	0.045	0.286	8.763	17.42
2	11760	6550	78.13	4593	10150	28860

1. 回帰モデル(続き)





1. 回帰モデル(続き)

- y_{n+1} の予測分布は...

$$\begin{aligned} \pi(y_{n+1}|x_{n+1}, y) \\ = \int f(y_{n+1}|\beta, \sigma^2, x_{n+1})\pi(\beta, \sigma^2|x, y)d\beta d\sigma^2 \end{aligned}$$

- ギブス・サンプラーの最後のステップで、得られた $(\hat{\beta}, \hat{\sigma}^2)$ のもとで $y_{n+1} | \hat{\beta}, \hat{\sigma}^2, x_{n+1} \sim N(x_{n+1}, \hat{\sigma}^2)$ を発生させればよい。

1. 回帰モデル(続き)

誤差項がt分布であるような回帰モデル

- 標準的な回帰モデルでは, 誤差項は $\epsilon_i \sim N(0, \sigma^2)$ と仮定した.
- これを正規分布ではなく, t分布にしたモデルも容易に扱うことができる.

$$\epsilon_i | \lambda_i \sim N(0, \lambda_i^{-1} \sigma^2), \quad \lambda_i \sim G\left(\frac{\nu_0}{2}, \frac{\nu_0}{2}\right)$$

とおくと (ν_0 は既知の定数), ϵ_i の周辺分布は自由度 ν_0 のt分布になることを利用する.

1. 回帰モデル(続き)

- 回帰モデルは

$$y|\beta, \sigma^2, \Lambda \sim N(X\beta, \sigma^2\Lambda^{-1})$$

$$y = (y_1, \dots, y_n), \quad X = (x_1, \dots, x_n)'$$

$$\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$$

と表現でき、尤度関数は

$$f(y|\beta, \sigma^2, \Lambda) \propto |\Lambda|^{1/2} \sigma^{-n} e^{-\frac{1}{2\sigma^2}(y-X\beta)'\Lambda(y-X\beta)}$$

- 事前分布

$$\beta \sim N(b_0, B_0), \sigma^2 \sim IG\left(\frac{n_0}{2}, \frac{n_0 S_0}{2}\right), \lambda_i \sim G\left(\frac{\nu_0}{2}, \frac{\nu_0}{2}\right)$$



1. 回帰モデル(続き)

- 事後分布

$$\beta|\sigma^2, \Lambda, y \sim N(b_1, B_1),$$

$$\sigma^2|\beta, \Lambda, y \sim IG\left(\frac{n_1}{2}, \frac{n_1 S_1}{2}\right)$$

$$\lambda_i|\beta, \sigma^2, y \sim G\left(\frac{\nu_0 + 1}{2}, \frac{\nu_0 + \sigma^{-2}(y_i - x_i'\beta)^2}{2}\right)$$



ただし

$$n_1 = n_0 + n,$$

$$n_1 S_1 = n_0 S_0 + (y - X\beta)' \Lambda (y - X\beta),$$

$$b_1 = B_1 (B_0^{-1} b_0 + \sigma^{-2} X' \Lambda y),$$

$$B_1^{-1} = B_0^{-1} + \sigma^{-2} X' \Lambda X$$

2. トービットモデル

トービットモデル(打ち切りのある回帰モデル)

- 被説明変数 y_i が,ある閾値を超えたときのみを観測される回帰モデル. 閾値を0とすれば,

$$y_i^* = \beta_1 x_{i1} + \cdots + \beta_K x_{iK} + \epsilon_i,$$

$$y_i = \max(0, y_i^*)$$

- y_i^* : 確率変数で被説明変数, $i=1, \dots, n$
- y_i : $y_i^* > 0$ のとき観測され, そうでないとき観測されない.
- $x_i = (x_{i1}, \dots, x_{iK})'$: 既知の定数で説明変数
- $\beta = (\beta_1, \dots, \beta_K)'$: 未知のパラメータ
- $\epsilon_i \sim N(0, \sigma^2)$ は確率変数で誤差項.



2. トービットモデル(続き)

- トービットモデルは

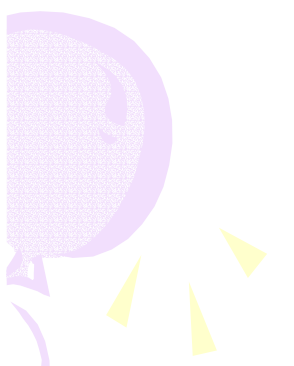
$$y^* | \beta, \sigma^2 \sim N(X\beta, \sigma^2 I)$$

$$y^* = (y_1^*, \dots, y_n^*), \quad X = (x_1, \dots, x_n)'$$

と表現でき、尤度関数は

$$f(y^* | \beta, \sigma^2) \propto \sigma^{-n} e^{-\frac{1}{2\sigma^2} (y^* - X\beta)' (y^* - X\beta)}$$

- 事前分布


$$\beta \sim N(b_0, B_0), \quad \sigma^2 \sim IG\left(\frac{n_0}{2}, \frac{n_0 S_0}{2}\right)$$



2. トービットモデル(続き)

- 事後分布

$$\beta | \sigma^2, y^* \sim N(b_1, B_1),$$

$$\sigma^2 | \beta, y^* \sim IG\left(\frac{n_1}{2}, \frac{n_1 S_1}{2}\right)$$

$$y_i^* | \beta, \sigma^2 \sim N(x_i' \beta, \sigma^2) I(y_i^* \leq 0) \text{ for } y_i = 0$$

ただし $n_1 = n_0 + n,$

$$n_1 S_1 = n_0 S_0 + (y^* - X\beta)'(y^* - X\beta),$$

$$b_1 = B_1(B_0^{-1}b_0 + \sigma^{-2}X'y^*),$$

$$B_1^{-1} = B_0^{-1} + \sigma^{-2}X'X$$



2. トービットモデル(続き)

例: 既婚女性の1年間の労働時間(y)を

1. 子供の人数(x_1),
2. 年齢(x_2),
3. 教育年数(x_3),
4. 夫の年収(x_4)

で説明するモデルを考える.

- 労働時間は、就職していない人は0時間になるので、トービットモデルを用いる。

2. トービットモデル(続き)

計算例(Ox言語)

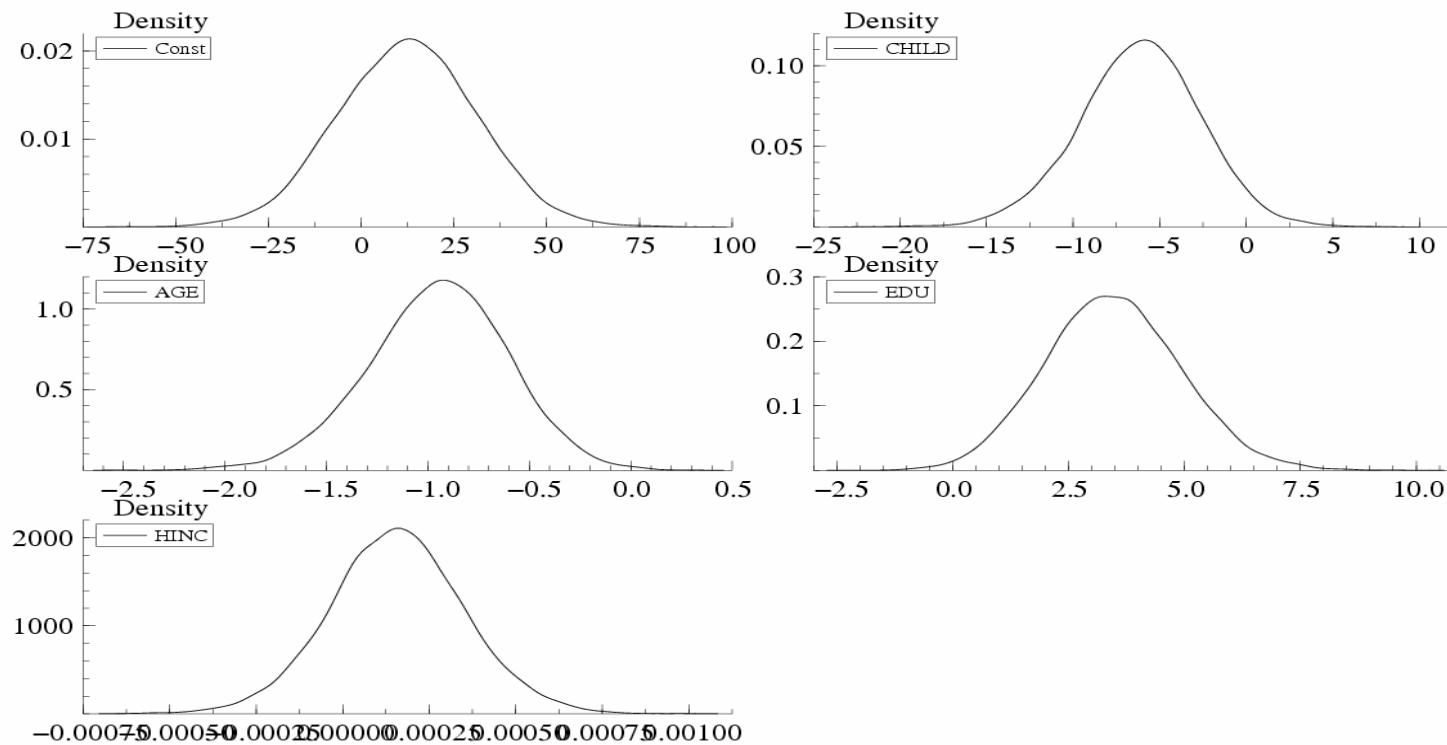
```
for(cn=-cburn;cn<crepeat;++cn){
  mB1=invert((1/dsigma2)*mx'*mx+invert(mB0));      # のサンプリング
  vb1=mB1*((1/dsigma2)*mx'*vystar+invert(mB0)*vb0);
  vb=vb1+choleski(mB1)*rann(cm,1);
  for(i=0;i<cnobs;++i){ #y*のサンプリング
    if(vy[i]==0){
      do{vystar[i]=mx[i][]*vb+rann(1,1)*sqrt(dsigma2);
        }while(vystar[i] >=0);
    }
    else{vystar[i]=vy[i];}}
  # 2のサンプリング
  da=0.5*(0.002+cnobs);db=0.5*(0.002+(vystar-mx*vb)'*(vystar-mx*vb));
  dsigma2=1/rangamma(1,1,da,db);
}
```

2. トービットモデル(続き)

計算結果

	事後平均	標準偏差	2.5%	97.5%
1	12.49	18.77	-24.03	49.01
2	-6.15	3.53	-13.38	0.56
3	-0.95	0.35	-1.66	-0.30
4	3.46	1.45	0.70	6.38
5	1.57E-4	1.95E-4	-2.31E-4	5.44E-4
2	679.6	215.5	369.3	1195.1

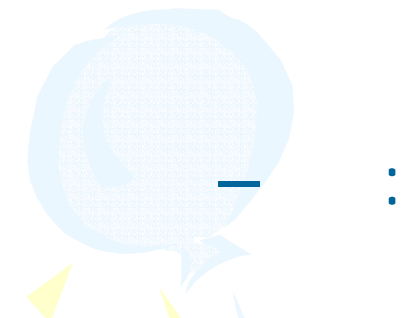
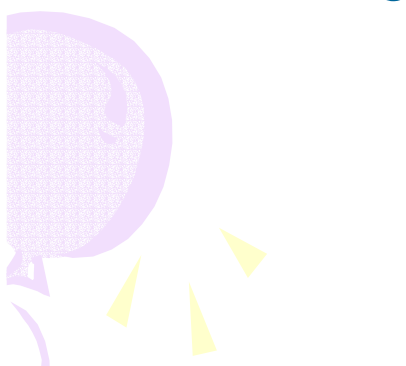
2. トービットモデル(続き)





3. プロビットモデル

プロビットモデルー選択行動のモデル

- 被説明変数 y_i が0または1といった離散的な値をとる.
 - 例: 新車を購入する($y_i = 1$), しない($y_i = 0$)である.
 - 例: ある顧客にローンを貸与する($y_i = 1$), しない($y_i = 0$)
- 
- 

3. プロビットモデル(続き)

- モデル

$$Pr(y_i = 1) = F(x_i' \beta), Pr(y_i = 0) = 1 - F(x_i' \beta)$$

- ここで $F(\cdot)$ は0から1の間をとる関数
- F に標準正規分布の分布関数 プロビットモデル
- F にロジスティック分布の分布関数 ロジットモデル

- 潜在変数 y_i^* を

$$y_i^* = x_i' \beta + \epsilon_i, \quad \epsilon_i \sim N(0, 1),$$

$$y_i = I(y_i^* > 0), \quad i = 1, \dots, n$$

と定義すると y^* が与えられたとき、標準的な回帰モデルで分散が1と既知のモデルとなる。



3. プロビットモデル(続き)

- プロビットモデルは

$$y^*|\beta \sim N(X\beta, I)$$

$$y^* = (y_1^*, \dots, y_n^*), \quad X = (x_1, \dots, x_n)'$$

と表現でき、尤度関数は

$$f(y^*|\beta) \propto e^{-\frac{1}{2}(y^* - X\beta)'(y^* - X\beta)}$$

- 事前分布

$$\beta \sim N(b_0, B_0)$$



3. プロビットモデル(続き)

- 事後分布

$$\beta|y^* \sim N(b_1, B_1),$$

$$y_i^*|\beta \sim N(x_i'\beta, 1)I(y_i^* \leq 0) \text{ for } y_i = 0$$

$$y_i^*|\beta \sim N(x_i'\beta, 1)I(y_i^* > 0) \text{ for } y_i = 1$$



ただし

$$n_1 = n_0 + n,$$

$$b_1 = B_1(B_0^{-1}b_0 + X'y^*),$$

$$B_1^{-1} = B_0^{-1} + X'X$$



3. プロビットモデル(続き)

例: 地方学校税に関する住民投票の行動分析のために, 有権者に対する調査. そのなかの95人のデータがPindyck and Rubinfeld (1998)にあり, このデータに対してプロビットモデルをあてはめる.

- Y: 学校税率引き上げに賛成のとき 1, 反対のとき 0 ¥¥
- X_1: 定数項
- X_2: 子供が私立学校に在籍しているとき 1, いないとき 0
- X_3: (私立または公立)学校の教師であるとき 1, ないとき 0
- X_4: 家計所得(年額)の対数値
- X_5: 財産税(年額)の対数値

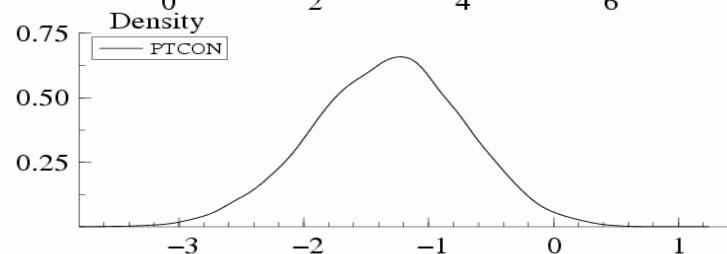
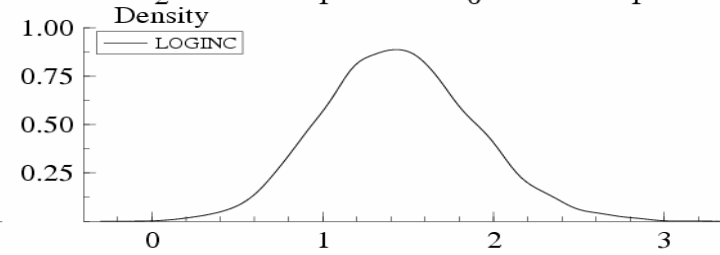
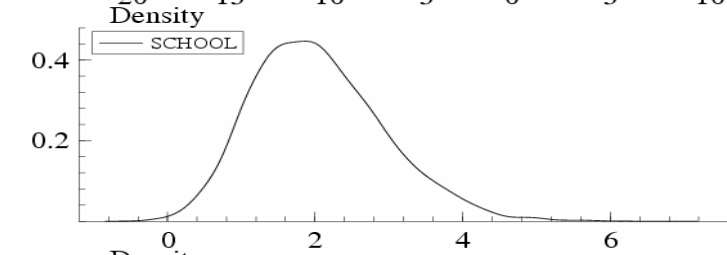
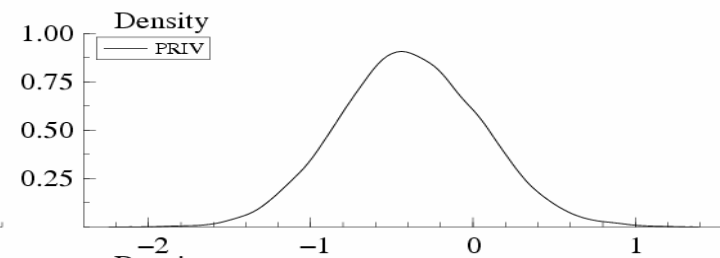
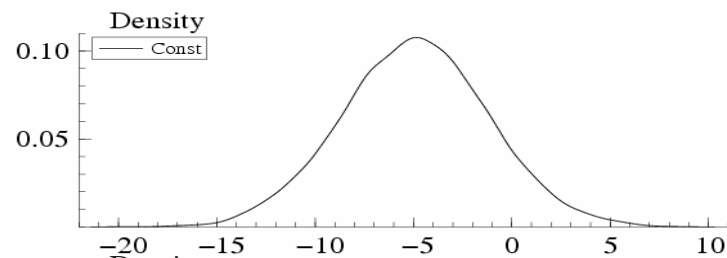


3. プロビットモデル(続き)

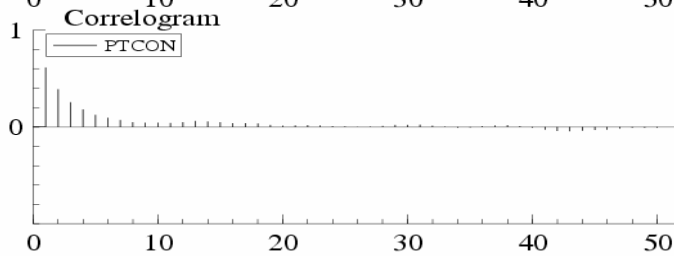
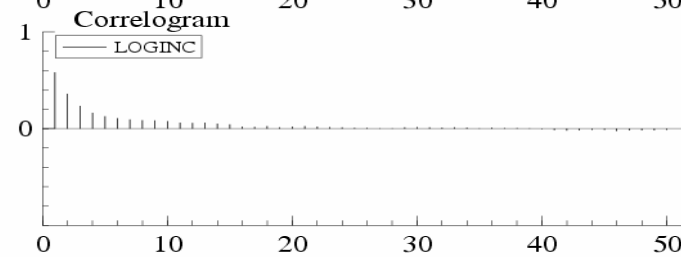
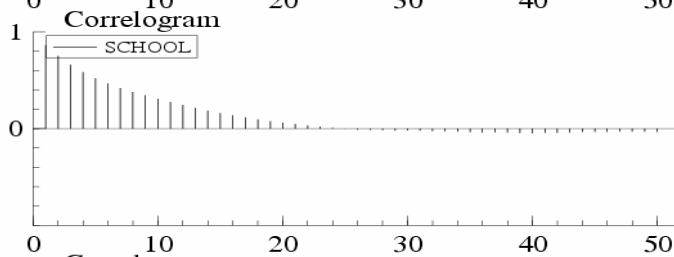
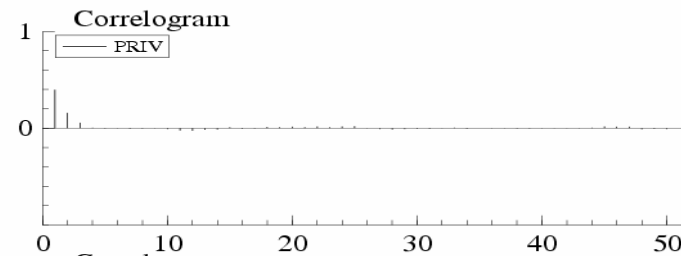
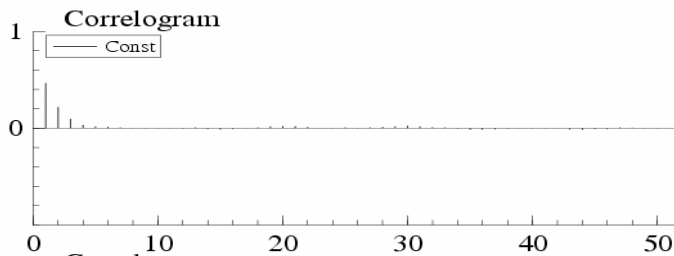
例: Ox言語の例

```
for(cn=-cburn;cn<crepeat;++cn){
# ~N(b1,B1)
vb1=mB1*(mx'*vystar+invert(mB0)*vb0);
vb=vb1+choleski(mB1)*rann(cm,1);
# y*のサンプリング
for(i=0;i<cnobs;++i){
  if(vy[i]==0){do{vystar[i]=mx[i][]*vb+rann(1,1);
    }while(vystar[i] >=0);}
  else{do{vystar[i]=mx[i][]*vb+rann(1,1);
    }while(vystar[i] <0);}
}
}
```

3. プロビットモデル(続き)



3. プロビットモデル(続き)



3. プロビットモデル(続き)

計算結果

	事後平均	標準偏差	2.5%	97.5%
1	-4.91	3.73	-12.3	2.43
2	-0.39	0.44	-1.23	0.48
3	2.05	0.89	0.55	4.00
4	1.44	0.45	0.62	2.36
5	-1.33	0.60	-2.55	-0.18



4. ロジットモデル

- ロジットモデルでは

$$Pr(y_i = 1) = \frac{\exp(x_i' \beta)}{1 + \exp(x_i' \beta)}, \quad Pr(y_i = 0) = \frac{1}{1 + \exp(x_i' \beta)},$$

- 事前分布

$$\beta \sim N(b_0, B_0)$$

- 尤度関数

$$f(y|\beta) = \prod_{i=1}^n \left(\frac{\exp(x_i' \beta)}{1 + \exp(x_i' \beta)} \right)^{y_i} \left(\frac{1}{1 + \exp(x_i' \beta)} \right)^{1-y_i}$$

4. ロジットモデル(続き)

- 事後分布

$$\pi(\beta|y) \propto \prod_{i=1}^n \{\exp(x'_i\beta)\}^{y_i} \{1 + \exp(x'_i\beta)\}^{-1} \\ \times \exp\left\{-\frac{1}{2}(\beta - b_0)' B_0^{-1}(\beta - b_0)\right\}$$

- MHアルゴリズム

- 提案分布は正規分布 $N(m, V)$. ただし,

$$m = \hat{\beta} + V \left. \frac{\partial g(\beta)}{\partial \beta} \right|_{\beta=\hat{\beta}}, \quad V = - \left\{ \left. \frac{\partial^2 g(\beta)}{\partial \beta \partial \beta'} \right\}^{-1} \right|_{\beta=\hat{\beta}}$$

$$g(\beta) = \sum_{i=1}^n y_i x'_i \beta - \log\{1 + \exp(x'_i \beta)\} - \frac{1}{2}(\beta - b_0)' B_0^{-1}(\beta - b_0)$$

である.

4. ロジットモデル(続き)

- 提案分布(正規分布, 密度関数は $(\beta | m, V)$ とおく)からの候補 β' を採択する確率は,

$$\alpha(\beta^{(i)}, \beta | y) = \min \left\{ \frac{\pi(\beta' | y) / \phi(\beta' | m, V)}{\pi(\beta^{(i)} | y) / \phi(\beta^{(i)} | m, V)}, 1 \right\}$$

となる.

- 比例定数の項は分子分母で共通するためキャンセルアウトする.



4. ロジットモデル(続き)

例: WinBUGSのコード

- モデルの指定

```
model {  
  for (i in 1:n) {  
    y[i] ~ dbern(p[i])  
    logit(p[i]) <- b[1] + b[2]*x[i,1]+ b[3]*x[i,2] + b[4]*x[i,3]  
    + b[5]*x[i,4]  
  }  
  #prior  
  for (j in 1:K) {b[j] ~ dnorm(0,0.001)}  
}
```

(WinBUGSにおいて採用されたサンプリングでは収束が遅かった)

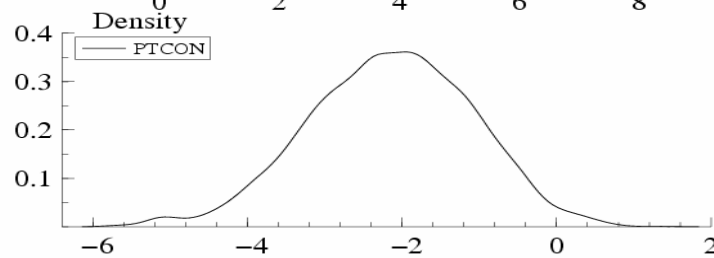
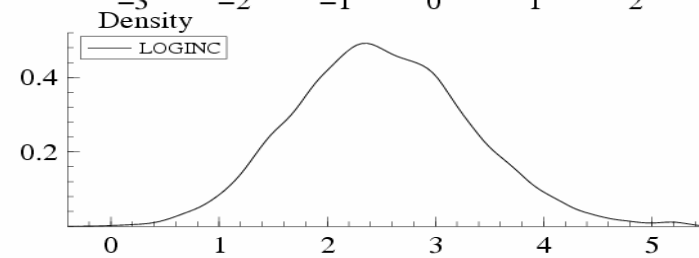
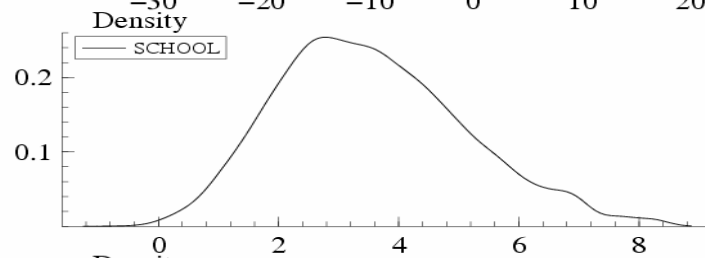
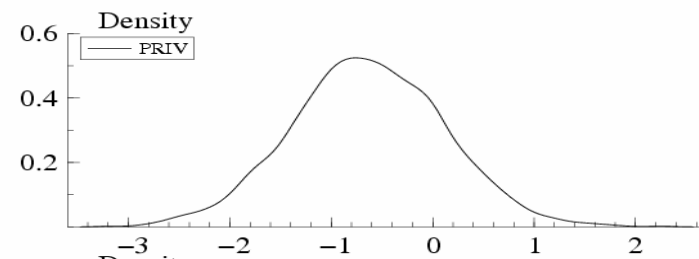
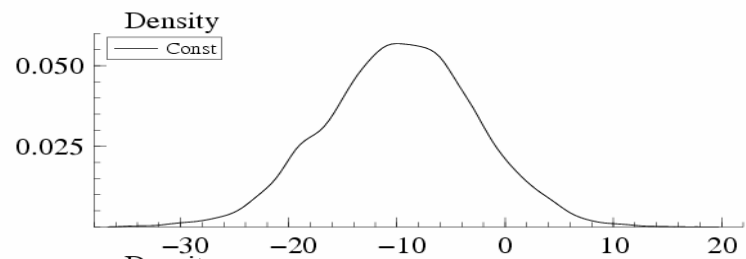
3. プロビットモデル(続き)

例: Ox言語の例

```
for(cn=-cburn;cn<crepeat; ++cn){
vb_o=vb;fLK(vb_o,&dl_o,0,0);      #現在の点での事後密度
dh_o=dlhat+v11hat'*(vb_o-vbhat)  #現在の点での提案密度
      +0.5*(vb_o-vbhat)'*ml2hat*(vb_o-vbhat);
vb_n=vb1+mB1root*rann(cm,1);    #候補点の提案
fLK(vb_n,&dl_n,0,0);            #候補点での事後密度
dh_n=dlhat+v11hat'*(vb_n-vbhat) #候補点での提案密度
      +0.5*(vb_n-vbhat)'*ml2hat*(vb_n-vbhat);
dfrac=exp(dl_n+dh_o-dl_o-dh_n);  #採択確率の計算
if(ranu(1,1)<=dfrac){vb=vb_n;}
}
```

MHの採択率は72.75%

4. ロジットモデル(続き)



4. ロジットモデル(続き)

計算結果

	事後平均	標準偏差	2.5%	97.5%
1	-9.66	6.91	-23.2	3.87
2	-0.64	0.76	-2.12	0.84
3	3.55	1.56	0.92	6.95
4	2.51	0.81	1.00	4.18
5	-2.15	1.06	-4.30	-0.20

5.見かけ上無関係な回帰モデル

Seemingly Unrelated Model

$$y_i = X_i\beta + \epsilon_i$$

- y_i : $m \times 1$ 確率変数ベクトルで被説明変数, $i=1, \dots, n$
 - 例: t 時点における m 個の危険資産の収益率など
- X_i : 既知の $m \times K$ で説明変数行列
- $\beta = (\beta_1, \dots, \beta_K)'$: 未知のパラメータ
- $\epsilon_i \sim N(0, \Sigma)$ は $m \times 1$ 確率変数ベクトルで誤差項.

5. SURモデル(続き)

- 事前分布

$$\beta \sim N(b_0, B_0), \quad \Sigma \sim IW(n_0, R_0)$$

$$\pi(\beta|b_0, B_0) \propto \exp\left\{-\frac{1}{2}(\beta - b_0)'B_0^{-1}(\beta - b_0)\right\}$$

$$\pi(\Sigma|n_0, R_0) \propto |\Sigma|^{-\frac{n_0+K+1}{2}} \exp\left\{-\frac{1}{2}\text{tr}(R_0^{-1}\Sigma^{-1})\right\}$$

- 尤度関数は

$$f(y|\beta, \Sigma) \propto |\Sigma|^{-n/2} e^{-\frac{1}{2}\sum_{i=1}^n (y_i - X_i\beta)' \Sigma^{-1} (y_i - X_i\beta)}$$



5. SURモデル(続き)

- 事後分布

$$\beta|\Sigma, y) \sim N(b_1, B_1), \quad \Sigma \sim IW(n_1, R_1)$$



- ただし

$$b_1 = B_1 \left(B_0^{-1} b_0 + \sum_{i=1}^n X_i' \Sigma^{-1} y_i \right),$$

$$B_1^{-1} = B_0^{-1} + \sum_{i=1}^n X_i' \Sigma^{-1} X_i,$$

$$n_1 = n_0 + n,$$

$$R_1^{-1} = R_0^{-1} + \sum_{i=1}^n (y_i - X_i \beta)(y_i - X_i \beta)'$$



5. SURモデル(続き)

例: General Electric と Westinghouse

- Gross investment (Y): 被説明変数
- Value of the Firm (V): 説明変数
- Stock of plant and equipment(K)
- $y_i = \beta_1 + \beta_2 V_i + \beta_3 K_i + \epsilon_i$
- Grunfeld Investment Data, 100 yearly observations. 1935-1954.



5. SURモデル(続き)

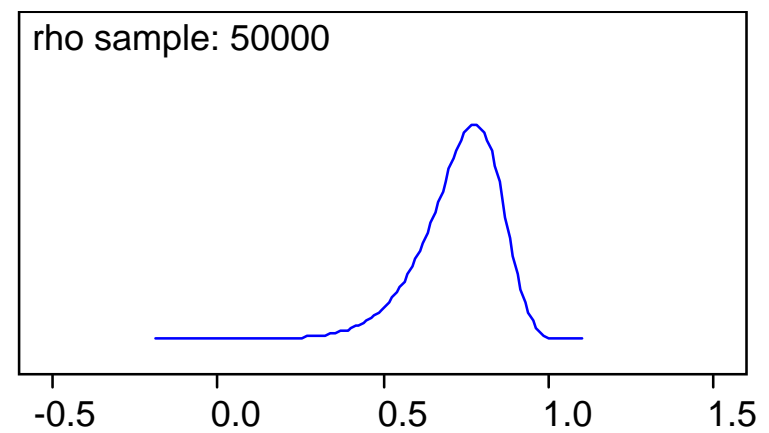
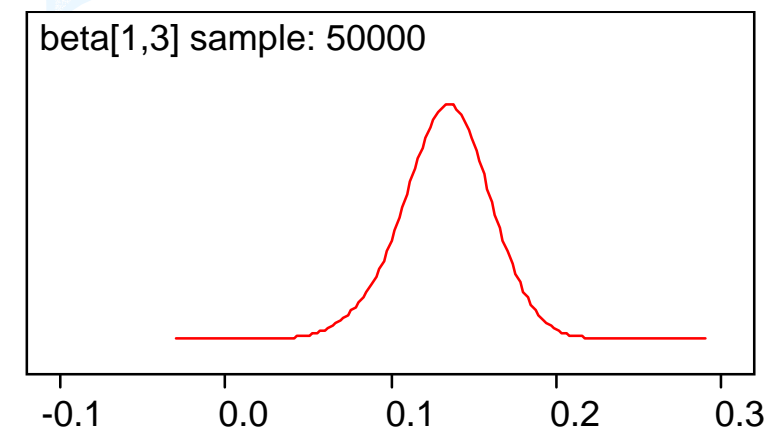
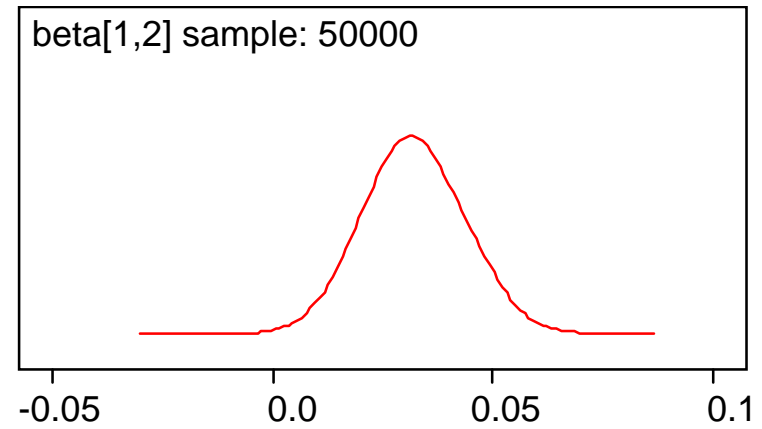
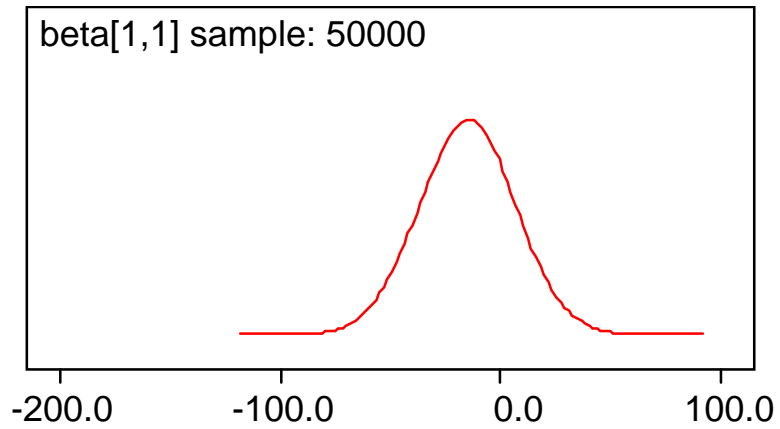


例: WinBUGSのコード

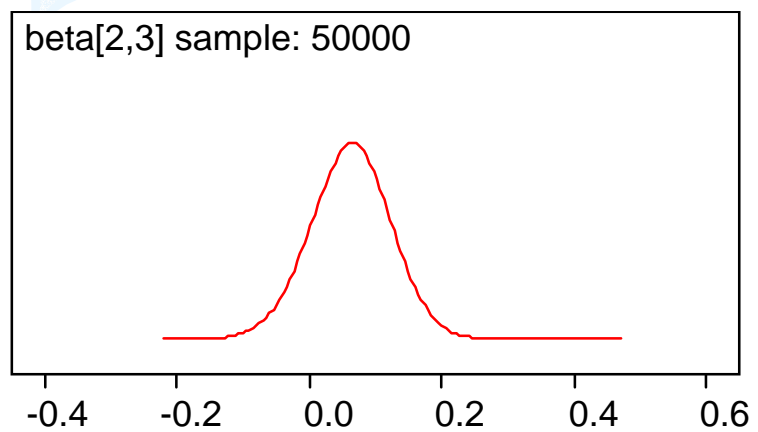
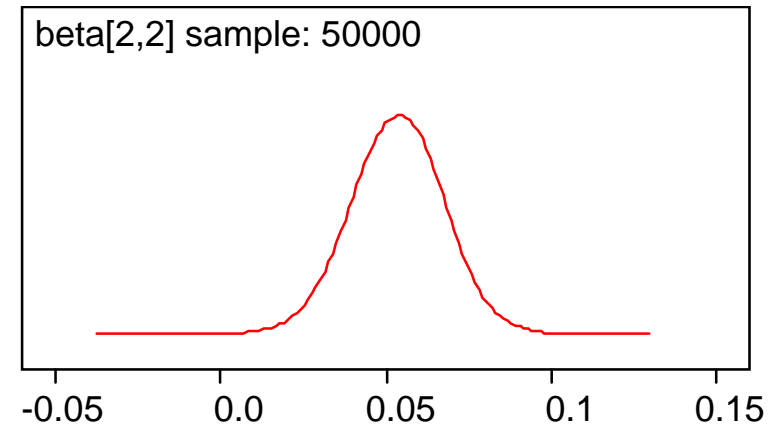
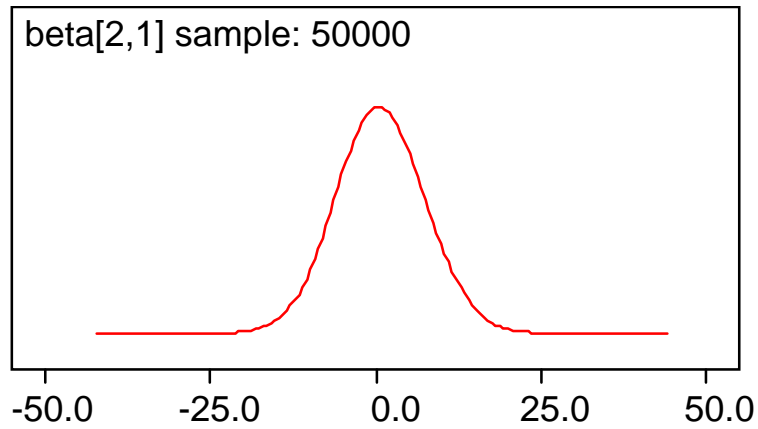
- モデルの指定

```
model {  
  for (t in 1:T) {  
    y[t,1:2] ~ dnorm(mu[t,],Omega[ , ]);  
    mu[t,1] <- beta[1,1]+beta[1,2]*v[t,1]+beta[1,3]*k[t,1]  
    mu[t,2] <- beta[2,1]+beta[2,2]*v[t,2]+beta[2,3]*k[t,2]  
  }  
  # priors on regression coefficients  
  for (i in 1:M) {for (j in 1:3) {beta[i,j] ~ dnorm(0,0.001)}}  
  Omega[1 : M , 1 : M] ~ dwish(R[ , ], 2)  
  # cross-correlation matrix of dimension M=2  
  Sigma[1:M,1:M] <- inverse(Omega[ , ])  
}
```


5. SURモデル(続き)



5. SURモデル(続き)



5.SURモデル(続き)

- 計算結果のまとめ

	事後平均	標準偏差	MC error	2.5%	中央値	97.5%
11	-14.21	20.78	0.65	-54.88	-14.05	26.65
12	0.03	0.01	3.7E-4	0.01	0.03	0.05
13	0.13	0.03	4.9E-4	0.08	0.14	0.18
21	1.13	6.67	0.21	-11.71	1.03	14.33
22	0.05	0.01	4.8E-4	0.03	0.05	0.08
23	0.06	0.06	0.001	-0.05	0.07	0.18
	0.74	0.12	0.001	0.46	0.76	0.90



参考文献

- Chen, Shao and Ibrahim(2000) Monte Carlo Methods in Bayesian Computation. Springer.
 - Gamerman, D. (1997) Markov Chain Monte Carlo. Chapman & Hall
 - Gelman, Carlin, Stern and Rubin(1995) Bayesian Data Analysis. Chapman & Hall
 - Ibrahim, Chen and Shinha (2001) Bayesian Survival Analysis. Springer.
 - Koop, G. (2003) Bayesian Econometrics. Wiley.
 - Liu, J.S. (2001) Monte Carlo Strategies in Scientific Computing. Springer.
 - 和合肇編「マルコフ連鎖モンテカルロによる計量分析」近刊
東洋経済
- 