確率的ボラティリティ変動モデル：分析法とモデルの発展

渡　部　敏　明

1. はじめに

近年，資産価格の時系列分析では，ボラティリティと呼ばれる2次のモーメントの変動に注目が集まっている．ボラティリティは投資リスクの指標であるとともに，オプション価格の決定要因でもある．そこで，もしそれが時間を通じて変動するのであれば，その変動をうまく捉えられるような時系列モデルを開発することは，単に研究者の間だけでなく，投資のリスク管理という観点から実務家にとっても重要である．ボラティリティの変動を明示的に定式化する時系列モデルとしては，これまでに，大きく分けて，2つのものが提案されている．1つは，Engle（1982）によって提案されたARCH（Autoregressive Conditional Heteroskedasticity）モデルとそれを発展させたモデル（本論文では，以下，そうしたモデルを総称して，ARCH型モデルと呼ぶ）であり[1)]，もう1つは，確率的ボラティリティ変動（Stochastic Volatility; 以下，略して，SV）モデルである．ARCH型モデルがパラメータの値を最尤法によって簡単に推定できるのに対して，SVモデルは尤度を解析的に評価するのが難しいため，パラメータの推定には最尤法に代る推定法が必要になる．

SVモデルの推定法にはこれまでさまざまな方法が提案されているが，その中で，特に注目を集めているものに，Jacquier，Polson and Rossi（1994）によって提案されたマルコフ連鎖モンテカルロ（Markov-chain Monte Carlo; 以下，略して，MCMC）法を用いたベイズ推定法がある[2)]．SVモデルのような尤度を解析的に求められないモデルの場合，パラメータの事後分布もベイズの定理を用いて解析的に求めることができない．そうしたモデルをベイズ推定する場合には，何らかの方法を用いてパラメータの値を事後分布からサンプリングし，サンプリングされた値を用いてパラメータを推定するという方法がとられる．解析的に求まらない未知の事後分布からのサンプリングを可能にしてくれるのが，MCMC法である[3)]．MCMC法とは1回前にサンプリングされた値に基づいて次の値をサンプリングする方法の総称であり，代表的なものにGibbs samplerとMetropolis-Hastings（MH）アルゴリズムがある．

SVモデルをベイズ推定する場合，パラメータだけでなく潜在変数であるボラティリティも同時事後分布からのサンプリングを行う．その際，ボラティリティは標本の大きさだけあるので，それをいかに効率的にサンプリングするかがポイントとなる．Jacquier，Polson and Rossi（1994）は，各期各期のボラティリティを別々にサンプリングするsingle-move samplerと呼ばれる方法を用いていたが，この方法を用いるとサンプリングされた値に高い自己相関が生じ，MCMC法の収束が遅い上，推定値の標準誤差が大きくなることがShephard and Pitt（1997）によって示されている．そこで，その後，Shephard and Pitt

（1997），Watanabe and Omori（2004a, b）らによって multi-move sampler, Kim, Shephard, and Chib（1998）によって mixture sampler と呼ばれるより効率的なボラティリティのサンプリング法が提案されている．本論文では，こうした MCMC 法を用いた SV モデルのベイズ推定法の最近の発展についてサーベイを行っている．

ベイズ統計学では，通常，モデル比較に周辺尤度（marginal likelihood）を用いるが，SV モデルの周辺尤度の計算方法やモデルの診断の方法も提案されているので，そうした方法についても簡単に解説している．また，SV モデルは推定が難しいため，ARCH 型モデルと比べると，これまでモデルの拡張はあまりなされてこなかった．しかし，近年，推定法の開発に伴い，SV モデルの拡張も行われるようになってきた．そこで，本論文では，そうした SV モデル自体の最近の発展についてもサーベイを行っている．

本論文の以下の構成は次の通りである．まず，次の第2節で，通常用いられる簡単な SV モデルについて解説する．続く第3節で，SV モデルの MCMC 法を用いたベイズ推定法を解説し，第4節で，SV モデルの最近の発展についてサーベイする．最後に第5節で，今後の課題について述べる．

2. SV モデル

通常用いられる簡単な SV モデルは，次の2式から構成される．

$$
\begin{aligned}
y_t &= \exp(h_t/2)\epsilon_t,\ \epsilon_t \sim i.i.d.N(0,1), && (1)\\
h_{t+1} &= \mu + \phi(h_t-\mu) + \eta_t, &&\\
&\qquad \eta_t \sim i.i.d.N(0,\sigma_\eta^2). && (2)
\end{aligned}
$$

ここで，y_t は t 期の価格変化率から平均と自己相関を除去したものである．（1）式は，y_t をボラティリティと呼ばれる非負の確率変数 exp（$h_t/2$）と過去と独立な標準正規分布に従う確率変数 ϵ_t の積として表している．（2）式は，ボラティリティの2乗の対数値 h_t が次数1の自己回帰モデル（AR（1）モデル）に従うものと仮定している[4]．さらに，誤差項 η_t は過去と独立な平均0，分散 σ_η^2 の正規分布に従い，ϵ_s ($s=1,\dots,T$) とも独立であると仮定する．株式市場では，株価が上がった日の翌日より下がった日の翌日の方がボラティリティがより上昇する傾向があることが知られており，そうしたボラティリティ変動の非対称性を捉えるためには，ϵ_t と η_t の間に相関を導入する必要があるが，そうした拡張については 4. 2 節を参照のこと．また，h_t の初期値 h_1 は h_t の無条件分布である平均0，分散 $\sigma_\eta^2/(1-\phi^2)$ の正規分布に従うものと仮定する．このモデルで推定すべき未知パラメータは（μ, ϕ, σ_η^2）であり，この内，重要なのは，h_t に対するショックの持続性を表すパラメータ ϕ である．本論文では，h_t は定常的であると考え，$|\phi|<1$ であると仮定する[5]．その場合，ϕ が1に近ければ近いほどボラティリティに対するショックの持続性が高いことになる．資産市場では，ボラティリティが上昇（低下）するとしばらくボラティリティの高い（低い）日が続くことが知られており，こうした現象をボラティリティ・クラスタリング（volatility clustering）と呼ぶ．このことから，ボラティリティに対するショックは持続性が高いことがわかる．実際，SV モデルを推定すると，ϕ の推定値には1に近い値が得られるのが常である[6]．

未知パラメータ（μ, ϕ, σ_η）をまとめて θ で表すと，SV モデルの尤度関数は次のように表される．

$$
\begin{aligned}
L(\boldsymbol{\theta}) &= f(y_1,\dots,y_T|\boldsymbol{\theta})\\
&= \int\cdots\int f(y_1,\dots,y_T|h_1,\dots,h_T)\\
&\qquad\qquad f(h_1,\dots,h_T|\boldsymbol{\theta})dh_1\dots dh_T,\\
&= \int\cdots\int\left[\prod_{t=1}^{T} f(y_t|h_t)\right]\left[f(h_1|\boldsymbol{\theta})\prod_{t=2}^{T}\right.
\end{aligned}
$$

$$f(h_t|h_{t-1},\boldsymbol{\theta})\Bigg]\,dh_1\ldots dh_T.$$

ここで,

$$\begin{aligned}
f(y_t|h_t) &= \frac{1}{\sqrt{2\pi\exp(h_t)}}\exp\left[-\frac{y_t^2}{2\exp(h_t)}\right],\\
f(h_t|h_{t-1},\boldsymbol{\theta}) &= \frac{1}{\sqrt{2\pi\sigma_\eta^2}}\\
&\qquad \exp\left[-\frac{\{h_t-\mu-\phi(h_{t-1}-\mu)\}^2}{2\sigma_\eta^2}\right],\\
f(h_1|\boldsymbol{\theta}) &= \frac{1}{\sqrt{2\pi\sigma_\eta^2/(1-\phi^2)}}\\
&\qquad \exp\left[-\frac{(h_1-\mu)^2}{2\sigma_\eta^2/(1-\phi^2)}\right].
\end{aligned}$$

この積分が解析的に解けないため，SV モデルのパラメータは最尤推定することが難しく，最尤法に代る推定法が必要になる．

3. マルコフ連鎖モンテカルロ法を用いたベイズ推定

3.1 パラメータのサンプリング

ベイズ推定法では，まず，未知パラメータ θ に適当な事前分布 $\pi(\theta)$ を設定する．従来のベイズ推定法は，事前分布をベイズの定理

$$f(\boldsymbol{\theta}|\boldsymbol{y})=\frac{f(\boldsymbol{y}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\int f(\boldsymbol{y}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta}} \qquad (3)$$

によってデータ $\boldsymbol{y}$ を観測した後の事後分布 $f(\boldsymbol{\theta}|\boldsymbol{y})$ に更新し，得られた事後分布に基づいてパラメータの値を推定するというものであった [7]．しかし，ベイズの定理（3）式の右辺にある $f(\boldsymbol{y}|\boldsymbol{\theta})$ は尤度であり，したがって，SV モデルのように尤度が解析的に求まらないモデルでは，事後分布をベイズの定理を使って解析的に求めることもできない．そうした場合には，何らかの方法によって事後分布から未知パラメータ θ の値をサンプリングし，得られた値に基づいてパラメータの値を推定するという方法がとられる．解析的に求まらない未知の事後分布からのサンプリングを可能にしてくれるのが MCMC 法である．MCMC 法は，通常のランダム・サンプリングと異なり，1 回前にサンプリングされた値に依存させて次の値をサンプリングする方法の総称であり，代表的なものに Gibbs sampler と Metropolis-Hastings（MH）アルゴリズムがある．

同時事後分布 $f(\boldsymbol{\theta}|\boldsymbol{y})$ をベイズの定理を使って解析的に求めることができないモデルでも，未知パラメータ $\boldsymbol{\theta}$ をいくつかのブロック（$\theta_1,\ldots,\theta_k$）に分けた場合の条件付事後分布 $f(\theta_i|\{\theta_j\}_{j\neq i},\boldsymbol{y})$（$i=1,\ldots,k$）はすべて求められ，かつ，そこからサンプリングを行えることは少なくない．（$\theta_1,\ldots,\theta_k$）はそれぞれ 1 変量であっても多変量であっても構わない．）こうした場合に用いられるのが，Gibbs sampler である．適当な初期値（$\theta_2^{(0)},\ldots,\theta_k^{(0)}$）からスタートして，まず，条件付事後分布 $f(\theta_1|\theta_2^{(0)},\theta_3^{(0)},\ldots,\theta_k^{(0)},\boldsymbol{y})$ から $\theta_1^{(1)}$ をサンプリングし，次に，条件付事後分布 $f(\theta_2|\theta_1^{(1)},\theta_3^{(0)},\ldots,\theta_k^{(0)},\boldsymbol{y})$ から $\theta_2^{(1)}$ をサンプリングする．これを繰り返し，最後に $f(\theta_k|\theta_1^{(1)},\theta_2^{(1)},\ldots,\theta_{k\text{-}1}^{(1)},\boldsymbol{y})$ から $\theta_k^{(1)}$ をサンプリングする．以上を第 1 ループと呼ぶことにする．この第 1 ループでサンプリングされた（$\theta_1^{(1)},\theta_2^{(1)},\ldots,\theta_k^{(1)}$）からスタートして，同様に第 2 ループを行い，（$\theta_1^{(2)},\theta_2^{(2)},\ldots,\theta_k^{(2)}$）をサンプリングする．以上を繰り返すと，第 l ループでは，（$\theta_1^{(l)},\theta_2^{(l)},\ldots,\theta_k^{(l)}$）がサンプリングされることになる．緩い制約条件の下で，$l\to\infty$ とすると，以上のようにしてサンプリングされた（$\theta_1^{(l)},\theta_2^{(l)},\ldots,\theta_k^{(l)}$）は同時事後分布 $f(\theta_1,\theta_2,\ldots,\theta_k|\boldsymbol{y})$ からサンプリングされた確率変数に分布収束することが知られている [8]．そこで，最初の M ループ（M は十分大きな値とする）でサンプリングされた値を捨て [9]，さらに N ループを行ってサンプリングされた（$\theta_1^{(l)}$,

$\theta_2^{(l)},\ldots,\ \theta_k^{(l)}$）（$l = M+1,\ M+2,\ldots,M+N$）は，同時事後分布$f(\theta_1, \theta_2,\ldots, \theta_k|\boldsymbol{y})$からサンプリングされた値と見なすことができる．

SVモデルの場合，$\boldsymbol{\theta} = (\mu,\ \phi,\ \sigma_\eta^2)$，$\boldsymbol{y} = (y_1,\ldots,\ y_T)$である．そこで，上のGibbs samplerを用いて同時事後分布$f(\mu,\phi,\sigma_\eta^2|y_1,\ldots,\ y_T)$からサンプリングするためには，例えば，$k=3$，$\theta_1=\mu$，$\theta_2=\phi$，$\theta_3=\sigma_\eta^2$）として，

$f(\mu\,|\,\phi,\ \sigma_\eta^2,\ y_1,\ldots,\ y_T)$,
$f(\phi\,|\,\mu,\ \sigma_\eta^2,\ y_1,\ldots,\ y_T)$,
$f(\sigma_\eta^2\,|\,\mu,\ \phi,\ y_1,\ldots,\ y_T)$,

から繰り返しサンプリングすればよいことになる．しかし，残念ながら，これらの条件付事後分布も解析的に求められない．ところが，潜在変数（$h_1,\ldots,\ h_T$）も条件に含めた

$$f(\mu\,|\,\phi,\ \sigma_\eta^2,\ h_1,\ldots,\ h_T,\ y_1,\ldots,\ y_T), \qquad (4)$$
$$f(\phi\,|\,\mu,\ \sigma_\eta^2,\ h_1,\ldots,\ h_T,\ y_1,\ldots,\ y_T), \qquad (5)$$
$$f(\sigma_\eta^2\,|\,\mu,\ \phi,\ h_1,\ldots,\ h_T,\ y_1,\ldots,\ y_T), \qquad (6)$$

は解析的に求められる．ここで，（$\mu\,|\,\phi,\ \sigma_\eta^2$）はSVモデルの（2）式だけに含まれるパラメータであることに注意しよう．（$y_1,\ldots,\ y_T$）は（1）式を通じて潜在変数（$h_1,\ldots,\ h_T$）の情報を与えてくれるので，潜在変数（$h_1,\ldots,\ h_T$）の値が未知の場合には，（$\mu,\ \phi,\ \sigma_\eta^2$）の分布を導出するのに（$y_1,\ldots,\ y_T$）と（1）式が必要となるが，（$h_1,\ldots,\ h_T$）の値が与えられると，（$y_1,\ldots,\ y_T$）も（1）式も必要なくなり，（$h_1,\ldots,\ h_T$）と（2）式だけを考えればよいことになる．その場合，（2）式は（$h_1,\ldots,\ h_T$）を観測値とする単なるAR（1）モデルになるので，（$h_1,\ldots,\ h_T$）を条件に加えた条件付事後分布（4）−（6）は解析的に求められるのである．また，（$h_1,\ldots,\ h_T$）が与えられると，（$y_1,\ldots,\ y_T$）は必要なくなるので，条件付事後分布（4）−（6）は，条件から（$y_1,\ldots,\ y_T$）を削除することができる．

事前分布として，通常，μには正規分布，σ_η^2には逆ガンマ分布が用いられる[10)]．

$\mu \sim N(\mu_0,\ \sigma_0^2)$，$\sigma_\eta^2 \sim IG(\nu_0/2,\ \delta_0/2)$，

ϕの事前分布には，ボラティリティが定常であるとの仮定の下，$|\phi|<1$の範囲で切断された切断正規分布，または（$1+\phi$）/2にベータ分布を仮定する．ベータ分布に従う確率変数は0から1までの値しかとれないので，（$1+\phi$）/2がベータ分布に従うと仮定すると，$0<(1+\phi)/2<1$より，$|\phi|<1$が満たされる．

こうした事前分布の下では，条件付事後分布（4），（6）はそれぞれ次のように計算される[11)]．

$$\mu\,|\,\phi,\ \sigma_\eta^2,\ h_1,\ldots,\ h_T, \sim N(\mu_1,\ \sigma_1^2) \qquad (7)$$
$$\sigma_\eta^2\,|\,\mu,\ \phi,\ h_1,\ldots,\ h_T, \sim IG(\nu_1/2,\ \delta_1/2,). \qquad (8)$$

ただし，

$$\begin{aligned}
\sigma_1^2 &= \frac{\sigma_0^2\sigma_\eta^2}{\sigma_0^2\{(T-1)(1-\phi)^2+1-\phi^2\}+\sigma_\eta^2}\\
\mu_1 &= \sigma_1^2\left\{\frac{(1-\phi^2)}{\sigma_\eta^2}h_1+\frac{(1-\phi)}{\sigma_\eta^2}\sum_{t=1}^{T-1}(h_{t+1}-\phi h_t)+\frac{\mu_0}{\sigma_0^2}\right\}\\
\nu_1 &= T+\nu_0\\
\delta_1 &= \delta_0+(h_1-\mu)^2(1-\phi^2)+\sum_{t=1}^{T-1}\{h_{t+1}-\mu-\phi(h_t-\mu)\}^2.
\end{aligned}$$

これらの条件付事後分布からは簡単にサンプリングできる[12)]．

ϕの条件付事後分布（5）は次のように計算される[13)]．

$$\begin{aligned}
&\ln\left(f(\phi\,|\,\mu,\ \sigma_\eta^2,\ h_1,\ldots,\ h_T)\right)\\
&= \text{定数}+\ln(\varphi(\phi))-\frac{1}{2\sigma_\eta^2}\sum_{t=1}^{T-1}\cdot
\end{aligned}$$

$$\{h_{t+1}-\mu-\phi(h_t-\mu)\}^2, \quad -1<\phi<1. \tag{9}$$

ここで，π（ϕ）をϕの事前分布とすると，

$$\ln(\varphi(\phi)) = \ln\pi(\phi) - \frac{(h_1-\mu)^2(1-\phi^2)}{2\sigma_\eta^2} + \frac{1}{2}\ln(1-\phi^2) \tag{10}$$

である．このような特殊な分布からサンプリングする方法に，Acceptance-Rejection（AR）アルゴリズムと MH アルゴリズムがある．

いま，確率密度関数f（x）からサンプリングしたいが，特殊な分布であるため直接サンプリングできないものとしよう．ただし，f（x）は，少なくとも（9）式の定数項のような基準化定数（normalizing constant）以外の部分は解析的に求められ，既知であるものとする．

AR アルゴリズムでは，まず，f（x）を近似していて，かつ，そこから直接サンプリングでき，取り得るすべてのxで優越条件（dominance condition）f（x）$\leq cg$（x）を満たしているような提案密度関数（proposal density function）g（x）と正の定数cを選択する．それを使って以下の AR アルゴリズムを実行すると，f（x）から 1 つの値がサンプリングされる．

AR アルゴリズム：

[1] 提案密度関数g（x）からサンプリングを行い，得られた値$x^{(\mathrm{proposal})}$を使って受容確率pを次のように計算する．

$$p=\frac{f(x^{(\mathrm{proposal})})}{cg(x^{(\mathrm{proposal})})}$$

[2]　[1] で得られた値$x^{(\mathrm{proposal})}$を確率pで受容し，確率$1-p$で棄却する．受容された，$x=x^{(\mathrm{proposal})}$とし，終了．棄却された場合には [1] へ戻る．

このサンプリングは 1 回前にサンプリングされた値に依存していないので，MCMC 法ではなく，ランダム・サンプリングである．また，この場合，受容確率$p=f$（$x^{(\mathrm{proposal})}$）$/cg$（$x^{(\mathrm{proposal})}$）が必ず 1 以下であることを保証するため，優越条件f（x）$\leq cg$（x）が必要になる．このアルゴリズムを用いる際に注意すべきことは，提案密度関数g（x）と定数cをf（x）モードの近辺で受容確率pが 1 に近くなるように選ぶことである．そうしないと，[2] で何度も棄却が続いて，サンプリングに時間がかかってしまう．優越条件を満たしつつ，モードの近辺で受容確率pが 1 に近くなるようにg（x）やcを選ぶのは一般的には難しい．例えば，モードの近辺で受容確率pが 1 に近くなるようにcを下げると，分布の裾の部分で優越条件が満たされなくなる可能性がある．これに対して，以下の MH アルゴリズムでは，優越条件は必要ない．そこで，優越条件を満たしつつ，モードの近辺で受容確率pが 1 に近くなるようにg（x）やcを選ぶのが難しい場合には MH アルゴリズムが用いられる．

MH アルゴリズムでも，まず，f（x）を近似していて，かつ，直接サンプリングできるような提案密度関数h（x）を選択する．ただし，既に述べたように，その際，AR アルゴリズムのような優越条件は必要ない．その上で，適当な初期値x_0からスタートして，以下の MH アルゴリズムを実行すれば，f（x）からN個の値（$x_1,\ldots,x_N$）をサンプリングできる．

MH アルゴリズム：

[1] $n=1$とする．

[2] 提案密度関数h（x）からサンプリングし，得られた値$x^{(\mathrm{proposal})}$を使って受容確率qを次のように計算する．

$$q = \min\left[\frac{f(x^{(\text{proposal})})h(x_{n-1})}{f(x_{n-1})h(x^{(\text{proposal})})}, 1\right]$$

[3] $x^{(\text{proposal})}$ を確率 q で受容し，確 $1-q$ で棄却する[14]．受容された場合には，$x_n = x^{(\text{proposal})}$ とする．棄却された場合には，$x_n = x_{n-1}$ とする．

[4] $n < N$ であれば，$n = n+1$ として，[2] に戻る．$n = N$ であれば，終了する．

これは1回前にサンプリングされた値 x_{n-1} を使ってサンプリングを行っているので，MCMC法である．MHアルゴリズムを用いる場合にも，ARアルゴリズム同様，提案密度関数 h（x）をいかに選ぶかがポイントになる[15]．このアルゴリズムでは，[3] で棄却されると1回前にサンプリングされた値をそのまま選ぶので，受容確率 q が低いと，続けて何度も同じ値をサンプリングしてしまう．そこで，やはり，モード近辺で q が1に近くなるように提案密度関数 h（x）を選ぶのが望ましい．

条件付事後分布（9）からサンプリングするためには，x $=\phi$，ln（f（x））を（9）式として MH アルゴリズムを実行すればよい．その際の提案密度関数 h（ϕ）の選択の方法はいくつか考えられるが，ここでは Chib and Greenberg（1994）で提案されている方法を紹介する．彼らは，（9）式で右辺の ln（φ（ϕ））を無視したものを定数＋ ln（h（x）），すなわち，

$$\begin{aligned}
&\ln f(\phi|\mu, \sigma_\eta^2, h_1, \dots, h_T) \\
&\approx \ \text{定数} - \frac{1}{2\sigma_\eta^2}\sum_{t=1}^{T-1}\{h_{t+1} - \mu - \phi(h_t - \mu)\}^2, \\
&= \ \text{定数} - \frac{\sum_{t=1}^{T-1}(h_t-\mu)^2}{2\sigma_\eta^2} \\
&\qquad \left\{\phi - \sum_{t=1}^{T-1}\frac{(h_{t+1}-\mu)(h_t-\mu)}{\sum_{t=1}^{T-1}(h_t-\mu)^2}\right\}^2, \\
&= \ \text{定数} + \ln(h(\phi)), \quad -1 < \phi < 1, \qquad (11)
\end{aligned}$$

としている[16]．そうすると，h（ϕ）は平均 $\mu_\phi = \sum_{t=1}^{T-1}(h_{t+1}-\mu)(h_t-\mu)/\sum_{t=1}^{T-1}(h_t-\mu)^2$，分散 $v_\phi = \sigma_\eta^2/\left\{\sum_{t=1}^{T}(h_t-\mu)^2\right\}$ の正規分布を（−1，1）の範囲を残して切断した切断正規分布になる．この切断正規分布からは簡単にサンプリングできる．まず，正規分布 N（μ_ϕ，v_ϕ）からサンプリングし，$|\phi| < 1$ あれば受容し，それ以外であれば棄却すればよい[17]．h（ϕ）をこのように選択すると，受容確率 q は，$q = \varphi$（$x^{(\text{proposal})}$ /φ（x_{n-1}）と簡単になる．ただし，φ（x）は（10）式で定義される．

3.2　ボラティリティのサンプリング

SV モデルの未知パラメータ（μ，ϕ，σ_η^2）の条件付事後分布は，条件に潜在変数（$h_1, \dots$，h_T）を加えると求まり，かつ，そこからサンプリングできることがわかった．そこで，Gibbs sampler を適用するには，潜在変数（$h_1, \dots$，h_T）も，（μ，ϕ，σ_η^2）同様，未知パラメータとして扱えばよい．ただし，そうすると，（$h_1, \dots$，h_T）も条件付事後分布からサンプリングしなければならない．（$h_1, \dots$，h_T）は標本の大きさ T だけあるので，このサンプリングを効率的に行わないと，膨大な時間がかかって事実上推定不可能になってしまう．そこで，MCMC 法を用いた SV モデルのベイズ推定では，（$h_1, \dots$，h_T）をいかに効率良くサンプリングするかがポイントとなる．これまでに提案されている（$h_1, \dots$，h_T）のサンプリング法には single-move sampler, multi-move sampler, mixture sampler の3つがある．

3.2.1　Single-move sampler

SV モデルの MCMC 法を用いたベイズ推定法を提案した Jacquier，Polson and Rossi（1994）は，各期の潜在変数を，

$f\ (h_t|\ \mu\ ,\ \phi\ ,\ \sigma^2_\eta,\ \ h_1,\ldots,\ h_{t-1},\ \ h_{t+1},\ldots,\ h_T,$
$y_1,\ldots,\ y_T)\ (t=1,\ldots,\ T)$ （12）

から1個ずつサンプリングしており，この方法は single-move sampler と呼ばれる．この条件付分布から効率的にサンプリングする方法はいくつか提案されているが[18]，この方法には大きな問題点がある．前節で述べたように，ボラティリティに対するショックは持続性が高いので，$(h_1,\ldots,\ h_T)$ は互いに強い相関がある．このように相関の高い変数を別々にサンプリングすると Gibbs sampler の収束の速度が遅いことが知られている．実際，Shephard and Pitt（1997）は，ϕ が1に近い場合にこの方法を用いると収束の速度が遅いことをシミュレーションによって示している．

3.2.2 Mixture sampler

そうした場合に，Gibbs sampler の収束の速度を高める1つの方法は，相関の高い変数をひとまとめにしてサンプリングを行うことである．つまり，SV モデルの場合，$(h_1,\ldots,\ h_T)$ を

$f\ (h_1,\ldots,\ h_T|\ \mu,\ \phi,\ \sigma^2_\eta,\ \ y_1,\ldots,\ y_T)$ （13）

から一度にサンプリングすればよい．これを行っているのが，Kim，Shephard and Chib（1998）の mixture sampler である．

この方法では，SV モデルの（1）式を，両辺を2乗して対数をとることにより，次のように書き換える．

$\ln\ (y^2_t)\ =h_t+z_t,\ \ z_t=\ln\ (\epsilon^2_t).$ （14）

そうすると，（14）式と（2）式は，h_t を状態変数，（14）式を観測方程式，（2）式を遷移方程式とする線形状態空間（linear state space）モデルになる[19]．線形状態空間モデルの中で，観測方程式の誤差項と遷移方程式の誤差項がいずれも正規分布に従うものを，線形ガウシアン状態空間（linear Gaussian state space）モデルと呼び，その場合には（13）からのサンプリングは簡単に行える．ところが，この場合の観測方程式（14）の誤差項 $z_t=\ln\ (\epsilon^2_t)$ の分布は正規分布ではない．mixture sampler では，それを次のような混合正規分布で近似する．

$$f(z_t)=\sum_{i=1}^{M}q_if_N(z_t|m_i,v_i). \qquad (15)$$

ただし，q_i は $\sum_{i=1}^{M}\ q_i=1$ を満たす正の定数，$f_N\ (z_t|m_i,\ v_i)$ は平均 m_i，分散 v_i の正規分布の確率密度関数を表す[20]．その上で，$1,\ldots,\ M$ のどれかの値をとり，$S_t\ =i$ となる確率 $Pr\ (S_t=i)$ が $q_i\ (i=1,\ldots,\ M)$ であるような確率変数 S_t を導入し，$(S_1,\ldots,\ S_T)$ も，パラメータや $(h_1,\ldots,\ h_T)$ 同様，条件付事後確率

$p\ (S_1,\ldots,\ S_T|\ \mu,\ \phi,\ \sigma^2_\eta,\ \ h_1,\ldots,\ h_T,\ \ y_1,\ldots,\ y_T)$

からサンプリングする．そうすると，$(h_1,\ldots,\ h_T)$ は，そこでサンプリングされた $(S_1,\ldots,\ S_T)$ の値を（13）の条件に加えた分布

$f\ (h_1,\ldots,\ h_T|\ \mu,\ \phi,\ \sigma^2_\eta,\ \ S_1,\ldots,\ S_T,\ \ y_1,\ldots,$
$y_T)$ （16）

からサンプリングすればよいことになる．$S_t=i$ が与えられると，z_t の分布は平均 m_i，分散 v_i の正規分布になる．そこで，$(S_1,\ldots,\ S_T)$ が与えられると，（14）（2）式は線形ガウシアン状態空間モデルになるので，Durbin and Koopman（2002）で提案されている simulation smoother を使えば，（16）からサンプリングすることができる．Simulation smoother とは，線形ガウシアン状態空間モデルの観測方程式または遷移方程式の誤差項をパラメータの値が与えられた下でサンプリングするアルゴリズムである．これを使っ

て h_t をサンプリングするためには，simulation smoother によって z_t をサンプリングした後，$h_t = \ln(y_t^2) - z_t$ とすればよい．

3.2.3 Multi-move sampler

（15）式のような変換を行わず，(1)，(2) 式を使って，条件付事後分布（13）から潜在変数 $(h_1, \ldots, h_T)$ を一度にサンプリングするのは，特に T が大きい場合には難しい．そこで，$(h_1, \ldots, h_T)$ を一度にサンプリングするのではなく，いくつかのブロックに分けて，1 つのブロックを一度にサンプリングするという方法を考えよう．例えば，$(h_t, \ldots, h_{t+k})$ が 1 つのブロックだとすると，それらを，

$$f(h_t, \ldots, h_{t+k} | \mu, \phi, \sigma_\eta^2, h_{t-1}, h_{t+k+1}, y_t, \ldots, y_{t+k}) \tag{17}$$

から一度にサンプリングするということである[21]．

しかし，$(h_t, \ldots, h_{t+k})$ には互いに相関があるので，(17) からのサンプリングも容易ではない．そこで，Shephard and Pitt (1997) は，潜在変数 $(h_t, \ldots, h_{t+k})$ ではなく，(2) 式の誤差項 $(\eta_{t-1}, \ldots, \eta_{t+k-1})$ を，

$$f(\eta_{t-1}, \ldots, \eta_{t+k-1} | \mu, \phi, \sigma_\eta^2, h_{t-1}, h_{t+k+1}, y_t, \ldots, y_{t+k}) \tag{18}$$

からサンプリングするという方法を提案している．$(\mu, \phi, \sigma_\eta^2)$ と h_{t-1} の値が与えられた下で，$(\eta_{t-1}, \ldots, \eta_{t+k-1})$ がサンプリングされると，(2) 式から，$(h_t, \ldots, h_{t+k})$ を逐次的に計算できる．この方法は，multi-move sampler もしくは block sampler と呼ばれる．

条件付分布（18）からサンプリングする方法として，Shephard and Pitt (1997) は，Tierney (1994) の提案した Acceptance-Rejection Metropolis-Hastings (ARMH) アルゴリズムを用いている[22]．このアルゴリズムは，既に説明した MH アルゴリズムの候補 $x^{(\text{proposal})}$ を AR アルゴリズムを用いてサンプリングするという方法である．そうすることで，MH アルゴリズムの提案密度関数が正しい密度関数 $f(x)$ をよりうまく近似するようになるので，MH アルゴリズムの受容確率が高まる．このアルゴリズムでも，MH アルゴリズム同様，AR アルゴリズムのような優越条件は必要ない．まず，AR アルゴリズムの提案密度関数 $g(x)$ と正の定数 c を選択し，適当な初期値 x_0 からスタートして，以下のアルゴリズムを実行すれば，$f(x)$ から N 個の値 $(x_1, \ldots, x_N)$ をサンプリングできる．

<u>ARMH アルゴリズム：</u>

[1] $n = 1$ とする．

[2] 提案密度関数 $g(x)$ からサンプリングを行い，得られた値 x を使って受容確率 p を次のように計算する．

$$p = \min\left[\frac{f(x)}{cg(x)}, 1\right]$$

[3] [2] で得られた値 x を確率 p で受容し，確率 $1 - p$ で棄却する．受容された場合には，$x^{(\text{proposal})} = x$ とおき，[4] に進む．棄却された場合には [2] へ戻る．

[4] 受容確率 q を以下のように計算する．

(a) $f(x_{n-1}) < cg(x_{n-1})$ ならば，$q = 1$，

(b) $f(x_{n-1}) \geq cg(x_{n-1})$ かつ $f(x^{(\text{proposal})}) < cg(x^{(\text{proposal})})$ ならば，

$$q = \frac{cg(x_{n-1})}{f(x_{n-1})},$$

(c) $f(x_{n-1}) \geq cg(x_{n-1})$ かつ $f(x^{(\text{proposal})}) \geq cg(x^{(\text{proposal})})$ ならば，

$$q = \min\left[\frac{f(x^{(\text{proposal})})g(x_{n-1})}{f(x_{n-1})g(x^{(\text{proposal})})}, 1\right].$$

[5] $x^{\text{(proposal)}}$ を確率 q で受容し，確率 $1-q$ で棄却する．受容された場合には，$x_n = x^{\text{(proposal)}}$ とする．棄却された場合には，$x_n = x_{n-1}$ とする．

[6] $n < N$ であれば，$n = n + 1$ として [2] に戻る．$n = N$ であれば，終了．

この内，[2]，[3] が AR アルゴリズム，[4]，[5] が MH アルゴリズムである．このアルゴリズムでは，優越条件 $f(x) \leq cg(x)$ は必要はないが，もし，取り得るすべての x でそれが満たされるならば，MH パートの受容確率は常に $q = 1$ となり，AR アルゴリズムでサンプリングされた値を常に受容するので，ARMH アルゴリズムは通常の AR アルゴリズムになる．

条件付分布（18）から（$\eta_{t-1}, \ldots, \eta_{t+k-1}$）をサンプリングするためには，$x = (\eta_{t-1}, \ldots, \eta_{t+k-1})$，（18）を $f(x)$ として，ARMH アルゴリズムを実行すればよい．その際，提案密度関数 $g(x)$ と定数 c を選択しなければならないが，効率的にサンプリングを行うためには，$g(x)$ および定数 c を AR パートおよび MH パートの受容確率 p，q ができるだけ 1 に近くなるように選ぶのが望ましい．そのため，Shephard and Pitt（1997）は，$g(x)$ を条件付事後分布（18）をうまく近似するような $k+1$ 変量正規分布として選んでいる．以下，彼らの $g(x)$ の選択の仕方およびそこからのサンプリングの方法について説明しよう．

まず，$t + k < T$ の場合を考えよう．この場合，条件付事後分布（18）の対数をとったものは次のように表せる[23]．

$$
\begin{aligned}
&\ln\left(f\left(\eta_{t-1}, \ldots, \eta_{t+k-1} \mid \mu, \phi, \sigma_\eta^2, h_{t-1}, h_{t+k+1}, y_t, \ldots, y_{t+k}\right)\right) \\
&= \text{定数} + \sum_{s=t}^{t+k} \ln(f(y_s|h_s)) \\
&\quad + \ln(f(h_{t+k+1}|\mu, \phi, \sigma_\eta^2, h_{t+k})) \\
&\quad + \sum_{s=t-1}^{t+k-1} \ln(f(\eta_s|\sigma_\eta^2)), \\
&= \text{定数} - \sum_{s=t}^{t+k} \left\{ \frac{h_s}{2} + \frac{y_s^2}{2}\exp(-h_s) \right\} \\
&\quad - \frac{1}{2\sigma_\eta^2}\{h_{t+k+1} - \mu - \phi(h_{t+k} - \mu)\}^2 \\
&\quad - \frac{1}{2\sigma_\eta^2} \sum_{s=t-1}^{t+k-1} \eta_s^2.
\end{aligned} \tag{19}
$$

Shephard and Pitt（1997）は，

$$
l(h_s) = -\frac{h_s}{2} - \frac{y_s^2}{2}\exp(-h_s) \tag{20}
$$

を $\hat{h}_s$ の回りで 2 次までテーラー展開することにより（19）式を以下のように近似したものを $\ln(cg(\eta_{t-1}, \ldots, \eta_{t+k-1}))$ としている．

$$
\begin{aligned}
&\ln\left(f\left(\eta_{t-1}, \ldots, \eta_{t+k-1} \mid \mu, \phi, \sigma_\eta^2, h_{t-1}, h_{t+k+1}, y_t, \ldots, y_{t+k}\right)\right) \\
&\approx \text{定数} + \sum_{s=t}^{t+k} \Big\{ l(\hat{h}_s) + (h_s - \hat{h}_s) l'(\hat{h}_s) \\
&\qquad + \frac{1}{2}(h_s - \hat{h}_s)^2 l''(\hat{h}_s) \Big\} \\
&\quad - \frac{1}{2\sigma_\eta^2}\{h_{t+k+1} - \mu - \phi(h_{t+k} - \mu)\}^2 \\
&\quad - \frac{1}{2\sigma_\eta^2} \sum_{s=t-1}^{t+k-1} \eta_s^2, \\
&= \text{定数} - \sum_{s=t}^{t+k-1} \frac{1}{2v_s}(\hat{y}_s - h_s)^2 \\
&\quad - \frac{1}{2\sigma_\eta^2} \sum_{s=t-1}^{t+k-1} \eta_s^2, \\
&= \ln(cg(\eta_{t-1}, \ldots, \eta_{t+k-1})).
\end{aligned} \tag{21}
$$

ここで，v_s，$\hat{y}_s$ は以下のように定義される．

$s = t, \ldots, t+k-1$ であれば，

$$
v_s = -\frac{1}{l''(\hat{h}_s)}, \tag{22}
$$

$$\hat{y}_s \quad = \quad \hat{h}_s + v_s l'(\hat{h}_s), \tag{23}$$

$s = t + k < T$ であれば，

$$v_s \quad = \quad \frac{\sigma_\eta^2}{\phi^2 - \sigma_\eta^2 l''(\hat{h}_s)}, \tag{24}$$

$$\hat{y}_s \quad = \quad \hat{h}_s + v_s \left[l'(\hat{h}_s) + \frac{\phi}{\sigma_\eta^2} \left\{ h_{t+k+1} - \mu - \phi(\hat{h}_s - \mu) \right\} \right]. \tag{25}$$

ただし，

$$l'(\hat{h}_s) \quad = \quad \frac{1}{2}\left\{ y_s^2 \exp(-\hat{h}_t) - 1 \right\} \tag{26}$$

$$l''(\hat{h}_s) \quad = \quad -\frac{y_s^2}{2}\exp(-\hat{h}_s) \tag{27}$$

である．

次に，$t + K = T$ の場合，すなわち，最後のブロックを考えよう．その場合，条件に h_{t+k+1} は含まれないので，（19）式を次のように書き換えなければならない．

$$\ln\left(f\left(\eta_{t-1}, \ldots, \eta_{t+k-1} \mid \mu, \phi, \sigma_\eta^2, h_{t-1}, y_t, \ldots, y_{t+\mathrm{k}}\right)\right) = \text{定数} - \sum_{s=t}^{t+k}\left\{\frac{h_s}{2} + \frac{y_s^2}{2}\exp(-h_s)\right\} - \frac{1}{2\sigma_\eta^2}\sum_{s=t-1}^{t+k-1}\eta_s^2. \tag{28}$$

そこで，この場合には，$s = t, \ldots, t + k$ すべてで，v_s，$\hat{y}_s$ は（22），（23）式で定義される．

以上をまとめると，v_s，$\hat{y}_s$ は，$s = t, \ldots, t + k - 1$ あるいは $s = t + k = T$ であれば（22），（23）式，$s = t + k < \mathrm{T}$ であれば（24），（25）式で定義される．ここで，このようにして定義される $\hat{y}_s$ を観測値，h_s を状態変数とする線形ガウシアン状態空間モデル

$$\hat{y}_s = h_s + \xi_s,\ \xi_s \sim i.i.d.N(0, v_s), \tag{29}$$

$$h_s = \mu + \phi\,(h_{s-1} - \mu) + \eta_s, \quad \eta_s \sim i.i.d.N(0, \sigma_\eta^2) \tag{30}$$

を考えよう[24]．提案密度関数 $g\,(\eta_{t-1}, \ldots, \eta_{t+k-1})$ はこの線形ガウシアン状態空間モデルにおける条件付事後分布

$$f\,(\eta_{t-1}, \ldots, \eta_{t+k-1} \mid \mu, \phi, \sigma_\eta^2, h_{t-1}, h_{t+\mathrm{k}+1}, y_t, \ldots, y_{t+\mathrm{k}}) \tag{31}$$

である[25]．また，そこからサンプリングするためには，Durbin and Koopman（2002）によっての提案されている simulation smoother を使えばよい．

テーラー展開を行う点（$\hat{h}_t, \ldots, \hat{h}_{t+k}$）は，（19）あるいは（28）式のモードとするのが望ましい．そうすれば，モードの周辺で，

$$f\,(\eta_{t-1}, \ldots, \eta_{t+k-1} \mid \mu, \phi, \sigma_\eta^2, h_{t-1}, h_{t+\mathrm{k}+1}, y_t, \ldots, y_{t+\mathrm{k}}) \approx cg\,(\eta_{t-1}, \ldots, \eta_{t+k-1})$$

となり，ARMH アルゴリズムの受容確率 p，q がいずれもモードの周辺で 1 に近くなる．そのためには，（$\hat{h}_t, \ldots, \hat{h}_{t+k}$）を次のように選べばよい．（$\hat{h}_t, \ldots, \hat{h}_{t+k}$）に適当な初期値を選んでやると，（23），（25）式より（$\hat{y}_t, \ldots, \hat{y}_{t+k}$）が計算できる．そこで，それを使って，（29），（30）式から成る線形ガウシアン状態空間モデルに対してカルマン・フィルターとスムーザーを実行すると，$E\,(h_{t+i} \mid \mu, \phi, \sigma_\eta^2, h_{t-1}, h_{t+k+1}, y_t, \ldots, y_{t+k})$（$i = 0, \ldots, k$）が求まる．それを（23），（25）式の（$\hat{h}_t, \ldots, \hat{h}_{t+k}$）に代入すると新たな（$\hat{y}_t, \ldots, \hat{y}_{t+k}$）が求まる．今度は，それを使い，（29），（30）式から成る線形状態空間モデルにおいて再びカルマン・フィルターとスムーザーを実行すると，新たな $E\,(h_{t+i} \mid \mu, \phi, \sigma_\eta^2, h_{t-1}, h_{t+k+1}, y_t, \ldots, y_{t+k})$（$i = 0, \ldots, k$）が求まる．これを数回繰り返すと，モードに近い（$\hat{h}_t, \ldots, \hat{h}_{t+k}$）が得

られるので，そこでテイラー展開を行えばよい．

このアルゴリズムでは，$(\eta_1, \ldots, \eta_T)$ をいくつかのブロックに分割する必要がある．$k_0 = 0$, $k_{K+1} = T$ として，$K+1$ 個のブロック $(\eta_{k_{i-1}+1}, \ldots, \eta_{k_i})$ $(i = 1, \ldots, K+1)$ に分割するものとしよう．Shephard and Pitt（1997）は，$(k_1, \ldots, k_K)$ をランダムに選んでいる．具体的には，U_i を [0，1] の一様分布からサンプリングし，

$$k_i = \mathrm{int}[\mathrm{T} \times \{ (i + U_i) / (K + 2) \}], \ i = 1, \ldots, K$$

としている．ここで，int[x] は，x に最も近い整数値を表している．このように，ブロックをランダムに選ぶと，n 回目のサンプリングで高い確率で棄却されるブロックがあったとしても，$n+1$ 回目のサンプリングでは，異なるブロックが選ばれるので，棄却が続いてサンプリングが行き詰まってしまうということを排除できる．

以上が multi-move sampler と呼ばれるアルゴリズムであるが，それを提案した Shephard and Pitt（1997）は，（19）式の

$$\ln(f(h_{t+k+1}|\mu, \phi, \sigma_\eta^2, h_{t+k})) = -\frac{1}{2\sigma_\eta^2}\{h_{t+k+1} - \mu - \phi(h_{t+k} - \mu)\}^2$$

の項を無視している．その結果，v_s，$\hat{y}_s$ を，$s = t + K < T$ の場合にも，(22)，(23) 式で定義している．この点を修正したのが，Watanabe and Omori（2004a, b）であり，Watanabe and Omori（2004a）では，さらに，Shephard and Pitt（1997）の方法をそのまま用いると，パラメータや潜在変数の推定値にバイアスが生じることが示されている．

3.2.4 Multi-move sampler と Mixture sampler の比較

Asai（2003）は，multi-move sampler と mixture sampler とを比較し，multi-move samp ler の方が収束が早いとの結果を報告している．

また，mixture sampler にはいくつか問題点がある．まず，z_t の正しい分布ではなく，混合正規分布で近似したものを使っているので，あくまでも近似的な方法にすぎないという点である．次に，$y_t = 0$ の場合，$\ln(y_t^2) = -\infty$ となり，計算ができなくなってしまうという点である．こうした問題を避けるため，Kim，Shephard and Chib（1998）は，$c = 0.001$ とし，（14）式の左辺を $\ln((y_t + c)^2)$ に置き換えているが，c の選択が推定結果に影響を与えないかどうか調べてみるべきであろう．

最後に，この方法を適用できるのは，線形状態空間モデルで表現できるモデルだけという点である．例えば，リスク・プレミアムを考慮して，（1）式を，

$$y_t = \mathrm{a} + \mathrm{b}\exp(h_t) + \exp(h_t/2)\,\epsilon_t$$

に置き換えたモデルは，線形状態空間モデルに書き換えることができない．また，SV モデルを取引高を含める形で発展させたモデルに動学的 2 変量分布混合モデル（次節参照）があるが，それも線形状態空間モデルでは表せない．mixture sampler は，そうしたモデルには適用できない．モデルの拡張性という観点からは，線形状態空間モデルへの変換を必要としない multi-move sampler の方が望ましい．

3.3 モデル比較

ベイズ統計学では，通常，モデル比較は事後オッズ比（posterior odds ratio）を用いて行われる．$\boldsymbol{y}_T = (y_1, \ldots, y_T)$ とすると，モデル M_i とモデル M_j の事後オッズ比は次のように定義される．

$$\mathrm{POR} = \frac{f(M_i|\boldsymbol{y}_T)}{f(M_j|\boldsymbol{y}_T)}. \tag{32}$$

この事後オッズ比が1を上回れば，モデルM_iの方が選択される．事後オッズ比は，以下のように表すことができる．

$$\mathrm{POR} = \frac{f(\boldsymbol{y}_T|M_i)}{f(\boldsymbol{y}_T|M_j)}\frac{f(M_i)}{f(M_j)}. \tag{33}$$

ここで，右辺の第1項$f(\boldsymbol{y}_T|M_i)/f(\boldsymbol{y}_T|M_j)$はベイズファクター（Bayes factor），第2項$f(M_i)/f(M_j)$は事前オッズ比（prior odds ratio）と呼ばれる．事前オッズ比は，通常，1に設定され，そうすると，事後オッズ比はベイズファクターと一致する．そこで，事後オッズ比の分子，分母の値を計算すれば，事後オッズ比を計算できる．ベイズファクターの分子（分母）は，モデルM_i（M_j）の周辺尤度（marginal likelihood）と呼ばれる．以下，Chib（1995）によって提案された周辺尤度の計算方法を説明する．

周辺尤度は，ベイズの定理

$$f(\boldsymbol{\theta}_i|M_i,\boldsymbol{y}_T) = \frac{f(\boldsymbol{y}_T|M_i,\boldsymbol{\theta}_i)f(\boldsymbol{\theta}_i|M_i)}{f(\boldsymbol{y}_T|M_i)}$$

の右辺の分母に相当する．ここで，θ_iはモデルM_iのパラメータを表し，左辺の$f(\theta_i|M_i,\boldsymbol{y}_T)$は事後密度，右辺の$f(\boldsymbol{y}_T|M_i,\theta_i)$は尤度，$f(\theta_i|M_i)$は事前密度である．そこで，周辺尤度は以下のように表すことができる．

$$f(\boldsymbol{y}_T|M_i) = \frac{f(\boldsymbol{y}_T|M_i,\boldsymbol{\theta}_i)f(\boldsymbol{\theta}_i|M_i)}{f(\boldsymbol{\theta}_i|M_i,\boldsymbol{y}_T)}. \tag{34}$$

したがって，周辺尤度を計算するためには，尤度$f(\boldsymbol{y}_T|M_i,\theta_i)$，事前密度$f(\theta_i|M_i)$，事後密度$f(\theta_i|M_i,\boldsymbol{y}_T)$の値をそれぞれ計算すればよい．（34）式は，$\theta_i$がいかなる値であっても成り立つが，Chib（1995）は，θ_iをその事後平均にすることを提案している．事前密度は簡単に計算できるので，以下，SVモデルの事後密度と尤度の計算方法について説明する．以下では，添字iや条件の中のM_iは省略する．

3.3.1　事後密度の評価

以下，パラメータの事後平均を$\hat{\theta} = (\hat{\mu},\hat{\phi},\hat{\sigma}_\eta^2)$で表す．$\theta = \hat{\theta}$における事後密度の値は次のように表すことができる．

$$f(\hat{\boldsymbol{\theta}}|\boldsymbol{y}_T) = f(\hat{\mu}|\boldsymbol{y}_T)f(\hat{\sigma}_\eta^2|\hat{\mu},\boldsymbol{y}_T) f(\hat{\phi}|\hat{\mu},\hat{\sigma}_\eta^2,\boldsymbol{y}_T) \tag{35}$$

右辺第1項は次のように表すことができる．

$$f(\hat{\mu}|\boldsymbol{y}_T) = \int f(\hat{\mu}|\boldsymbol{h}_T)f(\boldsymbol{h}_T|\boldsymbol{y}_T)d\boldsymbol{h}_T. \tag{36}$$

上で説明したGibbs samplerを行うと，$f(\boldsymbol{h}_T|\boldsymbol{y}_T)$から$\boldsymbol{h}_T$がサンプリングされる．それを（$\boldsymbol{h}_T^{(1)},\ldots,\boldsymbol{h}_T^{(M)}$）とすると，（36）式は以下のように計算できる．

$$f(\hat{\mu}|\boldsymbol{y}_T) \approx \frac{1}{M}\sum_{m=1}^{M} f(\hat{\mu}|\boldsymbol{h}_T^{(m)}). \tag{37}$$

（35）式の右辺第2項は次のように表される．

$$f(\hat{\sigma}_\eta{}^2|\hat{\mu},\boldsymbol{y}_T) = \int f(\hat{\sigma}_\eta^2|\hat{\mu},\boldsymbol{h}_T)f(\boldsymbol{h}_T|\hat{\mu},\boldsymbol{y}_T)d\boldsymbol{h}_T. \tag{38}$$

これを計算するためには，$\boldsymbol{h}_T$を$f(\boldsymbol{h}_T|\hat{\mu},\boldsymbol{y}_T)$からサンプリングしなければならないので，新たにGibbs samplerを行う必要がある．具体的には，μを$\hat{\mu}$に固定して，以下の条件付分布から繰り返しサンプリングを行う．

$f(\sigma_\eta^2|\phi.\ \hat{\mu},\boldsymbol{h}_T)$,

$f(\phi \mid \sigma^2_\eta, \hat{\mu}, \boldsymbol{h}_T)$,
$f(\boldsymbol{h}_T \mid \phi, \sigma^2_\eta, \hat{\mu}, \boldsymbol{y}_T)$.

ここでは，便宜上，$\boldsymbol{h}_T$ は $f(\boldsymbol{h}_T \mid \phi, \sigma^2_\eta, \hat{\mu}, \boldsymbol{y}_T)$ からサンプリングするものとして書いているが，multi-move sampler を用いる場合には，各ブロックごとにサンプリングする．そこで得られた（$\boldsymbol{h}_T^{(1)}, \ldots, \boldsymbol{h}_T^{(K)}$）を使って，（38）式は以下のように計算できる．

$$f(\hat{\sigma}_\eta{}^2 \mid \hat{\mu}, \boldsymbol{y}_T) \approx \frac{1}{K}\sum_{k=1}^{K} f(\hat{\sigma}_\eta{}^2 \mid \hat{\mu}, \boldsymbol{h}_T^{(k)}). \qquad (39)$$

最後に，第 3 項は，

$$f(\hat{\phi} \mid \hat{\mu}, \hat{\sigma}^2_\eta, \boldsymbol{y}_T) = \int f(\hat{\phi} \mid \hat{\mu}, \hat{\sigma}^2_\eta, \boldsymbol{h}_T) f(\boldsymbol{h}_T \mid \hat{\mu}, \hat{\sigma}^2_\eta, \boldsymbol{y}_T)\, d\boldsymbol{h}_T. \qquad (40)$$

と表されるが，この場合，$f(\hat{\phi} \mid \hat{\mu}, \hat{\sigma}^2_\eta, \boldsymbol{h}_T)$ の基準化定数がわからないので，上記のような方法で計算することができない．こうした場合には，Chib and Jeliazkov（2001）で提案されている以下のような方法を使う必要がある．

まず，

$$f(\phi \mid \hat{\mu}, \hat{\sigma}^2_\eta, \hat{\mu}, \boldsymbol{y}_T),\ f(\boldsymbol{h}_T \mid \phi, \hat{\mu}, \hat{\sigma}^2_\eta, \boldsymbol{y}_T),$$

から繰り返しサンプリングを行い，得られた値を $\phi^{(g)}$，$\boldsymbol{h}_T^{(g)}$（$g = 1, \ldots, G$）とする．次に，（11）式で与えられる MH アルゴリズムの提案密度関数 $h(\phi)$ を使って，

$$h(\phi \mid \hat{\mu}, \hat{\sigma}^2_\eta, \boldsymbol{h}_T),\ f(\boldsymbol{h}_T \mid \phi, \hat{\mu}, \hat{\sigma}^2_\eta, \boldsymbol{y}_T),$$

から繰り返しサンプリングを行い，得られた値を $\phi^{(j)}$，$\boldsymbol{h}_T^{(j)}$（$j = 1, \ldots, J$）とする．これらの値を使うと，（40）式は以下のように計算できる．

$$f(\hat{\phi} \mid \hat{\mu}, \hat{\sigma}^2_\eta, \boldsymbol{y}_T) = \frac{G^{-1}\sum_{g=1}^{G} q(\phi^{(g)}, \hat{\phi} \mid \hat{\mu}, \hat{\sigma}^2_\eta, \boldsymbol{h}_T^{(g)}) h(\hat{\phi} \mid \hat{\mu}, \hat{\sigma}^2_\eta, \boldsymbol{h}_T^{(g)})}{J^{-1}\sum_{j=1}^{J} q(\hat{\phi}, \phi^{(j)} \mid \hat{\mu}, \hat{\sigma}^2_\eta, \boldsymbol{h}_T^{(j)})}. \qquad (41)$$

ここで，$q(x, y)$ は，受容確率

$$q(x,y) = \min\left[\frac{f(y)h(x)}{f(x)h(y)}, 1\right]$$

である [26]．

3.3.2 尤度の評価

SV モデルの尤度は解析的には評価できないので，シミュレーションによって評価する方法が開発されている．そうした方法には，Danielsson（1994），Danielsson and Richard（1993）によって提案された AGIS（Accelerated Gaussian Importance Sampler）や Kim，Shephard and Chib（1998）によって提案された particle filter を使った方法がある．ここでは，後者の方法について簡単に説明する．

$\boldsymbol{y}_t =（y_1, \ldots, y_t）$とすると，尤度は，

$$f(\boldsymbol{y}_T \mid \boldsymbol{\theta}) = f(y_1 \mid \boldsymbol{\theta}) \prod_{t=1}^{T-1} f(y_{t+1} \mid \boldsymbol{y}_t, \boldsymbol{\theta}).$$

と表すことができ，さらに右辺の $f(y_{t+1} \mid \boldsymbol{y}_t, \boldsymbol{\theta})$ は以下のように表すことができる．

$$f(y_{t+1} \mid \boldsymbol{y}_t, \boldsymbol{\theta}) = \iint f(y_{t+1} \mid \boldsymbol{y}_t, h_{t+1}, \theta) f(h_{t+1} \mid \boldsymbol{y}_t, h_t, \theta) f(h_t \mid \boldsymbol{y}_t, \theta)\, dh_{t+1} dh_t.$$

これは，$f(h_t \mid \boldsymbol{y}_t, \theta)$ から h_t をサンプリングできれば，サンプリングされた値 $h_t^{(m)}$（$m = 1, \ldots, M$）を使って，次に $f(h_{t+1} \mid \boldsymbol{y}_t, h_t^{(m)}, \theta)$（これは，平均 $\mu + \phi(h_t^{(m)} - \mu)$，分散 σ^2_η の正規分布）から $h_{t+1}^{(m)}$（$m = 1, \ldots, M$）をサンプリングし，それを使って，

$$f(y_{t+1}|\boldsymbol{y}_t,\boldsymbol{\theta}) \approx \frac{1}{M}\sum_{m=1}^{M} f(y_{t+1}|h_{t+1}^{(m)})$$

として計算できる．

Particle filter とは，$t=0$ からスタートして逐次的に f（$h_t|\boldsymbol{y}_t, \theta$）から h_t をサンプリングするアルゴリズムである[27]．いま，f（$h_{t\text{-}1}|\boldsymbol{y}_{t\text{-}1}, \boldsymbol{\theta}$）から $h_{t\text{-}1}^{(m)}$（$m=1,\ldots,M$）がサンプリングされたとする．f（$h_t|\boldsymbol{y}_t, \boldsymbol{\theta}$）は，$h_{t\text{-}1}^{(m)}$（$m=1,\ldots,M$）を使って次のように表すことができる．

$$\begin{aligned} f(h_t|\boldsymbol{y}_t,\boldsymbol{\theta}) &\propto f(y_t|h_t,\boldsymbol{\theta})f(h_t|\boldsymbol{y}_{t-1},\boldsymbol{\theta}), \\ &= f(y_t|h_t,\boldsymbol{\theta})\int f(h_t|h_{t-1},\boldsymbol{\theta}) \\ &\quad f(h_{t-1}|\boldsymbol{y}_{t-1},\boldsymbol{\theta})dh_{t-1}, \\ &\approx f(y_t|h_t,\boldsymbol{\theta})\frac{1}{M}\sum_{m=1}^{M} f(h_t|h_{t-1}^{(m)},\boldsymbol{\theta}). \qquad (42) \end{aligned}$$

ここで，

$$\ln f(y_t|h_t,\boldsymbol{\theta}) = \text{定数} - \frac{1}{2}h_t - \frac{y_t^2}{2}\exp(-h_t)$$

であり，これは h_t の凸関数なので，右辺の定数項を除いた部分を $\ln f^*$（h_t）とし，$h_t=\hat{h}_t$ で1次のテーラー展開をすると，次の不等式が得られる[28]．

$$\begin{aligned} \ln f^*(h_t) &= -\frac{1}{2}h_t - \frac{y_t^2}{2}\exp(-h_t) \\ &\leq \frac{1}{2}h_t - \frac{y_t^2}{2}\exp(-\hat{h}_t)(1+\hat{h}_t-h_t) \\ &= \ln g(h_t). \qquad (43) \end{aligned}$$

この g（h_t）と（42）式の右辺の f（$h_t|h_{t\text{-}1}^{(m)}, \boldsymbol{\theta}$）の積は次のように表すことができる．

$$g(h_t)f(h_t|h_{t-1}^{(m)},\boldsymbol{\theta}) \propto \pi_m f_N(h_t|h_{t|t-1}^{(m)},\sigma_\eta^2).$$

ただし，f_N（$h_t|h_{t|t\text{-}1}^{(m)}$, σ_η^2）は平均 $h_{t|t\text{-}1}^{(m)}$，分散 σ_η^2 の正規分布の確率密度関数を表し，さらに，

$$\begin{aligned} h_{t|t-1}^{(m)} &= \mu+\phi(h_{t-1}^{(m)}-\mu) \\ &+\frac{\sigma_\eta^2}{2}\{y_t^2\exp(-\hat{h}_t)-1\}, \\ \pi_m &= \\ &\exp\left(-\frac{1}{2\sigma_\eta^2}[\{\mu+\phi(h_{t-1}^{(m)}-\mu)\}^2-h_{t|t-1}^{(m)2}]\right). \end{aligned}$$

そこで，（42）式はさらに次のように書き換えられる．

$$\begin{aligned} &f(h_t|\boldsymbol{y}_t,\boldsymbol{\theta}) \\ &\quad\propto f^*(h_t)\frac{1}{M}\sum_{m=1}^{M} f(h_t|\boldsymbol{h}_{t-1}^{(m)},\boldsymbol{\theta}), \\ &\quad\leq g(h_t)\frac{1}{M}\sum_{m=1}^{M} f(h_t|\boldsymbol{h}_{t-1}^{(m)},\boldsymbol{\theta}), \\ &\quad\propto \frac{1}{M}\sum_{m=1}^{M}\pi_m f_N(h_t|h_{t|t-1}^{(m)},\sigma_\eta^2). \end{aligned}$$

したがって，提案密度関数を混合正規分布

$$\sum_{m=1}^{M}\pi_m^* f_N(h_t|h_{t|t-1}^{(m)},\sigma_\eta^2),$$

受容確率を $p=f^*$（h_t）$/g$（h_t）として[29]，3.1節で説明した AR アルゴリズムを行えば，f（$h_t|\boldsymbol{y}_t, \boldsymbol{\theta}$）から h_t をサンプリングできる．ただし，$\pi_m^*=\pi_m/\sum_{m=1}^{M}\pi_m$．

3.4　モデルの診断

Particle filter により f（$h_t|\boldsymbol{y}_t, \boldsymbol{\theta}$）から $h_t^{(m)}$（$m=1,\ldots,M$）をサンプリングし，それを使って f（$h_{t+1}|\boldsymbol{y}_t, h_t^{(m)}, \boldsymbol{\theta}$）からサンプリングした $h_{t+1}^{(m)}$（$m=1,\ldots,M$）は，尤度の評価だけでなく，モデルの診断にも用いることができる．

y_{t+1} の実現値を y_{t+1}^{o2} とすると，$\boldsymbol{y}_t$ と $\boldsymbol{\theta}$ を条件としたときに $y_{t+1}^2 \leq y_{t+1}^{o2}$ となる確率は，

$$Pr(y_{t+1}^2 \le y_{t+1}^{o2}|\boldsymbol{y}_t, \boldsymbol{\theta}) =$$

$$\int\int Pr(y_{t+1}^2 \le y_{t+1}^{o2}|h_t)$$

$$f(h_{t+1}|h_t, \boldsymbol{\theta}) f(h_t|\boldsymbol{y}_t, \boldsymbol{\theta}) dh_{t+1} dh_t$$

と表されるので，$h_{t+1}^{(m)}$（$m = 1, \ldots, M$）を使って以下のように推定できる．

$$Pr(y_{t+1}^2 \le y_{t+1}^{o2}|\boldsymbol{y}_t, \boldsymbol{\theta}) \approx \frac{1}{M} \sum_{m=1}^{M} Pr(y_{t+1}^2 \le y_{i,t+1}^{o2}|h_{t+1}^{(m)}, \boldsymbol{\theta}).$$

この右辺を u_{t+1}^M とすると，モデルの定式化が正しければ，u_{t+1}^M は漸近的に互いに独立な [0, 1] の一様分布に従うことが知られている（Rosenblatt 1952）．そこで，u_{t+1}^M を標準正規分布の累積分布関数の逆関数 F^{-1}（•）によって変換した $u_{t+1}^M = F^{-1}$（u_{t+1}^M）は漸近的に互いに独立な標準正規分布に従う．そこで，n_{t+1}^M を計算し，それが互いに独立な標準正規分布に従っているかどうか検定を行うことによって，モデルの診断ができる．

4. SV モデルの発展

MCMC 法を用いたベイズ推定法はモデルを拡張しても適用可能なので，最近では SV モデルも，ARCH 型モデル同様，さまざまな形で拡張が行われるようになってきている．本節では，そうした SV モデルの発展についてサーベイを行う．

4.1 誤差項の分布

資産価格変化率は正規分布よりも裾の厚い分布に従っている（尖度が 3 を越える）ことが古くから知られている（Mandelbrot 1963，Fama 1965）．y_t が SV モデルに従っている場合，たとえ（1）式の誤差項 ϵ_t が正規分布であっても，y_t の尖度は 3 を越える [30]．しかし，だからと言って，ϵ_t の分布が標準正規分布で良いとは限らない．ϵ_t に正規分布以外の分布も当てはめ，どの分布が最もフィットが良いかを分析することは重要である．

Watanabe and Asai（2003）らは，ϵ_t の分布に標準正規分布，スチューデントの t 分布，一般化誤差分布（Generalized Error Distribution; GED），２つの正規分布の混合分布を当てはめ，MCMC 法を用いてベイズ推定している．さらに前節で説明した周辺尤度の値を計算し，TOPIX では t 分布が，円ドルレートでは２つの正規分布の混合分布が最も当てはまりが良いとの結果を得ている．ϵ_t の分布に標準正規分布以外の分布を当てはめるときには，分布のパラメータもサンプリングする必要がある．例えば，t 分布の自由度については，Watanabe（2001）が効率的なサンプリング法を提案している．

別の推定法を用いて同様の分析を行っているものに，Liesenfeld and Jung（2000）がある．そこでは，シミュレーションによる最尤法（Simulated Maximum Likelihood estimation; SML）が用いられている．また，SV モデルではなく，ARCH 型モデルを使って同様の研究を行っているものに，Bollerslev，Engle and Nelson（1994），Watanabe（2000b），渡部（2000, 2.4.2 節）がある．これらの研究では，2 つの正規分布の混合分布は分析されておらず，Bollerslev，Engle and Nelson（1994），Watanabe（2000b），渡部（2000, 2.4.2 節）では代りに一般化 t 分布を用いた分析がなされている．これらの分析ではすべて，t 分布が最も当てはまりが良いとの結果が得られている．

Chib，Nardari and Shephard（2002）では，誤差項の分布を変えるとともに，ジャンプを加えたモデルを MCMC 法を用いてベイズ推定している．また，Jacquier，Polson and Rossi（2004）では，誤差項の分布を変えるとともに，以下で説明するボラティリティ変動の非対称性を考慮したモデルを MCMC 法を用いてベイズ推定している．

4.2　非対称 SV モデル

株式市場では，株価が上がった日の翌日よりも下がった日の翌日の方がボラティリティが高まる傾向があることが知られている（Black 1976）[31]．SV モデルでこうしたボラティリティ変動の非対称性を捉えるためには，以下のように誤差項 ϵ_t と η_t の間に相関を導入すればよい．

$$\begin{pmatrix} \epsilon_t \\ \eta_t \end{pmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_\epsilon^2 & \rho\sigma_\epsilon\sigma_\eta \\ \rho\sigma_\epsilon\sigma_\eta & \sigma_\eta^2 \end{bmatrix}\right).$$

ここで，もし $\rho < 0$ であれば，株式市場で観測されるボラティリティ変動の非対称性と整合的であり，$\rho = 0$ であれば，通常の非対称性の無い SV モデルになる．

Yu（2004），Omori and Watanabe（2003），Omori，Chib，Shephard and Nakajima（2004）らは，こうした非対称 SV モデルを MCMC 法を用いてベイズ推定している．Yu（2004）はボラティリティのサンプリングに single-move sampler を用いており，この方法は既に述べたように効率的でない．前節で説明した multi-move sampler は非対称 SV モデルには適用できないので，Omori and Watanabe（2003）は，非対称 SV モデルの multi-move sampler を新たに提案している．また，Omori，Chib，Shephard and Nakajima（2004）は mixture sampler を拡張している．Jacquier，Polson and Rossi（2004）は，ϵ_t と η_t の間ではなく，ϵ_t と η_{t-1} の間に相関を導入したモデルを single-move sampler を用いてベイズ推定しているが，Yu（2004）は ϵ_t と η_t の間に相関を導入した方が当てはまりが良いとの結果を得ている．

非対称 SV モデルを他の推定法を用いて推定しているものには，Melino and Turnbull（1990）と Harvey and Shephard（1996）がある．Melino and Turnbull（1990）は一般化積率法（Generalized Method of Moments; GMM），Harvey and Shephard（1996）はカルマン・フィルターに基づく疑似最尤法（Quasi-Maximum Likelihood estimation; QML）を用いている[32]．GMM や QML は MCMC 法を用いたベイズ推定法と比べて推定量の効率性が低い（Jacquier，Polson and Rossi 1994）．

4.3　マルコフスイッチング SV モデル

So, Lam and Li（1998），Kalimipalli and Susmel（2004），Shibata and Watanabe（2004）らは，ボラティリティの平均 μ の値が一定ではなく，高い時期（$S_t = 1$）と低い時期（$S_t = 0$）があるものと考え，（2）式を以下のように拡張している．

$$h_t = \mu_0 + \mu_1 S_t + \phi\,(h_{t-1} - \mu_0 - \mu_1 S_{t-1}) + \eta_t. \tag{44}$$

彼らは，Hamilton（1989）に従い，S_t はマルコフ過程に従うものと仮定している．

このモデルをマルコフ連鎖モンテカルロ法を用いてベイズ推定する場合には，$\boldsymbol{S}_T = (S_1, \ldots, S_T)$ を $f\,(\boldsymbol{S}_T | \boldsymbol{h}_T, \boldsymbol{\theta})$ からサンプリングしなけらばならないが，これは容易である．$\boldsymbol{y}_t$ が与えられると，（4.3）式は Hamilton（1989）のマルコフスイッチングモデルになるので，Carter and Kohn（1994）や Chib（1996）の multi-move sampler により $\boldsymbol{S}_T = (S_1, \ldots, S_T)$ を $f\,(\boldsymbol{S}_T | \boldsymbol{h}_T, \boldsymbol{\theta})$ から1度にサンプリングできる．

上記の論文では，（44）式を用いた場合，ϕ の推定値が大幅に低下している．また，Kalimipalli and Susmel（2004）は予測パフォーマンスの比較により，Shibata and Watanabe（2004）は前節で説明した周辺尤度の比較やモデルの診断により，単純な SV モデルよりもマルコフスイッチング SV モデルの方が当てはまりが良いとの結果を得ている．

4.4　取引高の導入

ボラティリティと取引高の間には正の相関が

あることが知られている．SV モデルを取引高を含める形で拡張したモデルに，Tauchen and Pitts（1983），Andersen（1996）らによる動学的 2 変量分布混合（Dynamic Bivariate Mixture; DBM）モデルがある．このモデルは，1 日のうちの 1 回 1 回の取引で生じる価格変化率と取引高は独立であるが，1 日の取引回数が日々変動することにより，日々のボラティリティと取引高との間に正の相関が生まれるとするモデルである．Tauchen and Pitts（1983）モデルでは，第 t 日の取引回数 I_t が与えられると，価格変化率 y_t と取引高 v_t は次のような互いに独立な正規分布に従う．

$$y_t\,|I_t \sim N\ (0,\ \sigma_r^2 I_t),\ \ v_t\,|I_t \sim N\ (\mu_v I_t,\ \sigma_v^2 I_t).$$

Andersen（1996）モデルでは，価格変化率の分布は同じで，取引高の分布が次のようなポアソン分布になる．

$$v_t|I_t \sim Po\ (m_0 + m_1 I_t).$$

いずれも，ln（I_t）が AR（1）モデルに従うものとし，v_t を無視すると，SV モデルになるので，これらのモデルは SV モデルを取引高を含める形で拡張したモデルになっている．そこで，このモデルのパラメーターもやはり最尤推定することが困難であり，また，このモデルは線形状態空間モデルに変換することができないので，mixture sampler は使えない．Watanabe（2000a, 2003）は multi-move sampler を用いてベイズ推定を行い，日経 225 先物の日次データでは，Tauchen and Pitts（1983）モデルは価格変化率の 2 乗の自己相関を過小評価し，Andersen（1996）モデルは取引高の自己相関を過小評価するとの結果を得ている．Lamoureux and Lastrapes（1994）は積率法（Method of Moments; MM），Andersen（1996）は一般化積率法（Generalized Method of Moments; GMM），Liesenfeld（1998，2001），Lisenfeld and Richard（2003）はシミュレーションによる最尤法（Simulated Maximum Likelihood estimation; SML）を用いて推定を行っている．

Asai and Watanabe（2004b）は，ボラティリティと取引高とを 2 変量自己回帰モデルで定式化したモデルを MCMC 法を用いてベイズ推定し，さらにインパルス応答関数の 95％信用区間を求めて，ボラティリティと取引高の間の因果性の分析を行っている [33)]．

5. 今後の課題

本論文でサーベイしたように，SV モデルは，近年，推定法の開発が進み，それに伴って，モデルの拡張が行われるようになってきた．ところが，ARCH 型モデルと比べると，実際のデータへの応用はまだそれほど多いとは言えない．特に，実際のデータを用いて SV モデルやそれを発展させたモデルと ARCH 型モデルとを比較するという試みはこれまであまり行われておらず，今後の重要な課題と言えよう．

比較の方法としては，まず，SV モデル，ARCH 型モデル共に MCMC 法を用いてベイズ推定し，周辺尤度を比較するという方法が考えられる．ARCH 型モデルの MCMC 法を用いたベイズ推定法も，Bauwens and Lubrano（1998），Nakatsuma（2000），三井・渡部（2003）らによって提案されている [34)]．また，ARCH 型モデルは尤度が解析的に評価できるので，SV モデルより簡単に周辺尤度を計算できる．そうした分析を行っているものには，Kim，Shephard and Chib（1998），Watanabe and Asai（2003）があり，いずれも，Bollerslev（1986）の GARCH モデルの誤差項 ϵ_t に t 分布のような裾の厚い分布を仮定し，SV モデルの誤差項 ϵ_t に標準正規分布を仮定した場合には，GARCH モデルの方が周辺尤度が高くなる場合があるが，両方同じ分布を用いた場合には，SV モデルの方が周辺尤度が高くなるとの結果を得ている．

また，ボラティリティの予測パフォーマンスによる比較も重要である．その際，問題となるの

はボラティリティの真の値が未知であるということである．これまでよく行われてきたのは，y_t^2をボラティリティの代理変数と考えて，それとSV，ARCH型モデルのボラティリティの推定値とを比較するという方法である[35)]．Andersen and Bollerslev（1998）は，（1）式からわかるように，y_t^2の変動はexp（h_t）の変動だけでなく，ϵ_t^2の変動も含んでいるため，この方法だとARCH型モデルやSVモデルのボラティリティの予測パフォーマンスを過小評価してしまうことを指摘しており，y_t^2ではなく，日中データから計算されるrealized volatilityを使うことを提案している．また，Deb（1997）は，SVモデルからシミュレーションによって人工的に発生させたデータを用いてボラティリティを推定するというモンテカルロ実験を行い，SVモデルを一般化積率法（GMM）や疑似最尤法（QML）といった正しい尤度に基づかない簡単な方法で推定するよりも，ARCH型モデルを最尤推定することによりボラティリティを推定した方がパフォーマンスが良いとの結果を得ており，興味深い．

最後に，ボラティリティはオプション価格を決定する重要な要因なので，オプション価格を用いた比較も重要である．SVモデルをMCMC法を用いてベイズ推定し，オプション価格を計算しているものに，王（2004）がある[36)]．

（東京都立大学経済学部教授）

注

1）ARCH型モデルについて詳しくは，Bollerslev, Engle and Nelson（1994），渡部（2000）等を参照のこと．

2）その他の推定法については，Ghysels，Harvey and Renault（1996），Shephard（2004），渡部（2000）を参照のこと．

3）MCMC法について詳しくは，大森（2001），中妻（2003）を参照のこと．

4）（2）式を2次以上のARモデルやARMAモデルに拡張するのは容易である．

5）$\mu = 0$，$\phi = 1$と仮定するモデルもあり，そうしたモデルをランダム・ウオークSVモデルと呼ぶ．ランダム・ウオークSVモデルについて詳しくは，Harvey，Ruiz and Shephard（1994）やRuiz（1994）を参照のこと．また，SVモデルにおいて，$\phi = 1$かどうかを検定する方法については，So and Li（1999）やWright（1999）を参照のこと．

6）Jacquier，Polson and Rossi（1994）は，それまでのSVモデルを推定した文献をサーベイし，ϕの推定値には0.8から0.995までの値が得られているとしている．

7）こうした従来のベイズ統計学については，鈴木・国友（1989）を参照のこと．

8）制約条件については，大森（2001）を参照のこと．

9）この捨てる最初のM回のことを"burn-in"と呼ぶ．

10）σ_η^2が逆ガンマ分布に従うというのは，逆数$1/\sigma_\eta^2$がガンマ分布に従うということである．

11）導出については，渡部（2004）を参照のこと．

12）（8）のような逆ガンマ分布からサンプリングするには，ガンマ分布からサンプリングして逆数をとればよい．正規分布やガンマ分布といったよく知られた分布からのサンプリングについては，Ripley（1987）を参照のこと．

13）導出については，渡部（2004）を参照のこと．

14）[0, 1]の一様分布からサンプリングを行い，得られた値をuとして，$u < q$であれば受容，そうでなければ棄却すればよい．

15）MHアルゴリズムの提案密度関数の選び方について詳しくは，Chib and Greenberg（1995）を参照のこと．

16）（11）式の定数はすべて異なることに注意．

17）この方法で，$|\phi| < 1$の範囲に入る確率が低い場合には，他の方法を使う必要がある．他の方法については，渡部（2000，p. 97）を参照のこと．

18）Shephard and Pitt（1997），渡部（2000，pp. 98-100）を参照のこと．

19）状態空間モデルについて詳しくは，Harvey（1989），Durbin and Koopman（2001）を参照のこと．

20）$(M, q_1, \ldots, q_M, m_1, \ldots, m_M, v_1, \ldots, v_M)$ の選択については，Kim，Shephard and Chib（1998），Mahieu and Schotman（1998），Omori，Chib，Shephard and Nakajima（2004）を参照のこと．

21）ここで，$(h_1, \ldots, h_{t-2})$，$(h_{t+k+2}, \ldots, h_T)$，$(y_1, \ldots, y_{t-1})$，$(y_{t+k+1}, \ldots, y_T)$ は条件から削除されていることに注意されたい．(2) 式より，h_t は1期前の値にのみ依存するので，h_{t-1} と h_{t+k+1} が与えられると，$(h_t, \ldots, h_{t+k})$ の分布を導出するのに，$(h_1, \ldots, h_{t-2})$ と $(h_{t+k+2}, \ldots, h_T)$ は必要ない．また，$(y_1, \ldots, y_{t-1})$，$(y_{t+k+1}, \ldots, y_T)$ がもたらすのは $(h_1, \ldots, r_{t-1})$ と $(h_{t+k+1}, \ldots, h_T)$ の情報だけなので，それらも必要ない．

22）ARMH アルゴリズムについて詳しくは，Chib and Greenberg（1995），渡部（2000 3.5.3 節）参照．

23）導出については，渡部（2004）を参照のこと．

24）観測方程式（29）の分散 v_s は（22）または（24）式で定義されるが，(22)，(24) 式の分母の $l''(\hat{h}_s)$ は（27）式より非正なので，v_s は非負である．ただし，(27) 式より，$y_s = 0$ の場合には，$l''(\hat{h}_s) = 0$ なので，(22) 式より，$v_s = \infty$ となってしまう．そこで，0 または 0 に近い値を含んでいるようなデータの場合には，(22) 式の分母および（21）の $l''(\hat{h}_s)$ を，c を 0 に近い負の定数（例えば，-0.00001）として，min $[l''(\hat{h}_s), c]$ に置き換えればよい．

25）(29)，(30) 式から成る線形ガウシアン状態空間モデルの下で，条件付事後分布（31）を，(19) 式と同様に展開すれば，(21) が得られる．

26）ϕ を MH アルゴリズムではなく，ARMH アルゴリズムでサンプリングする場合には，この方法は使えない．その場合には，Chib and Jeliazkov（2005）で提案されている方法を使う必要がある．

27）Particle filter について詳しくは，Pitt and Shephard（1999）を参照のこと．

28）Kim, Shephard and Chib（1998）は，h_t の1期先予測値 $h_{t|t-1} = \mu + \phi\left(M^{-1}\sum_{m=1}^{M} h_{t-1}^{(m)} - \mu\right)$ でテーラー展開しているが，そうすると AR アルゴリズムの受容確率が低く，サンプリングに時間がかかることがある．h_t の事後平均 $\hat{h}_t$ でテーラー展開すると，そうした問題は生じない．

29）(43) 式より，受容確率 p が 1 を越えることはなく，優越条件は満たされている．

30）詳しくは，渡部（2004）を参照のこと．

31）なぜそのような非対称性が生じるのかを分析しているものに，Christie（1982），Wu（2001）がある．

32）SV モデルの QML 推定については，渡部（2000 3.4 節）を参照のこと．

33）インパルス応答関数はパラメータの非線形関数になっている．MCMC 法を用いたベイズ推定では，パラメータを事後分布からサンプリングするので，サンプリングされたパラメータの値を代入するだけでこうした関数の値も事後分布からサンプリングしていることになり，統計的分析が行える．SV モデルではないが，同じくパラメータの非線形関数になっているものに，Hansen = Jagannathan bound があり，それについて MCMC 法を用いて同様の分析を行っているものに，Gordon, Samson and Carmichael（1996）がある．

34）Asai and Watanabe（2004a）は，これらの方法を比較し，三井・渡部（2003）の方法が最も効率性が高いとの結果を得ている．

35）詳しくは，渡部（2000 2.3.3 節）を参照のこと．

36）三井・渡部（2003）は，GARCH モデルを MCMC 法を使ってベイズ推定し，オプション価格を計算している．

参考文献

王暁明（2004）「SV モデルを用いたオプション価格付けの実証研究」COE ディスカッションペーパー No. 23，東京都立大学経済学部．

大森裕浩（2001）「マルコフ連鎖モンテカルロ法の最近の展開」『日本統計学会誌』第 31 巻，第 3 号，pp. 305-344.

鈴木雪夫・国友直人（1989）『ベイズ統計学とその応用』東京大学出版会.

中妻照雄（2003）『ファイナンスのためのMCMC法によるベイズ分析』三菱経済研究所.

三井秀俊・渡部敏明（2003）「ベイズ推定法によるGARCHオプション価格付けモデルの分析」『日本統計学会誌』第33巻，第3号，pp. 307-324.

渡部敏明（2000）『ボラティリティ変動モデル』朝倉書店.

渡部敏明（2004）「マルチ・ムーブ・サンプラーを用いた確率的ボラティリティ変動モデルのベイズ推定法」和合肇編『マルコフ連鎖モンテカルロ法を用いた応用計量分析』第9章，東洋経済新報社，近刊.

Andersen, T. G. (1996) “Return Volatility and Trading Volume: An Information Flow Interpretation of Stochastic Volatility,” *Journal of Finance,* Vol. 51, pp. 169-204.

Andersen, T. G. and T. Bollerslev (1998) “Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts,” *International Economic Review,* Vol. 39, pp. 885-905.

Asai, M. (2003) “Comparison of MCMC Methods for Estimating Stochastic Volatility Models,” *Computational Economics,* forthcoming.

Asai, M. and T. Watanabe (2004a) “Comparison of MCMC Methods for Estimating GARCH Models,” *COE Discussion Paper Series*, No. 18, Faculty of Economics, Tokyo Metropolitan University.

——— (2004b) “Stock Return Volatility and Trading Volume: A Bayesian Impulse Response Analysis,” *COE Discussion Paper Series*, No. 19, Faculty of Economics, Tokyo Metropolitan University.

Bauwens, L. and M. Lubrano (1998) “Bayesian Inference on GARCH Models using the Gibbs Sampler,” *Econometrics Journal,* Vol. 1, pp. 23-46.

——— (2002) “Bayesian Option Pricing Using Asym metric GARCH,” *Journal of Empirical Finance,* Vol. 9, pp. 312-342.

Black, F. (1976) “Studies of Stock Market Volatility Changes,” *1976 Proceedings of the American Statistical Association, Business and Economic Statistics Section,* pp. 177-181.

Bollerslev, T. (1986) “Generalized Autoregressive Conditional Heteroskedasticity,” *Journal of Econometrics,* Vol. 31, pp. 307-327.

Bollerslev, T., R. F. Engle and D. B. Nelson (1994) “ARCH Models,” in Engle, R. F. and D. McFadden (eds.), *The Handbook of Econometrics 4,* Amsterdam: North-Holland, pp. 2959-3038.

Carter, C. K. and R. Kohn (1994) “On Gibbs Sampling for State Space Models,” *Biometrika,* Vol. 81, pp. 541-553.

Chib, S. (1995) “Marginal Likelihood from the Gibbs Output,” *Journal of the American Statistical Association,* Vol. 90, pp. 1313-1321.

——— (1996) “Calculating Posterior Distributions and Modal Estimates in Markov Mixture Models,” *Journal of Econometrics,* Vol. 75, pp. 79-97.

Chib, S. and E. Greenberg (1994) “Bayes Inference for Regression Models with ARMA (p, q) Errors,” *Journal of Econometrics,* Vol. 64, pp. 183-206.

——— (1995) “Understanding the Metropolis-Hastings Algorithm,” *American Statistician,* Vol. 49, pp. 327-335.

Chib, S., F. Nardari and N. Shephard (2002) “Markov Chain Monte Carlo Methods for Generalized Stochastic Volatility Models,” *Journal of Econometrics,* Vol. 108, pp. 281-316.

Chib, S. and I. Jeliazkov (2001) “Marginal Likelihood from the Metropolis-Hastings Output,” *Journal of the American Statistical Association,* Vol. 96, pp. 270-281.

——— (2005) “Accept-Reject Metropolis-Hastings Sampling and Marginal Likelihood Estimation,” *Statistica Neerlandica*, forthcoming.

Christie, A. A. (1982) “The Stochastic Behavior of Common Stock Variances: Value, Leverage, and Interest Rate Effects,” *Journal of Financial Economics,* Vol. 10, pp. 407-432.

Danielsson, J. (1994) "Stochastic Volatility in Asset Prices: Estimation with Simulated Maximum Likelihood," *Journal of Econometrics,* Vol. 64, pp. 375-400.

Danielsson, J. and J. F. Richard (1993) "Acceralated Gaussian Importance Sampler with Application to Dynamic Latent Variable Models," *Journal of Applied Econometrics,* Vol. 8, pp. 153-174.

Deb, P. (1997) "Finite Sample Properties of the ARCH Class of Models with Stochastic Volatility," *Economics Letters,* Vol. 55, pp. 27-34.

Durbin, J. and S. J. Koopman (2001) *Time Series Analysis by State Space Models,* Oxford: Oxford University Press.

——— (2002) "A Simple and Efficient Simulation Smoother for State Space Time Series Analysis," *Biometrika,* Vol. 89, pp. 603-616.

Engle, R. F. (1982) "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica,* Vol. 50, pp. 987-1007.

Fama, E. F. (1965) "The Behavior of Stock Market Prices," *Journal of Business,* Vol. 38, pp. 34-105.

Ghysels, E., A. C. Harvey and E. Renault (1996) "Stochastic volatility," in Rao, C. R. and G. A. Maddala (eds.), *Statistical Methods in Finance,* Amsterdam: North-Holland, pp. 119-191.

Gordon, S., L. Samson and B. Carmichael (1996) "Bayesian Estimation of Stochastic Discount Factors," *Journal of Business & Economic Statistics,* Vol. 14, pp. 412-420.

Hamilton, J. D. (1989) "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," *Econometrica,* Vol. 57, pp. 357-384.

Harvey, A. C. (1989) *Forecasting Structural Time Series Models and the Kalman Filter,* Cambridge: Cambridge University Press.

Harvey, A. C., E. Ruiz and N. Shephard (1994) "Multivariate Stochastic Variance Models," R*eview of Economic Studies,* Vol. 61, pp. 247-264.

Harvey, A. C. and N. Shephard (1996) "Estimation of an Asymmetric Stochastic Volatility Model for Asset Returns," *Journal of Business & Economic Statistics,* Vol. 14, pp. 429-434.

Jacquier, E., N. Polson and P. Rossi (1994) "Bayesian Analysis of Stochastic Volatility Models (with Discussion) " *Journal of Business & Economic Statistics,* Vol. 12, pp. 371-417.

——— (2004) "Bayesian Analysis of Stochastic Volatility Models with Fat-tails and Correlated Errors," *Journal of Econometrics,* Vol. 122, Issue 1, pp. 185-212.

Kalimipalli, M. and R. Susumel (2004) "Regime-switching Stochastic Volatility and Short-term Interest Rates," *Journal of Empirical Finance,* Vol. 11, pp. 309-329.

Kim, S., N. Shephard and S. Chib (1998) "Stochastic Volatility: Likelihood Inference and Comparison with ARCH models," *Review of Economic Studies,* Vol. 65, pp. 361-393.

Lamoureux, C. G. and W. D. Lastrapes (1994) "Endogenous Trading Volume and Momentum in Stock-Return Volatility," *Journal of Business & Economic Statistics,* Vol. 12, pp. 253-260.

Liesenfeld, R. (1998) "Dynamic Bivariate Mixture Models: Modeling the Behavior of Prices and Trading Volume," *Journal of Business & Economic Statistics,* Vol. 16, pp. 101-109.

——— (2001) "A Generalized Bivariate Mixture Model for Stock Price Volatility and Trading Volume," *Journal of Econometrics,* Vol. 104, pp. 141-178.

Liesenfeld, R. and R. C. Jung (2000) "Stochastic Volatility Models: Conditional Normality versus Heavy-Tailed Distributions," *Journal of Applied Econometrics,* Vol. 15, pp. 137-160.

Liesenfeld, R. and J. F. Richard (2003) "Estimation of Dynamic Bivariate Mixture Models: Comments on

Watanabe (2000)" *Journal of Business & Economics Statistics,* Vol. 21, pp. 570-576.

Mahieu, R. J. and P. C. Schotman (1998) "An Empirical Application of Stochastic Volatility Models," *Journal of Applied Econometrics,* Vol. 13, pp. 333-360.

Mandelbrot, B. (1963) "The Variation of Certain Speculative Prices," *Journal of Business,* Vol. 36, pp. 394-419.

Melino, A. and S. M. Turnbull (1990) "Pricing Foreign Currency Options with Stochastic Volatility," *Journal of Econometrics,* Vol. 45, pp. 239-265.

Nakatsuma, T. (2000) "Bayesian Analysis of ARMA-GARCH Models: A Markov Chain Sampling Approach," *Journal of Econometrics,* Vol. 95, pp. 57-69.

Omori, Y., S. Chib, N. Shephard and J. Nakajima (2004) "Stochastic Volatility with Leverage," *CIRJE Discussion Paper* F-297, Faculty of Economics, University of Tokyo.

Omori, Y. and T. Watanabe (2003) "Block Sampler and Posterior Mode Estimation for a Nonlinear and Non-Gaussian State-Space Model with Correlated Errors," *CIRJE Discussion Paper* F-221, Faculty of Economics, University of Tokyo.

Pitt, M. K. and N. Shephard (1999) "Filtering via Simulation: Auxiliary Particle Filters," *Journal of the American Statistical Association,* Vol. 94, pp. 590-599.

Ripley, B. D. (1987) *Stochastic Simulation,* New York: John Wiley & Sons.

Rosenblatt, M. (1952) "Remarks on a Multivariate Transformation," *Annals of Mathematical Statistics,* Vol. 23, pp. 470-472.

Ruiz, E. (1994) "Quasi-maximum Likelihood Estimation of Stochastic Volatility Models," *Journal of Econometrics*, Vol. 63, pp. 289-306.

Sandmann, G. and S. J. Koopman (1998) "Estimation of Stochastic Volatility Models via Monte Carlo Maximum Likelihood," *Journal of Econometrics,* Vol. 87, pp. 271-301.

Shephard, N. (2004) *Stochastic Volatility: Selected Readings,* Oxford: Oxford University Press.

Shephard, N. and M. K. Pitt (1997) "Likelihood Analysis of Non-Gaussian Measurement Time Series," *Biometrika*, Vol. 84, pp. 653-667.

Shibata, M. and T. Watanabe (2004) "Bayesian Analysis of Markov Switching Stochastic Volatility Models," mimeo.

So, M. K. P., K. Lam and W. K. Li (1998) "A Stochastic Volatility Model with Markov Switching," *Journal of Buisiness & Economic Statlistics,* Vol. 16, pp. 244-253.

So, M. K. P. and W. K. Li (1999) "Bayesian Unit-Root Testing in Stochastic Volatility Models," *Journal of Buisiness & Economic Statlistics,* Vol. 17, pp. 491-496.

Tauchen, G. and M. Pitts (1983) "The Price Variability-Volume Relationship on Speculative Markets," *Econometrica,* Vol. 51, pp. 485-505.

Tierney, L. (1994) "Markov Chains for Exploring Posterior Distributions (with Discussion)," *Annals of Statistics,* Vol. 22, pp. 1701-1762.

Watanabe, T. (2000a) "Bayesian Analysis of Dynamic Bivariate Mixture Models: Can They Explain the Behavior of Returns and Trading Volume ?" *Journal of Business & Economic Statistics,* Vol. 18, pp. 199-210.

——— (2000b) "Excess Kurtosis of Conditional Distribution for Daily Stock Returns: The Case of Japan," *Applied Economics Letters,* Vol. 7, pp. 353-355.

——— (2001) "On Sampling the Degree-of-Freedom of Student-*t* Disturbances," *Statistics and Probability Letters,* Vol. 52, pp. 177-181.

——— (2003) "The Estimation of Dynamic Bivariate Mixture Models: Reply to Liesenfeld and Richard Comments," *Journal of Business & Economic Statistics*, Vol. 21, pp. 577-580.

Watanabe, T. and M. Asai (2003) "Stochastic Volatility Models with Heavy-Tailed Distributions: A Bayesian Analysis," *COE Discussion Paper Series,* No. 1, Faculty of Economics, Tokyo Metropolitan University.

Watanabe, T. and Y. Omori (2004a) "A Multi-move Sampler for Estimating Non-Gaussian Time Series Models: Comments on Shephard and Pitt (1997)" *COE Discussion Paper Series*, No. 8, Faculty of Economics, Tokyo Metropolitan University.

———— (2004b) "A Multi-move Sampler for Estimating Non-Gaussian Time Series Models: Comments on Shephard & Pitt (1997)," *Biometrika,* Vol. 91, pp. 246-248.

Wright, J. H. (1999) "Testing for a Unit Root in the Volatility of Asset Returns," *Journal of Applied Econometrics,* Vol. 14, pp. 309-318.

Wu, G. (2001) "The Determinants of Asymmetric Volatility," *Review of Financial Studies,* Vol. 14, pp. 837-859.

Yu, J. (2004) "On Leverage in a Stochastic Volatility Model," *Journal of Econometrics,* forthcoming.