論　説

マルコフ連鎖モンテカルロ法

大　森　裕　浩

1　はじめに

ベイズ統計学では周辺事後分布や事後確率の計算が困難になることが多い。もとの密度関数自体が複雑でなくても共役でない事前分布を用いると事後分布の形が複雑になったり，また未知母数の個数が非常に多いときにはひとつの未知母数のための分布型が複雑なために他の未知母数も影響を受けたりして，周辺事後分布の積分の計算が困難になる。このような場合，これまでは一般に数値積分を用いたが，積分範囲をどこまでとるべきかという基準や積分する単位区間はどれだけ小さくすれば十分かという基準は存在しなかった。この方法は計算量の負荷が重いので，場合によっては数値積分も行えないことも多い。この問題を解決するために事後分布を確率的なシミュレーションによって求めるという手法が1990年代になって積極的にとりあげられ，その理論的な裏付けも急速に進められてきた。

その手法にはdata augmentation, substitution sampling, Gibbs sampling, Metropolis-Hastings Algorithmなどがあるが，現在はマルコフ連鎖モンテカルロ法と総称している。このモンテカルロ法は，はじめは独立な標本にもとづく非反復的な手法が主に用いられたが，1990年になるとGelfand and Smith (1990)が，Geman and Geman (1984)が画

像復元のために使用したGibbs Samplerを離散分布から連続分布に拡張し，周辺事後分布の推定を条件付分布にもとづいて反復的に求める手法が主流となってきた。ここではまず，第2章で非反復的モンテカルロ法，第3章で反復的モンテカルロ法をとりあげ，第4章で手法の収束の理論的結果と実際の判定法，第5章で応用分野について紹介する。現在，マルコフ連鎖モンテカルロ法に関する包括的な論文は存在しないが，理論的な結果についてはTierney (1994)がもっとも優れている。またGibbs SamplingのサーベイとしてGelfand (1994)，初歩的な説明としてCasella and George (1992)，少し古いがマルコフ連鎖モンテカルロ法，EMアルゴリズム，data augmentationをたくさんの例とともにまとめたものがTanner (1993)にある。初期の論文としてGelfand and Smith (1990)もdata augmentation，substitution sampling，Gibbs sampling，importance samplingについて比較的わかりやすくまとめている。計量経済学への応用についてはChib and Greenburg (1995)が参考になるだろう。

2 非反復的モンテカルロ法

2.1 基礎的な手法

2.1.1 Importance Sampling

X，Yを確率変数ベクトル，$g(x)$をXの周辺密度関数，$f(y|x)$を$X=x$が与えられたときのYの条件付き確率密度関数とする。もし$g(x)$から確率変数を簡単に発生することができるならば$h(y)=\int f(y|x)g(x)\,dx$を求めるために

(1) $g(x)$から互いに独立なn個の標本$x_1,\cdots,x_m$を発生させ，

(2)　$\widehat{h(y)}=\frac{1}{m}\sum_{i=1}^{m} f(y|x_i)$を計算すればよい。

$h(y)$からの標本を発生させるには$\widehat{h(y)}$の代わりに$f(y|x_i)$からy_iを発生させればよい。しかし，$g(x)$から標本を発生させることが難しい場合にはこの方法は適用できないのでGeweke (1989) は$g(x)$に近い，標本を発生させやすい分布$I(x)$を用いて，

(i)　$I(x)$から互いに独立なm個の標本$x_1^*,\cdots,x_m^*$を発生させる。

(ii)　$\widehat{h(y)}=\Sigma_{i=1}^{m} w_i f(y|x_i)/\Sigma_{j=1}^{m} w_j$，$w_i=g(x_i^*)/I(x_i^*)$，を計算する。

とすると$I(x)$のsupportが$g(x)$のsupportを含むとき，確率１で$\widehat{h(y)}$が$\widehat{h(y)}$に収束することを証明した。$h(y)$からの標本を発生させるには(ii)の代わりに

(ii)*　$f(y|x_i^*)$からy_i^*を発生させる（$i=1,\cdots,m$）。

(iii)　（$y_1,\cdots,y_m$）に確率（$w_1,\cdots,w_n$）を与えて標本（$y_1^*,\cdots,y_n^*$）を発生させる。

とすればよい。Smith and Gelfand (1992) は，これを重みつきブーストラップと呼びmが大きくなるに従い$h(y)$から発生される確率標本に対する近似がよくなることを証明した。

2.1.2　受容／棄却法

もしimportance samplingで$g(x)$から標本を発生させにくい場合には次のような受容／棄却法が用いられる。関数$I(x)$がすべてのxについて$g(x)<cI(x)$（ただし，cは定数）であるならば，

(1)　$I(x)$から確率標本x^*を発生させる。

(2)　（０，１）区間上の一様乱数uを発生させる。

(3)　もし，$u<g(x^*)/cI(x^*)$ならばx^*を確率標本として受容し，そうでなければ棄却して(1)へ戻る。

(4)　(3)で得られたx^*を用いて$f(y|x^*)$から確率標本y^*を発生させる。

とする。(1)から(3)の手順でx^*の分布は$g(x)$に従うことが証明されている（Ripley (1987)）。しかし，これは$I(x)$の$g(x)$に対する近似がよくなければ受容率が低くなり，効率の悪い乱数発生法になる。また分布のすそにおいて条件$g(x)<cI(x)$が満たされるかどうかは判定が難しいのでcの値を大きくしてしまう傾向にあるが，この場合も受容率が低くなり効率が悪くなる。

2.2 Sampling/Importance Resampling

Rubin (1987)は一般に公開されているデータを用いて統計的推測を簡単に行うために，欠損値を埋める方法としてSampling/Importance Resampling (SIR)方法を提案した。未知の母数をθ，観測されるデータをY，観測されないデータ（欠損値）をZとすると，事後分布$p(Z|Y)$から欠損値Zを発生させるというものである。まず，事後分布$p(\theta,Z|Y)$に対してよく近似していて標本を発生させやすい密度関数

$$h(\theta,Z|Y)=h(\theta|Y)h(Z|\theta,Y)$$

を選ぶ（$h(\theta|Y)$，$h(Z|\theta,Y)$は確率標本を発生させやすい密度関数でそれぞれ$p(\theta|Y)$，$p(Z|\theta,Y)$を近似）。そして

(1) $h(\theta,Z|Y)$からM個の標本(θ_j,Z_j) $(j=1,2,\cdots,M)$を発生させる。

(2) 重みw_jを次のように計算する。

$$w_j=w(\theta_j,z_j|Y)\propto\frac{p(Y,z_j|\theta_j)p(\theta_j)}{h(\theta_j,z_j|Y)},\quad j=1,2,\cdots,M.$$

(3) $(z_1,z_2,\cdots,z_M)$ の中からm $(<M)$ 個のZを確率$Pr(Z=z_j)=w_j/\Sigma_{i=1}^{M}w_i$で発生させる。

w_jの分布が非常に歪んでいるときにはSIRによる近似は非常に悪くなると考えられるのでhのとり方を再検討する必要がある。

もし$p(Z|\theta,Y)=h(Z|\theta,Y)$ならば$w_j$は$z_j$に依存しなくなるから

(1)′　$h(\theta|Y)$からM個の確率標本（$\theta_j, j=1,2,\cdots,M$）を発生させる。

(2)′　重みw_jを次のように計算する。

$$w_j = w(\theta_j|Y) \propto \frac{p(Y|\theta_j)p(\theta_j)}{h(\theta_j|Y)},\quad j=1,2,\cdots,M.$$

(3)′　（$\theta_1, \theta_2, \cdots, \theta_M$）の中から$m$（$<M$）個の$\theta$を確率$Pr(\theta=\theta_j)=w_j/\Sigma_{i=1}^{M} w_i$で発生させ，得られた$\theta_j^*$を用いて$Z$を$p(Z|\theta_j^*, Y)$から発生させる（$j=1,2,\cdots,m$）。

$M/m \to \infty$のとき，m個の（Z_j^*, θ_j^*）は

$$\frac{h(\theta,Z|Y)w(\theta,Z|Y)}{\int\int h(\theta,Z|Y)w(\theta,Z|Y)d\theta dZ} = \frac{p(Y,Z|\theta)p(\theta)}{\int\int p(Y,Z|\theta)p(\theta)d\theta dZ}$$

$$= p(Z,\theta|Y)$$

を満たす。Rubin (1987)ではM/mの値について，欠損している情報の割合に依存するが一般に20がよいであろうとしている。

Gelfand and Smith (1990)はRubin (1987)のSIRを修正して(1)，(2)はSIRと同じだが(3)の代わりに(3)*を次のように修正し事後分布$p(Z|Y)$と$p(\theta|Y)$を得られるようにした。

(3)*　$p(Z|Y)$と$p(\theta|Y)$を次のように計算する。

$$p(Z|Y) = \frac{\Sigma_{j=1}^{M} w_j p(Z|Y,\theta_j)}{\Sigma_{j=1}^{M} w_j},\quad p(\theta|Y) = \frac{\Sigma_{j=1}^{M} w_j p(\theta|Y,z_j)}{\Sigma_{j=1}^{M} w_j}$$

しかし，Gelfand and Smith (1990)はその応用例でsubstitution samplingやGibbs samplingなどの反復的な手法に比べて（反復するための適当な構造が存在する場合には）SIRは分布の近似について劣っており，またその精度は$h(\theta,Z|Y)$の選び方に大きく左右されると指摘している。

2.3 Poor Man's Data Augmentation，モンテカルロEMアルゴリズム

Wei and Tanner (1990)は非反復的方法としてPoor Man's Data Augmentationを紹介している。Poor Man'sというのは計算環境に制約があり，後で紹介する通常のdata augmentationが十分にできない人のためのアルゴリズムという意味である。data augmentationで$p(Z|Y)$からの確率標本を発生することができない場合，

(a) $p(\theta|Y)$のモードを求めて$\hat{\theta}$とし，$p(Z|Y,\hat{\theta})$を用いて確率標本を発生させる。

(b) 得られた$z_1, z_2, \cdots, z_m$を用いて現在の$p(\theta|Y)$への近似をθの事後分布の混合分布$g_i(\theta)$として更新する。

$$g_i(\theta)=\frac{1}{m}\sum_{j=1}^{m}p(\theta|z_j, Y).$$

あるいは，(b)のかわりに事後分布を

(b)′

$$g_i(\theta)=\frac{\Sigma_{j=1}^{m}w_j p(\theta|Z^{(j)},y)}{\Sigma_j^m w_j},\quad w_j=\frac{p(z^{(j)}|y)}{p(z^{(j)}|\hat{\theta},y)}$$

のように求める。もし$p(Z|Y)$を計算するのが難しいならば2次近似を用いて重みw_jを次のようにとればよい。

$$w_j=\frac{(det\Sigma^*)^{1/2}p(\theta_j^*|Y,z_j)}{p(\hat{\theta}|Y,z_j)},\quad \Sigma^*=\left[-(1/n)\frac{d^2}{d\theta^2}\log p(\theta|Y,Z)\Big|_{\theta^*}\right]^{-1}$$

ただしθ^*は$\log p(\theta|Y,Z)$の最大値を与えるθ，$\hat{\theta}$は$\log p(\theta|Y)$の最大値を与えるθ（事後分布のモード）である。

この(a)はdata augmentationで初期値を選ぶのに用いることもできる。

EMアルゴリズムは局外母数が存在するときに事後分布をもとめるためのアルゴリズムで

(E)　まずE-Step（Expectation，期待値を求めるステップ）として，

$$Q_{i+1}(\theta,\theta_i)=E_{Z|\theta_i,Y}[\mathrm{loglog}p(\theta|Z,Y)],$$

$$=\int \log(p(\theta|z,y))p(z|\theta_i,y)\,dz$$

のようにZに依存しないように対数事後分布の期待値をとり，

(M)　次にM-Step（Maximization，最大化を行うステップ）としてこれを最大にするθを求め，θ_{i+1}とする。

そして，再びE-Stepにもどって反復し続けるとある条件のもとで（定理4.2.2，Tanner (1993)）θの値がある値に収束する（ただし，局所最大値，局所最小値，鞍点のいずれかはわからない)。このE-Stepで積分を計算するのが困難な場合にモンテカルロ法を用いるのがモンテカルロEMアルゴリズムであり(E)の代わりに

(E)′　$p(Z|\theta_i,Y)$からm個のimputation，$z_1,z_2,\cdots,z_m$を発生させて

$$Q_{i+1}(\theta,\theta_i)=\frac{1}{m}\sum_{j=1}^{m}\log(p(\theta|z_j,y))$$

とする。

3　反復的モンテカルロ法―マルコフ連鎖モンテカルロ法

まず，マルコフ連鎖についてTierney (1994) に従って定義と記号を導入しよう。

定義3.0.1　不変分布πをもつ定常な（*time-homogeneous*）マルコフ連鎖とは，空間Eに値をとる確率変数列$X_n(n\geq 0)$で，推移核として次のようなPをもつものである（通常Eはk次元ユークリッド空間で，πはσ-有限な測度μに関して密度をもつと仮定される)。すべての可測集合Aについて

$$P(X_n, A) = P\ \{X_{n+1} \in A | X_0, \cdots, X_n\},\quad \pi(A) = \int \pi(dx) P(x, A)$$

X_0の分布は連鎖の初期分布であり，X_0が与えられたときのX_nの条件付き分布は，P^nを推移核Pをn回繰り返すことと定義すれば

$$P\ \{X_n \in A | X_0\} = P^n(X_0, A)$$

となる。

定義3.0.2　不変分布πは，すべての可測集合Aについて

$$\lim_{n\to\infty} P^n(x, A) = \pi(A) \text{ for } \pi-\text{almost all } x$$

を満たすとき，連鎖の均衡分布と呼ばれる。

定義3.0.3　不変分布πをもつマルコフ連鎖は初期状態にかかわらずπが正の確率を与える集合にはいる確率が正であるとき既約であるという。

定義3.0.4　マルコフ連鎖は，一定の時間間隔で必ず訪れる状態空間があるとき周期的であるといい，そうでないとき非周期的であるという。

定理3.0.1（Tierney (1994)）

マルコフ連鎖が(1)properな不変分布をもち(2)既約で(3)非周期的であるとき，πは唯一の不変分布であり，また連鎖の均衡分布である。

この$\pi(x)$が我々の求めたい周辺分布であり，マルコフ連鎖を反復することにより均衡分布として求めるのがマルコフ連鎖モンテカルロ法である。マルコフ連鎖モンテカルロ法には大きく分けてSubstitution sampling，Gibbs Sampling，Metropolis-Hastingsアルゴリズムの３種類とその混合型がある。これについて以下の節で説明をしていく。

3.1　Data AugmentationとSubstitution Sampling

Tanner and Wong (1987)によって紹介されたdata augmentationは，substitution samplingと基本的に同じもので，観測されるデータYを観測できないデータZによって増大させる手法である。YとZが与えられ

たとき事後分布$p(\theta|Y,Z)$が簡単に計算できるか，確率標本の発生が簡単であることを仮定しており，そのアルゴリズムは，次のように与えられる。

(a)　予測分布$p(Z|Y)$に対する現在の近似分布からm個の標本z_1, $z_2,\cdots,z_m$を発生させる。

(b)　得られた$z_j(j=1,\cdots,m)$を用いて，現在の$p(\theta|Y)$に対する近似を以下のように事後分布の混合分布$g_i(\theta)$として更新する。

$$g_i(\theta)=\frac{1}{m}\sum_{j=1}^{m}p(\theta|z_j,Y).$$

(a1)　(b)で得られた$g_i(\theta)$から確率標本θ^*を発生させる。

(a2)　(a1)でθ^*を用いて$p(Z|\theta^*,Y)$から標本zを発生させて（a1)に戻ることを繰り返す。標本がz_1, $\cdots$, z_mとm個得られたら(b)に戻る。

Rubin (1987b)は$z_1,z_2,\cdots,z_m$をmultiple imputationsと呼び，(a), (a1)，(a2)をimputationステップ，(b)を事後ステップと呼んだ。

このdata augmentationアルゴリズムの収束に関して次の正規条件の下で重要な結果がTanner and Wong (1987)に示されている。

正規条件：

$K(\theta,\phi)=\int p(\theta|Z,Y)p(Z|\phi,Y)dZ$が一様に有界で$\theta$に関して同程度連続である。また，いかなる$\theta_0\in\theta$についても$\theta_0$の開近傍$U$が存在し，すべての$\theta,\phi\in U$について$K(\theta,\phi)>0$である。

定理3.1.1　Kが正規条件を満たすとき

1．$g^*(\theta)=p(\theta|Z,Y)$が唯一の$g(\theta)=\int K(\theta,\phi)g(\phi)d\phi$を満たす確率密度関数である。

2．初期値g_0が$\sup_\theta|g_0(\theta)/p(\theta|Z,Y)|<\infty$を満たすならば，$i\to\infty$のとき

$$\int |g_i(\theta)-p(\theta|Z,Y)|\,d\theta \to 0$$

標本空間が有限な場合には，ある条件のもとで真の事後分布に$O(\log n)$回（ただしnはデータの個数）の繰り返しののちに収束することがRosenthal (1993)によって証明されている。

Substitution Sampingのアルゴリズムは

(a1) 予測分布$p(Z|Y)$に対する現在の近似分布からm個の確率標本z^*を発生させる。

(a2) 得られたz^*を用いて$p(\theta|z,Y)$から確率標本θ^*を発生させる。

(a3) (a2)で得られたθ^*を用いて$p(Z|\theta^*,Y)$から確率標本z^*を発生させて(a2)に戻る。(a2)，(a3)をi回反復して最後に得られた標本を$(\theta^{(i)},Z^{(i)})$ とし(a1)に戻る。これをm回繰り返しm個の$(\theta^{(i)},Z^{(i)})$ を得られた順に$(\theta^{(i)}_j,Z^{(i)}_j)$，$(j=1,2,\cdots,m)$ とする。

(b) 得られた$(\theta^{(i)}_j,Z^{(i)}_j)$，$(j=1,2,\cdots,m)$ を用いて次のように事後分布を推定する。

$$p(\hat{\theta}|Y)=\frac{1}{m}\sum_{j=1}^{m}p(\theta|Z^{(i)}_j)$$

(a1)～(a3)を反復することは，$m=1$のdata augmentationに等しいのでTanner (1993)はchained data augmentationと呼んでいる。またsubstitution samplingも事後分布がL^1収束することが知られている(Gelfand and Smith (1990))。

この例ではXとYの2変数であったがZを含む3変数にも容易に拡張できる。また，XとYの2変数の場合はsubstitution samplingの反復のしかたはGibbs samplingと同じである。Gelfand and Smith (1990)は，3変数以上で条件付密度関数が$f(\theta_i|\theta_j,j\neq i)$だけしか知られていない場合には，substitution samplingとGibbs samplingは本質的に収束までの

計算時間が同じであること，計算方法はGibbs samplingのほうが簡単であるが，$f(\theta_1|\theta_2)$，$f(\theta_1|\theta_3)$，$f(\theta_1|\theta_2,\theta_3)$などのようにより多くの条件付密度関数が利用可能なときにはsubstitution samplingが速く収束するだろうことを指摘している。実際の問題への応用では変数の数が多く条件付密度関数の情報が豊富ではないことが多いことから，どちらかといえばGibbs samplerのほうが有用であるといえるだろう。

3.2　Gibbs Sampler

3.2.1　定義

Geman and Geman (1984)によって導入された画像復元のためのGibbs samplingアルゴリズムがGelfand and Smith (1990)によって離散分布から連続分布へと拡張され，さまざまな応用への可能性を広げた。未知の母数を$\theta=(\theta_1,\cdots,\theta_m)'$，観測されたデータを$Y$とすると，事後分布$p(\theta|Y)$の周辺分布の計算が難しいときに次のように反復的に$\theta$を発生させることにより，周辺分布からの標本を得るのが(systematic scan) Gibbs samplingである。

(1)　まず初期値として（$\theta_1^{(0)},\theta_2^{(0)},\cdots,\theta_m^{(0)}$）を適当な分布から発生させる。

(2)　$\theta_1^{(1)}$を$p(\theta_1|\theta_2^{(0)},\cdots,\theta_m^{(0)},Y)$ から発生させる。

(3)　$\theta_2^{(1)}$を$p(\theta_2|\theta_1^{(1)},\theta_3^{(0)},\cdots,\theta_m^{(0)},Y)$から発生させる。

(4)　$\theta_3^{(1)}$を$p(\theta_3|\theta_1^{(1)},\theta_2^{(1)},\theta_4^{(0)},\cdots,\theta_m^{(0)},Y)$から発生させる。同様に$\theta_4^{(1)}$，$\cdots$，$\theta_m^{(1)}$を順次発生させていく。

(5)　一般に$\theta^{(i)}=$（$\theta_1^{(i)},\theta_2^{(i)},\cdots,\theta_m^{(i)}$）が得られたら，

(i)　$\theta_1^{(i+1)}$を$p(\theta_1|\theta_2^{(i)},\cdots,\theta_m^{(i)},Y)$から発生させる。

(ii)　$\theta_2^{(i+1)}$を$p(\theta_2|\theta_1^{(i+1)},\theta_3^{(i)},\cdots,\theta_m^{(i)},Y)$から発生させる。

(iii)　$\theta_3^{(i+1)}$を$p(\theta_3|\theta_1^{(i+1)},\theta_2^{(i+1)},\theta_4^{(i)},\cdots,\theta_m^{(i)},Y)$から発生させる。同様

に$\theta_4^{(i+1)}$, $\cdots, \theta_m^{(i+1)}$を順次発生させていく。

というように$\theta^{(i+1)} = (\theta_1^{(i+1)}, \theta_2^{(i+1)}, \cdots, \theta_m^{(i+1)})$ を得るという作業を$(i = 1, 2, 3, \cdots)$ と繰り返していく。

このとき$\theta^{(0)}$, $\theta^{(1)}, \theta^{(2)}, \cdots$は

$$K(\theta^{(i+1)}, \theta^{(i)}) = p(\theta_1^{(i+1)} | \theta_2^{(i)}, \cdots, \theta_m^{(i)}, Y) \prod_{j=2}^{m-1} p(\theta_j^{(i+1)} | \theta_1^{(i+1)}, \cdots,$$

$$\theta_1^{(i+1)}, \cdots, \theta_{j-1}^{(i+1)}, \theta_{j+1}^{(i)}, \cdots, \theta_m^{(i)}, Y)$$

$$\times p(\theta_m^{(i+1)} | \theta_1^{(i+1)}, \cdots, \theta_{m-1}^{(i+1)}, Y)$$

のように$K(\theta^{(i+1)}, \theta^{(i)})$を推移核とするマルコフ連鎖であり，十分大きな値$N$について$\theta^{(N)}$が事後分布$p(\theta | Y)$の確率標本になる。このプロセスを独立に$t$回繰り返して$\theta^{(N),j}$ $(j = 1, 2, \cdots, t)$ が得られたとき$(1/t)\Sigma_{j=1}^{t} f(\theta^{(N),j})$が$t \to \infty$のとき$E_\theta[f(\theta) | Y]$ に収束することが知られている。

3.3.2 格子点Gibbs sampler

Gibbs samplerを行うためには条件付分布から標本を簡単に発生できる必要があるが，それが困難な場合はどうしたらよいであろうか。この問題に対してRitter and Tanner (1992)は次のような格子点上で行うGibbs sampler (Griddy Gibbs sampler) を提案した。$p(\theta_j | \theta_1, \cdots, \theta_{j-1}, \theta_{j+1}, \cdots, \theta_m, Y)$から標本を抽出することが困難なとき，

(1) 密度関数$p(\theta_j | \theta_1, \cdots, \theta_{j-1}, \theta_{j+1}, \cdots, \theta_m, Y)$の値を$n$個の点$\theta_{j1}$, $\theta_{j2}, \cdots, \theta_{jn}$で計算し，その値を$w_k$ $(k = 1, 2, \cdots, n)$ とする。

(2) w_k $(k = 1, 2, \cdots, n)$ を用いて$p(\theta_j | \theta_1, \cdots, \theta_{j-1}, \theta_{j+1}, \cdots, \theta_m, Y)$の分布関数の近似を求める。

(3) (0，1) 区間上の一様乱数を発生させ(2)で得られた近似分布関数を用いて確率標本を発生させる。

つまり，確率標本の発生が困難な条件付分布についてはn個の適当な格

子点を選んで分布関数を階段関数や線形関数で近似するのである。近似分布関数の求めかたは

(i)　密度関数を確率関数$p(\theta_{jk})=w_k/\Sigma_{l=1}^{n}w_l$を持つ離散分布で近似する。

(ii)　あるいは，密度関数を，区間ごとに値が一定の密度関数

$$p(\theta_{jk})=\frac{w_k}{\Sigma_{i=1}^{n}w_i(a_{i+1}-a_i)},\ x_{jk}\in[a_k,\ a_{k+1}]\ \text{のとき}$$

とする。

などが考えられる。この方法は密度関数の定数項の値を計算する必要がないが格子点の選び方，格子点の数により近似精度が大きく左右される。このためRitter and Tanner (1992)では，近似分布関数から発生したデータの経験分布関数を用いて格子点の位置や数を柔軟に変化させていくAdaptive Grid/Grid-grower法を提案している。格子点を用いたGibbs samplerの収束に関する理論的な結果は得られていない。

3.3　Metropolis-Hastingsアルゴリズム

3.3.1　Metropolisアルゴリズムの発展

Metropolisアルゴリズムは，もともと統計物理学の分野で相互に作用しあう分子からなるシステムの均衡状態をシミュレーションするためにMetropolis, Rosenblush, Rosenblush, Teller and Teller (1953)によって導入されたものである。統計物理学ではある温度Tでシステムが均衡状態にあるとき，エネルギーがxである確率はボルツマン分布

$$\pi(x)=\frac{1}{Z(T)}\exp\left(-\frac{x}{k_bT}\right)$$

(ただし，$Z(T)=\int\exp(-x/k_bT)\,dx$，$k_b$は既知の定数でボルツマン定数と呼ばれる）に従うことが，ボルツマン定理によって知られている。

Metropolis *et al*. (1953) は，この均衡状態へシステムがたどりつくまでの過程を次のようにシミュレーションした。まず，温度はTで一定とし，システムがエネルギーxの初期状態にあるとする。

(1) ランダムに分子を選んで移動する，つまり新しい状態を選択する。

(2) この新しい状態のエネルギーyがxより小さければそのまま新しい状態にとどまり，yを新しいxとして(1)に戻る。

(3) そうでなければ，確率

$$\frac{\pi(y)}{\pi(x)}=\exp\left\{-\frac{(y-x)}{k_b T}\right\}$$

で新しい状態にとどまりyを新しいxとして（そうでなければもとの状態xへ戻り），(1)へ戻る。

そして，この過程を反復していくとやがてボルツマン定理による，ある温度Tにおけるシステムの均衡状態へたどりつくことを示したのである。このアルゴリズムはMetropolisアルゴリズムと呼ばれ，Hasting (1970) によってMetropolis-Hastingsアルゴリズムに一般化され，Peskun (1973) によって対称な推移行列Qが与えられたとき最適な性質をもつことが示された（定理2.1.1）。

その一方でKirkpatrick, Gellat and Vecchi (1983) により，Metropolisアルゴリズムはsimulated annealingとしても一般化された。もとのアルゴリズムでは，温度が一定という仮定があったが，温度一定のもとで均衡状態が達成されたら，温度を少し下げて再び均衡状態を達成するまで待って温度をさげるということを繰り返して最終的に，最小限のエネルギー状態へ凍結させるアルゴリズムをつくった（このアルゴリズムは，マルコフ連鎖として捉え直すことによって，温度の低下が十分緩慢であるとき収束することが証明されている)。simulated annealingは，もともと金属や液体を冷却する際に均衡状態を保ちながら徐々に冷却していくことによって最終的に安定的な物質や結晶体に凍結させる目的でつくら

れたが，冷却化との類似性から次のように目的関数の最小化（最大化）問題に応用されている。

(1) 初期温度としてT_0をとり，目的関数$f(\theta)$の初期パラメータθ_0とそのもとでの関数の値$f(\theta_0)$を計算する。

(2) ランダムにパラメータ空間のなかからθを選択し目的関数$f(\theta)$を計算する。

(3) （0，1）上の一様乱数Uを発生させて

$$U \leq \exp\left\{-\frac{(f(\theta)-f(\theta_0))}{T}\right\}$$

ならばθを新しい点として受け入れるが，そうでなければもとのθ_0を選択する。

(4) 受け入れた点の系列が均衡状態に達すると判断されるまで(2)，(3)を繰り返して行ったあと(5)へ進む。

(5) (4)で受け入れた最後の点で計算された目的関数の値が，最小値に収束したかどうかを判定する。もし，収束しなかった場合には，温度T_0を予め決めておいた分だけ低くし，その点を新たにθ_0として(2)へ戻る。

simulated annealingの方法についてはBrooks and Morgan (1995)やIngber (1993)を参照するとよい。

3.3.2　一般的な定義

均衡分布πがμに関して密度関数をもち，推移核Qを$Q(x,dy)=q(x,y)\mu(dy)$とする。また$E^+=\{x：\pi(x)>0\}$として$x\notin E^+$は$Q(x,E^+)=1$を満たしπは退化した分布ではないと仮定する。するとMetropolis-Hasting核P_{MH}は次のように定義される。

$$P_{MH}(x,dy)=p(x,y)\mu(dy)+r(x)\delta_x(dy),$$

ただし，

$$p(x, y) = \begin{cases} q(x,y)\alpha(x,y), & x \neq y\text{のとき} \\ 0, & x = y\text{のとき} \end{cases}$$

$$r(x) = 1 - \int p(x,y)\mu(dy),$$

$$\alpha(x, y) = \begin{cases} \min\left(\dfrac{\pi(y)q(y,x)}{\pi(x)q(x,y)},\ 1\right), & \pi(x)q(x,y) > 0\text{ のとき} \\ 1, & \pi(x)q(x,y) = 0\text{ のとき} \end{cases}$$

$\delta_x(\cdot)$はxの上で1をとり，それ以外では0である関数である。このMetropolis-Hastingsアルゴリズムは，次のように進められる。

(1) 現在の点が$X_n=x$であるとき分布関数Qを用いてyを発生させX_{n+1}の候補とする。

(2) (1)で得られたyを確率$\alpha(x, y)$でX_{n+1}として受け入れる。受け入れなかった場合には$X_{n+1}=x$とする。

(3) (1)に戻る。

ここではπがP_{MH}の不変分布になっているが，P_{MH}が既約かどうかはQやπに依存している。もし，P_{MH}が既約で$\pi(\{x : r(x) > 0\}) > 0$ならばアルゴリズムは非周期的となる。以下では$E=R^k$，$\mu$はルベーグ測度，$f$は$E$上の密度関数として具体例を見ていく。

3.3.3 酔歩連鎖

$q(x, y)=f(y-x)$のとき，$Y=x+Z$，$Z\sim f(z)$となり，推移核Qは酔歩過程となる。Tierney (1994)はfの例として一様分布や正規分布，多変量t分布，Split-t分布（Geweke（1989））をあげている。もし，$f(z)>0$，$z\in R^k$(正符号条件という)ならばP_{MH}は既約で非周期的となり，収束に関する節で説明するようにマルコフ連鎖は収束する。正符号条件を満たしていなくてもfが原点の近傍で正でありE^+が開集合で連結されて

いれば，同様に収束する。しかし，一般に収束速度が遅くなることが多いので，Zの分布の分散の最適な大きさについて研究がなされている (Gelman, Roberts and Gilks (1994), Roberts, Gelman and Gilks (1994), Roberts and Rosenthal (1995))。

3.3.4　独立連鎖

$q(x, y)=f(y)$のとき，新しい点の候補yは密度関数fによって毎回独立に発生させられることになり，推移核Qは独立連鎖となる。この場合，新しい候補の点yを受け入れる確率$\alpha(x,y)$が

$$\alpha(x, y)=\min\left\{\frac{w_y}{w_x}, 1\right\}, \quad w_x=\frac{\pi(x)}{f(x)}$$

であるが，w_xはImportance samplingで$I(x)=f(x)$としたときの重みであることから，xよりyの重みのほうが大きいときには必ずyを選択し，そうでないときには重みの比に応じて選択する形になっている。このように独立連鎖はimportance samplingと関係が深いのでTierney (1994)はfの選択も自由度の低い多変量t分布かSplit-t分布がよいのではないかと指摘している。独立連鎖ではE^+上でfがμに関してほとんと正であるならばP_{MH}は既約で非周期的となりその収束が証明される。

3.3.5　受容／棄却法の連鎖

独立連鎖の1つに受容／棄却法を用いた連鎖がある。これは独立連鎖で$f(x)=\pi(x)$としたいが$\pi(x)$からの確率標本の発生が難しいためすべてのxについて$\pi(x)<cI(x)$，(cは正の定数)を満たすような$I(x)$を使って受容／棄却法によって$\pi(x)$からの確率標本の発生を行うものでZeger and Karim (1991)に応用例がある。しかし，すでに述べたように受容／棄却法には$I(x)$の近似精度の問題，cの選択の問題がある。そこで，これを修正して$\pi(x)/cI(x)$を仮定せず，$f(x)\propto\min(\pi(x), cI(x))$として独

立連鎖を行えばよい。$f(x)$からの確率標本を発生するには，受容／棄却法を用いて

(1)　$I(x)$から確率標本x^*を発生させる。

(2)　（0，1）区間上の一様乱数uを発生させる。

(3)　もし，$u<f(x^*)/cI(x^*)$ならばx^*を確率標本として受容し，そうでなければ棄却して(1)へ戻る。

この場合，すべてのxについて$f(x)<cI(x)$であるから，分布のすそに関する問題は生じない。最初にcを決めるときに$\pi(x)$の分布の中心で$\pi(x)<cI(x)$となるようにしてやればよいだけである。しかし確率標本発生の効率は，$I(x)$の$\pi(x)$への近似精度に依存していることに変わりはない。収束に関する結果は，$f(x)\propto\min(\pi(x),\ cI(x))$として独立連鎖についての結果がそのままあてはまる。

3.3.6　自己回帰連鎖

自己回帰連鎖は酔歩連鎖と独立連鎖の丁度中間にあたるもので，$q(x,y)=f(y-\phi x)$，つまり，$Y=\phi x+Z$，$Z\sim f(z)$，(ただしϕは定数か行列)とするものであり，$\phi=0$であれば独立連鎖，$\phi=1$ならば酔歩連鎖になっている。

3.3.7　格子点連鎖

Gibbs samplerのところで紹介した格子点による方法では，格子点を十分細かくとらなければならなかったが，Tierney (1994)ではMetropolis-Hastingsアルゴリズムが，その問題を次のように克服することができると指摘している。簡単化のために

(1)　適当なm個の格子点$x_1,x_2,\cdots,x_m$を決める。

(2)　m個の格子点のなかから一つのx^*を確率関数$Pr(X=x_i)=\pi(x_i)/\Sigma_{j=1}^{m}\pi(x_j)$ $(i=1,2,\cdots,m)$を用いて(SIRのときのように)

選択する。

(3)　Zを$f(z)$から確率標本を発生させ$Y=x^*+Z$とする。

このとき$q(x,y)$，$\alpha(x,y)$はそれぞれ次のようになる。

$$q(x,\ y)=\sum_{i=1}^{m}\left(\frac{\pi(x_i)}{\sum_{j=1}^{m}\pi(x_j)}\right)f(y-x_i),$$

$$\alpha(x,y)=\min\left\{\frac{\pi(y)\Sigma\pi(x_i)f(x-x_i)}{\pi(x)\Sigma\pi(x_i)f(y-x_i)},\quad 1\right\}$$

この方法が効率的に機能するためには，格子点が分布の台に十分広がっていること，密度関数fのすそが十分広がっていることが必要である。格子点のとりかたについてよくわからない場合にはTierney (1994)は，また現在の標本の点xを中心にして両側にhずつの増分でm個の格子点（従って$x-mh, x-(m-1)h, \cdots, x-h, x, x+h, \cdots, x+(m-1)h, x+mh$の$2m+1$個）の格子点をとることを提案している。

3.4　手法の組み合わせ

手法を組み合わせにはGibbs Samplerのように条件づけを使うもの以外に，混合型と循環型がある。$P_1, P_2, \cdots, P_m$をマルコフ核とすると，混合型は正の確率$\alpha_1, \alpha_2, \cdots, \alpha_m$に従って毎回マルコフ核を確率的に選択する方法で，循環型は$P_1, P_2, \cdots, P_m, P_1, P_2, \cdots$のように順番に核を選択して$P_m$の次に再び$P_1$に戻って繰り返すという方法である。

混合型の場合，m個の核のひとつが既約で非周期的であれば混合核もまた既約で非周期的になり収束が保証されるが，循環型の場合には必ずしも既約で非周期的にならず，ケースバイケースで調べなくてはいけない。

4　収束

4.1　分布の収束

4.1.1　一般的な結果

マルコフ連鎖によって発生された系列の分布の収束に関する証明はTierney (1994) に次のように二つの定理と二つの系によってまとめられている。これらの証明はNummelin (1984) に大きく依存している。

定理4.1.1 (Theorem 1, Tierney (1994))

Pがπ-既約で$\pi P=\pi$を満たすとき，Pは正再帰的でありπはPの唯一の不変分布である、もしPがまた非周期的ならばπに関してほとんどすべてのxについて$\|P^n(x,\cdot)-\pi\|\to 0$ (ただし，$\|\cdot\|$は*total variation distance*を表わす)が成り立つ。もしPが*Harris*再帰的であるならば，この収束はすべてのxについて成り立つ。

同様な定理はChan (1993) のTheorem 1.1でも与えられている。Tierney (1994) のTheorem 1, Theorem 2, Corollary 1から

系4.1.1 Pがπ-既約で$\pi P=\pi$を満たし非周期的であるとする。もし，$P(x,\cdot)$がπに関してすべてのxについて絶対連続ならばすべてのxについて$\|P^n(x,\cdot)-\pi\|\to 0$が成り立つ。

4.1.2　Metropolis-Hastings核の収束

Metropolis-Hastings核P_{MH}の収束についてはTierney (1994) のTheorem 1, Theorem 2, Corollary 2から

系4.1.2 P_{MH}がπ-既約で非周期的ならば，すべてのxについて$\|P_{MH}^n(x,\cdot)-\pi\|\to 0$が成り立つ。

P_{MH}の既約性の条件には次がある。

補題4.1.1　(1)（Theorem 3 (ii), Roberts and Smith (1994)）．もし(i) $q(x,y)$ が π-既約であり，(ii) $q(x,y)=0$ と $q(y,x)=0$ が同値であるとき（つまりすべての $(x,y)\in E\times E$ について $\alpha(x,y)>0$ であるとき）P_{MH} は π-既約である。

(2)（Lemma 1.1, Mengersen and Tweedie (1994)）．もし，すべての $(x,y)\in E\times E$ について $\pi(y)>0$ が $q(x,y)>0$ を意味するならば，P_{MH} は π-既約である。

また，P_{MH} の非周期性の十分条件としては次がある。

補題4.1.2　(1)（Section 2.4, Nummelin (1984)）．P_{MH} は π-既約で $(\{x : r(x)>0\})>0$ ならば非周期的である。

(2)（Lemma 1.2, Mengersen and Tweedie (1994)）．もし $\pi(x)$ と $q(x,y)$ がすべての (x,y) について正で連続であるならば P_{MH} は非周期的であり，$\mu(C)>0$（μ はルベーグ測度）であるようなすべてのコンパクト集合 C はsmall setである。

(3)（Theorem 3 (i), Roberts and Smith (1994)）．もし $q(x,y)$ が非周期的か，あるいはある数 $n\geq 1$ について $Pr(X_n=X_{n-1})>0$ が成り立つならば P_{MH} は非周期的である。

これらをまとめるとMengersen and Tweedie (1994)のLemma 1.1, Lemma 1.2とTierney (1994)のTheorem 1, Corollary 2に基づいて

命題4.1.1 $\pi(x)$ と $q(x,y)$ がすべての (x,y) について正で連続であるならば P_{MH} はすべての x について $\|P^n_{MH}(x,\cdot)-\pi\|\to 0$ を満たす。

4.1.3　Gibbs Samplerの収束

Gibbs samplerにてついてはそれぞれのケースで系4.1.1を確かめることによって収束を導くことができる。Chan (1993) もGibbs samplerの収束の十分条件を次のように与えている。

定理4.1.2　（Theorem 1.2, Chan (1993)）．$\pi(x)$ が状態空間 E 上で

常に正であるならば，Gibbs Samplerによるマルコフ連鎖はエルゴード的である（従って収束する)。

Chan (1993)は，さらに状態空間E上で$\pi(x)>0$となるための十分条件を示している（Condition (C2))。Roberts and Smith (1994)も次のようなGibbs samplerの収束の十分条件を与えている。

補題4.1.3 (Lemma 1, Theorem 2, Roberts and Smith (1994))．

(1)状態空間の測度が離散測度であるとき，Gibbs Samplerの推移核Pは定義可能で非周期的である。従ってPがπ-既約ならばGibbs samplerは収束する。

(2)状態空間の測度がルベーグ測度であるとき，πが0で下に半連続で，$\int \pi(x)\,dx_i$がすべてのiについて局所有界であり，Eが連結されているならばPはπ-既約でGibbs samplerは収束する。

4.1.4 組み合わせアルゴリズムの収束

変数がたくさんあり，乱数発生をすべて同時にではなく1回に1変数ずつ発生させていくMetropolis-Hastingsアルゴリズムの収束については次のような結果が得られている。

定理4.1.3 (Theorem 1, Chan and Geyer (1994))。

乱数発生を1回に1変数ずつ発生させていくR^d上のP_{MH}が，ルベーグ測度に関して絶対連続である均衡分布πをもつとする。このときP_{MH}がHarris再帰的であるための十分条件は，すべての条件つき推移核（無条件推移核P_{MH}を含む）が，相当する条件つきの均衡分布（πの均衡分布）と測度に関して固定された変数のいかなる値に対しても既約であることである。

従って、さらにこのP_{MH}が非周期的であればすべてのxについて$\|P^n_{MH}(x,\cdot)-\pi\|\to 0$を満たす。

4.2　収束の速度

4.2.1　収束速度の種類

マルコフ連鎖が収束する速度に関する理論的な結果は，さらに強い仮定を加えることによって証明されている。正Harris再帰的で非周期的なマルコフ連鎖はエルゴード的であるといわれるが，次のようにエルゴード性には，通常のエルゴード性よりも強いものとして(1)二次のエルゴード性，(2)幾何エルゴード性，(3)一様エルゴード性があり，(3)は(2)を，(2)は(1)を，(1)は通常のエルゴード性を意味するという関係になっている。

(1)　二次のエルゴード性。S_Bを集合Bへの到達時間とする。不変分布πをもつエルゴード的な連鎖は$\pi(B)>0$であるすべての$B \in E$について$\int_B \pi(dx) E_x[S_B^2] < \infty$を満たすとき，二次のエルゴード性をもつ。このとき$\pi$に関してほとんどすべての$x$について次が成り立つ。

$$n\|P^n(x, \cdot) - \pi\| \to 0 .$$

(2)　幾何エルゴード性。不変分布πをもつエルゴード的なマルコフ連鎖は，非負のπ可積分な拡張された実数値関数$M(x)$と正の定数$r<1$が存在して，すべてのxについて次が成り立つとき幾何エルゴード性をもつという。

$$\|P^n(x, \cdot) - \pi\| \leq M(x) r^n$$

(3)　一様エルゴード性。不変分布πをもつエルゴード的なマルコフ連鎖は，正の定数Mと正の定数$r<1$が存在して，すべてのxについて次が成り立つとき一様にエルゴード性をもつという。

$$\sup_{x \in E} \|P^n(x, \cdot) - \pi\| \leq M r^n$$

一般に二次エルゴード性の条件を確かめるのは難しいので，通常，幾何エルゴード性か一様エルゴード性の条件について調べられている。

4.2.2 収束速度を証明する道具

収束の速度を証明するのに必ずといっていいほど用いられる，幾何エルゴード性や一様エルゴード性の十分条件がminorization条件とdrift条件である。

- [minorization条件] π-既約なPは，整数$m \geq 1$，定数$\beta > 0$，集合$C \in E$，E上の確率測度vについて

  $\pi(C) > 0$，$\beta v(\cdot) \leq P^m(x, \cdot)$ for all $x \in C$

  を満たすとき，minorization条件M（m, β, C, v）を満たすという。また，もしPがあるm, β, vについてminorization条件M（m, β, C, v）を満たすとき，集合CはPのsmall setであるという。
- [drift条件] X_nがエルゴード的であるとき，非負の実数値E-可測関数g，small set C，定数$r > 1$，整数$m \geq 1$が存在して

  $$\sup_{x \in C^c} E[rg(X_{n+m}) - g(X_n) | X_n = x] = \sup_{C^c}(rP^m g - g) < 0$$

  $$\sup_{x \in C} E[g(X_{n+m})\,;\, X_{n+m} \in C^c \,|\, X_n = x] = \sup_{C} P^m(1_{C^c} g) < \infty$$

  であるときdrift条件を満たすという。

命題4.2.1 （Proposition 1，Tierney (1994)，Theorem 2.1，Chan (1993)）．X_nがdrift条件を満たすとき，X_nは幾何エルゴード性をもち，すべてのxに対して$\|P^n(x, \cdot) - \pi\| \leq P^n(a + bg(x))$となるような$a, b$，$0 < p < 1$が存在する。従って$g$が有界であるならば，$X_n$は一様エルゴード性をもつ。

この命題のGibbs samplerへの応用例はChan (1993)を参照するとよい。幾何エルゴード性における収束速度の上限については，Meyn and Tweedie (1994)，Mengersen and Tweedie (1994)，Rosenthal (1994, 1995a, 1995b) などが一般的な条件のもとで導いているが，実用の目的には十分とはいえない（例外的にRosenthal (1995c) がJames-Stein推定量に関連するモデルの中で意味のある上限を示している)。一様エルゴード性における収束速度の上限については，次のような結果が得られている。

定理4.2.1　(Proposition 2, Tierney (1994), Theorem 1.3, Mengersen and Tweedie (1994)).マルコフ推移核Pが一様エルゴード性をもつための必要十分条件は状態空間E全体がsmall setであることである。Pがminorization条件$M(m, \beta, E, v)$ を満たすならば，

$$\|P^n(x, \cdot)-\pi\| \leq (1-\beta)^{[n/m]}$$

(必要十分条件の別の表現はMengersen and Tweedie (1994) を参照)。

4.2.3　Metropolis-Hastingsアルゴリズムの収束速度の証明

Metropolis-Hastingsアルゴリズムについての結果をTierney (1994) は次のようにまとめている。

系4.2.1 (Corollary 3, Tierney (1994)). P_{MH}において$\mu(E^+)<\infty$, で$q(x, y)$と$\pi(x)$がE^+で有界で0から離れているとする。また，vをμのE^+への制限に比例する測度とする。このとき，P_{MH}がminorization条件$M(1, \beta, E, v)$ を満たすならば，P_{MH}は一様エルゴード性をもつ。

特に独立連鎖のP_{MH}については

定理4.2.2　(Corollary 4, Tierney (1994), Theorem 2.1 Mengersen and Tweedie (1994).) 独立連鎖のP_{MH}は，ある$\beta>0$が存在して

$$\frac{f(y)}{\pi(y)} \leq \beta, \quad y \in E$$

ならばP_{MH}は一様エルゴード性をもち，

$$\|P_{MH}^n(x, \cdot)-\pi\| \leq (1-\beta)^{[n]}$$

である。

この他，Mengersen and Tweedie (1994) は$q(x, y)=q(x-y)=q(y-$

x)で状態空間がRのP_{MH}は一様エルゴード性をもたない（Theorem 3. 1）が，ある条件のもとで幾何エルゴード性をもつことを証明し(Theorem 3. 2)，Roberts and Tweedie (1994)は状態空間がR^nの場合への拡張をπが指数型分布族に属する場合に行っている（Theorem 2. 1）。

4.2.4 Gibbs Samplerの収束速度の証明

Geman and Geman (1984)は有限な状態空間について，均衡分布が正符号条件（Besag (1974)）を満たすときGibbs Samplerがピアソンの距離について幾何収束することを証明した。(同時密度関数が直積標本空間において常に正であるならば，すべての条件付き分布を知っているということが，一意に同時密度関数を定めるということになる（Besag (1974))。すでに述べたようにChan (1993)では幾何エルゴード性のための条件（Theorem 2. 1）が導出されており，Gibbs samplerへの応用例も豊富に与えられている。Roberts and Polson (1994)は定理4.2. 1の$m=1$の時の結果や(Lemma 2)，Gibbs samplerの収束を調べやすくするためにコンパクト性と連続性による条件を示している（Theorem 1, Corollay 2, 3)。

系4.2. 2 (Corollary 3, Roberts and Polson (1994))．もし，半連続な非負の関数$g_i: R^i \to R^+$, $1 \leq i \leq p$が

(1) $\pi(x_i^{(n)} | x_j^{(n)},\ 1 \leq j < i,\ x_k^{(n-1)},\ i < k \leq p) \geq g_i(x_1^n, \cdots, x_i^n)$

(2) 集合 $\{x_1^n, \cdots, x_i^n : g_i(x_1^n, \cdots, x_i^n) > 0\}$ は空集合ではない。

ならばGibbs Samplerは一様エルゴード性をもつ。

Shervish and Carlin (1992)はGibbs sampler, Tanner and Wong (1987)のData augmentationを含むSubstitution samplingの手法の収束について，測度$\mu(A) = \int_A (1/f_Y(y))\, d\lambda^p(y)$，(ただし$\lambda^p$は$\lambda^p = \lambda_1 \times \cdots \times \lambda_p$である直積測度）を定義し，この測度に関して二乗可積分な関数のヒルベルト空間Hを用いて，次のような結果を導いた（この結果はBaxter and

Rosenthal (1994)で一般のPのために拡張された)。

系4.2.3 (Corollary 1, Shervish and Carlin (1992)). もし，Substitution samplingの推移核Pが

$$\int\int |P(x,\ y)|^2 d\mu(x)\, d\mu(y) < \infty$$

で，$P(x,y) > 0$ が$\mu \times \mu$に関してほとんど確実に成り立つとする。また，π_0をμに関して絶対連続な分布関数で$d\pi_0 / d\mu = u_0 \in H$とするとき，$X_n$は幾何収束する。

Schervish and Carlin (1992)は，実際の応用では初期値の分布は特定の固定された値であることが多いが，この分布をπ_{-1}とみなして1回，連鎖を進めたπ_0は，結局仮定を満たすことを指摘している。同様な結果をLiu, Wong and Kong (1995)も証明している。

4.2.5 組み合わせアルゴリズムの収束速度

最後に，同じ不変分布をもつ推移核を組み合わせた場合，先に説明した混合型では推移核の一つが一様エルゴード性をもてば全体も一様エルゴード性をもつのに対して，循環型では推移核の一つが一様エルゴード性をもつだけでなくさらにminorization条件$M(m, \beta, E, v)$ をあるβとvについて満たすとき全体も一様エルゴード性をもつことが知られている (Proposition 3, 4, Tierney (1994))。たとえば，組み合わせの中に重み関数$w = \pi / f$が有界な独立連鎖のMetropolis核があれば全体として一様エルゴード性をもつ。

4.3 関数の期待値の収束

マルコフ連鎖モンテカルロ法を用いてπ可積分な関数fの期待値$E[f(x)]$ を求める場合，その推定量は

$$\hat{f}_n = \frac{1}{n} \sum_{i=1}^{n} f(x_i)$$

となる。エルゴード性をもつ系列を使った推定量の漸近的な性質は大数の法則に基づいて次のように得られている。

定理4.3.1 (Theorem 1, Roberts and Smith (1994)). 推移核Pがπ-既約で非周期的であり実関数fが$E|f(X)|<\infty$を満たすならば，確率1で$\hat{f}_i \to E[f(X)]$ である。

定理4.3.2 (Theorem 3, Tierney (1994)). X_nがエルゴード的で均衡分布πを持ち，実関数fが$E|f(X)|<\infty$を満たすならば，初期分布にかかわらず確率1で$\hat{f}_i \to E[f(X)]$ である。

定理4.3.3 (Theorem 4, Tierney (1994)). X_nが2次エルゴード的で均衡分布πを持ち，実関数fが有界ならば$\sigma(f)$が存在して，初期分布にかかわらず$\sqrt{n}(\hat{f}_n - E[f(X)]) \Rightarrow N(0, \sigma^2(f))$である。

定理4.3.4 (Theorem 5, Tierney (1994)). X_nが一様エルゴード的で均衡分布πを持ち，実関数fが$E|f(X)|^2<\infty$を満たすならば実数$\sigma(f)$が存在して，初期分布にかかわらず$\sqrt{n}(\hat{f}_n - E[f(X)]) \Rightarrow N(0, \sigma^2(f))$である。

4.4 収束の判定法

実際にマルコフ連鎖モンテカルロ法を応用する場合，反復を何回以上行えば均衡分布に収束するのかという問題が生じる。収束の速度の節で述べたように，収束に必要な反復回数を理論的に導く試みはなされているが，実用性に欠けているため現在の段階では，まず反復を行い得られた系列を用いてマルコフ連鎖が収束しているかどうかを検査するという方法がとられる。ここではそのなかで応用範囲にひろいと思われるものについてとりあげていく（収束の判定手法の包括的なサーヴェイや実験に基づく比較についてはCowles and Carlin (1994)を参照するとよい)。

4.4.1　Gelman and Rubin (1992)

Gelman and Rubin (1992)は，均衡分布が正規分布であるマルコフ連鎖を考えて，発生された確率標本の分布をt分布によって(従って分散の大きな分布で)近似し，その分散が正規分布の分散に近いかどうかによって連鎖の分布が収束しているかどうかを判断しようとした。具体的には次のようなステップを踏んで行われる。

(1)　1つの連鎖の長さが$2n$であるようなm（≥ 2）個の独立のマルコフ連鎖を発生させる。X_{ij}を第iの連鎖の第j番目の観測値とする。（$i=1,\cdots,m$，$j=1,\cdots,2n$）とする。

(2)　(1)で得られたデータで，各連鎖について最初のn個の分を初期値に依存するものとして捨てて，残りの後半のn個を用いて次の$\sqrt{R}$を計算する

$$\sqrt{R}=\sqrt{\left(\frac{n-1}{n}+\frac{m+1}{m}\frac{B/n}{W}\right)\frac{df}{df-2}},$$

$$\frac{B}{n}=\frac{\sum_{i=1}^{m}(\bar{x}_{i.}-\bar{x}_{..})^2}{m-1},\quad \bar{x}_{i.}=\frac{\sum_{j=n+1}^{2n}x_{ij}}{n},\quad \bar{x}_{..}=\frac{\sum_{i=1}^{m}\bar{x}_{i.}}{m}$$

$$W=\frac{\sum_{i=1}^{m}s_i^2}{m},\quad s_i^2=\frac{\sum_{j=n+1}^{2n}(x_{ij}-\bar{x}_{i.})^2}{n-1}$$

$$df=\frac{2\hat{V}^2}{\widehat{Var}(\hat{V})},\quad \hat{V}=\frac{n-1}{n}W+\frac{m+1}{mn}B$$

$$\widehat{Var}(\hat{V})=\left(\frac{n-1}{n}\right)^2\left(\frac{1}{m}\right)\frac{\sum_{i=1}^{m}(s_i^2-W)^2}{m-1}+\left(\frac{m+1}{mn}\right)^2\frac{2B^2}{m-1}$$

$$+\frac{2(m+1)(n-1)}{mn^2}\frac{n}{m}\left\{\frac{\sum_{i=1}^{m}(s_i^2-W)(\bar{x}_i^2-\overline{\bar{x}^2})}{m-1}\right.$$

$$-2\bar{x}_{..}\frac{\sum_{i=1}^{m}(s_i^2-W)(\bar{x}_{i.}-\bar{x}_{..})}{m-1}\Bigg\},\quad \overline{\bar{x}^2_{..}}=\frac{\sum_{i=1}^{m}\bar{x}^2_{i.}}{m}$$

ここでB/nは連鎖間での変動を，Wは各連鎖内の変動を表わしている。

Gelman and Rubin (1992)は，この$\sqrt{R}$が1に近いかどうかで収束を調べ，1に近い場合は収束していると考えて各連鎖の後半n個の系列をあわせて均衡分布の標本（最初のn個は初期値に依存していると考えられるため）とし，1に近くない場合には収束が起こっておらずに標本が初期値に依存していると考え，nの値を大きくして再び連鎖を発生させてから$\sqrt{R}$を計算することを提案した。これは基本的に均衡分布が一次元分布の場合に考えられたものであるが多次元分布の場合には各変数に対してこの$\sqrt{R}$を計算して総合的に収束の判定を行うことになる。

この方法はどんなマルコフ連鎖モデルにも応用が可能だが，正規分布の理論を基礎に考えられているため，均衡分布が正規分布とかなり異なる場合には注意が必要である。たとえばGelman and Rubin (1992)も指摘しているように均衡分布が2峰型であるときにm個の連鎖の初期値がすべて一方のモードの周辺でありnの値が小さいときにはそのモードの周辺ばかりで確率標本の発生が行われて収束したと誤認し，もう一方のモードを検出できないことが起こりうるからである。これを防ぐためには均衡分布の形状に関する情報をできるだけ得て，初期値も均衡分布の台に広く散らばるように発生させることが必要である。このGelman and Rubin (1992)の方法は，すでにSのプログラムとしてカーネギーメロン大学統計学部のStatlibで公開されており，電子メールでstatlib@stat.cmu.eduへsend itsim from Sという内容を送ればプログラムが自動的に送られてくる。

4.4.2 Raftery and Lewis (1992), (1994)

Raftery and Lewis (1992)では，Gelman and Rubin (1992)とは異なり，マルコフ連鎖モンテカルロ法は1つの長いマルコフ連鎖を発生させてそこからk個おきに標本を抽出するものと想定している。そして以下で説明するような問題において指定した推定の精度を満たすような連鎖の長さkの大きさ，初期値に依存している標本の系列の長さ（Burn-in期間と呼ばれる）を理論的に導いてくれる。

まずマルコフ連鎖Xの関数Uについて分布関数$Pr(U \leq u | Data) = q$（uは固定された点）の値の100s%信用区間を「点推定量±r」のように求める問題を考える。$Pr(U \leq u | Data)$の推定量を作るために$Z_t = I(U_t \leq u)$,（$t = 1, 2, \cdots$）（Iは定義関数）を定義する。しかし，Z_t自体はマルコフ連鎖ではないので，十分大きなkをとることによって$Z^{(k)}_t = Z_{1+(t-1)k}$を近似的にマルコフ連鎖とする（Raftery and Lewis (1992)では厳密な証明は与えられていない)。すると$Z^{(k)}_t$は，次のように均衡分布π，推移行列Pで与えられるマルコフ連鎖と考えられる。

$$\pi = (\pi_0,\ \pi_1) = (q,\ 1-q) = \left(\frac{\alpha}{\alpha+\beta},\ \frac{\beta}{\alpha+\beta}\right),\quad P = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}$$

ここで

$$P = \begin{pmatrix} \pi_0 & \pi_1 \\ \pi_0 & \pi_1 \end{pmatrix} + \frac{(1-\alpha-\beta)^m}{\alpha+\beta}\begin{pmatrix} \alpha & -\alpha \\ -\beta & \beta \end{pmatrix}$$

であることから$Z^{(k)}_t$の系列も初期値に依存するので，最初のm個を除く必要がある。$Z^{(k)}_t$の確率関数の値が均衡分布の確率関数の値±ϵになるようにmを定めると

$$m = \left\{ \log\left(\frac{\epsilon(\alpha+\beta)}{\max(\alpha, \beta)}\right)\right\} / \log(1-\alpha-\beta)$$

となるので$Z^{(k)}{}_t$を発生させて$m+1$個目からの標本を使う。すると$\bar{Z}=\Sigma_{t=m+1}^{m+n} Z^{(k)}{}_t/n$に中心極限定理を用いて$q$のための$100s$%信用区間が

$$Pr\ [\bar{Z}-r\leq q\leq \bar{Z}+r]=s,\quad r=z_{(1-s)/2}\sqrt{\frac{\alpha\beta(2-\alpha-\beta)}{n(\alpha+\beta)^3}}$$

（ただし，z_αは標準正規分布の上側$100(1-\alpha)$%点である）が与えられる。rが予め与えられた値になるようにするにはnを

$$n=\frac{z^2_{(1-s)/2}}{r^2}\frac{\alpha\beta(2-\alpha-\beta)}{(\alpha+\beta)^3}$$

ととればよい。$M=mk$，$N=nk$とすると，全体として$M+N$個の標本をマルコフ連鎖モンテカルロ法で抽出して，最初のM個を初期値に依存するものとして棄て，残りのN個を均衡分布からの標本とすれば良いことになる。

m，nの値を決めるためには，q,r,s,ϵについては実験者が与えてやらなければならないが，k,α,βについては観測値から推定しなくてはならない。そこで，Raftery and Lewis (1992) はまず試験的に与えられたq,r,s,ϵのもとで最小の$N=N_{\min}$でマルコフ連鎖モンテカルロ法を行うとした。$N=N_{\min}$となるのはZ_tが互いに独立な場合であると考えられるから，$1-\alpha=\beta=\pi_0=q$，$M=0$，$k=1$となり，$N_{\min}=\frac{z^2_{(1-s)/2}q(1-q)}{r^2}$で与えられる。この$N_{\min}$個のデータから$\alpha,\beta$を推定する。$k$についても，得られた観測値を用いて各$k$の値毎に１次のマルコフ過程か２次のマルコフ過程かをBIC基準によって選択し，１次のマルコフ過程を選択するような最初のkを求めるkとする（Raftery (1986)，Bishop，Fienberg and Holland (1975)）。この推定されたα，β，kを用いてM，Nを計算し必要ならば最初に発生させた$N_{\min}$の系列に付加的にマルコフ連鎖モンテカルロ法を行い，再度α,β,kを推定させM，Nを計算する，ということを繰り返して行けばよい。

Raftery and Lewis (1992) では，主として$q=0.025$，$s=0.95$，$\epsilon=0.01$

をとり，rについては$r=0.0125$（精度に厳しい場合には$r=0.005$）を勧めている。このうち$q=0.025$は恣意的であり，実際にはいくつかの異なるqに対して計算をしなくてはならないだろう。この手法は，すべてのマルコフ連鎖モンテカルロ法について有効であり，またプログラムがFORTRANとSで書かれたものが前述のカーネギーメロン大学のStatlibにあり，電子メールでstatlib@stat.cmu.eduあてにsend gibbsit from generalあるいはsend gibbsit from Sと送れば手に入れることができる。

4.4.3　Heidelberger and Welch (1983)

Heidelberger and Welch (1983) は「マルコフ連鎖が一様エルゴード性を持つ」という帰無仮説を検定する方法を与えている。もし，マルコフ連鎖$X_i(i=1,2,\cdots,n)$が一様エルゴード的であるならば，$S(0)$をスペクトル密度関数の推定値を0で評価したものとすると

$$B_n(t)=\frac{\sum_{i=1}^{[nt]}X_i-[nt]\bar{X}}{\sqrt{nS(0)}},\quad 0\leq t\leq 1$$

の分布が近似的にBrownian bridgeであることからCramér-von Mises統計量

$$\int_0^1 B_n^2(t)\,dt$$

（統計量は数値積分によって計算される）を用いて次のようなプロセスで収束の判定は行われる。

(1)　まず，反復回数の上限として$j_{\max}$，平均のための信頼区間の1/2の幅の（平均に対する）相対的な大きさをϵと決めて，試験的に$j_1=0.1j_{\max}$回だけ反復を行う。

(2)　反復したうちの後半50%を使ってスペクトル密度の推定値$S(0)$を計算し，検定を行う。

(3)　もし帰無仮説が棄却されたら，X_nの最初の10%を除いて検定を再

び行う。それでも棄却されれば，さらにまた同じ10%の長さだけ除いて検定を行うということを繰り返し，受容されれば(5)に行き，そうでなければ残りの長さがもとの50%になるまで続け，(4)に進む。

(4) 検定のための反復回数が十分でないため，さらに反復を行い合計でこれまでの1.5倍分の反復をし(ただし$j_{\max}$を超えない)，(2)に戻る。反復回数が$j_{\max}$でも棄却されたときには連鎖は収束していないと結論する。

(5) 検定に使ったすべてのX_n（n^*個とする）を用いて平均，スペクトル密度を推定し直して$\bar{X}^*, S^*(0)$とすると$z_{\alpha/2}\sqrt{S^*(0)/n^*}<\epsilon\bar{X}^*$ならば収束していると判定し作業を終了する。そうでなければ(4)のように反復をさらに行って(2)に戻る。

この方法はすべてのマルコフ連鎖に適用できるが，その検定は連鎖の初期値に依存する長さに対して$j_{\max}$が小さすぎると検出力が低いことが指摘されている。

4.4.4　Geweke (1992)

Geweke (1992)ではX_nをマルコフ連鎖，gをその関数としてgのスペクトル密度$S_g(\omega)$が存在し，$\omega=0$であるとき

$$\frac{g_n-E[g(X)]}{\sqrt{S_g(0)/n}}\Rightarrow N(0,1),\quad g_n=\frac{1}{n}\sum_{i=1}^{n}g(X_i)$$

であることを利用して，次のような手順で連鎖の収束を判定するものである。この仮定が満たされる限り，この方法はすべてのマルコフ連鎖に適用できる。

(1) X_i（$i=1,2,\cdots,n$）を発生させ，$g_{n_1}=\frac{1}{n_1}\Sigma_{i=1}^{n_1}g(X_i)$，$g_{n_2}=\frac{1}{n_2}\Sigma_{i=n-n_2+1}^{n}g(X_i)$と，スペクトル密度$Sgn_1(0)$，$Sgn_2(0)$を計算する(通常$n_1=0.1n$，$n_2=0.5n$とする)。

(2)　標準正規分布のz_αを上側100α%点，$Z=(g_{n_1}-g_{n_2})/\sqrt{Sgn_1(0)/n_1+Sgn_2(0)/n_2}$として$|Z|\leq z_{\alpha/2}$ならば収束をしていると判定する。

Cowles and Carlin (1994) は，この方法はスペクトルウインドウのとりかたに左右されやすいことを指摘している。

4.4.5　Liu and Liu (1993)

Liu and Liu (1993) では，Gibbs Samplerによるマルコフ連鎖の収束を調べるために，まず，ある分布に基づいて初期値をm個とりだしそこからそれぞれ別々に連鎖を開始して次のような統計量を計算する。二つの異なる連鎖，第i連鎖と第j連鎖のt回目の反復で得られたXとマルコフ推移核Kを用いて

$$U^{(i,j,t)}=\frac{\pi(X^{(j,t)})K(X^{(j,t-1)},X^{(i,t)})}{\pi(X^{(i,t)})K(X^{(j,t-1)},X^{(j,t)})}$$

とする。このとき$p^t(X)$を反復時点tにおける$\pi(X)$への近似分布とすると

$$E_0(U^{(i,j,t)})=Var_\pi\left(\frac{p_t(X)}{\pi(X)}\right)+1$$

であることから，Uが$p_t(X)$と$\pi(X)$の距離の不偏推定量になっている。そこでLiu and Liu (1993) は，各反復時点tで$m(m-1)$個のUの対数の平均$\overline{\log U_t}$を計算してtに対してプロットしてX軸に平行な直線に近づくならば$p_t(X)$が$\pi(X)$に収束すると考えた。ほかにもUの系列を$m/2$組のペアに分けてRubin and Gelman (1992) の方法を応用したり，各時点tでのUの累積分布関数をプロットしたりするなどが提案されている。この方法はπとKが明示的に求まる場合に有効で，Gibbs Samplerによるマルコフ連鎖などへ応用できる。

4.4.6 Gibbs Stopper

Ritter and Tanner (1992) は次のように現在の段階で予測された事後分布と，均衡分布である真の事後分布の比を用いてGibbs Samplerの収束を判定しようとした。つまり，

$$w_i = \frac{q(\theta^{(i)}_1, \theta^{(i)}_2, \cdots, \theta^{(i)}_d)}{g_i(\theta^{(i)}_1, \theta^{(i)}_2, \cdots, \theta^{(i)}_d)}$$

（ただし，$q \propto$ 真の事後分布，g_i は i 回目の反復のあとの予測された事後分布である）として，各反復時点 i で w_i の系列についてヒストグラムを作り，分布がある定数に退化するかどうかで収束を判定しようというものである。ここで，

$$g_i(\theta) = \int K(\theta^*, \theta) g_{i-1}(\theta^*) d\lambda(\theta^*)$$

であるが（K はマルコフ推移核，3.2.1節参照），積分が困難な場合には m 個の独立なGibbs samplerによる連鎖から，

$$\hat{g}_i(\theta) = \frac{1}{m} \sum_{j=1}^{m} K(\theta^i, \theta)$$

を用いてもよい。また，ひとつの長い連鎖を m 個の部分に分割してこれを g の推定にあててもよい。その場合，K を計算するために条件つき分布の比例定数項まで計算する必要がある。

4.4.7 再生時間によるアプローチ

Mykland, Luke and Yu (1995) は確率過程の再生時間を利用して，観測されたマルコフ連鎖が収束しているかどうかを図を用いて判断しようとした。この方法は，すべてのマルコフ連鎖に応用可能であるが，応用が簡単なのは主としてMetropolis-Hastingsアルゴリズムの独立連鎖である。

確率過程 $\{X_n : n = 0, 1, \cdots\}$ は，時点 T_i において（$T_0 \leq T_1 \leq \cdots$）そ

の将来の過程が過去の過程とは独立であり，かつ過去の過程と同じ分布に従うとき再生的であるといわれる。Mykland, Luke and Yu (1995)はモンテカルロ法によって発生させられるマルコフ連鎖が収束をしているならば，その再生時間は安定しているはずであり，再生時間と次の再生時間の間の長さであるtourの長さが安定していると考えた。そこでT_i/T_nをi/n ($i=1, 2, \cdots, n$) に対してプロットし，それが傾き1の直線上にそっていれば連鎖は収束しており，直線から乖離して不規則な動きを示していれば連鎖は収束していないと判断すればよいとした。このプロットはSRQ (Scaled Regeneration Quantile) プロットと呼ばれる。

ここでは独立連鎖のMetropolis-Hastingsアルゴリズムを例に考える。まず$q(x, y)=f(y)$はルベーグ測度上で定義されているものとし，初期値X_0を$f(y)\min(w(y)/c, 1)$に比例する密度関数をもつ分布から棄却法によって発生させる。次にマルコフ連鎖X_nの再生時間TをみつけるためにSplitting変数S_nという過程を定義する。S_nは独立連鎖の次の候補点X_{n+1}を発生させて棄却されたら$S_n=0$とし，受容されたらS_nを次のような$r(x, y)$を成功確率とするベルヌーイ確率変数とする。

$$r(x,\ y)=\begin{cases} \max\left\{\dfrac{c}{w(x)},\ \dfrac{c}{w(y)}\right\}, & w(x)>c\text{かつ}w(y)>c\text{のとき} \\ \max\left\{\dfrac{w(x)}{c},\ \dfrac{w(y)}{c}\right\}, & w(x)<c\text{かつ}w(y)<c\text{のとき} \\ 1, & \text{それ以外のとき} \end{cases}$$

Mykland, Luke and Yu (1995)は，$S_n=1$となる時点が確率過程の再生時間Tであることを証明した (Theorem 1, 2, 3)。もし$S_n=1$で$\Sigma_{j=1}^{n} S_j=i$ならば，$T_i=n$となる。こうして得られた$T_1, T_2, \cdots, T_k$を用いてSRQプロットを作り，収束の判定を図により行う (初期値は$T_0=0$を満たすようにとるので直線は原点から始まり(1, 1)で終わる)。tourの長

さの期待値は，$\int(\pi(x)/w(x))\,dx\Big/\left[\int \min\{\sqrt{c}/w(x),\ 1/\sqrt{c}\}\,\pi(x)\,dx\right]^2$で与えられる。独立連鎖のなかでも受容／棄却法を用いる場合には$f(x)\propto\min(\pi(x),cI(x))$として上の$r$の代わりに

$$r(x,\ y)=\begin{cases}1 & \pi(x)\leq cI(x)\text{あるいは}\pi(y)\leq cI(y)\text{のとき}\\[2ex] \min\left\{\dfrac{cI(x)}{\pi(x)},\ \dfrac{cI(y)}{\pi(y)}\right\} & \text{それ以外のとき}\end{cases}$$

とすればよい。$q(y,x)=q(x,y)$であるようなMetropolis-HastingsアルゴリズムやGibbs Samplerへの応用の仕方についてもMykland, Luke and Yu (1995) は言及しているが，すべての場合に簡単に適用できるとはいえない。

5 応用

マルコフ連鎖モンテカルロ法は，この数年の間にさまざまな分野に応用され始めている。以下に，そのなかで主なものをあげておく。

- 仮説検定・モデル選択　George and McCulloch (1993), Chipman (1994), Gelfand and Dey (1994), Kuo and Mallick (1994), Lewis and Raftery (1994), Raftery (1994) Raftery, Madigan and Hoeting (1994), Polson and Robert (1994), Carlin and Chib (1995), Chib (1995), Hoeting, Raftery. and Madigan (1995a), Hoeting, Raftery and Madigan (1995b), Verdinelli and Wasserman (1995).
- 時系列分析　Lamoureux and Zhou (1992), McCulloch and Tsay (1992), Chib and Greenburg (1993), Müller and Pole (1993), Jacquier, Polson and Rossi (1994), McCulloch and Tsay (1994), Müller, West and MacEachern (1994), Pai and Ravishanker

(1994), Chan and Ledolter (1995), Koop, Ley, Osiewalski and Steel (1995), West (1995a), West (1995b), West (1995c), Cargnoni, Mueller and West (1995).

- 生存解析 Dellaportas and Smith (1993), Doss (1994), Sinha, Tanner and Hall (1994), Aslanidou, Dey and Sinha (1995), Sargent (1995), Stangle and Greenhouse (1995), Kuo and Peng (1995).
- 一般線形モデル Zeger and Karim (1991), Dellaportas and Smith (1993), Breslow and Clayton (1993), Mallick and Gelfand (1994), Gelfand, Sahu and Carlin (1995), Mukhopadhyay and Gelfand (1995), Thomas (1995), Sahu and Gelfand (1995).
- 最尤法 Geyer and Thompson (1992), Gelfand and Carlin (1993), Geyer (1994).
- 状態空間モデル Carlin, Polson and Stoffer (1992), Carter and Kohn (1994), Früwirth-Schnatter (1994).
- プロビット・トービットモデル McCulloch and Rossi (1992), Cowles, Carlin and Connett (1993), Erkanli, Stangl and Müller (1993), Nobile (1995).
- センサーされたデータ・打ち切られたデータ Gelfand, Smith and Lee (1992), Hamada and Wu (1995).
- ノンパラメトリック・セミパラメトリックモデル Müller (1992a), Müller (1992b), Erkanli, Stangl and Müller (1993), Doss (1994), Doss, Huffer and Lawson (1994), Müller and Roeder (1994), Müller and Rosner (1994), MacEachern and Müller (1994a), MacEachern and Müller (1994b), Newton (1994), Eskobar and West (1995), Newton, Czado and Chappell (1995).
- Longitudinalデータ・パネルデータ Lange, Carlin and Gelfand

(1992), Gilks, Clayton, Spiegelhalter, Best, McNeil, Sharples and Kirby (1993), Gilks, Wang, Yvonnet and Coursaget (1993), Carlin (1994), Carlin and Goldman (1994).

- 正規分布に従うデータなど　Gelfand, Hills, Racine-Poon and Smith (1990), Gelfand, Sahu and Carlin (1995), Robert and Mengersen (1995).
- 二値回帰モデル　Gelfand and Kuo (1991), Albert and Chib (1993), Albert and Chib (1995), Basu and Mukhopadhyay (1994), Newton, Czado and Chappell (1995).
- クラスター分析　Bensmail, Celeux, Raftery and Robert (1995).

参考文献

Albert, J.H. (1995), "Bayesian Selection of Log-Linear Models", Technical Report, Department of Statistics, Bowling Green State University.

Albert, J.H. and Chib, S. (1993), "Bayesian Analysis of Binary and Polychotomous Response Data", *Journal of American Statistical Association*, 88, 669-679.

Albert, J.H. and Chib, S. (1995), "Bayesian Residual Analysis of Binary Response Regression Models", *Biometrika*, 82, 747-759.

Aslanidou, H., Dey, D.K. and Sinha, D. (1995), "Bayesian Analysis of Multivariate Survival Data Using Monte Calro Methods", Technical Report, Department of Statistics, University of Connecticut.

Barker, A.A. (1965), "Monte Calro Calculation of the Radial Distribution Function for a Protonelecton Plasma", *Australian Journal of Physics*, 18, 119-133.

Basu, S. and Mukhopadhyay, S. (1994), "Bayesian Analysis of a Random Link Function in Binary Response Regression", Technical Report, Department of Statistics, University of Connecticut.

Bensmail, H., Celeux, G., Raftery A.E. and Robert C.P. (1995), "Inference in Model-Based Cluster Analysis", Technical Report, Department of Statistics, University of Washington.

Baxter, J.R. and Rosenthal, J.S. (1994), "Rates of Convergence for Everywehere -positive Markov Chains", Technical Report 9406, University of Toronto, Department of Statistics.

Berger, J. and Oh, M.S. (1992), "Adaptive Importance Sampling in Monte Calro Integration", *Journal of Statistical Computing and Simulation*, 41, 143-168.

Besag, J., Green, P., Higdon, D. and Mengersen, K. (1995), "Bayesian Computation and Stochastic Systems (with comments)", *Statistical Science*, 10, 3-66.

Best, N.G., Cowles, M.K. and Vines, K. (1995), "CODA: Convergence Diagnosis and Output Analysis Software for Gibbs Sampling Output, Version 0.30.", Technical Report, MRC Biostatistics Unit, University of Cambridge.

Bishop, Y.M.M., Fienberg, S.E. and Holland, P.W. (1975), "Discrete Multivariate Analysis". Cambridge, Mass: MIT Press.

Breslow, N.E. and Clayton, D.G. (1993), "Approximate Inference in Generalized Linear Mixed Models", *Journal of American Statistical Association*, 88, 9-25.

Brooks, S.P. and Morgan, J.T. (1995), "Optimization Using Simulated Annealing", *The Statistician*, 44, 241-257.

Carlin, B.P. (1994), "Hierarchical Longitudinal Modeling with Application to HIV Progression", Unpublished manuscprit, University of Minnesota.

Carlin, B.P. and Chib, S. (1995), "Bayesian Model Choice via Markov Chain Monte Carlo Methods", *Journal of Royal Statistical Society*, *Ser. B*, 57, 473-484.

Carlin, B.P., Gelfand, A.E. and Smith A.F. (1992), "Hierarchical Bayesian Analysis of Chaingepoint Problems", *Applied Statistics*, 2, 389-405.

Carlin, B.P. and Goldman, A.I. (1994), "Bayesian Evaluation of CD4 Count as a Surrogate Endpoint in Patients with Advanced HIV Infection", Unpublished manuscript, University of Minnesota.

Carlin, B.P., Polson, N.G. and Stoffer, D.S. (1992), "A Monte Carlo Approach to Nonnormal and Nonlinear State-Space Modeling", *Journal of American Statistical Association*, 87, 493-500.

Cargnoni, C., Mueller, P. and West, M. (1995), "Bayesian Forecasting Multinomial Time Series through Conditionally Gaussian Dynamic Models", Working Paper, Institute of Statistics and Decision Sciences, Duke University.

Carter, C.K. and Kohn, R. (1994), "On Gibbs Sampling for State Space Models", *Biometrika*, 81, 541-543.

Casella, G. and George, E.I. (1992), "Explaining the Gibbs Sampler", *The American Statistician*, 46, 167-174.

Chan, K.S. (1993), "Asymptotic Behavior of the Gibbs Sampler", *Journal of American Statistical Association*, 88, 320-326.

Chan, K.S. and Geyer, C.J. (1994), Comment on "Markov Chains for Exploring Posterior Distributions (with discussion)", by Tierney, L., *Annals of Statistics*, 22, 1747-1758.

Chan, K.S. and Ledolter, J. (1995), "Monte Carlo EM Estimation for Time Series Models Involving Counts", *Journal of American Statistical Association*, 90, 242-252.

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

Chib, S. and Greenburg, E. (1994), "Bayes Inference in Regression Models with ARMA (p, q) Errors", *Journal of Econometrics*, 64, 183-206.

Chib, S. and Greenburg, E. (1995), "Understanding the Metropolis-Hastings Algorithm", *The American Statistician*, 49, 327-335.

Chib, S. and Greenburg, E. (1995), "Markov Chain Monte Calro Simulation Methods in Econometrics", Working Paper, Washington University.

Chimpan, H. (1994), "Bayesian Variable Selection with Related Predictors", Technical Report, Department of Statistics and Actuarial Science, University of Waterloo.

Cowles, M.K. and Carlin, B.P. (1994), "Markov Chain Monte Carlo Diagnostics: A Comparative Review", Technical Report, University of Minnesota, School of Public Health, Division of Biostatistics.

Cowles, M.K., Carlin, B.P. and Connett, J.E. (1993), "Bayesian Tobit Modeling of Longitudinal Ordinal Clinical Trial Compliance Data with Nonignorable Missingness", Technical Report, University of Minnesota, School of Public Health, Division of Biostatistics.

Dellaportas, P. and Smith, A.F.M. (1993), "Bayesian Inference for Generalized Linear and Proportional Hazards Models via Gibbs Sampling", *Applied Statistics*, 42, 443-459.

Dempster, A.P., Laird, N. and Rubin, D.B. (1977), "Maximum Likelihood from Incomplete Data via EM Algorithm", *Journal of Royal Statistical Society*, *Ser. B*, 39, 1-38.

Diebolt, J. and Robert, C.P. (1994), "Estimation of Finite Mixture Distributions through Bayesian Sampling", *Journal of Royal Statistical Society, Ser. B*, 56, 363-375.

Doss, H. (1994), "Bayesian Nonparametric Estimation for Incomplete Data Via Successive Substitution Sampling", *Annals of Statistics*, 22, 1763-1786.

Doss, H., Huffer, F.W. and Lawson, K.L. (1994), "Bayesian Nonparametric Estimation Via Gibbs Sampling for Coherent Systems with Redundancy", Unpublished manuscript, Ohio State University.

Erkanli, A.E., Stangl, D. and Müller, P. (1993), "A Bayesian Analysis of Ordinal Data Using Mixtures", Unpublished manusrcipt, Institute of Statistics and Decision Sciences, Duke University.

Eskobar, M.D. and West, M. (1995), "Bayesian Density Estimation and Inference Using Mixtures", *Journal of American Statistical Association*, 90, 577-588.

Frühwirth-Schnatter (1994), "Data Augmentation and Dynamic Linear Models", *Journal of Time Series Analysis*, 15, 183-202.

Gelfand, A.E. (1994), "Gibbs Sampling (A Contibution to the Encyclopedia of Statistical Sciences)", Technical Report, Department of Statistics, University of Connecticut.

Gelfand, A.E. and Carlin, B.P. (1993), "Maximum-likelihood Estimation for Constrained or Missing-data models", *Canadian Journal of Statistics*, 21, 303-311.

Gelfand, A.E. and Kuo, L. (1991), "Nonparametric Bayesian Bioassay Including Ordered Polytomous Response", *Biometrika*, 78, 657-666.

Gelfand, A.E., Hills, S.E., Racien-Poon, A. and Smith, A.F.M. (1990), "Sampling Based Approaches to Calculatng Marginal Densities", *Journal of American Statistical Association*, 85, 972-985.

Gelfand, A.E., Sahu, S.K. and Carlin, B.P. (1995), "Efficient Parameterizations for Normal Linear Mixed Models", *Biometrika*, 82, 479-488.

Gelfand, A.E., Sahu, S.K. and Carlin, B.P. (1995), "Efficient Parameterizations for Generalized Linear Mixed Models", In *Bayesian Statistics 5*, eds. J.M. Bernardo, J.O. Berger, AP. Dawid and A.F.M. Smith, Oxford, Oxford University Press.

Gelfand, A.E. and Sfridis, J.M. (1994), "Bayesian Analysis of Financial Event Studies

Data", Technical Report, Department of Statistics, University of Connecticut.

Gelfand, A.E. and Smith, A.F.M. (1990), "Sampling Based Approaches to Calculating Marginal Densities", *Journal of American Statistical Association*, 85, 398-409.

Gelfand, A.E. and Smith, A.F.M. (1991), "Gibbs Sampling for Marginal Posterior Expectations", *Journal of American Statistical Association*, 85, 398-409.

Gelfand, A.E. and Smith, A.F.M. and Lee, T. (1992), "Bayesian Analysis of Constrained Parameter and Truncated Data Problems Using Gibbes Sampling", *Journal of American Statistical Association*, 87, 523-532.

Gelman, A., Roberts, G.O. and Gilks, W.R. (1994), "Effecient Metropolis Jumping Rules", Research Rerport, Statistical Laboratory, University of Cambridge.

Gelman, A. and Rubin D.B. (1992), "Inference from Iterative Simulation Using Multiple Sequences (with discussion)", *Statistical Science*, 7, 457-511.

Geman, S. and Geman, D. (1984), "Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6, 721-741.

George, E.I. and McCulloch, R.E. (1993), "Variable Selection Via Gibbs Sampling", *Journal of American Statistical Association*, 88, 881-889.

Geweke, J. (1992), "Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments," In *Bayesian Statistics 4* eds. J.M. Bernardo, J. Berger, A.P. Dawid and A.F.M. Smith, Oxford, Oxford University Press, 169-193.

Geyer, C.J. (1994), "On the Convergence of Monte Carlo Maximum Likelihood Calculations", *Journal of Royal Statistical Society*, *Ser. B*, 56, 261-274.

Geyer, C.J. and Thompson, E.A. (1992), "Constrained Monte Carlo Maximum Likelihood for Dependent Data", *Journal of Royal Statistical Society*, *Ser, B*, 54, 657-699.

Geyer, C.J. and Thompson, E.A. (1995), "Annealing Markov Chain Monte Carlo with Applications to Ancestral Inference", *Journal of American Statistical Association*, 90, 909-920.

Gilks, W.R., Clayton, D.G., Spiegelhalter, D.J., Best, N.G., McNeil, A.J., Sharples, L.D. and Kirby, A.J. (1993), "Modelling Conplexity: Application of Gibbs Sampling in Medicine", *Journal of Royal Statistical Society*, *Ser. B*, 55, 39,52.

Gilks, W.R., Wang, C.C., Yvonnet, B. and Coursaget, P. (1993), "Random Effects Models for Longitudinal Data Using Gibbs Sampling", *Biometrics*, 49, 441-453.

Hamada, M. and Wu, C.F.J. (1995), "Analysis of Censored Data From Franctionated Experiments: A Bayesian Approach", *Journal of American Statistical Association*, 85, 398-409.

Hastings, W.K. (1970), "Monte Carlo Sampling Methods Using Markov Chains and their Applications", *Biometrika*, 57, 97-109.

Heidelberger, P. and Welch, P.D. (1983), "Simulation Run Length Control in the Presence of an Initial Transient", *Operations Research*, 31, 1109-1144.

Hoeting, J.A., Raftery, A.E. and Madigan, D. (1995a), "A Method for Simultaneous Variable Selection and Outlier Identification in Linear Regression Models", Unpublished manuscript, Department of Statistics, Corolado State University.

Hoeting, J.A., Raftery, A.E. and Madigan, D. (1995b), "Simultaneous Variable and Transformation Selection in Linear Regression Models", Unpublished manuscript, Department of Statistics, Corolado State University.

Ingber, L. (1993), "Simulated Annealing: Practice versus Theory", *Mathematical Computational Modelling*, 18, 29-57.

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

Kirkpatrick, S., Gelatt, C.D. and Vecchi, M.P. (1983), "Optimization by Simulated Annealing", *Science*, 220, 671-680.

Koop, G., Ley, E., Osiewalski., J. and Steel, M.F.J. (1995), "Bayesian Analysis of Long Memory and Persistence Using ARFIMA Models". To appear in *Journal of Econometrics*.

Kuo, L. and Mallick, B. (1994), "Variable Selection for Regression Models", Technical report, University of Connecticut.

Kuo, L. and Peng, F. (1995), "A Mixture Model Approach to the Analysis of Survival Data", Technical report, University of Connecticut.

Lamoureux, C.G. and Zhou, G. (1992), "Exact Densities for the Decomposition of Time Series into Permanent and Transitory Components Using the Gibbs Sampler", Unpublished manuscript, Washington University.

Lange N., Carlin, B.P. and Gelfand, A.E. (1992), "Hierachical Bayes Models for the Progression of HIV Infection Using Longitudinal CD4 T-Cell Numbers", *Journal of American Statistical Association*, 87, 615-632.

Lewis, S. and Raftery, A.E. (1994), "Estimating Bayes Factors Via Posterior Simulation with the Laplace-Metropolis Estimator", Unpublished manuscript, Department of Statistics, University of Washington.

Liu, C. and Liu, J.S. (1993), "Comment on Markov Chain Monte Carlo", *Journal of Royal Statistical Society*, *Ser. B*, 55, 82-83.

Liu, J.S., Wong, W.H. and Kong, A. (1994), "Covariance Structure of the Gibbs Sampler with Applications to the Comparison of Estimators and Augmentation Schemes", *Biometrika*, 81, 27-40.

Liu, J.S., Wong, W.H. and Kong, A. (1995), "Covariance Structure and Convergence Rate of the Gibbes Sampler with Various Scans", *Journal of Royal Statistical Society*, *Ser. B*, 57, 157-169.

MacEachern, S.N. and Müller, P. (1994a), "Estimating Mixture of Dirichlet Process Models", Unpublished manuscript, Institute of Statistics and Decision Sciences, Duke University.

MacEachern, S.N. and Müller, P. (1994b), "Efficient Estimation of Mixture of Dirichlet Process Models", Unpublished manusrcipt, Institute of Statistics and Decision Sciences, Duke University.

Mallick, B.K. and Gelfand, A.E. (1994), "Generalized Linear Models with Unknown Link Functions", *Biometrika*, 81, 237-245.

Mengersen, K.L., and Tweedie, R.L. (1994), "Rates of Convergence of the Hastings and Metropolis Algorithms", Unpublished manuscript, Corolado State University.

Mengersen, K. and Robert, C. (1995), "Testing for Mixtures: A Bayesian Entropic Approach", In *Bayesian Statistics 5*, eds. J.M. Bernardo, J.O. Berger, AP. Dawid and A.F.M. Smith, Oxford, Oxford University Press.

Metropolis, N., Rosenblush, A.W., Rosenblush, M.N., Teller, A.H. and Teller, E. (1953), "Equations of State Calculations by fast Computing machines", *Journal of Chemical Physics*, 21, 1087-1091.

Meyn, S.P. and Tweedie, R.L. (1994), "Computable Bounds for Convergence Rates of Markov Chains", *Annals of Applied Probability*, 4, 981-1011.

McCulloch, P. and Rossi, P.E. (1992), "An Exact Likelihood Analysis of the Multinomial Probit Model", Technical Report, Graduate School of Business, University of Chicago. To appear in *Journal of Econometrics*.

McCulloch, P. and Tsay, E.S. (1992), "Bayesian Inference of Trend- and Difference -Stationarity", Unpublished Manuscript, Graduate School of Business, University of Chicago.

McCulloch, P. and Tsay, E.S. (1994), "Bayesian Analysis of Autoregressive Time Series Via the Gibbs Sampler", *Journal of Time Series Analysis*, 15, 235-250.

Mukhopadhyay, S. and Gelfand, A.E. (1995), "Dirichlet Process Mixed generalized Linear Models", Technical Report, Department of Statistics, University of Connecticut.

Müller, P. (1992a), "Posterior Integration in Dynamic Models", Unpublished manusrcipt, Institute of Statistics and Decision Sciences, Duke University.

Müller, P. (1992b), "Alternatives to the Gibbs Sampling Scheme", Unpublished manusrcipt, Institute of Statistics and Decision Sciences, Duke University.

Müller, P., West, M. and MacEachern, S. (1994), "Bayesian Models for Non-linear Autoregressions", Unpublished manusrcipt, Institute of Statistics and Decision Sciences, Duke University.

Müller, P. and Pole, A. (1993), "Monte Carlo Posterior Integration in GARCH Models", Unpublished manusrcipt, Institute of Statistics and Decision Sciences, Duke University.

Müller, P. and Roeder, K. (1994), "A Bayesian Semiparametric Model for Case -Control Studies with Errors in Variables", Unpublished manusrcipt, Institute of Statistics and Decision Sciences, Duke University.

Müller, P. and Rosner, G. (1994), "A Semiparametric Bayesian Population Model with Hierarchical Mixture Priors", Unpublished manusrcipt, Institute of Statistics and Decision Sciences, Duke University.

Mykland, P., Tierney, L. and Yu, B. (1995), "Regeneration in Markov Chain Samplers", *Journal of American Statistical Association*, 90, 233-241.

Newton, M.A. (1994), "Computing with Priors That Support Identifiable Semiparametric Models", Technical Report, Department of Statistics, University of Wisconsin-Madison.

Newton, M.A., Czado, C. and Chappell, R. (1995), "Bayesian Inference for Semiparametric Binary Regression", Technical Report, Department of Statistics, University of Wisconsin-Madison.

Nobile, A. (1995), "A Hybrid Markov Chain for the Bayesian Analysis of the Multinomial Probit Model", Working Paper, Institute of Statistics and Decision Sciences, Duke University.

Nummelin, E. (1984), "General Irreducible Markov Chains and Nonnegative Operators", Camridge, U.K.: Cambridge University Press.

Pai, J.S. and Ravishanker, N. (1994), "Bayesian Analysis of Long Memory in U.S. GNP Using ARFIMA Models", Unpublished manuscript, Department of Statistics, University of Connecticut.

Peskun, P.H. (1973), "Optimum Monte-Carlo Sampling Using Markov Chains", *Biometrika*, 60, 607-612.

Polson, N.G. and Roberts, G.O. (1994), "Bayes Factors for Discrete Observations from Diffusion Processes", *Biometrika*, 81, 11-26.

Raftery, A.E. (1986), "A Note on Bayes Factors for Log-linear Contingency Tables with Vague Prior Information", *Journal of Royal Statistical Society, Ser. B*, 48, 249-250.

Raftery, A.E. (1994), "Hypothesis Testing and Model Selection Via Posterior Simulation", Unpublished manuscript, Department of Statittics, University of Washington.

Raftery, A.E. and Lewis, S. (1992), "How Many Iterations in the Gibbs Sampler?", In *Bayesian Statistics 4*, eds. J.M. Bernardo, J.O. Berger, AP. Dawid and A.F.M. Smith, Oxford, Oxford University Press, 763-773.

Raftery, A.E. and Lewis, S. (1994), "The Number of Iterations, Convergence Diagnostics and Generic Metropolis Algorithms", Unpublished manuscript, Department of Statistics, University of Washington.

Raftery, A.E., Madigan, D. and Hoeting, J.A. (1994), "Bayesian Model Averaging for Linear Regression Models", Unpublished manuscript, Department of Statistics, University of Washington.

Ripley, B. (1987), *Stochastic Simulation*, New York: Wiley.

Ritter, C. and Tanner, M.A. (1992), "Facilitating the Gibbs Sampler: The Gibbs Stopper and the Griddy Gibbs Sampler", *Journal of American Statistical Association*, 87, 861-868.

Robert, C.P. and Mengersen, K.L. (1995), "Reparameterization Issues in Mixture Modelling and Their Bearing on the Gibbs Sampler", Unpublished Manuscript, Colorado State University.

Roberts, G.O., Gelman, A. and Gilks, W.R. (1994), "Weak Convergence and Optimal Scaling of Random Walk Metropolis Algorithms", Research Rerport, Statistical Laboratory, University of Cambridge.

Roberts, G.O., and Polson, N.G. (1994), "On the Geometric Convergence of the Gibbs Sampler", *Journal of Royal Statistical Society*, *Ser. B*, 56, 377-384.

Roberts, G.O., and Tweedie, R.L. (1994), "Geometric Convergence and Central Limit Theorems for Multidimensional Hastings and Metropolis Algorithms", Unpublished manuscript, Corolado State University.

Roberts, G.O. and Rosenthal, J.S. (1995), "Optimal Scaling of Discrete Approximation to Langevin Diffusions", Research Report, Statisitcal Laboratory, University of Cambridge.

Roberts, G.O. and Smith, A.F.M. (1994), "Simple Conditions for the Convergence of the Gibbs Sampler and Metropolis-Hastings Algorithms", *Stochastic Processes and Their Applications*, 49, 207-216.

Rosenthal, J.S. (1993), "Rates fo Convergence for Data Augmentation on Finite Sample Spaces", *Annals of Applied Probability*, 3, 319-339.

Rosenthal, J.S. (1994), "Markov Chain Convergence: From Finite to Infinite", Technical Report, University of Toronto, Department of Statistics.

Rosenthal, J.S. (1995a), "Minorization Conditions and Convergence Rates for Markov Chain Monte Calro", *Journal of American Statistical Association*, 90, 558-566.

Rosenthal, J.S. (1995b), "Rates of Convergence for Gibbs Sampling for Variance Component Models", *Annals of Statistics*, 23, 740-761.

Rosenthal, J.S. (1995c), "Analysis of the Gibbs Sampler for a Model Related to James-Stein Estimators", to appear in *Statistics and Computing*.

Rubin, D.B. (1987a), Comment on "The Calculation of Posterior Distributions by Data Augmentation (with discussion)", by Tanner, M.A. and Wong, W.H., *Journal of American Statistical Association*, 82, 543-546.

Rubin, D.B. (1987b), *Multiple Imputation for Non-response in Surveys*, New York: Wiley.

Rubin, D.B. (1988), "Using the SIR Algorithm to Simulate Posterior Distributions", in *Bayesian Statistics 3*, eds. J.M. Bernardo, M.H. DeGroot, D.V. Lindley, and A.

F.M. Smith, Oxford, U.K.: Oxford University Press, 395-402.

Sahu, S.K. and Gelfand, A.E. (1995), "On Propriety of Posteriors and Bayesian Identifiability in Generalized Linear Models", Technical Report, University of Cambridge.

Sargent, D.J. (1995), "A General Framework for Random Effects Survival Analysis in the Cox Proportional Hazard Setting", Technical Report, Division of Biostatistics, School of Public Health, University of Minnesota.

Schervish, M.J. and Carlin, B.P. (1992), "On the Convergence of Successive Substitution Sampling", *Journal of Computational and Graphical Statistics*, 1, 111-127.

Sinha, D., Tanner, M.A. and Hall, W.J. (1994), "Maximization of the Marginal Likelihood of Grouped Survival Data", *Biometrika*, 81, 53-60.

Smith, A.F.M. and Gelfand, A.E. (1992), "Bayesian Statistics Without Tears", *American Statistican*, 46, 84-88.

Smith, A.F.M. and Roberts, G.O. (1993), "Bayesian Computation via the Gibbs Sampler and Related Markov Chain Monte Calro Methods", *Journal of Royal Statistical Society*, *Ser. B*, 55, 3-23.

Stangle, D.K. and Greenhouse, J.B. (1995), "Assessing Placebo Response Using Bayesian Hierachical Survival Models", Technical Report, Department of Statistics, Duke University.

Tanner, M.A. (1993), *Tools for Statistical Inference*, 2nd ed. Springer-Verlag.

Tanner, M.A. and Wong, W.H. (1987), "The Calculation of Posterior Distributions by Data Augmentation (with discussion)", *Journal of American Statistical Association*, 82, 528-550.

Thomas, W. (1995), "An Exploratory Diagnostic for Errors-in-varaibles Modeling in Generalized Linear Models", Technical Report, Division of Biotatistics, School of Public Health, University of Minnesota.

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

Verdinelli, I. and Wasserman, L. (1995), "Computing Bayes factors Using a Generalization of the Savage-Dickey Density Ratio", *Journal of American Statistical Association*, 90, 614-618.

Wei, G.C.G. and Tanner, M.A. (1990), "A Monte Carlo Implementation of the EM

Algorithm and the Poor Man's Data Augmentation Algorithm", *Journal of American Statistical Association*, 85, 699-704.

West, M. (1995a), "Time Series Decomposition and Analysis in a Study of Oxygen Isotope Records", Working Paper, Institute of Statistics and Decision Sciences, Duke University.

West, M. (1995b), "Bayesian Time Series", Working Paper, Institute of Statistics and Decision Sciences, Duke University.

West, M. (1995c), "Bayesian Inference in Cyclical Component Dynamic Linear Models", *Journal of American Statistical Association*, 90, 1301-1312.

Zeger, S.L. and Karim, M.R. (1991), "Generalized Linear Models with Random Effects: Gibb's Sampling Approach", *Journal of American Statistical Association*, 86, 79-86.

Zellner, A., and Min, C. (1995), "Gibbs Sampler Convergence Criteria", *Journal of American Statistical Association*, 90, 921-927.