オイコノミカ　第 54 巻　第 1 号，2017 年，pp. 3–13

Normal mixture のファイナンスへの応用

程　島　次　郎

要　旨

Normal mixture の分布のクラスは，非対称で裾が厚い分布を作ることができる．ファイナンスおよび金融のデータは，非対称で裾が厚い分布がその特徴であるといえる．そのため，normal mixture はファイナンスや金融のデータでよく使われてきたが，日本国内のファイナンス研究者には十分に周知されていないように思われる．そこで本稿では，normal mixture をファイナンスの実証に利用する場合に必要なことをサーベイする．具体的には，これまでの normal mixture を用いた実証研究の事例を紹介し，推定やモデルの検証方法，GARCH モデルを想定した場合の推定や検証方法などについて述べる．

1．はじめに

Normal mixture は，統計学では正常の観測値に異常値が含まれる場合のデータを表わす contaminated normal と同じ分布のクラスを作る．すなわち，確率 $1-\varepsilon$ で正規分布 $\mathrm{N}(\mu_1, \sigma_1^2)$（期待値 μ_1 分散 σ_1^2 の正規分布）が起き確率 ε で異なる正規分布 $\mathrm{N}(\mu_2, \sigma_2^2)$ が起きる場合が contaminated normal といわれる分布のもっとも簡単な事例である．確率 ε で起きる正規分布が異常値を表わす場合は，σ_2^2 を σ_1^2 とは異なる大きな値をもつか，μ_2 を μ_1 とかなり離れた値にする．このような contaminated normal と呼ばれる分布は，1950 年代や 1960 年代に盛んに研究された統計学の一分野であるロバスト統計学でよく使われる例である．

Contaminated normal を一般化すると，以下の K 個の正規分布に従う 1 次元の確率変数の normal mixture の cdf（累積確率分布関数）F(x) は，

$$\mathrm{F}(\mathrm{x})=\textstyle\sum_{i=1}^{K} p_i \Phi\left(\frac{x-\mu_i}{\sigma_i}\right) \tag{1}$$

ここで，Φ は N(0, 1) の cdf である．したがって，その pdf（確率密度関数）は

$$f(\mathrm{x})=\sum_{i=1}^{K}p_i\varphi(x;\ \mu_i,\sigma_i^2) \tag{2}$$

である．ただし，

$$\varphi(x;\ \mu_i,\sigma_i^2)=\frac{1}{\sqrt{2\pi\sigma_i^2}}e^{-\frac{(x-\mu_1)^2}{2\sigma_i^2}}$$

$$\sum_{i=1}^{K}p_i=1,\ 0\leq p_i\leq 1$$

である．(1) または (2) で定義される normal mixture は，K 個の離散的な normal mixture になり，非正規性や非対称性を持つ分布を作ることのできるフレキシブルな分布のクラスである．ファイナンスや金融のデータでは，裾の厚い分布や非対称な分布を示すことが多いので，normal mixture を用いたファイナンスや金融の実証研究がよく行われている．また，Hamilton (1989) や Hamilton and Susmel (1994) などの regime-switching モデルとも同様な特性を持っている．しかし，我が国のファイナンスや金融の実証研究では，normal mixture を用いた研究が数少ないように思われる．その原因は，我が国のファイナンスや金融の実証研究の層が薄いことやこれらの分野の研究者と統計学や計量経済学の研究者との交流が少ないことにあるように思われる．本稿では，このような現状を少しでも改善する 1 つの機会を提供できれば幸いである．

2. これまでの実証研究

Kon (1984) は，アメリカの株価収益率データに normal mixture を当てはめ，normal mixture の異なる成分での期待値と分散の違いが明らかになり，また t 分布よりもデータの説明力が高いことを示した．Zangari (1996) は，2 つの成分を用いた normal mixture を金融データに当てはめた．Hull and White (1998) は，同じく 2 つの成分の normal mixture を使って金融データに当てはめることを示唆した．また，normal mixture を用いて value at risk (VaR) を計算する問題に対しては, Alexander and Sheedy (2008) や Kamaruzzaman and Isa (2013) が行っている．Kamaruzzaman and Isa (2013) は，マレーシアの株式インデックスの月次および週次収益率に 2 つの成分の normal mixture を当てはめ，低いボラティリティーで通常の市場環境の安定レジームは高い確率で起き，高いボラティリティーでマイナスの期待値を持つ暴落している市場環境の暴落レジームは低い確率で起きていることを示した．たとえば，月次のマレーシア株式インデックスは，期待値が -0.2 で分散が 11.50^2 の暴落レジームが確率 0.22 で起き，期待値が 0.47 で分散が 4.75^2 の安定レジームが確率 0.78 で起きた．このように，2 つの成分の normal mixture は市場の特徴をほかの分布を想定した場合よりもうまく記述することができることを示している．また，Kamaruzzaman and Isa (2013) は，VaR と expected shortfall を 2 つの成分の normal mixture で計算し，それぞれデータと整合的か否かを検証す

るバックテストを行っていて，データと整合的という結果を得ている．

他方，金融データは，非対称性や非正規性以外にも，時間ととも変化するボラティリティやいわゆるボラティリティークラスタリングという現象があることが知られている．このような現象には，normal mixture は対応できていない．しかし，このような現象に対応するモデルである Engle（1982）によって提案された Autoregressive Conditional Heteroscedasticity（ARCH）モデルや Bollerslev（1986）によって提案された Generalized ARCH（GARCH）モデル，あるいは Taylor（1986）によって有名になった Stochastic Volatility（SV）モデルがある．そして，非対称あるいは非正規な分布に対応する normal mixture と時間とともに変化するボラティリティやボラティリティークラスタリングの現象を説明する ARCH/GARCH モデルあるいは SV モデルが併存するモデルが考えられた．Hsieh (1989), Vlaar and Palm (1993), Wong and Li（2001），Haas et al.（2004a），Alexander and Lazar（2006），Xu and Wirjanto（2010），Kamaruzzaman and Isa（2014）などは，normal mixture の下で GARCH モデルを導入した．一方，Mahieu and Schotman（1998），Kim et al.（1998），Xu and Knight（2009），Omori et al.（2007）などは，normal mixture の下で SV モデルを考察した．

Normal mixture GARCH モデルは，各成分で条件付き分散を想定し成分の条件付き分散の間のフィードバックも可能とし，歪度や尖度が時間とともに変化できるためデータにより当てはまり，また正規分布や t 分布を仮定した GARCH モデルよりもデータを説明できる．たとえば，Kamaruzzaman and Isa（2014）では，マレーシア株式インデックスの収益率に 2 成分の normal mixture で GARCH(1,1) モデルを当てはめ，以下の結果を得た．

表 1

レジーム固有の行動	マレーシア株式
安定レジーム	
確率	0.8240
期待値	0.4027
無条件分散	1.9156
α_1	0.0653
β_1	0.8926
暴落レジーム	
確率	0.1760
期待値	−1.8851
無条件分散	3.5383
α_2	0.2816
β_2	0.7969

ただし，確率とは（1）式での p_i の推定値を表わし，期待値は（1）式での μ_i の推定値をしめし，無条件分散は各レジームでの無条件分散をあらわし，α_i と β_i は，各成分の条件付き分散 σ_{it}^2 に GARCH(1,1) を仮定した式

$$\sigma_{it}^2 = \omega_i + \alpha_i \varepsilon_{t-1}^2 + \beta_i \sigma_{it-1}^2 \tag{3}$$

での α_i と β_i の推定値を表わす．表1の結果は，2つの成分の normal mixture を仮定しているが GARCH モデルは想定していない場合の Kamaruzzaman and Isa（2013）の結果とも整合的である．すなわち，安定レジームはより高い確率で起き，期待値は正で分散が低い，一方暴落レジームでは期待値は大きくマイナスになり分散は大きくなる．さらに，GARCH(1,1) のパラメータには，暴落レジームでは安定レジームと比べてマーケットのショックにより強く反応し（ α_2 が α_1 よりも大きな値をとり），他方ボラティリティーの持続性は暴落レジームのほうが少し小さい（ β_2 が β_1 よりも小さい値をとる）のでマーケットのショックは早く消える．

一方，Alexander and Lazar（2009）は，Kamaruzzaman and Isa（2014）と同様に2成分の normal mixture で GARCH(1,1) モデルを4つのヨーロッパの国の株式インデックスに当てはめているが，GARCH(1,1) プロセスに収益率がマイナスの場合にプラスの場合と異なるレバレッジ効果を導入している．彼らは，各成分の期待値が異なることとレバレッジ効果という2つの非対称性をモデルに導入している．その結果，1つの正規分布，1つの対称な t 分布，1つの非対称な t 分布，での GARCH(1,1) や，レバレッジ効果がない場合とある場合の GARCH(1,1) モデルなどの中で，レバレッジ効果を考慮した各成分の期待値が異なる非対称な2成分の normal mixture モデルでの GARCH(1,1) が，コールオプションから計算されるインプライドボラティリティでのボラティリティスキューを含めたモデルのあてはまりの良さの検証で最も良い結果が得られた．

このように，2つの成分を用いた normal mixture での GARCH(1,1) モデルは，株価や為替レートなどの金融データをうまく記述できている．このモデルは，Hamilton and Susmel（1994）で初めて導入され，Cai（1994），Gray（1996），Klaassen（2002），Haas et al.（2004b）などで修正された Markov switching GARCH モデルよりも推定が容易である．これらの Markov switching GARCH モデルでは，各成分の分散がお互いに依存し合っている点がモデルの推定を困難にしている．

一方，normal mixture での GARCH(1,1) モデルは，各成分の分散は，分析対象の収益率から計算される誤差項に依存しているために各成分の分散がお互いに依存し合っているため，推定が容易になる．すなわち，(3) 式の右辺で遅れのある条件付き分散は σ_{it-1}^2 だけであり σ_{jt-1}^2 $(j \neq i)$ が含まれておらず，ε_{t-1}^2 を通してのみ各成分の分散が関連している．

3. 推定方法

1次元の2成分の normal mixture の推定は，Hastie et al.（2001）で EM アルゴリズムを用いて推定する方法を示している．すなわち，

Expectation（E）ステップ：

$$r_t = \frac{p\varphi(x_t; \mu_2, \sigma_2^2)}{(1-p)\varphi(x_t; \mu_1, \sigma_1^2) + p\varphi(x_t; \mu_2, \sigma_2^2)} \qquad \mathrm{t}=1, \cdots, \mathrm{T}$$

Maximization（M）ステップ：

$$\mu_1 = \frac{\sum_{t=1}^{T}(1-r_t)x_t}{\sum_{t=1}^{T}(1-r_t)}, \ \sigma_1^2 = \frac{\sum_{t=1}^{T}(1-r_t)(x_t-\mu_1)^2}{\sum_{t=1}^{T}(1-r_t)}$$

$$\mu_2 = \frac{\sum_{t=1}^{T} r_t x_t}{\sum_{t=1}^{T} r_t}, \ \sigma_2^2 = \frac{\sum_{t=1}^{T} r_t (x_t-\mu_2)^2}{\sum_{t=1}^{T} r_t}$$

$$\mathrm{p} = \sum_{t=1}^{T} \frac{r_t}{T}.$$

また，成分や次元が一般的な場合の多次元の normal mixture に対する推測に関して，Chen and Tan（2009）が penalized likelihood method を提案している．彼らは，EM アルゴリズムを使った方法についても言及している．

Normal mixture の推定は，もっとも標準的な推定方法である最尤推定量（maximum likelihood estimator）は，問題があることが知られている．すなわち，normal mixture の尤度関数は有限でなく無限大になるため，最尤推定値が得られないことが起きる．そのため，Quandt and Ramsey（1978）は，MGF（moment generating function（積率母関数））に基づく方法で normal mixture のパラメータを推定することを提案した．1 次元の 2 成分の normal mixture の場合の MGF は，

$$\mathrm{g}(\mathrm{t}, \theta) = \mathrm{E}[e^{tX}] = (1-\mathrm{p})\exp(\mu_1 \mathrm{t} + \frac{1}{2}\sigma_1^2 t^2) + \mathrm{p}\exp(\mu_2 \mathrm{t} + \frac{1}{2}\sigma_2^2 t^2) \tag{4}$$

で与えられる．ただし，$\theta = (\mathrm{p}, \mu_1, \mu_2, \sigma_1^2, \sigma_2^2)$である．(4) の MGF に対応する
標本でのそれは，

$$g_n(\mathrm{t}, \mathrm{x}) = \frac{1}{n}\sum_{i=1}^{n}\exp(tx_i) \tag{5}$$

であり、大数の法則によって、$g_n(\mathrm{t}, \mathrm{x})$ は $\mathrm{g}(\mathrm{t}, \theta)$ に確率収束する．（4）と（5）の t についての m 個の所与のグリッドポイント $(t_1, \cdots, t_m)$ に対して，Quandt and Ramsey（1978）は以下の値

$$h_1(\mathrm{t}, \theta) = \sum_{j=1}^{m}(g_n(t_j, \mathrm{x}) - \mathrm{g}(t_j, \theta))^2 \tag{6}$$

を最小にすることを提案した．Quandt and Ramsey（1978）は，MGF 推定量の以下の漸近分布を示した．すなわち，$\sqrt{n}(\hat{\theta}_{MGF} - \theta)$ の漸近分布が $\mathrm{N}(0, \Psi_1)$ になる．ただし，$\hat{\theta}_{MGF}$ は，(6) を最小にする推定量で，$\Psi_1 = (A'A)^{-1}A'\Omega A(A'A)^{-1}$ である．ただし，A は mx5 の行列で(i, j)要素が $A_{ij} = \frac{\partial g(t_i, \theta)}{\partial \theta_j}$, i＝1, ⋯, m, j＝1, 2, ⋯, 5, Ω も mxmの行列で(i, j)要素が$\Omega_{ij} = \mathrm{g}(t_i + t_j, \theta) - \mathrm{g}(t_i, \theta)\mathrm{g}(t_j, \theta)$, i, j＝1, ⋯, mである．(6) を最小にするのは非線形最小 2 乗法と見なせるので，Schmidt（1982）は MGF 推定量よりも効率的な推定量である修正 MGF 推定量（MMGF）

を提案した．すなわち，MMGF 推定量は，

$$h_2(\mathrm{t},\ \theta)=\sum_{i=1}^{m}\sum_{j=1}^{m}[g_n(t_i,\ \mathrm{x})-\mathrm{g}(t_1,\ \theta))\Omega_{ij}^{-1}(g_n(t_j,\ \mathrm{x})-\mathrm{g}(t_j,\ \theta))] \tag{7}$$

を最小にする推定量である．ただし，Ω_{ij} は mxm の行列 Ω の (i, j) 要素が $\Omega_{ij}=\mathrm{g}(t_i+t_j,\ \theta)-\mathrm{g}(t_i,\ \theta)\mathrm{g}(t_j,\ \theta)$ である．Schmidt（1982）は，$\sqrt{n}(\hat{\theta}_{MMGF}-\theta)$ の漸近分布が $\mathrm{N}(0,\ \Psi_2)$ になることを示した．ただし，$\hat{\theta}_{MMGF}$ は，(7) を最小にする推定量で，$\Psi_2=(A'\Omega^{-1}A)^{-1}$ で，A は mx5 行列でその (i, j) 要素が$A_{ij}=\dfrac{\partial g(t_i,\ \theta)}{\partial\theta_j}$である．MGF と MMGF 推定量の問題は，グリッドポイントをいかに決めるのかという問題と MGF が有界ではないので収束に関する問題があるということである．

最尤推定と MGF および MMGF 推定量に問題があるため，Tran（1994）は特性関数（characteristic function（CF））に基づいて normal mixture を推定しようとした．その理由は，特性関数は積率母関数 MGF と違って，常に一様に有界であるが MGF は必ずしもそうでないことである．2 つの成分の normal mixture の特性関数は，

$$\mathrm{C}(\mathrm{t},\ \theta)=\mathrm{E}[e^{itX}]=(1-\mathrm{p})\exp(\mathrm{i}\mu_1\mathrm{t}-\frac{1}{2}\sigma_1^2t^2)+\mathrm{p}\exp(\mathrm{i}\mu_2\mathrm{t}-\frac{1}{2}\sigma_2^2t^2) \tag{8}$$

で与えられる．ただし，$\mathrm{i}=\sqrt{-1}$ である．$\exp(\mathrm{itx})=\cos(\mathrm{tx})+\mathrm{i}\sin(\mathrm{tx})$ であるので，(8) は以下のようになる：

$$\begin{aligned}\mathrm{C}(\mathrm{t},\ \theta)=&(1-\mathrm{p})\cos(\mu_1\mathrm{t})\exp(-\frac{1}{2}\sigma_1^2t^2)+\mathrm{p}\cos(\mu_2\mathrm{t})\exp(-\frac{1}{2}\sigma_2^2t^2)\\&+\mathrm{i}[(1-\mathrm{p})\sin(\mu_1\mathrm{t})\exp(-\frac{1}{2}\sigma_1^2t^2)+\mathrm{p}\sin(\mu_2\mathrm{t})\exp(-\frac{1}{2}\sigma_2^2t^2)]\end{aligned} \tag{9}$$

(9) の特性関数に対応する標本でのそれは，

$$C_n(\mathrm{t},\ \mathrm{x})=\frac{1}{n}\sum_{j=1}^{n}\exp(itx_j)=\frac{1}{n}\sum_{j=1}^{n}\cos(tx_j)+i[\frac{1}{n}\sum_{j=1}^{n}\sin(tx_j)] \tag{10}$$

である．大数の法則によって，$C_n(\mathrm{t},\ \mathrm{x})$ は $\mathrm{C}(\mathrm{t},\ \theta)$ に確率収束する．所与のグリッドポイント $(t_1,\ \cdots,\ t_m)$ に対して，Tran（1994）は，

$$\omega_1(t,\ \theta)=\sum_{j=1}^{m}(C_n(t_j,\ x)-\mathrm{C}(t_j,\ \theta))^2 \tag{11}$$

を θ に関して最小にすることにより θ を推定することを提案した．この場合，MGF 推定量の場合の Schmidt（1982）と同様に，

$$\omega_2(\mathrm{t},\ \theta)=\sum_{i=1}^{m}\sum_{j=1}^{m}[(C_n(t_i,\ \mathrm{x})-\mathrm{C}(t_i,\ \theta))\Omega_{ij}^{-1}(C_n(t_j,\ \mathrm{x})-\mathrm{C}(t_j,\ \theta))] \tag{12}$$

を最小にすることが考えられる．ここで，Ω_{ij} は Ω の (i, j) 要素で，Ω は，Tran（1994）や Yu（1998）で与えられている．この推定量を，DECF（discrete empirical characteristic function）推定量と呼ぶ．DECF 推定量は，Feuerverger and McDunnough（1981）により $\sqrt{n}(\hat{\theta}_{DECF}-\theta)$ の漸近分布が $N(0, \Psi_3)$ になることを示した．ここで，$\hat{\theta}_{DECF}$ は，（12）を最小にする DECF 推定量で，$\Psi_3=(A'\Omega^{-1}A)^{-1}$ で，A は mx5行列で，その (i, j) 要素が $A_{ij}=\dfrac{\partial c(t_i, \theta)}{\partial \theta_j}$ である (i=1, …m, j=1, …, 5)．DECF 推定量は，グリッドポイントをいかに決めるかという問題があり，また推定量を陽に解くことが困難であるという問題がある．

DECF 推定量の問題点を改善するのが CECF（continuous empirical characteristic function）推定量である．DECF 推定量で特性関数の真の値と標本での値を所与のグリッドポイントで最小にしたのに対して，CECF 推定量では 2 つの値の差を連続的なウエイト関数を使って連続的に最小にする．すなわち，CECF 推定量は，

$$D(\theta ; x)=\int_{-\infty}^{+\infty}|C_n(t, x)-C(t, \theta)|^2 w(t)dt \tag{13}$$

を θ に関して最小にする．ここで，w(t) は（13）の積分が存在するようなウエイト関数で，$w(t)=\exp(-bt^2)$ がよく使われる．ただし，b は非負の実数である．このウエイト関数を使うと，DGP が normal mixture の場合，（13）式と CECF 推定量の漸近分散が陽に導出できる．そのため，CECF 推定量と関連する計算上の問題が大幅に改善される．また，特性関数の真の値と標本での値の差を連続的に評価できる．それにより，グリッドポイントとグリッド間の距離の選択という問題が解決される．

もし DGP が(1)式で与えられる normal mixture で特性関数の真の値と標本での値の差が(13)であるとき，（13）式は，

$$\begin{aligned}D(\theta ; x)=&\frac{1}{n^2}\sqrt{\frac{\pi}{b}}\sum_{i=1}^{n}\sum_{j=1}^{n}\exp\left(-\frac{1}{4b}(x_i-x_j)^2\right)+\sum_{k=1}^{K}p_k^2\sqrt{\frac{\pi}{b+\sigma_k^2}}\\&+2\sum_{k=1}^{K}\sum_{h\neq k}^{K}p_k p_h\sqrt{\frac{\pi}{b+\frac{1}{2}(\sigma_k^2+\sigma_h^2)}}\exp\left(-\frac{(\mu_k-\mu_h)^2}{4b+2(\sigma_k^2+\sigma_h^2)}\right)\\&-\frac{2}{n}\sum_{k=1}^{K}\left[p_k\sqrt{\frac{\pi}{\frac{1}{2}\sigma_k^2+b}}\sum_{j=1}^{n}\exp\left(-\frac{(x_j-\mu_k)^2}{4b+2\sigma_k^2}\right)\right]\end{aligned} \tag{14}$$

で与えられる．特性関数に基づいた推定量の漸近特性は，Heathcote（1977）と Knight and Yu（2002）で確立された．（14）式を最小にする CECF 推定量 $\hat{\theta}_{CECF}$ については，$\sqrt{n}(\hat{\theta}_{CECF}-\theta)$ の漸近分布は $N(0, \Lambda^{-1}\Omega\Lambda^{-1})$ で与えられる．ここで，Λ と Ω は，(3K−1)x(3K−1) 行列でそれぞれ (i, j) 要素が

$$\Lambda_{ij}=E\left[\frac{\partial^2 D(\theta;x)}{\partial\theta_i\,\partial\theta_j}\right],\ \Omega_{ij}=E\left[\frac{\partial D(\theta;x)}{\partial\theta_i}\frac{\partial D(\theta;x)}{\partial\theta_j}\right]$$

で与えられる．Λ と Ω の表現は，Xu（2007）や Xu and Knight（2011）に示されている．

Xu and Knight（2011）は，以下の繰り返しのある推定方法を提案した．すなわち，

ステップ 1．b の初期値の b_0 を決める．

ステップ 2．b_0 と x のデータを（14）式に代入し，（14）式を最小にする $\hat{\theta}_0$ を計算する．

ステップ 3．$\hat{\theta}_0$ を漸近共分散行列に代入して，漸近共分散行列の推定値 M_0 を求める．さらに，trace(M_0) か行列式 det(M_0) を計算する．これらは b の関数であるので，trace(M_0) または行列式 det(M_0) を最小にする b_1 を求める．

ステップ 4．ステップ 2 と 3 を繰り返し，$|b_t - b_{t-1}| < 10^{-3}$ というような基準を満足するまで続ける．

Xu and Knight（2011）は，上記の方法が漸近的効率性を高めることを示している．

Normal mixture の下で GARCH モデルを仮定している場合の推定については，Alexander and Lazar（2006），Alexander and Lazar（2009），Kamaruzzaman and Isa（2014）などでは，Ox 3.30（Doornik（2002））と G@RCH 3.0（Laurent and Peters（2002））を用いた最尤法を使っている．一方，Xu and Wirjanto（2010）は，CECF 推定量を使って推定している．K 個の正規分布に従う 1 次元の確率変数の normal mixture で第 k 成分の t − 1 期までの情報が与えられた時の条件付き期待値が μ_k，条件付き分散が GARCH に従う $\sigma^2_{k,\,t}$ のとき，Xu and Wirjanto（2010）は，

$$\mathrm{C}(r,\ \theta)=E[e^{irX}]=\textstyle\sum_{k=1}^{k} p_k \exp(i\mu_k r-\frac{1}{2}\sigma^2_{k,\,t} r^2)$$

$$C_t(\mathrm{r},\ \mathrm{x})=\cos(\mathrm{r}x_t)+\mathrm{i}\sin(\mathrm{r}x_t)$$

のとき，

$$D_t(\theta\,;\mathrm{x})=\int_{-\infty}^{\infty}|C_t(r,\ x)-C(r,\ \theta)|^2 w(r)dr \tag{15}$$

は，以下の式と同値である；

$$\begin{aligned}D_t(\theta;\mathrm{x})=&\sqrt{\frac{\pi}{b}}+\textstyle\sum_{k=1}^{K} p_k^2\sqrt{\frac{\pi}{b+\sigma^2_{k,\,t}}}\\&-2\textstyle\sum_{k=1}^{K}(p_k\sqrt{\frac{\pi}{\frac{1}{2}\sigma^2_{k,\,t}+b}}\exp(-\frac{(x_t-\mu_k)^2}{4b+2\sigma^2_{k,\,t}}))\\&+2\textstyle\sum_{k\neq h} p_k p_h\sqrt{\frac{\pi}{b+\frac{1}{2}(\sigma^2_{k,\,t}+\sigma^2_{h,\,t})}}\exp(-\frac{(\mu_k-\mu_h)^2}{4b+2(\sigma^2_{k,\,t}+\sigma^2_{h,\,t})}).\end{aligned} \tag{16}$$

CECF推定量 $\hat{\theta}_0$ は, $\mathrm{D}(\theta)=\sum_{t=1}^{T} D_t(\theta;\mathrm{x})$ を θ に関して最小にすることにより得られ, $\sqrt{T}(\hat{\theta}-\theta)$ の漸近分布は $N(0,\ \Lambda^{-1}\Omega\Lambda^{-1})$ で与えられる．ただし，$\Lambda=\mathrm{E}[\frac{\partial^2 D(\theta)}{\partial\theta\partial\theta'}]$，$\Omega=\mathrm{E}[\frac{\partial D(\theta)}{\partial\theta}\frac{\partial D(\theta)}{\partial\theta'}]$ である．このように，CECF 推定量は目的関数を陽に示すことができるので推定は容易に実行できる．

Xu and Wirjanto（2010）は，推定に関して VaR，expected shortfall などのリスク測度を，2 つの成分の normal mixture で計算し，それぞれデータと整合的か否かを検証するバックテストを行っていて，GARCH がある場合とない場合の normal，t 分布などと比較して normal mixture で GARCH がある場合が最もデータと整合的という結果を得ている．

4. その他の問題点

以上の実証結果を見てみると，最近の normal mixture を用いた実証研究では，主に 2 成分の normal mixture を使う例が最も多いようである．初期の Kon（1984）などは例外的に 3 成分以上の normal mixture を考えているが, Haas et al.(2004a, b)や Alexander and Lazar(2006)などの結果では 3 成分の normal mixture では推定の際の計算の収束が得られにくいとか，予測の際にパラメータの数が多すぎることが知られている．この点，CECF 推定量ではどうなのかについては先行研究だけからは不明である．

また，normal mixture の場合に成分の数の決定は重要な問題であり，これに関しては Kasahara and Shimotsu（2015）が最近回帰モデルでの normal mixture の場合に，成分の検定に関する論文を書いているので，この結果などをファイナンスで使われる normal mixture のモデルに適応させて検定をすることが望ましいように思われる．

参考文献

Alexander, C. and Lazar, E. (2006). Normal mixture GARCH(1,1): applications to exchange rate modelling. *Journal of Applied Econometrics*, 21, 307–336.

Alexander, C. and Lazar, E. (2009). Modelling regime-specific stock price volatility. *Oxford Bulletin of Economics and Statistics*, 71, 761–797.

Alexander, C. and Sheedy, E. (2008). Developing a stress testing framework based on market risk models. *Journal of Banking and Finance*, 32, 2220–2236.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics*, 31, 309–328.

Cai, J. (1994). A Markov model of switching-regime ARCH. *Journal of Business & Economic Statistics*, 12, 309–316.

Chen, J. and Tan, X. (2009). Inference for multivariate normal mixtures. *Journal of Multivariate Analysis*, 100, 1367–1383.

Doornik, J. A. (2002). *Obeject-Oriented Matrix*

Programming Using Ox, 3rd ed., Timberlake Consultants Press and Oxford, London, available from: http:www.nuff.ox.ac.uk/Users/Doornik.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50, 987–1007.

Feuerverger, A. and McDunnough, P. (1981). On the efficiency of empirical characteristic function procedures. *Journal of the Royal Statistical Society*, 43, 147–156.

Gray, S. F. (1996). Modelling the conditional distribution of interest rates as a regime-switching process. *Journal of Financial Economics*, 42, 27–62.

Haas, M., Mittnik, S., and Paolella, M. S. (2004a). Mixed normal conditional heteroscedasticity. *Journal of Financial Econometrics*, 2, 211–250.

Haas, M., Mittnik, S., and Paolella, M. S. (2004b). A new approach to Markov switching GARCH models. *Journal of Financial Econometrics*, 2, 493–530.

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57, 357–384.

Hamilton, J. D. and Susmel, R. (1994). Autoregressive conditional heteroscedasticity and changes in regime. *Journal of Econometrics*. 64, 307–333.

Hastie, T., Tibshirani, R., and Friedman J. (2001) *The Elements of Statistical Learning*, New York: Springer Verlag, 272–275.

Heathcote, C. R. (1977). Integrated mean square error estimation of parameters. *Biometrika*, 64, 255–264.

Hsieh, D. A. (1989). Modeling heteroscedasticity in daily foreign-exchange rates. *Journal of Business and Economic Statistics*, 11, 307–317.

Hull, J. and White, A. (1998). Value at risk when daily changes in market variables are not normally distributed. *Journal of Derivatives*, 5, 9–19.

Kamaruzzaman, Z. A. and Isa, Z. (2013). Estimation of value at risk and conditional value at risk using normal mixture distributions model. *Proceedings of the 20th National Symposium at Mathematical Sciences*, AIP Conf. Proc. 1522, 1123–1131.

Kamaruzzaman, Z. A. and Isa, Z. (2014). Empirical application of normal mixture GARCH and value-at-risk estimation. *Proceedings of the 3rd International Conference on Mathematical Sciences*, AIP Conf. Proc. 1602, 453–459.

Kasahara, H. and Shimotsu, K. (2015). Testing the number of components in normal mixture regression models. *Journal of the American Statistical Association*, 110, 1632–1645.

Kim, S., N. Shephard, and S. Chib (1998). Stochastic volatility: likelihood inference and comparison with ARCH models. *Review of Economic Studies*, 45, 361–393.

Klaassen, F. (2002). Improving GARCH volatility forecasts with regime-switching GARCH. *Empirical Economics*, 27, 363–394.

Knight, J. L. and Yu, J. (2002). Empirical characteristic function in time series estimation. *Econometric Theory*, 18, 691–721.

Kon, S. J. (1984). Models of stock returns-a comparison. *Journal of Finance*, 39, 147–165.

Laurent, S. and Peters, J.-P. (2002). G@RCH2.2: an Ox package for estimating and forecasting various ARCH models. *Journal of Economic Surverys*, 16, 447–485.

Mahieu, R. and Schotman, P. (1998). An empirical application of stochastic volatility models. *Journal of Applied Econometrics*, 13, 333–359.

Omori, Y., Chib, S., Shephard, N., and Nakajima, J. (2007). Stochastic volatility with leverage: fast and efficient likelihood inference. *Journal of Econometrics*, 140, 425–449.

Quandt, R. E. and Ramsey, J. B. (1978). Estimating mixtures of normal distributions and switching regressions. *Journal of the American Statistical Association*, 73, 730–738.

Schmidt, P. (1982). An improved version of the Quandt-Ramsey MGF estimator for mixtures of normal distributions and switching regressions. *Econometrica*, 50, 501–516.

Taylor, S. J. (1986). *Modelling Financial Time Series*, Wiley: Chichester, UK.

Tran, K. (1994). Mixture, moment and information-three essays in econometrics, Ph. D. Thesis, The University of Western Ontario.

Vlaar, P. and Palm, F. (1993). The message in weekly exchange rates in the european monetary system: mean reversion, conditional heteroskedasticity and jumps. *Journal of Business and Economic Statistics*, 11, 351–360.

Wirjanto, T. S. and Xu, D. (2010). The applications of mixtures of normal distributions in empirical finance: a selected survey. Working paper, The University of Waterloo.

Wong, C. and Li, W. (2001). On a mixture autoregressive conditional heteroscedastic model. *Journal of the American Statistical Association*. 96, 982–995.

Xu, D. (2007). Asset returns, volatility and value-at-risk, Ph. D. Thesis, The University of Western Ontario.

Xu, D. and Knight, J. L. (2011). Continuous empirical characteristic function estimation of mixtures of normal parameters. *Econometric Reviews*, 30, 25–50.

Xu, D. and Knight, J. L. (2012). Stochastic volatility model under a discrete mixture of normal specification. forthcoming in *Journal of Economics and Finance*.

Xu, D. and Wirjanto, T. S. (2010). An empirical characteristic approach to VaR under a mixture-of-normal distribution with time-varying volatility. *Journal of Derivatives*, 18, 39–58.

Yu, J. (1998). Empirical characteristic function in time series estimation and a test statistic in financial modeling, Ph. D. Thesis, The University of Western Ontario.

Zangari, P. (1996). *An Improved Methodology for Measuring VaR*, RiskMetrics Monitor, Reuters/JP Morgan.