

要旨

本稿は Oxford-man Institute で提供されている世界各国の市場のデータから、リーマンショック以前の 2000 年 1 月 1 日から 2006 年 12 月 31 日までのデータを取り出し、ボラティリティについて分析を行ったものである。近年、収益率のボラティリティの精度のよい推定量として、Realized Volatility (RV) と Realized Kernel (RK) が注目を浴びている。本稿では RV および RK を組み入れたモデルに対して、ボラティリティの予測精度の比較を行った。その際、ボラティリティの長期記憶性を捉えるために RV に対して ARFIMAX モデルを適用した。また RV と RK におけるマイクロストラクチャ・ノイズの比較を行うため、Realized-GARCH(R-GARCH)モデルと Realized-SV(Realized-SV) モデルによる分析も行った。これら 3 つのモデルを使用して 20 期先のボラティリティを予測しその予測精度を比較した。その結果、R-SV モデルのパフォーマンスが一番良く、ARFIMAX モデルについては優れた結果が得られなかった。

Key word : Realized Volatility, R-SV model, R-GARCH model, ARFIMAX model

1. はじめに

ボラティリティとは、価格の変化率の分散(若しくは標準偏差)のことであり、オプションなどの派生証券の価格を決定する重要な要素の 1 つである。ボラティリティの特徴の 1 つとして「非対称性」と言われるものがある。これは株価が上がった日の翌日より株価が下がった日の翌日のほうがよりボラティリティが上昇する傾向がある事を意味している。この効果はバレッジ効果(Leverage effects)という。他にも重要な特徴として、ボラティリティの持続性があげられる。これは証券市場にショックが生じた場合に、いったん収益率のボラティリティが大きく(小さく)なるとその後しばらくボラティリティが大きい(小さい)状態が続くということを意味している。

ボラティリティは様々な場面において重要な変数である。しかしながら、真のボラティリティは直接観測することが出来ないため、ボラティリティを推定するための様々な方法が提案された。RV もそのような過程で提案されたボラティリティの推定量の一つである。RV は高頻度データと呼ばれる、分単位、あるいは秒単位の非常に高い頻度で観測されるデータを用いて計算される。その定義や性質として渡部(2009, p.1) は以下の様に説明している

“Realized Volatility は日中リターンの 2 乗を 1 日分足し合わせたものなので、ボラティリティのモデルに依存しない。また、Realized Volatility は計算に用いる日中リターンに時間間隔を 0 に近づけると、真のボラティリティに確率収束する(ただし、実際の資産価格はマイクロストラクチャ・ノイズを含むので必ずしもそうならない)。そこで、時間間隔の短い日中リターンを使って計算すれば、真のボラティリティの精度の高い推定量になる。さらに、Realized

* 慶應義塾大学経済学部 3 年長倉大輔研究会所属

** 慶應義塾大学経済学部 3 年長倉大輔研究会所属

*** 慶應義塾大学経済学部 3 年長倉大輔研究会所属

Volatility を用いて将来のボラティリティを予測すると、ARCH 型モデルを用いた場合と比べて、将来のボラティリティの予測パフォーマンスが高いことが明らかになっている(渡部・佐々木(2007, p.1)、渡部(2008, p.1))”。

RV はデータさえあれば複雑な計算を要せず簡単に計算できる一方、緩い過程のもとで一致性をもつ推定量であるため近年実証分析における応用が広がりつつあるが、その推定の際には 2 つの問題が存在する。1 つは、マーケットマイクロストラクチャ・ノイズの存在である。この問題点について渡部(2010, p.1)は

“もし資産価格にノイズがなければ、計算に用いる日中リターンを 0 に近づけると Realized Volatility は真のボラティリティに収束する。しかし、実際の資産価格はマイクロストラクチャ・ノイズを含むので、日中リターンを 0 に近づけると、Realized Volatility はマイクロストラクチャ・ノイズばかりを拾ってしまう。”

と述べている。また、2 つめとして挙げられるのは、取引外の時間帯のリターンの処理である。この点について渡部(2009)は

“東京証券取引所の取引時間は、9:00-11:00 (前場) と 12:30-15:00 (後場) なので、前日の 15:00-9:00(夜間)と 11:00-12:30(昼休み) の間は 1 分毎の価格が得られない。夜間と昼休みは時間間隔が長いので、それらのリターンをそのまま 2 乗して加えると、時間の離散化による誤差が無視できなくなる可能性がある。”

と述べている。RV を用いた推定には上記のような問題点があるので、何らかの対処が必要となる。本稿では、マイクロストラクチャ・ノイズの問題に対しては、後述する RK を用いる事や、R-GARCH モデルおよび R-SV モデルを用いることにより対応している。

R-GARCH モデルと R-SV モデルは収益率とボラティリティを同時にモデル化しているが、RV によりボラティリティの精度の良い推定値を得られる事から、これを真のボラティリティとみなして、RV を直接モデル化するというアプローチもある。RV には長期記憶性がある事が知られており、この性質を考慮して、本稿では RV を長期記憶モデルである、ARFIMAX モデルも推定した。実証分析において、全てのモデルにおいてレバレッジ効果が確認され、また ARFIMAZ モデルの推定結果によると RV に長期記憶過程が存在していることも示唆された。予測においては、R-SV モデルのパフォーマンスが一番優れており、ARFIMAX モデルは他の 2 つのモデルに比べてパフォーマンスが劣っていることが分かった。

2. モデル

この章では、従来のボラティリティ変動モデルを述べた後、高頻度データを用いた推定量である RV を取り入れたモデルを解説する。先行研究では、RV を用いたモデルの方が用いないモデルよりも予測精度が高いことが示されている。今回の論文では、RV を用いたモデルを中心に分析し、比較することにする。具体的には、R-GARCH、R-SV、ARFIMAX モデルの 3 つを分析する。

2.1 GARCH モデル

ボラティリティの変動モデルは、大きく Engle (1982)が提唱した ARCH (Autoregressive Conditional Heteroskedasticity) モデルから派生した GARCH (Generalized ARCH) モデルと Taylor(1986)の SV (Stochastic Volatility) モデルがある。まずは、この 2 つのモデルについて説明する。

資産価格の対数値を p_t とすると、リターン y_t は

$$y_t = p_t - p_{t-1} \quad (2.1)$$

と定義される。このとき、GARCH(p, q)モデルは次のように定義される。

$$y_t = \mu + \varepsilon_t \quad (2.2)$$

$$\varepsilon_t = \sigma_t z_t, \sigma_t > 0, z_t \sim i. i. d. (0, 1) \quad (2.3)$$

$$\sigma_t^2 = \omega + \sum_{i=1}^p \beta_i \sigma_{t-i}^2 + \sum_{j=1}^q \gamma_j \varepsilon_{t-j}^2 \quad (2.4)$$

$$\omega > 0, \beta_i, \gamma_j \geq 0 (i = 1, \dots, p; j = 1, \dots, q)$$

ここで、 z_t を平均 0 で分散 1 の同一分布に従う誤差項とする。以下 σ_t^2 をボラティリティと呼ぶ。GARCH モデルは、ボラティリティが過去のボラティリティと誤差項 ε_t の 2 乗に依存している。ここで、パラメータ $\omega, \beta_i, \gamma_j$ に非負条件を課しているのは、このモデルにおいてボラティリティが時間を通じて正になる事を保証するためである。

GARCH モデルには重要な問題点が存在する。前章でも述べたように、ボラティリティには前期のリターンがマイナスだった場合に、プラスだったときよりも今期のボラティリティの値が上昇するというレバレッジ効果が存在する。この GARCH モデルは、このレバレッジ効果を説明できないため、様々な改良がなされてきた。そのようなモデルの例として、Nelson (1991) の EGARCH や Glosten et al (1993) の GJR などがある。EGARCH モデルにおいて σ_t^2 は、次のように定式化される。

$$\ln \sigma_t^2 = \omega + \sum_{i=1}^p \beta_i [\ln \sigma_{t-i}^2 - \omega] + g(z_{t-1}) + \sum_{j=1}^q \psi_j g(z_{t-j-1}) \quad (2.5)$$

$$g(z_{t-j-1}) = \theta z_{t-j-1} + \gamma [|z_{t-j-1}| - E(|z_{t-j-1}|)]$$

$$(i = 1, \dots, p; j = 0, \dots, q)$$

このモデルではボラティリティを対数化することによって、パラメータに非負制約を課す必要がなくなる。また、この式のレバレッジ効果を表すパラメータは θ である。例えば、 z_{t-1} がマイナスのときに、 θz_{t-1} がプラスになったら、 $\ln \sigma_t^2$ が上昇するため、レバレッジ効果が起こっていることを示すことが出来る。つまり、 $\theta < 0$ であれば、レバレッジ効果が存在するという事になる。

2.2 Stochastic Volatility モデル

Stochastic Volatility (SV) モデルとは、確率的ボラティリティモデルの略称であり、次のように表される。

$$y_t = \varepsilon_t = \sigma_t z_t \quad (2.6)$$

$$\ln \sigma_t^2 = \omega + \varphi \ln \sigma_{t-1}^2 + \eta_t \quad (2.7)$$

$$z_t \sim i. i. d. (0, 1) \quad \eta_t \sim i. i. d. (0, \sigma_\eta^2)$$

$$\ln \sigma_1^2 \sim i. i. d. N \left(\frac{\omega}{1 - \varphi}, \frac{\sigma_\eta^2}{1 - \varphi^2} \right)$$

このモデルでは、ボラティリティの対数値を AR(1)過程で表現する。ここでは、 $|\varphi| < 1$ を仮定し、初期値である $\ln\sigma_1^2$ は平均が $\omega/(1-\varphi)$ 、分散が $\sigma_\eta^2/(1-\varphi^2)$ の正規分布に従うとした。このモデルでは、 $Cov(z_t, \eta_t) = \rho\sigma_\eta$ というように z_t (前期のリターンの誤差項)と η_t (今期のボラティリティの誤差項) の相関に ρ という定数を置くことによって、前期のリターンが今期のボラティリティに与える影響を捉えることが出来る。このとき、レバレッジ効果というのは、前期のリターンが負であると、今期のボラティリティが上昇するものである。つまり、 ρ の値が負であれば、レバレッジ効果が生じていると言う事が出来る。また、 $|\varphi| < 1$ であると、定常性を満たし、1 に近ければ近いほど持続性が存在することが分かる。

2.3 Realized Volatility

RV というのは、1 分、2 分ごとといった細かい時間間隔で観測される金融資産の収益率の 2 乗和として計算される。具体的には次のように定義される。第 t 日の n 個の日中リターンのデータ $\{r_{t-1+1/n}, r_{t-1+2/n}, \dots, r_t\}$ があるものとする。このとき、それらを 2 乗して足し合わせた

$$RV_t = \sum_{i=1}^n r_{t-1+i/n}^2 \quad (2.8)$$

を第 t 日の RV をいう。また、資産価格の対数値 $p(s)$ が以下のような拡散過程:

$$dp(s) = \mu(s)ds + \sigma(s)dW(s) \quad (2.9)$$

に従っているとしよう。ここで、 $W(s)$ はワイナー過程であり、 $\mu(s)$ は瞬時的なドラフト、 $\sigma(s)$ は瞬時的なボラティリティと呼ばれる。このとき、第 t 日のボラティリティ、 σ_t^2 、を $\sigma(s)^2$ の積分として次のように定義しよう。

$$\sigma_t^2 = \int_{t-1}^t \sigma^2(s)ds \quad (2.10)$$

これは瞬時的ボラティリティを積分したものなので、IV (Integrated Volatility) と呼ばれる。 RV_t は、 $n \rightarrow \infty$ の時、IV に確率収束する。よって、 n が十分に大きいと、 RV_t は真のボラティリティに対して精度の高い推定量になる。

RV には大きく 2 つの問題が存在する。1 つはマーケット・マイクロストラクチャ・ノイズ (以下ではマイクロストラクチャ・ノイズ) に起因する問題であり、もう 1 つは、取引時間外のリターンの取り扱いである。前者のマイクロストラクチャ・ノイズというのは、市場のマイクロ構造に起因し派生する観測価格に含まれる真の価格(効率的価格)からの観測誤差のことである。具体例としては、bid-ask bounce や非同時取引による効率的価格からの乖離があげられる。bid-ask bounce は、取引が売値や買値で成立するために真の取引価格から離れてしまい、負の自己相関が生じてしまうものである。また、非同時取引では多くの銘柄に関する情報を受け取ったときに取引頻度の高い銘柄の価格から反応して、それから順番に次に取引の高い銘柄の価格が反応することで、それらの平均価格のリターンに自己相関が発生してしまう事が問題になる。マイクロストラクチャ・ノイズが存在すると RV は真のボラティリティに確率収束しない。現在観測される対数価格を $\hat{p}(s)$ 、真の対数価格を $p(s)$ とし、マイクロストラクチャ・ノイズを $\tau(s)$ とする。これらの関係は以下のようにであると仮定しよう。

$$\begin{aligned} \hat{p}(s) &= p(s) + \tau(s), \\ \tau(s) &\sim WN(0, \sigma_\tau^2) \end{aligned} \quad (2.11)$$

このとき、先ほどの式を $t - \Delta$ から t までのリターンに直し、分散を計算すると次のようになる。

$$\text{Var}(\hat{p}(t) - \hat{p}(t - \Delta)) = \text{Var}(p(t) - p(t - \Delta)) + \text{Var}(\tau(t) - \tau(t - \Delta)) \quad (2.12)$$

ここで、右辺第 1 項は、時間間隔 Δ を小さくすると 0 に近づいていくが、それに対して、右辺第 2 項は、 Δ に依存せず、一定の値 $2\sigma_\tau^2$ になることが分かる。したがって、 Δ が小さくなるにつれて、真のボラティリティに比べてマイクロストラクチャ・ノイズの分散が大きくなってしまふ。つまり、時間間隔を短くすると、真のボラティリティに収束しやすい反面、マイクロストラクチャ・ノイズの影響を大きく受けてしまふ。先行論文では、この問題を解決するために適切な時間間隔を用いることが提案されており、それらによると、おおむね 5 分間隔のリターンを用いるとマイクロストラクチャノイズの影響を十分に軽減できる事が示されている。本稿では、それにしたがって 5 分ごとのリターンを用いることにする。その他の対処法として、Barndorff-Nielsen, Hansen, Lunde, and Shephard (2008, p.2-3) はマイクロストラクチャ・ノイズが存在する場合でも IV の一致推定量である RK を提案した。これは第 4 章に詳しく説明することにする。

もう 1 つの取引時間外のリターンの取り扱いというのは、夜間や昼休みなどで取引が行われていないときのリターンは短い時間間隔のリターンを計算することが出来ない事に起因する問題である。この結果、時間間隔が長くなつてまい、RV の離散化誤差が大きくなる可能性がある。そこで、Hansen and Lunde (2005, p.525-554) は、次のように夜間と昼休みを除いて計算した RV, \widehat{RV}_t , に定数 c を掛ける方法を提案している。

$$RV_t = c\widehat{RV}_{t,c} = \frac{\sum_{t=1}^T (R_t - \bar{R})^2}{\sum_{t=1}^T \widehat{RV}_t} \quad (2.13)$$

定数 c は日次リターンの標本分散と \widehat{RV} の標本平均の比を掛けている。この方法を用いると、RV の標本平均と日時リターンの標本分散が等しくなる。しかし、今回の分析ではこの取引時間外の考慮をしていないため、誤差が存在する可能性が高い。

2.4 Realized GARCH モデル

RV は説明変数としても様々なモデルに応用されている。ここでは GARCH モデルに RV を取り入れた Realized GARCH モデルについて説明する。Hansen, Huang, Shek (2011) を参考にすると、Realized GARCH (p, q) モデルは次のように定義される。

$$y_t = \mu + \varepsilon_t, \varepsilon_t = \sigma_t z_t \quad (2.14)$$

$$\ln \sigma_t^2 = \omega + \sum_{i=1}^p \beta_i \ln \sigma_{t-i}^2 + \sum_{j=1}^q \gamma_j \ln RV_{t-j} \quad (2.15)$$

$$\ln RV_t = \xi + \varphi \ln \sigma_t^2 + \tau(z_t) + u_t \quad (2.16)$$

$$\tau(z_t) = \tau_1 z_t + \tau_2 (z_t^2 - 1)$$

$$u_t \sim i.i.d. \quad E(u_t) = 0, \text{Var}(u_t) = \sigma_u^2$$

GARCH モデルと比較すると、ボラティリティが対数化されていて、誤差項 ε_t の 2 乗の代わりに対数化された RV が用いられている。それに加えて、RV を同期のボラティリティで説明している式がある。また、 $\tau(z_t)$ は、レバレッジ関数と呼ばれ、レバレッジ効果を表すパラメータを示している。例として、R-GARCH(1,1)は、

$$\ln \sigma_t^2 = \omega + \beta \ln \sigma_{t-1}^2 + \gamma \ln RV_{t-1} \quad (2.17)$$

$$\ln RV_t = \xi + \varphi \ln \sigma_t^2 + \tau_1 z_t + \tau_2 (z_t^2 - 1) + u_t \quad (2.18)$$

のように定義される。このとき、 z_{t-1} が負で、なおかつ $\tau_1 < 0, \gamma > 0$ 、のときに、 t 期のボラティリティがより上昇するため、レバレッジ効果を表現することが可能になる。また、 $\varphi = 1$ のとき、 $\ln RV_t$ と $\ln \sigma_t^2$ は比例していて、 ξ は先ほど述べたRVにおけるマイクロストラクチャ・ノイズの影響を捉えている。つまり、RVがボラティリティの不偏推定量であるには、 $\varphi = 1$ かつ $\xi = 0$ のときである。Hansen(2011, p.877-906)によると、 φ の推定値は1に近づいている。

2.5 Realized SV モデル

次にSVモデルにRVを取り入れたRealized SVモデルを説明する。Realized SVモデルは従来のSVモデルにRVの式を加えたモデルである。

$$\ln RV_t = \xi + \ln \sigma_t^2 + u_t \quad (2.19)$$

$$u_t \sim i.i.d. \quad E(u_t) = 0, \quad Var(u_t) = \sigma_u^2$$

この式は、大森、渡部 (2013, p.288) を参考にしたものであり、(2.18)式において、 $\varphi = 1$ 、 $\tau_1 = \tau_2 = 0$ 、としたもの等しい。この式は、今期のボラティリティの対数値が、今期のRVの対数値を説明しているものであり、R-GARCHモデルと同じように、 ξ はマイクロストラクチャ・ノイズを説明している定数項である。また、誤差項 u_t と z_t, η_t には相関がないと仮定している。大森、渡部 (2013, p.274) は、このモデルがこれまでのSVモデルよりもボラティリティの予測精度が高いことを示した。

2.6 ARFIMAX モデル

ある時系列過程における k 次の自己相関係数を $\rho(k)$ としよう。この時、

$$\sum_{k=1}^{\infty} |\rho(k)| < \infty \quad (2.20)$$

を満たしているとき、この過程は短期記憶過程に従うといわれる。それに対し、

$$\sum_{k=1}^{\infty} |\rho(k)| = \infty \quad (2.21)$$

で表される過程は長期記憶過程と呼ばれる。長期記憶過程と短期記憶過程をそれぞれグラフで表すと以下のようになる。

長期記憶過程では自己相関係数が幾何級数的に減衰していき、短期記憶過程では指数関数的に減衰していく。すなわち、長期記憶過程では自己相関係数の減衰スピードが極めて遅いため、自己相関係数の無限和がある一定の値に収束せず発散するということを示している。それに対し短期記憶過程では、減衰スピードが非常に速いため、ある一定の値に収束する。

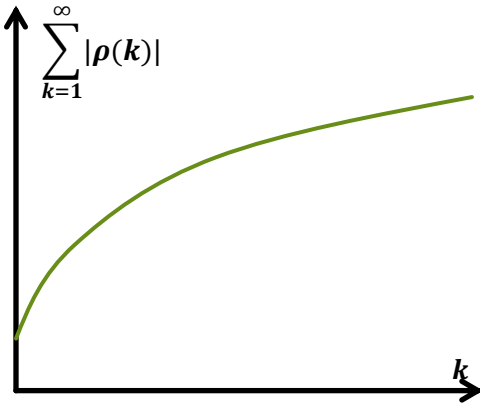


図 2-1. 長期記憶過程

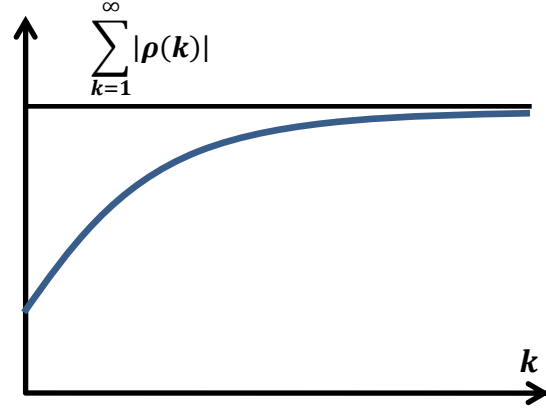


図 2-2. 短期記憶過程

GARCHモデルやSVモデルでは、パラメータが定常性を満たす限りボラティリティは短期記憶過程に従うが、実際のデータにおいては、リターン²の自己相関の減衰スピードは比較的遅いことから、ボラティリティは長期記憶過程に従うと考えられる。そこで、長期記憶過程のモデルとしてARFIMAXモデルを用いることにする。前述のR-GARCHモデルやR-SVモデルでは、リターンとボラティリティをモデル化しているのに対し、ARFIMAXモデルではRVの対数値をモデル化している。つまり、このモデルはRVの変動を分析するためのものである。

ARFIMAX (0, d, 1) モデルは以下のように表される。

$$(1 - L)^d \{ \ln RV_t - \mu_0 - \mu_1 |R_{t-1}| - \mu_2 D_{t-1}^- |R_{t-1}| \} = (1 + \theta L) u_t \quad (2.22)$$

$$u_t \sim i. i. d. N(0, \sigma_u^2)$$

$$D_{t-1}^- = \begin{cases} 1 & R_{t-1} < 0 \\ 0 & R_{t-1} \geq 0 \end{cases}$$

Lはラグオペレータであり、AR(∞)の影響を考慮することが出来る。また、 D_{t-1}^- は前期のリターンがマイナスなら1、プラスなら0を示すダミー変数である。このダミー変数を用いることにより、RVのレバレッジ効果を表すことが出来る。前期のリターンがマイナスなら、以下のような式に変形することが出来る。

$$\ln RV_t = \mu_0 + (\mu_1 + \mu_2) |R_{t-1}| + (1 - L)^{-d} (1 + \theta L) u_t \quad (2.23)$$

レバレッジ効果の場合、RVの値は大きくなっていないといけないため、 μ_2 が正でなければいけない。つまり、 μ_2 が正のときに、レバレッジ効果を確認することが出来る。また、パラメータdに関しては、次の表2-1を見てもらうと分かりやすい。

表 2-1. パラメータ d の説明

	d = 0	0 < d < 0.5	0.5 ≤ d < 1	d = 1
分散	有限 (定常)		無限 (非定常)	
自己相関の和	短期記憶	長期記憶		単位根

この表から見て分かるように、 $d = 0$ の場合は短期記憶過程で定常になり、 $d = 1$ の場合は単位根で非定常になる。また、 $0 < d < 1$ の場合には長期記憶過程に従っており、 $d < 0.5$ のときは定常、 $0.5 \leq d$ のときは非定常である。

これまで、様々なボラティリティ変動モデルを説明してきた。第 1 章でボラティリティや RV の性質を述べたが、これらの特徴を 1 つのモデルで表現することは難しい。表 2-2 には各モデルの特徴がまとめられている。この表から、今回分析するモデルを選択する際に大きく 2 つの点を考えた。1 つは長期記憶過程であり、それが観測できるのは ARFIMAX モデルしか存在しない。渡部、佐々木 (2006, p.39) は、RV が長期記憶過程に従っていると述べているので、その変動を確かめるためには ARFIMAX モデルを取り入れるのは当然であろう。しかし、ARFIMAX は他のモデルとは異なり、RV の変動を捉えるモデルなので、今回予測するボラティリティと少し違いが生じるかもしれない。しかし、RV がボラティリティの精度の良い推定量であるならば、ARFIMAX モデルも他のモデルと同様、ボラティリティの予測に役立つだろう。今回の予測評価の際にも真のボラティリティの代理変数として RV を用いている。もう 1 つは、RV に存在するマイクロストラクチャ・ノイズの影響であり、それが観測出来るのは、R-GARCH モデルと R-SV モデルである。そのため、RV と RK を比較する際にこのポイントは必要不可欠なので、これら 2 つのモデルを選ぶのは妥当である。

表 2-2. 各モデルの特徴

モデル	持続性		定常性	レバレッジ 効果	マイクロストラクチャ ・ノイズの影響
	短期	長期			
GARCH	○	×	○	×	×
EGARCH	○	×	○	○	×
SV	○	×	○	○	×
R-GARCH	○	×	○	○	○
R-SV	○	×	○	○	○
ARFIMAX	○	○	○	○	×

3. 推定方法

ここまで、ボラティリティ変動を捉えるモデルの候補を挙げてきたが、この章では、これらのモデルのパラメータの推定方法を述べていく。R-GARCH モデルや ARFIMAX モデルは、基本的には疑似最尤法 (QMLE) を用いて推定する。R-SV モデルに関しては、疑似最尤法以外にも一般化モーメント法 (GMM) やマルコフ連鎖モンテカルロ法 (MCMC) など様々な推定方法が存在するが本稿では、疑似最尤法によって推定を行う。

3.1 R-GARCH モデルの推定

ここで再び R-GARCH モデルを再掲すると、以下のようになる。

$$y_t = \varepsilon_t \quad (3.1)$$

$$\varepsilon_t = \sigma_t z_t, \sigma_t > 0, z_t \sim i. i. d. (0, 1) \quad (3.2)$$

$$\ln \sigma_t^2 = \omega + \sum_{i=1}^p \beta_i \ln \sigma_{t-i}^2 + \sum_{j=1}^q \gamma_j \ln RV_{t-j} \quad (3.3)$$

$$\ln RV_t = \xi + \varphi \ln \sigma_t^2 + \tau_1 z_t + \tau_2 (z_t^2 - 1) + u_t \quad (3.4)$$

$$u_t \sim i. i. d. (0, \sigma_u^2) \quad (3.5)$$

このモデルで推定するパラメータは、 $\theta = (\omega, \beta_i, \gamma_j, \xi, \varphi, \tau_1, \tau_2, \sigma_u^2)$ である。このとき、疑似最尤法では、(3.2) , (3.5) の誤差項は以下の (3.6) の正規分布に従っていると仮定して推定を行う。

$$z_t \sim i.i.d. N(0, 1), \quad u_t \sim i.i.d. N(0, \sigma_u^2) \quad (3.6)$$

上記のように仮定するが、実際の分布が異なっても疑似最尤推定によって一致推定できることが知られている。(疑似)尤度関数は次のような 2 つの尤度式を掛け合わせたモデルになる。

$$L = \prod_{t=1}^T [l_1(y_t | RV_1, \dots, RV_{t-1}) \times l_2(RV_t | RV_1, \dots, RV_{t-1}, y_t)] \quad (3.7)$$

$$l_1(y_t | RV_1, \dots, RV_{t-1}) = \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(-\frac{y_t^2}{2\sigma_t^2}\right) \quad (3.8)$$

$$l_2(RV_t | RV_1, \dots, RV_{t-1}, y_t) = \frac{1}{\sqrt{2\pi\sigma_u^2}} \exp\left(-\frac{u_t^2}{2\sigma_u^2}\right) \quad (3.9)$$

次のように推定を行う。

R-GARCH モデルの疑似最尤推定

パラメータ θ を $\theta_1 = (\omega, \beta_i, \gamma_j)$, $\theta_2 = (\xi, \varphi, \tau_1, \tau_2, \sigma_u^2)$ と分解する。

① y_t と σ_0^2, RV_0 が与えられる。

→ l_1 ((3.8)式に(3.3)式を代入) を疑似最尤法により、パラメータ $\theta_1 = (\omega, \beta_i, \gamma_j)$ を推定する。

② 推定したパラメータ $\theta_1 = (\omega, \beta_i, \gamma_j)$ を用いて、 σ_t^2 を求める。

③ 求めた σ_t^2 を用いて、 $z_t = y_t / \sigma_t$ を計算する。

④ y_t と σ_t^2, RV_t, z_t が与えられる。

→ l_2 ((3.9)式) を疑似最尤法により、パラメータ $\theta_2 = (\xi, \varphi, \tau_1, \tau_2, \sigma_u^2)$ を推定する。

このような方法で R-GARCH モデルのパラメータを推定することが出来る。

3.2 R-SV モデルの推定

ここでも、分かりやすくするため R-SV モデルを再掲する。

$$y_t = \varepsilon_t = \sigma_t z_t \quad (3.10)$$

$$\ln \sigma_t^2 = \omega + \varphi \ln \sigma_{t-1}^2 + \eta_t \quad (3.11)$$

$$\ln RV_t = \xi + \ln \sigma_t^2 + u_t \quad (3.12)$$

$$z_t \sim i.i.d. (0,1) \quad \eta_t \sim i.i.d. (0, \sigma_\eta^2) \quad u_t \sim i.i.d. (0, \sigma_u^2) \quad (3.13)$$

$$\ln \sigma_1^2 \sim i.i.d. N\left(\frac{\omega}{1-\varphi}, \frac{\sigma_\eta^2}{1-\varphi^2}\right)$$

疑似最尤法の場合、(3.9) の 3 つの誤差項が正規分布であると仮定され、次のようになる。

$$\begin{pmatrix} z_t \\ \eta_t \\ u_t \end{pmatrix} \sim i.i.d. N \begin{pmatrix} 1 & \rho\sigma_\eta & 0 \\ \rho\sigma_\eta & \sigma_\eta^2 & 0 \\ 0 & 0 & \sigma_u^2 \end{pmatrix} \quad (3.14)$$

先ほど述べたように、(3.10) からこの式はレバレッジ効果を含めた、R-SV モデルになっている。そして、尤度関数は次のような式になる。

$$L(\theta) = \int_0^\infty \cdots \int_0^\infty \left[\prod_{t=1}^T f(\varepsilon_t | \sigma_t^2) \right] f(\sigma_1^2 | \theta) \left[\prod_{t=1}^T f(\sigma_t^2 | \sigma_{t-1}^2; \theta) \right] d\sigma_1^2 \cdots d\sigma_T^2 \quad (3.15)$$

$$f(\varepsilon_t | \sigma_t^2) = \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(-\frac{\varepsilon_t^2}{2\sigma_t^2}\right) \quad (3.16)$$

$$f(\sigma_t^2 | \sigma_{t-1}^2; \theta) = \frac{1}{\sqrt{2\pi\sigma_\eta^2\sigma_t^2}} \exp\left(-\frac{(\ln\sigma_t^2 - \omega - \phi\ln\sigma_{t-1}^2)^2}{2\sigma_\eta^2}\right) \quad (3.17)$$

$$f(\sigma_1^2 | \theta) = \frac{1}{\sqrt{2\pi\sigma_\eta^2/(1-\phi^2)\sigma_1^2}} \exp\left(-\frac{(\ln\sigma_1^2 - \omega/1 - \phi)^2}{2\sigma_\eta^2/(1-\phi^2)}\right) \quad (3.18)$$

この尤度関数は解析的に解くことが不可能なので、様々な推定手法¹が考えられた。今回は疑似最尤法を用いることにする。この場合の R-SV モデルの推定方法は次のような手法で説明される。

R-SV モデルの疑似最尤推定

- ① (3.10)、(3.11) 式を状態空間モデルに変換する。
- ② カルマンフィルタを用いて、パラメータ $(\omega, \phi, \sigma_\eta^2)$ を推定する。
- ③ 平滑化を用いて、 σ_t^2 を計算する。
- ④ σ_t^2 と RV が与えられる。
→ u_t の尤度関数から、最尤法によりパラメータ (ξ, σ_u^2) を推定する。

カルマンフィルタや平滑化などの R-SV の詳しい推定方法の説明は補論 A で述べることにする。この方法のメリットは、MCMC 等の推定法と比べ、推定時間が短いことである。デメリットとしては、推定値の効率性が低いことがあげられる。というのも、疑似最尤法において、 z_t は標準正規分布に従っているが、実際は正規分布よりも裾が厚いことが知られており、正規分布よりも裾の厚い分布を仮定し推定する方が効率的と考えられるが、疑似最尤法の (3.16)~(3.18) の尤度関数は誤差項が正規分布でないと積分を含むため計算するのが難しい。本稿では推定効率の低下が想定されるものの、より簡便な方法である疑似最尤推定法を用いる。

¹ R-SV モデルの MCMC に関しては、大森・渡部(2013)を参照のこと

3.3 ARFIMAX モデルの推定

ARFIMAXモデルの推定は、渡部、佐々木 (2006, p.51) を参考にした。基本的には疑似最尤法を用いて推定する。

$$(1-L)^d \{ \ln(RV_t) - \mu_0 - \mu_1 |R_{t-1}| - \mu_2 D_{t-1}^- |R_{t-1}| \} = (1 + \theta L) u_t \quad (3.19)$$

の両辺に $(1 + \theta L)^{-1}$ をかけると、

$$(1 + \theta L)^{-1} (1-L)^d \{ \ln(RV_t) - \mu_0 - \mu_1 |R_{t-1}| - \mu_2 D_{t-1}^- |R_{t-1}| \} = u_t \quad (3.20)$$

のように次数無限大のAR model となる。

このままでは推定できないので、左辺に漸化式を用いると、

$$(1-L)^d = 1 - \sum_{j=1}^{\infty} \Phi_j L^j \quad (3.21)$$

$$\begin{aligned} \Phi_1 &= d + \theta, & \Phi_j &= \alpha_j - \theta \Phi_{j-1}, \\ \alpha_1 &= d, & \alpha_j &= \frac{j-d-1}{j} \alpha_{j-1} \quad (j \geq 2) \end{aligned}$$

と表すことができる。この証明については補論にて行うものとする。ここで、

$$\forall t \in Z, \forall t \leq 0, E \{ \{ \ln(RV_t) - \mu_0 - \mu_1 |R_{t-1}| - \mu_2 D_{t-1}^- |R_{t-1}| \} \} = 0$$

と仮定すると、(A)(B)式を合わせることで、 (u_1, u_2, \dots, u_T) を計算することができる。ここで、 (u_1, u_2, \dots, u_T) は、 $u_t \sim i.i.d. N(0, \sigma_u^2)$ であることから、次のように対数尤度を求めることができる。

$$\ln L = -\frac{T}{2} \ln(2\pi\sigma_u^2) - \frac{1}{2\sigma_u^2} \sum_{t=1}^T u_t^2 \quad (3.22)$$

右辺において、 $\min(\sum_{t=1}^T u_t^2)$ となるような、未知パラメータ $(d, \mu_0, \mu_1, \mu_2, \theta)$ を最尤法により求め、それを用いて (u_1, u_2, \dots, u_T) を計算し、

$$\sigma_u^2 = \frac{1}{T} \sum_{t=1}^T u_t^2 \quad (3.23)$$

このように標本分散として σ_u^2 を推定することができる

4. 分析結果

これまで、高頻度データである RV を取り入れた 3 つのボラティリティ変動モデルの説明と疑似最尤法の推定方法について述べてきたが、この章では、実際にデータを用いて分析した結果を述べていく。具体的な分析の流れとしては、

以下の表 4-1 に書いてある。まず、各株価指数のリターン、RV、RK などの分析を行い、その後、3つのモデルのパラメータを推定する。そして、その推定パラメータを用いて、ボラティリティ予測を行う。

4.1 データ

本稿ではオックスフォード大学の Oxford-Man Institute of Quantitative Finance の Realized Library (<http://realized.oxford-man.ox.ac.uk/>) のデータを使用した。本稿では世界各国の株価指数を分析対象とした。具体的な 14 つの株価指数に関しては、表 4-2 に載せている。リーマンショックが起こる以前である 2000 年 1 月 3 日～2006 年 12 月 29 日までのリターンと RV、RK のデータを用いた。また、RV と RK はマイクロストラクチャ・ノイズの影響を考慮して 5 分おきのデータを用いた。RK というのは、第 2 章で説明した RV に存在するマイクロストラクチャ・ノイズによる影響を抑えるために変換したものであり、次のように計算される。

$$RK = \gamma_0 + \sum_{h=1}^H k(x)(\gamma_h + \gamma_{-h}) \quad (4.1)$$

γ_0 は、RV の値、 $\gamma_h = \sum_{i=1}^n r_{t-1+i/n} r_{t-1+(i-h)/n}$ であり、 $k(x)$ の定義域は $[0,1]$ である。そして、そのデータから、R-GARCH、R-SV、ARFIMAX モデルの各パラメータを推定し、ボラティリティの予測を行った。図 4-1、4-2 は、各国株価指数のリターンと RV、RK が時系列プロットされている。そして、各国株価指数のリターンと RV とその対数値の平均、標準偏差、歪度、尖度、10 次までの Ljung-Box 統計量²が表 4-3 に書かれている。Ljung-Box 統計量とは、1 次から求めたいラグまでの自己相関が全て 0 かどうかを調べるために使われる統計量であり、全ての自己相関が 0 であることを帰無仮説として検定している。表 4-3 から、RV、RK の Ljung-Box 統計量は、有意水準 1% で棄却されているので、10 次までの自己相関が存在していることが分かる。また、尖度に関しては、正規分布の尖度は 3 であることが分かっている。しかし、全ての株価指数のリターンの尖度が 3 を超えているため、正規分布よりも裾が厚いことが分かる。歪度に関しても、0 から乖離しているため、ある程度の分布の非対称性が生じていることが分かる。

4.2 パラメータの分析結果

次に疑似最尤法によって各モデルのパラメータを推定した。表 4-4 は、R-GARCH モデルの推定パラメータ値、表 4-5 には R-SV モデルの推定パラメータ値、表 4-6 には ARFIMAX の推定パラメータ値が各々書かれている。まず、R-GARCH に関してだが、次数の p, q の決定においては、 $p = 1 \sim 2, q = 1 \sim 2$ の全 4 種類の組み合わせに対して AIC) を用いて決定した。R-GARCH の次数はほとんどのケースにおいて GARCH(1,2) が選ばれ、R-GARCH(2,1)、R-GARCH(1,1)、R-GARCH(2,2) の順で多かった。次にパラメータに関してであるが、表の 4-4 からすべてのケースで τ_1 がマイナスで γ がプラスであり、レバレッジ効果が存在することがわかった。また、RV と RK のそれぞれの場合において ξ の値を比較してみると、14 の指標のうち 10 の指標が RV より RK の ξ の方が 0 に近づいているため、マイクロストラクチャ・ノイズの影響を RK が抑えていることが分かる。 ξ の値にマイナスが多い理由としては、取引外の時間を考慮しないために、1 日のボラティリティを過小評価していることが考えられる。持続性に関しては、R-GARCH (1,1) を例に出すと、(2.19) 式に (2.18) 式を代入すると

$$\ln \sigma_t^2 = \omega + \beta \ln \sigma_{t-1}^2 + \gamma(\xi + \varphi \ln \sigma_{t-1}^2 + \tau_1 z_{t-1} + \tau_2(z_{t-1}^2 - 1) + u_{t-1}) \quad (4.2)$$

$$= \omega + \gamma\xi + (\beta + \gamma\varphi) \ln \sigma_{t-1}^2 + \gamma(\tau_1 z_{t-1} + \tau_2(z_{t-1}^2 - 1) + u_{t-1}) \quad (4.3)$$

² 詳しくは沖本 (2010, p.13)を参照

表 4-1. The step of analysis

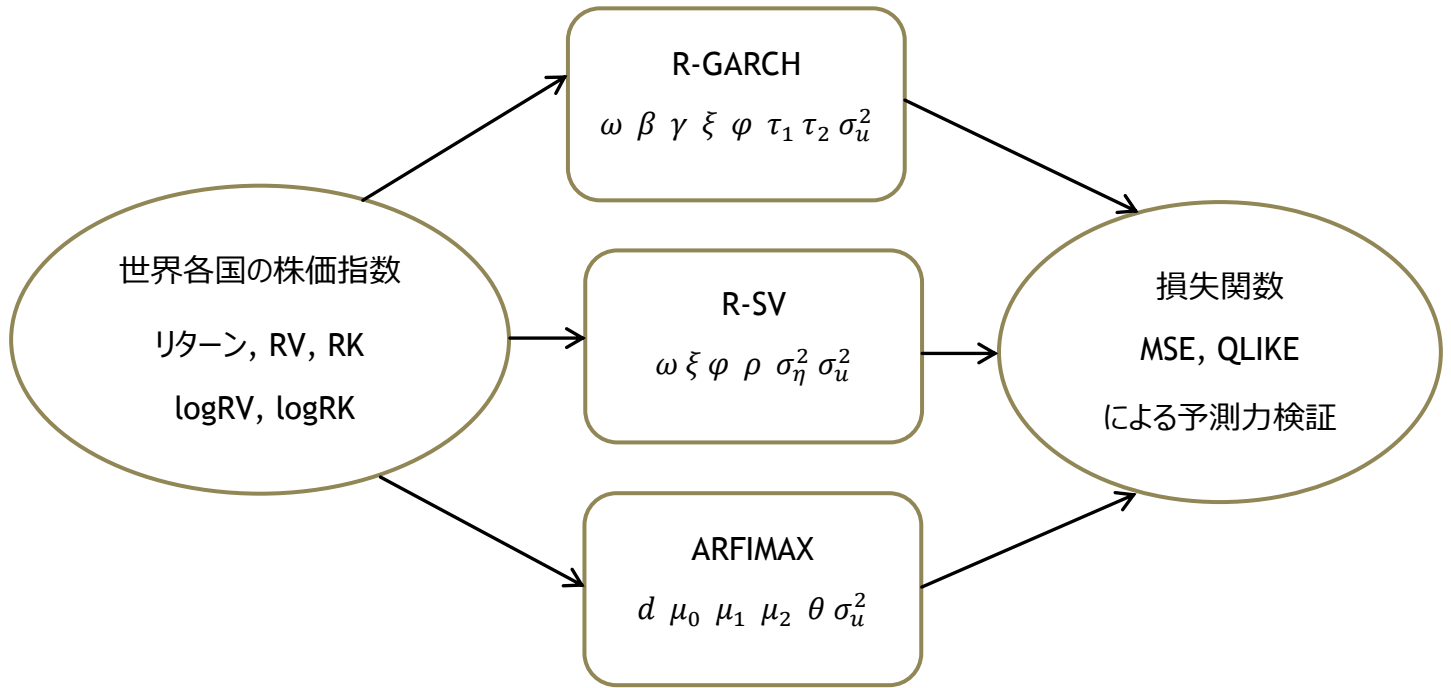


表 4-2. The analyzed index³

略称	名称	国
AEX	AEX指数	オランダ
AORD	ASX指数	オーストラリア
BVSP	ボベスパ指数	ブラジル
DJI	NYダウ	アメリカ合衆国
FCHI	CAC40指数	フランス
FTSEMIB	FTSE MIB指数	イギリス+イタリア
GDAXI	DAX指数	ドイツ
HSI	香港ハンセン指数	香港
IXIC	ナスダック総合	アメリカ合衆国
KS11	韓国総合	韓国
MXX	IPC指数	メキシコ
N225	日経225	日本
SSMI	SMI指数	スイス
STOXX50	ユーロダウ	欧州

³ Yahoo! Finance (<http://finance.yahoo.co.jp/>)より引用

図 4-1. 日次リターン

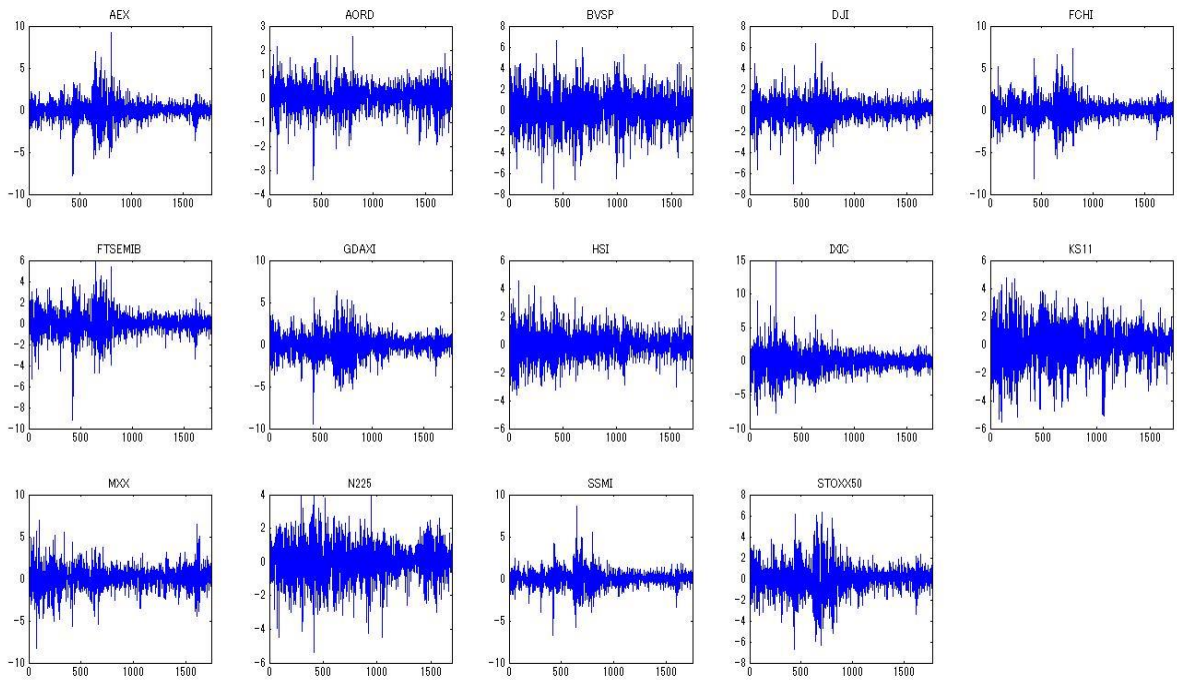
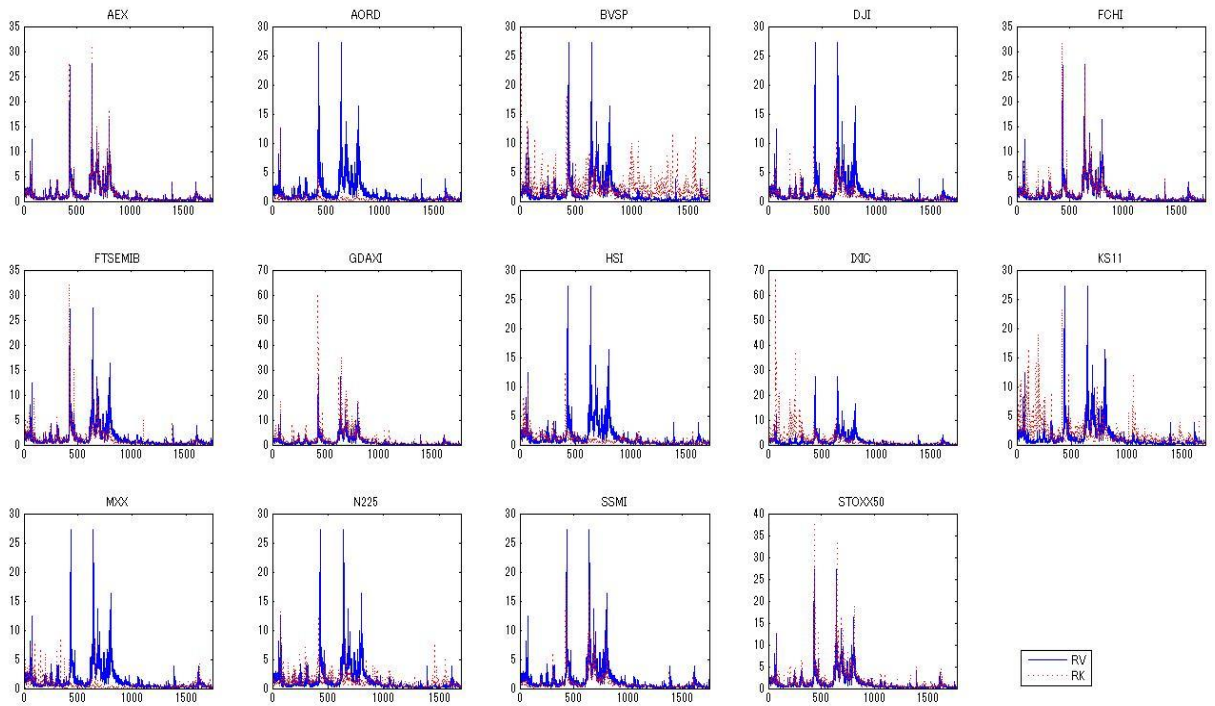


図 4-2. RV と RK



		平均	標準偏差	標準誤差	歪度	尖度	LB (10)	
AEX	Return	-0.0488	1.2419	0.0007	0.021	9.815	42.8	***
	RV	1.3295	2.1157	0.0012	5.180	43.850	6783.1	***
	RK	1.3578	2.2867	0.0013	5.564	49.567	6389.5	***
	logRV	-0.3286	1.0363	0.0006	0.507	2.921	11396.1	***
	logRK	-0.3433	1.0609	0.0006	0.521	2.977	11404.7	***
AORD	Return	0.0254	0.5868	0.0003	-0.420	5.154	5.0	***
	RV	0.2654	0.4458	0.0003	16.114	408.953	770.6	***
	RK	0.2764	0.4481	0.0003	15.244	372.600	853.6	***
	logRV	-1.7107	0.8051	0.0005	0.417	3.832	3582.2	***
	logRK	-1.6660	0.8065	0.0005	0.374	3.798	3653.5	***
BVSP	Return	0.0305	1.7661	0.0010	-0.228	3.704	13.6	***
	RV	1.9678	1.6078	0.0010	4.132	35.199	1423.8	***
	RK	2.1871	1.8202	0.0011	4.448	43.069	1507.9	***
	logRV	0.4659	0.6275	0.0004	0.224	3.812	2188.6	***
	logRK	0.5657	0.6370	0.0004	0.203	3.802	2314.0	***
DJI	Return	0.0145	1.0749	0.0006	-0.090	6.996	13.7	***
	RV	1.0099	1.3680	0.0008	4.746	35.334	4718.3	***
	RK	0.9016	1.0456	0.0006	3.884	24.594	6910.9	***
	logRV	-0.4521	0.9084	0.0005	0.370	3.249	8115.5	***
	logRK	-0.4949	0.8450	0.0005	0.356	3.077	9126.7	***
FCHI	Return	-0.0412	1.2490	0.0007	0.029	7.605	30.5	***
	RV	1.4122	2.0976	0.0012	6.303	67.125	5971.1	***
	RK	1.4314	2.1613	0.0012	6.145	63.462	5739.7	***
	logRV	-0.2059	1.0174	0.0006	0.264	2.704	11712.9	***
	logRK	-0.2176	1.0432	0.0006	0.255	2.672	11565.2	***
FTSEMIB	Return	-0.0528	1.1437	0.0007	-0.403	8.682	21.4	**
	RV	1.0744	1.8139	0.0010	8.663	127.253	4480.4	***
	RK	1.0750	1.7716	0.0010	7.209	90.274	4866.1	***
	logRV	-0.5063	1.0112	0.0006	0.459	2.770	11348.4	***
	logRK	-0.5679	1.0895	0.0006	0.350	2.576	11015.9	***
GDAXI	Return	-0.0670	1.4211	0.0008	-0.253	5.925	25.2	**
	RV	2.1956	3.2585	0.0018	5.160	47.124	6262.2	***
	RK	2.1228	3.3661	0.0019	6.469	76.822	5401.9	***
	logRV	0.1689	1.0933	0.0006	0.198	2.645	11565.1	***
	logRK	0.1289	1.0830	0.0006	0.265	2.743	11614.4	***
HSI	Return	-0.0197	0.9856	0.0006	-0.146	4.903	24.6	**
	RV	0.9697	1.0958	0.0006	6.117	69.439	3663.1	***
	RK	0.8542	0.9680	0.0006	5.893	64.975	4322.6	***
	logRV	-0.3699	0.7937	0.0005	0.243	3.135	8057.2	***
	logRK	-0.5063	0.8053	0.0005	0.258	3.065	8421.5	***
IXIC	Return	-0.0706	1.6857	0.0010	0.364	9.021	41.1	***
	RV	1.7813	2.6751	0.0015	6.312	71.741	4937.3	***
	RK	1.9398	3.1536	0.0018	8.189	126.642	4296.6	***
	logRV	-0.0066	1.0473	0.0006	0.303	2.628	11328.5	***
	logRK	0.0533	1.0613	0.0006	0.338	2.674	11443.8	***
KS11	Return	-0.0254	1.4155	0.0008	-0.324	4.061	37.3	***
	RV	1.8747	2.0107	0.0012	4.021	30.641	4029.5	***
	RK	1.9136	2.0576	0.0012	3.433	20.895	4833.8	***
	logRV	0.2751	0.8127	0.0005	0.264	3.081	8245.9	***
	logRK	0.2775	0.8387	0.0005	0.215	3.068	8785.6	***
MXX	Return	0.0651	1.3799	0.0008	-0.109	5.901	40.3	***
	RV	0.6675	1.1797	0.0007	7.471	86.117	948.8	***
	RK	0.5212	0.7268	0.0004	5.271	42.490	2043.3	***
	logRV	-0.9499	0.9421	0.0005	0.623	3.543	5555.3	***
	logRK	-1.1000	0.8740	0.0005	0.562	3.332	6343.3	***
N225	Return	-0.0454	1.1160	0.0007	-0.158	4.148	15.3	***
	RV	1.1447	0.9810	0.0006	4.401	51.276	3480.8	***
	RK	1.1739	1.0334	0.0006	3.633	28.725	3451.1	***
	logRV	-0.1480	0.7786	0.0005	-0.316	3.182	7393.6	***
	logRK	-0.1349	0.7908	0.0005	-0.273	3.208	6876.8	***
SSMI	Return	-0.0197	1.0262	0.0006	0.161	10.519	25.6	**
	RV	0.7800	1.2509	0.0007	5.123	39.388	6845.4	***
	RK	0.8250	1.3867	0.0008	5.772	50.529	8108.9	***
	logRV	-0.7718	0.8940	0.0005	0.982	3.820	11586.9	***
	logRK	-0.7305	0.9045	0.0005	0.969	3.830	11821.2	***
STOXX50	Return	-0.0301	1.3769	0.0008	-0.081	5.937	35.0	**
	RV	1.6979	2.6229	0.0015	5.313	46.735	6504.7	***
	RK	1.6850	2.6741	0.0015	5.668	52.982	6395.5	***
	logRV	-0.0733	1.0514	0.0006	0.324	2.934	9668.0	***
	logRK	-0.0942	1.0604	0.0006	0.343	2.904	10361.7	***
Average	Return	-0.0204	1.2478	0.0007	-0.117	6.598	26.5	***
	RV	1.2979	1.7517	0.0010	6.361	82.869	4344.2	***
	RK	1.3046	1.7786	0.0010	6.202	79.116	4651.5	***
	logRV	-0.3295	0.9156	0.0005	0.347	3.161	8696.0	***
	logRK	-0.3450	0.9214	0.0005	0.338	3.123	8910.6	***

LB 統計量の有意水準は p 値において、*...0.05、**...0.01、***...0.0001 で棄却される。

表 4-3. RV の基本データ

表 4-4. R-GARCH モデルの推定結果

		ω	β_1	β_2	γ_1	γ_2	ξ	ϕ	τ_1	τ_2	σ_u^2
AEX	RV	0.1723	0.2142		0.3626	0.4099	-0.9303	0.9384	-0.1085	0.1159	0.201
	RK	0.1723	0.2142		0.3626	0.4099	-0.9018	0.9415	-0.1081	0.1275	0.2026
AORD	RV	-0.051	0.2689	0.4002	0.3319		-1.2268	0.8811	-0.0157	0.0892	0.2881
	RK	-0.051	0.2689	0.4002	0.3324		-1.1354	0.8874	-0.0153	0.0925	0.2813
BVSP	RV	0.1696	0.2142		0.3623	0.4096	-2.601	0.7369	-0.0577	0.0767	0.2288
	RK	0.1692	0.2142		0.3623	0.4096	-2.4588	0.7499	-0.64	0.0875	0.2279
DJI	RV	0.0965	0.2626		0.2142	0.2498	-1.3406	1.3885	-0.4456	0.065	0.2913
	RK	0.1728	0.2142		0.3626	0.41	-1.2131	0.9103	-0.0714	0.0919	0.191
FCHI	RV	-0.0125	0.2966	0.2937	0.2689	0.3038	-0.7546	0.6766	-0.0131	0.0021	0.2174
	RK	-0.0127	0.2939	0.2909	0.2689	0.5766	-0.6926	0.6945	-0.0141	0.0029	0.2262
FTSEMIB	RV	0.0275	0.2287		0.4738	0.4356	-0.6135	0.8041	-0.0278	0.011	0.1855
	RK	0.0287	0.2292		0.4751	0.4367	-0.6146	0.8029	-0.0299	0.0113	0.2294
GDAXI	RV	-0.0488	0.2689	0.4218	0.324		-0.0264	0.9426	-0.0762	0.0536	0.1907
	RK	-0.0488	0.2689	0.4218	0.3236		-0.0425	0.9426	-0.079	0.0544	0.1807
HSI	RV	0.0956	0.2627		0.2142	0.2497	-1.2553	1.3713	-0.5048	0.245	0.2232
	RK	0.0971	0.2626		0.2142	0.2497	-1.1848	1.3869	-0.48	0.245	0.214
IXIC	RV	0.1713	0.2142		0.3624	0.4098	-0.8415	0.9501	-0.0507	0.0814	0.202
	RK	0.1711	0.2142		0.3624	0.4098	-0.795	0.9544	-0.0562	0.0851	0.1999
KS11	RV	-0.0468	0.3324	0.3581	0.2689		-0.1346	1.0963	-0.0578	0.2265	0.1817
	RK	-0.0468	0.3324	0.3581	0.2689		-0.0439	1.1078	-0.0661	0.2274	0.1759
MXX	RV	-0.0074	0.2785	0.3752	0.2689		-0.7535	1.1671	-0.0053	0.3013	0.3377
	RK	-0.0077	0.2689	0.3691	0.2773		-0.797	1.183	-0.0152	0.2464	0.271
N225	RV	-0.1157	0.7714		0.2992		0.115	0.7528	-0.0027	0.0032	0.1933
	RK	-0.1121	0.7601		0.307		0.0796	0.7677	-0.0033	0.0039	0.1789
SSMI	RV	0.1737	0.2142		0.3629	0.4103	-0.9039	0.9453	-0.0738	0.0936	0.1294
	RK	0.1736	0.2142		0.3629	0.4103	-0.9259	0.9422	-0.0766	0.102	0.1376
STOXX50	RV	-0.1683	0.8035		0.2028		0.4102	0.9389	-0.0517	0.0321	0.252
	RK	-0.1291	0.78		0.2662		0.2611	0.818	-0.0274	0.0071	0.2249
Average	RV	0.0326	0.3308	0.3698	0.3084	0.3598	-0.7755	0.9707	-0.1065	0.0998	0.2230
	RK	0.0412	0.3240	0.3680	0.3247	0.4141	-0.7475	0.9349	-0.1202	0.0989	0.2101

表 4-5. R-SV モデルの推定結果

		ω	ϕ	σ_η^2	ρ	ξ	σ_u^2
AEX	RV	-0.0020	0.9939	0.0146	-0.1551	-0.1741	0.3556
	RK					-0.1888	0.3564
AORD	RV	-0.0226	0.9820	0.0151	-0.1312	-0.5327	0.4110
	RK					-0.4880	0.4111
BVSP	RV	0.0229	0.9770	0.0125	-0.0787	-0.2850	0.3769
	RK					-0.1851	0.3732
DJI	RV	-0.0009	0.9975	0.0046	-0.0117	-0.1507	0.3614
	RK					-0.1934	0.2846
FCHI	RV	-0.0014	0.9932	0.0147	-0.1135	-0.1257	0.2596
	RK					-0.1374	0.2727
FTSEMIB	RV	-0.0014	0.9942	0.0155	-0.3649	-0.4025	0.3163
	RK					-0.4641	0.3558
GDAXI	RV	0.0010	0.9944	0.0144	-0.0101	-0.1116	0.2948
	RK					-0.1516	0.2943
HSI	RV	-0.0004	0.9982	0.0034	-0.1061	0.0827	0.3446
	RK					-0.0537	0.2976
IXIC	RV	0.0008	0.9985	0.0029	-0.1706	-0.5319	0.3228
	RK					-0.4720	0.3105
KS11	RV	0.0009	0.9976	0.0050	-0.2314	-0.2417	0.3364
	RK					-0.2394	0.3273
MXX	RV	0.0055	0.9836	0.0201	-0.0396	-1.3348	0.4225
	RK					-1.4848	0.3379
N225	RV	-0.0008	0.9904	0.0122	-0.1457	-0.2069	0.2640
	RK					-0.1938	0.2827
SSMI	RV	-0.0078	0.9851	0.0320	-0.0424	-0.3228	0.2326
	RK					-0.2815	0.2323
STOXX50	RV	0.0004	0.9946	0.0121	-0.0282	-0.2643	0.3774
	RK					-0.2853	0.3451
Average	RV	-0.0004	0.9914	0.0128	-0.1164	-0.3287	0.3340
	RK					-0.3442	0.3201

という式になるが、このときの $\beta + \gamma\phi$ がボラティリティの持続性を表している。つまり、 $|\beta + \gamma\phi| < 1$ のときにボラティリティは定常性を満たしている。そこで、実際のデータを代入してみると、ほとんどの指標はこの値が 1 以下であるが、1 に近いために、定常であるが持続性があることが分かる。

次に R-SV の推定結果を見る。最初の 4 つ ($\omega, \phi, \sigma_{\eta}^2, \rho$) のパラメータは、データはリターンのみ所与であったために、RV と RK で推定値の違いは現れない。全てのケースで ϕ の値は 0.9 を超えており高い持続性があることを示している。また、全てのケースで ρ の値はマイナスになっているため、R-GARCH モデルと同じく、レバレッジ効果が存在することが分かった。残りの 2 つの推定パラメータのうち、 ξ に注目すると、半分のケースで RK によって、マイクロストラクチャ・ノイズの影響を抑えることが分かった。しかし、これは R-GARCH の結果と比べると、RK がマイクロストラクチャ・ノイズを減少させたとは言い切れないだろう。平均を見てみると、RK の ξ の値の方が RV よりも 0 から乖離しているため、当初の前提よりも思わしくない結果であった。また、これも R-GARCH の結果と同じであるが、 ξ の値がマイナスなのは、取引外の時間を考慮していないために過小評価している。最後に ARFIMAX モデルの推定パラメータについて述べる。ARFIMAX の推定パラメータは表 4-6 にある。まず、 μ_2 の値をみってみると、HSI の RK 以外は正の値であり、レバレッジ効果が生じていることが分かった。またパラメータ d に関しては 9 つのケースにおいて $0.5 < d < 1$ に存在しているため、それらの RV は非定常な長期記憶過程であることが分かった。

表 4-6. ARFIMAX モデルの推定結果

		d	μ_0	μ_1	μ_2	θ	σ_{η}^2
AEX	RV	0.5748	0.1886	0.0377	0.0948	-0.2178	0.2150
	RK	0.5708	0.2117	0.0339	0.1013	-0.2271	0.2076
AORD	RV	0.4074	-1.2957	-0.0450	0.2254	-0.2336	0.3769
	RK	0.4149	-1.2387	-0.0451	0.2008	-0.2232	0.3753
BVSP	RV	0.4298	0.5808	0.0309	0.0676	-0.0898	0.2417
	RK	0.4335	0.7095	0.0266	0.0734	-0.0676	0.2405
DJI	RV	0.5095	-0.0029	-0.0637	0.1405	-0.2178	0.2674
	RK	0.5447	0.0120	-0.0476	0.1190	-0.2128	0.1956
FCHI	RV	0.5683	0.4695	0.0141	0.0986	-0.2486	0.1922
	RK	0.5627	0.4877	0.0202	0.1098	-0.2674	0.2086
FTSEMIB	RV	0.5627	0.4877	0.0202	0.1098	-0.2674	0.2086
	RK	0.5538	0.4374	0.0638	0.0969	-0.2580	0.2420
GDAXI	RV	0.5545	1.0425	0.0187	0.0684	-0.1953	0.2294
	RK	0.5644	1.0590	0.0254	0.0655	-0.2177	0.2223
HSI	RV	0.4737	0.4117	0.0288	0.0074	-0.2506	0.2100
	RK	0.4929	0.4063	0.0428	-0.0003	-0.2617	0.2023
IXIC	RV	0.5381	0.7392	0.0395	0.1060	-0.1875	0.1882
	RK	0.5486	0.9729	0.0451	0.1053	-0.2041	0.1866
KS11	RV	0.5500	1.2784	0.0797	0.0251	-0.2149	0.1895
	RK	0.5640	1.3996	0.0773	0.0320	-0.1994	0.1814
MXX	RV	0.4166	-0.5204	0.0808	0.0617	-0.1624	0.3933
	RK	0.4304	-0.6268	0.0602	0.0639	-0.1596	0.3079
N225	RV	0.4975	0.0865	0.0085	0.0912	-0.1450	0.2011
	RK	0.4912	0.0627	0.0246	0.0921	-0.1525	0.2199
SSMI	RV	0.6114	-0.4313	0.0290	0.0705	-0.2595	0.1441
	RK	0.6072	-0.3164	0.0373	0.0797	-0.2605	0.1517
STOXX50	RV	0.5185	0.3132	0.0116	0.0939	-0.2030	0.2962
	RK	0.5370	0.3537	0.0166	0.0948	-0.2156	0.2665
Average	RV	0.5152	0.2391	0.0208	0.0901	-0.2067	0.2395
	RK	0.5226	0.2808	0.0272	0.0882	-0.2091	0.2291

4.3 ボラティリティの予測

最後に各指標のボラティリティの予測を R-GARCH と R-SV、ARFIMAX の 3 つのモデルに加えて、そのモデルでの RV と RK の計 6 つで行う。まず、各モデルの予測なのだが、R-GARCH に関しては、パラメータの値と $t - 1$ 期までのデー

タが存在すれば、 t 期の予測値は求めることが出来る。2 つ目に R-SV モデルに関しては、大森、渡部 (2013) を参考にすると、1 期先の予測の $\ln\sigma_{t+1}^2$ の分布の平均 $\hat{\mu}_{t+1}$ と分散 $\hat{\sigma}_{t+1}^2$ は次のように与えられる。

$$\hat{\mu}_{t+1} = \omega + \varphi \ln\sigma_t^2 + \rho\sigma_\eta |y_t| \exp\left(-\frac{\ln\sigma_t^2}{2}\right) \quad (4.4)$$

$$\hat{\sigma}_{t+1}^2 = \sigma_\eta^2(1 - \rho^2) \quad (4.5)$$

そして、ボラティリティにするために、 $\exp(\ln\sigma_{t+1}^2)$ に変換することで、1 期先のボラティリティを予測することが出来る。

最後に ARFIMAX のモデルの 1 期先予測は、渡部、佐々木 (2006, p.57) に従うと、誤差項 u_t の分布を仮定する必要がある。今回は、 u_t を正規分布に仮定して推定しているので、予測の際も正規分布として仮定する。そうすると、 $\widehat{RV}_{t+1|t}$ は次のような式になる。

$$= \exp \left[\begin{array}{l} \widehat{RV}_{t+1|t} \\ \mu_0 + (\mu_1 + \mu_2 D_t^-) |y_t| - \sum_{k=1}^t \frac{d(d-1)\cdots(d-k-1)}{k} (-1)^k \\ \{\ln RV_{t-k} - \mu_0 - (\mu_1 + \mu_2 D_{t-k}^-) |y_{t-k}|\} + \theta \hat{u}_t + \frac{1}{2} \hat{\sigma}_u^2 \end{array} \right] \quad (4.6)$$

ここで、 \hat{u}_t は ARFIMAX モデルの誤差項 u_t の残差、 $\hat{\sigma}_u^2$ は残差分散を表す。

また、各モデルによって予測されたボラティリティを評価するために、次の 2 つの損失関数を定義する。

$$MSE = \frac{1}{T} \sum_{t=1}^T (\sigma_{t+1}^2 - \hat{\sigma}_{t+1|t}^2)^2 \quad (4.7)$$

$$QLIKE = \frac{1}{T} \sum_{t=1}^T \log \hat{\sigma}_{t+1|t}^2 + \frac{\sigma_{t+1}^2}{\hat{\sigma}_{t+1|t}^2} \quad (4.8)$$

σ_{t+1}^2 は、 $t+1$ 期の真のボラティリティの値であり、 $\hat{\sigma}_{t+1|t}^2$ は t 期の情報を用いて予測するボラティリティの予測値である。ところが、 $t+1$ 期の真のボラティリティの値というのは、観察されないため、従来ではリターン²の 2 乗を代理変数として用いてきた。しかし、RV が登場してきたことで、リターン²の代わりに RV を用いる先行論文が増えてきた。本稿でもこれに従って、真のボラティリティの値は $t+1$ 期の RV の値を使う。Patton (2011, p.246) によると真のボラティリティの代理変数として真のボラティリティの不偏推定量を用いた場合、上記の (4.4) と (4.5) 式の期待損失の大きさの順番は、真のボラティリティと代理変数どちらを用いても変わらない。この結果は推定効率とは関係ないので、5 分間隔の収益率により計算した RV はマイクロストラクチャノイズの影響がそれほどなく、真のボラティリティの不偏推定量に近い代理変数として RV を用いることが正当化される。この 2 つの損失関数が小さいモデルが予測精度の高いモデルとして評価される。

本稿では、先ほどの予測式を用いた 20 期先における RV と各モデルのボラティリティの予測値を損失関数の大小によって評価を行う。具体的には、各株価指数の標本数を T とおく。そこで、リターンを y_t 、RV を RV_t ($t = 1, \dots, T$) とすると、標本 $\{y_1, \dots, y_{T-219}, RV_1, \dots, RV_{T-219}\}$ から、予測値である $\hat{\sigma}_{T-199}^2$ の期待値を先ほどの式を用いて計算する。次に標本 $\{y_2, \dots, y_{T-218}, RV_2, \dots, RV_{T-218}\}$ から、予測値である $\hat{\sigma}_{T-198}^2$ の期待値を計算する。この作業を繰り返す。

返して、標本 $\{y_{T-200}, \dots, y_{T-20}, RV_{T-200}, \dots, RV_{T-20}\}$ から、予測値である $\hat{\sigma}_T^2$ の期待値を計算して終わりとする。そこで得られた $\{\hat{\sigma}_{T-199}^2, \dots, \hat{\sigma}_T^2\}$ から MSE と QLIKE の計算を行った。

これまで述べた手法を用いて求めた損失関数の値の結果が表 4-7 に書かれている。また、表 4-8 では MSE と QLIKE において最も値が低かったモデルをまとめた。まず、全体の株価指数の平均を見てみると、全体では MSE、QLIKE とともに R-SV モデルが優れていることが分かった。しかし、細かく見てみると、R-SV モデルと R-GARCH モデルの予測力は僅差であり、14 の指標のうち、7 つの指標において R-GARCH モデルが一番優れているという結論になった。一方で、ARFIMAX モデルはどの指標においても他の 2 つのモデルより予測力が劣っており、特に GDAXI や KS11 では他の 2 つもモデルよりも MSE や QLIKE が圧倒的に高いことが分かった。ARFIMAX のパフォーマンスが劣る原因としてあげられるのは、20 期先という予測期間が関係していると考えられる。RV は長期記憶過程に従うことが先ほどの結果によって明らかになったため、我々が今回予測した期間よりも長く予測期間を延ばしてみると短期記憶過程しか捉えることのできない R-SV、R-GARCH モデルよりも ARFIMAX モデルの方が優れたパフォーマンスを見せる可能性があるだろう。次に RV を用いたときと RK を用いたときの予測力の比較を述べていく。全体で見ると、RK を用いるよりも RV の方が予測力に置いては優れていた。しかし、AORD、N225、STOXX50 に関しては、RV よりも RK を用いたモデルが一番すぐれているという結果になった。渡部、大森 (2013, p.294-295) は、「RK を用いた方が RV よりも予測力に優れているならば、それは、RV のバイアスは R-SV モデルにおいての ξ だけでは完全捉えきれない。」と述べている。つまり、今回の結果でも前述した 3 つの株式指標においては RV のバイアスを R-SV、R-GARCH モデルでの ξ だけでは完全に捉えることが出来ていないことが判明した。また、2 つの損失関数の評価であるが、4 つの指標に置いて評価の違いが生じた。

以上の分析は、2 つほど問題が考えられる。1 つは分布の問題である。我々が 3 つのモデルを推定した際に、全て誤差項が正規分布であると仮定した上でパラメータを推定した。しかしながら、分析結果でこれらの分布は全て正規分布よりも裾が厚いことが確認された。そこで、リターンの分布は正規性を持つのかどうかを JB (Jarque-Bera) 統計量によって分析を行った。JB 統計量⁴とは、与えられたデータが正規分布に従うことを帰無仮説とする統計量である。表 4-9 には、その結果を載せている。表 4-9 から全ての指標が有意水準 1% で棄却されており、リターンの分布に正規性は存在しないことが分かる。つまり、今回の我々の論文はリターンの分布に正規分布を仮定して推定したため、推定パラメータの精度が低い。この問題を解決するためには、正規分布よりも裾の厚い分布を仮定する必要がある。正規分布よりも裾が厚い分布の例として、t 分布と GED (Generalized Error Distribution) が考えられる。2 つの分布の定義は、以下のように表す。

$$y_t = \mu + \sigma_t z_t \quad (4.9)$$

$$f(z) = \frac{1}{\sqrt{\pi(\nu-2)}} \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{z^2}{\nu-2}\right)^{-\frac{\nu+1}{2}}, \quad \nu > 0 \quad (4.10)$$

$$f(z) = \frac{\nu \exp\left(-\frac{1}{2} \left|\frac{z}{\lambda}\right|^\nu\right)}{2\lambda \left(1 + \frac{1}{\nu}\right) \Gamma\left(\frac{1}{\nu}\right)} \quad 0 < \nu \leq \infty \quad (4.11)$$

$$\lambda = \frac{1}{2^{\frac{1}{\nu}}} \sqrt{\frac{\Gamma\left(\frac{1}{\nu}\right)}{\Gamma\left(\frac{3}{\nu}\right)}} \quad (4.12)$$

⁴ 詳しくは中川 (2009, p.139~144) を参照

z_t を (4.9) 式のようにしたときに、(4.10) 式の t 分布に従うと、自由度 ν が大きいほど標準正規分布に近づくが、自由度が小さいと標準正規分布よりも裾が厚い分布になる。また、(4.11) 式の GED に従うとき、自由度 ν が 2 のときは標準正規分布に従うが、自由度 ν が 2 より大きいと標準正規分布よりも裾が厚く、2 より小さいと標準正規分布よりも裾が薄くなる。このような正規分布よりも裾が厚い分布で最尤法を行うと、推定パラメータの精度が上昇し、予測精度が上がる可能性が大いに考えられる。しかしながら、R-SV モデルの場合、このような分布に従った場合、最尤法を用いて解析することが非常に難しい。そのため、R-SV モデルの場合は、最尤法よりも MCMC を用いて推定する方が精度が高くなることが考えられる。つまり、R-SV モデルに関しては、推定方法におけるパラメータの精度の比較を行う必要がある。しかし、MCMC はパラメータが収束するのに時間を要するなどのデメリットも考えられる。

表 4-7. 各モデルの予測結果

		MSE	QLIKE			MSE	QLIKE			
AEX	R-GARCH	RV	0.1675	0.2631	HSI	R-GARCH	RV	0.1106	0.0944	
		RK	0.2685	0.4177			RK	0.2511	0.2603	
	R-SV	RV	0.1535	0.2581			R-SV	RV	0.1088	0.0590
		RK	0.2427	0.4039			RK	0.2621	0.2408	
	ARFIMAX	RV	0.5428	0.6453			ARFIMAX	RV	0.2145	0.1785
		RK	0.5333	0.6385			RK	0.1916	0.1538	
AORD	R-GARCH	RV	0.0621	-0.1423	IXIC	R-GARCH	RV	0.1172	0.3358	
		RK	0.0590	-0.1959			RK	1.4864	0.8424	
	R-SV	RV	0.0618	-0.1190			R-SV	RV	0.1186	0.3788
		RK	0.0611	-0.1518			RK	1.4388	0.8547	
	ARFIMAX	RV	0.1042	1.2491		ARFIMAX	RV	0.7650	0.6051	
		RK	0.1013	1.0497		RK	1.4643	0.7432		
BVSP	R-GARCH	RV	1.3653	1.4910	KS11	R-GARCH	RV	0.1551	0.5121	
		RK	3.0253	1.6638			RK	0.8416	0.8382	
	R-SV	RV	1.6166	1.5582			R-SV	RV	0.1739	0.4853
		RK	3.1457	1.7108			RK	0.8829	0.8283	
	ARFIMAX	RV	5.0435	1.6927		ARFIMAX	RV	4.9154	1.1435	
		RK	7.4943	1.7621		RK	6.4888	1.2063		
DJI	R-GARCH	RV	0.0907	0.0527	MXX	R-GARCH	RV	0.7644	0.7703	
		RK	0.1707	0.1873			RK	1.2960	0.9980	
	R-SV	RV	0.0496	-0.0487			R-SV	RV	0.2283	0.4529
		RK	0.1652	0.1656			RK	0.9234	0.8007	
	ARFIMAX	RV	0.1564	0.2728		ARFIMAX	RV	1.3850	1.5911	
		RK	0.1435	0.2577		RK	1.4803	2.2894		
FCHI	R-GARCH	RV	0.1631	0.5200	N225	R-GARCH	RV	0.4624	1.1017	
		RK	0.2867	0.5279			RK	0.4437	0.8545	
	R-SV	RV	0.1621	0.4946			R-SV	RV	0.4257	1.0650
		RK	0.2819	0.5029			RK	0.4048	0.8354	
	ARFIMAX	RV	0.9018	0.7136		ARFIMAX	RV	0.7194	0.9755	
		RK	0.8668	0.7157		RK	0.6768	0.9754		
FTSEMIB	R-GARCH	RV	0.0552	0.0201	SSMI	R-GARCH	RV	0.0635	0.0191	
		RK	0.2334	0.2351			RK	0.1073	0.1403	
	R-SV	RV	0.0695	0.0721			R-SV	RV	0.0803	0.0176
		RK	0.2379	0.2457			RK	0.1261	0.1472	
	ARFIMAX	RV	0.3544	0.2733		ARFIMAX	RV	0.1226	0.5109	
		RK	0.3205	0.2678		RK	0.1266	0.4429		
GDAXI	R-GARCH	RV	0.2484	0.5843	STOXX50	R-GARCH	RV	0.7501	1.3001	
		RK	0.5977	0.7667			RK	0.8022	0.7474	
	R-SV	RV	0.2190	0.5866			R-SV	RV	0.3964	0.9714
		RK	0.5721	0.7812			RK	0.5827	0.6822	
	ARFIMAX	RV	4.2334	1.0975		ARFIMAX	RV	1.5472	1.0102	
		RK	4.5761	1.1130		RK	1.3915	1.0070		
				Average	R-GARCH	RV	0.3268	0.4945		
						RK	0.7050	0.5917		
					R-SV	RV	0.2760	0.4451		
						RK	0.6662	0.5748		
					ARFIMAX	RV	1.5004	0.8542		
						RK	1.8468	0.9016		

表 4-8. 最も優れたモデル

	MSE	QLIKE
AEX		R-SV(RV)
AORD		R-GARCH(RK)
BVSP		R-GARCH(RV)
DJI		R-SV(RV)
FCHI		R-SV(RV)
FTSEMIB		R-GARCH(RV)
GDAXI	R-SV(RV)	R-GARCH(RV)
HIS		R-SV(RV)
IXIC		R-GARCH(RV)
KS11	R-GARCH(RV)	R-SV(RV)
MXX		R-SV(RV)
N225		R-SV(RK)
SSMI	R-GARCH(RV)	R-SV(RV)
STOXX50	R-SV(RV)	R-SV(RK)
Average		R-SV(RV)

2つ目の問題点は、RVの問題が考えられる。具体的には、取引時間外のRVを考慮する必要がある。我々の論文では、マイクロストラクチャ・ノイズに関しては、RVとRKを比較していたが、取引時間外の問題は考えていなかった。そのため、RVがより過小評価された可能性がある。そのため、第2章で述べたように、Hansen and Lunde (2005, p.525-554)などでRVを変換しなければならない。またその取引時間外の変換方法も正しい方法というのが存在しないため、試行錯誤する必要がある。

表 4-9. Jarque-Bera test

Index	JB		Index	JB	
AEX	3423.858	***	HIS	264.1101	***
AORD	390.5868	***	IXIC	2668.69	***
BVSP	49.5261	***	KS11	109.9627	***
DJI	1159.3	***	MXX	614.0158	***
FCHI	1564.92	***	N225	100.0576	***
FTSEMIB	2412.186	***	SSMI	4130.49	***
GDAXI	649.8282	***	STOXX50E	637.9532	***

有意水準は p 値において、*...0.1、**...0.05、***...0.01

5. まとめ

今回、我々はボラティリティ変動分析をテーマにこの論文の作成に取り組んだ。そこで、Realized Volatilityの存在に注目して、ボラティリティだけではなく、RVの性質を捉えることが出来るようなモデルを選択した。そのモデルが、R-GARCH、R-SV、ARFIMAXの3つであった。このモデルを用いて、パラメータの分析を用いた結果、今回の分析対象とした株価指数では、全てにボラティリティの持続性とRVの長期記憶性を確認することが出来た。また、ARFIMAXモデルで分析したハンセン指数のRK以外の全てにレバレッジ効果を確認することが出来た。そして、RVとRKでのマイクロストラクチャ・ノイズの影響の比較も行い、結果として、R-GARCHモデルにおいてRKの方がマイクロストラクチャ

ャ・ノイズの影響を抑える株価指標が多いことが分かった。ボラティリティの予測に関しては、全体としては R-SV モデルが優れていたものの、R-GARCH モデルとわずかな差であった。その一方で、ARFIMAX モデルが最も優れていた株価指標は存在せず、他の 2 つのモデルと比べて予測力が劣っていることが分かった。

参考文献

沖本竜義 (2010) 『計量時系列分析』朝倉書店

中川重和、橋口博樹、仁木直人 (2009) 「正規性検定統計量の提案」数理解析研究所講究録 第 1666 巻 p.139-144

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(2009)「Realized Volatility を用いた日経 225 オプション価格の導出」大阪証券取引所 先物・オプションレポート(Vol.21 No.3) p.1

渡部敏明(2010)「日経 225 の Realized Volatility-マイクロストラクチャ・ノイズと夜間・昼休みの調整-」大阪証券取引所 先物・オプションレポート(Vol.22 No.2) p.1

渡部敏明、佐々木浩二 (2006) 「ARCH 型モデルと “Realized Volatility” によるボラティリティ予測とバリュー・アット・リスク」日本銀行金融研究所 p.39-70

渡部敏明、長倉大輔 (2009) 「GARCH 型モデルと Realized Volatility を用いた TOPIX 日時リターンの非線形性の検証」『日本統計学会誌』第 39 巻 p.71-74

渡部敏明、大森裕浩 (2013) 「Realized Stochastic Volatility モデル—マルコフ連鎖モンテカルロ法を用いたベイズ分析—」『日本統計学会誌』第 42 巻 p.274-297

Barndorff-Nielsen, O., Hansen, P.R., Lunde A. and Shepard N. (2008). “Designing Realized Kernels to Measure the Ex Post Variation of Equity Prices in the Presence of Noise,” *Econometrica*, 2-3.
Engle, R.F. (1982). “Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation,” *Econometrica*, 988-990.

Glosten, L.R., R. Jagannathan, and D. Runkle (1993) “On the Relation between the Expected Value and the Volatility of Normal Excess Return on Stocks, *Journal of Finance*,” **48**, 1786-1789.

Hansen, P.R., and Huang, Z. and Shek, H. (2011). “Realized GARCH: A Joint Model for Returns and Realized Measures of Volatility,” *J. Appl. Econ*, **27**, 877-906

Hansen, P.R., and Lunde, A. (2005) “A Forecast Comparison of Volatility Models: Does Anything Best a GARCH(1,1)?,” *J. FinanceEcon*, **20**, 525-554

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

Nelson D.B. (1991) “Conditional Heteroskedasticity in Asset Returns: A New Approach,” *Econometrica*, 350-351.

Patton, A.J. (2011). “Volatility forecast comparison using imperfect volatility proxies,” *Journal of Econometrics*, **160**, 246-255.

Taylor, S.J. (1986) *Modeling Financial Time Series*, JohnWiley & Sons.

<Web>

Oxford-Man Institute Realized Library, Retrieved October 17, 2014 from <http://realized.oxford-man.ox.ac.uk/>

Yahoo! Finance (参照 2014/10/29)

<http://finance.yahoo.co.jp/>

6. 補論 A

補論 A は、SV モデルを疑似最尤推定する際に用いられる、線形状態空間モデルにおけるカルマンフィルタや平滑化によって、パラメータやボラティリティを推定するとともに、R-SV モデルのパラメータ推定を解説する。カルマンフィルタや平滑化に関しては、渡部(2000) を参考にしている。

A-1 線形状態空間モデルによるカルマンフィルタ

状態空間モデルというのは、観測された時系列データ y_t が存在し、観測されないパラメータ x_t の推移を説明するモデルである。状態空間モデルは、次のように表す。

$$y_t = \alpha + \beta x_t + u_t \quad (\text{A.1})$$

$$x_t = \omega + \varphi x_{t-1} + \eta_t \quad (\text{A.2})$$

$$u_t \sim i.i.d. N(0, \sigma_u^2), \quad \eta_t \sim i.i.d. N(0, \sigma_\eta^2) \quad (\text{A.3})$$

ここで、(A.1) を観測方程式、(A.2) を遷移方程式と呼ぶ。まず、 $Cov(u_t, \eta_t) = 0$ 、つまり SV モデルでいうと、レバレッジ効果を観測しないモデルを考える。ここで、 u_t と η_t が正規分布に従うと仮定すると、

$$f(y_t | x_t) = \frac{1}{\sqrt{2\pi\sigma_u^2}} \exp \left[-\frac{(y_t - \alpha - \beta x_t)^2}{2\sigma_u^2} \right] \quad (\text{A.4})$$

$$f(x_t | x_{t-1}) = \frac{1}{\sqrt{2\pi\sigma_\eta^2}} \exp \left[-\frac{(x_t - \omega - \varphi x_{t-1})^2}{2\sigma_\eta^2} \right] \quad (\text{A.5})$$

となる。ここで、求める条件付き尤度は、 $f(x_t | y_{t-1})$ であり、この期待値は次のように表せる。

$$E(x_t | y_{t-1}) = \int_{-\infty}^{\infty} x_t f(x_t | y_{t-1}) dx_t \quad (\text{A.6})$$

ここで、条件付き密度 $f(x_t | y_{t-1})$ は次のように書き換えられる。

$$\begin{aligned} f(x_t | y_{t-1}) &= \int_{-\infty}^{\infty} f(x_t, x_{t-1} | y_{t-1}) dx_{t-1} \\ &= \int_{-\infty}^{\infty} f(x_t | x_{t-1}, y_{t-1}) f(x_{t-1} | y_{t-1}) dx_{t-1} \\ &= \int_{-\infty}^{\infty} f(x_t | x_{t-1}) f(x_{t-1} | y_{t-1}) dx_{t-1} \end{aligned} \quad (\text{A.7})$$

次に観測値が与えられたときの x_t の条件付き密度 $f(x_t | y_t)$ は次のように変形される。

$$\begin{aligned} f(x_t | y_t) &= \frac{f(x_t | y_t, y_{t-1})}{f(x_t, y_t | y_{t-1})} \\ &= \frac{f(y_t | y_{t-1})}{f(y_t | x_t, y_{t-1}) f(x_t | y_{t-1})} \\ &= \frac{f(y_t | y_{t-1})}{f(y_t | x_t) f(x_t | y_{t-1})} \end{aligned} \quad (\text{A.8})$$

そして、(A.8) 式の分母は以下のように表すことができる。

$$\begin{aligned}
 f(y_t|y_{t-1}) &= \int_{-\infty}^{\infty} f(y_t, x_t|y_{t-1}) dx_t \\
 &= \int_{-\infty}^{\infty} f(y_t|x_t, y_{t-1})f(x_t|y_{t-1}) dx_t \\
 &= \int_{-\infty}^{\infty} f(y_t|x_t)f(x_t|y_{t-1}) dx_t
 \end{aligned} \tag{A.9}$$

このような式を用いると、初期値 $f(x_1|y_0)$ が与えられた上で、(A.8)、(A.9) 式により $f(x_1|y_1)$ が計算されて、それを使って (A.7) 式より $f(x_2|y_1)$ が計算される。この繰り返しにより、逐次的に $f(x_t|y_t)$ を求めることが可能になる。そこで、 $f(x_t|y_t)$ の平均と分散を α_t, P_t 、 $f(x_t|y_{t-1})$ の平均と分散を $x_{t|t-1}, P_{t|t-1}$ 、 $f(y_t|y_{t-1})$ の平均と分散を v_t, F_t とすると、(A.7)～(A.9) 式は以下ようになる。

$$x_{t|t-1} = \omega + \varphi x_{t-1} \tag{A.10}$$

$$P_{t|t-1} = \varphi^2 P_{t-1} + \sigma_\eta^2 \tag{A.11}$$

$$x_t = x_{t|t-1} + \frac{\beta P_{t|t-1}}{F_t} v_t \tag{A.12}$$

$$P_t = P_{t|t-1} - \frac{\beta^2 P_{t|t-1}^2}{F_t} \tag{A.13}$$

$$v_t = y_t - \alpha - \beta x_{t|t-1} \tag{A.14}$$

$$F_t = \beta^2 P_{t|t-1} + \sigma_u^2 \tag{A.15}$$

この (A.10)～(A.15) 式をカルマンフィルタと呼び、先ほども述べた通り、 $x_{1|0}, P_{1|0}$ を設定すれば、 $x_1, P_1, x_{2|1}, P_{2|1} \dots$ と求めることができる。

$Cov(u_t, \eta_t) = \rho$ の場合⁵は、(A.12)、(A.13) 式においての $\beta P_{t|t-1}$ が $\beta P_{t|t-1} + \rho$ になって、(A.15) 式が以下の式に変形される。

$$F_t = \beta^2 P_{t|t-1} + 2\rho + \sigma_u^2 \tag{A.16}$$

A-2 状態空間モデルにおけるパラメータの推定

このモデルにおいて、未知パラメータは、 $\theta = (\alpha, \beta, \omega, \varphi, \rho, \sigma_u^2, \sigma_\eta^2)$ である。このパラメータを推定するには、疑似最尤法が考えられる。そのときの尤度関数は以下のように表される。

$$L = f(y_1|\theta) \prod_{t=2}^T f(y_t|y_{t-1}; \theta) \tag{A.17}$$

$$f(y_t|y_{t-1}; \theta) = \frac{1}{\sqrt{2\pi F_t}} \exp\left[-\frac{v_t^2}{2F_t}\right] \tag{A.18}$$

カルマンフィルタで求めた F_t, v_t を (A.17) 式に代入してパラメータを推定する。

⁵ Harvey (1989, p112-113) 参照

A-3 SV モデルへの応用

先ほどの線形状態空間モデルを SV モデルに応用するために、(2.8) 式を 2 乗した後、対数化することによって、線形モデルにする。そのあとで、

$$y_t \equiv \ln \varepsilon_t^2, \quad x_t \equiv \ln \sigma_t^2, \quad u_t \equiv \ln z_t^2 - E[\ln z_t^2] \quad (\text{A.19})$$

と置くと、 $\ln z_t^2$ の分布が標準正規分布に従うとき、平均が -1.27 、分散が $\pi^2/2$ である⁶。そのため、このような状態空間モデルに変換される。

$$y_t = -1.27 + x_t + u_t \quad (\text{A.20})$$

$$x_t = \omega + \varphi x_{t-1} + \eta_t \quad (\text{A.21})$$

先ほどのモデルにおいて、 $\alpha = -1.27$ 、 $\beta = 1$ 、 $\sigma_u^2 = \pi^2/2$ と置かれている。なので、推定するパラメータは、 $\omega, \varphi, \sigma_\eta^2, \rho$ となる。

A-4 平滑化 (Smoothing)

平滑化とは、観測データ数である T 期のもとでの $x_{t|T}, P_{t|T}$ を求めることを平滑化と呼ぶ。 $f(x_t|y_T)$ の平均と分散を $x_{t|T}, P_{t|T}$ とすると、次のように求められる。

$$x_{t|T} = x_{t|t} + P_{t*} (x_{t+1|T} - x_{t+1|t}) \quad (\text{A.22})$$

$$P_{t|T} = P_{t|t} + P_{t*}^2 (P_{t+1|T} - P_{t+1|t}) \quad (\text{A.23})$$

$$P_{t*} = \varphi \frac{P_{t|t}}{P_{t+1|t}} \quad (\text{A.24})$$

また、この $x_{t|T}, P_{t|T}$ が与えられたことによって、 T 期のデータが与えられたことによるボラティリティは次のように求められる。

$$\sigma_{t|T}^2 = \exp \left(x_{t|T} + \frac{1}{2} P_{t|T} \right) \quad (\text{A.25})$$

A-5 R-SV モデルの推定方法

R-SV モデルに関しては、未知パラメータは、 $\theta = (\xi, \sigma_u^2)$ である。まず、A-1 から A-4 を用いて、 σ_t^2 を求める。 σ_t^2 と RV が与えられると、次のような尤度関数を定義することが出来る。

$$L = \prod_{t=1}^T h(RV_t | \sigma_t^2; \theta) \quad (\text{A.26})$$

$$h(RV_t | \sigma_t^2) = \frac{1}{\sqrt{2\pi\sigma_u^2}} \exp \left[-\frac{(\ln RV_t - \xi - \ln \sigma_t^2)^2}{2\sigma_u^2} \right] \quad (\text{A.27})$$

これを最尤推定することによって、R-SV のパラメータを求めることが出来る。

⁶ 証明に関しては渡部 (2000, p120-122)

6. 補論 B

補論 B では、ARFIMAX モデルの式の展開について解説をする。

ARFIMAX モデルにおける

$$(1 + \theta L)^{-1}(1 - L)^d = 1 - \sum_{j=1}^{\infty} \Phi_j L^j \quad (\text{B.1})$$

$$\Phi_1 = d + \theta, \Phi_j = \alpha_j - \theta \Phi_{j-1} \quad (\text{B.2})$$

$$\alpha_1 = d, \alpha_j = \frac{j-d-1}{j} \alpha_{j-1} \quad (j \geq 2)$$

という等式が成り立つことを証明する。

(B.1)の両辺に $(1 + \theta L)$ を掛けると、

$$(1 - L)^d = \left(1 - \sum_{j=1}^{\infty} \Phi_j L^j\right) (1 + \theta L) \quad (\text{B.3})$$

となり、

(B.3)の右辺は、

$$\begin{aligned} & \left(1 - \sum_{j=1}^{\infty} \Phi_j L^j\right) (1 + \theta L) \\ &= 1 + \theta L - (1 + \theta L)(\Phi_1 L + \Phi_2 L^2 + \Phi_3 L^3 + \Phi_4 L^4 + \dots) \\ &= 1 + \theta L - (\Phi_1 L + \Phi_2 L^2 + \Phi_3 L^3 + \Phi_4 L^4 + \dots) - \theta L(\Phi_1 L + \Phi_2 L^2 \\ & \quad + \Phi_3 L^3 + \Phi_4 L^4 + \dots) \\ &= 1 + \theta L - \Phi_1 L - (\Phi_2 + \theta \Phi_1)L^2 - (\Phi_3 + \theta \Phi_2)L^3 - (\Phi_4 + \theta \Phi_3)L^4 \\ & \quad - \dots \end{aligned} \quad (\text{B.4})$$

ここで、漸化式 $\Phi_1 = d + \theta, \Phi_j = \alpha_j - \theta \Phi_{j-1}$ を変形して、

$\Phi_j + \theta \Phi_{j-1} = \alpha_j$ となるので、これを利用すると、

$$\begin{aligned} & 1 + \theta L - \Phi_1 L - (\Phi_2 + \theta \Phi_1)L^2 - (\Phi_3 + \theta \Phi_2)L^3 - (\Phi_4 + \theta \Phi_3)L^4 - \dots \\ &= 1 - \alpha_1 L - \alpha_2 L^2 - \alpha_3 L^3 - \alpha_4 L^4 - \dots \\ &= 1 - dL - \frac{1-d}{2} dL^2 - \frac{2-d}{3} * \frac{1-d}{2} L^3 - \dots \\ &= \sum_{n=0}^{\infty} \frac{\prod_{k=2}^n (k-1-d)}{n!} L^n \end{aligned} \quad (\text{B.5})$$

また、(B.3)の左辺を maclaurin 展開すると、

$$(1 - L)^d = 1 + (-d) \frac{L}{1!} + d(d-1) \frac{L^2}{2!} + (-d)(d-1)(d-2) \frac{L^3}{3!} + \dots$$

$$= \sum_{n=0}^{\infty} \frac{\prod_{k=2}^n (k-1-d)}{n!} L^n \quad (\text{B.6})$$

となり、右辺と一致する。

したがって、(B.3)の等式は成り立ち、(B.1)も成り立つ。

6. 補論 C

ここではボラティリティを2つに分類した時にネーミングされる、ヒストリカル・ボラティリティ(以降 HV と示す)とインプライド・ボラティリティ(同様に IV と示す)の2つについて言及する。

HV とは、過去の一定期間の株式、為替、債券、コモディティといった原資産価格変動率を平均化することによって算出されるもので、言い換えると“歴史的変動率”であり、いわゆる標準偏差のことである。後述するブラックショールズ計算モデルにおいては、過去n日にわたって株価を観測、 s_i を第i日の終値とし、以下のように定められている。

$$u_i = \log \frac{s_i}{s_{i-1}}, \quad \bar{u} = u_i \text{の平均} \quad (\text{C.1})$$

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (u_i - \bar{u})^2} \quad (\text{C.2})$$

ここでは、 s が HV の数値を示している。日次のリターンを計測する一定期間の算出の際には、通常 20 日または 30 日を用いる。また日次の標準偏差を年率換算する際に乗じる因数として、祝日などで市場が閉まっている日があるので、年間の営業日を 250~260 日と考えて、通常、250~260 の平方根を乗じることが多い(iFinance,2014)。また、オプションにおいては、ブラックショールズ計算モデルのプレミアムの理論価格を求める際に、後述する IV を推測する事にも使われる。このモデルで求まる価格はあくまで理論的なもの・参考価格のようなものであるので、実際のマーケットで取引されている価格とは異なるので、混同しないように注意しなければならない。

この IV は、先程と同様に金利 r 、満期 T 、権利行使価格 K での価格 $C(K, T)$ によって以下のモデルで定められている。

$$C(K, T) = SN\left(\frac{\log(S/K) + (r + \sigma^2/2)T}{\sigma\sqrt{T}}\right) - Ke^{-rT}N\left(\frac{\log(S/K) + (r - \sigma^2/2)T}{\sigma\sqrt{T}}\right) \quad (\text{C.3})$$

ここで、 S は現在の株価であり、 $N(x)$ は標準正規分布の分布関数を示している。

これは主にオプションで使われる用語で、現在のオプション料（プレミアム）から将来の変動率を予測したものをいう。これは、市場参加者が今後のプレミアムがどのように変動するかを数値化したものであり、市場関係者の人気や期待度などの将来の予想が反映されている。オプショントレーダーが原資産価格の大きな変動を予想すると、オプションのプレミアムが高くなり、その結果 IV も高くなり、原資産相場がそれほど変動しないと予想された場合には、プレミアムが低下し、IV も低くなる。つまり、IV は HV のような純粋な統計データではなく、オプショントレーダーの心理を表す数値なのである。

ここに今回の動機がある。各国のボラティリティの推測をすることによって、その地域ごとのオプショントレーダーの心理状況を図ることが出来る。つまり、各国の国民性を市場データから分析する事が可能になる。

さて、ここで上述したブラックショールズ計算モデルについて触れることにする。

ブラックショールズ計算モデルは以下の様に示されているモデルである。

$$C = S \cdot N(d_1) - Xe^{-r\tau} \cdot N(d_2) \quad (\text{C.4})$$

$$P_t = -S_t N(-d_1) + Ke^{-r\tau} N(-d_2) \quad (\text{C.5})$$

$$d_1 = \frac{\ln\left(\frac{S}{X}\right) + \left(r + \frac{\sigma^2}{2}\right)\tau}{\sigma\sqrt{\tau}} \quad (\text{C.6})$$

$$d_2 = d_1 - \sigma\sqrt{\tau} \quad (\text{C.7})$$

このモデルではC, S, X, τ , r, σ^2 と変数が多く含まれている。それぞれ市場のオプションを買う権利の価値(コールプレミアム)、現在の株価、行使価格、期間、短期金利、変動率を示しており、式(C.4)は式(C.3)と同じものでIV(コールプレミアム)を、式(C.5)はプットオプション価格を表す。

このモデルは上記の計算に必要なデータ(株価、行使価格、期間、変動率、金利)を市場で簡単に入手出来るうえ、計算にかかる時間が非常に少ないという利点があるため、現在実務界では頻繁に使用されている。

また、このブラックショールズモデルを更に応用し、他の原資産にも適用できるようなモデルが作られている。しかし、本稿ではこのモデルを適用していない。その理由として挙げられるのは、“計測する期間の長さの違い”である。ブラックショールズモデルでは計測出来る期間はある一点の時期のみで非常に短い。そのため、今回我々が定めた2000年頭から2006年末までという長期間の計測には向いていない。そのため、変動率の σ が定まらず、正確なコールプレミアムが求められないのである。

Value at Risk(VaR)とは、統計的手法を用いることにより、市場の経営の際の予想最大損失額を算出する指標をいう。現在保有している資産(ポートフォリオ)を、将来のある一定期間保有すると仮定した場合に、ある一定の確率の範囲内(信頼区間)で、マーケットの変動によって、どの程度の損失を被る可能性があるかを計測したものである。たとえば、ある資産において期間100日での信頼水準95%・VaRが5億円であるという状態では、予想最大損失額は、100日間にわたり、95%の確率で5億円におさまるが、 $100 - 95 = 5\%$ の確率で5億円以上の損失となるということになる(JAMES D. MACBETH and LARRY J. MERVILLE, 1979)。

このVaRの最大の特徴は、ポートフォリオが多様な金融資産で構成されている場合でも、一定の確率における「予想最大損失額」という共通の尺度で比較・管理することが可能であることで、これによりリスク管理において非常に便利になる。しかし、VaRは過去のデータから求めた予想変動率を用いるため、継続的なデータを取得できないような商品には不向きで、リーマンショックのような特殊な場合は想定されていない。そのため、異常時におけるリスクは、別途ストレステストなどを用いなければならない。