

統計数理

第68巻第1号

(通巻131号)

PROCEEDINGS OF THE INSTITUTE OF STATISTICAL MATHEMATICS

目次

特集「複雑な依存構造を持つデータの統計解析法 —コピュラとその周辺—」

「特集 複雑な依存構造を持つデータの統計解析法 —コピュラとその周辺—」について
加藤 昇吾・吉羽 要直…………… 1
接合関数モデルにおける統計的推測 [総合報告]
塚原 英敦…………… 5
セミパラメトリックコピュラモデルにおけるダイバージェンスの性質 [原著論文]
清 智也・松本 和也…………… 25
非対称 t 接合関数の性質と統計的推定方法 —資産価格変動への応用— [研究詳解]
吉羽 要直…………… 45
Realized Stochastic Volatility モデル —拡張と日本の株価指数への応用— [研究詳解]
高橋 慎・大森 裕浩・渡部 敏明…………… 65
確率的依存構造をもつコピュラモデル —統計的推定方法と計量ファイナンスへの
応用— [研究詳解]
野澤 勇樹・中村 信弘…………… 87
コピュラを用いた CDO 価格付けモデルのリスク計測モデルへの拡張 [原著論文]
室町 幸雄…………… 107
連続変数で表される事件の接合関数を用いた生存分析 [原著論文]
福元 健太郎…………… 129
コピュラを用いた生存時間解析 —相関のあるエンドポイントとメタ分析の活用— [総合報告]
江村 剛志・道前 洋史…………… 147

サポートベクター回帰における感度分析による変数選択の有効性の検証
—都道府県別全死因死亡率の影響要因の分析— [原著論文]
田辺 和俊・鈴木 孝弘…………… 175

2020年6月

大学共同利用機関法人 情報・システム研究機構 統計数理研究所

〒190-8562 東京都立川市緑町10-3 電話 050-5533-8500(代)

本号の内容はすべて <https://www.ism.ac.jp/editsec/toukei/> からダウンロードできます

ISSN 0912-6112

統
計
数
理

PROCEEDINGS OF THE INSTITUTE OF STATISTICAL MATHEMATICS

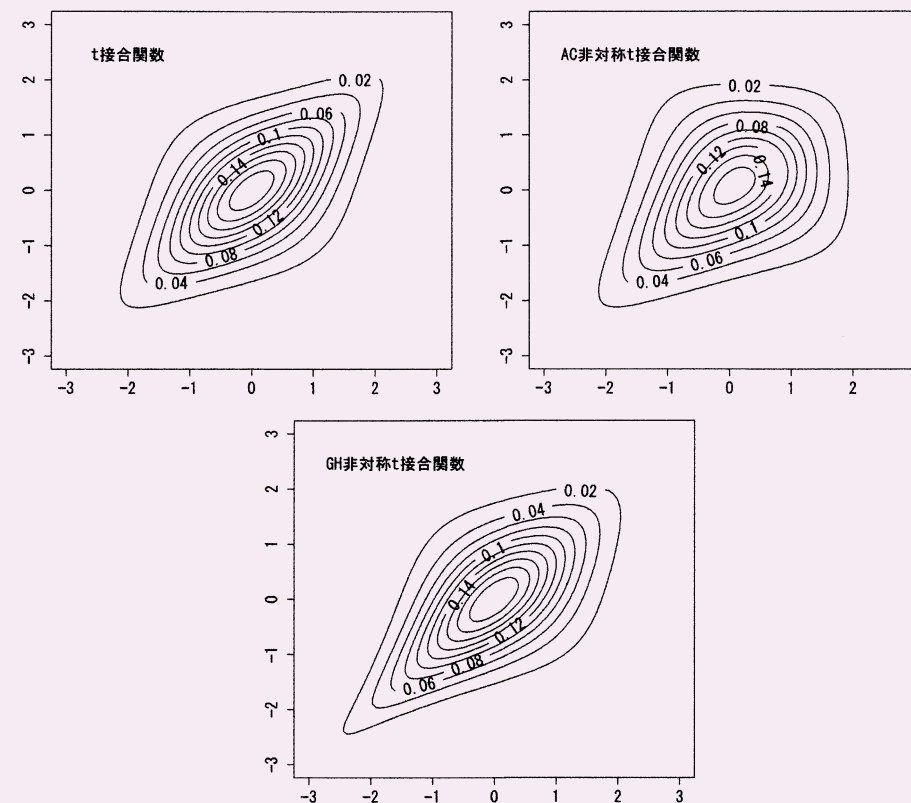
第68巻
第1号

2020

統計数理

Vol. 68, No. 1

PROCEEDINGS OF THE INSTITUTE OF STATISTICAL MATHEMATICS



統計数理研究所

統計数理

(年2回発行)

編集委員長 志村 隆彰

編集委員 ウ ステファン

坂田 綾香

島谷健一郎

庄 建倉

田中 未来

特集担当編集委員 加藤 昇吾

吉羽 要直 (東京都立大学)

編集室

池田 広樹

長嶋 昭子

脇地 直子

「統計数理」は、統計数理研究所における研究成果を掲載する統計数理研究所「彙報」として1953年に歴史を始め、1985年に誌名を変更し今の形となりました。現在は、統計数理研究所の研究活動に限らず、広く統計科学に関する投稿論文を掲載し、統計科学の深化と発展、そして統計科学を通じた社会への貢献を目指しています。

投稿を受け付けるのは、次の6種です。

- a. 原著論文
- b. 総合報告
- c. 研究ノート
- d. 研究詳解
- e. 統計ソフトウェア
- f. 研究資料

投稿された原稿は、編集委員会が選定・依頼した査読者の審査を経て、掲載の可否を決定します。投稿規程、執筆要項は、本誌最終頁をご参照ください。

また、上記以外にも統計科学に関して編集委員会が重要と認める内容について、編集委員会が原稿作成を依頼することがあります。

その他、「統計数理」に関するお問い合わせは、各編集委員にお願いします。

All communications relating to this publication should be addressed to associate editors of the Proceedings.

大学共同利用機関法人 情報・システム研究機構

統計数理研究所

〒190-8562 東京都立川市緑町10-3 電話050-5533-8500(代)

<https://www.ism.ac.jp/>

© The Institute of Statistical Mathematics 2020

印刷：笹氣出版印刷株式会社

PROCEEDINGS OF THE INSTITUTE OF STATISTICAL MATHEMATICS

Vol. 68, No. 1

Contents

Special Topic : Statistical Methods for Data with Complex Dependence Structure —Copulas and Related Topics—

On the Special Topic “Statistical Methods for Data with Complex Dependence Structure —Copulas and Related Topics—” Shogo KATO and Toshinao YOSHIBA	1
Statistical Inference in Copula Models Hideatsu TSUKAHARA	5
Properties of Divergence for Semiparametric Copula Models Tomonari SEI and Kazuya MATSUMOTO	25
Properties of Skew- t Copulas and Their Statistical Estimation —Application to Asset Returns— Toshinao YOSHIBA	45
Realized Stochastic Volatility Model —Extensions and Application to Japanese Stock Index— Makoto TAKAHASHI, Yasuhiro OMORI and Toshiaki WATANABE	65
A Copula Model with Stochastic Tail Dependence : Statistical Inference and Applications to Quantitative Finance Yuki NOZAWA and Nobuhiro NAKAMURA	87
An Extension of a CDO Pricing Model Using a Copula toward a Risk Evaluation Model Yukio MUROMACHI	107
Copula-based Continuous Event History Analysis Kentarō FUKUMOTO	129
Survival Analysis Using Copulas —Meta-analysis with Correlated Endpoints— Takeshi EMURA and Hirofumi MICHIMAE	147

Paper

Verification of the Effectiveness of Sensitivity Analysis as a Variable Selection in Support Vector Regression —Analysis of Factors Affecting Prefectural All-cause Mortality Rates— Kazutoshi TANABE and Takahiro SUZUKI	175
---	-----

June, 2020

Research Organization of Information and Systems

The Institute of Statistical Mathematics

10-3 Midori-cho, Tachikawa, Tokyo 190-8562, JAPAN

表紙の図は本誌 52 ページを参照

「特集 複雑な依存構造を持つデータの 統計解析法—コピュラとその周辺—」について

加藤 昇吾¹・吉羽 要直^{1,2} (オーガナイザー)

データ解析においては、変量間に非線形な関係があるデータや非対称に分布するデータなどの「複雑な依存構造を持つデータ」を扱わなければならないことがある。このようなデータは、金融・保険・医学・政治科学をはじめとする様々な学術分野で観測されている。しかし、古典的な統計学においては、線形回帰モデルや多変量正規分布のような単純な依存構造を持つモデルに基づく統計手法が主に研究され、これらの統計手法を複雑な依存構造を持つデータに応用すると誤った解析結果を与えてしまうことが知られている。

例えば、株価などの資産価格変動は、ブラックマンデーやリーマン・ショックのように急変動があるため、1変量正規分布では捉えられないと言われている。故に、こうした資産価格変動をリスクファクターとしてポートフォリオのリスクを把握するために多変量正規分布を用いることは、一般に適切ではない。そこで、リスクファクターに1変量正規分布ではない裾の重い分布などのより適切な分布を設定し、リスクファクター間には柔軟な依存構造を表現できるコピュラを用いて分析していくなどの統計解析法が必要となる。

本特集では、複雑な依存構造を持つデータのための統計解析法について、第一線で活躍する研究者による最新の研究結果およびレビューを提供する。本特集の副題にもあるように、本特集の論文では主にコピュラに関連した話題を扱っている。また、近年研究が進んでいる realized stochastic volatility モデルも議論している。

コピュラ(copula)は、非対称性・非線形性などの変量間の複雑な依存構造を表現できる確率分布族として、近年大きな注目を集めている。コピュラはその語源から、「接合分布関数」または「接合関数」とも呼ばれる(実際、塚原論文、吉羽論文、福元論文では「接合関数」を用いている)。コピュラは変量間の関係のみを表す確率分布族であるため、それぞれの変量の周辺分布はコピュラとは無関係に選ぶことができる。故にコピュラと周辺分布は自由に組み合わせることができ、それが複雑かつ柔軟なモデリングを可能としているのである。また、周辺分布やコピュラに経験分布を用いることで、セミパラメトリックまたはノンパラメトリックな統計解析を行うこともできる。さらに、モデリングのみならず、コピュラは変量間のある種の依存関係を測る尺度としても有用である。

コピュラに関して広く紹介している英語の文献としては、確率的な性質を詳しく議論している Nelsen (2006)や、より統計的な話題を紹介している Joe (2014)、計量経済学における応用を扱っている McNeil et al. (2015, 5章, 15.2節, 16.3節)が挙げられる。また、日本語の文献としては、統計的推測を含む入門として塚原(2008)、多変量コピュラに関する入門として室町(2014, 3章)、金融実務における解説として戸坂・吉羽(2005)、先に挙げた文献の初版の McNeil et al. (2005, 5章, 9.7節)の邦訳などがある。

¹ 統計数理研究所：〒190-8562 東京都立川市緑町10-3

² 東京都立大学大学院 経営学研究科：〒100-0005 東京都千代田区丸の内1-4-1

また、特集号で扱うもう 1 つの主要なテーマは、Takahashi et al. (2009) で提案された realized stochastic volatility モデルである。realized stochastic volatility モデルは、資産価格のボラティリティとして標準的になりつつある日中の高頻度リターンから計算される realized volatility と、代表的な資産価格変動モデルである確率ボラティリティ (stochastic volatility) モデルを組み合わせ、同時推定するモデルである。これにより、リターンとボラティリティの複雑な依存構造 (レバレッジ) を精緻に捉えようとしている。

本特集は 8 篇の招待論文から構成されている。塚原論文では、多変量分布から得られた標本を用いて周辺分布は未知のままコピュラ (接合関数) に関する推測を行う方法を網羅的に解説している。具体的には、各変量の順位のみからの推定が望ましくなる十分性を整理し、コピュラのノンパラメトリック推定、順位相関の漸近的性質、セミパラメトリック推定、適合度検定、リサンプリング法と、統計的推測に必要な理論的事項を網羅的に整理している。清・松本論文では、セミパラメトリックコピュラにおけるダイバージェンスの性質を考察している。多変量順位統計量によって定まる順位ダイバージェンスと Kullback-Leibler ダイバージェンスの局外パラメータに関する最小値として定義されるプロファイルダイバージェンスの関係について、興味深い性質を明らかにしている。吉羽論文では、資産価格変動の依存性が下落時により強くなる非対称性に注目し、金融実務でよく用いられる t コピュラの非対称 t コピュラへの拡張を整理している。非対称 t コピュラの定義と性質をまとめたうえで、最尤推定に必要な手続きと実証研究の結果を整理し、本邦の株価変動を推定した実証結果を検証している。高橋・大森・渡部論文では、前述の realized stochastic volatility モデルを整理したうえで、マルコフ連鎖モンテカルロ法を用いたベイズ推定法を示し、日経 225 株価指数に応用した結果を紹介している。野澤・中村論文では、コピュラのパラメータが確率的に変動する確率的コピュラのモデルの構築と数値計算による推定法を整理している。ファイナンス分野への応用事例として、各資産変動に高橋・大森・渡部論文で扱われた realized stochastic volatility モデルを取り込み、リターンとボラティリティの依存関係 (レバレッジ) を確率的コピュラで捉えた事例を挙げて説明している。さらに、ヴァインコピュラを通じて確率的コピュラを多変量に拡張し、為替ヘッジに適用した応用事例を報告している。室町論文では、多くの債務を集めそのデフォルト損失額に応じて証券化を行う CDO (Collateral Debt Obligations) の市場価格から導出されるインプライドコピュラをリスク計測モデルに拡張する際のモデリング手法について論じている。インプライドコピュラでは、価格付けのための確率測度で、僅かな確率で同時デフォルト確率が極端に高まる。これを、市場参加者の将来の環境激変への畏怖としてリスク計測のための確率測度に変換する際の取り扱いについて丁寧に論じている。福元論文では、連続変数で表される事件とその事件発生までの時間をコピュラで捉えた生存分析モデルを提案している。提案されたモデルで事件と時間を同時に分析することでバイアスを理論的に抑えられることを示したうえで、戦後英国の選挙の相対的タイミングと得票率の関係を示した先行研究を再分析している。江村・道前論文では、医療・生物学データベースの発展により複雑かつ膨大な生存時間データの解析の必要性が高まっていることに触れたうえで、コピュラを用いて患者の生存時間と無憎悪期間の 2 つの生存時間をモデリングする手法について総説している。個別医療への応用例として、遺伝子発現量と憎悪情報を用いて死亡率の動的予測を行う手法を紹介している。

本特集の概観からもわかるように、コピュラについては理論と応用の両面で、幅広く研究がなされている。さらに、応用については、以前から多くの研究がなされてきた金融・保険分野に加え、近年は政治科学・医学などの分野においても広がってきていることがわかる。コピュラを直接的に用いずに複雑な依存構造を表現した realized stochastic volatility モデルのようなモデルについても、今後ますます研究が発展していくであろう。本特集における論文を通して、読者がこれらの話題について、少しでも興味を持っていただければ、編集者としてはこれ

以上ない喜びである。

最後に、本特集の刊行にあたっては、執筆をお引き受けいただいた著者の皆様、有益なコメントをくださった査読者の皆様、手厚いサポートをしていただいた編集室の皆様に厚く御礼を申し上げたい。また、編集者が所属している統計数理研究所リスク解析戦略研究センターの山下智志センター長、栗木哲プロジェクトリーダーにも、本特集の企画に際し貴重なご助言をいただいたことにお礼申し上げたい。

参 考 文 献

- Joe, H. (2014). *Dependence Modeling with Copulas*, Chapman & Hall/CRC, Boca Raton, Florida.
- McNeil, A.J., Frey, R. and Embrechts, P. (2005). *Quantitative Risk Management: Concepts, Techniques, and Tools*, Princeton University Press, Princeton, New Jersey. (塚原英敦 他 訳 (2008). 『定量的リスク管理：基礎概念と数理技法』, 共立出版, 東京.)
- McNeil, A.J., Frey, R. and Embrechts, P. (2015). *Quantitative Risk Management: Concepts, Techniques, and Tools*, revised ed., Princeton University Press, Princeton, New Jersey.
- 室町幸雄 編 (2014). 『金融リスクモデリング—理論と重要課題へのアプローチ—』, 朝倉書店, 東京.
- Nelsen, R.B. (2006). *An Introduction to Copulas*, 2nd ed., Springer, New York.
- Takahashi, M., Omori, Y. and Watanabe, T. (2009). Estimating stochastic volatility models using daily returns and realized volatility simultaneously, *Computational Statistics and Data Analysis*, **53**, 2404–2426.
- 戸坂凡展, 吉羽要直 (2005). コピュラの金融実務での具体的な活用方法の解説, 金融研究, **24**(別冊2), 115–162 (<http://www.imes.boj.or.jp/japanese/kinyu/2005/kk24-b2-3.pdf> よりダウンロード可能).
- 塚原英敦 (2008). 接合分布関数(コピュラ)の理論と応用, 日本統計学会創立 75 周年記念出版, 『21 世紀の統計科学 第 3 巻: 数理・計算の統計科学』, 東京大学出版会, 東京 (<https://www.jss.gr.jp/book/books/> より増補版をダウンロード可能).

接合関数モデルにおける統計的推測

塚原 英敦[†]

(受付 2019 年 6 月 10 日；改訂 8 月 5 日；採択 8 月 7 日)

要 旨

本論文では、多変量分布 F からの標本を用いて、 F の接合関数に関する推測を、その周辺分布は未知のまま行う方法について解説する。この状況下では、 \mathcal{G} 十分性という基準に従えば、各座標ごとの順位から成るベクトルにのみ依存する方法が望ましいことになる。経験接合関数やその変形版は接合関数のノンパラメトリックな推定量であり、順位のみを通じてデータに依存する。その漸近的性質をまず簡潔に復習した後、順位相関係数のような接合関数の汎関数の推定量の性質を見ていく。さらに、接合関数が母数化されている場合に、そのパラメータをセミパラメトリックに推定するいくつかの方法を、擬似尤度推定量を含む順位近似 Z 推定量クラスにとりわけ重点を置いて考察する。続いて、独立性検定を含む一般的な適合度検定を手短かに再検討する。最後に、これまで述べた推定・検定法を実際に適用する際に必要不可欠である、経験接合関数に基づくいくつかのリサンプリング法について詳しく検討する。

キーワード：接合関数、 \mathcal{G} 十分性、経験接合関数、セミパラメトリック推定、適合度検定、ブートストラップ法。

1. はじめに

$\mathbf{X}_i := (X_{i1}, \dots, X_{id})$, $i = 1, \dots, n$ を独立に同一分布に従う d 次元確率ベクトルとし、その d 次元分布関数 F が連続であるとする。このとき、良く知られた Sklar の定理から F に対して一意的に定まる接合関数(copula) C を用いて

$$(1.1) \quad F(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d)), \quad (x_1, \dots, x_d) \in \mathbb{R}^d$$

と書ける。ここで、 F_1, \dots, F_d は 1 次元周辺分布関数である。よって、 d 次元連続分布関数の全体を \mathcal{P}_d と書くと、 \mathcal{P}_d は (F_1, \dots, F_d, C) という(無限次元)パラメータで母数化された族とみなすことができる。

各 $i \in \{1, \dots, n\}$ と $j \in \{1, \dots, d\}$ に対して、 X_{1j}, \dots, X_{nj} における X_{ij} の順位を $R_{ij,n}$ と書き(各座標ごとの順位)、 $\mathbf{R}_{i,n} := (R_{i1,n}, \dots, R_{id,n})$ とおく。すなわち、

$$(1.2) \quad R_{ij,n} = \sum_{k=1}^n \mathbb{1}\{X_{kj} \leq X_{ij}\}.$$

F_j は連続であるから、 $U_{ij} := F_j(X_{ij})$, $i \in \{1, \dots, n\}$ は $U(0, 1)$ 分布に従い、 (U_{i1}, \dots, U_{id}) , $i \in \{1, \dots, n\}$ は独立、そして分布関数が C である分布に従う。また、 $R_{ij,n}$ は U_{1j}, \dots, U_{nj} における U_{ij} の順位と一致することに注意する。

[†] 成城大学 経済学部：〒157-8511 東京都世田谷区成城 6-1-20

本稿では、 F_1, \dots, F_d が未知の場合に、データ X_1, \dots, X_n に基づく C に関する統計的推測手法について、これまでのこの分野での研究結果を概説する。次節では、順位ベクトルの \mathcal{G} 十分性を示して、順位に基づく方法がある種の望ましさの基準を満たすことを説明する。接合関数 C のノンパラメトリック推定量として経験接合関数を定義し、その漸近的性質を第 3 節で述べた上で、第 4 節では、順位相関係数などの接合関数の汎関数を考え、対応する経験接合関数プラグイン推定量の漸近的な結果を示す。第 5 節では、接合関数 C が C_θ と有限次元パラメータ θ で母数化されている場合に、 θ をセミパラメトリックに推定する方法をいくつか紹介する。第 6 節では、 C に関する様々な適合度検定を、 C がある接合関数族 \mathcal{C} に属するという帰無仮説の検定として統一的に整理して概観する。最後に第 7 節で、上記の推定・検定を行う際に必要なりサンプリング法について、これまでに提案されている手法を詳述する。接合関数に関する基本的な知識はある程度仮定して話を進めるが、それについては Nelsen (2006) や塚原 (2008) などを適宜参照していただきたい。

2. 順位統計量の \mathcal{G} 十分性

2.1 \mathcal{G} 十分性の一般論

まず、伝統的な統計的決定理論の枠組みを導入する。 \mathcal{X} を標本空間、 \mathcal{A} を \mathcal{X} 上の σ 集合体とし、可測空間 $(\mathcal{X}, \mathcal{A})$ 上の確率測度の族 $\{P_\theta\}_{\theta \in \Theta}$ が識別性条件を満たすとする ($\theta \neq \theta'$ ならば、 $P_\theta \neq P_{\theta'}$)。そして、標本空間 \mathcal{X} 上の変換群 \mathcal{G} が与えられているとする。すなわち、 \mathcal{G} は \mathcal{X} からそれ自身への(可測な)全単射の集合で、

- (i) $g_1, g_2 \in \mathcal{G}$ に対して、 $g_2 g_1 \in \mathcal{G}$ ($g_2 g_1$ は g_1 と g_2 の合成)
- (ii) $g \in \mathcal{G}$ に対して、 $g^{-1} \in \mathcal{G}$

を満たす。さらに、 \mathcal{G} から Θ 上の変換群 $\overline{\mathcal{G}}$ への準同型写像 $g \mapsto \overline{g}$ が存在し、

$$P_\theta(A) = P_{\overline{g}\theta}(gA)$$

が成り立つとする。このとき、確率分布族 $\{P_\theta\}_{\theta \in \Theta}$ は \mathcal{G} 不変であるという。

パラメータ空間 Θ は、ある Θ_1 と Θ_2 について直積空間 $\Theta = \Theta_1 \times \Theta_2$ と書けるものとし、その要素を $\theta = (\theta_1, \theta_2)$ と書く。 θ_1 が局外パラメータ (nuisance parameter) である場合に、 θ_2 に対する十分性 (部分十分性 (partial sufficiency)) をどう定義するかという問題は古くから議論されてきており、いくつかの概念が提案されている (Rémon, 1984 や Severini, 2000 参照)。ここでは群の考え方をういた流儀を述べて、接合関数モデルに適用する。

定義 2.1. (Barnard, 1963) $(\mathcal{X}, \mathcal{A})$ からある可測空間 $(\mathcal{T}, \mathcal{B})$ への可測写像、すなわち統計量 T が与えられたとする。このとき、 T が θ_1 未知の下で θ_2 に対して \mathcal{G} 十分 (\mathcal{G} -sufficient for θ_2 in the absence of knowledge of θ_1) であるとは、次の 3 つの条件が成り立つことをいう。

(i) $\overline{g} \in \overline{\mathcal{G}}$ は Θ_1 上推移的に作用し、かつ θ_2 は変えない。すなわち、任意の $\theta \in \Theta$ と $\overline{g} \in \overline{\mathcal{G}}$ に対して、 Θ_1 上の変換 \overline{g}' があって $\overline{g}\theta = (\overline{g}'\theta_1, \theta_2)$ が成り立つ。また、任意の $\theta_1, \tilde{\theta}_1 \in \Theta_1$ に対して、 $\theta_1 = \overline{g}'\tilde{\theta}_1$ となる \overline{g}' が存在する。

(ii) T は \mathcal{G} 不変である：すべての $g \in \mathcal{G}$, $x \in \mathcal{X}$ に対して、 $T(gx) = T(x)$ 。

(iii) S を任意の \mathcal{G} 不変な統計量とすると、 T を所与とした S の条件付き分布は θ_2 に依存しない。

この定義は Barnard (1963) で最初に与えられたとされるが、上記の形は Sprott (1965) に従った (この論文は、上の定義で (ii) から (iii) が自動的に成り立つ条件を与えている)。その骨子は、

θ_1 が未知であるということを, θ_2 に対する推測を行う場合に \mathcal{G} 不変な統計量に基づくべきであるという要請と捉え, その上で古典的な十分性の定義を当てはめるとのことである. データの縮約 (data reduction) という観点からいえば, まず不変性によって, 次に十分性による縮約を行うという考え方と解釈できる. これは, Hall et al. (1965, p. 579) による不変的十分性 (invariantly sufficient) という概念と一致する.

条件 (i) から, 写像 $\theta = (\theta_1, \theta_2) \mapsto \theta_2$ は変換群 $\overline{\mathcal{G}}$ の下での最大不変量となる. よって, \mathcal{G} 不変な統計量の分布は θ_2 にのみ依存する. この事実は Lehmann (1986, Theorem 6.3) に述べられているが, 不変性によりパラメータ空間も縮約され, θ_2 が共通の $\theta = (\theta_1, \theta_2)$ をすべて同一視することを意味する. この意味で θ_1 は局外パラメータということなのである.

上記のような群構造をもつモデルは, Barndorff-Nielsen et al. (1989) の流儀では変換モデル (transformation model) と呼ばれている. Severini (2000) では, 上記の設定および定義 2.1 (i) の状況を指して複合変換モデル (composite transformation model) と呼んでいる. Rémon (1984) は, (通常の意味での) 十分統計量の最大不変量による縮約であることを Barnard (1963) による \mathcal{G} 十分性の定義としているが, これでは十分性と不変性による縮約の順序が定義 2.1 とは異なっている. 縮約の結果が十分性と不変性の適用順序によらないための一般的条件は Hall et al. (1965) に与えられている (鍋谷, 1978, 5.4 節も参照). また, Severini (2000, Section 8.3.3) で最大不変量のことを \mathcal{G} 十分と呼んでいるが, これは Barnard (1963) の定義とは異なる.

上記の定義が適用される最も単純かつ古典的な例は以下のものである.

例 2.2. (正規分布族) $\mathcal{X} = \mathbb{R}^n$ とし, $\theta = (\mu, \sigma) \in \mathbb{R} \times \mathbb{R}_+$ に対して, P_θ を平均 μ , 分散 σ^2 の正規分布の n 個の直積とする (すなわち, 標本 X_1, \dots, X_n が i.i.d. $N(\mu, \sigma^2)$ ということである). 変換群 \mathcal{G} は平行移動群 (translation group), すなわち, $x = (x_1, \dots, x_n)$ に対して, $g_c x = (x_1 + c, \dots, x_n + c)$ とすると, $\overline{g}_c \theta = (\mu + c, \sigma)$ となる. そして, $S = [(n-1)^{-1} \sum_{i=1}^n (X_i - \overline{X})^2]^{-1/2}$ が μ 未知の下で σ に対して \mathcal{G} 十分となる.

2.2 接合関数モデルにおける順位の \mathcal{G} 十分性

独立性検定 (Blum et al., 1961) などのある種の問題では, C が興味の対象となるパラメータで, 周辺分布 F_j , $j \in \{1, \dots, d\}$ は局外パラメータとみなされる. このような場合, C に関する統計的推測において順位に基づく統計量を用いることの正当化は通常, 座標毎の狭義単調増加変換の下での不変性に基づいて議論されている. しかし, 単に不変だというだけでは順位に基づく方法の良さを示していることにはならない. そこで, 2.1 節において導入した定義 2.1 の意味での \mathcal{G} 十分性を一つの最適性と捉えて, それを順位統計量について示そう.

g_1, \dots, g_d を \mathbb{R} 上の連続な狭義単調増加関数で, $g_j(\mathbb{R}) = \mathbb{R}$, $j \in \{1, \dots, d\}$ を満たすものとし,

$$X = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1d} \\ x_{21} & x_{22} & \dots & x_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nd} \end{pmatrix} \in \mathbb{R}^{n \times d}$$

に対して, $g: \mathbb{R}^{n \times d} \rightarrow \mathbb{R}^{n \times d}$ を

$$gX = \begin{pmatrix} g_1(x_{11}) & g_2(x_{12}) & \dots & g_d(x_{1d}) \\ g_1(x_{21}) & g_2(x_{22}) & \dots & g_d(x_{2d}) \\ \vdots & \vdots & \ddots & \vdots \\ g_1(x_{n1}) & g_2(x_{n2}) & \dots & g_d(x_{nd}) \end{pmatrix}$$

で定義する．そして，このような変換全体を \mathcal{G} で書き表すと，これは $\mathbb{R}^{n \times d}$ 上の変換群となる．さらに， \mathcal{C}_d を d 次元接合関数全体， \mathcal{D} を 1 次元連続分布関数全体とし， $\Theta := \{\theta = (F_1, \dots, F_d, C) \in \mathcal{D}^d \times \mathcal{C}_d\}$ とおくと (上述の記法でいえば， $\Theta_1 = \mathcal{D}^d$ ， $\Theta_2 = \mathcal{C}$)， d 次元連続分布関数の全体 \mathcal{P}_d はパラメータ θ で母数化された族とみなせ， $\{P_\theta^{\otimes n}\}_{\theta \in \Theta}$ が標本空間 $\mathbb{R}^{n \times d}$ 上の確率分布族となる．

命題 2.3. 確率分布族 $\{P_\theta^{\otimes n}\}_{\theta \in \Theta}$ は \mathcal{G} 不変である．

証明. 対応する準同型写像 \bar{g} は

$$(2.1) \quad \bar{g}\theta = (F_1 \circ g_1^{-1}, \dots, F_d \circ g_d^{-1}, C)$$

で与えられるから，接合関数の単調変換下での不変性 (塚原, 2008, 定理 2.8) より主張は明らかである．□

$i \in \{1, \dots, n\}$ ， $j \in \{1, \dots, d\}$ に対して， x_{1j}, \dots, x_{nj} の中の x_{ij} の順位を r_{ij} とし，順位ベクトルを並べた行列を

$$R = R(X) := \begin{pmatrix} r_{11} & r_{12} & \dots & r_{1d} \\ r_{21} & r_{22} & \dots & r_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ r_{n1} & r_{n2} & \dots & r_{nd} \end{pmatrix} = (\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_d)$$

と書き，順位統計量と呼ぶ．

定理 2.4. 定義 2.1 の意味で，順位統計量 R は 1 次元周辺分布関数 (F_1, \dots, F_d) が未知の下で，接合関数 C に対して \mathcal{G} 十分である．

証明. 定義 2.1 の条件 (i)-(iii) を確かめればよい．準同型写像 \bar{g} は (2.1) で与えられており，(i) は明らかに満たされる．また，順位は狭義単調増加変換の下で不変だから， R が \mathcal{G} 不変であることも自明である．次に， S を任意の \mathcal{G} 不変な統計量とする． R を所与とするということは， (r_{1j}, \dots, r_{nj}) の逆順列を (s_{1j}, \dots, s_{nj}) とするとき， $\mathbb{R}^{n \times d}$ の $(n!)^d$ 個の順序錐による分割の中で，

$$\mathcal{K}_r := \mathcal{K}_{r_1} \times \dots \times \mathcal{K}_{r_d}, \quad \text{ただし } \mathcal{K}_{r_j} := \{\mathbf{x} \in \mathbb{R}^n : x_{s_{1j}} < \dots < x_{s_{nj}}\}$$

上に限定するというにほかならない．ところが，各 \mathcal{K}_{r_j} 内の任意の 2 点は適当な連続かつ狭義単調増加関数 g で互いに変換可能であるから， S はこの上で一定となる．よって， R を所与とした S の条件付き分布は C に依存しない．したがって，定理の主張が示された．□

次節以降では，接合関数 C に関する統計的推測の様々な方法について包括的に解説する．

注意 2.5. 上に述べた接合関数モデルに対しての順位統計量の部分十分性は，Hoff (2007) に簡単に述べられている．

3. 接合関数のノンパラメトリック推定

第 1 節で導入した設定を思い出そう： $\mathbf{X}_i := (X_{i1}, \dots, X_{id})$ ， $i = 1, \dots, n$ は d 次元連続分布関数 F からの iid 標本である． F が連続である場合には，その Sklar 分解 (1.1) における接合関数 C は

$$(3.1) \quad C(\mathbf{u}) = F(F_1^-(u_1), \dots, F_d^-(u_d)), \quad \mathbf{u} = (u_1, \dots, u_d) \in [0, 1]^d$$

で与えられることが容易にわかる。ここで、任意の 1 次元分布関数 H に対して、その分位関数(quantile)は $H^-(u) := \inf\{x \in \mathbb{R}: H(x) \geq u\}$ である。よって、上式における F および F_1^-, \dots, F_d^- を対応する経験分布関数・経験分位関数で置き換えれば、 C の自然なノンパラメトリック推定量が得られる。すなわち、 $j \in \{1, \dots, d\}$ に対して、

$$\mathbb{F}_{nj}(x_j) := \frac{1}{n} \sum_{i=1}^n \mathbb{1}\{X_{ij} \leq x_j\}, \quad \mathbb{F}_n(\mathbf{x}) := \frac{1}{n} \sum_{i=1}^n \mathbb{1}\{X_{i1} \leq x_1, \dots, X_{id} \leq x_d\}$$

と定義するとき、経験接合関数(Deheuvels, 1979) \mathbb{C}_n は

$$(3.2) \quad \mathbb{C}_n(\mathbf{u}) := \mathbb{F}_n(\mathbb{F}_{n1}^-(u_1), \dots, \mathbb{F}_{nd}^-(u_d))$$

で定義される。また、これを少し変形したバージョンである“順位に基づく経験接合関数”は次の式で定義される：

$$(3.3) \quad \tilde{\mathbb{C}}_n(\mathbf{u}) := \frac{1}{n} \sum_{i=1}^n \prod_{j=1}^d \mathbb{1}\left\{\frac{R_{ij,n}}{n+1} \leq u_j\right\}.$$

接合関数研究分野では、

$$(3.4) \quad \hat{U}_i := \frac{1}{n+1} \mathbf{R}_{i,n} = \left(\frac{R_{i1,n}}{n+1}, \dots, \frac{R_{id,n}}{n+1}\right), \quad i = 1, \dots, n$$

を擬似観測値(pseudo-observations)と呼ぶことが多い(Ghoudi and Rémillard, 2004)。よって、(3.3)は擬似観測値に基づく経験分布関数であるということが出来る。

同順位がない場合には、

$$(3.5) \quad \|\tilde{\mathbb{C}}_n - \mathbb{C}_n\|_\infty := \sup_{\mathbf{u} \in [0, 1]^d} |\tilde{\mathbb{C}}_n(\mathbf{u}) - \mathbb{C}_n(\mathbf{u})| \leq \frac{d}{n}$$

が成り立つことが容易にわかる。(3.2)と(3.3)はどちらも順位のみ依存するため、それらの確率分布は C を通してのみ F に依存する。

\mathbb{C}_n の漸近挙動を調べるために、経験接合過程(empirical copula process)を $\mathbb{G}_n := \sqrt{n}(\mathbb{C}_n - C)$ で定義する。 $[0, 1]^d$ 上の実数値有界関数全体に \sup ノルム $\|\cdot\|_\infty$ を備えたバナッハ空間を $\ell^\infty([0, 1]^d)$ で表す。そして、曲がりくねった矢印 \rightsquigarrow は、van der Vaart and Wellner (1996)で定義された意味での弱収束を表すとす。

経験接合関数が C の一様強一致推定量、すなわち $\|\mathbb{C}_n - C\|_\infty \rightarrow 0$, P -a.s. が成り立つことは Deheuvels (1979)らによって古くから知られていた。Fermanian et al. (2004)や Tsukahara (2005)は、 C の偏導関数 $\dot{C}_j = \partial C / \partial u_j$, $j \in \{1, \dots, d\}$ が $[0, 1]^d$ 全域で連続であるという条件の下で、経験接合過程 \mathbb{G}_n に対する収束定理を得ていたが、その仮定は Segers (2012)により裾依存性をもつ接合関数に対しても成り立つ条件に弱められた。

定理 3.1. (Segers, 2012) 各 $j \in \{1, \dots, d\}$ について、接合関数 C は集合 $\{\mathbf{u} \in [0, 1]^d: 0 < u_j < 1\}$ 上で連続な 1 次偏導関数 $\dot{C}_j(\mathbf{u}) = \partial C(\mathbf{u}) / \partial u_j$ をもつと仮定すると、 $\ell^\infty([0, 1]^d)$ において

$$\mathbb{G}_n \rightsquigarrow \mathbb{G}^C, \quad n \rightarrow \infty$$

が成り立つ。ここで、

$$\mathbb{G}^C(\mathbf{u}) := \mathbb{U}^C(\mathbf{u}) - \sum_{j=1}^d \dot{C}_j(\mathbf{u}) \mathbb{U}^C(\mathbf{1}, u_j, \mathbf{1})$$

であるが, $(\mathbf{1}, u_j, \mathbf{1})$ は第 j 成分が u_j で, その他のすべての成分が 1 の d 次元ベクトル, \mathbb{U}^C は C でピン止めされたブラウンシート (C -pinned Brownian sheet), すなわち連続な軌道をもつ平均 0 のガウス過程で, 共分散関数が

$$\text{Cov}\{\mathbb{U}^C(\mathbf{u}), \mathbb{U}^C(\mathbf{v})\} = C(\mathbf{u} \wedge \mathbf{v}) - C(\mathbf{u})C(\mathbf{v}), \quad \mathbf{u}, \mathbf{v} \in [0, 1]^d$$

で与えられるものである. また, $\mathbf{u} \wedge \mathbf{v}$ は成分毎に最小値をとったベクトル, すなわち, $\mathbf{u} \wedge \mathbf{v} = (\min(u_1, v_1), \dots, \min(u_d, v_d))$ である.

式(3.5)より, $\tilde{C}_n(\mathbf{u})$ を用いた経験接合過程 $\tilde{\mathbb{G}}_n := \sqrt{n}(\tilde{C}_n - C)$ についても同じ結果が成り立つことは明らかである.

C_n や \tilde{C}_n は階段関数であるから, 接合関数の性質の中で連続性や周辺分布の一様性は満たし得ない. これを満たす接合関数の推定量として, Segers et al. (2017) では次式で定義される経験ベータ接合関数が提案されている:

$$(3.6) \quad C_n^\beta(\mathbf{u}) = \frac{1}{n} \sum_{i=1}^n \prod_{j=1}^d F_{n, R_{ij, n}}(u_j), \quad \mathbf{u} \in [0, 1]^d$$

ここで, $u \in [0, 1]$ と $r \in \{1, \dots, n\}$ に対して,

$$F_{n, r}(u) = \sum_{s=r}^n \binom{n}{s} u^s (1-u)^{n-s} = \text{P}[\text{Bin}(n, u) \geq r]$$

はベータ分布 $B(r, n+1-r)$ の分布関数である. この経験ベータ接合関数は真の接合関数であるだけでなく, 平滑化パラメータが不要であること, リサンプリングが非常に容易であることなどの利点を兼ね備えている. また, Segers et al. (2017) では, 経験ベータ接合関数の小標本でのパフォーマンスが他の接合関数推定量と比べて良いこともモンテカルロ・シミュレーションで示されている. この経験ベータ接合関数に基づくリサンプリング法については, 第 7 節で後述する.

4. 接合関数の汎関数の推定

前節の漸近的な結果を用いることによって, 接合関数の汎関数の推定量に対する漸近分布も導くことができる. 最もよく知られているものは, 次の 2 つの順位相関係数であろう (Kendall and Gibbons, 1990):

$$(4.1) \quad \begin{aligned} \tau(C) &:= 4 \int_0^1 \int_0^1 C(u_1, u_2) dC(u_1, u_2) - 1 && \text{(Kendall の } \tau) \\ \rho(C) &:= 12 \int_0^1 \int_0^1 [C(u_1, u_2) - u_1 u_2] du_1 du_2 && \text{(Spearman の } \rho) \end{aligned}$$

それぞれの(伝統的な)標本版は

$$(4.2) \quad \begin{aligned} \hat{\tau} &:= \frac{2K}{n(n-1)} \sum_{i=1}^{n-1} \sum_{k=i+1}^n \text{sign}[(X_{k1} - X_{i1})(X_{k2} - X_{i2})], \\ \hat{\rho} &:= \frac{1}{n(n^2-1)} \left[12 \sum_{i=1}^n R_{i1, n} R_{i2, n} - 3n(n+1)^2 \right] = 1 - \frac{6 \sum_i (R_{i1, n} - R_{i2, n})^2}{n(n^2-1)} \end{aligned}$$

で定義される。これらは $\tau(\mathbb{C}_n)$ や $\rho(\mathbb{C}_n)$ とは若干異なる形となるが (Tsukahara, 2000), 漸近分布は同等となるため, 以下ではこれらのプラグイン推定量を考える。 C は連続微分可能で偏導関数 \dot{C}_j ($j = 1, 2$) をもつと仮定すると,

$$\sqrt{n}(\rho(\mathbb{C}_n) - \rho) = 12 \iint \sqrt{n}[C_n(u_1, u_2) - C(u_1, u_2)] du_1 du_2$$

は $12 \iint \mathbb{D}^C(u_1, u_2) du_1 du_2$ に弱収束する。この分布は平均 0 の正規分布であるが, その分散を書き下すのは面倒である。しかし, $C(u_1, u_2) = u_1 u_2 =: \lambda$, すなわち X_{i1} と X_{i2} , $i = 1, \dots, n$ が独立であれば, \mathbb{D}^λ はブラウン枕 (Brownian pillow) となるから, 直接的な計算により $12 \iint \mathbb{D}^\lambda(u_1, u_2) du_1 du_2$ の分散は 1 となることがわかる。これは標本 Spearman 順位相関係数の独立性帰無仮説の下での漸近分散に等しくなっていることも確かめることができる (Kendall and Gibbons, 1990 参照)。

一般に, $[0, 1]^3$ 上の任意の関数 J に対して,

$$S(C) := \iint J(u_1, u_2, C(u_1, u_2)) du_1 du_2$$

と定義し, 対応する標本版 $S(\mathbb{C}_n)$ を Spearman 型順位統計量と呼ぶことにする。すると, 上の $\rho(\mathbb{C}_n)$ は $J(u_1, u_2, w) = 12(w - u_1 u_2)$ の場合に対応する。この型の統計量に対する漸近分布は, Gaenssler and Stute (1987) で得られている。

定理 4.1. 関数 J は w に関して連続な偏導関数 $J^3(u, v, w)$ をもち, $\sup_{u, v, w} |J^3(u_1, u_2, w)| < \infty$ を満たすとする, $\sqrt{n}(S(\mathbb{C}_n) - S(C))$ は平均 0, 分散

$$\sigma^2(S) := \text{var} \left(\iint J^3(u_1, u_2, w) Z(u_1, u_2) du_1 du_2 \right)$$

の正規分布に分布収束する。ただし, $\sigma^2(S) > 0$ と仮定し, $Z(u_1, u_2)$ は

$$\begin{aligned} Z(u_1, u_2) := & \mathbb{1}_{\{\xi_1 \leq u_1, \xi_2 \leq u_2\}} - C(u_1, u_2) \\ & - \dot{C}_1(u_1, u_2)(\mathbb{1}_{\{\xi_1 \leq u_1\}} - u_1) - \dot{C}_2(u_1, u_2)(\mathbb{1}_{\{\xi_2 \leq u_2\}} - u_2) \end{aligned}$$

で定義される確率過程, (ξ_1, ξ_2) は分布関数 C に従う確率ベクトルである。

$T(\mathbb{C}_n)$ に対する漸近論はより複雑である。Spearman 型順位統計量の場合と同様に, $[0, 1]^3$ 上の関数 J に対して,

$$T(C) := \iint J(u_1, u_2, C(u_1, u_2)) dC(u_1, u_2)$$

とおき, $T(\mathbb{C}_n)$ を Kendall 型順位統計量と呼ぶことにする。 $J(u_1, u_2, w) = 4w - 1$ ととれば $\tau(\mathbb{C}_n)$ が得られる。この Kendall 型順位統計量の漸近分布は次の定理で与えられる。

定理 4.2. $J(u_1, u_2, w)$ は $(0, 1)^3$ 上で連続微分可能であるとし, 偏導関数

$$J^1(u_1, u_2, w) = \frac{\partial J(u_1, u_2, w)}{\partial u_1}, \quad J^2(u_1, u_2, w) = \frac{\partial J(u_1, u_2, w)}{\partial u_2}, \quad J^3(u_1, u_2, w) = \frac{\partial J(u_1, u_2, w)}{\partial w},$$

は一様に有界であると仮定する。このとき, $\sqrt{n}(T(\mathbb{C}_n) - T(C))$ は平均 0, 分散

$$\begin{aligned} \sigma^2(T) := & \text{var} \left(J(\xi_1, \xi_2, C(\xi_1, \xi_2)) + \iint [J^3(u_1, u_2, C(u_1, u_2))(\mathbb{1}_{\{\xi_1 \leq u_1, \xi_2 \leq u_2\}} - C(u_1, u_2)) \right. \\ & \left. + J^1(u_1, u_2, C(u_1, u_2))(\mathbb{1}_{\{\xi_1 \leq u_1\}} - u_1) + J^2(u_1, u_2, C(u_1, u_2))(\mathbb{1}_{\{\xi_2 \leq u_2\}} - u_2)] dC(u_1, u_2) \right) \end{aligned}$$

の正規分布に分布収束する。ただし、 $\sigma^2(T) > 0$ と仮定し、 (ξ_1, ξ_2) は定理 4.1 と同様に分布関数 C に従う確率ベクトルである。

証明は Shorack and Wellner (1986) 式の記法を用いて概略だけを述べておこう。まず、

$$\begin{aligned} & \sqrt{n}(T(\mathbb{C}_n) - T(C)) \\ &= \iint J^3(u_1, u_2, C(u_1, u_2)) \mathbb{G}_n(u_1, u_2) d\mathbb{C}_n(u_1, u_2) + \iint J(u_1, u_2, C(u_1, u_2)) d\mathbb{G}_n(u_1, u_2) + R_n \end{aligned}$$

と書ける。ここで、 $R_n = o_P(1)$ である。1 番目の積分は

$$\iint J^3(u_1, u_2, C(u_1, u_2)) \mathbb{G}^C(u_1, u_2) dC(u_1, u_2)$$

に収束する。2 番目の積分も $\iint J(u_1, u_2, C(u_1, u_2)) d\mathbb{G}^C(u_1, u_2)$ に収束する。定理 3.1 と部分積分を用いた直接的な計算によって、これら 2 つの確率変数間の共分散構造は

$$\iint J^3(u_1, u_2, C(u_1, u_2)) Z(u_1, u_2) dC(u_1, u_2)$$

($Z(u_1, u_2)$ は定理 4.1 で定義された確率過程である) と

$$\begin{aligned} & J(\xi_1, \xi_2, C(\xi_1, \xi_2)) - \iint J(u_1, u_2, C(u_1, u_2)) dC(u_2, u_2) \\ &+ \iint [J^1(u_1, u_2, C(u_1, u_2)) + J^3(u_1, u_2, C(u_1, u_2)) \dot{C}^1(u_1, u_2)] (\mathbb{1}_{\{\xi_1 \leq u_1\}} - u_1) dC(u_1, u_2) \\ &+ \iint [J^2(u_1, u_2, C(u_1, u_2)) + J^3(u_1, u_2, C(u_1, u_2)) \dot{C}^2(u_1, u_2)] (\mathbb{1}_{\{\xi_2 \leq u_2\}} - u_2) dC(u_1, u_2) \end{aligned}$$

との間の共分散構造と同じであることが示される。これら 2 つを加えるといくつかの項が相殺し合って、定理に述べた形の漸近分布が得られることがわかる。

注意 4.3. 以上の議論は大雑把であるが、より厳密には汎関数デルタ法 (van der Vaart and Wellner, 1996, Section 3.9) を用いて、ある種の関数微分可能性を示すことにより証明することもできる。他の依存性尺度については、Nelsen (2006, Chapter 5) とそこでの引用文献を参照せよ。

5. セミパラメトリック・モデルにおける推定

本節では、iid 標本 $\mathbf{X}_i, i \in \{1, \dots, n\}$ の母集団分布関数 F の接合関数 C が母数化されているセミパラメトリックモデルを考える：

$$F(x_1, \dots, x_d) = C_\theta(F_1(x_1), \dots, F_d(x_d)), \quad \theta \in \Theta \subset \mathbb{R}^m.$$

このモデルにおける推定問題とは、周辺分布関数 F_1, \dots, F_d が未知のまま θ を推定することである。

5.1 順位近似 Z 推定

C_θ が絶対連続で、 θ に関して微分可能な密度関数 c_θ をもつと仮定し、 $\dot{c}_\theta := (\partial c_\theta / \partial \theta_1, \dots, \partial c_\theta / \partial \theta_m)^\top$ とおくと、上記問題に対する尤度関数は、

$$\prod_{i=1}^n c_{\theta}(F_1(X_{i1}), \dots, F_d(X_{id})) f_1(X_{i1}) \cdots f_d(X_{id})$$

と書ける。もし F_1, \dots, F_d が既知であれば、尤度方程式

$$\sum_{i=1}^n \frac{\dot{c}_{\theta}(F_1(X_{i1}), \dots, F_d(X_{id}))}{c_{\theta}(F_1(X_{i1}), \dots, F_d(X_{id}))} = \mathbf{0}$$

の解が通常の最尤推定量となる。しかし、それらは未知であるため、スケールを変えて $n+1$ で割った経験分布関数で代用することになると、結局それは $(F_1(X_{i1}), \dots, F_d(X_{id}))$ を式(3.4)の擬似観測値 \hat{U}_i で置き換えることと同じになる。さらに、 $\dot{c}_{\theta}/c_{\theta}$ を一般の推定関数 $\phi(\cdot, \theta)$ で置き換えると、

$$(5.1) \quad \sum_{i=1}^n \phi[\mathbb{F}_{n1}(X_{i1}), \dots, \mathbb{F}_{nd}(X_{id})] = \mathbf{0}$$

という推定方程式が得られる。ただし、 $\phi(\cdot, \theta) = (\phi_1(\cdot, \theta), \dots, \phi_m(\cdot, \theta))$ は $(0, 1)^d$ 上の \mathbb{R}^m 値関数で、

$$\int \phi_k(\mathbf{u}, \theta) dC_{\theta}(\mathbf{u}) = 0, \quad k \in \{1, \dots, m\}$$

を満たさなければならない。このとき、方程式(5.1)の解 $\hat{\theta}_n^{RAZ}$ を θ の順位近似 Z 推定量と呼ぶ (Tsukahara, 2005)。そして、特に、 $\phi(\cdot, \theta) = \dot{c}_{\theta}/c_{\theta}$ のときには、(5.1)の解は擬似尤度推定量 (pseudo-likelihood estimator) と呼ばれている (Genest et al. 1995)。他の例については、Tsukahara (2005)を参照されたい。

この順位近似 Z 推定量 $\hat{\theta}_n^{RAZ}$ に対する漸近論を展開するために、次の定義をおく。

定義 5.1. (i) \mathcal{Q} は $[0, 1]$ 上の実数値連続関数 q で、 $(0, 1)$ 上では正の値をとり、 $1/2$ に関して対称、 $[0, 1/2]$ 上では増加関数で、 $\int_0^1 \{q(t)\}^{-2} dt < \infty$ を満たすものの全体である。

(ii) 関数 $r: (0, 1) \rightarrow (0, \infty)$ が \mathbf{u} 形であるとは、 $1/2$ に関して対称であり、 $(0, 1/2]$ 上で増加関数であることをいう。

(iii) $0 < \beta < 1$ と \mathbf{u} 形関数 r に対して、

$$r_{\beta}(t) = \begin{cases} r(\beta t) & \text{if } 0 < t \leq 1/2; \\ r\{1 - \beta(1 - t)\} & \text{if } 1/2 < t \leq 1. \end{cases}$$

とおく。ある 0 の近傍におけるすべての $\beta > 0$ に対して、 $(0, 1)$ 上 $r_{\beta} \leq M_{\beta} r$ となるような定数 M_{β} が存在するならば、 r を再生 \mathbf{u} 形関数と呼ぶ。そして、再生 \mathbf{u} 形関数の全体を \mathcal{R} と書く。

\mathcal{Q} と \mathcal{R} に属する関数で典型的な例は

$$q(t) = \{t(1-t)\}^{\alpha}, \quad 0 < \alpha < 1/2, \quad r(t) = D\{t(1-t)\}^{-\tau}, \quad \tau \geq 0, \quad D \geq 0$$

である。

一般的な正則条件として、次のような条件が必要である。

(A.1) 各 θ と $k \in \{1, \dots, m\}$ に対して、 $\phi_k(\cdot, \theta): (0, 1)^d \rightarrow \mathbb{R}$ は連続微分可能であり、その偏導関数を $\phi_k^j = \partial \phi_k / \partial u_j$, $j \in \{1, \dots, d\}$ としたとき、関数 $r_j \in \mathcal{R}$, $\tilde{r}_j \in \mathcal{R}$ と $q_j \in \mathcal{Q}$ ($j \in \{1, \dots, d\}$) が存在し、

$$|\phi_k(\mathbf{u}, \theta)| \leq \prod_{j=1}^d r_j(u_j), \quad |\phi_k^j(\mathbf{u}, \theta)| \leq \tilde{r}_j(u_j) \prod_{l \neq j} r_l(u_l), \quad j \in \{1, \dots, d\},$$

および

$$\int_{(0,1)^d} \left\{ \prod_{j=1}^d r_j(u_j) \right\}^2 dC_\theta(\mathbf{u}) < \infty,$$

$$\int_{(0,1)^d} \left\{ q_j(u_j) \tilde{r}_j(u_j) \prod_{l \neq j} r_l(u_l) \right\} dC_\theta(\mathbf{u}) < \infty, \quad j \in \{1, \dots, d\}$$

が成り立つ。

(A.2) θ_0 は写像 $\theta \mapsto \int \phi(\mathbf{u}, \theta) dC_{\theta_0}(\mathbf{u})$ の一意な零点である。

(A.3) 各 $j \in \{1, \dots, m\}$ に対して, $\phi_j(\mathbf{u}, \theta)$ は dC_θ に関して $(0, 1)^d$ 上 2 乗可積分である。

(A.4) $\phi(\cdot, \theta)$ は θ について微分可能で, そのヤコビ行列

$$\dot{\phi}(\mathbf{u}, \theta) = \left[\frac{\partial \phi_k(\mathbf{u}, \theta)}{\partial \theta_l} \right]_{m \times m}$$

は次の条件を満たす: $\dot{\phi}(\mathbf{u}, \theta)$ は \mathbf{u} と θ について連続であり, ある dC_θ 可積分関数 $h(\mathbf{u})$ が存在して, $\|\dot{\phi}(\mathbf{u}, \theta)\| \leq h$ となる。

(A.5) $m \times m$ 行列 $A_0 = \int \dot{\phi}(\mathbf{u}, \theta_0) dC_{\theta_0}(\mathbf{u})$ は非特異である。

一般によく用いられる接合関数の多くについて, 擬似尤度推定量に対応する $\phi(\cdot, \theta) = \dot{c}_\theta / c_\theta$ が上記の条件を満たすことは簡単に確かめられる。

以上の正則条件の下で, 順位近似 Z 推定量に対する漸近分布が得られる。

定理 5.2. 条件 (A.1)-(A.5) を仮定する。 $n \rightarrow \infty$ のとき, 推定方程式 (5.1) に対する解 $\hat{\theta}_n$ が存在し, θ の真値 θ_0 に収束する確率は 1 に近づく。さらに, $n \rightarrow \infty$ のとき, $\sqrt{n}(\hat{\theta}_n - \theta_0)$ は平均 0, 共分散行列 $A_0^{-1} \Sigma_0 A_0^{-1}$ の正規分布に分布収束する。ここで, ξ を分布関数 C_{θ_0} に従う確率ベクトルとすると,

$$\Sigma_0 = \text{var} \left\{ \phi(\xi, \theta_0) + \sum_{j=1}^d \int \phi^j(\mathbf{u}, \theta_0) (\mathbb{1}\{\xi_j \leq u_j\} - u_j) dC_{\theta_0}(\mathbf{u}) \right\}$$

である。

5.2 他の推定法

最小距離推定. 2つの接合関数 C と D の距離 $\gamma(C, D)$ が与えられているとき, 経験接合関数 C_n と C_θ の距離 $\gamma(C_n, C_\theta)$ を最小化する θ の値を推定値とする方法を最小距離推定法という。すなわち, 最小距離推定量は

$$\hat{\theta}_n^{\text{MD}} := \arg \min_{\theta} \gamma(C_n, C_\theta)$$

で定義される。 γ の例としては,

$$\gamma(C, D) = \int_{(0,1)^d} \{C(\mathbf{u}) - D(\mathbf{u})\}^2 d\mathbf{u} \quad (\text{Cramér-von Mises 距離})$$

$$\gamma(C, D) = \sup_{\mathbf{u} \in (0,1)^d} |C(\mathbf{u}) - D(\mathbf{u})| \quad (\text{Kolmogorov-Smirnov 距離})$$

$$\gamma(C, D) = \int_{(0,1)^d} \{\sqrt{c(\mathbf{u})} - \sqrt{d(\mathbf{u})}\}^2 d\mathbf{u} \quad (\text{Hellinger 距離})$$

などが考えられる。ここで、 c と d はそれぞれ C と D の接合密度関数を表す。そのため、Hellinger 距離を用いるためには、カーネル法などによって接合密度関数の推定を行う必要がある。これらの推定量はある種の頑健性をもつと期待されるが、その詳細については調べられていない。漸近分布は Tsukahara (2005) で得られている。

順位相関反転推定. $d = 2, m = 1$ の場合を考える。4 節で述べたように、Kendall と Spearman の母順位相関係数(4.1)は接合関数 C の汎関数として定義されるが、本節での設定のように C が母数化されて C_θ となっている場合には、 θ の関数となる。それらをそれぞれ $\tau(\theta)$ 、 $\rho(\theta)$ と書くことにする。ここで積率法との類似で、これらを標本順位相関係数(4.2)と等号で結び、それらを θ について解いて得られる推定量

$$\hat{\theta}_\tau = \tau^{-1}(\hat{\tau}), \quad \hat{\theta}_\rho = \rho^{-1}(\hat{\rho})$$

を順位相関反転推定量と呼ぶ(積率法推定量とも呼ばれる)。この方法は、 $\tau(\theta)$ や $\rho(\theta)$ が θ の簡単な関数である場合に有効であり(その例については、塚原, 2008 参照)、2次元の場合にしかな適用できない方法ではあるが、しばしばより複雑な最適化計算を要する推定方法の初期値として用いられることがある。 τ 反転推定量は、Oakes (1982) (Clayton 族) と Genest (1987) (Frank 族) で考察されている。

本節で紹介した推定方法は計算機言語 R でパッケージ ‘copula’ として実装されているが、その詳細についてはそのマニュアルである Hofert et al. (2018) や Hofert et al. (2019) を参照してほしい。Tsukahara (2005) や Kojadinovic and Yan (2010) で報告されているモンテカルロ・シミュレーションによる比較実験では、2次元の主だった接合関数族について、平均 2 乗誤差が小さいという意味で、擬似尤度推定量が最良の推定量であるという結論が得られている。Chen et al. (2006) では、篩 (sieve) 法によるセミパラメトリック最尤推定量がセミパラメトリックな意味で漸近的に最適な推定量であることが示されている。しかし、周辺分布がすべて未知の場合には、有限標本においてこの複雑な方法が擬似尤度推定量を凌駕するというようなシミュレーション結果は得られていない。

6. 適合度検定

本節では、 C に関する様々な検定を短く概観する(より詳細な解説論文としては、Fermanian, 2013 や Genest et al., 2009 がある)。 C に関する検定問題は、ある接合関数の族 \mathcal{C} について、帰無仮説 $C \in \mathcal{C}$ を対立仮説 $C \notin \mathcal{C}$ に対して検定するという形で表現できる。

- 例 6.1.** (i) 独立性の検定: $\mathcal{C} = \{\Pi\}$, $\Pi(\mathbf{u}) = \prod_{j=1}^d u_j$ は独立接合関数。
 (ii) 古典的な適合度検定: $\mathcal{C} = \{C_0\}$, C_0 は所与の接合関数。
 (iii) パラメータ族の適合度検定: $\mathcal{C} = \{C_\theta\}_{\theta \in \Theta}$ 。
 (iv) 対称性の検定: $\mathcal{C} = \{C: (1, \dots, d)$ の任意の置換 (π_1, \dots, π_d) に対して, $C(u_1, \dots, u_d) = C(u_{\pi_1}, \dots, u_{\pi_d})\}$ 。
 (v) 極値依存性の検定: $\mathcal{C} = \{C: \text{接合関数 } C_* \text{ が存在して } C = \lim_{n \rightarrow \infty} C_*^n(u_1^{1/n}, \dots, u_d^{1/n})\}$ 。
 (vi) アルキメデス型の検定: $\mathcal{C} = \{C: C(\mathbf{u}) = \phi^{-1}(\phi(u_1) + \dots + \phi(u_d))\}$, ϕ は d 単調な生成素}。

これらの検定問題に対する検定統計量の多くは、経験接合関数に基づくものである。(i)–(iii)については、5.2 節で最小距離推定量を考えた際に導入した、2つの接合関数 C と D の距離 $\gamma(C, D)$ を用いて、

$$\inf_{C \in \mathcal{C}} \gamma(\mathbb{C}_n, C)$$

という形の検定統計量が基本的である。\$\gamma\$ としては、5.2 節で挙げた Cramér-von Mises 距離、Kolmogorov-Smirnov 距離のほかに、積分で重みをつける

$$D_n := \int_{[0,1]^d} [\mathbb{C}_n(\mathbf{u}) - C(\mathbf{u})]^2 w_n(\mathbf{u}) d\mathbb{C}_n(\mathbf{u}), \quad D'_n := \int_{[0,1]^d} [\mathbb{C}_n(\mathbf{u}) - C(\mathbf{u})]^2 w_n(\mathbf{u}) d\mathbf{u}$$

といった Anderson-Darling 型のものも考えられる。Genest et al. (2009)では、Kolmogorov-Smirnov 距離に基づく検定は、多くの場合に検出力が他の検定よりも小さいことがシミュレーション実験により報告されている。

(iv)の検定として、\$d = 2\$ の場合には

$$\int_0^1 \int_0^1 [\mathbb{C}_n(u_1, u_2) - \mathbb{C}_n(u_2, u_1)]^2 d\mathbb{C}_n(u_1, u_2)$$

の形の検定統計量が自然に考えられる (Genest et al., 2012; Harder and Stadtmüller, 2017)は、\$d > 2\$ の場合への拡張を論じている。

(v)と(vi)の問題はより複雑であり、定義とは異なる \$\mathcal{C}\$ の便利な特徴づけが必要となる。(v)については、いわゆる Kendall 分布を用いる検定法や Pickands 従属性関数の推定量と経験接合関数に基づくものなどが提案されているが、詳しくは Bücher and Kojadinovic (2015)を参照されたい。(vi)については、\$d = 2\$ の場合に Bücher et al. (2012)が検定方式を提案しているが、\$d > 2\$ の場合への拡張は困難なようである。

これらの検定について、漸近的な帰無分布は経験接合過程 \$\sqrt{n}(\mathbb{C}_n - C)\$ の極限分布から導くことができることが多い。しかし、Fermanian (2013)にも述べられているように、漸近分布から \$p\$ 値、あるいは臨界値を求めることは 2 段階の近似を要することになるため実用的でない。よって、いくつかのリサンプリング法が提案されているのだが、これは次節の話題である。

7. リサンプリング法

以上見てきたように、接合関数の従属パラメータの区間推定や一般のクラスに属するかどうかの有意性検定を行うためには、何らかの形でリサンプリング法を適用しなければならないことが多い。本節では、経験接合過程のブートストラップ法についてこれまで提案されてきたものを再考察した上で、新たに我々が提案している経験ベータ接合関数に基づくリサンプリング手法についても簡単に述べる。

7.1 直接的ブートストラップ

\$(W_{n1}, \dots, W_{nn})\$ を多項分布 \$\text{Mult}(n; 1/n, \dots, 1/n)\$ に従う確率ベクトルとし、標本 \$\mathbf{X}_1, \dots, \mathbf{X}_n\$ とは独立であるとする。そして、

$$(7.1) \quad \mathbb{C}_n^*(\mathbf{u}) = \mathbb{F}_n^*(\mathbb{F}_{n1}^{*-}(u_1), \dots, \mathbb{F}_{nd}^{*-}(u_d)),$$

ただし

$$\mathbb{F}_n^*(\mathbf{x}) := \frac{1}{n} \sum_{i=1}^n W_{ni} \prod_{j=1}^d \mathbb{1}\{X_{ij} \leq x_j\},$$

$$\mathbb{F}_{nj}^*(x_j) := \frac{1}{n} \sum_{i=1}^n W_{ni} \mathbb{1}\{X_{ij} \leq x_j\}, \quad j \in \{1, \dots, d\}$$

とおく。

ブートストラップに関する漸近論のために、van der Vaart and Wellner (1996, Section 2.9) にならって、データを所与とした条件付き分布の確率収束を定義する。まず、

$$\text{BL}_1 := \{h: \ell^\infty([0, 1]^d) \rightarrow \mathbb{R}, \|h\|_\infty \leq 1 \text{ かつ} \\ \text{すべての } x, y \in \ell^\infty([0, 1]^d) \text{ に対して } |h(x) - h(y)| \leq \|x - y\|_\infty\}$$

とおく。 \hat{X}_n が W をランダムな重みとする、 $\ell^\infty([0, 1]^d)$ におけるブートストラップ過程の列であるとき、記号

$$\hat{X}_n \underset{W}{\overset{P}{\rightsquigarrow}} X, \quad n \rightarrow \infty$$

は

$$\sup_{h \in \text{BL}_1} |E_W h(\hat{X}_n) - E h(X)| \rightarrow 0 \quad (\text{外確率収束})$$

かつ、すべての $h \in \text{BL}_1$ に対して

$$E_W h(\hat{X}_n)^* - E_W h(\hat{X}_n)_* \xrightarrow{P} 0$$

が成り立つことを意味する。ここで、 E_W はデータ $\mathbf{X}_1, \dots, \mathbf{X}_n$ を所与とした、重み W に関する条件付き期待値を示す記号であり、 $h(\hat{X}_n)^*$ と $h(\hat{X}_n)_*$ はそれぞれ全データ $\mathbf{X}_1, \dots, \mathbf{X}_n, W$ に関する最小可測優関数 (minimal measurable majorant) と最大可測劣関数 (maximal measurable minorant) である。

Fermanian et al. (2004) は、 $\ell^\infty([0, 1]^d)$ において

$$(7.2) \quad \sqrt{n}(\mathbb{C}_n^* - \mathbb{C}_n) \underset{W}{\overset{P}{\rightsquigarrow}} \mathbb{G}_C, \quad n \rightarrow \infty$$

が成り立つことを示した。

いま (W_{n1}, \dots, W_{nn}) は上と同様の確率ベクトルとし、順位に基づく経験接合関数 (3.3) のブートストラップ版を

$$(7.3) \quad \tilde{\mathbb{C}}_n^*(\mathbf{u}) = \frac{1}{n} \sum_{i=1}^n W_{ni} \prod_{j=1}^d \mathbb{1} \left\{ \frac{R_{ij,n}^*}{n} \leq u_j \right\}$$

で定義する。ここで、

$$(7.4) \quad R_{ij,n}^* = \sum_{k=1}^n W_{nk} \mathbb{1} \{X_{kj} \leq X_{ij}\}.$$

である。ブートストラップ標本には高い確率で同順位が起こるため、 \mathbb{C}_n^* と $\tilde{\mathbb{C}}_n^*$ については、式 (3.5) の不等式は成り立たないが、

$$(7.5) \quad \sup_{\mathbf{u} \in [0, 1]^d} |\mathbb{C}_n^*(\mathbf{u}) - \tilde{\mathbb{C}}_n^*(\mathbf{u})| = O_p(n^{-1} \log n), \quad n \rightarrow \infty.$$

が成り立つことを示すことができる (証明は Kiriliouk et al., 2019 参照)。このため、経験接合関数として (3.2)、(3.3) のどちらをブートストラップしても、漸近的には同等となる。

7.2 偏導関数の推定値を用いた乗数ブートストラップ法

このアプローチは Remillard and Scaillet (2009) で提案されたものである。 Z_1, \dots, Z_n を、データとは独立な iid 確率変数で、 $E(Z_i) = 0$ 、 $\text{Var}(Z_i) = 1$ と

$$\|Z_i\|_{2,1} := \int_0^\infty \sqrt{\mathbb{P}(|Z_i| > x)} dx < \infty$$

を満たすものとする。そして、

$$(7.6) \quad \tilde{C}_n^Z(\mathbf{u}) := \frac{1}{n} \sum_{i=1}^n Z_i \prod_{j=1}^d \mathbb{1} \left\{ \frac{R_{ij}^{(n)}}{n} \leq u_j \right\}$$

と

$$(7.7) \quad \tilde{\beta}_n(\mathbf{u}) := \sqrt{n}[\tilde{C}_n^Z(\mathbf{u}) - \bar{Z}_n \tilde{C}_n(\mathbf{u})] = \frac{1}{\sqrt{n}} \sum_{i=1}^n (Z_i - \bar{Z}_n) \prod_{j=1}^d \mathbb{1} \left\{ \frac{R_{ij}^{(n)}}{n} \leq u_j \right\},$$

を定義する。ただし、 $\bar{Z}_n := n^{-1} \sum_{i=1}^n Z_i$ である。

Remillard and Scaillet (2009)への補足。Remillard and Scaillet (2009, p. 383) にちよつとした誤りが存在するため ($C_n(u) = \dots$ から始まるページ中段の式) ここで修正しておく。 $U_{ij} := F_j(X_{ij})$ とおくと、 (U_{i1}, \dots, U_{id}) , $i \in \{1, \dots, n\}$ は分布関数 C に従う iid 確率ベクトルとなる。さらに、

$$\mathbb{G}_n(\mathbf{u}) := \frac{1}{n} \sum_{i=1}^n \prod_{j=1}^d \mathbb{1}\{U_{ij} \leq u_j\}, \quad \mathbb{G}_{nj}(u_j) := \frac{1}{n} \sum_{i=1}^n \mathbb{1}\{U_{ij} \leq u_j\}$$

とおくと、 $\ell^\infty([0, 1]^d)$ において $\mathbb{U}_n^C := \sqrt{n}(\mathbb{G}_n - C) \rightsquigarrow \mathbb{U}^C$ が成り立つ。このとき、 $u_j \in [0, 1]$, $j \in \{1, \dots, d\}$ と $i \in \{1, \dots, n\}$ に対して

$$\{U_{ij} \leq \mathbb{G}_n^-(u_j)\} = \{X_{ij} \leq \mathbb{F}_n^-(u_j)\}, \text{ a.s.}$$

が成り立つことを見るのは容易である。よって、

$$\mathbb{G}_n(\mathbb{G}_{n1}^-(u_1), \dots, \mathbb{G}_{nd}^-(u_d)) = \mathbb{F}_n(\mathbb{F}_{n1}^-(u_1), \dots, \mathbb{F}_{nd}^-(u_d)), \text{ a.s.}$$

が結論される。ここで、右辺は定義より $\mathbb{C}_n(\mathbf{u})$ に等しいことに注意する。

ここで、

$$\mathbb{U}_n^Z(\mathbf{u}) := \frac{1}{\sqrt{n}} \sum_{i=1}^n Z_i \left(\prod_{j=1}^d \mathbb{1}\{U_{ij} \leq u_j\} - \mathbb{G}_n(\mathbf{u}) \right)$$

と定義すると、

$$\mathbb{U}_n^Z(\mathbf{u}) = \frac{1}{\sqrt{n}} \sum_{i=1}^n Z_i \left(\prod_{j=1}^d \mathbb{1}\{U_{ij} \leq u_j\} - C(\mathbf{u}) \right) - \bar{Z}_n \sqrt{n}(\mathbb{G}_n(\mathbf{u}) - C(\mathbf{u})),$$

となり、右辺第 2 項は一樣に 0 に確率収束する。ゆえに、 $\ell^\infty([0, 1]^d)$ において $\mathbb{U}_n^Z \rightsquigarrow \mathbb{U}^C$ が成立する。Remillard and Scaillet (2009) は、 \mathbb{U}^C と $\mathbb{U}^{C'}$ を 2 つの互いに独立な C でピン止めされたブラウンシートとして、“無条件の”収束 $(\mathbb{U}_n^C, \mathbb{U}_n^{C'}) \rightsquigarrow (\mathbb{U}^C, \mathbb{U}^{C'})$ を主張しているが、これは van der Vaart and Wellner (1996, Corollary 2.9.3), あるいはより直接的に Kosorok (2008, Corollary 10.3) を適用すれば直ちに得られる結果である。

$$(7.8) \quad \beta_n(\mathbf{u}) := \frac{1}{\sqrt{n}} \sum_{i=1}^n Z_i \left(\prod_{j=1}^d \mathbb{1}\{U_{ij} \leq \mathbb{G}_{nj}^-(u_j)\} - \mathbb{C}_n(\mathbf{u}) \right)$$

とくと、 $\beta_n(\mathbf{u}) = \mathbb{U}_n^Z(\mathbb{G}_{n1}^-(u_1), \dots, \mathbb{G}_{nd}^-(u_d))$ が成り立ち、(合成写像に関する)連続写像定理と $\|\mathbb{G}_{nj}^- - I\|_\infty \rightarrow 0$, a.s. (I は $[0, 1]$ 上の恒等関数) という事実、そして \mathbb{U}^C のパスの連続性から、 $\ell^\infty([0, 1]^d)$ において $\beta_n \rightsquigarrow \mathbb{U}^C$ が結論される。

Remillard and Scaillet (2009) では、下記の偏導関数推定量を用いた確率過程の無条件の収束を示しているが、ここではより興味深い条件付きの収束が成立することに注意する。すなわち、Kosorok (2008, Theorem 10.4) より、 $\ell^\infty([0, 1]^d)$ において、

$$\mathbb{U}_n^Z \xrightarrow[Z]{P} \mathbb{U}^C, \quad n \rightarrow \infty$$

であり、 $\|\mathbb{G}_{nj}^- - I\|_\infty \rightarrow 0$, a.s. が成り立つことより、 $\beta_n \xrightarrow[Z]{P} \mathbb{U}^C$ が従う。

各 $j \in \{1, \dots, d\}$ に対して、偏導関数 $\dot{C}_j(\mathbf{u}) = \partial C / \partial u_j(\mathbf{u})$ は

$$\hat{C}_j(\mathbf{u}) = \begin{cases} \frac{C_n(u_1, \dots, u_{j-1}, 2h, u_{j+1}, \dots, u_d)}{2h}, & u_j \in [0, h) \\ \frac{C_n(u_1, \dots, u_{j-1}, u_j + h, u_{j+1}, \dots, u_d) - C_n(u_1, \dots, u_{j-1}, u_j - h, u_{j+1}, \dots, u_d)}{2h}, & u_j \in [h, 1-h) \\ \frac{C_n(u_1, \dots, u_{j-1}, 1, u_{j+1}, \dots, u_d) - C_n(u_1, \dots, u_{j-1}, 1-2h, u_{j+1}, \dots, u_d)}{2h}, & u_j \in (1-h, 1] \end{cases}$$

で推定できる。 $h \asymp n^{-1/2}$ ととると、すべての \hat{C}_j , $j \in \{1, \dots, d\}$ が $[0, 1]^d$ 上で連続であるならば、 $\|\hat{C}_j - \dot{C}_j\|_\infty \xrightarrow{P} 0$ となる。最後に

$$\alpha_n^{\text{pdm}}(\mathbf{u}) := \beta_n(\mathbf{u}) - \sum_{j=1}^d \hat{C}_j(\mathbf{u}) \beta_n(\mathbf{1}, u_j, \mathbf{1})$$

とくと、 $\|\hat{C}_j - \dot{C}_j\|_\infty \xrightarrow{P} 0$ であるから、Slutsky の補題と連続写像定理により、 $\alpha_n^{\text{pdm}} \xrightarrow[Z]{P} \mathbb{G}^C$ が得られる。

上記の結果は標準の経験接合関数(3.2)に対するブートストラップ中心極限定理である。次に、順位に基づく経験接合関数(3.3)について、(7.7)で定義された $\tilde{\beta}_n$ についても同様の結果が成り立つことを示そう。まず、

$$(7.9) \quad \beta_n(\mathbf{u}) - \tilde{\beta}_n(\mathbf{u}) = \frac{1}{\sqrt{n}} \sum_{i=1}^n Z_i \left[\prod_{j=1}^d \mathbb{1}\{X_{ij} \leq \mathbb{F}_{nj}^-(u_j)\} - \prod_{j=1}^d \mathbb{1}\{\mathbb{F}_{nj}(X_{ij}) \leq u_j\} \right] - \frac{1}{\sqrt{n}} \sum_{i=1}^n Z_i (C_n(\mathbf{u}) - \tilde{C}_n(\mathbf{u})), \quad \text{a.s.}$$

となり、右辺の第2項は(3.5)により $O_p(n^{-1})$ である。さらに、(7.5)の証明と同様の議論により、(7.9)の右辺第1項の絶対値は $d/\sqrt{n} \max_{1 \leq i \leq n} |Z_i|$ によって押さえられる。 Z_1, Z_2, \dots は同じ分布に従い(ここでは独立性は不要である)、 $E(Z_i^2) < \infty$ を満たすから、

$$\max_{1 \leq i \leq n} \frac{|Z_i|}{\sqrt{n}} \xrightarrow{P} 0$$

となる。よって、(7.9)の右辺第1項は一様に $o_p(1)$ であり、 $\|\beta_n - \tilde{\beta}_n\|_\infty \xrightarrow{P} 0$ 、そして $\tilde{\beta}_n \xrightarrow[Z]{P} \mathbb{U}^C$

が成り立つ.

$$\tilde{\alpha}_n^{\text{pdm}}(\mathbf{u}) := \tilde{\beta}_n(\mathbf{u}) - \sum_{j=1}^d \hat{C}_j(\mathbf{u}) \tilde{\beta}_n(\mathbf{1}, u_j, \mathbf{1})$$

と定義すると, 上と同様の議論から $\tilde{\alpha}_n^{\text{pdm}} \xrightarrow[\mathbb{Z}]{\mathbb{P}} \mathbb{G}^C$ が得られる.

注意 7.1. 同じ論法を用いて, 一般の交換可能な重み (W_{n1}, \dots, W_{nn}) に対して,

$$\|\mathbb{C}_n^* - \tilde{\mathbb{C}}_n^*\|_\infty = o_p(n^{-1/2})$$

を示すことができる (c.f. (7.5)).

Bücher と Dette による修正. Bücher and Dette (2010)では, より使い勝手の良い次のような改良版が提案されている. ξ_1, \dots, ξ_n を, データとは独立な正の iid 確率変数で, $E(\xi_i) = \mu$, $\text{Var}(\xi_i) = \tau^2 > 0$ と $\|\xi_i\|_{2,1} < \infty$ を満たすものとし, $\bar{\xi}_n := n^{-1} \sum_{i=1}^n \xi_i$,

$$\mathbb{C}_n^\circ(\mathbf{u}) := \frac{1}{n} \sum_{i=1}^n \frac{\xi_i}{\bar{\xi}_n} \prod_{j=1}^d \mathbb{1}\{X_{ij} \leq \mathbb{F}_{nj}^-(u_j)\}$$

とおく. ξ_i を標準化して $Z_i = (\xi_i - \mu)/\tau$ と書くと,

$$\frac{\mu}{\tau} \sqrt{n}(\mathbb{C}_n^\circ(\mathbf{u}) - \mathbb{C}_n(\mathbf{u})) = \frac{\mu}{\bar{\xi}_n} \cdot \frac{1}{\sqrt{n}} \sum_{i=1}^n (Z_i - \bar{Z}_n) \prod_{j=1}^d \mathbb{1}\{X_{ij} \leq \mathbb{F}_{nj}^-(u_j)\} = \frac{\mu}{\bar{\xi}_n} \cdot \beta_n(\mathbf{u})$$

となる. よって, $\beta_n^\circ := \sqrt{n}(\mu/\tau)(\mathbb{C}_n^\circ - \mathbb{C}_n)$ は実際 β_n と近いことがわかる. 同様に,

$$\tilde{\mathbb{C}}_n^\circ := \frac{1}{n} \sum_{i=1}^n \frac{\xi_i}{\bar{\xi}_n} \prod_{j=1}^d \mathbb{1}\{\mathbb{F}_{nj}(X_{ij}) \leq u_j\}$$

とおけば,

$$\frac{\mu}{\tau} \sqrt{n}(\tilde{\mathbb{C}}_n^\circ(\mathbf{u}) - \tilde{\mathbb{C}}_n(\mathbf{u})) = \frac{\mu}{\bar{\xi}_n} \cdot \tilde{\beta}_n(\mathbf{u})$$

となり, $\tilde{\beta}_n^\circ := \sqrt{n}(\mu/\tau)(\tilde{\mathbb{C}}_n^\circ - \tilde{\mathbb{C}}_n)$ は $\tilde{\beta}_n$ とほぼ同等である.

Bücher and Dette (2010)の主張である $\beta_n^\circ \xrightarrow[\xi]{\mathbb{P}} \mathbb{U}^C$ は Kosorok (2008, Theorem 2.6)と $\|\mathbb{F}_{nj}^- - I\|_\infty \rightarrow 0$, a.s. を用いれば従う. また, $\tilde{\beta}_n^\circ \xrightarrow[\xi]{\mathbb{P}} \mathbb{U}^C$ も成り立つから,

$$\alpha_n^{\text{pdm}^\circ}(\mathbf{u}) := \beta_n^\circ(\mathbf{u}) - \sum_{j=1}^d \hat{C}_j(\mathbf{u}) \beta_n^\circ(\mathbf{1}, u_j, \mathbf{1})$$

と

$$\tilde{\alpha}_n^{\text{pdm}^\circ}(\mathbf{u}) := \tilde{\beta}_n^\circ(\mathbf{u}) - \sum_{j=1}^d \hat{C}_j(\mathbf{u}) \tilde{\beta}_n^\circ(\mathbf{1}, u_j, \mathbf{1})$$

が \mathbb{G}^C を条件付きで近似することがわかる.

7.3 他のリサンプリング法

Bücher and Dette (2010)では第3の方法も提案されているが、彼らのシミュレーション実験によれば、有限標本では偏導関数の推定値を用いた乗数ブートストラップ法が最良という結果が報告されているので、ここでは省略する。

例 6.1 (iii) の適合度検定に対しては、

$$\int_{[0,1]^d} [C_n(\mathbf{u}) - C_{\hat{\theta}_n}(\mathbf{u})]^2 dC_n(\mathbf{u})$$

の形の検定統計量が最も検出力が大きいとの大規模シミュレーション結果がある (Genest et al., 2009). その形から明らかなように、このタイプの統計量については、擬似観測値に基づくパラメトリック・ブートストラップ法が有効である。また、5.1 節で述べた順位近似 Z 推定量、あるいはその特殊ケースである疑似尤度推定量に基づく信頼区間を構成する場合にも、同方法を適用することが可能である。

最後に、(3.6) 式で定義した経験ベータ接合関数 C_n^β を用いたリサンプリング法を簡潔に説明しよう。実際、分布関数 C_n^β に従う \mathbf{V} を発生させるのは極めて簡単である：順位 $R_{ij,n} = r_{ij}$, $j = 1, \dots, d$ を所与として、

1. $\{1, \dots, n\}$ から無作為に1つを選び、 I とする。
2. それとは独立に $V_j^\# \sim B(r_{Ij}, n+1-r_{Ij})$, $j \in \{1, \dots, d\}$ を発生させる。
3. $\mathbf{V}^\# = (V_1^\#, \dots, V_d^\#)$ とおく。

このアルゴリズムを n 回繰り返せば、データ $\mathbf{X}_1, \dots, \mathbf{X}_n$ を所与として、 C_n^β からの n 個の独立な確率ベクトル $\mathbf{V}_i^\# = (V_{i1}^\#, \dots, V_{id}^\#)$, $i = 1, \dots, n$ を発生させることができる。経験ベータ接合関数は経験接合関数の平滑化版と考えることができるので、上のリサンプリング法は一種の平滑化ブートストラップ (smoothed bootstrap) とみなすこともできる (Efron, 1982 や Shao and Tu, 1995, Section 3.5 を参照せよ)。この方法では同順位が起こることがないため、安定的にリサンプリングすることができる。Kiriliouk et al. (2019) では、この方法に関する漸近的妥当性を証明し、いくつかの区間推定や対称性の検定について包括的なモンテカルロ・シミュレーションを行って比較分析しており、 C_n^β を用いたリサンプリング法は多くの場合に有望な選択肢となり得ることが示されている。

謝 辞

査読者の有益なコメントに感謝する。また、本研究は JSPS 科研費 JP18H00836 の助成を受けたものである。

参 考 文 献

- Barnard, G. A. (1963). Some logical aspects of the fiducial argument, *Journal of the Royal Statistical Society (B)*, **25**, 111–115.
- Barndorff-Nielsen, O. E., Blæsild, P. and Eriksen, P. S. (1989). *Decomposition and Invariance of Measures, and Statistical Transformation Models*, Springer-Verlag, Berlin, Heidelberg.
- Blum, J. R., Kiefer, J. and Rosenblatt, M. (1961). Distribution free tests of independence based on the sample distribution function, *Annals of Mathematical Statistics*, **32**, 485–498.
- Bücher, A. and Dette, H. (2010). A note on bootstrap approximations for the empirical copula process,

- Statistics and Probability Letters*, **80**, 1925–1932.
- Bücher, A. and Kojadinovic, I. (2015). An overview of nonparametric tests of extreme-value dependence and of some related statistical procedures, *Extreme Value Modeling and Risk Analysis: Methods and Applications* (eds. D. Day and J. Yan), 377–398, Chapman & Hall/CRC, Boca Raton, Florida.
- Bücher, A., Dette, H. and Volgushev, S. (2012). A test for archimedeanity in bivariate copula models, *Journal of Multivariate Analysis*, **110**, 121–132.
- Chen, X., Fan, Y. and Tsyrennikov, V. (2006). Efficient estimation of semiparametric multivariate copula models, *Journal of the American Statistical Association*, **101**, 1228–1240.
- Deheuvels, P. (1979). La fonction de dépendance empirique et ses propriétés: Un test non paramétrique d'indépendance, *Académie Royale de Belgique, Bulletin de la Classe des Sciences, 5e série*, **65**, 274–292.
- Efron, B. (1982). *The Jackknife, the Bootstrap and Other Resampling Plans*, Society for Industrial and Applied Mathematics, Philadelphia.
- Fermanian, J.-D. (2013). An overview of the goodness-of-fit test problem for copulas, *Copulae in Mathematical and Quantitative Finance* (eds. P. Jaworski, F. Durante and W. K. Härdle), 61–89, Springer-Verlag, Berlin, Heidelberg.
- Fermanian, J.-D., Radulović, D. and Wegkamp, M. J. (2004). Weak convergence of empirical copula processes, *Bernoulli*, **10**, 847–860.
- Gaenssler, P. and Stute, W. (1987). *Seminar on Empirical Processes*, Birkhäuser, Basel.
- Genest, C. (1987). Frank's family of bivariate distributions, *Biometrika*, **74**, 549–555.
- Genest, C., Ghoudi, K. and Rivest, L.-P. (1995). A semiparametric estimation procedure of dependence parameters in multivariate families of distributions, *Biometrika*, **82**, 543–552.
- Genest, C., Rémillard, B. and Beaudoin, D. (2009). Goodness-of-fit tests for copulas: Review and a power study, *Insurance: Mathematics and Economics*, **44**, 199–213.
- Genest, C., Nešlehová, J. and Quessy, J.-F. (2012). Tests of symmetry for bivariate copulas, *Annals of the Institute of Statistical Mathematics*, **64**, 811–834.
- Ghoudi, K. and Rémillard, B. (2004). Empirical processes based on pseudo-observations II: The multivariate case, *Asymptotic Methods in Stochastics: Festschrift for Miklós Csörgő*, Fields Institute Communications, Vol. 44 (eds. L. Horváth and B. Szyszkowicz), 381–406, American Mathematical Society, Providence, Rhode Island.
- Hall, W. J., Wijsman, R. A. and Ghosh, J. K. (1965). The relationship between sufficiency and invariance with applications in sequential analysis, *Annals of Mathematical Statistics*, **36**, 575–614.
- Harder, M. and Stadtmüller, U. (2017). Testing exchangeability of copulas in arbitrary dimension, *Journal of Nonparametric Statistics*, **29**, 40–60.
- Hofert, M., Kojadinovic, I., Maechler, M., Yan, J. and Nešlehová, J. G. (2018). Package 'copula', R package version 0.999-19, <http://CRAN.R-project.org/package=copula>.
- Hofert, M., Kojadinovic, I., Mächler, M. and Yan, J. (2019). *Elements of Copula Modeling with R*, Springer International Publishing, Cham, Switzerland.
- Hoff, P. D. (2007). Extending the rank likelihood for semiparametric copula estimation, *Annals of Applied Statistics*, **1**, 265–283.
- Kendall, M. and Gibbons, J. D. (1990). *Rank Correlation Methods*, 5th ed., Oxford University Press, New York.
- Kiriliouk, A., Segers, J. and Tsukahara, H. (2019). On some resampling procedures with the empirical beta copula, <http://arxiv.org/abs/1905.12466>.
- Kojadinovic, I. and Yan, J. (2010). Comparison of three semiparametric methods for estimating dependence parameters in copula models, *Insurance: Mathematics and Economics*, **47**, 52–63.

- Kosorok, M. R. (2008). *Introduction to Empirical Processes and Semiparametric Inference*, Springer Science+Business Media Inc., New York.
- Lehmann, E. L. (1986). *Testing Statistical Hypotheses*, 2nd ed., John Wiley & Sons, New York.
- 鍋谷清治 (1978). 『数理統計学』, 共立出版, 東京.
- Nelsen, R. B. (2006). *An Introduction to Copulas*, 2nd ed., Springer Science+Business Media Inc., New York.
- Oakes, D. (1982). A model for association in bivariate survival data, *Journal of the Royal Statistical Society (B)*, **44**, 414–422.
- Rémillard, B. and Scaillet, O. (2009). Testing for equality between two copulas, *Journal of Multivariate Analysis*, **100**, 377–386.
- Rémon, M. (1984). On a concept of partial sufficiency: L-sufficiency, *International Statistical Review*, **52**, 127–135.
- Segers, J. (2012). Asymptotics of empirical copula processes under nonrestrictive smoothness assumptions, *Bernoulli*, **18**, 764–782.
- Segers, J., Sibuya, M. and Tsukahara, H. (2017). The empirical beta copula, *Journal of Multivariate Analysis*, **155**, 35–51.
- Severini, T. A. (2000). *Likelihood Methods in Statistics*, Oxford University Press, Oxford.
- Shao, J. and Tu, D. (1995). *The Jackknife and Bootstrap*, Springer-Verlag, New York.
- Shorack, G. R. and Wellner, J. A. (1986). *Empirical Processes with Applications to Statistics*, John Wiley & Sons, New York.
- Sprott, D. A. (1965). Transformations and sufficiency, *Journal of the Royal Statistical Society (B)*, **27**, 479–485.
- Tsukahara, H. (2000). *Empirical copulas and some applications*, Research Report No. 27, The Institute for Economic Studies, Seijo University, <http://www.seijo.ac.jp/research/economics/publications/reserch-report/jtmo420000000mul-att/green27.pdf>.
- Tsukahara, H. (2005). Semiparametric estimation in copula models, *Canadian Journal of Statistics*, **33**, 357–375 (Erratum: *ibid.* (2011). **39**, 734–735).
- 塚原英敦 (2008). 接合分布関数(コピュラ)の理論と応用, 『21世紀の統計科学 III: 数理・計算の統計科学』(北川源四郎・竹村彰通 編), 111–146, 東京大学出版会, 東京.
- van der Vaart, A. W. and Wellner, J. A. (1996). *Weak Convergence and Empirical Processes: With Applications to Statistics*, Springer-Verlag, New York.

Statistical Inference in Copula Models

Hideatsu Tsukahara

Faculty of Economics, Seijo University

In this paper, we provide an exposition of statistical inference procedures on the copula associated with a multivariate distribution from which we have a random sample, without any knowledge of the marginal distributions. According to the \mathcal{G} -sufficiency criterion, the inference should be based on vectors of coordinatewise ranks only. The empirical copula and its variants are nonparametric estimators of copula, and depend on the data only through their ranks. Their asymptotic properties are briefly reviewed. Then estimation of functionals of the copula such as rank correlation coefficients are discussed. Semiparametric estimation of the dependence parameter of copula is also considered with a special focus on rank approximate Z -estimators which includes the well-known pseudo-likelihood estimators. After a short review of general goodness-of-fit testing on copulas, we conclude the paper with a detailed examination of some resampling schemes based on the empirical and related copulas, which are indispensable for carrying out the aforementioned inference procedures in practice.

セミパラメトリックコピュラモデルにおける ダイバージェンスの性質

清 智也[†]・松本 和也[†]

(受付 2019 年 5 月 31 日；改訂 9 月 27 日；採択 10 月 4 日)

要 旨

多次元の量的データに対し、コピュラとしてパラメトリックモデルを仮定し、周辺分布には何も仮定をおかない統計モデルのことをセミパラメトリックコピュラモデルという。本論文ではこのモデルにおけるダイバージェンスの性質を考察する。特に、多変量順位統計量の周辺分布によって定まる順位ダイバージェンスと、Kullback–Leibler ダイバージェンスの局外パラメータに関する最小値として定義されるプロファイルダイバージェンスの関係を調べる。また区分一様コピュラとガウスコピュラの場合について具体的な計算結果を示す。

キーワード：コピュラ，最適輸送理論，情報幾何，ダイバージェンス，複合変換モデル，ホロノミック勾配法。

1. はじめに

d 次元の確率密度関数 $c(x)$ ($x = (x_1, \dots, x_d) \in [0, 1]^d$) がコピュラ密度であるとは、全ての 1 次元周辺密度が $[0, 1]$ 上の一様密度になることと定義する。Sklar の定理から、 \mathbb{R}^d 上の任意の確率密度関数 $p(x) = p(x_1, \dots, x_d)$ は

$$(1.1) \quad p(x_1, \dots, x_d) = c(F_1(x_1), \dots, F_d(x_d)) \prod_{i=1}^d F'_i(x_i)$$

という形で一意的に表すことができる。ここで c はコピュラ密度、 F_i は x_i の周辺分布関数を表す。また F'_i は F_i の導関数である。コピュラ密度 c にパラメトリックモデルを仮定し、周辺分布 F_i には仮定をおかない統計モデルのことをセミパラメトリックコピュラモデル (semiparametric copula model) という。より正確な定義は式 (3.2) で与える。

セミパラメトリックコピュラモデルのパラメータ推定については様々な研究がある。Klaassen and Wellner (1997) は 2 変量ガウスコピュラの場合に擬似最尤推定量および正規スコア順位相関係数が漸近有効推定量を与えることを示した。Genest and Werker (2002) は、ガウスコピュラ以外では擬似最尤推定量が必ずしも漸近有効でないことを指摘した。Chen et al. (2006) は周辺分布の推定に sieve を用いることで漸近有効推定量を構成した。Tsukahara (2005) は順位統計量のみ依存した推定量のクラスを与え、その漸近的性質を明らかにした。順位統計量のみ依存した漸近有効推定量の構成は重要な未解決問題となっている。ガウスコピュラの部分モデルに限れば、Hoff et al. (2014) が漸近有効情報量の特徴付け、Segers et al. (2014) が実際に漸

[†] 東京大学大学院 情報理工学系研究科：〒113-8656 東京都文京区本郷 7-3-1

近有効推定量を構成した。その他、コピュラに関する総合的な解説は塚原 (2011) を参照されたい。

本論文ではセミパラメトリックコピュラモデルにおけるダイバージェンスの性質を調べる。Kullback–Leibler ダイバージェンスを規範として、順位ダイバージェンスとプロファイルダイバージェンスという 2 つの量を定義する。これらはそれぞれ順位尤度 (Hoff, 2007) とプロファイル尤度の母集団特性を表したものである。いずれも一般には陽に計算できないが、区分一様コピュラに限れば具体的に計算を進めることができる。その結果、正則条件のもとで順位ダイバージェンスがプロファイルダイバージェンスに収束することが示される (定理 2)。またガウスコピュラの場合、プロファイルダイバージェンスの計算は有限次元の最適化問題に帰着され、順位ダイバージェンスは多次元正規分布の象限確率で表される。後者はホロノミック勾配法を利用して計算できる。以上の結果は推定量の性能を知るための手がかりを与えるものと考えられる。

論文の構成は以下の通りである。まず 2 節では簡単な例を通してセミパラメトリックコピュラモデルにおけるダイバージェンスを論ずる必要性を確認する。3 節では実際にダイバージェンスを定義し、基本的な定理を与える。4 節と 5 節ではそれぞれ区分一様コピュラとガウスコピュラの場合について具体的な計算結果を示す。最後に 6 節で今後の課題を述べる。

2. 簡単な例

簡単な例として、次のような 2 次元の区分一様分布を考えてみよう。正方形領域 $[0, 1]^2$ を 4 つの小さい正方形に分割し、それぞれの領域で定数となるようなコピュラ密度

$$c(x_1, x_2) = \begin{cases} 1.8 & \text{if } (x_1, x_2) \in [0, \frac{1}{2}]^2 \cup [\frac{1}{2}, 1]^2, \\ 0.2 & \text{otherwise} \end{cases}$$

を考える (図 1(a))。また 1 次元の分布関数として

$$(2.1) \quad F_1(\xi) = F_2(\xi) = \begin{cases} \frac{2}{3}\xi & \text{if } \xi \in [0, \frac{3}{4}), \\ \frac{1}{2} + 2(\xi - \frac{3}{4}) & \text{if } \xi \in [\frac{3}{4}, 1] \end{cases}$$

を選び、式 (1.1) によって密度 $p(x_1, x_2)$ を定義すれば、

$$p(x_1, x_2) = \begin{cases} 0.8 & \text{if } (x_1, x_2) \in [0, \frac{3}{4}]^2, \\ 0.8/3 & \text{if } (x_1, x_2) \in ([0, \frac{3}{4}) \times [\frac{3}{4}, 1]) \cup ([\frac{3}{4}, 1] \times [0, \frac{3}{4})) \\ 7.2 & \text{if } (x_1, x_2) \in [\frac{3}{4}, 1]^2, \end{cases}$$

となる (図 1(b))。 $c(x_1, x_2)$ と $p(x_1, x_2)$ は周辺分布を変えただけであり、式 (1.1) のコピュラ部分は共通であることに注意する。

ところが、いま $u(x_1, x_2) = 1$ を一様密度 (独立コピュラ) とすれば、 u と c の間の Kullback–Leibler ダイバージェンスは

$$\begin{aligned} \text{KL}(u, c) &= \int_{[0,1]^2} u(x) \log \frac{u(x)}{c(x)} dx \\ &= \frac{1}{2} \log \frac{1}{1.8} + \frac{1}{2} \log \frac{1}{0.2} \\ &\approx 0.511 \end{aligned}$$

となるのに対し、 u と p の間のダイバージェンスは

コピュラモデルにおけるダイバージェンスの性質

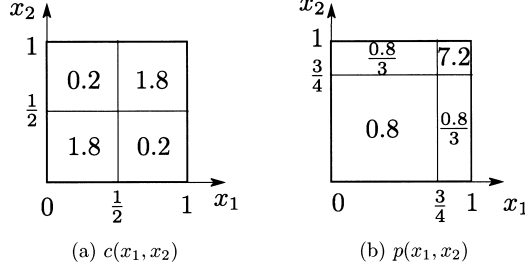


図 1. 周辺密度は異なるがコピュラ部分は等しい 2 つの密度関数. 各領域内の数値は密度関数の値を表す. 一様密度からの Kullback–Leibler ダイバージェンスは $c(x_1, x_2)$ よりも $p(x_1, x_2)$ の方が小さい.

$$\begin{aligned} \text{KL}(u, p) &= \left(\frac{3}{4}\right)^2 \log \frac{1}{0.8} + 2 \left(\frac{3}{4}\right) \left(\frac{1}{4}\right) \log \frac{1}{0.8/3} + \left(\frac{1}{4}\right)^2 \log \frac{1}{7.2} \\ &\approx 0.498 \end{aligned}$$

となる. つまり, u から見て c よりも p の方が少し近いことが分かる.

このように, 周辺分布を動かすことでダイバージェンスの値は変化する. その最小値を本稿ではプロファイルダイバージェンスと呼ぶ. 一方, 次節で詳しく見るように, セミパラメトリックコピュラモデルの情報は順位統計量に集約されていると考えることができる. そこで順位統計量に基づいて定義されるダイバージェンスを順位ダイバージェンスと呼ぶ. これらの関係を明らかにすることが本研究の主題である.

3. 順位ダイバージェンスとプロファイルダイバージェンス

ここではまず, セミパラメトリックコピュラモデルが座標ごとの変数変換に関して不変であることを説明する (Hoff, 2007; Hoff et al., 2014). よって複合変換モデル (付録 A) として扱うことができる. この観点から順位ダイバージェンスとプロファイルダイバージェンスという 2 つのダイバージェンスを定義する. また両者の関係を定理として与える.

3.1 セミパラメトリックコピュラモデル

セミパラメトリックコピュラモデルを改めて定義しよう. $[0, 1]^d$ 上の正值確率密度関数の全体を \mathcal{P} と表す. ただし密度関数 $p(x)$ が正值であるとはほとんどいたるところで $p(x) > 0$ という意味とする. また座標ごとの変数変換 $T(x) = (T_1(x_1), \dots, T_d(x_d))$ のうち, 各 $T_i : [0, 1] \rightarrow [0, 1]$ が単調増加, 全単射かつ T_i も T_i^{-1} も絶対連続であるようなものの全体を \mathcal{T} と表す. たとえば式 (2.1) で考えた周辺分布の組 (F_1, F_2) は \mathcal{T} の元である. \mathcal{T} は関数の合成に関して群をなす. 密度 $p \in \mathcal{P}$ と変換 $T \in \mathcal{T}$ に対し, 密度の押し出し $T_*p \in \mathcal{P}$ を

$$(3.1) \quad (T_*p)(x_1, \dots, x_d) = p(T_1^{-1}(x_1), \dots, T_d^{-1}(x_d)) \prod_{i=1}^d (T_i^{-1})'(x_i)$$

によって定義する. これは p に従う確率変数ベクトルを T で変数変換したときに得られる密度関数である. この写像 $(T, p) \mapsto T_*p$ によって \mathcal{T} は \mathcal{P} に作用する. この作用に関する軌道 (同値

類)を $[p] = \{T_*p \mid T \in \mathcal{T}\}$ と表す. Sklar の定理より, 各軌道にはコピュラ密度がただ一つ存在する. つまり, コピュラを指定することは軌道を指定することと同じである.

ここではコピュラと変換群の整合性を考えて $[0, 1]^d$ 上の密度関数に限定したが, \mathbb{R}^d 上の密度関数を扱うことももちろん可能である. その場合, \mathbb{R}^d から $(0, 1)^d$ への変換をあらかじめ一つ決めておけば, 結果として \mathbb{R}^d から \mathbb{R}^d への変数変換を扱うことができる.

以上の記法のもとで, セミパラメトリックコピュラモデルは,

$$(3.2) \quad \mathcal{M} = \{T_*c_\theta \mid \theta \in \Theta, T \in \mathcal{T}\}$$

と定義される. ただし $\{c_\theta \mid \theta \in \Theta\}$ はパラメトリックなコピュラ密度の族である. パラメトリックコピュラのいろいろな例は Nelsen (2006) を参考にされたい. 本稿では 4 節と 5 節でそれぞれ区分一様コピュラとガウスコピュラを扱う.

モデル \mathcal{M} は \mathcal{T} の作用に関して複合変換モデルとなる. 興味のあるパラメータは θ である. つまり, 個々の分布ではなく軌道に興味がある. そこで, 2 つの軌道 $[p], [q]$ の間のダイバージェンスを定義したい.

3.2 順位ダイバージェンス

$p \in \mathcal{P}$ を真の密度とし, p に従うランダム標本 $X = (x_{ti})_{1 \leq t \leq n, 1 \leq i \leq d}$ を考える. ここで n はサンプルサイズである. 連続分布を考えているので, 各 i に対して $\{x_{ti}\}_{t=1}^n$ は相異なる値を取るものと仮定してよい. 次の命題は 1 次元の場合にはよく知られており (たとえば Eaton, 1983; 吉田, 2006), 多次元の場合にも同様に示される. 最大不変量の定義は付録 A を参照せよ.

命題 1 (Hoff, 2007). セミパラメトリックコピュラモデルにおける最大不変量は, 多変量順位統計量

$$r_{ti} = \#\{s \in \{1, \dots, n\} \mid x_{si} \leq x_{ti}\}, \quad 1 \leq t \leq n, \quad 1 \leq i \leq d,$$

である. ただし $\#A$ は集合 A の要素数を表す. また行列表記で $R = (r_{ti})$ と表す.

以下では多変量順位統計量を単に順位統計量と呼ぶ. 真の密度が p のとき, 順位統計量 R の周辺分布を $\bar{p}_n(R)$ と表し, 順位尤度(rank likelihood)と呼ぶ. R の取り得る値は有限個(正確には $(n!)^d$ 個)であるから, \bar{p}_n は離散分布であることに注意しよう.

順位尤度は高次元の積分によって記述することができる. 実際, 順位統計量 R と矛盾しない標本 $X \in \mathbb{R}^{n \times d}$ の集合を $\{R(X) = R\}$ と表せば, 密度 p に対する順位尤度は

$$(3.3) \quad \bar{p}_n(R) = \int_{\{R(X)=R\}} \prod_{t=1}^n p(x_{t1}, \dots, x_{td}) dX$$

と書ける. 4 節で述べる通り, 区分一様分布の場合には積分を含まない形で書き下すことができる.

順位尤度を用いて, 順位ダイバージェンスを次のように定義する.

定義 1. 密度関数 $p, q \in \mathcal{P}$ と標本サイズ n に対し,

$$(3.4) \quad \begin{aligned} D_n([p], [q]) &= \frac{1}{n} \text{KL}(\bar{p}_n, \bar{q}_n) \\ &= \frac{1}{n} \sum_R \bar{p}_n(R) \log \frac{\bar{p}_n(R)}{\bar{q}_n(R)} \end{aligned}$$

を順位ダイバージェンス (rank divergence) という.

$D_n([p], [q])$ が well-defined であることは、 $\bar{p}_n(R)$ が $[p]$ の元の選び方に依存しないことから言える。また式(3.4)の右辺において n で割っているのは、順位統計量 R が $O(n)$ の情報を持っていることを考慮したものである。Kullback–Leibler ダイバージェンスの単調性と加法性 (Kullback, 1959) から

$$(3.5) \quad D_n([p], [q]) \leq \frac{1}{n} \text{KL}(p_n, q_n) = \text{KL}(p, q)$$

が成り立つ。ここで p_n, q_n はそれぞれ p, q からの標本 $X \in \mathbb{R}^{n \times d}$ の同時密度関数を表す。

順位ダイバージェンスは正值性(識別性)を満たさない。すなわち $[p] \neq [q]$ であっても $D_n([p], [q]) = 0$ となる場合がある。実際、次のような例が存在する。2次元コピュラ密度 p で対称性 $p(x_1, x_2) = p(1 - x_1, x_2)$ を満たし、かつ一様でないものを選ぶ(たとえば $p(x_1, x_2) = 1 + \cos(2\pi x_1) \cos(2\pi x_2)$ など)。このとき式(3.3)と p の対称性より $\bar{p}_2(R)$ は離散一様分布となる。しかし一様密度 u の順位尤度 $\bar{u}_2(R)$ も離散一様分布となるから $D_2([p], [u]) = 0$ となり、正值性が成り立たない。

一方、「全ての n に対して $D_n([p], [q]) = 0$ 」となるような $[p] \neq [q]$ は(正則条件のもとで)存在しない。これは後の定理1と定理2より示される。

3.3 プロファイルダイバージェンス

複合変換モデル(付録A)の類推により、次のダイバージェンスを考えることは自然である。

定義 2. 密度関数 $p, q \in \mathcal{P}$ に対してプロファイルダイバージェンスを

$$(3.6) \quad \begin{aligned} \tilde{D}([p], [q]) &= \inf_{T, U \in \mathcal{T}} \text{KL}(T_*p, U_*q) \\ &= \inf_{T \in \mathcal{T}} \text{KL}(T_*p, q) \end{aligned}$$

と定義する。2番目の等号は Kullback–Leibler ダイバージェンスの不変性 $\text{KL}(T_*p, T_*q) = \text{KL}(p, q)$ より従う。

p, q がともにコピュラ密度であったとしても一般には $\tilde{D}([p], [q]) < \text{KL}(p, q)$ となる。2節で挙げた例はそのような例である。

$\tilde{D}([p], [q]) = \text{KL}(p, q)$ が成り立つための p, q の条件を調べておこう。計算の便宜上、写像 $T \in \mathcal{T}$ による密度 $q \in \mathcal{P}$ の引き戻しを

$$(3.7) \quad (T^*q)(x) = q(T_1(x_1), \dots, T_d(x_d)) \prod_{i=1}^d T'_i(x_i)$$

と定義する。これは式(3.1)で定義した押し出しの逆演算である。 $\text{KL}(T_*p, q) = \text{KL}(p, T^*q)$ とすることに注意しよう。

命題 2. p, q は $(0, 1)^d$ において連続的微分可能な(コピュラ密度とは限らない)確率密度関数とする。このとき $\tilde{D}([p], [q]) = \text{KL}(p, q)$ が成り立つための必要条件は

$$(3.8) \quad \partial_i \log p_i(x_i) = \mathbb{E}_p[\partial_i \log q(x)|x_i], \quad i = 1, \dots, d,$$

である。ただし p_i は p の周辺密度、 ∂_i は x_i による偏微分、 $\mathbb{E}_p[\cdot]$ は密度 p に関する条件付き期待値を表す。さらに、 q が対数凹(log-concave)であれば式(3.8)は十分条件にもなる。

証明. 変分法による。 $T_i(x_i) = x_i + \delta T_i(x_i)$ とおき、 $\delta T_i(x_i)$ は開区間 $(0, 1)$ 内にコンパクト

なサポートを持つ滑らかな関数とする. $\text{KL}(p, T^*q)$ を δT_i について 1 次まで展開すると

$$(3.9) \quad \text{KL}(p, T^*q) = \int p(x) \log \frac{p(x)}{q(T(x)) \prod_i T'_i(x_i)} dx \\ \simeq \text{KL}(p, q) - \sum_i \int p(x) (\partial_i \log q(x)) \delta T_i(x_i) dx - \sum_i \int p(x) \delta T'_i(x_i) dx$$

$$(3.10) \quad = \text{KL}(p, q) + \sum_i \int \{-p(x) \partial_i \log q(x) + \partial_i p(x)\} \delta T_i(x_i) dx \\ = \text{KL}(p, q) + \sum_i \int p_i(x_i) \{E_p[-\partial_i \log q(x)|x_i] + \partial_i \log p_i(x_i)\} \delta T_i(x_i) dx_i$$

となる. ただし (3.10) の等号は部分積分と境界条件 $\delta T_i(0) = \delta T_i(1) = 0$ による. よって停留条件 (3.8) を得る. また q が対数凹ならば, \mathcal{T} 上の汎関数 $T \mapsto \text{KL}(p, T^*q)$ は式 (3.9) より凸になるので, 停留条件が満たされればそれが最適解となる. \square

上の証明の最後で, 集合 \mathcal{T} の凸性, および汎関数 $T \mapsto \text{KL}(p, T^*q)$ の凸性を用いた. これらは最適輸送理論の文脈では displacement convexity と呼ばれる性質である (McCann, 1997; Villani, 2003). また命題 2 は p, q のサポートが $[0, 1]^d$ でない場合にも同様に成立する.

命題 2 において特に q を $[0, 1]^d$ 上の一様密度とおけば, 式 (3.8) は $\partial_i p_i = 0$, つまり p がコピュラ密度であることと同値である. またこのとき $\text{KL}(p, q) = \int p(x) \log p(x) dx$ はエントロピーの -1 倍である. こうして次の結果を得る.

系 1. $p \in \mathcal{P}$ がコピュラ密度であるための必要十分条件は, p が軌道 $[p]$ の中でエントロピー最大となることである.

3.4 主結果

順位ダイバージェンスとプロファイルダイバージェンスに関する主定理を示す. 式 (3.5) より, 両者には一般に $D_n([p], [q]) \leq \tilde{D}([p], [q])$ という関係が成り立つことに注意しよう.

まずプロファイルダイバージェンスの正值性に関して次の定理が成り立つ.

定理 1. p, q は正值のコピュラ密度関数とし, q は上に有界かつ上半連続とする. このとき $\tilde{D}([p], [q]) = \text{KL}(p, T^*q)$ を満たす $T \in \mathcal{T}$ が存在する. 特に $p \neq q$ ならば $\tilde{D}([p], [q]) > 0$ となる.

以下の証明には最適輸送理論の帰結が用いられる. そこでは最小化問題 $\tilde{D}([p], [q]) = \inf_{T \in \mathcal{T}} \text{KL}(T^*p, q)$ が輸送写像 T に関するエネルギー最小化問題と解釈される.

証明. $p = q$ の場合は T を恒等写像とすればよい. 以下 $p \neq q$ とする. プロファイルダイバージェンスの定義より, $\text{KL}(p_m, q)$ が $\tilde{D}([p], [q])$ に収束するような密度関数の列 $\{p_m\}_{m=1}^\infty \subset [p]$ が取れる. $\{p_m\}$ は確率分布列としてタイトであるから, 最初から弱収束先が存在すると仮定してよい. ところで仮定より $q(x)$ は有界だから $q(x) \leq M$ とすれば

$$\int p_m(x) \log p_m(x) dx = \text{KL}(p_m, q) + \int p_m(x) \log q(x) dx \\ \leq \text{KL}(p_m, q) + \log M$$

となる. よって

$$\liminf_{m \rightarrow \infty} \int p_m(x) \log p_m(x) dx \leq \tilde{D}([p], [q]) + \log M < \infty$$

が成り立つ. すると p_m の弱収束先は絶対連続となる (McCann, 1997, Corollary 3.5). この収束先の密度関数を p_∞ とおけば

$$\liminf_{m \rightarrow \infty} \int p_m(x) \log p_m(x) dx \geq \int p_\infty(x) \log p_\infty(x) dx$$

が成り立つ (McCann, 1997, Lemma 3.4). また $q(x)$ の上半連続性より

$$\liminf_{m \rightarrow \infty} \int p_m(x) \log(1/q(x)) dx \geq \int p_\infty(x) \log(1/q(x)) dx$$

も成り立つ (van der Vaart, 2000). 以上から

$$\tilde{D}([p], [q]) = \lim_{m \rightarrow \infty} \int p_m(x) \log \frac{p_m(x)}{q(x)} dx \geq \int p_\infty(x) \log \frac{p_\infty(x)}{q(x)} dx > 0$$

を得る. 最後の不等号は p_∞ と q が異なる密度であることによる. 実際に $p_\infty \neq q$ であることは以下のようにして示される. まず $p_\infty \notin \mathcal{P}$ の場合 (p_∞ のサポートが $[0, 1]^d$ の真部分集合の場合) は明らかに $p_\infty \neq q$ である. よって $p_\infty \in \mathcal{P}$ と仮定してよい. このとき $p_\infty \in [p]$ であることを示せば十分である ($q \notin [p]$ より). さて p_m の弱収束先が p_∞ であったから, p_m の周辺分布も p_∞ の周辺分布に弱収束する. すると, $p_m = (T_m)_* p$ を満たす写像列 $T_m \in \mathcal{T}$ は, ある $T \in \mathcal{T}$ に一様収束する (T は p_∞ の周辺分布から定まる). したがって $p_m = (T_m)_* p$ は $T_* p$ に弱収束する (例えば吉田, 2006 の定理 1.54). よって $p_\infty = T_* p \in [p]$ が従う. 以上で定理 1 が示された. □

定理 1 における q の有界性は証明の都合によるものであり, 必要ではない. 実際, ガウスコピュラの場合はこの仮定が満たされないが, プロファイルダイバージェンスの正值性は直接確認できる (5 節). 定理の条件を緩めることは今後の課題である.

次の結果は有限次元の複合変換モデルに対して成り立つ事実 (付録 A) の類推である.

定理 2. p, q は正値かつ連続なコピュラ密度関数とし, ある正の定数 C_0, C_1 が存在して任意の $x \in [0, 1]^d$ に対して $C_0 \leq p(x), q(x) \leq C_1$ を満たすものとする. 真の密度が p のとき, $n \rightarrow \infty$ のもとで確率 1 で

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \frac{\bar{p}_n(R)}{\bar{q}_n(R)} = \tilde{D}([p], [q])$$

が成立する. また順位ダイバージェンスはプロファイルダイバージェンスに収束する:

$$\lim_{n \rightarrow \infty} D_n([p], [q]) = \tilde{D}([p], [q]).$$

以下の証明には, 4 節で示す区分一様コピュラに関する結果を用いる.

証明. 2次元の場合を証明する. まず2次元の区分一様密度を考える (式 (4.5) 参照). $p = p_\theta$, $q = p_\phi$ とおく. ここで $\theta = (\theta_{ij}), \phi = (\phi_{ij}) \in \mathbb{R}^{I \times J}$ であり, θ_{ij} および ϕ_{ij} は長方形領域 $A_{ij} = [\frac{i-1}{I}, \frac{i}{I}] \times [\frac{j-1}{J}, \frac{j}{J}]$ における定数の確率密度を表す. 真の密度が p であるとして $\bar{q}_n(R)$ の漸近形を求める. 定理 3 より

$$\bar{q}_n(R) = (IJ)^{-n} \sum_{\sigma} \sum_{\tau} \frac{1}{\prod_i \sigma_i! \prod_j \tau_j!} \prod_i \prod_j \phi_{ij}^{n_{ij}(R, \sigma, \tau)}$$

となる. ただし $n_{ij} = n_{ij}(R, \sigma, \tau)$ は長方形領域 A_{ij} に入る観測値の度数を表し, $\sigma = (\sigma_i), \tau = (\tau_j)$ は周辺度数を表す. $\hat{\pi}_{ij} = n_{ij}/n$ とおくと,

$$\log \bar{q}_n(R) = -n \log(IJ) + \log \left(\sum_{\sigma} \sum_{\tau} \frac{\prod_i \prod_j \phi_{ij}^{n\hat{\pi}_{ij}}}{\prod_i (n\hat{\pi}_{i+})! \prod_j (n\hat{\pi}_{+j})!} \right)$$

となる．ここで全ての非負整数 m に対して $m \log m - m \leq \log m! \leq (m+1) \log(m+1) - m$ が成り立つ(たとえば Feller (1968) の II.9 節)．また σ, τ の取り得る値の個数はそれぞれ $\binom{n+I-1}{n}$, $\binom{n+J-1}{n}$ であり，これは n に関して多項式オーダーである．よってラプラス近似を適用でき，

$$\frac{1}{n} \log((n!)^2 \bar{q}_n(R)) = \sup_{\sigma, \tau} \left(\sum_i \sum_j \hat{\pi}_{ij} \log \phi_{ij} - \sum_i \hat{\pi}_{i+} \log(I\hat{\pi}_{i+}) - \sum_j \hat{\pi}_{+j} \log(J\hat{\pi}_{+j}) \right) + o(1)$$

が言える．ここで $o(1)$ は $n \rightarrow \infty$ のもとで R に関して一様に 0 に収束する項である．さて経験測度 $\hat{P} = n^{-1} \sum_{t=1}^n \delta_{(x_{t1}, x_{t2})}$ を用いると，ある $T \in \mathcal{T}$ が存在して $\hat{\pi}_{ij} = \hat{\pi}_{ij}^T := \hat{P}(T^{-1}(A_{ij}))$ と書ける．ただし A_{ij} は上で定義した長方形領域である．さらに T は各 $T^{-1}(A_{ij})$ において線形としてよい．このとき写像 T を動かすことと σ, τ を動かすことは同値となる．よって

$$\frac{1}{n} \log((n!)^2 \bar{q}_n(R)) = \sup_T \left(\sum_i \sum_j \hat{\pi}_{ij}^T \log \phi_{ij} - \sum_i \hat{\pi}_{i+}^T \log(I\hat{\pi}_{i+}^T) - \sum_j \hat{\pi}_{+j}^T \log(J\hat{\pi}_{+j}^T) \right) + o(1)$$

が得られる．Glivenko–Cantelli の定理(たとえば van der Vaart, 2000)から，確率 1 で， T について一様に $\hat{\pi}_{ij}^T$ は $\pi_{ij}^T := P(T^{-1}(A_{ij}))$ に収束する．ここで P は真の確率分布，すなわち $P(dx) = p(x)dx$ である．よって確率 1 で

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log((n!)^2 \bar{q}_n(R)) = \sup_T \left(\sum_i \sum_j \pi_{ij}^T \log \phi_{ij} - \sum_i \pi_{i+}^T \log(I\pi_{i+}^T) - \sum_j \pi_{+j}^T \log(J\pi_{+j}^T) \right)$$

となる．一方， T が各 $T^{-1}(A_{ij})$ において線形であるから， $x \in T^{-1}(A_{ij})$ のとき $T'_1(x_1) = 1/(I\pi_{i+}^T)$ などが言えるので，

$$\begin{aligned} \text{KL}(p, T^*q) &= \int p(x) \log p(x) dx + \sum_i \sum_j \int_{T^{-1}(A_{ij})} p(x) \log \frac{1}{\phi_{ij} T'_1(x_1) T'_2(x_2)} dx \\ &= \int p(x) \log p(x) dx - \sum_i \sum_j \pi_{ij}^T \log \phi_{ij} + \sum_i \pi_{i+}^T \log(I\pi_{i+}^T) + \sum_j \pi_{+j}^T \log(J\pi_{+j}^T) \end{aligned}$$

となる．以上から

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log((n!)^2 \bar{q}_n(R)) = \int p(x) \log p(x) dx - \inf_T \text{KL}(p, T^*q)$$

が示された． $\bar{p}_n(R)$ についても同じことが言えるので，結局

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} \log \left(\frac{\bar{p}_n(R)}{\bar{q}_n(R)} \right) &= -\inf_T \text{KL}(p, T^*p) + \inf_T \text{KL}(p, T^*q) \\ &= \inf_T \text{KL}(p, T^*q) \\ &= \check{D}([p], [q]) \end{aligned}$$

が得られる．最後の等号は命題 3 による．

p, q が区分一様とは限らない有界連続密度の場合は，いったん密度を区分一様密度で近似すればよい．つまり，任意の $\varepsilon > 0$ に対して $(1-\varepsilon)p_\varepsilon \leq p \leq (1+\varepsilon)p_\varepsilon$ を満たす区分一様密度 p_ε

が存在するので (q も同様), これに基づいて順位尤度を評価すれば

$$\frac{1}{n} \log \frac{\bar{p}_n(R)}{\bar{q}_n(R)} \leq \frac{1}{n} \log \frac{(1+\varepsilon)^n \bar{p}_{\varepsilon,n}(R)}{(1-\varepsilon)^n \bar{q}_{\varepsilon,n}(R)} \rightarrow \log \frac{1+\varepsilon}{1-\varepsilon} + \tilde{D}([p_\varepsilon], [q_\varepsilon]) \quad (n \rightarrow \infty)$$

が成り立つ (下側の評価も同様). この不等式で $\varepsilon \rightarrow 0$ とすればよい.

順位ダイバージェンスの収束は有界収束定理より従う. 実際, 仮定から $C_0 \leq p(x) \leq C_1$ であるから, $C_0^n \leq \bar{p}_n(R) \leq C_1^n$ が成り立つ. $\bar{q}_n(R)$ についても同様なので $n^{-1} \log(\bar{p}_n(R)/\bar{q}_n(R))$ は有界である. \square

この定理により, セミパラメトリックコピュラモデルの漸近的な性質はプロファイルダイバージェンスが担っていると考えられることができる.

4. 区分一様コピュラ

前節で定義した順位ダイバージェンスやプロファイルダイバージェンスは一般には陽に計算できない. しかし区分一様分布やガウス分布に限れば計算を進めることができる. 本節では区分一様分布を考える. 2次元に限定して議論を進めるが, 多次元でも同様の結果が成り立つ. なお区分一様分布はチェス盤分布 (chessboard distribution) とも呼ばれる (Ghosh and Henderson, 2001).

4.1 プロファイルダイバージェンス

I, J を正の整数とし, $[0, 1]^2$ を $I \times J$ 個の小長方形に分けてそれぞれ $A_{ij} = \left[\frac{i-1}{I}, \frac{i}{I}\right) \times \left[\frac{j-1}{J}, \frac{j}{J}\right)$ ($1 \leq i \leq I, 1 \leq j \leq J$) とおく. 密度関数 p が区分一様 (piecewise uniform) であるとは,

$$p(x_1, x_2) = p_{ij} \quad \text{if } (x_1, x_2) \in A_{ij}, \quad 1 \leq i \leq I, \quad 1 \leq j \leq J$$

と書けることとする. ここで p_{ij} は $\sum_{i=1}^I \sum_{j=1}^J p_{ij}/IJ = 1$ を満たす正の実数である. 区分一様密度がコピュラ密度であるための必要十分条件は

$$\sum_{i=1}^I \frac{p_{ij}}{I} = 1 \quad (1 \leq j \leq J), \quad \sum_{j=1}^J \frac{p_{ij}}{J} = 1 \quad (1 \leq i \leq I)$$

となることである. したがって区分一様コピュラ密度全体は $(I-1)(J-1)$ 次元のパラメトリックモデルとなる. 一見すると分割表モデルと同じだが, 2節で既に論じたように, ダイバージェンスの構造は異なるものとなる.

次の命題から, q が区分一様である限り, プロファイルダイバージェンス $\tilde{D}([p], [q])$ の計算は有限次元の最適化問題に帰着されることが分かる.

命題 3. p を任意のコピュラ密度, q を区分一様コピュラ密度とし, 領域 A_{ij} における $q(x)$ の値を q_{ij} とする. このとき $\tilde{D}([p], [q]) = \text{KL}(p, T^*q)$ を達成する $T = (T_1, T_2) \in \mathcal{T}$ は

$$(4.1) \quad \xi_i = T_1^{-1}(i/I), \quad \eta_j = T_2^{-1}(j/J) \quad (1 \leq i \leq I-1, \quad 1 \leq j \leq J-1)$$

を節点とする区分線形変換 T_1, T_2 で与えられる. ここで ξ_i, η_j は次の最小化問題の解である:

$$(4.2) \quad \begin{aligned} \text{Minimize} \quad & \sum_i \sum_j \left(\int_{\xi_{i-1}}^{\xi_i} \int_{\eta_{j-1}}^{\eta_j} p(x) dx \right) \log \frac{1}{q_{ij}} \\ & + \sum_i (\xi_i - \xi_{i-1}) \log(\xi_i - \xi_{i-1}) + \sum_j (\eta_j - \eta_{j-1}) \log(\eta_j - \eta_{j-1}) \end{aligned}$$

subject to $0 = \xi_0 < \xi_1 < \cdots < \xi_I = 1$, $0 = \eta_0 < \eta_1 < \cdots < \eta_J = 1$.

この目的関数は $\text{KL}(p, T^*q)$ と定数差を除いて等しい.

証明. $\{\xi_i\}, \{\eta_j\}$ を固定し, 式(4.1)を満たす T の中で $\text{KL}(p, T^*q)$ が最小となるのは区分線形変換であることを示す. このとき q の区分一様性から $q(T(x))$ は T の選び方によらないので

$$(4.3) \quad \begin{aligned} \text{KL}(p, T^*q) &= \int p(x) \log \frac{p(x)}{q(T(x))T'_1(x_1)T'_2(x_2)} dx \\ &= (\text{const.}) - \int_0^1 \log T'_1(x_1) dx_1 - \int_0^1 \log T'_2(x_2) dx_2 \end{aligned}$$

となる. ただしコピュラ密度の条件 $\int p(x) dx_2 = \int p(x) dx_1 = 1$ を用いた. さて, 一般に $t(0) < t(1)$ を固定したもとで $-\int_0^1 \log t'(x) dx$ を最小化する単調増加関数 $t(x)$ は一次式である. 実際, $-\log$ の凸性から $-\int_0^1 \log t'(x) dx \geq -\log \int_0^1 t'(x) dx = -\log(t(1) - t(0))$ となり, 等号は $t''(x) = 0$ のとき成立する. 以上から, T_1, T_2 は区分線形である. 式(4.2)の目的関数は, $x_1 \in [\xi_{i-1}, \xi_i]$ のとき $T'_1(x_1) = 1/(I(\xi_i - \xi_{i-1}))$ などに注意すれば, 式(4.3)より得られる. \square

式(4.2)の最適化問題は一般に非凸であり, 解の一意性は成り立たない. しかし定理 1 より解は必ず存在する.

以下, $I = J = 2$ の場合を詳しく見てみよう. すなわち次の形のコピュラ密度を考える:

$$c_\theta(x) = \begin{cases} 1 + \theta & \text{if } x \in [0, 1/2]^2 \cup [1/2, 1]^2, \\ 1 - \theta & \text{otherwise.} \end{cases}$$

ただし $-1 < \theta < 1$ とする. 2 節で考えたコピュラ密度もこの形であった. $\theta = 0$ の場合は一様密度となる.

実は, 次の命題に示すような「分岐現象」が成り立つ. ξ_1, η_1 の定義は上と同様である.

命題 4. $\theta > 0$ とする. このとき一様密度 u と c_θ の間のプロファイルダイバージェンスは

$$\tilde{D}([u], [c_\theta]) = \begin{cases} \text{KL}(u, c_\theta) & \text{if } 0 < \theta \leq \tanh(1), \\ \text{KL}(u, T^*c_\theta) & \text{if } \tanh(1) < \theta < 1 \end{cases}$$

となる. ここで $T = (T_1, T_2)$ は区分線形変換であり, その節点は $\xi_1 = \eta_1 = \xi$ または $1 - \xi$ で与えられる. ただし ξ は次の方程式の一意的な解である:

$$(4.4) \quad \xi = \frac{1}{2} \left(1 + \frac{\log \frac{\xi}{1-\xi}}{\log \frac{1+\theta}{1-\theta}} \right), \quad \frac{1}{2} < \xi < 1.$$

証明. 式(4.2)の目的関数を $f(\xi_1, \eta_1)$ とおく. f の停留点の方程式は

$$\eta_1 = \frac{1}{2} \left(1 + \frac{\log \frac{\xi_1}{1-\xi_1}}{\log \frac{1+\theta}{1-\theta}} \right), \quad \xi_1 = \frac{1}{2} \left(1 + \frac{\log \frac{\eta_1}{1-\eta_1}}{\log \frac{1+\theta}{1-\theta}} \right)$$

となる. この方程式は $|\log \frac{1+\theta}{1-\theta}| \leq 2$ すなわち $|\theta| \leq \tanh(1)$ のときは $(\xi_1, \eta_1) = (1/2, 1/2)$ という一意解を持ち, $|\theta| > \tanh(1)$ のときは $(1/2, 1/2)$ の他に対称な 2 つの解 $(\xi, \xi), (1 - \xi, 1 - \xi)$ を持つことが示される. ただし ξ は式(4.4)の解である. また f のヘッセ行列は

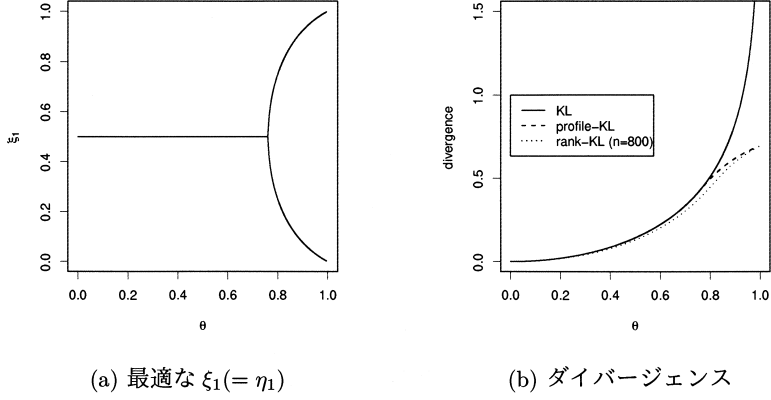


図 2. 分岐現象. (a) $\theta = \tanh(1)$ を境にして, 最適な写像 T は切り替わる. (b) ダイバージェンス $KL(u, c_\theta)$, 順位ダイバージェンス $D_n([u], [c_\theta])$ ($n = 800$), プロファイルダイバージェンス $\bar{D}([u], [c_\theta])$ を θ の関数としてプロットしたもの.

$$\begin{pmatrix} \frac{1}{\xi_1(1-\xi_1)} & 2 \log \frac{1-\theta}{1+\theta} \\ 2 \log \frac{1-\theta}{1+\theta} & \frac{1}{\eta_1(1-\eta_1)} \end{pmatrix}$$

となる. 特に $\theta > \tanh(1)$ ならば $\xi_1 = \eta_1 = 1/2$ は極小点とならず, $\xi_1 = \eta_1 = \xi$ および $\xi_1 = \eta_1 = 1 - \xi$ が極小点となる. \square

式(4.4)を θ について解くと

$$\theta = \frac{\left(\frac{\xi}{1-\xi}\right)^{1/(2\xi-1)} - 1}{\left(\frac{\xi}{1-\xi}\right)^{1/(2\xi-1)} + 1}, \quad \frac{1}{2} < \xi < 1$$

となる. たとえば $\xi = \frac{3}{4}$ とすると, $\theta = 0.8$ となる. これが 2 節で与えた例である.

命題 4 から, 一様密度からのダイバージェンスに関して, $\theta = \tanh(1)$ を境として分岐現象が生ずる. この様子を図示したものが図 2 である. なお, $\theta \rightarrow 1$ のとき $\bar{D}([u], [c_\theta])$ は $\log 2$ に近づくことも示せる. 特にプロファイルダイバージェンスは有界となる. これは, $\theta \rightarrow 1$ のとき $KL(u, c_\theta) \rightarrow \infty$ となることに比べると対照的である.

4.2 区分一様コピュラの順位尤度

区分一様コピュラのもとで, 順位尤度を求める. 区分一様コピュラモデル

$$(4.5) \quad p_\theta(x_1, x_2) = \theta_{ij} \quad \text{if } (x_1, x_2) \in A_{ij}$$

を考える. ただし A_{ij} は前項で定義した小領域であり, θ_{ij} は $\sum_i \theta_{ij}/I = 1, \sum_j \theta_{ij}/J = 1$ を満たすものとする. このモデルに従うランダム標本 $X = \{(x_{t1}, x_{t2})\}_{t=1}^n$ の同時密度関数は

$$(4.6) \quad (IJ)^{-n} \prod_{i=1}^I \prod_{j=1}^J \theta_{ij}^{n_{ij}}$$

となる. ここで n_{ij} は各 A_{ij} に入った観測値の度数である. 度数分布 (n_{ij}) は (θ_{ij}) の十分統計量となる. ただし実際には X は観測されず順位統計量 R だけが観測されるので, (n_{ij}) は潜在変数となる.

度数分布 (n_{ij}) の周辺度数を $\sigma_i = n_{i+} = \sum_j n_{ij}$, $\tau_j = n_{+j} = \sum_i n_{ij}$ とおく. $\sigma = (\sigma_i)$, $\tau = (\tau_j)$ が与えられたもとでは, R が決まれば (n_{ij}) が確定する. これを $n_{ij} = n_{ij}(R, \sigma, \tau)$ と表す.

定理 3. 区分一様コピュラ密度に対し, 順位尤度は

$$(4.7) \quad \bar{p}_n(R) = (IJ)^{-n} \sum_{\sigma} \sum_{\tau} \frac{1}{\prod_i \sigma_i! \prod_j \tau_j!} \prod_i \prod_j \theta_{ij}^{n_{ij}(R, \sigma, \tau)}$$

となる. ここで σ, τ は周辺度数全体を動くものとする.

証明. $\bar{p}_n(R)$ を単に $p(R)$ と表す. 式(4.6)より, $N := (n_{ij})$ の周辺分布は

$$p(N) = \frac{n!}{\prod_i \prod_j n_{ij}!} (IJ)^{-n} \prod_i \prod_j \theta_{ij}^{n_{ij}},$$

で与えられる. 次に N を固定したもとの R の条件付き分布 $p(R|N)$ を求める. N は θ に関して十分統計量であるから, $p(R|N)$ は θ には依存しない. したがって, X が一様分布から生成されたと仮定して $p(R|N)$ を計算すればよい. さて σ, τ は N だけで決まるから,

$$\begin{aligned} p(R|N) &= p(R|N, \sigma, \tau) \\ &= \frac{p(R, N|\sigma, \tau)}{p(N|\sigma, \tau)} \\ &= \begin{cases} \frac{p(R|\sigma, \tau)}{p(N|\sigma, \tau)} & \text{if } N = N(R, \sigma, \tau), \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

となる. いま X が一様分布からのサンプルであるとしているから, $p(N|\sigma, \tau)$ は超幾何分布となり, また $p(R|\sigma, \tau) = p(R) = 1/(n!)^2$ となる. 以上から

$$\begin{aligned} p(R) &= \sum_N p(R|N)p(N) \\ &= \sum_{\sigma} \sum_{\tau} p(R|\sigma, \tau) \frac{1}{p(N|\sigma, \tau)} p(N) \Bigg|_{N=N(R, \sigma, \tau)} \\ &= \sum_{\sigma} \sum_{\tau} \frac{1}{(n!)^2} \frac{n! \prod_i \prod_j n_{ij}!}{\prod_i \sigma_i! \prod_j \tau_j!} \frac{n!}{\prod_i \prod_j n_{ij}!} (IJ)^{-n} \prod_i \prod_j \theta_{ij}^{n_{ij}} \Bigg|_{N=N(R, \sigma, \tau)} \end{aligned}$$

となり, これを整理すれば式(4.7)が得られる. \square

図 2(b)にある順位ダイバージェンスのグラフは定理 3 の結果とモンテカルロ法を利用して計算したものである.

5. ガウスコピュラの場合

ガウスコピュラとは, 座標ごとの変数変換によって多変量正規分布となるようなコピュラのことである. 2つのガウスコピュラの間のプロファイルダイバージェンスは簡単な形になる. また順位尤度は多次元正規分布の象限確率の計算に帰着される. これらの結果を述べる.

なお, 3 節で注意した通り, \mathbb{R}^d 上の 2つの密度関数の間のダイバージェンスは $(0, 1)^d$ 上の密度関数に帰着させて定義する. しかし結果として \mathbb{R}^d 上の変数変換を考えればよいことになるので $(0, 1)^d$ を意識する必要はない.

5.1 プロファイルダイバージェンス

ガウスコピュラに対してはプロファイルダイバージェンスが有限次元の凸最適化によって計算できる。これを示そう。平均ベクトル 0 、共分散行列 Σ の正規分布の密度関数を ϕ_Σ と記す。またベクトル u を対角成分とする対角行列を $\text{diag}(u)$ と表す。

命題 5. P, Q を d 次正定値対称行列とし、 $p = \phi_P, q = \phi_Q$ とおく。このとき $\tilde{D}([p], [q]) = \text{KL}(p, T^*q)$ を満たす T は線形変換 $T(x) = \text{diag}(u)x$ で与えられる。ここで $u = (u_i)$ は次の凸最適化問題の解である：

$$(5.1) \quad \begin{aligned} \text{Minimize} \quad & - \sum_i \log u_i + \frac{1}{2} \text{tr}(Q^{-1} \text{diag}(u) P \text{diag}(u)) \\ \text{subject to} \quad & u_1, \dots, u_d > 0. \end{aligned}$$

この目的関数は $\text{KL}(p, T^*q)$ と定数差を除いて等しい。

証明. q は対数凹である。よって、 $T(x) = \text{diag}(u)x$ が命題 2 の停留条件

$$\partial_i \log p_i(x_i) = \mathbb{E}_p[(\partial_i \log(T^*q))(x)|x_i]$$

を満たすことをチェックすれば十分である。左辺は $\partial_i \log p_i(x_i) = -P_{ii}^{-1}x_i$ である。右辺は

$$\begin{aligned} \mathbb{E}_p[(\partial_i \log(T^*q))(x)|x_i] &= - \sum_j u_j (Q^{-1})_{ij} u_j \mathbb{E}_p[x_j|x_i] \\ &= - \sum_j u_j (Q^{-1})_{ij} u_j P_{ji} P_{ii}^{-1} x_i \\ &= -P_{ii}^{-1} x_i \end{aligned}$$

となる。最後の等号は式(5.1)の停留条件による。式(5.1)が実際に一意解を持つことは Marshall and Olkin (1968) で示されている。□

$d = 2$ の場合は、密度 p, q に対応する相関係数をそれぞれ ρ_p, ρ_q とおけば、

$$\tilde{D}([p], [q]) = \log \frac{1 - \rho_p \rho_q}{\sqrt{(1 - \rho_p^2)(1 - \rho_q^2)}}$$

となる。これは対称なダイバージェンスである。また、Fisher の Z 変換 $\rho = \tanh z$ を施すと、

$$\tilde{D}([p], [q]) = \log \cosh(z_p - z_q)$$

となり、平行移動に関する不変性が成り立つ。

ガウスコピュラに限れば、プロファイルダイバージェンスから導かれる計量は、Segers et al. (2014) によって求められている有効情報量に一致する。これを確かめよう。ただし一般にモデル $\{p_\theta \mid \theta \in \mathbb{R}^m\}$ に対し、ダイバージェンスから導かれる計量を

$$\tilde{g}_{ij} = - \left. \frac{\partial^2 \tilde{D}([p_\theta], [p_\phi])}{\partial \theta_i \partial \phi_j} \right|_{\theta=\phi}, \quad i, j \in \{1, \dots, m\},$$

と定義する (Eguchi, 1983; Amari, 1985)。

定理 4. 共分散行列のパラメトリックモデル $\{P = P_\theta \mid \theta \in \mathbb{R}^m\}$ に対し、計量は

$$(5.2) \quad \tilde{g}_{ij} = \frac{1}{2} \text{tr}\{P^{-1}(\partial_i P - \Pi(\partial_i P))P^{-1}(\partial_j P - \Pi(\partial_j P))\}$$

で与えられる。ただし $\partial_i = \partial/\partial\theta_i$ とし、射影 $\Pi: \mathbb{R}^{d \times d} \rightarrow \mathbb{R}^{d \times d}$ は以下で定義される：

$$(5.3) \quad \Pi(A) = P \operatorname{diag}(b(A)) + \operatorname{diag}(b(A))P,$$

$$(5.4) \quad b(A) = (P^{-1} \circ P + I)^{-1}(P^{-1} \circ A)1_d, \quad 1_d = (1, \dots, 1)^\top \in \mathbb{R}^d.$$

$A \circ B$ は行列 A, B の要素ごとの積(アダマール積)を表す。

証明. 多変量正規分布のフィッシャー情報行列が

$$g_{ij} = \frac{1}{2} \operatorname{tr}(P^{-1}(\partial_i P)P^{-1}(\partial_j P))$$

であることはよく知られている。また一般に式(A.2)の意味でのプロファイルダイバージェンスに対して、計量は

$$\tilde{g}_{ij} = E[\{\partial_i \ell - \Pi(\partial_i \ell)\}\{\partial_j \ell - \Pi(\partial_j \ell)\}]$$

となる(Barndorff-Nielsen and Jupp (1988), Theorem 7.2)。ここで $\ell = \log p_{\theta, \nu}$ であり、 Π は局外接空間 $\operatorname{span}\{\frac{\partial \ell}{\partial \nu}\}$ への直交射影である。接ベクトルの対応 $\partial_i P \mapsto \partial_i \ell$ は線形同型であるから、式(5.2)が得られる。あとは直交射影 Π を求めればよい。いま局外パラメータ u の入り方は $\operatorname{diag}(u)P \operatorname{diag}(u)$ であり、 $u = 1_d$ における接ベクトルは

$$\partial_{u_k}(\operatorname{diag}(u)P \operatorname{diag}(u)) = \operatorname{diag}(e_k)P + P \operatorname{diag}(e_k), \quad k = 1, \dots, d,$$

である。ただし e_k は第 k 単位ベクトルである。よって直交射影 Π は式(5.3)の形となる。また $\Pi(A) = PB + BP, B = \operatorname{diag}(b)$ とおくと、任意の対角行列 $C = \operatorname{diag}(c)$ に対して

$$\begin{aligned} 0 &= \frac{1}{2} \operatorname{tr}\{P^{-1}(PC + CP)P^{-1}(A - (PB + BP))\} \\ &= \sum_k c_k (P^{-1}A)_{kk} - \sum_k c_k b_k - \sum_k \sum_l c_k b_l (P^{-1})_{kl} P_{kl} \end{aligned}$$

となる。よって $(P^{-1} \circ A)1_d - (I + P^{-1} \circ P)b = 0$ となり、式(5.4)を得る。□

5.2 ガウスコピュラの順位尤度

ガウスコピュラの場合、順位尤度の式(3.3)が正規分布の象限確率で表される。これは積分領域 $\{R(X) = R\}$ が

$$x_{t(s,i),i} < x_{t(s+1,i),i}, \quad s \in [n-1], \quad i \in [d]$$

と表されることによる。ただし $t(s,i)$ は順位が $r_{ti} = s$ となる観測値番号 t を表す。また $[n] = \{1, \dots, n\}$ と略記した。結果として次の定理が得られる。

定理 5. 共分散行列 Σ の d 次元正規分布に対する順位尤度は

$$\bar{p}_n(R|\Sigma) = \frac{1}{n^{d/2} |\Sigma|^{(n-1)/2} |B|^{1/2}} \int_{\mathbb{R}_+^{(n-1)d}} \phi_{B^{-1}}(w) dw$$

で与えられる。ただし $\mathbb{R}_+ = (0, \infty)$ であり、 $\phi_{B^{-1}}$ は平均ベクトル 0 、共分散行列 B^{-1} の多次元正規分布の密度関数、 B は次で定義される $(n-1)d$ 次の正定値対称行列である：

$$B_{(r-1)d+i, (s-1)d+j} = (\Sigma^{-1})_{ij} \left(\sum_{t=1}^n I_{\{r_{ti} \leq r, r_{tj} \leq s\}} - \frac{rs}{n} \right), \quad r, s \in [n-1], \quad i, j \in [d].$$

証明. $S = \Sigma^{-1}$ と記す。変数変換 $v_i = x_{t(n,i),i}$, $w_{ri} = x_{t(r+1,i),i} - x_{t(r,i),i}$ を考える。それぞれ

れの積分範囲は $v_i \in \mathbb{R}$, $w_{ri} > 0$ となる. 逆変換は $x_{ti} = v_i - \sum_{r=1}^{n-1} I_{\{r_{ti} \leq r\}} w_{ri}$ で与えられる. I は定義関数である. よって X の同時密度関数の指数部は,

$$\begin{aligned}
 \sum_{t=1}^n \sum_{i=1}^d \sum_{j=1}^d S_{ij} x_{ti} x_{tj} &= \sum_{t=1}^n \sum_{i=1}^d \sum_{j=1}^d S_{ij} \left(v_i - \sum_{r=1}^n I_{\{r_{ti} \leq r\}} w_{ri} \right) \left(v_j - \sum_{s=1}^n I_{\{r_{tj} \leq s\}} w_{sj} \right) \\
 &= \sum_i \sum_j \left(n S_{ij} v_i v_j - 2n S_{ij} m_i v_j + \sum_r \sum_s S_{ij} I_{\{r_{ti} \leq r\}} I_{\{r_{tj} \leq s\}} w_{ri} w_{sj} \right)
 \end{aligned}$$

となる. ただし

$$m_i = \frac{1}{n} \sum_{t=1}^n \sum_{r=1}^{n-1} I_{\{r_{ti} \leq r\}} w_{ri} = \frac{1}{n} \sum_{r=1}^{n-1} r w_{ri}$$

とおいた. よって $(v_i), (w_{ri})$ の同時密度関数の指数部は

$$n \sum_i \sum_j S_{ij} (v_i - m_i)(v_j - m_j) + \sum_i \sum_j \sum_r \sum_s B_{(r-1)d+i, (s-1)d+j} w_{ri} w_{sj}$$

となる. v_i について積分をすれば結果が得られる. \square

多次元正規分布の象限確率はホロノミック勾配法によって高精度に計算できることが知られており (Koyama and Takemura, 2015), R のパッケージとして公開されている (Koyama et al., 2014). よって定理 5 を用いて順位尤度および順位ダイバージェンスを計算することができる. ホロノミック勾配法の計算量はホロノミックランクと呼ばれる量が目安となる. m 次元正規分布の象限確率におけるホロノミックランクは 2^m であり, サンプルサイズが大きくなると急激に計算量は大きくなる.

表 1 および表 2 は共分散行列が

$$(5.5) \quad \Sigma = \Sigma(\theta) = \begin{pmatrix} 1 & \theta \\ \theta & 1 \end{pmatrix}, \quad \theta \in \{0.0, 0.1, \dots, 0.9, 0.95\},$$

で与えられる 2 次元ガウスコピュラについて, それぞれ $n = 2$ と $n = 3$ の場合の順位尤度を求めたものである. ただし対称性から順位統計量 R の第 1 列 (第 1 成分の順位統計量) を

表 1. $n = 2$ の場合の順位尤度. ただし $r_{t1} = t$ ($\forall t$) とし, 結果を $n!$ 倍している.

θ	.00	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95
$(r_{t2}) = (1, 2)$.5000	.5319	.5641	.5970	.6310	.6667	.7048	.7468	.7952	.8564	.8989
(2, 1)	.5000	.4681	.4359	.4030	.3690	.3333	.2952	.2532	.2048	.1436	.1011

表 2. $n = 3$ の場合の順位尤度. ただし $r_{t1} = t$ ($\forall t$) とし, 結果を $n!$ 倍している.

θ	.00	.10	.20	.30	.40	.50	.60	.70	.80	.90	.95
$(r_{t2}) = (1, 2, 3)$.1667	.1918	.2196	.2509	.2866	.3280	.3773	.4381	.5179	.6359	.7302
(1, 3, 2)	.1667	.1780	.1880	.1964	.2028	.2067	.2068	.2017	.1875	.1549	.1212
(2, 1, 3)	.1667	.1780	.1880	.1964	.2028	.2067	.2068	.2017	.1875	.1549	.1212
(2, 3, 1)	.1667	.1542	.1407	.1262	.1109	.0947	.0777	.0597	.0408	.0210	.0106
(3, 1, 2)	.1667	.1542	.1407	.1262	.1109	.0947	.0777	.0597	.0408	.0210	.0106
(3, 2, 1)	.1667	.1439	.1230	.1037	.0859	.0693	.0537	.0391	.0253	.0123	.0061

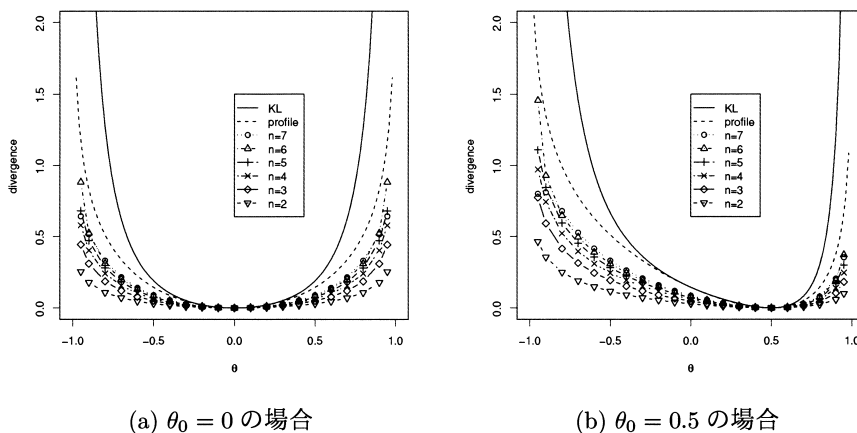


図 3. ガウスコピュラの順位ダイバージェンス $D_n([p_{\theta_0}], [p_{\theta}])$ を θ の関数としてプロットしたものの。

$(1, \dots, n)$ に固定し、代わりに確率の値を $n!$ 倍してある。結果として表の各列の和は 1 となる。

図 3 は同じく 2 次元のガウスコピュラで $\theta_0 = 0$ および $\theta_0 = 0.5$ の場合について、順位ダイバージェンス $D_n([p_{\theta_0}], [p_{\theta}])$ を θ の関数としてプロットしたものである。ここで p_{θ} は式 (5.5) の共分散行列を持つ 2 次元正規分布の密度関数、すなわち $\phi_{\Sigma(\theta)}$ を表す。サンプルサイズは $n = 2$ から $n = 7$ までとした。同図にはプロファイルダイバージェンス $\hat{D}([p_{\theta_0}], [p_{\theta}])$ および通常の Kullback-Leibler ダイバージェンス $KL(p_{\theta_0}, p_{\theta})$ も示した。理論上、 $KL \geq \hat{D} \geq D_n$ という関係があり、これは図においても確かに成立している。

θ を固定したとき、 $n = 2$ から $n = 7$ に増加するにしたがって、順位ダイバージェンスも増加する傾向がある。ただし $n = 7$ のときだけ $n \leq 6$ の曲線と交差している。これは $n = 7$ における計算精度が十分でない可能性が考えられる。今後の検討課題としたい。なお $n = 7$ の場合に要した計算時間は 3 時間程度であった。また計算機環境は Amazon Web Services (32 コア、インスタンスタイプ r3.8xlarge) であり、ソフトウェアは R、ホロノミック勾配法のパッケージとして hgm (Koyama et al., 2014) を用いた。

6. 今後の課題

本稿では順位ダイバージェンスとプロファイルダイバージェンスを定義し、前者は後者に収束することを述べた (定理 2)。しかしここではコピュラ密度関数が有界であることが仮定されており、多くの実用的なコピュラモデルには適用できない。そのような場合に定理が成り立つかどうかを調べる必要がある。この点について査読者より、Huber (1967) の 4 節にある領域分割の方法が役立つのではないかと指摘があった。すなわち、定理 2 の証明で使われている区分一様コピュラの領域分割を等分割にしないという方法である。いまの目的では $[0, 1]^d$ の境界に近付くにつれて分割を細かくすることが自然である。そのためには無限個に領域分割した区分一様コピュラの挙動を明らかにせねばならない。後の論理は定理 2 の証明と同様であると考えられる。

定理 4 ではガウスコピュラの場合に、プロファイルダイバージェンスから有効情報量が求められることを示した。このことは一般のコピュラモデルに拡張できるものと考えられるが、まだ示せていない。コピュラモデルの有効情報量は Sturm-Liouville 型の方程式で表され、普通

は陽に求められない(Bickel et al., 1993).

また本稿では最初から周辺分布 F_1, \dots, F_d に関する情報は全く得られない状況を仮定した. 現実問題としては周辺分布自体にも多少の興味はあるだろうし, 何らかの事前情報を入れる方が自然な場合も想定される. Chen et al. (2006)の推定量はその一例と解釈できる. このような場合に適切なダイバージェンスを導くことも一つの課題である.

確率分布の間には Kullback–Leibler ダイバージェンス以外にも様々なダイバージェンスが定義される. それらに対応する順位ダイバージェンスやプロファイルダイバージェンスを形式的に定義することは可能である. ただし変数変換不変性や加法性が成り立たない場合には注意深く検討する必要があるだろう.

謝 辞

本稿に対して多くの有益なコメントをくださった編集委員および査読者の皆様に深く感謝いたします. 本研究は JSPS 科研費 JP17K00044 および JP26108003 の助成を受けたものです. また, 研究の発端となる話題を提供してくれた慶應義塾大学(当時)の橋本研一氏に感謝します.

付 録

A. 複合変換モデル

複合変換モデル(composite transformation model)とは, 変数変換に対して不変な統計モデルのことであり, 本稿では次のように定義する. 標本空間 \mathcal{X} 上の確率分布族 $\{P_{\theta, \nu} \mid \theta \in \Theta, \nu \in N\}$ を考え, θ は興味のあるパラメータ, ν は局外パラメータを表すものとする. いま群 G が存在して, G は \mathcal{X} と N にそれぞれ作用し, かつ

$$x \sim P_{\theta, \nu}, \quad g \in G \quad \Rightarrow \quad gx \sim P_{\theta, g\nu}$$

という性質を満たすものとする. また G の N への作用は推移的であると仮定する. このときこの確率分布族 $\{P_{\theta, \nu}\}$ を複合変換モデルという(Barndorff-Nielsen and Jupp, 1988). 不変確率モデルとも呼ばれる(Eaton, 1983; 紙屋, 1999). 群 G の作用によって \mathcal{X} は互いに排反な軌道(orbit)に分解される. この軌道と一対一に対応する統計量を最大不変量(maximal invariant)という. 最大不変量の分布は θ のみに依存する.

例 1. 例として正規分布 $N(\mu, \sigma^2)$ に従うランダム標本 $x = (x_1, \dots, x_n)$ を考える. 標本をスケール変換して $x_i \mapsto ax_i$ ($a > 0$) とすると, パラメータは (μ, σ) から $(a\mu, a\sigma)$ に変換される. よって $N(\mu, \sigma^2)$ はスケール変換に関して複合変換モデルになっている. ここで興味のあるパラメータは $\theta = \mu/\sigma$ であり, 局外パラメータは $\nu = \sigma$ である. また観測値 $(x_1, \dots, x_n) \in \mathbb{R}^n$ を含む軌道は $\{(ax_1, \dots, ax_n) \mid a > 0\}$ となる. 最大不変量はたとえば $w(x) = (x_1/\hat{\sigma}, \dots, x_n/\hat{\sigma})$ が挙げられる. ただし $\hat{\sigma}^2$ は標本分散を表す. $w(x)$ の分布は θ のみに依存する. なお, 最小十分統計量 $(\bar{x}, \hat{\sigma})$ に対する最大不変量は $\bar{x}/\hat{\sigma}$ であり, これは Student の t 統計量の定数倍である(Cox and Hinkley, 1974, Example 5.16).

分布 $P_{\theta, \nu}$ (密度 $p_{\theta, \nu}$) に従う標本サイズ n のランダム標本 $x = (x_1, \dots, x_n)$ を考える. パラメータ θ に関する推測は最大不変量 $w = w(x)$ に基づいて行うのが自然である. そこで w の周辺密度関数を $\bar{p}_\theta(w)$ とおき, 2つの密度関数 $\bar{p}_{\theta_1}, \bar{p}_{\theta_2}$ の間のダイバージェンスを考える:

$$(A.1) \quad D_n(\theta_1, \theta_2) = \frac{1}{n} \int \bar{p}_{\theta_1}(w) \log \frac{\bar{p}_{\theta_1}(w)}{\bar{p}_{\theta_2}(w)} dw.$$

これを周辺ダイバージェンスと呼ぶことにしよう．右辺を n で割っているのは w が $O(n)$ の情報を持っているためである．一方で，局外パラメータに依存しないダイバージェンスとして

$$(A.2) \quad \tilde{D}(\theta_1, \theta_2) := \inf_{\nu_2 \in N} \text{KL}(p_{\theta_1, \nu_1}, p_{\theta_2, \nu_2}) \quad (\forall \nu_1)$$

が考えられる．右辺が ν_1 に依存しないのは Kullback–Leibler ダイバージェンスの不変性による．この \tilde{D} をプロファイルダイバージェンスと呼ぶことにする．Barndorff-Nielsen and Jupp (1988) の Section 4 では profile discrimination information と呼ばれている．次の命題は周辺分布をラプラス近似することによって得られ，実質的には同論文の Section 6 で示されている．

命題 6. 真のパラメータが θ のとき，標準的な仮定のもとで，

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \frac{\bar{p}_\theta(w)}{\bar{p}_\phi(w)} = \lim_{n \rightarrow \infty} D_n(\theta, \phi) = \tilde{D}(\theta, \phi)$$

が成り立つ．

例 2 (つづき)．先の正規分布の例では，周辺ダイバージェンスは非心 t 分布の間のダイバージェンスとなり，陽には求まらない．しかしプロファイルダイバージェンスは

$$\begin{aligned} \tilde{D}(\theta_1, \theta_2) &= \inf_{\sigma_2 > 0} \left(-\frac{1}{2} \log \frac{\sigma_1^2}{\sigma_2^2} - \frac{1}{2} + \frac{(\theta_1 \sigma_1 - \theta_2 \sigma_2)^2 + \sigma_1^2}{2\sigma_2^2} \right) \\ &= -\log \left(\frac{\theta_1 \theta_2 + \sqrt{\theta_1^2 \theta_2^2 + 4(\theta_1^2 + 1)}}{2(\theta_1^2 + 1)} \right) - \frac{\theta_1 \theta_2}{2} \left(\frac{\theta_1 \theta_2 + \sqrt{\theta_1^2 \theta_2^2 + 4(\theta_1^2 + 1)}}{2(\theta_1^2 + 1)} \right) + \frac{\theta_2^2}{2} \end{aligned}$$

と陽に求められる．特に $\theta_1 = 0$ のときは $\tilde{D}(0, \theta_2) = \theta_2^2/2$ ，また $\theta_2 = 0$ のときは $\tilde{D}(\theta_1, 0) = (1/2) \log(\theta_1^2 + 1)$ となり，著しく非対称性を持ったダイバージェンスになっていることが分かる．これは t 検定における水準と検出力の関係の非対称性に対応する．このようにプロファイルダイバージェンスを調べれば，複合変換モデルの構造をある程度把握することができる．

参 考 文 献

- Amari, S. (1985). *Differential-Geometrical Methods in Statistics*, Springer, New York.
- Barndorff-Nielsen, O. E. and Jupp, P. E. (1988). Differential geometry, profile likelihood, L-sufficiency and composite transformation models, *The Annals of Statistics*, **16**(3), 1009–1043.
- Bickel, P. J., Klaassen, C. A. J., Ritov, Y. and Wellner, J. A. (1993). *Efficient and Adaptive Estimation for Semiparametric Models*, Johns Hopkins University Press, Baltimore.
- Chen, X., Fan, Y. and Tsyrennikov, V. (2006). Efficient estimation of semiparametric multivariate copula models, *Journal of the American Statistical Association*, **101**(475), 1228–1240.
- Cox, D. R. and Hinkley, D. V. (1974). *Theoretical Statistics*, Chapman and Hall, London.
- Eaton, M. L. (1983). *Multivariate Statistics — A Vector Space Approach*, Wiley, New York.
- Eguchi, S. (1983). Second order efficiency of minimum contrast estimators in a curved exponential family, *The Annals of Statistics*, **11**(3), 793–803.
- Feller, W. (1968). *An Introduction to Probability Theory and Its Applications*, 3rd ed., John Wiley & Sons, New York.
- Genest, C. and Werker, B. J. M. (2002). Conditions for the asymptotic semiparametric efficiency of an omnibus estimator of dependence parameters in copula models, *Distributions with Given*

- Marginals and Statistical Modelling* (eds. C. M. Cuadras and J. A. R. Lallena), 103–112, Kluwer Academic, Dordrecht.
- Ghosh, S. and Henderson, S. G. (2001). Chessboard distributions and random vectors with specified marginals and covariance matrix, *Operations Research*, **50**(5), 820–834.
- Hoff, P. D. (2007). Extending the rank likelihood for semiparametric copula estimation, *The Annals of Applied Statistics*, **1**(1), 265–283.
- Hoff, P. D., Niu, X. and Wellner, J. A. (2014). Information bounds for Gaussian copulas, *Bernoulli*, **20**(2), 604–622.
- Huber, P. J. (1967). The behavior of maximum likelihood estimates under nonstandard conditions, *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, 221–233.
- 紙屋英彦 (1999). 不変確率モデルの特徴付け, *統計数理*, **47**(1), 63–69.
- Klaassen, C. A. J. and Wellner, J. A. (1997). Efficient estimation in the bivariate normal copula model: Normal margins are least favourable, *Bernoulli*, **3**(1), 55–77.
- Koyama, T. and Takemura, A. (2015). Calculation of orthant probabilities by the holonomic gradient method, *Japan Journal of Industrial and Applied Mathematics*, **32**, 187–204.
- Koyama, T., Nakayama, H., Ohara, K., Sei, T. and Takayama, N. (2014). Software packages for holonomic gradient method, *Mathematical Software—ICMS 2014* (eds. H. Hong and C. Yap), Lecture Notes in Computer Science, Vol. 8592, 706–712, Springer, Berlin.
- Kullback, S. (1959). *Information Theory and Statistics*, John Wiley and Sons, New York.
- Marshall, A. W. and Olkin, I. (1968). Scaling of matrices to achieve specified row and column sums, *Numerische Mathematik*, **12**, 83–90.
- McCann, R. J. (1997). A convexity principle for interacting gases, *Advances in Mathematics*, **128**, 153–179.
- Nelsen, R. B. (2006). *An Introduction to Copulas*, 2nd ed., Springer, New York.
- Segers, J., van den Akker, R. and Werker, B. J. M. (2014). Semiparametric Gaussian copula models: Geometry and efficient rank-based estimation, *The Annals of Statistics*, **42**(5), 1911–1940.
- Tsukahara, H. (2005). Semiparametric estimation in copula models, *The Canadian Journal of Statistics*, **33**(3), 357–375.
- 塚原英敦 (2011). 接合分布関数 (コピュラ) の理論と応用, 『21 世紀の統計科学』 (北川源四郎, 竹村彰通編), 第 3 巻, 101–140, 日本統計学会, <http://park.itc.u-tokyo.ac.jp/atstat/iss75shunen/>.
- van der Vaart, A. W. (2000). *Asymptotic Statistics*, Cambridge University Press, Cambridge.
- Villani, C. (2003). *Topics in Optimal Transportation*, American Mathematical Society, Providence.
- 吉田朋広 (2006). 『数理統計学』, 朝倉書店, 東京.

Properties of Divergence for Semiparametric Copula Models

Tomonari Sei and Kazuya Matsumoto

Graduate School of Information Science and Technology, The University of Tokyo

A semiparametric copula model is a statistical model in which the copula is assumed to be parametric and the marginal distribution is arbitrary. In this paper, we investigate the divergence of the model. In particular, we establish a relationship between the rank divergence induced from the marginal distribution of the multivariate rank statistic and the profile divergence defined by infimum of the Kullback–Leibler divergence with respect to the nuisance parameter. We also obtain formulas for piecewise uniform and Gaussian copulas.

非対称 t 接合関数の性質と統計的推定方法

—資産価格変動への応用—

吉羽 要直^{1,2}

(受付 2019 年 5 月 31 日 ; 改訂 11 月 14 日 ; 採択 11 月 15 日)

要 旨

非対称 t 接合関数 (skew- t copula) は、多変量非対称 t 分布に内在する接合関数である。多変量非対称 t 分布は非対称性の導入方法により、いくつかの異なる分布が提案されている。本稿では代表的な多変量非対称 t 分布に内在する非対称 t 接合関数の定義と性質をまとめ、非対称 t 接合関数の最尤推定に必要な手続きを整理する。また、非対称 t 接合関数の実証研究の結果を整理し、本邦の株価変動を推定した実証結果を示す。こうした実証結果から、裾依存性の度合いと上下での依存性の違いを調整できる非対称 t 接合関数が資産価格変動の表現に有効であることを指摘し、今後の課題を述べる。

キーワード：接合関数，多変量非対称 t 分布，裾依存係数。

1. はじめに

金融ポートフォリオのリスク量の把握などに際して、リスクファクター間の依存構造は一般的に接合関数 (copula) で表現される (接合関数の基本的な理論と応用手法は塚原, 2008 などを参照)。特に金融リスク管理実務においては、順位相関の程度を相関行列パラメータで柔軟に表現できる正規接合関数 (Gaussian/Normal copula) か t 接合関数 (Student- t copula) で表現されることが多い。正規接合関数は、線形相関行列を用いたリスクファクター間の依存性について、個々のリスクファクターの変動 (周辺分布) に関しては正規分布以外の分布も許容した拡張表現となっており、推定等は容易である。しかしながら、ファクターとなる資産価格変動の裾での依存性は低くなるという特徴を有しているため、デフォルトなどを含めストレス状況での資産変動の記述には向かず、証券化商品の価格が暴落したリーマンショック時には批判の対象となった。一方、 t 接合関数は、ファクターとなる資産価格変動の裾での強い依存性を表現できるため、ストレス状況を考慮した金融リスクの把握に実務上広く用いられている。しかしながら、ファクターとなる資産価格変動の依存性が上下対称になるという制約があり、非現実的な設定となってしまう。Yoshihara (2018) では、こうした背景から非対称な裾依存性を表現できる非対称 t 接合関数 (skew- t copula) を利用するためにその最尤推定法を具体的に実装し、日米欧の株価収益率に適用した実証結果を示した。本稿ではまず代表的な多変量非対称 t 分布に内在する非対称 t 接合関数の定義と性質をまとめる。そのうえで、非対称 t 接合関数の推定法の概要を整理し、資産価格変動等に応用した実証結果を整理する。

¹ 東京都立大学大学院 経営学研究科：〒100-0005 東京都千代田区丸の内 1-4-1

² 統計数理研究所 客員：〒190-8562 東京都立川市緑町 10-3

非対称 t 接合関数は、正規接合関数や t 接合関数と同様に、多変量の非対称 t 分布に内包される陰な接合関数 (implicit copula) である。Kotz and Nadarajah (2004) で示されているように多変量の非対称 t 分布には様々な提案があることから、様々な非対称 t 接合関数を定義できる。先行研究ではこれまで非対称 t 接合関数として 3 つの提案がなされている。1 つ目は、Demarta and McNeil (2005) によって提案されたもので、Blaesild and Jensen (1981) が提案した多変量の一般化双曲型 (generalized hyperbolic, GH) の非対称 t 分布に内包される接合関数である。本稿では、この多変量非対称 t 分布を GH 非対称 t 分布と呼び、内包される接合関数を GH 非対称 t 接合関数 (GH skew- t copula) と呼ぶ。2 つ目は、Smith et al. (2012) が提案し電力市場価格のモデリング等に適用したもので、Sahu et al. (2003) による多変量非対称 t 分布に内包される接合関数である。3 つ目は、Joe (2006) がその利用を提唱しているもので、Azzalini and Capitanio (2003) による多変量の非対称 t 分布に内包される接合関数である。本稿では、この多変量非対称 t 分布を AC 非対称 t 分布と呼び、内包される接合関数を AC 非対称 t 接合関数 (AC skew- t copula) と呼ぶ。Joe (2006) は AC 非対称 t 接合関数の表現について言及し応用上の可能性を述べているものの、具体的な推定法については論じていない。AC 非対称 t 分布は、統計学上は最も利用されている多変量非対称 t 分布であるが、その接合関数の推定方法や実証分析についてはあまり研究が進んでいなかったことから Yoshiba (2018) では AC 非対称 t 接合関数に焦点を当ててその推定法を論じている。

以下、2 節では上記の 3 つの非対称 t 接合関数を中心にその性質をまとめる。Yoshiba (2018) で取り上げた AC 非対称 t 、GH 非対称 t の接合関数のほか、Sahu et al. (2003) の多変量非対称 t 分布に内包される接合関数についても、その性質を比較してまとめる。また、多変量非対称 t 分布については、Jones (2002)、Ferreira and Steel (2007)、Rosco et al. (2011) の提案にも言及する。3 節では最尤推定法の実装について必要な接合密度の表現と推定手法を整理する。4 節では非対称 t 接合関数を利用した実証研究の結果を整理するとともに、本邦の 12 年間の株価日次収益率に対して推定した実証結果を示す。5 節では本稿で解説した研究をまとめ、今後の発展の可能性について述べる。

2. 非対称 t 接合関数の性質

2.1 3 種類の d 変量標準非対称 t 分布

まず、変量数を d として、1 節で示した 3 種類の非対称 t 接合関数の基となっている d 変量非対称 t 分布に従う確率ベクトル $\mathbf{X} \in \mathbb{R}^d$ の表現と密度関数を示す。この際、導出される接合関数は基となる分布の位置パラメータ $\mathbf{l}^\top = (l_1, \dots, l_d)$ や尺度パラメータ $\mathbf{s}^\top = (s_1, \dots, s_d)$ に依らないため、ここでは $\mathbf{l}^\top = (0, \dots, 0)$ 、 $\mathbf{s}^\top = (1, \dots, 1)$ という標準分布を考える。

2.1.1 GH 非対称 t 分布

d 変量の標準 GH 非対称 t 分布に従う確率ベクトル \mathbf{X} は

$$(2.1) \quad \mathbf{X} = \gamma V^{-1} + \frac{\mathbf{Z}}{\sqrt{V}}$$

で与えられる。ただし、 $\gamma \in \mathbb{R}^d$ で、 V はガンマ分布 $G(\nu/2, \nu/2)$ に従い、 \mathbf{Z} は V とは独立に d 変量の正規分布 $N_d(\mathbf{0}, \Psi)$ に従う。ここで、 γ は歪みを表すパラメータベクトルである。この分布は、Barndorff-Nielsen (1977) が提案した一般化双曲型分布の多変量版 (Blaesild and Jensen, 1981) になっており、Demarta and McNeil (2005) で示されているように密度関数は次式で与えられる。

$$(2.2) \quad f(\mathbf{x}; \Psi, \gamma, \nu) = \frac{2^{\frac{2-(\nu+d)}{2}} K_{\frac{\nu+d}{2}}(\eta(\mathbf{x}; \Psi, \gamma, \nu)) \exp(\mathbf{x}^\top \Psi^{-1} \gamma)}{(\pi\nu)^{d/2} \Gamma(\nu/2) \sqrt{|\Psi|} (\eta(\mathbf{x}; \Psi, \gamma, \nu))^{-\frac{\nu+d}{2}} \left(1 + \frac{\mathbf{x}^\top \Psi^{-1} \mathbf{x}}{\nu}\right)^{\frac{\nu+d}{2}}}.$$

ただし, $\eta(\mathbf{x}; \Psi, \gamma, \nu) = \sqrt{(\nu + \mathbf{x}^\top \Psi^{-1} \mathbf{x}) \gamma^\top \Psi^{-1} \gamma}$ であり, $K_\lambda(\cdot)$ は第3種の修正ベッセル関数である. 確率ベクトル(2.1)の期待値と分散共分散行列は

$$(2.3) \quad E[\mathbf{X}] = \frac{\nu}{\nu-2} \gamma, \quad \text{cov}[\mathbf{X}] = \frac{\nu}{\nu-2} \Psi + \frac{2\nu^2}{(\nu-2)^2(\nu-4)} \gamma \gamma^\top$$

与えられ, $\nu \leq 4$ では分散や共分散は発散する.

GH 非対称 t 分布の性質については, 単変量が中心ではあるものの Aas and Haff (2006) で詳しく記述されているほか, 増田 (2002) ではより広く GH 分布の性質について纏めている. $\nu \rightarrow \infty$ では d 変量の(標準)正規分布に収束し, $\gamma = \mathbf{0}$ ならば(対称な) d 変量の t 分布に帰着する. $\gamma = \mathbf{0}$ ならば $\eta(\mathbf{x}; \Psi, \mathbf{0}, \nu) = 0$ であり, 増田 (2002) の(A.8)に示されているように, 第3種の修正ベッセル関数は $\eta \rightarrow 0$ で $K_\lambda(\eta) \sim \Gamma(\lambda) 2^{\lambda-1} \eta^{-\lambda}$ となることから, 密度関数(2.2)は次式のように尺度行列を Ψ とする自由度 ν の d 変量 t 密度関数 $t_{d,\nu}(\mathbf{x}; \Psi)$ に帰着する.

$$(2.4) \quad t_{d,\nu}(\mathbf{x}; \Psi) = \frac{\Gamma((\nu+d)/2)}{(\pi\nu)^{d/2} \Gamma(\nu/2) \sqrt{|\Psi|}} \left[1 + \frac{\mathbf{x}^\top \Psi^{-1} \mathbf{x}}{\nu}\right]^{-\frac{\nu+d}{2}}.$$

2.1.2 Sahu et al. (2003)の非対称 t 分布

Sahu et al. (2003)の非対称 t 分布に従う確率ベクトル \mathbf{X} は,

$$(2.5) \quad \mathbf{X} = \frac{D|\mathbf{W}| + \mathbf{Z}}{\sqrt{V}}$$

与えられる. ただし, $D = \text{diag}(\tilde{\delta}_1, \dots, \tilde{\delta}_d)$ は歪みパラメータの行列, V はガンマ分布 $G(\nu/2, \nu/2)$ に従う確率変数, \mathbf{Z} は V とは独立に d 変量正規分布 $N_d(0, \Sigma)$ に従う確率ベクトル, $\mathbf{W} = (W_1, \dots, W_d)^\top$ は V および \mathbf{Z} とは独立に d 変量正規分布 $N_d(0, I)$ に従う確率ベクトルである. また, Σ は \mathbf{Z} の相関行列, I は $d \times d$ の単位行列である. GH 非対称 t 分布に従う確率ベクトルが(2.1)のように t 分布に従う確率変数 Z/\sqrt{V} を生成してから, そこで用いたガンマ確率変数 V を用いて歪みのファクター γV^{-1} を加えて生成されているのに対し, Sahu et al. (2003)の非対称 t 分布に従う確率ベクトル \mathbf{X} は, まず \mathbf{Z} に歪みファクター $D|\mathbf{W}|$ を加えて非対称正規確率変数を生成した後に, \sqrt{V} で除することによって生成されているのが特徴である. 確率ベクトル(2.5)の密度関数は, Sahu et al. (2003)の(11)式で与えられているように, 尺度行列を $I - D(\Sigma + D^2)^{-1}D$ とする自由度 $\nu + d$ の d 変量 t 分布に従う確率ベクトル \mathbf{U} を用いて,

$$(2.6) \quad g(\mathbf{x}; \Sigma, D, \nu) = 2^d t_{d,\nu}(\mathbf{x}; \Sigma + D^2) \Pr[\mathbf{U} \leq D(\Sigma + D^2)^{-1} \mathbf{x}]$$

与えられる. 確率ベクトル(2.5)の期待値と分散共分散行列は, $\tilde{\delta} = (\tilde{\delta}_1, \dots, \tilde{\delta}_d)^\top$ として,

$$(2.7) \quad E[\mathbf{X}] = \left(\frac{\nu}{\pi}\right)^{1/2} \frac{\Gamma((\nu-1)/2)}{\Gamma(\nu/2)} \tilde{\delta}$$

$$(2.8) \quad \text{cov}[\mathbf{X}] = \frac{\nu}{\nu-2} (\Sigma + D^2) - \frac{\nu}{\pi} \left\{ \frac{\Gamma((\nu-1)/2)}{\Gamma(\nu/2)} \right\}^2 D^2$$

与えられ, $\nu \leq 2$ では分散や共分散は発散する. $\nu \rightarrow \infty$ では d 変量の非対称正規分布に収束し, $\text{diag} D = \mathbf{0}$ ならば密度関数(2.6)で $\Pr[\mathbf{U} \leq \mathbf{0}] = 1/2^d$ となり(対称な) d 変量の t 分布に帰着する. $\text{diag} D = \mathbf{0}$ でかつ $\nu \rightarrow \infty$ ならば d 変量の標準正規分布に帰着する.

2.1.3 AC 非対称 t 分布

AC 非対称 t 分布に従う確率ベクトル \mathbf{X} は, Sahu et al. (2003) と同様に, 非対称正規確率ベクトル $\mathbf{Y} = (Y_1, \dots, Y_d)^\top$ を生成した後に, \mathbf{Y} とは独立にガンマ分布 $G(\nu/2, \nu/2)$ に従う V の平方根で除することによって, 次式のように与えられる.

$$(2.9) \quad \mathbf{X} = \frac{\mathbf{Y}}{\sqrt{V}}.$$

ここで, Y_j ($j = 1, \dots, d$) は

$$(2.10) \quad Y_j = \delta_j |Z_0| + \sqrt{1 - \delta_j^2} Z_j, \quad Z_0 \sim N(0, 1), \quad \mathbf{Z} = (Z_1, \dots, Z_d)^\top \sim N_d(0, \Psi)$$

で与えられる. Z_0, \mathbf{Z}, V は互いに独立である. Sahu et al. (2003) の確率ベクトルを与えた (2.5) 式において, 右辺の分子を \mathbf{Y} で書き表すと $Y_j = \delta_j |W_j| + Z_j$ となっており, 非対称性を与えるファクターが $|Z_0|$ と共通しているか, $|W_j|$ と独立に与えているかの違いがあることがわかる. 非対称性を与える 1 つの共通ファクター $|Z_0|/\sqrt{V}$ は 1 変量 t 分布に従うことから, d 変量の標準 AC 非対称 t 分布の密度関数は,

$$(2.11) \quad g(\mathbf{x}; \Omega, \boldsymbol{\alpha}, \nu) = 2t_{d,\nu}(\mathbf{x}; \Omega) T_{1,\nu+d} \left(\boldsymbol{\alpha}^\top \mathbf{x} \sqrt{\frac{\nu+d}{\mathbf{x}^\top \Omega^{-1} \mathbf{x} + \nu}} \right)$$

と表せる. ただし, $\boldsymbol{\alpha} \in \mathbb{R}^d$ は歪みを表すパラメータであり, $T_{1,\nu}(\cdot)$ は自由度 ν の 1 変量 t (累積) 分布関数である. 尺度行列 Ω は, (2.10) で用いた相関行列 Ψ を用いて, 次式のように表される.

$$(2.12) \quad \Omega = \Delta(\Psi + \boldsymbol{\zeta} \boldsymbol{\zeta}^\top) \Delta, \quad \Delta \equiv \text{diag}(\sqrt{1 - \delta_1^2}, \dots, \sqrt{1 - \delta_d^2}).$$

ここで, $\boldsymbol{\zeta} = (\zeta_1, \dots, \zeta_d)^\top \in \mathbb{R}^d$ は元々の歪みパラメータ $\boldsymbol{\delta} = (\delta_1, \dots, \delta_d)^\top$ と各成分が 1 対 1 に対応する別表現の歪みパラメータで

$$(2.13) \quad \zeta_j = \frac{\delta_j}{\sqrt{1 - \delta_j^2}}, \quad j = 1, \dots, d$$

で定義される. また, (2.11) での歪みを表すパラメータ $\boldsymbol{\alpha}$ は, 次式のように表される.

$$(2.14) \quad \boldsymbol{\alpha} = \frac{\Omega^{-1} \boldsymbol{\delta}}{\sqrt{1 - \boldsymbol{\delta}^\top \Omega^{-1} \boldsymbol{\delta}}} = \frac{\Delta^{-1} \Psi^{-1} \boldsymbol{\zeta}}{\sqrt{1 + \boldsymbol{\zeta}^\top \Psi^{-1} \boldsymbol{\zeta}}}.$$

確率ベクトル \mathbf{X} は (2.9), (2.10) に従って構成可能であるが, より簡便な方法として,

$$(2.15) \quad R = \begin{pmatrix} 1 & \boldsymbol{\delta}^\top \\ \boldsymbol{\delta} & \Omega \end{pmatrix}$$

という $(d+1) \times (d+1)$ の相関行列 R を持つ $d+1$ 変量正規分布ベクトル $\tilde{\mathbf{Z}}$ を $\tilde{\mathbf{Z}}$ とは独立な $V \sim G(\nu/2, \nu/2)$ の平方根で除して自由度 ν の t 分布に従う確率ベクトル $(\tilde{X}_0, \tilde{X}_1, \dots, \tilde{X}_d)^\top$ を生成し,

$$(2.16) \quad \mathbf{X} = \begin{cases} (\tilde{X}_1, \dots, \tilde{X}_d)^\top & \text{if } \tilde{X}_0 \geq 0 \\ -(\tilde{X}_1, \dots, \tilde{X}_d)^\top & \text{if } \tilde{X}_0 < 0 \end{cases}$$

で \mathbf{X} を定める方法もある. 本稿でも 3 節の分析では (2.16) の表現を用いる. 確率ベクトル (2.9) の期待値と分散共分散行列は

$$(2.17) \quad E[\mathbf{X}] = \left(\frac{\nu}{\pi}\right)^{1/2} \frac{\Gamma((\nu-1)/2)}{\Gamma(\nu/2)} \boldsymbol{\delta}$$

$$(2.18) \quad \text{cov}[\mathbf{X}] = \frac{\nu}{\nu-2} \Omega - \frac{\nu}{\pi} \left\{ \frac{\Gamma((\nu-1)/2)}{\Gamma(\nu/2)} \right\}^2 \boldsymbol{\delta} \boldsymbol{\delta}^\top$$

与えられ、2.1.2 節の (2.8) と同様に、 $\nu \leq 2$ では分散や共分散は発散する。 $\nu \rightarrow \infty$ では d 変量の非対称正規分布に収束し、 $\boldsymbol{\delta} = \mathbf{0}$ ならば $\boldsymbol{\alpha} = \mathbf{0}$ で、密度関数 (2.11) で $T_{1,\nu+d}(0) = 1/2$ となり、(対称な) d 変量の t 分布に収束する。 $\boldsymbol{\delta} = \mathbf{0}$ でかつ $\nu \rightarrow \infty$ ならば d 変量の標準正規分布に収束する。

Sahu et al. (2003) と Azzalini and Capitanio (2003) の非対称 t 分布に従う確率変数の構成のように潜在変数の正負に応じて分布を歪ませる方法は、非対称分布の構成方法としては最も一般的なものであり、どちらも Branco and Dey (2001) で提案された多変量非対称楕円分布の一般的な構成方法に準じている。

フィッシャー情報行列について、Arellano-Valle (2010) は、多変量 AC 非対称 t 分布が多変量非対称正規分布と違い、 $\boldsymbol{\alpha} = \mathbf{0}$ で特異にならないということを指摘している。すなわち、多変量 AC 非対称 t 分布は多変量非対称正規分布よりも推定精度が高まりやすいことを示唆している。多変量 AC 非対称 t 分布や関連する分布については Azzalini (2013) に詳しく記述されている。

2.2 導出される接合関数

一般に、 $\boldsymbol{\theta}$ をパラメータとする d 変量の分布関数 $G(\mathbf{x}; \boldsymbol{\theta})$ に内包される接合関数 $C_G(u_1, \dots, u_d; \boldsymbol{\theta}_C)$ は、各 j ($j = 1, \dots, d$) 変量の周辺分布の分位点関数(分布関数の逆関数)を $G_j^{-1}(x_j; \boldsymbol{\theta}_j)$ として、

$$(2.19) \quad C_G(u_1, \dots, u_d; \boldsymbol{\theta}_C) = G(G_1^{-1}(u_1; \boldsymbol{\theta}_1), \dots, G_d^{-1}(u_d; \boldsymbol{\theta}_d); \boldsymbol{\theta})$$

と定義される。すなわち、接合関数の特定には、基となる d 変量の分布を特定するとともに、その周辺分布を特定する必要がある。そこで、ここでは 2.1 節で導入した 3 種類の d 変量標準非対称 t 分布の周辺分布を特定することで、 $\boldsymbol{\theta}$ 、 $\boldsymbol{\theta}_j$ ($j = 1, \dots, d$)、 $\boldsymbol{\theta}_C$ を定めて、導出される接合関数を特定する。

2.2.1 GH 非対称 t 分布の周辺分布

(2.1) より第 j 成分の周辺分布は、

$$(2.20) \quad X_j = \gamma_j V^{-1} + \frac{Z_j}{\sqrt{V}}$$

与えられるため、密度関数は (2.2) で $d = 1$ を与えたものに相当し、

$$(2.21) \quad f_j(x; \gamma_j, \nu) = \frac{2^{\frac{1-\nu}{2}} K_{\frac{\nu+1}{2}}(\sqrt{\gamma_j^2(\nu+x^2)}) \exp(\gamma_j x)}{\Gamma\left(\frac{\nu}{2}\right) \sqrt{\pi\nu} (\sqrt{\gamma_j^2(\nu+x^2)})^{-\frac{\nu+1}{2}} \left(1 + \frac{x^2}{\nu}\right)^{\frac{\nu+1}{2}}}$$

与えられる。したがって、導出される接合関数は、 $\boldsymbol{\theta}_C = \boldsymbol{\theta} = (\gamma, \Psi, \nu)$ 、 $\boldsymbol{\theta}_j = (\gamma_j, \nu)$ ($j = 1, \dots, d$) で特定される。

2.2.2 Sahu et al. (2003) の非対称 t 分布の周辺分布

(2.5) より第 j 成分の周辺分布は、

$$(2.22) \quad X_j = \frac{\tilde{\delta}_j |W_j| + Z_j}{\sqrt{V}}$$

で与えられ、密度関数は

$$(2.23) \quad g_j(x; \zeta_j, \nu) = 2t_{1,\nu}(x; 1 + \tilde{\delta}_j^2) T_{1,\nu+1} \left(\frac{\tilde{\delta}_j}{\sqrt{1 + \tilde{\delta}_j^2}} x \sqrt{\frac{(\nu+1)(1 + \tilde{\delta}_j^2)}{x^2 + \nu(1 + \tilde{\delta}_j^2)}} \right)$$

で与えられる. $t_{1,\nu}(x; \sigma^2)$ は(2.4)を用いて定義される尺度パラメータが σ で自由度 ν の 1 変量 t 分布の密度関数である. したがって、導出される接合関数は、 $\theta_C = \theta = (\tilde{\delta}, \Sigma, \nu)$, $\theta_j = (\tilde{\delta}_j, \nu)$ ($j = 1, \dots, d$) で特定される.

2.2.3 AC 非対称 t 分布の周辺分布

(2.9), (2.10)より第 j 成分の周辺分布は、

$$(2.24) \quad X_j = \frac{\delta_j |Z_0| + \sqrt{1 - \delta_j^2} Z_j}{\sqrt{V}}$$

で与えられ、密度関数は

$$(2.25) \quad g_j(x; \zeta_j, \nu) = 2t_{1,\nu}(x) T_{1,\nu+1} \left(\zeta_j x \sqrt{\frac{\nu+1}{x^2 + \nu}} \right)$$

で与えられる. ここで、 $t_{1,\nu}(\cdot)$ は自由度 ν の 1 変量 t 分布の密度関数で、(2.4)を用いて $t_{1,\nu}(x) = t_{1,\nu}(x; 1)$ と定義される. ζ_j は(2.13)で定義される. (2.11)より $\theta = (\alpha, \Omega, \nu)$, (2.25)より $\theta_j = (\zeta_j, \nu)$ ($j = 1, \dots, d$) と与えられる. θ_C については、 α が(2.14), ζ_j が(2.13)で定義されていることを踏まえると、 $\theta_C = (\delta, \Omega, \nu)$ で導出される接合関数を特定できるため、Yoshiba (2018)では導出される接合関数をそのように特定している.

なお、Kollo and Petter (2010)は早い段階で AC 非対称 t 接合関数の推定を行ったと主張している. しかし、Kollo and Petter (2010)では、(2.25)の ζ_j を α_j とし、周辺分布の特定を誤っているため、その後の議論が誤った方向に進んでいる点に留意されたい.

2.3 裾依存係数

2 変量の累積分布関数 $F(x_1, x_2)$ に対し、各変量の周辺分布関数を $F_j(\cdot)$ ($j = 1, 2$) とすると、下側・上側裾依存係数 λ_L, λ_U は

$$(2.26) \quad \begin{aligned} \lambda_L &= \lim_{u \rightarrow 0^+} \Pr[X_1 \leq F_1^{-1}(u) | X_2 \leq F_2^{-1}(u)], \\ \lambda_U &= \lim_{u \rightarrow 1^-} \Pr[X_1 \geq F_1^{-1}(u) | X_2 \geq F_2^{-1}(u)] \end{aligned}$$

で定義され、接合関数 $C(u_1, u_2)$ を用いると

$$(2.27) \quad \lambda_L = \lim_{u \rightarrow 0^+} \frac{C(u, u)}{u}, \quad \lambda_U = \lim_{u \rightarrow 1^-} \frac{1 - 2u + C(u, u)}{1 - u}$$

と表せる. (2.26)から明らかなように裾依存係数は極値での 2 変量 $\mathbf{X} = (X_1, X_2)$ の依存関係を示すもので、 \mathbf{X} を多変量で考えると金融ポートフォリオのリスク管理などではシステミックリスクを捉える重要な概念に相当している. Fung and Seneta (2010)は、 \mathbf{X} を価格変動とすると、下側裾依存係数は極端な価格急落は市場クラッシュを捉えるものであり、価格、収益率、信用リスクのモデリングで重要な概念であるとして、GH 非対称 t 分布と AC 非対称 t 分布の下側裾依存係数 λ_L を整理している. ここではそれらの裾依存係数について上側裾依存係数も含めて確認する.

2.3.1 GH 非対称 t 分布の場合

GH 非対称 t 分布については, Banachewicz and van der Vaart (2008) が Theorem 2.1 で示した特殊ケースの場合として, Fung and Seneta (2010) が Theorem 1 で下側裾依存係数 λ_L を次のように整理している.

(1) $\gamma_1 = \gamma_2 = 0$ の場合,

$$\lambda_L = 2T_{1,\nu+1} \left(-\sqrt{\frac{(\nu+1)(1-\rho)}{1+\rho}} \right)$$

(2) $\gamma_1 > 0$ で $\gamma_2 > 0$ の場合, $\lambda_L = 0$

(3) $\gamma_1 < 0$ で $\gamma_2 < 0$ の場合, $\lambda_L = 1$

(4) $\gamma_1 > 0$ で $\gamma_2 < 0$ の場合, $\lambda_L = 0$

(5) $\gamma_1 = 0$ で $\gamma_2 > 0$ の場合, $\lambda_L = 0$

(6) $\gamma_1 = 0$ で $\gamma_2 < 0$ の場合,

$$\lambda_L = \int_0^1 \left(1 - \Phi \left(\left(\frac{2^{\nu/2} \Gamma((\nu+1)/2)}{2\sqrt{\pi}} \right)^{1/\nu} u^{1/\nu} \right) \right) du.$$

上記の結果から, $\gamma_1 = \gamma_2 = \gamma$ とすると, $\gamma < 0$ であれば $\lambda_L = 0$, $\gamma > 0$ であれば $\lambda_L = 1$ と γ の正負で上下双方向の裾依存係数が極端に変化することがわかる. λ_L は γ_1, γ_2 に応じて変化するため, 必要に応じて $\lambda_L(\gamma_1, \gamma_2)$ と記述することにする.

上側裾依存係数 $\lambda_U(\gamma_1, \gamma_2)$ については, Banachewicz and van der Vaart (2008) で示されているが, 下側裾依存係数 $\lambda_L(\gamma_1, \gamma_2)$ から導出可能である. (2.1) において, \mathbf{Z} は d 変量標準正規分布であることから, $\Pr[\mathbf{Z}/\sqrt{V} \leq -x] = \Pr[\mathbf{Z}/\sqrt{V} \geq x]$ が成立する. したがって

$$(2.28) \quad \Pr[\gamma V^{-1} + \mathbf{Z}/\sqrt{V} \geq x] = \Pr[-\gamma V^{-1} + \mathbf{Z}/\sqrt{V} \leq -x]$$

となり, 周辺分布についても同様の関係が成立するため,

$$(2.29) \quad \lambda_U(\gamma_1, \gamma_2) = \lambda_L(-\gamma_1, -\gamma_2)$$

となる.

2.3.2 AC 非対称 t 分布の場合

Fung and Seneta (2010) と Padoan (2011) は独立に 2 変量 AC 非対称 t 分布の裾依存係数を与えている. 特に Padoan (2011) は, 下側裾依存係数 λ_L だけでなく上側裾依存係数 λ_U についても

$$(2.30) \quad \lambda_L = F_{EST}(-a_{2,1}; \alpha_2 \sqrt{1-\rho^2}, -\tau_1, \nu+1) + F_{EST}(-a_{1,2}; \alpha_1 \sqrt{1-\rho^2}, -\tau_2, \nu+1)$$

$$(2.31) \quad \lambda_U = 2 - F_{EST}(a_{2,1}; \alpha_2 \sqrt{1-\rho^2}, \tau_1, \nu+1) - F_{EST}(a_{1,2}; \alpha_1 \sqrt{1-\rho^2}, \tau_2, \nu+1)$$

であることを示している. ここで, $F_{EST}(\cdot)$ は 1 変量の拡張(標準)非対称 t 分布の累積分布関数で

$$F_{EST}(x; \alpha, \tau, \nu) = \int_{-\infty}^x t_{1,\nu}(z) \frac{T_{1,\nu+1} \left((\alpha z + \tau) \sqrt{\frac{\nu+1}{\nu+z^2}} \right)}{T_{1,\nu}(\tau/\sqrt{1+\alpha^2})} dz$$

と定義され, (2.30) に含まれる $a_{2,1}, a_{1,2}, \tau_1, \tau_2$ は

$$a_{2,1} = \left\{ \left(\frac{T_{1,\nu+1}(-\zeta_2\sqrt{\nu+1})}{T_{1,\nu+1}(-\zeta_1\sqrt{\nu+1})} \right)^{1/\nu} - \rho \right\} \sqrt{\frac{\nu+1}{1-\rho^2}}$$

$$a_{1,2} = \left\{ \left(\frac{T_{1,\nu+1}(-\zeta_1\sqrt{\nu+1})}{T_{1,\nu+1}(-\zeta_2\sqrt{\nu+1})} \right)^{1/\nu} - \rho \right\} \sqrt{\frac{\nu+1}{1-\rho^2}}$$

$$\tau_1 = \sqrt{\nu+1}(\alpha_1 + \rho\alpha_2), \tau_2 = \sqrt{\nu+1}(\alpha_2 + \rho\alpha_1)$$

で定義される。

2.3.3 AC 非対称 t 接合関数と GH 非対称 t 接合関数

図 1 は、周辺分布を標準正規分布、接合関数を AC および GH の非対称 t 接合関数としたときの 2 変量密度関数の等高線を描いたものである。自由度パラメータは $\nu = 3$ 、相関パラメータは $\rho = 0.5$ とし、歪みパラメータは AC 非対称 t 接合関数では $\delta_1 = \delta_2 = \delta$ 、GH 非対称 t 接合関数では $\gamma_1 = \gamma_2 = \gamma$ という等しい歪みパラメータで比較している。具体的には $\delta = -0.7$ 、 $\gamma = -0.2$ という設定である。この設定の下では、AC 非対称 t 接合関数 ($\delta = -0.7$) では (2.30) より $\lambda_L \cong 0.327$ 、 $\lambda_U \cong 0.085$ となる。一方、GH 非対称 t 接合関数 ($\gamma = -0.2$) では 2.3.1 節で示した結果から、 $\lambda_L = 1$ 、 $\lambda_U = 0$ であることがわかる。

図 2 は、 $\delta_1 = \delta_2 = \delta$ とし、 $\rho = 0.5$ で AC 非対称 t 分布の下側および上側裾依存係数をプロットした図である。歪みパラメータ δ が小さくなると、下側裾依存係数が大きくなることがわかる。下側と上側の裾依存係数の差は、自由度 ν が小さくなるほど大きくなることがわかる。

2.4 その他の多変量非対称 t 分布

R の rugarch パッケージなど、ファイナンスの時系列データ分析では単変量の非対称 t 分布として Fernández and Steel (1998) の分布が用いられることが多い。これは、歪みパラメータ ξ

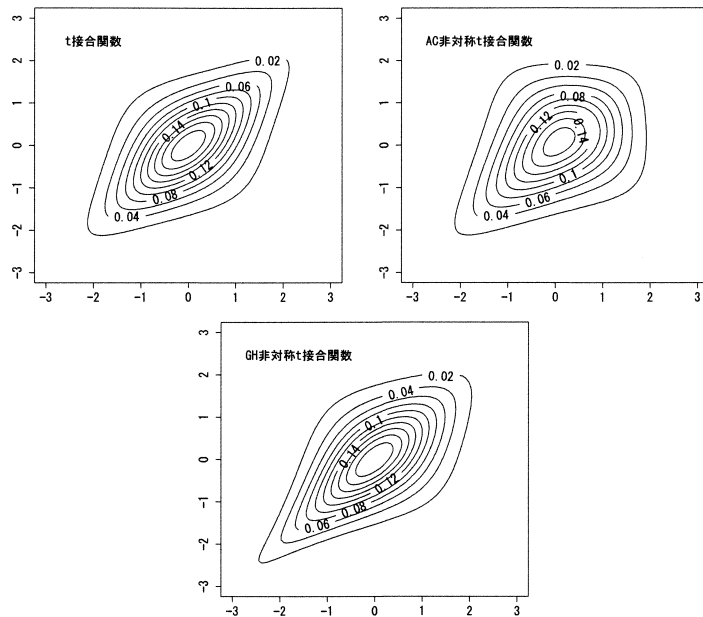


図 1. 周辺分布を標準正規分布、接合関数を t 、AC 非対称 t ($\delta = -0.7$)、GH 非対称 t ($\gamma = -0.2$) の各接合関数としたときの 2 変量密度関数の等高線 ($\rho = 0.5$ 、 $\nu = 3$)。

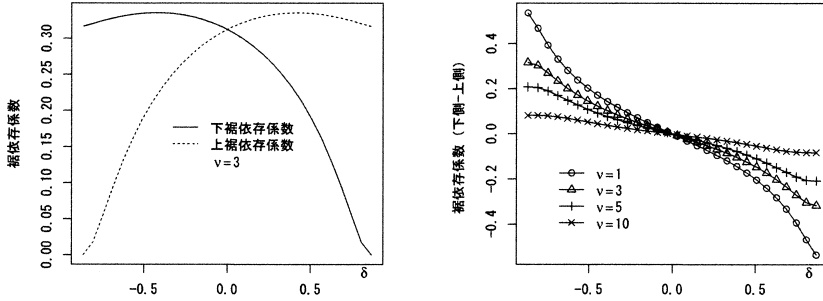


図 2. 2 変量 AC 非対称 t 分布 ($\delta_1 = \delta_2 = \delta$, $\rho = 0.5$) の下側・上側裾依存係数 (左図は $\nu = 3$ での下側依存係数と上側依存係数の値について, 右図は $\nu = 1, 3, 5, 10$ で下側依存係数から上側依存係数を差し引いた値について, 横軸を δ としてプロットしたもの).

と自由度 ν の 1 変量 t 密度関数 $t_{1,\nu}(x)$ を用いて, 密度関数 $f(x|\xi)$ を

$$(2.32) \quad f(x|\xi) = \frac{2\xi}{\xi^2 + 1} \{t_{1,\nu}(x/\xi)1_{[0,\infty)}(x) + t_{1,\nu}(\xi x)1_{(-\infty,0]}(x)\} = \frac{2\xi}{\xi^2 + 1} t_{1,\nu}(x\xi^{-\text{sign}(x)})$$

と定義した非対称 t 分布である. Bauwens and Laurent (2005) は, 周辺密度を (2.32) とする多変量非対称 t 分布を構築し, GARCH モデルと組み合わせてポートフォリオのリスク計測への利用を提唱し, 3 変量の為替ポートフォリオ, 株式ポートフォリオでの実証分析を行っている. Bauwens and Laurent (2005) は, 成分間で独立な $\mathbf{x} = (x_1, \dots, x_d)^\top$ を用いているが, Ferreira and Steel (2007) ではさらに, 非特異行列 $A \in \mathbb{R}^{d \times d}$, $\boldsymbol{\mu} = (\mu_1, \dots, \mu_d) \in \mathbb{R}^d$ を用いて,

$$(2.33) \quad \mathbf{y} = A^\top \mathbf{x} + \boldsymbol{\mu}$$

という線形変換を施すことにより, 相関のある d 変量非対称 t 分布を構成している. 確率ベクトル (2.33) の同時密度は,

$$(2.34) \quad f(\mathbf{y}|\xi) = |A|^{-1} \prod_{j=1}^d f((\mathbf{y} - \boldsymbol{\mu})^\top A_j^{-1} |\xi_j)$$

と与えられる. ここで, A_j^{-1} は A^{-1} の第 j 列ベクトルであり, $|A|$ は A の行列式の絶対値である.

Jones (2002) は, Jones and Faddy (2003) で提案された単変量非対称 t 分布を d 変量に拡張した非対称 t 分布を提案している. Rosco et al. (2011) は, 対称な分布に対して非対称性と裾の重さを与える sinh-arcsinh 変換の一部を用いて, 対称な t 分布に非対称性を与えている. 単変量の議論が中心であるが, d 変量 t 分布に従う確率ベクトル $\mathbf{T}_\nu = (T_{\nu,1}, \dots, T_{\nu,d})^\top$ に対して, 非対称性 $\boldsymbol{\epsilon} = (\epsilon_1, \dots, \epsilon_d)^\top$ を与え,

$$(2.35) \quad T_{\nu,j} = S_{\epsilon_j}(T_{\epsilon_j,\nu,j}) = \sinh(\sinh^{-1}(T_{\epsilon_j,\nu,j} - \epsilon_j)), \quad j = 1, \dots, d$$

として d 変量非対称 t 分布に従う確率ベクトル $\mathbf{T}_{\boldsymbol{\epsilon},\nu} = (T_{\epsilon_1,\nu,1}, \dots, T_{\epsilon_d,\nu,d})^\top$ を構成することを提案している. 本小節で取り上げた d 変量非対称 t 分布について, 応用可能性を踏まえて内在する接合関数の性質や推定法を考察することは今後の課題である.

3. 非対称 t 接合関数の統計的推定方法

3.1 導出される接合密度と最尤推定法

一般に, d 変量確率分布 G に内在する接合関数 (2.19) の密度 $c_G(u_1, \dots, u_d; \boldsymbol{\theta}_G)$ は,

$$(3.1) \quad c_G(u_1, \dots, u_d; \boldsymbol{\theta}_C) = \frac{\partial^d C_G(u_d, \dots, u_d; \boldsymbol{\theta}_C)}{\partial u_1 \dots \partial u_d} = \frac{g(\boldsymbol{x}; \boldsymbol{\theta})}{\prod_{j=1}^d g_j(x_j; \boldsymbol{\theta}_j)}$$

ただし, $\boldsymbol{x} = (x_1, \dots, x_d)^\top$ で

$$(3.2) \quad x_j = G_j^{-1}(u_j; \boldsymbol{\theta}_j), \quad j = 1, \dots, d$$

と定義されることになる.

したがって, 周辺分布について一旦推定した後, その分布関数を用いて N 個の観測データ $\{\boldsymbol{u}_1, \dots, \boldsymbol{u}_N\}$, $\boldsymbol{u}_i \in [0, 1]^d$, $i = 1, \dots, N$ が得られたとすると, この標本に対する対数尤度 $\ell(\boldsymbol{\theta}_C; \boldsymbol{u}_1, \dots, \boldsymbol{u}_N)$ は, (3.1) の接合密度を用いて

$$(3.3) \quad \ell(\boldsymbol{\theta}_C; \boldsymbol{u}_1, \dots, \boldsymbol{u}_N) = \sum_{i=1}^N \ln c_G(\boldsymbol{u}_i; \boldsymbol{\theta}_C)$$

で定義されることになる. このように周辺分布のパラメータを推定してから接合関数のパラメータを推定する方法は 2 段階推定と呼ばれる (詳細と漸近効率性については Joe, 2005 を参照). すなわち, 最尤推定法を適用するのに必要な点は, (a) 想定する d 変量非対称 t 分布の密度, (b) その周辺分布の密度および (c) 周辺分布の分位点関数 (3.2) の 3 点である.

3.1.1 GH 非対称 t 分布

想定する d 変量非対称 t 分布の密度は (2.2), 周辺分布の密度は (2.21) で与えられる. 最尤推定法の適用に必要な残る分位点関数 (3.2) は, 周辺密度 (2.21) を数値積分して累積分布関数 $G(x)$ に直し, 求めたい水準 $q \in (0, 1)$ に対し, $G(x_q) = q$ を満たすような x_q を Newton 法などで探索するのが, 正統的な方法である. ただし, この手法は $[0, 1]$ の観測データが $\{u_{1j}, \dots, u_{Nj}\}$ と観測数 N が大きいときには計算に時間を要してしまうため, 何らかの近似を行うのが実務上は現実的である.

Christoffersen et al. (2012) は, (2.20) に従って 10 万個の乱数を発生させ, その経験分位点を求めることで分位点関数 (3.2) の近似を行った. これに対し, Yoshihara (2018) は, まず, $\gamma_j = \gamma$ とみなす成分の集合を J とし, $j \in J$ について,

$$u_{\min} = \min_{i=1, \dots, N, j \in J} u_{ij}, \quad u_{\max} = \max_{i=1, \dots, N, j \in J} u_{ij}$$

と定義したうえで, 想定される区間 $[x_{\min}, x_{\max}]$ を上記の正統的な方法で $x_{\min} = G^{-1}(u_{\min})$, $x_{\max} = G^{-1}(u_{\max})$ で定めている. 次に, 区間 $[x_{\min}, x_{\max}]$ を $m = 100 \sim 150$ 分割し ($x_{\max} = x_0 < x_1 < \dots < x_m < x_{m+1} = x_{\max}$), 累積分布関数の値 p_1, \dots, p_m を数値積分で求めている. このとき, $p_0 = u_{\min}, p_{m+1} = u_{\max}$ となる. 最後に, $\{(x_0, p_0), (x_1, p_1), \dots, (x_m, p_m), (x_{m+1}, p_{m+1})\}$ の表から, 分割点以外の値は単調補間して, 分位点 $x = G^{-1}(p)$ の値を求めることを提案している. これにより, Christoffersen et al. (2012) の方法よりも精度良く, かつ, 高速に評価できることを確認している.

3.1.2 Sahu et al. (2003) の非対称 t 分布

最尤法で必要な d 変量非対称 t 分布の密度 (2.6) には $\Pr[U \leq D(\Sigma + D^2)^{-1}\boldsymbol{x}]$ が含まれており, d 変量の t 分布の累積分布関数を評価することになるため, 密度を解析的に計算することは困難である. すなわち, Sahu et al. (2003) の非対称 t 分布から導出される接合関数については, 最尤法を適用することは難しい. Smith et al. (2012) は, V や W を条件付けることによって d 変量正規分布の表現にすることで, 導出される接合関数について, マルコフ連鎖モンテカルロ法を用いてバイズ更新でパラメータ推定を行う手法を提示している. Sahu et al. (2003) でもマ

ルコフ連鎖モンテカルロ法を用いた分布の推定法が提案されている。

Smith et al. (2012)のアルゴリズムは、最尤法によるパラメータ推定ではないものの、密度比を利用しながらパラメータ更新を行うため、更新されたパラメータ θ_j ($j = 1, \dots, d$) に対して (3.2) で与えられる周辺分布の分位点評価が必要となる。この点に関し、Smith et al. (2012) は 3~4 回の反復の Newton 法によって十分な精度が得られると記述している。Newton 法は、反復回数を予め定めずに、解の更新幅が収束してきたときに $G(x) = p$ の解 x が得られたとすることが多い。Smith et al. (2012)の方法は、上述の Yoshiba (2018)の手法のように観測データ全体 $\{u_{1j}, \dots, u_{Nj}\}$, $j \in J$ に対して精度の良い近似分位点を求めるのではなく、各観測点 u_{ij} に対する妥当な近似分位点を求めるものと考えられる。

3.1.3 AC 非対称 t 分布

想定する d 変量非対称 t 分布の密度は (2.11)、周辺分布の密度は (2.25) で与えられる。最尤推定法の適用に必要な残る分位点関数 (3.2) は、3.1.1 節で論じた GH 非対称 t 分布の分位点関数と同様である。Yoshiba (2018) は、 δ_j を共通とする変量 $j \in J$ のデータ $\{u_{1j}, \dots, u_{Nj}; j \in J\}$ に対し、想定される区間 $[x_{\min}, x_{\max}]$ を 100~150 分割し、高速かつ精度良く設定したパラメータでの AC 非対称 t 分布の分位点を求めるアルゴリズムを提案している。

3.2 相関行列パラメータの更新

最尤法は、尤度を最大化するようにパラメータを推定する方法であるが、尤度の最大化の過程では設定できないパラメータの領域に更新しないように工夫を施す必要がある。多変量非対称 t 接合関数で留意すべきパラメータは、相関行列のパラメータである。具体的には、GH 非対称 t 接合関数では (2.1) の Z に想定している相関行列 Ψ である。AC 非対称 t 接合関数では (2.15) で定義された拡大相関行列 R である。

Yoshiba (2018) は、 Ψ や R といった相関行列が相関行列としての性質、すなわち、非負定値対称行列で対角要素が 1 という性質を満たすように、相関行列を Cholesky 分解したうえで三角関数を使って θ_{ij} と再パラメータ化する手法を提案している。この方法は、Lewandowski et al. (2009) や Joe (2014) で示されているように、確率ベクトルの第 1 成分から第 $j-1$ 成分を定数としたときの第 i 成分と第 j 成分 ($j < i$) の偏相関係数 $\rho_{ij;1:(j-1)}$ について、 $\cos \theta_{ij} = \rho_{ij;1:(j-1)}$ と再パラメータ化していることに相当している。

3.3 自由度パラメータの更新

一般に、 t 分布や非対称 t 分布の自由度パラメータ ν について、他のパラメータと同様に尤度最大化の過程での更新パラメータとする方法もあるが、整数の自由度に固定してその他のパラメータを可変とした場合の最大尤度を比較し、最も最大尤度が大きくなる自由度を選択するという方法もある。 t 接合関数や非対称 t 接合関数でも同様である。

Yoshiba (2018) は自由度パラメータを他のパラメータと同様の更新パラメータとしている。一方、戸坂・吉羽 (2005) では、 t 接合関数の推定に際して、自由度 ν を 3~20 までの整数の値で固定した最大尤度をプロットし、自由度を推定している。Joe and Sang (2016) では、尤度が自由度 ν に対してフラットであると論じ、GH 非対称 t 接合関数と AC 非対称 t 接合関数の最尤推定をする際に自由度 ν を 5 刻み、すなわち、 $\nu = 5, 10, 15, \dots$ と推定して最大尤度を比較するアプローチを採用している。

AC 非対称 t 接合関数では、整数の自由度に限定すると、対数尤度 (3.3) の計算が高速になる。それは、自由度が 1 と 2 の場合の 1 変量 AC 非対称 t 分布の累積分布関数は解析的に表現され、Jamalizadeh et al. (2009) が示した再帰式によって、整数自由度の累積分布関数は高速に計算さ

れ, Newton 法での分位点の計算も高速になるためである. Joe and Sang (2016)が参照している Yoshiba (2018)のプレプリント版 Yoshiba (2015)では, 自由度が実数の場合の対数尤度の計算と自由度が整数の場合の対数尤度の計算時間が比較されている.

ここで, 自由度 ν , 歪みパラメータ ζ の 1 変量 AC 非対称 t 分布の累積分布関数を $G(x; \nu, \zeta)$ と置く. Jamalizadeh et al. (2009)では,

$$(3.4) \quad G(x; 1, \zeta) = \frac{1}{\pi} \left\{ \tan^{-1}(x) + \cos^{-1} \left(\frac{\zeta}{\sqrt{(1+\zeta^2)(1+x^2)}} \right) \right\}$$

$$(3.5) \quad G(x; 2, \zeta) = \frac{1}{2} - \frac{1}{\pi} \tan^{-1}(\zeta) + \frac{x}{\sqrt{2+x^2}} \left\{ \frac{1}{2} + \frac{1}{\pi} \tan^{-1} \left(\frac{\zeta x}{\sqrt{2+x^2}} \right) \right\}$$

が示され, $\nu > 1$ に対して, Theorem 1 で

$$(3.6) \quad G(x; \nu + 1, \zeta) = G \left(\sqrt{\frac{\nu-1}{\nu+1}} x; \nu - 1, \zeta \right) + \frac{\Gamma(\nu/2)(\nu+1)^{(\nu-1)/2}}{\sqrt{\pi}\Gamma((\nu+1)/2)} \\ \times \frac{x}{(\nu+1+x^2)^{\nu/2}} G \left(\frac{\sqrt{\nu}\zeta x}{\sqrt{\nu+1+x^2}}; \nu, 0 \right)$$

という再帰式を示している. $G(x; \nu, 0)$ は通常の t 分布関数のため, 高速に計算可能である.

3.4 実装

本節で示した Yoshiba (2018)の計算方法については, AC 非対称 t 接合関数における整数自由度での接合密度, 尤度計算を含めて, Yoshiba (2018)の Supplemental material に統計分析ソフト R での実装コードを示している.

4. 実証分析

4.1 非対称 t 接合関数を用いた実証研究

Ammann and Süß (2009)は, 金融時系列の非対称な依存構造に注目し, 2 変量の GH 非対称 t 接合関数を用いて, GARCH(1,1) でフィルタリングした株価の日次収益率と対応するインプライド・ボラティリティ指標について AR-GARCH(1,1) でフィルタリングした日次変化幅との依存構造を調べている. ボラティリティ指標変動の AR 次数は, 0~3 次についてベイズ情報量規準で選択している. 株式市場としては S&P500(ボラティリティ指標は VIX), Nasdaq 100(ボラティリティ指標は VXN), DAX(ボラティリティ指標は VDAXNew), DJ Euro STOXX 50(ボラティリティ指標は VSTOXX)の 4 市場で分析している. 分析の結果, いずれの市場でも t 接合関数よりも GH 非対称 t 接合関数の方が赤池情報量規準, ベイズ情報量規準のいずれでも低く, 良好な結果が得られた. また, 歪みパラメータ γ_j については, 株価収益率に対しては負, ボラティリティ変動幅に対しては正の推定値を得た¹⁾.

Smith et al. (2012)は, AC 非対称 t 分布と似た構成である Sahu et al. (2003)の非対称 t 分布に内在する接合関数を用いて, 電力市場価格などの分析を行っている. 具体的には, オーストラリアの 5 つの地域の 2007 年 1 月~2010 年 2 月の日次の電力価格について, 周辺分布には t や非対称 t 分布, 接合関数には t , 非対称 t , 非対称正規などを当てはめて分析を行った. ベイズの交差検証を行った結果, 周辺分布を非対称 t , 接合関数を非対称 t としたモデルの優位性が示された. 歪みパラメータ δ_j は地域ごと $j = 1, \dots, 5$ に異なる設定でモデリングしており, 推定値の正負は地域によって区々であった. 電力価格は日々の需要の高まりにより急騰することもある点は他の資産価格変動と異なる点であり, その結果が反映されていると考えられる.

Smith et al. (2012)はさらに、周辺分布が離散分布である事例として、15の有名な Web サイトの 2007年5月1日の閲覧件数分布を非対称 t 接合関数と t 接合関数でモデリングした結果を比較し、非対称 t 接合関数の優位性を示している。

Christoffersen et al. (2012)は、GH 非対称 t 接合関数について、Engle (2002)のモデルに倣い、相関行列パラメータ Ψ について、前期のパラメータに回帰する部分と一定水準に回帰する部分とで動的にモデル化したうえで、16か国の先進国と17か国の発展途上国にグルーピングし、週次株価変動について各グループでの相関行列パラメータ Ψ の時間変化などを分析している。各国の株価の週次収益率については、誤差項を非対称 t 分布とする AR(2)-GARCH(1,1)でモデル化している。また、各グループの各時点で GH 非対称 t 接合関数を設定する際、歪みパラメータ γ については各グループで共通しているものと仮定してパラメータを最尤推定し、いずれのグループ、観測期間でも歪みパラメータ γ は概ね負で有意という結果を得ている。

Joe and Sang (2016)は、多変量のモデリングに際し、まず、変量全体をいくつかのグループに分けて、グループ内の接合関数のモデリングにグループ間の接合関数のモデリングを繋げることにより、全体の接合関数をモデリングする手法を提案している。具体例として、銀行・保険、製薬・化学、消費財の3つのグループの企業から4~5社の株価を選択し、計14社の個別株価の2011~2012年の日次収益率について、各変量を AR(1)-GARCH(1,1)でフィルタリングしてから上記の多変量モデリングを行った結果を示している。接合関数には、自由度を5刻みにした AC 非対称 t 、GH 非対称 t 、 t の接合関数を用いて推定し、最も尤度の高かった自由度を選択している。推定結果については、正規化したスコアの推定相関行列と経験相関行列との違いや赤池情報量規準で評価を行っている。評価の結果、相関行列の誤差という観点では、3つのグループをそれぞれ自由度 15, 10, 10 の t 接合関数でモデリングし、自由度 10 の t 接合関数でまとめ上げるモデリングの誤差が小さかった。自由度 15, 10, 10 の AC 非対称 t 接合関数でモデリングし、自由度 10 の t 接合関数でまとめ上げるモデリングは、相関行列の誤差の観点で最も良いそのモデルとほぼ変わらず、赤池情報量規準の観点では選択されることを示している。

Yoshida (2018)は、提案した非対称 t 接合関数の最尤推定の応用事例として、日経 225, S&P500, DAX の3変量の日次株価収益率について、正規、 t 、AC 非対称 t 、非対称正規、GH 非対称 t (歪みパラメータは共通 $\langle \delta_j = \delta, \gamma_j = \gamma \rangle$) の各接合関数を最尤法により推定し、赤池情報量規準やベイズ情報量規準でモデル選択を行っている。計測期間は2005年4月~2015年3月の10年間と後半の5年間の双方とし、各変量については、フィルタリングしない収益率のほか、GARCH(1,1)やEGARCH(1,1)でフィルタリングしてから接合関数を当てはめた推定を行っている。推定の結果、いずれの場合も AC 非対称 t が赤池情報量規準でもベイズ情報量規準でも選択された。自由度パラメータについては、フィルタリングしないもの、GARCH(1,1)でのフィルタリング、EGARCH(1,1)でのフィルタリングの順で大きくなっており、周辺分布のモデリングが精緻になるほど、裾依存性は低下することが示されている。一方で、歪みパラメータ δ はモデリングが精緻になるほど、負の有意性が強まっていることが示されている。

夷藤・中村 (2019)は、Christoffersen et al. (2012)のモデルを修正したうえで、Yoshida (2018)の提案した非対称 t 接合関数の推定法を利用して、新興国債市場の実証分析を行っている。具体的には、4か国の新興国債インデックスの日次リターンについて、各変量については ARMA-GARCH(1,1)でモデル化し、変量間の依存関係については正規、 t 、GH 非対称 t (共通及び個々の歪みパラメータ)の接合関数を用いてモデル化している。なお、AR 項の次数は 0~1、MA 項の時数は 0~2 のモデルを検討し、赤池情報量規準でモデル選択している。資産運用面でのパフォーマンスの計測を念頭に置いているため、為替ヘッジしない場合と為替ヘッジする場合とも比較している。その結果、為替ヘッジしない場合においては、歪みパラメータ γ_j を共通 ($\gamma_j = \gamma$) とした GH 非対称 t 接合関数が赤池情報量規準で選択され、推定された歪みパ

ラメータ γ は有意に負であることが示されている。一方、為替ヘッジをした場合は、 t 接合関数が赤池情報量規準で選択されており、GH 非対称 t 接合関数を推定した場合の歪みパラメータ γ は有意にならないことが示されている。なお、相関行列パラメータ Ψ が時間変化しない静的なモデルに対して、正規、 t 、GH 非対称 t 、AC 非対称 t (歪みパラメータについては共通と個々の双方) の各接合関数を推定し、赤池情報量規準でモデル選択を行った結果、為替ヘッジしない場合には歪みパラメータが共通の AC 非対称 t 接合関数が選択され、為替ヘッジをした場合は t 接合関数が選択されたとの結果を得ている。

これらの先行研究から、本稿で取り上げた 3 種類の非対称 t 接合関数での実証分析において、非対称性は有意になるケースが多く、モデル選択では非対称 t 接合関数が選択されるケースが多いことがわかる。非対称性については、株価などの金融資産価格変動を対象とした場合は負になることが多く、電力市場価格などの実物資産の変動を対象とすると符号は状況によることがわかる。

4.2 本邦の株価変動に対する推定

4.1 節の先行研究を踏まえ、本邦の 3 変量の株価変動に対して AR(1)-GARCH(1,1) でフィルタリングし、その残差に対して、AC 非対称 t 、GH 非対称 t 、 t 、非対称正規、正規の各接合関数を推定し、先行研究との整合性を確認する。株価の日次収益率 r_t に対して AR(1)-GARCH(1,1) は、

$$(4.1) \quad \begin{aligned} r_t &= \mu + a_1 r_{t-1} + \epsilon_t = \mu + a_1 r_{t-1} + \sigma_t z_t, \\ \sigma_t^2 &= \omega + b_1 \epsilon_{t-1}^2 + c_1 \sigma_{t-1}^2 \end{aligned}$$

で表される。基準化された誤差項 z_t には Fernández and Steel (1998) の 1 変量非対称 t 分布 (2.32) を想定する (歪みパラメータ ξ 、自由度パラメータ ν)。非対称 t と非対称正規の接合関数について、非対称性パラメータは各成分で共通のものと個別に設定するものを想定し、パラメータの推定値の符号と有意性を確認し、情報量規準で接合関数を選択する。3 変量の株価としては、東証 33 業種別株価指数から、銀行業、保険業、証券業 (証券・商品先物取引業) の 3 業種を選択する。観測期間は、国際金融危機を含む 2006 年 10 月 2 日～2018 年 9 月 28 日の 12 年間とする (標本サイズは $N=2,941$)。

まず、各業種の日次収益率に対して、R の rugarch パッケージを用いて (4.1) のパラメータ推定を行った結果は表 1 のとおりである。ここで、標準誤差は White (1982) の頑健な標準誤差を用いた。いずれの業種でも歪みパラメータは $\xi > 1$ となっており、正に歪んでいることがわかる。また、自由度パラメータ ν は 7～8 程度と比較的小さく、裾の厚い分布になっていることがわかる。 μ や ω は誤差を考慮すると有意性がない一方、 a_1, b_1, c_1 は保険業の a_1 を除き有意にゼロから異なっていることがわかる。

次に、表 1 で求められたパラメータを (4.1) に代入して得られた残差系列 z_{ij} , $i = 1, \dots, N, j =$

表 1. 周辺分布の推定結果.

	μ	a_1	ω	b_1	c_1	ξ	ν
銀行業	3.19E-04	0.0822	4.77E-06	0.0831	0.9051	1.1145	7.2482
標準誤差	2.89E-04	0.0171	5.20E-06	0.0212	0.0282	0.0317	0.9287
保険業	4.77E-04	0.0175	5.32E-06	0.0793	0.9105	1.0337	8.3885
標準誤差	3.24E-04	0.0188	6.94E-06	0.0249	0.0332	0.0259	1.2846
証券業	1.26E-04	0.0776	4.89E-06	0.0821	0.9116	1.0832	6.9597
標準誤差	3.61E-04	0.0176	6.64E-06	0.0263	0.0317	0.0256	0.9065

1, 2, 3 に対して, Yoshiba (2018) と同様に, $\mathbf{u}_i = (u_{i1}, u_{i2}, u_{i3})$ として, 接合関数推定用の疑似標本 $\{\mathbf{u}_1, \dots, \mathbf{u}_N\}$ を, 各 j 成分 ($j = 1, 2, 3$) に対して

$$(4.2) \quad u_{ij} = \frac{1}{N+1} \sum_{k=1}^N 1_{\{z_{kj} \leq z_{ij}\}}, \quad i = 1, \dots, N$$

のように経験分布関数の一種を用いて求める. 疑似標本 $\{\mathbf{u}_1, \dots, \mathbf{u}_N\}$ に対して AC 非対称 t , GH 非対称 t , t , 非対称正規, 正規の各接合関数を推定した結果は, 表 2, 3 のように与えられ

表 2. 接合関数(共通非対称性)の推定結果.

	AC 非対称 t	GH 非対称 t	t	非対称正規	正規
ρ_{21}	0.8327	0.7811	0.7874	0.8482	0.7868
標準誤差	0.0106	0.0070	0.0064	0.0084	0.0055
ρ_{31}	0.8639	0.8217	0.8266	0.8726	0.8207
標準誤差	0.0087	0.0058	0.0052	0.0071	0.0046
ρ_{32}	0.7967	0.7338	0.7411	0.8106	0.7334
標準誤差	0.0128	0.0084	0.0076	0.0105	0.0068
δ, γ	-0.5909	-0.2574		-0.6699	
標準誤差	0.0581	0.0673		0.0402	
ν	7.6484	7.5062	7.8678		
標準誤差	0.7506	0.7768	0.7870		
対数尤度	3249.49	3249.22	3238.48	3155.66	3147.79
AIC	-6488.98	-6488.43	-6468.96	-6303.31	-6289.58
BIC	-6459.05	-6458.50	-6445.01	-6279.37	-6271.62

表 3. 接合関数(個別非対称性)の推定結果.

	AC 非対称 t	GH 非対称 t	非対称正規
ρ_{21}	0.8296	0.7815	0.8456
標準誤差	0.0116	0.0069	0.0095
ρ_{31}	0.8674	0.8214	0.8701
標準誤差	0.0095	0.0058	0.0070
ρ_{32}	0.7936	0.7342	0.8214
標準誤差	0.0143	0.0082	0.0107
δ_1, γ_1	-0.6145	-0.2711	-0.6386
標準誤差	0.0709	0.0533	0.0465
δ_2, γ_2	-0.5478	-0.2390	-0.8189
標準誤差	0.0755	0.0517	0.0819
δ_3, γ_3	-0.6211	-0.2671	-0.6761
標準誤差	0.0740	0.0523	0.0389
ν	7.6399	7.5349	
標準誤差	0.7560	0.6495	
対数尤度	3249.92	3249.65	3156.93
AIC	-6485.85	-6485.31	-6301.86
BIC	-6443.94	-6443.40	-6265.95

る。赤池情報量規準(AIC), ベイズ情報量規準(BIC)のいずれでも共通非対称性の AC 非対称 t が選択されており, 非対称性のパラメータは, GH 非対称 t や非対称正規の各接合関数も含めて, 非対称性の共通・個別の設定に依らず, 有意に負であることがわかる。相関パラメータ ρ_{ij} は 0.8 前後と高く, 自由度パラメータ ν は 7.5 程度と低いため, 下側の裾依存性が強いことがわかる。

5. おわりに

本稿では, 先行研究で提案されている 3 種類の非対称 t 接合関数を中心に, 多変量非対称 t 分布とその接合関数の性質をまとめたうえで, 最尤法による推定方法が実装可能なものとして AC 非対称 t 接合関数と GH 非対称 t 接合関数での実装方法をまとめた。実証分析としては, 株価や債券価格といった資産収益率に応用した事例を中心に整理し, 本邦の株価変動に対する推定結果を示した。資産収益率に対する実証結果からは, 共通歪みパラメータの AC 非対称 t 接合関数の当てはまりが良好な場合が多いことが示された。

相関行列パラメータが Engle (2002) のモデルのように動的に変化するモデルについては, 先行研究は GH 非対称 t 接合関数での事例にとどまっている。しかし, 夷藤・中村 (2019) の分析で相関行列パラメータが変化しない静的なモデルにおいて, GH 非対称 t 接合関数よりも AC 非対称 t 接合関数が選ばれていたことを勘案すると, 相関行列パラメータ Ω あるいは R が動的に変化する AC 非対称 t 接合関数のモデルが有効になる可能性も高い。そうした観点で, 動的な AC 非対称 t 接合関数の統計的推定方法の実装が今後期待される。

注.

- ¹⁾ Ammann and Süß (2009) の (10) 式で定義されている 2 変量の GH 非対称 t 接合関数の表記は誤っているため (2 変量 GH 非対称 t 分布の周辺分布と GARCH(1, 1) モデルの誤差分布として想定した t 分布を混同), 得られた結果が本来の結果と異なっている可能性がある。

謝 辞

匿名の査読者と編集委員の方には, 表記面で不十分だった原稿に対し, 的確なご示唆を頂きました。ここに記して深謝致します。

参 考 文 献

- Aas, K. and Haff, I. H. (2006). The generalized hyperbolic skew Student's t -distribution, *Journal of Financial Econometrics*, **4**(2), 275–309.
- Ammann, M. and Süß, S. (2009). Asymmetric dependence patterns in financial time series, *The European Journal of Finance*, **15**(7-8), 703–719.
- Arellano-Valle, R. B. (2010). On the information matrix of the multivariate skew- t model, *Metron*, **68**(3), 371–386.
- Azzalini, A. (2013). *The Skew-normal and Related Families*, Cambridge University Press, Cambridge.
- Azzalini, A. and Capitanio, A. (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t -distribution, *Journal of the Royal Statistical Society Series B*, **65**(2), 367–389.
- Banachewicz, K. and van der Vaart, A. (2008). Tail dependence of skewed grouped t -distributions, *Statistics & Probability Letters*, **78**(15), 2388–2399.

- Barndorff-Nielsen, O. E. (1977). Exponentially decreasing distributions for the logarithm of particle size, *Proceedings of the Royal Society of London Series A*, **353**(1674), 401–419.
- Bauwens, L. and Laurent, S. (2005). A new class of multivariate skew densities, with application to generalized autoregressive conditional heteroscedasticity models, *Journal of Business & Economic Statistics*, **23**(3), 346–354.
- Blaesild, P. and Jensen, J. L. (1981). Multivariate distributions of hyperbolic type, *Statistical Distributions in Scientific Work*, Vol. 4, 45–66, D. Reidel Publishing Company, Dordrecht.
- Branco, M. D. and Dey, D. K. (2001). A general class of multivariate skew-elliptical distributions, *Journal of Multivariate Analysis*, **79**(1), 99–113.
- Christoffersen, P., Errunza, V., Jacobs, K. and Langlois, H. (2012). Is the potential for international diversification disappearing? A dynamic copula approach, *Review of Financial Studies*, **25**(12), 3711–3751.
- Demarta, S. and McNeil, A. J. (2005). The t copula and related copulas, *International Statistical Review*, **73**(1), 111–129.
- Engle, R. (2002). Dynamic conditional correlation, *Journal of Business & Economic Statistics*, **20**(3), 339–350.
- Fernández, C. and Steel, M. F. J. (1998). On Bayesian modeling of fat tails and skewness, *Journal of the American Statistical Association*, **93**(441), 359–371.
- Ferreira, J. T. A. S. and Steel, M. F. J. (2007). A new class of skewed multivariate distributions with applications to regression analysis, *Statistica Sinica*, **17**(2), 505–529.
- Fung, T. and Seneta, E. (2010). Tail dependence for two skew t distributions, *Statistics & Probability Letters*, **80**(9–10), 784–791.
- 夷藤 翔, 中村信弘 (2019). ダイナミック非対称 t コピュラを用いた新興国債市場の相互依存構造に関する研究, *ジャフイー・ジャーナル*, **17**, 45–66.
- Jamalzadeh, A., Khosravi, M. and Balakrishnan, N. (2009). Recurrence relations for distributions of a skew- t and a linear combination of order statistics from a bivariate- t , *Computational Statistics & Data Analysis*, **53**(4), 847–852.
- Joe, H. (2005). Asymptotic efficiency of the two-stage estimation method for copula-based models, *Journal of Multivariate Analysis*, **94**(2), 401–419.
- Joe, H. (2006). Discussion of “Copulas: Tales and facts”, by Thomas Mikosch, *Extremes*, **9**(1), 37–41.
- Joe, H. (2014). *Dependence Modeling with Copulas*, Chapman & Hall/CRC, Boca Raton.
- Joe, H. and Sang, P. (2016). Multivariate models for dependent clusters of variables with conditional independence given aggregation variables, *Computational Statistics & Data Analysis*, **97**, 114–132.
- Jones, M. C. (2002). Multivariate t and beta distributions associated with the multivariate F distribution, *Metrika*, **52**(3), 215–231.
- Jones, M. C. and Faddy, M. J. (2003). A skew extension of the t -distribution, with applications, *Journal of the Royal Statistical Society, Series B (Statistical Methodology)*, **65**(1), 159–174.
- Kollo, T. and Pettere, G. (2010). Parameter estimation and application of the multivariate skew t -copula, *Copula Theory and Its Applications*, 289–298, Springer, Heidelberg.
- Kotz, S. and Nadarajah, S. (2004). *Multivariate t Distributions and Their Applications*, Cambridge University Press, Cambridge.
- Lewandowski, D., Kurowicka, D. and Joe, H. (2009). Generating random correlation matrices based on vines and extended onion method, *Journal of Multivariate Analysis*, **100**(9), 1989–2001.
- 増田弘毅 (2002). GIG 分布と GH 分布に関する解析, *統計数理*, **50**(2), 165–199.
- Padoan, S. A. (2011). Multivariate extreme models based on underlying skew- t and skew-normal distributions, *Journal of Multivariate Analysis*, **102**(5), 977–991.
- Rosco, J. F., Jones, M. C. and Pewsey, A. (2011). Skew t distributions via the sinh-arcsinh transformation, *Test*, **20**(3), 630–652.

- Sahu, S. K., Dey, D. K. and Branco, M. D. (2003). A new class of multivariate skew distributions with applications to Bayesian regression models, *Canadian Journal of Statistics*, **31**(2), 129–150.
- Smith, M. S., Gan, Q. and Kohn, R. J. (2012). Modelling dependence using skew t copulas: Bayesian inference and applications, *Journal of Applied Econometrics*, **27**(3), 500–522.
- 戸坂凡展, 吉羽要直 (2005). コピュラの金融実務での具体的な活用方法の解説, *金融研究*, **24**(別冊第 2), 115–162.
- 塚原英敦 (2008). 接合分布関数(コピュラ)の理論と応用, 『21 世紀の統計科学 III 数理・計算の統計科学』(北川源四郎, 竹村彰通 編), 第 5 章, 111–146, 東京大学出版会, 東京.
- White, H. (1982). Maximum likelihood estimation of misspecified models, *Econometrica*, **50**(1), 1–25.
- Yoshihara, T. (2015). Maximum likelihood estimation of skew- t copulas with its applications to stock returns, Research Memo., No.1195, The Institute of Statistical Mathematics, Tokyo.
- Yoshihara, T. (2018). Maximum likelihood estimation of skew- t copulas with its applications to stock returns, *Journal of Statistical Computation and Simulations*, **88**(13), 2489–2506.

Properties of Skew- t Copulas and Their Statistical Estimation —Application to Asset Returns—

Toshinao Yoshiba^{1,2}

¹Graduate School of Management, Tokyo Metropolitan University

²Visiting Professor, The Institute of Statistical Mathematics

The skew- t copula is the copula that is implicit in a multivariate skew- t distribution. There are various types of multivariate skew- t distributions, depending on the way that skewness is incorporated into the multivariate Student- t distribution. First, we summarize the representative multivariate skew- t distributions and show the procedures for applying the maximum likelihood estimation to the three types of skew- t copulas. Next, we refer to empirical studies for the applications of skew- t copulas and show estimation results for three indices from TOPIX33 Sector Indices to indicate the effectiveness of skew- t copulas in representing asset returns. Finally, we conclude by describing future research tasks.

Realized Stochastic Volatility モデル

— 拡張と日本の株価指数への応用 —

高橋 慎¹・大森 裕浩²・渡部 敏明³

(受付 2019 年 6 月 1 日 ; 改訂 11 月 5 日 ; 採択 11 月 6 日)

要 旨

近年、資産価格のボラティリティの推定には日中の高頻度リターンから計算される Realized Volatility (RV) が用いられることが多い。しかし、RV にはマーケット・マイクロストラクチャー・ノイズや夜間や昼休みなどの市場が閉まっている時間帯によってバイアスが生じることが知られている。こうした RV のバイアスを考慮して、日次リターンと日次 RV の同時モデル化が提案されている。このモデルは Stochastic Volatility モデルに RV を加えて拡張したものである。Realized Stochastic Volatility モデルと呼ばれる。Stochastic Volatility モデル同様、Realized Stochastic Volatility モデルは尤度の解析的評価が難しいため、マルコフ連鎖モンテカルロ法 (Markov chain Monte Carlo, MCMC) を用いたベイズ推定法が多く用いられる。本稿では、こうした Realized Stochastic Volatility モデルとその MCMC を用いたベイズ推定法について解説する。また、日次リターン分布やボラティリティの定式化の拡張についても解説する。さらに、日経 225 株価指数に応用し、推定結果を説明する。

キーワード：マルコフ連鎖モンテカルロ法, Realized Volatility, Stochastic Volatility モデル。

1. はじめに

金融資産のリターンを予測することは困難であることが知られているが、その分散を予測することはある程度可能であることが知られている。なぜならリターンの分散(ボラティリティ)は、ある時期には大きい値が続き、またある時期には小さい時期が続くからであり、この現象はボラティリティ・クラスタリングと呼ばれている。ボラティリティの予測を行うことはオプション価格や Value-at-Risk (VaR) の予測、ポートフォリオの最適化など金融リスク管理において重要な意味を持つ (Andersen et al., 2013)。ボラティリティ・クラスタリングという現象を説明する統計的モデルには一般化自己回帰不均一分散 (Generalized Autoregressive Conditional Heteroskedasticity, GARCH) モデル (Engle, 1982; Bollerslev, 1986) と確率的ボラティリティ変動 (Stochastic Volatility, SV) モデル (Taylor, 1986) があり、多くの実証研究が行われてきた(初期研究について詳しくは、例えば Shephard, 1996 や渡部, 2000 を参照されたい)。それらの先行研究において、特に SV モデルは現実のデータへのあてはまりにおいて GARCH モデルより

¹ 法政大学 経営学部：〒 102-8160 東京都千代田区富士見 2-17-1

² 東京大学大学院 経済学研究科：〒 113-0033 東京都文京区本郷 7-3-1

³ 一橋大学 経済研究所：〒 186-8603 東京都国立市中 2-1

優れていることが確認されている。しかし、SV モデルは日々のリターンの分散は観測できない潜在変数として表現しており、その推定精度や予測精度の改善には限界があった。

一方で、近年において利用可能になった資産価格の高頻度データを用いて、日中リターンの 2 乗和として計算される Realized Volatility (RV) は、ボラティリティの観測変数として、推定や予測の精度を改善できるのではないかと注目を集めるようになった (Andersen and Bollerslev, 1998; Barndorff-Nielsen and Shephard, 2002)。RV は、計量経済モデルには依存しないモデル・フリーのボラティリティ推定量として魅力的な推定量であるものの、さまざまな仮定を満たすことが前提となる。しかし、その仮定は現実には満たされておらず、RV はバイアスをもつ (RV の統計的性質について詳しくは、Andersen et al., 2001a; Andersen et al., 2001b, 2003; 渡部, 2007; Andersen et al., 2010 を参照されたい)。例えば、連続な値をとると仮定されている価格が現実には離散的な値をとっていることなどで生じるマーケット・マイクロストラクチャー・ノイズ (MMN) により、RV にバイアスが生じることが知られている。また、夜間や昼休みなど取引のない時間帯の高頻度のリターンを無視すると、そうした時間帯を含む 1 日のボラティリティを推定する場合には、過小評価してしまう (MMN については Campbell et al., 1997; O'Hara, 1995; Hasbrouck, 2007; O'Hara, 2015 を、MMN の RV への影響については Hansen and Lunde, 2006; Bandi and Russell, 2006, 2008; Ubukata and Watanabe, 2014 など参照されたい)。

この問題を克服するために考えられたのが、リターンの SV モデルに RV の情報を追加した同時モデルである Realized SV (RSV) モデルである。Takahashi et al., 2009 によって最初に導入された RSV モデルは、SV モデルに対してさらに RV に関する観測方程式を加えたものであり、その際に RV のバイアス自体もモデル化する。RV のバイアスを補正する推定量は数多く存在するが、バイアスの補正をモデルの中で克服するというアプローチはそれまでになかった。Dobrev and Szerszen (2010)、Koopman and Scharth (2013) は RSV モデルを拡張しており、また、Hansen et al. (2012)、Hansen and Huang (2016) は、同様のアイデアを GARCH モデルに応用した Realized GARCH モデルを提案している。

RSV モデルは、ボラティリティ・クラスタリングと呼ばれる時系列構造と、ボラティリティの非対称性またはレバレッジ効果 (後述) と呼ばれる日次リターンとボラティリティとの依存構造を、2 つの観測方程式と 1 つの状態方程式により記述するモデルである。本稿では、この RSV モデルの解説、推定法の紹介および日本の株価指数への応用を行う。特に推定法についてはマルコフ連鎖モンテカルロ法 (MCMC) を用いたベイズ推定法について紹介する。SV モデルや RSV モデルにおいてはボラティリティを観測されない状態変数とするため、最尤法を行うには状態変数を積分して尤度を求めることが必要である。しかし実際には多重積分の計算は実行困難であり、積分を回避して問題を解くためにベイジアン・アプローチをとり、MCMC 法を用いたシミュレーションによる方法でベイズ推定を行うことが多いのである。

本稿の以下の構成は次の通りである。まず、次の第 2 節で、SV モデルおよび RSV モデルとそれらの MCMC を用いたベイズ推定法について解説する。続く第 3 節で、RSV モデルの拡張について紹介する。第 4 節では、RSV モデルの日本の株価指数への応用を行う。最後に、本稿のまとめを行う。

2. Realized Stochastic Volatility モデルとその推定法

2.1 Stochastic Volatility モデル

SV モデルは、被説明変数の分散変動を説明する代表的なモデルの一つであり、危険資産のリターンによくあてはまることが多くの実証分析において知られている。まず p_t を第 t 日の資産価格とし、第 t 日のリターンを $y_t = \log p_t - \log p_{t-1}$ と定義する。このとき SV モデルは、対

数ボラティリティ h_t を所与として以下のように表現される。

$$(2.1) \quad y_t = \epsilon_t \exp(h_t/2), \quad t = 1, \dots, T,$$

$$(2.2) \quad h_{t+1} = \mu + \phi(h_t - \mu) + \eta_t, \quad t = 1, \dots, T-1, \quad |\phi| < 1,$$

$$\begin{pmatrix} \epsilon_t \\ \eta_t \end{pmatrix} \sim N \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho\sigma_\eta \\ \rho\sigma_\eta & \sigma_\eta^2 \end{pmatrix} \right).$$

ただし、 $h_1 \sim N(\mu, \sigma_\eta^2/(1-\phi^2))$ とする。対数ボラティリティ h_t は観測できないので潜在変数とし、定常な 1 階の自己回帰過程 ($|\phi| < 1$) に従うと仮定する。これは、現実のデータにおいてボラティリティ・クラスティングという、分散が大きい時期には暫く大きい時期が続き、分散が小さい時期には小さい時期が暫く続く現象を反映している。また、誤差項 $(\epsilon_t, \eta_t)'$ は 2 変量正規分布に従うと仮定し、 h_t を所与とした時の y_t と h_{t+1} の相関係数 ρ は、非対称性またはレバレッジ効果と呼ばれる。株式リターンの実証分析においては、 t 時点におけるリターン y_t が減少すると $t+1$ 時点における対数ボラティリティ h_{t+1} が増加することから、 ρ は負の値をとることが知られている。

SV モデルでは観測方程式 (2.1) は非線形ガウス型なので、パラメータ $\theta = (\mu, \phi, \sigma_\eta^2, \rho)$ を所与とする尤度関数の計算を行うことは難しい。このため多くの研究において、シミュレーションを用いるマルコフ連鎖モンテカルロ法によるベイズ推定が行われており、本稿もこれに従う。基本的な SV モデルのベイズ推定の詳細については、紙幅の都合上割愛するが、大森・渡部 (2013) を参照されたい。

2.2 Realized Volatility

SV モデルでは h_t は潜在変数であり、したがってボラティリティ $\exp(h_t)$ も潜在変数であった。これに対して近年では、日中の高頻度データを用いて計算される RV がボラティリティの推定量として提案されている。以下では、その定義、有用性と問題点について見ていくこととする。

まず、真の日次ボラティリティの定義を行う。資産価格の対数値を $p(s)$ とし、それが連続時間の拡散過程

$$(2.3) \quad dp(s) = \mu(s)ds + \sigma(s)dW(s)$$

に従うとする。ここで $\mu(s)$, $\sigma^2(s)$, $W(s)$ をそれぞれ瞬時的なドリフト、ボラティリティ、ウィナー過程とする (ただし、ファイナンスの他の文献では $\sigma^2(s)$ ではなく $\sigma(s)$ をボラティリティと呼ぶことも多い)。このとき、第 t 日の真のボラティリティは、

$$(2.4) \quad IV_t = \int_t^{t+1} \sigma^2(s)ds$$

と定義され、Integrated Volatility (IV) とも呼ばれる。RV は、この IV_t の推定量で、次のように定義される。いま、第 t 日の日中の n 個のリターンデータ $\{r_t, r_{t+1/n}, \dots, r_{t+(n-1)/n}\}$ が与えられているとする。このとき、それらを 2 乗して足し合わせた

$$(2.5) \quad RV_t = \sum_{i=0}^{n-1} r_{t+i/n}^2$$

を第 t 日の RV という。式 (2.5) で定義される RV_t は、 $n \rightarrow \infty$ のとき IV_t に確率収束する。

ところが、実際の高頻度データは RV_t が IV_t に収束するための条件を必ずしも満たしておらず、バイアスが存在することが知られている。例えば、現実の高頻度の資産価格は MMN を含

んでいるので、その結果としてリターンに自己相関が生じる問題や、夜間や昼休みなど取引がない時間帯が存在するという問題がある。

MMN については、取引間隔が短くなればなるほどノイズの占める相対的な割合が多くなるので、RV を計算する際の時間間隔をどのようにとればよいかということが問題となり、多くの提案がなされている (Aït-Sahalia et al., 2005; Bandi and Russell, 2006, 2008). あるいは他の推定量として、Two-Scale または Multi-Scale 推定量 (Zhang et al., 2005; Zhang, 2006) や、カーネル関数を使って対数価格差の 2 乗累積値をスムージングした Realized Kernel (RK) (Barndorff-Nielsen et al., 2008; Barndorff-Nielsen et al., 2009) などが提案されている。これらの方法について詳しくは、Aït-Sahalia and Mykland (2009), Ubukata and Watanabe (2014) などを参照されたい。一方、取引がない時間帯の取り扱いについては、最初に昼休みと夜間を除いて計算した $RV_t^{(o)}$ に、日次リターン y_t の標本分散と $RV_t^{(o)}$ の標本平均の比を乗じる方法がある (Hansen and Lunde, 2005)。

高頻度データから計算される RV や RK などのボラティリティの推定量を、まとめて Realized Measure (RM) と呼ぶ。上記のどのような方法も決定的な解決方法となるわけではなく、また提案された推定量の良い性質が得られるための仮定が現実になり立つとは限らない。このため、第 2.3 節の Realized SV (RSV) モデルでは RM におけるバイアスをモデル・パラメータとすることにより、バイアスの存在に統計的モデルの枠組みの中で対処していくこととする。RM について詳しくは、渡部 (2007), Andersen and Benzoni (2009), McAleer and Medeiros (2008) などを参照されたい。

2.3 Realized SV モデル

Takahashi et al. (2009) では、日次リターンの SV モデルに RM の観測方程式を追加し、バイアスを考慮に入れた同時モデルを提案した。このモデルは Realized SV (RSV) モデルと呼ばれており、以下では RSV モデルとそのモデル・パラメータの MCMC による推定方法について述べる。

第 2.1 節の式(2.1), (2.2) のように y_t を第 t 日の日次リターン、 h_t を真のボラティリティの対数値とし、さらに x_t を RM の対数値とすると、RSV モデルは以下のように表される。

$$(2.6) \quad y_t = \exp(h_t/2)\epsilon_t, \quad t = 1, \dots, T,$$

$$(2.7) \quad h_{t+1} = \mu + \phi(h_t - \mu) + \eta_t, \quad t = 1, \dots, T-1, \quad |\phi| < 1,$$

$$(2.8) \quad x_t = \xi + h_t + u_t, \quad t = 1, \dots, T,$$

$$(2.9) \quad \begin{pmatrix} \epsilon_t \\ \eta_t \\ u_t \end{pmatrix} \sim N \left(\mathbf{0}, \begin{pmatrix} 1 & \rho\sigma_\eta & 0 \\ \rho\sigma_\eta & \sigma_\eta^2 & 0 \\ 0 & 0 & \sigma_u^2 \end{pmatrix} \right).$$

式(2.8)における $\xi + u_t$ は、MMN 等によって生じる、RM の対数値 $x_t = \log RM_t$ と真のボラティリティの対数値 h_t との乖離を表す。 ξ は、 $\log RM_t$ に含まれるバイアスを表し、 $\xi = 0$ はバイアスがないことを意味する。MMN によって RV_t が IV_t を過大評価するなら (Hansen and Lunde, 2006 は RV_t が IV_t を過小評価する可能性もあることを示している)、RM として RV_t を用いたとき $\xi > 0$ となることが予想される。一方、夜間や昼休みなど取引のない時間帯を無視して RM を計算すると、そうした時間帯を含む 1 日のボラティリティを過小評価するので、 $\xi < 0$ となることが予想される。したがって MMN による正のバイアスよりも、夜間や昼休みなど取引のない時間帯を無視したことによる負のバイアスが大きい時に、 ξ は負の値になると考えられる。また、 u_t は ξ では捉えることのできない MMN 等によって生じる観測誤差

である。日次リターンの符号と MMN に相関があるとは考えにくいので、 u_t と η_t は無相関と仮定する。一般に、 u_t と ϵ_t は独立ではない (Hansen and Lunde, 2006; Koopman and Scharth, 2013) が、ここでは簡単化のためそれらは無相関と仮定する。なお、式(2.8)は、より一般的に $x_t = \xi + \psi h_t + u_t$ とすることもできるが、ボラティリティの予測精度が必ずしも改善しないことなどが実証分析において知られているので、本稿においては $\psi = 1$ であると仮定する。

2.4 RSV モデルの MCMC を用いたベイズ推定

RSV モデルのパラメータは $\theta = (\mu, \phi, \sigma_\eta^2, \rho, \xi, \sigma_u^2)$ で、観測値を $\mathbf{y} = \{y_t, x_t\}_{t=1}^T$ 、また潜在変数を $\mathbf{h} = \{h_t\}_{t=1}^T$ で表すこととする。モデルのパラメータに関する事前分布については、 μ, ξ には正規分布、 ϕ には確率密度関数が $\pi_\phi(\phi)$ であるような確率分布、 ρ には確率密度関数が $\pi_\rho(\rho)$ であるような確率分布、 $\sigma_\eta^2, \sigma_u^2$ には逆ガンマ分布を仮定する。

$$\begin{aligned} \mu &\sim N(\mu_0, \sigma_0^2), & \phi &\sim \pi_\phi(\phi), & \rho &\sim \pi_\rho(\rho), & \sigma_\eta^2 &\sim IG(n_0/2, S_0/2). \\ \xi &\sim N(m_\xi, s_\xi^2), & \sigma_u^2 &\sim IG(n_u/2, S_u/2). \end{aligned}$$

$(\mu_0, \sigma_0^2, n_0, S_0, m_\xi, s_\xi^2, n_u, S_u)$ はハイパーパラメータで、分析者によって与えられる定数である。これらの同時事前確率密度関数を $\pi(\theta)$ 、 θ を所与としたときの \mathbf{y} と \mathbf{h} の同時確率密度関数を $f(\mathbf{y}, \mathbf{h}|\theta)$ で表せば、 θ と \mathbf{h} の同時事後確率密度関数 $\pi(\theta, \mathbf{h}|\mathbf{y})$ は

$$\begin{aligned} &\pi(\theta, \mathbf{h}|\mathbf{y}) \\ &\propto f(\mathbf{y}, \mathbf{h}|\theta)\pi(\theta) \\ &\propto \frac{1}{\sigma_u^T} \exp \left\{ -\frac{1}{2} \sum_{t=1}^T \left[h_t + \bar{y}_t^2 + \frac{(x_t - \xi - h_t)^2}{\sigma_u^2} \right] \right\} \\ &\quad \times \frac{\sqrt{1-\phi^2}}{\sigma_\eta^T (1-\rho^2)^{\frac{T-1}{2}}} \exp \left\{ -\frac{1}{2\sigma_\eta^2} \left[(1-\phi^2)\bar{h}_1^2 + \frac{1}{1-\rho^2} \sum_{t=1}^{T-1} (\bar{h}_{t+1} - \phi\bar{h}_t - \rho\sigma_\eta \bar{y}_t)^2 \right] \right\} \\ &\quad \times \frac{1}{\sigma_\eta^{n_0+2} \sigma_u^{n_u+2}} \exp \left\{ -\frac{1}{2} \left[\frac{(\mu - \mu_0)^2}{\sigma_0^2} + \frac{S_0}{\sigma_\eta^2} + \frac{(\xi - m_\xi)^2}{s_\xi^2} + \frac{S_u}{\sigma_u^2} \right] \right\} \pi_\phi(\phi)\pi_\rho(\rho) \end{aligned}$$

となる。ただし、 $\bar{h}_t = h_t - \mu$ 、 $\bar{y}_t = \exp(-h_t/2)y_t$ である。Takahashi et al. (2009) では、まずモデル・パラメータと潜在変数の初期値を設定し、以下のように MCMC によるベイズ推定を行う。

- (1) $\mu|\phi, \sigma_\eta^2, \rho, \xi, \sigma_u^2, \mathbf{h}, \mathbf{y}$ からサンプリングする。
- (2) $\phi|\sigma_\eta^2, \rho, \mu, \xi, \sigma_u^2, \mathbf{h}, \mathbf{y}$ からサンプリングする。
- (3) $(\sigma_\eta^2, \rho)|\phi, \mu, \xi, \sigma_u^2, \mathbf{h}, \mathbf{y}$ からサンプリングする。
- (4) $\xi|\phi, \sigma_\eta^2, \rho, \mu, \sigma_u^2, \mathbf{h}, \mathbf{y}$ からサンプリングする。
- (5) $\sigma_u^2|\phi, \sigma_\eta^2, \rho, \mu, \xi, \mathbf{h}, \mathbf{y}$ からサンプリングする。
- (6) $\mathbf{h}|\theta, \mathbf{y}$ からサンプリングする。

各ステップの具体的なサンプリングについては、Takahashi et al. (2009) および大森・渡部 (2013) を参照されたい。

3. Realized Stochastic Volatility モデルの拡張

3.1 日次リターン分布の拡張

SV モデルや RSV モデルを用いて Value-at-Risk (VaR) や期待ショートフォールなど分布の裾

のリスクを計測する場合、日次リターンの誤差項 ϵ_t の分布も重要である。金融資産の日次リターンは通常、尖度が有意に 3 を超え、正規分布よりも裾が厚いことが知られている。第 2.1 節の SV モデルや第 2.3 節の RSV モデルでは、日次リターンの誤差項 ϵ_t の分布として正規分布を仮定しているが、 ϵ_t の分布が正規分布であっても、ボラティリティが変動するのであれば、日次リターン y_t の尖度は 3 を超える（証明については渡部, 2000 の 1.4 節を参照されたい）。しかし、多くの場合、日次リターンの裾の厚さはボラティリティの変動だけでは十分に捉えられないので、渡部 (2005), Omori et al. (2007) では、SV モデルの ϵ_t の分布に t 分布を用いている。正規分布や t 分布は左右対称であるが、近年では非対称な分布も用いられている。非対称な分布にはいくつかあるが、SV モデルでは、中島・大森 (2011), Nakajima and Omori (2012) が Aas and Haff (2006) が提案した一般化双曲非対称 t 分布 (generalized hyperbolic skew Student's t distribution), Abanto-Valle et al. (2015) が Azzalini and Capitanio (2003) が提案した非対称 t 分布を用いている。RSV モデルでは、Takahashi et al. (2016) が一般化双曲非対称 t 分布を用いているので、以下このモデルについて説明する（一般化双曲非対称 t 分布を含むより一般的な分布に、Barndorff-Nielsen (1977) が導入した一般化双曲分布があるが、Prause, 1999; Aas and Haff, 2006; Nakajima and Omori, 2012 では一般化双曲分布はパラメータの推定が難しいことが指摘されており、本稿では一般化双曲非対称 t 分布を説明する）。

一般化双曲非対称 t 分布は以下のように定義される。まず、 ϵ_t が標準正規分布に従い、 z_t が以下の逆ガンマ分布に従うとする。

$$(3.1) \quad z_t \sim IG\left(\frac{\nu}{2}, \frac{\nu}{2}\right), \quad \mu_z = E[z_t] = \frac{\nu}{\nu-2}, \quad \sigma_z^2 = \text{Var}[z_t] = \frac{2\nu^2}{(\nu-2)^2(\nu-4)}.$$

このとき、 $\beta(z_t - \mu_z) + \sqrt{z_t}\epsilon_t$ の分布を一般化双曲非対称 t 分布と呼ぶ。

一般化双曲非対称 t 分布は、2つのパラメータ (β, ν) によって、左右非対称かつ裾の厚い分布を表現できる。図 1 には、いくつかのパラメータの値に対する一般化双曲非対称 t 分布の密度関数が描かれている。図 1(i) は、 $\nu = 10$ と固定したときの $\beta = 0, -1, -2$ に対する密度関数である。 β の値が負の場合、その絶対値が大きくなるにつれて、分布の歪みと裾の厚さのどちらも大きくなっている。一方、図 1(ii) は、 $\beta = -2$ と固定したときの $\nu = 15, 10, 5$ に対する密度

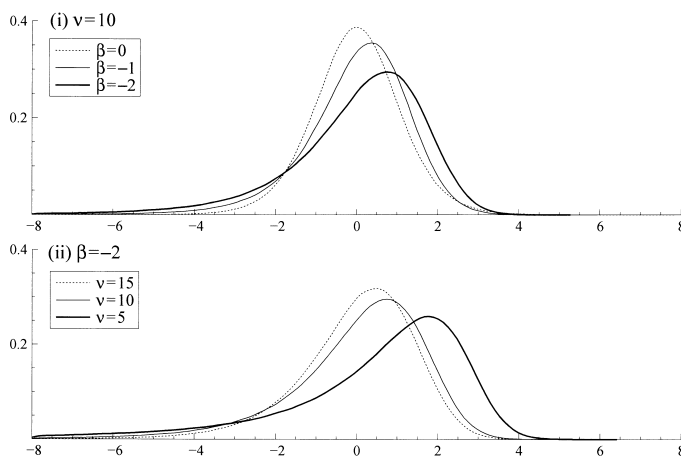


図 1. 一般化双曲非対称 t 分布の密度関数. (i) $\nu = 10$ と固定し $\beta = 0, -1, -2$ の場合.
(ii) $\beta = -2$ と固定し $\nu = 15, 10, 5$ の場合.

関数である。 ν の値が小さくなるほど、裾の厚さが大きくなり、歪み具合も大きくなっている。一般化双曲非対称 t 分布は、 $\beta = 0$ の場合、自由度 ν のスチューデントの t 分布となる。また、 $\nu \rightarrow \infty$ の場合、 $z_t = \mu_z$ となるので、 β の値によらず正規分布となる。以上から、 ν は裾の厚さ（または尖度）を、 β は左右の非対称性（または歪度）を、それぞれ表すパラメータと考えられるが、分布の非対称性と裾の厚さは、2つのパラメータの値の組によって決まることに注意が必要である。

Takahashi et al. (2016)では、式(2.6)の ϵ_t を、分散を1に基準化した $\frac{\beta(z_t - \mu_z) + \sqrt{z_t} \epsilon_t}{\sqrt{\beta^2 \sigma_z^2 + \mu_z}}$ に置き換えることにより、以下のモデルに拡張している。

$$(3.2) \quad y_t = \frac{\beta(z_t - \mu_z) + \sqrt{z_t} \epsilon_t}{\sqrt{\beta^2 \sigma_z^2 + \mu_z}} \exp(h_t/2), \quad t = 1, \dots, T,$$

$$(3.3) \quad x_t = \xi + h_t + u_t, \quad t = 1, \dots, T,$$

$$(3.4) \quad h_{t+1} = \mu + \phi(h_t - \mu) + \eta_t, \quad t = 0, \dots, T-1.$$

ここで、誤差項 $(\epsilon_t, u_t, \eta_t)$ の分布は式(2.9)と同じである。

Takahashi et al. (2016)は、このモデルのMCMCを用いたベイズ推定法を提案しており、以下この方法を簡単に説明する。このモデルのパラメータは第2.3節のRSVモデルのパラメータに一般化双曲非対称 t 分布のパラメータ (β, ν) が加わり、 $\theta = (\mu, \phi, \sigma_\eta^2, \rho, \xi, \sigma_u^2, \beta, \nu)$ となる。また潜在変数は $\mathbf{h} = \{h_t\}_{t=1}^T$ に一般化双曲非対称 t 分布の $\mathbf{z} = \{z_t\}_{t=1}^T$ が加わる。パラメータ $(\mu, \phi, \sigma_\eta^2, \rho, \xi, \sigma_u^2)$ については第2.3節のRSVモデルと同じ事前分布を仮定する。新たなパラメータについては、 ν の事前分布を $\pi_\nu(\nu)$ 、 β の事前分布に以下の正規分布を仮定する。

$$(3.5) \quad \beta \sim N(m_\beta, s_\beta^2).$$

これらの同時事前確率密度関数を $\pi(\theta)$ 、 θ を所与としたときの $\mathbf{y}, \mathbf{h}, \mathbf{z}$ の同時確率密度関数を $f(\mathbf{y}, \mathbf{h}, \mathbf{z}|\theta)$ で表せば、 $\theta, \mathbf{h}, \mathbf{z}$ の同時事後確率密度関数 $\pi(\theta, \mathbf{h}, \mathbf{z}|\mathbf{y})$ は

$$\begin{aligned} & \pi(\theta, \mathbf{h}, \mathbf{z}|\mathbf{y}) \\ & \propto f(\mathbf{y}, \mathbf{h}, \mathbf{z}|\theta) \pi(\theta) \\ & \propto \frac{(\beta^2 \sigma_z^2 + \mu_z)^{\frac{T}{2}}}{\sigma_u^T} \exp \left\{ -\frac{1}{2} \sum_{t=1}^T \left[\log z_t + h_t + \tilde{y}_t^2 + \frac{(x_t - \xi - h_t)^2}{\sigma_u^2} \right] \right\} \\ & \quad \times \frac{\sqrt{1 - \phi^2}}{\sigma_\eta^T (1 - \rho^2)^{\frac{T-1}{2}}} \exp \left\{ -\frac{1}{2\sigma_\eta^2} \left[(1 - \phi^2) \bar{h}_1^2 + \frac{1}{1 - \rho^2} \sum_{t=1}^{T-1} (\bar{h}_{t+1} - \phi \bar{h}_t - \rho \sigma_\eta \tilde{y}_t)^2 \right] \right\} \\ & \quad \times \frac{(\nu/2)^{\frac{T\nu}{2}}}{\Gamma(\nu/2)^T} \exp \left\{ -\frac{1}{2} \sum_{t=1}^T \left[(\nu - 2) \log z_t + \frac{\nu}{z_t} \right] \right\} \\ & \quad \times \frac{1}{\sigma_\eta^{n_0+2} \sigma_u^{n_u+2}} \exp \left\{ -\frac{1}{2} \left[\frac{(\mu - \mu_0)^2}{\sigma_0^2} + \frac{S_0}{\sigma_\eta^2} + \frac{(\xi - m_\xi)^2}{s_\xi^2} + \frac{S_u}{\sigma_u^2} \right] \right\} \pi_\phi(\phi) \pi_\rho(\rho) \pi_\nu(\nu) \end{aligned}$$

となる。ただし、

$$\tilde{y}_t = \frac{\sqrt{\beta^2 \sigma_z^2 + \mu_z} \bar{y}_t - \beta(z_t - \mu_z)}{\sqrt{z_t}}$$

である。Takahashi et al. (2016)では、まずモデル・パラメータと潜在変数の初期値を設定し、以下のようにMCMCによるベイズ推定を行う。

- (1) $\mu|\phi, \sigma_\eta^2, \rho, \xi, \sigma_u^2, \beta, \nu, \mathbf{h}, \mathbf{z}, \mathbf{y}$ からサンプリングする.
- (2) $\phi|\sigma_\eta^2, \rho, \mu, \xi, \sigma_u^2, \beta, \nu, \mathbf{h}, \mathbf{z}, \mathbf{y}$ からサンプリングする.
- (3) $(\sigma_\eta^2, \rho)|\phi, \mu, \xi, \sigma_u^2, \beta, \nu, \mathbf{h}, \mathbf{z}, \mathbf{y}$ からサンプリングする.
- (4) $\xi|\phi, \sigma_\eta^2, \rho, \mu, \sigma_u^2, \beta, \nu, \mathbf{h}, \mathbf{z}, \mathbf{y}$ からサンプリングする.
- (5) $\sigma_u^2|\phi, \sigma_\eta^2, \rho, \mu, \xi, \beta, \nu, \mathbf{h}, \mathbf{z}, \mathbf{y}$ からサンプリングする.
- (6) $\mathbf{h}|\theta, \mathbf{z}, \mathbf{y}$ からサンプリングする.
- (7) $\beta|\phi, \sigma_\eta^2, \rho, \mu, \xi, \sigma_u^2, \nu, \mathbf{h}, \mathbf{z}, \mathbf{y}$ からサンプリングする.
- (8) $\nu|\phi, \sigma_\eta^2, \rho, \mu, \xi, \sigma_u^2, \beta, \mathbf{h}, \mathbf{z}, \mathbf{y}$ からサンプリングする.
- (9) $\mathbf{z}|\theta, \mathbf{h}, \mathbf{y}$ からサンプリングする.

ステップ(1)–(6)はRSVモデルと同様にサンプリングできる. ステップ(7)–(9)の具体的なサンプリングについては, Takahashi et al. (2016)を参照されたい.

一般化双曲非対称 t 分布では, 分布の裾の厚さと非対称性の双方が, モデルのデータへの適合度や予測精度の改善に貢献すると考えられる. 非対称性が貢献しているかどうかはパラメータ β が 0 と異なるかどうかにより判断することが可能であり, 裾の厚さについては自由度 ν の大きさによりある程度は貢献度を測ることができる. Takahashi et al. (2016)は, 日経 225 と S&P500 のデータについて, 一般化双曲非対称 t 分布を用いることにより, モデルのデータへの適合度が改善されることを示していることに加えて, ボラティリティと VaR および期待ショートフォールの予測精度が改善されることを示している. また, Nugroho and Morimoto (2014)は TOPIX のデータについて, Nugroho and Morimoto (2016)は 6 つの株価指数(日経 225, S&P500, FTSE, Nasdaq100, DAZ, DJIA)について, それぞれ一般化双曲非対称 t 分布を用いることによりモデルの適合度が改善されることを示している. これらの実証結果では, パラメータ β が負となる事後確率が非常に高くなる一方, 自由度 ν の事後平均は 22 から 32 と比較的大きく, 日次リターン分布の非対称性を考慮することの重要性が示唆されている.

3.2 ボラティリティの定式化の拡張

3.2.1 長期記憶性

これまで多くの実証分析において, 高頻度データを用いて計算される RV には長期記憶性が存在すると指摘されている(例えば, Andersen et al., 2001a). 定常な確率過程 z_t の自己相関関数 $\rho(s) = \text{Corr}(z_t, z_{t-s})$ が $\sum_{s=0}^{\infty} |\rho(s)| = \infty$ であるとき, z_t は長期記憶定常過程と呼ばれる. その長期記憶性をもつ代表的な確率過程に, 自己回帰実数積分移動平均(Autoregressive Fractionally Integrated Moving Average, ARFIMA)過程があり, h_t の ARFIMA(p, d, q) 過程は次のように定義される.

$$(3.6) \quad (1-L)^d \Phi(L)(h_{t+1} - \mu) = \Theta(L)\eta_t, \quad t = 0, 1, \dots, T.$$

ただし, η_t は誤差項でホワイトノイズ, L は $L^i h_t = h_{t-i}$ となるようなラグ・オペレータ, $\Phi(L) = 1 - \phi_1 L - \dots - \phi_p L^p$, $\Theta(L) = 1 - \theta_1 L - \dots - \theta_q L^q$ である. $d = 1$ であるときには, ARIMA($p, 1, q$) 過程となり, 非定常である. 一般に

$$(3.7) \quad (1-L)^d = 1 + \sum_{j=1}^{\infty} \frac{d(d-1)\cdots(d-j+1)}{j!} (-L)^j = 1 + \sum_{j=1}^{\infty} \gamma_j L^j,$$

$$(3.8) \quad \gamma_0 = 1, \quad \gamma_{j+1} = \frac{j-d}{j+1} \gamma_j, \quad j \geq 0,$$

であり, パラメータ d が長期記憶性を説明するパラメータ節約的なモデルとなっている.

ARFIMA モデルは $|d| < 0.5$ の場合に定常になるが、RV に ARFIMA モデルを適用した実証分析では、多くの場合、 d の推定値は $0 < d < 0.5$ であることが知られている (Andersen et al., 2003; Giot and Laurent, 2004; 渡部・佐々木, 2006)。そこで、 $0 < d < 0.5$ もしくは $|d| < 0.5$ という制約を加えて推定することも多い (Koopman et al., 2005; Nishino and Kakamu, 2013) が、そうした制約を課さないと d の推定値が 0.5 を超えることもある (Ubukata and Watanabe, 2014; Shirota et al., 2014)。

SV モデルにおいても、潜在変数 h_t に ARFIMA モデルを用いた研究が複数ある (Breidt et al., 1998; So, 2002; Ruiz and Veiga, 2008) が、特に Shirota et al. (2014) は RV の情報を観測方程式として以下のように明示的にとりいれている。

$$(3.9) \quad y_t = \exp(h_t/2)\epsilon_t, \quad t = 1, 2, \dots, T,$$

$$(3.10) \quad x_t = \xi + h_t + u_t, \quad t = 1, 2, \dots, T,$$

$$(3.11) \quad (1-L)^d \Phi(L)(h_{t+1} - \mu) = \Theta(L)\eta_t, \quad t = 0, 1, \dots, T-1.$$

ここで、誤差項 $(\epsilon_t, u_t, \eta_t)$ の分布は式 (2.9) と同じである。Shirota et al. (2014) は、モデル・パラメータの推定を行うために、まず Chan and Palma (1998) と同様に ARFIMA 過程を十分な数の有限次元の $MA(q^*)$ 過程で近似し ($q^* > \sqrt{T}$)、状態空間モデルのためのベイズ推定を効率的な MCMC アルゴリズムで行っている。提案されたモデルは、1996 年から 2009 年の日次 S&P500 株価指数リターンとその RV に適用され、RSV モデルや superposition モデル (第 3.2.2 節) よりも、あてはまりもボラティリティ予測精度も良いことが示されている。ただし、RV の長期記憶性は構造変化による見せかけの長期記憶性の可能性がある。長期記憶モデルではないが、構造変化を捉えるモデルとして、高橋 (2014) では、RSV モデルにおいて式 (2.7) の μ を平滑推移 (smooth transition) 関数によって時変にした平滑推移 RSV モデルを提案している。

3.2.2 Superposition モデル

ARFIMA モデルのほかに、対数ボラティリティの長期記憶性を表現する方法としてしばしば用いられるのが superposition モデルである。superposition モデルでは、対数ボラティリティが K 個の独立な ARMA 過程に従う対数ボラティリティの和として表現される。SV モデルにおいては superposition モデルのあてはまりが良いことが先行研究において知られている (Omori et al., 2007) が、Dobrev and Szerszen (2010) はこれを RSV モデルに拡張して、対数ボラティリティが 2 個の独立な AR(1) 過程に従う対数ボラティリティの和として表現されるモデルとし、また日次リターンにジャンプを表現する項も追加した。さらに Koopman and Scharth (2013) は、対数ボラティリティが K 個の独立な AR(1) 過程に従う対数ボラティリティの和として表現される、より一般的な以下のようなモデルを提案した。

$$(3.12) \quad y_t = \exp\left(\frac{\mu + \sum_{j=1}^K \alpha_{jt}}{2}\right)\epsilon_t, \quad t = 1, 2, \dots, T,$$

$$(3.13) \quad x_{it} = \xi_i + \sum_{j=1}^K \alpha_{jt} + u_{it}, \quad t = 1, 2, \dots, T, \quad i = 1, \dots, p,$$

$$(3.14) \quad \alpha_{j,t+1} = \phi_j \alpha_{jt} + \eta_{jt}, \quad \eta_{jt} \sim N(0, \sigma_{\eta_j}^2 / (1 - \phi_j^2)) \quad j = 1, \dots, K, \quad t = 1, \dots, T-1.$$

Koopman and Scharth (2013) は高頻度データから計算される複数の RM ($RM_{i,t}$) を導入し、さらにリターンの誤差項 ϵ_t と観測誤差 u_{it} の相関を考慮するなどの拡張を行い、パラメータの推定法として 2 段階推定法を提案した。superposition モデルは、比較的少ないパラメータ数で長期記憶性を表現することができるが、 K の値や α_{jt} の ARMA 過程の次数をどのように決める

かは明らかではない. さらに RM の観測方程式の個数 p を増やしすぎると, そのバイアスを補正すべき日次リターンの観測方程式の相対的なウェイトが小さくなってしまうので, K の値を変えることで予測の評価が改善できるかどうかを精査していく必要がある. 提案手法は 1993 年から 2010 年までの, 9 つの米国の株式収益率のデータに, RK などの 3 つの RM ($p = 3$) とともに適用され, RK の予測において他のモデルよりも精度が良いことを示している. 石原 (2015) は, Koopman and Scharth (2013) のモデルの式(3.13), (3.14)に曜日, 月, 休日前と後のダミーを加えて推定を行い, 日経 225 のボラティリティと複数の RM の暦効果について分析を行っている.

3.2.3 Box-Cox 変換

RM の観測方程式では, 非負制約をはずすために $x_{it} = \log RM_{it}$ のように RM_{it} の対数変換を行った. その際, 観測方程式では ξ というパラメータを導入することでバイアス補正を行った. これに対してより広いクラスの変換を考えてバイアスの補正を行う拡張が, Nugroho and Morimoto (2014), Nugroho and Morimoto (2016), Zheng and Song (2014) により行われている. Zheng and Song (2014) は, Koopman and Scharth (2013) の superposition モデルにおいて, 式(3.13)の $x_{it} = \log RM_{it}$ を拡張して下記のような Box-Cox 変換を考えた.

$$(3.15) \quad x_{it} = \begin{cases} (RM_{it}^{\lambda_i} - 1)/\lambda_i, & \text{if } \lambda_i \neq 0, \\ \log RM_{it}, & \text{if } \lambda_i = 0. \end{cases}$$

提案モデルは, 2003 年から 2012 年までの, 5 つの日次株価指数リターンについて, RV と RK の 2 つの実現尺度を用いて適用されているが, ボラティリティの予測精度については株価指数やボラティリティの代理変数に何を選ぶかに依存している. Nugroho and Morimoto (2014) では, RSV モデルのリターンの分布を一般化双曲 t 分布や非心 t 分布に拡張し, RV に指数変換, Yeo-Johnson 変換, modulus 変換を適用し, また Nugroho and Morimoto (2016) では Box-Cox 変換を適用している.

4. 日経 225 株価指数への応用

4.1 データと基本統計量

本節では, SV モデルと RSV モデルを日経 225 株価指数に応用する. サンプル期間は 2007 年 1 月 4 日から 2016 年 12 月 30 日までの 10 年間で, サンプルサイズは 2,449 である. 日次リターン (%) は各営業日の終値の対数階差により算出した. また, 日経 225 株価指数の 1 分ごとの価格の対数階差として日中リターン (%) を計算し, バイアスを考慮した RK を算出した (詳しくは大森・渡部, 2013 および Ubukata and Watanabe, 2014 を参照されたい). 大森・渡部 (2013) では, 通常の RV を用いた RSV モデルの場合, MMN によるバイアスを完全には捉えきれないことが示されている. そこで, 本稿では, 式(2.8), (3.3)の x_t としてバイアスを考慮した RK の対数値を用いて, RSV モデルを推定する.

図 2 には日次リターンと RK の対数値 ($\log RK$) の推移が, 図 3 にはヒストグラムが描かれている. また, 表 1 にはこれらの基本統計量が計算されている. まず, 日次リターンを見ると, 平均は 0 から有意に乖離していない. また, LB(10) は 1 次から 10 次までの自己相関がすべて 0 であるという帰無仮説を検定するための Ljung and Box (1978) 統計量の p 値であり, Diebold (1988) の方法により分散不均一性を調整している (詳しくは渡部, 2000, 1.5.1 節を参照されたい). この統計量の p 値によると, 日次リターンに自己相関がないという帰無仮説は有意水準 10% でも棄却されない. したがって, 日次リターンは平均も自己相関も有意でないので, SV モデルの式(2.1)や RSV モデルの式(2.6)の y_t として, 日次リターンをそのまま用いることがで

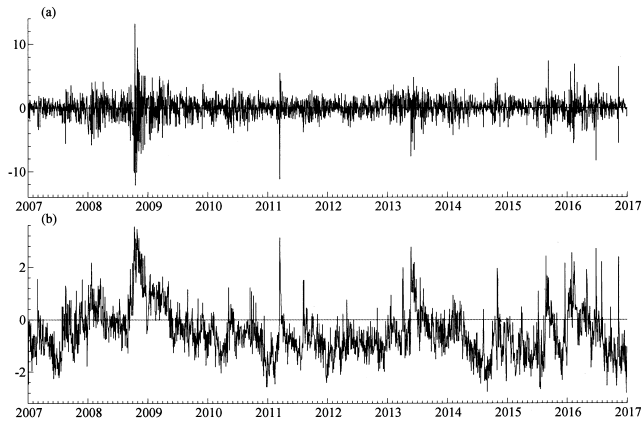


図 2. (a) 日次リターン (%) と (b) RK の対数値の推移.

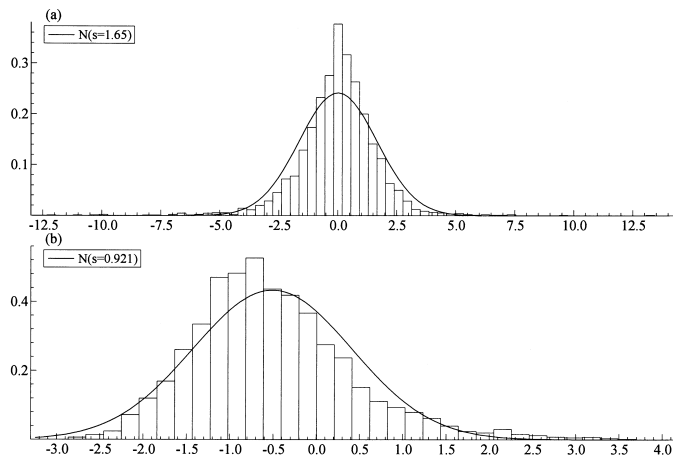


図 3. (a) 日次リターン (%) と (b) RK の対数値のヒストグラム. 実線は, それぞれの平均値と標準偏差に対応した正規分布の確率密度関数.

きる. さらに, 歪度は有意に負となっており, 尖度は正規分布の 3 よりも有意に大きい. JB は分布が正規分布に従うという帰無仮説を検定するための Jarque and Bera (1987) の p 値であり, 日次リターンが正規分布に従うという帰無仮説は有意水準 1% でも棄却される. 図 3(a) のヒストグラムからも, 日次リターンが正規分布から乖離していることは明らかである. 第 3.1 節で述べたように, 日次リターンの裾の厚さはボラティリティの変動だけでは十分に捉えられない場合があり, 有意な非対称性も観測されている. 以下では, 日次リターンの分布として, 正規分布だけでなく一般化双曲非対称 t 分布を用いた SV モデルおよび RSV モデルを推定する.

次に, RK の対数値を見ると, LB(10) から自己相関がないという帰無仮説は有意水準 1% で棄却される. これはボラティリティ・クラスタリングと整合的である. また, 歪度は有意に正となっており, 尖度は正規分布の 3 よりも有意に大きい. JB の値からも, RK の対数値が正規分布に従うという帰無仮説は有意水準 1% で棄却される. 図 3(b) のヒストグラムからも, RK の対数値が正規分布から乖離していることがわかる. したがって, RK の対数値の分布につい

表 1. 日次リターン(%)と RK の対数値の基本統計量. サンプル期間は 2007 年 1 月 4 日～2016 年 12 月 30 日. サンプルサイズは 2,449. 括弧内の数値は平均の標準誤差を表し, 歪度と尖度の標準誤差はそれぞれ 0.049 と 0.099 である. JB は分布が正規分布に従うという帰無仮説を検定するための Jarque and Bera(1987)統計量の p 値を表す. LB(10) は 1 次から 10 次までの自己相関がすべて 0 であるという帰無仮説を検定するための Ljung and Box(1978)統計量の p 値で, Diebold (1988)の方法により分散不均一性を調整している.

	平均	標準偏差	歪度	尖度	最小値	最大値	JB	LB(10)
日次リターン	0.004 (0.033)	1.652	-0.492	10.360	-12.111	13.235	0.00	0.95
RK の対数値	-0.508 (0.019)	0.921	0.833	4.339	-2.809	3.560	0.00	0.00

でも正規分布からの拡張も考えられるが, このような拡張は今後の課題として, RK の対数値には正規分布を用いた RSV モデルを推定する.

4.2 各モデルの推定結果

まず, 日次リターンに正規分布を用いた SV モデルの各パラメータの事前分布は以下のよう

$$\mu \sim N(0, 10), \quad \frac{\phi+1}{2} \sim B(20, 1.5), \quad \sigma_{\eta}^2 \sim IG(2.5, 0.025), \quad \frac{\rho+1}{2} \sim B(1, 2).$$

ここで, $B(a, b)$ はパラメータ (a, b) のベータ分布を表す. これらのパラメータについては, RSV モデルでも同じ事前分布を用いた. また, RSV モデル固有のパラメータについては, 以下の事前分布を設定した.

$$\xi \sim N(0, 1), \quad \sigma_u^2 \sim IG(2.5, 0.1).$$

さらに, 日次リターンの誤差項の分布に用いる一般化双曲非対称 t 分布のパラメータの事前分布については, 以下のよう

$$\beta \sim N(0, 1), \quad \nu \sim G(5, 0.5)I(\nu > 4).$$

ここで, $I(\cdot)$ は指示関数であり, $\nu > 4$ は式(3.1)の σ_z^2 を有限にするため, すなわち, 日次リターンの 4 次モーメントを有限にするためである. また, t 分布を用いる場合は, $\beta = 0$ と固定した.

表 2 と表 3 には, 日次リターンの誤差項に正規分布, t 分布, 一般化双曲非対称 t 分布を用いた SV モデル(それぞれ SV-N, SV-T, SV-GH-ST とする)と RSV モデル(それぞれ RSV-N, RSV-T, RSV-GH-ST とする)の推定結果がまとめられている. 表には, 5,000 回のサンプルを burn-in として捨て, それ以降の 20,000 回のサンプルを使って計算した各パラメータの事後平均, 事後標準偏差, 95% 信用区間, CD 統計量の p 値, IF(非効率性因子)の値を掲載している. 平均と標準偏差は, それぞれ各パラメータの burn-in 以降の 20,000 個のサンプルの標本平均と標準偏差として計算した. 95% 信用区間は, サンプルを大きさの順に並べ替え, 上 2.5% と下 2.5% の値として求めた. CD 統計量は, サンプルの分布が事後分布に収束しているかどうかを検定する Geweke (1992)によって提案された収束診断(Convergence Diagnostic, CD)統計量である. IF は, 標本平均の標準誤差をランダム・サンプリングと同じにするためにランダム・サンプリングの何倍の回数サンプリングを行う必要があるかを表す非効率性因子(Inefficiency Factor, IF)である(CD 統計量と IF の具体的な計算については, 例えば大森・渡部, 2013 を参照されたい). なお, 以下の実証分析の結果については, σ_{η} や σ_u をパラメータとして結果を

表 2. SV モデルのパラメータの推定結果. サンプル期間とサイズは表 1 と同様. MCMC により 25,000 回サンプリングを行い, 最初の 5,000 回を burn-in として捨て, 残りの 20,000 回のサンプルを使って推定した. 平均と標準偏差はサンプルの標本平均と標本標準偏差として計算した. 95% 信用区間は, サンプルを大きさの順に並べ替え, 上 2.5% と下 2.5% の値として求めた. CD は, burn-in 以降の 20,000 個のサンプルの分布が事後分布に収束しているかどうかを検定する Geweke (1992) の CD (convergence diagnostic) 統計量の p 値である. IF は, 標本平均の標準誤差をランダム・サンプリングと同じにするためにランダム・サンプリングの何倍の回数サンプリングを行う必要があるかを表す非効率性因子 (inefficiency factor) である.

パラメータ	平均	標準偏差	95% 信用区間	CD	IF
SV-N					
ϕ	0.9479	0.0088	[0.9293, 0.9639]	0.105	28.40
σ_{η}	0.2754	0.0248	[0.2356, 0.3312]	0.135	214.07
ρ	-0.5284	0.0546	[-0.6222, -0.4086]	0.305	135.90
μ	0.5991	0.1045	[0.4041, 0.8150]	0.397	51.87
SV-T					
ϕ	0.9597	0.0081	[0.9420, 0.9739]	0.744	111.69
σ_{η}	0.2361	0.0251	[0.1906, 0.2910]	0.621	188.69
ρ	-0.6723	0.0515	[-0.7637, -0.5613]	0.311	77.78
μ	0.6285	0.0995	[0.4334, 0.8247]	0.051	5.77
ν	16.5857	3.7422	[11.0953, 26.2200]	0.285	232.80
SV-GH-ST					
ϕ	0.9623	0.0070	[0.9474, 0.9748]	0.518	112.62
σ_{η}	0.2404	0.0219	[0.2010, 0.2860]	0.550	202.78
ρ	-0.7528	0.0491	[-0.8450, -0.6495]	0.637	130.89
μ	0.5533	0.0983	[0.3597, 0.7476]	0.954	5.09
ν	19.1282	3.8742	[12.6679, 28.2960]	0.408	296.93
β	-0.9512	0.2794	[-1.6743, -0.5149]	0.936	226.57

表示する.

CD 統計量の p 値は, RSV-T モデルの σ_{η} と σ_u がそれぞれ 10% と 5% をわずかに下回るが, それ以外のすべてのパラメータについて 10% を上回っており, 推定に使った burn-in 以降の 20,000 回のサンプルが事後分布に収束しているという帰無仮説は, 少なくとも有意水準 1% で受容される. また, 非効率性因子の値は SV モデルよりも RSV モデルが低くなっており, SV-N モデルと RSV-N モデルに共通するパラメータ ($\mu, \phi, \sigma_{\eta}, \rho$) すべてで, SV モデルよりも RSV モデルの方が標準偏差が小さくなっている. これらの結果から, RV をデータとして加えることによりサンプリングがより効率的になると考えられる.

すべてのモデルに共通するパラメータ ($\mu, \phi, \sigma_{\eta}, \rho$) の推定結果を見ると, μ については, すべてのモデルで同様の推定値が得られている. ϕ については, SV モデルよりも RSV モデルの事後平均が若干低いものの, すべてのモデルで同様の推定値が得られている. すなわち, ϕ の事後平均と 95% 信用区間はいずれも 1 に近く, 日経 225 株価指数のボラティリティにも高い持続性があることを示している. ρ については, SV モデルよりも RSV モデルの推定値が 0 に近くになっているが, すべてのモデルについて事後平均は負で, 95% 信用区間も 0 を下回っている. この結果から, 日経 225 株価指数のボラティリティも, 株価が上がった翌日より下がった翌日の方がより上昇する傾向があることがわかる.

表 3. RSV モデルのパラメータの推定結果. サンプル期間とサイズおよび推定の方法は表 2 と同様.

パラメータ	平均	標準偏差	95% 信用区間	CD	IF
RSV-N					
ϕ	0.9337	0.0085	[0.9164, 0.9497]	0.865	15.60
σ_η	0.2800	0.0139	[0.2542, 0.3080]	0.787	36.45
ρ	-0.3204	0.0389	[-0.3985, -0.2454]	0.385	20.76
μ	0.5633	0.0893	[0.3891, 0.7407]	0.540	3.94
ξ	-1.0290	0.0362	[-1.1011, -0.9577]	0.954	19.51
σ_u	0.4127	0.0110	[0.3910, 0.4345]	0.469	25.27
RSV-T					
ϕ	0.9356	0.0082	[0.9187, 0.9511]	0.281	11.90
σ_η	0.2757	0.0135	[0.2506, 0.3039]	0.096	31.57
ρ	-0.3321	0.0399	[-0.4107, -0.2548]	0.169	18.60
μ	0.5907	0.0913	[0.4094, 0.7719]	0.370	12.44
ν	26.3082	5.8274	[17.6581, 40.6667]	0.933	142.63
ξ	-1.0533	0.0381	[-1.1293, -0.9788]	0.266	65.67
σ_u	0.4158	0.0109	[0.3942, 0.4374]	0.042	20.36
RSV-GH-ST					
ϕ	0.9379	0.0081	[0.9214, 0.9531]	0.663	14.12
σ_η	0.2701	0.0133	[0.2460, 0.2973]	0.515	25.77
ρ	-0.3572	0.0421	[-0.4397, -0.2758]	0.312	18.60
μ	0.5857	0.0897	[0.4122, 0.7637]	0.162	4.26
ν	27.2934	5.3025	[18.5698, 39.0077]	0.900	124.10
β	-0.6027	0.2975	[-1.2752, -0.0845]	0.256	33.63
ξ	-1.0550	0.0382	[-1.1318, -0.9813]	0.353	25.53
σ_u	0.4198	0.0110	[0.3984, 0.4417]	0.659	13.69

また, RSV モデルに固有のパラメータ (ξ, σ_u) の推定結果を見ると, RM のバイアスを表すパラメータ ξ の事後平均は, すべての RSV モデルでおよそ -1 と負の値で, 95% 信用区間も 0 を下回っている. RK でマイクロストラクチャー・ノイズによるバイアスがほぼ除去されているとすると, ξ の事後平均が負となっているのは, 夜間と昼休みを無視したために 1 日のボラティリティを過小評価していることによるものと考えられる. σ_η については, すべての RSV モデルで同様の推定値が得られている.

さらに, t 分布および一般化双極非対称 t 分布のパラメータ (ν, β) の推定結果を見ると, 裾の厚さを表す ν については, SV モデルよりも RSV モデルの推定値が大きくなっている. 一方, 非対称性を表す β については, SV モデルよりも RSV モデルの推定値が 0 に近くなっているが, どちらのモデルでも事後平均は負で, 95% 信用区間も 0 を下回っている. したがって, RM を加えることにより, 日次リターンの誤差項の分布について裾の厚さと非対称性の程度がどちらも小さくなっていると考えられる. この結果の解釈については, 以下の周辺尤度によるモデル比較でより詳しく述べる.

4.3 周辺尤度によるモデル比較

各モデルのデータへの適合度を比較するため, Chib (1995) の方法に従い, 対数周辺尤度とそ

表 4. SV モデルと RSV モデルの対数周辺尤度. サンプル期間とサイズは表 1 と同様. 括弧内の数値は標準誤差を表す. 尤度関数部分については, 3,000 個の粒子を用いた補助粒子フィルタによる推定を 10 回行い, その平均値と標準誤差を求めた. 事後確率密度関数部分については, Chib and Greenberg (1995) と Chib and Jeliazkov (2001) に従い, 3,000 回の追加サンプリングにより求めた.

モデル	尤度	事前密度	事後密度	対数周辺尤度
SV-N	-4252.15 (0.46)	-1.55 10.46 (0.18)	-4264.16 (0.50)	
SV-T	-4248.85 (0.80)	-4.17 7.57 (0.14)	-4260.59 (0.82)	
SV-GH-ST	-4232.52 (1.02)	-6.31 6.11 (0.37)	-4244.94 (1.09)	
RSV-N	-6113.08 (0.83)	-3.70 18.86 (0.08)	-6135.64 (0.83)	
RSV-T	-6121.38 (0.27)	-10.33 14.23 (0.12)	-6145.93 (0.30)	
RSV-GH-ST	-6118.44 (0.86)	-11.68 13.56 (0.20)	-6143.68 (0.88)	

表 5. SV モデルと RSV モデルのボラティリティ ($\exp(h_t/2)$) の推定値で基準化した日次リターンの基本統計量. サンプル期間とサイズおよび各基本統計量の計算方法は表 1 と同様.

	平均	標準偏差	歪度	尖度	最小値	最大値	JB	LB(10)
SV-N	0.014 (0.019)	0.922	-0.210	2.711	-3.168	2.619	0.00	0.40
SV-T	0.009 (0.019)	0.939	-0.291	3.082	-4.008	3.312	0.00	0.35
SV-GH-ST	0.001 (0.019)	0.941	-0.368	3.188	-4.353	3.186	0.00	0.23
RSV-N	0.038 (0.019)	0.963	-0.094	2.606	-3.068	2.534	0.00	0.26
RSV-T	0.036 (0.019)	0.953	-0.098	2.632	-3.140	2.532	0.00	0.25
RSV-GH-ST	0.034 (0.019)	0.952	-0.115	2.643	-3.228	2.478	0.00	0.25

の標準誤差を計算した結果が, 表 4 にまとめられている. 尤度関数部分については, 3,000 個の粒子を用いた補助粒子フィルタ (Pitt and Shephard, 1999; Omori et al., 2007) による推定を 10 回行い, その平均値と標準誤差を求めた. また, 事後確率密度関数部分については, Chib and Greenberg (1995) と Chib and Jeliazkov (2001) に従い, 3,000 回の追加サンプリングにより求めた. なお, 周辺尤度の計算方法には他にも Geweke (1999) の修正調和平均(modified harmonic mean)がよく用いられるが, この方法では, 完全な尤度ではなく潜在変数を条件とする条件付き尤度を用いると, 周辺尤度の推定値にバイアスが生じることを Chan and Grant (2015) が示しているのので, 注意が必要である.

表 4 の対数周辺尤度によると, SV モデルでは一般化双曲非対称 t 分布, RSV モデルでは正規分布が選択される. 表 2 と表 3 を比べると, 一般化双曲非対称 t 分布では, 裾の厚さを表す ν の推定値は RSV モデルの方が高く, 分布の非対称性を表す β の推定値は SV, RSV とともに負の値になっているが, RSV の方がより 0 に近い値になっており, 周辺尤度の結果と整合的である. これは, リターンが極端な値になった時に RK が大きな値になるので, RSV モデルでは真のボラティリティも大きくなり, 真のボラティリティで割って基準化したリターンの分布の裾が薄くなるものと解釈できる. また, 特にリターンが極端な負の値になった時に RK が大きな値になるので, RSV モデルでは基準化したリターンの分布の非対称性が小さくなるものと解釈できる.

表 5 は, 各モデルの真のボラティリティ $\exp(h_t/2)$ の事後平均で基準化した日次リターンの基本統計量であり, 図 4 はそれらのヒストグラムである. ヒストグラムからは, すべてのモデルについて正規分布との乖離に差はないように見える. 基本統計量を見ると, すべての SV モ

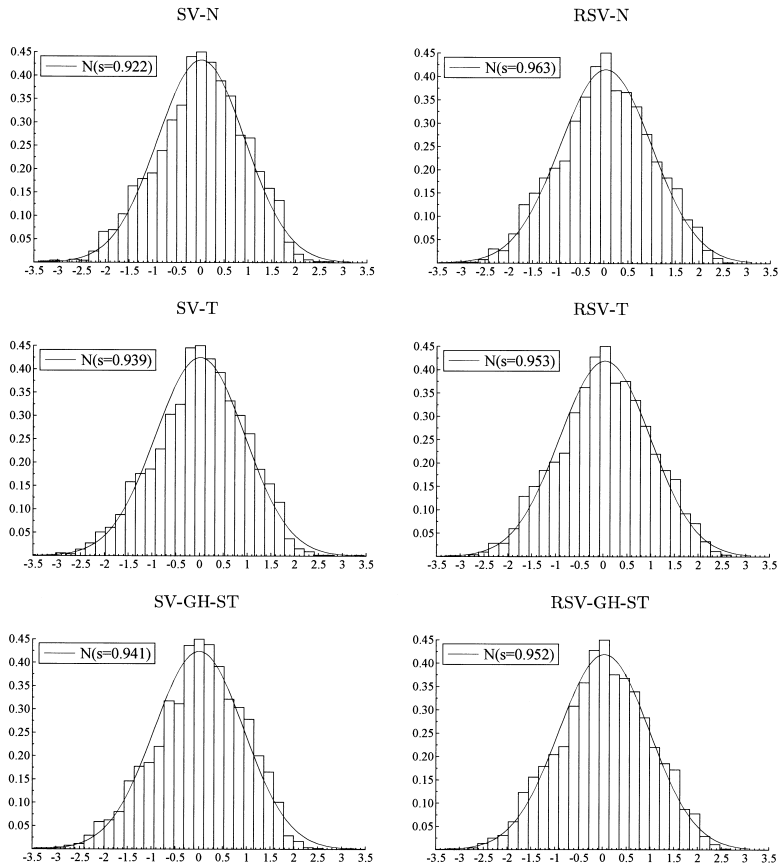


図 4. SV モデルと RSV モデルのボラティリティ ($\exp(h/2)$) の推定値で標準化した日次リターンのヒストグラム。実線は、それぞれの平均値と標準偏差に対応した正規分布の確率密度関数。横軸の範囲は $[-3.5, 3.5]$ ，縦軸の範囲は $[0, 0.45]$ 。

デルについて歪度は有意に負で、SV-GH-ST モデルについては尖度も有意に 3 より高くなっている。JB の値から標準化したリターンが正規分布に従うという帰無仮説は有意水準 1% で棄却される。したがって、誤差項には一般化双曲非対称 t 分布が適していることがわかる。一方、RSV モデルについては、正規分布を用いた場合に、歪度は最も 0 に近く有意に乖離していない。また、すべての RSV モデルについて尖度は有意に 3 より小さくなっている。JB の値からは標準化したリターンが正規分布に従うという帰無仮説は有意水準 1% で棄却されるが、SV モデルと比べると標準化したリターンの歪度は 0 に近く、正規分布に近い。これらの結果は、上記の周辺尤度の結果と整合的である。ただし、これは本稿で用いたデータでの結果であり、一般的に RSV モデルにすると誤差項の分布は正規分布で良いということではないので、注意されたい。

5. おわりに

本稿では、RSV モデルおよびその拡張モデルとそれらの MCMC を用いたベイズ推定につい

て解説した。また、日経 225 株価指数の日次リターンと RK を用いて、リターンの誤差項の分布を一般化双極非対称 t 分布に拡張した SV モデルと RSV モデルを推定し、周辺尤度によるモデル比較を行った。その結果、SV モデルについては、リターンの裾の厚さと非対称性を捉えるために一般化双極非対称 t 分布が有効である一方、RSV モデルについては分布の拡張は必ずしも必要ではないということが明らかになった。ただし、これは本稿で用いたデータでの結果であり、一般的に RSV モデルにすると誤差項の分布は正規分布で良いということではない。

RSV モデルは、本稿で解説したモデル以外にも様々な拡張が提案されている。例えば、日次リターン分布の裾の厚さと非対称性を捉えるために、一般化双極非対称 t 分布の他にも Azzalini and Capitanio (2003) や Fernández and Steel (1998) が提案した非対称 t 分布を用いることも考えられる。また、本稿では 1 変量モデルを考えたが、RSV モデルの多変量モデルへの拡張も重要であり、盛んに研究が行われている (Shirota et al., 2017; Kurose and Omori, 2020; Yamauchi and Omori, 2019)。RSV モデルにはまだまだ改良の余地があり、上記の拡張も含めて今後も発展が期待される。

謝 辞

本稿は科学研究費補助金基盤研究 (A) 17H00985, 19H00588 と一橋大学社会科学高等研究院を通じて文部科学省から助成を受けている。本稿の作成にあたり、編集委員と匿名の査読者から貴重なコメントを頂いた。また、本稿で用いた日経 225 株価指数の RK のデータは、Ubukata and Watanabe (2014) に基づいて生方雅人氏 (明治学院大学) に計算して頂いた。ここに記して深く感謝の意を表したい。

参 考 文 献

- Aas, K. and Haff, I. H. (2006). The generalized hyperbolic skew Student's t -distribution, *Journal of Financial Econometrics*, **4**(2), 275–309.
- Abanto-Valle, C. A., Lachos, V. H. and Dey, D. K. (2015). Bayesian estimation of a skew-Student- t stochastic volatility model, *Methodology and Computing in Applied Probability*, **17**(3), 721–738.
- Aït-Sahalia, Y. and Mykland, P. A. (2009). Estimating volatility in the presence of market microstructure noise: A review of the theory and practical considerations, *Handbook of Financial Time Series* (eds. T. G. Andersen, R. A. Davis, J.-P. Kreiß and T. Mikosch), 577–598, Springer-Verlag, Berlin.
- Aït-Sahalia, Y., Mykland, P. A. and Zhang, L. (2005). How often to sample a continuous-time process in the presence of market microstructure noise, *Review of Financial Studies*, **18**(2), 351–416.
- Andersen, T. G. and Benzoni, L. (2009). Realized volatility, *Handbook of Financial Time Series* (eds. T. G. Andersen, R. A. Davis, J.-P. Kreiß and T. Mikosch), 555–575, Springer-Verlag, Berlin.
- Andersen, T. G. and Bollerslev, T. (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts, *International Economic Review*, **39**(4), 885–905.
- Andersen, T. G., Bollerslev, T., Diebold, F. X. and Ebens, H. (2001a). The distribution of realized stock return volatility, *Journal of Financial Economics*, **61**(1), 43–76.
- Andersen, T. G., Bollerslev, T., Diebold, F. X. and Labys, P. (2001b). The distribution of realized exchange rate volatility, *Journal of the American Statistical Association*, **96**(453), 42–55.
- Andersen, T. G., Bollerslev, T., Diebold, F. X. and Labys, P. (2003). Modeling and forecasting realized volatility, *Econometrica*, **71**(2), 579–625.
- Andersen, T. G., Bollerslev, T. and Diebold, F. X. (2010). Parametric and nonparametric volatility measurement, *Handbook of Financial Econometrics* (eds. Y. Aït-Sahalia and L. P. Hansen), Chapter 2, 67–138, North Holland, Amsterdam.

- Andersen, T. G., Bollerslev, T., Christoffersen, P. F. and Diebold, F. X. (2013). Financial risk measurement for financial risk management, *Handbook of the Economics of Finance* (eds. G. M. Constantinides, M. Harris and R. M. Stulz), Vol.2, Part B, Chapter 17, 1127–1220, North Holland, Amsterdam.
- Azzalini, A. and Capitanio, A. (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution, *Journal of the Royal Statistical Society B*, **65**(2), 367–389.
- Bandi, F. M. and Russell, J. R. (2006). Separating microstructure noise from volatility, *Journal of Financial Economics*, **79**(3), 367–389.
- Bandi, F. M. and Russell, J. R. (2008). Microstructure noise, realized volatility, and optimal sampling, *Review of Economic Studies*, **75**(2), 655–692.
- Barndorff-Nielsen, O. E. (1977). Exponentially decreasing distributions for the logarithm of particle size, *Proceedings of the Royal Society A*, **1674**(353), 401–419.
- Barndorff-Nielsen, O. E. and Shephard, N. (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models, *Journal of Royal Statistical Society B*, **64**(2), 253–280.
- Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A. and Shephard, N. (2008). Realized kernels in practice: Trades and quotes, *Econometrica*, **76**(6), 1482–1536.
- Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A. and Shephard, N. (2009). Designing realized kernels to measure the ex post variation of equity prices in the presence of noise, *Econometrics Journal*, **12**(3), C1–C32.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics*, **31**(3), 307–327.
- Breidt, F. J., Crato, N. and Lima, P. D. (1998). The detection and estimation of long memory in stochastic volatility, *Journal of Econometrics*, **83**(1-2), 325–348.
- Campbell, J. Y., Lo, A. W. and MacKinlay, A. C. (1997). *The Econometrics of Financial Markets*, Princeton University Press, Princeton.
- Chan, J. C. C. and Grant, A. L. (2015). Pitfalls of estimating the marginal likelihood using the modified harmonic mean, *Economics Letters*, **131**, 29–33.
- Chan, N. H. and Palma, W. (1998). State space modeling of long-memory processes, *Annals of Statistics*, **26**(2), 719–740.
- Chib, S. (1995). Marginal likelihood from the Gibbs output, *Journal of the American Statistical Association*, **90**(432), 1313–1321.
- Chib, S. and Greenberg, E. (1995). Understanding the Metropolis-Hastings algorithm, *American Statistician*, **49**(4), 327–335.
- Chib, S. and Jeliazkov, I. (2001). Marginal likelihood from the Metropolis-Hastings output, *Journal of the American Statistical Association*, **96**(453), 270–281.
- Diebold, F. X. (1988). *Empirical Modeling of Exchange Rate Dynamics*, Springer-Verlag, Berlin.
- Dobrev, D. P. and Szerszen, P. J. (2010). The information content of high-frequency data for estimating equity return models and forecasting risk, FEDS Working Papers, No.2010-45.
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, *Econometrica*, **50**(4), 987–1008.
- Fernández, C. and Steel, M. F. J. (1998). On Bayesian modeling of fat tails and skewness, *Journal of the American Statistical Association*, **93**(441), 359–371.
- Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments, *Bayesian Statistics 4* (eds. J. M. Bernardo, J. O. Berger, A. Dawid and A. Smith), 169–193, Oxford University Press, Oxford.
- Geweke, J. (1999). Using simulation methods for Bayesian econometric models: Inference, development, and communication, *Econometric Reviews*, **18**(1), 1–73.
- Giot, P. and Laurent, S. (2004). Modeling daily value-at-risk using realized volatility and ARCH type

- models, *Journal of Empirical Finance*, **11**(3), 379–398.
- Hansen, P. R. and Huang, Z. (2016). Exponential GARCH modeling with realized measures of volatility, *Journal of Business & Economic Statistics*, **34**(2), 269–287.
- Hansen, P. R. and Lunde, A. (2005). A forecast comparison of volatility models: Does anything beat a GARCH(1,1)?, *Journal of Applied Econometrics*, **20**(7), 873–889.
- Hansen, P. R. and Lunde, A. (2006). Realized variance and market microstructure noise, *Journal of Business & Economic Statistics*, **24**(2), 127–161.
- Hansen, P. R., Huang, Z. and Shek, H. (2012). Realized GARCH: A joint model of returns and realized measures of volatility, *Journal of Applied Econometrics*, **27**(6), 877–906.
- Hasbrouck, J. (2007). *Empirical Market Microstructure*, Oxford University Press, New York.
- 石原庸博 (2015). 一般化した Realized Stochastic Volatility モデルの推定—日経 225 収益率・実現ボラティリティの暦効果への応用—, *経済研究*, **66**(1), 1–18.
- Jarque, C. M. and Bera, A. K. (1987). A test for normality of observations and regression residuals, *International Statistical Review*, **55**(2), 163–172.
- Koopman, S. J. and Scharth, M. (2013). The analysis of stochastic volatility in the presence of daily realized measures, *Journal of Financial Econometrics*, **11**(1), 76–115.
- Koopman, S. J., Jungbacker, B. and Hol, E. (2005). Forecasting daily variability of the S&P 100 stock index using historical, realised and implied volatility measurements, *Journal of Empirical Finance*, **12**(3), 445–475.
- Kurose, Y. and Omori, Y. (2020). Multiple-block dynamic equicorrelations with realized measures, leverage and endogeneity, *Econometrics and Statistics*, **13**, 46–68.
- Ljung, G. M. and Box, G. E. P. (1978). On a measure of lack of fit in time series analysis, *Biometrika*, **65**(2), 297–303.
- McAleer, M. and Medeiros, M. C. (2008). Realized volatility: A review, *Econometric Reviews*, **27**(1–3), 10–45.
- 中島上智, 大森裕浩 (2011). 一般化双曲非対称 t 分布を用いた確率的ボラティリティ変動モデルの推定と株価収益率データへの応用, *日本統計学会誌*, **40**(2), 61–88.
- Nakajima, J. and Omori, Y. (2012). Stochastic volatility model with leverage and asymmetrically heavy-tailed error using GH skew Student's t -distribution, *Computational Statistics & Data Analysis*, **56**(11), 3690–3704.
- Nishino, H. and Kakamu, K. (2013). Bayesian Whittle estimation of ARFIMA model, *Advances and Applications in Statistics*, **37**(2), 149–170.
- Nugroho, D. B. and Morimoto, T. (2014). Realized non-linear stochastic volatility models with asymmetric effects and generalized Student's t -distributions, *Journal of Japan Statistical Society*, **44**(1), 83–118.
- Nugroho, D. B. and Morimoto, T. (2016). Box-Cox realized asymmetric stochastic volatility models with generalized Student's t -error distributions, *Journal of Applied Statistics*, **43**(10), 1906–1927.
- O'Hara, M. (1995). *Market Microstructure Theory*, Blackwell, Oxford.
- O'Hara, M. (2015). High frequency market microstructure, *Journal of Financial Economics*, **116**(2), 257–270.
- 大森裕浩, 渡部敏明 (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推定—, *日本統計学会誌*, **42**(2), 273–303.
- Omori, Y., Chib, S., Shephard, N. and Nakajima, J. (2007). Stochastic volatility with leverage: Fast likelihood inference, *Journal of Econometrics*, **140**(2), 425–229.
- Pitt, M. K. and Shephard, N. (1999). Filtering via simulation: Auxiliary particle filters, *Journal of the American Statistical Association*, **94**(446), 590–599.
- Prause, K. (1999). The generalized hyperbolic model: Estimation, financial derivatives and risk measures, PhD Dissertation, University of Freiburg, Freiburg.
- Ruiz, E. and Veiga, H. (2008). Modeling long-memory volatilities with leverage effect: A-LMSV versus

- FIEGARCH, *Computational Statistics & Data Analysis*, **52**(6), 2846–2862.
- Shephard, N. (1996). Statistical aspects of ARCH and stochastic volatility, *Time Series Models in Econometrics, Finance and Other Fields* (eds. D. R. Cox, D. V. Hinkley and O. E. Barndorff-Nielsen), 1–67, Chapman & Hall, New York.
- Shirota, S., Hizu, T. and Omori, Y. (2014). Realized stochastic volatility with leverage and long memory, *Computational Statistics & Data Analysis*, **76**(C), 618–641.
- Shirota, S., Omori, Y., Lopes, H. F. and Piao, H. (2017). Cholesky realized stochastic volatility model, *Econometrics and Statistics*, **3**(1), 34–59.
- So, M. K. P. (2002). Bayesian analysis of long memory stochastic volatility models, *Sankhyā: The Indian Journal of Statistics, Series B*, **64**(1), 1–10.
- 高橋 慎 (2014). Smooth transition Realized Stochastic Volatility モデル, 日本統計学会誌, **44**(1), 41–60.
- Takahashi, M., Omori, Y. and Watanabe, T. (2009). Estimating stochastic volatility models using daily returns and realized volatility simultaneously, *Computational Statistics & Data Analysis*, **53**(6), 2404–2426.
- Takahashi, M., Watanabe, T. and Omori, Y. (2016). Volatility and quantile forecasts by Realized Stochastic Volatility models with generalized hyperbolic distribution, *International Journal of Forecasting*, **32**(2), 437–457.
- Taylor, S. J. (1986). *Modelling Financial Time Series*, Wiley, New York.
- Ubukata, M. and Watanabe, T. (2014). Pricing Nikkei 225 options using Realized Volatility, *Japanese Economic Review*, **65**(4), 431–467.
- 渡部敏明 (2000). 『ボラティリティ変動モデル』, 朝倉書店, 東京.
- 渡部敏明 (2005). マルチ・ムーブ・サンプラーを用いた確率的ボラティリティ変動モデルのベイズ推定, 『ベイズ計量経済分析 マルコフ連鎖モンテカルロ法とその応用』(和合肇 編), 第 9 章, 259–294, 東洋経済新報社, 東京.
- 渡部敏明 (2007). Realized Volatility—サーベイと日本の株式市場への応用, 経済研究, **58**(4), 352–373.
- 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, 金融研究, **25**(2), 39–74.
- Yamauchi, Y. and Omori, Y. (2019). Multivariate stochastic volatility model with realized volatilities and pairwise realized correlations, *Journal of Business and Economic Statistics* (in press).
- Zhang, L. (2006). Efficient estimation of stochastic volatility using noisy observations: A multi-scale approach, *Bernoulli*, **12**(6), 1019–1043.
- Zhang, L., Mykland, P. A. and Ait-Sahalia, Y. (2005). A tale of two time scales: Determining integrated volatility with noisy high-frequency data, *Journal of the American Statistical Association*, **100**(472), 1394–1411.
- Zheng, T. and Song, T. (2014). A Realized Stochastic Volatility model with Box-Cox transformation, *Journal of Business and Economic Statistics*, **32**(4), 593–605.

Realized Stochastic Volatility Model —Extensions and Application to Japanese Stock Index—

Makoto Takahashi¹, Yasuhiro Omori² and Toshiaki Watanabe³

¹Faculty of Business Administration, Hosei University

²Graduate School of Economics, University of Tokyo

³Institute of Economic Research, Hitotsubashi University

Realized volatility (RV), which is the sum of squared intraday returns over a certain interval (such as a day), has widely been used as an estimator of the financial volatility. In the real market, however, the presence of non-trading hours and market microstructure noise in transaction prices may create bias in the RV. Taking account of this bias, several studies propose modeling daily returns and RV simultaneously. The resultant model, based on the stochastic volatility (SV) model, is called the realized stochastic volatility (RSV) model. Because the likelihood of the RSV model, as well as the SV model, is difficult to evaluate analytically, the Bayesian estimation method via the Markov chain Monte Carlo (MCMC) is often used. In this article, we explain the RSV model, its Bayesian estimation method via MCMC, and several extensions of the RSV model. Further, we apply the models to the Nikkei 225 stock index and explain the estimation results.

確率的依存構造をもつコピュラモデル

—統計的推定方法と計量ファイナンスへの応用—

野澤 勇樹¹・中村 信弘²

(受付 2019 年 5 月 31 日；改訂 2020 年 2 月 5 日；採択 2 月 6 日)

要 旨

コピュラ関数の依存構造が確率的に変動するような確率的コピュラのモデル構築とその代表的な統計的推定方法をサーベイする。依存構造の確率的な変動の記述には潜在変数を内包する状態方程式を用いることから、数値計算による尤度評価が必要となる。本論文ではこれらの手法についてまとめるとともに、確率的コピュラのヴァインコピュラを通じた多次元化への応用について拡張する。また、ファイナンス分野への応用事例として、時変レバレッジを持つコピュラと時変の依存構造パラメータを持つコピュラのモデルを紹介する。時変の依存構造パラメータを持つコピュラのモデルを為替ヘッジへの適用した例について報告する。

キーワード：確率的依存構造、確率的コピュラ、ヴァインコピュラ。

1. 序論

コピュラモデルは依存関係を記述するのに簡便で柔軟性のある優れた統計的手法である。従来、依存関係を記述するコピュラ関数は静的なものかまたは動的であっても確率的要素を含まない決定論的なものが用いられてきた。しかし、標本期間をローリングしながら長い期間で分析してみると、依存関係が確率的に時変する金融時系列を多く目にする。実際、代表的金融資産である株式と債券の相関は概ね負の値をとるが、経時的にその大きさは変化している。そのため、長期投資では、資産配分比率を適切に変えていく必要がある。依存関係に定常で動的時変構造がある場合はそれを利用したリスク管理、リスクヘッジが可能であり、より最適な資産配分比率を計算することに利用できる可能性がある。本稿では、確率的コピュラのファイナンスへの応用を紹介する。

確率ボラティリティ(SV)モデルは、リスク資産のリターンの変動が確率的に変動する現象をモデル化するために、よく用いられる。コピュラモデルの周辺モデルにも用いられることが多い。SVモデルの実証研究によると、リターンとボラティリティの攪乱項どうしの相関¹⁾は、大抵の場合、負の値で推定される。リスク資産の分散変動に対して投資家が要求するリスクプレミアム(分散リスクプレミアム, VRP)とリスク資産の実現リターンの間には密接な関係があり、Bollerslev et al. (2009, 2014)はVRPによるリターンの予測可能性に焦点を当てて研究している。国際間比較の実証分析の結果によると、S&P500のVRPによるリター

¹ ステート・ストリート・グローバル・アドバイザーズ株式会社：〒105-6325 東京都港区虎ノ門1-23-1 虎ノ門ヒルズ森タワー 25 階

² 一橋大学大学院 経営管理研究科：〒101-8439 東京都千代田区一ツ橋 2-1-2 学術総合センター

ン予測が最も高くなり、日経平均では最も低くなることが報告されている。SVモデルを用いたNakamura (2017)の研究では、レバレッジ係数がVRPによるリターンの予測可能性に重要な役割を演じていることを明らかにした。SVモデルで推定したレバレッジ係数の強い(負で絶対値が大きい)リスク資産ほど、VRPによるリターンの予測可能性が高くなることを理論的に示すことができる。リターンとボラティリティの依存構造に確率的コピュラモデルを適用すると、レバレッジ係数の確率的変動を捉えることができるようになり、あるリスク資産で、いつVRPによるリターンの予測力が高くなるのか分析できる可能性がある。

Frazzini and Pedersen (2014)は、株式市場で個別株の市場インデックスに対するCAPMベータを計測し、期待リターンとの関係を分析したところ、低ベータの資産ほど実現リターンが高くなるという現象を報告している。通常ファイナンスの資産価格理論は、投資家がとったリスクに見合った対価である期待リターンが得られることを説くが、Frazzini and Pedersen (2014)で報告された現象はその逆であり、ベータ・アノマリーと呼ばれている。現在、色々な解釈が試みられているが、その中の一つに、ベータ自身の変動に対して投資家が要求するリスクプレミアムの時系列構造から現象の解明を試みる研究(Boloorforoosh et al., 2020)もある。CAPMでは個別株ベータ $\beta_i = \rho_{i,m}\sigma_i/\sigma_m$ ($\rho_{i,m}$ は市場インデックスと個別株の相関; σ_i, σ_m は個別株と市場インデックスのそれぞれのボラティリティ)で表されるが、Asness et al. (2020)では、 $\rho_{i,m}$ が小さい程、実現リターンが高くなる傾向にあると報告している。個別株、市場インデックス各々の周辺モデルに、例えば、SV(確率ボラティリティ)モデル、それらの依存関係の確率変動を確率的コピュラで記述し、ベータの変動リスクプレミアムを分析したり、潜在変数のダイナミクスを推定して得られる情報が、リターンに如何に織り込まれるか分析することで、この現象にアプローチできると考えられる。

リスク管理の分野では、テール・リスクのモデルが盛んに研究されている。最近、Adrian and Brunnermeier (2016)で、金融機関のシステミックを分析する概念としてCoVaRという量が提唱されている。個々の金融機関の損失率 l_i が VaR_p (信頼水準 p) レベルにあるときに、その事象が金融機関全体に与える最大損失率(信頼水準 q)として定義される。この量を測定するためには、個々の金融機関の損失率と金融システム全体の損失率の間の相互依存構造をモデル化する必要がある。監物 (2017)では、2変量確率的コピュラを用いて、これらの量の動的変化を考慮したリスク尺度を研究している。

アセット・アロケーションでは、多変量の資産クラスの依存構造を記述する必要があるが、その際、ペアコピュラの組み合わせで構成されるヴァイン(vine)コピュラは、柔軟な依存構造をモデル化することが可能である。多変量コピュラからサンプリングできれば、各周辺モデルを通じて各資産の将来リターンをモンテ・カルロ法で生成することができるため、例えば、CVaRを最小化する確率計画法(Rockafellar and Uryasev, 2000)により、下方リスクを抑制したポートフォリオを作ることができる。最近、流行している投資手法に、リスクパリティ投資(Roncalli, 2013)がある。これはポートフォリオの全リスクを各資産で等しく分担するように設計された投資手法で、最適投資比率は共分散行列だけで決定される。これを拡張してポートフォリオの全下方リスクを各資産で等しく分担するように設計することも可能で、テール・リスクパリティ(TRP)投資と呼ばれている。この場合は共分散行列に加えて、各資産の期待リターンの推定値が必要となる。Boudt et al. (2013)は、TRPの他にテール・リスクに関するポートフォリオ最適化の幾つかのバリエーションを研究している。これらの投資手法はいずれも資産間の依存関係を線形相関で捉えているが、それを確率的コピュラ関数で表現すると、より多様で時変性のある依存構造を記述できる拡張の余地がある。

現物と先物の2変量で、現物の価格変動をヘッジする際に、先物に関するベータが利用される。また、外貨建て資産の為替レートのヘッジに関しても同様に、外貨建て資産に関するベ-

タが利用される。このベータを上述のようなモデルを用いて推定することで動的ベータをヘッジ比率に利用できる。確率のコピュラを通じた外貨建て資産に関するヘッジ効率に関しては、Nozawa and Nakamura (2015)で分析されている。

本稿の構成は以下の通り。第2節では、確率のコピュラの定義、第3節では確率的ヴァインコピュラについて記述する。第4節では、確率のコピュラの代表的な統計的推定方法を解説する。第5節では、確率のコピュラのファイナンス分野における様々の応用事例を紹介し、最後の第6節は、結論と今後の課題にあてられる。

2. 確率のコピュラ

2.1 確率のコピュラの定義

Sklar の定理により、周辺分布関数 $(F_i)_{i=1:n}$ をもつ連続 n 変量分布関数 F に対して

$$(2.1) \quad \Pr(X_1 \leq x_1, \dots, X_n \leq x_n) = F(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n))$$

を満たす関数 C が一意に存在する。このコピュラ関数と呼ばれる関数 C は、 $F(x_i) = u_i \in [0, 1]$ とすると、 $\mathbf{u} = (u_i)_{i=1:n}$ を定義域として $C(u_1, \dots, u_n) = F(F_1^{-1}(u_1), \dots, F_n^{-1}(u_n))$ と表される。この同時分布関数から確率密度関数は $c(u_1, \dots, u_n) = \partial^n C(u_1, \dots, u_n) / \partial u_1 \cdots \partial u_n$ で与えられ、 X_i の確率密度関数を f_i 、同時分布関数 F の確率密度関数を f とすると、これらは次の関係式に従う。

$$(2.2) \quad f(x_1, \dots, x_n) = c(F_1(x_1), \dots, F_n(x_n)) \prod_{i=1}^n f_i(x_i)$$

コピュラ関数は各変数の依存関係を表現しており、その依存関係を特徴付けるパラメータを Λ とすると、確率のコピュラ関数はその量が確率的に変動する関数 Λ_t として定義されたものである。

3. 確率的ヴァイン(vine)コピュラ

多変量の確率のコピュラとして、ヴァイン(vine)コピュラ型のもの(Aas et al., 2009)を取り上げる。

ヴァインコピュラとは、多変数分布関数をペアコピュラ関数のみの積に分解して表現するモデルである。最も基本的なヴァインコピュラは、Bedford and Cooke (2002)によって提案された正則ヴァイン(R-vine)である。 n 個の確率変数 $\{X_1, X_2, \dots, X_n\}$ の正則ヴァインとは、 $(n-1)$ 個の木(tree; 連結非巡回グラフ) $\{T_1, \dots, T_{n-1}\}$ の集合であり、次のような入れ子構造条件を満たすものである。

- (1) $T_1 = (N_1, E_1)$ は頂点(node) 集合 $N_1 = \{1, \dots, n\}$ とこれらの頂点を繋ぐ辺(edge) 集合 E_1 からなる連結された木である。
- (2) $T_i (i = 2, \dots, n-1)$ は、頂点集合 $N_i = E_{i-1}$ 、辺集合 E_i とする木である。すなわち、 $T_i = (E_{i-1}, E_i)$ 。
- (3) T_{i-1} の2つの辺で、それらが共通の頂点を共有するときのみ、次の木 T_i で連結される。

最後の条件は近接(proximity)条件と言われる。

R-ヴァインの木 T_i の各辺 $(e \in E_i)$ の端点を $j(e), k(e)$ とし、条件付き変数の集合を $D(e)$ とするに、ペアコピュラの確率密度関数 $c_{j(e), k(e) | D(e)}$ を割り当て、周辺確率密度関数を $f_k (k = 1, \dots, n)$ とすると、R-ヴァインで表現された同時確率密度は

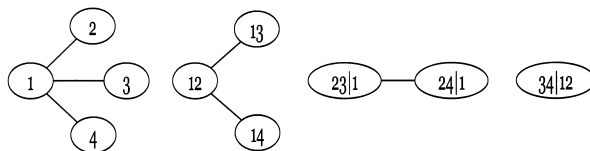


図 1. C (Canonical) ヴァインコピュラの連結構造.

$$(3.1) \quad \prod_{k=1}^n f(x_k) \prod_{j=1}^{n-1} \prod_{e \in E_j} c_{j(e), k(e)|D(e)}(F(x_{j(e)}|\mathbf{x}_{D(e)}), F(x_{k(e)}|\mathbf{x}_{D(e)}))$$

となる.

R-ヴァインに含まれる特別なヴァインとして, C-ヴァイン (canonical vine; 図 1) がある. これは各 T_i で, $n-i$ 本の辺で結ばれている頂点が 1 つだけあるものとして定義される. C-ヴァインの各木の辺の総数は $n(n-1)/2$ であるため, C-ヴァインで表現された同時確率密度は

$$(3.2) \quad \prod_{k=1}^n f(x_k) \prod_{j=1}^{n-1} \prod_{i=1}^{n-j} c_{j, i+j|1, \dots, j-1}(F(x_j|x_1, \dots, x_{j-1}), F(x_{i+j}|x_1, \dots, x_{j-1}))$$

となる.

n 変数の R-ヴァインで, 木の構造を決めるにあたり, Nápoles et al. (2010) では $\binom{n}{2} \times (n-2)! \times 2^{\binom{n-2}{2}}$ 通りのパターンから次のようにして適切なグラフを選択する方法が考えられている. 木を構成する各辺に, 対応するペア (X_i, X_j) の依存尺度 δ_{ij} (裾依存係数や Kendall の τ など) を割り付け, 全域木 (spanning tree; すべての頂点を閉路なしで結ぶグラフ) \mathcal{E} の中から, 依存尺度の絶対値の合計が最大になるような木を選択する.

$$\max \sum_{e_{ij} \in E, \mathcal{E} = \{E\}} |\delta_{ij}|.$$

Heinen and Valdesogo (2009), Brechmann and Czado (2011) では, C-ヴァインコピュラを用いた非正規・非線形 CAPM の拡張 (canonical vine autoregressive, CAVA) の研究を行っている. 階層構造をもつ個別銘柄, 業種セクター, 市場インデックスの依存関係は, C-ヴァインコピュラで適切に表現することができる. 岩永 (2015) では, 日本株市場における CAVA モデルの実証分析を行っている.

4. 確率的コピュラの統計的推定方法

4.1 フィルタリング, ベイズ推定による確率的コピュラの推定

推定方法に関して, 潜在変数を含む時系列モデルとなるため, フィルタリングやベイズ推定が用いられる. フィルタリングの場合は, 非線形モデルとなることが多いので, Liu and West (2001) の粒子フィルターや, 効率的重点サンプリング法 (Efficient Importance Sampling, EIS) (Liesenfeld and Richard, 2003, 2006; Richard and Zhang, 2007; Hafner and Manner, 2012) (次節で詳述), 尤度を粒子フィルターで計算し, それを最大化するような方法などがとられる. ベイズ推定の場合は, 次のような事後確率密度からパラメータを 1 つずつサンプリングしていく従来型の MCMC に基づくもの Almeida and Czado (2012) がある. 時刻 t における周辺分布を通じて観測データを一様化したコピュラベースの n_u 次元の観測データを $\mathbf{u}_t \in [0, 1]^{n_u}$ とし, $\mathbf{u}_{1:T} := \{\mathbf{u}_t\}_{t=1}^T$ と表現する. また, 依存構造パラメータを駆動する n_x 次元の潜在変数を \mathbf{x}_t と

し, $\mathbf{x}_{1:T} := \{\mathbf{x}_t\}_{t=1}^T$ と表現する. このとき, $\mathbf{u}_{1:T}$ と $\mathbf{x}_{1:T}$ の同時密度関数は以下のように記述できる.

$$f(\mathbf{u}_{1:T}, \mathbf{x}_{1:T} | \boldsymbol{\theta}) = \prod_{t=1}^T f(\mathbf{u}_t | \mathbf{x}_t) \cdot f(\mathbf{x}_1 | \boldsymbol{\theta}) \prod_{t=2}^T f(\mathbf{x}_t | \mathbf{x}_{t-1}, \boldsymbol{\theta}).$$

ここで, $f(\mathbf{u}_t | \mathbf{x}_t)$ を推定したい確率的コピュラの確率密度関数 $c(\mathbf{u}_t | \mathbf{x}_t)$ とする. $f(\mathbf{x}_t | \mathbf{x}_{t-1}, \boldsymbol{\theta})$ は, 潜在変数の推移確率である. また, 最近, 普及し始めている HMC (Hamiltonian Monte-Carlo) 法 (Neal, 2011) などでも推定が可能である. HMC 法は, 潜在変数の値を粒子の座標と考え, 潜在変数の十分なミキシングが起こるように, 仮想的な運動エネルギー $K(p)$ を与え²⁾, 潜在変数の事後分布 $f(\boldsymbol{\theta} | \mathbf{y})$ の $-\log(f(\boldsymbol{\theta} | \mathbf{y}))$ をポテンシャルエネルギー $U(\theta)$ と見做し, その和 $H(\theta, p) := K(p) + U(\theta)$ (Hamiltonian) から導かれる粒子の運動方程式 (Hamilton-Jacobi 正準方程式) をリープ・フロッグ (leap-frog) 法により数値的に解くことで, 粒子の座標値と運動量をサンプリングする. Metropolis-Hastings (MH) 法の一種であるが³⁾, 位相空間 (θ, p) 内で, サンプリングにより移動した点が採択される受容率は, エネルギー保存則により高くなる. 多変数の場合, 多次元位相空間内での遷移となり, リープ・フロッグ法の離散近似の精度, 位相空間内の移動距離などをうまく選ぶ必要があり, この HMC 法は, 事後分布の高い場所を行き過ぎ, 無駄な U-ターンが起こることを抑制するような Hoffman and Gelman (2014) によるアルゴリズム (NUTS (No-U-Turn Sampler) 法) と組み合わせてソフトウェア Stan (Stan Development Team, 2020) に実装されている. 本稿では, HMC 法を用いて後述するファイナンスの応用事例の推定を行う.

HMC でのモデル選択では, Watanabe (2010) による AIC のベイズ推定版の一つである WAIC (widely applicable information criterion) などが用いられる. 新しい標本データ $\{\tilde{y}_i\}_{i=1}^n$ に対するモデルの適合度は, 各データ点での予測確率密度の事後平均の対数の和 $\sum_{i=1}^n \log \mathbb{E}_{\text{post}}[f(\tilde{y}_i | \boldsymbol{\theta})]$ を計算して求められる. ここで, $\boldsymbol{\theta}$ は推定されたモデル・パラメータや潜在変数であり, その事後分布 $p_{\text{post}}(\boldsymbol{\theta}) = p(\boldsymbol{\theta} | \mathbf{y}_{1:T})$ (ここで, $\mathbf{y}_{1:T} = \{y_t\}_{t=1}^T$ はモデル推定に使った標本データ) からサンプリングした MCMC 標本 $\{\boldsymbol{\theta}_s\}_{s=1}^S$ を用いて, 予測確率密度の事後平均 $\mathbb{E}_{\text{post}}[f(\tilde{y}_i | \boldsymbol{\theta})]$ は $(1/S) \sum_{s=1}^S f(\tilde{y}_i | \boldsymbol{\theta}_s)$ と近似される.

情報量規準の計算では, モデルの推定に用いたデータを再び利用してモデルの適合度を計算するため, オーバーフィッティングによる過大推定バイアスの補正が必要となる. WAIC の場合, その補正には, 実効的パラメータ数に関する各データの対数確率密度の事後分散の和, $p_{\text{WAIC}} = \sum_{t=1}^T \mathbb{V}_{\text{post}}[\log f(y_t | \boldsymbol{\theta})]$ が使われる. ここで, $\mathbb{V}_{\text{post}}[\log f(y_t | \boldsymbol{\theta})]$ は, 各データ点で MCMC 標本 $\{\boldsymbol{\theta}_s\}_{s=1}^S$ を用いて計算される標本分散 $V_{s=1}^S[\log f(y_t | \boldsymbol{\theta}_s)]$ (ここで, $V_{s=1}^S[x_s] = \frac{1}{S-1} \sum_{s=1}^S (x_s - \bar{x})^2$) で近似される. これらから, WAIC は AIC と同様にバイアス補正した適合度を -2 倍した次の量で定義される.

$$(4.1) \quad \text{WAIC} = -2 \sum_{t=1}^T \log \mathbb{E}_{\text{post}}[f(y_t | \boldsymbol{\theta})] + 2p_{\text{WAIC}}.$$

この量が最小となるモデルが情報量規準の観点から選択されることになる.

4.2 効率的重点サンプリングによる確率的コピュラの推定

コピュラに対して確率変動を導入したのは Hafner and Manner (2012) である. 推定方法には, 効率的重点サンプリング (EIS) と呼ばれる, Liesenfeld and Richard (2003) によって考案された重点サンプリングのひとつが尤度評価に用いられている.

Hafner and Manner (2012) は 2 次元のペアコピュラに対してその依存構造パラメータに確率

変動を導入し、依存構造パラメータ δ が潜在変数 $x_t (t = 1, \dots, T)$ に駆動されると仮定し、 $\delta(x_t)$ として潜在変数 x_t の関数として与えている。このもとで、コンピュータベースの観測方程式と潜在変数の挙動を記述する状態方程式を次のように与える。

$$(4.2) \quad C(u_t, v_t | \delta(x_t)), \quad x_t = \alpha_x + \beta_x x_{t-1} + \sigma_x \eta_t \quad (t = 1, \dots, T).$$

ここで、Hafner and Manner (2012) は、観測データであるリターンの周辺分布モデルを SV モデルとし、 u_t と v_t は一様化変数で、周辺分布を通じて観測データを変換したコンピュータベースの観測データである。なお、状態方程式の攪乱項は標準正規分布に従い、 $\eta_t \sim \mathcal{N}(0, 1)$ である。

$\theta = (\alpha_x, \beta_x, \sigma_x)$ とすると、このモデルの観測データ $U_{1:T} = \{u_t\}_{t=1}^T$ と $V_{1:T} = \{v_t\}_{t=1}^T$ に関する尤度は、AR(1) の構造を持つ潜在変数 $X_{1:T} = \{x_t\}_{t=1}^T$ をモデルに内包することから、式 (4.2) の状態方程式より、

$$\begin{aligned} \mathcal{L}(U_{1:T}, V_{1:T}; \theta) &= \int f(U_{1:T}, V_{1:T}, X_{1:T}; \theta) dX_{1:T} \\ &= \int \prod_{t=1}^T f(u_t, v_t, x_t | X_{1:t-1}; \theta) dX_{1:T} \\ &= \int \prod_{t=1}^T c(u_t, v_t | x_t; \theta) p(x_t | x_{t-1}; \theta) dX_{1:T} \end{aligned}$$

として潜在変数に関する積分によって表される。ここで $c(u_t, v_t | x_t; \theta)$ は依存構造の確率密度関数、 $p(x_t | x_{t-1}; \theta)$ は状態方程式の確率密度関数を表し、 $p(x_1 | x_0; \theta) = p(x_1; \theta)$ とする。この尤度関数は、潜在変数系列の存在により、解析解の導出が困難であるため、数値計算によって評価される。まず、モンテ・カルロ法によって積分評価を行う場合、式 (4.2) より、 $X_T^{(k)} = \{x_t^{(k)}\}_{t=1}^T$ を状態方程式の確率密度関数である $p(x_t^{(k)} | x_{t-1}^{(k)}; \theta)$ の正規分布から K 個のサンプリングをすることで、

$$(4.3) \quad \hat{\mathcal{L}}(U_{1:T}, V_{1:T}; \theta) = \frac{1}{K} \sum_{k=1}^K \prod_{t=1}^T c(u_t, v_t | x_t^{(k)}; \theta)$$

で尤度を得る。しかしながら、Liesenfeld and Richard (2003) によれば、状態方程式のナチュラル・サンプラーである $p(x_t^{(k)} | x_{t-1}^{(k)}; \theta)$ は観測データの情報に乏しく、推定が非効率とされている。そこで Hafner and Manner (2012) は、Liesenfeld and Richard (2003) によって考案された重点サンプリングのひとつである EIS によって推定を行っている。EIS では、観測データの情報を取り込むための補助パラメータ $a_t (t = 1, \dots, T)$ と、補助パラメータを有する補助サンプラー $m(x_t | X_{1:t-1}; a_t) (t = 1, \dots, T)$ を用いることで尤度関数を次のように変形する。

$$\mathcal{L}(U_{1:T}, V_{1:T}; \theta) = \int \prod_{t=1}^T \left[\frac{c(u_t, v_t | x_t; \theta) p(x_t | x_{t-1}; \theta)}{m(x_t | X_{1:t-1}, a_t)} \right] \prod_{t=1}^T m(x_t | X_{1:t-1}, a_t) dX_{1:T}.$$

この尤度を数値的に評価するために、 $t = 1, \dots, T$ について a_t を補助パラメータとする補助サンプラー $m(x_t | X_{1:t-1}; a_t)$ から $\tilde{x}_t^{(k)}(a_t)$ を発生させ、

$$(4.4) \quad \tilde{\mathcal{L}}(U_{1:T}, V_{1:T}; \theta) = \frac{1}{K} \sum_{k=1}^K \prod_{t=1}^T \left[\frac{c(u_t, v_t | \tilde{x}_t^{(k)}(a_t); \theta) p(\tilde{x}_t^{(k)}(a_t) | \tilde{x}_{t-1}^{(k)}(a_{t-1}); \theta)}{m(\tilde{x}_t^{(k)}(a_t) | \tilde{X}_{1:t-1}^{(k)}(a_{t-1}); a_t)} \right]$$

を得る。ここで $k = 1, \dots, K$ はサンプリング数である。

さて、重点サンプリングの効率を向上させるためには、補助サンプラー m と補助パラメー

タ a の決定が重要となる．ここで，補助サンプラーを構成する補助関数 k_x を補助サンプラー m の密度カーネルとして導入し，

$$m(x_t|X_{1:t-1}; a_t) = \frac{k_x(x_t, x_{t-1}; a_t)}{g(x_{t-1}; a_t)}, \quad g(x_{t-1}; a_t) = \int k_x(x_t, x_{t-1}; a_t) dx_t$$

とする．さらに，補助関数 k_x について，状態方程式の密度関数 $p(x_t^{(k)}|x_{t-1}^{(k)}; \theta)$ が正規分布であることを考慮して， $a_t := (a_{1t}, a_{2t})$ ， $\zeta(x_t; a_t) := \exp(a_{1t}x_t + a_{2t}x_t^2)$ なる別の補助関数を導入し， $k_x(x_t, x_{t-1}; a_t) = p(x_t|x_{t-1}; \theta)\zeta(x_t; a_t)$ とする． $p(x_t|x_{t-1}; \theta)$ が正規分布の確率密度関数であることから， $\zeta(x_t; a_t)$ を上式の形にすることで，補助サンプラーは引き続き正規分布となる．この設定のもと，式 (4.4) は，

$$(4.5) \quad \tilde{\mathcal{L}}(U_{1:T}, V_{1:T}; \theta) = \frac{1}{K} \sum_{k=1}^K \prod_{t=1}^T \left[\frac{c(u_t, v_t | \tilde{x}_t^{(k)}(a_t); \theta) g(\tilde{x}_{t-1}^{(k)}(a_{t-1}); a_t)}{\zeta(\tilde{x}_t^{(k)}(a_t); a_t)} \right]$$

のように書き換えられる．

さて，補助パラメータ a_t の推定であるが，表現の簡単のため， k 番目のサンプリングによって得られる尤度 $\mathcal{L}^{(k)}$ を以下で解説する．式 (4.5) について， $g(\tilde{x}_{t-1}^{(k)}; a_t)$ に含まれる潜在変数 x_t の時点の違いに着目し，便宜上 $g(x_T(a_T); a_{T+1}) = 1$ を満たす a_{T+1} を想定することで，

$$\begin{aligned} \mathcal{L}^{(k)} &= \prod_{t=1}^T \left[\frac{c(u_t, v_t | \tilde{x}_t^{(k)}(a_t); \theta) g(\tilde{x}_{t-1}^{(k)}(a_{t-1}); a_t)}{\zeta(\tilde{x}_t^{(k)}(a_t); a_t)} \right] \\ &= \frac{1}{g(\tilde{x}_T^{(k)}(a_T); a_{T+1})} \prod_{t=1}^T \left[\frac{c(u_t, v_t | \tilde{x}_t^{(k)}(a_t); \theta) g(\tilde{x}_t^{(k)}(a_t); a_{t+1})}{\zeta(\tilde{x}_t^{(k)}(a_t); a_t)} \right] \end{aligned}$$

と変形できる．上式より，観測データの情報を補助パラメータ a_t に取り込むために， $g(x_T; a_{T+1}) = 1$ を仮定して $t = T, \dots, 1$ と時刻を遡るように x_t に関する最小二乗推定を繰り返すことで更新する．

パラメータ θ を所与として，尤度評価の一連の流れをまとめると，

(1) 与えられた状態方程式のパラメータ θ のもとで，ナチュラル・サンプラー $p(x_t|x_{t-1}; \theta)$ から K 個の潜在変数系列 $\{\tilde{x}_t^{(k)}\}_{t=1}^T (k = 1, \dots, K)$ を発生させる．

(2) 時刻 $t = T, \dots, 1$ と時間を遡るように時点 t における対数尤度をもとにした次式の回帰を繰り返し，補助パラメータ $\{a_t\}_{t=1}^T$ の推定値を得る．

$$\tilde{a}_t = \arg \min_{a_t} \sum_{k=1}^K [\log c(u_t, v_t | \tilde{x}_t^{(k)}; \theta) + \log g(\tilde{x}_t^{(k)}; \hat{a}_{t+1}) - c_t - \log \zeta(\tilde{x}_t^{(k)}; a_t)]^2$$

ここで c_t は定数項を表し， a_{T+1} は $g(x_T; a_{T+1}) = 1$ を満たす．

(3) 推定した補助パラメータ $\{\tilde{a}_t\}_{t=1}^T$ を元に，補助サンプラー $m(x_t|X_{1:t-1}; \tilde{a}_t)$ から潜在変数系列 $\{\tilde{x}_t^{(k)}(\tilde{a}_t)\}_{t=1}^T (k = 1, \dots, K)$ を発生させる．

(4) 補助パラメータ $\{\tilde{a}_t\}_{t=1}^T$ が収束するまで (2) と (3) を繰り返す．

(5) 収束した補助パラメータ $\{\check{a}_t\}_{t=1}^T$ による補助サンプラー $m(\check{x}_t|X_{1:t-1}; \check{a}_t)$ から潜在変数系列 $\{\check{x}_t^{(k)}(\check{a}_t)\}_{t=1}^T (k = 1, \dots, K)$ を生成し，尤度評価を行う．

なお，上で $a_t = (a_{1t}, a_{2t})$ と $\zeta(x_t; a_t) = \exp(a_{1,t}x_t + a_{2,t}x_t^2)$ と定義したことにより，この回帰式は x_t と x_t^2 を説明変数とする線形回帰となり，計算負荷の削減につながる．また，Liesenfeld and Richard (2003)によれば，この一連のプロセスの中での潜在変数の発生には，共通の乱数

(common random numbers, CRNs)を用いる必要がある点に注意が必要である。確率的依存構造パラメータの推定においては、Hafner and Manner (2012)は、100 から 200 の潜在変数の系列数 K があれば十分と報告している。

5. 確率的コピュラの応用

5.1 2 変量確率的コピュラの応用：確率的レバレッジ

SV モデルのレバレッジ構造に静的コピュラを導入する先行研究は Nakamura (2011) や Mehrabian (2012) で試みられている。SV モデルで、確率的コピュラを導入するには、例えば、高頻度データから計算される実現分散 (RV) で強化したより現実的な SV モデル (Takahashi et al., 2009) が有効であり、これを用いて説明する。

$$\begin{aligned}
 r_t &= \mu_r + e^{h_t/2} \sqrt{\frac{\nu_r - 2}{\nu_r}} \frac{\epsilon_t}{\sqrt{z_{r,t}}}, \quad z_{r,t} \sim \mathcal{G}\left(\frac{\nu_r}{2}, \frac{\nu_r}{2}\right), \\
 h_{t+1} &= \mu_h + \phi_h(h_t - \mu_h) + \sigma_h \frac{\eta_t}{\sqrt{z_{h,t}}}, \quad z_{h,t} \sim \mathcal{G}\left(\frac{\nu_h}{2}, \frac{\nu_h}{2}\right), \\
 (5.1) \quad f(r_t, h_{t+1}|h_t) &= f(r_t|h_t)f(h_{t+1}|h_t)c(U_t, V_t|R_t), \quad R_t = \tanh(x_t/2) \\
 \log RV_t &= \xi + h_t + \sigma_{RV}\zeta_t, \\
 x_t &= \mu_x + \phi_x(x_{t-1} - \mu_x) + \sigma_x \epsilon_{x,t}.
 \end{aligned}$$

ここで、 $c(U_t, V_t|R_t)$ は動的依存係数 R_t をもつ確率的コピュラの確率密度関数である。 $U_t = N_1(\epsilon_t), V_t = N_1(\eta_t) (N_1(\cdot): 1$ 次元標準正規確率分布関数)、攪乱項の (ϵ_t, η_t) には確率的正規コピュラの依存構造を入れ、簡単のため、その他のノイズ $\zeta_t, \epsilon_{x,t}$ はそれぞれ独立な標準正規分布に従うと仮定する。リターンと SV の攪乱項は正規分布よりは t 分布を適用したほうが当てはまりがよいので、ガンマ分布 $\mathcal{G}(\cdot, \cdot)$ に従う z_r, z_h を用いて上記のような設定にした。レバレッジを表すコピュラには正規コピュラを用い、依存係数は確率相関 $R(x) = \tanh(x/2)$ とした。

Bollerslev et al. (2009, 2014) は、分散リスクプレミアム (VRP) がリターンの予測可能性に説明力をもつとする理論、実証論文である。Nakamura (2017, 2018) の研究では、予測可能性に重要な役割を担っているのが、レバレッジ係数であることを明らかにしている。そこで、日経平均と S&P500 に対して、このモデル (5.1) を適用し、レバレッジ係数の動的変化をそれぞれ推定してみよう³⁾。

図 2 は、確率的レバレッジ相関係数の変化を表している。この値は負の領域で変化するが、株式指数が大きく下落した、所謂、リスクオフの時期には、レバレッジが強く (絶対値が大きく) なることが分かる。また、確率相関の平均回帰レベル $\rho_\infty = \tanh(\mu_x/2)$ は日経平均、S&P500 でそれぞれ -0.3273 , -0.4677 と推計され、日経平均のほうがレバレッジがより弱い (負で絶対値が小さい) ことがわかる。このことは、Bollerslev et al. (2014) で実証された VRP によるリターンの予測可能性で、S&P500 に比べて、日経平均が劣る理由の一つと考えられる。このように、確率的コピュラを使うことでレバレッジ係数が強まる時点が同定できるため、予測力が相対的に高まっている時期を知ることができる。レバレッジが静的な SV モデルでは、平均化した負の一定値として推定され、市場のリスクオン・オフとの関連性が不明であるが、動的モデルを使うことで、リスク管理や投資に活用できる可能性があるであろう。

5.2 確率的コピュラの応用：両側裾依存係数の確率的コピュラ

Nozawa and Nakamura (2016) では、株式指数とそれに対応するボラティリティ・インデッ

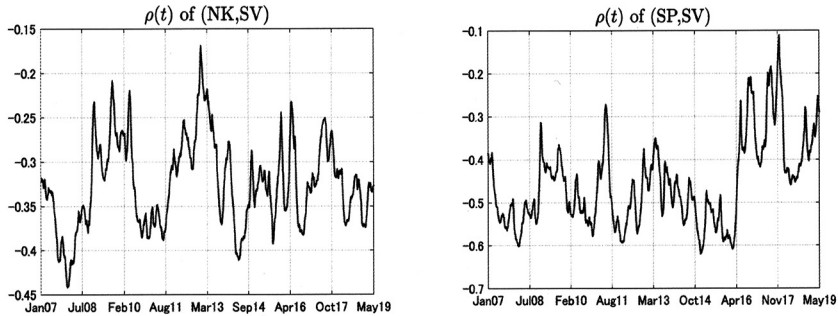


図 2. 確率的正規コピュラで表現したレバレッジパラメータの相関係数の変化(左図は日経平均, 右図は S&P500, 分析期間は 1/3/2007-5/31/2019, MCMC はバーンイン期間 1000 回, 試行回数 5000 回を採用).

クスを対象に依存構造の実証研究を行っている。依存構造の記述については、正規コピュラや t コピュラなどの楕円コピュラ以外に、上下裾依存について異なる 2 つのパラメータを有する BB7 コピュラ⁴⁾についても言及し、さらにそれぞれのパラメータに対して独立な確率変動を持たせるよう拡張している。

この研究では、inference functions for margins (IFM) と呼ばれる 2 段階推定を採用 (Joe, 2014 を参照) している。第 1 段階では、各データの周辺分布の推定を単変量のモデルで行う。第 2 段階では、先に推定した周辺分布を通じて、観測データを一様分布に従う変数へと変換し、コピュラの依存構造の推定を行う。

まず単変量の周辺分布であるが、ここでは、各指数のリターンの系列のボラティリティが市場環境に応じて変化する特徴を表現するため、株式リターンに SV モデルを適用している。SV モデルは状態空間モデルによる記述となり、観測データを含む観測方程式と、潜在変数の推移を表現する状態方程式の 2 本から構成される。観測方程式の攪乱項は、各指数のリターンや差分の分布の裾の厚さと、市場の上昇と下落についての非対称性を考慮して、一般化双曲型非対称 t 分布 (generalized hyperbolic skewed t , GHST) を仮定している⁵⁾。また、状態方程式の攪乱項を、正規分布ではなく t 分布を仮定し、潜在変数の推移に裾の厚さを持たせ大幅なボラティリティ変動を表現する。さらに、株式指数については、負のリターンとボラティリティの上昇が同時に観測される傾向がある。したがって、リターンを表現する観測方程式とボラティリティを表現する状態方程式の攪乱項の間に相関を仮定する。これらを踏まえて、この研究の SV モデルは、

$$(5.2) \quad \begin{aligned} y_t &= \mu_y + e^{h_t/2} \epsilon_t, & \epsilon_t &\sim \mathcal{GHST}(\nu_1, \beta_{\text{GH}}), \\ h_{t+1} &= \mu_h + \phi_h(h_t - \mu_h) + \sigma_h \eta_t, & \eta_t &\sim \mathcal{T}(\nu_2, 0, 1), \end{aligned}$$

となる。ここで $\mathcal{GHST}(\nu_1, \beta_{\text{GH}})$ は GHST 分布を表し、裾の厚さのパラメータは $\nu_1 > 4$ であり、歪度パラメータは β_{GH} で、歪度が正(負)の場合、 $\beta_{\text{GH}} > 0$ ($\beta_{\text{GH}} < 0$) となる。 $\mathcal{T}(\nu_2, 0, 1)$ は自由度が $\nu_2 > 4$ の標準 t 分布を表す。また、ボラティリティのリターンに対する非対称性を表すために、相関 ρ を攪乱項 ϵ_t と η_t の間に仮定する。この相関を用いることで、観測方程式の攪乱項と状態方程式の攪乱項は次のように表現できる。

$$(5.3) \quad \epsilon_t = \beta_{\text{GH}}(z_{1,t} - \mu_{z1}) + \sqrt{z_{1,t}} e_{1,t}, \quad \mu_{z1} = \frac{\nu_1}{\nu_1 - 2}, \quad \eta_t = \frac{\rho e_{1,t} + \sqrt{1 - \rho^2} e_{2,t}}{\sqrt{z_{2,t}/\nu_2}}.$$

表 1. 依存構造パラメータと潜在変数の変換.

コピュラの種類	パラメータ定義域	依存構造パラメータ
正規	$\delta_t(x_t) \in (-1, 1)$	$\delta_t(x_t) = \tanh(x_t/2)$
t	$\delta_t(x_t) \in (-1, 1), \nu > 4$	$\delta_t(x_t) = \tanh(x_t/2)$
BB7	$\delta_t^L(x_t) > 0$	$\delta_t^L(x_t^L) = \exp(x_t^L)$
	$\delta_t^U(x_t) \geq 1$	$\delta_t^U(x_t^U) = \exp(x_t^U) + 1$

ここで, $z_{1,t} \sim IG(\nu_1/2, \nu_1/2)$, $z_{2,t} \sim G(\nu_2/2, 1/2)$ および $e_{1,t}, e_{2,t} \sim \mathcal{N}(0, 1)$ は互いに独立な確率変数であり, $IG(\cdot)$ と $G(\cdot)$ はそれぞれ逆ガンマ分布とガンマ分布を表す. また, 各株式指数に対応するボラティリティ・インデックスの周辺分布については, 指数のリターンではなく, 変分を観測データとして式 (5.2) のモデルを当てはめている.

周辺分布を適切に特定した後は依存構造の推定である. 周辺分布のモデルを通じて, 観測可能なデータを周辺分布 $F(r_t)$ によって一様分布 $U(0, 1)$ に従う観測データ $u_t = F(r_t)$ ($t = 1, \dots, T$) へと変換する. データ系列が d 次元ある場合, 一様化変数としての観測データ $\{u_{i,t}\}_{t=1}^T$ と $\{u_{j,t}\}_{t=1}^T$ ($i, j \in \{1, \dots, d\}$) の間の確率的ペアコピュラは次の状態方程式で表される.

$$(5.4) \quad C_{ij}(u_{i,t}, u_{j,t}; \delta_{ij,t}(x_{ij,t})), \quad x_{ij,t+1} = \mu_{xij} + \phi_{xij}(x_{ij,t} - \mu_{xij}) + \sigma_{xij}\eta_{ij,t}, \quad \eta_{ij,t} \sim \mathcal{N}(0, 1).$$

ここで $\{x_{ij,t}\}_{t=1}^T$ ($t = 1, \dots, T$) は潜在変数であり, 確率的依存構造パラメータ $\delta_{ij,t}(x_{ij,t})$ を駆動する.

依存構造の推定では, 正規コピュラ, tコピュラと BB7 コピュラをペアコピュラとして依存構造に時変性を持たせ, 依存構造パラメータが潜在変数によって駆動されると仮定している. 潜在変数は $\{x_t\}_{t=1}^T$ とし, 式 (5.4) に規定される AR(1) 構造を仮定する. 表 1 は, 潜在変数と依存構造パラメータの変換をまとめたものである. BB7 コピュラは 2 つの異なる依存構造パラメータを有し, それらは $\delta^L > 0$ と $\delta^U \geq 1$ である. ここで, 異なる独立の潜在変数がそれぞれのパラメータを駆動するとし, $\{x_t^L\}_{t=1}^T$ と $\{x_t^U\}_{t=1}^T$ とする. BB7 コピュラのパラメータ δ^L と δ^U はそれぞれ下方および上方裾依存係数に関連し, $\lambda^L = 2^{-1/\delta^L}$ と $\lambda^U = 2 - 2^{1/\delta^U}$ として変換される. それぞれの裾依存係数は $0 < \lambda^L < 1$ と $0 \leq \lambda^U < 1$ の範囲をとり, 1 に近いほど裾依存が強いことを表す.

Nozawa and Nakamura (2016) は, ドイツの株式指数として DAX, 米国の株式指数として S&P, およびこれらのボラティリティ・インデックスである VDAX と VIX の日次データを用い, ヴァインコピュラを適用して 4 次元の依存構造を表現している. データ期間は 2010 年 12 月末から 2014 年 12 月末で, 主な市場の攪乱イベントとして欧州債務危機を含む. 時間 t における株式指数 S_t の日次対数リターンを $y_t = 100 \times (\log(S_t) - \log(S_{t-1}))$ として計測し, ボラティリティ・インデックスは日次変化 $y_t = \Delta V_t = V_t - V_{t-1}$ を用いている.

推定方法は, 先に述べた IFM と呼ばれる 2 段階推定を採用している. また, 用いられているモデルでは周辺分布のボラティリティおよび依存構造を駆動する潜在変数を含むことから, HMC 法を採用し, 数値計算によるパラメータ推定を行なっている. 推定に際しては, Stan (Stan Development Team, 2020) を用い, はじめの 2000 サンプルをバーンイン期間として破棄し, その後の 20000 サンプルを推定に用いている.

周辺分布モデルでは SV モデルを想定しており, この周辺分布モデルには 3 つの異なる潜在変数が含まれる. 具体的には, ボラティリティを駆動する h_t , 観測方程式の裾の厚さと歪度に関連する $z_{1,t}$, そして状態方程式の攪乱項の裾の厚さに関連する $z_{2,t}$ である. HMC 法では, 推

定するパラメータセット $\theta_m = (\mu_y, \beta_{GH}, \nu_1, \mu_h, \phi_h, \sigma_h, \nu_2, \rho)$ について事前分布を設定する必要がある⁶⁾。各パラメータが満たす定義域を考慮してそれぞれの事前分布を下のように定め、また事前分布に含まれるハイパーパラメータは以下の通り設定している。

$$\mu_{yi}, \mu_{hi}, \beta_{GH} \sim \mathcal{N}(0, 3^2), \quad \frac{\phi_{ri} + 1}{2}, \frac{\phi_{hi} + 1}{2}, \frac{\rho_i + 1}{2} \sim \mathcal{B}(10, 2),$$

$$\sigma_{hi}^2 \sim \mathcal{IG}(10, 2), \quad \nu_1 - 4, \nu_2 - 4 \sim \mathcal{G}(15, 5).$$

周辺分布の推定の後、依存構造の推定を行う。この研究ではヴァインコピュラを採用しているため、各系列についてペアコピュラを推定しながら依存構造を積み上げる。確率的ペアコピュラを駆動する潜在変数は式 (5.2) の AR(1) に従うと仮定する。また、HMC 法による推定では、そのパラメータ μ_{xij} , ϕ_{xij} および σ_{xij} について事前分布を設定する必要があるが、その定義域を考慮して、それぞれ正規分布、ベータ分布、逆ガンマ分布に従うと仮定している。事前分布のハイパーパラメータは全てのペアコピュラで共通とし、

$$\mu_{xij} \sim \mathcal{N}(0, 3^2), \quad \frac{\phi_{xij} + 1}{2} \sim \mathcal{B}(10, 2), \quad \sigma_{xij}^2 \sim \mathcal{IG}(10, 2),$$

としている⁷⁾。

また、ヴァインの構造は複数存在するが、ここでは C-ヴァインを採用している。C-ヴァインを構成する各ペアコピュラの確率的依存構造パラメータは、潜在変数を通じて表 1 のように時変性を仮定している。また市場間の依存構造の推定に際して、次の 2 通りを考えている。1 つ目はドイツから米国への影響、2 つ目は米国からドイツへの影響である。このうち、ここではドイツから米国への影響について紹介する。添え字 i と j がとる $\{1, 2, 3, 4\}$ はそれぞれ $\{\text{DAX}, \text{VDAX}, \text{S\&P500}, \text{VIX}\}$ に割り当てられ、第 1 変数である DAX を C-ヴァインの軸として、4 変数ヴァインコピュラの時点 t の対数尤度は次式のように与えられる。

$$l_t(\theta_c) = l_{12}(u_{1,t}, u_{2,t}; \delta_{12,t}) + l_{13}(u_{1,t}, u_{3,t}; \delta_{13,t}) + l_{14}(u_{1,t}, u_{4,t}; \delta_{14,t})$$

$$+ l_{23|1}(u_{2|1,t}, u_{3|1,t}; \delta_{23,t}) + l_{24|1}(u_{2|1,t}, u_{4|1,t}; \delta_{24,t})$$

$$+ l_{34|12}(u_{3|12,t}, u_{4|12,t}; \delta_{34,t}).$$

時点は $t = 1, \dots, T$ とし、 θ_c は式 (5.4) のうち依存構造パラメータを駆動する潜在変数の状態方程式のパラメータセットを表す。

ここで、依存構造パラメータが確率的に変動する場合、ヴァインコピュラの構築について注意が必要である。具体的には、第 2 枝以降のペアコピュラについては、データ k ($k = \{\{1\}, \{12\}\}$) に条件付けされた観測データ $\{u_{i|k,t}\}_{t=1}^T$ の h 関数⁸⁾ を通じたデータの変換が必要となるが、 h 関数は依存構造パラメータ $\delta_{ik,t}(x_{ik,t})$ を含み、このパラメータは潜在変数 x_t を内包することから、潜在変数 x_t をある値で固定する必要がある。この研究では、HMC 法による潜在変数の推定の過程で得られる事後平均をもって値を固定している。すなわち、

$$(5.5) \quad \bar{x}_t = \frac{1}{N} \sum_{n=1}^N (x_t^{(n)})$$

によって潜在変数 x_t の値を代替している。周辺分布が SV モデルの場合も同様である。SV モデルによってフィルターにかけられた観測データ $\hat{\epsilon}_{i,t}$ を $u_{i,t} = F(\epsilon_{i,t}; h_{i,t})$ によって一様化変数 $\hat{u}_{i,t}$ へ変換することが必要である。ポラテリティを駆動する潜在変数 h_t を固定することで周辺分布を通じた変換が可能となるが、この潜在変数を HMC 法における事後平均をもって値を固定している。

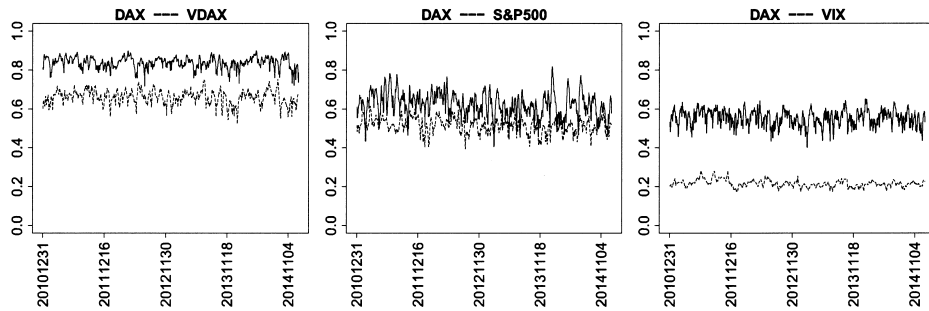


図 3. ドイツ市場から米国市場への影響を計測する BB7 コピュラの依存構造パラメータ(データ期間: 12/31/2010 – 12/31/2014). 実線と破線はそれぞれ下側と上側裾依存係数 λ_t^L と λ_t^U を表す. 左からそれぞれ DAX と VDAX, DAX と S&P500, DAX と VIX の依存構造を表す.

モデルの当てはまりについては, Watanabe (2010) による WAIC を用いて評価している. ヴァインコピュラ全体の WAIC については, それぞれのペアコピュラの WAIC の合計値として比較をしている. 結果としては, 前述の 4 つのコピュラのうち, BB7 コピュラの WAIC が他のコピュラと比較して当てはまりが良いと報告されている. ここから考えることは, 株式リターンとボラティリティ・インデックス変化の依存構造に上下非対称性が存在することである. 以下では BB7 コピュラに基づく確率的コピュラに焦点を当てる.

図 3 は上方および下方裾依存 λ_t^U と λ_t^L の変動を示している. 対象はドイツ市場から米国市場への影響で, 同日のリターンデータに基づく. 左と右のチャートは, DAX とボラティリティ・インデックスの依存構造である. 両チャートとも, 下側裾依存係数が上側裾依存係数よりも常に上位で推移している様子が見られる. ボラティリティ・インデックス変化についてはコピュラ・ベースの観測データの上下を反転させた反転データとしていることを考慮すると, DAX の下落とボラティリティ・インデックスの上昇が同時に発生する傾向があり, またその度合いが, DAX の上昇時よりも強い傾向が示唆されている. 中央のチャートは DAX と S&P500 の関係を表している. 上方および下方裾依存パラメータがともに同水準で推移していることから, 市場の上昇, 下落について上下対称の依存構造であることがわかる.

5.3 確率的依存構造をもつコピュラモデルによる為替ヘッジに対する考察

Nozawa and Nakamura (2015) では, 日本円をベース通貨とするファンドの視点から, 依存構造モデルの違いによる通貨ヘッジの効率性をグローバル株式およびグローバル債券について考察している.

単変量の資産リターン $\{r_t\}_{t=1}^T$ のモデルとして, リターン分布の歪度を考慮するとともに, 資産リターンの下落時にボラティリティが上昇する特徴を組み込む. これを表現するため, Skewed-t 分布を攪乱項に持つ AR(1)-GJR-GARCH(1,1) を採用し, 以下のように表す.

$$\begin{aligned}
 (5.6) \quad r_t &= \mu_r + \alpha_r(r_{t-1} - \mu_r) + \epsilon_t, \\
 \epsilon_t &= \sigma_t z_t, \\
 \sigma_t^2 &= \phi_0 + \phi_1 \epsilon_{t-1}^2 + \phi_2 \mathbb{I}_{(-\infty, 0)}(\epsilon_{t-1}) \epsilon_{t-1}^2 + \phi_3 \sigma_{t-1}^2.
 \end{aligned}$$

ここで σ_t は時点 t までの情報が反映された条件付ボラティリティで, $\mathbb{I}_{(-\infty, 0)}(\epsilon_{t-1})$ を $\epsilon_{t-1} < 0$ で 1 をとり, $\epsilon_{t-1} \geq 0$ で 0 をとる定義関数とすることでボラティリティの非対称性を表現する.

また、 $z_t \sim ST(z_t; \nu, \gamma)$ は Fernández and Steel (1998) の skewed-t 分布に従う攪乱項で、平均 0、分散 1 を取り、その確率密度関数は、

$$f_{ST}(z_t; \nu, \gamma) = \frac{2}{\gamma + \gamma^{-1}} [t_\nu(\gamma z_t) \mathbb{I}_{(-\infty, 0)}(z_t) + t_\nu(\gamma^{-1} z_t) \mathbb{I}_{[0, +\infty)}(z_t)]$$

で表される。ここで、 $t_\nu(\cdot)$ は自由度 $\nu > 4$ を持つ標準 t 分布であり、歪度を司るパラメータ γ は正値をとり、 $\gamma = 1$ のとき分布は対称となり、 $0 < \gamma < 1$ のとき歪度が負、 $\gamma > 1$ のとき正となる。依存構造の推定に際しては、コピュラベースの観測方程式と潜在変数の挙動を記述する状態方程式を式 (5.4) として与え、依存構造パラメータの定義域を考慮して、潜在変数 x_t を表 1 にまとめられた関数を通じて変換し $\delta_{ij,t}$ を得る。また対象とするコピュラは楕円コピュラに属する正規コピュラと t コピュラとしている⁹⁾。

多通貨の為替ヘッジ戦略にあたり、多次元の動的な依存構造を把握するために、この研究では確率的ヴァインコピュラによる推定を行なっている。ひとつのシステムとして組み入れる資産数は、円建て外貨資産のリターン、ドル円およびユーロ円の 3 つであり、3 次元をヴァインコピュラで表現する。また、ヴァインの構築方法としては C-ヴァインを用いており、円建て外貨資産を C-ヴァインの第 1 変数として議論している。ここからは、添え字の $i \in \{1, 2, 3\}$ をそれぞれ円建て外貨資産、ドル円およびユーロ円に割り当てる。ここで、時変依存構造パラメータ $\delta_{ij,t}$ ($i, j \in \{1, 2, 3\}$) は表 1 のように定義する。

ここで用いられるデータは、債券指数は Citi 米国国債インデックスと欧州国債インデックス、株式指数は S&P 500 と Eurostoxx 50、為替ヘッジに用いられるのは 1 ヶ月フォワード為替レート、そしてロンドン 4 時の米ドル円とユーロ円のスポット為替レートである。ここでは、彼らの分析期間のうち 2008 年末から 2014 年末について紹介する。この期間について、前 5 年間をイン・サンプル期間、後 1 年間をアウト・オブ・サンプル期間として設定している。

また、モデルの当てはまりについては WAIC を通じて評価し、t コピュラによるヴァインが正規コピュラと比較してパフォーマンスが良好であったとの報告をしている。

ヴァインコピュラを推定した後、日本円をベース通貨とするファンドの視点から為替ヘッジの効率性を検証している。具体的には、グローバル債券およびグローバル株式に対し、米ドル円とユーロ円のヘッジ効率を観察している。

推定したモデルに基づき、為替のヘッジ比率について次のように設定する。時刻 t におけるヘッジ付きポートフォリオのリターン r_t^H を、ヘッジなしポートフォリオのリターン $r_{1,t}$ と、2 つの為替リターンベクトル $r_t^F = (r_{2,t}, r_{3,t})^\top$ 、そしてヘッジ比率ベクトル $\beta_t = (\beta_{2,t}, \beta_{3,t})^\top$ によって次のように定義する。

$$(5.7) \quad r_t^H = r_{1,t} - \beta_t^\top r_t^F.$$

また、時刻 t におけるヘッジなし外貨資産リターンは、米国建て資産とユーロ建て資産への投資比率を w として

$$(5.8) \quad r_{1,t} = w^{\text{US}} r_t^{\text{US}} + w^{\text{EU}} r_t^{\text{EU}}$$

とし、円建て外貨資産のリターン r_t^i は、 $i \in \{\text{US}, \text{EU}\}$ として、 $r_t^i = \log(P_t^{a,i} P_t^{s,i} / P_{t-1}^{a,i} P_{t-1}^{s,i})$ とする。なお、 $P_t^{a,i}$ をグローバル資産指数、 $P_t^{s,i}$ を為替レートとし、米国とユーロのウェイト $w = (w^{\text{US}}, w^{\text{EU}})^\top$ は期を通じて一定とし $w^{\text{US}} = w^{\text{EU}} = 0.5$ とし、等ウェイトとする。ヘッジ比率 β_t は $0 \leq \beta_t \leq w$ とすることで、投機的な買いやオーバーヘッジを禁止している。

ヘッジ比率の計算は、周辺分布モデルである式 (5.6) と依存構造モデルである式 (5.4) をもとにシミュレーションによって生成された一期先リターンにより計算する。具体的には、 $t+1$ 期の各リターン $r_{t+1}^{(k)} = (r_{1,t+1}^{(k)}, r_{2,t+1}^{(k)}, r_{3,t+1}^{(k)})$ ($k = 1, \dots, K$) をコピュラモデルから生成し、期

待ショートフォール最小化問題の解として $t+1$ 期のヘッジ比率 β_{t+1} を決定する．ポートフォリオの損失 $L = -r_{t+1}^H$ について数値的に VaR を定め、次の最適化問題を解く．

$$(5.9) \quad \min_{\beta_{t+1} \in \mathbb{R}^2} \text{ES}_{p,t+1}(\beta_{t+1}) \quad \text{s.t.} \quad 0 \leq \beta_{t+1} \leq w.$$

ここで、期待ショートフォールの推定値は

$$\widehat{\text{ES}}_{p,t+1}(\beta_{t+1}) = \widehat{\text{VaR}}_{p,t+1}(\beta_{t+1}) + \frac{1}{pK} \sum_{k=1}^K \max(-r_{t+1}^{H(k)}(\beta_{t+1}) - \widehat{\text{VaR}}_{p,t+1}(\beta_{t+1}), 0)$$

で与えられ、シミュレーションに基づくヘッジ後リターン $r_{t+1}^{H(k)}(\beta_{t+1})$ は式 (5.7) によって求め、また、シミュレーションに基づく $\widehat{\text{VaR}}_{p,t+1}$ は、 p 分位点について $r_{t+1}^{H(k)}(\beta_{t+1})$ ($k = 1, \dots, K$) より求める．

ヘッジ比率は、 $t+1$ 時点のヘッジ比率を t 時点までの市場情報をもとに予測して算出する．推定したモデルに基づく 1 日先予測により期待ショートフォールを式 (5.9) を通じて計算し、最小化ヘッジ比率 β_{t+1} を決定する．なお、計算負荷が高いため、予測はイン・サンプル期間のモデルのパラメータに基づき、期先予測においてはモデルのパラメータを再度推定せずに固定する．イン・サンプル期間の終点を T_{in} とし、アウト・オブ・サンプル期間である $t = T_{\text{in}} + 1, \dots, T_{\text{out}}$ において市場情報を更新する．なお、シミュレーションでは 10000 のサンプルパスを発生させてヘッジ効率を評価している．

ヘッジなしおよびヘッジ付きポートフォリオリターンをもとに、Hilal et al. (2011) を参考に分散削減比率、VaR 削減比率、および期待ショートフォール削減比率によってヘッジ効率を評価している．それぞれの指標は次のとおりである．

$$\text{HPM}_1 = 1 - \frac{\text{Var}(r^H)}{\text{Var}(r^{\text{UH}})}, \quad \text{HPM}_2 = 1 - \frac{\widehat{\text{VaR}}_p(r^H)}{\widehat{\text{VaR}}_p(r^{\text{UH}})}, \quad \text{HPM}_3 = 1 - \frac{\widehat{\text{ES}}_p(r^H)}{\widehat{\text{ES}}_p(r^{\text{UH}})}.$$

ここで p は VaR と期待ショートフォールの分位点を表し、添え字の H と UH はそれぞれヘッジ付とヘッジ無しリターンを表す．なお、これらの指数が高位であるほどヘッジが効率的であることを示す．アウト・オブ・サンプル期間の分散は当該期間のシミュレーション・リターンに基づく標本分散、VaR および期待ショートフォールについては、当該期間のシミュレーション・リターンに基づく平均値として次のように定めている．

$$\widehat{\text{VaR}}_p = \frac{1}{T_{\text{out}} - T_{\text{in}}} \sum_{t=T_{\text{in}}+1}^{T_{\text{out}}} \text{VaR}_{p,t}, \quad \widehat{\text{ES}}_p = \frac{1}{T_{\text{out}} - T_{\text{in}}} \sum_{t=T_{\text{in}}+1}^{T_{\text{out}}} \text{ES}_{p,t}.$$

結果は表 2 の通りである．グローバル債券については、 t コピュラに基づくヘッジが正規コピュラに基づくヘッジよりも効率的であるとの結果が、特に VaR と期待ショートフォールの

表 2. ヘッジ・パフォーマンス評価.

	グローバル債券				グローバル株式			
	正規コピュラ		t コピュラ		正規コピュラ		t コピュラ	
	$p = 5\%$	$p = 1\%$	$p = 5\%$	$p = 1\%$	$p = 5\%$	$p = 1\%$	$p = 5\%$	$p = 1\%$
HPM 1	0.8188	0.8466	0.8865	0.8838	0.3730	0.4049	0.3831	0.4014
HPM 2	0.2941	0.2001	0.6677	0.6528	0.1653	0.2109	0.2705	0.2584
HPM 3	0.3640	0.3723	0.6618	0.6483	0.2033	0.1972	0.2556	0.2564

指標 HPM 2 と HPM 3 から得られたと報告している。確率的な裾依存構造を適切に捕捉することが、特にテール・リスクのヘッジ効率向上につながると考えられる。また、グローバル債券とグローバル株式のリスク削減効果を比較すると、特に期待ショートフォールについて債券に対する為替ヘッジの効果がより高いことが報告されている。

6. 結論と今後の課題

確率的コピュラモデルに関する統計的推定と関連するトピック、特に、ファイナンスへの応用事例を中心に解説を行った。確率的コピュラでは依存構造のパラメータが潜在変数の関数として与えられるため静的コピュラのような最尤推定を行うことができない。そのため、フィルタリングやベイズ推定法などの統計的推定方法を用いることになる。本稿では EIS, HMC 法を中心に解説した。ファイナンスにおける応用事例に関しては、枚挙に暇はないが、動的依存関係が重要な例として、ダイナミックヘッジ、アセット・アロケーション、資産価格理論におけるベータ・アノマリーなどを取り上げた。

今後の課題に関しては、高次元の確率的ヴァインコピュラの推定では、一般的には次元の呪いで推定は容易ではないが、事前に 2 変量の確率的ヴァインコピュラを逐次的にできるだけ精緻に推定しておき、パラメータのレンジを押さえた上で、同時推定を行うのが薦められる。リスク管理やアセット・アロケーションでは多変量の依存構造のモデル化が特に重要である。Yoshida (2018) の多変量 Skewed-t コピュラの最尤推定法の研究や、夷藤・中村 (2019) における Engle (2002) の DCC (Dynamic Conditional Correlation) モデルを組み合わせた多変量非対称 t コピュラによる資産運用への応用研究などを多変量確率的ヴァインコピュラに拡張するのも今後の興味深い課題であろう。

注.

- 1) レバレッジ係数と呼ばれる。
- 2) $K(p) = p^2/(2m)$; ここで m は粒子の質量で、 p は粒子の運動量、通常 m は 1 と取るが、Fisher の情報行列と関係する量である。
- 3) 株式指数の RV は Oxford-Man Institute から入手可能である。
- 4) 付録参照。Joe (2014) に詳しい。
- 5) GHST については Aas and Haff (2006) が詳しい。なお、GHST には様々なパラメータ設定があるが、この研究では Nakajima and Omori (2012) を参照している。
- 6) AR モデルのパラメータ μ_h , ϕ_h および σ_h の事前分布については、たとえば Almeida and Czado (2012) を参照。その他の事前分布は、パラメータの上限と下限を考慮して設定した。具体的には、上限と下限について、 $(-\infty, +\infty)$ のパラメータは μ_h と同様に正規分布を、上限と下限の両方を有するパラメータは ϕ_h と同様にベータ分布を、下限のみ有するパラメータはガンマ分布を仮定してそれぞれの定義域を考慮した。
- 7) t コピュラの推定には自由度 ν が含まれるが、推定の安定性のため、このパラメータについてはまず依存構造パラメータが時不変である静的 t コピュラで最尤推定し、そこで得られた自由度の推定値を用いている。また、BB7 コピュラは負の依存構造を表現できないことから、静的正規コピュラの前最尤推定で負の関係性が示唆されたペアについては第 2 変数を $1 - u_j$ とした第 2 軸反転アルキメディアン・コピュラを推定している。
- 8) 本稿に関連するコピュラの h 関数については付録を参照。
- 9) ただし、t コピュラの自由度パラメータ ν は、推定の安定性のため、依存構造パラメータが時不変の静的コピュラモデルの前最尤推定によって求めた結果を用いている。

謝 辞

本論文作成にあたり、日本銀行金融機構局の吉羽要直博士ならびに匿名のレフェリーの方から有益なコメントをいただいた。ここに深い感謝の意を示す次第である。なお、本稿の内容は野澤勇樹が属するステート・ストリート・グローバル・アドバイザーズ株式会社あるいはそのグループ会社の見解を示すものではなく、個人の見解である。The opinions expressed here by the Author, Yuki Nozawa, are solely Author's opinions and do not reflect opinions of State Street Global Advisors (Japan) Co., Ltd. or its group entities.

付録：各種ペアコピュラの補足説明

ペアコピュラ $C(u, v)$ を組み合わせるヴァインコピュラの構築では $h(v|u) := \partial C(u, v)/\partial u$ で定義される h 関数での変換が必要となる。本論文で採用した各種ペアコピュラとその h 関数は以下のように与えられる。

A.1 楕円コピュラの h 関数

Joe (2014) より、正規コピュラと t コピュラの h 関数はそれぞれ

$$h^{\text{Normal}}(v|u; \delta) = \Phi\left(\frac{\Phi^{-1}(v) - \delta\Phi^{-1}(u)}{\sqrt{1-\delta^2}}\right),$$

$$h^{\text{Student}}(v|u; \delta, \nu) = t_{\nu+1}\left(\frac{t_{\nu}^{-1}(v) - \delta t_{\nu}^{-1}(u)}{\sqrt{(\nu + t_{\nu}^{-1}(u)^2)(1-\delta^2)/(\nu+1)}}\right)$$

で与えられる。ここで $\Phi(\cdot)$ と $t_{\nu}(\cdot)$ はそれぞれ標準正規分布と自由度 ν の t 分布の分布関数を表す。

A.2 アルキメディアン・コピュラの h 関数

2変量クレイトンコピュラでは、裾依存パラメータを $\alpha(\geq 0)$ とすると、コピュラ分布関数

$$C^{\text{Clayton}}(u, v; \alpha) = (u^{-\alpha} + v^{-\alpha} - 1)^{1/\alpha}$$

から

$$h^{\text{Clayton}}(v|u; \alpha) = u^{-\alpha-1}(u^{-\alpha} + v^{-\alpha} - 1)^{-1-1/\alpha}.$$

2変量グンベルコピュラでは、裾依存パラメータを $\delta(\geq 1)$ とすると、コピュラ分布関数

$$C^{\text{Gumbel}}(u, v; \delta) = \exp[-\{(-\log u)^{\delta} + (-\log v)^{\delta}\}^{1/\delta}]$$

から

$$h^{\text{Gumbel}}(u|v; \delta) = C^{\text{Gumbel}}(u, v; \delta) \cdot \frac{1}{u} \cdot (-\log u)^{\delta-1} \{(-\log u)^{\delta} + (-\log v)^{\delta}\}^{1/\delta-1}.$$

と計算される。

A.3 BB7 コピュラ

Joe (2014) より、BB7 コピュラのコピュラ分布関数は以下のように表される。

$$C^{\text{BB7}}(u, v; \delta^U, \delta^L) = 1 - (1 - [(1 - \bar{u}^{\delta^U})^{-\delta^L} + (1 - \bar{v}^{\delta^L})^{-\delta^U} - 1]^{-1/\delta^L})^{1/\delta^U}$$

ここで, $\bar{u} = 1 - u$, $\bar{v} = 1 - v$ である. また, 確率密度関数は, $x = (1 - [1 - u]^{\delta^U})^{-\delta^L} - 1$, $y = (1 - [1 - v]^{\delta^U})^{-\delta^L} - 1$ とおいて,

$$c^{\text{BB7}}(u, v; \delta^U, \delta^L) = [1 - (x + y + 1)^{-1/\delta^L}]^{1/\delta^U - 2} (x + y + 1)^{-1/\delta^L - 2} [(x + 1)(y + 1)]^{1+1/\delta^L} \\ \cdot [\delta^U (\delta^L + 1) - (\delta^U \delta^L + 1)(x + y + 1)^{-1/\delta^L}] ([1 - u][1 - v])^{\delta^U - 1}$$

として表され, h 関数は $\partial C^{\text{BB7}}(u, v; \delta^U, \delta^L) / \partial u$ より,

$$h^{\text{BB7}}(v|u; \delta^U, \delta^L) = [1 - (x + y + 1)^{-1/\delta^L}]^{1/\delta^U - 1} (x + y + 1)^{-1/\delta^L - 1} (x + 1)^{1+1/\delta^L} [1 - u]^{\delta^U - 1}$$

として表される.

参 考 文 献

- Aas, K. and Haff, I. H. (2006). The generalized hyperbolic skew Student's t -distribution, *Journal of Financial Econometrics*, **4**(2), 275–309.
- Aas, M., Czado, C., Frigessi, A. and Bakken, H. (2009). Pair-copula constructions of multiple dependence, *Insurance: Mathematics and Economics*, **44**(2), 182–198.
- Adrian, T. and Brunnermeier, M. K. (2016). CoVaR, *American Economic Review*, **106**(7), 1705–1741.
- Almeida, C. and Czado, C. (2012). Efficient bayesian inference for stochastic time-varying copula models, *Computational Statistics and Data Analysis*, **56**, 1511–1527.
- Asness, A., Frazzini, A., Gormsen, N. and Pedersen, L. (2020). Betting against correlation: Testing theories of the low-risk effect, *Journal of Financial Economics* (to appear), <https://doi.org/10.1016/j.jfineco.2019.07.003>.
- Bedford, T. and Cooke, R. M. (2002). Vines — A new graphical model for dependent random variables, *Annals of Statistics*, **30**, 1031–1068.
- Bollerslev, T., Tauchen, G. and Zhou, H. (2009). Expected stock returns and variance risk premia, *Review of Financial Studies*, **22**, 4463–4492.
- Bollerslev, T., Marrone, J., Xu, L. and Zhou, H. (2014). Stock return predictability and variance risk premia: Statistical inference and international evidence, *Journal of Financial and Quantitative Analysis*, **49**, 633–661.
- Bloorforoosh, A., Christoffersen, P., Fournier, M. and Gouriéroux, C. (2020). Beta risk in the cross-section of equities, *The Review of Financial Studies* (to appear), <https://doi.org/10.1093/rfs/hhz139>.
- Boudt, K., Carl, P. and Peterson, B. (2013). Asset allocation with conditional value-at-risk budgets, *Journal of Risk*, **15**, 39–68.
- Brechmann, E. and Czado, C. (2011). Extending the CAPM using pair copulas: The regular vine market sector model, *Proceedings of 4th Workshop on Vine Copula Distributions and Applications*, Munich, May 11–12.
- Engle, R. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models, *Journal of Business and Economic Statistics*, **20**(3), 339–350.
- Fernández, S. and Steel, M. F. J. (1998). On Bayesian modelling of fat tails and skewness, *Journal of the American Statistical Association*, **93**, 359–371.
- Frazzini, A. and Pedersen, L. (2014). Betting against beta, *Journal of Financial Economics*, **111**, 1–25.
- Hafner, C. M. and Manner, H. (2012). Dynamic stochastic copula models: Estimation inference and applications, *Journal of Applied Econometrics*, **27**, 269–295.
- Heinen, A. and Valdesogo, A. (2009). Asymmetric CAPM dependence for large dimensions: The canonical vine autoregressive model, Working Paper, Departamento de Estadística, Universidad Carlos III de Madrid, Spain.

- Hilal, S., Poon, S. and Tawn, J. (2011). Hedging the black swan: Conditional heteroskedasticity and tail dependence in S&P500 and VIX, *Journal of Banking and Finance*, **35**, 2374–2387.
- Hoffman, M. and Gelman, A. (2014). The no-u-turn sampler: Adaptively setting path lengths in hamiltonian monte carlo, *Journal of Machine Learning Research*, **15**, 1593–1623.
- 夷藤翔, 中村信弘 (2019). ダイナミック非対称 t コピュラを用いた新興国債市場の相互依存構造に関する研究, *ジャファイ・ジャーナル*, **17**, 45–66.
- 岩永育子 (2015). ヴァインコピュラを用いた CAPM の非正規・非線形への拡張: 日本株式市場における実証分析, *ジャファイ・ジャーナル*, 69–113.
- Joe, H. (2014). *Dependence Modeling with Copulas*, Chapman and Hall/CRC, New York.
- 監物輝夫 (2017). CoVaR によるシステミック・リスク計測: 確率的コピュラによる比較分析, *ジャファイ・ジャーナル*, 10–48.
- Liesenfeld, R. and Richard, J. (2003). Univariate and multivariate volatility models: Estimation and diagnostics, *Journal of Empirical Finance*, **10**, 505–531.
- Liesenfeld, R. and Richard, J. (2006). Classical and bayesian analysis of univariate and multivariate stochastic volatility models, *Econometric Reviews*, **25**, 335–360.
- Liu, J. and West, M. (2001). Combined parameters and state estimation in simulation-based filtering, *Sequential Monte Carlo Methods in Practice* (eds. A. Doucet, N. de Freitas and N. Gordon), 197–223, Springer, New York.
- Mehrabian, A. (2012). The leverage effect in stochastic volatility: New models, Bayesian inference, and model selection, PhD Thesis, Department of Mathematics & Statistics, McGill University, Montréal, Canada.
- Nakajima, J. and Omori, Y. (2012). Stochastic volatility model with leverage and asymmetrically heavy-tailed error using GH skew Student's t -distribution, *Computational Statistics and Data Analysis*, **56**, 3690–3704.
- Nakamura, N. (2011). Copula-based asymmetric leverage in stochastic volatility models — Particle filtering approach —, *Proceedings of the 35th Meeting of the Japanese Association of Financial Econometrics and Engineering*, Vol. 35 Summer, 241–252.
- Nakamura, N. (2017). Asset return predictability and dynamics of return and variance risk premia, *Proceedings of the 48th meeting of the Japanese Association of Financial Econometrics and Engineering*, Vol. 48 Winter, 24–35.
- Nakamura, N. (2018). Non-affine and non-reduced form approach to pricing of VIX and VVIX: Quadratic diffusion model, *Proceedings of the 49th meeting of the Japanese Association of Financial Econometrics and Engineering*, Vol. 49 Summer.
- Nápoles, O., Cooke, R. M. and Kurowicka, D. (2010). About the number of vines and regular vines on n nodes, Working Paper, Delft Institute of Applied Mathematics, Delft University of Technology, *The Netherlands*.
- Neal, R. M. (2011). MCMC using Hamiltonian dynamics, *Handbook of Markov Chain Monte Carlo* (eds. S. Brooks, A. Gelman and G. Jones), 113–162, Chapman and Hall/CRC, New York.
- Nozawa, Y. and Nakamura, N. (2015). Dynamic hedging strategy with stochastic vine copulas, *Proceedings of the 43rd Meeting of the Japanese Association of Financial Econometrics and Engineering*, Vol. 43 Summer, 168–179.
- Nozawa, Y. and Nakamura, N. (2016). Estimation of stochastic dependence structures between equity markets and volatility indices using stochastic copulas, *Proceedings of the 44th meeting of the Japanese Association of Financial Econometrics and Engineering*, Vol. 44 Winter, 228–238.
- Richard, J. and Zhang, W. (2007). Efficient high-dimensional importance sampling, *Journal of Econometrics*, **141**, 1385–1411.
- Rockafellar, R. and Uryasev, S. (2000). Optimization of conditional value-at-risk, *Journal of Risk*, **2**, 21–41.
- Roncalli, T. (2013). *Introduction to Risk Parity and Budgeting*, Chapman and Hall/CRC, New York.

- Stan Development Team (2020). *Stan User's Guide 2.22*, https://mc-stan.org/docs/2_22/stan-users-guide-2_22.pdf.
- Takahashi, M., Omori, Y. and Watanabe, T. (2009). Estimating stochastic volatility models using daily returns and realized volatility simultaneously, *Computational Statistics and Data Analysis*, **53**, 2404–2426.
- Watanabe, S. (2010). Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory, *Journal of Machine Learning Research*, **11**, 3571–3594.
- Yoshihara, T. (2018). Maximum likelihood estimation of skew- t copulas with its applications to stock returns, *Journal of Statistical Computation and Simulation*, **88**, 2489–2506.

A Copula Model with Stochastic Tail Dependence: Statistical Inference and Applications to Quantitative Finance

Yuki Nozawa¹ and Nobuhiro Nakamura²

¹State Street Global Advisors (Japan) Co., Ltd.

²Graduate School of Business Administration, Hitotsubashi University

We survey model structures of stochastic copulas in which the dependence structures stochastically vary, as well as statistical estimation methods for these dependence structures. Because dependence structures in stochastic copulas are described by state equations that incorporate latent variables, likelihood evaluation requires numerical calculation. In this survey, we summarize these methodologies and discuss the application of stochastic copulas to multivariate dependence structures through vine copulas. We also introduce some applications to the field of finance, including time-varying copula models with leverage and copula models with time-varying dependence parameters. As an example of the latter models, we report a currency hedging model for time-varying dependence parameters in copula models.

コピュラを用いた CDO 価格付けモデルの リスク計測モデルへの拡張

室町 幸雄[†]

(受付 2019 年 6 月 6 日；改訂 2020 年 1 月 22 日；採択 1 月 23 日)

要 旨

2007~08 年の世界金融危機以前から多くの金融リスク計測モデルが開発されてきたが、金融危機の予兆は捉えられなかった。それは既存モデルが過去データを分析する純粋な統計学的モデルであったためである。本稿では、もともと CDO の価格付けのために提案されたインプライドコピュラを用いたプライシングモデルのリスク計測への拡張を提案する。そのモデルではデフォルト時刻は条件付独立と仮定し、デフォルト強度の分布を CDO 価格から推定するが、既存の推定結果は僅かな確率で同時デフォルト確率が極端に高まることを示唆していた。そのような結果をリスク計測に活用するのが本稿のモデルで、過去データだけでなく現在の時価情報も活用することで市場参加者の将来の環境激変への畏怖をリスク量に反映できる。本稿では既存理論に沿って具体的なリスク計測モデルを構築していくが、測度変換の重要性とその取扱いには特に詳細に議論する。このモデルではデフォルト確率全体に影響するファクターの存在が重要で、その分布次第では既存モデルでは見られない結果が現れる。例えば CDO のリスク計測では、ほとんどのトランシェで小確率で大規模損失が発生しうること、大規模損失の内訳が実際のデフォルト損失ではなく CDO 価格の暴落であることを示した。これらは金融危機時に見られた現象と整合的である。

キーワード：金融リスク管理，統計学的モデル，インプライドコピュラ，条件付独立，観測確率，リスク中立確率。

1. はじめに

2007 年以降の米国のサブプライム問題に端を発した世界金融危機では、世界中の金融機関で巨大損失が発生し、幾つもの主要金融機関がデフォルトした。それまでも金融機関はリスクの計測と管理に注力し、RiskMetrics™ (Morgan Guaranty Trust Company and Reuters Ltd., 1996), CreditMetrics™ (Gupton et al., 1997) などのさまざまな金融リスク計測モデルを開発してきたが、金融危機の予兆を捉えることはできなかった。後日、それはどのモデルも過去データを分析して将来のリスクを評価する純粋に統計学的なモデルであったためだと批判され、フォワードルッキングなリスク評価の重要性が指摘され、過去データに依らないシナリオに基づくストレステストが注目されるようになった。しかし、過去データを全く用いずにシナリオを構築すると既存の統計学的モデルと組み合わせた議論ができないという問題がある。一方、金融商品の市場価格、特にデリバティブの価格にはフォワードルッキングな情報が含まれると

[†] 東京都立大学大学院 経営学研究科：〒192-0397 東京都八王子市南大沢 1-1

考えられており、例えば、金融危機前の CDO (Collateral Debt Obligations) のトランシェ価格には、Hull and White (2006) や Brigo et al. (2010) の分析結果に見られるように、将来の巨大損失の発生を示唆すると解釈できるデータも含まれていた。例えば、CDO の各トランシェ価格から逆算されたインプライドコリレーション、特にベースコリレーションは、値が高いほど CDO に含まれる多くの裏付資産が同時にデフォルトする確率が高いことを示唆するが、これらの論文の分析によると、トランシェの優先度が高くなるほどベース・コリレーションの値は高くなっている。また、後述する Hull and White (2006) のインプライドコピュラの論文のモデルによるトランシェ価格の分析結果は、非常に低い確率で、極端に裏付資産のデフォルト強度が同時に高まる事象の発生を明示している。

本稿では、デリバティブ価格に内在する将来の損失発生に関するフォワードルッキングな情報を読み出して活用できる統計学的リスク計測モデルを提案する。もとにする考え方は、Hull and White (2006) がインプライドコピュラを提案した論文で使われたモデルと、Kijima and Muromachi (2000) による将来価値ベースのリスク計測のフレームワークである。Kijima and Muromachi (2000) は、プライシング(価格付け)モデルと金融規制で扱われるリスク計測モデルの違いを明確にして、両者を整合的に行うモデルを構築するための理論的なフレームワークを提案した。一般に、変数間の相互依存関係をコピュラ(接合関数)で表現したとき、測度変換を行うとコピュラは変化するが、Kijima and Muromachi (2000) が論文で取り上げた例はシンプルで、提案したフレームワークに関してコピュラの視点からは特に言及されていなかった。本稿では、Hull and White (2006) で例示されたプライシングモデルを Kijima and Muromachi (2000) に沿ってリスク計測モデルへ拡張する。これは、ガウシアンコピュラモデルよりも巨大なテイルリスクを表現できる Hull and White (2006) のモデルをベースにすることで、既存のリスク計測モデルでは得られない新しい結果の導出を狙ったものである。また、リスク計測モデルへの拡張をコピュラの視点からも言及することで、Kijima and Muromachi (2000) のフレームワークの解釈の拡大も試みる。

論文の構成は以下である。2 節ではもとにする 2 つの理論を概説し、3 節でリスク計測モデルを提案する。4 節ではモデルによる数値例を紹介し、5 節でまとめる。

2. もとにする既存理論

2.1 1-ファクターガウシアンコピュラモデルとインプライドコピュラ

Hull and White (2006) のインプライドコピュラは、当時活発に取りざされていた CDO トランシェ、特に CDS (Credit Default Swap) からなるポートフォリオを裏付資産とする合成 CDO (synthetic CDO) の市場価格を合理的に説明するためのモデルである。その説明の前に、標準的な CDO のプライシングモデルである 1-ファクターガウシアンコピュラモデル (one-factor Gaussian copula model, 本稿では OFGCM と略) から説明する。

n 個の異なる参照資産を考え、資産 j に対応する潜在変数を

$$(2.1) \quad X_j = \rho_j M + \sqrt{1 - \rho_j^2} Z_j, \quad j = 1, \dots, n$$

と表現する。ここで、定数 $\rho_j \in (-1, 1)$ はファクターローディング、 M と $Z_j, j = 1, \dots, n$ は互いに独立な標準正規分布に従う確率変数である。資産 i と資産 j のコリレーションは $\rho_i \rho_j$ と表現される。このとき X_j も標準正規分布に従うので、 F_j を τ_j の分布関数、 F_j^{-1} を F_j の逆関数、 Φ を標準正規分布の分布関数として、 j 番目の資産のデフォルト時刻 τ_j を

$$(2.2) \quad \tau_j = F_j^{-1}(\Phi(X_j)), \quad j = 1, \dots, n$$

で定義すると、(2.1)と(2.2)から、

$$(2.3) \quad F_j(t|M) \equiv P\{\tau_j \leq t|M\} = P\{X_j \leq \Phi^{-1}(F_j(t))|M\} = \Phi\left(\frac{\Phi^{-1}(F_j(t)) - \rho_j M}{\sqrt{1 - \rho_j^2}}\right)$$

が得られる。このモデルは OFGCM と呼ばれ、確率変数 M が具体的に与えられたとき、個々の資産の時刻 t までのデフォルトは(2.3)で与えられる条件付デフォルト確率で独立に発生する。

しかし、OFGCM、特にマーケットで広く使われた全資産の ρ_j を等しい($\rho_j = \rho$)と仮定する OFGCM では、同じ CDO に由来する複数のトランシェの市場価格を同時に説明することはできなかった。代わりに、各トランシェの市場価格を OFGCM で再現する ρ^2 がインプライドコリレーションと呼ばれ、デリバティブ市場におけるインプライドボラティリティと同様の役割を CDO 市場で果たしてきた。一方、全てのトランシェ価格を統一的に説明しようとして多くのモデルが提案されたが、その一つがインプライドコンピュータである。

本稿では、インプライドコンピュータによるデフォルトモデルを、条件付デフォルト確率がノンパラメトリックな分布を持つ、デフォルト時刻が条件付独立なモデルと定義する。条件となる確率変数の実現値が与えられたとき、資産ごとに条件付デフォルト確率が決まり、その確率で互いに独立にデフォルトが発生すると考えて、その確率変数の分布を、CDO の全トランシェの理論価格が市場価格と整合的になるように決定する。Hull and White (2006)が具体例として示したのは、 n 個の裏付資産すべてのデフォルト強度が等しく、かつ時間的に一定な場合である。彼らは強度が λ_k となる確率を π_k , $k = 1, \dots, L$ として、 L 個の組 $(\lambda_k, \pi_k)_{k=1, \dots, L}$ からなる λ の離散確率分布を複数のトランシェ価格から推定した。得られたデフォルト強度は非常に裾の長い分布になり、低確率だが非常に高い強度も現れた。理論的にはインプライドコンピュータをさらに拡張した形で定義することも可能であるが、本稿では Hull and White (2006)のオリジナルの論文に沿って上記の定義を採用する。

2.2 リスク計測モデルへの拡張のフレームワークと本稿のモデル

Kijima and Muromachi (2000)は、ポートフォリオの金利リスクと信用リスクを統合評価するフレームワークを提案した。リスク計測モデルは、ある将来時点の価値の分布をもとにその時点における損益を扱う将来価値ベースのモデルと、ある将来時点までに発生しうる潜在的な損失額の分布を扱うモデルに分けられ、市場リスク計測モデルは前者に、多くの信用リスク計測モデルは後者に属する。Kijima and Muromachi (2000)は前者に属し、金利やデフォルト強度などのリスクファクター(資産価格に影響を与える変数)を確率微分方程式などの確率モデルで表現し、現在価値と将来価値には無裁定価格を使い、モンテカルロシミュレーションを用いてリスクホライズン(リスク計測期間の最終時点) T におけるポートフォリオの将来価値分布を構築し、VaR(Value at Risk)やES(Expected Shortfall)などのリスク量を算出する。

このフレームワークの顕著な特徴は2つの確率測度、観測確率(実確率、統計的確率ともいう)と価格付けのための確率(リスク中立確率など)の両方を用いる点である。リスク計測では、将来シナリオは現実に発生しうるシナリオを扱うので、シナリオ生成には観測確率を使用する。一方、現在の状況から現在価値を、将来時点の状況からその時点の将来価値を算出する際は、リスク中立確率のような価格付けのための確率を使用して無裁定価格を算出する。そのため、ポートフォリオのリスク計測の実行手順は以下のようになる。

(1) 現時点 $t = 0$ におけるリスクファクター値を初期値として、リスクホライズン $t = T$ までのリスクファクターの将来シナリオを、与えられた確率モデルをもとにモンテカルロシミュレーションで多数発生させる。ここでは観測確率を使用する。

(2) 各シナリオごとに、時刻 T におけるリスクファクター値をもとに全資産の無裁定価格を算出し、ポートフォリオの将来価値を算出する。価格評価にモンテカルロ法を使う場合、時刻 T 以降の将来シナリオをリスク中立確率などの価格付けのための確率下で生成して、時刻 T 以降のキャッシュフロー列を算出し、その時刻 T における割引価値の期待値を無裁定価格とする。

(3) 手順(1)–(2)を多数回繰り返す、全シナリオのポートフォリオの将来価値から将来価値分布を作成してリスク量を算出する。得られたリスク量は観測確率下の値である。

プライシングモデルでは価格付けのための確率、例えばリスク中立確率下のリスクファクターの確率過程のみ与えればよい。なぜならば、無裁定価格はリスク中立確率における将来キャッシュフローの割引現在価値の期待値で与えられるからである。一方、プライシングモデルに多少の変更を行い、例えば将来キャッシュフローの割引現在価値の確率分布を求め、VaRなどのリスク量を求めたとしても、それらは市場参加者によるリスク調整が加味されたりリスク中立確率下でみたリスク量に過ぎないので、金融リスク管理で必要とされる観測確率下のリスク量にはならない。例えば、社債価格などから求めたインプライドデフォルト確率から直接 VaR を計算できたとしても、その値はリスク中立確率下の VaR になり、リスク管理実務で求められる観測確率下の VaR ではない。インプライドデフォルト確率を既存のリスク計測モデルにデフォルト確率として入力して VaR を算出しても同様である。CDO のトランシェ価格から推定されるインプライドコリレーションを用いる場合も同様だが、さらに、インプライドコリレーションにはスマイル、すなわち、ポートフォリオ全体のどの部分(どのトランシェ)を扱うかによって値が変わるとい現象が見られるので、インプライドコリレーションを用いてリスクを評価しても、選択したトランシェ部分のみにバイアスをかけた評価しか行うことができないので、やはりリスク管理の場で使える適切なリスク量は得られない。

将来価値ベースのリスク計測で観測確率下のリスク量を算出するには、観測確率下とリスク中立確率下の確率過程、あるいは観測確率下(またはリスク中立確率下)の確率過程と測度変換、を与えて上記の手順(1)–(3)を踏む必要がある。そのため、既存のプライシングモデルと整合的なリスク計測を行うためには、Kijima and Muromachi (2000) のフレームワークに沿ってプライシングモデルをリスク計測モデルへと拡張することが必要となる。

3. 提案モデル

本節では、Hull and White (2006) のモデルに Kijima and Muromachi (2000) のフレームワークを適用してリスク計測モデルを構築する。どちらもデフォルト時刻の条件付独立性を前提にしているため、モデルとしての相性は非常に良い。ただし、Bielecki and Rutkowski (2002) や Bielecki et al. (2009) によると、一般に、確率変数の条件付独立性は測度変換に対して保存されるとは限らない。そこで、条件付独立性が保存される条件とその意味を示すとともに、Kijima and Muromachi (2000) では議論されなかったコンピュータの観点からのリスク計測モデルへの拡張についても記述する。まずは一般論から始め、固有の設定や仮定を段階的に導入して提案モデルを説明し、数値例で使用する金融商品の評価方法を示す。

3.1 設定

時刻を $t \geq 0$ で表し、現時点を $t = 0$ 、リスクホライズンを $T > 0$ とする。観測確率を P として、フィルター付確率空間 $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, P)$ を考える。ポートフォリオは相異なる n 個の資産からなり、各資産の保有量は $t \in [0, T]$ で変わらないものとする。以下では、十分に先の時点(ポートフォリオ中の資産から生じる最も先のキャッシュフローの発生時点より先の時点)

$T^* > T$ までリスク中立確率 \tilde{P} が唯一つ存在すると仮定する.

資産 j のデフォルト時刻 $\tau_j, j = 1, \dots, n$ を正值確率変数, 1_A を定義関数(事象 A が真のとき $1_A = 1$, 偽のとき $1_A = 0$)として, デフォルト過程を $H_j(t) \equiv 1_{\{\tau_j \leq t\}}, j = 1, \dots, n$ で定義し, デフォルト過程により生成されるフィルトレーションを $\mathcal{H}_t^j = \sigma(H_j(s), 0 \leq s \leq t)$, $\mathcal{H}_t = \mathcal{H}_t^1 \vee \dots \vee \mathcal{H}_t^n$ とする. 詳細は後述するが, 金利とデフォルト確率を決定する確率過程から生成されるフィルトレーションを \mathcal{G}_t とし, $\mathcal{F} = \mathcal{G} \vee \mathcal{H}$, すなわち任意の $t \in [0, T^*]$ に対して $\mathcal{F}_t = \mathcal{G}_t \vee \mathcal{H}_t$ が成り立つとしてサブフィルトレーション・アプローチを用いる.

観測確率 P の下で, 時刻 t における瞬間的なデフォルトフリーな短期金利 $r(t)$ は

$$(3.1) \quad dr(t) = \mu_0(t)dt + \sigma_0(t)dz_0(t), \quad t \geq 0$$

に従うと仮定する. ただし, $\mu_0(t)$ と $\sigma_0(t)$ は \mathcal{F}_t -適合過程で, $z_0(t)$ は P 下における標準ブラウン運動である. デフォルトリスクに関しては誘導型モデルを用いることにして,

$$(3.2) \quad M_j(t) \equiv H_j(t) - \int_0^t (1 - H_j(s-))h_j(s)ds, \quad 0 \leq t \leq T^*, \quad j = 1, \dots, n$$

を観測確率 P 下で \mathcal{F} -マルチンゲールにする非負の \mathcal{F}_t -可予測過程 $h_j(t)$ を資産 j のデフォルト強度とする. 一般に $h_j(t)$ は \mathcal{F}_t -可予測過程で,

$$(3.3) \quad dh_j(t) = \mu_j(t)dt + \sigma_j(t)dz_j(t), \quad t \geq 0, \quad j = 1, \dots, n$$

に従うと仮定する. ただし, $\mu_j(t)$ と $\sigma_j(t)$ は \mathcal{F}_t -適合過程, $z_j(t)$ は P 下の標準ブラウン運動で, $dz_j(t)dz_k(t) = c_{jk}dt, j, k = 0, \dots, n$, ただし, $k \neq j$ で, c_{ij} は相関係数とする.

ラドンニコデイル微分 $d\tilde{P}/dP$ を \mathcal{F}_{T^*} -可測な確率変数とし, 密度過程を

$$\rho(t) \equiv E \left[\frac{d\tilde{P}}{dP} \middle| \mathcal{F}_t \right], \quad 0 \leq t \leq T^*$$

で定義すると, 適当な技術的条件のもとで, \mathcal{F}_t -可予測な確率過程 $\beta_j(t), j = 0, \dots, n$ と $\kappa_j(t), j = 1, \dots, n$ を用いて,

$$(3.4) \quad \rho(t) = 1 + \int_0^t \rho(s-) \left(- \sum_{j=0}^n \beta_j(s)dz_j(s) + \sum_{j=1}^n \kappa_j(s)dM_j(s) \right), \quad 0 \leq t \leq T^*$$

と一意に表現できる. このとき,

$$(3.5) \quad d\tilde{z}_j(t) = dz_j(t) + \beta_j(t)dt, \quad 0 \leq t \leq T^*, \quad \tilde{z}_j(0) = 0, \quad j = 0, \dots, n$$

で定義される $\tilde{z}_j(t)$ は \tilde{P} 下における標準ブラウン運動,

$$(3.6) \quad \tilde{M}_j(t) \equiv H_j(t) - \int_0^t (1 - H_j(s-))(1 + \kappa_j(s))h_j(s)ds, \quad 0 \leq t \leq T^*, \quad j = 1, \dots, n$$

は \tilde{P} 下における \mathcal{F} -マルチンゲールになる. (3.5) の $\beta_j(t)$ は $z_j(t)$ に対するリスクの市場価格で, (3.6) は $\tilde{h}_j(t) \equiv (1 + \kappa_j(t))h_j(t)$ が \tilde{P} 下のデフォルト強度であることを意味している.

以下では, 本稿固有の設定や仮定を順次導入しながらモデルを説明する. $z_{n+1}(t) \equiv (z_0(t), \dots, z_n(t))$, $\mathcal{G}'_t \equiv \sigma(z_{n+1}(s) : 0 \leq s \leq t)$ を定義し, 経済環境の信用リスク全般への影響を表す変数として新たな正值確率変数 ν (multiplier と呼ぶ) を導入して, $\mathcal{G} \equiv \mathcal{G}' \vee \sigma(\nu)$, すなわち任意の $t \in [0, T^*]$ に対して $\mathcal{G}_t \equiv \mathcal{G}'_t \vee \sigma(\nu)$ が成り立つとする.

本稿では取り扱い易さを考えて, デフォルト時刻 τ_j が P 下でも \tilde{P} 下でも \mathcal{G} -条件付独立とな

るようにモデルを構築する。また、 G_t が与えられたとき、まずは (i) G'_t に含まれる情報から、 ν に関して平均化された G' -条件付デフォルト確率を確定し、(ii) その確率と ν の実現値から (最も詳細に条件付けされた) G -条件付デフォルト確率を確定する、という二段階構造をとる。この構造をとることで、第一段階の部分のモデルパラメータは既知の手法で推定できるので、推定上の課題は第二段階のみに限定できる。

3.1.1 第一段階のモデリング

第一段階はフィルトレーション $G' \equiv (G'_t)_{t \geq 0}$ 下のモデリングで、まずは以下を仮定する。

仮定 1. $r(t)$, $h_j(t)$, $j = 1, \dots, n$ は G'_t -可予測過程とする。

仮定 2. $d\tilde{P}/dP$ は ν に依存しない。また、 $\beta_j(t)$, $j = 0, \dots, n$, $\kappa_j(t)$, $j = 1, \dots, n$ は G'_t -可予測過程とする。

一般に $h_j(t)$ は \mathcal{F}_t -可予測過程であり、Davis and Lo (2001) など扱われているデフォルト伝播モデルでは、 $h_j(t)$ が他資産のデフォルト時刻 τ_k , $k \neq j$ に依存する形で表現される。一方、仮定 1 では、観測確率 P 下のデフォルト確率は経済状況などには依存するが他資産の状態には依存しないので、本稿のモデルは Davis and Lo (2001) のようなデフォルト伝播モデルと異なる。また、一般にラドンニコディム微分 $d\tilde{P}/dP$ は ν にも依存するので、密度過程 $\rho(t)$ は (3.4) と異なる表現になるが、仮定 2 の前半より、 $\rho(t)$ は (3.4) で与えられる。さらに、仮定 2 の後半で $\kappa_j(t)$ を G'_t -可予測としたので、 $\tilde{h}_j(t)$ も G'_t -可予測となり、 \tilde{P} 下のデフォルト確率も経済状況などに依存するが他資産の状態には依存しない。

これらの仮定より、 G'_t -可予測過程 $h_j(t)$, $\tilde{h}_j(t)$ に対して (3.2) と (3.6) が成立し、 $h_j(t)$ と $\tilde{h}_j(t)$ はそれぞれ P 下と \tilde{P} 下で ν に関して平均化されたデフォルト強度とみなすことができ、 G'_u , $0 \leq u$ を与えたときの P 下の条件付デフォルト確率が得られる。具体的には、時刻 t までの G' -条件付デフォルト確率を $F_j(t|G'_u) \equiv P\{\tau_j \leq t|G'_u\}$ で定義すると、 G'_u , $0 \leq t \leq u$ が与えられたとき、

$$(3.7) \quad F_j(t|G'_u) = 1 - P\{\tau_j > t|G'_u\} = 1 - \exp\left\{-\int_0^t h_j(w)dw\right\}$$

となり、 $\{\tau_j > s\}$ の下で G'_u が与えられたときの時刻 s における時刻 t までの G' -条件付フォワードデフォルト確率を $F_j(s, t|G'_u) \equiv P\{\tau_j \leq t|\tau_j > s, G'_u\}$, $0 \leq s \leq t \leq u$ で定義すると、

$$(3.8) \quad F_j(s, t|G'_u) = \frac{P\{s < \tau_j \leq t|G'_u\}}{P\{\tau_j > s|G'_u\}} = 1 - \exp\left\{-\int_s^t h_j(w)dw\right\}$$

となる。一方、 G'_u , $0 \leq u \leq t$ が与えられたときは、条件付期待値の連鎖公式より、

$$(3.9) \quad \begin{aligned} F_j(t|G'_u) &= P\{\tau_j \leq t|G'_u\} = E[P\{\tau_j \leq t|G'_t\}|G'_u] \\ &= 1 - \exp\left\{-\int_0^u h_j(w)dw\right\} E\left[\exp\left\{-\int_u^t h_j(w)dw\right\}\middle|G'_u\right] \end{aligned}$$

が得られる。同様に、 $\{\tau > s\}$ の下で G'_u , $0 \leq u \leq s \leq t$ が与えられたとき、

$$(3.10) \quad \begin{aligned} F_j(s, t|G'_u) &= \frac{P\{s < \tau_j \leq t|G'_u\}}{P\{\tau_j > s|G'_u\}} = \frac{E[P\{s < \tau_j \leq t|G'_t\}|G'_u]}{E[P\{\tau_j > s|G'_t\}|G'_u]} \\ &= 1 - \frac{E\left[\exp\left\{-\int_u^t h_j(w)dw\right\}\middle|G'_u\right]}{E\left[\exp\left\{-\int_u^s h_j(w)dw\right\}\middle|G'_u\right]} \end{aligned}$$

が得られる。リスク中立確率 \tilde{P} 下でも, $\tilde{F}_j(t|\mathcal{G}'_u)$, $\tilde{F}_j(s,t|\mathcal{G}'_u)$ を同様に定義すると (3.7)–(3.11) に対応する式が得られる。

もう少し具体的にモデルを考える。観測確率 P 下で, $r(t)$ は Hull-White モデル (Hull and White, 1990)

$$(3.11) \quad dr(t) = (b_0(t) - a_0r(t))dt + \sigma_0dz_0(t), \quad t \geq 0$$

に従うと仮定する。ここで, a_0, σ_0 は正定数, $b_0(t)$ は時刻 t の確定関数である。また, $h_j(t)$ も Hull-White モデル

$$(3.12) \quad dh_j(t) = (b_j(t) - a_jh_j(t))dt + \sigma_jdz_j(t), \quad t \geq 0, \quad j = 1, \dots, n$$

に従い, a_j, σ_j は正定数, $b_j(t)$ は時刻 t の確定関数とする。(3.12)は, 将来の $h_j(t)$ が正の確率で負値になるので本来は強度モデルとして不適切であるが, 解析的取り扱いが容易なため, また, σ_j が小さければモデルの問題点が目立たないので, 理論や実務でしばしば使われる。さらに, $l_j(t) \equiv \kappa_j(t)h_j(t)$ をリスクプレミアム調整率と呼び,

$$(3.13) \quad \tilde{h}_j(t) = h_j(t) + l_j(t), \quad 0 \leq t \leq T^*, \quad j = 1, \dots, n$$

と表現する。一般に $l_j(t), j = 1, \dots, n$ は \mathcal{F}_t -可予測過程, 仮定 1 と仮定 2 を適用するならば \mathcal{G}'_t -可予測過程になるが, カリブレーションを容易にするため, Kijima and Muromachi (2000) で使われた次の仮定を用いる。

仮定 3. $\beta_0(t), l_j(t), j = 1, \dots, n$ は時刻 t の確定関数で, $\beta_j(t) = 0, j = 1, \dots, n$ とする。さらに, $l_j(t), j = 1, \dots, n$ は t で微分可能とする。

信用リスクの理論では $\tilde{h}_j(t) = (1 + \kappa_j(t))h_j(t)$ を使う方が標準的であるが, 実務では強度の差 $\tilde{h}_j(t) - h_j(t)$ を利回り格差 (スプレッド) と同形式で表現する方が馴染みやすい。それが (3.13) の表現である。このとき, $r(t)$ は \tilde{P} 下で Hull-White モデル

$$(3.14) \quad dr(t) = (\phi_0(t) - a_0r(t))dt + \sigma_0d\tilde{z}_0(t), \quad 0 \leq t \leq T^*$$

に従い, $\tilde{h}_j(t)$ も Hull-White モデル

$$d\tilde{h}_j(t) = (\phi_j(t) - a_j\tilde{h}_j(t))dt + \sigma_jd\tilde{z}_j(t), \quad 0 \leq t \leq T^*, \quad j = 1, \dots, n$$

に従う。仮定 3 の後半より, $\phi_0(t) = b_0(t) - \beta_0(t)\sigma_0$ と $\phi_j(t) = b_j(t) + a_jl_j(t) + dl_j(t)/dt$ は時刻 t の確定的な関数である。(3.14)より, 時刻 t における満期 $\tau, \tau > t$ の割引債価格は,

$$(3.15) \quad v_0(t, \tau) = \tilde{E} \left[\exp \left\{ - \int_t^\tau r(s)ds \right\} \middle| \mathcal{G}'_t \right] = A_0(t, \tau)e^{-B(a_0, \tau-t)r(t)}$$

で与えられ, また, \mathcal{G}'_t が与えられたときの条件付生存確率は, $\{\tau_j > t\}$ の下で,

$$(3.16) \quad \tilde{P}\{\tau_j > \tau | \mathcal{G}'_t\} = \tilde{E} \left[\exp \left\{ - \int_t^\tau \tilde{h}_j(s)ds \right\} \middle| \mathcal{G}'_t \right] = A_j(t, \tau)e^{-B(a_j, \tau-t)\tilde{h}_j(t)}$$

で与えられる。ただし, $\tilde{E}[\cdot]$ は \tilde{P} 下の条件付期待値演算子で, $j = 0, \dots, n$ に対して,

$$A_j(t, \tau) \equiv \exp \left\{ \frac{\sigma_j^2(\tau - t - 2B(a_j, \tau - t) + B(2a_j, \tau - t))}{2a_j^2} - \int_t^\tau \phi_j(u)B(a_j, \tau - u)du \right\},$$

$$B(a, t) \equiv \frac{1 - e^{-at}}{a}$$

である．観測確率 P 下でも (3.16) に対応する式が成り立つ．

3.1.2 第二段階のモデリング

第二段階はフィルトレーション $\mathcal{G} \equiv (\mathcal{G}_t)_{t \geq 0}$ 下のモデリングである．まず， \mathcal{G} -条件付デフォルト確率に関しても， \mathcal{G}' -条件付デフォルト確率と同様の式が成り立つ． \mathcal{G} -条件付デフォルト確率は $F_j(t|\mathcal{G}_u) \equiv P\{\tau_j \leq t|\mathcal{G}_u\}$, $0 \leq t \leq u$, $\{\tau_j > s\}$ の下で \mathcal{G}_u が与えられたときの時刻 s における時刻 t までの \mathcal{G} -条件付フォワードデフォルト確率は

$$(3.17) \quad F_j(s, t|\mathcal{G}_u) \equiv P\{\tau_j \leq t|\tau_j > s, \mathcal{G}_u\} = \frac{P\{s < \tau_j \leq t|\mathcal{G}_u\}}{P\{\tau_j > s|\mathcal{G}_u\}}, \quad 0 \leq s \leq t \leq u,$$

である．一方， \mathcal{G}_u , $0 \leq u \leq t$ が与えられたときは，条件付期待値の連鎖公式より，

$$(3.18) \quad F_j(t|\mathcal{G}_u) = P\{\tau_j \leq t|\mathcal{G}_u\} = E[P\{\tau_j \leq t|\mathcal{G}_t\}|\mathcal{G}_u],$$

$\{\tau > s\}$ の下で \mathcal{G}_u , $0 \leq u \leq s \leq t$ が与えられたとき，

$$(3.19) \quad F_j(s, t|\mathcal{G}_u) = \frac{P\{s < \tau_j \leq t|\mathcal{G}_u\}}{P\{\tau_j > s|\mathcal{G}_u\}} = \frac{E[P\{s < \tau_j \leq t|\mathcal{G}_t\}|\mathcal{G}_u]}{E[P\{\tau_j > s|\mathcal{G}_t\}|\mathcal{G}_u]}$$

となる．リスク中立確率 \tilde{P} の下でも (3.17)–(3.19) と同様の式が成り立つ．

確率変数 ν に関しては以下を仮定する．

仮定 4. ν は \mathcal{G}' および \mathcal{H} と独立な確率変数で， ν の分布は確率測度 P や \tilde{P} に依存しない．

仮定 5. $\{\tau_j > t\}$ の下で， \mathcal{G} -条件付フォワードデフォルト確率は，

$$(3.20) \quad F_j(t, \tau|\mathcal{G}_u) = F_j(t, \tau|\mathcal{G}'_u \vee \sigma(\nu)) = \nu F_j(t, \tau|\mathcal{G}'_u), \quad 0 \leq t < \tau \leq u, \quad j = 1, \dots, n$$

$$(3.21) \quad \tilde{F}_j(t, \tau|\mathcal{G}_u) = \tilde{F}_j(t, \tau|\mathcal{G}'_u \vee \sigma(\nu)) = \nu \tilde{F}_j(t, \tau|\mathcal{G}'_u), \quad 0 \leq t \leq \tau \leq u \leq T^*, \quad j = 1, \dots, n$$

で与えられる．

P 下および \tilde{P} 下の \mathcal{G} -条件付フォワードデフォルト確率は ν の実現値に応じて一律に ν 倍になるので， ν は経済状態の悪化具合と解釈できる．上述の ν に関する仮定には観測や理論による裏付けはないが，もしもこれらの仮定を認めれば， \tilde{P} 下の分布は CDO などの市場価格から推定可能であり，本稿のようにリスク計測の議論へと繋げることができる．もしも仮定 4 が成り立てば， $d\tilde{P}/dP$ は ν に依存しないので仮定 2 の前半は不要である．仮定 4 の前半部の ν と \mathcal{G}' の独立性，特に ν が金利と独立という仮定の緩和や，仮定 4 を外した一般的な場合における ν の分布の推定方法の確立は今後の課題である．なお，確率変数 ν を資産 j ごとに ν_j ，あるいは確率過程 $\nu(t)$ または $\nu_j(t)$ としなかったのは，観測データからの推定をますます困難にすると考えたからである．

以下では， ν は離散分布に従うものとして，その確率関数を $\eta_i \equiv P\{\nu = \nu_i\}$, $i = 1, \dots, K$ と書く． ν とその分布は， $\eta_i > 0$, $i = 1, \dots, K$ ，及び

$$(3.22) \quad \sum_{i=1}^K \eta_i = 1, \quad E[\nu] = \sum_{i=1}^K \nu_i \eta_i = 1,$$

$$0 \leq \nu \leq \frac{1}{\max_{\mathcal{G}'_{\tau, j}} F_j(0, \tau|\mathcal{G}'_{\tau})},$$

を満たさなければならない．

3.1.3 同時分布のモデリング

3.1.2 節まででデフォルト時刻 τ_j の周辺分布をモデル化した。 \mathcal{G} が与えられたときの τ_j の条件付周辺分布は具体的に与えられるので、 \mathcal{G} に関してその期待値をとることで τ_j の無条件周辺分布が得られる。残るは τ_j の同時分布である。Sklar の定理によると、同時分布は周辺分布関数とコピュラで表現できる。本稿のようにサブフィルトレーション・アプローチをとる場合は、 \mathcal{G} が与えられたときの $\tau_j, j = 1, \dots, n$ の条件付同時分布の依存関係を示す条件付コピュラを用いるのが適している。Kijima and Muromachi (2000) ではデフォルト時刻は条件付独立であったが、これはデフォルト時刻の条件付コピュラを積コピュラ(独立コピュラともいう)としたことに相当する。理論上は任意のコピュラを条件付コピュラとして使用できるので、彼らのフレームワークには条件付コピュラを選択という拡張が可能である。例えば、条件付コピュラに共単調コピュラを選択すれば $\tau_j, j = 1, \dots, n$ の依存関係は最も強くなることが期待できる。また、 \mathcal{G} に関する期待値計算を行って得られる $\tau_j, j = 1, \dots, n$ の同時分布の依存関係を表すのが通常のコピュラ(無条件コピュラ)なので、条件付コピュラと無条件コピュラは一致するとは限らない。例えば Kijima and Muromachi (2000) では、条件付コピュラとして積コピュラを選択しているが、無条件同時分布には依存性が生じているので無条件コピュラは積コピュラではない。

さらに、測度変換を定めるラドンニコディム微分は $d\tilde{P}/dP = \rho(T^*)$ で、密度過程 $\rho(t)$ は一般に \mathcal{F}_t -適合過程であるので、 P 下ではデフォルト伝播を考慮しないモデルであっても、 \tilde{P} 下ではデフォルト伝播を考慮するモデルになることもあり(Kusuoka, 1999 を参照)、また、その逆もありえる。このようなことは、 \mathcal{G} に含まれない確率過程・確率変数が密度過程 $\rho(t)$ に含まれる場合に発生しうる。そしてこれらのことは、一般に、 P 下の条件付コピュラと \tilde{P} 下の条件付コピュラが一致するとは限らないことを示している。

このような視点からみると、 P 下でも \tilde{P} 下でも条件付独立な Kijima and Muromachi (2000) のモデルはどちらの確率下でも条件付コピュラに積コピュラを用いた例である。これは実務的な取り扱い易さ、パラメータの推定し易さの点からも支持できる提案だが、それがどのようなモデルであるかは明言されていなかった。本稿では、測度変換で条件付独立性が保存されるための条件を具体的に挙げ、その意味を説明し、3.1.2 節までで述べてきたモデルに適用してリスク計測モデルへと拡張する。

仮定 6. 観測確率 P の下でデフォルト時刻は \mathcal{G} -条件付独立である。

このとき、 \mathcal{G}_T が与えられたときのデフォルト時刻の条件付同時分布は、

$$(3.23) \quad P\{\tau_1 \leq t_1, \dots, \tau_n \leq t_n | \mathcal{G}_T\} = \prod_{j=1}^n P\{\tau_j \leq t_j | \mathcal{G}_T\}, \quad 0 \leq \max_j t_j \leq T,$$

与えられ、(3.23) に条件付期待値の連鎖公式を用いると、

$$(3.24) \quad P\{\tau_1 \leq t_1, \dots, \tau_n \leq t_n\} = E \left[\prod_{j=1}^n P\{\tau_j \leq t_j | \mathcal{G}_T\} \right], \quad 0 \leq \max_j t_j \leq T,$$

が得られる。ここで、 $P\{\tau_j \leq t_j | \mathcal{G}_T\} = F_j(0, t_j | \mathcal{G}_T)$ である。

定理 3.1. 仮定 1, 仮定 2, 仮定 6 が成り立つとき、デフォルト時刻はリスク中立確率 \tilde{P} の下で \mathcal{G} -条件付独立である。

証明は付録 5 を参照されたい。定理 3.1 より、 \tilde{P} の下でも (3.23)–(3.24) と同様の式が成り立つので、 \tilde{P} 下のデフォルト時刻の同時分布は P 下と同様に構築できる。

定理 3.1 は測度変換で条件付独立性が保存されるための十分条件である。この条件はデフォルト強度と測度変換がデフォルト時刻に依存しないことを意味するので、大雑把に言えば、この条件を満たすのはどちらの確率下でもデフォルト伝播のような事象を考慮しないモデルである。なお、本稿の証明では ν の存在を無視する形で示したが、その理由は容易に計算可能なモデルを提供するためである。ラドンニコディム微分が ν に依存しても、 $\sigma(\nu)$ が \mathcal{G} に含まれていれば測度変換しても条件付独立性は保存される。しかし、測度変換前後の ν の分布に依存することになるので、本稿よりも複雑なモデルになる。

2 節の最後でも述べたように、プライシングモデルである Hull and White (2006) のモデルと、リスク計測モデルである本稿のモデルは、前者のデフォルト強度 λ と後者の multiplier ν がともにデフォルト確率を変動させる確率変数であるという点では似ているが、前者は \tilde{P} 下のみのモデルであるのに対し、後者は二つの確率測度 P と \tilde{P} にまたがるモデルであり、モデルの構成は全く異なる。また、デフォルト確率の表現の柔軟性も大きく異なる。Hull and White (2006) は全資産でデフォルト強度と分布 $(\lambda_k, \pi_k)_{k=1, \dots, L}$ を共通としたが、実際には個々に異なるはずである。一方、本稿のモデルでは \mathcal{G}' -条件付デフォルト確率と multiplier ν を使うことで資産によるデフォルト強度の違いに対応できる。しかし、その代償は 2 つ生じる。一つは条件 (3.22) の右側の不等号の追加だが、実務的には $\nu F_j(0, T | \mathcal{G}'_T)$ の代わりに $\min(\nu F_j(0, T | \mathcal{G}'_T), 1)$ を使えばよいのであまり問題はない。もう一つは信用リスクの標準的な数式との関係である。multiplier ν を (3.20) や (3.21) のようにデフォルト確率に掛ける形でなく、デフォルト強度に掛ける形で導入すれば、信用リスクの標準理論でデフォルト強度 $h_j(t)$ を $\nu h_j(t)$ に置換するだけなので数理的には理解し易い。しかしその場合、 $\nu h_j(t)$ の ν に関する平均は $h_j(t)$ とは一致せず、また、デフォルト確率は ν に指数関数的に依存するようになるので、強度が資産ごとに異なる場合、 ν の分布と $h_j(t)$, $j = 1, \dots, n$ の同時推定が必要になるかもしれない等、パラメータ推定が困難になることが懸念される。しかし、どちらが推定しやすいか、もっともらしいモデルになるかは実際に比較してみなければわからない。なお、Hull and White (2010) は強度が資産により異なる場合のモデルも提案しているが、パラメトリックで制約が強く、期間構造も表現しづらいので、本稿では採用しなかった。

3.2 デフォルトリスクのある債券の評価

企業 j が発行した満期 τ , $t < \tau < T^*$ の割引社債に関しては Jarrow and Turnbull (1995) の設定、すなわち、企業 j が τ までにデフォルトしないときは満期 τ に 1 円、デフォルトした時は満期 τ に δ_j 円 (δ_j は一定) を受け取るとし、さらに簡単化のために次を仮定する。

仮定 7. $c_{0j} = 0$, $j = 1, \dots, n$, すなわち $r(t)$ と $h_j(t)$, $j = 1, \dots, n$ は互いに独立である。

T をリスクホライズンとする。 $\mathcal{F}_T = \mathcal{G}_T \vee \mathcal{H}_T$, $0 \leq T \leq T^*$ が与えられたとき、資産 j を満期 τ のデフォルトリスクのある割引社債とすると、その時刻 T における価格は

$$(3.25) \quad v_j(T, \tau) = v_0(T, \tau) [\delta_j + 1_{\{\tau_j > T\}} (1 - \delta_j) \tilde{P}\{\tau_j > \tau | \mathcal{G}_T\}]$$

で与えられる。ここで $v_0(T, \tau)$ には (3.15) を用い、 $\{\tau_j > T\}$ の下で

$$(3.26) \quad \tilde{P}\{\tau_j > \tau | \mathcal{G}_T\} = 1 - \tilde{F}_j(T, \tau | \mathcal{G}_T) = 1 - \nu \left(1 - \tilde{E} \left[\exp \left\{ - \int_T^\tau \tilde{h}_j(s) ds \right\} \middle| \mathcal{G}'_T \right] \right)$$

は (3.16) と同様に算出できて、 $\tilde{h}_j(T) = h_j(T) + \ell_j(T)$ の関数となる。

リスク計測には観測確率 P 下における時刻 T までのデフォルト損失額の評価も必要である。上記の仮定の下で \mathcal{G}_T が与えられたとき、資産 j のデフォルトは確率

$$(3.27) \quad F_j(0, T | \mathcal{G}_T) = \nu F_j(0, T | \mathcal{G}'_T) = \nu \left(1 - \exp \left\{ - \int_0^T h_j(s) ds \right\} \right)$$

で互いに独立に発生する。この確率は ν とサンプルパス $(h_j(s))_{0 \leq s \leq T}$ に依存する。

クーポン債は、同じ企業が発行した割引社債のポートフォリオとみなして同様に考える。ただし、 T 時点の将来価値で考えるので、 T までに受け取ったクーポンも時刻 T の将来価値に換算する必要があるが、その評価方法の一例を 3.3 節で述べる。

3.3 CDO トランシェの評価

CDO トランシェの無裁定価格は Hull and White (2006) の方法を参考に算出する。CDO の仕組みやトランシェの無裁定価格の導出などに関しては室町 (2007) を参照されたい。

想定元本が定数 g 、回収率が定数 δ の相異なる n 個の負債を参照債務とする CDS (Credit Default Swap) からなるポートフォリオをアセットプールとする満期 T_M の CDO を考える。アセットプール全体を 1 としたとき、アタッチメントポイント a_L 、デタッチメントポイント a_H ($0 \leq a_L < a_H \leq 1$) の損失をカバーするトランシェを考える。アセットプールに j 番目のクレジットイベントが発生したときの残存想定元本 $A_j(a_L, a_H)$ は、

$$A_j(a_L, a_H) = \begin{cases} (a_H - a_L)ng, & j < m(n_L) \\ a_Hng - j(1 - \delta)g, & m(n_L) \leq j < m(n_H) \\ 0, & m(n_H) \leq j \end{cases}$$

と書ける。ここで、 $n_L = a_L n / (1 - \delta)$ 、 $n_H = a_H n / (1 - \delta)$ 、 $m(x)$ は x を越える最小の自然数である。当該トランシェの無裁定価格はリスク中立確率 \tilde{P} の下における将来キャッシュフローの割引現在価値の期待値の総和で与えられ、将来キャッシュフローは (1) プレミアムレグ、(2) デフォルトレグ、(3) 経過利息からなる。 $A(t)$ を時刻 t における当該トランシェの残存想定元本、 s をプレミアム (またはスプレッド)、 $t_i, i = 1, \dots, I$ を定期的な利払日、 $t_I = T_M$ を満期とし、 $t_0 = 0$ とする。簡単化のため、クレジットイベントは利払日間の中央の日 (期央と呼ぶ) のみ発生し、デフォルトレグも経過利息も即日支払われると仮定する。このとき、現時点 $t = 0$ におけるトランシェ価格は、

$$V = s \sum_{i=1}^I (t_i - t_{i-1}) v_0(0, t_i) \tilde{E}[A(t_i)] + \sum_{i=1}^I \left[\frac{s(t_i - t_{i-1})}{2} - 1 \right] v_0 \left(0, \frac{t_i + t_{i-1}}{2} \right) (\tilde{E}[A(t_{i-1})] - \tilde{E}[A(t_i)])$$

で与えられる。詳細は室町 (2007) を参照されたい。トランシェ価格は利払日の残存想定元本の期待値 $\tilde{E}[A(t_i)]$ 、 $i = 1, \dots, I$ で表現され、 $\tilde{E}[A(t_i)]$ は時刻 t_i におけるアセットプールの損失額分布から算出できる。損失額分布に関しては、例えば Hull and White (2004) が提案したバケツ法 (bucketing method) を用いれば、十分な精度の分布が容易に算出できる。

トランシェの現在価値と時点 T における将来価値の計算には相違点が 3 つある。第一はアセットプールの損失額分布の計算法である。現在価格では将来の ν の分布を考慮して計算を行う。具体的には、 $\nu = \nu_i$ により与えられるフォワードデフォルト確率に応じて算出されるアセットプールの条件付損失額分布が確率 η_i で発生するとして、 η_i による加重平均として得られる無条件損失額分布から価格を計算する。将来価値では、ある ν_i が実現したときの条件付損失額分布が価格付けに使われる。第二はアセットプールである CDS ポートフォリオの構成である。現在価格の計算では当初設定した CDS ポートフォリオそのものを使用するが、将来価値の計算では既にデフォルトした CDS を除外した CDS ポートフォリオを使用する。第三は、将来価値では時刻 T までにやり取りするキャッシュフローの価値も含める点である。具体

的には、(a)時刻 T までのデフォルトレグ、(b)時刻 T までのプレミアムレグ、(c)時刻 T までの経過利息で、受取時点から時刻 T までデフォルトフリー金利で運用されると考えれば、それらの時刻 T における価値は、それぞれ

$$\sum_{0 < \frac{t_{i-1}+t_i}{2} \leq T} \frac{A(t_{i-1})-A(t_i)}{v_0\left(\frac{t_{i-1}+t_i}{2}, T\right)}, \quad \sum_{0 < t_i \leq T} \frac{s(t_i-t_{i-1})A(t_i)}{v_0(t_i, T)}, \quad \sum_{0 < \frac{t_{i-1}+t_i}{2} \leq T} \frac{s(t_i-t_{i-1})(A(t_{i-1})-A(t_i))}{2v_0\left(\frac{t_{i-1}+t_i}{2}, T\right)}$$

で与えられる。本稿では、時刻 T におけるトランシェの価値に、時刻 T までに受け取るキャッシュフローの T 時点の価値を加え、時刻 T までのデフォルトレグの T 時点の価値を引いた値を、 T 時点のトランシェの将来累積価値 (Future Cumulative Value) と定義する。なお、金融危機前の取引慣行にあわせてエクイティのみアップフロント払いとする。

3.4 モンテカルロシミュレーションによるリスク計測の手順

上述した確率モデルをもとに、リスクホライズン T におけるリスク量 (VaR や ES など) をモンテカルロシミュレーションで推定する手順は以下である。

(1) 観測確率 P 下で、(3.1) と (3.3) から乱数を用いて $(r(s), h_j(s), j = 1, \dots, n)_{0 \leq s \leq T}$ のサンプルパスを、確率関数 η から ν を生成する。

(2) サンプルパスごとに時刻 T までのデフォルトを評価する。時刻 T の情報のみわかればよい場合は、資産 j ごとに (3.27) から $F_j(0, T|\mathcal{G}_T)$ を算出し、乱数を使ってデフォルトか否かを独立に判定する。デフォルト時刻の情報が欲しい場合は、この判定を短い時間区間ごとに時刻 T まで行い、おおよそのデフォルト時刻を算出する。

(3) 各資産の時刻 T における将来価値を算出する。

(a) 割引社債ポートフォリオの場合、(3.25)–(3.26) を用いて時刻 T の将来価値を算出する。将来価値は $(r(T), h_j(T), \nu, \text{時刻 } T \text{ でデフォルトか否か})$ に依存する。

(b) CDO トランシェの場合、3.3 節で述べた方法で将来累積価値を算出する。

(4) 時刻 T における全資産の将来価値を合算し、ポートフォリオの将来価値を算出する。

(5) 十分な数のシナリオが得られたら、この手順を終了して統計処理に移り、リスク量を算出する。シナリオ数が不十分なときは手順(1)に戻る。

4. 数値例

本節では、CDO および割引社債ポートフォリオのリスク計測の事例を示す。なお、計算の単純化のため、本節では \mathcal{G}' -条件付デフォルト強度 $h_j(t), j = 1, \dots, n$ は時刻 t の確定的な関数と仮定する。この仮定は信用リスク評価の実務では標準的で、(3.12) で $\sigma_j = 0$ とおくことに相当する。このとき、 $F_j(t, \tau|\mathcal{G}'_t), \tilde{F}_j(t, \tau|\mathcal{G}'_t)$ などの確率は確定関数となり、

$$F_j(t, \tau|\mathcal{G}'_t) = \nu \frac{F_j(\tau) - F_j(t)}{1 - \nu F_j(t)} \quad \text{under } P, \quad \tilde{F}_j(t, \tau|\mathcal{G}'_t) = \nu \frac{\tilde{F}_j(\tau) - \tilde{F}_j(t)}{1 - \nu \tilde{F}_j(t)} \quad \text{under } \tilde{P}.$$

のように簡単に計算できる。

4.1 合成 CDO のリスク計測事例

4.1.1 設定

以下で述べるパラメータは、モデルの特徴が明確に現れやすい金融危機前の数値を参考に設定した。執筆時 (2019 年) のような超低金利環境にはそぐわないことに注意されたい。

表 1. multiplier ν の分布.

ν	0.1	0.2	0.3	0.5	0.6	0.7	0.8	0.9	1.0	1.1	1.2	1.3	1.4	1.5
η (%)	0.01	0.015	0.04	0.055	0.08	0.13	0.15	0.175	0.12	0.09	0.04	0.01	0.005	0.001

ν	1.7	1.8	1.9	2.0	2.1	2.2	2.3	2.4	4.1	4.2	4.3	4.4	4.5	
η (%)	0.002	0.004	0.007	0.01	0.008	0.004	0.003	0.001	0.003	0.008	0.013	0.011	0.005	

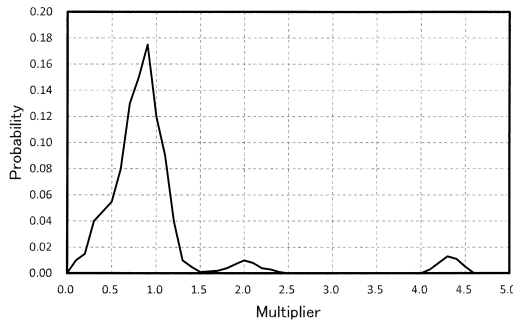


図 1. multiplier ν の分布.

金利モデル(3.11)に関しては、 $b_0(t)$ を定数とし、 $b_0(t)/a_0 = 0.03$ (平均回帰水準 3.0%)、 $a_0 = 0.1$ 、 $\sigma_0 = 0.01$ (年率 1%)とする。CDO の満期は $\tau = 5$ 年とし、想定元本 10、回収率 40% の 125 個の CDS をアセットプールとする。CDS の格付けは A1, A2, B, C1, C2, C3 の 6 段階で、観測確率 P 下のデフォルト確率はそれぞれ年率で 0.1%, 0.2%, 0.5%, 1.0%, 2.0%, 3.0% とし、A1 格が 40 資産、A2 格が 40 資産、B 格が 30 資産で、C1, C2, C3 はそれぞれ 5 資産ずつとする。現時点の格付別フォワードレートカーブはすべてフラットで、デフォルトフリー金利は 3.0% とし、上記格付けの順にそれぞれ 3.1%, 3.2%, 3.5%, 4.0%, 5.0%, 6.0% とする。格付別のデフォルト確率の期間構造はフォワードレートカーブと回収率から算出する。トランシェは 6 つで、デタッチメントポイントをそれぞれ 3%, 6%, 9%, 12%, 22%, 100% とする。CDO の総想定元本は $125 \times 10 = 1250$ で、1% は 12.5 に相当し、例えばエクイティトランシェ [0, 3%](トランシェをアタッチメントポイント a_p とデタッチメントポイント d_p で $[a_p, d_p]$ と表現する)の想定元本は 37.5 に相当する。

表 1 と図 1 に、本稿で使用する multiplier ν の分布を示す。図 1 の横軸は ν 、縦軸は発生確率である。モデルの特徴を明確に示すため、右裾を長く設定し、例えば $P\{\nu > 3\} = 4.0\%$ とした。この分布は実際の CDO 価格からの推定値ではないが、Hull and White (2006) や Brigo et al. (2010) による推定結果も同様に裾の長い、多峰形の分布であった。各トランシェのスプレッドは現時点 ($t = 0$) での契約価値がゼロになるフェアスプレッドに設定した。具体的には、エクイティから順に、40.7%, 169.0bp, 69.9bp, 21.6bp, 0.3bp である。なお、エクイティ [0, 3%] はアップフロント払い(ランニングスプレッドは年率 5%)とした。

リスクホライズンは $T = 1$ 年、シミュレーションは 50 万回とし、各トランシェの $T = 1$ での将来価値には将来累積価値を用いた。なお、モンテカルロシミュレーションによるサンプルパスでは各利払日(四半期毎)にシナリオを生成し、ある CDS がある時間区間でデフォルトが発生した場合、期央でデフォルトしたものとして評価した。

4.1.2 将来累積価値の分布とリスク量

図2と図3に、 $T=1$ における各トランシェの将来累積価値の分布関数を示す。横軸は将来累積価値、縦軸は分布関数である。将来累積価値はトランシェ購入者側からみた価値で、デフォルト発生件数が多いほど低くなり、 ν の値が高いほど将来のトランシェ価格が低下するので低くなる。これらの図より、多くのトランシェの分布関数は左裾が長く、10%以下の部分で左裾が長く伸び始め、特に5.0%以下で一段と伸びる。これらの結果は ν の分布を明確に反映し、大規模損失が低確率で発生することを示唆している。

各トランシェのリスク量を表2にまとめる。ここで、 $100\alpha\%$ -VaR ($0 < \alpha < 1$)は将来累積価値の平均値と $100(1-\alpha)$ パーセント点の差、 $100\alpha\%$ -ES ($0 < \alpha < 1$)は将来累積価値の平均値と $100(1-\alpha)$ パーセント点以下の部分の条件付期待値の差である。表2によると、 $[0, 3\%]$ と $[3, 6\%]$ (ジュニアメザニン)では95%-VaRが標準偏差の2倍以上と正規分布に比べて高い値をとる。一方、 $[6, 9\%]$ (シニアメザニン)、 $[9, 12\%]$ (シニア)、 $[12, 22\%]$ (ファーストスーパーシニア)の95%-VaRは標準偏差以下あるいは同程度であるが、97.5%-VaRと99%-VaRは急上昇し、それ以上の信頼水準ではあまり上昇しない。また、99%以上のVaRは $[0, 3\%]$ トランシェよりも $[3, 6\%]$ と $[6, 9\%]$ のトランシェの方が高く、大損失リスクがより高いことになるが、これは $[0, 3\%]$ トランシェのアップフロント払いがリスクを低減させるためである。金融危機以降、実

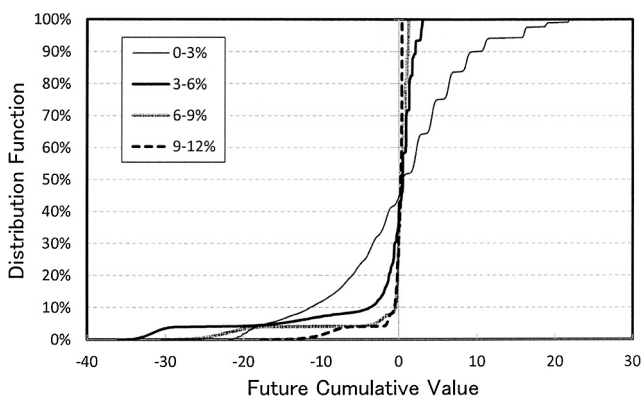


図2. 各トランシェの将来累積価値の分布関数(1).

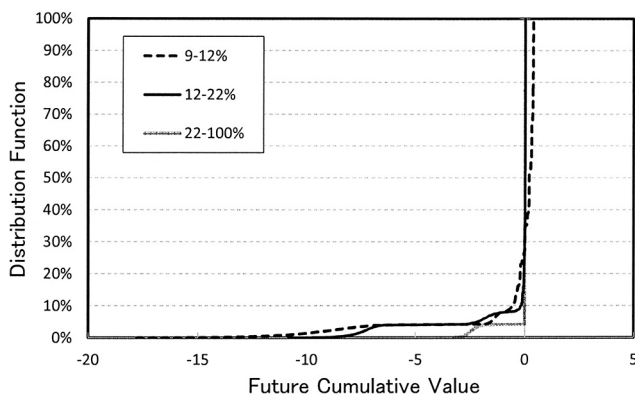


図3. 各トランシェの将来累積価値の分布関数(2).

表 2. 各トランシェの主なリスク量の推定値.

[attachment, detachment]	[0%, 3%]	[3%, 6%]	[6%, 9%]	[9%, 12%]	[12%, 22%]	
Initial Face Value	37.5	37.5	37.5	37.5	125	
Average	0.135	-1.468	-0.621	-0.302	-0.381	
Standard Deviation	8.507	6.970	4.584	1.995	1.523	
VaR	95.0%	17.224	14.181	2.433	1.232	1.796
	97.5%	19.611	29.772	20.894	8.386	6.848
	99.0%	20.550	31.268	23.551	10.448	7.493
	99.5%	21.013	31.961	24.891	11.708	7.939
	99.9%	21.877	33.157	27.364	14.398	8.996
ES	95.0%	19.578	27.731	18.319	7.787	6.186
	97.5%	20.554	31.450	23.699	10.464	7.540
	99.0%	21.210	32.144	25.272	12.170	8.143
	99.5%	21.511	32.696	26.378	13.335	8.595
	99.9%	21.935	33.721	28.415	15.792	9.668

表 3. 各トランシェの実際のデフォルト損失額の分布.

Loss	[0, 3%]	[3, 6%]	[6, 9%]
0	299,159	499,778	500,000
4.5	0	162	0
6	142,437	0	0
10.5	0	45	0
12	41,153	0	0
16.5	0	13	0
18	11,202	0	0
22.5	0	2	0
24	3,864	0	0
30	1,460	0	0
36	493	0	0
37.5	222	0	0

際の市場ではエクイティ以外のトランシェもアップフロント払いに変更されたが、それがリスクの低減に奏功したことがこの結果から推測される。

表 3 は、最も劣後する 3 トランシェの実際のデフォルトによる損失額のみ分布である。表 3 より、実際のデフォルト損失の分布の形状は将来累積価値の分布と全く異なり、実際のデフォルト損失はそれほど多くない。つまり、将来累積価値に見られる損失の大半はトランシェ価格の下落によるもので、それは特に [3, 6%] と [6, 9%] のメザニンで顕著である。このような傾向は金融危機時に CDO 市場で見られた現象と整合的である。

巨大損失発生事象は CDO の価格付けではアルマゲドンファクターと呼ばれてきた。本稿のモデルを用いれば、価格付けの世界のアルマゲドンファクターをリスク管理の世界の事象として議論できる。ただし、発生確率や損失額の規模はモデル構築上のさまざまな仮定に依存することに注意されたい。

4.2 割引社債ポートフォリオのリスク計測事例

4.2.1 設定

異なる企業が発行した 1,000 銘柄の割引社債を考える。すべて満期は 5 年、額面は 10、回収率は 40% で、1,000 銘柄のうち 400 銘柄が A2 格、300 銘柄が B 格、200 銘柄が C1 格、100 銘柄が C2 格とし、各格付けの特性は 4.1 節と同じとする。一方、本節では特に信用リスクを顕著に示すため σ_0 を 4.1 節の 1/10 にする。すなわち、(3.11) のパラメータは $b_0(t)/a_0 = 0.03$ 、 $a_0 = 0.1$ 、 $\sigma_0 = 0.001$ とおく。また、リスクホライズンは $T = 1$ 年とし、シミュレーションは 100 万回とする。

4.2.2 将来価値の分布とリスク量

図 4 はリスクホライズン $T = 1$ 年後の割引社債ポートフォリオの将来価値の分布(ヒストグラム)で、図 5 は分布関数で縦軸が確率、横軸が将来価値である。また、表 4 に標準偏差、VaR、ES などの主なリスク量をまとめる。ここでは $100\alpha\%$ -VaR ($0 < \alpha < 1$) は将来価値の平均値と $100(1 - \alpha)$ パーセント点の差、 $100\alpha\%$ -ES ($0 < \alpha < 1$) は将来価値の平均値と $100(1 - \alpha)$ パーセント点以下の部分の条件付期待値の差である。 $\sigma_0 = 0.001$ なのでこの数値例の金利リスクは小さく、幅広い分布が得られるのは専ら信用リスクの効果である。

図 4 は ν の分布(図 1)を反転させた形によく似ている。図 5 では、分布関数の左裾が 10% 以下で左に伸び、特に 4% 付近で更に伸びるが、これらはそれぞれ図 1 の 2.0 付近と 4.0 付近のピークに対応する。表 4 のリスク量もこれらを反映し、95% から 99% に見られる VaR の上昇

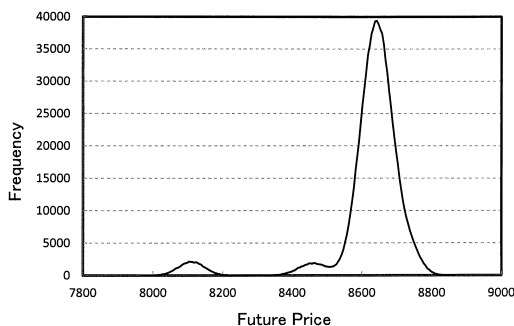


図 4. 割引社債ポートフォリオの将来価値の分布(ヒストグラム).

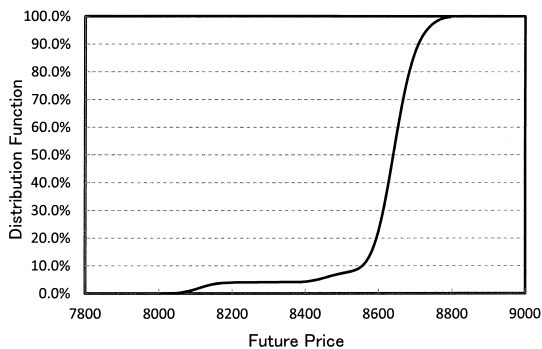


図 5. 割引社債ポートフォリオの将来価値の分布関数.

表 4. 割引社債ポートフォリオの主なリスク量の推定値.

Average		8,618.4
Standard Deviation		120.1
VaR	95.0%	186.5
	97.5%	496.5
	99.0%	533.7
	99.5%	552.0
	99.9%	581.9
ES	95.0%	450.0
	97.5%	531.3
	99.0%	556.3
	99.5%	570.3
	99.9%	596.7

は ν の 4.0 付近のピークの存在に対応する.

5. おわりに

本稿では、Hull and White (2006) のインプライドコピュラの論文の中で提示された、ノンパラメトリックな分布を持つデフォルト強度とデフォルト時刻の条件付独立性を用いた CDO のプライシングモデルを、Kijima and Muromachi (2000) の将来価値ベースのリスク計測のフレームワークに沿って拡張した新しいリスク計測モデルを提案した。純粋な統計学的モデルである既存のリスク計測モデルはヒストリカルデータしか使用しないが、このモデルではデリバティブの市場価格から抽出される情報も活用することで、市場参加者の将来の環境激変への畏怖をリスク計測に定量的に反映できる。具体的には、デフォルト確率が multiplier ν というノンパラメトリックな分布を持つ確率変数に依存すると考えて、その分布を CDO などの市場価格から推定することで、市場参加者のテイルリスクへの評価を ν の分布を通してリスク計測に反映できるようにしたところ、既存のリスク計測モデルでは把握しきれないテイルリスクを表現できることを示した。例えば、提案モデルで CDO のリスクを計測したところ、優先度の高いトランシェでも大規模損失が小さな確率で発生することが示唆された。また、大規模損失の内訳が実際のデフォルトによる損失ではなくトランシェ価格の暴落であるという、金融危機時の現実と整合的な結果が得られた。

しかし、モデルで用いる仮定には問題も残されている。特に、仮定 4 は P 下の ν の分布の推定が現状では難しいと考えて課しただけに過ぎない。また、CDO 価格から推定された ν の分布を別の市場、例えば社債ポートフォリオに対して適用することにも疑問は残る。CDO 市場と社債市場でテイルリスクの評価が共通であるという確証はないからである。

提案モデルの結果は多くの仮定、特に ν の分布に強く依存するものの、通常の統計学的モデルの結果に、市場価格からインプライされる極端に大きな損失発生事象を組み込んだ結果を出力可能である。価格に内在する情報をフォワードルッキングと考えれば、提案モデルはフォワードルッキングなリスク評価を行えるモデルであり、さらに、大損失事象を価格からインプライされるフォワードルッキングなストレスシナリオと解釈すれば、統計学的モデルにストレステストを組み込んだものと見做せるだろう。多くの課題は残るが、この種のリスク計測モデルの検討もリスク管理の高度化に向けた一つの方向性であろう。

謝 辞

本研究は JSPS 科研費基盤研究 (B) No.16H03123 の助成を受けたものである。本稿の最終版の作成にあたり、匿名の査読者と編集委員から多くの有益かつ好意的なコメントをいただいたことに厚く感謝する。

付録 A : 定理 3.1 の証明

簡単化のため資産数 $n = 2$ とする。観測確率 P 下でデフォルトフリーな短期金利 $r(t)$ は (3.1) に、デフォルト強度 $h_i(t)$, $i = 1, 2$ は (3.3) に従うとして、 $z_3(t) = (z_0(s), z_1(s), z_2(s))^T$ とし、仮定 1, 仮定 2, 仮定 6 の下で考える。まず、

$$M_i(t) = H_i(t) - \int_0^t (1 - H_i(s-))h_i(s)ds, \quad i = 1, 2,$$

は P 下の \mathcal{F} -マルチンゲールである。ラドンニコディム微分 $d\tilde{P}/dP$ は \mathcal{F}_{T^*} -可測な正値確率変数で、密度過程 $\rho(t)$ を局所有界な左極限右連続な確率過程とすると、 $\rho(t)$ は P -局所マルチンゲールであり、ある \mathcal{G}_t -可予測過程 $\beta_3(t) = (\beta_0(t), \beta_1(t), \beta_2(t))^T$ と $\kappa_i(t)$, $i = 1, 2$ を用いて、

$$(A.1) \quad \rho(t) = 1 + \int_0^t \rho(s-) \left(-\beta_3^\top(s)dz_3(s) + \sum_{i=1}^2 \kappa_i(s)dM_i(s) \right), \quad 0 \leq t \leq T^*,$$

と一意に表現できるとする。このとき (A.1) の解は、 $\varepsilon(\cdot)$ を Doléans-Dade の指数として、

$$(A.2) \quad \rho(t) = \rho_0(t)\rho_1(t)\rho_2(t),$$

$$(A.3) \quad \rho_0(t) \equiv \varepsilon \left(-\int_0^t \beta_3^\top(s)dz_3(s) \right) = \exp \left\{ -\int_0^t \beta_3^\top(s)dz_3(s) - \frac{1}{2} \int_0^t \beta_3^\top(s)\beta_3(s)ds \right\},$$

$$(A.4) \quad \rho_i(t) \equiv \varepsilon \left(\int_0^t \kappa_i(s)dM_i(s) \right)$$

で与えられる。ここで、 $\rho_i(t)$ は \mathcal{F}^i -マルチンゲールである。

$0 < T \leq T^*$ として、フィルトレーション $\mathcal{F}_t^i \equiv \mathcal{G}_T \vee \mathcal{H}_t^i$, $i = 1, 2$ と $\bar{\mathcal{F}}_t \equiv \mathcal{G}_T \vee \mathcal{H}_t^1 \vee \mathcal{H}_t^2$ を定義する。まず、 \mathcal{G}_T が与えられたとき、ベイズの定理と (A.2)-(A.4) より、

$$(A.5) \quad \tilde{P}\{\tau_1 > t | \mathcal{G}_T\} = \frac{E^P [1_{\{\tau_1 > t\}} \frac{d\tilde{P}}{dP} | \mathcal{G}_T]}{E^P [\frac{d\tilde{P}}{dP} | \mathcal{G}_T]} = \frac{E^P [1_{\{\tau_1 > t\}} \rho_1(T)\rho_2(T) | \mathcal{G}_T]}{E^P [\rho_1(T)\rho_2(T) | \mathcal{G}_T]}$$

となる。 τ_1 と τ_2 は \mathcal{G} -条件付独立なので、 $\rho_1(t)$ と $\rho_2(t)$ も \mathcal{G} -条件付独立である。さらに、 $\rho_i(t)$ は \mathcal{F}^i -マルチンゲールで $E^P[\rho_i(T) | \mathcal{G}_T] = E^P[\rho_i(T) | \mathcal{F}_0^i] = \rho_i(0) = 1$ となることから、

$$(A.6) \quad E^P[\rho_1(T)\rho_2(T) | \mathcal{G}_T] = E^P[\rho_1(T) | \mathcal{G}_T]E^P[\rho_2(T) | \mathcal{G}_T] = 1$$

である。また、 $\bar{\mathcal{F}}_t$ が与えられたとき、 τ_1 と τ_2 の \mathcal{G} -条件付独立性から、

$$(A.7) \quad E^P[\rho_1(T)\rho_2(T) | \bar{\mathcal{F}}_t] = E^P[\rho_1(T) | \bar{\mathcal{F}}_t]E^P[\rho_2(T) | \bar{\mathcal{F}}_t] = \rho_1(t \wedge \tau_1)\rho_2(t \wedge \tau_2)$$

が成立するので、

$$(A.8) \quad E^P[\rho_1(T)\rho_2(T) | \mathcal{F}_t^1] = E^P[E^P[\rho_1(T)\rho_2(T) | \bar{\mathcal{F}}_t] | \mathcal{F}_t^1] = \rho_1(t \wedge \tau_1)E^P[\rho_2(t \wedge \tau_2) | \mathcal{F}_t^1] \\ = \rho_1(t \wedge \tau_1)E^P[\rho_2(t \wedge \tau_2) | \mathcal{G}_T] = \rho_1(t \wedge \tau_1)$$

が得られ、同様に $E^P[\rho_1(T)\rho_2(T)|\mathcal{F}_t^2] = \rho_2(t \wedge \tau_2)$ も得られる。(A.5), (A.6), (A.8)より,

$$(A.9) \quad \begin{aligned} \tilde{P}\{\tau_1 > t|\mathcal{G}_T\} &= E^P[E^P[1_{\{\tau_1 > t\}}\rho_1(T)\rho_2(T)|\mathcal{F}_t^1]|\mathcal{G}_T] = E^P[1_{\{\tau_1 > t\}}\rho_1(t)|\mathcal{G}_T] \\ &= \bar{\rho}_1(t)E^P[1_{\{\tau_1 > t\}}|\mathcal{G}_T] = \bar{\rho}_1(t)P\{\tau_1 > t|\mathcal{G}_T\} \end{aligned}$$

と表現できる。ここで、 $\bar{\rho}_i(t)$, $i = 1, 2$ は $t \in [0, T]$ で $1_{\{\tau_i > t\}}\bar{\rho}_i(t) = 1_{\{\tau_i > t\}}\rho_i(t)$ を満たす \mathcal{G}_t -適
合過程である。同様にして、 $\tilde{P}\{\tau_2 > t|\mathcal{G}_T\} = \bar{\rho}_2(t)P\{\tau_2 > t|\mathcal{G}_T\}$ も得られる。

一方、ベイズの定理、(A.2)-(A.4), (A.6)-(A.7), P 下でのデフォルト時刻の \mathcal{G} -条件付独立性より、 $0 \leq u \leq t \leq T$ に対して、

$$(A.10) \quad \begin{aligned} \tilde{P}\{\tau_1 > t, \tau_2 > u|\mathcal{G}_T\} &= \frac{E^P[1_{\{\tau_1 > t, \tau_2 > u\}}\frac{d\tilde{P}}{dP}|\mathcal{G}_T]}{E^P[\frac{d\tilde{P}}{dP}|\mathcal{G}_T]} = E^P[1_{\{\tau_1 > t, \tau_2 > u\}}\rho_1(T)\rho_2(T)|\mathcal{G}_T] \\ &= E^P[E^P[1_{\{\tau_1 > t\}}1_{\{\tau_2 > u\}}\rho_1(T)\rho_2(T)|\bar{\mathcal{F}}_t]|\mathcal{G}_T] \\ &= E^P[1_{\{\tau_1 > t\}}1_{\{\tau_2 > u\}}\rho_1(t)\rho_2(t \wedge \tau_2)|\mathcal{G}_T] \\ &= E^P[1_{\{\tau_1 > t\}}\rho_1(t)|\mathcal{G}_T]E^P[1_{\{\tau_2 > u\}}\rho_2(t \wedge \tau_2)|\mathcal{G}_T] \end{aligned}$$

となる。 $\rho_i(t)$ の \mathcal{F}^i -マルチンゲール性と $\bar{\rho}_i(t)$ を用いると、

$$\begin{aligned} E^P[1_{\{\tau_2 > u\}}\rho_2(t \wedge \tau_2)|\mathcal{G}_T] &= E^P[E^P[1_{\{\tau_2 > u\}}\rho_2(t \wedge \tau_2)|\mathcal{F}_u^2]|\mathcal{G}_T] = E^P[1_{\{\tau_2 > u\}}\rho_2(u)|\mathcal{G}_T] \\ &= \bar{\rho}_2(u)E^P[1_{\{\tau_2 > u\}}|\mathcal{G}_T] = \bar{\rho}_2(u)P\{\tau_2 > u|\mathcal{G}_T\} \end{aligned}$$

となる。 $E^P[1_{\{\tau_1 > t\}}\rho_1(t)|\mathcal{G}_T]$ も同様に扱うと、(A.10)は、

$$(A.11) \quad \tilde{P}\{\tau_1 > t, \tau_2 > u|\mathcal{G}_T\} = \bar{\rho}_1(t)\bar{\rho}_2(u)P\{\tau_1 > t|\mathcal{G}_T\}P\{\tau_2 > u|\mathcal{G}_T\}$$

となるので、(A.9)と(A.11)より、

$$(A.12) \quad \tilde{P}\{\tau_1 > t, \tau_2 > u|\mathcal{G}_T\} = \tilde{P}\{\tau_1 > t|\mathcal{G}_T\}Q\{\tau_2 > u|\mathcal{G}_T\}$$

が得られる。同様にして $0 \leq t \leq u \leq T$ の場合も (A.12) が得られるので、 τ_1 と τ_2 は確率測度 \tilde{P} の下でも \mathcal{G} -条件付独立である。□

参 考 文 献

- Bielecki, T. R. and Rutkowski, M. (2002). *Credit Risk: Modelling, Valuation and Hedging*, Springer-Verlag, Berlin Heidelberg.
- Bielecki, T. R., Jeanblanc, M. and Rutkowski, M. (2009). *Credit Risk Modelling*, Osaka University CSFI Lecture Note Series 2, Osaka University Press, Osaka.
- Brigo, D., Pallavicini, A. and Torresetti, R. (2010). *Credit Models and the Crisis: A Journey into CDOs, Copulas, Correlations and Dynamic Models*, John Wiley & Sons, Chichester.
- Davis, M. and Lo, V. (2001). Infectious defaults, *Quantitative Finance*, **1**, 382-387.
- Gupton, G. M., Finger, C. C. and Bhatia, M. (1997). CreditMetrics™ — Technical Document, RiskMetrics Group, New York, <http://www.riskmetrics.com/research.html>.
- Hull, J. and White, A. (1990). Pricing interest-rate-derivative securities, *Review of Financial Studies*, **3**, 573-592.
- Hull, J. and White, A. (2004). Valuation of a CDO and n-th to default CDS without Monte Carlo simulation, *Journal of Derivatives*, **12**(2), 8-23.
- Hull, J. and White, A. (2006). Valuing credit derivatives using an implied copula approach, *Journal of Derivatives*, **14**(2), 8-28.

- Hull, J. and White, A. (2010). An improved implied copula model and its application to the valuation of bespoke CDO tranches, *Journal of Investment Management*, **8**(3), 11-31.
- Jarrow, R. A. and Turnbull, S. M. (1995). Pricing derivatives on financial securities subject to credit risk, *Journal of Finance*, **50**, 53–86.
- Kijima, M. and Muromachi, Y. (2000). Evaluation of credit risk of a portfolio with stochastic interest rate and default processes, *Journal of Risk*, **3**(1), 5–36.
- Kusuoka, S. (1999). A remark on default risk models, *Advances in Mathematical Economics*, **1**, 69–82.
- Morgan Guaranty Trust Company and Reuters Ltd. (1996). RiskMetrics™—Technical Document, Morgan Guaranty Trust Company of New York, New York, <https://www.msci.com/documents/10199/5915b101-4206-4ba0-ae2-3449d5c7e95a>.
- 室町幸雄 (2007). 『信用リスク計測と CDO の価格付け』, 朝倉書店, 東京.

An Extension of a CDO Pricing Model Using a Copula toward a Risk Evaluation Model

Yukio Muromachi

Graduate School of Management, Tokyo Metropolitan University

Many financial risk evaluation models were developed before the worldwide financial crisis, but none of them could predict the crisis. This is because all of these models are purely statistical, meaning that they are based on analyses of historical data and therefore cannot predict crises that have never happened. In this article, we consider the CDO pricing model used in the paper that proposed the implied copula, and propose extending it to a risk evaluation model. In the pricing model, assuming that the default times are conditionally independent, the non-parametric distribution of conditional default probabilities are estimated from the market prices of CDO tranches, and the estimated distributions in previous studies show that the default probabilities will increase dramatically with some small probabilities. It is our model that will apply such estimated results to evaluating the risk of a portfolio. By using not only the historical data but also the market prices of CDOs and derivatives, market participants' potential fears of future catastrophic loss can be reflected in the risk evaluation. In this article, we show how to extend the pricing model toward a risk evaluation model according to some known theoretical results, and in particular, discuss in detail the importance of the change of measure and its mathematical description. The important tool of our model is the factor that has a large influence on conditional default probabilities for all entities. We show some numerical results that can hardly be obtained from the existing models, and that are consistent with the remarkable features seen in the financial crisis.

連続変数で表される事件の接合関数を用いた 生存分析

福元 健太郎[†]

(受付 2019 年 6 月 3 日；改訂 12 月 11 日；採択 12 月 16 日)

要 旨

ある事件がいつ起きるか(時間変数)に影響する要因を調べるには、生存分析が使われることが多い。しかしそれだけでなく、どんな事件が起きるか(事件変数)にも関心がある場合がある。しかも往々にして、時間変数と事件変数は統計的に独立ではない。もし事件変数が名義変数であれば従属的競合リスク・モデルが使われる。本稿は、事件変数が連続変数である場合を扱う、接合関数を利用したモデルを提案する。そしてモンテ・カルロ・シミュレーションにより、時間変数と事件変数を別々に分析すると(打ち切りがある場合は特に)バイアスが生じるので、本稿の連続変数事件接合関数生存分析モデルを用いて両者を同時に分析する必要があることを示す。最後に、戦後英国の選挙の相対的タイミングと与党の得票率との関係を検討した先行研究のデータを再分析する。選挙が遅いと与党に有利ではあるが、選挙を早くしても必ずしも与党に不利になる訳ではない、という非対称な従属性が明らかになる。

キーワード：生存時間分析，生存解析，タイミング，従属性，選挙，イギリス。

1. はじめに

ある事件がいつ起きるか(時間変数)に影響する要因を調べるには、生存分析が使われることが多い。それだけでなく、どんな事件が起きるか(事件変数)にも関心がある場合がある。しかも往々にして、時間変数と事件変数は統計的に独立ではない。例えば、選挙がいつ行われるかと、選挙結果がどうなるかは、後述するように密接に関連する。しかし、本稿で明らかにするように、時間変数と事件変数を別々に分析するとバイアスが生じるので、両者を同時に分析する必要がある。もし事件変数が名義変数であれば(例えば選挙事由が、議会解散か、内閣不信任か、任期満了か)、従属的競合リスク・モデルが使われる(Fukumoto, 2009; Gordon, 2002)。しかし事件変数が連続変数や順序変数で表される場合の分析手法はこれまでなかった。Fukumoto (2015)で事件が順序変数である場合(例えば、与党過半数か、過半数を得た政党無しか、野党第 1 党過半数か)を扱っているので、本稿は事件が連続変数である場合(例えば与党の得票率)を扱うモデルとして、連続変数事件接合関数生存分析(Copula-based Continuous Event History Analysis, 以下 CCEHA)モデルを提案する¹⁾。

本稿は時間変数と事件変数の従属性をモデル化するにあたり、接合関数(後述)を利用する。接合関数の利点の 1 つは、時間変数や事件変数の周辺分布が何であってもよいというモジュール性である。もう 1 つは、時間変数と事件変数の従属性は、既存の二変量分布を超えて、実

[†] 学習院大学 法学部：〒171-8588 東京都豊島区目白 1-5-1

多様なものを表現できる、という柔軟性である。しかも、従属性は対称的でなくてもよい。例えば、本稿が明らかにするように、選挙が遅いと与党に有利ではあるが、選挙を早くしても与党に不利になる訳ではない、という非対称な従属性をも扱える。

本稿は次のように構成される。先ず、CCEHA モデルを提示する。次いで、モンテ・カルロ・シミュレーションを行う。時間変数と事件変数に従属性がある時に(かつ場合によっては標本に打ち切りがある時に)、時間モデルと事件モデルを別々に推定すると、事件モデルの推定にバイアスが生じるが、CCEHA モデルはそれに対処し得る。その後、Smith (2004) が分析した戦後英国の選挙データを CCEHA モデルで再分析する。CCEHA モデルは、元の分析よりも多くの情報を、同じデータから引き出す。最後にモデルの長所と今後の課題を展望する。

2. モデル

2.1 設定

2.1.1 記法

ある観測単位(例えば議会)が時点 $t_0 = 0$ から始まって時点 $T > 0$ まで存続し、そこで事件 E (例えば選挙における与党得票率) を起こして終わるとする。ここで T も E も確率変数かつ連続変数かつ潜在変数であり、それぞれの周辺分布の累積分布関数を

$$T \sim F_T(t | \theta_T), \quad E \sim F_E(e | \theta_E)$$

と表す。 θ_T と θ_E はそれぞれの媒介変数ベクトルである。これらの分位変数は当然に標準一様分布 \mathcal{U} に従う。

$$U_T \equiv F_T(T | \theta_T) \sim \mathcal{U}, \quad U_E \equiv F_E(E | \theta_E) \sim \mathcal{U}.$$

T と E の同時分布の累積分布関数は、これら分位変数の関数 $C(\cdot, \cdot)$ で表すことができる。

$$(T, E) \sim F_{TE}(t, e | \theta) \equiv C(u_T, u_E | \theta_C).$$

ここで θ_C は T と E の従属性を決める媒介変数であり、 $\theta \equiv (\theta'_T, \theta'_E, \theta_C)'$ とする。この関数 $C(\cdot, \cdot)$ を接合関数と呼ぶ²⁾。シュクラールの定理により、 $C(\cdot, \cdot)$ は一意に定まる。ここで $F_T(\cdot)$ と $F_E(\cdot)$ は 1 回微分可能、 $F_{TE}(\cdot, \cdot)$ は(よって $C(\cdot, \cdot)$ も) 2 回微分可能であると想定し、確率密度関数を次のように表記する。

$$\begin{aligned} \frac{\partial F_T(t | \theta_T)}{\partial t} &\equiv f_T(t | \theta_T), & \frac{\partial F_E(e | \theta_E)}{\partial e} &\equiv f_E(e | \theta_E), \\ \frac{\partial^2 F_{TE}(t, e | \theta)}{\partial t \partial e} &\equiv f_{TE}(t, e | \theta), & \frac{\partial^2 C(u_T, u_E | \theta_C)}{\partial u_T \partial u_E} &\equiv c(u_T, u_E | \theta_C). \end{aligned}$$

2.1.2 接合関数

本稿は 5 種類の接合関数を検討する。最も基本的なのは正規接合関数である。

$$C_G(u_T, u_E | \theta_{C(G)}) \equiv \Phi_2\{\Phi^{-1}(u_T), \Phi^{-1}(u_E) | \theta_{C(G)}\}.$$

ここで $\Phi(\cdot)$ と $\Phi_2(\cdot, \cdot)$ はそれぞれ 1 変量と 2 変量の標準正規分布の累積分布関数である。 $-1 < \theta_{C(G)} < 1$ は 2 変量の相関係数である。図 1 はいくつかの接合関数の密度を等高線及び濃淡で示したものである。いずれもケンドール相関係数が、後出の実データの値(約 0.325)に等しくなっている。図 1 の左上図が正規接合関数であり、従属性はやや強い($\theta_{C(G)} \equiv 0.488$)。

次はファーリー・グンベル・モルゲンシュテルン(FGM)接合関数である。

$$C_{FGM}(u_T, u_E | \theta_{C(FGM)}) \equiv u_T u_E \{1 + \theta_{C(FGM)}(1 - u_T)(1 - u_E)\}.$$

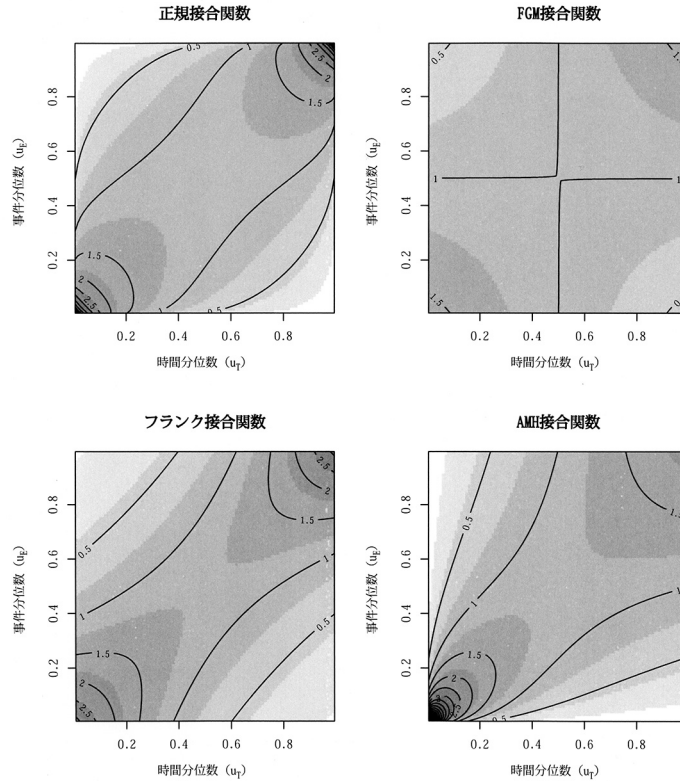


図 1. 接合関数の密度の例.

ここで $-1 < \theta_{C(FGM)} < 1$ である. 従属性は弱い(図 1, 右上図, $\theta_{C(FGM)} \cong 1.462$).

以下では, アルキメデス接合関数と呼ばれる接合関数の種類を紹介する. これはある生成関数 $\varphi(\cdot)$ に対して, 次のように定義される.

$$C_A\{u_T, u_E \mid \varphi(\cdot)\} \equiv \varphi^{-1}\{\varphi(u_T) + \varphi(u_E)\}.$$

その一例はフランク接合関数で, その生成関数は

$$\varphi_F(z \mid \theta_{C(F)}) \equiv -\log \left\{ \frac{\exp(-z\theta_{C(F)}) - 1}{\exp(-\theta_{C(F)}) - 1} \right\}$$

である($\theta_{C(F)}$ が取り得る値の範囲に制約はない). 従属性はやや弱い(図 1, 左下図, $\theta_{C(F)} \cong 3.205$).

もう 1 つはクレイトン接合関数で, その生成関数は

$$\varphi_K(z \mid \theta_{C(K)}) \equiv (z^{-\theta_{C(K)}} - 1)\theta_{C(K)}^{-1}$$

である. ここで $\theta_{C(K)} > 0$ とする. この接合関数は順相関しか扱えず, 図 2 の左上図 ($\theta_{C(K)} \cong 0.962$) の通り, 従属性が左裾(左上図の左下隅)で強く右裾(左上図の右上隅)で弱い場合に適切である³⁾. しかしクレイトン接合関数の関連接合関数を用いれば, その他の従属性も扱える. 関連接合関数の 1 つは生存接合関数であり, 一般に $C(u_T, u_E)$ に対して

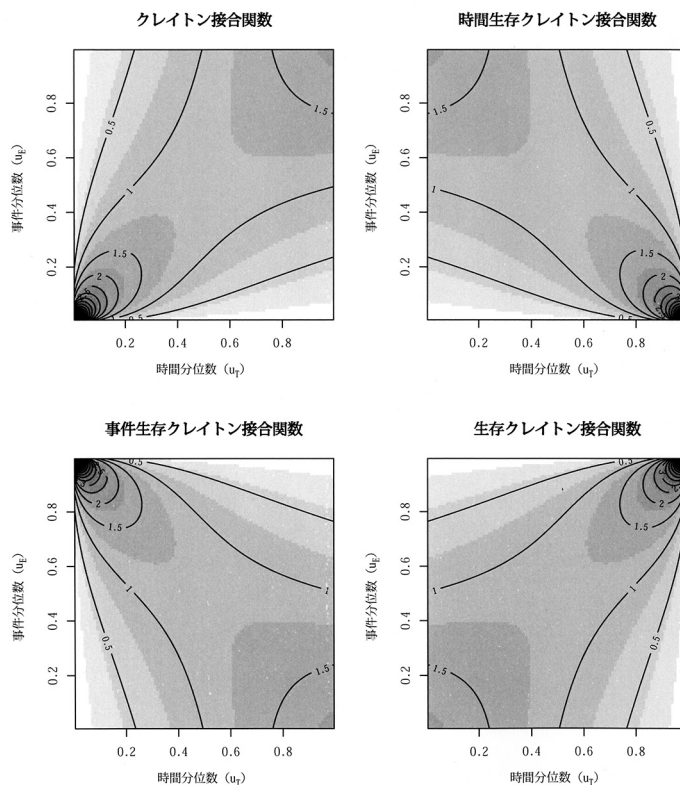


図 2. クレイトン接合関数とその関連接合関数の密度の例.

$\bar{C}^{(TE)}(u_T, u_E) \equiv u_T + u_E - 1 + C(1 - u_T, 1 - u_E)$ と定義される. クレイトン接合関数の生存接合関数の従属性は, 右裾で強く左裾で弱くなる (図 2, 右下図). 別の関連接合関数として $\bar{C}^{(T)}(u_T, u_E) \equiv u_E - C(1 - u_T, u_E)$ がある (本稿では時間生存接合関数と呼ぶ). クレイトン接合関数の場合 (図 2, 右上図), 従属性は事件変数の左裾と時間変数の右裾 (右上図の右下隅) で強く, 事件変数の右裾と時間変数の左裾 (右上図の左上隅) で弱いという, 逆相関となる. さらに別の関連接合関数として $\bar{C}^{(E)}(u_T, u_E) \equiv u_T - C(u_T, 1 - u_E)$ も考えられる (本稿では事件生存接合関数と呼ぶ). クレイトン接合関数の場合 (図 2, 左下図), 従属性は事件変数の右裾と時間変数の左裾 (左下図の左上隅) で強く, 事件変数の左裾と時間変数の右裾 (左下図の右下隅) で弱い⁴⁾.

最後にアリ・ミハイル・ハク (AMH) 接合関数は生成関数として

$$\varphi_{AMH}(z | \theta_{C(AMH)}) \equiv \log \left\{ \frac{1 - \theta_{C(AMH)}(1 - z)}{z} \right\}$$

を用いて導出される (図 1, 右下図, $\theta_{C(AMH)} \equiv 0.987$). ここで $-1 < \theta_{C(AMH)} < 1$ である. AMH 接合関数についても 3 つの関連接合関数を検討する.

2.2 推定

ここからは観測単位 i の添字を明記する. 打ち切り時点を \bar{t}_i とする.

2.2.1 事件を観測する場合の尤度

もし $T_i = t_i \leq \bar{t}_i$ であれば、時点 $T_i = t_i$ において事件 $E_i = e_i$ を観測する。そうすると i の尤度は $T_i = t_i$ と $E_i = e_i$ で評価した同時分布密度に比例するから

$$\begin{aligned} \mathcal{L}_i(\boldsymbol{\theta} \mid t_i, e_i) &\propto p(T_i = t_i, E_i = e_i \mid \boldsymbol{\theta}) \\ &= f_{TE}(t_i, e_i \mid \boldsymbol{\theta}) \\ &= \left. \frac{\partial^2 F_{TE}(t, e \mid \boldsymbol{\theta})}{\partial t \partial e} \right|_{t=t_i, e=e_i} \\ &= \left. \frac{\partial u_T}{\partial t} \right|_{t=t_i} \left. \frac{\partial u_E}{\partial e} \right|_{e=e_i} \left. \frac{\partial^2 C(u_T, u_E \mid \theta_C)}{\partial u_T \partial u_E} \right|_{u_T=u_{T,i}, u_E=u_{E,i}} \\ &= f_T(t_i \mid \boldsymbol{\theta}_T) f_E(e_i \mid \boldsymbol{\theta}_E) c(u_{T,i}, u_{E,i} \mid \theta_C) \end{aligned}$$

が導かれる。これは、2つの周辺分布の密度と接合関数の密度の積であり、本モデルのモジュール性をよく表している。すなわち、時間の周辺分布 ($f_T(\cdot)$) 及び・又は事件の周辺分布 ($f_E(\cdot)$) 及び・又は両者の従属性 ($c(\cdot, \cdot)$) を、相互に無関係に変えることができる。例えば、 $f_T(\cdot)$ として対数正規分布、ワイブル分布のどちらを選んでもよいし、そのいずれの場合でも、 $f_E(\cdot)$ として正規分布、 t 分布のどちらを選んでもよいし、以上4通りのどの場合でも、 $c(\cdot, \cdot)$ として正規、FGM、フランク、クレイトン、AMHなどの接合関数から何を選んでもよい。ここに挙げたメニューだけでも、 $2 \times 2 \times 5 = 20$ 通りのモデルがある。

2.2.2 存続が打ち切られる場合の尤度

もし $T_i > \bar{t}_i$ であれば、打ち切り時点 \bar{t}_i までに事件 E_i もそれが起きる時点 T_i も観測しない。これを存続が打ち切られると言う。便宜上、 $t_i = \bar{t}_i, e_i = \phi$ と表記しよう。この場合 i の尤度は打ち切り時点 \bar{t}_i 以降のいずれかの時点 T_i で何らかの事件 E_i が起こる確率に比例する。

$$\begin{aligned} \mathcal{L}_i(\boldsymbol{\theta} \mid t_i, e_i) &\propto \Pr(T_i > \bar{t}_i \mid \boldsymbol{\theta}) \\ &= \int_{\bar{t}_i}^{\infty} \int_{-\infty}^{\infty} f_{TE}(t, e \mid \boldsymbol{\theta}) dt de \\ &= 1 - F_T(\bar{t}_i \mid \boldsymbol{\theta}_T). \end{aligned}$$

2.2.3 最尤法

全観測数を N とすれば、全尤度は

$$\mathcal{L}(\boldsymbol{\theta} \mid \mathbf{t}, \mathbf{e}) = \prod_{i=1}^N \mathcal{L}_i(\boldsymbol{\theta} \mid t_i, e_i)$$

である(ここで $\mathbf{t} \equiv (t_1, t_2, \dots, t_N)'$ 、 \mathbf{e} も同様)。この値を最大化する $\boldsymbol{\theta}$ を推定値 $\hat{\boldsymbol{\theta}}$ とする。このように本稿は、時間モデルと事件モデルを同時に推定する CCEHA モデルを提唱する。

3. シミュレーション

時間変数と事件変数の間に従属性があると、時間モデルと事件モデルを別々に推定することがバイアスをもたらす。そのことを示すため、以下のようにモンテ・カルロ・シミュレーションを行う。

3.1 モデルの特定化

3.1.1 時間モデル

$\boldsymbol{\theta}_T = (\alpha_T, \beta_T)'$ とし, 説明変数ベクトル \boldsymbol{x}_T を用いた線形予測値を $\mu_T = \boldsymbol{x}_T' \boldsymbol{\beta}_T$ で表す. 2つの分布を考える. 1つは, 対数正規分布である.

$$F_{T(LN)}(t | \boldsymbol{\theta}_T, \boldsymbol{x}_T) \equiv \Phi \left\{ \frac{\log(t) - \mu_T}{\exp(\alpha_T)} \right\}.$$

もう1つはワイブル分布である.

$$F_{T(W)}(t | \boldsymbol{\theta}_T, \boldsymbol{x}_T) \equiv 1 - \exp\{-\exp(\mu_T)t^{\exp(\alpha_T)}\}.$$

\boldsymbol{x}_T は定数項 $x_{T0} = 1$ と2つの共変量 x_{T1}, x_{T2} から成るとし, 媒介変数の値を次のように定める⁵⁾.

$$\boldsymbol{x}_T = (x_{T0}, x_{T1}, x_{T2})', \quad \boldsymbol{\beta}_T = (\beta_{T0}, \beta_{T1}, \beta_{T2})' = (-1, 1, -1)', \quad \alpha_T = 0.$$

3.1.2 事件モデル

$\boldsymbol{\theta}_E = (\alpha_E, \beta_E)'$ とし, 説明変数ベクトル \boldsymbol{x}_E を用いた線形予測値を $\mu_E = \boldsymbol{x}_E' \boldsymbol{\beta}_E$ で表す. E は正規分布に従うとする.

$$F_{E(N)}(e | \boldsymbol{\theta}_E, \boldsymbol{x}_E) \equiv \Phi \left\{ \frac{e - \mu_E}{\exp(\alpha_E)} \right\}.$$

\boldsymbol{x}_E は定数項 $x_{E0} = 1$ と2つの共変量 x_{E1}, x_{E2} から成るとし, 媒介変数の値を次のように定める.

$$\boldsymbol{x}_E = (x_{E0}, x_{E1}, x_{E2})', \quad \boldsymbol{\beta}_E = (\beta_{E0}, \beta_{E1}, \beta_{E2})' = (1, -1, 1)', \quad \alpha_E = 2.$$

3.1.3 従属性

時間変数と事件変数の従属性をモデル化するものとして2つの接合関数を検討する. 1つは正規接合関数である. $\theta_{C(G)}$ の値として, 0.762(順相関), 0(無相関), -0.762(逆相関)の3つを考える. もう1つはクレイトン接合関数である. $\theta_{C(K)} = 1$ とする.

3.2 データ発生

標準一様分布 U から3つの乱数 $u_T, \bar{u}, \underline{u}$ を発生させる. 次に以下のような変数を作る.

$$t = F_T^{-1}(u_T), \quad \bar{t} = F_T^{-1}(\bar{u}^{\exp(\psi)}), \quad \underline{t} = F_T^{-1}(\underline{u}^{\exp(4)}).$$

ここで打ち切り調整媒介変数 ψ は0(半分の観測が打ち切り)もしくは-10(ほとんど打ち切り無し)である. \underline{t} は左側切断時点を表す. もし $t \leq \underline{t}$ であれば, この存続はそもそも観測されない. $t > \bar{t}$ の場合, この存続は打ち切られる. 以上に該当しない, つまり $\underline{t} < t \leq \bar{t}$ であれば, 次の段階に進んで, 事件変数 e を発生させる.

一般に t に条件づけられた E の累積分布関数 $F_{E|T}(e | t)$ は導出するのが往々にして難しいが,

$$F_{E|T}(e | t) = \frac{\partial C(u_T, u_E)}{\partial u_T} \equiv C_{E|T}(u_E | u_T)$$

が成り立つため, 接合関数の微分さえ導出できれば容易である. 正規接合関数の場合は

$$C_{E|T}(u_E | u_T) = \Phi \left\{ \frac{\Phi^{-1}(u_E) - \theta_{C(G)} \Phi^{-1}(u_T)}{\sqrt{1 - \theta_{C(G)}^2}} \right\}$$

であり⁶⁾、クレイトン接合関数の場合は

$$C_{E|T}(u_E | u_T) = (u_T^{-\theta_{C(K)}} + u_E^{-\theta_{C(K)}} - 1)^{-\frac{1}{\theta_{C(K)}} - 1} u_T^{-\theta_{C(K)} - 1}$$

である。以上を踏まえ、乱数 $u_{E|T}$ を標準一様分布 U から発生させて、まず $u_E = C_{E|T}^{-1}(u_{E|T} | u_T)$ を計算し、ついで $e = F_E^{-1}(u_E)$ を作る⁷⁾。

3.3 推定

以上のように CCEHA モデルに基づいて作成したデータを、本稿はまず CCEHA モデルで推定する⁸⁾。CCEHA モデルは真のデータ発生過程を想定しているから、当然に推定の性能は良い。関心があるのは、時間モデルと事件モデルを別々に推定するとどうなるかである。

まず時間モデルの媒介変数を最尤法で推定する(ここでは生存分析になる)。観測単位 i の事件が観測された場合の尤度は

$$\mathcal{L}_{T,i}(\theta_T | t_i, \mathbf{x}_{T,i}) \propto \frac{f_T(t_i | \theta_T, \mathbf{x}_{T,i})}{1 - F_T(t_i | \theta_T, \mathbf{x}_{T,i})}$$

である。存続が打ち切られた場合の尤度は

$$\mathcal{L}_{T,i}(\theta_T | t_i, \mathbf{x}_{T,i}) \propto \frac{1 - F_T(t_i | \theta_T, \mathbf{x}_{T,i})}{1 - F_T(t_i | \theta_T, \mathbf{x}_{T,i})}$$

である。推定値 $\hat{\theta}_T$ は $\prod_i \mathcal{L}_{T,i}(\theta_T | t_i, \mathbf{x}_{T,i})$ を最大化するものとして求められる。

次に、事件が観測された観測単位のみを用いて、以下の尤度を最大化するものとして、事件モデルの媒介変数を推定する(ここでは線形回帰分析になる。以下、単純事件モデルと呼ぶ)。

$$\prod_{i:e_i \neq \phi} \mathcal{L}_{E,i}(\theta_E | e_i, \mathbf{x}_{E,i}) \propto \prod_{i:e_i \neq \phi} f_E(e_i | \theta_E, \mathbf{x}_{E,i}).$$

さらに、事件モデルがなるべく CCEHA モデルと対等になるように、時間モデルの残差を説明変数に加えることも考える。まず、時間モデルの推定値に基づく時間変数の中央値を \hat{t}_i と表記する。具体的には、時間変数が対数正規分布に従う場合、

$$\hat{t}_i \equiv \exp(\hat{\mu}_{T,i}),$$

ワイブル分布に従う場合、

$$\hat{t}_i \equiv \left\{ -\frac{\log(1 - 0.5)}{\exp(\hat{\mu}_{T,i})} \right\}^{\exp(-\hat{\alpha}_T)}$$

となる。ここで、対数化した時間変数の実現値から中央値を引いた差を、時間モデルの残差と定義する。

$$\hat{e}_{T,i} \equiv \log(t_i) - \log(\hat{t}_i).$$

これを共変量に加えたものを誤差修正事件モデルとする。

$$\mathbf{x}_{EC,i} = (x_{E0,i}, x_{E1,i}, x_{E2,i}, \hat{e}_{T,i})', \quad \beta_{EC} = (\beta_{E0}, \beta_{E1}, \beta_{E2}, \beta_{TE})'.$$

以上のように時間モデルと事件モデルを別々に推定する手順を分離推定、これに対して CCEHA モデルによる推定を同時推定と呼ぼう。

3.4 シミュレーションの手順

観測数は $N = 1,000$ である。共変量 $x_{T1,i}, x_{T2,i}, x_{E1,i}, x_{E2,i}$ は、シミュレーション全体を通じて 1 回だけ、標準正規分布から乱数発生させた。時間変数の分布(対数正規かワイブル

か), 接合関数(正規かクレイトンか), 従属性媒介変数(正規接合関数の場合に $\theta_{C(G)}$ が正か負か 0 か, クレイトン接合関数の場合は常に $\theta_{C(K)} = 1$), 打ち切り調整媒介変数(打ち切り有り ($\psi = 0$)か無し ($\psi = -10$)か)を選ぶ.

1 回のルーティンで, t と e から成るデータセットを無作為に発生させる. 同時推定と分離推定を行い, 推定値を保存する. ヘシアン行列が特異で計算できない(すなわち推定値が収束しない)場合は, そのデータセットを廃棄する. このルーティンを 1,000 組の推定値が得られるまで繰り返す⁹⁾.

3.5 結果

時間モデルの係数 β_T は, 分離推定による推定値も不偏であり, かつ CCEHA による推定値と同じくらい効率的であったので, 以下では省略する. 事件モデルの係数 β_E については, 分離推定による推定値は, 時間変数と事件変数の従属性を考慮していないために, CCEHA による推定値とそのままでは比較できない. そこで本稿では, $x_{T1} = x_{T2} = x_{E1} = x_{E2} = 0$ の場合の, u_T の各値に対する E の中央値

$$\hat{e}(u_T, \hat{\theta}) \equiv F_{E|T}^{-1}(0.5 | u_T, \hat{\theta}_E, \mathbf{x}_E, \hat{\theta}_C)$$

について, 両モデルを比べる. 計算方法は 3.2 節と同じ要領で $u_{E|T} = 0.5$ とすればよい.

図 3 は時間変数が対数正規分布に従い, 正規接合関数が用いられた場合である. 6 つの図から成る. 左列は打ち切りがない場合 ($\psi = -10$), 右列は半数が打ち切られた場合 ($\psi = 0$) である. 上段, 中段, 下段はそれぞれ接合関数が順相関 ($\theta_{C(G)} > 0$), 無相関 ($\theta_{C(G)} = 0$), 逆相関 ($\theta_{C(G)} < 0$) を表す場合である. 全ての図で, 横軸は t , 縦軸は $\hat{e}(u_T, \cdot)$ を表す¹⁰⁾. 太点線と太実線はそれぞれ CCEHA モデルと誤差修正事件モデルによる 1,000 組の $\hat{\theta}$ に基づく $\hat{e}(u_T, \hat{\theta})$ の平均値を表す. 斜線部と陰影部は同様に 1,000 組の $\hat{e}(u_T, \hat{\theta})$ の 2.5% 及び 97.5% 分位数に挟まれた区間(以下「95% 区間」と呼ぶ)である. 真の値 $\hat{e}(u_T, \theta)$ は, CCEHA モデルによる $\hat{e}(u_T, \hat{\theta})$ の平均値(太点線)に重なるため, 改めて図示はしていない. CCEHA モデルが真のデータ発生過程であることを考えれば, これは驚くに当たらない. 単純事件モデルは, 時間モデルの残差 \hat{e}_T を考慮していないので, 1,000 組の $\hat{\theta}$ に基づく $\hat{\mu}_E$ の平均値は水平な細実線で表されている. 水平な細点線は, 事件変数の周辺分布の真の平均値 $\mu_E = \beta_{E0} = 1$ を表す. 垂直な細点線は t の真の中央値 $F_T^{-1}(0.5 | \theta_T, \mathbf{x}_T)$ を表す.

まず, 時間変数と事件変数の間に順相関があり, 標本の一部が打ち切られている場合(右上図), 単純事件モデルによる推定値の平均値(水平な細実線)は(真の値を表す水平な細点線と比べて)下向バイアスがかかり, 符号が真実とは逆で負になる. その理由は以下の通りである. 順相関があれば, 時間変数が大きいほど, 事件変数は大きくなるが, 打ち切られることも多くなる. よって, 潜在的な事件変数は大きいほど観測されにくく, 観測された事件変数の周辺分布は下向きに歪む. 従って, 観測された事件変数の平均値は, 潜在的な(真の)事件変数の平均値より低くなる. 同様に, 逆相関でかつ打ち切りがあれば, 単純事件モデルによる推定値は上向バイアスがかかる(右下図). 相関がないか, 打ち切りがなければ, 単純事件モデルは不偏である. こうした事情はヘックマン選択モデルと似ている.

また CCEHA モデル(太点線・斜線部)と誤差修正事件モデル(太実線・陰影部)はほとんど同じ結果である(そのため太点線は太実線に覆われて見えない). 両者は今の設定では同じモデルだからである. 時間変数と事件変数の間に順相関がある場合(上段), 時間変数が大きいほど, 事件変数は大きくなる. 逆相関の場合(下段), 時間変数が大きいほど, 事件変数は小さくなる. 相関がないと(中段), 時間変数が長くても, 事件変数の平均値は変わらないが, 95% 区間は広がる. 事件変数の 95% 区間は打ち切りがある方(右列)がない場合(左列)よりも広がる.

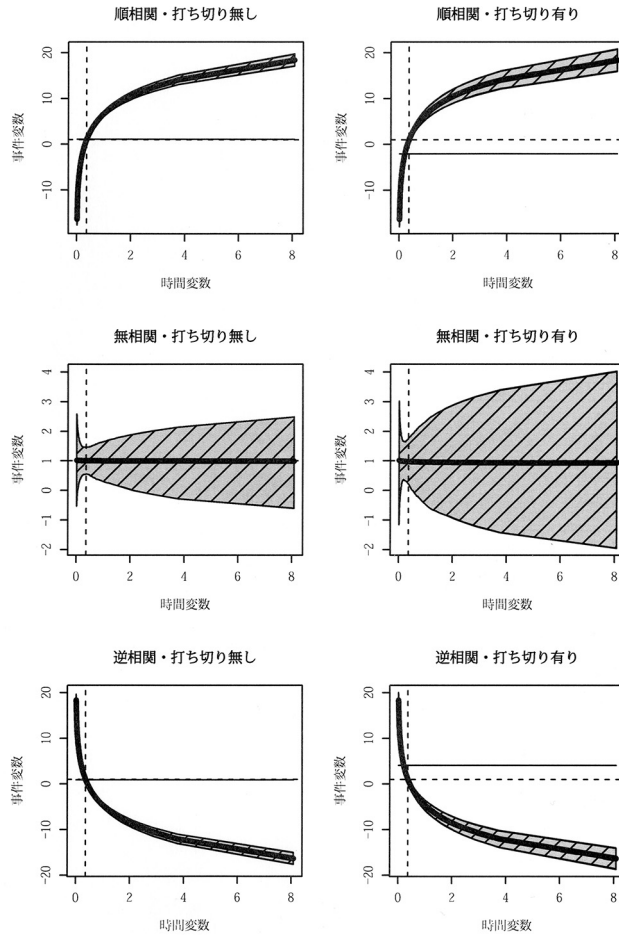


図 3. 正規接合関数と対数正規分布する時間変数を使った場合のモンテ・カルロ・シミュレーション：CCEHA モデル (太点線・斜線部)，誤差修正事件モデル (太実線・陰影部)，単純事件モデル (水平な細実線)，事件変数の周辺分布の中央値 (水平な細点線)，時間変数の中央値 (垂直な細点線)。

何故なら，事件変数が観測される標本の大きさは，前者の方が後者より小さいからである．興味深いことに，分離推定 (陰影部) は同時推定 (斜線部) と同じくらい効率的である。

もっとも，誤差修正事件モデルが CCEHA モデルと代替可能であるのは，時間変数が対数正規分布で正規接合関数が当てはまるという，これまで検討してきた極めて特殊な状況下に限られる．図 4 は時間変数の分布が対数正規からワイブルに変わった場合を示している (正規接合関数はそのまま)．すると，時間変数と事件変数の間に正 (または負) の相関があれば (上段または下段)，打ち切りが有っても (左列) 無くても (右列)，誤差修正事件モデルの推定値 (太実線・陰影部) さえ (太点線・斜線部と比べて) 下向 (または上向) バイアスがかかる．もっとも，時間変数と事件変数が互いに独立であれば (図 4 の中段)，誤差修正事件モデルの推定値は不偏であり (太点線と太実線が重なる)，95% 区間 (陰影部) は CCEHA モデルのそれ (斜線部) よりも狭い．なお，時間変数と事件変数の間に従属性があり，かつ打ち切られた標本があれば (図 4 の

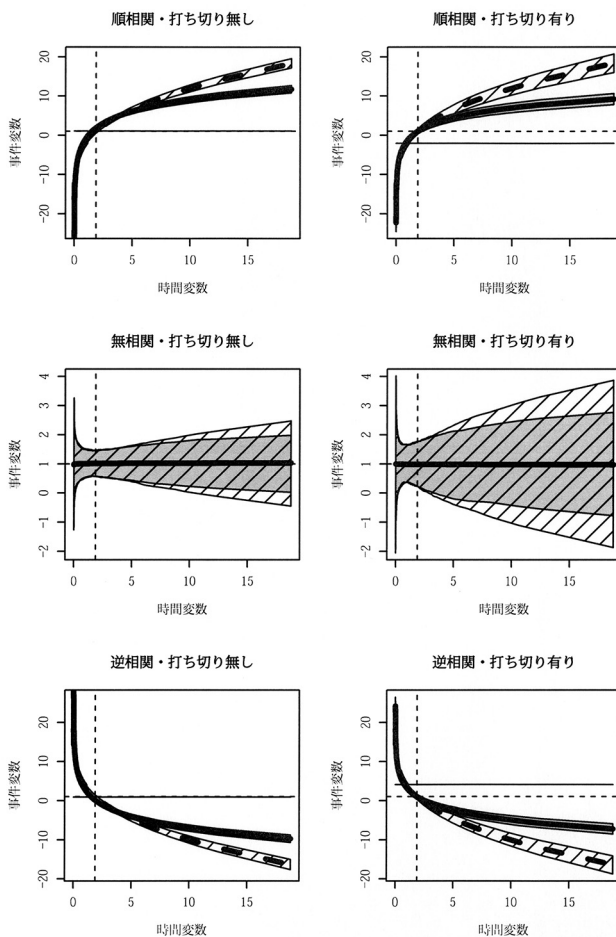


図 4. 正規接合関数とワイブル分布する時間変数を使った場合のモンテ・カルロ・シミュレーション：CCEHA モデル(太点線・斜線部)，誤差修正事件モデル(太実線・陰影部)，単純事件モデル(水平な細実線)，事件変数の周辺分布の中央値(水平な細点線)，時間変数の中央値(垂直な細点線)。

右上図と右下図)，単純事件モデルの推定値にバイアスがあるのは，図 3 の場合と同様である。

最後に，正規接合関数をクレイトン接合関数に変えたらどうだろうか．時間変数は再び対数正規分布するとしよう．図 5 に結果を示す(クレイトン接合関数の場合，時間変数と事件変数の間の相関は正しかない)．誤差修正事件モデルの推定値は先ほどとは逆に上向バイアスがかかる．これは次のような理由による．クレイトン接合関数の場合，時間変数と事件変数がともに小さい場合は順相関だが，両変数とも大きいとほとんど無相関になる(図 2 の左上図)．しかし誤差修正事件モデルは時間変数と事件変数がともに小さい場合の順相関を，両変数とも大きい場合にまで誤って外挿してしまう．そのため時間変数が大きい場合の上向バイアスを生んでしまう．打ち切られた標本があれば(右図)，単純事件モデルの推定値にバイアスがあるのは，図 3 の場合と同様である。

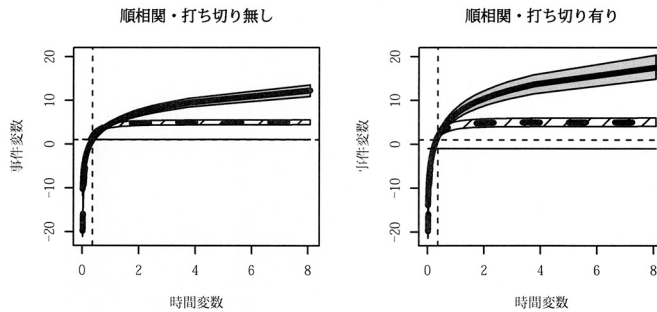


図 5. クレイトン接合関数と対数正規分布する時間変数を使った場合のモンテ・カルロ・シミュレーション：CCEHA モデル(太点線・斜線部)，誤差修正事件モデル(太実線・陰影部)，単純事件モデル(水平な細実線)，事件変数の周辺分布の中央値(水平な細点線)，時間変数の中央値(垂直な細点線)。

4. 実例

4.1 モデルの特定化

本節は CCEHA モデルを用いて Smith (2004) のデータを再分析する¹¹⁾。彼は、1945 年から 2001 年までの英国における選挙のタイミングと結果の関係を分析している。彼の議論によれば、与党はこれから経済状況が悪化することが予想される場合、その時になってから選挙をすると負けるので、先手を打って早めに議院を解散して選挙に持ち込む。しかし有権者もそれがわかるので、結局、予想よりも早く選挙が行われる場合、与党の得票率は落ちる、と主張している。

まず Smith (2004, p. 104, モデル 3.3) は、ワイブル分布を用いた時間モデルにより、選挙のタイミングを検討している。個々の議院は、直前の選挙から始まり、首相が選挙を行うことを発表するという「事件」により終わる¹²⁾。後述する日次時間変数変数を考慮するため、議院を日単位に分割したものを観測単位とする。直前の選挙からの日数が時間変数 (T) である。総観測数は 19,742 (日) であるが、議院の数は 16 であり、そのうち 2 つは打ち切られている¹³⁾。よって、議院開始日以外の $19,742 - 16 = 19,726$ 日は左側切断されており、事件が起きていない $19,742 - 14 = 19,728$ 日は右側打ち切りされている。説明変数 (x_T) は、残任期間(残任日数を 365 で除した商)、残任期間の二乗項、与野党支持率差(直近の月次世論調査における、与党へ投票する意向を持つ人の百分率から、主要野党へ投票する意向を持つ人の百分率を引いた差)、与野党議席差(与党議席数から主要野党議席数を引いた差)である (Smith, 2004, p. 96)¹⁴⁾。

次に Smith (2004, p. 141, モデル 4.5) は、正規分布を用いた(誤差修正ではなく単純)事件モデルにより、選挙の結果を分析している。事件変数 (E) は、与党支持率の変化(「選挙で実際に与党に投票した有権者の百分率から、選挙が発表される直前の世論調査で与党に投票する意向を表明した有権者の百分率を」引いた差 (Smith, 2004, p. 130)) である。選挙の相対的なタイミングを説明変数に含めるため、著者は累積危険率比と彼が呼ぶ値を計算している。これは「選挙が発表される前 30 日間の累積危険率を、それに先立つ 5 ヶ月間の累積危険率で除した比」である (Smith, 2004, p. 140)。すなわち、直前の選挙から s 日目の、先ほどの時間モデルから推定された危険率を \hat{h}_s と記し、直前の選挙から t 日目の累積危険率比を $\sum_{s=t-30}^t \hat{h}_s / \sum_{s=t-183}^{t-31} \hat{h}_s$ と定義する。危険率が上がり始めて 1 ヶ月も経たない頃だと、分子は既に大きいが分母はまだ小さいので、累積危険率比は大きな値となる。よって相対的に早い選挙の指標になると主張し

ている。この他の説明変数(x_E)は、時間モデルの説明変数(x_T)から残任期間の二乗項を除き、新首相(同じ議会で過去 100 日以内に首相が変わったか否かのダミー変数)、保守党ダミー変数を加えたものである(Smith, 2004, p. 96)。

CCEHA モデルを上述の Smith (2004)のデータに適用する。周辺分布は Smith (2004)の分析と同じで、時間変数はワイブル分布、事件変数は正規分布に従う。時間変数と事件変数の接合関数には生存 AMH 接合関数を用いる(この接合関数を選んだ理由は後述する)。説明変数は 2 組を検討する。1 組では、Smith (2004)の分析と同じ説明変数を用いる。もう 1 組では、累積危険率比を 2 つの理由から外す。1 つには、こうしたアドホックな尺度を拵えなくても、CCEHA モデルは相対的タイミングが事件変数に及ぼす影響を既に考慮に入れているからである。もう 1 つの理由は、累積危険率比が実のところ何を意味しているのか、よくわからないからである。例えば、もし累積危険率の「指数」の比であれば、生存率の比の逆数という意味を持つが、ここではそうになっていない。何故過去 1 ヶ月とそれに先立つ 5 ヶ月間を選んだのかも恣意的である。そして何より、そもそも累積危険率比が相対的タイミングを測定しているのか疑わしい。確かに危険率が上がり始めた直後の累積危険率比は高い。しかし、危険率が上がってかなり経った後だけでなく、危険率が上がる前も、累積危険率比は低くなる(前者は分子も分母も大きくなるため、後者は分子も分母も小さくなるため)。すると累積危険率比が小さいからといって、相対的なタイミングが遅いとは限らず、むしろあまりにも早すぎることもあり得る。

4.2 結果

本稿は分離推定により元の分析結果をほぼ再現することに成功した(表 1 の左 2 列(生存分析)と表 2 の左 2 列(単純な線形回帰))¹⁵⁾。CCEHA は同時推定だが、結果は時間モデル(表 1)と事件モデル(表 2)に分けて報告している(いずれも中 2 列が累積危険率比を制御した場合、右 2 列が制御しなかった場合)。時間モデルは、分離推定と同時推定で係数の推定値はほとんど同じである¹⁶⁾。事件モデルは、分離推定によれば残任期間や与野党議席率差の係数が有意でないが、同時推定によると有意になる、といった違いが見られる。

本稿で重要なのは、時間変数と事件変数の従属性である。一方で分離推定によれば、累積危険率比の係数が有意に負なので、選挙の相対的なタイミングが早いほど与党の得票率が減る、あるいは逆にタイミングが遅いと与党の得票率が増える、という単調な従属性が示唆される。他方、同時推定により得られた従属性の媒介変数の推定値 $\hat{\theta}_C = 0.987$ に基づいて生存 AMH 接合関数の密度を図示したのが、図 6 の左図である¹⁷⁾。さらに図 6 の右図で、推定された分位数である \hat{u}_T と \hat{u}_E の散布図を示す。ここで \hat{u}_T は選挙の相対的なタイミングを表す。右上隅に(重なって見えにくい)4 つの観測が集中していることは、左図の右上隅で密度が高くなって

表 1. 英国戦後総選挙のタイミングの分析。

モデル	分離推定		同時推定 有り		同時推定 無し	
	点推定値	標準誤差	点推定値	標準誤差	点推定値	標準誤差
与野党支持率差	0.130	0.033	0.131	0.029	0.138	0.028
与野党議席率差	-0.017	0.006	-0.014	0.005	-0.016	0.005
残任期間	-4.647	1.027	-4.462	1.025	-4.826	1.005
残任期間 ²	1.215	0.401	1.203	0.413	1.417	0.368
定数項	-3.590	5.204	-4.476	4.276	-6.345	4.249
$\log(\alpha_T)$	1.748	0.541	1.793	0.465	1.998	0.371

表 2. 英国戦後総選挙の結果(与党支持率の変化)の分析.

モデル 累積危険率比	分離推定 有り		同時推定 有り		同時推定 無し	
	点推定値	標準誤差	点推定値	標準誤差	点推定値	標準誤差
残任期間	0.307	0.748	1.309	0.642	2.478	0.869
累積危険率比	-4.212	1.419	-4.232	1.033		
与野党支持率差	-0.250	0.048	-0.302	0.037	-0.340	0.057
与野党議席率差	0.006	0.011	0.021	0.009	0.031	0.013
保守党	2.117	1.374	1.541	0.994	2.456	1.586
新首相	-2.054	1.706	0.134	1.566	1.463	2.150
定数項	1.597	3.122	-1.435	2.532	-7.097	3.171
α_E	0.402	0.189	0.358	0.182	0.693	0.179

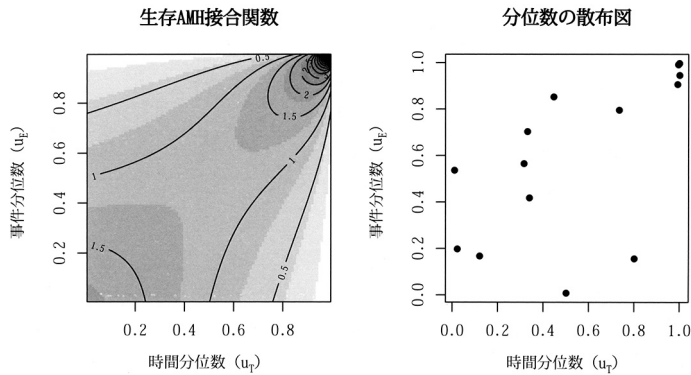


図 6. 英国戦後総選挙の時間変数と事件変数.

いることと呼応している。逆に左下隅は、それほど密度が高くなく(左図)観測も集中していない(右図)。ここからわかることは、選挙までの時間がかかる(時間変数が大きくなる)ほど与党に有利になる(事件変数が大きくなる)が、逆に選挙を早めに行うからといって必ずしも与党に不利に働くわけではないという、非対称な従属性が見られることである。言い方を変えれば、忍耐強い与党は報われると言えるが、せっかちな与党が罰されるとは限らない。こうしたニュアンスに富んだ選挙のタイミングと結果の関係は CCEHA モデルなしでは明らかにし得ないだろう。

最後に、生存 AMH 接合関数を用いた理由を述べておこう。表 3 は、様々な接合関数を用いた CCEHA モデルの赤池情報量規準を比べたものである(左列が累積危険率比を制御した場合、右列が制御しなかった場合)。最後の行は分離推定で時間モデルと単純事件モデルの赤池情報量規準値を合計した値である。事件生存クレイトン接合関数を用いた場合の赤池情報量規準が最小であるが、実は最適化が失敗している(推定が収束していない)。よって本稿は 2 番目に赤池情報量規準が小さい生存 AMH 接合関数を選んだ。これは分離推定の赤池情報量規準の総和(最終行)よりも小さい。従って、赤池情報量規準の観点からも、CCEHA による同時推定は分離推定よりも望ましい。

表 3. 様々な接合関数を用いた場合の CCEHA の赤池情報量規準(最終行のみ分離推定の合計値).

接合関数	累積危険率比		
	有り	無し	
正規	98.2	103.3	
FGM	98.0	103.1	
フランク	98.1	102.9	
クレイトン 標準	99.9	104.7	
	時間生存	99.9	104.7
	事件生存	85.3	88.9
	生存	98.4	101.2
AMH	標準	99.4	102.8
	時間生存	98.4	103.8
	事件生存	98.4	102.9
	生存	94.6	101.3
分離推定	97.9	102.7	

5. おわりに

本稿で提示した、連続変数事件接合関数生存分析モデルを適用できる研究課題は枚挙に暇がない。疾病が発見されるのが遅いほど、疾病は進行しているだろうか？ ストライキが長引くほど、高い賃上げが獲得できるだろうか？ 政党が野に下っている時間が長いほど、閣僚ポストが少なくても入閣するだろうか？

これまで、2変数の同時分布をモデル化する際は、2変量正規分布や2変量ワイブル分布のように、限られた周辺分布に対して、特定の従属性しか扱うことができなかった。あるいは生存分析の文脈では、共通脆弱性などを導入すると、周辺分布が変わってしまうし、脆弱性の分布はもとの周辺分布と共役である必要があった。それに対して接合関数の利点の1つは、各変数の周辺分布とは無関係に、変数間の従属性を扱えるという点にある。つまり、各変数の周辺分布が、例えば指数関数族のような何らかの特定のクラスに属さなくてもよいし、互いに共役分布である必要もない (Fukumoto, 2020)。従って、周辺分布や同時分布を、統計的便宜ではなく、研究対象の実態に合わせて選ぶことができるのである。

接合関数のもう1つの利点は、非対称な従属性を表現できる、ということである。本稿でも示したように、一方の変数が増えると他方の変数も増えるからと言って、一方の変数が減れば他方の変数も減るとは限らない。そうした微妙な従属性をも、接合関数は扱える。

接合関数の適用事例は、海外に目を向ければ、政治学に限っても皆無ではないが (Braumoeller et al., 2018; Chiba et al., 2015a, 2015b)、まだ広く使われているとは言い難い。本稿が社会科学に対し、接合関数の利点を伝え、広めることに貢献できれば幸いである。

注.

- ¹⁾ Fukumoto (2020)は複数の存続の従属性を軸に生存分析全般を概観する中に本稿のモデルも位置づけている。
- ²⁾ 接合関数全般の説明は、例えば Nelsen (2006)や Trivedi and Zimmer (2007)を参照。
- ³⁾ 本稿で相関とは、スピアマン順位相関やケンドール順位相関を指すことにする。詳しくは例えば Trivedi and Zimmer (2007, pp. 16, 22–23)を参照。

- 4) なお最初に紹介した3つの接合関数($C_G, C_{FGM}, C_A\{\cdot, \cdot | \varphi_F(\cdot)\}$)は、点対称なので生存接合関数に一致する($\overline{C}^{(TE)}(\cdot, \cdot | \theta_C) = C_G(\cdot, \cdot | \theta_C)$). その他の関連接合関数も媒介変数の正負を逆にすれば一致する($\overline{C}^{(T)}(\cdot, \cdot | \theta_C) = \overline{C}^{(E)}(\cdot, \cdot | \theta_C) = C_G(\cdot, \cdot | -\theta_C)$). 従って関連接合関数を別途検討する必要はない.
- 5) $\alpha_T = 0$ なので指数分布に帰着するが、後に実例でワイブル分布を用いるので、ここでワイブル分布を導入しておく.
- 6) なおここでは E の周辺分布は正規分布なので $\Phi^{-1}(u_E) = (e - \mu_E) / \exp(\alpha_E)$ である.
- 7) この乱数発生方法については Nelsen (2006, p. 41)を参照. なお Trivedi and Zimmer (2007, p. 113)は必ずしも正確ではないように見受けられる.
- 8) 左側切断があるので, \mathcal{L}_i を $1 - F_T(t_i | \theta_T, \mathbf{x}_{T,i})$ で除した商を, 観測単位 i の尤度とした.
- 9) 本稿における分析は R Core Team (2019, Version 3.5.3)を用いた. なお本稿の分析を再現するファイル一式を著者のサイト <https://www-cc.gakushuin.ac.jp/~e982440/research/replication.htm> で公開する.
- 10) t は媒介変数の真の値 θ_T を用いて $F_T^{-1}(u_T | \theta_T, \mathbf{x}_T)$ から計算している.
- 11) 元のデータと統計ソフト STATA のコードは, 2008年2月19日に <http://www.nyu.edu/gsas/dept/politics/faculty/smith/electiontimingwebpage.html> からダウンロードした. これらを公開し使用を許諾している著者の Smith 氏に感謝する.
- 12) 厳密には, この後に議会が解散され, さらにその後に選挙が行われる. なお5年の任期が満了したことはない.
- 13) 1979年に終わる議会は内閣不信任可決によるため (Smith, 2004, p. 93). 2001年から始まる議会は観測期間終了のため.
- 14) 残任期間は任期の5年から生存時間 ($t/365$) を引いた差であるから, ワイブル分布を修正したベースライン・ハザードを想定していることになる. なお Smith (2004, p. 96)には与野党議席「率」差と書かれているが, データは与野党議席「数」差である.
- 15) 但し, 元の分析結果 (Smith, 2004, p. 104, モデル 3.3, p. 141, モデル 4.5) と比べて本稿の再分析の方が, 時間モデルでは定数項の点推定値がやや大きくなっており, 事件モデルでは標準誤差がおしなべて小さくなっている. また Smith (2004, p. 104, モデル 3.3) の p は本稿の α_T に等しいが, 表 1 では実際に推定している $\log(\alpha_T)$ を報告している.
- 16) 定数項の点推定値は異なるが, 全て標準誤差が大きく, いずれにしても大きな問題ではないだろう.
- 17) 累積危険率比を制御した場合である. 実際には $\log\{(1 + \theta_C)/(1 - \theta_C)\}$ を推定しており, その点推定値と標準誤差はそれぞれ 5.007 と 2.337 である. 累積危険率比を制御しない場合はそれぞれ 4.174 と 2.154 であるが, 含意はあまり変わらない.

謝 辞

本稿は, “What Happens Depends on When It Happens: Continuous or Ordered Event History Analysis,” paper prepared for the Annual Meeting of American Political Science Association, Toronto, September 3–6, 2009 の主として前半部分をもとにしている (後半部分は Fukumoto (2015) として公刊した). 草稿は, the Annual Summer Meetings of the Society for Political Methodology, University of Michigan, July 9–13, 2008 and Yale University, July 23–25, 2009, 日本行動計量学会 (埼玉大学, 2010年9月22–25日), 応用統計学会 (統計数理研究所, 2010年5月20–21日), 研究集会「時空間現象データに対する統計科学モデルの構築及び解析に関する組織的研究」 (沖縄県青年会館, 2008年11月27–28日), 研究集会「官庁統計データの公開にお

ける諸問題の研究」(統計数理研究所, 2008年10月23–24日)で報告された。また本研究は、JSPS 科研費 JP20330023, 安倍能成記念教育基金学術研究助成金を受けており、そのことに謝意を表す。さらにコメントをいただいた、有賀賢一, 大森裕浩, Frederick Boehmke, Robert J. Franzese Jr., Adam Glynn, Jay Goodliffe, Jude Hays, Gary King, John Londregan の諸氏及び担当編集委員と匿名査読者に感謝する。

参 考 文 献

- Braumoeller, B. F., Marra, G., Radice, R. and Bradshaw, A. E. (2018). Flexible causal inference for political science, *Political Analysis*, **26**(1), 54–71.
- Chiba, D., Martin, L. W. and Stevenson, R. T. (2015a). A copula approach to the problem of selection bias in models of government survival, *Political Analysis*, **23**(1), 42–58.
- Chiba, D., Metternich, N. W. and Ward, M. D. (2015b). Every story has a beginning, middle, and an end (but not always in that order): Predicting duration dynamics in a unified framework, *Political Science Research and Methods*, **3**(3), 515–541.
- Fukumoto, K. (2009). Systematically dependent competing risks and strategic retirement, *American Journal of Political Science*, **53**(3), 740–759.
- Fukumoto, K. (2015). What happens depends on when it happens: Copula-based ordered event history analysis of civil war duration and outcome, *Journal of the American Statistical Association*, **110**(509), 83–92.
- Fukumoto, K. (2020). Duration analysis, *The SAGE Handbook of Research Methods in Political Science and International Relations* (eds. L. Curini and R. J. Franzese Jr.), 659–678, Sage, London.
- Gordon, S. (2002). Stochastic dependence in competing risks, *American Journal of Political Science*, **46**(1), 200–217.
- Nelsen, R. B. (2006). *An Introduction to Copulas*, 2nd ed., Springer, New York.
- R Core Team (2019). *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna.
- Smith, A. (2004). *Election Timing*, Cambridge University Press, Cambridge, UK.
- Trivedi, P. K. and Zimmer, D. M. (2007). *Copula Modeling: An Introduction for Practitioners*, now Publishers, Boston.

Copula-based Continuous Event History Analysis

Kentaro Fukumoto

Department of Political Studies, Faculty of Law, Gakushuin University

Scholars are interested not just in when an event happens but also in what kind of event happens. Moreover, the latter can depend on the former. If an event variable is nominal, the dependent competing risks approach is available. The model of the present paper, copula-based continuous event history analysis (CCEHA), takes advantage of copulas to model dependence between time and continuous event variables. Monte Carlo simulation illustrates that separate estimation of time and event models creates bias in the estimate of the event model when time and the event are not independent of each other (and some durations are censored). By contrast, CCEHA does not have this problem. We reanalyzed a dataset on election timing and outcome in the postwar U.K. and found an asymmetric positive dependence: later elections helped the governing party, whereas earlier elections did not necessarily hurt it.

コピュラを用いた生存時間解析

— 関連のあるエンドポイントとメタ分析の活用 —

江村 剛志¹・道前 洋史²

(受付 2019 年 4 月 24 日；改訂 2020 年 1 月 18 日；採択 1 月 20 日)

要 旨

ここ 10 年間で医学・生物学関連データベースの公開が急速に発展し、個々の患者の詳細かつ正確な情報が誰にでも利用出来るようになった。これらデータベースでは、患者の生存期間、無増悪期間、腫瘍径、遺伝子発現量など患者の極めて詳細な情報が記録されている。加えて、複数のデータソースを統合して解析する「メタ分析」の重要性がここ 10 年間で高まり、複雑かつ膨大な生存期間データを古典的な生存時間解析法だけで解析することは不十分になりつつある。本稿では、コピュラを用いて 2 つの生存期間変数をモデリングする手法について総説する。モデルの実践的な有用性を示すため、より具体的に患者の生存期間と無増悪期間の相関をコピュラでモデル化する手法を考える。また複数のデータソースを統合して、患者の生存期間と術後無増悪期間を同時解析する「メタ分析」法のためのモデルを紹介する。モデルのパラメータを推定する際に、生存期間と無増悪期間が半競合リスク関係にあることを考慮した上で最尤法を用いる必要性を解説する。最後に、個別医療の問題に関連した応用例の一つとして、遺伝子発現量と増悪情報を用いて死亡率の動的予測を行う手法を紹介する。

キーワード：遺伝子発現量，個別医療，競合リスク，動的予測，臨床試験，Cox 比例ハザードモデル。

1. はじめに

生存時間解析 (Survival analysis) は、最も歴史のある学術分野の 1 つでその起源は生命表が考案された 1600 年代にまで遡る。生存時間解析で使用されている統計手法は極めて多岐に及び、いまなお多くの研究者が学術研究を行っていることは生存時間解析のハンドブック (Klein et al., 2014) から理解できる。生存時間解析は現実問題にも幅広く応用されており、医学研究における癌患者の生存率の計算法、工場で生産される部品の信頼性の計算法、政府統計または保険数理における生命表の計算法など多くの統計手法の数理的基礎を成す学術分野といえる。

医学研究における生存時間解析では、年齢や性別などの共変量が生存期間にどのように影響するのかを Cox 回帰法 (Cox, 1972) で調べることが頻繁に行われている。Cox 回帰法は生存時間解析で最も広く使用されてきた統計手法の一つであり、これからも医学研究に必要不可欠な手法であることは疑う余地が無い。

しかしながら、ここ 10 年間で医学・生物学関連の公開データベース (癌ゲノムのデータベー

¹ 長庚大学 情報管理学系：〒333 桃園市龟山区文化一路 259 号 (台湾)；takeshiemura@gmail.com

² 北里大学 薬学部：〒108-8641 東京都港区白金 5-9-1

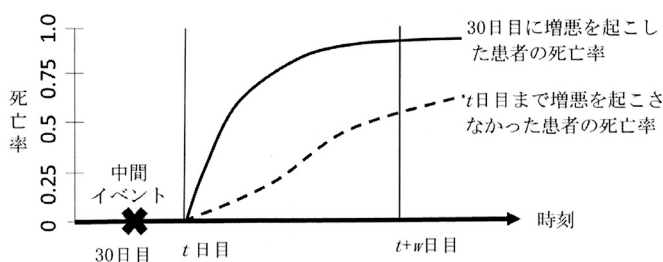


図 1. 術後 t 日目から開始する患者の死亡率の動的予測。

ス GEO や TCGA など)や統計解析フリーソフト (R など)が急速に発展し、個々の患者の詳細かつ正確な情報が誰にでも利用出来るようになった。これらデータベースでは、患者の生存期間、無増悪期間、腫瘍径、遺伝子発現量など患者の極めて詳細な情報が記録されている。加えて、複数のデータソースを統合して解析する「メタ分析」の重要性がここ 10 年間で高まり、複雑かつ膨大な生存期間データを古典的な生存時間解析法だけで解析することは不十分になりつつある。このような患者の個別データを有効利用し、個別医療の進展に貢献できる統計手法を開発することは生物・医学統計学者の重要な使命でもある。

本稿では、コンピュータを用いて 2 つの生存期間変数をモデリングする手法について総説する。モデルの実践的な有用性を示すため、より具体的に患者の生存期間と無増悪期間の相関をコンピュータでモデル化する手法を考える。また複数のデータソースを統合して、患者の生存期間と無増悪期間を同時解析する「メタ分析法」のためのモデルを紹介する。モデルのパラメータを推定する際、生存期間と無増悪期間が半競合リスク関係 (Fine et al., 2001)にあることを考慮した上で最尤法を用いる必要性を解説する。最後に、個別医療の問題に関連した応用例の一つとして、遺伝子発現量と増悪情報を用いて死亡率の動的予測を行う手法を紹介する (Emura et al., 2018)。すなわち、患者個々の術後経過と遺伝子情報を考慮した死亡率の個別予測式をメタ分析のモデルから導出し、患者レベルの死亡予測を行う手法 (図 1)を紹介する。このような動的予測の手法は多く提案されている (Sène et al., 2016; Proust-Lima et al., 2014; Rondeau et al., 2017 など)が、本稿で紹介する手法はメタ分析を使用する点と、コンピュータを利用する点で既出のものとは大きく異なる。

本稿は以下のように構成されている。2 節では生存時間解析における重要な用語「エンドポイント」を解説し、なぜコンピュータを利用するのかを癌の臨床研究の側面から説明する。3 節では、生存時間解析におけるコンピュータのモデルを定式化し、提案されているいくつかのモデルと統計手法を紹介する。4 節では Frailty モデルを用いたメタ分析の考え方を紹介する。5 節では Frailty モデルにコンピュータを組み合わせた統計モデルを紹介し、エンドポイント間の半競合リスク関係を解説する。6 節では個別医療の問題に関連した動的予測の手法を紹介する。5 節、6 節では卵巣癌患者のデータを用いた事例を用いて、統計手法の利用法を解説する。

2. 生存期間のエンドポイント

医学研究における生存時間解析の基本的な目的の 1 つは、特定の治療や予後因子が、患者の生存期間にどれだけ影響するか調べることである。医学研究において、生存期間はエンドポイントと呼ばれ、具体的には次のような複数の定義がある。

- 全生存期間 (OS: Overall survival)

- 無増悪期間 (TTP: Time to tumor progression)
- 無増悪生存期間 (PFS: Progression-free survival)
- 無病生存期間 (DFS: Disease-free survival)

増悪とは癌の再発、腫瘍増大、遠隔転移などを意味する。無増悪期間とは増悪がなかった期間である。例えば、手術後に残った直径 1cm の腫瘍がそのまま増大せず、かつ遠隔転移が見られなければ、無増悪とみなされる。

エンドポイントの定義は上記以外にもあるが、いずれにせよ定義が明確であり、かつ臨床的に重要な転帰を表すものでなければならない (Pazdur, 2008; Piedbois and Crowell, 2008; Le Tourneau et al., 2009)。医学研究者や臨床医は、これらエンドポイントから重要なものを選び、個々の患者に対して追跡・観察する。

治療効果や予後因子がこれらエンドポイントにどのように影響するのかを調べるためには通常 Cox 比例ハザードモデル (Cox, 1972) を用いる。例えば、生存期間を応答変数、治療群と予後因子を共変量とし、部分尤度法により共変量がハザード(瞬間死亡率)に与える影響を調べることが出来る。このような共変量の効果を調べる解析は、医学研究において極めて頻繁に行われている。

全生存期間は癌の研究において最も重要かつ明確に定義できるエンドポイントであり、真のエンドポイント (true endpoint) またはゴールドスタンダードエンドポイント (Pazdur, 2008; Shi and Sargent, 2009; Michiels et al., 2009; Oba et al., 2013) と呼ばれる。例えば、頭頸部癌の化学療法や放射線療法の効果を評価する際の最も重要なエンドポイントは全生存期間である (Michiels et al., 2009; Le Tourneau et al., 2009)。

無増悪期間、無増悪生存期間、無病生存期間などのエンドポイントはしばしば代替エンドポイント (surrogate endpoint) とよばれ、全生存期間に次いで重要なエンドポイントである。代替エンドポイントを使用する場合、“増悪”の判定は臨床医の恣意性が伴うため、RECIST ガイドラン (Eisenhauer et al., 2009) などに従ってすすめることが推奨されている。

多くの臨床研究では、患者から 2 つのエンドポイント (例えば、全生存期間と無増悪生存期間) を観察するためのフォローアップを行う。その後、得られたデータをもとに、Cox 回帰法などで治療効果や予後因子が各エンドポイントに与える影響を解析する。しかしながら、単純な Cox 回帰では 1 つのエンドポイントを応答変数とした周辺解析しか行うことができず、2 つ以上のエンドポイントを同時解析する場合には不十分である。

例えば、ある治療が生存期間と無増悪期間の両方に与える影響に興味があるとする。この場合、生存期間を応答変数とした Cox 回帰と、無増悪期間を応答変数とした周辺 Cox 回帰を別々に実行することが考えられるが、これらの 2 つの解析結果からは、全生存期間と無増悪期間の相関の推定値は得られない。無増悪期間が全生存期間とどのように関連しているかを調べるには相関の解析が必要であり、このような相関の情報は抗癌剤開発の観点からも重要である (Sherrill et al., 2008; Michiels et al., 2009; Rondeau et al., 2015)。加えて、無増悪期間が全生存期間の代替エンドポイントであることを実証したい場合や、無増悪期間を全生存期間の予後予測に利用する場合に、相関の推定値が必要であることを考慮すると、周辺 Cox 回帰法はデータを有効に活用しているとはいえない。

3. コピュラの生存時間モデル

コピュラは 2 つの確率変数間の相関構造をあらわす関数である。コピュラ概念は数学者 Abe Sklar により導入された (Sklar, 1959)。コピュラの数学的な定義と性質は Nelsen (2006) の著書に良くまとめられている。これは数学書であるが、大学初年度の微積分の知識で読める良

書である。以降では、コピュラを生存時間モデルに応用するための基礎知識や用語を概説する。

3.1 2次元生存時間モデル

コピュラは2つの生存期間変数間の相関構造をモデリングするために用いることが出来る。いま X と Y を生存期間変数とし、 $\mathbf{Z}_1 = (Z_{11}, \dots, Z_{1p_1})$ と $\mathbf{Z}_2 = (Z_{21}, \dots, Z_{2p_2})$ を X と Y それぞれに関連する共変量とする。ここで \mathbf{Z}_1 と \mathbf{Z}_2 は $\mathbf{Z}_1 = \mathbf{Z}_2$ であっても良い。また $S_X(x|\mathbf{Z}_1) = \Pr(X > x|\mathbf{Z}_1)$ と $S_Y(y|\mathbf{Z}_2) = \Pr(Y > y|\mathbf{Z}_2)$ を X と Y の生存関数とする。このとき、 X と Y の2次元生存関数を次のように定義する。

$$(3.1) \quad \Pr(X > x, Y > y|\mathbf{Z}_1, \mathbf{Z}_2) = C_\theta\{S_X(x|\mathbf{Z}_1), S_Y(y|\mathbf{Z}_2)\}.$$

ここで C_θ はコピュラ (Sklar, 1959; Nelsen, 2006) で、 θ は X と Y の相関の強さを定めるパラメータである。コピュラの種類は無数にあるが、本稿で主に扱うコピュラは次の3つである。

- 独立型: $C(u, v) = uv$, $0 \leq u, v \leq 1$,
- クレイトン型 (Clayton, 1978):

$$C_\theta(u, v) = (u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}, \quad \theta > 0, \quad 0 \leq u, v \leq 1,$$

- ゲンベル型 (Gumbel, 1960):

$$C_\theta(u, v) = \exp[-\{(-\log u)^{\theta+1} + (-\log v)^{\theta+1}\}^{\frac{1}{\theta+1}}], \quad \theta \geq 0, \quad 0 \leq u, v \leq 1.$$

独立型のコピュラを選ぶと、 $\Pr(X > x, Y > y|\mathbf{Z}_1, \mathbf{Z}_2) = S_X(x|\mathbf{Z}_1)S_Y(y|\mathbf{Z}_2)$ となり、 X と Y は共変量を所与したときの条件付独立なモデルとなる。コピュラの定義により、上記3つの C_θ はどれもコピュラであるから、 C_θ は2次元分布関数であり、その周辺分布は $[0, 1]$ 上の一様分布である。実際、周辺分布は $C_\theta(u, 1) = u$ と $C_\theta(1, v) = v$ であり、これらは $[0, 1]$ 上の一様分布の分布関数である。

C_θ が2次元分布関数であるから、 $\Pr(U \leq u, V \leq v) = C_\theta(u, v)$ となる一様分布の確率変数のペア (U, V) を考えることが出来る。このとき、確率変数のペア (X, Y) を変換 $X = S_X^{-1}(U|\mathbf{Z}_1)$ と $Y = S_Y^{-1}(V|\mathbf{Z}_2)$ で定義すると、 (X, Y) の分布は式 (3.1) を満足することがわかる。これは、確率変数のペア (U, V) が相関構造のみを定め、 $S_X(\cdot|\mathbf{Z}_1)$ と $S_Y(\cdot|\mathbf{Z}_2)$ が周辺構造のみを定めていることを意味する。この周辺構造と相関構造が分離できることは統計モデリングの観点から極めて有用である。

上記の3つのコピュラの例は、どれも下記で定義するアルキメデス型コピュラ (Archimedean copula) に属することがわかる。いま、関数 $\phi_\theta: [0, 1] \mapsto [0, \infty]$ を連続な減少関数で $\phi_\theta(1) = 0$ を満足するとする。このとき、アルキメデス型コピュラは

$$C_\theta(u, v) = \phi_\theta^{-1}\{\phi_\theta(u) + \phi_\theta(v)\},$$

で定義される。ここで ϕ_θ は生成素 (generator) と呼ばれ、その逆関数を ϕ_θ^{-1} で表す。クレイトン型は $\phi_\theta(t) = (t^{-\theta} - 1)/\theta$, $\theta > 0$ である。その極限は $\lim_{\theta \rightarrow 0} \phi_\theta(t) = -\log(t)$ であり、これは独立型の生成素になっていることがわかる。ゲンベル型は $\phi_\theta(t) = \{-\log(t)\}^{\theta+1}$ で生成され、 $\theta = 0$ が独立型の生成素になる。すなわち、クレイトン型とゲンベル型では極限が $\lim_{\theta \rightarrow 0} C_\theta(u, v) = uv$ となっているので、独立型に帰着する。表 1 にコピュラとその生成素をまとめた。

3.2 ケンドール順位相関係数

ケンドール順位相関係数(以下、ケンドール相関と記す)は X と Y の相関を表す指標である。

表 1. コピュラの例.

コピュラの型	パラメータ	生成素 $\phi_\theta(t)$	ケンドール相関 τ_θ	オッズ比 $r_\theta(s) = -3\dot{\phi}_\theta(s)/\phi_\theta(s)$
独立型	なし	$-\log(t)$	0	1
クレイトン型	$\theta > 0$	$(t^{-\theta} - 1)/\theta$	$\theta/(\theta + 2)$	$1 + \theta$
グンベル型	$\theta \geq 0$	$\{-\log(t)\}^{\theta+1}$	$\theta/(\theta + 1)$	$1 - \theta/\log(s)$

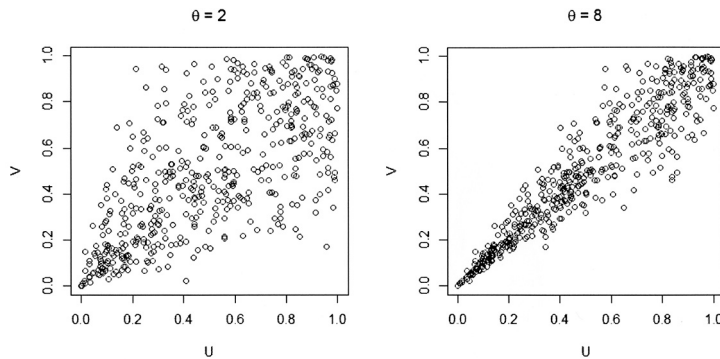


図 2. 500 組の (U, V) をクレイトン型コピュラから生成し、散布図にしたもの.

モデル (3.1) のもとで、 X と Y のケンドール相関はコピュラ C_θ のみに依存することがわかり、

$$\tau_\theta = 4 \int_0^1 \int_0^1 C_\theta(u, v) dC_\theta(u, v) - 1$$

と書ける. ケンドール相関は周辺分布の構造 $S_X(x|Z_1)$ と $S_Y(y|Z_2)$ に依存しない. よってコピュラ C_θ の型を定めれば、 τ_θ はパラメータ θ によって定まる. アルキメデス型コピュラでは、ケンドール相関は生成素を使って次のように計算できる.

$$\tau_\theta = 1 + 4 \int_0^1 \frac{\phi_\theta(t)}{\dot{\phi}_\theta(t)} dt, \quad \dot{\phi}_\theta(t) = d\phi_\theta(t)/dt.$$

この式を用いると、クレイトン型では $\tau_\theta = \theta/(\theta + 2)$ 、グンベル型では $\tau_\theta = \theta/(\theta + 1)$ となり、ともに無相関 $\tau_0 = 0$ から正の完全相関 $\tau_\infty = 1$ までとりうる.

図 2 は 500 組の (U, V) をクレイトン型から生成し、散布図にしたものである. 図から U と V との間に正の相関があることがわかる. 相関の程度は $\theta = 2$ ($\tau_\theta = 0.5$) よりも $\theta = 8$ ($\tau_\theta = 0.8$) が強いことがわかる. また、 $U \approx 0$ と $V \approx 0$ 付近でより強い相関を示し、これは下側裾従属性 (lower tail dependence) と呼ばれるクレイトン型コピュラが有する特徴的な現象である. 一方、グンベル型では $U \approx 1$ と $V \approx 1$ 付近でより強い相関を示し、上側裾従属性 (upper tail dependence) が現れる.

モデル (3.1) のもとで X と Y を生成するには、 $X = S_X^{-1}(U|Z_1)$ と $Y = S_Y^{-1}(V|Z_2)$ で変換すれば良い. 変換 S_X^{-1} と S_Y^{-1} は減少関数であるため、変換により裾従属の位置は反転する. すなわち、クレイトン型では X と Y 間に上側裾従属、グンベル型では X と Y 間に下側裾従属があらわれる.

3.3 オッズ比による相関

生存時間解析においては、 2×2 クロス集計表と関連した相関の指標が、推定や予測の局面で有用となることがある。例えば、死亡と増悪の関連をみるとき、死亡と増悪のあり・なしを 2×2 クロス集計表に集計し、そのオッズ比を計算して相関の度合いを調べることがある。このオッズ比の考え方は Clayton (1978) が提案した 2 次元生命表に遡り、Oakes (1989) がより明確に定式化した。

いま X を無増悪期間、 Y を全生存期間とし、モデル (3.1) を仮定する。オッズ比関数 (英語では Cross-ratio function) を次のように定義する。

$$\frac{\Pr(X = x, Y = y | Z) \Pr(X > x, Y > y | Z)}{\Pr(X = x, Y > y | Z) \Pr(X > x, Y = y | Z)} = R_\theta \{S_X(x | Z_1), S_Y(y | Z_2)\}$$

左辺はオッズ比の定義であり、表 2 の 2×2 クロス集計表に基づく。右辺は、モデル (3.1) から導出された結果であり、

$$R_\theta(u, v) = \frac{C_\theta^{[1,1]}(u, v) C_\theta(u, v)}{C_\theta^{[1,0]}(u, v) C_\theta^{[0,1]}(u, v)}$$

はコピュラのオッズ比関数である。ここで

$$C_\theta^{[1,0]}(u, v) = \frac{\partial}{\partial u} C_\theta(u, v), \quad C_\theta^{[0,1]}(u, v) = \frac{\partial}{\partial v} C_\theta(u, v),$$

はコピュラの条件付分布関数である。すなわち $U = u$ の条件下の V の分布関数は $\Pr(V \leq v | U = u) = C_\theta^{[1,0]}(u, v)$ であり、 $V = v$ の条件下の U の分布関数は $\Pr(U \leq u | V = v) = C_\theta^{[0,1]}(u, v)$ である。また、

$$C_\theta^{[1,1]}(u, v) = \frac{\partial^2}{\partial u \partial v} C_\theta(u, v)$$

はコピュラの密度関数を表す。

関数 $R_\theta(u, v)$ は点 (u, v) における局所的な相関を表し、次のように解釈できる。

- $R_\theta(u, v) > 1$; 局所正相関,
- $0 < R_\theta(u, v) < 1$; 局所負相関,
- $R_\theta(u, v) = 1$; 局所独立.

独立型コピュラのもとで $R_\theta(u, v) = 1$ が全ての $0 \leq u \leq 1$ と $0 \leq v \leq 1$ で成立する。クレイトン型コピュラではオッズ比関数が定数 $R_\theta(u, v) = 1 + \theta$ となり、これは原論文 Clayton (1978) に由来する。アルキメデス型コピュラのオッズ比関数は次のように書ける

$$R_\theta(u, v) = r_\theta \{C_\theta(u, v)\}, \quad r_\theta(s) = -s \ddot{\phi}_\theta(s) / \dot{\phi}_\theta(s), \quad \ddot{\phi}_\theta(t) = d^2 \phi_\theta(t) / dt^2.$$

オッズ比関数 R_θ が相対リスク (Relative risk) として解釈できることは生存時間解析において極めて有用である。例えば、 X を無増悪期間、 Y を全生存期間とし、増悪が生存期間に与える影響に興味があるとする。ある時刻 x において増悪したか否かの情報を共変量と考え、次の 2 つの条件付ハザード関数を定義する。

表 2. 2×2 クロス集計表。A, B, C, D はイベントが起こる頻度、オッズ比は AD/(BC)。

	死亡; $Y = y$	生存; $Y > y$
増悪あり; $X = x$	A	B
増悪なし; $X > x$	C	D

- $\lambda_Y(y|X = x, \mathbf{Z}) = \Pr(y \leq Y < y + dy | Y \geq y, X = x, \mathbf{Z})/dy$
時刻 x で増悪したときの条件付ハザード,
- $\lambda_Y(y|X > x, \mathbf{Z}) = \Pr(y \leq Y < y + dy | Y \geq y, X > x, \mathbf{Z})/dy$
時刻 x 以前に増悪しなかったときの条件付ハザード.

相対リスクは2つの条件付ハザード関数の比で定義される;

$$\frac{\lambda_Y(y|X = x, \mathbf{Z})}{\lambda_Y(y|X > x, \mathbf{Z})} = R_\theta \{S_X(x|\mathbf{Z}_1), S_Y(y|\mathbf{Z}_2)\}.$$

右辺を導出するには、モデル(3.1)の仮定が必要である。 $R_\theta > 1$ であれば、増悪したことが死亡リスクを上昇させることを意味する。クレイトン型コピュラでは、相対リスクが定数 $1 + \theta$ なので、これは比例ハザード性を意味する。また、 $\theta > 0$ なので、増悪と死亡が任意の時刻 (x, y) において局所正相関をもつモデルを表す。

Clayton (1978) は、データを表2の 2×2 クロス集計表にまとめ、そのオッズ比を計算することにより θ を推定することを提唱している。この手法は X と Y が共に右側打ち切りを受けている場合にも応用でき、さらに周辺分布の構造 $S_X(x|\mathbf{Z}_1)$ と $S_Y(y|\mathbf{Z}_2)$ に依存しないセミ・パラメトリックな推定法である。この手法はアルキメデス型コピュラに一般化可能であり (Emura et al., 2010), さらに半競合リスクデータ (Wang, 2003) や従属切断データ (Emura and Wang, 2010; Emura et al., 2011) にも適用可能である。

3.4 応用例

コピュラは生存時間のモデルに広く用いられてきたが、その中でもとりわけ臨床医学研究において広く利用されているのが Burzykowski et al. (2001) による2段階推定法 (two-step method) である。代替エンドポイントを用いる臨床試験の策定にあたり、現時点ではコピュラを用いた2段階推定法 (Burzykowski et al., 2005) が広く用いられている。最近、この2段階推定法を実行する R パッケージ (Rotolo et al., 2018) が開発され、利用が容易になった。この R パッケージではクレイトン型、グンベル型、プラケット型コピュラが利用できる。

いま、 i 番目の研究単位から得られた j 番目の患者に対して、 X_{ij} を無増悪期間、 Y_{ij} を全生存期間とする。Burzykowski et al. (2001) は X_{ij} と Y_{ij} の周辺ハザードに

$$\begin{cases} \lambda_{X_{ij}}(t|Z_{ij}) = \lambda_{X_i}(t) \exp(\beta_{X_i} Z_{ij}) \\ \lambda_{Y_{ij}}(t|Z_{ij}) = \lambda_{Y_i}(t) \exp(\beta_{Y_i} Z_{ij}) \end{cases}$$

のような Cox モデルを仮定した。ここで Z_{ij} は治療群を表す変数、 β はその回帰係数、 $\lambda_i(\cdot)$ は i 番目の研究単位の基準ハザード関数である。さらに、 X_{ij} と Y_{ij} の同時分布を、

$$\Pr(X_{ij} > x, Y_{ij} > y) = C_\theta \left[\exp \left\{ - \int_0^x \lambda_{X_{ij}}(t|Z_{ij}) dt \right\}, \exp \left\{ - \int_0^y \lambda_{Y_{ij}}(t|Z_{ij}) dt \right\} \right]$$

で定義し、 C_θ はコピュラとした。パラメータ推定の方法はいくつかある。例えば、まず通常の Cox 回帰法 (部分尤度法) やワイブル回帰法で $(\beta, \lambda_i(\cdot))$ を推定してから、次に θ を推定する手法がある。他にも、 $(\theta, \beta_{X_i}, \beta_{Y_i}, \lambda_{X_i}(\cdot), \lambda_{Y_i}(\cdot))$ を同時推定する全尤度法がある。さらに基準ハザードの共通性 $\lambda_{X_i}(\cdot) = \lambda_{X_0}(\cdot)$ と $\lambda_{Y_i}(\cdot) = \lambda_{Y_0}(\cdot)$ を仮定して推定する場合もある。いずれにせよ、独立な右側打ち切りのある生存期間データに適用可能である。詳細は Burzykowski et al. (2001, 2005), Rotolo et al. (2018) を参照されたい。

上記手法において、エンドポイント間の相関を考慮しているにもかかわらず推定の段階において、独立右側打ち切りしか想定していない点に注意すべきである。例えば、無増悪期間を

X_{ij} として β_{X_i} を Cox 回帰法(部分尤度法)で推定する場合, 死亡を独立右側打ち切りとみなして解析していることになる. しかし死亡による打ち切りは, 無増悪期間と明らかに独立でないため, 多くの患者が早期に(増悪する前に)死亡している場合には回帰推定値に大きなバイアスが入る. この問題は従属打ち切りの問題として知られ, 打ち切り数が多いほど, また相関が強いほどバイアスが大きくなることが知られている (Emura and Chen, 2016, 2018).

この従属打ち切り問題を統計的に正しく扱うには, データ構造を半競合リスク (Fine et al., 2001)の枠組みで捉えた統計手法を用いると良い. すなわち, 死亡は増悪の観測を打ち切るため競合リスクであるが, 増悪は死亡の観測を打ち切ることがないため競合リスクでないとする考えである. 半競合リスクデータの枠組みで正しく回帰推定値を得る場合には, 部分尤度法 (Cox 回帰法)ではなく, 全尤度を利用した推定法 (Chen, 2012; Emura et al., 2017)を利用することになる. 半競合リスクデータを扱う統計手法は 5 節で再考する.

上記以外にも, コピュラは極めて多くの生存時間解析法で利用されてきた. 例えば, クラスタ生存時間解析 (Rotolo et al., 2013; Emura et al., 2017; Peng et al., 2018), 競合リスク解析 (Lo and Wilke, 2010; de Uña-Álvarez and Veraverbeke, 2013, 2017; Shih and Emura, 2018; Shih et al., 2019; Zhou et al., 2019; Emura et al., 2019c; Wang et al., 2020), 半競合リスク解析 (Wang, 2003; Chen, 2012; Peng, 2019; Peng and Xiang, 2019), 従属打ち切り下での生存時間解析 (Emura and Michimae, 2017; Moradian et al., 2019; Emura and Chen, 2016, 2018), 従属切断下での生存時間解析 (Chaieb et al., 2006; Emura and Wang, 2012; Emura and Murotani, 2015; Emura and Pan, 2020), ヴァインコピュラを用いた多変量生存時間データ (Barthel et al., 2018), 繰り返しイベント時間解析などである (Ling et al., 2016; Li et al., 2019, 2020).

4. メタ分析の活用

メタ分析とは, 複数の研究単位から集められたデータを統合し, より一般的な母集団に対する推論を行う手法である. メタ分析は出版された複数の研究論文から要約統計量(相対リスクなど)を抽出して統合解析を行う手法を指すこともあるが, 本稿でいうメタ分析とは患者個人のデータを用いる, いわゆる IPD メタ分析 (Individual-patient data meta-analysis)を指す. 複数の研究単位から集めた患者の個別データに基づく統合解析は, 全生存期間と代替エンドポイントの相関解析に不可欠とされている (Burzykowski et al., 2001, 2005; Shi and Sargent, 2009; Rotolo et al., 2018; Sofeu et al., 2019, 2020).

IPD メタ分析を通じて全生存期間とその代替エンドポイントの相関を推定する研究が数多く行われている. 大腸癌 (Buyse et al., 2008), 頭頸部癌 (Michiels et al., 2009), 胃癌 (Oba et al., 2013), 結腸癌 (Alonso and Molenberghs, 2008)の患者において全生存期間と無増悪生存期間の間に強い正相関があることが報告されている一方, 乳癌の患者においては正相関がそれほど強くないことが報告されている (Michiels et al., 2016). このように全生存期間と無増悪生存期間の間に意味のある相関があるかどうかは癌種や進行度などにより異なる. これらのメタ分析において, 相関の推定にはコピュラが用いられてきた(3.4 節).

推定効率の点からも, 2つ以上のエンドポイントをコピュラでモデル化した同時解析には利点がある. 実際, Riley (2009)は, 連続変量の多変量メタ分析において, 同時モデルで推定した回帰係数が, 周辺モデルで推定したものより推定精度が良く(推定量の平均二乗誤差が小さく)なるとしている.

ほとんどのメタ分析においては, 研究単位間の異質性を考慮した解析が行われる. 一般的に, 異なる時期, 異なる地域・国で行われた研究を統合する場合, 試験間で, 参加者の背景, エンドポイントや予後因子の定義など, さまざまな要因が厳密には異なることが一般的であ

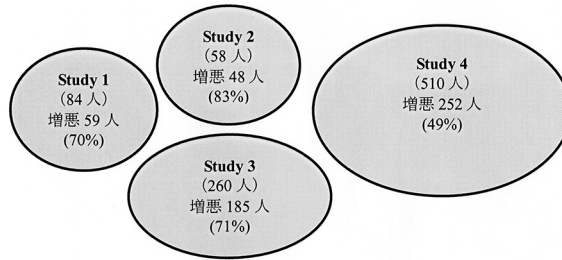


図 3. 卵巣癌患者を 4 つのデータソースから集めたメタ分析のデータ (Ganzfried et al., 2013; Emura et al., 2018). データは R パッケージ *joint.Cox* (Emura, 2019) で入手可能.

る. また, 設備が整い経験豊富なスタッフが揃った国は, そうでない国より, 患者の死亡や増悪リスクが低いと考えられる. ほとんどのメタ分析の例において, このような研究単位間の異質性は, 患者から観測されている共変量で説明しきれないことが多く, これを観測不能なランダム効果として解析モデル内に取り入れる. このような解析モデルをランダム効果モデル (Random-effects model) と呼ぶ.

図 3 は 4 つの研究単位から集められた患者をもとにしたメタ分析のためのデータの例を示している. データは合計 912 人の外科手術後の卵巣癌患者を追跡調査し, 増悪があったかどうかを調べたものである. 図 3 から増悪率が研究単位間で大きく異なることがわかる. 例えば, 最も高い研究単位で 83% (Study 2), 最も低いもので 49% である (Study 4). この増悪率の差異は追跡期間にもよるため, Study 2 が Study 1 より増悪率が高いとは断言できず, 増悪率の異質性はデータから容易に観測できるものではない.

生存時間解析において, ランダム効果モデルに相当するものが, Frailty モデルである. とりわけ Shared frailty といわれるモデルにおいては, 研究単位内の全ての患者で共通した (Share された) 観測されない因子が存在することを仮定し, この因子の分散パラメータで異質性の度合いをモデル化する (Duchateau et al., 2002; Rondeau et al., 2015).

図 3 のデータを解析する統計手法を簡単に紹介する. G 個の研究単位を考え, i 番目の研究単位が N_i 人の患者を含むとする ($i = 1, 2, \dots, G$). i 番目の研究単位の j 番目の患者 ($j = 1, 2, \dots, N_i$) に対し, X_{ij} をイベント時間, \mathbf{Z}_{ij} を共変量ベクトルと定義する. 図 3 の例では, $G = 4$ で $N_1 = 84, N_2 = 58, N_3 = 260, N_4 = 510$ であり, 合計 912 人となる.

Shared frailty モデルでは, i 番目の研究単位の j 番目の患者 ($j = 1, 2, \dots, N_i$) のハザード関数を次のようにモデル化する

$$\lambda_{ij}(t|u_i, \mathbf{Z}_{ij}) = u_i \lambda_0(t) \exp(\beta' \mathbf{Z}_{ij}).$$

ここで $\beta = (\beta_1, \dots, \beta_p)'$ は未知の回帰係数, $\lambda_0(\cdot)$ は未知の基準ハザード関数, $u_i > 0$ ($i = 1, 2, \dots, G$) は観測不能な Frailty 項である. 上記モデルにおいて, Frailty 項 u_i が全ての患者で共有されていることが Shared frailty と呼ばれる由来である.

Frailty 項 $u_i > 0$ は確率変数の実現値とみなす. 以下では, 最も広く使用される平均, 分散のガンマ分布を用いる. すなわち, 確率密度関数が

$$(4.1) \quad f_\eta(u) = \frac{1}{\Gamma(1/\eta)\eta^{1/\eta}} u^{1/\eta-1} \exp\left(-\frac{u}{\eta}\right), \quad \eta > 0, \quad u > 0$$

であると仮定する. 他の研究単位と比較して高リスクな研究単位は $u_i > 1$, 低リスクな研究単

位は $0 < u_i < 1$ で与えられ、その分散 η は研究単位間の異質性の度合いを決めるパラメータである。極限は $\eta \rightarrow 0$ 研究単位間に異質性が無いことを意味する。実現値 u_i は観測することが出来ないが、その分散 η はデータから推定することが出来る。

パラメータ $(\eta, \beta, \lambda_0(\cdot))$ を最尤法で推定する様々な方法が提案されており、そのうち Frailty 分布としてガンマ分布や対数正規分布を使用する場合は、R のパッケージ等を利用して最尤推定量の計算が可能である (Hirsch and Wienke, 2012; Rondeau and Gonzalez, 2005)。

実データではイベント時間 X_{ij} は右側打ち切りにより観測出来ないケースが多い。そのため、右側打ち切り時間 C_{ij} を定義し、観測される時間を $T_{ij} = \min(X_{ij}, C_{ij})$ とする。そのため i 番目の研究単位の j 番目の患者に関するデータは $(T_{ij}, \delta_{ij}, \mathbf{Z}_{ij})$ となり、 $\delta_{ij} = I(T_{ij} = X_{ij})$ はイベントが起こったか否かを示す指標である。

パラメータ $(\eta, \beta, \lambda_0(\cdot))$ の推定はデータ $\{(T_{ij}, \delta_{ij}, \mathbf{Z}_{ij}); i = 1, 2, \dots, G; j = 1, 2, \dots, N_i\}$ が与えられたときの尤度関数を構成し、対数尤度関数を $(\eta, \beta, \lambda_0(\cdot))$ に関して最大化することによって行う (詳しくは Emura et al., 2019b の 2.5 節を参照)。

5. メタ分析による 2 つのエンドポイントの同時解析

本節では 2 つのエンドポイントを同時に解析するための方法を紹介する。とりわけ、無増悪期間と全生存期間を同時解析し、その関連を調べるために患者の個別データ (Individual-Patient-Data, IPD) を用いたメタ分析を考える。このような分析法は「IPD メタ分析」と呼ばれており、要約データに基づくメタ分析 (個別データを用いない分析) と本質的に異なる。3 節で示したようにエンドポイント間の相関をコンピュータでモデル化し、4 節で示したように研究単位間の異質性を Frailty 項でモデル化するために Joint frailty-copula モデル (Emura et al., 2017) を導入する。

5.1 エンドポイントの測定と半競合リスク

複数のエンドポイントを扱う統計手法では、エンドポイントがどのような打ち切りを受けているか明確にする必要がある。これは、データ構造を明らかにするだけでなく、尤度関数を正しく構成するために必要なプロセスである。

G 個の研究単位を考え、 i 番目の研究単位が N_i 人の患者を含むとする ($i = 1, 2, \dots, G$)。 i 番目の研究単位の j 番目の患者 ($j = 1, 2, \dots, N_i$) に対し、

- X_{ij} = 無増悪期間,
- D_{ij} = 全生存期間,
- C_{ij} = 右側打ち切り期間 (独立かつ無情報な打ち切りを仮定),

と定義する。患者は 3 つの変数 X_{ij} , D_{ij} , C_{ij} を有しているが、どれが観測されるかは、患者の状態や打ち切り期間によって決まる。例えば、患者の状態が安定していれば、 X_{ij} と D_{ij} は観測されず、打ち切り期間 C_{ij} のみが観測される。この場合、打ち切り期間 C_{ij} までに増悪と死亡が起こらなかった情報が記録される。

図 4 で患者がとりうる 4 つのケース (A~D) を示す。ケース A では、患者が増悪し、その後死亡しているので、 X_{ij} と D_{ij} が観測される。ケース B では、患者は増悪したが、死亡する前に打ち切りとなっているので、 X_{ij} と C_{ij} が観測される。すなわち、 D_{ij} が C_{ij} によって右側打ち切りされている。ケース C では、患者が増悪せずに死亡しているので、 D_{ij} のみが観測される。この場合、 X_{ij} が D_{ij} によって右側打ち切りされている。 X_{ij} と D_{ij} に相関があるので、この打ち切りは通常の「独立打ち切り」でないことに注意する。ケース D では、 X_{ij} も D_{ij} も観

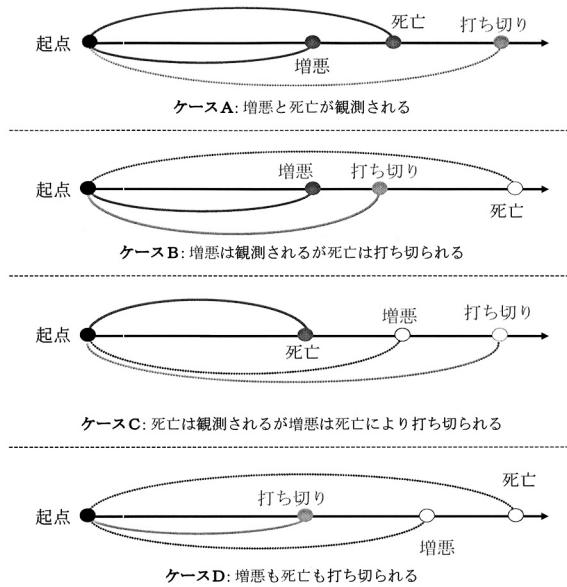


図 4. 患者がとりうる 4 つのケース (A~D). 観測されるイベントは●, 観測されないイベントは○で表す. 観測される期間は実線, 観測されない期間は点線で表す.

測されず, 打ち切り期間 C_{ij} のみが観測される. すなわち, X_{ij} も D_{ij} も右側打ち切りされている.

まとめると i 番目の研究単位の j 番目の患者 ($j = 1, 2, \dots, N_i$) に対して観測されるデータは $(T_{ij}, T_{ij}^*, \delta_{ij}, \delta_{ij}^*, Z_{1,ij}, Z_{2,ij})$ と書ける. ここで,

- $T_{ij} = \min(X_{ij}, D_{ij}, C_{ij})$; 最初に起こったイベントの時間,
- $\delta_{ij} = I(T_{ij} = X_{ij})$; 増悪の指標 (無増悪 = 0; 増悪 = 1),
- $T_{ij}^* = \min(D_{ij}, C_{ij})$; 最後に起こったイベントの時間,
- $\delta_{ij}^* = I(T_{ij}^* = D_{ij})$; 死亡の指標 (生存 = 0; 死亡 = 1),
- $Z_{1,ij}$; 増悪に関連する p_1 次元共変量,
- $Z_{2,ij}$; 死亡に関連する p_2 次元共変量.

ペア $(\delta_{ij}, \delta_{ij}^*)$ は上記 A, B, C, D の 4 つのケースを区別するための情報であり, $(\delta_{ij}, \delta_{ij}^*) = (1, 1), (1, 0), (0, 1), (0, 0)$ のいずれかを取りえる (表 3). 例えば, ケース A は $(\delta_{ij}, \delta_{ij}^*) = (1, 1)$ に対応し, この場合 $T_{ij} = X_{ij}$ と $T_{ij}^* = D_{ij}$ が観測される.

上記のような構造をもつ生存期間データは「半競合リスクデータ」と呼ばれている (Fine et al., 2001). これは, 死亡のような終着イベント (terminal event) が増悪のような非終着イベント (non-terminal event) の競合リスクになっていることに由来する. ここで「終着イベント」とは, それが起こった後の一切の情報が得られなくなるようなイベントを指す. 死亡は明らかに終着イベントである. 増悪後には死亡時刻などの情報が得られるため, 増悪は終着イベントでない. 半競合リスクデータでは増悪だけでなく, より一般の中間イベント (非終着イベント) を考えることが出来る. 半競合リスクデータの統計解析法の最新の動向は Peng (2019) を参照されたい.

表 3. 患者がとりうる 4 つのケース (A~D).

	最初のイベント	最後のイベント	T_{ij}	T_{ij}^*	δ_{ij}	δ_{ij}^*
ケース A	増悪	死亡	X_{ij}	D_{ij}	1	1
ケース B	増悪	打ち切り	X_{ij}	C_{ij}	1	0
ケース C	死亡	死亡	D_{ij}	D_{ij}	0	1
ケース D	打ち切り	打ち切り	C_{ij}	C_{ij}	0	0

5.2 Joint frailty-copula モデル

Rondeau et al. (2015) は研究単位間の異質性を考慮した半競合リスクデータに基づくメタ分析の手法を提案している. 彼女らが提案している Joint frailty モデルにおいて, Frailty 項 $u_i > 0$ は $\text{Var}_\eta(u_i) = \eta$ のガンマ分布のモデル (4.1) に従うと仮定し, 分散 η は研究単位間の異質性の度合いを決めるパラメータである. X_{ij} と D_{ij} の条件付 ($u_i, \mathbf{Z}_{1,ij}, \mathbf{Z}_{2,ij}$ を与えたとき) ハザード関数が次のような Joint frailty モデルに従うと仮定する.

$$\begin{cases} r_{ij}(t|u_i) = u_i r_0(t) \exp(\beta'_1 \mathbf{Z}_{1,ij}) \\ \lambda_{ij}(t|u_i) = u_i^\alpha \lambda_0(t) \exp(\beta'_2 \mathbf{Z}_{2,ij}). \end{cases}$$

ここで β は $\mathbf{Z}_{\cdot,ij}$ の効果を決める未知パラメータで, $r_0(\cdot)$ と $\lambda_0(\cdot)$ は未知の基準ハザード関数である. 他の研究単位と比較して高リスクな研究単位は $u_i > 1$, 低リスクな研究単位は $0 < u_i < 1$ で与えられ, パラメータ α は増悪率と死亡率の異質性の差を表す. 例えば, $\alpha = 0$ は増悪率に異質性があるが死亡率には異質性がないモデル, $\alpha = 1$ は増悪率と死亡率に同等の異質性があるモデルに対応する. Rondeau et al. (2015) のモデルにおいて重要な仮定は, ($u_i, \mathbf{Z}_{1,ij}, \mathbf{Z}_{2,ij}$) が与えられたもとで, X_{ij} と D_{ij} が条件付独立となることである.

Emura et al. (2017) は条件付独立の仮定を除いたコピュラのモデルである Joint frailty-copula モデルを提案した. C_θ をコピュラとしたとき, Joint frailty-copula モデルは次の式で与えられる;

$$(5.1) \quad \begin{cases} r_{ij}(t|u_i) = u_i r_0(t) \exp(\beta'_1 \mathbf{Z}_{1,ij}) \\ \lambda_{ij}(t|u_i) = u_i^\alpha \lambda_0(t) \exp(\beta'_2 \mathbf{Z}_{2,ij}) \\ \Pr(X_{ij} > x, D_{ij} > y|u_i) = C_\theta[S_{X_{ij}}(x|u_i), S_{D_{ij}}(y|u_i)]. \end{cases}$$

ここで θ は未知パラメータである. このモデルにおいて, X_{ij} と D_{ij} の生存関数は次のように書ける

$$\begin{cases} S_{X_{ij}}(x|u_i) = \exp\{-u_i R_0(x) \exp(\beta'_1 \mathbf{Z}_{1,ij})\} \\ S_{D_{ij}}(y|u_i) = \exp\{-u_i^\alpha \Lambda_0(y) \exp(\beta'_2 \mathbf{Z}_{2,ij})\}. \end{cases}$$

ここで $R_0(x) = \int_0^x r_0(t) dt$ と $\Lambda_0(y) = \int_0^y \lambda_0(t) dt$ は累積基準ハザード関数である.

コピュラ C_θ は X_{ij} と D_{ij} の相関構造を定める. 独立型 $C_\theta(v, w) = vw$ は Rondeau et al. (2015) のモデルに帰着する. クレイトン型ではケンドール相関は $\tau_\theta = \theta/(\theta + 2)$ であり, グンベル型ではケンドール相関は $\tau_\theta = \theta/(\theta + 1)$ である. Burzykowski et al. (2001) のコピュラモデルと同じ原理で, ケンドール相関は, 無増悪期間が全生存期間の代替エンドポイントになりえるかどうかの指標の 1 つである.

Joint frailty-copula モデル内の未知パラメータは観測されたデータ $\{(T_{ij}, T_{ij}^*, \delta_{ij}, \delta_{ij}^*, \mathbf{Z}_{1,ij}, \mathbf{Z}_{2,ij}), i = 1, 2, \dots, G, j = 1, 2, \dots, N_i\}$ を用いて最尤法で推定することが出来る. まず次のよう

な記号を導入する；

$$R_{ij}(t) = R_0(t) \exp(\beta'_1 \mathbf{Z}_{1,ij}), \quad r_{ij}(t) = dR_{ij}(t)/dt = r_0(t) \exp(\beta'_1 \mathbf{Z}_{1,ij}),$$

$$\Lambda_{ij}(t) = \Lambda_0(t) \exp(\beta'_2 \mathbf{Z}_{2,ij}), \quad \lambda_{ij}(t) = d\Lambda_{ij}(t)/dt = \lambda_0(t) \exp(\beta'_2 \mathbf{Z}_{2,ij}).$$

モデル(4.1)と(5.1)のもとで、対数尤度関数は次のように構成される；

$$\begin{aligned} \ell(\alpha, \eta, \theta, \beta_1, \beta_2, r_0, \lambda_0) &= \sum_{i=1}^G \left[\sum_{j=1}^{N_i} \{ \delta_{ij} \log r_{ij}(T_{ij}) + \delta_{ij}^* \log \lambda_{ij}(T_{ij}^*) \} \right. \\ &\quad \left. + \log \int_0^\infty \left\{ u^{m_i + \alpha m_i^*} \prod_{j=1}^{N_i} \psi_\theta [uR_{ij}(T_{ij}), u^\alpha \Lambda_{ij}(T_{ij}^*)]^{\delta_{ij}} \psi_\theta^* [uR_{ij}(T_{ij}), u^\alpha \Lambda_{ij}(T_{ij}^*)]^{\delta_{ij}^*} \right. \right. \\ &\quad \left. \left. \times \Theta_\theta [uR_{ij}(T_{ij}), u^\alpha \Lambda_{ij}(T_{ij}^*)]^{\delta_{ij} \delta_{ij}^*} D_\theta [uR_{ij}(T_{ij}), u^\alpha \Lambda_{ij}(T_{ij}^*)] \right\} f_\eta(u) du \right], \end{aligned}$$

ここで $m_i = \sum_{j=1}^{N_i} \delta_{ij}$ は i 番目の研究単位で増悪した患者数, $m_i^* = \sum_{j=1}^{N_i} \delta_{ij}^*$ は i 番目の研究単位で死亡した患者数であり, $D_\theta[s, t] = C_\theta[\exp(-s), \exp(-t)]$,

$$\psi_\theta[s, t] = \frac{D_\theta^{[1,0]}[s, t]}{D_\theta[s, t]}, \quad \psi_\theta^*[s, t] = \frac{D_\theta^{[0,1]}[s, t]}{D_\theta[s, t]}, \quad \Theta_\theta[s, t] = \frac{D_\theta^{[1,1]}[s, t] D_\theta[s, t]}{D_\theta^{[1,0]}[s, t] D_\theta^{[0,1]}[s, t]},$$

$D_\theta^{[1,0]}[s, t] = -\partial D_\theta[s, t]/\partial s$, $D_\theta^{[0,1]}[s, t] = -\partial D_\theta[s, t]/\partial t$, $D_\theta^{[1,1]} = \partial^2 D_\theta[s, t]/\partial s \partial t$ はコンピュータを微分して得られる. 対数尤度の導出は Emura et al. (2017, 2019b) を参照されたい.

例えばクレイトン型では, $D_\theta[s, t] = A_\theta(s, t)^{-1/\theta}$, $\psi_\theta[s, t] = \exp(\theta s)/A_\theta(s, t)$, $\psi_\theta^*[s, t] = \exp(\theta t)/A_\theta(s, t)$, $\Theta_\theta[s, t] = 1 + \theta$ とおき, ここで $A_\theta(s, t) = \exp(\theta s) + \exp(\theta t) - 1$ である. これらを用いて対数尤度を計算する.

基準ハザード関数 $r_0(\cdot)$ と $\lambda_0(\cdot)$ のモデル化はいくつか提案されている. Emura et al. (2017) はスプライン関数でモデル化し, 罰則付き最尤法で推定することを提案している. スプライン関数は単調増加(減少)型, 凸型, 凹型などのハザードを柔軟にモデル化できる (Shih and Emura, 2020). Peng et al. (2018) と Peng (2019) は完全にノンパラメトリックなモデルを提案している. Li et al. (2019) はパラメトリックなモデルを考え, パラメータに事前分布を仮定したベイズ法を用いている. 篠原 (2019) と Wu et al. (2020) はワイブル分布を用いている. Schneider et al. (2020) でもワイブル分布と区分的指数分布のモデルを考えているが, Rondeau et al. (2015) と同様に, 条件付独立性を仮定している. スプライン関数を用いる場合, 推定値の標準誤差(SE)と信頼区間(CI)は罰則付き対数尤度のヘッセ行列から計算でき (Emura et al., 2017, 2019b), R パッケージ *joint.Cox* (Emura, 2019) で実行可能である. ノンパラメトリック最尤法を用いる場合, ヘッセ行列を利用すると SE を過小評価する傾向があるため, 層化ブートストラップ法が推奨されている (Peng, 2019).

5.3 卵巣癌データへの適用例

Ganzfried et al. (2013) の卵巣癌データを用いて, 2つのエンドポイントを同時に解析する方法を説明する. Ganzfried et al. (2013) は IPD メタ分析を用いて, 遺伝子 *CXCL12* の発現量の全生存期間に与える影響が有意であることを示した. 本稿では, Joint frailty-copula モデルの適用例として, *CXCL12* 遺伝子が2つのエンドポイント(無増悪期間と全生存期間)に与える影

表 4. 卵巣癌データ (Ganzfried et al., 2013; Emura et al., 2017). 研究単位は GEO (Gene Expression Omnibus) アクセス番号をあらわす. イベント発生率 (%) は研究単位ごとに特定のイベント (増悪, 死亡, 打ち切り) の発生率を計算した.

研究単位	サンプル数	観測されたイベント数 (イベント発生率 %)		
		増悪 ($\delta_y = 1$)	死亡 ($\delta_y^* = 1$)	打ち切り ($\delta_y^* = 0$)
GSE17260	$N_1 = 110$	76 (69%)	46 (42%)	64 (58%)
GSE30161	$N_2 = 58$	48 (83%)	36 (62%)	22 (38%)
GSE9891	$N_3 = 278$	185 (67%)	113 (41%)	165 (59%)
TCGA	$N_4 = 557$	266 (48%)	290 (52%)	267 (48%)
合計	$\sum_{i=1}^4 N_i = 1003$	575 (57%)	485 (48%)	518 (52%)

表 5. 卵巣癌データにおける無増悪期間と全生存期間を同時解析した結果. *CXCL12* 遺伝子発現量の相対リスクを考えている. 相対リスクが 1 以上とは, *CXCL12* 値が高いと増悪 (死亡) リスクが上昇することを意味する.

コピュラ	クレイトン型	独立型
	推定値 (95% CI)	推定値 (95% CI)
増悪の相対リスク ; $\exp(\beta_1)$	1.22 (1.13-1.32)	1.24 (1.14-1.35)
死亡の相対リスク ; $\exp(\beta_2)$	1.18 (1.08-1.29)	1.17 (1.07-1.29)
異質性の分散 ; η	0.033 (0.006-0.187)	0.028 (0.004-0.180)
相関パラメータ ; θ	2.35 (1.90-2.90)	0.00 (固定)
ケンドール相関 ; $\tau = \theta / (\theta + 2)$	0.54 (0.49-0.59)	0.00 (固定)

響の解析事例を紹介する.

Ganzfried et al. (2013) のデータは多くの研究単位を含む IPD メタ分析用のデータであるが, 多くの研究単位では 1 つのエンドポイントしか測定されていない. そのため, Emura et al. (2017) で行われたように, 無増悪期間と全生存期間の両方が測定されている 4 つの研究単位から 1,003 人の卵巣癌患者を取り出した (表 4). この取り出したデータは R パッケージ *joint.Cox* に収められている. Ganzfried et al. (2013) に従い, *CXCL12* 遺伝子の発現量のみを共変量とした.

卵巣癌データを Joint frailty-copula のクレイトン型コピュラモデルに当てはめた結果を表 5 に示す. *CXCL12* 遺伝子が全生存期間に与える影響は有意 (P 値 < 0.05) であり, その相対リスク (Relative Risk) は $RR = 1.18$ (95% CI: 1.08-1.29) である. また, *CXCL12* 遺伝子が無増悪期間に与える影響も有意 (P 値 < 0.05) であり, その影響は全生存期間のそれよりも大きい ($RR = 1.22$, 95% CI: 1.13-1.32). このことは, *CXCL12* の発現量は卵巣癌患者の術後増悪を予測するバイオマーカーとなりうることを示している. これら相対リスクの値は, *CXCL12* の発現量が SD 1 つ分増加したときのリスク増加を表す. コピュラのパラメータ推定値 ($\hat{\theta} = 2.35$, 95% CI: 1.90-2.90) は, 増悪と死亡の間に中程度の相関があることを示す ($\tau = 0.54$, 95% CI: 0.38-0.70). このことは, 増悪で死亡をある程度予測できることを意味する.

独立型コピュラを適用した結果を表 5 に示す. この場合, 増悪と死亡間の相関を考慮できないため, このモデルの妥当性は弱い. しかし, 相対リスクの推定値は, クレイトン型のものと比べてほとんど変化がない. これは, 相対リスクの推定値はコピュラの誤特定に関してある程

度の頑健性を持つことを示している。このような頑健性は Emura et al. (2017) のシミュレーション解析でも報告されている。

6. 動的予測と個別医療

6.1 生存期間の予測

生存時間解析における重要なテーマの1つが、生存期間の予測法の開発である。過去の患者の生存期間データをもとにして、新たな患者の生存期間を予測するためのモデル構築に関する手法は数多く提案されており、これは統計学の手法に加えて機械学習の手法も多い。例えば、乳癌 (Gómez et al., 2016; Shukla et al., 2018)、卵巣癌 (Yoshihara et al., 2010; Emura et al., 2018)、膀胱癌 (Guinney et al., 2017)、その他の癌種 (Kim et al., 2018; Choi et al., 2018) について、患者の生存期間の予測式が提案されている。

正確な予測式を構築することは、患者本人やその主治医に対してより正確な人生設計や治療計画を考える礎となるため、個別医療 (personalized medicine) の中核となる。個別医療の定義によると、それは遺伝子 (分子) 情報を基に健康増進を図る医療である (Schleiden et al., 2013; Hayes et al., 2014)。これは遺伝子 (分子) 情報の活用が患者の予後予測に重要な役割を果たすためである。

Waldron et al. (2014) は卵巣癌患者の大規模な IPD メタ分析用データを用い、遺伝子を利用した 14 種類の予測法の予測精度を比較した。彼らは、12 種類の予測法が全生存期間の予測に有意 (P 値 < 0.05) に働くことを示したが、同時にその予測能力はそれほど高くない (C 統計量 $= 0.56 - 0.60$) ことも指摘した。彼らはまた、これら生存期間予測法が臨床現場で実際に有効利用されるためには、さらなる改良や検証が必要であることも指摘している。

生存期間予測法を改良するためのアイデアは数多くあるが、以下では動的予測法とコンピュータを IPD メタ分析用データに適用した予測手法を考える。

6.2 動的予測

動的予測 (dynamic prediction) では、通常の共変量 (年齢、腫瘍径など) に加えて、中間イベント (増悪など) の情報を予測因子に取り入れる (van Houwelingen and Putter, 2011)。治療開始時点 (例えば手術日や投与開始日) を起点とする通常の予測の場合、手術日に測定される年齢や腫瘍径を共変量として扱い、手術日から時刻 w 後 (例えば $w = 5$ 年) に死亡する確率を予測する。動的予測では、手術後に経過観察した後にしか得られない増悪イベントの情報などを取り入れるため、起点を術後 (例えば術後 t 日目) に設定し、時刻 t まで生存した患者が、時刻 $t + w$ までに死亡する (または生存している) 確率を予測する (Thompson et al., 2007)。

いま D を全生存期間、 X を無増悪期間、 Z を遺伝子情報を含む共変量とする。共変量の値 Z は起点 $t = 0$ (例えば手術日) に測定され、その後変化しないと仮定する (時間依存性共変量は考えない)。

図 5 はで 4 人の患者 (患者 1 ~ 患者 4) に対して動的予測を考えている。患者 1, 2 は時刻 t 以前に死亡しているので、予測の対象とはならない。患者 3, 4 は時刻 t において生存しているので、予測を行うことが出来る。患者 3 は時刻 t 以前に増悪しているので、増悪時刻 $X = x$ で条件付けを行った死亡率を予測する。患者 4 は、時刻 t 以前には増悪していないので $X > t$ で条件付けを行った死亡率を予測する。

動的予測を実践するにあたり、予測時刻 t は前もって決めておく必要がある。例えば、時刻 t は術後 3 ヶ月ごとにとるなどと決めておく。この予測スケジュールは患者の状態によって変更してはならず、また患者の死亡や増悪に関連のある事象に依存してはならない。例えば、患

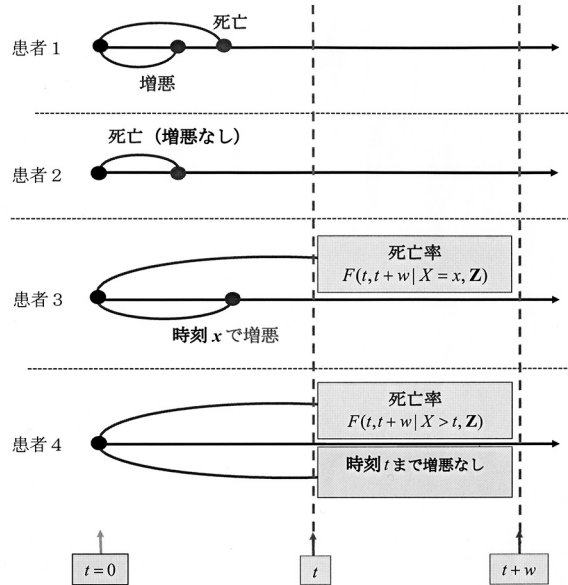


図 5. 増悪情報に基づく死亡率の動的予測. 患者 1, 2 は時刻 t 以前に死亡しているため, 動的予測の対象とならない. 患者 3, 4 に対する動的予測は, 時刻 t 以前に観測されている増悪情報に基づいて行われる.

者が何らかの不調を訴えて来院した時刻や, 増悪した時刻を t に設定すると, 妥当な予測は行えなくなる.

動的予測を定式化するために次のような 2 つの関数を利用する.

(i) 条件付死亡率関数 (Conditional Failure Function)

動的予測の基礎となるのが, 次の 2 つの条件付死亡率関数

$$F(t, t+w | X > t, \mathbf{Z}) = \Pr(D \leq t+w | D > t, X > t, \mathbf{Z}),$$

$$F(t, t+w | X = x, \mathbf{Z}) = \Pr(D \leq t+w | D > t, X = x, \mathbf{Z}), \quad x \leq t$$

である. 2 つの式にある条件 $\{D \geq t\}$ は, 予測の対象である患者が時刻 t において生存していることを前提としている. 1 つ目の式にある条件 $\{X > t\}$ は, 時刻 t においてまだ増悪していないことを意味する. 2 つ目の式にある条件 $\{X = x\}$ は, 時刻 t 以前に増悪したことを意味する. この場合, 増悪した時刻 $X = x$ は既知なので予測式に組み入れることができる. 条件付死亡率関数は, これら条件下で, 患者が時刻 t から $t+w$ の間に死亡する確率である. 通常の (動的でない) 予測では予測時刻を $t=0$ とし, $w=5$ 年などに設定して予測を行う. 予測時刻を $t=0$ と設定すると, 増悪の情報は不明なので, 予測式は $F(t, t+w | \mathbf{Z}) = \Pr(D \leq t+w | D > t, \mathbf{Z})$ のみである.

増悪した患者は, そうでない患者に比べて, その死亡率が上昇する. したがって, 直感的に不等式 $F(t, t+w | X = x, \mathbf{Z}) > F(t, t+w | X > t, \mathbf{Z})$ が全ての $w > 0$ に対して成立すると考えられる (図 1). X と D が正相関を持つ場合, この不等式は通常成立する (数学的に成立するかどうかはモデルに依存). X と D が (\mathbf{Z} を与えたときの) 条件付独立であれば, 等式 $F(t, t+w | X = x, \mathbf{Z}) = F(t, t+w | X > t, \mathbf{Z})$ が成立することは明らかである.

(ii) 条件付ハザード関数 (Conditional Hazard Function)

条件付ハザード関数は、条件付死亡率関数を用いて次のように表現できる。

$$\lambda(t|X = x, \mathbf{Z}) = F(t, t + dw|X = x, \mathbf{Z})/dw,$$

$$\lambda(t|X > x, \mathbf{Z}) = F(t, t + dw|X > x, \mathbf{Z})/dw.$$

これらの式により、条件付ハザード関数は、上記の条件付死亡率関数において w が微小な値 dw をとる場合と解釈できる。2つの条件付ハザード関数の比をとると、 X と D の関連を表す指標である次のようなオッズ比関数(3.3節)となる。

$$\frac{\lambda(t|X = x, \mathbf{Z})}{\lambda(t|X > x, \mathbf{Z})} = \frac{\Pr(X = x, D = t|\mathbf{Z})\Pr(X > x, D > t|\mathbf{Z})}{\Pr(X = x, D > t|\mathbf{Z})\Pr(X > x, D = t|\mathbf{Z})}, \quad t > 0, \quad x > 0.$$

オッズ比が1より大きいときは正相関を表し、オッズ比が1より小さいときは負相関を表す。オッズ比が1のときは X と D は (\mathbf{Z} を与えたときの)条件付独立である。

Clayton (1978)が提案した次のセミパラメトリックモデル

$$\Pr(X > x, D > t|\mathbf{Z}) = [S_X(x|\mathbf{Z})^{-\theta} + S_D(t|\mathbf{Z})^{-\theta} - 1]^{-\frac{1}{\theta}}, \quad t > 0, \quad x > 0, \quad \theta > 0$$

を考えると、このモデルの下では、2つの条件付ハザード関数の関係をオッズ比定数 ($\theta + 1$) を用いて

$$\lambda(t|X = x, \mathbf{Z}) = (\theta + 1)\lambda(t|X > x, \mathbf{Z}), \quad t > 0, \quad x > 0.$$

と表すことができる。これは、患者の増悪イベント $\{X = x\}$ があつたときと、そうでないとき $\{X > x\}$ の相対リスクが $(\theta + 1)$ となるクレイトン型 (Clayton, 1978)コピュラモデルである(3.3節を参照)。Day et al. (1997)はクレイトン型モデルにランドマーク解析を使った動的予測を提案している。

動的予測において、条件付死亡率関数、条件付ハザード関数のどちらも有用な予測指標である。時刻 t を予測時刻と呼び、これは臨床医が患者に対して予測を始める時刻である。これら予測指標をデータから推定する場合、通常何らかのモデルを利用する。van Houwelingen and Putter (2011)はランドマーク解析法を採用している。これは予測時刻 t 以前に収集されたデータを基に条件付 Cox 回帰法を当てはめて予測モデルを構築するものである。ランドマーク解析法では、 t を複数選び、複数の条件付 Cox モデルを構築したときにモデル間の一致性が成立しないなどのいくつかの諸問題がある。詳細については van Houwelingen and Putter (2011)と Suresh (2018)に詳しく議論されている。

ランドマーク解析法の諸問題の解決策として、結合モデル (Joint model) を用いた動的予測式が提案されている (Proust-Lima et al., 2014; Rondeau et al., 2017; Sène et al., 2016 など)。これらのモデルにおいては、Frailty 項を用いて応答変数間の相関を考慮している。Frailty 項の替わりに、コピュラで相関を考慮するモデルは Emura et al. (2018), Suresh et al. (2019)などを参照されたい。

6.3 Joint frailty-copula モデルを利用した動的予測

以下では、Frailty 項とコピュラで構成した Joint frailty-copula モデルによる動的予測法を紹介する (Emura et al., 2018)。この手法では、全生存期間と無増悪期間の相関を解析する際に用いる IPD メタ分析データを利用する。このため、Frailty 項は相関を考慮するというよりも、研究単位間の異質性をモデル化するために導入されている。

無増悪期間の生存関数を $S_X(x|u) = \Pr(X > x|u, \mathbf{Z})$ 、全生存期間の生存関数を $S_D(y|u) =$

$\Pr(D > y|u, \mathbf{Z})$ で定義し, ここで u は共変量 \mathbf{Z} で説明しきれない Frailty 項である. u の分布は, 式(4.1)で与えられるように, 平均 1, 分散 η のガンマ分布であるとする. Joint frailty-copula モデルの仮定をおくと,

$$\Pr(X > x, D > y|u, \mathbf{Z}) = C_\theta[S_X(x|u), S_D(y|u)],$$

となり, $C_\theta(v, w)$ はコピュラ, θ は X と D の相関を表すパラメータである. このモデルの下で, 条件付死亡率関数は次のように表せる.

(I) 時刻 t 以前に患者が増悪していない場合

$X > t$ で条件付けを行い, 条件付死亡率関数は

$$\begin{aligned} F(t, t+w|X > t, \mathbf{Z}) &= \Pr(D \leq t+w|D > t, X > t, \mathbf{Z}) \\ &= \frac{\int_0^\infty (C_\theta[S_X(t|u), S_D(t|u)] - C_\theta[S_X(t|u), S_D(t+w|u)])f_\eta(u)du}{\int_0^\infty C_\theta[S_X(t|u), S_D(t|u)]f_\eta(u)du} \end{aligned}$$

となる. ただし $f_\eta(u)$ は式(4.1)で与えられるガンマ分布の確率密度関数である. また, 条件付ハザード関数は

$$\lambda(t|X > t, \mathbf{Z}, u) = \lambda_D(t|u) \frac{S_D(t|u)C_\theta^{[0,1]}[S_X(t|u), S_D(t|u)]}{C_\theta[S_X(t|u), S_D(t|u)]}$$

となる. ここで $\lambda_D(t|u) = -\partial \log S_D(t|u)/\partial t$, $C_\theta^{[0,1]}(v, w) = \partial C_\theta(v, w)/\partial w$ である.

(II) 時刻 t 以前に患者が増悪している場合

増悪時刻 $X = x$ で条件付けを行い, 条件付死亡率関数は

$$\begin{aligned} F(t, t+w|X = x, \mathbf{Z}) &= \Pr(D \leq t+w|D > t, X = x, \mathbf{Z}) \\ &= \frac{\int_0^\infty (C_\theta^{[1,0]}[S_X(x|u), S_D(t|u)] - C_\theta^{[1,0]}[S_X(x|u), S_D(t+w|u)])\lambda_X(x|u)S_X(x|u)f_\eta(u)du}{\int_0^\infty C_\theta^{[1,0]}[S_X(x|u), S_D(t|u)]\lambda_X(x|u)S_X(x|u)f_\eta(u)du} \end{aligned}$$

となる. ここで $\lambda_X(x|u) = -\partial \log S_X(x|u)/\partial x$ である. また, 条件付ハザード関数は

$$\lambda(t|X = x, \mathbf{Z}, u) = \lambda_D(t|u) \frac{S_D(t|u)C_\theta^{[1,1]}[S_X(x|u), S_D(t|u)]}{C_\theta^{[1,0]}[S_X(x|u), S_D(t|u)]}$$

となる. ここで $C_\theta^{[1,0]}(v, w) = \partial C_\theta(v, w)/\partial v$, $C_\theta^{[1,1]}(v, w) = \partial^2 C_\theta(v, w)/\partial v \partial w$ である.

上式の導出は Emura et al. (2018, 2019b)を参照されたい. IPD メタ分析用データなどを用いて条件付死亡率関数や条件付ハザード関数内のパラメータは推定することができる(詳しくは Emura et al., 2018, 2019b を参照).

条件付死亡率関数は Frailty 項 u で積分している. これは予測すべき患者の Frailty 項 u が観測不能なため, u の分布で平均化していることを意味する. 積分は数値的に計算する. 簡便な方法であるが, 積分する代わりに平均値 $u = 1$ で置き換える方法もある.

条件付ハザード関数は Frailty 項 u で条件付けを行っている. u は未知なので積分する方が望ましいが, 条件付けを行うことにより, 次のような数学的性質が利用できる.

$$\frac{\lambda(t|X = t, \mathbf{Z}, u)}{\lambda(t|X > t, \mathbf{Z}, u)} = R_\theta[S_X(t|u), S_D(t|u)].$$

ここで R_θ は 3.3 節で既出のオッズ比関数である. クレイトン型コピュラでは $R_\theta(v, w) = 1 + \theta$ が $\{X = t\}$ の $\{X > t\}$ に対する相対リスクとなる. 先に $\lambda(t|X > t, \mathbf{Z}, u)$ を計算すれば, それに $1 + \theta$ をかけることにより $\lambda(t|X = t, \mathbf{Z}, u)$ が得られる. 予測に使用する Frailty 項は $u = 1$

とする。

動的予測を臨床現場で実践する場合、まだ死亡していない新たな患者を対象にする。ここで「新たな」とは、すでに死亡・生存の状態が確定しているデータ内の患者と区別するための語である。現実にもそのような患者を有していなくても、仮想的な患者を想定し、どのように予測のプロセスが実行されるのかを考察することは重要である。

まず患者が有している共変量情報を $\mathbf{Z} = (\mathbf{Z}_1, \mathbf{Z}_2, \text{CC}_1, \text{CC}_2)$ とし、これらは：

- \mathbf{Z}_1 ; 増悪に関連する p_1 次元の臨床的共変量,
- \mathbf{Z}_2 ; 全生存に関連する p_2 次元の臨床的共変量,
- $\text{CC}_1 = w_1 V_1 + \dots + w_{q_1} V_{q_1}$; 増悪に関連する遺伝子発現量の複合共変量,
- $\text{CC}_2 = \varpi_1 W_1 + \dots + \varpi_{q_2} W_{q_2}$; 全生存に関連する遺伝子発現量の複合共変量,

と定義する。ここで (V_1, \dots, V_{q_1}) は q_1 個の増悪に関連する遺伝子発現量, (W_1, \dots, W_{q_2}) は q_2 個の全生存に関連する遺伝子発現量である。遺伝子の個数 q_1 (または q_2) と重み w_k (または ϖ_k) はデータから計算される (Emura et al., 2018, 2019b)。全ての遺伝子発現量は、遺伝子ごとに平均 0, 標準偏差 1 で標準化されていると仮定する。複合共変量 (compound covariate) の基本的な考え方については Matsui (2006), Emura et al. (2012, 2019a, 2019b) を参照されたい。

データを Joint frailty-copula モデルに当てはめることにより、無増悪期間と全生存期間の生存関数が次のように推定される。

$$\begin{aligned} \hat{S}_X(t|u) &= \exp \left\{ -u \hat{R}_0(t) \exp \left(\hat{\beta}'_1 \mathbf{Z}_1 + \hat{\gamma}_1 \frac{\text{CC}_1 - \hat{\mu}_1}{\hat{\sigma}_1} \right) \right\}, \\ \hat{S}_D(t|u) &= \exp \left\{ -u^\alpha \hat{\Lambda}_0(t) \exp \left(\hat{\beta}'_2 \mathbf{Z}_2 + \hat{\gamma}_2 \frac{\text{CC}_2 - \hat{\mu}_2}{\hat{\sigma}_2} \right) \right\}. \end{aligned}$$

ここで $\hat{R}_0(t) = \int_0^t \hat{r}_0(x) dx$ と $\hat{\Lambda}_0(t) = \int_0^t \hat{\lambda}_0(x) dx$ は推定された累積基準ハザード関数, $(\hat{\theta}, \hat{\eta}, \hat{\beta}_1, \hat{\beta}_2, \hat{\gamma}_1, \hat{\gamma}_2, \hat{r}_0, \hat{\lambda}_0)$ は推定値されたパラメータの値である。 $\hat{\mu}$ は CC の平均, $\hat{\sigma}$ は CC の標準偏差である。

これら推定値を条件付死亡率関数や条件付ハザード関数の未知パラメータに代入することにより、新たな患者への動的予測が可能となる。例えば、条件付死亡率関数は次のように計算される。

$$\hat{F}(t, t+w | X > t, \mathbf{Z}) = \frac{\int_0^\infty (C_\theta^\alpha[\hat{S}_X(t|u), \hat{S}_D(t|u)] - C_\theta[\hat{S}_X(t|u), \hat{S}_D(t+w|u)]) f_{\hat{\eta}}(u) du}{\int_0^\infty C_\theta[\hat{S}_X(t|u), \hat{S}_D(t|u)] f_{\hat{\eta}}(u) du},$$

$$\begin{aligned} \hat{F}(t, t+w | X = x, \mathbf{Z}) &= \frac{\int_0^\infty (C_\theta^{[1,0]}[\hat{S}_X(x|u), \hat{S}_D(t|u)] - C_\theta^{[1,0]}[\hat{S}_X(x|u), \hat{S}_D(t+w|u)]) \hat{\lambda}_X(x|u) \hat{S}_X(x|u) f_{\hat{\eta}}(u) du}{\int_0^\infty C_\theta^{[1,0]}[\hat{S}_X(x|u), \hat{S}_D(t|u)] \hat{\lambda}_X(x|u) \hat{S}_X(x|u) f_{\hat{\eta}}(u) du}. \end{aligned}$$

6.4 卵巣癌患者への適用例

Emura et al. (2018) は Ganzfried et al. (2013) の 912 人の卵巣癌患者をもとに、次のような複合共変量を得た。

- 増悪に関連する 158 個の遺伝子 (P 値 < 0.001) ;

$$\text{CC}_1 = (0.249 \times \text{CXCL12}) + (0.235 \times \text{TIMP2}) + (0.222 \times \text{PDPN}) + \dots + (-0.152 \times \text{MMP12}).$$

- 死亡に関連する 128 個の遺伝子 (P 値 < 0.001) ;

$$CC_2 = (0.237 \times NCOA3) + (0.223 \times TEAD1) + (0.263 \times YWHAB) + \dots + (-0.157 \times KCNH4).$$

ここで遺伝子発現量は遺伝子名 (*CXCL12* など) で表し, 全患者で平均 0, 標準偏差 1 となるよう標準化してある. CC_1 の平均は $\hat{\mu}_1 = 0.338$, 標準偏差は $\hat{\sigma}_1 = 10.468$ であり, CC_2 の平均は $\hat{\mu}_2 = 0.222$, 標準偏差は $\hat{\sigma}_2 = 7.894$ であった. P 値が 0.001 より小さい遺伝子を選択する基準はマイクロアレイ解析でしばしば利用される (Matsui et al., 2012).

データを Joint frailty-copula モデルに当てはめた結果,

$$\Pr(X > x, D > y | u) = C_{\hat{\theta}}[\hat{S}_X(x|u), \hat{S}_D(y|u)] = [\hat{S}_X(x|u)^{-\hat{\theta}} + \hat{S}_D(y|u)^{-\hat{\theta}} - 1]^{-1/\hat{\theta}},$$

となり, ここで $\hat{\theta} = 1.9$ (ケンドール相関 $\hat{\tau} = 0.49$),

$$\hat{S}_X(x|u) = \exp \left\{ -u \hat{R}_0(t) \exp \left(\hat{\gamma}_1 \frac{CC_1 - \hat{\mu}_1}{\hat{\sigma}_1} \right) \right\},$$

$$\hat{S}_D(y|u) = \exp \left\{ -\hat{\Lambda}_0(t) \exp \left(\hat{\beta}_2 Z_2 + \hat{\gamma}_2 \frac{CC_2 - \hat{\mu}_2}{\hat{\sigma}_2} \right) \right\},$$

となった. ここで Z_2 は (0 または 1) をとる術後腫瘍径 (≤ 1 cm or > 1 cm) である. 腫瘍径が無増悪期間に与える影響は有意でなかった (P 値 > 0.05) ので $\hat{S}_X(x|u)$ のモデルから除いてある. 回帰推定値は $\hat{\gamma}_1 = 0.39$, $\hat{\beta}_2 = 0.16$, $\hat{\gamma}_2 = 0.44$ であった. 基準ハザード関数の推定値は, $M \cdot(t)$ を M スプライン基底関数 (Emura et al., 2017, 2019b) とすると,

$$\hat{r}_0(t) = d\hat{R}_0(t)/dt = 0.85 \times M_1(t) + 2.14 \times M_2(t) + 0 \times M_3(t) + 0.07 \times M_4(t) + 0 \times M_5(t),$$

$$\hat{\lambda}_0(t) = d\hat{\Lambda}_0(t)/dt = 0.17 \times M_1(t) + 1.05 \times M_2(t) + 1.24 \times M_3(t) + 0.27 \times M_4(t) + 0 \times M_5(t),$$

$t \in [0, 6420]$ であり, 6420 (日) は全患者の最長追跡日数である. 異質性分散の推定値は $\text{Var}(u) = \hat{\eta} = 0.04$ であった. ステージなどの他の共変量もいくつか測定されているが, これらを追加してもモデルの当てはまりは改善しなかった.

次に, 以下のような仮想的な二人の卵巣癌患者を考え, どのように予測のプロセスが実行されるのかを説明する.

- 患者 1; 高リスク遺伝子 ($CC_1 = 10$, $CC_2 = 10$), 腫瘍径 > 1 cm ($Z_2 = 1$).
- 患者 2; 低リスク遺伝子 ($CC_1 = -10$, $CC_2 = -10$), 腫瘍径 ≤ 1 cm ($Z_2 = 0$).

例えば患者 2 に対して条件付死亡率関数 $\hat{F}(t, t+w | X = x, \mathbf{Z})$ を $t = 1000$, $t+w < 6420$, $x = 600$ の設定で計算するときは, R パッケージ *joint.Cox* を利用し, 「付録」にあるコードを実行すれば良い. CC を計算するには, 患者の 158 個の遺伝子を調べ, 重み付け和を計算する必要がある. ここではそのプロセスを省略した.

図 6 は二人の患者の条件付死亡率の推移を表す. $t = 500$ 日目を起点に予測を始めると, 患者 1 は患者 2 よりも高い死亡率を示す (図 6 の左パネル). これは患者 1 が患者 2 より高リスク遺伝子を持ち, かつ腫瘍径が大きいことに由来する. 図 6 の左パネルにおいては予測時刻 $t = 500$ 日目において両患者ともに増悪がないと仮定した. 時刻を進め, $t = 1000$ 日目を起点とし, 患者 2 だけに増悪があった ($x = 600$ 日目) と仮定すると, 患者 2 の死亡率が急激に上昇し, 患者 1 の死亡率を上回る (図 6 の右パネル). このことは, 増悪の死亡に与える影響が, 遺伝子と腫瘍径の死亡に与える影響よりも強いことを示している.

図 7 は二人の卵巣癌患者の条件付ハザード関数 $\hat{\lambda}(t | X = t, \mathbf{Z}, u)$ と $\hat{\lambda}(t | X > t, \mathbf{Z}, u)$ を $u = 1$ で計算したものを示している. 患者 1 は危険因子 (高リスク遺伝子と腫瘍径 > 1 cm) を持つため, 患者 2 よりも高いハザード値を示していることがわかる. 他方, 患者 2 は時間とともに自然に減少するハザード値を示している. 同一の患者で, 増悪した場合としなかった場合を比較

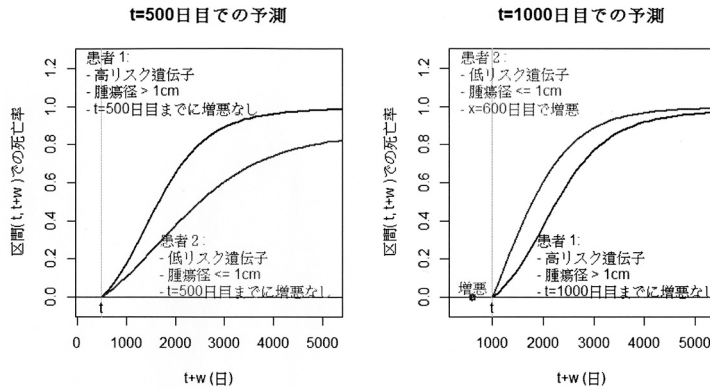


図 6. 二人の卵巣癌患者の条件付死亡率の推移.

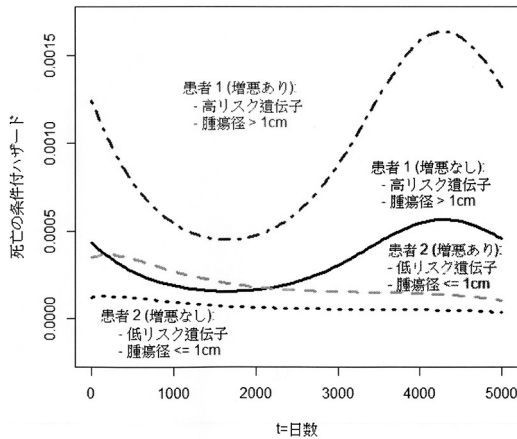


図 7. 二人の卵巣癌患者の条件付ハザードの推移.

するときに、これら 2 つのハザード間に比例関係が存在し、

$$\hat{\lambda}(t|X = t, \mathbf{Z}, u) = (1 + \hat{\theta})\hat{\lambda}(t|X > t, \mathbf{Z}, u) = 2.9\hat{\lambda}(t|X > t, \mathbf{Z}, u)$$

と表せることに注意する。これはクレイトン型コピュラの仮定から導かれるものである。したがって、増悪した場合は、増悪しなかった場合と比較してハザードが約 3 倍になることがわかる。逆に増悪しなければ、患者 2 の死亡リスクはかなり低下することがわかる。また、増悪しなかった場合でも、患者 1 のように危険因子(高リスク遺伝子と腫瘍径 > 1 cm)を持つ患者は、将来的に死亡リスクが上昇する可能性があることがわかる。

7. 今後の課題

コピュラの型は無数にあるため、研究者はその目的に応じて最も適したコピュラを選択することができる。しかし実際には、生存期間において変数間の相関をモデリングするためにはクレイトン型のコピュラが最も広く利用されてきた。本稿で示したようにその数学的に簡潔な関数形、ケンドール相関やオッズ比の特徴からその利便性・有用性は理解されるであろう。また

相関パラメータが相対リスクとして解釈できる医学上の重要性も強調した。このような理由から、本稿でも主にクレイトン型コピュラを用いて解説してきたが、これは母集団がクレイトン型コピュラに従うことを必ずしも意味していない。母集団の相関構造をよりよく近似するコピュラを用いたい場合は、クレイトン型、ゲンベル型、フランク型など特徴が異なるいくつかのコピュラ間でモデル選択を行うことが望ましい。しかしながら、尤度や予測モデルの計算が複雑になることから、現時点では5節、6節の手法はRパッケージ *joint.Cox* においてクレイトン型のみ利用可能となっている。

6節においては、患者の生存期間の予測にコピュラをどのように利用するかを、動的予測法の枠組みで解説した。患者の生存期間を正確に予測することは、個別医療の観点からも重要な課題であるため、現場の臨床医や医師が利用できる状態にすることが重要である。1つの取り組みとして、臨床医向けに動的予測の利用法ウェブ・アプリケーションを作成する手法を解説した論文を執筆している (Emura et al., 2020)。

国内における上記動的予測の手法は、間質性肺炎合併肺癌患者へ応用された例がある (高見澤, 2018)。またこの動的予測の手法をワイブル分布で行うことにより、長期予測法を実現する方法も提案されている (篠原, 2019)。

謝 辞

本特集号の企画者である吉羽要直先生と加藤昇吾先生には、執筆の機会を頂いたことに心より感謝申し上げます。また改稿に当たり、有益なコメントをいただいた査読者と編集委員の方々に感謝申し上げます。最後に、図1をご提供頂いた篠原紗也華さんに感謝申し上げます。

参 考 文 献

- Alonso, A. and Molenberghs, G. (2008). Evaluating time to cancer recurrence as a surrogate marker for survival from an information theory perspective, *Statistical Methods in Medical Research*, **17**(5), 497-504.
- Barthel, N., Geerdens, C., Killiches, M., Janssen, P. and Czado, C. (2018). Vine copula based likelihood estimation of dependence patterns in multivariate event time data, *Computational Statistics & Data Analysis*, **117**, 109-127.
- Burzykowski, T., Molenberghs, G., Buyse, M., Geys, H. and Renard, D. (2001). Validation of surrogate end points in multiple randomized clinical trials with failure time end points, *Applied Statistics*, **50**(4), 405-422.
- Burzykowski, T., Molenberghs, G. and Buyse, M. (eds.) (2005). *The Evaluation of Surrogate End-points*, Springer, New York.
- Buyse, M., Burzykowski, T., Michiels, S. and Carroll, K. (2008). Individual-and trial-level surrogacy in colorectal cancer, *Statistical Methods in Medical Research*, **17**, 467-475.
- Chaieb, L., Rivest, L. P. and Abdous, B. (2006). Estimating survival under a dependent truncation, *Biometrika*, **93**(3), 655-669.
- Chen, Y. H. (2012). Maximum likelihood analysis of semicompeting risks data with semiparametric regression models, *Lifetime Data Analysis*, **18**, 36-57.
- Choi, J., Oh, I., Seo, S. and Ahn, J. (2018). G2Vec: Distributed gene representations for identification of cancer prognostic genes, *Scientific Reports*, **8**(1), 13729.
- Clayton, D. G. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence, *Biometrika*, **65**(1), 141-151.

- Cox, D. R. (1972). Regression models and life-tables (with discussion), *Journal of the Royal Statistical Society, Series B*, **34**, 187-220.
- Day, R., Bryant, J. and Lefkopoulou, M. (1997). Adaptation of bivariate frailty models for prediction, with application to biological markers as prognostic indicators, *Biometrika*, **84**(1), 45-56.
- de Uña-Álvarez, J. and Veraverbeke, N. (2013). Generalized copula-graphic estimator, *TEST*, **22**(2), 343-360.
- de Uña-Álvarez, J. and Veraverbeke, N. (2017). Copula-graphic estimation with left-truncated and right-censored data, *Statistics*, **51**(2), 387-403.
- Duchateau, L., Janssen, P., Lindsey, P., Legrand, C., Nguti, R., et al. (2002). The shared frailty model and the power for heterogeneity tests in multicenter trials, *Computational Statistics & Data Analysis*, **40**(3), 603-620.
- Eisenhauer, E. A., Therasse, P., Bogaerts, J., Schwartz, L. H., Sargent, D., et al. (2009). New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1), *European Journal of Cancer*, **45**(2), 228-247.
- Emura, T. (2019). joint.Cox: Joint frailty-copula models for tumour progression and death in meta-analysis, *CRAN*.
- Emura, T. and Chen, Y. H. (2016). Gene selection for survival data under dependent censoring, a copula-based approach, *Statistical Methods in Medical Research*, **25**(6), 2840-2857.
- Emura, T. and Chen, Y. H. (2018). *Analysis of Survival Data with Dependent Censoring, Copula-based Approaches*, JSS Research Series in Statistics, Springer, Singapore.
- Emura, T. and Michimae, H. (2017). A copula-based inference to piecewise exponential models under dependent censoring, with application to time to metamorphosis of salamander larvae, *Environmental and Ecological Statistics*, **24**(1), 151-173.
- Emura, T. and Murotani, K. (2015). An algorithm for estimating survival under a copula-based dependent truncation model, *TEST*, **24**(4), 734-751.
- Emura, T. and Pan, C. H. (2020). Parametric likelihood inference and goodness-of-fit for dependently left-truncated data, a copula-based approach, *Statistical Papers*, **61**(1), 479-501.
- Emura, T. and Wang, W. (2010). Testing quasi-independence for truncation data, *Journal of Multivariate Analysis*, **101**, 223-239.
- Emura, T. and Wang, W. (2012). Nonparametric maximum likelihood estimation for dependent truncation data based on copulas, *Journal of Multivariate Analysis*, **110**, 171-188.
- Emura, T., Lin, C. W. and Wang, W. (2010). A goodness-of-fit test for Archimedean copula models in the presence of right censoring, *Computational Statistics & Data Analysis*, **54**, 3033-3043.
- Emura, T., Wang, W. and Hung, H. N. (2011). Semi-parametric inference for copula models for dependently truncated data, *Statistica Sinica*, **21**, 349-367.
- Emura T., Chen, Y. H. and Chen, H. Y. (2012). Survival prediction based on compound covariate under Cox proportional hazard models, *PLoS One*, **7**(10), e47627, DOI:10.1371/journal.pone.0047627.
- Emura, T., Nakatochi, M., Murotani, K. and Rondeau, V. (2017). A joint frailty-copula model between tumour progression and death for meta-analysis, *Statistical Methods in Medical Research*, **26**(6), 2649-2666.
- Emura, T., Nakatochi, M., Matsui, S. and Michimae, H. and Rondeau, V. (2018). Personalized dynamic prediction of death according to tumour progression and high-dimensional genetic factors: Meta-analysis with a joint model, *Statistical Methods in Medical Research*, **27**(9), 2842-2858.
- Emura, T., Matsui, S. and Chen, H. Y. (2019a). compound.Cox: Univariate feature selection and compound covariate for predicting survival, *Computer Methods and Programs in Biomedicine*, **168**, 21-37.
- Emura, T., Matsui, S. and Rondeau, V. (2019b). *Survival Analysis with Correlated Endpoints, Joint*

- Frailty-Copula Models*, JSS Research Series in Statistics, Springer, Singapore.
- Emura, T., Shih, J. H., Ha, I. D. and Wilke, R. A. (2019c). Comparison of the marginal hazard model and the sub-distribution hazard model for competing risks under an assumed copula, *Statistical Methods in Medical Research*, DOI: 10.1177/0962280219892295.
- Emura, T., Michimae, H. and Matsui, S. (2020). Building web applications for personalized risk prediction using R packages joint.Cox and Shiny (submitted).
- Fine, J. P., Jiang, H. and Chappell, R. (2001). On semi-competing risks data, *Biometrika*, **88**, 907-920.
- Ganzfried, B. F., Riester, M., Haibe-Kains, B., Risch, T., Tyekucheva, S., et al. (2013). Curated ovarian data: Clinically annotated data for the ovarian cancer transcriptome, *Database*, Article ID bat013, DOI:10.1093/database/bat013.
- Gómez, I., Ribelles, N., Franco, L., Alba, E. and Jerez, J. M. (2016). Supervised discretization can discover risk groups in cancer survival analysis, *Computer Methods and Programs in Biomedicine*, **136**, 11-19.
- Guinney, J., Wang, T., Laajala, T. D., Winner, K. K., Bare, J., et al. (2017). Prediction of overall survival for patients with metastatic castration-resistant prostate cancer: Development of a prognostic model through a crowdsourced challenge with open clinical trial data, *The Lancet Oncology*, **18**(1), 132-142.
- Gumbel, E. J. (1960). Distributions de valeurs extremes en plusieurs dimensions, *Publications de l'Institut de statistique de l'Université de Paris*, **9**, 171-173.
- Hayes, D. F., Markus, H. S., Leslie, R. D. and Topol, E. J. (2014). Personalized medicine: Risk prediction, targeted therapies and mobile health technology, *BMC Medicine*, **12**(1), 37.
- Hirsch, K. and Wienke, A. (2012). Software for semiparametric shared gamma and log-normal frailty models: An overview, *Computer Methods and Programs in Biomedicine*, **107**(3), 582-597.
- Kim, M., Oh, I. and Ahn, J. (2018). An improved method for prediction of cancer prognosis by network learning, *Genes*, **9**, 478.
- Klein, J. P., Van Houwelingen, H. C., Ibrahim, J. G. and Scheike, T. H. (eds.) (2014). *Handbook of Survival Analysis*, CRC Press, New York.
- Le Tourneau, C., Michiels, S., Gan, H. K. and Siu, L. L. (2009). Reporting of time-to-event end points and tracking of failures in randomized trials of radiotherapy with or without any concomitant anticancer agent for locally advanced head and neck cancer, *Journal of Clinical Oncology*, **27**(35), 5965-5971.
- Li, Z., Chinchilli, V. M. and Wang, M. (2019). A Bayesian joint model of recurrent events and a terminal event, *Biometrical Journal*, **61**(1), 187-202.
- Li, Z., Chinchilli, V. M. and Wang, M. (2020). A time-varying Bayesian joint hierarchical copula model for analysing recurrent events and a terminal event: An application to the Cardiovascular Health Study, *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, **69**(1), 151-166.
- Ling, M., Hu, T. and Sun, J. (2016). Cox regression analysis of dependent interval-censored failure time data, *Computational Statistics & Data Analysis*, **103**, 79-90.
- Lo, S. M. and Wilke, R. A. (2010). A copula model for dependent competing risks, *Journal of the Royal Statistical Society, Series C*, **59**(2), 359-376.
- Matsui, S. (2006). Predicting survival outcomes using subsets of significant genes in prognostic marker studies with microarrays, *BMC Bioinformatics*, **7**, 156.
- Matsui, S., Simon, R., Qu, P., Shaughnessy, J. D., Barlogie, B. and Crowley, J. (2012). Developing and validating continuous genomic signatures in randomized clinical trials for predictive medicine, *Clinical Cancer Research*, **18**(21), 6065-6073.
- Michiels, S., Le Maître, A., Buyse, M., Burzykowski, T., Maillard, E., et al. (2009). Surrogate end-

- points for overall survival in locally advanced head and neck cancer: Meta-analyses of individual patient data, *Lancet Oncology*, **10**(4), 341-350.
- Michiels, S., Pugliano, L., Marguet, S., Grun, D., Barinoff, J., et al. (2016). Progression-free survival as surrogate end point for overall survival in clinical trials of HER2-targeted agents in HER2-positive metastatic breast cancer, *Annals of Oncology*, **27**(6), 1029-1034.
- Moradian, H., Larocque, D. and Bellavance, F. (2019). Survival forests for data with dependent censoring, *Statistical Methods in Medical Research*, **28**(2), 445-461.
- Nelsen, R. B. (2006). *An Introduction to Copulas*, 2nd ed., Springer, New York.
- Oakes, D. (1989). Bivariate survival models induced by frailties, *Journal of the American Statistical Association*, **84**, 487-493.
- Oba, K., Paoletti, X., Alberts, S., Bang, Y. J., Benedetti, J., et al. (2013). Disease-free survival as a surrogate for overall survival in adjuvant trials of gastric cancer: A meta-analysis, *Journal of the National Cancer Institute*, **105**(21), 1600-1607.
- Pazdur, R. (2008). Endpoints for assessing drug activity in clinical trials, *The Oncologist*, **13**, 19-21.
- Peng, M. (2019). *Analysis of Complex Survival Data Subject to Semi-competing Risks*, Doctoral Thesis, Nanyang Technological University, Singapore.
- Peng, M. and Xiang, L. (2019). Joint regression analysis for survival data in the presence of two sets of semi-competing risks, *Biometrical Journal*, DOI: 10.1002/bimj.201800137.
- Peng, M., Xiang, L. and Wang, S. (2018). Semiparametric regression analysis of clustered survival data with semicompeting risks, *Computational Statistics & Data Analysis*, **124**, 53-70.
- Piedbois, P. and Crosswell, M. J. (2008). Surrogate endpoints for overall survival in advanced colorectal cancer: A clinician's perspective, *Statistical Methods in Medical Research*, **17**(5), 519-527.
- Proust-Lima, C., Séne, M., Taylor, J. M. and Jacqmin-Gadda, H. (2014). Joint latent class models for longitudinal and time-to-event data: A review, *Statistical Methods in Medical Research*, **23**(1), 74-90.
- Riley, R. D. (2009). Multivariate meta-analysis: The effect of ignoring within-study correlation, *Journal of the Royal Statistical Society, Series A*, **172**, 789-811.
- Rondeau, V. and Gonzalez, J. R. (2005). frailtypack: A computer program for the analysis of correlated failure time data using penalized likelihood estimation, *Computer Methods and Programs in Biomedicine*, **80**(2), 154-164.
- Rondeau, V., Pignon, J. P. and Michiels, S. (2015). A joint model for dependence between clustered times to tumour progression and deaths: A meta-analysis of chemotherapy in head and neck cancer, *Statistical Methods in Medical Research*, **24**(6), 711-729.
- Rondeau, V., Mauguen, A., Laurent, A., Berr, C. and Helmer, C. (2017). Dynamic prediction models for clustered and interval-censored outcomes: Investigating the intra-couple correlation in the risk of dementia, *Statistical Methods in Medical Research*, **26**(5), 2168-2183.
- Rotolo, F., Legrand, C. and Van Keilegom, I. (2013). A simulation procedure based on copulas to generate clustered multi-state survival data, *Computer Methods and Programs in Biomedicine*, **109**(3), 305-312.
- Rotolo, F., Paoletti, X. and Michiels, S. (2018). surrosurv: An R package for the evaluation of failure time surrogate endpoints in individual patient data meta-analyses of randomized clinical trials, *Computer Methods and Programs in Biomedicine*, **155**, 189-198.
- Schleiden, S., Klingler, C., Bertram, T., Rogowski, W. H. and Marckmann, G. (2013). What is personalized medicine: Sharpening a vague term based on a systematic literature review, *BMC Medical Ethics*, **14**(1), 55.
- Schneider, S., Demarqui, F. N., Colosimo, E. A. and Mayrink, V. D. (2020). An approach to model clustered survival data with dependent censoring, *Biometrical Journal*, **62**(1), 157-174.

- Sène, M., Taylor, J. M., Dignam, J. J., Jacqmin-Gadda, H. and Proust-Lima, C. (2016). Individualized dynamic prediction of prostate cancer recurrence with and without the initiation of a second treatment: Development and validation, *Statistical Methods in Medical Research*, **25**(6), 2972-2991.
- Sherrill, B., Amonkar, M., Wu, Y., Hirst, C., Stein, S, et al. (2008). Relationship between effects on time-to-disease progression and overall survival in studies of metastatic breast cancer, *British Journal of Cancer*, **99**, 1542-1548.
- Shi, Q. and Sargent, D. J. (2009). Meta-analysis for the evaluation of surrogate endpoints in cancer clinical trials, *International Journal of Clinical Oncology*, **14**(2), 102-111.
- Shih, J. H. and Emura, T. (2018). Likelihood-based inference for bivariate latent failure time models with competing risks under the generalized FGM copula, *Computational Statistics*, **33**(3), 1293-1323.
- Shih, J. H. and Emura, T. (2020). Penalized Cox regression with a five-parameter spline model (submitted).
- Shih, J. H., Lee, W., Sun, L. H. and Emura, T. (2019). Fitting competing risks data to bivariate Pareto models, *Communications in Statistics — Theory and Methods*, **48**(5), 1193-1220.
- 篠原紗也華 (2019). Dynamic prediction using joint frailty-copula model: Long-term prediction by Weibull distribution, 薬学学士論文, 北里大学薬学部薬学科臨床統計学教室.
- Shukla, N., Hagenbuchner, M., Win, K. T. and Yang, J. (2018). Breast cancer data analysis for survivability studies and prediction, *Computer Methods and Programs in Biomedicine*, **155**, 199-208.
- Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges, *Publications de l'Institut de Statistique de L'Université de Paris*, **8**, 229-231.
- Sofeu, C., Emura, T. and Rondeau, V. (2019). One-step validation method for surrogate endpoints in multiple randomized cancer clinical trials with failure-time endpoints, *Statistics in Medicine*, **38**(16), 2928-2942.
- Sofeu, C., Emura, T. and Rondeau, V. (2020). A joint frailty-copula model for meta-analytic validation of failure time surrogate endpoints in clinical trials (in revision).
- Suresh, K. (2018). *Dynamic Methods for the Prediction of Survival Outcomes using Longitudinal Biomarkers*, Ph.D Dissertation, University of Michigan Library.
- Suresh, K., Taylor, J. M., Tsodikov, A. (2019). A Gaussian copula approach for dynamic prediction of survival with a longitudinal biomarker, *Biostatistics*, DOI: 10.1093/biostatistics/kxz049.
- 高見澤勇 (2018). 化学療法が対象となる間質性肺炎合併肺がん患者の死亡率の動的予測: Joint frailty-copula モデルの臨床データへの適用, *RIMS Kôkyûroku*, No.2091, 31-37.
- Thompson, R. H., Leibovich, B. C., Lohse, C. M., Cheville, J. C., Zincke, H., et al. (2007). Dynamic outcome prediction in patients with clear cell renal cell carcinoma treated with radical nephrectomy: The D-SSIGN score, *The Journal of Urology*, **177**(2), 477-480.
- Yoshihara, K., Tajima, A., Yahata, T., Kodama, S., Fujiwara, H., et al. (2010). Gene expression profile for predicting survival in advanced-stage serous ovarian cancer across two independent datasets, *PLoS One*, **5**(3), e9615.
- van Houwelingen, H. C. and Putter, H. (2011). *Dynamic Prediction in Clinical Survival Analysis*, CRC Press, New York.
- Waldron, L., Haibe-Kains, B., Culhane, A. C., Riester, M., Ding, J., et al. (2014). Comparative meta-analysis of prognostic gene signatures for late-stage ovarian cancer, *Journal of the National Cancer Institute*, **106**(5), dju049.
- Wang, W. (2003). Estimating the association parameter for copula models under dependent censoring, *Journal of the Royal Statistical Society, Series B*, **65**(1), 257-273.

- Wang, Y. C., Emura, T., Fan, T. H., Lo, S. and Wilke, R. (2020). Likelihood-based inference for a frailty-copula model based on competing risks failure time data, *Quality and Reliability Engineering International*, DOI: 10.1002/qre.2650.
- Wu, B. H., Emura, T. and Michimae, H. (2020). Meta-analysis of individual patient data with semi-competing risks under the Weibull joint frailty-copula model, *Computational Statistics*, DOI: 10.1007/s00180-020-00977-1.
- Zhou, Y., Lu, Z., Shi, Y. and Cheng, K. (2019). The copula-based method for statistical analysis of step-stress accelerated life test with dependent competing failure modes, *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, **233**(3), 401-418.

Survival Analysis Using Copulas —Meta-analysis with Correlated Endpoints—

Takeshi Emura¹ and Hirofumi Michimae²

¹Department of Information Management, Chang Gung University

²School of Pharmacy, Kitasato University

With the rapid development this decade of open databases for biomedical researchers, we have gained access to complex and yet accurate patient-level information. For cancer patients in particular, these databases record information on individual patients, including overall survival time, time-to-tumor progression, tumor size, and gene expressions. In addition, there has been significant development of meta-analytical methodologies for analyzing data from different sources. To fully utilize such complex survival data, it is insufficient to apply the classical tools used in survival data analysis, such as Cox regression. In this paper, we review copula-based statistical methods for analyzing two survival time variables, namely overall survival time and time-to-tumor progression. We also review the joint frailty-copula model for individual patient data (IPD) meta-analysis methods, which account for the heterogeneity of patients from different sources. Regarding the construction of an appropriate likelihood function for a given dataset, we explain the importance of acknowledging the semi-competing risks relationship between overall survival time and time-to-tumor progression. Finally, we introduce a dynamic prediction method for overall survival time according to gene expressions and tumor progression, which may contribute to the development of personalized medicine.

サポートベクター回帰における感度分析による 変数選択の有効性の検証

—都道府県別全死因死亡率の影響要因の分析—

田辺 和俊¹・鈴木 孝弘²

(受付 2019 年 6 月 1 日；改訂 8 月 23 日；採択 12 月 26 日)

要 旨

本研究では、筆者らが開発したサポートベクター回帰における変数選択としての感度分析法を都道府県別死亡率の要因分析に適用し、その有効性の検証を行った。わが国では少子高齢化社会の次に到来するのが多死社会とされ、医師不足、終末期医療、孤独死、死に場所難民、放置空き家等、さまざまな社会問題の発生が危惧されている。地域の自治体が多死化対策を行うためには、多数の要因の中から死亡率に重大な影響を与える要因の解明とその相対的影響度の推定が重要になる。そこで、各都道府県の全死因年齢調整死亡率を目的変数、その影響要因の候補として生活習慣、医療・福祉、社会・経済の3分野の指標56種を説明変数とし、サポートベクター回帰により解析し、感度分析により影響要因を探索した。その結果、都道府県別死亡率を統計的に有意な精度で再現する11種の影響要因が得られ、喫煙率や高齢単身率だけでなく、これまで未検証の社会福祉士数が死亡率に大きな影響を与えることを見出した。また、地域の多死化対策への試みとして、死亡率が最高の青森県について、死亡率が最小の長野県との対比において影響要因の指標値に基づいて死亡率低減策を考察した。以上の結果から、広範な分野の多数の説明変数の中から感度分析により影響要因を探索する解析手法の有効性を実証した。

キーワード：サポートベクター回帰、変数選択、感度分析、全死因死亡率、影響要因。

1. はじめに

わが国では平均寿命が延び続ける一方で少子化に歯止めがかかっていないが、少子高齢化社会の次に到来する問題として危惧されているのが多死社会である。「平成29年版高齢社会白書」によれば、2010年に約120万人であった年間死亡者数は、団塊世代が80歳代後半となる2030年代には160万人を超え、2040年には最多の168万人に達すると予想されている。その影響として、医師不足、終末期医療、孤独死、死に場所難民、放置空き家等、さまざまな社会問題の発生が危惧されている。そのため、国は2025年をめどに地域包括ケア体制の構築に取り組み始めた。このシステムでは、本格的な高齢社会において地域全体で高齢者を支えるため、保健・医療・福祉関係者や、地域の各種団体や住民が連携し、介護、予防、医療、生活支援、住まいのサービスを一体化して提供していく地域づくりに取り組むとしている。

¹ 東洋大学 現代社会総合研究所：〒112-8606 東京都文京区白山 5-28-20

² 東洋大学 経済学部：〒112-8606 東京都文京区白山 5-28-20

ところで、20 世紀後半以降、人類の健康や病気が遺伝等の先天的要因だけでなく、経済、社会、文化、環境等の後天的要因によっても影響されることが広く認められるようになった。WHO の報告書 (2016) “The determinants of health” は健康格差の社会的影響要因として社会格差、ストレス、社会的排除、労働、失業、社会的支援等の 10 項目を挙げている。わが国の健康格差については、所得や失業等の経済・労働要因、学歴等の教育要因、病院や医師等の医療資源要因の影響が大きいとされている (福田・今井, 2007; Kagamimori et al., 2009; 日本公衆衛生学会, 2011; 橋本, 2012; 近藤, 2014)。日本学術会議は健康の社会格差の現状に関する報告書 (2011) を発表し、保健医療福祉政策に健康格差を考慮するよう提言した。

このように死亡率には非常に多くの要因が影響すると考えられるため、国や地域自治体が多死化対策を行うためには、多数の要因の中から死亡率に重大な影響を与える要因の解明とその相対的影響度の推定が重要になる。その手法として、地域の死亡率のデータを目的変数、いくつかの指標を説明変数として重回帰分析する生態学的分析手法 (Rothman, 2012) がある。この手法を用いて、これまでは特定地域の死亡率を解析した研究は多いが、地域別の死亡率について要因の解明を行った研究は少ない。

たとえば、Araki and Murata (1986) は都道府県別の全死亡率について学歴、所得、失業率等を説明変数に用いて重回帰分析した。Fukuda et al. (2004) は全国の市区町村の全死亡率を対象とし、説明変数に失業率、学歴、所得等を用いて解析した。大坪 他 (2009) は医師数、病床数、高齢世帯数等を説明変数に用いて分析した。しかし、これまでの研究では特定の分野の比較的少数の説明変数を用いて線形回帰分析 (OLS) により解析しているため、回帰決定係数が低く、死亡率の要因解明は未だ十分ではない。

本研究では、地域の多死化対策に資する情報提供を目的として、都道府県別の死亡率を目的変数とし、多種多様な指標を説明変数として非線形重回帰分析により解析し、死亡率の影響要因を解明する実証研究を試みた。本研究の意義は、先行研究での説明変数の不足を考慮して広範囲の多種多様な説明変数の中から影響要因を探索したこと、先行研究で多用されている線形重回帰分析の不十分さを改善するために、非線形回帰分析の 1 手法であるサポートベクター回帰 (SVR) を適用したこと、および多数の説明変数の中から影響要因を探索する変数選択の手法として、筆者らが開発した感度分析法を適用しその有効性を検証したことの 3 点である。

2. データと方法

2.1 死亡率データ

死亡率の影響要因を重回帰分析によって解明する手法を設計する際、3 つの検討すべき点がある。第 1 は目的変数に用いる死亡率の内容である。「平成 28 年人口動態統計」による日本人の死因は表 1 のように男女で異なり、また、年齢階層でも異なる。これらの死因の要因として喫煙、飲酒、肥満等の生活習慣の影響が大きいことが判明している。そのため、疾患別の死亡率について、生活習慣指標等を説明変数に用いて回帰分析を行った研究は多い。たとえば、筆者らは肺がん死亡率の解析から喫煙率や脂質摂取量等が重大要因であることを見出した (田辺他, 2016)。

しかし、地域自治体にとってはそのような多種類の死因に関する影響要因の情報は多死化対策の観点からは実施上の困難さがあり、むしろ、できるだけ簡潔かつ包括的な情報の方が施策には有用であると考えられる。そこで、目的変数には「平成 29 年人口動態調査特殊報告」による 2015 年の都道府県別の全死因年齢調整死亡率 (以下、死亡率) の男女平均を用いることにした。死因別の死亡率でなく全死因の死亡率の採用については、死因により影響要因に違いがあると考えられることから、また、死亡率の男女平均の採用については、男女間で死因に違いが

表 1. 性別・死因別の年齢調整死亡率の全死因死亡率に対する割合。平成 28 年人口動態統計。

	男女合計		男		女	
	死因	割合(%)	死因	割合(%)	死因	割合(%)
1	悪性新生物	28.5	悪性新生物	32.6	悪性新生物	24.2
2	心疾患	15.1	心疾患	13.8	心疾患	16.5
3	肺炎	9.1	肺炎	9.7	老衰	11.0
4	脳血管疾患	8.4	脳血管疾患	7.8	脳血管疾患	8.9
5	老衰	7.1	老衰	3.4	肺炎	8.5
6	不慮の事故	2.9	不慮の事故	3.3	不慮の事故	2.6
7	腎不全	1.9	自殺	2.2	腎不全	2.0
8	自殺	1.6	慢性閉塞性肺疾患	1.9	大動脈瘤及び解離	1.4
9	大動脈瘤及び解離	1.4	腎不全	1.8	血管性等の認知症	1.3
10	肝疾患	1.2	肝疾患	1.5	アルツハイマー病	1.3

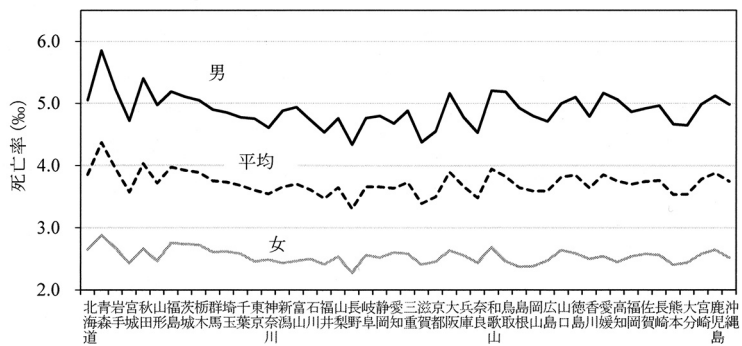


図 1. 都道府県別の全死因年齢調整死亡率(2015 年)。平成 29 年人口動態調査特殊報告。

あることから、疫学的観点からは異論があると思われる。しかし、本研究では地域の多死化対策に有用な要因を探ることが目的であるため、男女別の全死因年齢調整死亡率の単純平均を採用した。

図 1 に示すように、男女の死亡率の地域別傾向はほぼ類似しており、両者間の相関係数は 0.758 と高い。これは表 1 のように、男女の死因の 1 位と 2 位を占めているがんと心疾患の割合の合計が 4 割を超えているためと考えられる。したがって、目的変数に死亡率の男女平均を採用することは妥当と考えられ、男女平均の死亡率について解析して得られる影響要因は、男性と女性のそれぞれの影響要因を反映していると推測される。

男女平均死亡率の地域差をみると、青森県が 4.37% と突出して高く、最少(3.31%)の長野県との差は大きい。地域別の傾向では、宮城県を除く東北 5 県の死亡率が全般的に高く、隣接の北関東も高い。一方、南関東と中部地方は全国平均を下回るが、西日本は全国平均を超える府県が多い。また、近畿地方のように死亡率が高い府県と低い府県が隣接している地域もある。このように複雑な地域格差は多種多様な要因が死亡率に関与している可能性を示唆する。

2.2 説明変数

重回帰分析の第 2 の検討点は説明変数の選定であり、一般に 2 つのアプローチがある。1 つは目的変数に関する理論または仮説を立て、それに基づく比較的少数の説明変数を設定し、多くの場合、説明変数を固定して重回帰分析を行う理論主導アプローチである。この手法の目的は、設定した理論や仮説が目的変数データをどの程度説明するかの検証に重点が置かれる。しかし、研究者により理論や仮説が異なるため、解析で得られた結果が異なることが多い。また、

説明変数が少数なため、一般に回帰分析の決定係数が低く、目的変数データを十分再現できないことが多い。

一方、これと逆に、目的変数との関連にはこだわらず、入手可能な多数の説明変数を用いて解析し、その中から変数選択法を用いて目的変数を再現する必要最小限の説明変数を探索するデータマイニング的アプローチが考えられる。この手法は、目的変数に関する原理や機構が未解明であるが、社会政策上の要請から目的変数の予測システムの開発が急務であるような問題の場合に採用される。この問題例としては化学物質の発がん性予測を検証した筆者らの研究(Tanabe et al., 2010)がある。したがって、この手法の目的は解析に用いた既存データだけでなく、未知データに対しても高精度の結果を与える予測システムを開発することであり、有効な説明変数をいかに探索するかに力点が置かれる。しかし、得られた影響要因のすべてについて目的変数との関連性を必ずしも説明できない場合がありうる。

本研究ではこれら2つのアプローチの中間的手法を取る。すなわち、死亡率との関連が想定され、これまで検証されてきた多数の指標を説明変数に用いて重回帰分析により一括解析し、それらの中から変数選択法により死亡率の影響要因を探索する実証的手法である。そこで、説明変数としてこれまで検証された指標は100種以上に上るが、本研究ではそれらの中で検証例が多いものとして、喫煙・摂取食品・肥満等の生活習慣指標、医師・保健師・健診等の医療・福祉指標、所得・学歴・就業等の社会・経済指標を選定した。さらに、健康への効果が報告されているが、これまで回帰分析の説明変数として検証されていない数種の新規指標を追加した。

以上の方法により選定した説明変数56種を表2に示す。これらの都道府県別データは死亡率のデータが2015年時点であるため、各種政府統計の2010年～2015年の調査データを使用した。総数が多いため、それらの指標の記述統計量の一部のみ表3に示したが、分布の非対称性が高く、線形回帰(OLS)では統計的に有意な結果を得ることが困難と推測される変数が多い。なお、指標の単位が異なり、また下記の感度分析のために、各指標は最小と最大が0と1になるよう正規化して解析に用いた。

これらの説明変数を採用した理由を簡潔に記すと、生活習慣分野では、死亡率の中で大きな割合を占めるがんや、心・脳血管疾患(表1)等の主因とされている喫煙、飲酒、塩分、肥満、高血圧、ストレスなどを用いた。さらに、これらの疾患予防に効果が期待される歩数、スポーツなども追加した。医療・福祉分野では、死亡率の低下が期待できる医療費、医師、保健師、健診、がん検診(胃・肺・大腸・子宮・乳がんの5種の平均受診率)を採用した。さらに、特に高齢者の健康に貢献すると思われる指標として、老人医療費、訪問介護員、老人ホーム、介護施設などを採り上げた。社会・経済分野では、高齢ほど死亡率が高いことから、高齢者に関連する指標を採用した。失業者や貧困者は健康意識が低く死亡率が高いとされることから、所得や就業に関連する指標を採用した。また、多雪地域は多種の面での住民サービスの低下が死亡率の向上を招くと考えられることから、年間の降雪日を採用し、また、長野県のように山間地域が多い県民は足腰が鍛えられることで死亡率が低いとされていることから、標高を追加した。さらに、人口集中地区ほど医療や福祉が充実し、死亡率が低いと考えられることから、人口密度や都市化率を追加した。

2.3 解析方法

死亡率の影響要因探索における第3の検討事項は重回帰分析の手法である。死亡率要因分析の先行研究では、最も一般的な線形重回帰分析(OLS)が多用されてきた。しかし、表3に示すように、本研究の説明変数の中には、死亡率に対して非線形関係にあるものが多いため、OLSでは統計的に有意な結果を得ることが困難と思われる。

先行研究では、このような非線形関係に対処するために、一部の説明変数の2乗の項の追加

表 2. 説明変数の定義、およびデータの出所。出所)国生：国民生活基礎調査(2013)、国健：国民健康栄養調査(2012)、特健：特定健康診査(2010)、社統：社会生活統計指標(2013)、家計：家計調査(2013)、社基：社会生活基本調査(2014)、衛行：衛生行政報告(2014)、社福：社会福祉振興・試験センター(2015)、社人：社会・人口統計体系(2014)、筆者：田辺・鈴木(2018)、賃金：賃金構造基本統計調査(2010)。

分野	説明変数	定義(単位)	出所
生活習慣	喫煙	喫煙者の割合(%)	国生
	飲酒	飲酒者の割合(%)	国生
	エネルギー	1日の平均エネルギー摂取量(Kcal)	国健
	脂質	1日の平均脂質摂取量(g/日)	国健
	野菜	1日の平均野菜摂取量(g/日)	国健
	塩分	1日の平均食塩摂取量(g/日)	国健
	規則食事	規則正しく食事している人の割合(%)	国生
	バランス食事	バランスの取れた食事をしている人の割合(%)	国生
	薄味食事	薄味の食事をしている人の割合(%)	国生
	適量食事	食べ過ぎないようにしている人の割合(%)	国生
	健康無関心	日頃健康のために特に何もしていない人の割合(%)	国生
	肥満	BMI (Body Mass Index)の平均(kg/m ²)	国健
	高血圧	収縮期血圧 140mmHg 以上の該当者の割合(%)	特健
	メタボリックシンドローム	メタボリックシンドローム該当者の割合(%)	特健
	ストレス	悩みやストレスのある人の割合(%)	国生
	歩数	1日の平均歩数(歩)	国健
	スポーツ	スポーツの年間行動者率(%)	社統
	交際	世帯当たりの交際費年間支出金額(千円)	家計
	趣味	趣味・娯楽の高齢行動者の割合(%)	社基
	医療・福祉	保健医療費	世帯当たりの保健医療費年間支出金額(千円)
老人医療費		被保険者当たりの後期高齢者医療費(千円)	社統
老人福祉費		高齢人口当たりの老人福祉費(千円)	社統
医師		人口当たりの医療施設に従事する医師数(人)	社統
保健師		人口当たりの保健師数(人)	衛行
社会福祉士		人口当たりの社会福祉士数(人)	社福
介護福祉士		人口当たりの介護福祉士数(人)	社福
訪問介護員		人口当たりの訪問介護員数(人)	社人
老人ホーム		高齢人口当たりの老人ホーム数(所)	社統
老人福祉センター		高齢人口当たりの老人福祉センター数(所)	社統
介護老人福祉施設		高齢人口当たりの介護老人福祉施設数(所)	社統
介護療養型医療施設		高齢人口当たりの介護療養型医療施設数(所)	社統
高齢生活保護		高齢人口当たりの生活保護高齢者数(人)	社統
健診		健診の受診率(%)	国生
がん検診	5種のがん検診の平均受診率(%)	国生	
社会・経済	高齢世帯	高齢者のいる世帯の割合(%)	社統
	単身世帯	単身世帯の割合(%)	社統
	高齢単身	高齢単身世帯の割合(%)	社統
	学歴	平均就学年数(年)	社統
	社会教育費	人口当たりの社会教育費(千円)	社統
	公民館	人口当たりの公民館数(館)	社統
	高齢者講座	人口当たりの高齢者講座数(講座)	社統
	高齢ボランティア	高齢ボランティア活動者の割合(%)	社基
	家計収入	世帯の平均月間収入(千円)	家計
	貯蓄	世帯の貯蓄現在高(千円)	社統
	負債	貯蓄総額に対する負債額の比率	社統
	就業	就業者の割合(%)	社統
	高齢就業	高齢就業者の割合(%)	社統
	失業	完全失業率(%)	社統
	貧困率	貧困層の比率を示す指数(%)	筆者
	残業	超過労働時間(時間)	賃金
	標高	都道府県所在地の標高(m)	社統
	気温	年間の平均気温(℃)	社統
	日照	年間の日照時間(時間)	社統
	降雪	年間の平均降雪日数(日)	社統
人口密度	可住地面積当たりの人口密度(人/km ²)	社統	
都市化	人口集中地区人口割合(%)	社統	

表 3. 説明変数の記述統計. *死亡率との単相関係数.

分野	説明変数	最小	平均	最大	標準偏差	歪度	相関係数*
生活習慣	喫煙	16.3	20.5	26.6	2.1	0.62	0.474
	飲酒	22.8	26.4	31.2	1.9	0.31	0.182
	エネルギー	1,843	2,059	2,143	46.2	-2.54	-0.140
	脂質	52.0	58.8	64.0	2.5	-0.27	-0.315
	野菜	233	280	344	21	0.67	0.024
	塩分	8.6	10.0	10.9	0.5	-0.36	0.078
	規則食事	44.0	52.6	58.9	3.2	-0.45	-0.043
	バランス食事	30.8	35.9	41.7	2.4	-0.08	-0.384
	薄味食事	26.0	31.6	35.9	2.2	-0.25	0.062
	適量食事	32.5	38.4	41.5	1.8	-0.86	-0.005
	健康無関心	13.2	16.0	18.1	1.2	-0.51	0.333
	肥満	22.7	23.3	24.4	0.4	0.91	0.374
	高血圧	16.3	19.6	24.2	1.9	0.70	0.389
	メタボリックシンドローム	12.8	14.7	17.7	1.1	0.66	0.332
	ストレス	44.0	47.5	51.1	1.8	0.17	-0.390
	歩数	5,743	7,034	8,048	514	0.14	-0.354
	スポーツ	11.2	27.0	49.9	8.1	0.30	-0.250
	交際	81	140	185	24	-0.19	-0.231
	趣味	22.2	26.3	30.7	2.3	-0.21	-0.544
	医療・福祉	保健医療費	81	124	167	19	0.75
老人医療費		745	926	1,182	108	0.39	-0.096
老人福祉費		154	210	255	24	-0.28	0.150
医師		223	309	423	45	0.54	-0.156
保健師		21.0	43.0	64.4	10.8	-0.22	-0.034
社会福祉士		109	157	216	24	0.26	-0.547
介護福祉士		756	1,275	1,772	255	-0.07	0.306
訪問介護員		30	74	179	34	1.14	0.012
老人ホーム		35	65	143	23	1.17	0.283
老人福祉センター		1.0	6.9	13.2	2.9	0.48	0.074
介護老人福祉施設		1.0	5.0	16.1	3.9	1.28	-0.015
介護療養医療施設		12.9	23.1	36.0	4.5	0.24	0.201
高齢生活保護		6.2	23.9	59.7	12.5	1.10	0.283
健診		54.9	62.4	70.0	3.7	-0.09	-0.257
がん検診	23.3	31.3	43.0	4.0	0.54	-0.030	
社会・経済	高齢世帯	30.9	44.8	55.8	5.3	-0.44	0.253
	単身世帯	25.5	31.5	47.3	4.1	1.28	-0.013
	高齢単身	8.3	11.4	16.5	1.9	0.71	0.327
	学歴	8.9	9.8	10.7	0.4	-0.06	-0.121
	社会教育費	6.0	12.2	26.3	4.3	1.51	-0.113
	公民館	6.1	188	724	146	1.58	-0.104
	高齢者講座	20	449	1,225	304	0.64	0.151
	高齢ボランティア	14.1	24.2	33.5	4.1	-0.30	-0.486
	家賃収入	420	531	633	57	-0.13	-0.232
	貯蓄	5,747	14,497	19,669	3,082	-0.56	-0.446
	負債	0.16	0.33	0.85	0.11	2.54	0.212
	就業	50	55	59	2.6	-0.26	-0.184
	高齢就業	17.9	22.7	28.7	2.1	0.21	-0.200
	失業	2.9	4.2	6.3	0.7	0.49	0.404
	貧困率	10.2	12.9	16.9	1.4	0.72	0.295
	残業	2.1	11.3	40.0	9.3	2.00	-0.112
	標高	2.4	41.9	371	73	3.01	-0.265
	気温	9.3	16.2	24.1	2.4	-0.22	-0.234
	日照	1,622	1,921	2,188	160	-0.02	-0.174
	降雪	0.0	26.2	114	27.8	1.54	0.347
人口密度	239	1,357	9,605	1,757	3.41	-0.166	
都市化	24.2	52.5	98.4	18.9	0.90	-0.200	

や、対数変換を行った論文がある。また、説明変数間の交絡効果に対処するために2変数の積の項を追加して解析した論文がある。さらに、説明変数間に相関の高い組がある場合、OLSでは多重共線性効果により回帰分析が不安定になるため、高相関の組の一方を削除して解析した論文がある。さらに、本研究のように説明変数が目的変数のデータ数より多い場合、OLSではそのままの解析は不可能なため、主成分分析等の手法を用いて説明変数を減らして解析した論文がある。しかし、これらの対処は ad hoc 的なものであり、完全な解決策とはいえない。

本研究では、これらの問題を解決するために非線形重回帰分析の手法としてサポートベクターマシン(SVM)(大北, 2005; 小野田, 2007; 阿部, 2011)を適用した。SVMは、説明変数の数値に対してカーネルと呼ぶ非線形関数を用いて学習パターンを別の空間(超平面)に写像し、その空間で線形回帰を行う。この操作により、説明変数の元の数値での非線形回帰が可能になり、目的変数と説明変数の間の任意の関係に対して高精度の回帰結果が得られる。さらに、SVMでは、①説明変数間の交絡効果が予想される場合でもこの効果が自動的に考慮されるため、変数の積の項の追加は不要、②変数間に強い相関がある場合でも解析可能であり、多重共線性問題は生じない、③目的変数のデータ数以上の説明変数を用いた解析が可能、などの利点がある。以上のSVMの多くの利点はカーネル回帰という手法の採用によるものである(赤穂, 2008)。

SVMのソフトウェアはLIBSVM ver. 3.11(Chang and Lin, 2011)の回帰機能(SVR)を、カーネル関数はradial basis function(RBF)を用いた。多数の説明変数の中から影響要因を探索するためにはSVRモデルと説明変数の最適化を行う必要がある。本研究では、前者については、LIBSVMの ϵ SVRのオプションパラメータのうち、最も重要な g (RBF関数 $\exp(-g*|u-v|^2)$ における係数)、 c (cost)および p (終了のしきい値)(これらのパラメータの詳細についてはChang and Lin, 2011 参照)の最適化を1個抜き交差検証法(LOOCVT)により行った。

後者に関しては、回帰分析では一般に説明変数の中に有効でないものがあると過学習状態に陥り、学習データに対する誤差は減少するが、予測データについての誤差は増大するため、必要最小限の説明変数を抽出する変数選択が必要である。本研究では迅速な変数選択法として感度分析法を採用した。この感度分析法は、目的変数に対する各説明変数の感度を計算し、感度の低い変数を順次削除しながらSVRモデルを最適化し、目的変数の予測値と実測値の平均二乗誤差が最小となる組み合わせを探索する方法である。筆者らはこの感度分析法の有用性を別のタイプの問題において検証した(Tanabe et al., 2013)が、本研究ではデータの特性が異なる問題において有用性をさらに詳細に検証した¹⁾。

この感度分析法の具体的な方法は、感度を推定する説明変数以外の変数はそれぞれ平均に固定し、当該変数のみ実際の数値に設定したときの死亡率のSVR出力値の変化をその変数の感度として算出する。その際、多くの説明変数の場合にSVRの出力値は説明変数の実測値にほぼ線形に変化することが認められるため、実測値と出力値の線形回帰から算出される回帰直線の傾きをその説明変数の感度として推定する。

そこで、以下の手順により影響要因の探索を行った(操作の流れを図2に示す)。

- ① 1個抜き交差検証法(LOOCVT)を用いてSVRモデルを最適化するために、LIBSVMのCVT機能パラメータ v を47(都道府県の数)に設定し、47行57列の学習データを用いてSVRのモデルパラメータ(g, c, p)をそれぞれ0.001~1000の範囲でグリッドサーチし、全都道府県について死亡率の予測値と実測値との平均誤差(RMSE)が最小になる最適条件を探索する。
- ② 各説明変数の感度を求めるために、当該変数は実際の数値のまま、その他の変数は全都道府県の平均値に設定したデータセットを最適モデルに入力して出力値を求め、当該変数の実測値を説明変数、出力値を目的変数とする単回帰分析を行い、回帰直線の傾きをその変

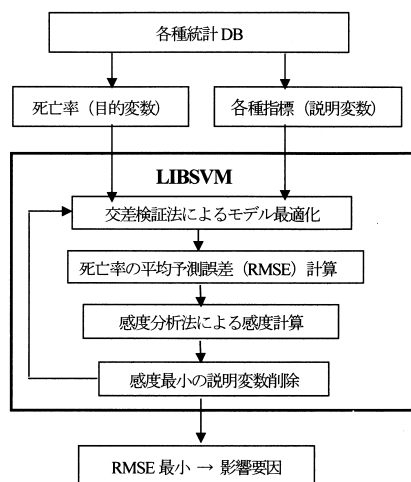


図 2. 死亡率の影響要因探索の流れ図.

数の感度とする.

- ③ 全説明変数の中で感度の絶対値の最も小さい変数を取り除き, ①②の操作を繰り返し, RMSE が最小になる説明変数の組み合わせを死亡率の影響要因とする.

3. 結果

以上の方法により, 56 種の説明変数の中から影響要因を探索した結果, 11 種の変数を用いた場合に死亡率の予測値と実測値との RMSE が最小となり, その前後で RMSE は増加した. この 11 種の変数による最適モデルの自由度調整済回帰決定係数 (AR^2) は 0.740 であり, 危険率 1% で有意と判定される.

この SVR の結果と比較するために, 線形回帰分析 (OLS) を行った. 先ず, 56 種の説明変数のデータについて, 非線形性に対処するために対数変換を行い, さらに最小 0, 最大 1 に正規化した. 次に, これら 56 種のデータを 28 種の 2 群にランダムに分割し, それぞれの群で OLS を行い, その結果から各群で有意性の低い (p -値の高い) 各 8 変数を取り除いた. さらに, 両群を合体させた 40 種の説明変数により OLS を行い, 変数減少法による変数選択を実行した (使用ソフトは (株) エスミの Excel 多変量解析 ver.7.0). その結果, AR^2 は 0.677 となり, SVR とほぼ同程度の回帰結果が得られた. しかし, この際に選択された変数は日照, 健康無関心, 介護福祉士, 趣味, 高齢ボランティア, 歩数, 標高, 保健師, 老人福祉費, 高齢者講座, 貧困率, 公民館, 老人福祉センター, 介護老人福祉施設の 14 種であり, SVR で寄与率の上位になった喫煙や高齢単身などが選択されないのは疑問である. さらに, このうちの日照, 介護福祉士, 歩数, 標高, 高齢者講座, 貧困率, 老人福祉センター, 介護老人福祉施設の 8 変数については, 偏回帰係数の符号が想定とは逆となった. ここで, 得られた有意な変数の偏回帰係数の符号に関して想定との整合性を判定したが, 筆者らは先行研究や疫学調査などの結果から, 各説明変数について死亡率に及ぼす影響の方向を想定している.

以上の結果から OLS の結果は受け入れできないと判定され, 一方, SVR では回帰決定係数が低く, また 11 種の影響要因の感度の符号が筆者らの想定と整合することから, 本研究の SVR と感度分析の有効性, およびその解析による要因の信頼性が実証できたと判断される.

表 4. 影響要因の内訳, および死亡率に対する感度と寄与率.

	影響要因	分野	感度			寄与率 (%)
			危険要因	抑制要因	(誤差)	
1	社会福祉士	医療・福祉		-0.231	(0.032)	20.2
2	喫煙	生活習慣	0.211		(0.031)	16.8
3	高齢单身	社会・経済	0.177		(0.022)	11.9
4	がん検診	医療・福祉		-0.157	(0.021)	9.3
5	健康無関心	生活習慣	0.153		(0.019)	8.9
6	塩分	生活習慣	0.140		(0.018)	7.4
7	保健師	医療・福祉		-0.139	(0.018)	7.3
8	肥満	生活習慣	0.112		(0.018)	4.7
9	趣味	生活習慣		-0.111	(0.017)	4.7
10	交際	生活習慣		-0.109	(0.017)	4.5
11	都市化	社会・経済		-0.106	(0.016)	4.2

次に, 得られた各要因の死亡率への相対的影響度について考察するために, 影響要因 i の感度 S_i (前頁の操作②で求めたもの) から式

$$(3.1) \quad C_i(\%) = \frac{S_i^2}{\sum_{i=1}^{11} S_i^2} \times 100$$

により死亡率に対する寄与率 C_i を推定した. 影響要因の内訳, 死亡率に対する感度とその推定誤差, および寄与率を表 4 に示す. 得られた影響要因の感度の誤差は感度分析過程における感度値のばらつきから推定したが, その大きさを見ると, 感度はいずれも信頼できる精度で推定されていると考えられる.

また, 感度分析で得られた各要因の感度は, 他の変数は固定し, 当該変数のみ数値を変化させた時の死亡率の変化から求めたので, 死亡率に対する当該要因の正味の影響度を表わしている. したがって, 感度が正の 5 種の要因 (喫煙, 高齢单身, 健康無関心, 塩分, 肥満) は死亡率増加に寄与する危険要因であり, 感度が負の 6 種の要因 (社会福祉士, がん検診, 保健師, 趣味, 交際, 都市化) は死亡率減少に寄与する抑制要因であると解釈できる.

4. 考察

4.1 変数選択における感度の符号について

表 4 に示した 11 種の影響要因のうち, 感度が正の危険要因 5 種と, 感度が負の抑制要因 6 種についての結果は, 各要因の死亡率に対する影響の予測方向と整合するが, この点について補足しておく. 本研究では, 図 2 に示したように, 感度分析による変数選択において感度最小の変数を順次削除していき, 死亡率の平均予測誤差 (RMSE) が最小になったときの説明変数を死亡率の影響要因としたが, 実際の削除すべき説明変数の選択においては, 感度の数値だけでなく, その符号も考慮している. これは通常の線形重回帰分析 (OLS) においても遭遇する問題であるが, 実例でこの問題に対する筆者らの対処法を示そう.

本研究で使用した都道府県別の死亡率データを目的変数とし, 簡単のために, 説明変数は喫煙率と魚介類摂取量の 2 変数のみとする. OLS では

$$(4.1) \quad \text{死亡率} = a \times \text{喫煙率} + b \times \text{魚介類摂取量}$$

として, 係数 a と b を最小二乗法で計算すると, a, b は正の数値になる. また, 本研究で用いた SVR の感度分析の結果でも a, b とも符号は正になる.

周知のように喫煙は肺がんや心疾患, 脳血管疾患などの重大疾患 (表 1) の重要原因であると

されることから、係数 a が正符号になることは妥当である。一方、係数 b の符号については、死亡率国内 1 位の青森県は魚介類の摂取量も 1 位で、死亡率最下位の長野県の摂取量が低く、47 都道府県での死亡率と魚介類摂取量との相関係数が 0.418 と正符号であることから、係数 b も正符号になることは妥当な結果であるようにみえる。しかし、魚介類に多く含まれる DHA や EPA などの栄養成分は、悪玉コレステロールや中性脂肪を下げ、動脈硬化の予防に有効とされていることから、魚介類の摂取は死亡率の低下に働く可能性が高いと考えられている。したがって、OLS と SVR での係数 b の符号が正であるという結果は想定外であり、説明変数を多数に増やしてもこの結論は変わらない。

このような回帰係数の符号が想定と異なる場合、OLS の解説書などでは多重共線性問題として、説明変数間の相関分析を行い、高相関の組の一方を削除する処理を勧めていることが多い。しかし、47 都道府県での喫煙率と魚介類摂取量との相関係数は 0.352 と低く、多重共線性問題ではないと考えられる。そこで本研究では、表 2 に挙げた説明変数について予備的な SVR の感度分析を行い、想定とは異符号の説明変数は不适当(厳密にいえば、本研究のような都道府県別の死亡率の影響要因を探索する場合には)として削除した。このような前処理を施すことにより本研究で得られた影響要因 11 種に関しては、栄養学、疫学、社会学などの観点から、危険要因であるか、抑制要因であるかの点で問題がない結果を得ることができたと考えられる。

4.2 感度と相関係数

影響要因を含む全説明変数について、死亡率との単相関係数と、死亡率に対する感度との散布図を図 3 に示す。この図をみると、全説明変数の中で感度が高いものとして選ばれた影響要因の中には相関係数の絶対値が低いものがある。また逆に、説明変数の中には相関係数が高いにもかかわらず、感度が低く影響要因にならなかったものが多数ある。さらに、感度と相関係数が異符号の変数も多数ある。これらの結果は、感度が死亡率に対する説明変数個々の正味の感度を示すのに対し、死亡率との相関係数には他の説明変数の寄与が含まれているためであるが、先行研究における解析に関して 2 つの問題点を示唆する。

第 1 は死亡率との相関分析に基づく説明変数の影響度の考察である。これまでは都道府県の死亡率との相関係数に基づいて各説明変数の影響度を分析した研究が多い。しかし、本研究の結果はこのような相関分析で得られた先行研究の結果には疑問があることを示唆する。第 2 は説明変数の選定であり、死亡率と相関の高い指標を説明変数に選び重回帰分析を行った先行研究が多い。しかし、このような方法では選定された説明変数以外に死亡率に大きな影響を与え

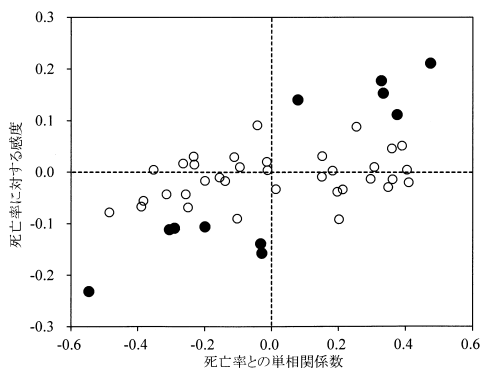


図 3. 全説明変数の死亡率との単相関係数 vs 感度の散布図(●:影響要因, ○:非影響要因).

る変数を見落とす可能性がある。たとえば、本研究で影響要因となったがん検診は死亡率との相関係数が -0.030 、保健師は -0.034 と非常に低いため、それらを説明変数として見落とす可能性が高いと考えられる。

以上を総括すると、先行研究では説明変数の選定や死亡率との影響度に関する考察において疑問がある場合が考えられる。これに対し、本研究では死亡率に対する正味の感度に基づいて広範な分野の多数の説明変数の中から探索した影響要因が死亡率を統計的に有意な精度で再現する結果が得られたことから、本研究の解析法の有効性が実証されたと考えられる。

4.3 影響要因に関する全般的考察

表4の11種の影響要因を分野別に分類すると、生活習慣分野が6要因、医療・福祉分野が3要因、社会・経済分野が2要因であり、影響要因が3分野にまたがっている。このことは多くの先行研究のように、特定の分野の比較的少数の説明変数を用いた解析では、その他の分野の影響が大きい説明変数が欠落している可能性があり、得られた死亡率の影響要因について信頼性に疑問があることを示唆する。たとえば、Fukuda et al. (2004)は失業率、学歴、所得等を中心とする社会経済指標を、大坪 他 (2009)は医師数、病床数等の医療資源指標を、また、林 他 (2009)は喫煙、飲酒、食塩摂取量等の健康・栄養関連指標をそれぞれ用いて解析しているが、これらの結果はいずれも回帰係数が低く、それらの結果には疑問の余地がある。これに対し、本研究では広範な分野の多数の説明変数を用いて非線形重回帰により解析したことで、影響要因に関して先行研究より信頼性の高い結果が得られたと考えられる。

影響要因の寄与率を分野別に合計すると、生活習慣47.0%、医療・福祉36.9%、社会・経済16.1%となる。海外では死亡率や平均寿命等の健康に関連する社会経済的要因はSES (Socio-Economic Status)と呼ばれ、中でも三大要因とされる学歴、所得、職業と健康との関係が多く研究されている (Adler and Newman, 2002)。日本人については所得、失業、学歴等の経済・社会要因、病院や医師等の医療要因の影響が大きいとされている (福田・今井, 2007; Kagamimori et al., 2009; 日本公衆衛生学会, 2011; 橋本, 2012; 近藤, 2014)。しかし、これまでは本研究のように広い分野の多数の説明変数を用いて解析した研究はない。

4.4 生活習慣分野の影響要因

影響要因を分野別に考察すると、生活習慣分野では、上記のように人間の死亡には生活習慣が大きな影響を与えるとされていることから、喫煙、飲酒等の19種の説明変数をとり上げた。その内、影響要因となったのは喫煙、健康無関心、塩分、肥満、趣味、交際の6種であり、その寄与率の合計は3分野の中で最大である。この結果に対し、これまで喫煙や飲酒等の生活習慣の影響を調べた研究はあるが、上記のように説明変数の分野が限定的な研究がほとんどのため、生活習慣要因の影響が全体の中でどの程度を占めているかを明らかにしたものはない。

本研究で見出された影響要因の中でも、喫煙が寄与率16.8%で影響要因2位となったことが注目される。この結果は日本人の死因の上位を占めるがん、心疾患、肺炎、脳血管疾患(表1)のいずれにも喫煙が悪影響を及ぼすとされていることによく対応している。死亡率に対する喫煙の影響を調べた先行研究としては、林 他 (2009)は都道府県別死亡率について喫煙、飲酒、食品摂取量等の健康・栄養関連指標を用いて解析し、歩数や塩分は影響要因になるが、喫煙はならないとした。片山他(2012)は都道府県別乳がん死亡率について喫煙、ボランティア活動、教育施設等の説明変数を用いて解析し、喫煙の影響が大きいことを示した。しかし、本研究のように都道府県別全死因死亡率について多種多様な説明変数を用いて解析し、喫煙の影響が大きいことを示した先行研究はない。

しかし、喫煙に関してはその単独要因としての影響だけでなく、他の要因と複合した関連要

因としての影響の大きさが注目される。すなわち、喫煙、健康無関心(要因 5 位)、塩分(6 位)、肥満(8 位)および医療・福祉分野のがん検診(4 位)を統合すると健康関心要因とみなすことができ、これら 5 要因の寄与率の合計は 47.2% に達する。このうちで、健康無関心者のデータは「国民生活基礎調査」において、日頃健康のために実行している事柄は何か、という質問に対し、規則正しく食事している、適度に運動している、十分に睡眠している、煙草を吸わない、などの項目には回答せず、特に何もしていない、と回答した人の割合として集計した。一方、これまで健康への関心度と健康指標との関連を調べた研究としては、森本・丸山(2001)は地域住民や高齢者等を対象に喫煙、飲酒、運動などの生活習慣と各種疾患との相関を解析し、健康意識の低い人ほど慢性疾患の発症率が高いことを示した。また、熊谷 他(2008)は地域住民を対象として主観的精神健康度と身体健康度の関連を解析した。しかし、都道府県別死亡率に対する健康関心度の影響を解析した先行研究はない。

4.5 医療・福祉分野の影響要因

医療・福祉分野については、直接的に人の死亡に関連することから、医療費、医療施設数・従事者数等の医療資源指標を中心に 15 種の説明変数を取り上げた。その結果、影響要因となったのは社会福祉士、がん検診、保健師の 3 種であり、その寄与率の合計は 36.9% で、死亡率に大きな影響を与えることを示している。先行研究でも医療分野の指標が死亡率に対して影響要因になることを示した研究が多い。大坪 他(2009)は全国の市区町村を対象として各種医療資源等の指標を説明変数に用いて重回帰分析を行い、医師数や救急告示病院数が死亡率に有意な影響を与えることを示した。また、五十嵐 他(2014)は全死因の都道府県別自宅死亡割合と医療社会的指標の相関を分析し、病床数が少なく、在宅医療へのアクセスが良い都道府県で自宅死亡割合が高い傾向にあることを示した。しかし、先行研究では用いられた説明変数の範囲が限定的であるため、医療・福祉分野の影響要因が死亡率にどの程度の寄与を及ぼすかについては明らかにされていない。

本研究の解析で得られた影響要因 4 位のがん検診が寄与率 9.3% で死亡率に影響しているという結果は、表 1 のように現在のわが国ではがんによる死亡が全死因の 28.5% を占めている実態に対応しており、がん検診の有効性を示している。蔵満 他(2014)は北海道の 179 市町村の健康寿命について保健医療環境や社会経済的指標を説明変数に用いて重回帰分析を行い、肺がん検診率が健康寿命の延伸に関連していることを示した。しかし、全都道府県の全死因死亡率に対するがん検診の効果を検証した先行研究はない。

しかし、本研究の結果の中でもっとも注目すべきは、これまで検証されていなかった社会福祉士数が寄与率 20.2% で第 1 位の影響要因になり、各種要因の中で死亡率に対してもっとも大きな影響を与える結果が得られたことである。2006 年に改正された介護保険法により全国に設置された地域包括支援センターには、保健師、社会福祉士、介護福祉士等主任ケアマネジャーの 3 職種が配置されている。社会福祉士は医療・福祉等の施設において病気やケガなどにより日常生活を送るのが難しい人に対して助言や指導を行う職業専門職であり、その活動が地域住民の健康に機能している状況については多くの報告がある。しかし、これまで社会福祉士が地域の健康指標の改善に及ぼす影響を統計的に検証した研究はない。蔵満 他(2014)は北海道の 179 市町村の健康寿命について保健医療環境や社会経済的指標を説明変数に用いて重回帰分析を行い、保健師が健康寿命の延伸に関連していることを示したが、社会福祉士は検証していない。この要因についてはこれまで実証例がないため、今後の検証研究が待たれる。

4.6 社会・経済分野の影響要因

社会・経済分野では 3 分野の中で最多の 21 種の説明変数を用いて解析した結果、高齢単身

と都市化の2種が影響要因で、寄与率16.1%で死亡率に影響していることが見出された。この中で注目すべきは、高齢単身が3位の11.9%という高率で死亡率に影響している結果が得られたことである。高齢単身の割合と死亡率等との関連を解析した先行研究としては、尾形他(2011)は慢性透析患者数と高齢単身者数や病院数等との相関を解析し、吉本他(2011)は救急搬送時の転倒数と高齢単身率との相関を解析している。しかし、本研究のように都道府県別死亡率と高齢単身世帯数との関連を解析した先行研究はない。

この高齢単身についても、その単独要因としての影響だけでなく、他の要因と複合した関連要因としての影響の大きさが注目される。すなわち、生活習慣分野の趣味、交際、および社会・経済分野の都市化との統合であり、これらに高齢単身を加えた4要因は死亡率に対して共通する関連性があり、それは社会との関連性・孤立性である。「高齢社会白書」によれば、近年の日本では職場、地域、家庭における人間関係が希薄化し、社会的に孤立する人が増え、健康への影響が懸念されている(近藤, 2004; 内閣府, 2018)。藤本他(2011)はこのような社会的孤立と健康との関連性について実証的研究を行い、社会・経済分野の三大要因である所得、学歴、職業だけでなく、社会的孤立も健康に影響している可能性があることを示した。また、安梅他(2006)は農村に居住する高齢者について社会とのかかわり状況を示す社会関連性指標を用いて死亡率との関連を回帰分析し、活動参加、趣味、積極性等の社会関連性が乏しいと死亡率が高いという結果を得た。しかし、本研究で全都道府県の死亡率の解析から得られた社会連携関連4要因の寄与率の合計が25.3%に達することから、社会的孤立が健康に大きく影響していることを実証したのは本研究が初めてである。

以上を総括すると、本研究では47都道府県別の死亡率について56種の説明変数を用いて非線形重回帰分析した結果、死亡率を統計的に有意な精度で再現する11種の影響要因が得られ、その内でも、これまで未検証の社会福祉士が死亡率にもっとも大きな影響を与えること、健康関心要因や社会関連要因といった複合要因も大きな影響を与えること等、新たな結果が得られ、広範な分野の多数の説明変数の中から感度分析により影響要因を探索する解析手法の有効性を実証した。

4.7 地域の多死化対策の試論

最後に、本研究の結果に基づいて地域の多死化対策に資する情報提供の可能性を考察してみる。対象地域として、国内で死亡率が長年、最高の青森県と、近年、最低の長野県を取り上げると、両県での影響要因11種の国内順位は表5のようになる。ここで各要因の危険度順位は、感度が正の危険要因については指標値の降順、符号が負の抑制要因については昇順の順位である。この表を見ると、長野県は塩分摂取量以外は全体的に危険度順位が低い要因が多く、死亡率の全国最低に対応している。これに対し、青森県は危険度順位の高い要因が多く、特に寄与率の高い社会福祉士と喫煙の危険度順位の高さが青森県の死亡率の国内最高に大きく関与していることが明らかである。

青森県では現在、健康づくりの推進や自殺対策などの短命県返上への模索を実施しているが、本研究の影響要因の結果を基に青森県の死亡率改善策を考えると、まず、社会福祉士の増員と、喫煙率や塩分摂取を中心とする健康意識の向上が考えられる。青森県の社会福祉士数(人口当たり)は全国で3番目に少ない123人であり、これは最多の新潟県の6割弱にとどまる。社会福祉士の充実が青森県のみならず、全国の自治体の死亡率低減に費用対効果の高い対策であると考えられる。また、喫煙率24.9%は北海道の26.6%に次ぐ全国2位であり、最下位の奈良県の16.3%とは大きな開きがある。

青森県は降雪日数が2位の雪国県であり、屋外活動率の低さなどが肥満(5位)につながり、飲酒率1位、喫煙率2位、塩分摂取6位などの健康意識の低さ、無関心につながり、最終的に

表 5. 影響要因の指標値の最小、最大と該当県、および青森県と長野県の指標値と危険度順位. *危険度順位: 危険要因は指標値の降順, 抑制要因は指標値の昇順.

影響要因	影響	最小		最大		青森県		長野県	
		指標値	該当県	指標値	該当県	指標値	順位*	指標値	順位*
社会福祉士	抑制	109	茨城	216	新潟	123	3	162	28
喫煙	危険	16.3	奈良	26.6	北海道	24.9	2	19.3	31
高齢单身	危険	8.3	滋賀	16.5	高知	12.1	14	10.5	29
がん検診	抑制	23.3	大阪	43.0	山形	31.9	28	35.4	40
健康無関心	危険	13.2	長野	18.1	青森	18.1	1	13.2	47
塩分	危険	8.6	沖縄	10.9	長野	10.5	6	10.9	1
保健師	抑制	21.0	神奈川	64.4	山梨	41.9	18	62.2	45
肥満	危険	22.7	京都	24.4	福島	23.9	5	23.4	17
趣味	抑制	22.2	青森	30.7	奈良	22.2	1	27.0	28
交際	抑制	81	青森	185	高知	81	1	178	45
都市化	抑制	24.2	島根	98.4	東京	46.6	23	34.2	7

死亡率の高さを招いていることが推測される。そこで、青森県の根本的な死亡率低減対策には健康意識の向上が重要であると考えられ、そのための一つの方策として第一次産業からの産業構造の転換を提案する。同県が雪国県である点を考慮すると、IoT、ビッグデータ、AIなどICT(情報通信技術)産業の推進が有効と思われる。山形県なども同様の施策を進めているが、その見本になるのは自然条件が似ている北欧のフィンランドとエストニアであり、両国の新興IT企業は世界に進出している。青森県内の大学や高専の情報関係学部と奨学金制度の拡充強化やIT企業の誘致に成功すれば、教育水準の改善が図られ、これらを通じて健康意識の向上、ひいては死亡率の低下が期待できよう。

5. 結論

本研究では、地域の多死化対策に有用な情報を提供することを目的に、都道府県の全死因死亡率の男女平均値を目的変数とし、生活習慣、医療・福祉、社会・経済の3分野の指標56種を説明変数に用い、サポートベクター回帰により影響要因を探索した。その結果、社会福祉士、喫煙、高齢单身、がん検診、健康無関心、塩分、保健師、肥満、趣味、交際、都市化の11種が見出された。このうち、喫煙を中心とする健康関心関連要因、高齢单身を中心とする社会連携関連要因、およびこれまで未検証の社会福祉士要因が死亡率に大きな影響を与えることを明らかにした。これらの結果から、非線形重回帰分析における変数選択としての感度分析法により影響要因を探索する解析手法の有効性を実証した。また、地域の多死化対策への試みとして、死亡率が最高の青森県について死亡率低減策を検討した。

今後の課題として、本研究で用いた手法は生態学的研究であるため、生態学的誤謬(Ecological Fallacy)の問題がある。すなわち、都道府県別の死亡率の解析から得られた要因は個人の死因に関連付けられるものではなく、都道府県の死亡率差を説明するものにすぎない。地域の多死化対策により有用な情報を得るためには、時系列データや個人単位のマイクロデータ等の各種データを利用した総合的な解析を行う必要がある。

また、本研究で見出された11種の影響要因はそれらが単独に死亡率に影響するだけでなく、喫煙を中心とする健康関心関連要因、および高齢单身を中心とする社会連携関連要因のように、幾つかの要因が統合して死亡率に影響している可能性があることが推測された。このような階層構造を介した健康指標への影響分析に関しては、共分散構造分析手法やマルチレベルモデルを用いたアプローチがあり、これらの手法を用いた死亡率の要因分析は今後の課題である。

以上、本研究では都道府県単位の死亡率について多種多様な説明変数の中から影響要因を探

索した結果、都道府県の死亡率の影響要因を見出すことができ、地域の多死化対策やコホート研究等に役立つ参考情報が得られた。しかし、地域の多死化対策にさらに密接した情報を取得するためには、都道府県内の市区町村単位の死亡率について同様の解析を行うことで、より地域の実態に即した情報が得られることを付記しておく。

注.

- 1) 筆者らがサポートベクター回帰における変数選択のための感度分析法を開発、検証した既報(Tanabe et al., 2013)においては、使用データの特性の点において本研究とは異なっており、このことが本研究における感度分析法の検証の理由である。すなわち、既報においては化学物質の発がん性予測システムを開発するために、1500種以上の化学物質発がん性データを目的変数としてサポートベクター回帰による影響要因の探索を試みた。しかし、現時点では化学物質の発がん性に関しては原理や機構が未解明であるため、入手可能な1500種以上の説明変数の中から感度分析による変数選択法を用いて影響要因を探索した。その際、採用した説明変数の大半は連続変数ではなく離散変数であり、しかも、その数値0, 1, 2, ... に対して該当化学物質の頻度がたとえば1000, 100, 10, 1, 0, 0, ... のように極端に非対称な分布形状をしている変数が多かった。これに対して、本研究で検証した都道府県別の全死因死亡率に対する説明変数はすべて連続変数であり、多少の非対称分布を示す変数もあるが、大半の変数は対称性の高い分布を示している。筆者らとしてはこのようなデータに対して感度分析法による変数選択の有効性を検証したいと考えたのが本研究における感度分析法の検証の理由である。

参 考 文 献

- 阿部重夫 (2011). 『パターン認識のためのサポートベクトルマシン入門』, 森北出版, 東京.
- Adler, N. E. and Newman, K. (2002). Socioeconomic disparities in health: Pathways and policies, *Health Affairs*, **21**(3), 60-76.
- 赤梅昭太郎 (2008). 『カーネル多変量解析非線形データ解析の新しい展開』, 岩波書店, 東京.
- 安梅勅江, 篠原亮次, 杉澤悠圭, 伊藤澄雄 (2006). 高齢者の社会関連性と生命予後: 社会関連性指標と7年間の死亡率の関係, *日本公衆衛生雑誌*, **53**(9), 681-687.
- Araki, S. and Murata, K. (1986). Social life factors affecting the mortality of total Japanese population, *Social Science and Medicine*, **23**, 1163-1169.
- Chang, C.-C. and Lin, C.-J. (2011). LIBSVM-a library for support vector machines, <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>.
- 賃金構造基本統計調査 (2010). <https://www.e-stat.go.jp/stat-search/files?page=1&toukei=00450091&tstat=000001011429>.
- 衛生行政報告 (2014). <https://www.mhlw.go.jp/toukei/saikin/hw/eisei/16/>.
- 藤本健太郎, 濱野強, 藤澤由和 (2011). 社会的孤立と健康, *エストレーラ*, (205), 9-14.
- 福田吉治, 今井博久 (2007). 日本における「健康格差」研究の現状, *保健医療科学*, **56**(2), 56-62.
- Fukuda, Y., Nakamura, K. and Takano, T. (2004). Wide range of socioeconomic factors associated with mortality among cities in Japan, *Health Promotion International*, **19**, 177-187.
- 橋本英樹 (2012). 健康格差の実証研究—方法論的課題と展望, *医療と社会*, **22**(1), 5-17.
- 林美美, 横山徹爾, 吉池信男 (2009). 都道府県別にみた健康・栄養関連指標の状況と総死亡および疾患別死亡率, *日本公衆衛生雑誌*, **56**, 633-644.
- 五十嵐美幸, 佐藤一樹, 清水恵, 菅野雄介, 菅野喜久子, 川原礼子, 宮下光令 (2014). がん死亡および全死

- 因の都道府県別自宅死亡割合と医療社会的指標の地域相関分析, *Palliative Care Research*, **9**(2), 114-21.
- 人口動態調査特殊報告 (2017). <https://www.e-stat.go.jp/stat-search/files?page=1&toukei=00450013&tstat=000001102115>.
- 人口動態統計 (2016). <https://www.mhlw.go.jp/toukei/saikin/hw/jinkou/kakutei16/index.html>.
- Kagamimori, S., Gaina, A. and Nasermoaddeli, A. (2009). Socioeconomic status and health in the Japanese population, *Social Science and Medicine*, **68**, 2152-2160.
- 家計調査 (2013). https://www.e-stat.go.jp/stat-search/files?page=1&layout=datalist&toukei=00200561&tstat=000000330001&cycle=7&year=20170&month=0&tclass1=000000330001&tclass2=000000330019&tclass3=000000330020&result_back=1.
- 片山佳代子, 助友裕子, 黒沢美智子, 横山和仁, 岡本直幸, 稲葉裕 (2012). 都道府県別乳がん死亡率と教育系ファシリティとの関連: ソーシャル・キャピタルの視点から, 厚生指標, **59**, 26-34.
- 国民健康栄養調査 (2012). https://www.e-stat.go.jp/stat-search/files?page=1&layout=datalist&toukei=00450171&kikan=00450&tstat=000001041744&cycle=7&tclass1=000001111535&survey=%E5%81%A5%E5%BA%B7&result_page=1&second2=1.
- 国民生活基礎調査 (2013). <https://www.e-stat.go.jp/stat-search/files?page=1&layout=datalist&toukei=00450061&tstat=000001031016&cycle=7&tclass1=000001064999&tclass2=000001065383&tclass3=000001065392&second2=1>.
- 近藤克則 (2004). 人間関係と健康, 公衆衛生, **68**(3), 224-228.
- 近藤克則 (2014). 健康の社会的決定要因と医療経済・政策学, 医療経済研究, **26**(2), 81-99.
- 熊谷幸恵, 森岡郁晴, 吉益光一, 富田容枝, 宮井信行, 宮下和久 (2008). 主観的な精神健康度と身体健康度, 社会生活満足度および生きがい度との関連性: 性およびライフステージによる検討, 日本衛生学雑誌, **63**, 636-641.
- 蔵満美奈, 木村宣哉, 藤田直人, 河原田まり子 (2014). 健康寿命の延伸と地域保健活動との関連—北海道市町村を対象にした生態学的的研究—, 日本公衆衛生看護学会誌, **2**(1), 20-28.
- 森本兼襄, 丸山総一郎 (2001). ライフスタイルと心身の健康, 心身医学, **41**, 241-251.
- 内閣府 (2018). 平成 29 年版高齢社会白書, https://www8.cao.go.jp/kourei/whitepaper/w-2017/zenbun/29pdf_index.html.
- 日本学術会議パブリックヘルス科学分科会 (2011). わが国の健康の社会格差の現状理解とその改善に向けて, <http://www.scj.go.jp/ja/info/kohyo/pdf/kohyo-21-t133-7.pdf>.
- 日本公衆衛生学会公衆衛生モニタリング・レポート委員会 (2011). 高齢者における社会経済的要因と健康, 日本公衆衛生雑誌, **58**(7), 564-568.
- 尾形聡, 西慎一, 若井建治, 井関邦敏, 椿原美治 (2011). 慢性透析患者の地域差の要因, 日本透析医学会雑誌, **44**, 681-688.
- 小野田崇 (2007). 『サポートベクターマシン』, オーム社, 東京.
- 大北剛 (訳) (2005). 『サポートベクターマシン入門』, 共立出版, 東京.
- 大坪浩一, 山岡和枝, 横山徹爾, 高橋邦彦, 丹後俊郎 (2009). 標準化死亡比の経験的ベイズ推定量に基づく医療資源と死亡との関連: 全国の市区町村を対象として, 日本公衆衛生雑誌, **56**(2), 101-110.
- Rothman, K. J. (2012). *Epidemiology — An Introduction*, Oxford University Press, Oxford.
- 社会福祉振興・試験センター (2015). 社会福祉士・介護福祉士・精神保健福祉士の都道府県別登録者数, http://www.sssc.or.jp/touroku/pdf/pdf_t04_3.pdf.
- 社会・人口統計体系 (2014). <https://www.e-stat.go.jp/regional-statistics/ssdsview>.
- 社会生活基本調査 (2014). <https://www.e-stat.go.jp/stat-search/files?page=1&toukei=00200533&tstat=000001095335>.
- 社会生活統計指標—都道府県の指標 (2013). <https://www.e-stat.go.jp/stat-search/files?page=1&layout=datalist&toukei=00200502&tstat=000001095336&cycle=0&tclass1=000001095337>.
- 田辺和俊, 鈴木孝弘 (2018). 都道府県の相対的貧困率の計測と要因分析, 日本労働研究雑誌, **60**(2-3),

45-58.

- Tanabe, K., Lučić, B., Amić, D., Kurita, T., Kaihara, M., Onodera, N. and Suzuki, T. (2010). Prediction of carcinogenicity for diverse chemicals based on substructure grouping and SVM modeling, *Molecular Diversity*, **14**(4), 789-802.
- Tanabe, K., Kurita, T., Nishida, K., Lučić, B., Amić, D. and Suzuki, T. (2013). Improvement of carcinogenicity prediction performances based on sensitivity analysis in variable selection of SVM models, *SAR and QSAR in Environmental Research*, **24**(7), 565-580.
- 田辺和俊, 鈴木孝弘, 中川晋一 (2016). サポートベクター回帰による都道府県別肺がん死亡率の関連要因に関する検討, *保健医療科学*, **65**(6), 598-610.
- 特定健康診査 (2010). <https://www.mhlw.go.jp/bunya/shakaihoshoh/iryouseido01/info02a-2.html>.
- World Health Organization (2016). The determinants of health, <http://www.who.int/hia/evidence/doh/en/>.
- 吉本好延, 三木章江, 浜岡克伺, 大山幸綱, 河野淑子, 佐藤厚 (2011). 救急搬送における高齢者の転倒の標準化発生比と社会経済状態の関連, *日本公衆衛生雑誌*, **58**, 183-189.

Verification of the Effectiveness of Sensitivity Analysis as a Variable
Selection in Support Vector Regression
—Analysis of Factors Affecting Prefectural All-cause Mortality Rates—

Kazutoshi Tanabe¹ and Takahiro Suzuki²

¹Institute of Social Sciences, Toyo University

²Department of Economics, Toyo University

The sensitivity analysis method as a variable selection in support vector regression has been applied to the search of factors affecting prefectural all-cause mortality rates, and its effectiveness has been verified. In Japan, a multi-death society will come after an aging society with a declining birthrate, and various social problems such as shortage of doctors, end-of-life care, lonely death, death place refugees, vacant houses are feared. In order for local governments to take countermeasures against mortality, it is important to clarify the factors that have a significant effect on mortality among many factors and estimate their relative impact. All-cause age-adjusted mortality rates of 47 prefectures were used as an objective variable, and 56 indices in lifestyle habit, medical care/welfare, and society/economy fields were employed as potential explanatory variables. Factors related to the mortality rates were searched by applying a support vector regression technique to these data, and their sensitivities to the mortality rates were estimated by employing a sensitivity analysis method. Eleven kinds of factors which reproduce the observed mortalities of 47 prefectures with an accuracy of statistical significance level were obtained. It is found that rates of social workers not examined in previous studies as well as those of smoking habit and of elderly singles highly contribute to the observed mortality rates. Countermeasures for decreasing the mortality rate of Aomori prefecture showing the highest rate in Japan were proposed on the basis of the affecting factors obtained in this study. From these results, the effectiveness of the sensitivity analysis method as a variable selection in support vector regression has been demonstrated.