

テンソル解析を用いた死因別将来死亡率の同時推定

鈴木 孝太郎* 松山 直樹†

2020年1月6日投稿

2020年1月22日受理

概要

本研究では死因別死亡率の同時推定手法を提案する。一般に死因別死亡率推定は死因間の相互作用を反映出来ず、合算した全死因死亡率の将来推定が意味を持たないことが知られている。このため死亡率解析で標準的な Lee-Carter モデルを準用し、年齢・時間・死因・性別の4軸をもつ4階テンソル形式のデータに対して、死亡率の改善が軸によらず単一の医療水準の進歩でドライブされるとみなした単一時系列モデルを提案する。この手法により、死因間の相関関係を明らかにすると共に、全死因死亡率と整合的な死因別死亡率の外挿を実現した。

キーワード：Lee-Carter モデル、死因別死亡率、テンソル解析、WHO 死亡率データベース

1 研究の背景

1.1 経済価値評価の要請

近年保険会社は経済価値ベースでの保険負債評価を推し進めている。その背景には各国の監督当局が担任する新しい資本規制の存在があり、EUにおけるソルベンシー II や IAIS の公表した ICS ver2.0 などが代表的な規制である。ここでいう負債の経済価値とは規制によって細部の定義が異なるが、

1. 負債が将来生み出す期待キャッシュフローの割引現在価値
2. その期待値を超えるリスクマージン

の2つを基本として計算される概念という点では概ね一致する。これらの規制により保険会社の財務状況を監督官庁が適切に把握し、保険会社のリスク管理の高度化をより一層推進させるという狙いがある。このような国際的な動向を受け、日本でも経済価値ベースでのソルベンシー規制の導入の動きがある。

現行の保険会計では、死亡率や金利を定数とみなすことで伝統的な保険数理計算の手法に則って、満期までの保険負債を契約時点で評価することが出来た。一方で経済価値ベースで評価する場合、保険負債の時価評価、すなわち評価時点での死亡率や金利のリアルタイムな情報が必要になり、死亡率や金利の将来にわたる超長期の時間変動を適切にモデリングすることが現行会計と比較してより重要視されるようになる。

ここでいう「適切なモデリング」に明確な定義はないが、超長期の外挿を前提としたモデリングでは予測結果が妥当なものであるか、定量的な統計的モデル評価指標のみならず、学際的議論であったり、定性的な経験

* 明治大学大学院 先端数理科学研究科 E-mail: cs181011(at)meiji.ac.jp

† 明治大学 総合数理学部 E-mail: ma2yama(at)meiji.ac.jp

による議論も必要不可欠である。そのため複雑な統計モデル（もしくは機械学習のようなアルゴリズム的手法）による高い精度の予測よりも解釈可能性・可読性・納得性といった要素も加味されたモデルの方が望ましい。

1.2 リレーショナルモデルの既存研究

本論文では将来死亡率予測をテーマに扱うが、将来死亡率予測は上記の経済価値ベースでの保険負債評価のために限らず、保険のプライシングや死亡リスクの管理、年金制度運営、人口推計など、将来の死亡によって変動するキャッシュフローの制御や意思決定をする上で非常に重要であり、そのためかねてより多くのモデルが提案されてきた。その中でも代表的な統計モデルの1つとして Lee-Carter モデルが知られている。Lee-Carter モデルは年齢 x 、暦年 t の死亡率 $m_{x,t}$ を

$$\log m_{x,t} = \alpha_x + \beta_x \kappa_t$$

と表現するモデルの総称であり、いくつかのバリエーションが知られている。原論文である Lee and Carter[1992] で提案されたモデルは、誤差項 $\varepsilon_{x,t}$ を i.i.d. として

$$\begin{aligned} \log m_{x,t} &= \alpha_x + \beta_x \kappa_t + \varepsilon_{x,t} \\ \varepsilon_{x,t} &\sim N(0, \sigma^2) \end{aligned}$$

という正規ノイズを伴った双線形回帰型のモデルである。しかし、分散が年齢に依らず一定であり分散不均一性 (heteroscedasticity) を表現できないことから Brouhns et al.[2002] では

$$\begin{aligned} D_{x,t} &\sim \text{Po}(E_{x,t} m_{x,t}) \\ \log m_{x,t} &= \alpha_x + \beta_x \kappa_t \end{aligned}$$

という双線形ポアソン回帰型のモデルを提案しており、こちらも Lee-Carter モデルの代表的な形として知られている。ここで $D_{x,t}$ 、 $E_{x,t}$ はそれぞれ年齢 x 、暦年 t における死亡者数、リスク対応延べ年数 (Exposure) である。

Lee-Carter モデルの構造はパラメータの解釈可能性と推定の容易性に優れていることから、原論文によって提案されてから現在に至るまで多くの実証と拡張がなされ、実務においても利用されている。ここで、各パラメータは α_x が「対数死亡率の平均的な水準」、 κ_t が「死亡率の改善度」、そして β_x が「死亡率改善 κ_t の年齢別感応度」と解釈されることが多い。特に、 κ_t は通常ノイズを伴いつつも単調に下がっていくことから、医療水準の向上を表すと解釈される場合が多い。

しかしながら、Lee-Carter モデルは可読性が高い反面、予測精度が低いことが指摘されている。また信頼区間を過小推定することも報告されており、顕在化されたモデルリスクの1つとして知られている (例えば、Lee[2000], Cairns[2000] を参照)。

この2つの問題点に対してそれぞれ解決方法が提案されている。

まず予測精度を高める方法として、時間依存項を増やすモデルが Renshaw and Haberman[2003] により提案されている。このモデルは

$$\log m_{x,t} = \alpha_x + \sum_{k=1}^K \beta_x^{(k)} \kappa_t^{(k)}$$

と表現されるモデルであり、 $K = 1$ のときは通常の Lee-Carter モデルに一致する。これは Lee-Carter モデルが行列の特異値分解を用いた低ランク近似をベースとしたモデルであることの類推であり、第 K 特異ベクトルまでの情報を用いて行列の近似を行うことに対応している。

また、信頼区間の過小推定を解決する方法に関しては、過分散モデル (Lee-Carter モデルの平均構造を保ちつつ分散のみ大きくするモデル) が提案されている。これは大きく分けて2つのモデルが提案されており、1

つは Renshaw and Haberman[2006] により提案されたモデルで、上述のポアソン型 Lee-Carter モデルに対して $V[D_{x,t}] = \phi E_{x,t}$ という構造を付加し、 ϕ によって分散の大きさをコントロールするモデルである。もう 1 つは Li et al.[2009] により提案されたモデルで、これもポアソン型 Lee-Carter モデルに対する拡張ではあるが、この論文では年齢・時間以外の要素で死亡率が異なるとし、

$$\begin{aligned} D_{x,t} | w_x &\sim \text{Po}(w_x E_{x,t} m_{x,t}) \\ w_x &\sim \Gamma(\theta_x^{-1}, \theta_x^{-1}) \\ \log m_{x,t} &= \alpha_x + \beta_x \kappa_t \end{aligned}$$

という母数混合型モデルを提案した。ここで w_x は平均 1、分散 θ_x となるように選ばれており、これによって Lee-Carter モデルの平均構造を保っている。また、 $D_{x,t}$ は w_x を積分消去することで負の二項分布になることが知られており、このことから負の二項型の Lee-Carter モデルと呼ばれることもある。

これらのモデル、もしくはその組み合わせによって上述の予測精度と信頼区間の問題は解決されるが、新たな問題点がそれぞれ指摘できる。まず時間依存項を増やすモデルにおいては、追加の時間依存項が (Lee-Carter モデルで推定される κ_t が概ね線型であることと比較して) 非常に非線型であり統計的に予測が難しく、また実務面においてもその解釈が難しいという欠点がある。また過分散モデルにおいては、平均構造が変わっていないため中短期の予測に関して改善が行われれないという点が指摘できる。

これらの問題から、上述の「適切なモデリング」という点では拡張された Lee-Carter モデルを保険実務に直接適用して良いかは議論の必要がある。

1.3 死因別モデリングの必要性和困難性

一方で、単なる死亡率の予測を行うだけでなく、死因別死亡率の予測を行うことも保険商品の開発やプライシング、リスク管理であったり、政策の意思決定などにおいて有用である。従来のそういった死因別死亡率の分析については、前述の Lee-Carter モデルを死因別死亡者数のデータに当てはめ、パラメータの解釈を行うといった方法が知られている。すなわち、前述の死亡率に死因分類 c を含めた $m_{x,t,c}$ に対して

$$\log m_{x,t,c} = \alpha_{x,c} + \beta_{x,c} \kappa_{t,c}$$

というモデルを当てはめていることになる。しかし、このモデル (死因別 Lee-Carter モデルと呼ぶ) は

1. 死因によっては Lee-Carter モデルの当てはまりが悪いこと。
2. パラメータ総数の多さにより推定が不安定になってしまうこと。
3. データ数が極端に少ないデータでは推定が不安定になってしまうこと。
4. 個別に推定することにより死因間の相互作用を反映出来ず、
合算した全死因死亡率の将来推定が意味を持たないこと。

の 4 点が欠点として挙げられる。特に、死因別に見たときに Lee-Carter モデルが当てはまらない問題は Girosi and King[2007] でも指摘されており、著者らは論文中で死因別に Lee-Carter モデルを当てはめることに対して、「日常的に用いるべきではない」と表現している。

これらの問題の中でも、特にモデルの当てはまりを改善する方法として、上述の時間依存項を増やすモデルを死因別データに援用するという方法が挙げられる。すなわち

$$\log m_{x,t,c} = \alpha_{x,c} + \sum_{k=1}^K \beta_{x,c}^{(k)} \kappa_{t,c}^{(k)}$$

というモデルである。これを前述の死因別データに当てはめれば当てはまりの問題は解決される。しかし、このモデルに対しても

1. K をどこまで大きく取れば良いかの判断が難しいこと。
2. Lee-Carter モデルより更にパラメータ数が多くなってしまうこと。
3. 推定された複数の $\kappa_{t,c}^{(k)}$ ($k = 1, \dots, K$) の解釈が困難であること。

といった欠点が指摘できる。

ところで、死因別死亡率推定は過去に日本の公的な人口推定にも用いられていたという過去がある（石井 [2017] 参照）。当時は統計モデル的な手法ではなく Gompertz モデルのような数学的関数のフィッティングによる予測が行われており、古くから死因別死亡率の予測が全死因死亡率の予測に有用であると認識されていたと言える。現在は死因別死亡率に基づく人口推定は行われていないが、これには下記のような理由があげられている。

1. パラメータの多さとデータの少なさにより推定精度が不安定であること。
2. Lee-Carter モデルに代表される統計モデルが台頭してきたこと。
3. 死因別死亡率推定は死亡率改善を過小推定すること。

3 つめに関して言及すると、これは Wilmoth [1995] により報告された内容であり、いくつかの仮定のもとで将来予測値と観測値の間に正のバイアスが生じることを理論的に示したものである。

2 本研究の目的

前章を踏まえ、本研究の目的を 2 つ提示する。1 つは、経済価値評価のために頑健な死亡率の長期外挿を可能とすることである。もう 1 つは死因別死亡率を全死因死亡率と整合的に推定することである。最初の目的から時系列成分を 1 つに制限し Lee-Carter モデルの修正を行うことが要請され、2 つ目の目的からテンソル型モデルであることが要請される。本論文で提案するモデルは、Li et al. [2009] で言及されている「年齢・暦年以外の死亡率決定要因」として死因を選択したものと考えることも出来るため、先行研究の結果と整合的である。

またテンソル型モデルは、本論文では具体的に年齢・時間・死因・性別の 4 つの軸をもつ 4 階テンソル形式のデータに対するモデルである。本研究では κ_t を医療水準を表すパラメータと解釈し、死亡率の改善は長期的に見ると性別や死因、年齢を越えて医療水準の向上にドライブされると仮定し、性別も含めたテンソルデータに対し単一時系列のモデルを当てはめた。ただしここで「単一時系列」とは死亡率変化のファクターが 1 つという意味で、死因別や年齢別の死亡率トレンドが同一であるというわけではないことに注意されたい。

なお、テンソル形式のデータを死亡率推定に用いた先行研究としては野村 [2018] が挙げられる。野村 [2018] は小地域 i ごとの死亡率 $m_{x,t,i}$ に対して

$$\log m_{x,t,i} = \alpha_x + \sum_{k=1}^K \beta_x^{(k)} \kappa_t^{(k)} \gamma_i^{(k)}$$

を当てはめた。これはテンソルの高次特異値分解（テンソル分解）と呼ばれる手法を用いている。テンソル分解には CP 分解や Tucker 分解が基本的な手法として知られており、野村 [2018] はその中でも CP 分解による方法を提案しているが、本研究では単一時系列による推定及びパラメータ数の節約による推定の安定を目標としているため、テンソル分解よりもシンプルなモデルを提案する。

また本論文では上記の Wilmoth [1995] の示した内容が現在のデータ・推定手法に対しても当てはまるかどうか確認し、言及する。

3 提案モデル

3.1 モデルの定式化

本研究では年齢 x , 暦年 t , 死因 c , 性別 s の死亡率 $m_{x,t,c,s}$ に対して下記のモデルを提案する。

$$\log(1 - m_{x,t,c,s}) = \beta_{x,c,s}\kappa_t + \varepsilon_{x,t,c,s} \quad (1)$$

ここでパラメータは $\beta_{x,c,s}$, κ_t であり最尤法により推定する。また $\varepsilon_{x,t,c,s}$ はこのモデルの誤差項であるが、本研究では

1. 年齢・死因・性別ごとの相関構造を独立に記述でき、特に死因間の相関構造を推定できること
2. ニュートン法に代表される勾配法を用いないため安定した推定が出来ること

という2点から、誤差分布にテンソル正規分布 TN を仮定した（テンソル正規分布については付録 B を参照）。モデルをテンソル表記すると下記の通りであり、こちらの表記の方が誤差の構造が捉えやすいという特徴がある。

$$\begin{aligned} \mathcal{M}_t &= \mathcal{B}\kappa_t + \mathcal{E}_t \\ \mathcal{E}_t &\sim \text{TN}(0, \Sigma_{\text{age}}, \Sigma_{\text{CoD}}, \Sigma_{\text{sex}}) \end{aligned}$$

ただし、 $\mathcal{M} = (\log(1 - m_{x,t,c,s}))$, $\mathcal{B} = (\beta_{x,c,s})$, $\mathcal{E}_t = (\varepsilon_{x,t,c,s})$ であり、 Σ_{age} , Σ_{CoD} , Σ_{sex} はそれぞれ年齢 (age)・死因 (Cause of Death)・性別 (sex) ごとの分散共分散行列である。

本モデルは外挿性能を重視する観点から Lee-Carter モデルの平均水準 α_x にあたる項をもたないが、理由については第 6 節で後述する。また、通常の Lee-Carter モデルやその拡張モデルが死亡率の対数値に対するモデルであるのに対して、提案モデルはその死因の生存率に対するモデルとなっている。これは、死因別の観測死亡者数が 0 のときデータから直接計算される相対数死亡率が $-\infty$ になってしまうためである。

3.2 推定方法

提案モデルは式の形から β と κ によって完全に変数が分離されている。従って「年齢・死因・性別」と「暦年」の2つの軸になるようにテンソルの行列化を行うことで、原論文の Lee-Carter モデルと同じ特異値分解を用いた推定が可能である（テンソル計算に関する記号は付録 A を参照）。すなわちデータテンソル \mathcal{M} を標準的な特異値分解によって

$$\mathcal{M}_{(2)} = USV^T$$

と分解し

$$\begin{aligned} \kappa'_t &= U[t, 1] \\ \beta'_q &= S[1, 1]V[q, 1] \end{aligned}$$

とした後にモデルの識別性の問題から $\beta_0 = \sum_q \beta'_q$ として

$$\begin{aligned} \kappa_t &= \beta_0 \kappa'_t \\ \beta_q &= \frac{\beta'_q}{\beta_0} \end{aligned}$$

と修正し、ベクトル β_q を再度テンソル化すれば良い。ここで q はテンソル $\beta_{x,c,s}$ をベクトル化した後の添え字であり、

$$q = x + (c - 1)N_{\text{age}} + (s - 1)N_{\text{age}}N_{\text{CoD}}$$

とすればいい。ただし N_{age} , N_{CoD} はそれぞれ年齢・死因の各軸（モード）の次元である。

また、分散共分散行列は特定の要素に注目しその他の分散共分散行列は既知とすると解析的に最尤推定量が求まるため、データから平均構造を引いた残差テンソル \mathcal{R}_t に対して次のような交互最適化で求めた（導出は付録 B を参照）。

$$\Sigma_i = \frac{N_i}{NT} \sum_{t=1}^T \mathcal{R}_{t,(i)} \Sigma_{-i}^{-1} \mathcal{R}_{t,(i)}^T \quad (i = \text{age, CoD, sex})$$

ここで $\mathcal{R}_{t,(i)}$ は軸 i に対するテンソル \mathcal{R}_t の行列化であり、 Σ_{-i}^{-1} はそれぞれ

$$\Sigma_{-\text{age}}^{-1} = \Sigma_{\text{sex}}^{-1} \otimes \Sigma_{\text{CoD}}^{-1}, \quad \Sigma_{-\text{CoD}}^{-1} = \Sigma_{\text{age}}^{-1} \otimes \Sigma_{\text{sex}}^{-1}, \quad \Sigma_{-\text{sex}}^{-1} = \Sigma_{\text{CoD}}^{-1} \otimes \Sigma_{\text{age}}^{-1}$$

であり、 \otimes は行列のクロネッカー積である。ただし、今回はコーホート効果などの年齢間の相関や性別間の相関は考えず、 Σ_{age} , Σ_{sex} は対角行列とし、 Σ_{CoD} は一般の分散共分散行列とした。

4 推定結果

4.1 使用データ

本研究では WHO の提供する WHO mortality database からデータを用意した。暦年は 1995 年から 2015 年まで、年齢は 0 歳、1 歳から 4 歳、5 歳から 9 歳、（以降 5 歳区分）、80 歳から 84 歳、85 歳以上と不均一にまとめられている。死因に関しては元のデータでは国際的な死因分類である ICD10 に従って分類されているが、各項目が非常に詳細であるため、厚生労働省が提示している ICD10 準拠基本分類表の 22 区分にデータを圧縮した上で分析を行った。その区分は別表の通りである。なお、特殊目的用のコードとその区分があるため実際の死因は 19 分類まで圧縮された。その結果、今回使用するデータでは次元が 19（年齢） \times 21（暦年） \times 19（死因） \times 2（性別）の 4 階テンソルとなった。

4.2 解析結果

提案モデルにより推定した結果の一例を図 1-4 に表す。ここで横軸は上述の年齢区分、縦軸は対数死亡率 $\log m_{x,t,c,s}$ である。モデル $\log(1 - m_{x,t,c,s})$ と出力が異なることに注意されたい。

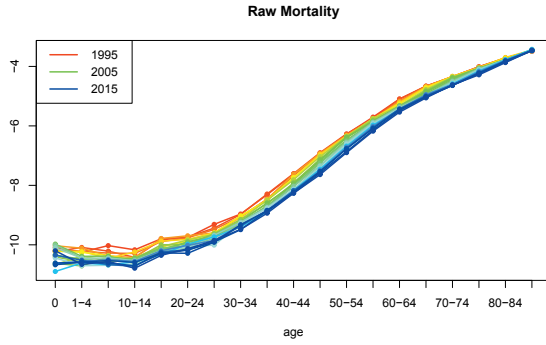


図1 男性：新生物（観測値）

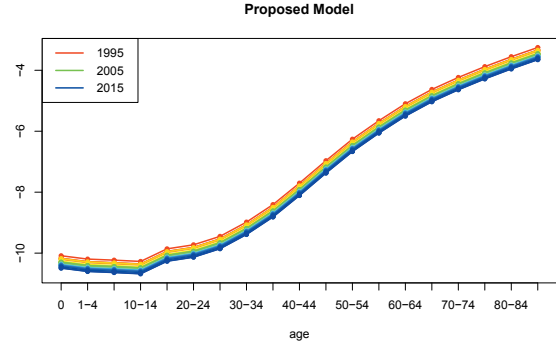


図2 男性：新生物（推定値）

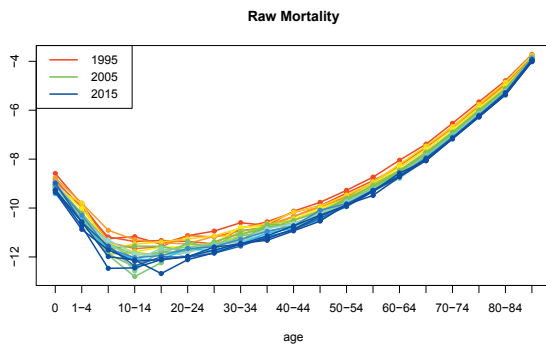


図3 女性：呼吸器系の疾患（観測値）

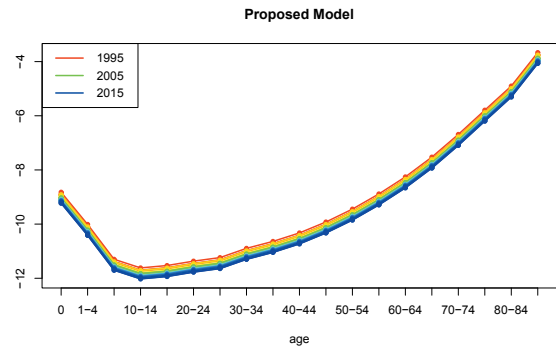


図4 女性：呼吸器系の疾患（推定値）

この図から、提案モデルが死因別死亡率の特徴を良く表現できていることがわかる。実際、推定された平均構造とデータの乖離をフロベニウスノルムで測ると、norm percent は 91.06% であり、非常に高い精度でデータを表現できていることが分かった。ここで norm percent は観測値 \mathcal{X}_{obs} と推定値 \mathcal{X}_{est} に対して以下で定義される。

$$\text{norm percent} = \left(1 - \frac{|\mathcal{X}_{est} - \mathcal{X}_{obs}|_F}{|\mathcal{X}_{obs}|_F} \right) \times 100 \quad (2)$$

次に分散構造の推定結果を図 5,6 に示す。

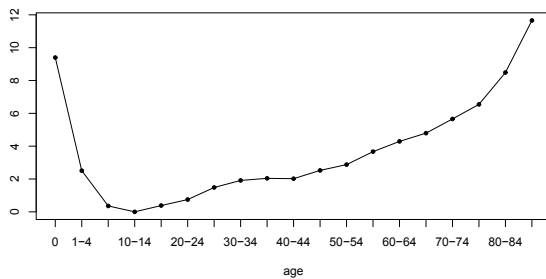


図5 年齢別標準偏差（対数スケール）

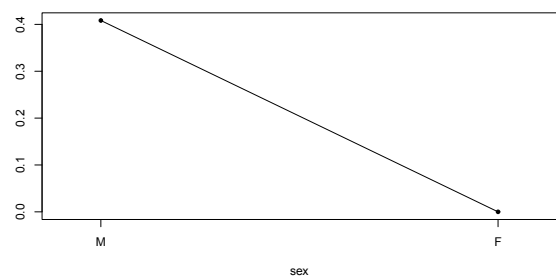


図6 男女別標準偏差（対数スケール）

ここで、 $\Sigma_{age}, \Sigma_{sex}$ は対角行列に制限しているため対角成分をプロットしている。またテンソル正規分布における分散共分散行列はスケール一意性を持たず値を直接評価できず相対評価しかできないため、どちらの図

も最小値を1に基準化した上で対数スケールでプロットしている。相対評価という観点でこの図を見ると、0歳と高年齢層の分散が大きいことがわかった。これは図1の0歳と高年齢層を見てわかるように、回帰誤差の大きさと観測データのブレの大ききの2つに原因があると考えられる。また男性の方が女性よりも分散が大きいこともわかった。

次に推定された時系列成分 κ_t を図7に表す。

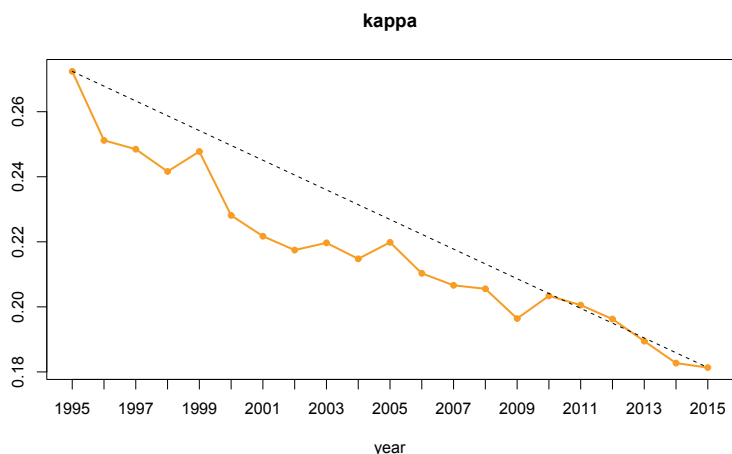


図7 推定された κ_t

端点同士を点線で繋いでいるが、この図から κ_t は通常 Lee-Carter モデルの文脈で仮定されるような直線的ではなく、逡減的な減少トレンドを持つことが示唆された。最後に提案モデルと死因別 Lee-Carter モデルの AIC を表1に示す。

表1 AIC の比較

モデル	対数尤度	パラメータ数	AIC
提案モデル	163577	954	-325246
死因別 LC	4884	2204	-5360

このように提案モデルは死因別 Lee-Carter モデルと比較してパラメータ数も削減しつつ AIC を大幅に削減することが出来た。

4.3 死因間の相関構造

ここではテンソル正規分布の推定により得られる死因間の相関構造について言及する。テンソル正規分布を用いずに死因間の相関を得ようとしたとき、得られた観測対数死亡率から平均構造を引いた残差テンソル $r_{x,t,c,s}$ を死因の軸に関して行列化し、各列が独立に得られた標本だとみなせば分散共分散行列を得ることが可能である。しかし、この方法は年齢間や性別間の分散の違いを無視しているという欠点があり、一方で年齢・性別ごとに分散共分散行列を計算すると、こちらは相対的なデータの少なさにより推定が不安定になるという欠点がある。テンソル正規分布はそれら2つの欠点を解消する方法として有用であり、本モデルに採用した理由の1つであることは前節で述べた通りである。

図8に Σ_{CoD} より得られる偏相関行列を表す。偏相関は注目する2つの死因以外の影響を取り除いた相関であり、単なる相関行列よりも真の相関概念に近いためこちらを用いた。また(偏)相関行列は分散共分散行列

の一種の規格化であるため、分散共分散行列の一意性の問題はなく、大小比較ではなく直接数値を評価することが可能である。

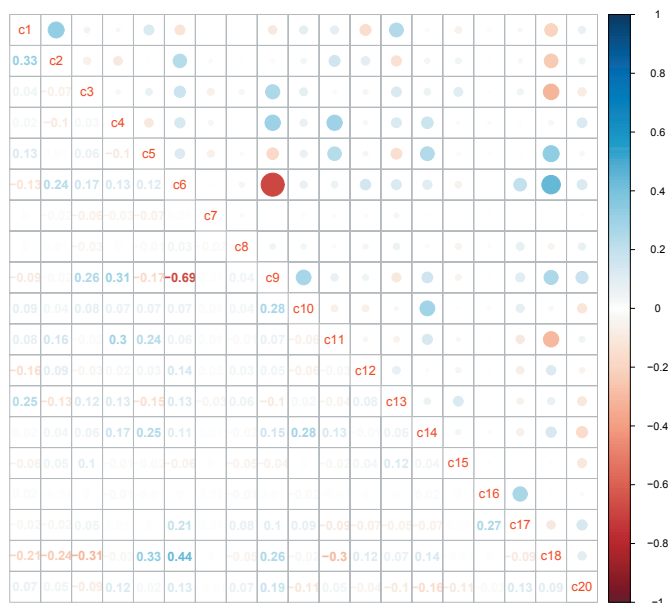


図8 死因間の偏相関プロット

図8によると、死因間には正の相関と負の相関が点在しているが、比較的高相関と言えるものは死因分類の6番目と9番目である「神経系の疾患」と「循環器系の疾患」の相関(-0.69)のみであった。「神経系の疾患」は髄膜炎や脳炎、パーキンソン病、アルツハイマー病、てんかん、筋ジストロフィー、脳性麻痺などを含む死因であり、「循環器系の疾患」はリウマチ熱や心不全、虚血性心疾患、脳血管疾患などを含む死因である。

このような結果が得られた原因を調べるために残差の分析を行う。推定値と観測値の残差をプロットした図が図9であり、確かに中心付近では右下がりの傾向が読み取れた。またこの残差の時間推移をプロットした図が図10である。2つの残差は2007年を境に上下が反転し、その後「神経系の疾患」の残差(c6)は負の方向に拡大し、「循環器系の疾患」の残差(c9)は正の方向に拡大していることがわかった。このような逆方向への動きが相関に寄与していると考えられるが、これらの死因間に負の相関が発生する定性的な理由・解釈を与えることはできなかった。この部分に関しては医学的視点からの研究が望まれる。ただし今回死因の分類に用いたICD10は、2007年に(ICDのバージョンを変えないマイナーチェンジの中で)大きな変更が加えられており、その影響によって2007年前後のトレンド変化が起きている可能性があるが、残差のトレンドの向きや大きさの理由・解釈を得ることはできなかった。

なおここでは年齢・性別の影響を排除するため、観測値から推定値を引いた残差テンソル \mathcal{R}_t に対してそれぞれの分散共分散行列の逆行列をかけた $\mathcal{R}_t^{\text{CoD}} = \mathcal{R}_t \times_1 \Sigma_{\text{age}}^{-1} \times_3 \Sigma_{\text{sex}}^{-1}$ という変換を行っている。図9はそうして得られた残差 $\mathcal{R}_t^{\text{CoD}}$ (から「神経系の疾患」と「循環器系の疾患」を抽出したもの)をそのままプロットしているが、図10では死因間の比較のためにどちらも平均0、分散1にスケールした上で、年齢軸と性別軸の平均を取った残差のプロットをしている。

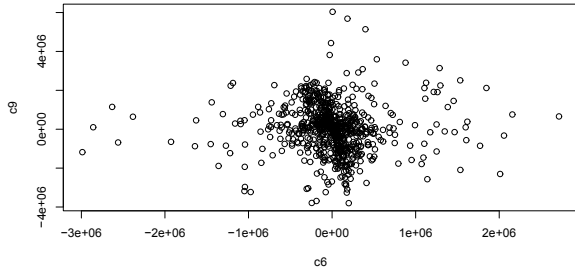


図9 残差プロット

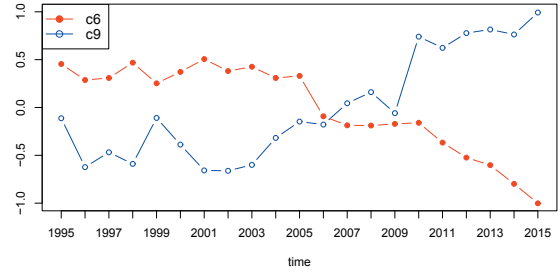


図10 平均残差時間推移プロット

5 将来推定への適用

5.1 κ_t の外挿

前節で、提案モデルが比較対象とした死因別 Lee-Carter モデルと比べパラメータ数を削減しながら高精度な推定が行えていることが分かった。この節では、推定した κ_t の外挿により死亡率の将来推定を行う。しかし、そのまま将来の外挿を行うと確率とは言えない値に発散してしまうため、制約条件を課す。死因別死亡率の満たす性質 $0 \leq m_{x,t,c,s} < 1$ より、誤差項を無視すると

$$\begin{aligned}
 0 \leq m_{x,t,c,s} < 1 &\iff 0 < 1 - m_{x,t,c,s} \leq 1 \\
 &\iff \log(1 - m_{x,t,c,s}) \leq 0 \\
 &\iff \beta_{x,c,s} \kappa_t \leq 0
 \end{aligned} \tag{3}$$

という式が得られる。パラメータ $(\beta_{x,c,s}, \kappa_t)$ 間には -1 倍の自由度，すなわち

$$\beta'_{x,c,s} = -\beta_{x,c,s}, \quad \kappa'_t = -\kappa_t$$

という変換で得られた $(\beta'_{x,c,s}, \kappa'_t)$ もモデルの形を変えないため，(3) 式は $\beta_{x,c,s} \leq 0, \kappa_t > 0$ としても一般性を損なわない。そこで本研究では得られた κ_t の対数を取り， $\log \kappa_t$ に対して ARIMA モデルを当てはめる。

AIC に基づいて ARIMA モデルのモデル選択を行ったところ次の ARIMA(0,1,1) が選択された。

$$\log \kappa_t = \log \kappa_{t-1} + d + \epsilon_t + \theta \epsilon_{t-1} \tag{4}$$

推定されたパラメータを表 3 に示す。

表 2 推定パラメータ

パラメータ	数値	標準誤差
d	-0.0188	0.0033
θ	-0.5220	0.2628
誤差分散	0.0008265	-

推定された κ_t の予測平均を図 8 に示す。

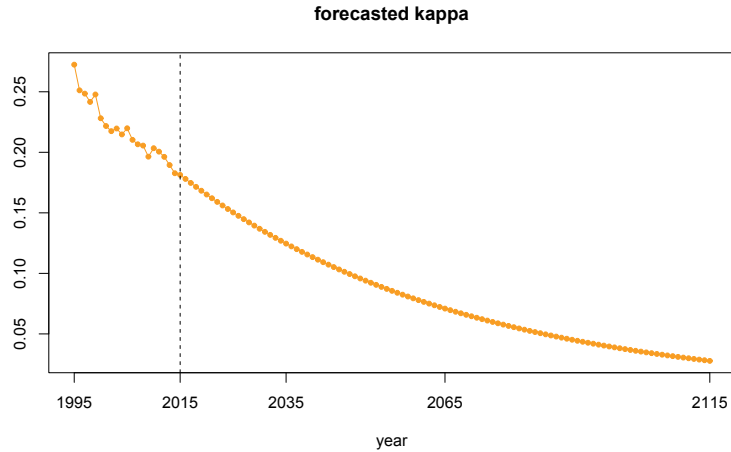


図 11 推定された κ_t とその外挿

ここで、2015 年までは (1) 式で定義される提案モデルによって推定したものであり図 7 と同じものである。そして 2015 年から 100 年後の 2115 年までが上記 (4) 式で推定したものであり、Lee and Carter[1992] と同様に二段階推定のプロセスを採っている。

5.2 外挿された κ_t を用いた将来死亡率の推定

次に外挿により将来推定された κ_t を用いた将来死亡率の推定値を図 12,13 に示す。ただし図 1-4 と同様、横軸は年齢区分、縦軸は対数死亡率 $\log m_{x,t,c,s}$ である。

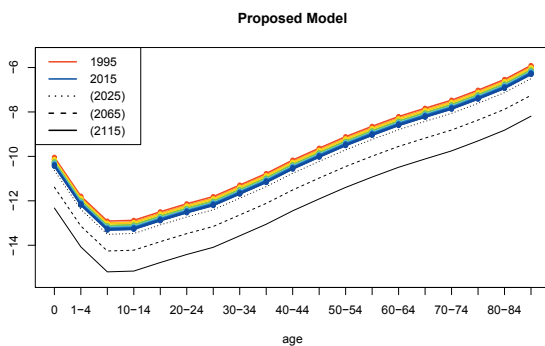


図 12 男性：内分泌，栄養及び代謝疾患

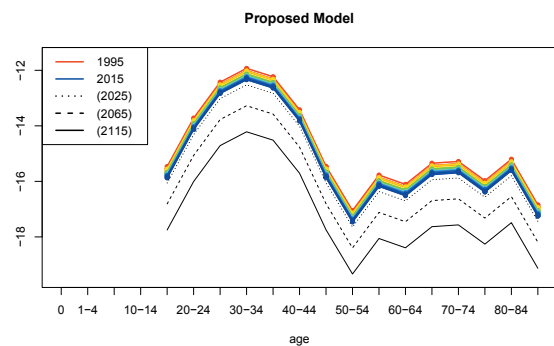


図 13 女性：妊娠，分娩及び産褥

このように、提案モデルは超長期の予測においてもグラフの概形を崩すことなく推定が行えることがわかった。ここで図 13 の若年齢層は、死亡率が 0 であり対数死亡率が計算できないため表示されておらず、観測死亡率の特徴を良く捉えていると言える。

5.3 死因別死亡率の過小推定に関して

最後に、冒頭で述べた通り Wilmoth[1995] の示した「死因別将来推定は死亡率改善を過小推定する」という内容に関して確認する。ここでは 2010 年までのデータを用いて、保持しているデータの最新年である 2015 年までの推定を行い、推定結果と観測値の比較を行う。ただし、複数の死因間での比較を行うため、2010 年の死

亡率を1に基準化した上で比較を行い、性別は男性のみとした。その結果が下記の図14-17である。またこの比較方法は石井 [2017] に即した方法である。

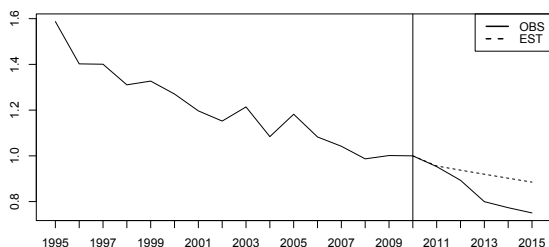


図14 感染症及び寄生虫症

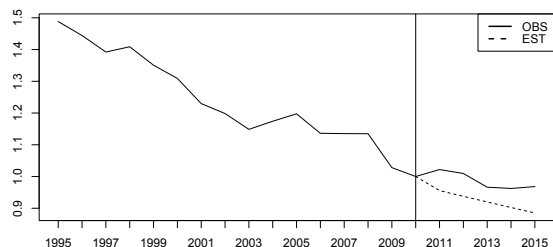


図15 新生物<腫瘍>

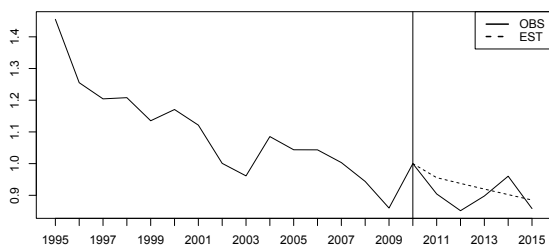


図16 先天奇形，変形及び染色体異常

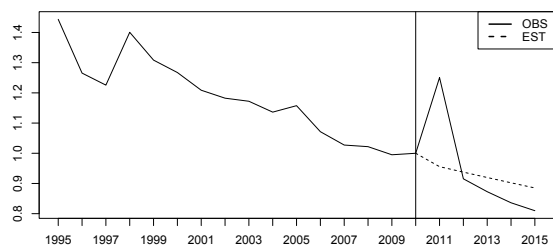


図17 傷病及び死亡の外因

Wilmoth[1995]の主張通りであれば、どの死因に関しても推定結果よりも観測値の方が小さくなる（すなわち実線が下に位置する）はずであるが、そのような結果は得られなかった。

6 時間平均項の取扱いに関する検討

本研究は Lee-Carter モデルのアイデアをテンソルデータに応用したモデルを提案したが、提案モデルには Lee-Carter モデルの時間平均項 α_x に対応する項を設けていない。本節では時間平均項の有無による推定結果の差について言及する。

本節で比較するモデルは

$$\log(1 - m_{x,t,c,s}) = \alpha_{x,c,s} + \beta_{x,c,s} \kappa_t + \varepsilon_{x,t,c,s}$$

という式で表されるモデルであり、 $\alpha_{x,c,s}$ の有無以外はすべて同条件とする。推定結果を表3に示す。

表3 AICの比較

モデル	対数尤度	パラメータ数	AIC
提案モデル	163577	954	-325246
比較モデル	168500	1676	-333648

この表から、提案モデルの AIC よりも比較モデルの AIC の方が値が小さい。すなわち AIC による定量的な判断に基づくと $\alpha_{x,c,s}$ を含む比較モデルの方が優れていると結論付けられる。しかし、この比較モデルか

ら推定された将来死亡率を図 18-21 に示す。ただし図 1-4,12,13 と同様、横軸は年齢区分、縦軸は対数死亡率 $\log m_{x,t,c,s}$ である。

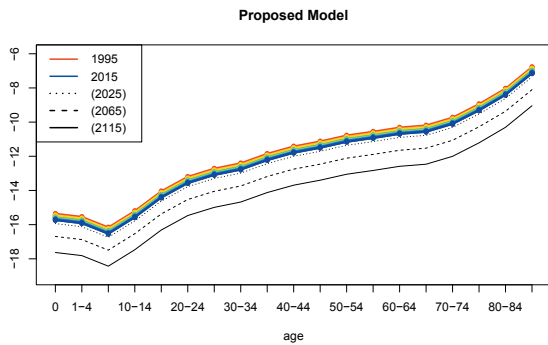


図 18 男性：精神及び行動の障害（提案モデル）

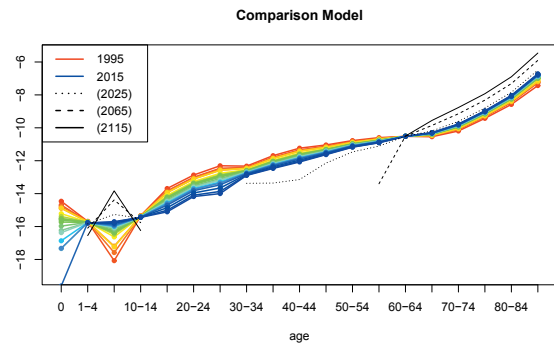


図 19 男性：精神及び行動の障害（比較モデル）

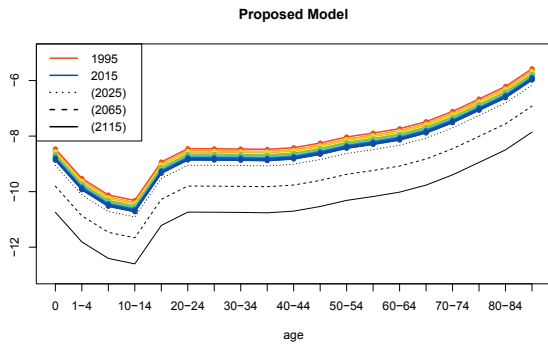


図 20 女性：傷病及び死亡の外因（提案モデル）

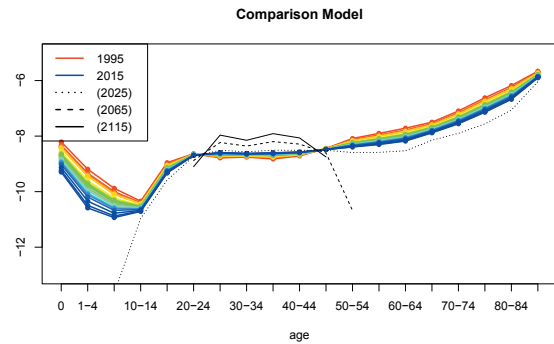


図 21 女性：傷病及び死亡の外因（比較モデル）

このように、 $\alpha_{x,c,s}$ を含めた場合の将来死亡率は、現在観測されているグラフの概形を大きく崩してしまうことがわかった。特に、死亡率が上昇し続ける年齢層と減少し続ける年齢層に分かれてしまう現象が確認でき、図に示していない死因も含めほとんどの死因で 50 年後の予測の段階で死亡率が 0 に収束してしまった。したがってこのモデルによるキャッシュフロー予測や意思決定は合理的でないと判断した。

こうなってしまう原因として $\beta_{x,c,s}$ の符号安定性が挙げられる。図 20,21 の推定に用いた $\beta_{x,c,s}$ をプロットするとそれぞれ図 22,23 のようになる。

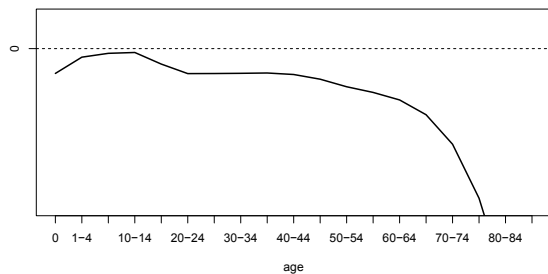


図 22 女性：傷病及び死亡の外因（提案モデルの β ）

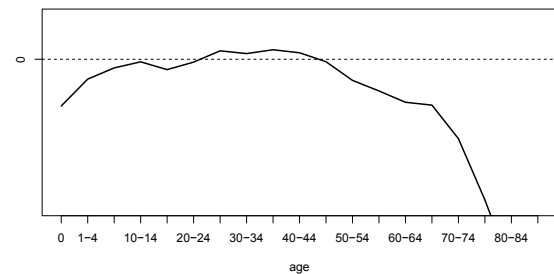


図 23 女性：傷病及び死亡の外因（比較モデルの β ）

図を見て明らかなように提案モデルは $\beta_{x,c,s}$ が常に負であるのに対し、比較モデルでは $\beta_{x,c,s}$ の一部に正の

部分が現れる。 κ_t は概ね単調減少であるため、死亡率の低下・上昇は $\beta_{x,c,s}$ の符号に依存し、 $\beta_{x,c,s}$ が正の部分は死亡率が年々上昇していくと推定される（これは図 21,23 を見ても明らかである）。したがって $\beta_{x,c,s}$ の符号が不安定であるから比較モデルは形状が安定しない、言い換えると $\beta_{x,c,s}$ の符号安定性が提案モデルの形状頑健性を高めている理由の 1 つだと言える。

7 全死因死亡率と Lee-Carter モデルとの比較

本節では提案モデルにより得られた死因別推定値を合算し全死因死亡率にした上で、通常の Lee-Carter モデルとの比較を行う。すなわち推定された死因別死亡率 $m_{x,t,c,s}$ に対して $\log \sum_c m_{x,t,c,s}$ を Lee-Carter モデルと比較する。しかし、誤差項を含めると分布が陽に表せないため平均構造のみで比較を行った。推定結果を図 24-26 に、観測値との norm percent による比較結果を表 4 に表す。ここで norm percent の定義は式 (2) と同様である。

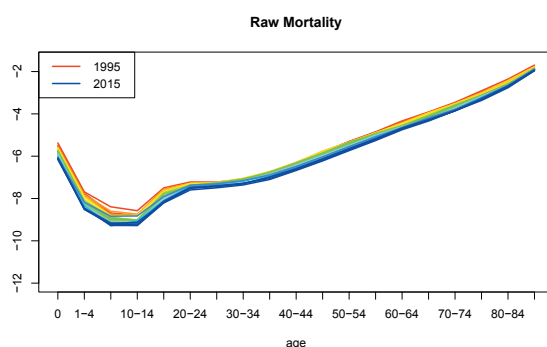


図 24 観測値

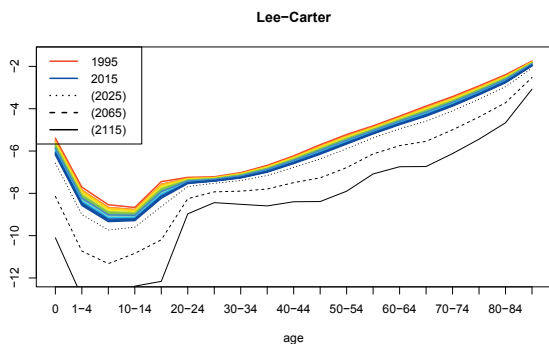


図 25 Lee-Carter モデル

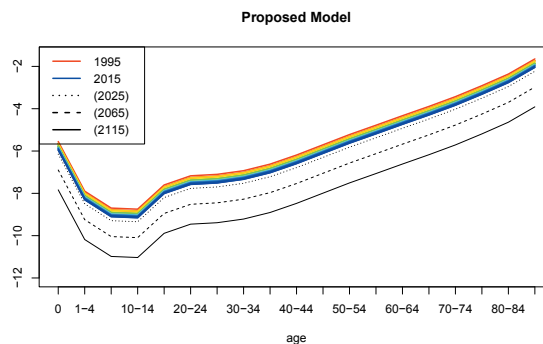


図 26 提案モデル

表 4 norm percent の比較

モデル	norm percent
提案モデル	75.49
Lee-Carter モデル	84.37

これらの図・表から、提案モデルは全体死亡率に対して通常の Lee-Carter モデルほどのデータ説明力を持たないものの 75% という比較的高い数値を得ることが出来た。また将来予測の観点では、Lee-Carter モデルが前節の比較モデルと同じように超長期の予測においてグラフの概形を崩してしまうのに対して、提案モデルは全死因死亡率の予測においてもグラフの概形を崩さず安定した推定値を得ることが出来た。

8 まとめと課題

本研究では Lee-Carter モデルとテンソル解析のアイデアを基に、年齢・暦年に加え死因・性別のデータを内包した 4 階テンソルのデータに対して新たな死亡率モデルの提案を行った。死因別死亡率を独立ではなく同時に推定することで年齢・性別によらない死因間の相関関係を含んだ推定をすることができ、全死因死亡率と整合的な結果を得ることが出来た。加えて長期の外挿における形状頑健性を確保するとともに、Wilmoth の指摘する死因別死亡率改善の過小推定問題を回避することが出来た。これにより死因ごとのトレンドや将来予測と同時に全死因死亡率のトレンドや将来予測を行うことができ、保険商品の開発や死亡率を用いる企業意思決定に応用出来ると考えている。

将来予測や全死因死亡率の比較では、本研究は平均構造にのみ注目したが実際は分散構造により多少のズレが生じるので、シミュレーション等を通じてそのズレがどのような影響を与えるかを調べる必要がある。特に、死因別の対数死亡率 $\log(1 - m_{x,t,c,s})$ は正規分布と仮定しているものの、死因合算後の対数死亡率 $\log \sum_c m_{x,t,c,s}$ はもはや正規分布に従っていないので、裾の重さ等の分布形に関する情報も調べる必要がある。今後の課題としたい。また今回は推定アルゴリズムの安定性・計算速度に鑑みてテンソル正規分布を仮定したが、複雑な相関を表現する方法として周辺分布にポアソン分布や二項分布を用い、相関構造にコンピュータを用いる方法も考えられるのでそういったモデルも検討していきたい。ただしそういった複雑なモデルが分布の形状頑健性を保つかどうか、実務に耐えうる解釈可能性や計算速度、推定の安定性をもつかどうかは未知の部分なので、今後の課題としてそちらも調査したい。

謝辞

本論文の投稿にあたり、有益なアドバイスを頂いた匿名査読者に感謝する。

参考文献

- [1] 石井太 [2017], 「日本の将来推計人口」, <http://www.actuaries.jp/lib/meeting/reikai29-7-siryu.pdf>, (2020年1月5日閲覧)
- [2] 野村俊一 [2018], 「テンソル分解を利用した Lee-Carter モデルの拡張と都道府県別生命表解析」, 『JARIP 研究発表大会 2018 予稿集』
- [3] Booth, H., Maindonald, J., and Smith, L. [2002], "Applying Lee-Carter under conditions of variable mortality decline", *Population studies*, 56(3), 325-336.
- [4] Brouhns, N., Denuit, M., and Vermunt, J. K. [2002], "A Poisson log-bilinear regression approach to the construction of projected lifetables", *Insurance: Mathematics and economics*, 31(3), 373-393.
- [5] Cairns, A. J. [2000], "A discussion of parameter and model uncertainty in insurance", *Insurance: Mathematics and Economics*, 27(3), 313-330.
- [6] Cairns, A. J., Blake, D., and Dowd, K. [2006], "A two - factor model for stochastic mortality with parameter uncertainty: theory and calibration", *Journal of Risk and Insurance*, 73(4), 687-718.
- [7] Girosi, F., and King, G. [2007], "Understanding the Lee-Carter mortality forecasting method", *Gking. Harvard. Edu.*
- [8] Hunter, G., and Luis, C. [2018], "An expectation-maximization algorithm for the matrix normal distribution with an application in remote sensing", *Journal of Multivariate Analysis*, 167, 31-48.
- [9] Kolda, T. G., and Bader, B. W. [2009], "Tensor decompositions and applications", *SIAM review*, 51(3), 455-500.
- [10] Lee, R. D., and Carter, L. R. [1992], "Modeling and forecasting US mortality", *Journal of the American statistical association*, 87(419), 659-671.
- [11] Lee, R. [2000], "The Lee-Carter method for forecasting mortality, with various extensions and applications", *North American actuarial journal*, 4(1), 80-91.
- [12] Li, J. S. H., Hardy, M. R., and Tan, K. S. [2009], "Uncertainty in mortality forecasting: an extension to the classical Lee-Carter approach", *ASTIN Bulletin: The Journal of the IAA*, 39(1), 137-164.
- [13] Manceur, A. M., and Dutilleul, P. [2013], "Maximum likelihood estimation for the tensor normal distribution: Algorithm, minimum sample size, and empirical bias and dispersion", *Journal of Computational and Applied Mathematics*, 239, 37-49.
- [14] Renshaw, A. E., and Haberman, S. [2003], "Lee-Carter mortality forecasting with age-specific enhancement", *Insurance: Mathematics and Economics*, 33(2), 255-272.
- [15] Renshaw, A. E., and Haberman, S. [2006], "A cohort-based extension to the Lee-Carter model for mortality reduction factors", *Insurance: Mathematics and economics*, 38(3), 556-570.
- [16] Wilmoth, J. R. [1995], "Are mortality projections always more pessimistic when disaggregated by cause of death?", *Mathematical Population Studies*, 5(4), 293-319.

付録 A テンソルの性質

A.1 テンソルの導入

本論文ではテンソルをプログラミングにおける多次元配列と同様に扱う。すなわち、ベクトルが1次元配列、行列が2次元配列として扱われることの類推として n 次元配列に対応するテンソル \mathcal{X} を

$$\mathcal{X} = (x_{i_1, i_2, \dots, i_n}) \in \mathbb{R}^{N_1 \times N_2 \times \dots \times N_n}$$

として扱う。これは 2×2 行列 A を $A = (a_{i,j})$ や $A \in \mathbb{R}^{2 \times 2}$ と表記することに対応しており、 n をテンソルの階数、 (N_1, N_2, \dots, N_n) をテンソルの次元といい、各添字をテンソルの軸またはモードという。

テンソルを数学的に厳密に定義するにはテンソル積による方法や多重線型写像を用いた方法などが知られているが、本論文ではテンソルの厳密な定義までは言及しない。

A.2 テンソルの演算

テンソルの式変形で頻繁に用いられるベクトル化 (vectorization) 作用素と行列化 (matricization) 作用素を定義する。ただし略記として $N = \prod_{k=1}^n N_k$, $N_{\neg m} = N/N_m$ を使い、総和の範囲も省略することがある。

まずテンソルのベクトル化作用素 $\text{vec} : \mathbb{R}^{N_1 \times N_2 \times \dots \times N_n} \rightarrow \mathbb{R}^N$ は

$$x_{i_1, i_2, \dots, i_n} = [\text{vec } \mathcal{X}]_i$$

ただし $i = 1 + \sum_{k=1}^n (i_k - 1) \prod_{l=1}^{k-1} N_l$

を満たすように定義される。ただし $k = 1$ のとき $\prod_{l=1}^0 N_l = 1$ と解釈する。また逆に、 i と (i_1, i_2, \dots, i_n) は一対一対応であるため、(添字の動きうる範囲での) 任意の i について $[\text{vec } \mathcal{X}]_i = x_{i_1, i_2, \dots, i_n}$ を満たす (i_1, i_2, \dots, i_n) が存在する。これはベクトル化作用素がテンソルを一列に並び替えているだけであることから明らかである。

次にテンソルのモード m に関する行列化作用素 $\bullet_{(m)} : \mathbb{R}^{N_1 \times N_2 \times \dots \times N_n} \rightarrow \mathbb{R}^{N_m \times N_{\neg m}}$ は、ベクトル化作用素を用いて

$$\mathcal{X}_{(m)} = (\text{vec}(\mathcal{X}_{i_m=1}), \text{vec}(\mathcal{X}_{i_m=2}), \dots, \text{vec}(\mathcal{X}_{i_m=N_m}))^T$$

と定義する。ここで $\mathcal{X}_{i_m=k}$ はテンソルの軸 m に関するスライスであり

$$\mathcal{X}_{i_m=k} = (x_{i_1, \dots, i_{m-1}, k, i_{m+1}, \dots, i_n}) \in \mathbb{R}^{N_{\neg m}}$$

である。これはプログラミングにおいてデータセットの特定の列を抽出することに対応している。

次にテンソルのノルムを定義する。テンソルのフロベニウスノルム $\|\bullet\|_F : \mathbb{R}^{N_1 \times N_2 \times \dots \times N_n} \rightarrow \mathbb{R}$ は

$$\|\mathcal{X}\|_F = \left(\sum_{i_1, i_2, \dots, i_n} x_{i_1, i_2, \dots, i_n}^2 \right)^{\frac{1}{2}}$$

と定義され、行列のフロベニウスノルムと同様に全要素の2乗和の平方根である。

今までは1つのテンソルに対する作用素であったが、テンソルと行列の間の積としてテンソルのモード積という演算が知られている。テンソルのモード m に対するモード積 $\bullet \times_m \bullet : \mathbb{R}^{N_1 \times \dots \times N_m \times \dots \times N_n} \times \mathbb{R}^{R_m \times N_m} \rightarrow \mathbb{R}^{N_1 \times \dots \times R_m \times \dots \times N_n}$ は次の式で与えられる演算である。

$$[\mathcal{X} \times_m A]_{i_1, \dots, r_m, \dots, i_n} = \sum_{i_m=1}^{N_m} a_{r_m, i_m} x_{i_1, \dots, i_m, \dots, i_n}$$

また、記号 $\times_{i=1}^n A_i$ を $\times_1 A_1 \times_2 \cdots \times_n A_n$ の略記とする。

ここでテンソルのモード積とベクトル化作用素を結びつける次の公式が存在する。ただし \otimes は行列のクロネッカー積である。

命題 1.

$$\text{vec} \left(\mathcal{X} \times_{i=1}^n A_i \right) = (A_N \otimes \cdots \otimes A_1) \text{vec } \mathcal{X}$$

[証明] テンソル $\mathcal{X} \times_{i=1}^n A_i$ の添字を整理する。行列 A_k の要素を $a_{r_k, i_k}^{(k)}$ で表すと、モード積の定義より

$$\left[\mathcal{X} \times_{i=1}^n A_i \right]_{r_1, \dots, r_n} = \sum_{i_1, \dots, i_n} x_{i_1, \dots, i_n} a_{r_1, i_1}^{(1)} \cdots a_{r_n, i_n}^{(n)}$$

次にクロネッカー積の定義から

$$\begin{aligned} a_{r_1, i_1}^{(1)} \cdots a_{r_n, i_n}^{(n)} &= [A_N \otimes \cdots \otimes A_1]_{r, i} \\ \text{ただし } r &= 1 + \sum_{k=1}^n (r_k - 1) \prod_{l=1}^{k-1} R_l, \\ i &= 1 + \sum_{k=1}^n (i_k - 1) \prod_{l=1}^{k-1} N_l \end{aligned}$$

また、上式の i と (i_1, \dots, i_n) 、そして r と (r_1, \dots, r_n) がそれぞれ一対一対応であることを用いると添字の動きうる範囲の任意の r に対して

$$\begin{aligned} [(A_N \otimes \cdots \otimes A_1) \text{vec } \mathcal{X}]_r &= \sum_i [A_N \otimes \cdots \otimes A_1]_{r, i} [\text{vec } \mathcal{X}]_i \\ &= \sum_{i_1, \dots, i_n} x_{i_1, \dots, i_n} a_{r_1, i_1}^{(1)} \cdots a_{r_n, i_n}^{(n)} \end{aligned}$$

を満たす (r_1, \dots, r_n) が存在する。

よって任意の r で

$$[(A_N \otimes \cdots \otimes A_1) \text{vec } \mathcal{X}]_r = \sum_{i_1, \dots, i_n} x_{i_1, \dots, i_n} a_{r_1, i_1}^{(1)} \cdots a_{r_n, i_n}^{(n)} = \left[\text{vec} \left(\mathcal{X} \times_{i=1}^n A_i \right) \right]_r$$

であり、すなわち

$$\text{vec} \left(\mathcal{X} \times_{i=1}^n A_i \right) = (A_N \otimes \cdots \otimes A_1) \text{vec } \mathcal{X}$$

が成り立つ。 □

同様にモード積と行列化作用素の間にも下記のような関係が成り立つが証明は省略する。

命題 2.

$$\left(\mathcal{X} \times_{i=1}^n A_i \right)_{(m)} = A_m \mathcal{X}_{(m)} A_{-m}^T$$

$$\text{ただし } A_{-m} = A_n \otimes \cdots \otimes A_{m+1} \otimes A_{m-1} \otimes \cdots \otimes A_1$$

付録 B テンソル正規分布

B.1 テンソル正規分布の定義

\mathcal{X} を n 階テンソルの空間 $\mathbb{R}^{N_1 \times N_2 \times \dots \times N_n}$ に値をとる確率変数とし、次のような密度関数を持つとする。

$$f(\mathcal{X}) = (2\pi)^{-\frac{N}{2}} \prod_{i=1}^n |\Sigma_i|^{-\frac{N}{2N_i}} \exp \left[-\frac{1}{2} \left\| (\mathcal{X} - \mathcal{M}) \times_{i=1}^n U_i \right\|_F^2 \right]$$

このとき \mathcal{X} は平均 \mathcal{M} 、分散 $\Sigma_1, \Sigma_2, \dots, \Sigma_n$ のテンソル正規に従うと言い、

$$\mathcal{X} \sim \text{TN}(\mathcal{M}, \Sigma_1, \Sigma_2, \dots, \Sigma_n)$$

と表す。ここで U_i は $U_i^T U_i = \Sigma_i^{-1}$ を満たす行列である（テンソル正規分布の詳細は Manceur et al.[2013] を参照）。

B.2 多変量正規分布との関係

テンソル正規分布は多変量正規分布を用いて次のように記述できることが知られている。

命題 3.

$$\begin{aligned} \mathcal{X} \sim \text{TN}(\mathcal{M}, \Sigma_1, \Sigma_2, \dots, \Sigma_n) &\iff x \sim \text{N}(m, \Sigma) \\ \text{ただし } x = \text{vec } \mathcal{X}, \quad m = \text{vec } \mathcal{M}, \quad \Sigma = \Sigma_n \otimes \dots \otimes \Sigma_1 \end{aligned} \quad (\text{B.1})$$

[証明] まずフロベニウスノルムをモード積とクロネッカー積の性質を用いて変形すると

$$\begin{aligned} \left\| (\mathcal{X} - \mathcal{M}) \times_{i=1}^n U_i \right\|_F^2 &= \left\| \text{vec} \left((\mathcal{X} - \mathcal{M}) \times_{i=1}^n U_i \right) \right\|_F^2 \\ &= \left\| (U_n \otimes U_{n-1} \otimes \dots \otimes U_1) \text{vec}(\mathcal{X} - \mathcal{M}) \right\|_F^2 \\ &= (x - m)^T (U_n \otimes U_{n-1} \otimes \dots \otimes U_1)^T (U_n \otimes U_{n-1} \otimes \dots \otimes U_1) (x - m) \\ &= (x - m)^T (U_n^T U_n \otimes U_{n-1}^T U_{n-1} \otimes \dots \otimes U_1^T U_1) (x - m) \\ &= (x - m)^T (\Sigma_n^{-1} \otimes \Sigma_{n-1}^{-1} \otimes \dots \otimes \Sigma_1^{-1}) (x - m) \\ &= (x - m)^T \Sigma^{-1} (x - m) \end{aligned} \quad (\text{B.2})$$

よってテンソル正規分布のカーネル（正規化項を除いた部分）を計算すると、上式より

$$\exp \left[-\frac{1}{2} \left\| (\mathcal{X} - \mathcal{M}) \times_{i=1}^n U_i \right\|_F^2 \right] = \exp \left[-\frac{1}{2} (x - m)^T \Sigma^{-1} (x - m) \right]$$

右辺は多変量正規分布 $\text{N}(m, \Sigma)$ のカーネルである。したがって

$$\mathcal{X} \sim \text{TN}(\mathcal{M}, \Sigma_1, \Sigma_2, \dots, \Sigma_n) \iff x \sim \text{N}(m, \Sigma)$$

が成り立つ。 □

B.3 パラメータの最尤推定量

$\mathcal{X} \sim \text{TN}(\mathcal{M}, \Sigma_1, \Sigma_2, \dots, \Sigma_n)$ を仮定する. 観測データ $\mathcal{X}_1, \dots, \mathcal{X}_T$ の対数尤度は

$$\log \mathcal{L} = \sum_{t=1}^T \left(-\frac{N}{2} \log 2\pi + \sum_{i=1}^n -\frac{N}{2N_i} \log |\Sigma_i| - \frac{1}{2} \left\| (\mathcal{X}_t - \mathcal{M}) \times_{i=1}^n U_i \right\|_F^2 \right) \quad (\text{B.3})$$

$$= -\frac{NT}{2} \log 2\pi + \sum_{i=1}^n -\frac{NT}{2N_i} \log |\Sigma_i| - \frac{1}{2} \sum_{t=1}^T \left\| (\mathcal{X}_t - \mathcal{M}) \times_{i=1}^n U_i \right\|_F^2 \quad (\text{B.4})$$

と書ける.

B.3.1 平均

ベクトル化した上で尤度方程式を求める. 記号は (B.1) と共通とし, (B.2) を用いると

$$\begin{aligned} \frac{\partial}{\partial m} \log \mathcal{L} &= -\frac{1}{2} \frac{\partial}{\partial m} \sum_{t=1}^T \left\| (\mathcal{X}_t - \mathcal{M}) \times_{i=1}^n U_i \right\|_F^2 \\ &= -\frac{1}{2} \frac{\partial}{\partial m} \sum_{t=1}^T (x_t - m)^T \Sigma^{-1} (x_t - m) \\ &= -\Sigma^{-1} \sum_{t=1}^T (x_t - m) \end{aligned}$$

したがって

$$\frac{\partial}{\partial m} \log \mathcal{L} = 0 \iff m = \frac{1}{T} \sum_{t=1}^T x_t \iff \mathcal{M} = \frac{1}{T} \sum_{t=1}^T \mathcal{X}_t$$

すなわち, テンソル正規分布においても標本平均が最尤推定量になる.

B.3.2 分散共分散行列

テンソル正規分布は分散共分散行列を軸の数だけ持つところが多変量正規分布との大きな違いであり, 全ての分散共分散行列の尤度方程式を連立させた解を解析的に求めることが出来ない. そこで交互最適化の枠組みを適用し, 他のパラメータを固定したうえで, 1つの分散共分散行列に対して尤度方程式を解くことを考える. $\mathcal{R}_t = \mathcal{X}_t - \mathcal{M}$ として, Σ_k の尤度方程式を求めるためにモード k での行列化を行うと

$$\begin{aligned} \frac{\partial}{\partial \Sigma_k} \log \mathcal{L} &= -\frac{NT}{2N_k} \frac{\partial}{\partial \Sigma_k} \log |\Sigma_k| - \frac{1}{2} \frac{\partial}{\partial \Sigma_k} \sum_{t=1}^T \left\| \mathcal{R}_t \times_{i=1}^n U_i \right\|_F^2 \\ &= -\frac{NT}{2N_k} \Sigma_k^{-T} - \frac{1}{2} \sum_{t=1}^T \frac{\partial}{\partial \Sigma_k} \left\| U_k \mathcal{R}_{t,(k)} U_{-k}^T \right\|_F^2 \\ &= -\frac{NT}{2N_k} \Sigma_k^{-T} - \frac{1}{2} \sum_{t=1}^T \frac{\partial}{\partial \Sigma_k} \text{Tr} \left[U_{-k} \mathcal{R}_{t,(k)}^T U_k^T U_k \mathcal{R}_{t,(k)} U_{-k}^T \right] \\ &= -\frac{NT}{2N_k} \Sigma_k^{-T} - \frac{1}{2} \sum_{t=1}^T \frac{\partial}{\partial \Sigma_k} \text{Tr} \left[U_{-k} \mathcal{R}_{t,(k)}^T \Sigma_k^{-1} \mathcal{R}_{t,(k)} U_{-k}^T \right] \\ &= -\frac{NT}{2N_k} \Sigma_k^{-T} + \frac{1}{2} \sum_{t=1}^T \Sigma_k^{-T} \mathcal{R}_{t,(k)} U_{-k}^T U_{-k} \mathcal{R}_{t,(k)}^T \Sigma_k^{-T} \\ &= -\frac{NT}{2N_k} \Sigma_k^{-T} + \frac{1}{2} \sum_{t=1}^T \Sigma_k^{-T} \mathcal{R}_{t,(k)} \Sigma_k^{-1} \mathcal{R}_{t,(k)}^T \Sigma_k^{-T} \end{aligned}$$

したがって

$$\begin{aligned}
\frac{\partial}{\partial \Sigma_k} \log \mathcal{L} = 0 &\iff \frac{NT}{2N_k} \Sigma_k^{-T} = \frac{1}{2} \sum_{t=1}^T \Sigma_k^{-T} \mathcal{R}_{t,(k)} \Sigma_{-k}^{-1} \mathcal{R}_{t,(k)}^T \Sigma_k^{-T} \\
&\iff \Sigma_k \frac{NT}{N_k} = \sum_{t=1}^T \mathcal{R}_{t,(k)} \Sigma_{-k}^{-1} \mathcal{R}_{t,(k)}^T \\
&\iff \Sigma_k = \frac{N_k}{NT} \sum_{t=1}^T \mathcal{R}_{t,(k)} \Sigma_{-k}^{-1} \mathcal{R}_{t,(k)}^T
\end{aligned}$$

この式と適当に与えた初期値によって分散共分散行列を逐次的に求めていく。しかし、テンソル正規分布の分散共分散行列は積演算によって結ばれているため、スケール変換の一意性を持たない。そこで分散共分散行列の値を $n-1$ 個固定して（例えば $\Sigma_1[1,1] = 1$, $\Sigma_2[1,1] = 1$ とする等）その後

$$x \sim N(m, s\Sigma)$$

を $\mathcal{M}, \Sigma_1, \dots, \Sigma_n$ が既知のもとで s について最適化するという方法が知られている（Hunter and Luis[2018]を参照）。このときの s の推定量は

$$\frac{\partial}{\partial s} \log \mathcal{L} = 0 \iff s = \frac{1}{NT} \left\| \mathcal{R}_t \times_{i=1}^n U_i \right\|_F^2$$

で与えられる。

また、本論文では分散共分散行列を対角行列に制限し推定を行っているが、今回はその場合の推定式を導出することはせずに、上記の繰り返し計算の途中で非対角成分に 0 を代入することによって対角行列を得た。式で書くとすれば下記の通りである。

$$\Sigma_k = \text{diag} \left(\frac{N_k}{NT} \sum_{t=1}^T \mathcal{R}_{t,(k)} \Sigma_{-k}^{-1} \mathcal{R}_{t,(k)}^T \right)$$

別表

別表1 ICD10 準拠基本分類表

第1章	感染症及び寄生虫症
第2章	新生物
第3章	血液及び造血器の疾患並びに免疫機構の障害
第4章	内分泌、栄養及び代謝疾患
第5章	精神及び行動の障害
第6章	神経系の疾患
第7章	眼及び付属器の疾患
第8章	耳及び乳様突起の疾患
第9章	循環器系の疾患
第10章	呼吸器系の疾患
第11章	消化器系の疾患
第12章	皮膚及び皮下組織の疾患
第13章	筋骨格系および結合組織の疾患
第14章	腎尿路生殖器系の疾患
第15章	妊娠、分娩及び産じょく〈褥〉
第16章	周産期に発生した病態
第17章	先天奇形、変形及び染色体異常
第18章	症状、徴候及び異常臨床所見・異常検査所見で他に分類されないもの
第19章	損傷、中毒及びその他の外因の影響
第20章	傷病及び死亡の外因
第21章	健康状態に影響を及ぼす要因及び保健サービスの利用
第22章	特殊目的用コード

Joint estimation of cause-specific future mortality using tensor analysis

Kotaro Suzuki Naoki Matsuyama

Abstract

In this article, we propose a new method for joint estimation of cause-specific mortalities. In generally, estimation of cause-specific mortalities cannot reflect interactions between causes of death, therefore the total of them does not make sense. For this reason, we propose a Lee-Carter type cause-specific mortalities model using single time development factor and tensor structure with four modes: age, time, cause of death, and sex. This method clarifies the correlation between causes of death and achieves robust extrapolation of cause-specific mortalities consistent with all-cause mortality.