

参照母集団データを用いた保険契約ポートフォリオの将来死亡率推計 ～Two population model にもとづくアプローチ～

西尾 和記

【要旨】

保険会社の経験データから Lee-Carter モデルにもとづいて直接的に将来の死亡率を推計することはデータのボリュームが不十分であるため困難である場合が多い。このため、国民の将来死亡率推計値に会社の経験値にもとづく調整を加えて、保険契約ポートフォリオの将来死亡率推計を行うことが通常の方法となる。本稿では、two population model、特に relative approach にもとづく保険契約ポートフォリオの将来死亡率推計のためのモデル設計手法を検討する。検討対象のモデルとして、参照母集団については Haberman and Renshaw [2011]等 で取り上げられているものを、小さい母集団との差には先行研究における代表的なモデルを選定する。これらを組み合わせた候補となるモデルに関して、提案するモデル選択基準にもとづく比較評価を行う。各モデルの比較検討を行うため、適用例として都道府県別人口の将来死亡率推計問題を扱う。本稿で検討するアプローチは、リスク管理およびソルベンシー評価のための内部モデル等のための将来死亡率前提の作成において有用であることが期待される。

【キーワード】

死亡リスク、長寿リスク、確率論的死亡モデル、multi population models、保険会社、リスク管理、内部モデル、都道府県データ

第1章 はじめに

死亡率の将来トレンドの推計は、保険会社、年金ファンドの将来給付額計算や国の将来人口推計等における応用を念頭に、世界各国でその研究が長年に亘って行われてきている。Booth and Tickle [2008]によると、死亡率の将来推計手法は大きく3つに分類される。1つ目は Lee-Carter モデルに代表される外挿法である。ここでは、過去のトレンドが将来も継続することが仮定され、年齢パターンおよび時間的なトレンドに見出される規則性が将来死亡率の計算に用いられる。2つ目は死因別に疫学的なモデルを利用して予測を行う方法である。3つ目は専門家の主観的な意見に基づく予測である。現時点において、これらのうちどれかが完璧な方法ということではなく、必要に応じて種々の観点から総合的な判断をすることとなる。本稿では、保険契約ポートフォリオに係る死亡率アサンプション設定のための利用を念頭に、外挿法による死亡率の将来推計に焦点を当てる。

外挿法に関する代表的なモデルは、1992年に発表された Lee-Carter モデルである(Lee and Carter [1992])。少ないパラメータで死亡率の過去実績によく当てはまることから、国の将来死亡率推計にも用いられている。時系列モデルを用いて外挿するため、確率論的なサンプルパスの生成が可能という特徴がある。Lee-Carter モデルについては、過去からの死亡トレンド変化への対応やデータへの適合度および予測性能の向上等を目的として種々の改良が研究されている(Lee [2000])。コホート効果を考慮したモデル等の拡張モデルのうち代表的なものは Cairns et al. [2009, 2011]、Haberman and Renshaw [2011]等において整理され、モデルが満たすべき要件を踏まえた比較評価が行われている。近年ではモデル構造を一般的な視点から捉え直し、新たな死亡率モデルを構築するための方法論 (general procedure)を探求する試みもある(Hunt and Blake [2014])。

これらは単一の母集団を対象として死亡率の将来予測を行う single population model に関する取組みであるが、ここ10年間においては類似したトレンド或いは相関を有する複数の母集団の死亡率を統合的に扱う multi population model の研究が注目を集めている。このアイデアは Carter and Lee [1992]にさかのぼるが、具体的な定式化は Li and Lee [2005]で与えられている。そこでは、ある母集団の男女それぞれに対して Lee-Carter モデルを個別に適用すると、時間の経過とともに男女の将来時点の死亡率に過去に観測されたことがない水準の格差が生じる課題が挙げられている。これを解決するため、Li and Lee [2005]は相関を有する各母集団を統合した全体の死亡率トレンドをまず求め、そこに各母集団のトレンドとの差に関する調整を加えて、個別の母集団の死亡率推計を行う方法を提示している。その後、Li and Lee [2005]に続いて、既存の single population model の拡張等、様々な観点からの multi population model に関する研究が行われてきている(Dowd et al. [2011]、Haberman et al. [2014]、Li et al. [2015])。

保険契約ポートフォリオに関する死亡率の将来推計モデルを構築する際には、保険会社の経験データのボリューム等にモデルの性能が依存することに注意が必要である。すなわち、件数が少ない場合にはデータにおいて死亡率が安定しないため、適切に実態を把

握することが困難となる。このような状況においては、multi population model の手法が有効である。小さい母集団の将来死亡率推計を目的とした代表的なモデルとして、Jarner and Kryger [2011]による SAINT(Spread Adjusted InterNational Trend)モデルが挙げられる。ここでは、小さい母集団と長期的な死亡率トレンドを共有する規模が大きい参照母集団を frailty モデルでモデリングし、小さい母集団との差を多変量時系列モデルで捉え、それらを統合して小さい母集団の死亡率推計を行う方法(relative approach)をとっている。そして、19 カ国のプーリングデータを参照母集団のデータとした上で、デンマークの将来死亡率推計に応用している。Wan and Bertschi [2015]は、同様のアプローチで参照母集団を Plat モデル、小さい母集団との差を拡張 Lee-Carter モデルで捉える方法により、スイスの年金ファンドの死亡率推計に応用している。Villegas et al. [2017] は、relative approach に関する一般的な定式化を示すとともに、長寿デリバティブのベシスリスク評価への応用を念頭にモデル評価を行っている。具体的には、Cairns et al. [2009] 等の single population model に関するモデル選択基準に修正を加えた基準のもと、two population model の既存モデルに relative approach の枠組みで扱うための修正を加えたモデルの比較評価を行っている。

本稿では、上記の一連の研究を踏まえて、リスク管理およびソルベンシー評価のための内部モデル等への応用を目的とした relative approach にもとづく保険契約ポートフォリオの将来死亡率推計モデルの設計手法を検討する。また、都道府県データを例として、モデル設計に係る一連のステップの数値例を示すとともに、各モデルの比較評価に関する議論を行う。検討対象とするモデルは、参照母集団については Haberman and Renshaw [2011] 等で取り上げられているものから選定し、小さい母集団との差には先行研究をもとに選定したいいくつかのモデルを候補とする。したがって、先行研究で扱われていない two population model も対象としている。モデル選択基準としては、Villegas et al. [2017]のベシスリスク評価のための基準を修正したものを提案する。なお、参照母集団として日本国民を対象に将来死亡率推計モデルを検討する際は、日本が置かれた状況等を考慮して、小松[2002]や石井[2008, 2013]で Lee-Carter モデルに対して適用されているようなモデル修正を適用することも考えられるが、これらについては本稿の議論の範囲外とする。

本稿の構成は次のとおりである。第2章では、本稿で対象とする確率論的死亡モデルの概要を与える。第3章では、保険契約ポートフォリオに関する将来死亡率推計モデル構築のためのデータの準備について議論した後、本稿の目的を踏まえたモデル選択基準を導入する。また、本基準にもとづくモデリングの手順について述べる。第4章では、小さい母集団として都道府県データを例に、第3章のモデル設計手法の適用結果を示すとともに各モデルの比較評価に関する議論を行う。最後に第5章で本稿をまとめる。

第2章 モデル概要

2.1 記法

対象とする母集団の暦年 t における満年齢 x 歳の死亡数を表す確率変数を D_{xt} 、観測値を

d_{xt} 、平均人口またはリスク対応生存延べ年数(exposure-to-risk)を E_{xt} とする。暦年 t における満年齢 x 歳の死亡率の期待値を $m_{xt} = E(D_{xt}/E_{xt})$ と記載する。また、時刻 t に x 歳の者が t から $t + 1$ の間に死亡する確率を q_{xt} とする。死力 μ_{xt} は満年齢 x 歳および暦年 t において一定と仮定する。これより、 $q_{xt} = 1 - \exp(-\mu_{xt}) = 1 - \exp(-m_{xt})$ となる。

Two population modelに係る記法としては、参照母集団に関する諸数値に対して ref、目標ポートフォリオに対して port の添え字を用いるものとする。例えば、前述の記法のもと、参照母集団および目標ポートフォリオの暦年 t における満年齢 x 歳の死亡率の期待値はそれぞれ m_{xt}^{ref} 、 m_{xt}^{port} となる。なお、特段母集団を区別する必要がない場合には添え字は記載しない。

2. 2 Single population model

本節では、Cairns et al. [2009]、Haberman and Renshaw [2011]等で挙げられる代表的な single population model の概要を与える。本稿では、以下のうち M1、LC2、M2、M3、M5*、M6*、M6**を参照母集団のモデルの候補として検討する。これらに関しては、Brouhns et al. [2002]と同様に死亡数 D_{xt} がポアソン分布に従うことを前提とする。

$$D_{xt} \sim \text{Poisson}(m_{xt}E_{xt}) \quad (1)$$

この確率分布の前提のもと、 m_{xt} の対数を予測子で表現したモデルを検討する。CBD および拡張モデルについては、オリジナルの論文と同様に死亡数 D_{xt} が二項分布に従うことを前提とし、 q_{xt} のロジットに関するモデルを記載している。M1 等でも二項分布アプローチとすることは多くの論文で示されるように可能であるが、本稿では 3.1 節で述べる想定データ(日本版死亡データベース)との整合性を考慮して、ポアソン分布アプローチを採用する。(詳細は Hunt and Blake [2015]を参照されたい。)

なお、以下では識別性の問題について述べているが、識別性を確保するために設定している制約条件は記載のもの以外とすることも可能である。この場合、トレンド項の変化により時系列モデルのパラメータが変わり、将来死亡率の推計値に影響を与えることがあるため、一定の留意が必要である。(この点についての議論は Cairns et al. [2011]等を参照されたい。)

(1) Lee-Carter モデル(M1)

代表的な single population model である Lee-Carter モデルは次で与えられる。

$$\log m_{xt} = \alpha_x + \beta_x^{(1)} \kappa_t^{(1)} \quad (2)$$

右辺第 1 項の α_x は、時点によらない母集団の年齢別の死亡率形状を表す。第 2 項は死亡率のトレンド項であり、 $\kappa_t^{(1)}$ によってトレンドを時点 t の 1 パラメータで表現し、 $\beta_x^{(1)}$ で年齢毎の変化率の違いを調整する。右辺は任意の実数 c_1 および $c_2 \neq 0$ に対して、変換

$$\left(\alpha_x, \beta_x^{(1)}, \kappa_t^{(1)}\right) \rightarrow \left(\alpha_x + c_1 \beta_x^{(1)}, \frac{1}{c_2} \beta_x^{(1)}, c_2 \left(\kappa_t^{(1)} - c_1\right)\right) \quad (3)$$

のもとで不変となることから、識別性を確保するためにモデルのパラメータについて制約条件を設定する必要がある。Lee and Carter [1992]は次を制約条件としている。

$$\sum_x \beta_x^{(1)} = 1, \sum_t \kappa_t^{(1)} = 0 \quad (4)$$

この制約条件は、上記(3)で

$$c_1 = \frac{1}{n} \sum_t \kappa_t^{(1)}, c_2 = \sum_x \beta_x^{(1)} \quad (5)$$

とした変換を、モデルをデータに当てはめることで得られたパラメータの推定値に対して行うことによって課すことができる。ここで、 n はデータにおける期間要素の数を表す定数である。(以下、同様とする。)

将来の死亡率は、トレンド項 $\kappa_t^{(1)}$ を外挿することで予測値を算出する。トレンド項 $\kappa_t^{(1)}$ のモデリングにはARIMAモデルが用いられる。特に、ドリフト項付ランダムウォーク

$$\kappa_t^{(1)} = \delta + \kappa_{t-1}^{(1)} + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma_\varepsilon^2) \quad (6)$$

が実績によく当てはまり、パラメータ数が少ないモデルであることから主に用いられている。本稿においても同様に(6)を用いるモデルとする。 δ はドリフト項、 ε_t は分散を σ_ε^2 とする正規ホワイトノイズ過程である。

以降の議論では断りのない限り、これらのオリジナルのモデルを検討の対象とする。なお、Lee-Carterモデルに関しては、対象とする母集団の特性等に応じて更なるモデル修正が加えられることがある(Lee [2000])。日本の2017年の将来推計人口における将来生命表の作成においては、モデルのベースとなるLee-Carterモデルについて、 α_x を直近4年間の対数死亡率の平均とする(オリジナルのモデルでは、制約条件(4)の $\kappa_t^{(1)}$ に関する条件を(2)に適用することで $\alpha_x = \frac{1}{n} \sum_t \log m_{xt}$ となる)、 $\kappa_t^{(1)}$ の外挿に時系列モデルではなくある非線形関数を当てはめるといった修正が行われている。

(2) 拡張 Lee-Carter モデル (LC2)

Lee-Carterモデルは、発表された当時には特異値分解によるパラメータ推定を基礎として、特異値分解の第1成分のみを残した算式としていたが、その後、データへの適合度を改善するために高次の項を加えることによる拡張が検討された(Booth et al. [2002])。ただし、この場合モデルが複雑になり、複数の $\kappa_t^{(i)}$ 項の外挿が必要となることから、予測を目的とする場合は時系列モデルの妥当性に注意を払う必要がある。拡張 Lee-Carterモデルに関して、Renshaw and Haberman [2003]では2次の拡張が取り扱われている。

$$\log m_{xt} = \alpha_x + \beta_x^{(1)} \kappa_t^{(1)} + \beta_x^{(2)} \kappa_t^{(2)} \quad (7)$$

ここでも、先ほどと同様に識別性の問題があり、実際、右辺は任意の実数 c_1, c_2 および $c_3, c_4 \neq 0$ に対して、次の変換のもとで不変となる。

$$\begin{aligned} & (\alpha_x, \beta_x^{(1)}, \kappa_t^{(1)}, \beta_x^{(2)}, \kappa_t^{(2)}) \\ & \rightarrow \left(\alpha_x + c_1 \beta_x^{(1)} + c_2 \beta_x^{(2)}, \frac{1}{c_3} \beta_x^{(1)}, c_3 (\kappa_t^{(1)} - c_1), \frac{1}{c_4} \beta_x^{(2)}, c_4 (\kappa_t^{(2)} - c_2) \right) \end{aligned} \quad (8)$$

Renshaw and Haberman [2003]では、次を制約条件としている。

$$\sum_x \beta_x^{(1)} = 1, \sum_t \kappa_t^{(1)} = 0, \sum_x \beta_x^{(2)} = 1, \sum_t \kappa_t^{(2)} = 0 \quad (9)$$

この条件は

$$c_1 = \frac{1}{n} \sum_t \kappa_t^{(1)}, c_2 = \frac{1}{n} \sum_t \kappa_t^{(2)}, c_3 = \sum_x \beta_x^{(1)}, c_4 = \sum_x \beta_x^{(2)} \quad (10)$$

とした(8)の変換により課される。

死亡率予測において、トレンド項 $\kappa_t^{(1)}, \kappa_t^{(2)}$ のモデリングにはARIMAモデルを用いることが考えられる。本稿ではモデル比較を行う観点から、これらを多変量ドリフト項付ランダムウォークでモデリングする。実際に適用する時系列モデル等には更なる検討の余地があると考えられる(近藤[2013])。

(3) Renshaw-Haberman モデル (M2)

Lee-Carter モデルに関してコホート効果を考慮に入れたモデルの一般化が Renshaw and Haberman [2006]において行われている。

$$\log m_{xt} = \alpha_x + \beta_x^{(1)} \kappa_t^{(1)} + \beta_x^{(0)} \gamma_{t-x} \quad (11)$$

右辺第3項がコホート効果に関する項であり、添え字の $c = t - x$ は生年である。このモデルは生年の項を右辺に加えているため、パラメータ推定の安定性が問題となる。パラメータ推定については後述するが、 $\beta_x^{(0)} = 1$ としたモデルの方が安定性の観点から扱いやすいことが知られている。本稿ではこのように簡便化した次のモデルを扱う。

$$\log m_{xt} = \alpha_x + \beta_x^{(1)} \kappa_t^{(1)} + \gamma_{t-x} \quad (12)$$

前述のモデルと同様に、任意の実数 c_1, c_2 および $c_3 \neq 0$ に対して、右辺は

$$\begin{aligned} & (\alpha_x, \beta_x^{(1)}, \kappa_t^{(1)}, \gamma_{t-x}) \\ & \rightarrow \left(\alpha_x + c_1 \beta_x^{(1)} + c_2, \frac{1}{c_3} \beta_x^{(1)}, c_3 (\kappa_t^{(1)} - c_1), \gamma_{t-x} - c_2 \right) \end{aligned} \quad (13)$$

の変換のもとで不変となる。ここでは次を制約条件とする。

$$\sum_x \beta_x^{(1)} = 1, \sum_t \kappa_t^{(1)} = 0, \sum_c \gamma_c = 0 \quad (14)$$

この条件は(13)において

$$c_1 = \frac{1}{n} \sum_t \kappa_t^{(1)}, c_2 = \frac{1}{n+k-1} \sum_c \gamma_c, c_3 = \sum_x \beta_x^{(1)} \quad (15)$$

とすることで課することができる。式中の k は、データにおける年齢要素の数を表す定数である。(以下、同様とする。)

本モデルの死亡率予測においては、コホート項の外挿方法が論点となる。これに対して、Cairns et al. [2011]では ARIMA(p,d,q)モデル

$$\Delta^d \gamma_c = \delta_0 + \sum_{j=1}^p \phi_j \Delta^d \gamma_{c-j} + \xi_c + \sum_{k=1}^q \delta_k \xi_{c-k}, \xi_c \sim N(0, \sigma_\gamma^2) \quad (16)$$

によりモデリングを行っている。ここで、 Δ^d は d 階の差分演算子、 ϕ, δ は ARIMA モデルの係数パラメータである。本稿では、当該モデルによりコホート項をモデリングする。なお、 $\kappa_t^{(1)}$ は他と同様にドリフト項付ランダムウォークでモデリングする。

(4) APC モデル (M3)

Renshaw-Haberman モデルにおいて、さらに $\beta_x^{(1)} = 1$ としたモデルは APC(age-period-cohort)モデルと呼ばれる。

$$\log m_{xt} = \alpha_x + \kappa_t^{(1)} + \gamma_{t-x} \quad (17)$$

このモデルは Renshaw-Haberman モデルと比較してパラメータのロバスト性に優れるといった特徴がある。任意の実数 $\varphi_1, \varphi_2, c_1$ に対して、APC モデルは変換

$$(\alpha_x, \kappa_t^{(1)}, \gamma_{t-x}) \rightarrow (\alpha_x + \varphi_1 - \varphi_2 x, \kappa_t^{(1)} + \varphi_2 t, \gamma_{t-x} - \varphi_1 - \varphi_2(t-x)) \quad (18)$$

$$(\alpha_x, \kappa_t^{(1)}, \gamma_{t-x}) \rightarrow (\alpha_x + c_1, \kappa_t^{(1)} - c_1, \gamma_{t-x})$$

のもとで不変となる。識別性は次の制約条件を設定することで確保可能である。

$$\sum_t \kappa_t^{(1)} = 0, \sum_c \gamma_c = 0, \sum_c c \gamma_c = 0 \quad (19)$$

最後の制約条件は、 γ_c に線形トレンドが含まれないとするものである。変換(18)において次のとおりパラメータを設定することで当該制約条件を課することができる。まず、 φ_1, φ_2 は最小二乗法で $\gamma_{t-x} = \varphi_1 + \varphi_2(t-x)$ の線形回帰を行った結果得られる推定値を適用する。 c_1 は他と同様に次のとおりとなる。

$$c_1 = \frac{1}{n} \sum_t \kappa_t^{(1)} \quad (20)$$

APC モデルの識別問題については文献によって異なるアプローチが見られるが、ここでは Villegas et al. [2015]の方法を用いている。

$\kappa_t^{(1)}$ および γ_c に関する外挿は、Renshaw-Haberman モデルと同様とする。

(5) P-spline モデル (M4)

Currie et al. [2004] において P-spline (penalized spline) によって死亡率曲線を表現するモデルが提案されている。

$$\log m_{xt} = \sum_{i,j} \theta_{ij} B_{ij}^{ay}(x, t) \quad (21)$$

このモデルはサンプルパスを生成できないことから本稿では対象外とする。(モデルが満たすべき要件については後述する。)

(6) CBD モデルおよび拡張モデル (M5～M8)

Cairns、Blake および Dowd は英国の高齢者死亡率によく当てはまるモデルとして CBD モデルを提案している(Cairns et al. [2006])。P-spline を除く前述のモデルとの主たる違いは、ある年の隣り合う年齢間では(コホート項の影響を除いて)死亡率が滑らかになることを基本とし、トレンドの年齢調整項を関数で設定する点である。このモデル思想の違いにより、これらは Lee-Carter 派生モデル(M1～M3、LC2)と区別して扱われる。以下は、上から順に CBD モデルおよび拡張モデルである。((22)～(25)のモデルをそれぞれ M5～M8 と呼ぶ。)

$$\text{logit } q_{xt} = \kappa_t^{(1)} + (x - \bar{x})\kappa_t^{(2)} \quad (22)$$

$$\text{logit } q_{xt} = \kappa_t^{(1)} + (x - \bar{x})\kappa_t^{(2)} + \gamma_{t-x} \quad (23)$$

$$\text{logit } q_{xt} = \kappa_t^{(1)} + (x - \bar{x})\kappa_t^{(2)} + \{(x - \bar{x})^2 - \hat{\sigma}_x^2\}\kappa_t^{(3)} + \gamma_{t-x} \quad (24)$$

$$\text{logit } q_{xt} = \kappa_t^{(1)} + (x - \bar{x})\kappa_t^{(2)} + (x_c - x)\gamma_{t-x} \quad (25)$$

ここで、 \bar{x} は平均年齢、 $\hat{\sigma}_x^2$ は $(x - \bar{x})^2$ の平均である。(いずれも元データの年齢範囲についての単純平均である。)また、 x_c は定数であり、推定が必要なパラメータである。M7 は M6 に 2 次の年齢調整項を加えたモデル、M8 は M6 に高齢になるとともにコホート効果が消滅するような修正を加えたモデルとなっている。なお、左辺については、前述のとおり、オリジナルの論文と同様に q_{xt} のロジットとしている。

これらについては、CBD モデル(M5)を除いて識別性の問題がある。通常設定される制約条件およびパラメータ変換は付録に示すとおりである。CBD 派生モデルはいずれも基本的に高齢者死亡率の当てはめを目的としている。本稿では、若齢を含めた幅広い

年齢範囲を扱うために、これらは参照母集団のモデルを検討する上で対象外とする。

(7) 混合モデル (M5*~M8*, M6**)

Plat [2009]において、CBD モデルに Lee-Carter モデルの持つ要素を組み合わせたモデルが提案されている。両者の特徴を合わせた混合モデルは若齢も含めた幅広い年齢範囲の死亡率を扱うのに適したモデルとなっている。Plat によるモデルは M6 を拡張した次の(27)であるが、Haberman and Renshaw [2011]において、M5、M7、M8 に対する拡張モデルも取り上げられている。

$$\log m_{xt} = \alpha_x + \kappa_t^{(1)} + (\bar{x} - x)\kappa_t^{(2)} + (\bar{x} - x)^+\kappa_t^{(3)} \quad (26)$$

$$\log m_{xt} = \alpha_x + \kappa_t^{(1)} + (\bar{x} - x)\kappa_t^{(2)} + (\bar{x} - x)^+\kappa_t^{(3)} + \gamma_{t-x} \quad (27)$$

$$\log m_{xt} = \alpha_x + \kappa_t^{(1)} + (\bar{x} - x)\kappa_t^{(2)} + (\bar{x} - x)^+\kappa_t^{(3)} + \{(\bar{x} - x)^2 - \hat{\sigma}_x^2\}\kappa_t^{(4)} + \gamma_{t-x} \quad (28)$$

$$\log m_{xt} = \alpha_x + \kappa_t^{(1)} + (\bar{x} - x)\kappa_t^{(2)} + (\bar{x} - x)^+\kappa_t^{(3)} + (x_c - x)\gamma_{t-x} \quad (29)$$

ここで、 $(\bar{x} - x)^+ = \max\{(\bar{x} - x), 0\}$ である。なお、式(26)~(29)では $(\bar{x} - x)$ の符号が、式(22)~(25)のものから反転している。以降では、(26)~(29)のモデルをそれぞれ M5*~M8*と呼ぶ。

M6*の場合、 $\kappa_t^{(3)}$ の項は若齢(最大で 40、50 歳まで)の死亡率のダイナミクスが大きく変動しうることを捉えるためのものである。高齢のみに興味がある場合は取り除くことができることが Plat [2009]において述べられている。(以下の(30)のモデルを M6**と呼ぶ。)

$$\log m_{xt} = \alpha_x + \kappa_t^{(1)} + (\bar{x} - x)\kappa_t^{(2)} + \gamma_{t-x} \quad (30)$$

本稿では、一般的に用いられる Plat モデルより複雑なモデルは対象外とし、M5*、M6*、M6**の 3 つのモデルのみを検討する。

これらについて、本稿で適用する制約条件等は付録に示すとおりである。トレンド項 $\kappa_t^{(i)}$ の時系列モデルは多変量ドリフト項付ランダムウォークとする。

2. 3 Two population model (relative approach)

前節の single population model では、一定規模以上の国を対象とする場合は十分なデータが確保されるため、信頼性のある将来死亡率推計を行うことが可能である。しかし、十分なデータがない母集団を扱う際は、実績死亡率が不安定となり適切にトレンドを捉えることが困難である。また、ある国の男女のような強く関係性がある母集団についてそれぞれ独立にモデルを作成し長期の将来予測を行う際には、過去の経験に反して予測結果が互いに発散してしまうといった問題が生じる。これらの問題に対して有用なモデルの 1 つとして multi population model がある。本稿では、multi population model にもとづくリスク管理およびソルベンシー評価のための内部モデル等への応用を目的とした保険会社の

将来死亡率推計について検討する。特に、Jarner and Kryger [2011]、Wan and Bertschi [2015]、Villegas et al. [2017]で用いられている relative approach による two population model に関するモデル設計手法をベースとする。この手法は、比較的理解しやすく、また実装環境がある程度整備されていることから、実務家にとって有用なアプローチであると考えられる。なお、目標ポートフォリオは参照母集団と比較して相対的に小さい母集団を想定する。

(1) Relative approach におけるモデリング

Villegas et al. [2017] では、参照母集団に対して当てはめた死亡率のモデル値を所与として、目標ポートフォリオの死亡率モデルを次のとおり定式化している。

$$D_{xt}^{\text{port}} \sim \text{Poisson}(m_{xt}^{\text{port}} E_{xt}^{\text{port}})$$

$$\log m_{xt}^{\text{port}} - \log m_{xt}^{\text{ref}} = \alpha_x^{\text{port}} + \sum_{j=1}^M \beta_x^{(j,\text{port})} \kappa_t^{(j,\text{port})} + \gamma_{t-x}^{\text{port}} \quad (31)$$

前述の single population model とは異なり、本モデルでは対数死亡率の差をモデリングしている。右辺の第1項から第3項はそれぞれ、基準となる死亡率水準、死亡率トレンド、コホート効果の違いを表している。 M は死亡率トレンドの違いを記述する年齢・期間項の個数であり、同論文では一般的に参照母集団の年齢・期間項の個数以下とされている。式変形すると、死亡率の比が次のように3つの要素の積となる。

$$\frac{m_{xt}^{\text{port}}}{m_{xt}^{\text{ref}}} = \exp(\alpha_x^{\text{port}}) \cdot \prod_{j=1}^M \exp(\beta_x^{(j,\text{port})} \kappa_t^{(j,\text{port})}) \cdot \exp(\gamma_{t-x}^{\text{port}}) \quad (32)$$

基準となる死亡率の調整を行いたい場合、右辺第1項のみとすることも可能である。一方、リスク管理等での応用を念頭に置くと、確率論的評価を視野に入れたモデルとなっていることが望ましい。右辺第2項は差に関するプロセスリスクを織り込むものであり、リスク評価をより精緻にすることが期待される。右辺第3項のコホート項は適合度を高めるものの、モデルに織り込むことを検討する際には目標ポートフォリオの実績データの年数の十分性等に注意する必要がある。

(2) 本稿で検討するモデル

参照母集団のモデルとして前節の age-period-cohort 型のものを選択した場合、目標ポートフォリオの基準となる年齢別の対数死亡率は $\alpha_x^{\text{ref}} + \alpha_x^{\text{port}}$ 、コホート効果に関する項は $\gamma_{t-x}^{\text{ref}} + \gamma_{t-x}^{\text{port}}$ となる。一方、トレンドに関する項は双線形であるため、 β_x と κ_x のそれぞれについて一般的には ref と port の和で表すことができない。ただし、ref と port において β_x が共通する場合は $\kappa_t^{\text{ref}} + \kappa_t^{\text{port}}$ と整理することが可能である。

Single population model におけるトレンドの年齢調整項についてのモデル思想を踏ま

えて、対数死亡率の差に関して表 1 のモデルを検討対象とする。この表は参照母集団のモデルグループごとに、候補モデルを列挙したものである。各列は参照母集団のモデルが Lee-Carter 派生モデルと混合モデルの場合に対応し、それぞれにおいて差に関するモデルの年齢調整項 $\beta_x^{(j,port)}$ をノンパラメトリックおよびパラメトリックなものとしている。Lee-Carter 派生モデルのグループに対しては、年齢調整項を $\beta_x^{(j,ref)}$ とした差に関するモデルを加えている。これは Kleinow [2015]の共通年齢効果(Common Age Effect)モデルに対応するものである。すなわち、各母集団のトレンドの年齢調整項に同様の傾向がある場合、共通のものを用いることでパラメータ数の削減を図る。Kleinow [2015]は relative approach を対象としたものではないが、アプローチの趣旨に鑑み、共通の年齢調整項とした方法を検討する。なお、M3 および混合モデルでは標準的に共通の年齢調整項となっている。

表 1 からの候補となるモデルの選択においては、前述のとおり、年齢・期間項の個数の意味で port の方が ref よりも簡素なモデルとなるようにする。すなわち、参照母集団のモデルを LC2 とする場合は、差のモデルは M0、M1、LC2、M1^(CAE)、LC2^(CAE)を候補とする。これは、包含される port の方が ref よりもモデルで適切に捉えられる情報量が一般的に少ないためと解釈できる。

表 1 : $\log m_{xt}^{port} - \log m_{xt}^{ref}$ に関する候補モデル

参照母集団のモデル	
M1, LC2, M2, M3	M5*, M6*, M6**
M0: α_x^{port}	M0: α_x^{port}
M1: $\alpha_x^{port} + \beta_x^{(1,port)} \kappa_t^{(1,port)}$	M5: $\kappa_t^{(1,port)} + (\bar{x} - x) \kappa_t^{(2,port)}$
M2: $\alpha_x^{port} + \beta_x^{(1,port)} \kappa_t^{(1,port)} + \gamma_{t-x}^{port}$	M6: $\kappa_t^{(1,port)} + (\bar{x} - x) \kappa_t^{(2,port)} + \gamma_{t-x}^{port}$
LC2: $\alpha_x^{port} + \beta_x^{(1,port)} \kappa_t^{(1,port)} + \beta_x^{(2,port)} \kappa_t^{(2,port)}$	M5**: $\alpha_x^{port} + \kappa_t^{(1,port)} + (\bar{x} - x) \kappa_t^{(2,port)}$
M3: $\alpha_x^{port} + \kappa_t^{(1,port)} + \gamma_{t-x}^{port}$	M6**: $\alpha_x^{port} + \kappa_t^{(1,port)} + (\bar{x} - x) \kappa_t^{(2,port)} + \gamma_{t-x}^{port}$
M1 ^(CAE) : $\alpha_x^{port} + \beta_x^{(1,ref)} \kappa_t^{(1,port)}$	
M2 ^(CAE) : $\alpha_x^{port} + \beta_x^{(1,ref)} \kappa_t^{(1,port)} + \gamma_{t-x}^{port}$	
LC2 ^(CAE) : $\alpha_x^{port} + \beta_x^{(1,ref)} \kappa_t^{(1,port)} + \beta_x^{(2,ref)} \kappa_t^{(2,port)}$	

(3) 時系列ダイナミクス

目標ポートフォリオの時系列ダイナミクスについては、Jarner and Kryger [2011]、Wan and Bertschi [2015]、Villegas et al. [2017]と同様の前提を置く。すなわち、参照母集団と目標ポートフォリオの死亡率トレンドが発散しないとし、差に関するモデルのトレンド項とコホート項は定常過程とする。それぞれ次のとおり VAR(1)および AR(1)とする。

$$\begin{aligned}\kappa_t^{\text{port}} &= \Phi_0 + \Phi_1 \kappa_{t-1}^{\text{port}} + \varepsilon_t^{\text{port}}, \varepsilon_t^{\text{port}} \sim N(0, \Sigma^{\text{port}}), \kappa_t^{\text{port}} = (\kappa_t^{(1,\text{port})}, \dots, \kappa_t^{(M,\text{port})})^T \\ \gamma_c^{\text{port}} &= \Psi_0 + \Psi_1 \gamma_{c-1}^{\text{port}} + \xi_c^{\text{port}}, \xi_c^{\text{port}} \sim N(0, \sigma_{\text{port}}^2)\end{aligned}\tag{33}$$

ここで、 Φ_0 と Φ_1 は M 次元ベクトルおよび $M \times M$ 行列であり、 Σ^{port} は $M \times M$ の分散共分散行列である。 κ_t^{ref} 、 γ_c^{ref} 、 κ_t^{port} 、 γ_c^{port} はそれぞれ独立と仮定する。

第3章 モデル設計

3.1 データの準備

(1) 参照母集団

本稿では日本国民を主に想定し、参照母集団に対するモデル作成のための基礎データを国立社会保障・人口問題研究所の日本版死亡データベースとする。これは国際的な死亡データベースである HMD (Human Mortality Database)の方法論に準拠して作成されている。死亡率の分母データはリスク対応生存延べ年数に対応していることから、前述のモデルとデータは整合的なものになっている。なお、HMD を基礎データとしても以下は同様の議論となる。

(2) 目標ポートフォリオ

保険契約ポートフォリオの将来死亡率推計モデルを構築するにあたって、モデリングの対象となるポートフォリオ区分を設定する必要がある。その際には、ポートフォリオ内の被保険者毎のリスク特性を同等に揃えることが基本となる。区分する要素は、商品種類(死亡保障、第3分野、個人年金等の他、定期、終身、養老等)、診査基準、経過年数、性別、保険料の払方、販売チャンネル等が考えられる。一方、区分における契約件数とのバランスに留意が必要である。十分なデータ量とならない場合には、上記の他、到達年齢を5歳刻みでまとめるなどの対応が考えられる。また、トレンドを把握するために一定のデータ蓄積期間が必要であることも考慮する。

3.2 モデル選択基準

第2章で挙げたモデルの候補から、リスク管理およびソルベンシー評価のための内部モデル等への応用を目的とした保険契約ポートフォリオの将来死亡率推計に用いるモデルを選定するために、モデル選択基準にもとづくモデルの比較評価を行う。以下で示す基準は、Villegas et al. [2017]でベースリスク評価用の two population model に関するモデル選択基準として用いられているものを本稿の目的を踏まえて修正したものである。モデルが満たすべき性質は次のとおりである。

<モデルの理論的性質に関する基準>

- 異なる年齢における死亡率の対前年比の変動に不完全な相関を生成する

2. サンプルパスの生成と予測区間の計算が可能である
3. シミュレーションにおいてパラメータの不確実性を取り込むことが可能である
4. 必要であればコホート効果を考慮することが可能である
5. 実務家が利用可能な標準的な統計的手法を用いて容易に実装できる
6. 透明性が高く、モデルの仮定、制約やアウトプットを利用者が理解することができ、非専門家に容易に説明が可能である

<データに当てはめた後のモデルが満たすべき基準>

7. 幅広い目標ポートフォリオ母集団に対して、参照母集団および目標ポートフォリオ母集団の両方の過去データに合理的な適合度を示す
8. 幅広い目標ポートフォリオ母集団に対して、死亡率の差、比のような 2 つの母集団に關与する指標または平均寿命に合理的な適合度を示す
9. 相対的に節約的なモデルである
10. 単一母集団および 2 つの母集団の指標について尤もらしく合理的な中心プロジェクションを生成する
11. 単一母集団および 2 つの母集団の指標のプロジェクションにおいて、尤もらしく合理的な不確実性の予測水準を生成し、それが過去の変動水準に沿ったものとなっている
12. データの期間および使用される年齢範囲に対してロバストなパラメータ推定値およびモデル予測値を生成する

ここでは候補としたモデルについて 1 から 6 を議論し、7 から 12 は次節で評価方法を示す。まず、基準 1 が満たされない場合は、どの年齢においても一律同じ対前年比で死亡率が改善することとなり、過去実績と不整合になる。この場合、たとえば、将来のある年の死亡保険金支払の合計に係る不確実性は実態より過度となりうる。M1、M2、M3 は時系列モデルで外挿するコホートの初期年齢を除いて完全な相関を有するため、結果に与える影響には留意が必要である。基準 2 については、M4 のみ要件を満たしていなかったため除外している。基準 3 はブートストラップ法を用いることにより対応可能であるが、時系列プロセスに紐づくパラメータリスクは評価できない点に留意が必要である。基準 4 に関しては、日本では昭和一桁世代の男性の死亡率がその前後と比較して高いと言われている。なお、日本の将来推計人口の死亡率モデルではコホート効果はモデルに織り込まれていない。基準 5 および 6 については、いずれのモデルも透明性が比較的高いものと考えられ、また標準的な統計ソフトのパッケージを用いて容易に実装可能である。以上の観測を踏まえて基準 1 を満たさないものがあるものの、M1 のベンチマークとしての意義も考慮してこの段階ではモデルを絞り込まないこととする。

3. 3 モデリングの手順

(1) 参照母集団のモデリング (STEP1)

はじめに参照母集団のモデルを設計する。本稿では幅広い年齢範囲を対象とすることから M1、LC2、M2、M3、M5*、M6*、M6**を候補とする。第2章の式(1)のとおり死亡数がポアソン分布に従う前提のもと、最尤法によりモデルのパラメータ推定を行う。対数尤度は次のとおりとなる。

$$l(\theta) = \sum_{x,t} w_{xt} \{d_{xt}^{\text{ref}} \log(m_{xt}^{\text{ref}} E_{xt}^{\text{ref}}) - m_{xt}^{\text{ref}} E_{xt}^{\text{ref}} - \log(d_{xt}^{\text{ref}}!)\} \quad (34)$$

ここで、 w_{xt} は (x,t) セルへの重みであり、データが欠損しているときに0、それ以外で1をとる。対数尤度の最大化はNewton-Raphson法により行う。なお、LC2、M2に関しては解の収束性の問題がある。本稿では、LC2、M2に対してNewton-Raphson法を実行する際は α_x 、 $\beta_x^{(1)}$ 、 $\kappa_t^{(1)}$ の初期値としてM1の解を与える。また、M2についてはコホート項に対してHunt and Villegas [2015]の近似制約条件を適用する。

パラメータ推定値のうちコホート項に関しては、当てはめに用いるデータの年数に注意が必要である。たとえば、最新暦年のデータの最も若いコホートについては、データが1年分しか利用できないため、コホート効果を推定するためには不十分と考えられる。本稿では、Cairns et al. [2009]等と同様に3年以下の観測値のコホートは当てはめにおいては除外するものとする。コホート項の時系列モデルであるARIMA(p,d,q)のモデル選択はBICにもとづいて行う。

以上のもと、次のとおり、モデル選択基準に沿って検証を行う。

- 過去データへの適合度等の確認(基準7,9)

情報量規準および残差プロットを用いた適合度等の確認を行う。情報量規準としては、 $AIC = 2v - 2l$ (赤池情報量規準)および $BIC = v \log K - 2l$ (ベイズ情報量規準)を各モデルについて計算し、数値を比較する。ここで、 v はモデルの有効パラメータ数、 l は最大対数尤度、 K は観測データ数を表す。情報量規準は適合度およびモデルの複雑度を考慮したものとなっており、数値が小さいほどよい。

残差プロットは適合度の確認を行うものであり、当てはめ後のモデルが過去データの特徴を適切に捉えられているかを検証する。死亡数がポアソン分布に従う前提のもと、プロットを行うスケール逸脱残差(scaled deviance residuals)は次のとおりである。ただし、 \hat{d}_{xt} は最尤法により定まる死亡数のモデル値である。

$$r_{xt} = \text{sign}(d_{xt} - \hat{d}_{xt}) \sqrt{\frac{\text{dev}(x,t)}{\hat{\phi}}} \quad (35)$$

ここで、逸脱度は

$$\text{dev}(x, t) = 2 \left[d_{xt} \log \left(\frac{d_{xt}}{\hat{d}_{xt}} \right) - (d_{xt} - \hat{d}_{xt}) \right] \quad (36)$$

であり、 $\hat{\phi}$ は次で与えられる。

$$\hat{\phi} = \frac{1}{K - v} \sum_{x,t} w_{xt} \text{dev}(x, t) \quad (37)$$

- 将来予測の尤もらしさの確認(基準 10, 11)

モデリングの目的は将来死亡率の予測であるため、過去データへの適合度等だけでなく、トレンドに沿って尤もらしい予測結果となっていることが重要である。この評価は主観を伴うものであるが、ここでは各モデルの代表年齢における死亡率およびコホート生存率推移をプロットすることで確認する。

- 元データに対するロバスト性の確認(基準 12)

モデルの当てはめに用いるデータの変更によりパラメータの推定値および予測結果が大きく変わらないことは、モデルの信頼性に関して重要な要素である。ここでは、異なる暦年範囲におけるパラメータ推定値およびモデル予測値の比較を行うことで元データに対するロバスト性の確認を行う。

本稿では、モデル予測結果の元データに対するロバスト性の確認を、バックテストによる予測性能の検証および死亡率の将来予測値のファンチャートの確認により行う。バックテストでは、予測対象を実績データの最終年の死亡率とし、予測に用いる元データとして暦年範囲の長さを固定した上で 1 年ずつ最新年に範囲をずらしていったものを適用した場合の予測性能の安定性を検証する。これは Dowd et al. [2010]で提案されている contracting horizon backtests に対応しているが、実務において保険契約ポートフォリオの将来死亡率推計を行う際に、元データの暦年範囲の長さを固定するケースを想定して、ここではこの方法を採用している。(なお、Dowd et al. [2010]では上記以外に 3 つのバックテスト手法が提案されている。) 予測性能の評価指標には、Jarner and Kryger [2011]と Wan and Bertschi [2015]で使用されている指標のうち次を用いる。

$$G_1 = \frac{\sum_{x,t} (d_{xt} - \hat{m}_{xt} E_{xt})}{\sum_{x,t} d_{xt}} \quad (38)$$

$$G_2 = \frac{\sum_{x,t} |d_{xt} - \hat{m}_{xt} E_{xt}|}{\sum_{x,t} d_{xt}} \quad (39)$$

$$MAPE = \text{mean} \frac{|d_{xt} - \hat{m}_{xt} E_{xt}|}{d_{xt}} \quad (40)$$

ここで、 d_{xt} と E_{xt} はそれぞれ暦年 t における満年齢 x 歳の死亡数およびリスク対応生存延べ年数の観測値、 \hat{m}_{xt} は死亡率の予測値の期待値である。 G_1 は偏り、 G_2 は近さを表す指標となっている。3 つ目は平均絶対誤差率(Mean Absolute Percentage Error)

である。各指標の詳細については、Jarner and Kryger [2011]および Wan and Bertschi [2015]を参照されたい。

上記の確認結果にもとづきモデルの選定を行うが、通常すべての基準において最良となるモデルはない。これを踏まえてここでは複数のモデルを選定する。

(2) 目標ポートフォリオ母集団のモデリング (STEP2)

選定された参照母集団の当てはめ後の死亡率および将来予測値を所与のものとして、表 1 のうち対応する目標ポートフォリオとの差に関するモデルのパラメータ推定を行う。推定は(31)の前提のもと、参照母集団と同様に最尤法によって行う。識別問題の制約条件の設定等は参照母集団と同様とする。コホート項の除外も同様に行う。

- **過去データへの適合度等の確認(基準 7, 8, 9)**

STEP1 と同様にAICとBICによる確認を行う。残差プロットによる検証は、残差の偏りに関するパターンがデータ量の少なさに起因する死亡率変動で隠れることがあるため行わない。ここでは、残差プロットに代わり、目標ポートフォリオの各年齢の死亡率から計算した生存率の実績とモデル値を比較することにより、モデルの過去データへの当てはまりの良さを確認する。

- **将来予測の尤もらしさの確認(基準 10, 11)**

参照母集団と同様に、各モデルの代表年齢における死亡率およびコホート生存率推移のプロットを確認する。

- **元データに対するロバスト性の確認(基準 12)**

参照母集団と同様に、異なる暦年範囲のデータに対するパラメータ推定値およびモデル予測値の比較を行うことで元データについてロバスト性の確認を行う。

第 4 章 都道府県別人口の将来死亡率推計への適用

前章のモデル設計手法に関する数値例として、比較的規模が小さい都道府県別人口の将来死亡率推計問題を扱う。国民全体よりも健康状態が良好と想定される被保険者群団と比較して、都道府県別人口は国民全体に相対的に近い特性を有すると考えられるが、都道府県別の平均寿命に差異があるように死亡率に一定の傾向の違いがあることから、この例は手法の保険契約ポートフォリオへの適用時の有用性を示唆するものと考えられる。

以下の数値例の作成においては、主に R の StMoMo パッケージを用いている。なお、本パッケージは single population model を対象として作成されており、現時点においては multi population model に幅広く対応していない。一方、relative approach に関しては 2 段階のステップに分けてモデルを作成することから、データの当てはめ等の大部分の計算を本パッケージの関数で実行可能となっている。StMoMo ではパラメータ推定に R の

gnm(generalized nonlinear model)パッケージが用いられているが、数値例の作成では目標ポートフォリオデータの当てはめにおいて $\log m_{xt}^{ref}$ をオフセット項とした。VAR(1)のカリブレーションと予測には、本パッケージに加えて vars および tsDyn パッケージを用いた。

4. 1 使用するデータについて

目標ポートフォリオとする都道府県は、人口数が全都道府県中最少であることから鳥取県とする。データは日本国民と同様に国立社会保障・人口問題研究所の日本版死亡データベースのものを使用する。

日本国民、鳥取県ともに、20歳から100歳の男性を対象とする。観測期間は、日本国民については1976年から2015年とする。鳥取県に関しては、保険会社データを念頭に、日本国民より短く1996年から2015年とする。なお、対象データにおける2015年のリスク対応生存延べ年数は日本国民で約5000万人、鳥取県で約22万人となっている。

4. 2 目標ポートフォリオに対して Lee-Carter モデルを適用した場合

はじめに、次節で適用する手法の必要性を確認するため、鳥取県のデータに対して直接 Lee-Carter モデルを当てはめた場合のパラメータ推定値および死亡率の将来予測値のファンチャート(予測区間を50%、80%、95%としたものをプロットする。以降も同様とする。)を、それぞれ図1、2に示す。

年齢調整項のパラメータ推定値に大きなバラつきが見られる。年齢調整項は各年齢の死亡率トレンドに影響し、図2においては25歳の死亡率の将来予測値に関して水準が高齢と逆転するといった妥当性に乏しい結果となっている。年齢調整項の平滑化により一定の修正は可能であるが、小さい群団のモデリングにおいて Lee-Carter モデルを直接適用することが不十分である場合があることが分かる。

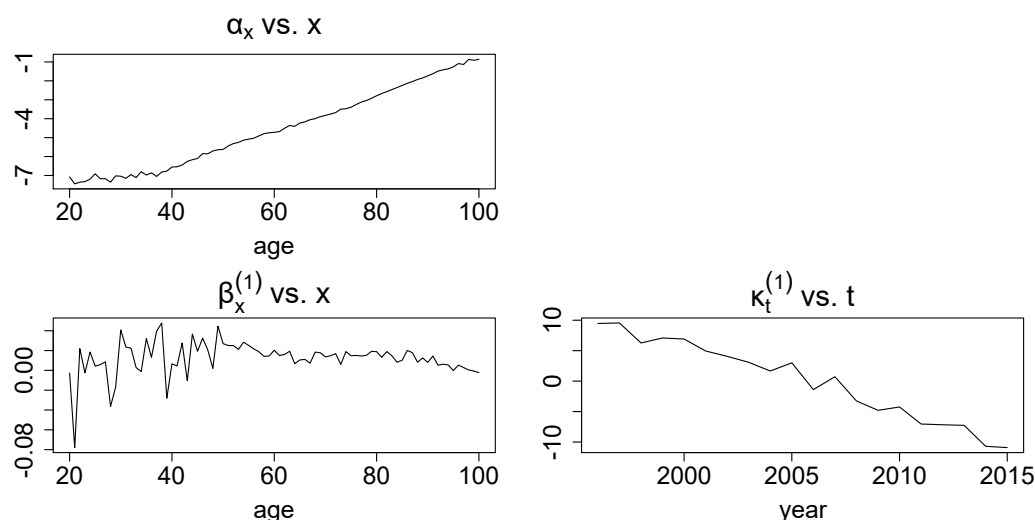


図1 鳥取県データに Lee-Carter モデルを当てはめた時のパラメータ推定値

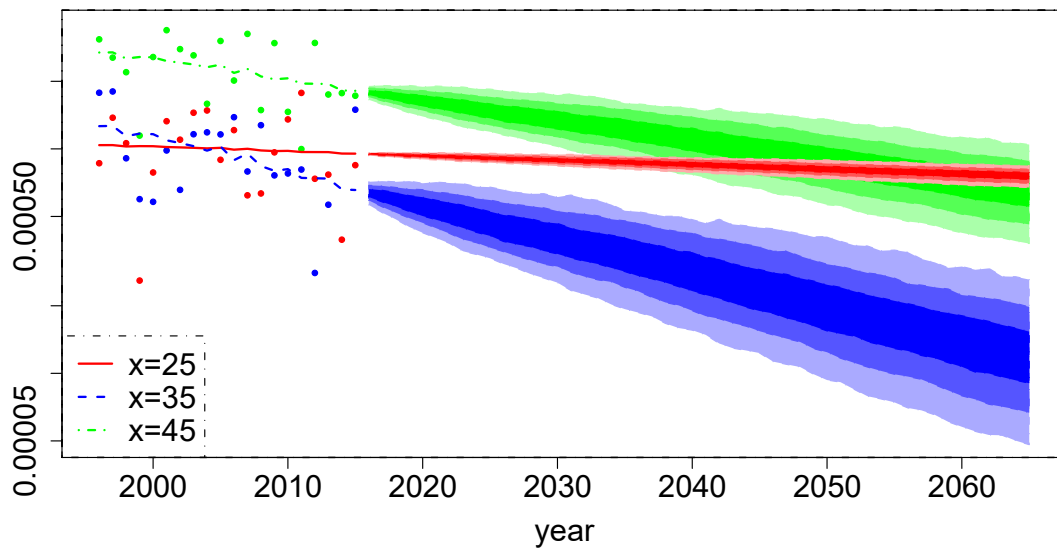


図2 鳥取県データに Lee-Carter モデルを当てはめた時の死亡率に係るファンチャート (縦軸は対数スケール)

4. 3 参照母集団のモデリング

この問題に対応するために、死亡率に関して鳥取県と相関があり規模が大きい日本国民のデータを活用する。はじめに、M1、LC2、M2、M3、M5*、M6*、M6**に前述の日本国民データを当てはめると、それぞれのパラメータ推定値は図7の実線のようにになる。いずれにおいても十分と考えられるデータ量があるため、前節の問題は生じていない。

各モデルのデータへの適合度等は表2のとおりとなる。パラメータ数が相対的に多いLC2、M2、M6*に関しては比較的当てはまりがよく、AICおよびBICも低くなっている。その他については、M3とM5*が高いAICおよびBICを示している。

表2 各モデルの参照母集団データへの適合度と節約性

	パラメータ数	対数尤度(順位)	AIC(順位)	BIC(順位)
M1	200	-27,772(5)	55,945(5)	57,161(5)
LC2	317	-24,143(3)	48,920(3)	50,847(3)
M2	312	-21,935(2)	44,495(2)	46,392(2)
M3	232	-31,232(6)	62,928(6)	64,338(6)
M5*	198	-32,380(7)	65,156(7)	66,360(7)
M6*	309	-20,434(1)	41,486(1)	43,365(1)
M6**	270	-26,768(4)	54,077(4)	55,719(4)

次に、各モデルの残差プロットを示す。図3は暦年を横軸、年齢を縦軸にとったヒートマップである。ランダムなバラつきが示されているほどモデルの当てはまりがよいと判

断するが、M1、LC2、M5*には対角方向にややパターンが見られる。これはモデルで捉えられていないコホート効果が存在するためと考えられる。図4は、横軸に年齢、暦年、生年をとった場合の散布図であるが、M1とLC2においては前述の昭和一桁代の高い死亡率が残差として残っていることが確認できる。一方、M5*についてはそれ以外の生年での残差の偏りが見られる。これはモデルがコホート効果以外の実績に含まれるパターンを捉えられておらず、その影響が表れているためと考えられる。

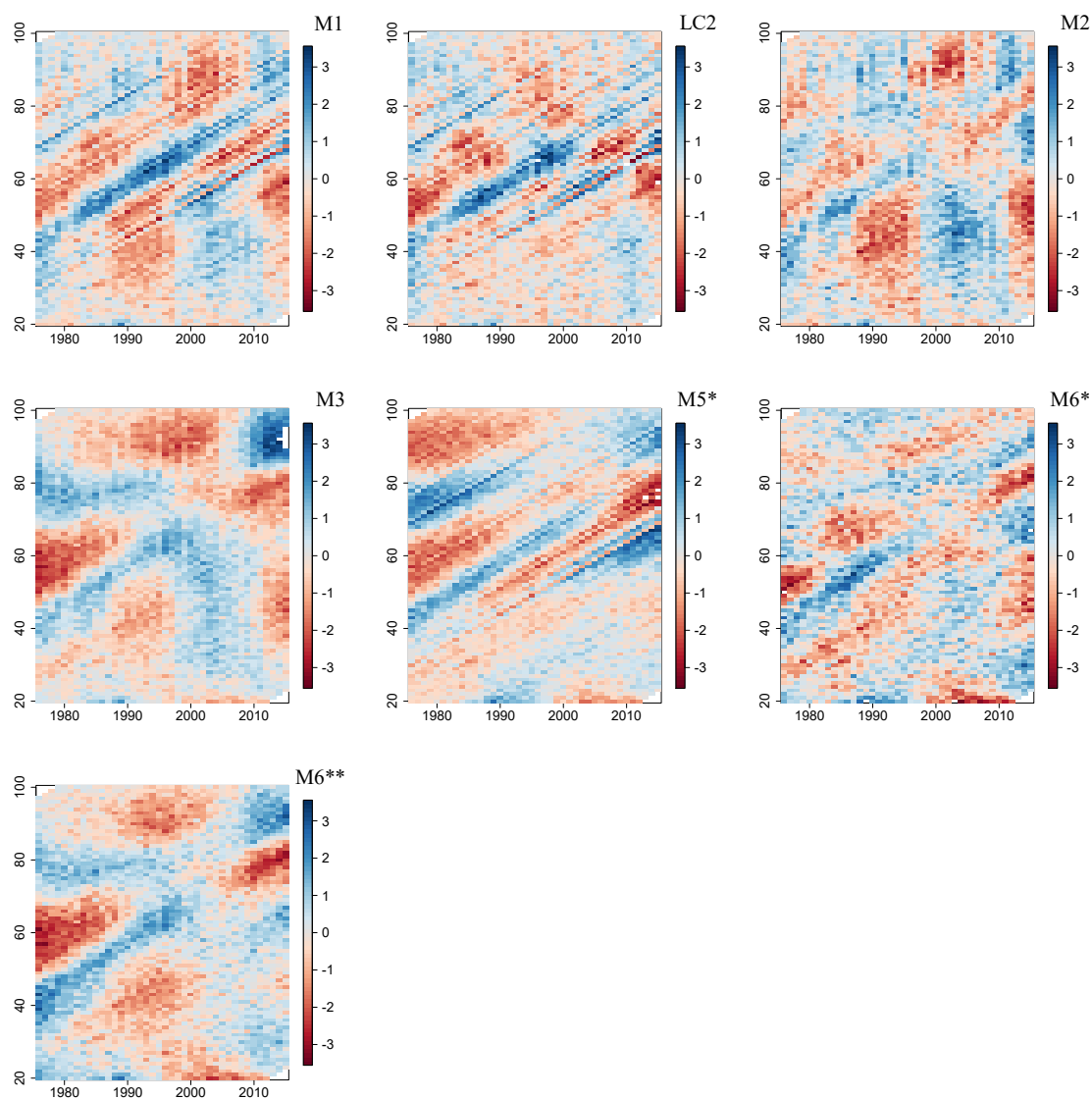


図3 日本人男性データ(20-100歳、1976-2015年)に当てはめた各モデルの残差のヒートマップ(横軸：暦年、縦軸：年齢)

将来予測値の尤もらしさを確認するために、モデル別の死亡率に関するファンチャートを図5に示す。M1、LC2では、将来に向かって過去トレンドが反映された結果となっ

ており、高齢においてはLC2の方がやや各年での変動が大きい。M2はコホート項の影響が表れており、M1と比べて異なる死亡率改善トレンドが見られる。また、若齢において変動幅が大きくなっている。M3については、過去トレンドとやや乖離した年齢があるものの大きな問題はないように思われる。M5*は年齢調整項の絶対値が60歳を中心として大きくなるモデル構造から、25歳と85歳での変動幅が大きくなっている。M6*はコホート項の影響で明らかに過去トレンドから大きく乖離した結果となっている。M6**では、若齢での死亡率変動を表現する項を除いたことにより、全体的な不自然さが軽減されている。M6*、M6**では死亡率の将来の上昇が見られるが、これは図7の実線のグラフで

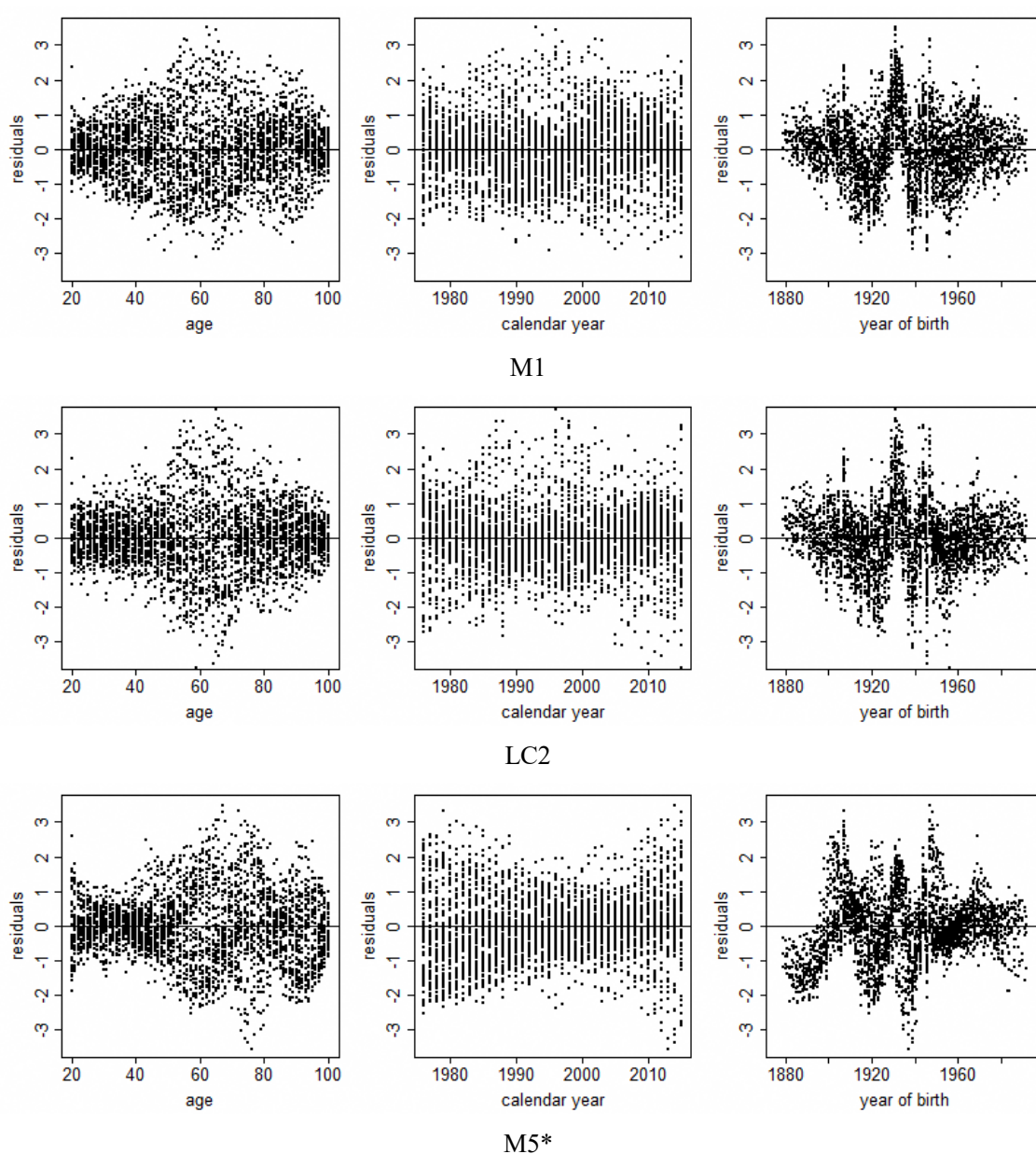


図4 日本人男性データ(20-100歳、1976-2015年)に当てはめた各モデルの残差の散布図

示されるように 1950 年から 1960 年あたりの生年からコホート項の値が右肩上がりとなっていることによるものである。これらのコホートについてはデータに若齢の実績しか含まれず、コホート効果を適切に捉えられていない可能性が考えられる。実際、Wan and Bertschi [2015]では、65 歳未満のデータのみから推定されたコホート項の安定性の問題が指摘されている。これより、図 5 に 2015 年に 65 歳未満のコホートに関するコホート項を 0 とした場合の死亡率のモデル値および将来予測値をプロットする。(M6*(2)、M6**(2) がそれぞれに該当する。)そこでは死亡率の将来上昇の要素は取り除かれているものの、コホート項を除いた部分の過去実績への当てはまりが悪くなる等の影響が見られる。

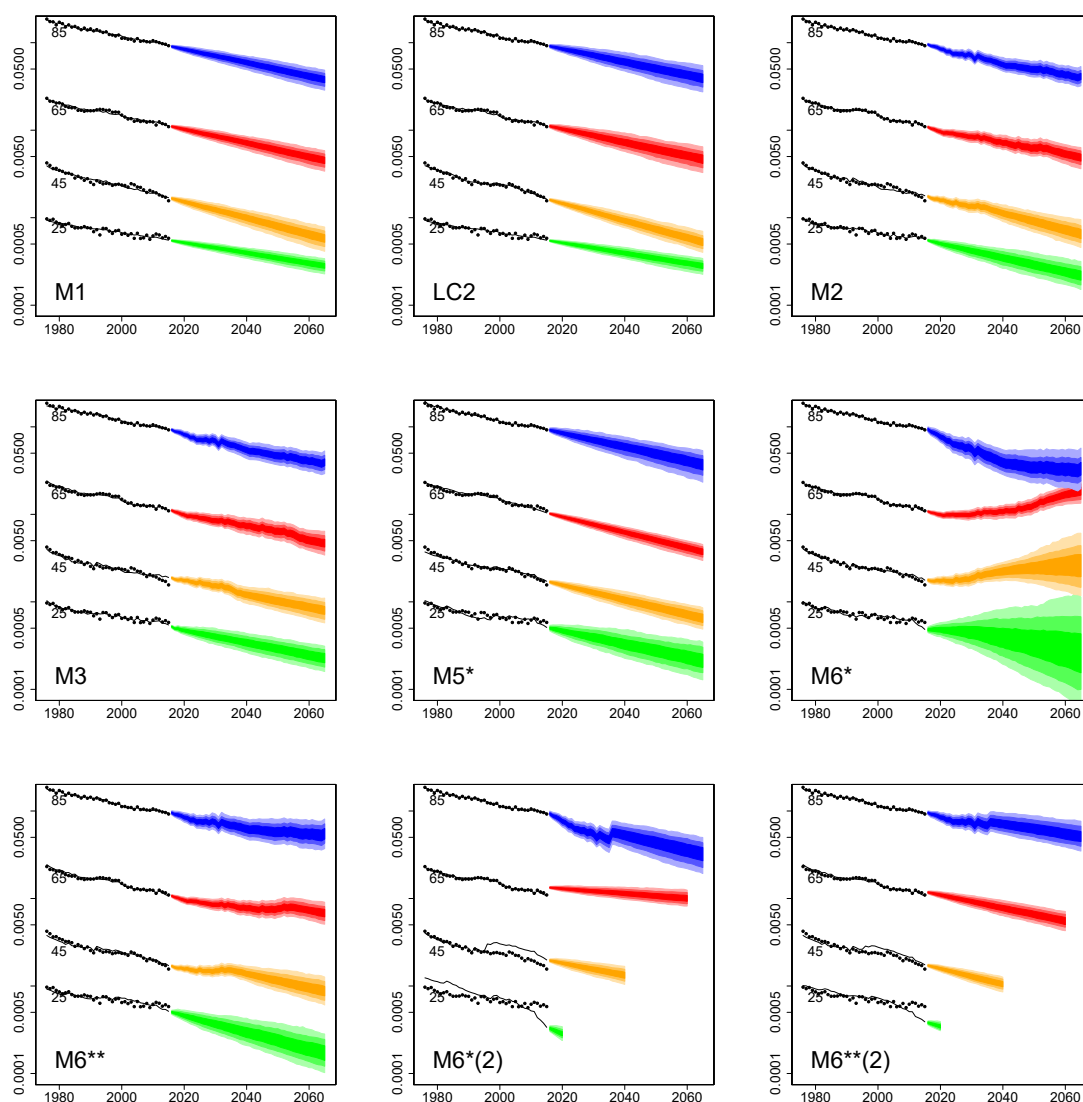


図 5 日本人男性の代表年齢別の将来死亡率のファンチャート : M6*(2)と M6**(2)は M6*, M6**において、2015 年に 65 歳未満のコホートに関するコホート項を 0 とした場合の死亡率のモデル値、将来予測値(縦軸は対数スケール)

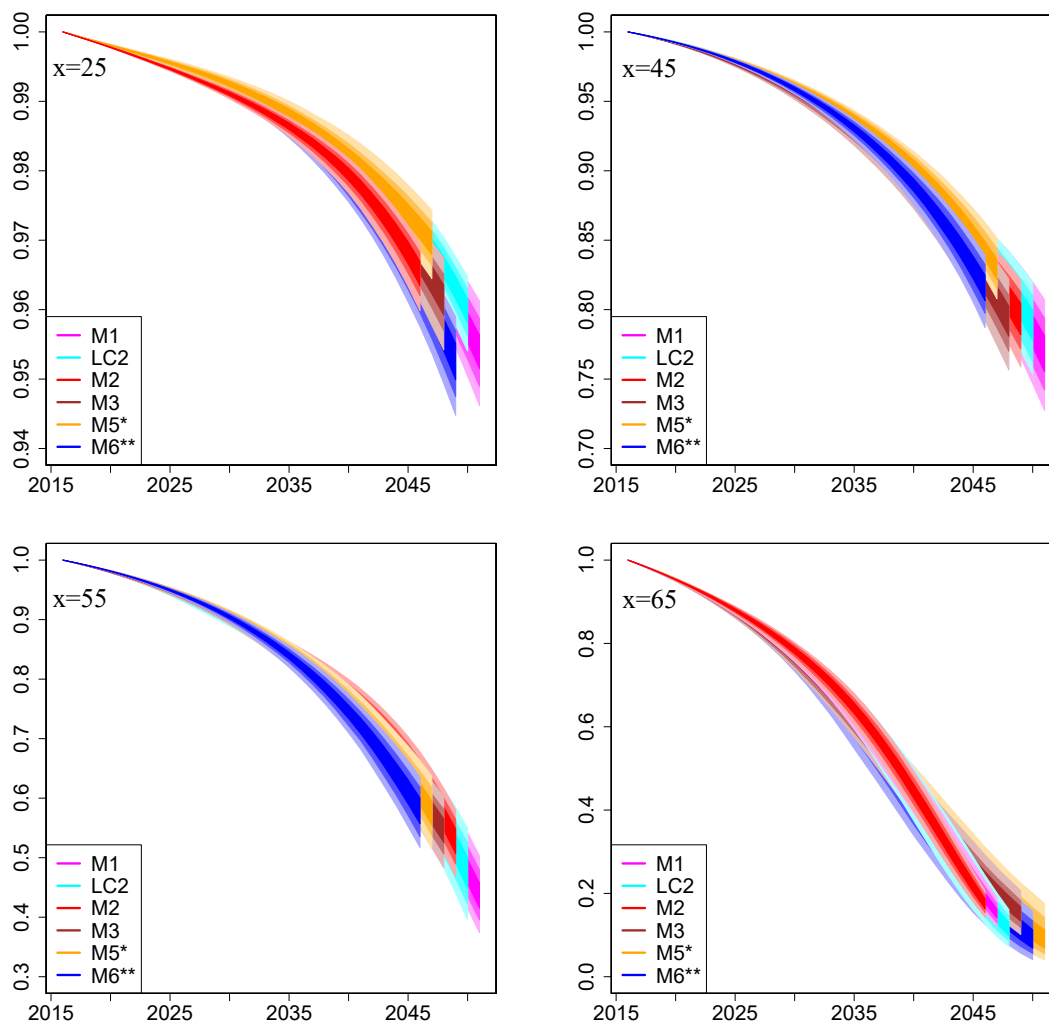


図6 日本人男性のうち2016年に25、45、55、65歳のコホートの35年間の生存率推定値のファンチャート

若齢の変動幅が特に大きい M6*を除いた特定コホートの将来生存率推移のファンチャートは図6のとおりである。モデルによって差異はあるものの、極端に外れたものはない。2016年に25歳および45歳のコホートに関しては、M2、M3、M6**で生存率が相対的に低くなっている。これらはコホート項を含むモデルである。

最後に、元データに対するロバスト性を確認する。図7はパラメータ推定値について、元データの期間を短く1996年から2015年にした場合を重ねたものである。M1やLC2のトレンド項について値が概ね平行にずれているのは、制約条件による影響と考えられる。各モデルのパラメータ推定値を見ると、コホート項を含むモデルであるM2、M3、M6*、M6**において、程度の違いはあるもののロバスト性がやや低いように思われる。これらのうちM3については、将来推計にほとんど影響のないコホートで差が生じている。それぞれの元データのもとでの将来死亡率推移の比較を行ったものが図8である。水色のグ

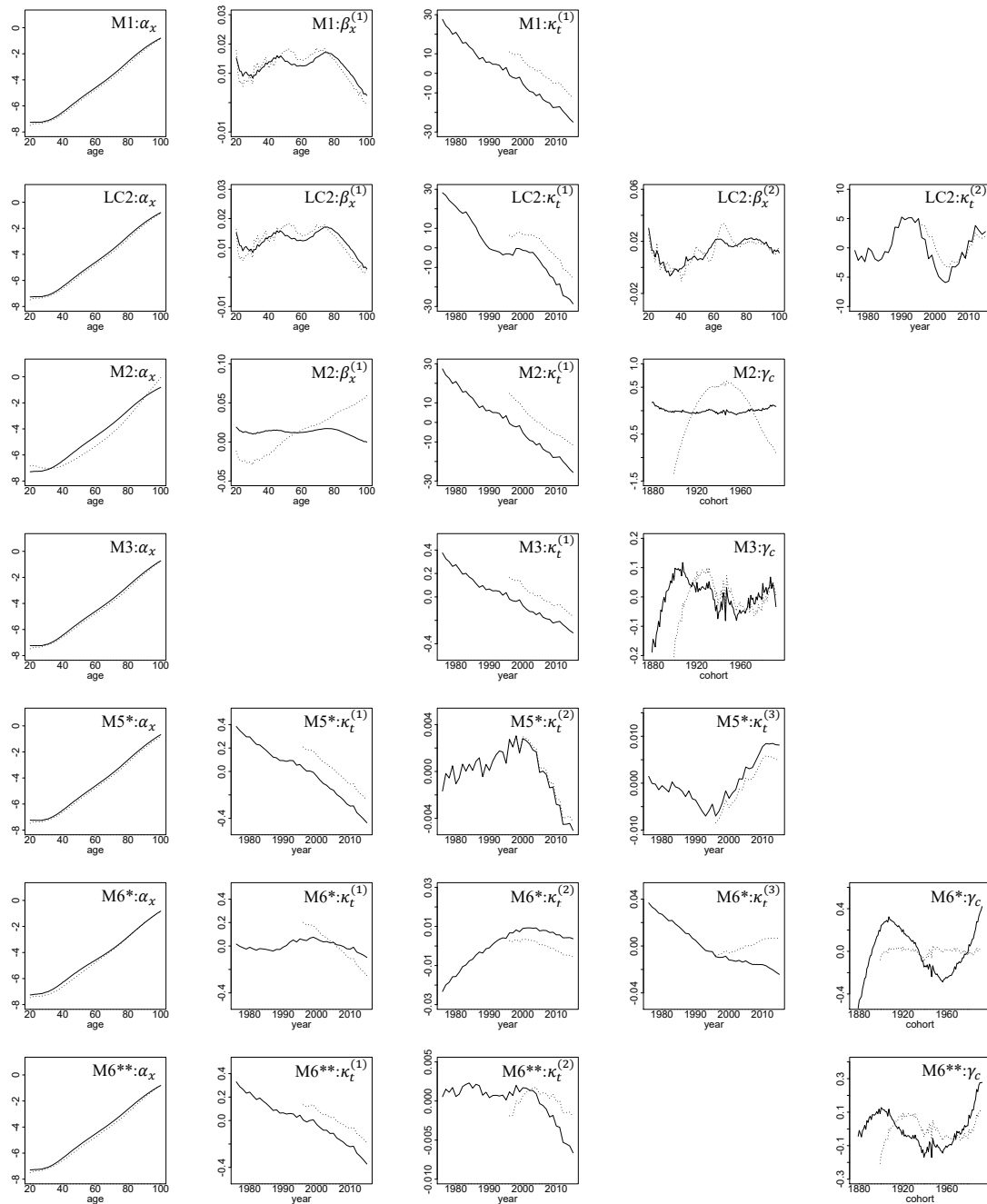


図7 日本人男性データに当てはめた各モデルのパラメータ推定値の比較：実線は元データを1976-2015年、点線は1996-2015年とした場合

ラフが1996年から2015年のデータにもとづくものであるが、パラメータ推定値で示されたようにM2、M6*はロバスト性に問題がある。M6**も少し結果に違いが見られる。バックテストによる予測性能を比較した結果は図9のとおりである。ここでは、元データの暦年範囲の長さを40年とし、1995年末から2014年末に将来推計を実施した場合の2015年の予測値に関する予測性能をプロットしている。(すなわち、元データの暦年範囲はそ

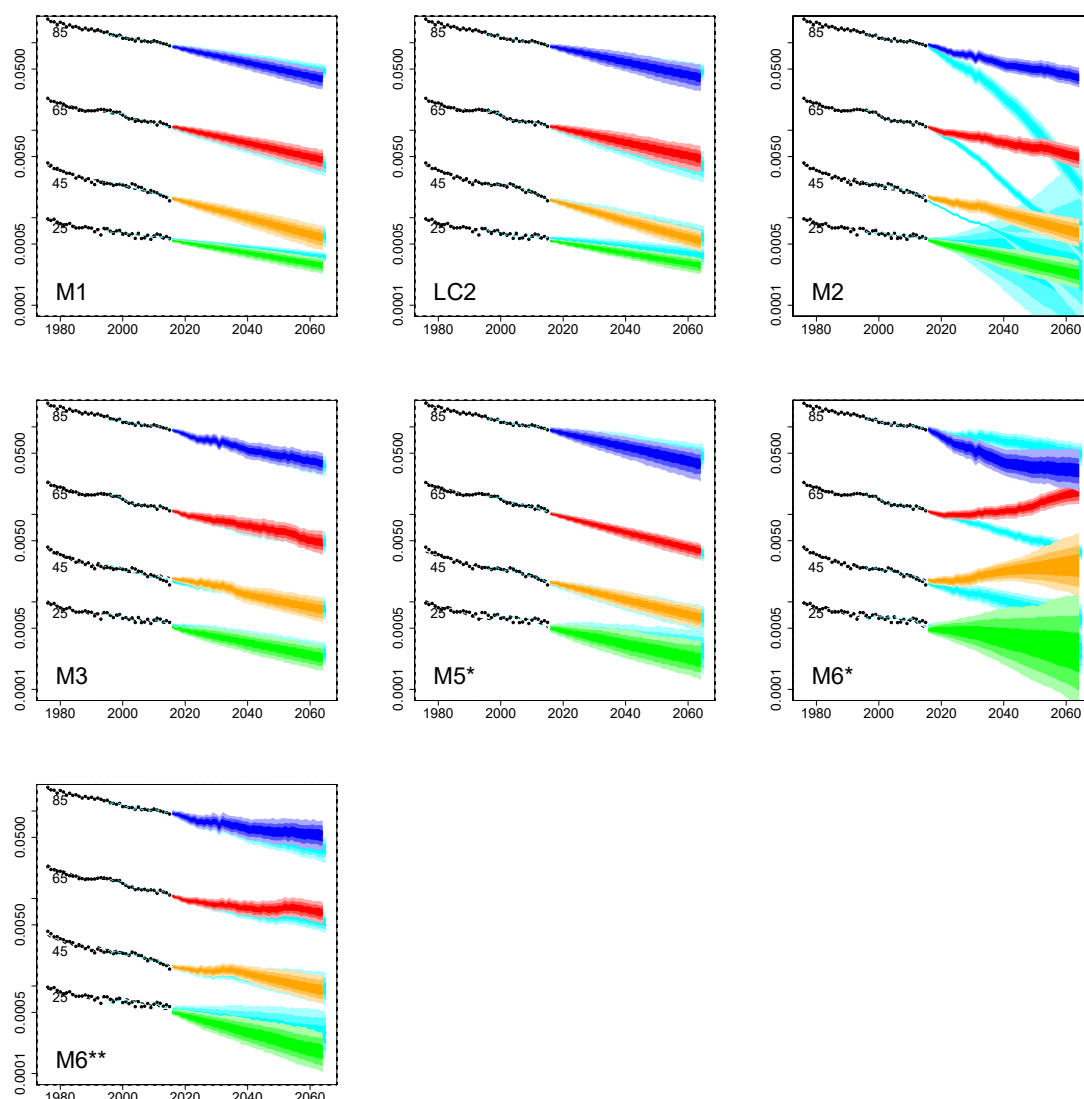


図 8 日本人男性の代表年齢別の将来死亡率のファンチャートの比較：元データを 1976-2015 年とした場合の死亡率を 1996-2015 年とした死亡率(水色)に重ねている(縦軸は対数)

れぞれ 1956-1995 年、1957-1996 年、…、1975-2014 年である。) 全体として、直近の予測になるにつれて性能改善が見られ、各モデルは予測実施時期毎の元データのの違いに対して一定程度ロバストであると解釈できる。ただし、M6**については、いずれの指標においてもグラフが不安定な変動を示しており、ロバスト性に問題があるものと考えられる。また、M6*の G_1 のグラフでは 2000 年頃まで予測の方が実績を上回っており、他のモデルと異なる傾向となっている。予測性能の水準に関しては、特に予測期間が長いほど、M1、LC2、M5*と比べて M2、M3、M6*の方が悪くなっている。これらはコホート項を含むモデルであることから、両グループの性能の差はコホート項の外挿に影響を受けているものと考えられる。

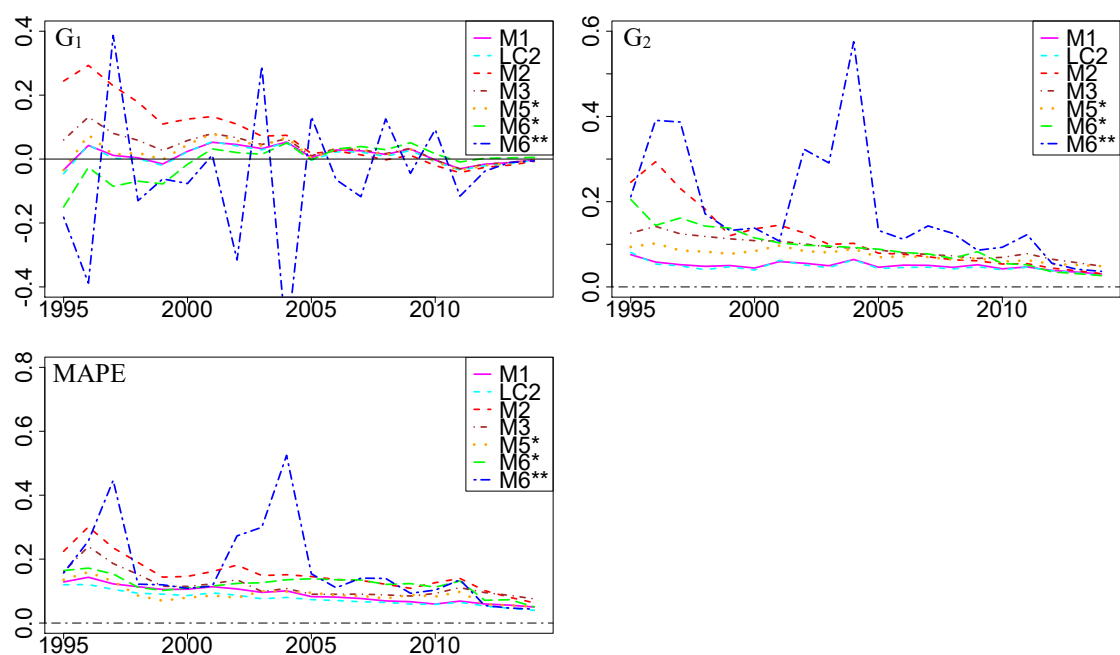


図 9 日本人男性データに当てはめた各モデルのバックテストにもとづくロバスト性の検証：3つのグラフは各グラフの左上に記載の評価指標に対応しており、横軸は元データのうちの直近の年を表している(元データの暦年範囲は左から順に 1956-1995 年、1957-1996 年、…、1975-2014 年)

以上をまとめたものを表 3 に整理する。これより、すべての基準で優位な結果を示すモデルはないものの、大きな問題が見られなかった M1、LC2、M3 をここでは選定し、参照母集団のモデルとする。(基準 1 は前述のとおりだが、参考のため記載している。)

表 3 日本人男性データに関する死亡率モデルの比較評価：基準 7,9 の列の(A,B)における A と B はそれぞれ BIC の順位と残差プロットの評価

参照 モデル	基準 1 (年齢間の相関)	基準 7,9 (適合度・節約性)	基準 10,11 (尤もらしさ)	基準 12 (ロバスト性)
M1	×	(5, △)	○	○
LC2	○	(3, △)	○	○
M2	×	(2, ○)	○	×
M3	×	(6, △)	○	○
M5*	○	(7, ×)	△	○
M6*	○	(1, ○)	×	×
M6**	○	(4, △)	△	×

表4：目標ポートフォリオの候補モデル

差に関するモデル ($\log m_{xt}^{\text{port}} - \log m_{xt}^{\text{ref}}$)	参照母集団のモデル		
	M1	LC2	M3
M0: α_x^{port}	✓	✓	✓
M1: $\alpha_x^{\text{port}} + \beta_x^{(1,\text{port})} \kappa_t^{(1,\text{port})}$	✓	✓	✓
LC2: $\alpha_x^{\text{port}} + \beta_x^{(1,\text{port})} \kappa_t^{(1,\text{port})} + \beta_x^{(2,\text{port})} \kappa_t^{(2,\text{port})}$		✓	
M3: $\alpha_x^{\text{port}} + \kappa_t^{(1,\text{port})} + \gamma_{t-x}^{\text{port}}$			✓
M1 ^(CAE) : $\alpha_x^{\text{port}} + \beta_x^{(1,\text{ref})} \kappa_t^{(1,\text{port})}$	✓	✓	✓
LC2 ^(CAE) : $\alpha_x^{\text{port}} + \beta_x^{(1,\text{ref})} \kappa_t^{(1,\text{port})} + \beta_x^{(2,\text{ref})} \kappa_t^{(2,\text{port})}$		✓	

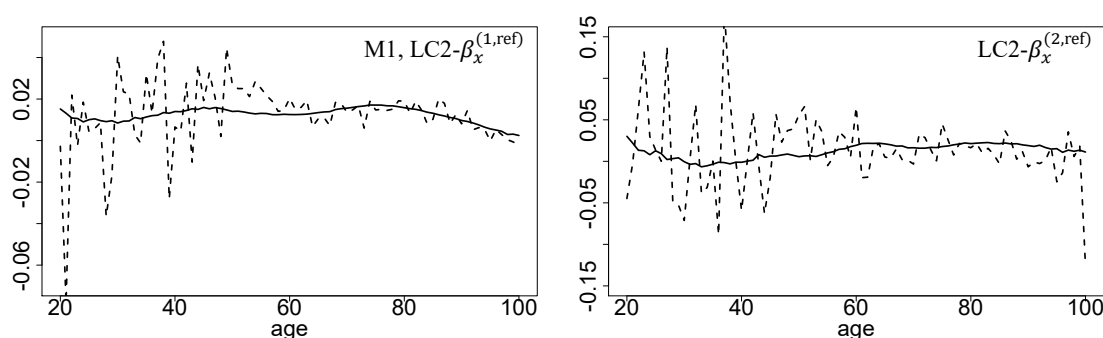


図10 国民全体に当てはめた場合(実線)と鳥取県の場合(破線)のパラメータ推定値

4. 4 目標ポートフォリオ母集団のモデリング

前節で選定したモデルに当てはめた参照母集団に係る死亡率のモデル値を所与のものとして、目標ポートフォリオのモデル検討を行う。2.3節で挙げた候補にしたがって、検討するモデルは表4のとおりとする。ただし、モデル比較の観点からM3-M1も対象としている。(参照母集団および目標ポートフォリオとの差について、それぞれM3、M1を適用する場合のモデルをM3-M1と表記する。他の組合せの表記についても同様とする。) 差のモデルとしてM1^(CAE)およびLC2^(CAE)を用いることの合理性を示すために、図10に国民全体と鳥取県データに対してM1、LC2を当てはめた場合の $\beta_x^{(1,\text{ref})}$ および $\beta_x^{(2,\text{ref})}$ を重ねてプロットする。鳥取県データでは大きなバラつきが見られるが、それらを除くと国民全体のパラメータ推定値と近い傾向があるように思われる。これより、M1^(CAE)とLC2^(CAE)とすることには一定の合理性があると考えられる。

各モデルのデータへの適合度等は表5のとおりである。参照母集団の場合と異なり、対数尤度と情報量規準の順位に相違があることから、差のモデルをM1、LC2、またはM3とすることは過剰なパラメータ数となっているものと考えられる。一方で、基準項のみのM0や共通の年齢調整項とすることは、パラメータ数を削減し予測においては望ましい選択であることがAICおよびBICによって示唆されている。

情報量規準に加えて、目標ポートフォリオの各年齢の死亡率から計算した生存率の実

表 5 各モデルの参照母集団データへの適合度と節約性

参照モデル	差分モデル	パラメータ数 [差分モデル]	対数尤度(順位) [差分モデル]	AIC(順位) [差分モデル]	BIC(順位) [差分モデル]
M1	M0	81	-4,644(3)	9,450(1)	9,886(1)
	M1	180	-4,564(1)	9,488(3)	10,457(3)
	M1 ^(CAE)	100	-4,630(2)	9,460(2)	9,999(2)
LC2	M0	81	-4,645(5)	9,452(1)	9,888(1)
	M1	180	-4,561(2)	9,483(4)	10,451(4)
	LC2	277	-4,484(1)	9,523(5)	11,014(5)
	M1 ^(CAE)	100	-4,631(4)	9,463(2)	10,001(2)
	LC2 ^(CAE)	119	-4,620(3)	9,478(3)	10,119(3)
M3	M0	81	-4,644(4)	9,451(1)	9,887(1)
	M1	180	-4,560(1)	9,481(3)	10,450(3)
	M3	192	-4,568(2)	9,520(4)	10,553(4)
	M1 ^(CAE)	100	-4,630(3)	9,461(2)	10,000(2)

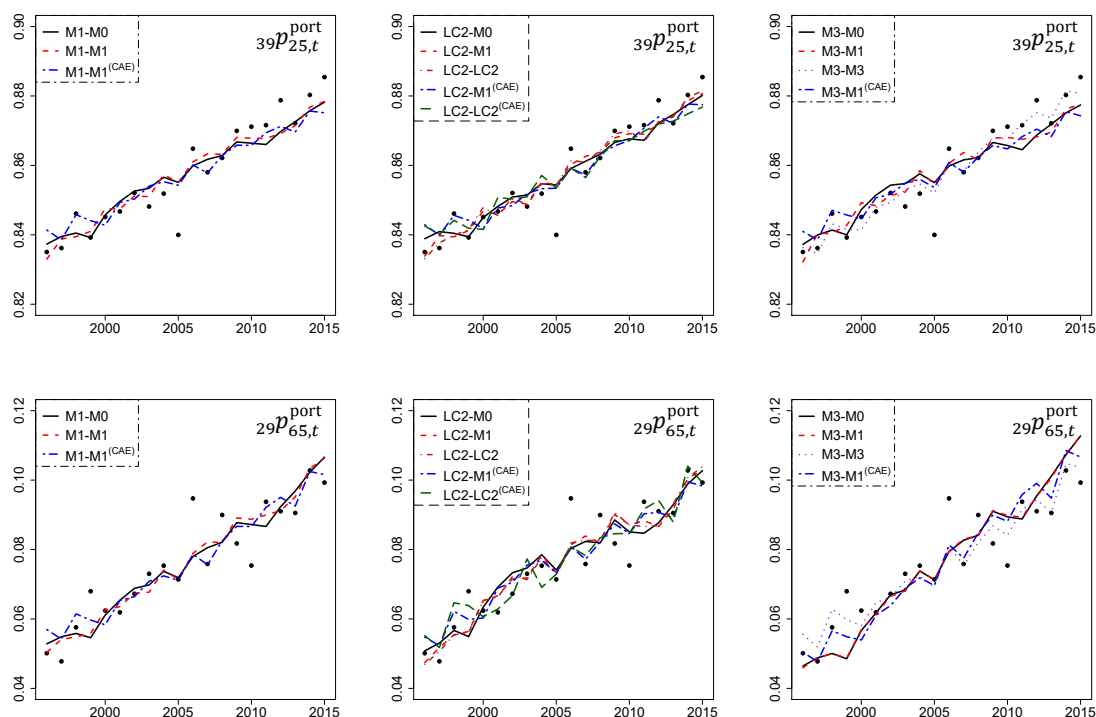


図 11 暦年別の目標ポートフォリオにおける生存率実績(黒点)とモデル当てはめ値(折れ線)の比較

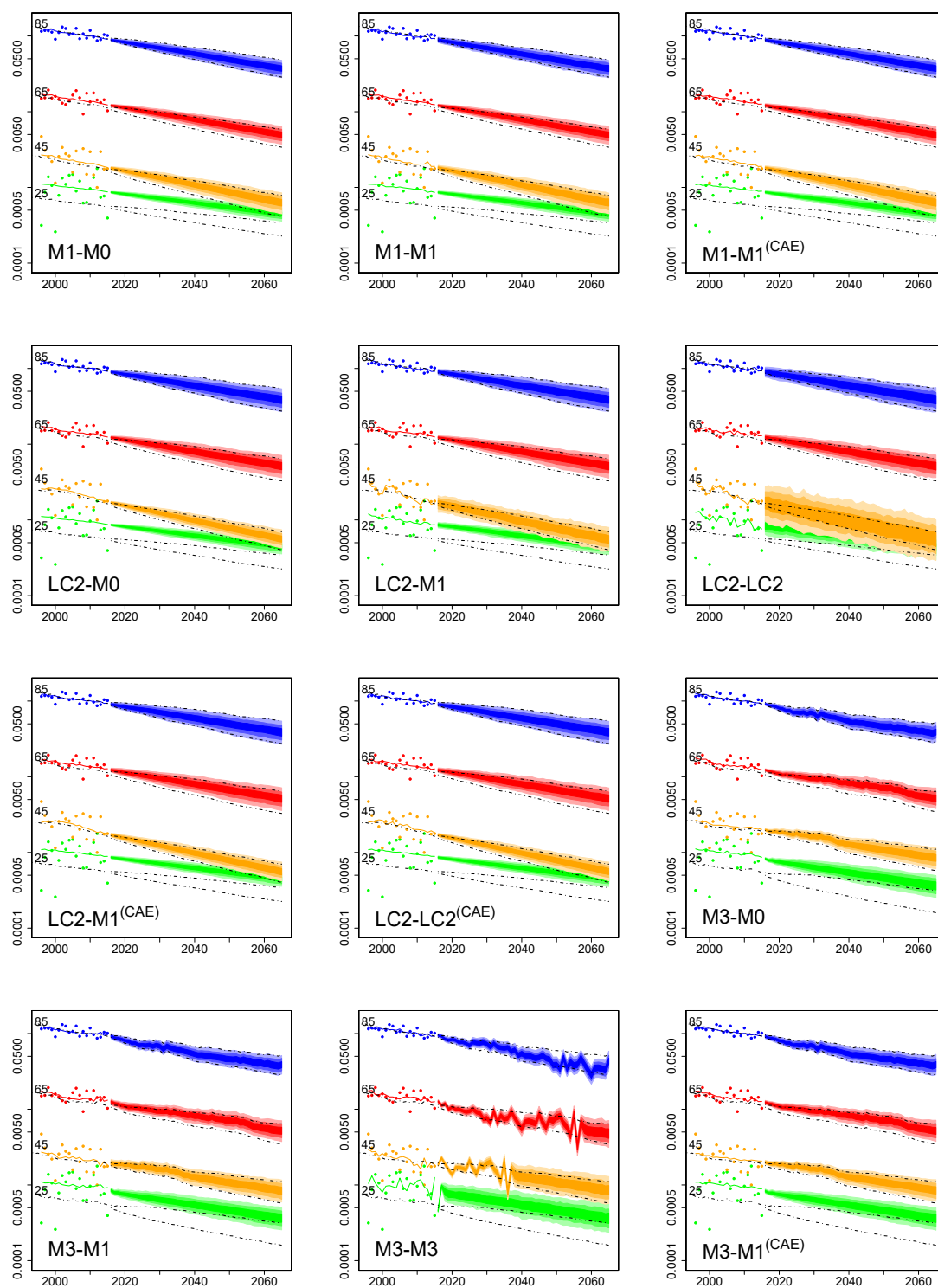


図 12 目標ポートフォリオの代表年齢別の将来死亡率のファンチャート(縦軸は対数)

績とモデル値を比較することにより、モデルの過去データへの当てはまりを確認する。図 11 の上段は ${}_{39}p_{25,t}^{\text{port}} = \prod_{x=25}^{64}(1 - q_{xt}^{\text{port}})$ 、下段は ${}_{29}p_{65,t}^{\text{port}} = \prod_{x=65}^{94}(1 - q_{xt}^{\text{port}})$ の 1996 年から 2015 年までの推移を表している。いずれのモデルにおいても実績によく適合していることが確認できる。なお、死亡率の変換には、 $q_{xt} = 1 - \exp(-m_{xt})$ の関係を用いている。

モデル別の目標ポートフォリオについての死亡率に関するファンチャートは図 12 のとおりである。グラフの一点鎖線は前節で算出した参照母集団の死亡率の 2015 年までのモデル値および将来の 95%予測区間を表す。グラフより、まず両母集団の死亡率水準差がモデルによって捉えられていることが確認できる。各モデルを比較すると、差のモデルが M1 および LC2 の場合、モデルが死亡率の変動を一定程度捉えられていることが分かる。ただし、LC2 については、将来死亡率の変動がやや過大となっているものと思われる。これは、時系列モデルで予測している $\kappa_t^{(j,\text{port})}$ がモデルに複数含まれているためと考えられる。M3-M3 において、著しい死亡率の変動が見られる。これは、図 16 の実線で示されている直近のコHORT項の形状が順次反映されているためである。これらに関しては、尤もらしさの観点から疑問が残るところである。その他については、これらのグラフからだけではモデルの違いによる影響は特段見られない。

特定コHORTの将来生存率推移のファンチャートは図 13 のとおりである。25 歳から 80 歳までの推移を示す上段の LC2-LC2 と M3-M3 の傾向が少し他と外れている他は、特に

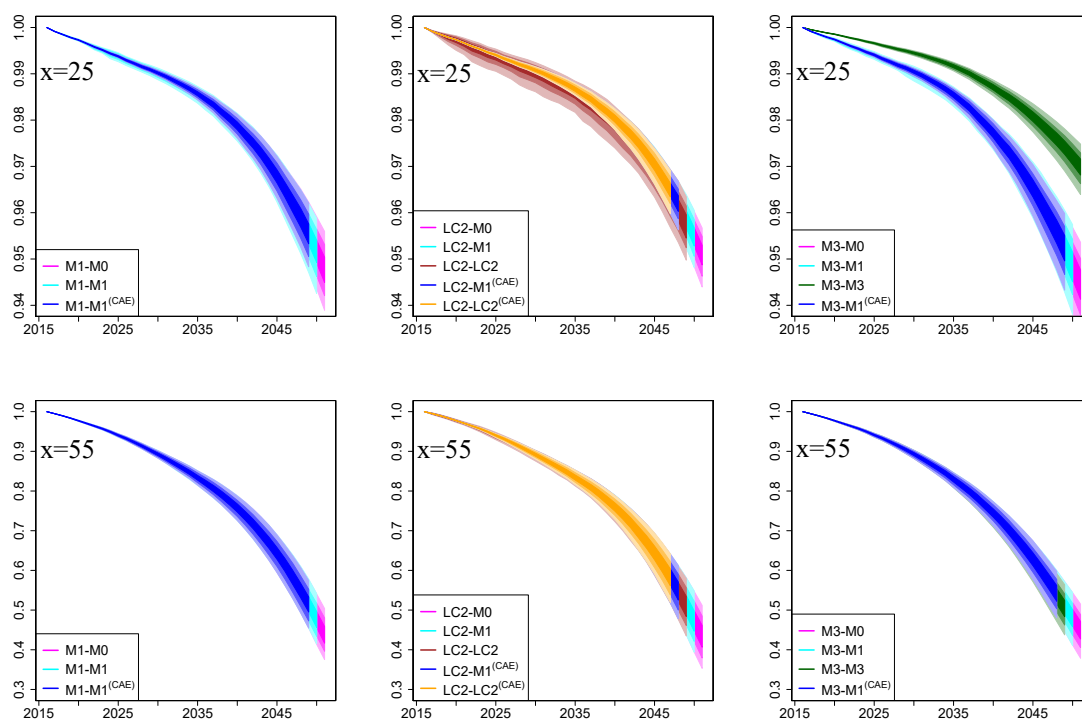


図 13 目標ポートフォリオのうち 2016 年に 25、55 歳のコHORTの 35 年間の生存率推定値のファンチャート

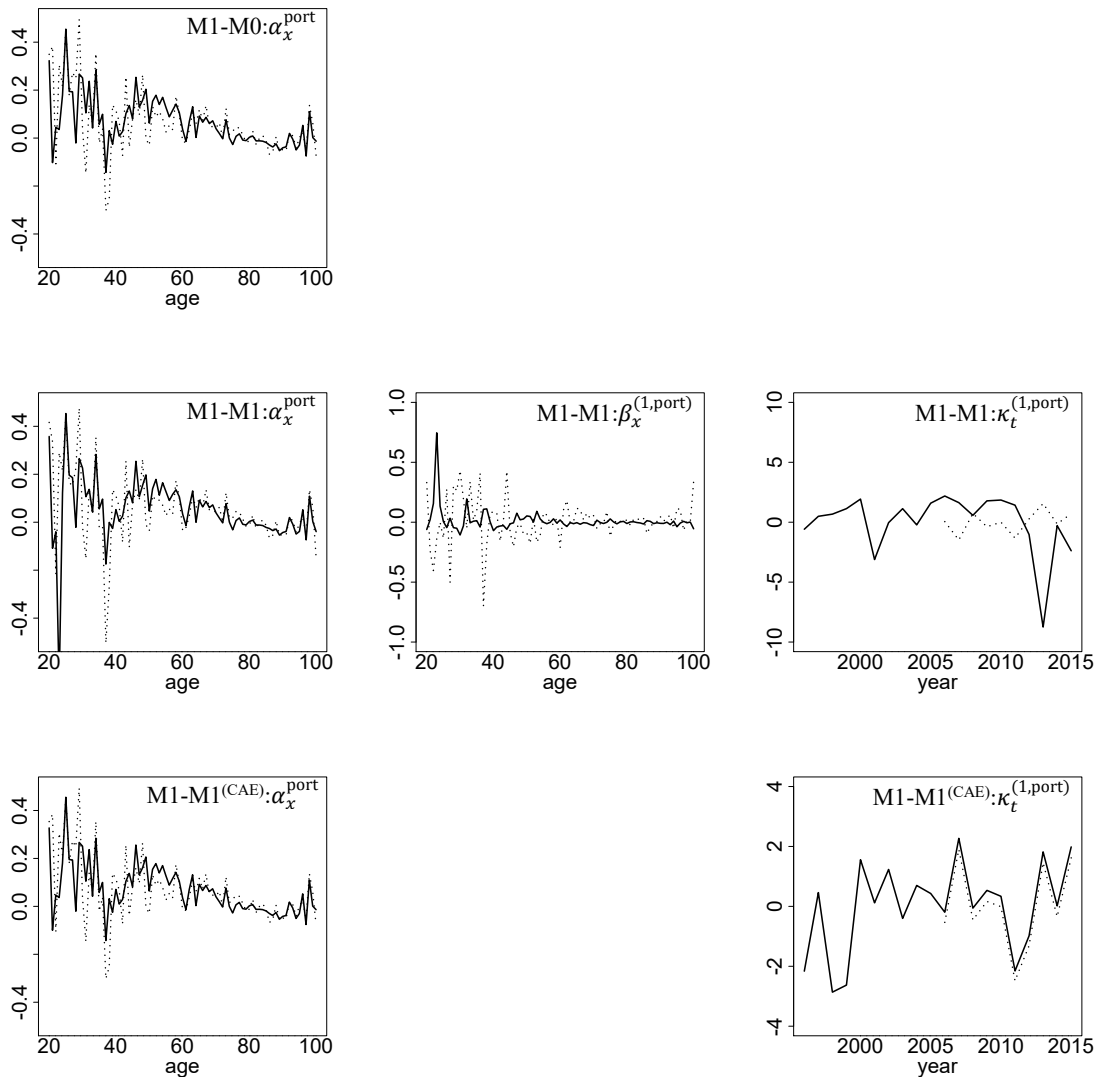


図 14 目標ポートフォリオのデータに当てはめた各モデルのパラメータ推定値の比較：実線は元データを 1996-2015 年、点線は 2006-2015 年とした場合

差異は見られない。差のモデルを M1 とした場合の将来死亡率の変動については、生存率を計算する過程で平均化されているものと考えられる。

元データを変更した場合のロバスト性を確認するため、図 14-16 に 2006 年から 2015 年のデータを用いた場合のパラメータ推定値と比較するためのグラフを示す。これらより、差のモデルを M1、LC2 としたものについてはロバスト性の問題が見られる。その他はロバスト性に関して特段の問題はない。図 17 では将来死亡率推移の比較を行っているが、ここでも差のモデルを M1、LC2 とした場合に変動の幅が元データによって変化していることが確認できる。一方、その他ではパラメータ推定値の比較と同様に問題は見られない。なお、これらのグラフから、適切なモデル選択のもとでは目標ポートフォリオのデータを 10 年に短縮しても都道府県データに関しては将来推計に大きな問題がないことも分かる。

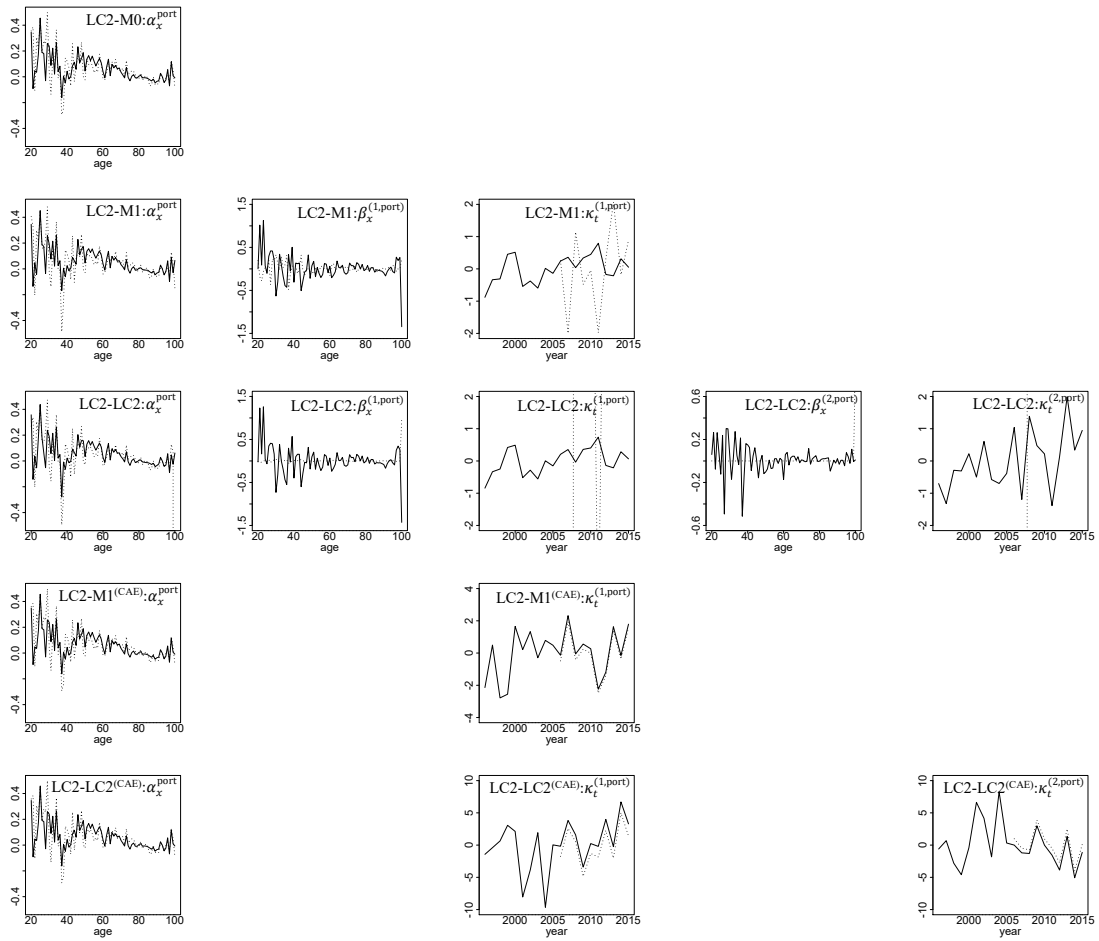


図 15 目標ポートフォリオのデータに当てはめた各モデルのパラメータ推定値の比較：実線は元データを 1996-2015 年、点線は 2006-2015 年とした場合

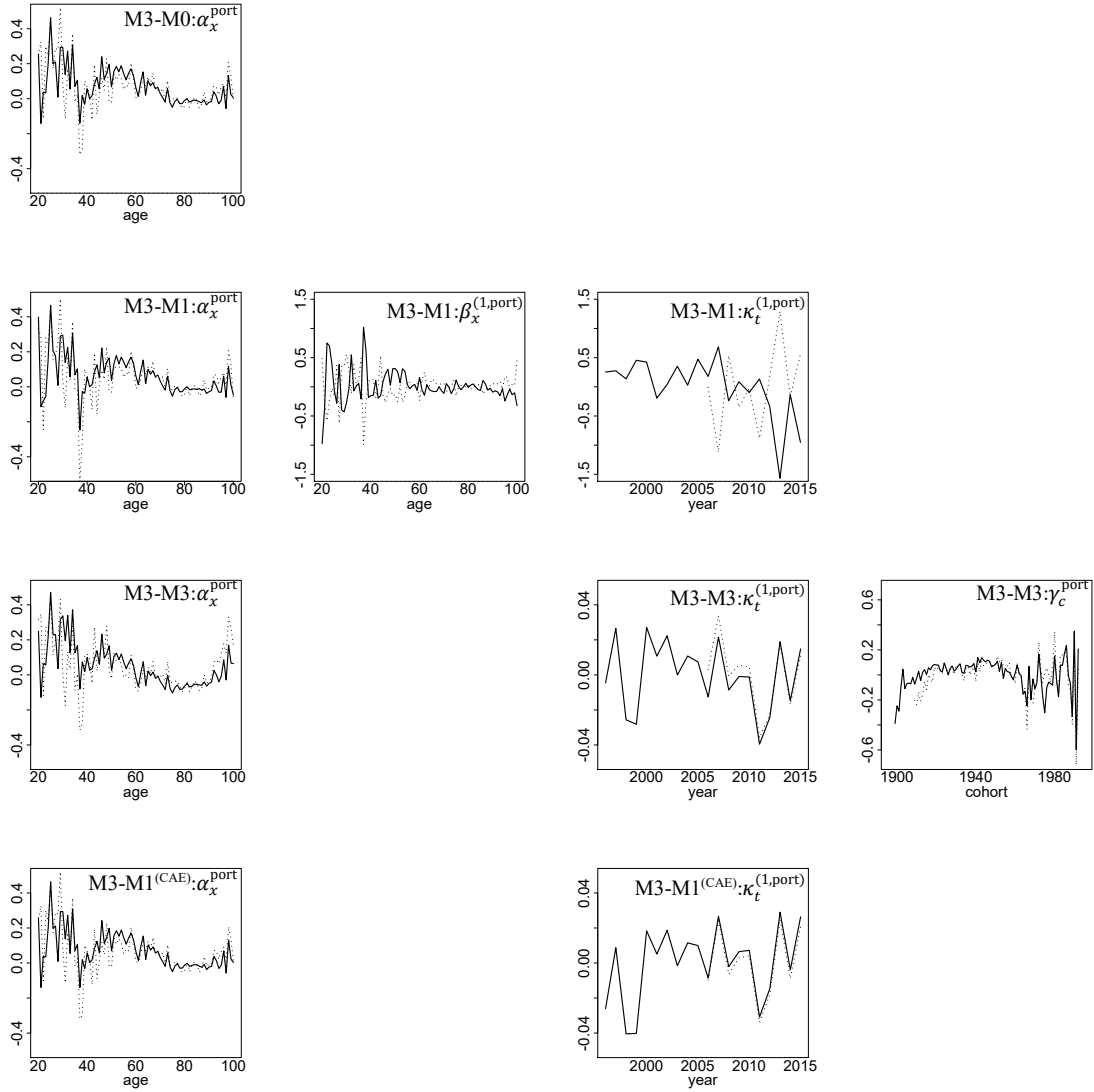


図 16 目標ポートフォリオのデータに当てはめた各モデルのパラメータ推定値の比較：
 実線は元データを 1996-2015 年、点線は 2006-2015 年とした場合

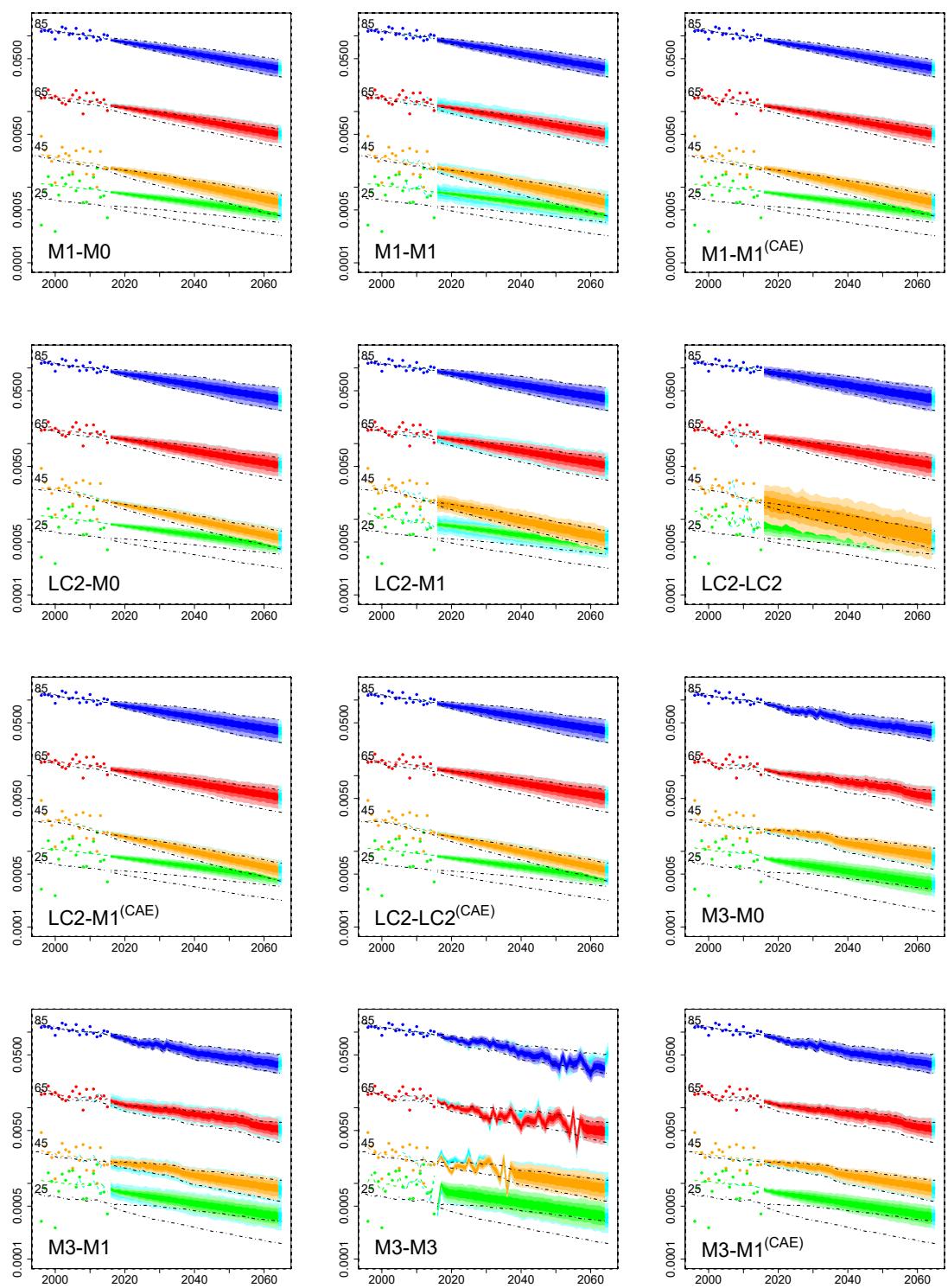


図 17 目標ポートフォリオの代表年齢別の将来死亡率のファンチャートの比較：元データを 1996-2015 年とした場合の死亡率を 2006-2015 年とした死亡率(水色)に重ねている(縦軸は対数)

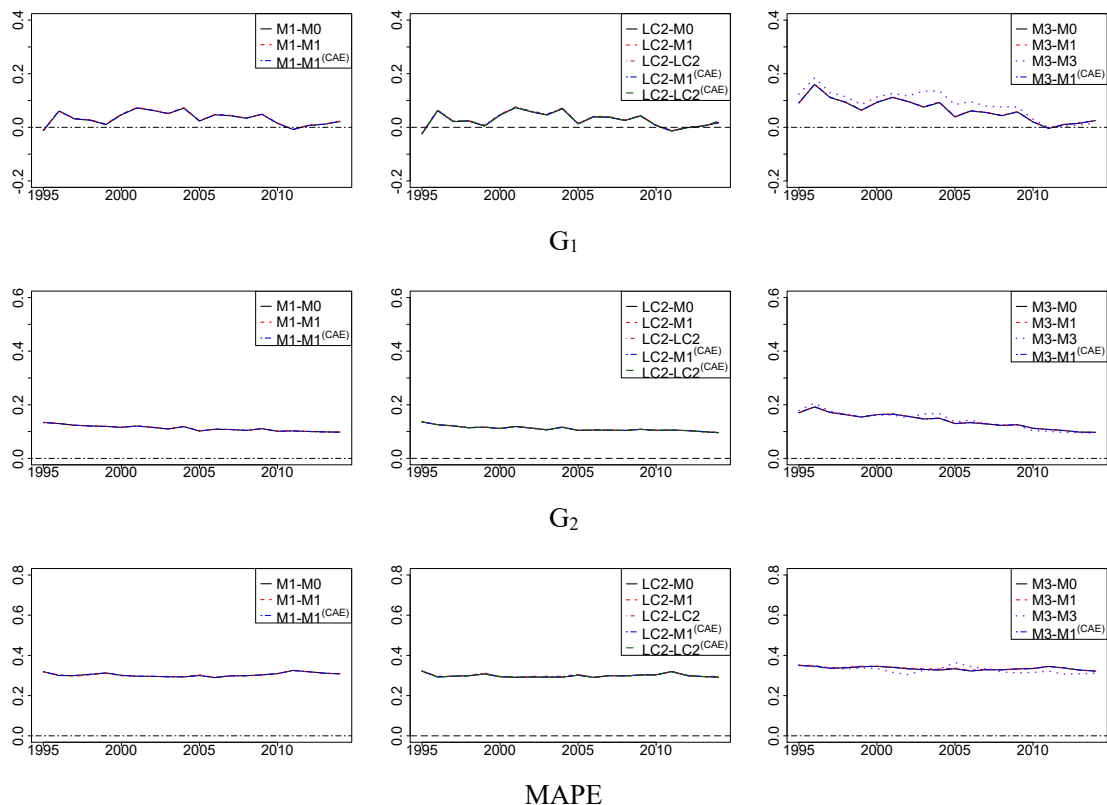


図 18 日本人男性データに当てはめた各モデルのバックテストにもとづくロバスト性の検証：グラフは上段が G_1 、中段が G_2 、下段が MAPE に対応しており、横軸は元データのうちの直近の年を表している

バックテストによる中心プロジェクションの予測性能を比較した結果を図 18 に示す。ここでは、前節と同様に 2015 年の死亡率の予測に関する予測性能を評価している。予測に用いる元データについては、参照母集団に関しては暦年範囲の長さを 40 年とし、目標ポートフォリオについては同範囲の後半の 20 年のデータを用いている。すなわち、1995 年末実施の将来推計における元データは、参照母集団については 1956-1995 年、目標ポートフォリオについては 1976-1995 年となり、それ以降の年では各範囲を 1 年ずつ進めたものとなる。図 18 の各グラフの結果を見ると、参照母集団のモデルが共通している場合は評価指標のプロットがほぼ一致しており、M3-M3 のみ同グラフ内の他モデルとプロットにずれが見られる。これは、本数値例で扱うモデルの範囲では、目標ポートフォリオと参照母集団の差分モデルが M3 の場合を除き、参照母集団のモデルが同じであれば、死亡率モデルは基本的に同様の中心プロジェクションを生成することを示唆している。なお、図 18 の各グラフの評価指標のプロットは、左の列から順にそれぞれ図 9 の M1、LC2、M3 のプロットを概ね平行移動させたようなものとなっている。これより評価指標は安定しており、中心プロジェクションに関してロバスト性に特段の問題はないものと考えられる。

表 6 鳥取県データに関する死亡率モデルの比較評価：基準 7-9 の列の(A,B)における A と B はそれぞれ参照モデルと差分モデルについての BIC の順位

参照 モデル	差分 モデル	基準 1 (年齢間の相関)	基準 7-9 (適合度・節約性)	基準 10,11 (尤もらしさ)	基準 12 (ロバスト性)
M1	M0	×	(5, 1)	○	○
	M1	△	(5, 3)	◎	×
	M1 ^(CAE)	×	(5, 2)	○	○
LC2	M0	○	(3, 1)	○	○
	M1	○	(3, 4)	◎	×
	LC2	○	(3, 5)	△	×
	M1 ^(CAE)	○	(3, 2)	○	○
	LC2 ^(CAE)	○	(3, 3)	○	○
M3	M0	×	(6, 1)	○	○
	M1	△	(6, 3)	◎	×
	M3	×	(6, 4)	×	△
	M1 ^(CAE)	×	(6, 2)	○	○

4. 5 総括

これまでの議論をまとめた鳥取県データに関する死亡率モデルの比較評価を表 6 に整理する。すべての基準で優位な結果を示すものはないものの、LC2-M0、LC2-M1^(CAE)および LC2-LC2^(CAE)が総合的に優れたモデルと評価できる。なお、これらは先行研究で扱われていないモデルである。過去実績における死亡率の変動の捕捉性能が LC2-M1 とした場合より劣るものの、その他の基準についてはいずれも良好な結果となっている。ベンチマークとして代表的な multi population model である Li and Lee [2005]の改良共通要因モデル (Augmented Common Factor model)に該当する M1-M1 と比較すると、やはり死亡率の変動が小さく、また年齢毎の死亡率に差異が見られることから、モデルリスクの影響には注意をする必要がある(図 13)。

これら 3 つのうちいずれを採用するかは、モデルの利用目的によって異なるものと考えられる。モデル検討の初期段階では、相対的に節約的な LC2-M0 を用いることが基本となる。特に、死亡率に関してオプション性等を考慮する必要がない前提で保険負債の最良推計を計算する場合には、LC2-M0 の中心プロジェクションを利用すれば十分であろう。一方、リスク評価等においては、LC2-M1^(CAE)および LC2-LC2^(CAE)を用いた関係指標に関する計算をあわせて行い、モデルによる影響を確認する必要がある。また、複数の保険契約ポートフォリオの計算結果を統合して全体的なリスク評価等を行う場合には、LC2-M0 ではポートフォリオ毎のトレンドの違いを考慮することができず、ポートフォリオ間で死亡率変動に完全な相関が生じることに注意する必要がある。このため、相関のモデリン

グに求められる精緻さ等に応じて、LC2-M1^(CAE)および LC2-LC2^(CAE)を用いる必要性を検討することが考えられる。LC2-M1^(CAE)と LC2-LC2^(CAE)の比較を行う上では、前述のリスク指標における数値的な影響の他、時系列モデルの複雑度の違いも重要となる。利用可能なデータの年数によっては $\kappa_t^{(1,port)}$ と $\kappa_t^{(2,port)}$ の間の相関をそもそも捉えられないといったことが考えられる。なお、保険会社データにおいては、被保険者群団という特殊性から $\beta_x^{(i,ref)}$ に関して共通年齢効果が見られない可能性もある。

上記の将来死亡率推計においては簡便的にプロセスリスクのみを加味したシミュレーションとしているが、ブートストラップ法を用いることによりパラメータリスクを織り込むことも可能である。特に、小さい母集団を扱う際には信頼区間および予測区間の計算において一定の影響を与えることが想定されるため、実務においてモデル検討を行う際には考慮することが望ましい。代表的な手法については、Brouhns et al. [2005]、Koissi et al. [2006]、Renshaw and Haberman [2008]を参照されたい。

第5章 結論

本稿では、リスク管理およびソルベンシー評価のための内部モデル等への応用を目的とした relative approach にもとづく保険契約ポートフォリオの将来死亡率推計モデルの設計手法の基礎的な検討を行った。数値例として、比較的規模が小さい都道府県別人口の将来死亡率推計問題を扱い、最後に各モデルの比較評価、および総合的に他のモデルよりも優位性を示した LC2-M0、LC2-M1^(CAE)、LC2-LC2^(CAE)のモデル選択に関する議論を行った。これらのモデルは著者の知る限り multi population model に関する先行研究で扱われておらず、また近藤[2013]で示されているように LC2 は日本の死亡率のモデリングにおいて有用と考えられることから、更なる研究の余地がある。

特に、本稿では参照母集団に関して時系列モデルを多変量ドリフト項付ランダムウォークとするなど簡便的なモデリングを行っているが、これはモデルの比較評価のために前提を揃えているためであり、実務に応用する際にはデータを入念に分析した上で追加のモデル修正を行うことが考えられる。例えば、日本国民の死亡率のモデリングに際しては、小松[2002]、石井[2008, 2013]で Lee-Carter モデルに対して適用されているモデル修正の考え方が基本となるであろう。将来推計人口における将来生命表の作成では、死亡率改善が緩やかになるという仮定のもと、時系列モデルによる外挿に代わり、非線形関数の当てはめが用いられている。この方法はサンプルパスを生成することができないため本稿のモデル選択基準を満たさないが、日本における保険契約ポートフォリオの将来死亡率推計モデルの構築においては、同様の仮定のもとサンプルパスを生成可能な外挿方法の検討の余地があると考えられる。

その他の将来研究の方向性としては、保険会社データにもとづく手法の有効性の検証や、本手法の適用が適切でないボリュームおよび年数のデータに対するモデリング手法の検討が挙げられる。(後者のデータに関する考察は Villegas et al. [2017]を参照されたい。)

参考文献

- [1] 石井太 [2008], 「近年のわが国の死亡動向に適合した将来生命表推計モデルの研究－年齢シフトモデルの開発－」, 人口問題研究, 64(3): 28-44
- [2] 石井太 [2013], 「死亡率曲線の自由な方向への変化を表現する数理モデルとわが国の将来生命表への応用」, 人口問題研究, 69(3): 3-26
- [3] 国立社会保障・人口問題研究所 [2017], 「日本の将来推計人口（平成 29 年推計）報告書」, 人口問題研究資料, 336
- [4] 国立社会保障・人口問題研究所, 「日本版死亡データベース」, <http://www.ipss.go.jp/p-toukei/JMD/index.asp> (2017 年 10 月 25 日ダウンロード)
- [5] 小松隆一 [2002], 「リレーショナル・モデルによる日本の将来生命表作成の試み」, 人口問題研究, 58(3): 3-14
- [6] 近藤勇志 [2013], 「死亡率の将来推定手法に関する一考察 ～第 2 因子を用いた Lee-Carter モデル～」, 日本アクチュアリー会会報, 66: 204-233
- [7] Booth, H., Maindonald, J. and Smith, L. [2002], “Applying Lee-Carter under conditions of variable mortality decline”, *Population Studies*, 56(3): 325-336
- [8] Booth, H. and Tickle, L. [2008], “Mortality modelling and forecasting: A review of methods”, *Annals of Actuarial Science*, 3: 3-43
- [9] Brouhns, N., Denuit, M. and Van Keilegom, I. [2005], “Bootstrapping the Poisson log-bilinear model for mortality forecasting”, *Scandinavian Actuarial Journal*, (3): 212-224
- [10] 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
- [11] Cairns, A. J. G., Blake, D. and Dowd, K. [2006], “Two-factor model for stochastic mortality with parameter uncertainty: Theory and calibration”, *Journal of Risk and Insurance*, 73(4): 687-718
- [12] Cairns, A. J. G., Blake, D., Dowd, K., Coughlan, G. D., Epstein, D. and Khalaf-Allah, M. [2011], “Mortality density forecasts: An analysis of six stochastic mortality models”, *Insurance: Mathematics and Economics*, 48(3): 355-367
- [13] Cairns, A. J. G., Blake, D., Dowd, K., Coughlan, G. D., Epstein, D., Ong, A. and Balevich, I. [2009], “A quantitative comparison of stochastic mortality models using data from England and Wales and the United States”, *North American Actuarial Journal*, 13(1): 1-35
- [14] Carter, L. R. and Lee, R. D. [1992], “Modeling and forecasting US sex differentials in mortality”, *International Journal of Forecasting*, 8(3): 393-411
- [15] Currie, I. D., Durban, M. and Eilers, P. H. [2004], “Smoothing and forecasting mortality rates”, *Statistical Modelling*, 4(4): 279-298
- [16] Dowd, K., Cairns, A. J. G., Blake, D., Coughlan, G. D., Epstein, D. and Khalaf-Allah, M. [2010], “Backtesting stochastic mortality models: An ex-post evaluation of multi-period-

- ahead density forecasts”, *North American Actuarial Journal*, 14(3): 281-298
- [17] Dowd, K., Cairns, A. J. G., Blake, D., Coughlan, G. D. and Khalaf-Allah, M. [2011], “A gravity model of mortality rates for two related populations”, *North American Actuarial Journal*, 15(2): 334-356
- [18] Haberman, S., Kaishev, V. K., Millosovich, P., Villegas, A. M., Baxter, S., Gaches, A., Gunnlaugsson, S. and Sison, M. [2014], “Longevity basis risk: A methodology for assessing basis risk”, *Institute and Faculty of Actuaries Sessional Paper*
- [19] Haberman, S. and Renshaw, A. [2011], “A comparative study of parametric mortality projection models”, *Insurance: Mathematics and Economics*, 48(1): 35-55
- [20] Hunt, A. and Blake, D. [2014], “A general procedure for constructing mortality models”, *North American Actuarial Journal*, 18(1): 116-138
- [21] Hunt, A. and Blake, D. [2015], “On the structure and classification of mortality models”, *Pension Institute Working Paper*
- [22] Hunt, A. and Villegas, A. M. [2015], “Robustness and convergence in the Lee-Carter model with cohorts”, *Insurance: Mathematics and Economics*, 64: 186-202
- [23] Jarner, S. F. and Kryger, E. M. [2011], “Modelling adult mortality in small populations: The SAINT model”, *ASTIN Bulletin*, 41(2): 377-418
- [24] Kleinow, T. [2015], “A common age effect model for the mortality of multiple populations”, *Insurance: Mathematics and Economics*, 63: 147-152
- [25] Koissi, M.-C., Shapiro, A., and Hognas, G. [2006], “Evaluating and extending the Lee-Carter model for mortality forecasting: Bootstrap confidence interval”, *Insurance: Mathematics and Economics*, 38(1): 1-20
- [26] Lee, R. D. [2000], “The Lee-Carter method for forecasting mortality, with various extensions and applications”, *North American Actuarial Journal*, 4(1): 80-93
- [27] Lee, R. D. and Carter, L. R. [1992], “Modeling and forecasting U. S. mortality”, *Journal of the American Statistical Association*, 87(419): 659-671
- [28] Li, J. S.-H., Zhou, R. and Hardy, M. [2015], “A step-by-step guide to building two-population stochastic mortality models”, *Insurance: Mathematics and Economics*, 63: 121-134
- [29] Li, N. and Lee, R. D. [2005], “Coherent mortality forecasts for a group of populations: An extension of the Lee-Carter method”, *Demography*, 42(3): 575-594
- [30] Plat, R. [2009], “On stochastic mortality modeling”, *Insurance: Mathematics and Economics*, 45(3): 393-404
- [31] Renshaw, A. and Haberman, S. [2003], “Lee-Carter mortality forecasting with age-specific enhancement”, *Insurance: Mathematics and Economics*, 33(2), 255-272
- [32] Renshaw, A. 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
- [33] Renshaw, A. and Haberman, S. [2008], “On simulation-based approaches to risk measurement

in mortality with specific reference to Poisson Lee-Carter modelling”, *Insurance: Mathematics and Economics*, 42(2): 797-816

- [34] Villegas, A. M., Haberman, S., Kaishev, V. K. and Millosovich, P. [2017], “A comparative study of two population models for the assessment of basis risk in longevity hedges”, *ASTIN Bulletin*, 47(3): 631-679
- [35] Villegas, A. M., Kaishev, V. and Millosovich, P. [2015], “StMoMo: An R package for stochastic mortality modelling”, *7th Australasian Actuarial Education and Research Symposium*
- [36] Wan, C. and Bertschi, L. [2015], “Swiss coherent mortality model as a basis for developing longevity de-risking solutions for Swiss pension funds: A practical approach”, *Insurance: Mathematics and Economics*, 63: 66-75

付録 識別問題への対応

第2章のCBD派生モデルおよび混合モデルの識別問題への対応のために設定する制約条件、および当該制約条件を課すためのパラメータの変換を以下に示す。

- M6 モデル :

モデルの識別性は次を制約条件とすることで確保できる。

$$\sum_c \gamma_c = 0, \sum_c c\gamma_c = 0 \quad (41)$$

最小二乗法で $\gamma_{t-x} = \varphi_1 + \varphi_2(t-x)$ の線形回帰を行うことで得られる φ_1, φ_2 をパラメータとした変換

$$\left(\kappa_t^{(1)}, \kappa_t^{(2)}, \gamma_{t-x} \right) \rightarrow \left(\kappa_t^{(1)} + \varphi_1 + \varphi_2(t - \bar{x}), \kappa_t^{(2)} - \varphi_2, \gamma_{t-x} - \varphi_1 - \varphi_2(t - x) \right) \quad (42)$$

によって、この制約条件を課すことができる。

- M7 モデル :

モデルの識別性は次を制約条件とすることで確保できる。

$$\sum_c \gamma_c = 0, \sum_c c\gamma_c = 0, \sum_c c^2\gamma_c = 0 \quad (43)$$

最後の制約条件は γ_c に2次のトレンドが包含されないとするものである。最小二乗法で $\gamma_{t-x} = \varphi_1 + \varphi_2(t-x) + \varphi_3(t-x)^2$ の多項式回帰を行うことで得られる $\varphi_1, \varphi_2, \varphi_3$ をパラメータとした変換

$$\left(\kappa_t^{(1)}, \kappa_t^{(2)}, \kappa_t^{(3)}, \gamma_{t-x}\right) \rightarrow \left(\kappa_t^{(1)} + \varphi_1 + \varphi_2(t - \bar{x}) + \varphi_3\{(t - \bar{x})^2 + \hat{\sigma}_x^2\}, \right. \\ \left. \kappa_t^{(2)} - \varphi_2 - 2\varphi_3(t - \bar{x}), \kappa_t^{(3)} + \varphi_3, \gamma_{t-x} - \varphi_1 - \varphi_2(t - x) - \varphi_3(t - x)^2\right) \quad (44)$$

によって、この制約条件を課すことができる。

- M8 モデル :

モデルの識別性は次を制約条件とすることで確保できる。

$$\sum_c \gamma_c = 0 \quad (45)$$

この制約条件は

$$c_1 = \frac{1}{n+k-1} \sum_c \gamma_c \quad (46)$$

とした変換

$$\left(\kappa_t^{(1)}, \kappa_t^{(2)}, \gamma_{t-x}\right) \rightarrow \left(\kappa_t^{(1)} + c_1(x_c - \bar{x}), \kappa_t^{(2)} - c_1, \gamma_{t-x} - c_1\right) \quad (47)$$

によって、課すことができる。

- M5*モデル :

モデルの識別性は次を制約条件とすることで確保できる。

$$\sum_t \kappa_t^{(1)} = 0, \sum_t \kappa_t^{(2)} = 0, \sum_t \kappa_t^{(3)} = 0 \quad (48)$$

この制約条件は

$$c_1 = \frac{1}{n} \sum_t \kappa_t^{(1)}, c_2 = \frac{1}{n} \sum_t \kappa_t^{(2)}, c_3 = \frac{1}{n} \sum_t \kappa_t^{(3)} \quad (49)$$

とした変換

$$\left(\alpha_x, \kappa_t^{(1)}, \kappa_t^{(2)}, \kappa_t^{(3)}\right) \quad (50)$$

$$\rightarrow \left(\alpha_x + c_1 + c_2(\bar{x} - x) + c_3(\bar{x} - x)^+, \kappa_t^{(1)} - c_1, \kappa_t^{(2)} - c_2, \kappa_t^{(3)} - c_3\right)$$

によって、課すことができる。

- M6*モデル :

モデルの識別性は次を制約条件とすることで確保できる。

$$\sum_t \kappa_t^{(1)} = 0, \sum_t \kappa_t^{(2)} = 0, \sum_t \kappa_t^{(3)} = 0, \sum_c \gamma_c = 0, \sum_c c\gamma_c = 0, \sum_c c^2\gamma_c = 0 \quad (51)$$

最小二乗法で $\gamma_{t-x} = \varphi_1 + \varphi_2(t-x) + \varphi_3(t-x)^2$ の多項式回帰を行うことで得られる $\varphi_1, \varphi_2, \varphi_3$ をパラメータとした変換

$$\begin{aligned} (\alpha_x, \kappa_t^{(1)}, \kappa_t^{(2)}, \kappa_t^{(3)}, \gamma_{t-x}) &\rightarrow (\alpha_x + \varphi_1 - \varphi_2x + \varphi_3x^2, \kappa_t^{(1)} + \varphi_2t + \\ &\varphi_3(t^2 - 2\bar{x}t), \kappa_t^{(2)} + 2\varphi_3t, \kappa_t^{(3)}, \gamma_{t-x} - \varphi_1 - \varphi_2(t-x) - \varphi_3(t-x)^2) \end{aligned} \quad (52)$$

および(49)と同様に計算した c_1, c_2, c_3 をパラメータとした変換

$$\begin{aligned} (\alpha_x, \kappa_t^{(1)}, \kappa_t^{(2)}, \kappa_t^{(3)}, \gamma_{t-x}) \\ \rightarrow (\alpha_x + c_1 + c_2(\bar{x} - x) + c_3(\bar{x} - x)^2, \kappa_t^{(1)} - c_1, \kappa_t^{(2)} - c_2, \kappa_t^{(3)} - c_3, \gamma_{t-x}) \end{aligned} \quad (53)$$

によって、この制約条件を課すことができる。

● M6**モデル：

モデルの識別性は次を制約条件とすることで確保できる。

$$\sum_t \kappa_t^{(1)} = 0, \sum_t \kappa_t^{(2)} = 0, \sum_c \gamma_c = 0, \sum_c c\gamma_c = 0 \quad (54)$$

最小二乗法で $\gamma_{t-x} = \varphi_1 + \varphi_2(t-x)$ の線形回帰を行うことで得られる φ_1, φ_2 をパラメータとした変換

$$\begin{aligned} (\alpha_x, \kappa_t^{(1)}, \kappa_t^{(2)}, \gamma_{t-x}) \\ \rightarrow (\alpha_x, \kappa_t^{(1)} + \varphi_1 + \varphi_2(t - \bar{x}), \kappa_t^{(2)} + \varphi_2, \gamma_{t-x} - \varphi_1 - \varphi_2(t-x)) \end{aligned} \quad (55)$$

および c_1, c_2 を

$$c_1 = \frac{1}{n} \sum_t \kappa_t^{(1)}, c_2 = \frac{1}{n} \sum_t \kappa_t^{(2)} \quad (56)$$

とした変換

$$(\alpha_x, \kappa_t^{(1)}, \kappa_t^{(2)}, \gamma_{t-x}) \rightarrow (\alpha_x + c_1 + c_2(\bar{x} - x), \kappa_t^{(1)} - c_1, \kappa_t^{(2)} - c_2, \gamma_{t-x}) \quad (57)$$

によって、この制約条件を課すことができる。

(住友生命保険相互会社 主計部数理室)

Mortality projections for portfolio of insurance contracts with reference population data: Two population model approach

Kazunori Nishio

Abstract

Lee-Carter model, which is one of typical stochastic mortality models, is not directly applicable to insurance portfolios due to insufficient experience data volume of insurance companies in many cases. Thus, it is a conventional method to apply portfolio experience factors to projected mortality rates of the country population for solving insurance portfolio mortality projection problems. This paper investigates the model design for insurance portfolio mortality projections based on the relative approach of two population models. We select models for reference population from those featured in several papers including Haberman and Renshaw [2011] and models for the difference between small population and reference population from typical previous research. For these compounded candidate models, comparative assessment is made based on the proposed model selection criteria. We consider a mortality projection problem for population of administrative divisions of Japan as an application example to compare each model. The approach investigated in this paper is expected to be useful in preparing for the assumptions of future mortality rates in risk management and internal models for solvency assessment.

Keywords: mortality risk, longevity risk, stochastic mortality models, multi population models, insurance companies, risk management, internal models, data of administrative divisions of Japan