

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～

近藤 勇志

<要旨>

昨今、保険業界では、国際会計基準の導入や経済価値ベースのソルベンシー規制など様々な場面において、資産や負債を経済価値ベースで評価することが求められるようになった。その負債の評価に大きな影響を与える死亡率についても経済価値ベースで評価することが求められるようになったため、死亡率の将来を推定するモデルの必要性が高まっている。

死亡率の将来推定手法としては、各国のアクチュアリーや人口統計の専門家などにおいて様々な研究が行われており、米国において考案された Lee-Carter モデル (1992) がそのモデルの簡便性から、日本をはじめ各国で広く研究が行われている。日本においても小松 (2002) や及川 (2006)、石井 (2006) (2008) などが Lee-Carter モデルを一部改定して日本の死亡率に適用したモデルを公表しており、小松や石井の死亡率推計モデルは国立社会保障・人口問題研究所が発表している「日本の将来推計人口」においても使用されている。

経済価値ベースの死亡率推計においてもそれらの推計手法を用いることが考えられるが、保険会社の経験死亡率を使って自社の状況をより適切に反映させようとした場合、米国や日本といった国民全体の死亡率データと比べて、相対的にデータ量が少ないことや観測期間の短さから、Lee-Carter モデルをそのまま適用すると推計の精度が低くなる可能性がある。

本論文においては、保険会社における経験死亡率のような相対的にデータ量が少なく観察期間も短い死亡率データに対して Lee-Carter モデルを適用する場合の留意点について考察した。そして、その対応策の一つとして Lee-Carter モデルにおける特異値分解の第2因子を用いたモデルについて検討をおこない、日本の死亡率に適用し、そのモデルの有効性について確認した。

以上

<目次>

1. はじめに
2. 代表的な死亡率推定モデル
 - 2.1. Lee-Carter モデル
 - 2.2. Lee-Carter モデルの修正モデル
 - 2.3. CMI のプロトタイプモデル
 - 2.4. その他の死亡率推定モデル
3. 保険会社の死亡率推定として Lee-Carter モデルを適用する場合の課題
 - 3.1. 課題
 - 3.2. 小さな群団に適用する場合の課題
4. 第2因子を用いた Lee-Carter モデル
 - 4.1. 死亡率データについて
 - 4.2. 第1因子のみを用いた Lee-Carter モデルについて
 - 4.3. 第2因子を用いた Lee-Carter モデルの検討
 - 4.4. 第2因子を用いた Lee-Carter モデルの推定結果
5. まとめ
 - 6.1. まとめ
 - 6.2. 今後の課題

(参考) 山口県の死亡率データを用いた将来死亡率の推計

(参考.1.) 山口県の死亡率推計を用いた将来死亡率の推計

(参考.2.) 山口県の死亡率推計を用いた将来死亡率の推計結果

<参考文献>

1. はじめに

昨今、保険業界では、国際会計基準の導入や経済価値ベースのソルベンシー規制など様々な場面において、資産や負債を経済価値ベースで評価することが求められるようになった。その負債の評価に大きな影響を与える死亡率についても経済価値ベースで評価することが求められるため、将来死亡率を推定するモデルの必要性が高まっている。

経済価値ベースの保険負債評価においては、偏りのない将来キャッシュフローの推計が必要であり、死亡率についても、偏りのない将来推定が必要である。日本をはじめ世界の多くの国では死亡率は改善傾向にあるため、死亡率の偏りのない将来推計においては改善傾向を反映させることが必要である。

死亡率の改善傾向を反映させた将来推定モデルとしては、Lee-Carter モデルや CMI のプロトタイプモデルなどが代表的である。Lee-Carter モデルは 1992 年に米国で発表されたモデルであり、シンプルなモデル、少ないパラメータで死亡率の改善傾向を表したモデルであるため、世界各国の研究機関等で利用されている。日本においても国立社会保障・人口問題研究所が発表している「日本の将来推計人口」において使用されている。CMI のプロトタイプモデルは英国アクチュアリー会の継続死亡率調査部会 (CMI) が 2009 年 6 月に Working Paper38 において公表した死亡率予測モデルであり、理解と説明が容易であるため、ユーザーの見解と目的にあわせて予測を修正できることが特徴である。

本論文においては、保険会社の死亡率推定として、日本の将来推計人口などにも使用されている Lee-Carter モデルを適用した場合の課題について考察する。

2. 代表的な死亡率推定モデル

2.1. Lee-Carter モデル (Lee and Carter, 1992)

Lee-Carter モデルとは、暦年 t における年齢 x 歳の中央死亡率 $m_{x,t}$ を、

$$\log m_{x,t} = a_x + b_x \cdot k_t + \varepsilon_{x,t}$$

という算式でモデル化したものであり、各パラメータは実際の死亡率データより推定される。各パラメータの意義は以下のとおりである。

- $m_{x,t}$: 暦年 t における年齢 x 歳の中央死亡率
- a_x : 観察期間における対数死亡率の平均値
- b_x : k_t が変化するときの年齢別感応度
- k_t : 暦年 t における死亡率の改善ファクター
- $\varepsilon_{x,t}$: 平均 0 の誤差項

Lee-Carter モデルはこのようにシンプルなパラメータで表現したモデルであり、死亡率の改善ファクターである k_t を将来推定することにより、将来の死亡率を推定することが出来る。

(1) パラメータの推定

各パラメータは、最小二乗法により決定される。すなわち、実際の死亡率データとパラメータの制約条件

$$\sum_{x=0}^{\omega} b_x = 1, \quad \sum_{t=0}^{T-1} k_t = 0$$

のもと、誤差項の二乗和

$$\sum_{x=0}^{\omega} \sum_{t=0}^{T-1} \varepsilon_{x,t}^2 = \sum_{x=0}^{\omega} \sum_{t=0}^{T-1} (\log m_{x,t} - a_x - b_x \cdot k_t)^2$$

を最小化することによってパラメータが決定される。

まず、 a_x については制約条件から

$$a_x = \frac{1}{T} \cdot \sum_{t=0}^{T-1} \log m_{x,t}$$

という推定値を得る。次に

$$z_{x,t} = \log m_{x,t} - a_x$$

とおくと、上記の最小化問題は、制約条件のもとで

$$\sum_{x=0}^{\omega} \sum_{t=0}^{T-1} (z_{x,t} - b_x \cdot k_t)^2$$

を最小化することに帰着する。これは、 $z_{x,t}$ を成分とする $(\omega + 1) \times T$ 行 Z 列の特異値分解

$$Z = U \cdot S \cdot {}^t V = \left(\sum_{i=1}^r s_i \cdot u_{xi} \cdot v_{ti} \right)_{x=0 \sim \omega, t=0 \sim T-1}$$

s_i : 行列 Z の特異値分解で $s_1 \geq s_2 \geq s_3 \geq \dots \geq s_r$

r : 行列 Z の階数

を行うことで解決する。すなわち、上記特異値分解から

$$b_x \cdot k_t = s_1 \cdot u_{x1} \cdot v_{t1}$$

$$\varepsilon_{x,t} = \sum_{i=2}^r s_i \cdot u_{xi} \cdot v_{ti}$$

を得ることができ、 b_x, k_t については制約条件から定まる（詳細は及川（2006）を参照）。

(2) k_t の将来推定

Lee and Carter は米国の死亡率にこのモデルを適応することにより、 k_t を時系列と見なして、時間の経過と共に概ね直線的に推移することを発見した。

$$k_t = k_{t-1} - \alpha + \varepsilon'_t \quad (\alpha \geq 0 : \text{定数、} \varepsilon'_t : \text{誤差項})$$

具体的には Lee and Carter は k_t の将来推定を、自己回帰和分移動平均モデル ARIMA(0, 1, 0) を用いて推定した。

(3) Lee-Carter モデルの特徴と課題

Lee-Carter モデルは以下のような特徴があるため、世界各国で幅広く使用されている。

- ✓ 改善度を一次元のパラメータ k_t で記述できるので平易
- ✓ 改善度 k_t が近似的には時間の経過と共に直線的に推移
- ✓ 特異値分解を用いることで計算が容易
- ✓ 各パラメータで観測データが集約できるため、例えば男女別、国別といった比較も可能

しかしながら、Lee-Carter モデルには以下のような課題があるとされており、様々な工夫や修正モデルの開発、研究がなされている。

- ✓ 特異値分解において第 1 因子しか使用しておらず、第 2 因子以下が使用されていない
- ✓ 誤差項 $\varepsilon_{x,t}$ は分散が均一な正規分布と仮定しており、各年齢での人口の大きさなどは考慮されていない
- ✓ 将来の k_t が直線であるが、日本においては死亡率の改善率は必ずしも直線ではない

- ✓ 将来の k_t が時点 t のみに依存し、年齢 x に依存しない

2.2. Lee-Carter モデルの修正モデル

Lee-Carter モデルはその簡明性から世界各国で幅広く使用されており、国の死亡率推定などにも用いられているが、死亡率の状況は各国で異なるため、様々な研究者がその国の死亡率の状況にあった Lee-Carter モデルの開発なども行っている。

ここでは、Lee-Carter モデルの応用例の一部について、日本の研究者を中心に紹介する。なお、海外では Booth et al. (2002) が様々な角度から Lee-Carter モデルに修正を加えることを試みている。

(1) 小松モデル：小松隆一（2002）

Lee and Carter は米国の死亡率予測として k_t の推定に ARIMA モデルを採用し、死亡率の改善が今後も直線的に継続する仮定したモデルであるが、日本の死亡率の改善状況は低下傾向にあるため、日本の死亡率予測として k_t の推定に ARIMA モデルを採用すると、過度に死亡率が改善する可能性がある。

小松（2002）は日本の死亡率の改善状況を反映させるため、 k_t に特定の値に向かって漸近していく指数関数と、低下傾向が徐々に弱まるものの限りなく低下し続ける対数関数を使用し推定した（以下、「小松モデル」）。

$$\log m_{x,t} = a_x + b_x \cdot k_t + \varepsilon_{x,t}$$

$$a_x = \frac{1}{2} \cdot \sum_{t=T-1}^T \log m_{x,t} \quad : \text{最近の死亡率を反映}$$

b_x : Lee-Carter モデルにおける b_x の移動平均

k_t : Lee-Carter モデルにおける k_t に対して関数を当てはめて推定

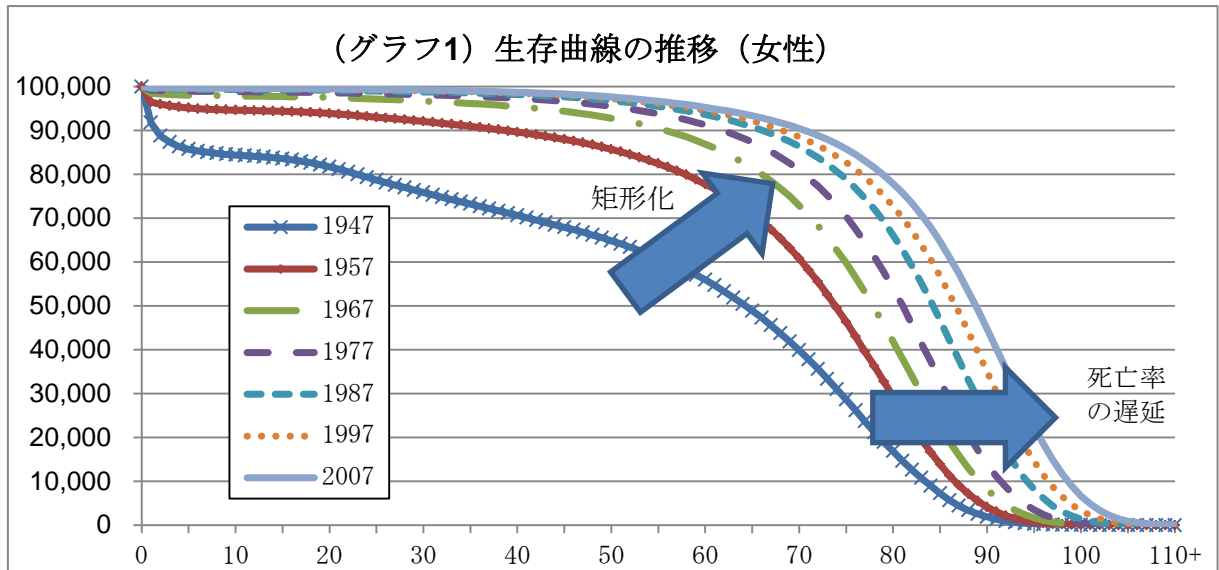
$$\cdot \text{指数関数 } k_t = \alpha_1 + \alpha_2 \exp\left(\frac{t-\alpha_4}{\alpha_3}\right) \text{ と}$$

$$\cdot \text{対数関数 } k_t = \beta_1 + \beta_2 \ln(t + \beta_3)$$

$\varepsilon_{x,t}$: 誤差項

(2) 年齢シフトモデル：石井太（2006）（2008）

小松モデルは日本の死亡率の改善傾向を反映したモデルであり、国立社会保障・人口問題研究所が発表した「日本の将来推計人口（平成 14 年 1 月推計）」においても使用されているが、我が国の近年の死亡率の改善傾向、死亡率の遅延という傾向を十分に反映できていない。石井（2006）は、死亡率の遅延、すなわち死亡率曲線のシフトを表現できるように、年齢変換を行ってから Lee-Carter モデルを適用するように改善したモデルを提案した（以下、「年齢シフトモデル」）。



石井は死亡率の遅延を反映させるため、シフティング・ロジスティック・モデル (Bongaarts (2005)) を用いて年齢のシフト量を推定している。つまり、まず死力についてロジスティック曲線

$$\mu_{x,t} = \frac{\alpha_t}{1 + \alpha_t \exp(\beta_t x)} + \gamma_t : \text{死力}$$

への当てはめを行い、年齢シフト量を定めるパラメータ $S_t = -\frac{\log \alpha_t}{\beta_t}$ と、曲線の勾配を表すパラメータ β_t を推定する。

死亡率のオリジナルの年齢 x 、年齢シフト後を y とし、 $x = f(y)$ を以下の通り定義し、年齢シフト後の $\hat{m}_{y,t} (= m_{f(y),t})$ の対数値に対して Lee-Carter モデルを適用する。

$$f(y) = \begin{cases} y & (y \leq B_1) \\ \left\{ \frac{\beta_{t_0} (B_2 - S_{t_0}) + S_t - B_1}{\beta_t} \right\} \frac{y - B_1}{B_2 - B_1} + B_1 & (B_1 \leq y \leq B_2) \\ \frac{\beta_{t_0}}{\beta_t} (y - S_{t_0}) + S_t & (B_2 \leq y \leq S_{t_0}) \\ y - S_{t_0} + S_t & (S_{t_0} \leq y) \end{cases}$$

年齢シフトを考慮することにより、Lee-Carter モデルがもつ簡明性を保ちつつ、小松モデルの死亡率推計が表現できていない、死亡率の遅延を表現している。なお、この石井の年齢シフトモデルは国立社会保障・人口問題研究所が発表した「日本の将来推計人口 (平成 18 年 12 月推計、平成 24 年 1 月推計)」において使用されている。

(3) その他

上記 2 つのモデルは「日本の将来推計人口」において使用された死亡率モデルであるが、上記のモデル以外にも様々な研究者が Lee-Carter モデルを修正し、日本の死亡率に適用している。

例えば、小暮厚之・長谷川知弘 (2005) は、Lee-Carter モデルにおける誤差項 $\epsilon_{x,t}$ の分散が年齢や暦年に関わらず均一な正規分布と仮定していることに疑問を感じ、誤差項 $\epsilon_{x,t}$ の分散に年齢や暦

年によるばらつきを組み込んだモデル（ポアソン対数双線形モデル）を提案した。

また、及川桂（2006）は、死亡率の改善状況は同じ年齢でも世代によって違うことから、改善率を到達年齢ではなく生年に焦点をあてた死亡率の推定モデルを提案した。

2.3. CMI のプロトタイプモデル

英国アクチュアリー会の継続死亡率調査部会（CMI）は2009年6月にWorking Paper38において、以下の特徴を持つ死亡率予測モデルを公表した（以下、「CMI のプロトタイプモデル」）。

- ✓ 死亡率トレンドの最新の経験を反映
- ✓ 理解と説明が容易
- ✓ ユーザーの見解と目的にあわせて予測を修正することができるような自主性を許容
- ✓ 定期的見直しが可能

CMI のプロトタイプモデルは死亡率改善率に焦点を当てて将来の死亡率予測をするモデルであり、以下のパラメータを設定し、将来の死亡率を予測する。

- ✓ 基準となる死亡率
- ✓ 初期の死亡率改善率
- ✓ 長期間の死亡率改善率
- ✓ 初期の死亡率改善率から長期間の死亡率改善率への収束過程

なお、死亡率の改善率 $m_{x,t}$ は「到達年齢要素」と「コーホート要素」に分割される。

$$m_{x,t} = 1 - \frac{q_{x,t}}{q_{x,t-1}} = m_{x,t}^A - m_{t-x,x}^C$$

到達年齢要素 $m_{x,t}^A$: t年における到達年齢 x 歳の死亡率改善率

コーホート要素 $m_{t-x,x}^C$: t-x年生まれの、x年後の死亡率改善率

初期の死亡率改善率は、この記号を使うと実績の最終年Tに対して以下の算式になる。

$$m_{x,T} = 1 - \frac{q_{x,T}}{q_{x,T-1}} = m_{x,T}^A - m_{T-x,x}^C$$

長期間の死亡率改善率は、Kを予測期間とすると以下の算式になる。

$$m_{x,T+K} = m_{x,T+K}^A - m_{T+K-x,x}^C$$

この算式を使用すると、T+K年における到達年齢 x 歳の死亡率 $q_{x,T+K}$ は以下の式で表される。

$$q_{x,T+K} = q_{x,T} \times (1 - m_{x,T+1}) \times (1 - m_{x,T+2}) \cdots \times (1 - m_{x,T+K-1})$$

また、初期の死亡率改善率と長期間の死亡率改善率は2つの要素「到達年齢要素」と「コーホート要素」ごとに収束する。すなわち、

到達年齢要素 : $m_{x,T+t}^A \rightarrow m_{x,T+K}^A$ (t → K)

コーホート要素 : $m_{T-x,x+t}^C \rightarrow m_{T-x,x+K}^C$ (t → K)

と、それぞれの要素で収束していく。

2.4. その他の死亡率推定モデル

Lee-Carter モデルや CMI のプロトタイプモデル以外にも世界では様々な死亡率推定モデルが提唱されており、代表的なモデルをここで簡潔に紹介する。なお、Cairns 等 (2009) はそれらのモデルの特徴の整理と比較をおこなっている。

(1) P-spline モデル (CMI の Working Paper20 を参照)

P-spline モデルは、年齢および暦年、あるいは年齢および出生年によるスプライン曲線群の重ね合わせで死亡率曲面を表現するものである。

$$\log m_{x,t} = \sum_{i,j} \theta_{i,j} B_{i,k}(x) B_{i,l}(t)$$

$B_{i,k}(x), B_{j,l}(t)$ はそれぞれ k 次および l 次の B スプライン基底関数で表されたスプライン曲線である。一般に 3 次スプライン関数を用いられることが多い。パラメータ $\theta_{i,j}$ を推定する際には最尤法ではなくペナルティ付最尤法を用いることで、死亡率曲線の滑らかさを考慮に入れながらパラメータの推定を行っている。

(2) CBD モデル : Cairns-Blake-Dowd (2006)

CBD モデルは高齢層の過去の死亡率データと良くフィットしていると言われ、英国での議論のひとつの軸となっているモデルである。

$$\text{logit}(q_{x,t}) := \log\left(\frac{q_{x,t}}{1-q_{x,t}}\right) = \kappa_t^{(1)} + \kappa_t^{(2)}(x - \bar{x})$$

$\kappa_t^{(1)}$: 全体的な死亡率の水準

$\kappa_t^{(2)}$: 死亡率の年齢間の傾き

\bar{x} : 対象としている集団の平均年齢

3. 保険会社の死亡率推計として Lee-Carter モデルを適用する場合の課題

3.1. 課題

本論文においては保険会社の将来死亡率を推計するモデルとして、日本の将来推計人口においても使用されている Lee-Carter モデルを適用した場合の課題について考察する。

まず、保険会社の死亡データを用いて推定する場合、保険会社の死亡率データには選択効果が含まれており、特に契約直後の死亡率は、国民全体の死亡率より低くなる。また、保険会社に死亡率データを使用する場合は、解約や更新などによるリスク濃縮なども考慮する必要がある。

さらに保険会社の死亡率データは米国や日本といった国民全体の死亡率データと比べると小さな群団であり、また観測期間も短いことから、Lee-Carter モデルをそのまま適用すると推計の精度が低くなる可能性がある。

選択効果やリスク濃縮の反映についても重要な課題であるが、本論文においては小さな群団に適用する場合の課題について考察する。

3.2. 小さな群団に適用する場合の課題

Lee-Carter モデルはパラメータの設定において特異値分解の第1因子のみを使用しているため、小さな群団に Lee-Carter モデルを適用すると第1因子の寄与度が小さくなり、推定の精度が落ちる場合がある。

日本の死亡率に Lee-Carter モデルを適用した場合における各因子の寄与度は以下のとおりであり、第1因子の寄与度は第2因子以降の寄与度と比べると大きいものの、第2因子以降が無視できるほど第1因子の寄与度が十分高いとは言えない。

なお、石井（2006）も日本の死亡率の特徴である、高齢期への年齢シフトといった死亡状況の変化をとらえるためには、第2因子まで合わせた推計値が必要ではないかと示唆している。

（表1）＜日本の死亡率（1970～2005年）（男性）に適用した場合の寄与度＞

	1	2	3	4	5	6	7	8	9	10	20
固有値	25.9	3.0	1.2	1.0	0.9	0.8	0.8	0.7	0.6	0.5	0.3
寄与度(%)	60.1	7.0	2.8	2.4	2.1	1.9	1.9	1.6	1.4	1.3	0.8
累積寄与度(%)	60.1	67.1	69.9	72.3	74.4	76.3	78.2	79.7	81.1	82.4	92.4

日本の死亡率においても第1因子の寄与度が十分高いとは言えないが、小さな群団では第1因子の寄与度がさらに下がり、推定の精度が落ちることがあるため、第2因子を用いた Lee-Carter モデルについて考察する。

4. 第2因子を用いた Lee-Carter モデル

4.1. 死亡率データについて

まず、日本の死亡率に Lee-Carter モデルを適用して考察するが、基礎データとしては The Human Mortality Database (以下 HMD) に収録されている日本の粗中央死亡率($m_{x,t}$)を使用する。

HMD とは死亡率の研究者に対して、プロジェクトが独自に計算した死亡率や生命表を、これらの計算に用いた入力データとともに提供している死亡データベースである。HMD はカリフォルニア大学バークレー校 (UCB) とマックスプランク人口研究所 (MPIDR) の研究チームが共同で立ち上げたプロジェクトであり、2013 年 1 月現在で 37 の国または地域が対象となっている。

(URL : <http://www.mortality.org/>)

死亡率や生命表は通常、人口や死亡という生データそのものではなく、これらのデータを加工し精査することによって作成される。そのため、異なる生命表を比較する場合、この加工の仕方を踏まえて比較する必要があるが、その加工の違いは多岐にわたるため、非常に困難である。そこで HMD では死亡率、生命表の作成方法について、すべての国・地域または時代に統一的に適用する方法論を確立し、取りまとめている。

日本においては厚生労働省が発表している完全生命表や簡易生命表などがあるが、それらにも作成方法や最終年齢などに違いがあるため、本論文においては作成方法等が統一されている HMD の死亡率データを使用する。

ただし、1995 年については阪神大震災の影響があることから、1995 年の死亡率は 1994 年と 1996 年の和半とした。

4.2. 第1因子のみを用いた Lee-Carter モデルについて

まず、第1因子のみを用いた Lee-Carter モデル (小松モデル) を日本の死亡率 (1970 年~2005 年) に適用し、将来の死亡率を推定する。なお、小松モデルにおいては、 a_x は直近 2 年平均であるが、ここでは直近 5 年平均とした。直近 2 年平均とすると、どちらかの年に特異な死亡率が含まれていた場合、その特異な死亡率が a_x に影響を与えることがある。また、小松モデルの後に発表された石井の年齢シフトモデルにおいても直近 5 年平均としている。

$$\log m_{x,t} = a_x + b_x \cdot k_t + \varepsilon_{x,t}$$

$$a_x = \frac{1}{5} \cdot \sum_{t=T-4}^T \log m_{x,t} \quad : \text{最近の死亡率を基準にするため}$$

b_x : Lee-Carter モデルにおける b_x の移動平均

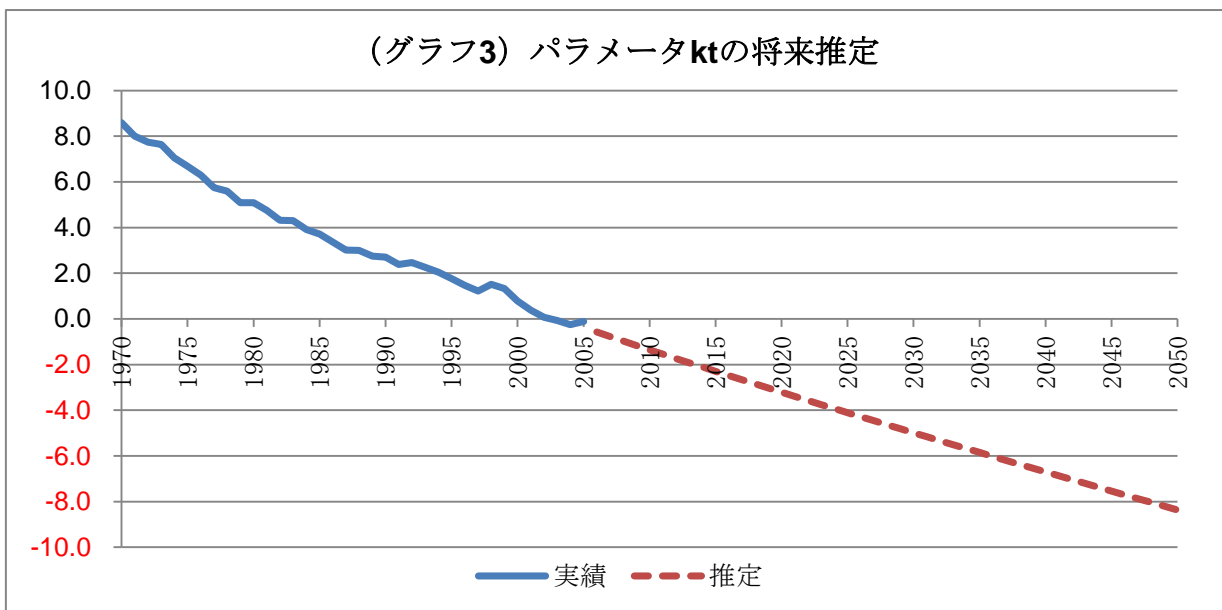
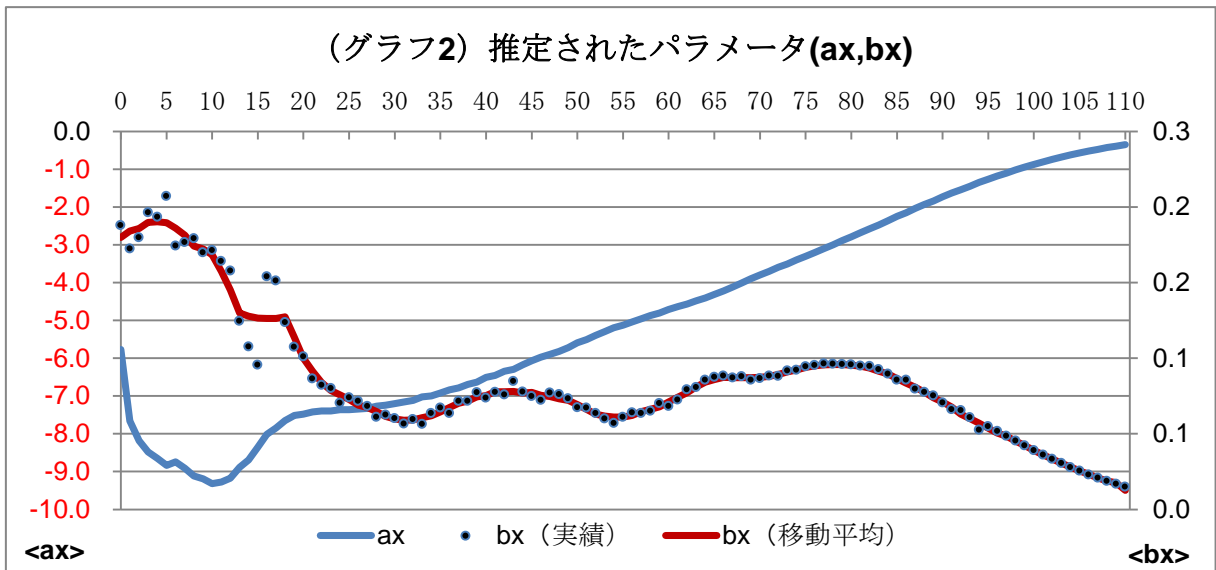
k_t : Lee-Carter モデルにおける k_t に対して関数を当てはめて将来の k_t を推定

$$\cdot \text{指数関数 } k_t = \alpha_1 + \alpha_2 \exp\left(\frac{t-\alpha_4}{\alpha_3}\right)$$

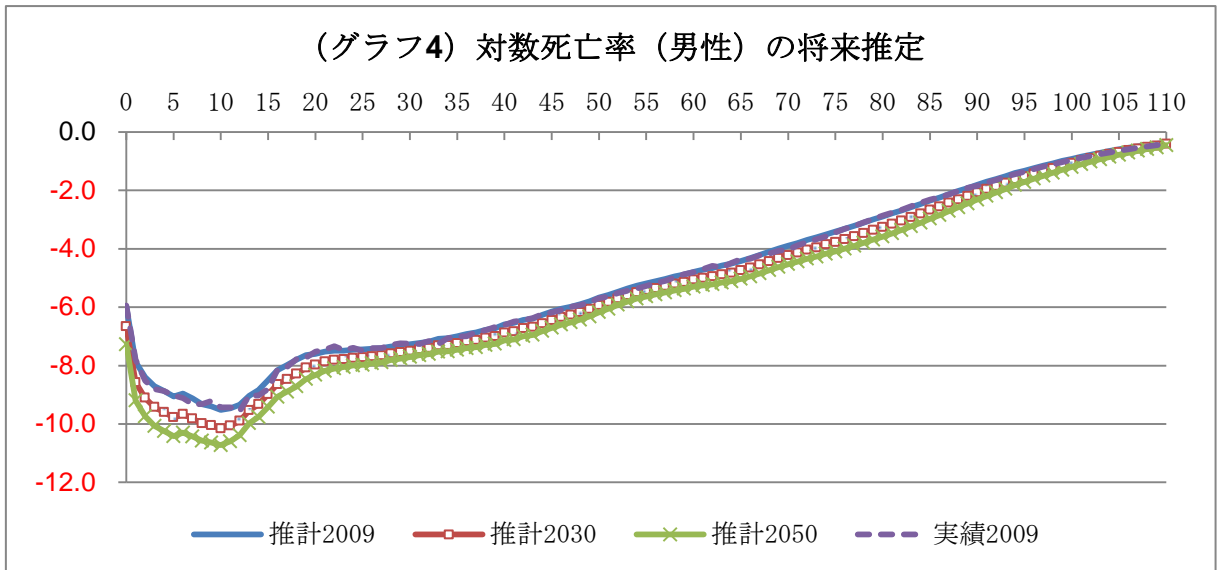
$$\cdot \text{対数関数 } k_t = \beta_1 + \beta_2 \ln(t + \beta_3)$$

$\varepsilon_{x,t}$: 誤差項

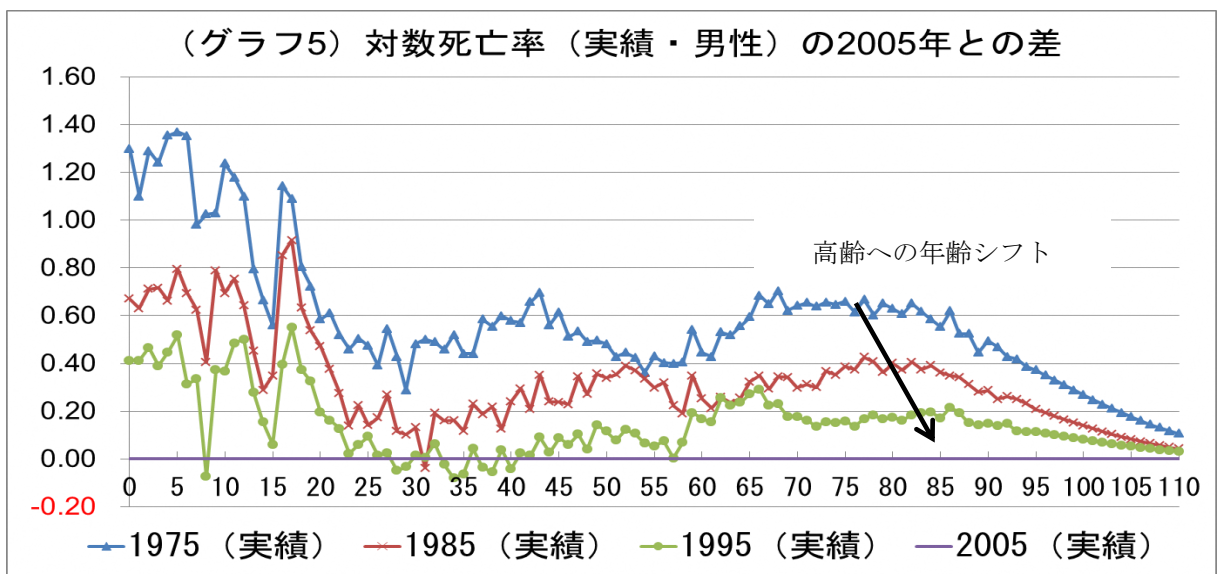
各パラメータは以下のように推定される。なお、パラメータの推定においては統計ソフト R を使用している。



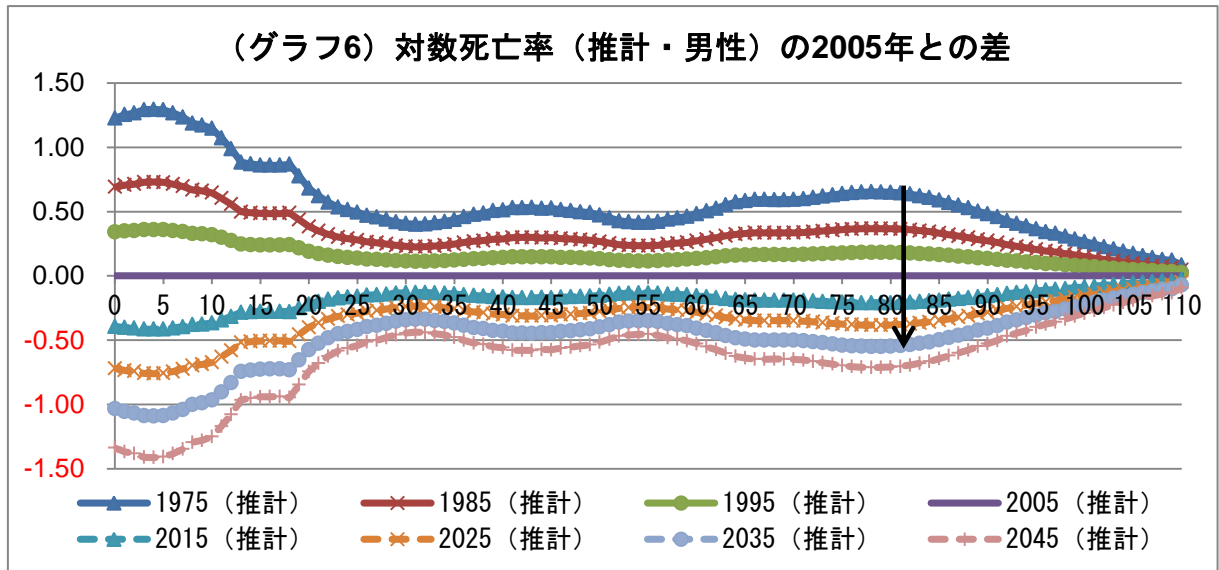
このように推定されたパラメータを用いて死亡率の将来推定を行うと、以下のような結果となった。2005年までの死亡率データを用いて2009年の死亡率を推定した値と2009年の実績値を比較すると、ほぼ推定できているように見える。また、2030年、2050年の死亡率は2005年の死亡率より改善していることが分かる。



次に死亡率の遅延について確認する。まず、実績値を用いて 2005 年の対数死亡率との差をグラフ化して並べると、高齢層において年齢シフトしていることが分かる。



次に、第1因子を用いた Lee-Carter モデルにおいて推定した対数死亡率の 2005 年との差をグラフ化してみる。実績値で見たような高齢層における年齢シフトを確認することができない。

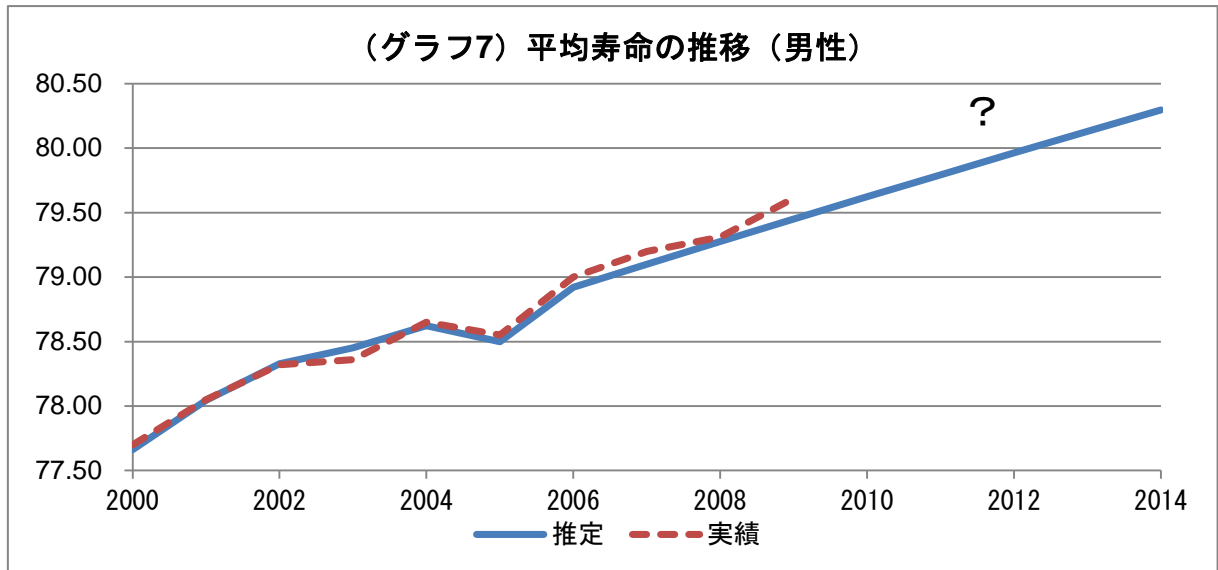


次に、平均寿命について確認する。2005年までのデータをもとに第1因子を用いた Lee-Carter モデルで2006年から2009年までを推定した平均寿命と、平均寿命の2006年から2009年までの実績値を比較すると、特に男性において、推定値のほうが実績値より平均寿命が短くなっている。

(表2) <平均寿命の実績と推計>

		2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
男性	実績	77.70	78.05	78.32	78.36	78.65	78.55	79.00	79.20	79.31	79.61
	推計	77.66	78.05	78.33	78.45	78.62	78.50	78.92	79.10	79.28	79.45
	(差)	0.04	0.00	-0.01	-0.09	0.03	0.05	0.08	0.10	0.03	0.16
女性	実績	84.58	84.91	85.21	85.30	85.57	85.49	85.79	85.98	86.04	86.42
	推計	84.66	84.97	85.13	85.19	85.71	85.50	85.68	85.91	86.14	86.36
	(差)	-0.08	-0.06	0.08	0.11	-0.14	-0.01	0.11	0.07	-0.10	0.06

2009年の推計においても平均寿命を短く推定しており、さらに遠い将来を推定すると、その差が大きくなる可能性がある。



これは、第1因子のみを用いたモデルでは日本の高齢層における年齢シフトを十分に反映できておらず、将来の平均寿命を短く推定している可能性がある。

石井（2006）も日本の死亡率の特徴である、高齢期への年齢シフトといった死亡状況の変化をとらえるためには、第2因子まで合わせた推計値が必要ではないかと示唆している。これらのことから、日本の死亡率に当てはめて第2因子を用いた Lee-Carter モデルを検討する。

4.3. 第2因子を用いた Lee-Carter モデルの検討

特異値分解における第2因子を用いるため、第2因子より定まるパラメータを $b_x^{(2)}$ 、 $k_t^{(2)}$ とすると、モデルとしては

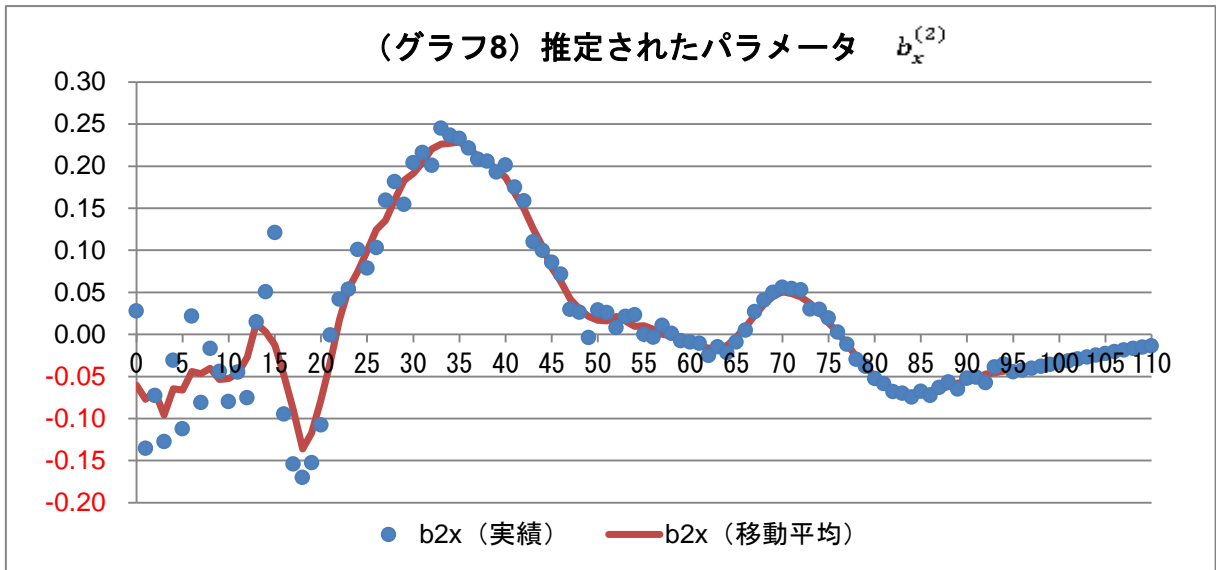
$$\log m_{x,t} = a_x + b_x^{(1)} \cdot k_t^{(1)} + b_x^{(2)} \cdot k_t^{(2)} + \varepsilon_{x,t}$$

$b_x^{(1)}$ 、 $k_t^{(1)}$ ：第1因子より定まるパラメータ

$b_x^{(2)}$ 、 $k_t^{(2)}$ ：第2因子より定まるパラメータ

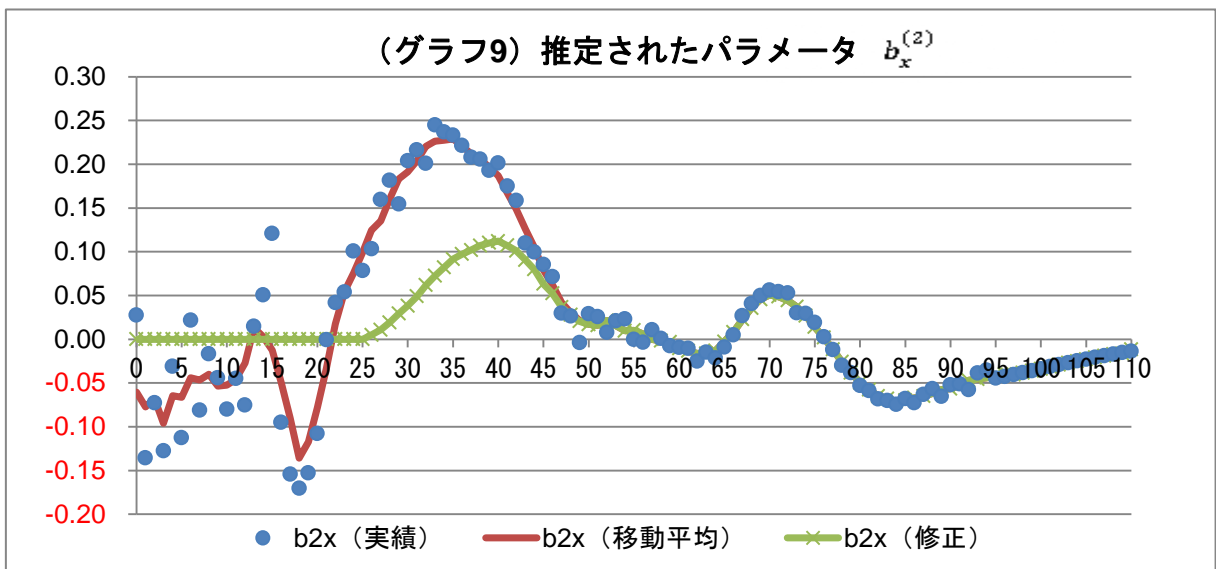
と定義することができる。なお、 a_x 、 $b_x^{(1)}$ 、 $k_t^{(1)}$ は 4.2. におけるモデルと同様に設定した。

$b_x^{(2)}$ 、 $k_t^{(2)}$ の設定方法について日本の死亡率に適用して検討する。まず、 $b_x^{(1)}$ と同様に特異値分解で定まる値 $b_x^{(2)}$ (実績) に対して移動平均を取ると以下のようになる。移動平均後においても、若年の変動が激しく安定していない。

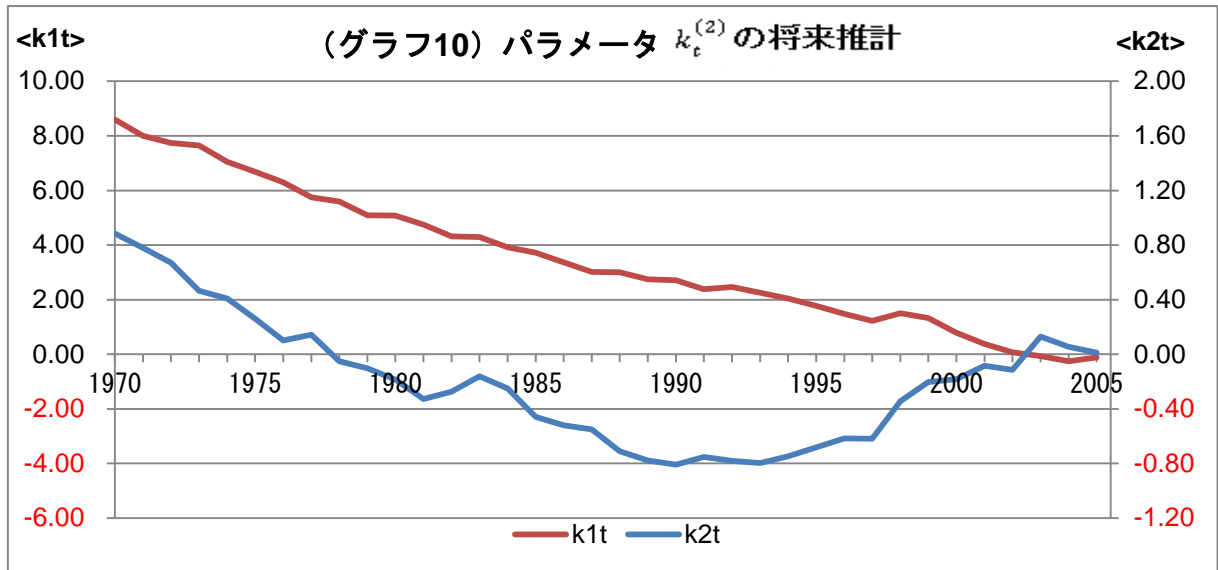


石井の年齢シフトモデルにおいては、年齢シフトは高齢のみであり、若年は年齢シフトを抑制している。同様に本モデルの $b_x^{(2)}$ の設定においても、若年の $b_x^{(2)}$ は変動が大きいため、25歳以下についてはモデルへの反映を行わず、50歳以上においては $b_x^{(2)}$ の移動平均をそのまま使用することとし、その間の年齢層では補間によりモデルへの反映を行った。

$$\hat{b}_x^{(2)} = \begin{cases} 0 & : \text{if } x \leq 25 \\ b_x^{(2)} \times \frac{(x - 25)}{25} & : \text{if } 25 < x \leq 50 \\ b_x^{(2)} & : \text{if } 50 < x \end{cases}$$

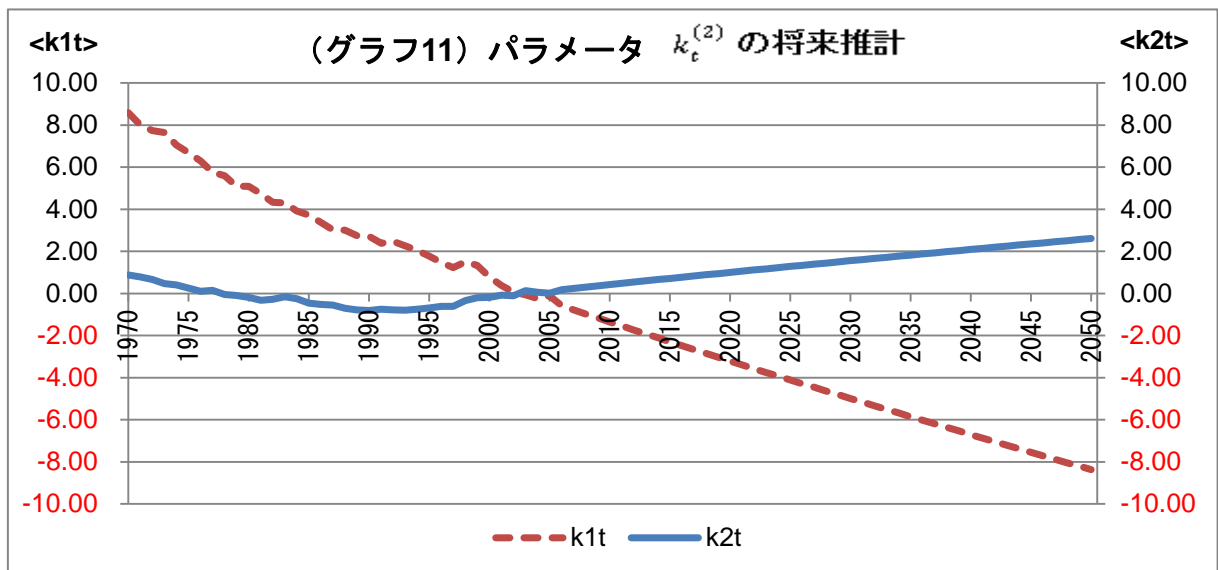


次に、 $k_t^{(2)}$ の将来推定について検討する。まず、特異値分解により定まる $k_t^{(2)}$ と $k_t^{(1)}$ の推移を比較すると、以下のとおりである。



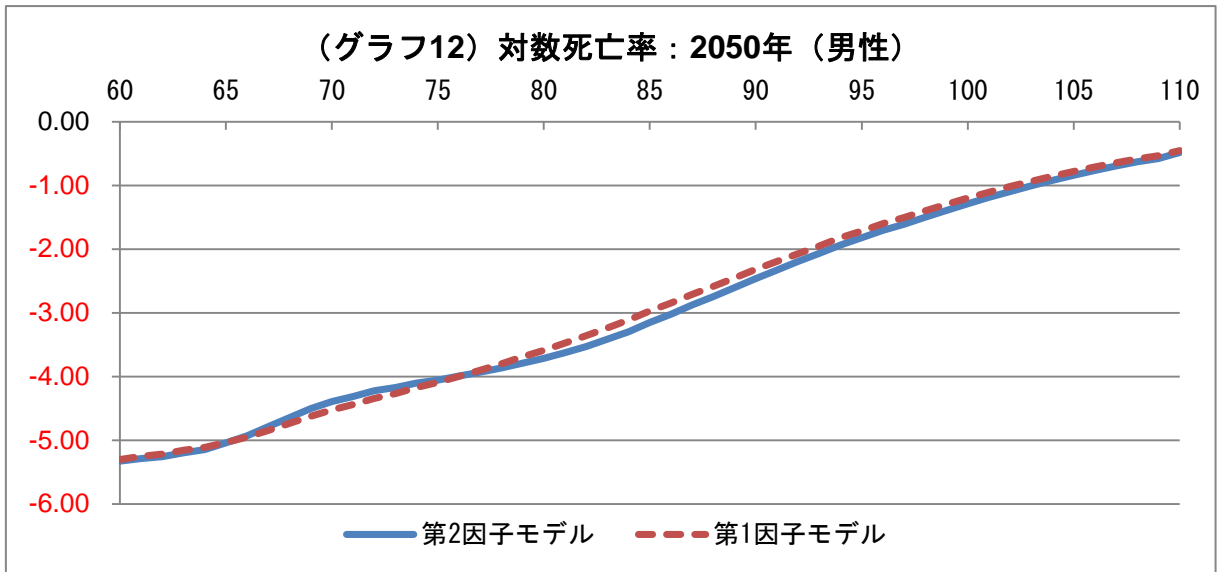
$k_t^{(2)}$ は1990年から反転し、上昇傾向にあるため、今後も上昇すると仮定して推定するが、時系列モデル等で推計すると第2因子による影響が大きくなりすぎるため、過去10年間の $k_t^{(1)}$ との線形関係を用いて将来推定する。なお、石井の年齢シフトモデルにおいてもシフト量を決めるパラメータ S_t の将来推定を過去10年間の k_t との線形関係を用いて将来推定している。

$$\hat{k}_t^{(2)} = \alpha \cdot k_t^{(1)} + \beta$$

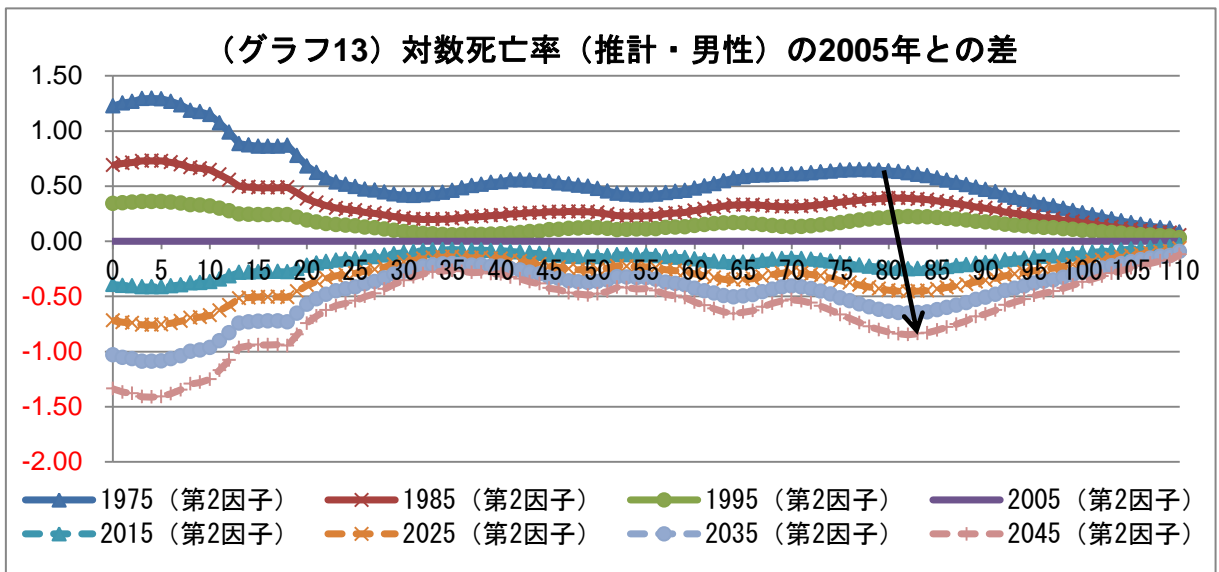


4.4. 第2因子を用いた Lee-Carter モデルの推定結果

このようにして定まるパラメータを用いて将来推定した死亡率と第1因子のみを用いて将来推定した死亡率を比較すると、以下のようになる。



第1因子のみでは表現出来なかった高齢層の年齢シフトについて、第2因子を用いたモデルで確認すると、グラフにおける高齢層における山が右にシフトしており、高齢層の年齢シフトが確認できる。

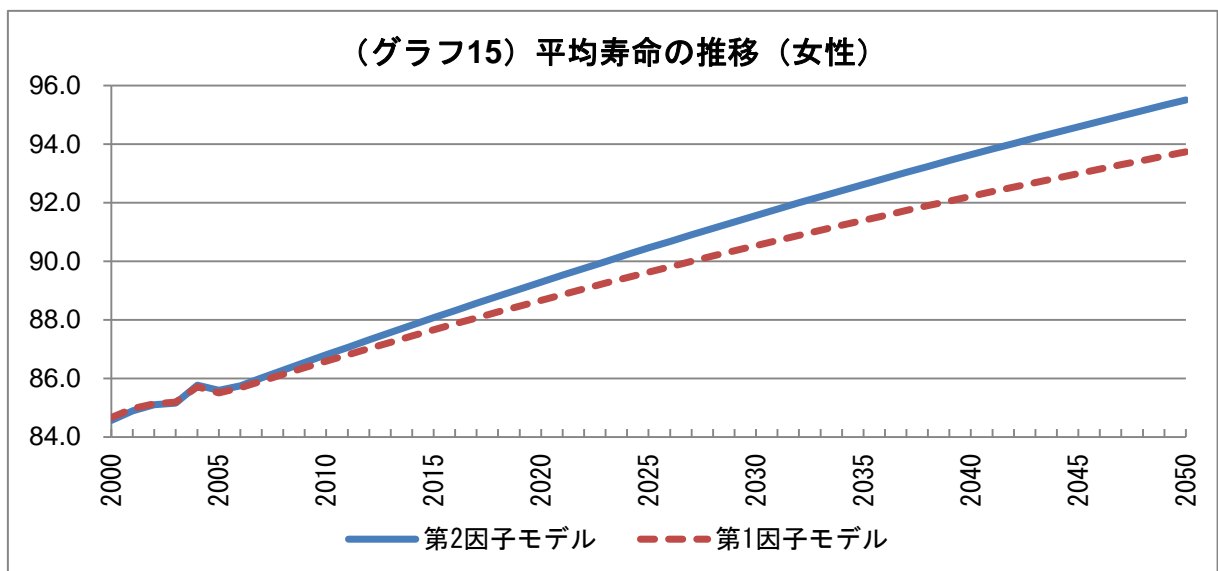
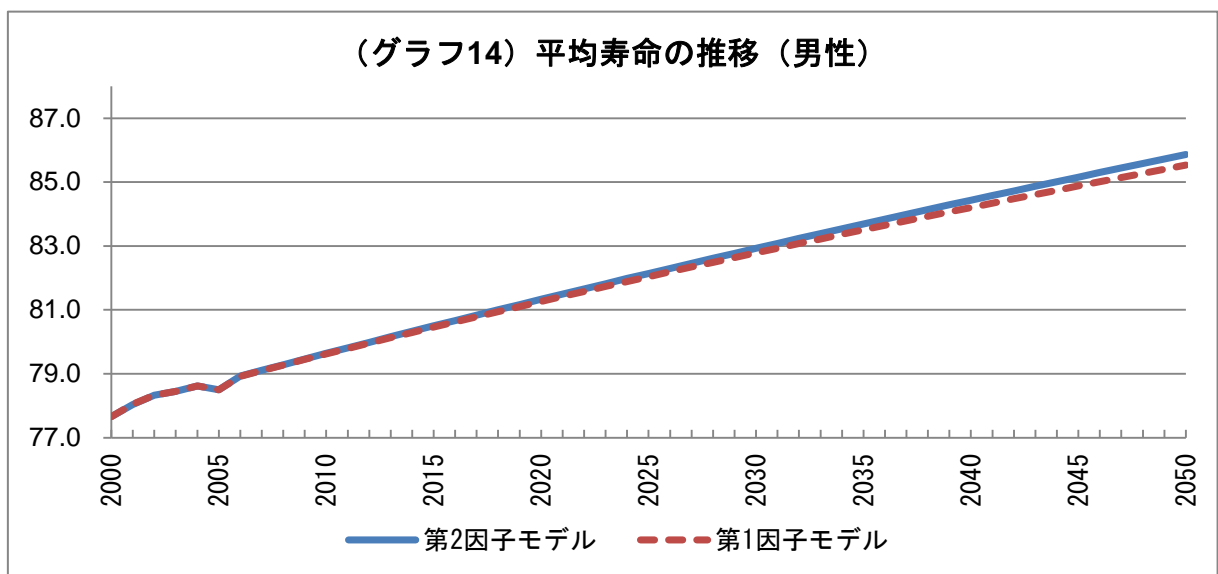


2006年から2009年までの平均寿命の推計値と実績値を比較すると、第2因子を用いたモデルのほうが、第1因子を用いたモデルより概ね近い値を示した。

		2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
男性	実績	77.70	78.05	78.32	78.36	78.65	78.55	79.00	79.20	79.31	79.61
	第2因子	77.66	78.04	78.33	78.46	78.62	78.50	78.93	79.11	79.29	79.46
	(差)	0.04	0.01	-0.01	-0.10	0.03	0.05	0.07	0.09	0.02	0.15
	第1因子	77.66	78.05	78.33	78.45	78.62	78.50	78.92	79.10	79.28	79.45
	(差)	0.04	0.00	-0.01	-0.09	0.03	0.05	0.08	0.10	0.03	0.16

		2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
女性	実績	84.58	84.91	85.21	85.30	85.57	85.49	85.79	85.98	86.04	86.42
	第2因子	84.57	84.90	85.10	85.16	85.77	85.60	85.75	86.01	86.28	86.54
	(差)	0.01	0.01	0.11	0.14	-0.20	-0.11	0.04	-0.03	-0.24	-0.12
	第1因子	84.66	84.97	85.13	85.19	85.71	85.50	85.68	85.91	86.14	86.36
	(差)	-0.08	-0.06	0.08	0.11	-0.14	-0.01	0.11	0.07	-0.10	0.06

また、将来（2050年）の平均寿命については、第2因子を用いたモデルは、第1因子のみを用いたモデルより平均寿命が長くなるという推定結果になった。特に女性について顕著である。



5. まとめ

5.1. まとめ

Lee-Carter モデルはシンプルなモデルであり、少ないパラメータで死亡率の改善傾向を表したモデルであるため、世界各国の研究機関等で使用されている。しかしながら、Lee-Carter モデルは特異値分解の第1因子のみを用いたモデルであるため、小さな群団に適用すると第1因子の寄与度が下がり、推定の精度が落ちる可能性がある。

日本の死亡率推計のような国全体の将来推計の場合は、第2因子以降の影響は小さいかもしれないが、保険会社が自社データを用いて将来推計をする場合や、さらにリスク細分化商品の死亡率を推計する場合などには、第2因子以降の影響が無視できないと思われる。そのため、本論文では特異値分解の第2因子を用いたモデルを提案した。なお、モデルの検討においては石井の年齢シフトモデルのパラメータの設定方法などを参考に設定している。

モデルの検討においては、まず日本の死亡率に適用してパラメータの設定を試みた。その結果、第2因子を用いた Lee-Carter モデルは、第1因子のみを用いたモデルと比較すると、第1因子のみのモデルでは表現できなかった高齢層の年齢シフトが表現でき、平均寿命についても概ね第2因子を用いたモデルのほうが実績に近い値を推定できていた。日本の死亡率データに適用した場合については、第2因子を用いたモデルの有効性について、一定程度、確認できた。

5.2. 今後の課題

本論文においては、第2因子のパラメータを石井の年齢シフトモデルと関連付けて推定をおこない、その有効性について一定程度確認できたが、以下の点について課題があると考えている。

まず、パラメータの設定方法について課題があると考えている。本論文では、石井の年齢シフトモデルと関連付けて推定をおこなったが、設定方法を定性的な位置づけも踏まえてさらに検討する必要がある。例えば、 $b_x^{(2)}$ は特異値分解より定まる値をそのまま使用するのではなく、若年についてはその値の変動が激しいため、反映を抑えるように修正した。若年は変動が激しいため、そのまま使用するの適切でないのかもしれないが、反映を抑える修正の仕方について検討が必要である。また $k_t^{(2)}$ については $k_t^{(1)}$ との線形関係を用いて推計したが、 $k_t^{(2)}$ そのものを推計する手法についても検討が必要である。

次に、モデルの有効性の評価についても課題があると考えている。本論文ではモデルの評価について、実績値と推計値の比較や、グラフなどを用いておこなったが、モデルが有効であるかどうかは、統計的手法を用いて評価するのが望ましいため、モデルの評価の仕方について検討することが必要である。

また、本論文においては、第2因子を用いたモデルを検討するために、日本の死亡率に適用してパラメータを検討したが、保険会社の経験死亡率のような小さな群団データにおいても本モデルが適用できるかの確認も大きな課題である。

小さな群団としては例えば、都道府県の死亡率データなどがあるが、都道府県の死亡率データは日本の死亡率データと比べると整備されていない、もしくは公表されているデータが数年分の都道府県が多く、分析するには十分ではなかった。都道府県の死亡率データについては、現在、国立社会保障・人口

問題研究所において、HMD の日本版として JMD (Japanese Mortality Database) の構築が行われており、JMD には都道府県別の死亡率についても収録が予定されているため、今後、このデータを使用することにより、小さな群団に対する本モデルの有効性を確認することが出来る可能性があるため、注視していきたい。(URL : <http://www.ipss.go.jp/p-toukei/JMD/index.html>)

(注) 都道府県の死亡率データとしては山口県が平成 12 年から平成 23 年までの 12 年間分の死亡率データを公表しており、確認できた範囲では最も長い期間にわたる死亡率データを公表している。しかしながら死亡率が 90 歳までであること、他の都道府県の死亡率データに適用した場合との比較ができていないこと、また山口県の死亡率の推移が日本の死亡率の推移(高齢における死亡率の遅延等)と異なる可能性があることなどから、本モデルを小さな群団に適用した場合の有効性については確認することが出来なかったが、参考として、山口県の死亡率データに本モデルを適用した場合の将来推定について記述したので、ご確認いただきたい。

(参考) 山口県の死亡率データを用いた将来死亡率の推計

(参考.1) 山口県の死亡率データを用いた将来死亡率の推定

都道府県別の生命表としては、厚生労働省が昭和40年から平成17年までに合計9回、作成しているが、5年ごとであり各年の生命表ではない。各年の生命表としては山口県が平成12年から平成23年まで合計12回、作成しているため、本論文においては山口県の生命表を使用する（山口県、保健統計年報より）。

なお、山口県の人口は約143万人（平成25年1月1日現在）であり、平均寿命は厚生労働省の平成23年の調査において男性が79.11歳、女性が85.27歳である。

		平成10年	平成11年	平成12年	平成13年	平成14年	平成15年	平成16年	平成17年	平成18年	平成19年	平成20年	平成21年	平成22年	平成23年
男	全国	77.19	77.16	77.72	78.07	78.32	78.36	78.64	78.56	79.00	79.19	79.29	79.59	79.55	79.44
	山口県	76.49	76.52	77.03	77.42	77.83	77.99	77.86	78.11	78.11	78.29	78.79	78.68	79.03	79.11
女	全国	83.82	84.01	84.60	84.93	85.23	85.33	85.59	85.52	85.81	85.99	86.05	86.44	86.30	85.90
	山口県	83.44	83.39	84.61	84.33	84.26	84.31	84.56	85.63	84.69	84.83	84.99	84.92	86.07	85.27

※山口県の平成12、17、22年は都道府県別生命表（厚生労働省）、平成12、17、22年以外は簡易生命表（県厚政課）による。

※全国の平成12、17、22年は完全生命表（厚生労働省）、平成12、17、22年以外は簡易生命表（厚生労働省）による。

本モデルを山口県の死亡率データに適用するにあたり、死亡率データは2000年から2011年までの12年間のデータを使用し、 a_x は2007年から2011年までの5年平均とする。なお、女性の死亡率において一部、負値となっている年齢があるため、将来死亡率の推計は男性のみ行った。

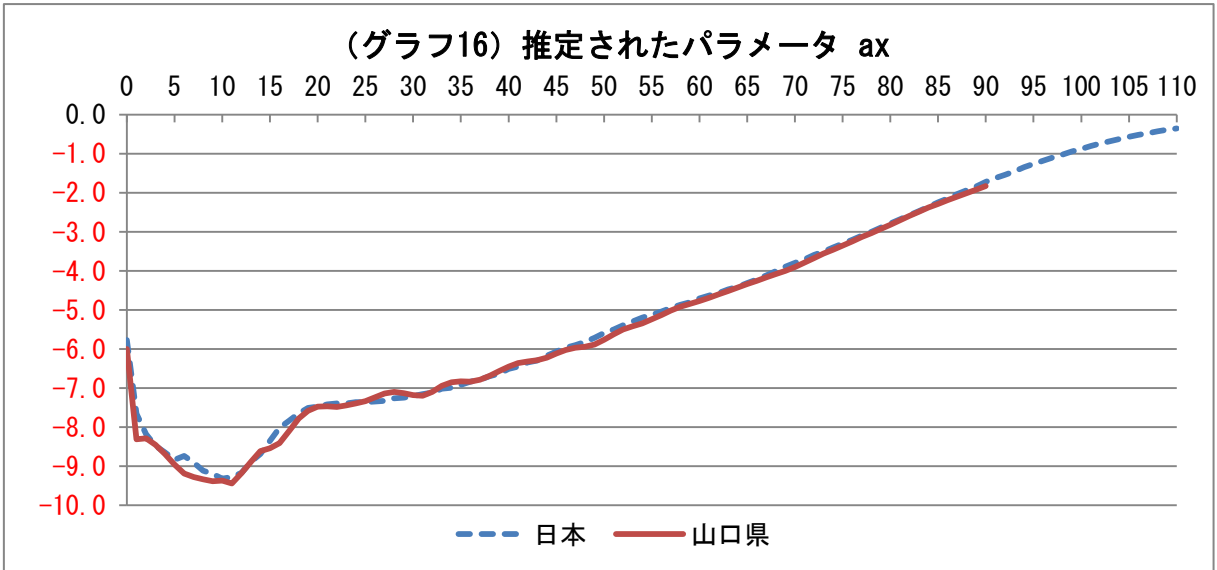
これらのデータを用いて特異値分解を行った場合の、第1因子、第2因子の寄与度は以下のとおりであり、国の死亡データを用いた時と比べ、第1因子の寄与度が小さくなっている。

(表1) <山口県の死亡率(2000~2011年)(男性)に適用した場合の寄与度>

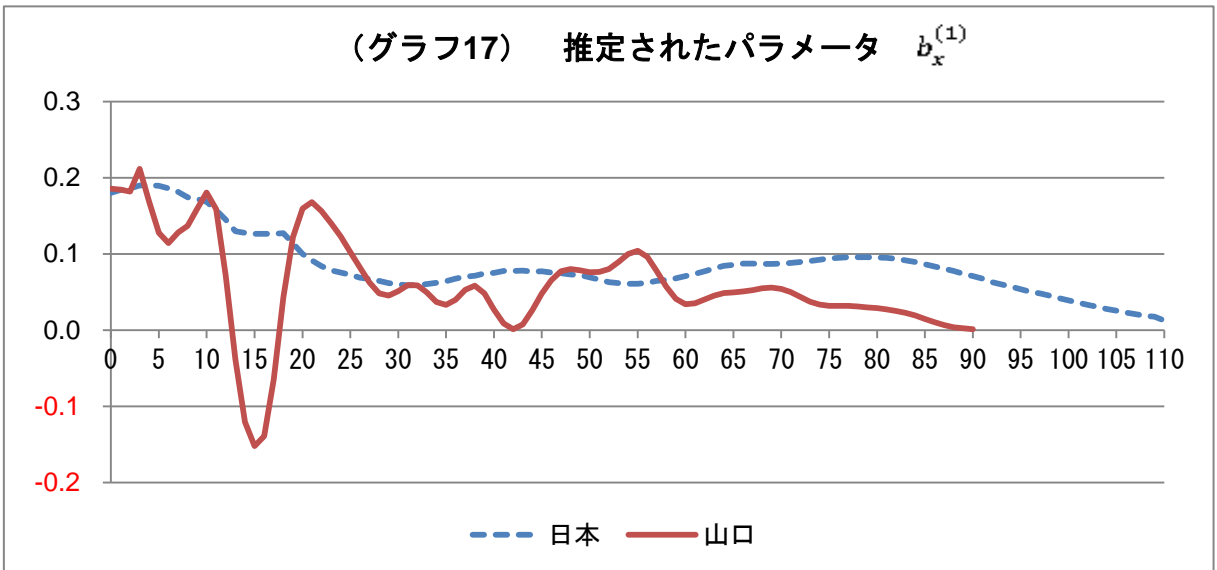
	1	2	3	4	5	6	7	8	9	10	11
固有値	5.0	4.0	3.5	2.7	2.3	2.0	1.7	1.7	1.3	1.1	0.9
寄与度(%)	18.9	15.3	13.5	10.3	8.8	7.5	6.6	6.5	5.1	4.2	3.4
累積寄与度(%)	18.9	34.2	47.7	58.0	66.8	74.3	80.9	87.3	92.4	96.6	100.0

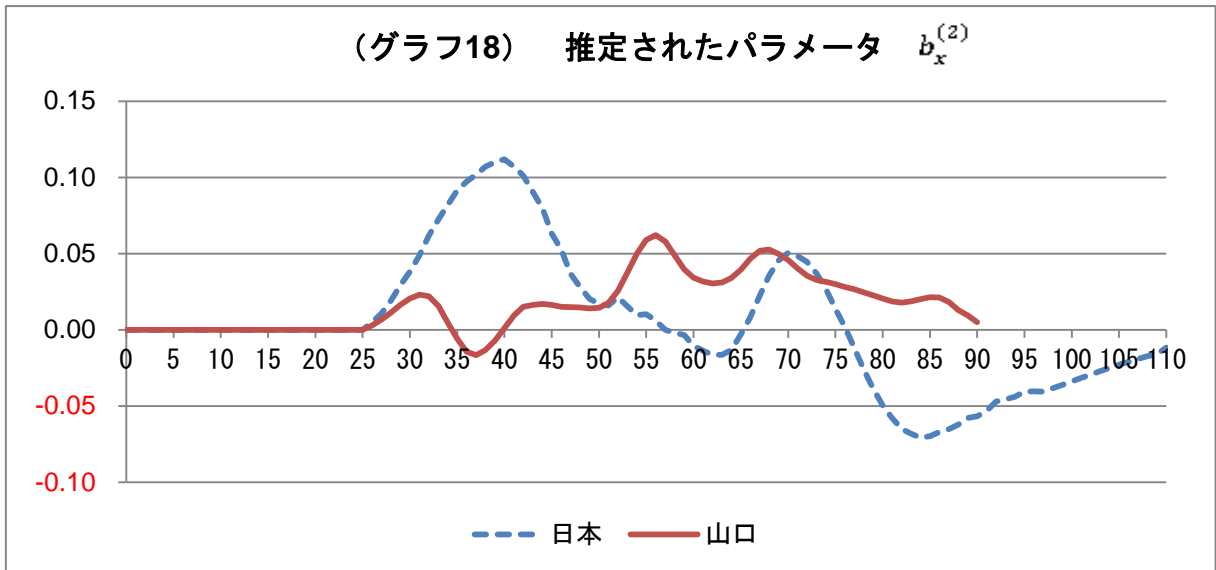
山口県の死亡率データを用いて設定した各パラメータ、 a_x 、 $b_x^{(1)}$ 、 $k_t^{(1)}$ 、 $b_x^{(2)}$ 、 $k_t^{(2)}$ は以下のとおりである。

まず、 a_x については、対数死亡率の5年平均であるため、日本に死亡率に適用した場合と傾向そのものには共通している。

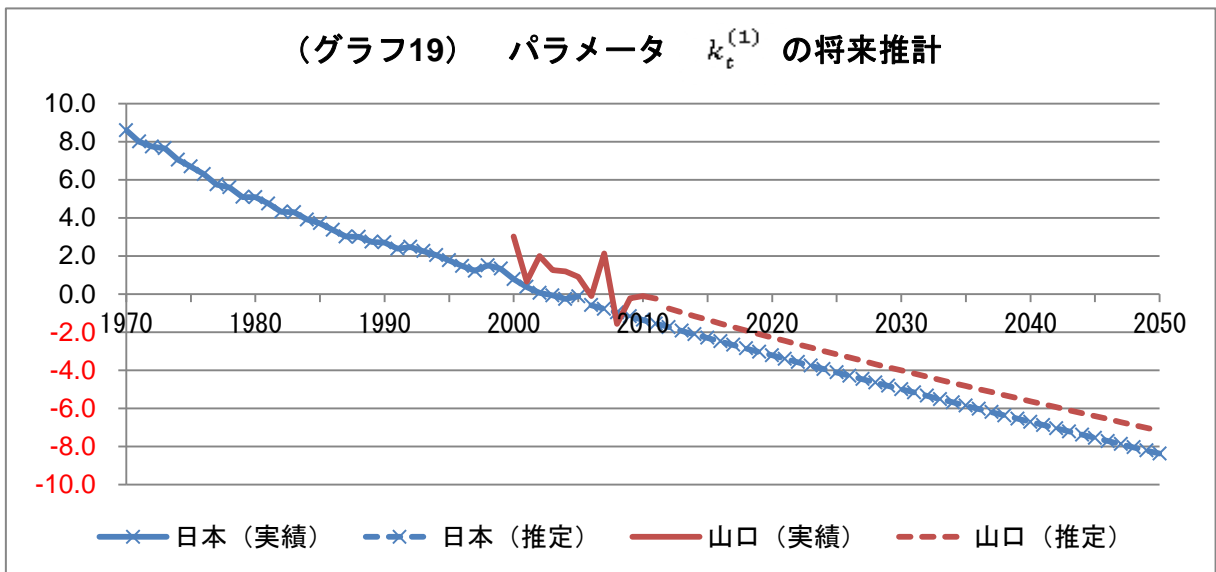


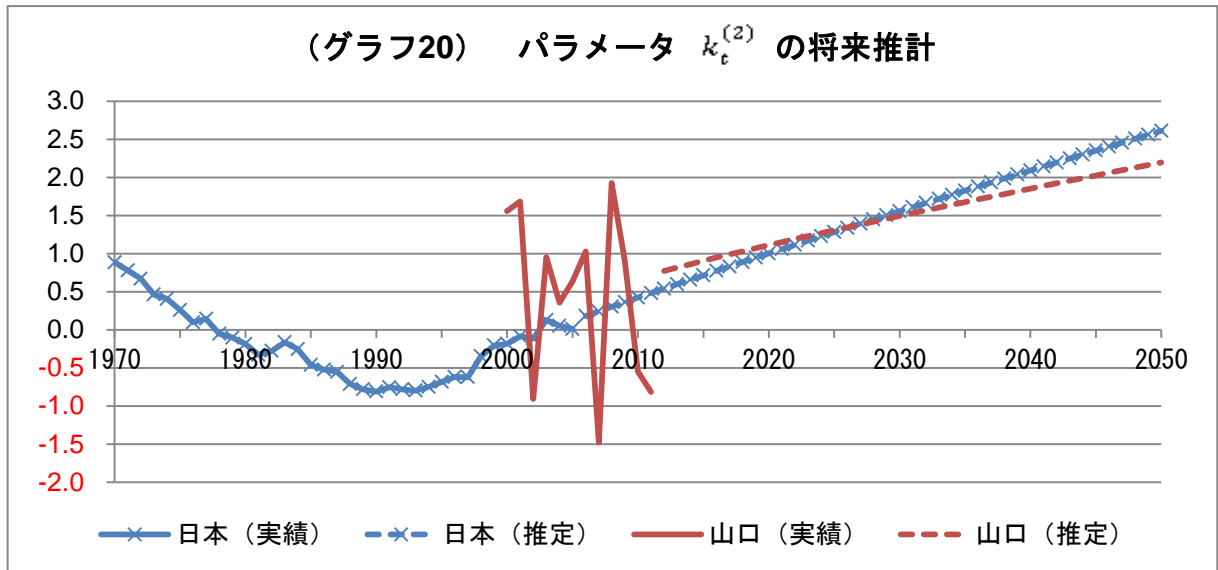
次に、 $b_x^{(1)}$, $b_x^{(2)}$ については日本の死亡率に適用した場合と異なる年齢がある。 $b_x^{(1)}$ は若年部分の違いが顕著である。また、 $b_x^{(2)}$ は高齢部分で符号が異なっているのが、大きな違いである。





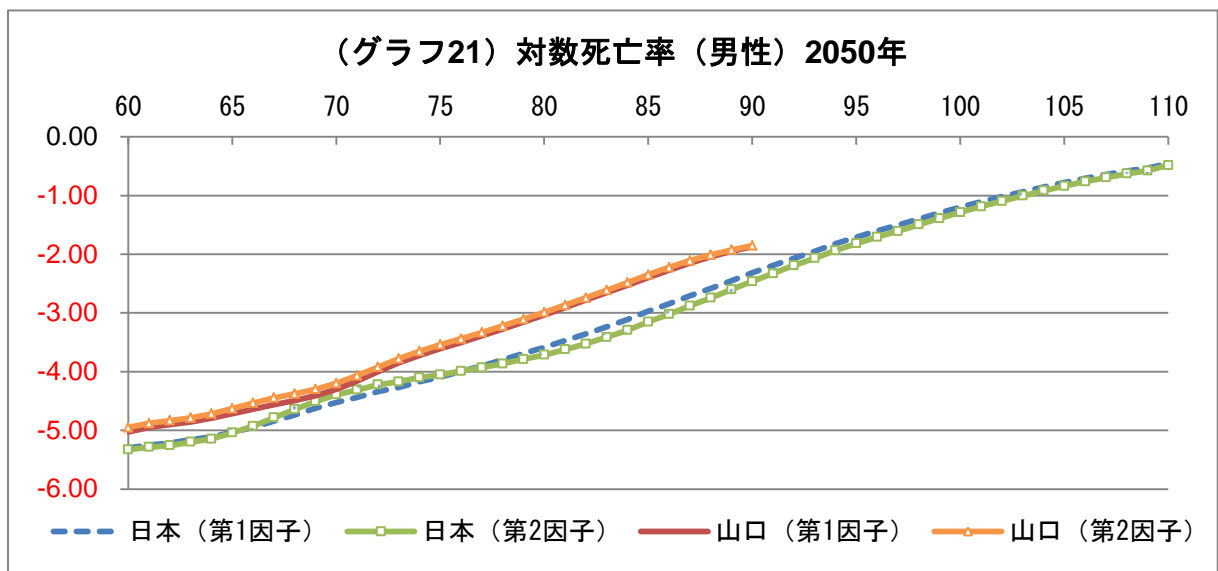
つぎに、将来推定に関するパラメータ $k_t^{(1)}$, $k_t^{(2)}$ については、将来推定は、 $k_t^{(1)}$, $k_t^{(2)}$ ともに概ね日本の死亡率に適用した場合と共通している。実績値については、 $k_t^{(2)}$ の変動が日本に適用した場合と比べると大きくなっている。山口県の死亡率データの期間が2000年から2011年までと短かったため、第2因子のトレンドまでうまく捉えることができなかった可能性がある。





(参考. 2.) 山口県の死亡率データを用いた将来死亡率の推定結果

山口県の死亡率データに本モデルを適用した場合の将来死亡率については以下のとおり。70歳前後において第1因子のみを用いたモデルの対数死亡率のほうが、第2因子を用いたモデルの対数死亡率より僅かに低くなっているが、日本の死亡率データに適用した場合と比べると、死亡率に大きな差は見られなかった。 $k_t^{(2)}$ の傾向については日本の死亡率データに適用した場合と大きな差はなかったが、 $b_x^{(2)}$ の特に高齢部分が日本の死亡率データに適用した場合と異なり、マイナスとなっていないため、死亡率の推定値に差が生じなかったものと思われる。

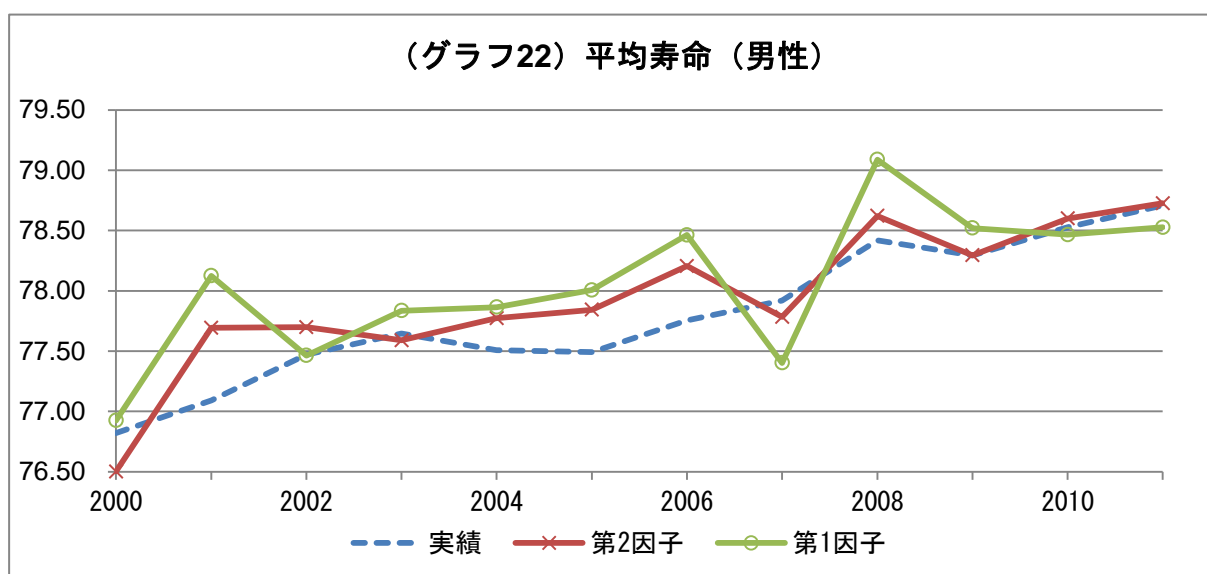


次に平均寿命についてそれぞれもモデルの値と実績値とで比較する。第2因子を用いたモデルの平均寿命のほうが第1因子のみのモデルの平均寿命より、概ね実績値に近くなっている。

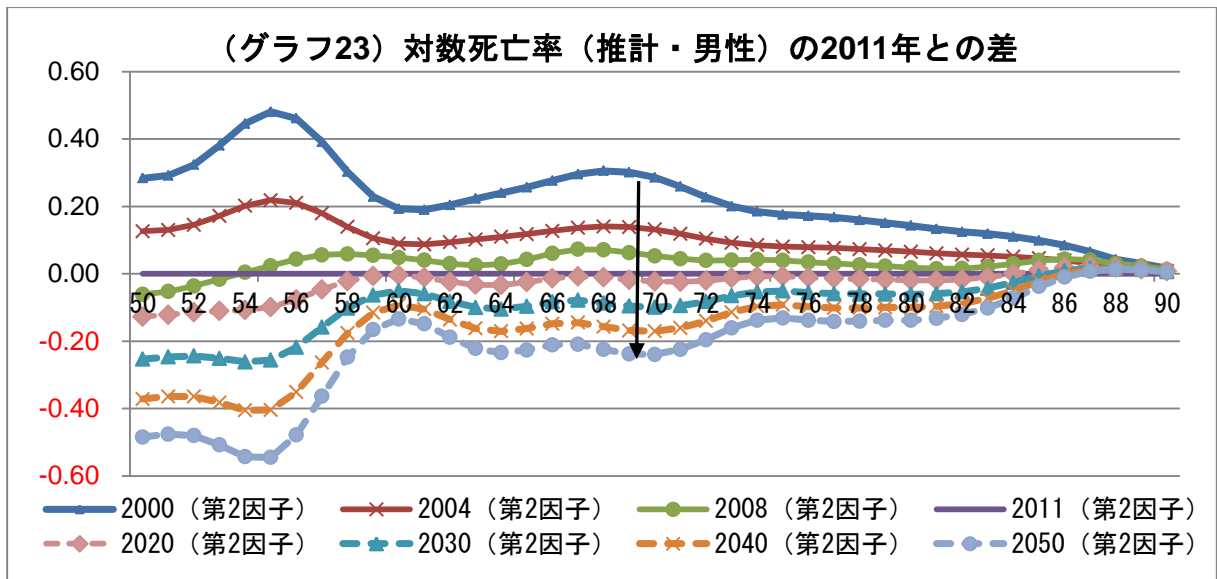
なお、ここでの実績値は開示されている平均寿命の実績値と異なる。開示されている死亡率は90歳までであるが、平均寿命は開示されていない90歳超の死亡率も含めて計算されていると思われるため、

比較のために開示されている 90 歳までの死亡率をもとに計算した平均寿命を用いた。

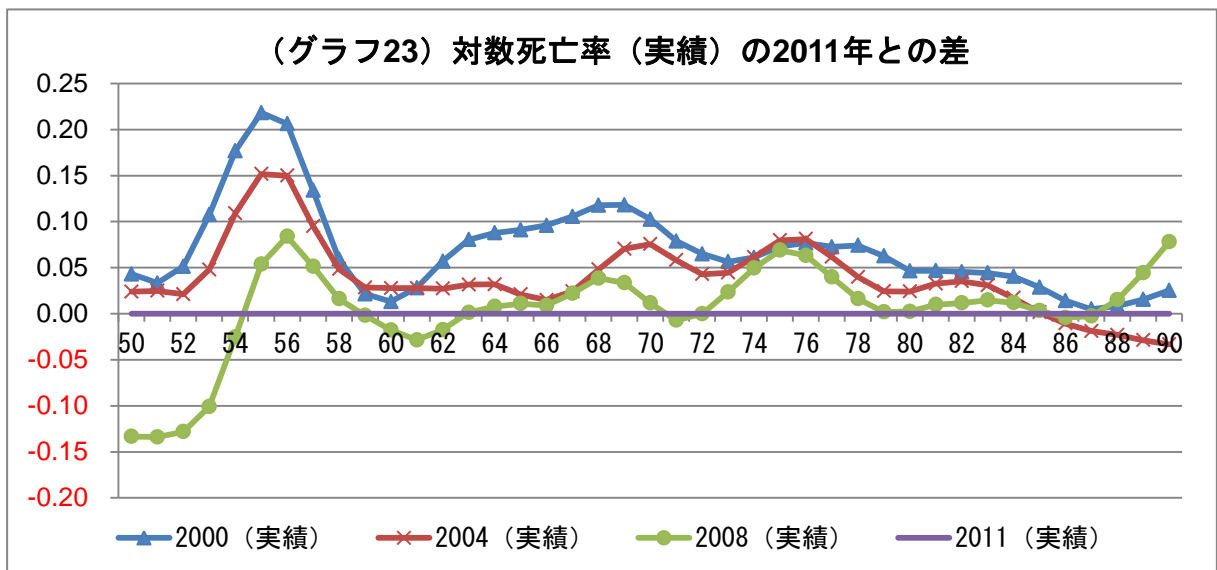
		2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
男性	実績	76.82	77.09	77.47	77.65	77.51	77.49	77.76	77.92	78.42	78.29	78.53	78.71
	第 2 因子	76.50	77.69	77.70	77.59	77.77	77.84	78.21	77.78	78.62	78.29	78.60	78.73
	(差)	0.32	-0.60	-0.23	0.06	-0.26	-0.35	-0.45	0.14	-0.20	-0.00	-0.07	-0.02
	第 1 因子	76.92	78.12	77.46	77.84	77.86	78.01	78.46	77.40	79.09	78.52	78.47	78.53
	(差)	-0.10	-1.03	0.01	-0.19	-0.36	-0.52	-0.71	0.52	-0.67	-0.23	0.06	0.18



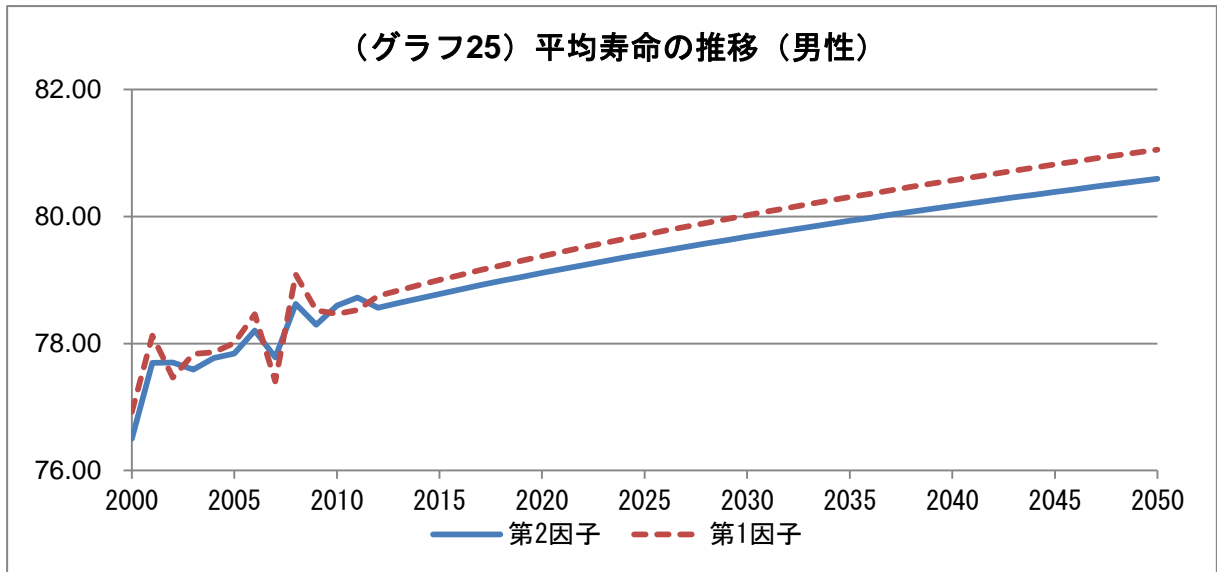
次に日本の死亡率データにおいて本モデルを適用した場合に確認できた高齢層の年齢シフトについて、山口県の死亡率データに適用した場合にも確認できるか確認する。2011 年の対数死亡率（男性）との差をグラフ化すると下記のグラフ 23 となるが、日本の死亡率とは異なり、山口県の死亡率データに本モデルを適用した場合においては高齢層の年齢シフトは確認できなかった（グラフ 13 と比較）。



これは本モデルが高齢層の年齢シフトを表現できなかったことより、山口県の死亡率データそのものに原因があると考えられる。山口県の死亡率データは観察年度が少ないこともあるが、実際のデータにおいても高齢層の年齢シフトは確認できない。また、山口県の平均寿命は男性が39位、女性が37位（平成22年都道府県別生命表の概況より）と、全国の都道府県の中で低位であるため、日本の死亡率の改善傾向と少し異なるのかもしれない。



最後に、平均寿命の将来推計を確認する。日本の死亡率においては第2因子を用いたモデルによって推定した平均寿命のほうが第1因子のみを用いて推定した平均寿命より長かったが、山口県に適用すると第2因子を用いたモデルのほうが短くなった。これは、 $b_x^{(2)}$ の高齢部分の符号が日本の死亡率に適用した場合と異なるからであると思われる。



(明治安田生命 商品部)

<参考文献>

- [1] Andrew J.G. Cairns, David Blake, and Kevin Dowd. (2006) ” A Two-Factor Model for Stochastic Mortality with Parameter Uncertainty” Theory and Calibration. Journal of Risk and Insurance, Vol. 73, pp. 687-718.
- [2] Andrew J.G. Cairns, David Blake, Kevin Dowd, Guy D. Coughlan, David Epstein, Alen Ong and Igor Balevich. (2009) ” A Quantitative Comparison of Stochastic Mortality Models Using Data From England and Wales and the United States” North American Actuarial Journal, Vol. 13, pp. 1-35.
- [3] Bongaarts, J. (2005) ” Long-range Trends in Adult Mortality: Models and Projection Methods” Demography, Vol. 42, NO. 1, pp. 23-49.
- [4] Booth, H., Maindonald, J., and Smith, L. (2002) ” Applying Lee-Carter under conditions of variable mortality decline.” Population Studies, Vol. 56, NO. 3, pp. 325-336.
- [5] Lee, R.D. and Carter, L.R. (1992) ” Modeling and forecasting U.S. mortality,” Journal of the American Statistical Association, 87, 659-671
- [6] Masakazu Ozeki (2005) ” Application of Mortality Models to Japan”
- [7] 赤松賢 (2010) 「死亡率に関する一考察～Prototype CMI Mortality Projection Model を使った死亡率予測について～」、日本アクチュアリー会会報第 63 号第 2 分冊、P67-107
- [8] 浅野幸弘 (2012) 「長生きリスクと年金運用」、日本経済新聞出版社
- [9] 池辺八洲彦 等 (2009) 「現代線形代数-分離定理を中心として-」、共立出版
- [10] 石井太 (2006) 「わが国の平均寿命の動向と死亡率推計モデルの検討」、人口問題研究所、62-3、P21-30
- [11] 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—」、人口問題研究所、64-3、P28-44
- [12] 石井太 (2010) 「寿命研究と Human Mortality Database」、人口問題研究所、66-3、P80-87
- [13] 及川桂 (2006) 「将来死亡率推計に関する一考察」日本アクチュアリー会会報第 59 号第 2 分冊、P1-28
- [14] 公益社団法人日本年金数理人会 (2013) 「退職給付会計基準における死亡率に関する報告書」
- [15] 小暮厚之・長谷川知弘 (2005) 「将来生命表と統計モデリング：Lee-Carter 法とその拡張—ヒューマンセキュリティへの基礎研究—」、総合政策学ワーキングペーパーシリーズ NO. 71
- [16] 小暮厚之・長谷川知弘 (2008) 「21 世紀の統計科学、第 I 巻、第 3 部、第 8 章 生命表の統計学—死亡率予測モデルとその年金リスク評価への応用—」、東京大学出版会、P182-203
- [17] 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の試み」、人口問題研究所、58-3、P3-14

Recently, it is sometimes required to evaluate assets and liabilities on economic value basis in the IFRS or solvency regime.

Models used to project future mortality rate are more important because assumption of mortality has a big impact on economic value base liability.

Projection method has been studied by many actuaries and demographers in each country.

Especially, the Lee-Carter model is one of the most famous models because of its simplicity.

In Japan, Komatsu, Oikawa and Ishii revised the Lee-Carter model partially and applied it to project Japanese mortality rate, respectively.

Komatsu or Ishii model was also used in the 'projection of future Japanese population' published by the National Institute of Population and Social Security Research.

It is considered to apply the models mentioned above to project mortality rate on economic value basis. However, compared to be applied to national mortality data, the original Lee-Carter model does not suit well when insurance company reflects its mortality experience to evaluate its insurance liability, because company's mortality experience data are relatively small and observed for a short term.

In this paper, it is considered what are important views when the Lee-Carter model is applied to relatively small and short-term observed mortality experience data, such as insurance company's data.

In addition, I introduce a revised model which contains a second eigen value in the singular value decomposition and exam its effectiveness in application to Japanese mortality rate.