

死亡率推計へのモデル生命表の応用について

石井太・堀口侑

はじめに

死亡は年齢によってその発生頻度が大きく異なることから、死亡率などの生命表関数を単純な法則や規則を用いて簡約化したものが死亡の年齢パターンのモデリングである。死亡の年齢パターンモデルには、大きく分けて、(1) 数学的関数によるモデル、(2) 数表によるモデル、(3) リレーショナルモデルの3種類があり、従来、モデル生命表と呼ばれるものは、Coale-Demney のモデル生命表など数表によるモデルを指すことが多かった。しかしながら、現在では、コンピュータを用いることにより、モデルにある程度複雑な計算過程が含まれるものであっても、ユーザーが必要なパラメータを設定しさえすれば、推計結果を比較的簡単に得ることができようになっており、モデル生命表の提示にあたって、数表による形式よりも、いくつかのパラメータを投入することで生命表関数を得られるリレーショナル・モデルを基本とする方式の方がユーザーにとっても活用がしやすい状況になっている。このような背景から、近年、リレーショナル・モデルを基本とするモデル生命表がいくつか提案されている。

そこで、本研究では、死亡率のモデリングと、近年のリレーショナル・モデルを基本とするモデル生命表についてレビューを行うとともに、これらを日本の死亡率推計に適合させるための修正を行ったモデル開発に関する試みについて述べることにしたい。

1 死亡率のモデリング

人口動態事象は年齢によってその頻度が大きく異なるものがあるが、死亡もその代表例である。一般に、年齢別に死亡率を見ると精密に観察を行うことができるが、多くの率を取り扱う必要性が生じるため、複雑さも増大することとなる。したがって、死亡率などの死亡を表現する生命表関数を、単純な法則や規則を用いて簡約化して表すことができると便利であり、これを実現するのが死亡の年齢パターンのモデリングである。

死亡の年齢パターンモデルには、大きく分けて、(1) 数学的関数によるモデル、(2) 数表によるモデル、(3) リレーショナルモデルの3種類がある。(1) の数学的関数によるモデルとは、年齢の数学的関数によって生命表関数を表現するものであり、死亡法則 (law of mortality) とも言われる。これには多くの先行研究があるが、死力を年齢の指数関数で表す Gompertz モデル (Gompertz 1825) はその代表例である。また、これに定数項を加えた Gompertz-Makeham モデル (Makeham 1860) は、現在、厚生労働省が作成する完全

生命表・簡易生命表で用いられている。一方、これらの指数関数では超高年層での死亡率を過大推定してしまうことがしばしば観察されることから、加齢に応じて死力の傾きが通減するロジスティック関数で死力を表現する Perks (1932)、Beard (1971) のような先行研究がある。

(2) の数表によるモデルは、経験に基づくいくつかの数表によって生命表関数を表現するものである。Coale-Demeny のモデル生命表 (Coale and Demeny 1983) がその代表例であり、生命表の形状 (シェイプ) を 4 種類 (North, South, East and West) に分け、それぞれに 25 のレベルから成るモデル的な生命表を用意して様々な生命表を表現している。また、国連のモデル生命表 (United Nations 1982) もその例であり、4 つの地域パターン (Latin American pattern, Chilean pattern, South Asian pattern, Far Eastern pattern) と一般的なパターン (General pattern) の 5 種類の形状について、それぞれ男女別に 35~75 年の e_0 に対応する生命表を示している。

数学的関数によるモデルは、生命表関数を数少ないパラメータで表現でき、豊富な情報を簡約化して表現できるという点で優れているが、一方で、実際の死亡の年齢パターンは複雑であり、必ずしも既知の数学的関数で表されるとは限らない。また、これを改善するために複数の数学的関数を組み合わせていくと、多くのパラメータが必要となっかえて複雑性を増してしまうという問題点がある。これに対し、数表によるモデルは、実際に経験された年齢パターンから作成されることから、このような数学的関数によるモデルのような制約はない。しかしながら、この方法では、それぞれのレベルや形状の種類に対応した数表を用意しなければならないため、レベルや形状を細分化していくと、膨大な数の数表が必要となり、やはり複雑性が增大してしまうという問題点がある。

(3) のリレーショナルモデルは、この二つのモデルの長所を採り入れたモデルともいうことができ、経験に基づく (数表による) 標準的な年齢パターンと、そこからの変化に関するパラメータを組み合わせて死亡パターンを表現するものである。リレーショナルモデルの代表例としては、Brass (1971) により開発されたブラス・ロジットシステムが挙げられる。ブラス・ロジットシステムは、 l_x をロジット変換した関数 $Y_x = \log \left(\frac{l_x}{1-l_x} \right)$ を考え、ある基準となるパターン Y_x^s を用いて、任意の Y_x^a が、 $Y_x^a = \alpha + \beta Y_x^s$ と表されるとするモデルである。ここで、 α がレベル、 β が形状を表すパラメータとなっている。

また、リーとカーターによって開発された生命表のリレーショナルモデル (リー・カーター・モデル, LC モデル) は、年齢別死亡率を、標準となる年齢パターン、死亡の一般的水準 (死亡指数)、死亡指数の動きに対する年齢別死亡率変化率および誤差項に分解することで、死亡の一般的水準の変化に応じて年齢ごとに異なる変化率を記述するモデルであり、現在国際機関や各国が行う将来推計においては標準的なモデルとして広く用いられている (Lee and Carter 1992)。

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

ここで、

$\log m_{x,t}$: 対数死亡率

a_x : 対数死亡率の標準的な年齢パターン

k_t : 死亡水準 (死亡指数)

b_x : k_t が変化する時の年齢別死亡率の変化

$\epsilon_{x,t}$: 平均 0 の残差項

である。

さて、この死亡モデルの分類であるが、最終的な表現はそれぞれ異なっているものの、その作成過程においては互いにオーバーラップする部分も存在している。例えば、Coale and Demeny (1983) では、生命表の形状である 4 種類 (North, South, East and West) と性別に、 ${}_nq_x$ と $\log_{10}(10000{}_nq_x)$ を e_{10} で回帰し、

$${}_nq_x = A_x + B_x e_{10}$$

$$\log_{10}(10000{}_nq_x) = A'_x + B'_x e_{10}$$

という形で回帰係数 A_x, B_x などとして求め、これを基に各レベルの生命表関数を作成している。ここで、この推定されたパラメータ A_x, B_x は、リー・カーター・モデルの a_x, b_x と本質的には同じ役割を果たしている。また、高齢部分の死亡率の作成では、数学的関数によるモデルである Gompertz モデルが仮定されている。

また、United Nations (1982) では、5 種類の形状 i に含まれる各種生命表 j の死亡率 ${}_nq_x$ のロジット

$${}_nD_x^{ij} = \text{logit}{}_nq_x = \frac{1}{2} \ln \left(\frac{{}_nq_x}{1 - {}_nq_x} \right)$$

に対して、主成分分析を行って第 1 主成分に対応した項を取ることで、

$${}_nY_x^{ij} = {}_n\bar{Y}_x^i + a_{1j}U_{1x}$$

という形 (ただし、 ${}_nY_x^{ij}$ はモデル化された死亡率のロジット、 ${}_n\bar{Y}_x^i$ は形状 i のクラスターに含まれる ${}_nY_x^{ij}$ の平均) で表すことにより数表を作成しており、この手続きはリー・カーター・モデルと同じものとなっている。

このように、数表によるモデルであっても、その作成過程では、リレーショナル・モデルや数学的関数によるモデルが用いられていることがあることがわかる。

現在では、コンピュータを用いることにより、モデルにある程度複雑な計算過程が含まれるものであっても、ユーザーが必要なパラメータを設定しさえすれば、推計結果を比較的簡単に得ることができようになっており、モデル生命表の提示にあたって、数表による形式よりも、いくつかのパラメータを投入することで生命表関数を得られるリレーショナル・モデルを基本とする方式の方がユーザーにとっても活用がしやすい状況になっている。そこで、このようなリレーショナル・モデルを基本とするモデル生命表作成の試みについて次に述べることにしたい。

2 リレーショナル・モデルを用いたモデル生命表

近年、リレーショナル・モデルを基本とするモデル生命表がいくつか提案されているが、本稿ではその中から、Wilmoth et al. (2012)、Clark (2019) についてレビューを行う。

2.1 Wilmoth らの flexible model

Wilmoth et al. (2012) は、間接推定のための新たなモデル生命表 (flexible two-dimensional mortality model, 以下、flexible model と呼ぶ) を開発して提案している。

このモデルは、

$$\log(m_x) = a_x + b_x h + c_x h^2 + \nu_x k$$

という形を取り、 h は $\log({}_5q_0)$ であり死亡率のレベルを表すパラメータ、 k は通常 $(-2, 2)$ の範囲の値を取り、通常のパターンからの乖離を示すシェイプに関するパラメータである。 a_x, b_x, c_x, ν_x は、Human Mortality Database(HMD) における 719 の生命表に基づき、ウエイト付き最小二乗法と特異値分解を用いて推定されている。

この flexible model は、少ない情報から様々な年齢パターンを効率的に表現することができ、これまでに提案されているモデル生命表よりも優れた、あるいは少なくとも同程度のパフォーマンスを発揮するとされている。

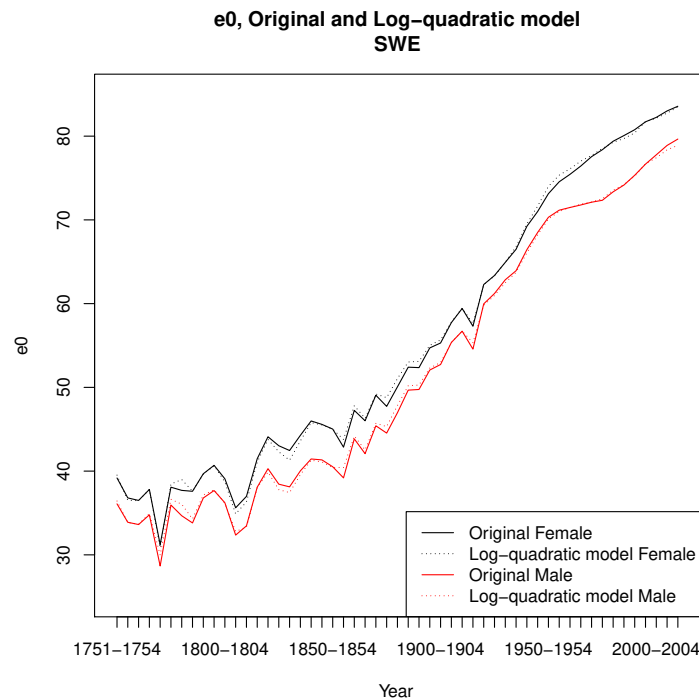


図1 flexible model による e_0 の推定 (スウェーデン)

図1は、この flexible モデルを用いてスウェーデンの e_0 を推定した結果と、実績の e_0 と一緒にグラフにしたものである。実線が実績の e_0 を、点線が flexible model によって推定された e_0 を示しているが、極めて長期間の推定にも関わらず、両者は概ね一致していることが観察できる。

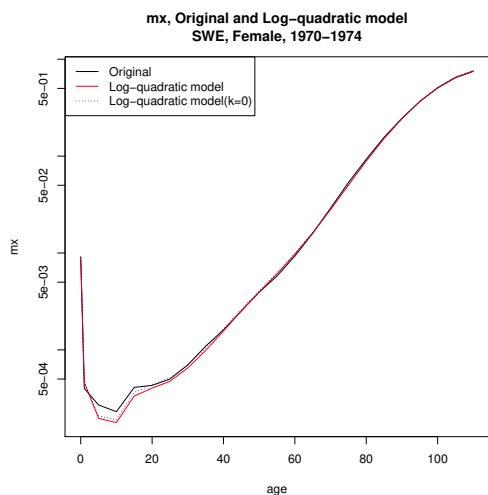


図2 flexible model による m_x の推定 (スウェーデン, 女性, 1970-1974 年)

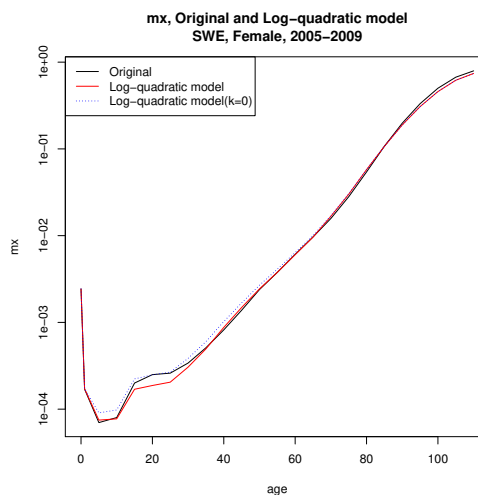


図3 flexible model による m_x の推定 (スウェーデン, 女性, 2005-2009 年)

また、図2、3は、スウェーデン女性の、1970-1974年、2005-2009年の m_x についての推定を行った結果を示したものである。ここで、黒の実線が実績の m_x 、赤の実線が flexible model による推定値、青の点線は flexible model において、 $k = 0$ という制約の下に推定した推定値となっているが、いずれの推定値についても実績との当てはまりがよいことが観察される。このように、flexible model は極めて長期的な死亡率の動向を、その形状も含めてモデル化することに成功していることがわかる。

なお、このような優れたモデル生命表の開発には、HMD プロジェクトによって、一定のクオリティが保証された同一形式の膨大な量の生命表がデータベース化されたことも大きく貢献しているといえよう。

2.2 Clark の SVD-comp model

一方、Clark (2019) は、特異値分解を用いて死亡率のモデリングを行った、“SVD Component model”(以下、SVD-Comp model と略す) を提案した。

今、 \mathbf{Q}_z を性 $z \in \{\text{female, male}\}$ の死亡率を格納した $A \times L$ 行列であるとする。ただし、 A は年齢階級の数、 L は生命表の数を表し、 $l \in \{1, 2, \dots, L\}$ で、各生命表を添字を付けて表す。 \mathbf{Q}_z を特異値分解することにより、

$$\text{SVD}(\mathbf{Q}_z) = \mathbf{U}_z \mathbf{S}_z \mathbf{V}_z^T = \sum_{i=1}^{\rho} s_{zi} \mathbf{u}_{zi} \mathbf{v}_{zi}^T$$

という表現を得る。ただし、 \mathbf{U}_z は左特異ベクトル \mathbf{u}_{zi} を列方向に並べた行列、 \mathbf{V}_z は右特異ベクトル \mathbf{v}_{zi} を列方向に並べた行列、 \mathbf{S}_z は特異値を成分とする対角行列であり、 $\rho = \text{rank}(\mathbf{Q}_z)$ である。 \mathbf{v}_{zi} の第 l 成分を v_{zli} と書くとき、性 z 、生命表 l の年齢階級別死亡確率 \mathbf{q}_{zl} は、特異値分解における c 項目まで ($c \leq \rho$) の和を用いて、

$$\mathbf{q}_{zl} \approx \sum_{i=1}^c v_{zli} \cdot s_{zi} \mathbf{u}_{zi}$$

と書けることになる。Clark (2019) では、 $c = 4$ で、HMD に含まれる死亡確率の近似に十分であるとしている。

また、flexible model のように、 ${}_5q_0$ や ${}_{45}q_{15}$ を用いてパラメータ推定を行う観点から、 $s_{zi} \mathbf{u}_{zi}$ を固定する一方で、 v_{zli} を ${}_5q_0$ や ${}_{45}q_{15}$ によって回帰し、その係数を固定することによって、モデル生命表の推定を行う方法を提案している。

このように、Clark (2019) のモデルはリー・カーター・モデルよりも高次の特異値に対応する成分を用いながら、flexible model のような、 ${}_5q_0$ や ${}_{45}q_{15}$ を用いて推定のできるモデル生命表を提案したものと位置付けることができよう。

3 わが国の死亡率推計のためのモデル生命表

第 2.1 節においては、flexible model がスウェーデンの死亡率を長期にわたって表現することが可能であることを見た。それでは、このモデルは日本の死亡率推計にも有用だろうか。

図 4 は、この flexible モデルを用いて日本の e_0 を推定した結果と、実績の e_0 と一緒にグラフにしたものである。実線が実績の e_0 を、点線が flexible model によって推定された e_0 を示している。日本のデータに当てはめた場合、スウェーデンとは異なり、1980 年代以降で実績との間に乖離が生じていることが観察される。

また、図 5、6 は、日本女性の、1970-1974 年、2005-2009 年の m_x についての推定を行った結果を示したものである。これを見ると、1970-74 年については、モデルの実績値への当てはまりはかなりよいことが観察されるのに対し、2005-2009 年では、黒の実線で示された実績値よりも、flexible model の推定値がかなり上回っていることがわかる。図 4 における e_0 の乖離は、このような高齢死亡率におけるモデル当てはめの問題点がある原因となっていることが考えられ、特に近年の日本の死亡率に対して、flexible model を用いることはあまり適切ではないと考えられる。

そこで、flexible model を日本の死亡率に合うように修正し、日本の死亡率に対するモデル生命表の開発を試みたものが、筆者らによる堀口 (2021) である。堀口 (2021) では、先に示した日本の高齢死亡率に対する flexible model の問題点の構造を調べる観点から、日本版死亡データベースに記載されている都道府県別生命表を用い、flexible model による推定値と実績値の乖離を観察したところ、高齢層でベル状の年齢パターンを持ってお

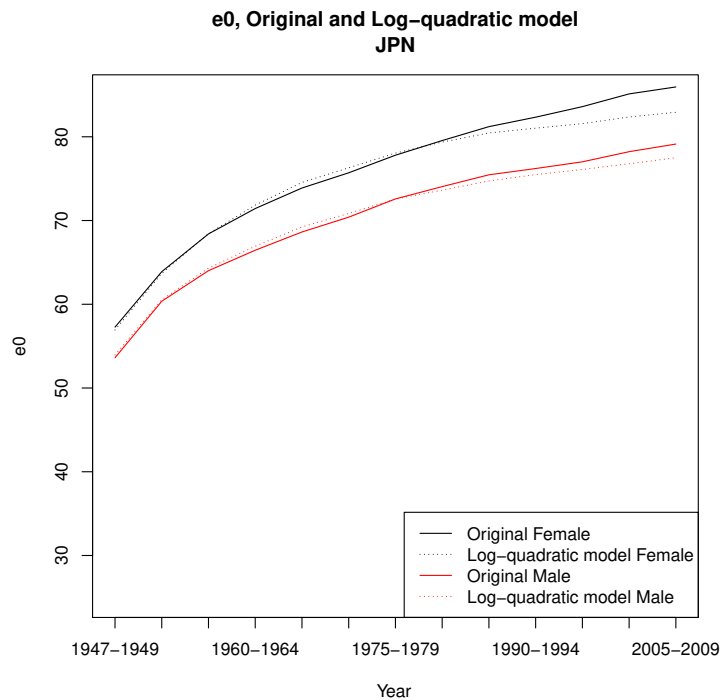


図4 flexible model による e_0 の推定 (日本)

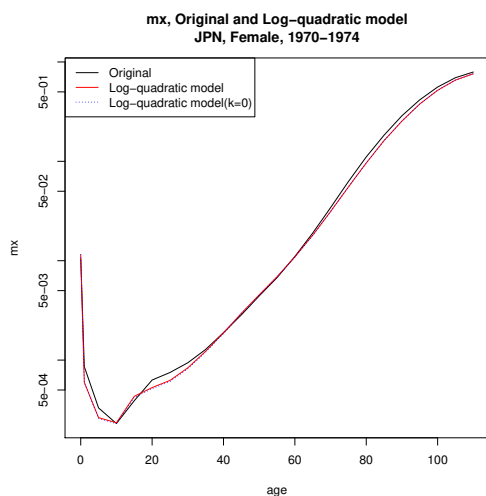


図5 flexible model による m_x の推定 (日本, 女性, 1970-1974年)

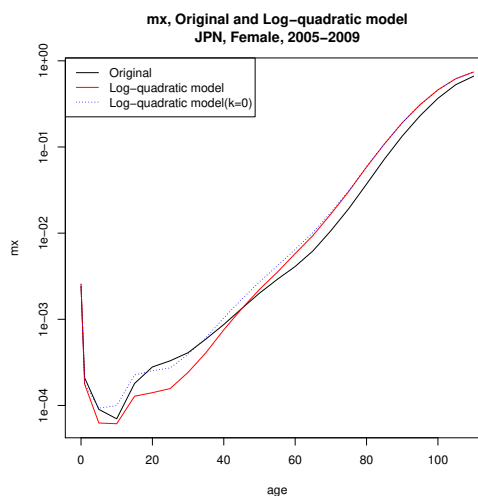


図6 flexible model による m_x の推定 (日本, 女性, 2005-2009年)

り、実績の平均寿命が伸長するほどその山が高くなるような傾向が見られた。

そこで、flexible model と実績の平均二乗誤差 (MSE) を計算し、 e_{65} との関係を示したものが図7である。これを見ると、MSE は e_{65} が大きくなるほど大きい傾向が観察される。

この観察に基づき、flexible model と実績の乖離を特異値分解し、その第一特異値に対応する項でこの乖離を近似することで flexible model を改善することが可能であると考え

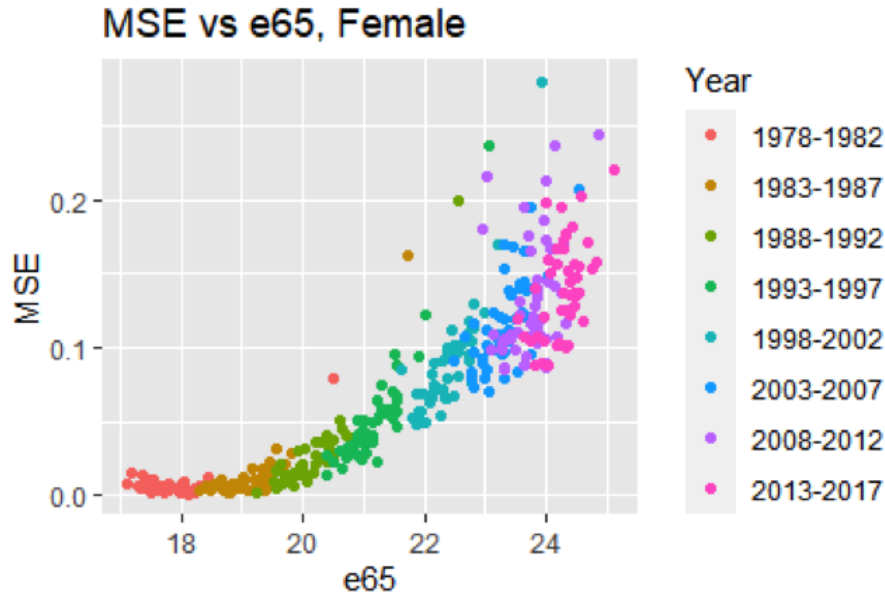


図7 e_{65} と flexible model と実績の MSE との関係 (女性) (堀口 (2021) より引用)

られる。そこで、堀口 (2021) では、これを修正項として新たにモデルに付加することにより、flexible model を修正することを提案した。以下、これを修正モデルと呼ぶ。

修正モデルは以下のような式で表される。

$$\log(m_x) = a_x + b_x h + c_x h^2 + \nu_x k + u_x r$$

ここで、右辺第4項までが flexible model に相当する部分であり、修正モデルではこれに第5項 $u_x r$ が追加されている。ここで、 a_x, b_x, c_x, ν_x については flexible model のパラメータをそのまま用い、 u_x については、特異値分解を用いて推定されたもので固定することとする。そして、この修正モデルを用いて、生命表の推計をするためには、まず、flexible model と同様に、 $h = \log_5 q_0$ とし、さらに k を $\log_{45} q_{15}$ を再現するような値として数値解析的に求める。次に、 r を e_{65} を再現するような値として同じく数値解析的に求めることにより、生命表関数が推定可能となる。

近年、日本では ${}_5q_0$ や ${}_{45}q_{15}$ はかなり低いレベルとなっており、平均寿命の伸長は高齢死亡率改善によるところが大きいことから、 e_{65} をある程度適切に推計することができれば、この修正モデルを用いることにより全ての生命表関数の推計が可能となる。

そこで、修正モデルを市区町村レベルの生命表推計に応用する例として、埼玉県和光市の女性 (2015 年) の死亡率推計を試みた結果が図8である。ここでは、2015 年の e_{65} , ${}_{45}q_{15}$, ${}_5q_0$ について、 e_{65} については、和光市の 2010 年 e_{65} の実績値に、埼玉県の女性 (JMD) の e_{65} の 2003-2007 年から 2008-2012 年への増加率を乗じて推計し、 ${}_{45}q_{15}$, ${}_5q_0$ については和光市の 2010 年実績値を固定して用いている。図からも明らかな通り、 e_{65} を増加率を用いて単純に推定しただけであっても、実績値をかなりよく表現できていることがわかる。このように、修正モデルは、自治体が将来人口推計を行う際に必要となる生

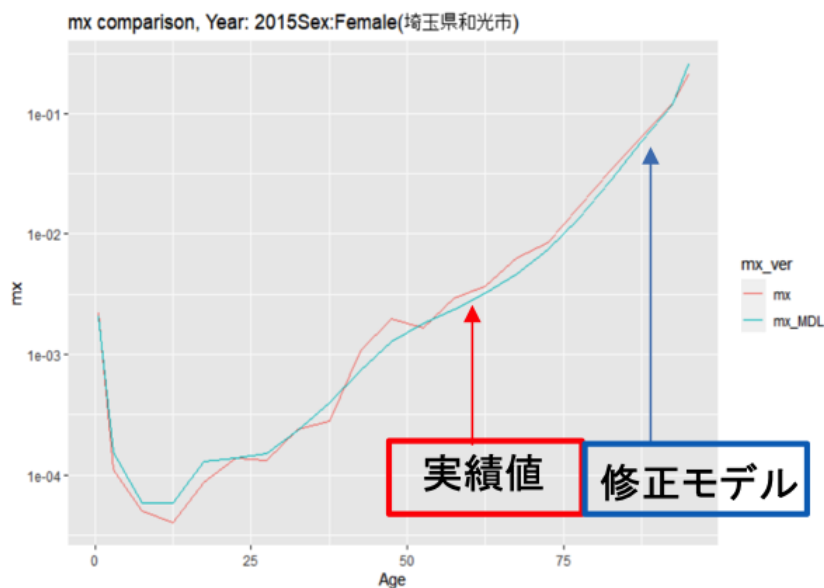


図8 死亡率推計結果 (埼玉県和光市女性, 2015年) (堀口 (2021) より引用)

残率の設定などに極めて有用であると考えられる。

おわりに

本研究では、死亡率のモデリングと、近年のリレーショナル・モデルを基本とするモデル生命表についてレビューを行うとともに、これらを日本の死亡率推計に適合させるための修正を行ったモデル開発に関する試みについて述べた。

具体的には、死亡の年齢パターンモデルとして、(1) 数学的関数によるモデル、(2) 数表によるモデル、(3) リレーショナルモデルの3種類が存在するが、近年、リレーショナル・モデルを基本とするモデル生命表がいくつか提案されており、その例として、Wilmoth et al. (2012) の flexible model と Clark (2019) の SVD-comp model に関するレビューを行った。また、flexible model を日本の死亡率に合うように修正し、日本の死亡率に対するモデル生命表の開発を試みた堀口 (2021) の修正モデルについて述べ、埼玉県和光市の女性 (2015年) の死亡率を修正モデルを用いて推計した例を観察することで、このモデルは自治体が将来人口推計を行う際に必要となる生残率の設定などに極めて有用であると考えられることを見た。

このように、わが国の死亡率推計に有用であると考えられる修正モデルであるが、問題点もないわけではない。現在の修正モデルは flexible model のパラメータを用いることで推計を行っており、若年層の死亡パターンは概ね flexible model によって規定されていると考えられるが、特に近年の日本では若年死亡のレベルも諸外国に比べてかなり低いものとなっており、flexible model に基づく若年死亡パターンが日本での表現に必ずしも優れているとは言えない面があることである。また、修正モデルはあくまでも都道府県別生命

表に基づいてモデリングされているが、自治体としてはさらに規模の小さい市区町村での活用も必要とされることから、このような小規模自治体における死亡率推計の安定性についてはさらなる検討が必要であると考えられる。このような問題については、今後の課題としていきたい。

謝辞

本研究は、厚生労働行政推進調査事業費補助金（政策科学総合研究事業（政策科学推進研究事業））「国際的・地域的視野から見た少子化・高齢化の新潮流に対応した人口分析・将来推計とその応用に関する研究」（課題番号：H29-政策-指定-003，研究代表者：石井太）、および「長期的人口減少と大国際人口移動時代における将来人口・世帯推計の方法論的発展と応用に関する研究」（課題番号：20AA2007，研究代表者：小池司朗）による助成を受けた。

参考文献

- 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>.
- 堀口侑 (2021) 「日本のモデル生命表の開発と地域別生命表推計への応用」, 日本人口学会第73回大会口頭報告(予定), 東京大学, 2021年6月.
- Beard, R. E. (1971) "Some aspects of theories of mortality, cause of death analysis, forecasting and stochastic processes", in *Biological Aspects of Demography*, London: Taylor & Francis Ltd.
- Brass, W. (1971) "On the Scale of Mortality", in W. Brass ed. *Biological Aspects of Demography*: Taylor and Francis Ltd, pp. 69–110.
- Clark, S. J. (2019) "A general age-specific mortality model with an example indexed by child mortality or both child and adult mortality", *Demography*, Vol. 56, No. 3, pp. 1131–1159.
- Coale, A. J. and P. Demeny (1983) *Regional Model Life Tables and Stable Populations, 2nd Edition*, New York: Academic Press.
- Gompertz, B. (1825) "On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies", *Philosophical Transactions of the Royal Society of London*, Vol. 115, pp. 513–583.
- Human Mortality Database. University of California, Berkeley (USA) and Max Planck Institute for Demographic Research (Germany). Available at www.mortality.org or www.humanmortality.de.
- Lee, R. and L. Carter (1992) "Modeling and Forecasting U.S. Mortality", *Journal of the American Statistical Association*, Vol. 87, No. 419, pp. 659–675.
- Makeham, W. M. (1860) "On the Law of Mortality and the Construction of Annuity Tables", *The Assurance Magazine, and Journal of the Institute of Actuaries*, Vol.

- 8, No. 6, pp. 301–310.
- Perks, W. (1932) “On some experiments on the graduation of mortality statistics”, *Journal of the Institute of Actuaries*, Vol. 63, pp. 12–40.
- United Nations (1982) *Model Life Tables for Developing Countries, Population Studies No.77*.
- Wilmoth, J., S. Zureick, V. Canudas-Romo, M. Inoue, and C. Sawyer (2012) “A flexible two-dimensional mortality model for use in indirect estimation”, *Population studies*, Vol. 66, No. 1, pp. 1–28.