

LC-VARモデルと死亡率予測・リスク評価

井川 孝之 総合研究大学院大学 統計科学専攻 5年一貫制博士課程5年

1. はじめに

1990年代前半に開発されたLee-Carter(LC)モデルは国際的に死亡率推計やリスク評価等に用いられている。LCモデルを日本の死亡率へ適用すると、年齢と時代(期間)を軸にとった残差局面にはうねりのようなものが観察される。ここでは、この残差構造を解析し開発したLC-VARモデルを紹介し、英米の死亡データへ適用した場合の適合度を示す。また、LC-VARモデルの死亡率予測・リスク評価への応用例を示す。

2. LCモデルと既存の拡張モデル

(1) LCモデル

年齢 x についての平均的な対数死亡率を表すパラメータ α_x 、年 t についての死亡の一般的な水準を表すパラメータ(死亡指数) κ_t 、 κ_t に係る年齢別パラメータ β_x とし、中央死亡率を $m_{x,t}$ 、 $\varepsilon_{x,t}$ を残差項とすると、LCモデルは次式の通り表される。

$$\ln m_{x,t} = y_{x,t} = \alpha_x + \beta_x \kappa_t + \varepsilon_{x,t} \quad (2.1)$$

$$\kappa_t = \lambda + \kappa_{t-1} + \omega_t \quad (2.2)$$

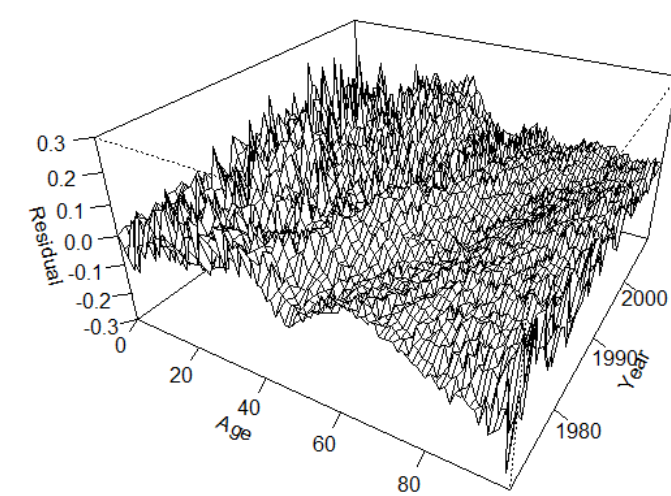
κ_t にはARIMAモデル等の時系列モデルが適用され、それにより期間の経過に伴う死亡率の低下傾向を表せることがLCモデルの特徴の1つとなっている。上式(2.2)はドリフト付ランダムウォークモデルを適用した場合を表す。パラメータ推定にあたっては、識別可能とするため $\sum \beta_x = 1, \sum \kappa_t = 0$ の制約が置かれる。

(2) LCモデルの推定と残差

カリフォルニア大学とマックスプランク人口研究所等が共同で構築しているHuman Mortality Database (HMD) の1971-2009年の男・死亡数を用いて、LCモデルのパラメータを観察死亡数 $D_{x,t}$ がポアソン分布に従うとし最尤推定すると、年齢と時代(期間)を軸に取った残差局面は図1の通りとなる。残差局面にはうねりのようなものが観察され、LCモデルでは捉えきれない効果がある可能性がある。

Willets(2004)は、英国の死亡率と同様、日本の死亡率においても生年コーホート別の死亡率改善効果が見られるとしているが、Ishii(2008)は、死亡が遅延する効果があると考えLCモデルを修正し、日本の将来推計人口へ利用している。

図1 LCモデルの残差 (男・1971-2009年)



(3) 既存のLC拡張モデル

コーホート効果を考慮した既存モデルとして、以下のものがある。

a. Renshaw and Haberman (RH) モデル

LCモデルを拡張しコーホート効果を織り込んだ以下のRHモデルが提案されている。コーホート効果を表す生年コーホート別のパラメータを l_c 、 l_c に係る年齢別パラメータを $\beta_x^{(2)}$ とし、パラメータ推定は、識別可能とするため、一定の条件が置かれる。

$$\ln m_{x,t} = \alpha_x + \beta_x^{(1)} \kappa_t + \beta_x^{(2)} l_{c,t-x} + \varepsilon_{x,t} \quad (2.3)$$

b. APCモデル

APCモデルはLCモデルより前に開発され、医療統計や社会調査等において利用されてきた。APCモデルは、総平均を年齢の主効果に組み入れ、以下の通り記述される。パラメータ推定にあたっては、識別可能とするため、一定の条件が置かれる。

$$\ln m_{x,t} = \alpha_x + \kappa_t + l_c + \varepsilon_{x,t} \quad (2.4)$$

3. LC-VARモデル

日本の死亡率へLCモデルを適用した場合の残差構造を踏まえ策定したLC-VARモデルは以下の通りである。

$$M_t = A_x + B_x \cdot \kappa_t + Z_t \quad (3.1)$$

$$M_t = \begin{bmatrix} \ln m_{0t} \\ \ln m_{1t} \\ \vdots \\ \ln m_{ot} \end{bmatrix} \quad A_x = \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_o \end{bmatrix} \quad B_x = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_o \end{bmatrix} \quad \kappa_t = \lambda + \kappa_{t-1} + \eta_t \quad \eta_t \sim N(0, \sigma_\eta)$$

$$Z_t = \sum_{L=1}^q C_L Z_{t-L} + W_t \quad C_L = \begin{bmatrix} C_{(0,0)}^L & \cdots & C_{(0,o)}^L \\ \vdots & \ddots & \vdots \\ C_{(o,0)}^L & \cdots & C_{(o,o)}^L \end{bmatrix}$$

$$Z_t = \begin{bmatrix} \varepsilon_{0t} \\ \varepsilon_{1t} \\ \vdots \\ \varepsilon_{ot} \end{bmatrix} \quad W_t = \begin{bmatrix} W_{0t} \\ W_{1t} \\ \vdots \\ W_{ot} \end{bmatrix} \quad w_{it} \sim N(0, \sigma_i) \quad \text{cov}(w_{jt}, w_{it}) = 0$$

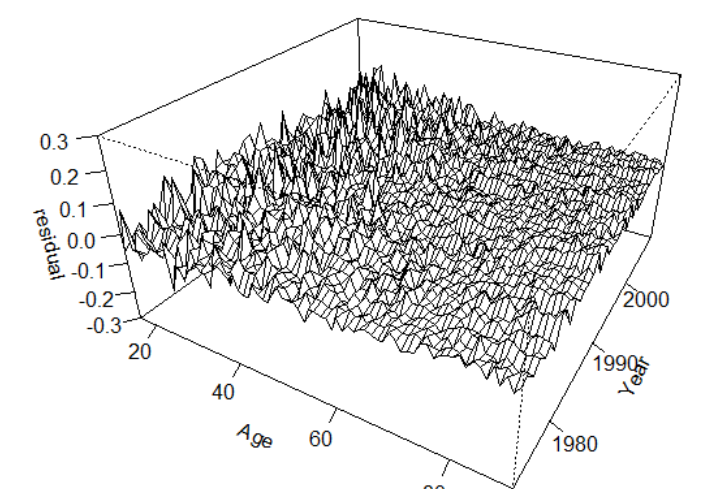
4. 各モデルの推定結果と比較

前述のHMDの1971-2009年、1951-2009年の男・死亡数(14-90歳)のデータについて、各モデルを最尤推定した場合のBIC、標準化誤差の分散は、下表の通りである。RHモデルの推定及びAPCモデルの推定は、Cairns et al (2009)と同様の識別条件を設定した。

表1 各モデルの推定結果

対象データ	モデル	パラメータ数	BIC	標準化誤差の分散
1971-2009年 14-90歳 男	LC	191	56,697	9.09
	APC	229	50,503	6.86
	RH	381	39,523	2.62
	LC-VAR(1)	303	39,051	2.58
	LC-VAR(2)	378	38,591	2.26
1951-2009年 14-90歳 男	LC	211	126,098	18.40
	APC	269	123,119	18.22
	RH	421	65,673	4.17
	LC-VAR(1)	323	64,085	3.89
	LC-VAR(2)	398	64,610	3.96

図2 LC-VAR(1) モデルの残差(VAR対象年齢16-90歳・1971-2009年)



5. 英米の死亡データへの適用

英国・米国の死亡データ(HMDの1951-2009年の男・死亡数(14-90歳))へ各モデルを適用した場合の適合度は表2・表3の通りである。米国については、LC-VARモデルが良く適合している。

表2 英国の死亡データ(男)

モデル	BIC	標準化誤差の分散
LC	72,661	6.70
APC	62,661	4.33
RH	54,319	2.14
LC-VAR(1)	56,760	2.87
LC-VAR(2)	60,023	3.60

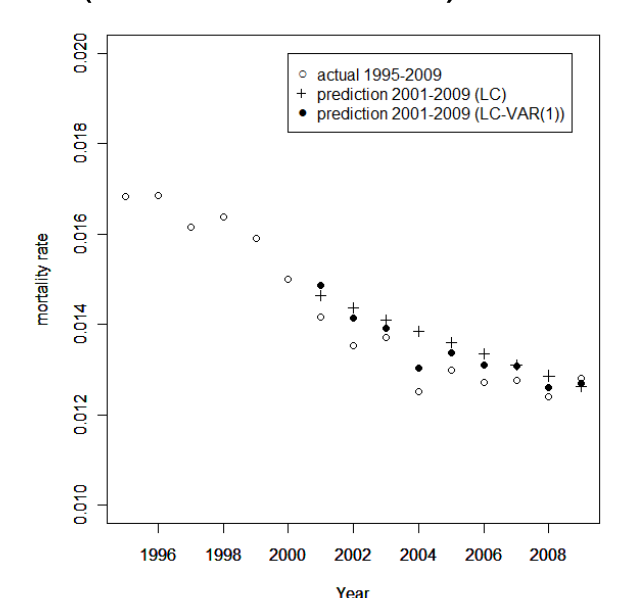
表3 米国の死亡データ(男)

モデル	BIC	標準化誤差の分散
LC	158,240	24.80
APC	150,756	23.14
RH	92,892	9.66
LC-VAR(1)	80,677	6.75
LC-VAR(2)	84,480	7.64

6. 死亡率予測

図3は、1971-2000年の男・死亡数のデータ(14-90歳)に基づき推定したLC-VAR(1)モデルにより2001-2009年の65歳死亡率を予測し実際の値と比較したバックテストの結果である。LCモデルの短所である短期予測の悪さがLC-VAR(1)モデルにより改善されることが分かる。

図3 バックテストの結果 (男・65歳死亡率)

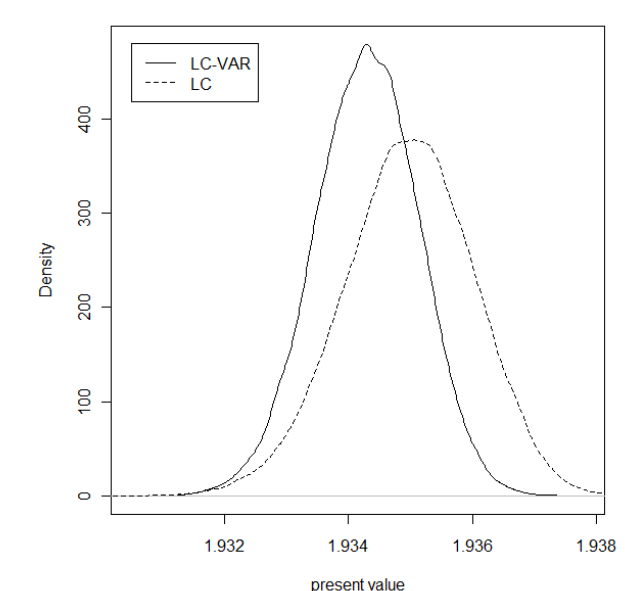


7. リスク評価への応用

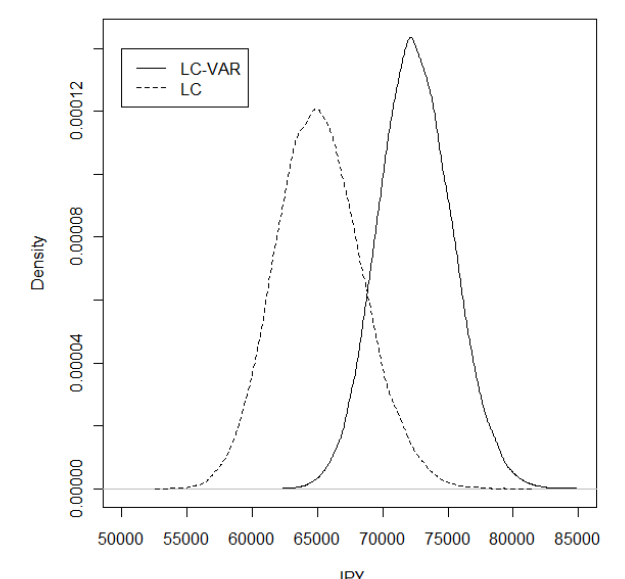
LC-VARモデルを年金債務評価及び一時払保険料の評価へ応用しLCモデルによる場合と比較した例を図4に示す。年金給付1円あたりの現価(年金現価率)の分布は、年金債務のリスクをVaRやCTE等により評価する際の基礎とできる。年金支給期間が短期の場合はLC-VARモデルとLCモデルによる結果の差異が大きい。一時払保険料の分布についても、短期の契約期間で保険金額が大きい場合は、両モデルによる結果の差異は大きなものとなっている。

図4 LC-VARモデルの応用例

年金現価の分布



一時払保険料の分布



8. おわりに

LCモデルの残差構造解析により策定したLC-VARモデルは、日本や米国の死亡データに対し、LCモデルや既存のLC拡張モデルより良好な適合度を提供します。LC-VARモデルは、LCモデルと比較し、良好な短期予測を提供し、リスク評価への応用も可能である。リスク評価において、個々人の死亡確率の異質性を考慮すること等は課題となろう。

文献

- [1] Lee, R. D., L. Carter (1992). "Modeling and forecasting the time series of US mortality," *Journal of the American Statistics Association* 87: 659-671.
- [2] Renshaw, A. E., S. Haberman (2006). "A cohort-based extension to the Lee-Carter model for mortality reduction factors," *Insurance: Mathematics and Economics* 38: 556-570.
- [3] Cairns, A. J. G., D. Blake, K. Dowd, G. D. Coughlan, D. Epstein, A. Ong, I. Balevich (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.
- [4] 井川孝之, 川崎能典(2012). Lee-Carterモデルの残差構造の解析とモデリング, 2012年度統計関連学会連合大会講演報告集, 128