Extensions to the Lee-Carter model, including risk measurement in the age-period and age-period-cohort versions of the model.

Arthur E. Renshaw and Steven Haberman.

Cass Business School, City University, London, EC1Y 8TZ, UK.

To be presented at meetings with the CMI Bureau in Edinburgh and London.
April 2008.

Abstract

We describe the age-period-cohort version of the Lee-Carter model and how it can be extended to incorporate heterogeneity through modeling of the scale parameter. In applications, it is important to be able to estimate measures of uncertainty – for example, prediction intervals. Risk measurement involving the repeated fitting of the age-period-cohort parametric structure to mortality rates is not practical due to the slow rate of convergence of the iterative fitting algorithm. We present some key findings from a comparative study of three such boot-strapping methods, which have been described in the literature and have been applied to the basic age-period Lee-Carter parametric structure. We identify and correct for the limited prominence given to the simulation of the forecast error in the period component of the model structure, treated as a time series. We then discuss the implications of this correction for age-period-cohort modeling. In the talk, we will present several numerical examples in order to illustrate the main points.

Notation:

Mortality data:-

$\{(d_{xt}, e_{xt}) : t = t_1, t_2, ..., t_n, x = x_1, x_2, ..., x_k\}$

d_{xt} - reported deaths, age *x*, period *t*

e_{xt} - matching exposure to risk of death

ω_{xt} - 0/1 (empty cell) indicator weights

Possible targets:-

μ_{xt} - force of mortality

q_{xt} - probability of death

ϕ_{xt} - dispersion parameter

Modeling I:

Target	μ_{xt}
Response	D_{xt}
$E(D_{xt})$	$e_{xt}\mu_{xt}$
$Var(D_{xt})$	$\phi_{xt} \dfrac{V\{E(D_{xt})\}}{\omega_{xt}}$
Link	$\log E(D_{xt}) = \eta_{xt}$
Predictor	$\eta_{xt} = \log e_{xt} + \log \mu_{xt}$
Structure	$LC : \log \mu_{xt} = \alpha_x + \beta_x \kappa_t$
	$M : \log \mu_{xt} = \alpha_x + \beta_x^{(0)} \iota_{t-x} + \beta_x^{(1)} \kappa_t$
Offset	$LC : \log e_{xt}$
	$M : \log e_{xt} + \alpha_x$

Distribution	*Variance function*	*Dispersion*
Poisson	$V(u) = u$	$\phi_{xt} = 1 \ \forall \ (x,t)$
Negative binomial	$V(u) = u + \lambda_x u^2$	$\phi_{xt} = 1 \ \forall \ (x,t)$
Joint Poisson/gamma	$V(u) = u$	$\log \phi_{xt} = \varsigma_x$

Modeling II:

Target	q_{xt}
Response	D_{xt}
$E(D_{xt})$	$e^i_{xt} q_{xt}$

Var $\left(D_{xt}\right)$	$\phi_{xt}\dfrac{V\left\{E\left(D_{xt}\right)\right\}}{\omega_{xt}}$
Links	*comp. l-l*: $\log\left\{-\log\left(1-q_{xt}\right)\right\}=\eta_{xt}$
	log-odds: $\log\left(\dfrac{q_{xt}}{1-q_{xt}}\right)=\eta_{xt}$
	probit: $\Phi^{-1}\left(q_{xt}\right)=\eta_{xt}$
Structure	$LC:\eta_{xt}=\alpha_x+\beta_x\kappa_t$
	$M:\eta_{xt}=\alpha_x+\beta_x^{(0)}\iota_{t-x}+\beta_x^{(1)}\kappa_t$
Offset	$M:\alpha_x$

Distribution	*Variance function*	*Dispersion*
Binomial	$V(u)=u\left(1-u/e_{xt}^{i}\right)$	$\phi_{xt}=1\ \forall\ (x,t)$
Joint binomial/gamma	$V(u)=u\left(1-u/e_{xt}^{i}\right)$	$\log\phi_{xt}=\varsigma_x$

Comments:

(i) The emphasis in this presentation is on the targeting of μ_{xt} rather than q_{xt}.

(ii) The use of the complementary log-log link would appear to be the natural choice in the context of mortality rate modeling, given the approximate relationship, $-\log\left(1-q_{xt}\right)\approx\mu_{xt}$, between q_{xt} and μ_{xt}. It has also played a prominent role in the construction of static life tables by the CMI Bureau.

Findings:

(i) Life expectancy projections using Poisson log-link modeling to target μ_{xt} are found to be in close agreement with life expectancy projections using binomial complementary log-log link modeling to target q_{xt}.

References:

Renshaw & Haberman (2008c); Forfar, McCutcheon & Wilkie (1988).

Model Fitting:

By optimizing the model deviance or log-likelihood.

Define:-

r_{xt} - deviance residual

Model Extrapolation:

By time series forecasting applied to the period (and cohort) model component(s).

Statistics of Interest:

These include

$e_x(t)$ - life expectancy

$a_x(t)$ - fixed rate annuity

requiring projected mortality rates

$q_{xt} \approx 1 - \exp(-\mu_{xt})$ or $\mu_{xt} \approx -\log(1 - q_{xt})$.

Computation is either by fixed cohort using

$$e_x(t) = \frac{\sum_{i \geq 1} l_{x+i}(t+i)\left\{1 - \tfrac{1}{2} q_{x+i,t+i}\right\}}{l_x(t)}$$

$$a_x(t) = \frac{\sum_{i \geq 1} l_{x+i}(t+i)\nu^i}{l_x(t)},\ \nu \text{ - discount factor}$$

where

$$l_{x+1}(t+1) = \{1 - q_{xt}\} l_x(t),$$

or by fixed period *t*.

Simulating Prediction Intervals (PIs)

Context: Poisson LC modeling with random walk period component.

Algorithm A (semi-parametric bootstrap)

For simulations $m = 1, 2, ..., M$

1. $\forall\ x, t$
   simulate responses d^*_{xt} by sampling $Poi(\hat{d}_{xt})$,
   preserving any empty data cells.
2. Obtain estimates $\hat{\alpha}^*_{x,m}, \hat{\beta}^*_{x,m}, \kappa^*_{t,m}$ by fitting $d^*_{xt,m}$,
   same structure.
3. For $k = 1, 2, ..., K$
   set $\dot{\kappa}^*_{t_n+k,m} = \hat{\kappa}_{t_n} + k\hat{\theta}^*_m$.
4. Compute statistics of interest.

Algorithm C (residual bootstrap)

As in A above, subject to the Stage 1 replacement with

1. $\forall\ x, t$

a. randomly sample $r^*_{xt,m}$ from $\{r_{xt}\}$ with replacements, preserving any empty data cells.
b. map $r^*_{xt,m} \mapsto d^*_{xt,m}$.

Algorithm B (parametric Monte-Carlo)

Define $\boldsymbol{\psi}^T = (\boldsymbol{\alpha}^T_x, \boldsymbol{\beta}^T_x, \boldsymbol{\kappa}^T_t)$ - a basis vector of parameters. Then as in A above, subject to the stage 1 and 2 replacement with

1. Simulate a vector of N(0,1) errors $\boldsymbol{\varepsilon}^*$.
2. Obtain estimates $\boldsymbol{\psi}^* = \hat{\boldsymbol{\psi}} + \sqrt{\phi}\ \boldsymbol{\vartheta}\boldsymbol{\varepsilon}^*$, where $\boldsymbol{\vartheta}$ - Cholesky factorization of the variance-covariance matrix.

References:

Brouhns, Denuit & Vermunt (2002); Brouhns, Denuit & van Keilegom (2005); Koissi, Shapiro & Hognas (2006); Renshaw & Haberman (2008a).

Comments:

(i) The algorithms adapt to other modeling distributions and time series models.

(ii) It is not practical to adapt the simulation algorithms to age-period-cohort structures M, because of the slow rate of convergence of the iterative model fitting algorithm.

Findings:

(i) PIs generated using Algorithm B, are shown to be highly dependent on the basis vector chosen, and the particular constraints used to ensure that the model is identifiable and hence are not considered further.

(ii) Wider PIs are obtained by switching from the Poisson setting to a negative binomial or joint Poisson modeling setting.

(iii) Life expectancy PIs generated by these methods, (random walk time series), are found to be appreciably narrower when compared with their theoretical counterparts (using Denuit (2007)).

(iv) The application of Stage 3 in Algorithm A is shown not to capture the full magnitude of the forecast error in the time series.

References:
Renshaw & Haberman (2008a); Li, Hardy & Tan (2006); Denuit (2007).

Bootstrapping both aspects of the random walk

Recall the period component time series

$$\{\kappa_t : t = 1, 2, ..., t_n\} = \{\kappa_t\}_1^{t_n}$$

For the random walk with drift parameter θ get

$$\Delta\kappa_t = \kappa_t - \kappa_{t-1} = \theta + \varepsilon_t,\ \varepsilon_t \sim N(0, \sigma^2)\ iid,\ t = 2, 3, ..., t_n .$$

1) Estimates are

$$\hat{\theta} = \frac{\sum_{t=2}^{t_n} \Delta\kappa_t}{n-1} = \frac{\kappa_{t_n} - \kappa_1}{n-1},\ \hat{\sigma}^2 = \frac{\sum_{t=2}^{t_n} \hat{\varepsilon}^2}{n-1}$$

with residuals and adjusted residuals

$$\hat{\varepsilon}_t = \Delta\kappa_t - \hat{\theta},\ r_t = \hat{\varepsilon}_t \Big/ \sqrt{\frac{n-2}{n-1}} .$$

CIs for θ and σ^2 based on theory follow. To bootstrap CI's use

Algorithm A1

For simulations $m = 1, 2, ..., M$

1. For $t = 2, 3, ..., t_n$
   (a) randomly sample $r_{t,m}^*$ from $\{r_t\}$ with replacement
   (b) set $\Delta\kappa_{t,m}^* = \hat{\theta} + r_{t,m}^*$
2. Obtain estimates $\hat{\theta}_m^*,\ \hat{\sigma}_m^{2*}$.

2) Forecasting *k* periods-ahead get

$$\dot{\kappa}_{t_n+k} = \kappa_{t_n} + k\theta + \sum_{j=1}^{k} \varepsilon_{t_n+j},\ \varepsilon_{t_n+j} \sim N(0,\sigma^2),\ iid$$

for which

$$E\left[\dot{\kappa}_{t_n+k} \mid \{\kappa_t\}_1^{t_n}\right] = \kappa_{t_n} + k\theta,\ Var\left[\dot{\kappa}_{t_n+k} \mid \{\kappa_t\}_1^{t_n}\right] = k\sigma^2 .$$

PIs for $\dot{\kappa}_{t_n+k}$ based on theory follow. To bootstrap PIs, we try

<u>Algorithm A2</u>

For simulations $m = 1,2,...,M$

1. For $k = 1,2,...,K$
   - (a) randomly sample $z^*_{k,m}$ from $N(0,1)$
   - (b) set $\dot{\kappa}^*_{t_n+k,m} = \kappa_{t_0} + k\hat{\theta} + \sqrt{k}\hat{\sigma}z^*_{k,m}$.

The following is also of interest

<u>Algorithm A3</u>

For simulations $m = 1,2,...,M$

1. For $t = 2,3,...,t_n$
   - (a) randomly sample $r^*_{t,m}$ from $\{r_t\}$ with replacement
   - (b) set $\Delta\kappa^*_{t,m} = \hat{\theta} + r^*_{t,m}$.
2. Obtain estimates $\hat{\theta}^*_m,\ \hat{\sigma}^{2*}_m$.
3. For $k = 1,2,...,K$
   set $\dot{\kappa}^*_{t_n+k,m} = \kappa_{t_0} + k\hat{\theta}^*_m$.

Comments:

(i) In A2, it is sufficient, but not necessary to sample $z^*_{k,m}$ once only for each *m*, in which case $z^*_{k,m} = z^*_m$.

(ii) The A2 approach to simulating predictions in $\{\boldsymbol{\kappa}_t\}_1^{t_n}$ differs from that used in Algorithms A, B, C, hence our interest in A3.

Findings:

(i) PIs generate using A2 are found to be in close agreement with their theoretical equivalents.

(ii) PIs generated using A3 are found to understate the equivalent PIs generated under A2 (and theoretical PIs).

References:

Davison & Hinkley (2006); Renshaw & Haberman (2008b).

Simulating PIs: Reformulation

Context: Poisson LC modeling with random walk period component

The following are of interest:

Algorithm A4

For simulations $m = 1, 2, ..., M$

1. For $k = 1, 2, ..., K$

   (a) randomly sample z_m^* from $N(0,1)$

   (b) $\dot{\kappa}^*_{t_n+k,m} = \kappa_{t_0} + k\hat{\theta} + \sqrt{k}\hat{\sigma}z^*_m$.

2. Compute statistics of interest.

Algorithm A5A

For simulations $m = 1, 2, ..., M$

1. $\forall\ x, t$

   Simulate responses d^*_{xt} by sampling $Poi(\hat{d}_{xt})$,

   preserving any empty data cells.

2. Obtain estimates $\hat{\alpha}^*_{x,m}, \hat{\beta}^*_{x,m}, \kappa^*_{t,m}$ by fitting $d^*_{xt,m}$, same structure.

3. Obtain $\hat{\theta}^*_m$, $\hat{\sigma}^{2*}_m$ as in A1.

For simulations $n = 1, 2, ..., N$

4. Randomly sample z^*_{mn} from $N(0,1)$

5. For $k = 1, 2, ..., K$ set $\dot{\kappa}^*_{t_n+k,m} = \kappa_{t_0} + k\hat{\theta} + \sqrt{k}\hat{\sigma}z^*_{mn}$

6. Compute statistics of interest.

Algorithm A5C

As for A5A above, subject to the appropriate change in Stage 1, in accordance with Algorithm C.

Comments:

(i) Algorithm A4 merely replicates the forecast error in the period component (correctly) while preserving the fitted LC structure and parameter estimates, consistent with the Denuit (2007) theoretical approach based on the distribution of random future life expectancies under the LC structure.

(ii) In addition to correctly bootstrapping the prediction error in the time series, extra provision for bootstrapping the Poisson LC model error is included in Algorithms A5A and A5C.

(iii) Here, (unlike Algorithm A2), it is necessary (as well as sufficient) that sampling from $N(0,1)$ is independent of k (Stage 1a- A4, Stage 4- A5A & A5C) when the statistics of interest are computed by fixed cohort.

Findings:

(i) We provide empirical evidence of the close agreement of (matching) future life expectancy PIs, irrespective of the method of construction, using Algorithms A4, A5A, A5C and the Denuit (2007) theoretical approach.

(ii) The close agreement between simulated and theoretical life expectancy PIs, lends plausibility to similarly constructed bootstrap PIs of other statistics of interest, such as fixed rate annuities, currently not otherwise available by theory.

(iii) The close agreement between simulated PIs under A4 compared with A5A & A5C is indicative of the dominance of the forecast error over and above the model fitting error: a conclusion reached in the original Lee-Carter 1992 paper.

(iv) The use of Algorithm A4 does not require the repeated fitting of the model structure and hence provides a practical means of constructing plausible PIs in the age-period-cohort structured model, where convergence of the fitting algorithm is notoriously slow. We illustrate this.

(v) We have generalized the random walk time series (Stage 1b- A4, Stage 5- A5A & A5C) to an ARI time series, while further generalization to an ARIMA time series is possible.

References:

Renshaw & Haberman (2008b); Denuit (2007); Lee & Carter (1992)

References

Brouhns, N., Denuit, M. and Vermunt, J.K., 2002. Measuring the longevity risk in mortality projections. *Bulletin of the Swiss Association of Actuaries*, 105-130.

Brouhns, N., Denuit, M. and van Keilegom, 2005. Bootstrapping the Poisson log-bilinear model for mortality forecasting. *Scandinavian Actuarial Journal*, 3, 212-224.

Davison, A.C. and Hinkley, D.V. 2006. Bootstrap methods and their application. Cambridge University Press.

Denuit, M. 2007. Distribution of the random future life expectancies in log-bilinear mortality projection models. *Lifetime Data Analysis*, 13, 381-397.

Forfar, D.O., McCutcheon, J.J. and Wilkie, A.D. (1988). On graduation by mathematical formula. *Journal of the Institute of Actuaries*, 115, 1-135.

Koissi, M-C., Shapiro, A.F. and Hognas, G. 2006. Evaluating and extending the Lee-Carter model for mortality forecasting: bootstrap confidence intervals. *Insurance: Mathematics and Economics*, 38, 1-20.

Lee, R.D. and Carter, L., 1992. Modeling and forecasting the time series of US mortality. *Journal of the American Statistical Association (with discussion).* 87, 659-671.

Li, S-H, Hardy, M.R. and Tan, K.S. 2006. Uncertainty in mortality forecasting: an extension of the classical Lee-Carter approach. University of Waterloo, Ontario N2L 3GI (email iipr@icarus.math.uwaterloo.ca).

Renshaw, A.E. and Haberman, S., 2006. A cohort-based extension to the Lee-Carter model for mortality reduction factors. *Insurance: Mathematics and Economics*, 38, 556-570.

Renshaw, A.E. and Haberman, S., 2008a. On simulation-based approaches to risk measurement in mortality with specific reference to Poisson Lee-Carter modeling. *Insurance: Mathematics and Economics*. (To appear).

Renshaw, A.E. and Haberman, S., 2008b. A note on the distribution of random future life expectancies in Poisson log-bilinear mortality projections. (In preparation).

Renshaw, A.E. and Haberman, S., 2008c. On simulation-based approaches to risk measurement in mortality with specific reference to binomial Lee-Carter modeling. Presented to Society of Actuaries' Living to 100 Symposium. Orlando, Florida, USA.