The ilc package: Iterative Lee-Carter

Han Lin Shang

Research School of Finance, Actuarial Studies and Applied Statistics,
Australian National University

Presentation prepared for the useR2015
Aalborg, 1st July 2015

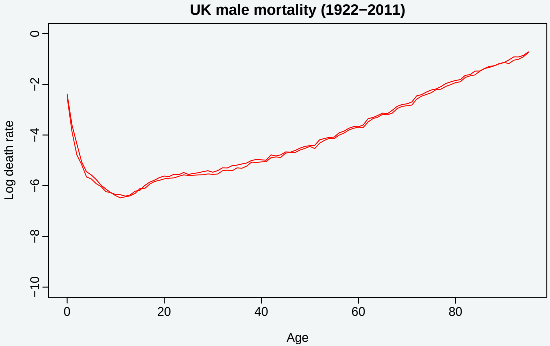
Acknowledgement: Zoltan Butt and Steven Haberman (Cass Business School, London)

Motivating example: UK male log mortality rates

Force of mortality is

$$\widehat{\mu}_{x,t} = \widehat{m}_{x,t} = \frac{y_{x,t}}{e_{x,t}}$$

where $y_{x,t}$ and $e_{x,t}$ represent the number of deaths and corresponding central exposure for any given age group at year t. Obtain a $n \times p$ matrix to represent age and time dimensions



Lee-Carter model under Gaussian error

LC model is defined as

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

1. α_x represents a constant age-specific pattern

Lee-Carter model under Gaussian error

LC model is defined as

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

1. α_x represents a constant age-specific pattern
2. κ_t measures the trend in mortality over time

Lee-Carter model under Gaussian error

LC model is defined as

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

1. α_x represents a constant age-specific pattern
2. κ_t measures the trend in mortality over time
3. β_x measures the age-specific deviations of mortality change from the overall trend

Lee-Carter model under Gaussian error

LC model is defined as

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

1. α_x represents a constant age-specific pattern
2. κ_t measures the trend in mortality over time
3. β_x measures the age-specific deviations of mortality change from the overall trend
4. $\varepsilon_{x,t}$ are assumed to be $N\left(0, \sigma^2\right)$ random effects by age and time

Objectives of *methods*

- Implement an iterative regression method for analysing age-period mortality rates via generalised Lee-Carter (LC) model

Objectives of *methods*

- Implement an iterative regression method for analysing age-period mortality rates via generalised Lee-Carter (LC) model
- Use Generalised Linear Model (GLM) model (Renshaw and Haberman 2006)

Objectives of *methods*

- Implement an iterative regression method for analysing age-period mortality rates via generalised Lee-Carter (LC) model
- Use Generalised Linear Model (GLM) model (Renshaw and Haberman 2006)
- Develop and implement a stratified LC model for the measurement of the additive effect on the log scale of an explanatory factor (other than age and time)

Objectives of *methods*

- Implement an iterative regression method for analysing age-period mortality rates via generalised Lee-Carter (LC) model
- Use Generalised Linear Model (GLM) model (Renshaw and Haberman 2006)
- Develop and implement a stratified LC model for the measurement of the additive effect on the log scale of an explanatory factor (other than age and time)
- Produce forecasts of age-specific mortality rates and life expectancy

Objectives of *ilc package*

- Extend LC model based on the Gaussian error structure to Poisson

Objectives of *ilc package*

- Extend LC model based on the Gaussian error structure to Poisson
- Instead of singular value decomposition, consider a regression model based on Poisson likelihood maximisation

Objectives of *ilc package*

- Extend LC model based on the Gaussian error structure to Poisson
- Instead of singular value decomposition, consider a regression model based on Poisson likelihood maximisation
- ilc package contains methods for the analysis of a class of six log-linear models (capturing age, period, cohort) in the GLM with Poisson errors

Features of the ilc package

1. To assess goodness of fit of the regression, estimation routines support a range of residual diagnostic plots

Features of the ilc package

1. To assess goodness of fit of the regression, estimation routines support a range of residual diagnostic plots
2. Allows preliminary data corrections, to replace missing data cells, but also to eliminate potential outliers that might result from data inaccuracies

Features of the ilc package

1. To assess goodness of fit of the regression, estimation routines support a range of residual diagnostic plots
2. Allows preliminary data corrections, to replace missing data cells, but also to eliminate potential outliers that might result from data inaccuracies
3. Includes two simple methods of “closing-out” produces to correct the original data at very old ages before the application of the model

Features of the ilc package

1. To assess goodness of fit of the regression, estimation routines support a range of residual diagnostic plots
2. Allows preliminary data corrections, to replace missing data cells, but also to eliminate potential outliers that might result from data inaccuracies
3. Includes two simple methods of "closing-out" produces to correct the original data at very old ages before the application of the model
4. ilc package integrates with the *demography* and *forecast* packages

Features of the ilc package

1. To assess goodness of fit of the regression, estimation routines support a range of residual diagnostic plots
2. Allows preliminary data corrections, to replace missing data cells, but also to eliminate potential outliers that might result from data inaccuracies
3. Includes two simple methods of “closing-out” produces to correct the original data at very old ages before the application of the model
4. ilc package integrates with the *demography* and *forecast* packages
5. ilc package has improved inspection and graphical visualisation of mortality data and regression output

Lee-Carter model under Poisson error

1. LC parameters can be estimated by maximum likelihood methods based on Poisson error distribution

Lee-Carter model under Poisson error

1. LC parameters can be estimated by maximum likelihood methods based on Poisson error distribution
2. Assuming that age- and period-specific number of deaths are independent realisations from a Poisson distribution with parameters

$$\mathsf{E}[Y_{x,t}] = e_{x,t}\mu_{x,t}, \quad \mathsf{Var}[Y_{x,t}] = \phi\mathsf{E}[Y_{x,t}]$$

where ϕ is a measure of over-dispersion to allow for heterogeneity

Lee-Carter model under Poisson error

1. LC parameters can be estimated by maximum likelihood methods based on Poisson error distribution
2. Assuming that age- and period-specific number of deaths are independent realisations from a Poisson distribution with parameters

$$\mathsf{E}[Y_{x,t}] = e_{x,t}\mu_{x,t}, \quad \mathsf{Var}[Y_{x,t}] = \phi \mathsf{E}[Y_{x,t}]$$

   where ϕ is a measure of over-dispersion to allow for heterogeneity
3. GLM model of the response variable $Y_{x,t}$ with log-link and non-linear parameterized predictor:

$$\eta_{x,t} = \ln(\widehat{y}_{x,t}) = \underbrace{\ln(e_{x,t})}_{\text{offset}} + \alpha_x + \beta_x \kappa_t$$

Estimation algorithm

1. Maximum likelihood point estimates under the GLM approach are obtained at the minimum value of the total deviation, given by

$$D\left(y_{x,t}, \widehat{y}_{x,t}\right) = \sum_{x,t} \mathsf{dev}(x,t) = \sum_{x,t} 2\omega_{x,t} \left\{ y_{x,t} \ln \frac{y_{x,t}}{\widehat{y}_{x,t}} - (y_{x,t} - \widehat{y}_{x,t}) \right\} \quad (1)$$

where $\mathsf{dev}(x,t)$ are the deviance residuals that depend on a set of prior weights $\omega_{x,t}$

Estimation algorithm

1. Maximum likelihood point estimates under the GLM approach are obtained at the minimum value of the total deviation, given by

$$D\left(y_{x,t}, \widehat{y}_{x,t}\right) = \sum_{x,t} \mathsf{dev}(x,t) = \sum_{x,t} 2\omega_{x,t}\left\{y_{x,t}\ln\frac{y_{x,t}}{\widehat{y}_{x,t}} - (y_{x,t} - \widehat{y}_{x,t})\right\} \tag{1}$$

   where $\mathsf{dev}(x,t)$ are the deviance residuals that depend on a set of prior weights $\omega_{x,t}$

2. Resort to an iterative Newton-Raphson method applied to the deviance function (1) . We use the iterative procedure:

Estimation algorithm

1. Maximum likelihood point estimates under the GLM approach are obtained at the minimum value of the total deviation, given by

$$D\left(y_{x,t}, \widehat{y}_{x,t}\right) = \sum_{x,t} \mathsf{dev}(x,t) = \sum_{x,t} 2\omega_{x,t}\left\{ y_{x,t} \ln \frac{y_{x,t}}{\widehat{y}_{x,t}} - (y_{x,t} - \widehat{y}_{x,t}) \right\} \quad (1)$$

   where $\mathsf{dev}(x,t)$ are the deviance residuals that depend on a set of prior weights $\omega_{x,t}$

2. Resort to an iterative Newton-Raphson method applied to the deviance function (1) . We use the iterative procedure:
   1. Set starting values $\widehat{\beta}_x$

Estimation algorithm

1. Maximum likelihood point estimates under the GLM approach are obtained at the minimum value of the total deviation, given by

$$D\left(y_{x,t}, \widehat{y}_{x,t}\right) = \sum_{x,t} \mathsf{dev}(x,t) = \sum_{x,t} 2\omega_{x,t} \left\{ y_{x,t} \ln \frac{y_{x,t}}{\widehat{y}_{x,t}} - (y_{x,t} - \widehat{y}_{x,t}) \right\} \tag{1}$$

   where $\mathsf{dev}(x,t)$ are the deviance residuals that depend on a set of prior weights $\omega_{x,t}$

2. Resort to an iterative Newton-Raphson method applied to the deviance function (1) . We use the iterative procedure:
   1. Set starting values $\widehat{\beta}_x$
   2. Given $\widehat{\beta}_x$, update $\widehat{\alpha}_x$ and $\widehat{\kappa}_t$

Estimation algorithm

1. Maximum likelihood point estimates under the GLM approach are obtained at the minimum value of the total deviation, given by

$$D\left(y_{x,t}, \widehat{y}_{x,t}\right) = \sum_{x,t} \mathsf{dev}(x,t) = \sum_{x,t} 2\omega_{x,t} \left\{ y_{x,t} \ln \frac{y_{x,t}}{\widehat{y}_{x,t}} - (y_{x,t} - \widehat{y}_{x,t}) \right\} \tag{1}$$

   where dev(x,t) are the deviance residuals that depend on a set of prior weights $\omega_{x,t}$

2. Resort to an iterative Newton-Raphson method applied to the deviance function (1) . We use the iterative procedure:
   1. Set starting values $\widehat{\beta}_x$
   2. Given $\widehat{\beta}_x$, update $\widehat{\alpha}_x$ and $\widehat{\kappa}_t$
   3. Given $\widehat{\kappa}_t$, update $\widehat{\alpha}_x$ and $\widehat{\beta}_x$

Estimation algorithm

1. Maximum likelihood point estimates under the GLM approach are obtained at the minimum value of the total deviation, given by

$$D\left(y_{x,t}, \widehat{y}_{x,t}\right) = \sum_{x,t} \text{dev}(x,t) = \sum_{x,t} 2\omega_{x,t} \left\{ y_{x,t} \ln \frac{y_{x,t}}{\widehat{y}_{x,t}} - (y_{x,t} - \widehat{y}_{x,t}) \right\} \tag{1}$$

   where dev(x,t) are the deviance residuals that depend on a set of prior weights $\omega_{x,t}$

2. Resort to an iterative Newton-Raphson method applied to the deviance function (1) . We use the iterative procedure:
   1. Set starting values $\widehat{\beta}_x$
   2. Given $\widehat{\beta}_x$, update $\widehat{\alpha}_x$ and $\widehat{\kappa}_t$
   3. Given $\widehat{\kappa}_t$, update $\widehat{\alpha}_x$ and $\widehat{\beta}_x$
   4. Compute $D(y_{x,t}, \widehat{y}_{x,t})$

Estimation algorithm

1. Maximum likelihood point estimates under the GLM approach are obtained at the minimum value of the total deviation, given by

$$D\left(y_{x,t}, \widehat{y}_{x,t}\right) = \sum_{x,t} \text{dev}(x,t) = \sum_{x,t} 2\omega_{x,t}\left\{y_{x,t}\ln\frac{y_{x,t}}{\widehat{y}_{x,t}} - \left(y_{x,t} - \widehat{y}_{x,t}\right)\right\} \tag{1}$$

   where $\text{dev}(x,t)$ are the deviance residuals that depend on a set of prior weights $\omega_{x,t}$

2. Resort to an iterative Newton-Raphson method applied to the deviance function (1) . We use the iterative procedure:
   1. Set starting values $\widehat{\beta}_x$
   2. Given $\widehat{\beta}_x$, update $\widehat{\alpha}_x$ and $\widehat{\kappa}_t$
   3. Given $\widehat{\kappa}_t$, update $\widehat{\alpha}_x$ and $\widehat{\beta}_x$
   4. Compute $D(y_{x,t}, \widehat{y}_{x,t})$
   5. Repeat the updating cycle; stop when $D(y_{x,t}, \widehat{y}_{x,t})$ converges

Inclusion of cohort effect

1. Basic LC model can be extended to include an additional bilinear term, containing a second period effect or a cohort effect

Inclusion of cohort effect

1. Basic LC model can be extended to include an additional bilinear term, containing a second period effect or a cohort effect
2. Force of mortality by a generalised structure is given as

$$\mu_{x,t} = \exp\left(\alpha_x + \underline{\beta_x^{(0)} l_{t-x}} + \beta_x^{(1)} \kappa_t\right)$$

where α_x: main age profile; l_{t-x}: cohort effect; κ_t: period

Lee-Carter model with additional covariates

1. Additional factor depends on the size and nature of the mortality experience, such as geographical, socio-economic or race differences

Lee-Carter model with additional covariates

1. Additional factor depends on the size and nature of the mortality experience, such as geographical, socio-economic or race differences
2. Consider a cross-classified mortality experience observed over age x, period t and an extra variate g made up of $(k \times n \times l)$ data cells

Lee-Carter model with additional covariates

1. Additional factor depends on the size and nature of the mortality experience, such as geographical, socio-economic or race differences
2. Consider a cross-classified mortality experience observed over age x, period t and an extra variate g made up of $(k \times n \times l)$ data cells
3. Stratified LC model is given by

$$\eta_{x,t,g} = \ln(\widehat{y}_{x,t,g}) = \underbrace{\ln(e_{x,t,g})}_{\text{offset}} + \alpha_x + \underline{\alpha_g} + \beta_x \kappa_t,$$

where α_g measures the relative differences between the age-specific log mortality profiles among subgroups defined by the extra variate g

Forecasting

1. Forecasting mortality in the LC family of models is based on time series prediction of the calendar time dependent parameters (l_{t-x}, κ_t)

Forecasting

1. Forecasting mortality in the LC family of models is based on time series prediction of the calendar time dependent parameters (l_{t-x}, κ_t)
2. Mortality rate forecasts can be written as

$$\widehat{\mu}_{x,n+h} = \exp\left(\widehat{\alpha}_x + \widehat{\beta}_x^{(0)}\widehat{l}_{n+h-x} + \widehat{\beta}_x^{(1)}\widehat{\kappa}_{n+h}\right)$$

where $\widehat{l}_{n+h-x}$ and $\widehat{\kappa}_{n+h}$ represent the forecast cohort and period effects

Forecasting

1. Forecasting mortality in the LC family of models is based on time series prediction of the calendar time dependent parameters (l_{t-x}, κ_t)
2. Mortality rate forecasts can be written as

$$\widehat{\mu}_{x,n+h} = \exp\left(\widehat{\alpha}_x + \widehat{\beta}_x^{(0)}\widehat{l}_{n+h-x} + \widehat{\beta}_x^{(1)}\widehat{\kappa}_{n+h}\right)$$

where $\widehat{l}_{n+h-x}$ and $\widehat{\kappa}_{n+h}$ represent the forecast cohort and period effects

3. Random walk with drift, ARIMA(0,1,0), is used to forecast period effect (κ_t), expressed as

$$\kappa_t = \kappa_{t-1} + d + e_t$$

where d measures the drift and e_t represents the white noise

Forecasting

1. Forecasting mortality in the LC family of models is based on time series prediction of the calendar time dependent parameters (l_{t-x}, κ_t)
2. Mortality rate forecasts can be written as

$$\widehat{\mu}_{x,n+h} = \exp\left(\widehat{\alpha}_x + \widehat{\beta}_x^{(0)}\widehat{l}_{n+h-x} + \widehat{\beta}_x^{(1)}\widehat{\kappa}_{n+h}\right)$$

   where $\widehat{l}_{n+h-x}$ and $\widehat{\kappa}_{n+h}$ represent the forecast cohort and period effects
3. Random walk with drift, ARIMA(0,1,0), is used to forecast period effect (κ_t), expressed as

$$\kappa_t = \kappa_{t-1} + d + e_t$$

   where d measures the drift and e_t represents the white noise
4. In the cohort effects, forecasts revert to the fitted parameters when h falls within the available data range

CMI data

1. CMI data contains the mortality experience of male life office pensioners retiring at or after normal retirement age

CMI data

1. CMI data contains the mortality experience of male life office pensioners retiring at or after normal retirement age
2. Data is made up of observed central exposure and deaths for ages 50-108 from 1983 to 2003

R demo of explanatory plots

1. Plot mortality rates, population-at-risk, death counts for any age group and year

```
>insp.dd(dd.cmi.pens,age=50:80,year=1985:1990)
>insp.dd(dd.cmi.pens,what='pop',age=70:100,year=1988:1993)
>insp.dd(dd.cmi.pens,what='deaths',age=seq(100),
year=1980:2010)
```

R demo of explanatory plots

1. Plot mortality rates, population-at-risk, death counts for any age group and year

```
>insp.dd(dd.cmi.pens,age=50:80,year=1985:1990)
>insp.dd(dd.cmi.pens,what='pop',age=70:100,year=1988:1993)
>insp.dd(dd.cmi.pens,what='deaths',age=seq(100),
year=1980:2010)
```

2. Produce simple plots (i.e., without legend) of log- or untransformed rates:

```
>plot(dd.cmi.pens)
>plot(dd.cmi.pens,transform=FALSE)
```

1. Produce annotated plots of log or original rates:

```
>plot_dd(dd.cmi.pens, xlim=c(40, 110),
lpar = list(x.int = -0.2, y.int = 0.9, cex = 0.85))
>plot_dd(dd.cmi.pens, year=1985:1995, transform=FALSE)
>plot_dd(dd.cmi.pens, year=1995:1997, transform=FALSE,
lty=1:3, col=1:3)
```

R demo of explanatory plots

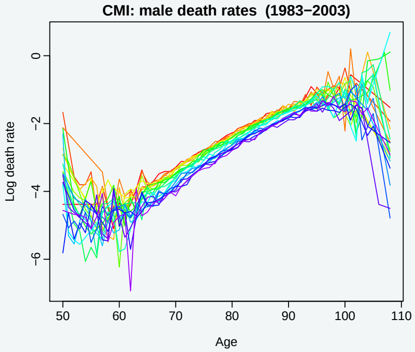
1. Produce annotated plots of log or original rates:

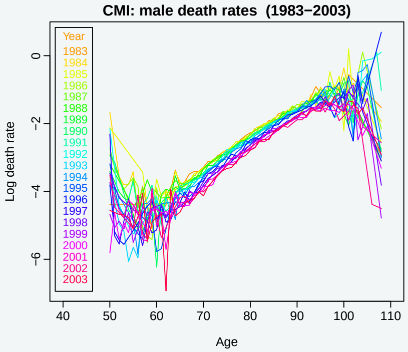
```
>plot_dd(dd.cmi.pens, xlim=c(40, 110),
lpar = list(x.int = -0.2, y.int = 0.9, cex = 0.85))
>plot_dd(dd.cmi.pens, year=1985:1995, transform=FALSE)
>plot_dd(dd.cmi.pens, year=1995:1997, transform=FALSE,
lty=1:3, col=1:3)
```

2. Deal with missing data

```
# without correction of empty cells
>tmp.d = extract.deaths(dd.cmi.pens, ages=55:100)
# empty cells are filled using perk model
>tmp.d = extract.deaths(dd.cmi.pens, ages=55:100,
fill='perks')
```

Explanatory plots: dealing with missing values

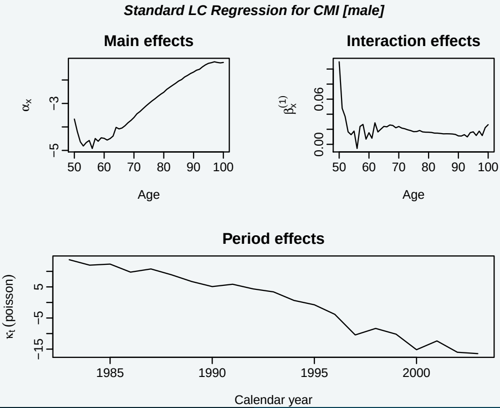




R demo for estimation of LC model

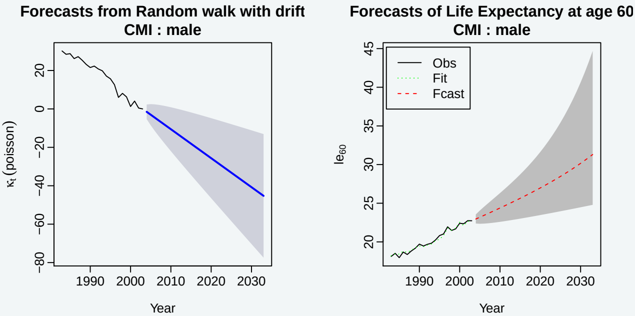
Estimate the base LC model with Poisson errors

```
>mod6 = lca.rh(dd.cmi.pens, mod = 'lc', interpolate=TRUE)
>coef(mod6); plot(mod6)
>fitted_plot(mod6); residual_plot(mod6)
```



Plots of fitted models

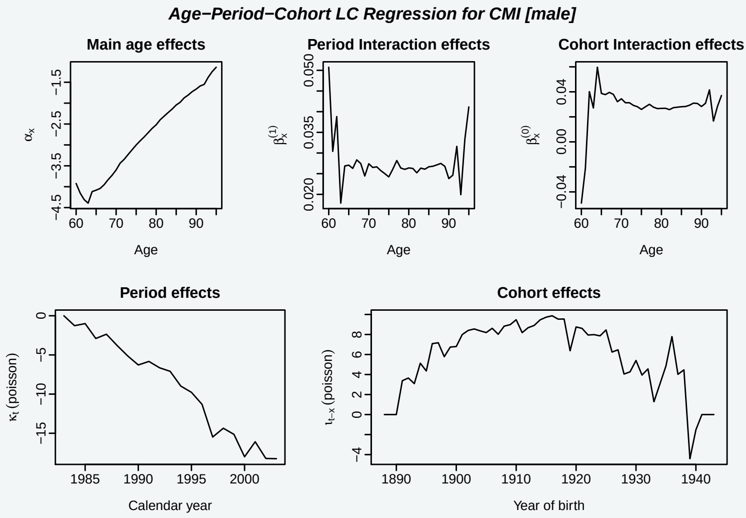
```
>forc6 = forecast(mod6, h = 20, jump = 'fit', level = 90,
shift=FALSE)
>plot_dd(forc6, xlim=c(45,100), lpar=list(x.int=-0.2,
y.int=0.9, cex=0.95))
>le6 = life.expectancy(forc6, age=60)
>flc.plot(mod6, at=60, h=30, level=90)
```



R demo for estimation of age-period-cohort model

```
>mod1 = lca.rh(dd.cmi.pens, age=60:95, mod = "m",
restype='deviance', dec.conv=3)
>coef(mod1)
>plot(mod1)
```

Age-period-cohort plot



R demo for stratified LC model

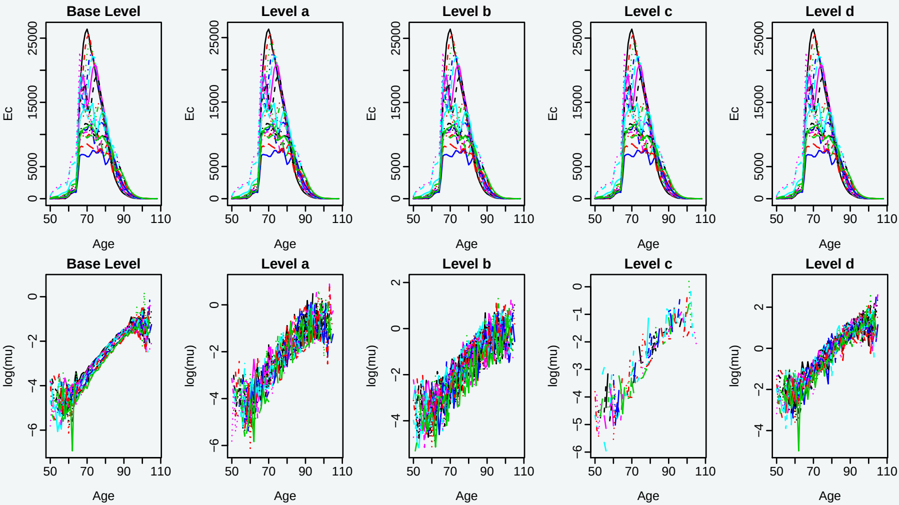
1. For stratified LC model, ilc package introduces a special class of data object that holds information about the grouping factors and aggregate data of number of deaths, central exposures and mortality rates

R demo for stratified LC model

1. For stratified LC model, ilc package introduces a special class of data object that holds information about the grouping factors and aggregate data of number of deaths, central exposures and mortality rates
2. Taking the CMI experience as the base data, produce a randomly stratified mortality data

```
>rfp.cmi = dd.rfp(dd.cmi.pens, rfp = c(0.5,1.2,-0.7,2.5))
>matplot(rfp.cmi$age, rfp.cmi$pop[,,1], type='l',
xlab='Age', ylab='Ec', main = 'Base Level')
>matplot(rfp.cmi$age, rfp.cmi$pop[,,2], type='l',
xlab='Age', ylab='Ec', main = 'Base Level')
```

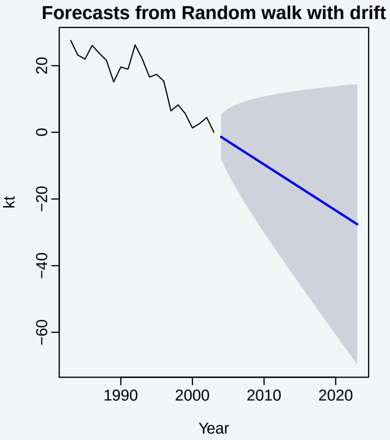
Plots of stratified Lee-Carter

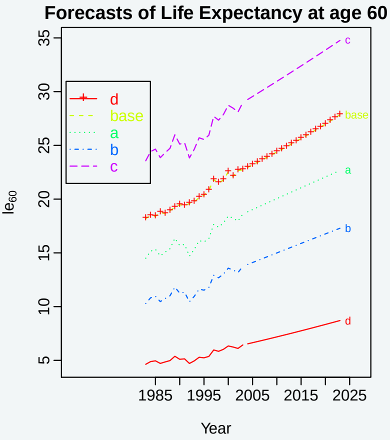


R demo for estimating and forecasting of stratified LC model

```
>rfp.cmi = dd.rfp(dd.cmi.pens, rfp = c(0.5, 1.2, -0.7, 2.5))
>mod6e = elca.rh(rfp.cmi, age=50:100, interpolate=TRUE,
dec.conv=3, verbose=TRUE)
>coef(mod6e)
>mod6ef = forecast.lca(mod6e, h = 20, level = 90, jump=‘fit’,
shift=FALSE)
>plot(mod6ef$kt, ylab=‘kt’, xlab=‘Year’)
>matfle.plot(mod6e$lca, mod6, at=60, label=‘RFP CMI’, h=20)
```

Plot of forecast life expectancy





Conclusion

1. ilc package implements the Lee-Carter model with Gaussian and Poisson errors

Conclusion

1. ilc package implements the Lee-Carter model with Gaussian and Poisson errors
2. ilc package implements additional five models discussed in Renshaw and Haberman (2006)

Conclusion

1. ilc package implements the Lee-Carter model with Gaussian and Poisson errors
2. ilc package implements additional five models discussed in Renshaw and Haberman (2006)
3. Stratified Lee-Carter model allows users to include additional covariates (other than age and time)

References

- Renshaw, A. E., and S. Haberman, 2003. Lee-Carter mortality forecasting with age-specific enhancement. *Insurance: Mathematics and Economics*, 33, 255-272.

References

- Renshaw, A. E., and S. Haberman, 2003. Lee-Carter mortality forecasting with age-specific enhancement. *Insurance: Mathematics and Economics*, 33, 255-272.
- Hyndman, R. J., and H. L. Shang, 2010. Rainbow plots, bagplots and boxplot for functional data, *Journal of Computational and Graphical Statistics*, 19(1), 29-45.

References

- Renshaw, A. E., and S. Haberman, 2003. Lee-Carter mortality forecasting with age-specific enhancement. *Insurance: Mathematics and Economics*, 33, 255-272.
- Hyndman, R. J., and H. L. Shang, 2010. Rainbow plots, bagplots and boxplot for functional data, *Journal of Computational and Graphical Statistics*, 19(1), 29-45.
- H. L. Shang, 2011. rainbow: An R package for visualizing functional time series, *The R Journal*, 3(2), 54-59.

Thank you! · Tak