Lee Carter

Franco Marais

10 July 2021

The demography package allows you to access data from directly from the Human Mortality Database. The StMoMo package allows one to fit a variety of mortality models.

```
library(StMoMo) # used to fit models
library(demography) # used to access data from the mortality database
```

We import the data for *England and Whales*. Since we're aiming to fit the Poisson and Negative Binomial versions of the Lee-Carter model, we'll be using the central mortality data.

```
EWdata <- hmd.mx(country = "GBRTENW", username = "francochess@gmail.com",
                 password = "1616331654", label = "GBRTENW")

# central exposure data
EWMaleCenData <- StMoMoData(EWdata, series = "male")
EWFemaleCenData <- StMoMoData(EWdata, series = "female")
```

```
ages <- 19:108
years <- 1950:2018
```

Poisson

We fit the Poisson Lee-Carter model to both the male and female data.

```
LCpoisMale <- fit(lc(link = "log"), data = EWMaleCenData, ages.fit = ages, years.fit = years)
LCpoisFemale <- fit(lc(link = "log"), data = EWFemaleCenData, ages.fit = ages, years.fit = years)
```

Negative Binomial

LCnegbinom_algorithm.R contains the code to fit the negative binomial version of the Lee-Carter model. The code is an implementation of the algorithm described in *Uncertainty in Mortality Forecasting: An Extension to the Classical Lee-Carter Approach* in Li2009.

```
# functions to fit negative binomial version of Lee-Carter model
source(paste0(getwd(), "/LCnegbinom_algorithm.R"))
```

We gather some of the StMoMo output for convenience. We also set up a list of the Poisson maximum likelihood estimates to use as the initial parameters for the Negative Binomial algorithm iteration.

```
# death-, exposure- and missing data weight matrix
dxt <- LCpoisMale$data$Dxt
dxt <- dxt[as.integer(rownames(dxt))%in%ages, as.integer(colnames(dxt))%in%years]
Ext <- LCpoisMale$data$Ext
Ext <- Ext[as.integer(rownames(Ext))%in%ages, as.integer(colnames(Ext))%in%years]
wxt <- LCpoisMale$wxt

# fixing layout for Julia read-in
ax <- c(LCpoisMale$ax)
bx <- c(LCpoisMale$bx)
kt <- c(LCpoisMale$kt)
names(ax) <- ages
names(bx) <- ages
names(kt) <- years
# Poisson maximum likelihood estimates of parameters
LCpoisMale <- list(ax=ax, bx=bx, kt=kt)
```

We use the Poisson maximum likelihood parameters as the initial parameters for the negative binomial algorithm for faster convergence.

Note that the dispersion parameters are estimated using the method of moments

$$E(\lambda_x) = Var\left(\frac{D_{xt}}{\hat{D_{xt}}}\right)$$

We found that the dispersion parameters, especially at the youngest and oldest ages, take some time to converge, while the base Lee-Carter parameters converge reasonably quickly. Hence, we run the model for 2 iterations so that the base Lee-Carter parameters can converge reasonably at most ages and the dispersion parameters can be reestimated according to the new parameters. We then start iteration again using the new initial parameters.

```
# using Poisson maximum likelihood parameters as starting parameters for iteration
LCnbMale_init <- LCnegbinom(LCpoisMale, dxt, Ext, wxt, maxiter=2)
LCnbMale <- LCnegbinom(LCnbMale_init, dxt, Ext, wxt, maxiter=1000)
```

We do the same for fitting the female data.

```
# death-, exposure- and missing data weight matrix
dxt <- LCpoisFemale$data$Dxt
dxt <- dxt[as.integer(rownames(dxt))%in%ages, as.integer(colnames(dxt))%in%years]
Ext <- LCpoisFemale$data$Ext
Ext <- Ext[as.integer(rownames(Ext))%in%ages, as.integer(colnames(Ext))%in%years]
wxt <- LCpoisFemale$wxt

# fixing layout for Julia read-in
ax <- c(LCpoisFemale$ax)
bx <- c(LCpoisFemale$bx)
kt <- c(LCpoisFemale$kt)
names(ax) <- ages
names(bx) <- ages
names(kt) <- years
# Poisson maximum likelihood estimates of parameters
LCpoisFemale <- list(ax=ax, bx=bx, kt=kt)
```

```
# using Poisson maximum likelihood parameters as starting parameters for iteration
LCnbFemale_init <- LCnegbinom(LCpoisFemale, dxt, Ext, wxt, maxiter=2)
LCnbFemale <- LCnegbinom(LCnbFemale_init, dxt, Ext, wxt, maxiter=1000)
```

We fit random walks with drift to κ_t of each model. We note that the Negative Binomial Lee-Carter models will have broader forecast intervals (as expected) due to the larger standard deviations of the random walk models.

```
# Fitting Random Walks with drift
arima(diff(as.vector(LCpoisMale$kt)), order = c(0, 0, 0))
```

```
##
## Call:
## arima(x = diff(as.vector(LCpoisMale$kt)), order = c(0, 0, 0))
##
## Coefficients:
##       intercept
##         -0.9441
## s.e.     0.2084
##
## sigma^2 estimated as 2.953:  log likelihood = -133.3,  aic = 270.6
```

```
arima(diff(as.vector(LCpoisFemale$kt)), order = c(0, 0, 0))
```

```
##
## Call:
## arima(x = diff(as.vector(LCpoisFemale$kt)), order = c(0, 0, 0))
##
## Coefficients:
##       intercept
##         -1.1510
## s.e.     0.3067
##
## sigma^2 estimated as 6.395:  log likelihood = -159.58,  aic = 323.15
```

```
arima(diff(LCnbMale$kt), order = c(0, 0, 0))
```

```
##
## Call:
## arima(x = diff(LCnbMale$kt), order = c(0, 0, 0))
##
## Coefficients:
##       intercept
##         -0.9435
## s.e.     0.2166
##
## sigma^2 estimated as 3.191:  log likelihood = -135.94,  aic = 275.87
```

```
arima(diff(LCnbFemale$kt), order = c(0, 0, 0))
```

```
## 
## Call:
## arima(x = diff(LCnbFemale$kt), order = c(0, 0, 0))
## 
## Coefficients:
##          intercept
##            -1.1401
## s.e.        0.3212
## 
## sigma^2 estimated as 7.017:  log likelihood = -162.73,  aic = 329.47
```