Mortality modeling and forcasting

February 9, 2016

Introduction

- The life table plays an important role in actuarial sciences.
- It is one of the oldest topics in the field of statistics.
- Important in any area where birth, death or illness may happen.
- Oldest documented life table is Halley's in 1693.
- It is not limited to human beings.

Halley's life table (1693)

Age. Curt.	Per-ſons.	Age. Curt.	Per-ſons	Age. Curt.	Per-ſons	Age. Curt.	Per-ſons	Age. Curt.	Per-ſons	Age. Curt.	Per-ſons
1	1000	8	680	15	628	22	585	29	539	36	481
2	855	9	670	16	622	23	579	30	531	37	472
3	798	10	661	17	616	24	573	31	523	38	463
4	760	11	653	18	610	25	567	32	515	39	454
5	732	12	646	19	604	26	560	33	507	40	445
6	710	13	640	20	598	27	553	34	499	41	436
7	692	14	634	21	592	28	546	35	490	42	427
Age. Curt	Per-ſons.	Age. Curt.	Per-ſons	Age. Curt.	Per-ſons	Age Curt	Per-ſons	Age. Curt.	Per-ſons	Age. Curt.	Per-ſons
43	417	50	346	57	272	64	202	71	131	78	58
44	407	51	335	58	262	65	192	72	120	79	49
45	397	52	324	59	252	66	182	73	109	80	41
46	387	53	313	60	242	67	172	74	98	81	34
47	377	54	302	61	232	68	162	75	88	82	28
48	367	55	292	62	222	69	152	76	78	83	23
49	357	56	282	63	212	70	142	77	68	84	20

Age.	Perſons.
7	5547
14	4584
21	4270
28	3964
35	3604
42	3178
49	2709
56	2194
63	1694
70	1204
77	692
84	253
100	107
	34000
	Sum Total.

Types of life tables

- Cohort (dynamic, generational): Records the actual mortality within a group of individuals over a periode of time starting from birth to last death in the group.
- Period (static, current): Considers the mortality of an entire population at a specific point of time.

Cohort life table

- Suppose we can follow a group, e.g 100.000, of people from a certain group (Female/Male, Non-Smoker/Smoker . . .) born on the same day, then we can record, at each birthday, the number of group members alive and the number of those dying within the next year.
- Then the ratio of these two quantities gives us the probability of dying at age x.
- In practice this is difficult (even impossible) to do.

Period life table

- Instead period life tables are much more commonly used.
- These tables are constructed based on an evaluation of the mortality experience of persons from all ages in a short period of time (typically from 1 to 5 years).
- This evaluation can be based on census information provided at regular intervals.

Columns of a life table

- lx: The number of persons alive at age x.
- dx: The number of persons dead in the interval (x, x+1).
- Lx: The total number of person-years lived in the interval (x, x+1).
- qx: Probability of dying at age x.
- mx: Mortality rate at age x. mx = dx/Lx.
- Tx: Total number of person-years lived by the group from age x til ω, where ω is the ultimate age.
- ex: The residual life expectancy of persons alive at age x.

Construction of a life table (Period)

- $q(x)$ (or $m(x)$) from census data.
- $q(x) = 1 - \exp(-m(x))$
- $l(0) = 100.000$
- $d(x) = l(x)*q(x)$
- $l(x+1) = l(x) - d(x)$
- $L(x) = l(x) + a(x)*d(x)$ (usually $a(x)=0.5$ except at extreme ages).
- $T(x)$ = Sum from age x to ω of the $L(x)$ column.
- $e(x) = T(x)/l(x)$

Conclusion

Knowing $m(x)$, we can calculate all the other quantities.

The force of mortality $\mu_{x,t}$

- Let $D_{x,t}$ denote the number of deaths of people aged x in year t.
- Let $E_{x,t}$ denote the exposure of age x in year t. This is just L(x) in year t.

Then we can define the force of mortality $\mu_{x,t}$ as:

$$\mu_{x,t} = D_{x,t}/E_{x,t}$$

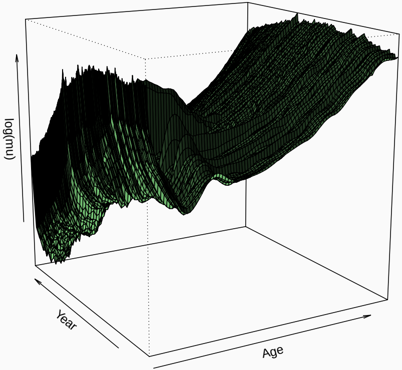
The force of mortality $\mu_{x,t}$

Take for example the data from HMD of “France (Total Population)”, then $\mu_{x,t}$ can be calculated as follows:

```r
Mu<-DEATH[,3:5]/EXPOSURE[,3:5]
head(Mu)
```

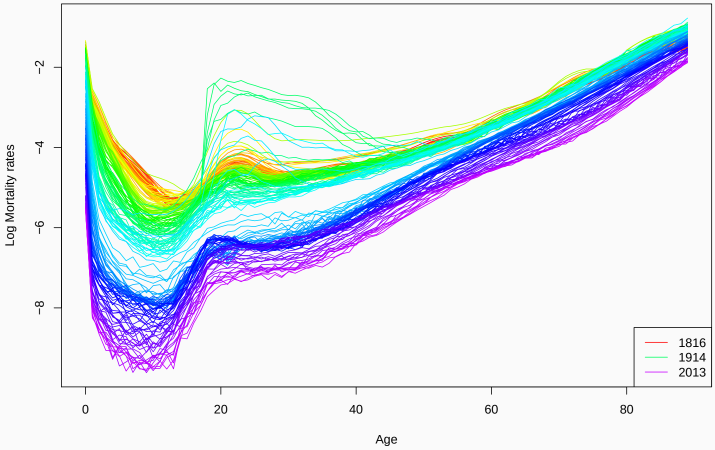
```
##       Female       Male      Total
## 1 0.18698611 0.22293069 0.20534411
## 2 0.04670190 0.04666953 0.04668535
## 3 0.03392803 0.03430644 0.03412039
## 4 0.02291217 0.02315458 0.02303538
## 5 0.01599465 0.01607464 0.01603530
## 6 0.01383446 0.01363534 0.01373288
```

The force of mortality $\mu_{x,t}$

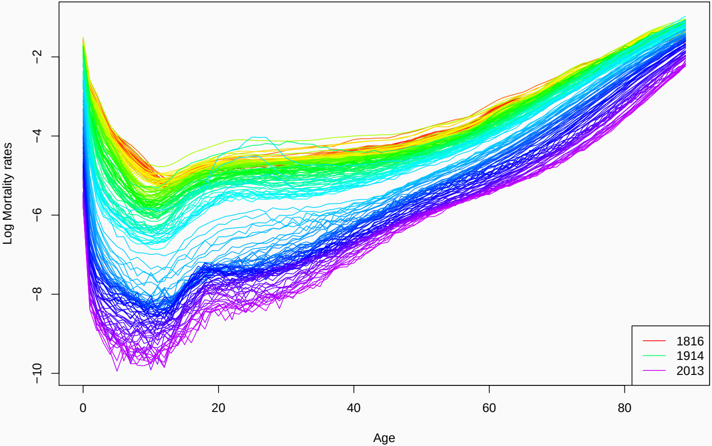


Force of mortality as a function of time and age

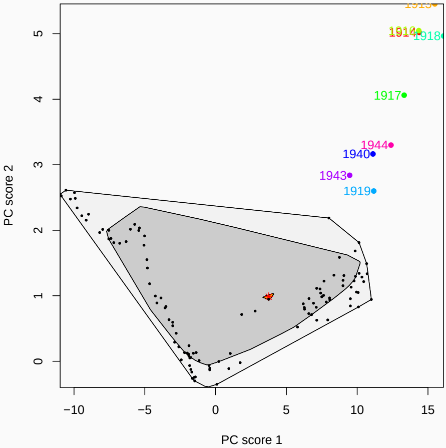
The force of mortality $\mu_{x,t}$ as a function of t (Male)



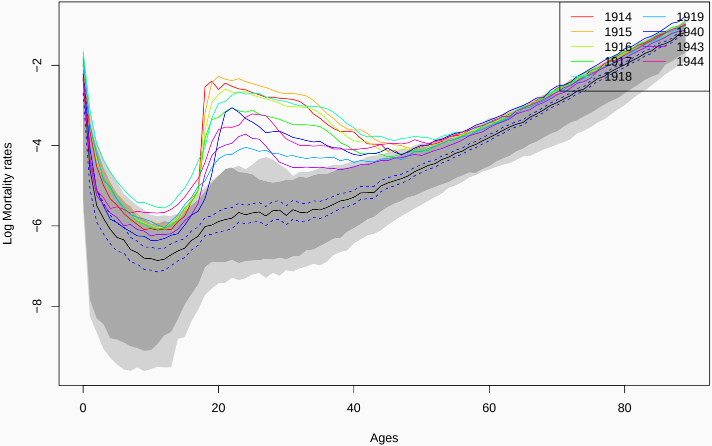
The force of mortality $\mu_{x,t}$ as a function of t (Female)



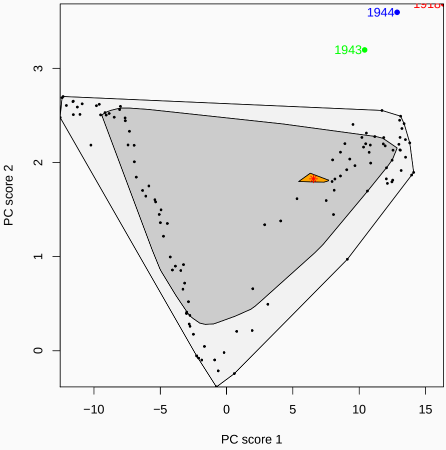
Detecting outliers in the $\mu_{x,t}$ series (Male)



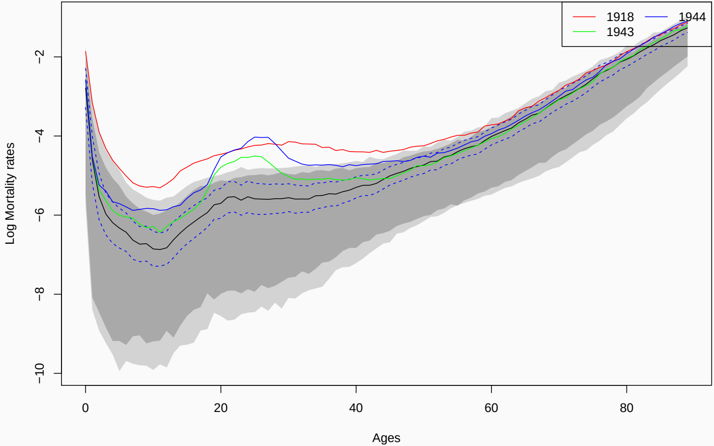
Detecting outliers in the $\mu_{x,t}$ series (Male)



Detecting outliers in the $\mu_{x,t}$ series (Female)



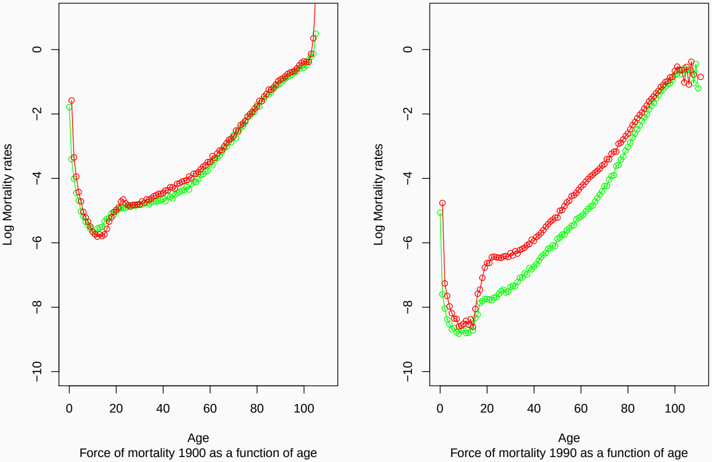
Detecting outliers in the $\mu_{x,t}$ series (Female)



Comparing $\mu_{x,t}$ for t=1900 and t=1990

```
year = 1900

D1900 = DEATH[DEATH$Year == year, ]
E1900 = EXPOSURE[EXPOSURE$Year == year, ]
Mu1900 = D1900[, 3:5]/E1900[, 3:5]



year = 1990

D1990 = DEATH[DEATH$Year == year, ]
E1990 = EXPOSURE[EXPOSURE$Year == year, ]
Mu1990 = D1990[, 3:5]/E1990[, 3:5]
```

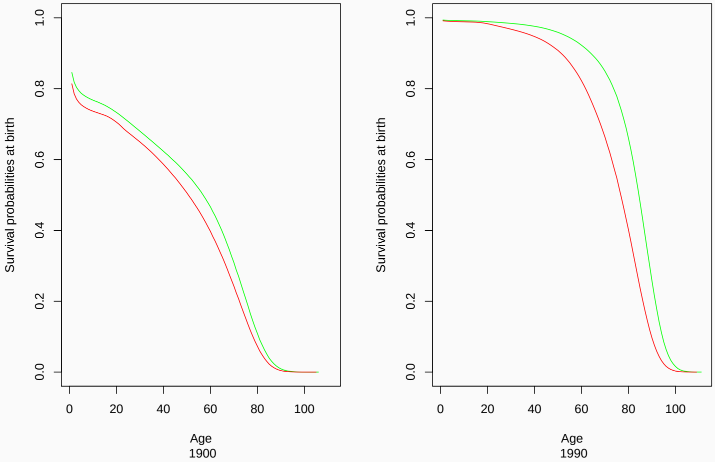
Comparing $\mu_{x,t}$ for t=1900 and t=1990



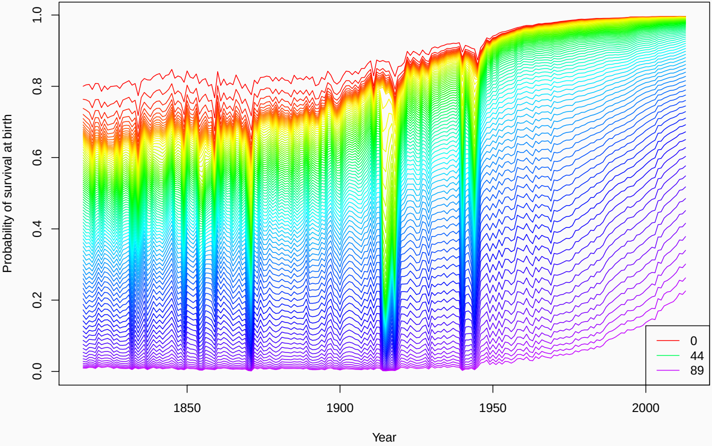
Comparing $p_t(x, x+1)$ for t=1900 and t=1990

```
PM1900=PF1900=matrix(0,111,111)
PM1990=PF1990=matrix(0,111,111)



for(x in 0:110){

  PM1900[x+1,1:(111-x)]=exp(-cumsum(Mu1900[(x+1):111,2]))
  PF1900[x+1,1:(111-x)]=exp(-cumsum(Mu1900[(x+1):111,1]))

  PM1990[x+1,1:(111-x)]=exp(-cumsum(Mu1990[(x+1):111,2]))
  PF1990[x+1,1:(111-x)]=exp(-cumsum(Mu1990[(x+1):111,1]))
}
```

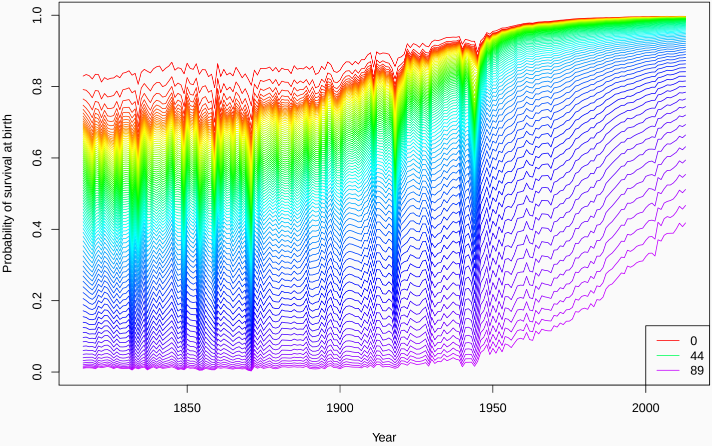
Comparing $p_t(x, x+1)$ for t=1900 and t=1990



Comparing $p_t(x, x+h)$ from t=1900 to t=1990 (Male)



Comparing $p_t(x, x+h)$ from t=1900 to t=1990 (Female)



General remarks on $\mu_{x,t}$

- $x \longmapsto \log(\mu_{x,t})$ has a very similar shape for different t's, and it is (almost) linear from age 30 to 80.
- $t \longmapsto \log(\mu_{x,t})$ has decreasing trend across (almost) all ages.
- In other words, $\log(\mu_{x,t})$ has fixed characteristics for all t's and changing ones through the years.

The Lee-Carter model (1992)

$$\log(\mu_{x,t}) = \alpha_x + \beta_x \kappa_t + \epsilon_{x,t}$$

- α_x: Age specific component independent of time.
- κ_t: Time varying parameter indicating the general level of mortality through the years.
- β_x : Represents the effect of κ_t on each age.
- $\epsilon_{x,t}$: i.i.d noise.

Due to overparameterization, identification assumptions are imposed e.g:

$$\sum_x \beta_x = 1, \qquad \sum_t \kappa_t = 0$$

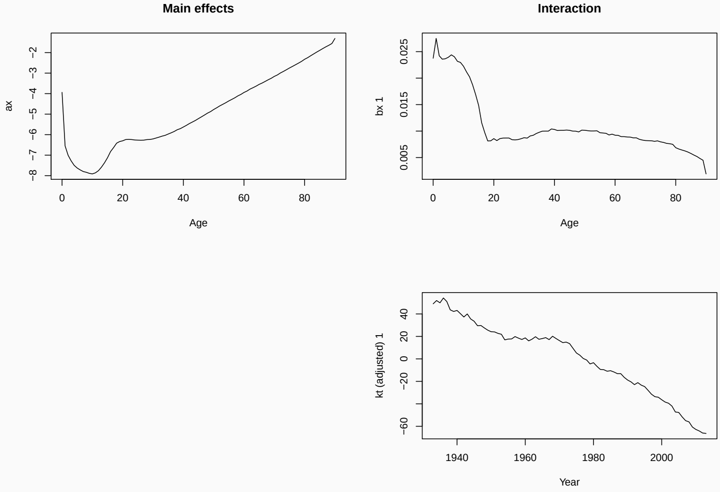
Estimation is done via Likelihood estimation under certain assumptions.

Estimation of the Lee-Carter model using demography package

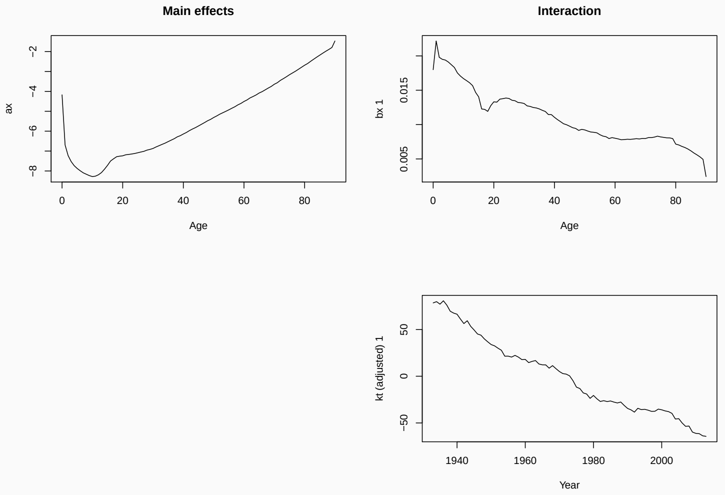
- The demography package has built-in functions for the estimation and forcasting of demographic statistics using the Lee-Carter model.
- In addition, one can use the function **hmd.mx** to fetch data directly from the *Human Mortality Database*.

```r
library(demography)
usa<-hmd.mx("USA","aymane10@hotmail.com","STK4500","USA")

usa.90<-extract.ages(usa,ages = 0:90 )

lc.male<- lca(usa.90, series = "male")
lc.female<- lca(usa.90, series = "female")
```

Estimates of α_x, β_x and κ_t (male)



Estimates of α_x, β_x and κ_t (female)



Forcasting in the Lee-Carter model

Forcasting in the Lee-Carter model is easiliy done in the **demography** package

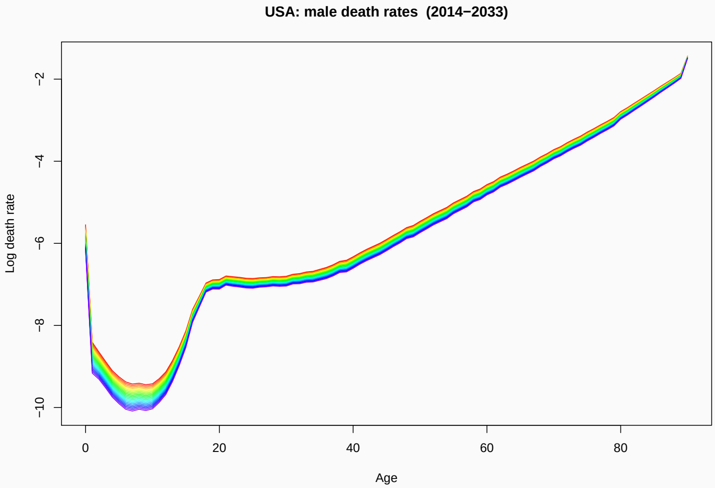
```
forecast.lc.male <- forecast(lc.male, h=20)
forecast.lc.female <- forecast(lc.female, h=20)
```

This based on the assumption that κ_t is a time series of the form:

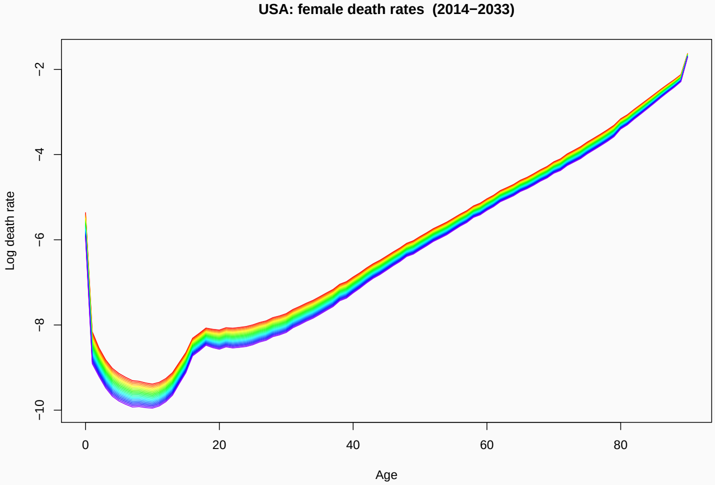
$$\kappa_t = \theta + \kappa_{t-1} + \xi_t$$

and ξ_t are i.i.d normally distributed r.v with mean 0 and variance σ^2.

Forcasting in the Lee-Carter model



Forcasting in the Lee-Carter model



Estimation and forcastng of κ_t

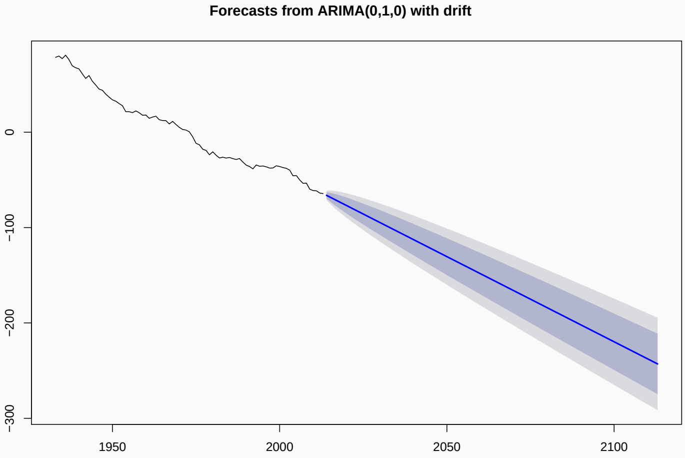
```r
kappat <- lc.female$kt
smoothkt <- ets(kappat)
forecastkt <- forecast(smoothkt, h=100)
```

or more concretely

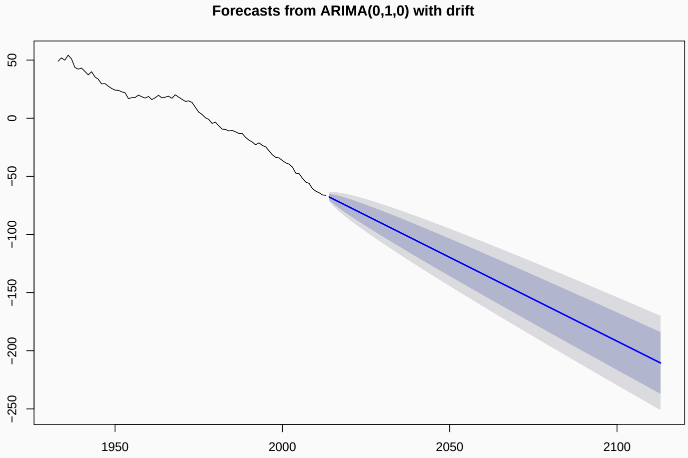
```r
(fittedkt <- auto.arima(kappat))
```

```
## Series: kappat
## ARIMA(0,1,0) with drift
##
## Coefficients:
##         drift
##       -1.7861
## s.e.   0.2766
##
```

Estimation and forecasting of κ_t (Female)



Estimation and forecasting of κ_t (Male)



The LifeMetrics functions

- The problem with the **demography** package is that it implements only the Lee-Carter model.
- Although the Lee-Carter model is widely used in practice, there are other models that have improved the ideas in L-C.
- For example, the Renshaw-Haberman model (2006) adds another component that takes into account the year of birth (i.e cohort effect)

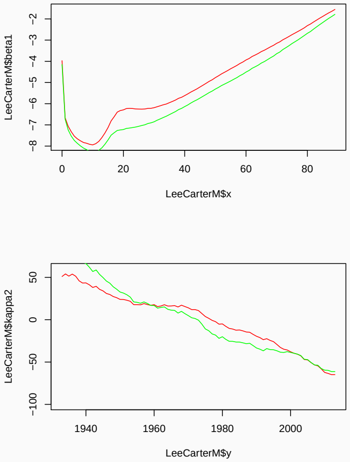
$$\log(\mu_{x,t}) = \alpha_x + \beta_x \kappa_t + \delta_x \gamma_{t-x} + \epsilon_{x,t}$$

- The LifeMetrics package has several models that are built-in.
- In addition one can simulate different senarios based on the forecasted component.

```r
DEATH_90<- deaths[deaths$Age<90,]
EXPOSURE_90<- exposures[exposures$Age<90,]

Ages <- unique(DEATH_90$Age)
Years <- unique(DEATH_90$Year)

ages <- seq(0,length(Ages)-1)
years <- seq(Years[1],Years[length(Years)])

n <- length(ages)
m <- length(years)
EXPOSUREM_tx <- t(matrix(EXPOSURE_90[,4],n,m))
EXPOSUREF_tx <- t(matrix(EXPOSURE_90[,3],n,m))
DEATHM_tx <- t(matrix(DEATH_90[,4],n,m))
DEATHF_tx <- t(matrix(DEATH_90[,3],n,m))
```

```r
source("C:/Users/Amin/Documents/R/fitModels.r")



LeeCarterM <- fit701(xv=ages,yv=years,
                     etx=EXPOSUREM_tx,dtx=DEATHM_tx,
                     wa=WeightFun)
LeeCarterF <- fit701(xv=ages,yv=years,
                     etx=EXPOSUREF_tx,dtx=DEATHF_tx,
                     wa=WeightFun)
```

LifeMetrics



Forecasting with LifeMetrics

```r
source("C:/Users/Amin/Documents/R/simModels.r")



forcastsLC <- sim2001(LeeCarterM$x,LeeCarterM$y,
                      LeeCarterM$beta1,LeeCarterM$beta2,
                      LeeCarterM$kappa2,
                      nsim = 10000,
                      nyears = 60,
                      tmax = 40,
                      x0 = 40)
```

Effect of increased longevity on a life insurance annuity

- Using the forcasted time series κ_t, we can forecast practically any demographical or actuarial quantity we need.
- For example, we can calculate the projected PV, at different points in the future, of a deferred temporary life annuity with term 30 issued for a person aged 40 and deferred for 30 years

$$\sum_{k=30}^{30+30-1} \frac{1}{(1+i)^k} p(x, x+k)$$

Effect of increased longevity on a life insurance annuity