Annuities Modeling with R

Giorgio Alfredo Spedicato, Ph.D C.Stat ACAS

29th June 2015

Intro

- The use of lifecontingencies R package (Spedicato 2013) to evaluate annuities will be introduced.
- lifecontingencies package provides a broad set of tools to perform actuarial calculations on life insurances. It requires a recent version of R (>=3.0), the markovchain package (Giorgio Alfredo Spedicato 2015) and Rcpp (Eddelbuettel 2013).
- This presentation is focused on actuarial evaluation of annuities present value distribution from the perspective of a Pension Fund.

- As a memo, the present value of a life annuity on a policyholder aged x is a random variable that depends on the future lifetime, K_x, future cash flows c_j and future (real) discount factors $v_j = \frac{1+r_j}{1+i_j} - 1$.
- It can be written as $\ddot{a}_{\overline{K_x+1}|} = \sum_{j=0}^{K_x} \left(v^j * c_j \right)$.
- Therefore it is a function of three random variables: the residual life time, K_t, the nominal (future) rates of interest, r_t, and the (future) inflation rates i_t, where $t \in \left[0, \frac{1}{k}, \ldots, K_t, \right]$.

- We will calculate APV from a classical central baseline scenario (defined life table and fixed financials hyphotheses). We will then allow for variability on these hyphoteses.
- In particular we will allow for process and parameter variance on the demographical side of calculation. Stochastic interest rate and inflation scenarios will be simulated as well.
- Slides will contain key code snippets, figures and tables. The dedicated repository (`https://bitbucket.org/spedygiorgio/rininsurance2015`) collect the full code needed to replicate the results.

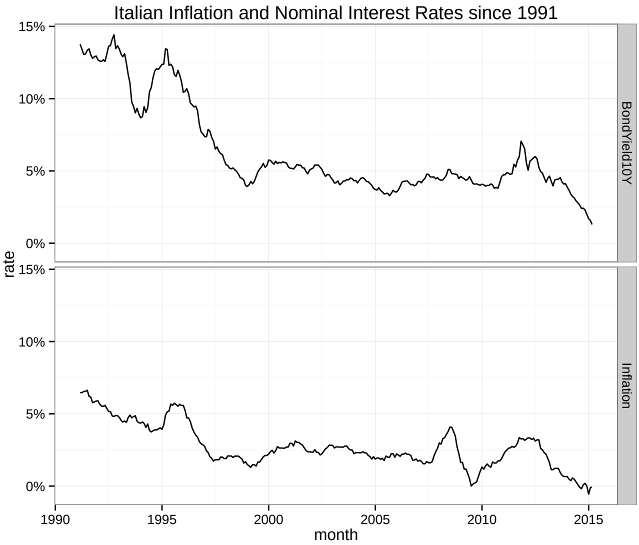
Collecting data

Collecting Financial from FRED

- We will retrieve economics from Federal Reserve Economic Data using quantmod package (Ryan 2015).
- Demographic tables used to project life tables will come from Human Mortality Database (University of California and Demographic Research 2015).

```
#retrieving Italian financial data from FRED
library(quantmod)
getSymbols(Symbols = 'ITACPIALLMINMEI',src='FRED') #CPI
getSymbols(Symbols='IRLTLT01ITM156N',
src='FRED') #Nominal IR (10Y Bond Yield)
```

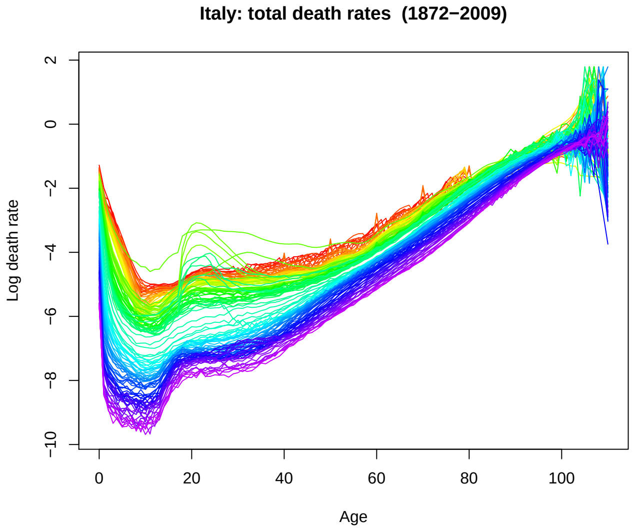
- we can observe a certain degree of structural dependency between i and r, that suggests the use of cointegration analysis for modeling.



Collecting demographics from HMD

- Retrieving data from Human Mortality Database (HMD) is easy as well

```
library(demography)
italyDemo<-hmd.mx("ITA", username="yourUN",
password="yourPW")
```


Italy: total death rates (1872−2009)
Log death rate
Age
2
0
−2
−4
−6
−8
−10
0
20
40
60
80
100

Defining baseline actuarial hyphotheses

- Our baseline hyphoteses are:
- retiree age: $x = 65$, cohort 1950.
- yearly pension benefit: 1, CPI indexed.
- frequency of payment of $\frac{1}{k}$ payments $k = 12$.
- mortality model: Lee Carter.
- no survivors benefit: no reversionary annuity component.
- only one interest rate (r): 10 year governemnt bond yield.
- Baseline inflation (i) and nominal interest rate (r) calculated as straight average of 10 Years values between 1999 and 2014 (Euro Currency Era). This yields a real discount factor around 2.4%.

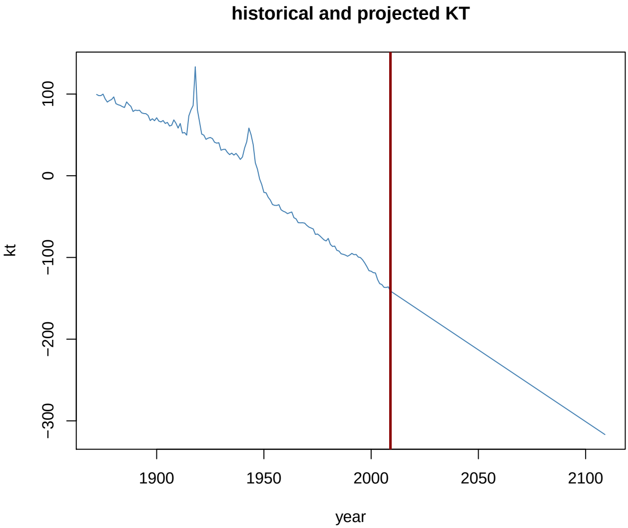
- xts package (Ryan and Ulrich 2014) functions have been used to calculate baseline figures.

```r
#use xts functions to calculate yearly averages
finHyph<-apply(apply.yearly(x =
italy.fin["1999/2014",2:1], FUN="mean"),2,"mean")
infl<-finHyph[1];nomRate<-finHyph[2]
realRate<-(1+nomRate)/(1+infl)-1
names(realRate)<-"RealRate"
c(nomRate, infl, realRate)
```

```
## BondYield10Y     Inflation      RealRate
##   0.04520074    0.02073512    0.02396862
```

Projecting mortality

- Mortality dynamics will follow Lee Carter model (Lee and Carter 1992): $q_{x,t} = e^{a_x + b_x * k_t + \varepsilon_{x,t}}$,
- We will start projecting a central scenario for mortality on which a baseline life table can be derived for the 1950 cohort.
- demography (Rob J Hyndman et al. 2011) and forecast (Hyndman and Khandakar 2008) packages calibrate and perform Lee Carter' model projections respectively.


historical and projected KT
kt
100
0
−100
−200
−300
1900
1950
2000
2050
2100
year

-The code below generates the matrix of prospective life tables

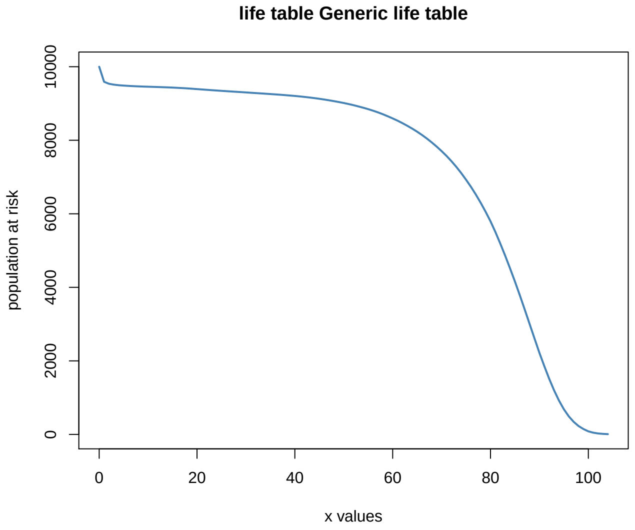
```
#getting and defining the life tables matrix
mortalityTable<-exp(italy.leecarter$ax+
italy.leecarter$bx%*%t(kt.full))
rownames(mortalityTable)<-seq(from=0, to=103)
colnames(mortalityTable)<-seq(from=1872,
to=1872+dim(mortalityTable)[2]-1)
```

- The function below returns the yearly death probabilities for a given birth cohort.

```
getCohortQx<-function(yearOfBirth)
{
  colIndex<-which(colnames(mortalityTable)
                ==yearOfBirth) #identify
  #the column corresponding to the cohort
  #definex the probabilities from which
  #the projection is to be taken
  maxLength<-min(nrow(mortalityTable)-1,
               ncol(mortalityTable)-colIndex)
  qxOut<-numeric(maxLength+1)
  for(i in 0:maxLength)
    qxOut[i+1]<-mortalityTable[i+1,colIndex+i]
  #fix: we add a fictional omega age where
  #death probability = 1
  qxOut<-c(qxOut,1)
  return(qxOut)
}
```

- Now we can get a baseline projected life table for a 1950 born.

```
#generate the life tables
qx1950<-getCohortQx(yearOfBirth = 1950)
lt1950<-probs2lifetable(probs=qx1950,type="qx",
name="Baseline")
at1950Baseline<-new("actuarialtable",x=lt1950@x,
lx=lt1950@lx,interest=realRate)
```



life table Generic life table
population at risk
0
2000
4000
6000
8000
10000
0
20
40
60
80
100
x values

Variability assessment

Process variance

- It is now easy to assess the expected curtate lifetime, $\mathring{e}_{65}$ and the APV of the annuity $\ddot{a}_{65}^{12}$, for $x = 65$.

```
#curtate lifetime @x=65
ex65.base<-exn(object = at1950Baseline,x = 65)
names(ex65.base)<-"ex65"
#APV of an annuity, allowing for fractional payment
ax65.base<-axn(actuarialtable = at1950Baseline,
x=65,k = 12);names(ax65.base)<-"ax65"
c(ex65.base,ax65.base)
```

```
##     ex65     ax65
## 18.72669 14.95836
```

- It is possible to simulate the process variance (due to “pure” chance on residual life time) of both K_t and $\ddot{a}_{\overline{K_x+1}|} = \sum_{j=0}^{K_x} \left(v^j * c_j \right)$.

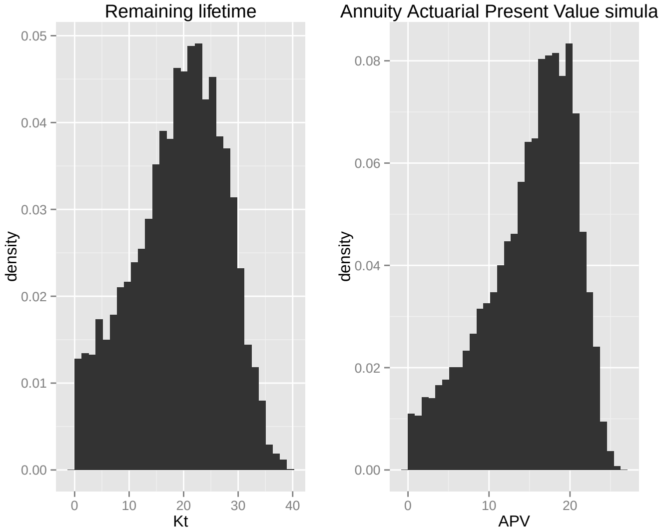
```
#simulate curtate lifetime
lifesim<-rLife(n=1e4,object=at1950Baseline,x=65,k=12,type = "Kx")
c(mean(lifesim),sd(lifesim))
```

```
## [1] 19.057092  8.263257
```

```
#simulate annuity present value
annuitysim<-rLifeContingencies(n=1e4,lifecontingency="axn",
object=at1950Baseline,x=65,k=12,parallel=TRUE)
c(mean(annuitysim),sd(annuitysim))
```

```
## [1] 14.943719  5.548229
```

Process Variance simulation



Stochastic life tables

- Estimation variance due to the demographic component arises from uncertainty in demographic assumption (stochastic life tables).
- We take into account this source of variability generating random life tables by randomizing k_t projections and by sampling from $\varepsilon_{x,t}$ on the Lee Carter above mentioned formula.

- The code that follows integrates the simulation of parameter and process variance on the demographic component.

```
###nesting process variance and parameter variance
numsim.parvar=31
numsim.procvar=31
### simulating parameter variance
tablesList<-list()
for(i in 1:numsim.parvar) tablesList[[i]] <-
  tableSimulator(lcObj = italy.leecarter,
  kt.model = italy.kt.model,
  coort = 1950,
  type = 'simulated',
  ltName = paste("table",i),
  rate=realRate
)
### simulating process variance
lifesim.full<-as.numeric(sapply(X=tablesList, FUN="rLife",n=numsim.proc
annuitysim.full<-as.numeric(sapply(X=tablesList, FUN="rLifeContingencie
```

Varying financial assumptions

Changing nominal interest rate assumption

- We assume only one financial asset: 10 years Gvt Bond.
- Interest rate dynamic follows a Vasicek model, $dr_t = \lambda * (r_t - \mu) * dt + \sigma * dW_t$.
- Therefore parameters can be estimated by least squares from the equation $S_t = a * S_{t-1} + b + \varepsilon_t$, as shown below. We have used the approach shown in `http://www.sitmo.com/article/calibrating-the-ornstein-uhlenbeck-model/`.

$$\lambda = -\frac{\ln a}{\delta}$$

$$\mu = \frac{b}{1-a}$$

$$\sigma = sd(\varepsilon)\sqrt{-\frac{2 * \ln(a)}{\delta * (1 - a^2)}}$$

- The function that follows calibrates the Vasicek interest rate mode on Italian data since 1991.

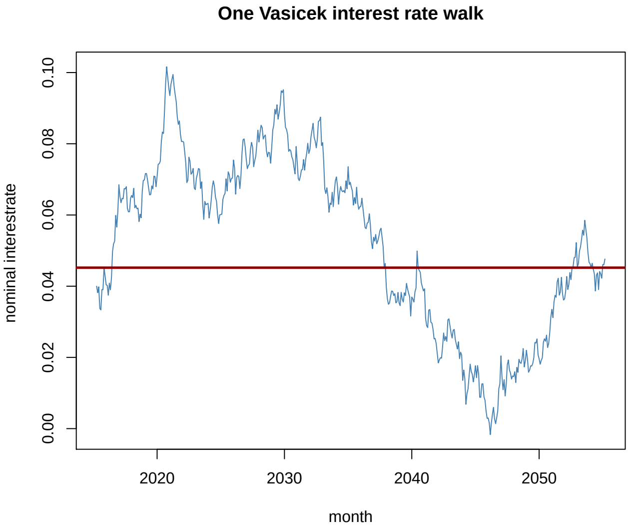
```
calibrateVasicek<-function(x, delta=1) {
  y<-as.zoo(x)
  x<-lag(y,-1,na.pad=TRUE)
  myDf<-data.frame(y=y,x=x)
  linMod <- lm(y~x,data=myDf)
  a<-coef(linMod)[2]
  b<-coef(linMod)[1]
  lambda<- as.numeric(-log(a)/delta)
  mu <- as.numeric(b/(1-a))
  sigma <-  as.numeric(((-2*log(a))/(delta*(1-a^2)))^.5*sd(linMod$resid
  out<-list(lamdba=lambda, mu=mu, sigma=sigma)
  return(out)
}
```

- The calibrated parameters follows. However the long term average has been judgmentally fixed to 4.52%.
- They will be used to simulate Vasicek - driven interest rates paths.

```
#calibrate
italianRatesVasicek<-calibrateVasicek(x =as.numeric(italy.fin[,1]),
delta=1/12)
italianRatesVasicek
```

```
## $lamdba
## [1] 0.1129055
##
## $mu
## [1] 0.01704393
##
## $sigma
## [1] 0.0102649
```

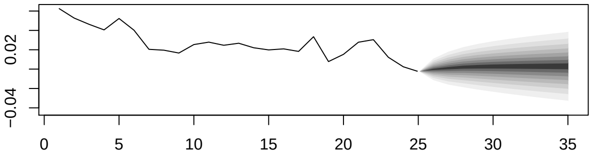
```
#simulate
randomInterestRateWalk<-ts(VasicekSimulations(M = 1,N = 40*12,r0 = 0.04
italianRatesVasicek$sigma ,dt = 1/12)[,1],start=c(2015,4),frequency=12)
```


One Vasicek interest rate walk
nominal interestrate
0.00
0.02
0.04
0.06
0.08
0.10
2020
2030
2040
2050
month

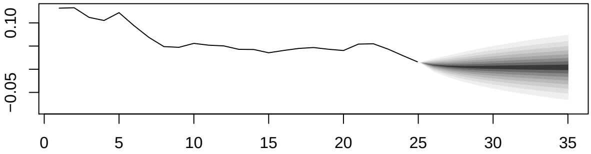
Modeling inflation dynamic

- The dependency between nominal rates and inflation is well know. Cointegration analysis is needed to avoid spurious regression.
- We have assumed the their relationship to be linear $inf_t = \alpha + \beta * r_t + \gamma_t$.
- Cointegration analysis on annualized data was peformed using package vars (Pfaff 2008). One cointegrating relationship was found to be quite well supported by empirical data.

Fanchart for variable Inflation



Fanchart for variable BondYield10Y

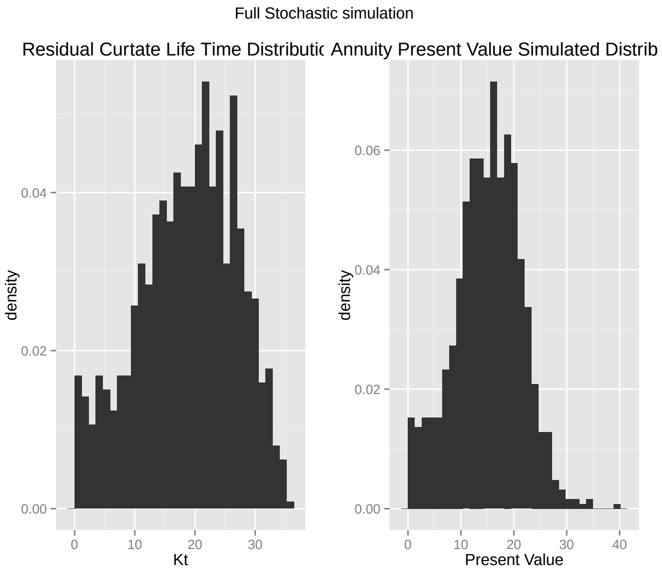


Pulling all together

This algorithm simulatrs the distribution of annuity's:

1. simulate K_t.
2. Project the fixed cash flow vector $\frac{1}{k}$ for all t in $0, \frac{1}{k}, \frac{2}{k}, \ldots, K_t$.
3. Simulate a r_t path.
4. Simulate a i_t path by the relation $i_t = \alpha + \beta * r_t + \gamma_t$.
5. Determine real interest rate vector and discount.

- The figure below shows the distribution of the K_t and $\ddot{a}_{\overline{K+1}|} = \sum_{j=0}^{K_x} \left(v^j * c_j \right)$ after allowing for process and parameter variance (for the demographic component).



- Expected value changes moderately and standard deviations increases slightly for K_x and $\ddot{a}_{\overline{K_x+1}|} = \sum_{j=0}^{K_x} \left(v^j * c_j \right)$ distributions.

```
#central scenario remark
c(ex65.base,ax65.base)
```

```
##     ex65     ax65
## 18.72669 14.95836
```

```
#assessing impact on expected value
apply(final.simulation,2,mean)/apply(processVarianceSimulations,2,mean)
```

```
## residualLifetime       annuityAPV
##        0.9836045        1.0722882
```

```
#assessing impact on volatility
apply(final.simulation,2,sd)/apply(processVarianceSimulations,2,sd)
```

```
## residualLifetime       annuityAPV
##         1.173815         1.125270
```

Other package's features

- Even if not discussed in the presentation, it is worth to remark that lifecontingencies package allows to calculate also the impact of reversionary annuity benefits (see below the calculation for $a_{\overline{xy}}$. Multiple decrements will be fully supported within 2015.

```
#baseline benefit
ax65.base
```

```
##     ax65
## 14.95836
```

```
#allowign for reversionary on 62
axn(actuarialtable = at1950Baseline,x = 65,k=12)+
 (axn(actuarialtable = at1950Baseline,x = 62,k=12)-
 axyzn(tablesList = list(at1950Baseline,at1950Baseline),
 x = c(65,62),status = "joint",k = 12))
```

```
## [1] 18.85357
```

Bibliography I

Eddelbuettel, Dirk. 2013. *Seamless R and C++ Integration with Rcpp*. New York: Springer.

Giorgio Alfredo Spedicato. 2015. *Markovchain: Discrete Time Markov Chains Made Easy*.

Hyndman, Rob J, and Yeasmin Khandakar. 2008. "Automatic Time Series Forecasting: The Forecast Package for R." *Journal of Statistical Software* 26 (3): 1–22. http://ideas.repec.org/a/jss/jstsof/27i03.html.

Lee, Ronald D, and Lawrence R Carter. 1992. "Modeling and Forecasting US Mortality." *Journal of the American Statistical Association* 87 (419). Taylor & Francis Group: 659–71.

Pfaff, Bernhard. 2008. "VAR, SVAR and SVEC Models: Implementation Within R Package vars." *Journal of Statistical Software* 27 (4). http://www.jstatsoft.org/v27/i04/.

Rob J Hyndman, Heather Booth, Leonie Tickle, and John Maindonald. 2011. *Demography: Forecasting Mortality, Fertility, Migration and Population Data*. http://CRAN.R-project.org/package=demography.

Ryan, Jeffrey A. 2015. *Quantmod: Quantitative Financial Modelling Framework*. http://CRAN.R-project.org/package=quantmod.

Bibliography II

Ryan, Jeffrey A., and Joshua M. Ulrich. 2014. *Xts: EXtensible Time Series*. http://CRAN.R-project.org/package=xts.

Spedicato, Giorgio Alfredo. 2013. "The lifecontingencies Package: Performing Financial and Actuarial Mathematics Calculations in R." *Journal of Statistical Software* 55 (10): 1–36. http://www.jstatsoft.org/v55/i10/.

University of California, and Max Planck Institute for Demographic Research. 2015. "Human Mortality DataBase." Available at www.mortality.org or www.humanmortality.de (data downloaded on 14th June 2015).