

ESTIMAÇÃO DO TEMPO DE VIDA FUTURO ALEATÓRIO INDIVIDUAL PARA GRUPOS DE INDIVÍDUOS

Cristaine Silva Corrêa¹

Resumo: Para muitos problemas de simulação estocástica de populações no tempo se faz necessária a estimação do tempo de vida futuro dos indivíduos de forma também estocástica. O presente trabalho se dedica a apresentar uma função desenvolvida no software livre R para a estimação deste tempo de vida futuro a partir de matrizes de idades e de sexo dos indivíduos, assumindo, para cada sexo, uma distribuição de mortalidade descrita por uma tabela de vida diferente. Os resultados se mostram consistentes em termos de distribuições de probabilidades e velocidade computacional, sendo, portanto, úteis para diversos trabalhos futuros acerca dessa temática.

Palavra-chave: Tempo de vida futuro; microssimulação; tabela de vida

Abstract: For many problems of stochastic simulation of populations over time it is necessary to estimate the future life span of individuals in a stochastic way. The present work is devoted to presenting a function developed in the free software R for the estimation of this future life time from matrices of ages and sex of the individuals, assuming, for each sex, a distribution of mortality described by a life table different. The results are consistent in terms of probability distributions and computational speed, and are therefore useful for several future studies on this subject.

Key-words: Future life time; Microsimulation; Life table.

1. Introdução

Tabelas de vida são instrumentos utilizados para retratar o nível e a estrutura de mortalidade de uma população (PRESTON; HEUVELINE; GUILLOT, 2001), ou de qualquer outra ação que caracterize saída dos indivíduos de uma população, como a aprovação em um teste que marque a saída dos indivíduos da população de estudantes, ou a aposentadoria, que marca a saída da população de trabalhadores. Pelas funções da tabela de vida é possível estimar o tempo esperado de vida futuro dos indivíduos (BOWERS et al., 1997), medida útil para projeções populacionais, atuariais e para políticas públicas.

Entre os métodos de projeção populacionais, um dos mais complexos é a microssimulação, um método probabilístico que projeta indivíduo a indivíduo, separadamente, por meio de sorteios aleatórios (CORRÊA, 2014; ZHAO, 2006), em vez de projetar a população como um todo por meio de uma distribuição média, como fazem os métodos de macrossimulação (VOS; PALLONI, 1989). A microssimulação é particularmente útil para projeção de situações complexas, que envolvem muitas variáveis que interagem entre si afetando o resultado da população total, e para a projeção de redes pessoais, já que todos os resultados são armazenados e é possível identificar as relações estabelecidas no tempo entre os indivíduos (CORRÊA, 2014; VOS; PALLONI, 1989). Uma de suas vantagens é que, após um grande número de repetições com números aleatórios diferentes, possibilitar a estimação de intervalos de confiança para os resultados (CORRÊA, 2014; VOS; PALLONI, 1989; ZHAO, 2006). Contudo, o tempo de processamento de uma microssimulação pode ser elevado em comparação aos outros métodos de projeção populacional (CORRÊA, 2014).

¹Departamento de Demografia e Ciências Atuariais da UFRN - criscorrea@ccet.ufrn.br

Diante dessa situação, o objetivo deste artigo é apresentar o desenvolvimento de um algoritmo eficiente para estimar tempo de vida futuro individual para grupos de indivíduos, um dos procedimentos mais demorados da microssimulação. Esse algoritmo foi desenvolvido em R, um software livre e aberto, de forma que pode ser utilizado e melhorado por outros pesquisadores. A função desenvolvida pode ser aplicada para qualquer matriz de indivíduos com idades e sexos diversificados que possam ser agrupados por idade e sexo iguais (ou outra característica de interesse e diferenciação pela TV). Portanto, os resultados desse trabalho são úteis tanto para projeções por microssimulação quanto para qualquer outro método que se utilize da estimação do tempo futuro em determinada população em suas estimativas.

2. Revisão de literatura

No R há alguns pacotes destinados a funções demográficas e projeções populacionais.

O pacote LifeTable é útil para construção de tábuas de vida baseada em algumas informações sobre a mortalidade de uma região e em tabelas modelos, mas não apresenta nenhuma função para estimação do tempo de vida futuro individual (SHARROW; SEVCIKOVA, 2012).

O pacote Lifecontingencies conta com a função rLife para a estimação do tempo de vida futuro para n indivíduos com idade x partindo de uma tabela de vida declarada como da classe lifetable (SPEDICATO; KAINHOFER; OWENS, 2014). Contudo, essa função não se aplica a uma matriz de vetores com idades diferentes ou sexos diferentes de uma só vez, de forma que, se utilizada nesse caso, teria que ser aplicado para cada célula da matriz individualmente, aumentando o tempo de processamento.

O pacote demography, do R, também se dedica a funções demográfica e projeta populações no tempo (HYNDMAN, 2014). Esse pacote conta, ainda, com as funções lca, que projeta a fecundidade ou a mortalidade com o método de Lee e Carter, um método probabilístico, mas que estima a função média, e não tempos de vida individuais a partir de uma função média já projetada, como requer a microssimulação. Além disso, as funções pop.sim e simulate.fmforecast, que simula a população futura por reamostragem pressupondo uma função de mortalidade, fecundidade e migração o fazem para o total da população ou de grupos etários, e não individualmente.

A função apresentada neste trabalho, portanto, contribui para o conjunto de funções já desenvolvido e disponível no R ao poder ser aplicada uma população caracterizada por idade e sexo oferecendo resultados individuais, insumos essenciais para as microssimulações.

3. Metodologia

Tábuas de vida representam a experiência de mortalidade de uma coorte hipotética e apresentam, para cada idade x , funções associadas às probabilidades de estar vivo nesta população (BOWERS *et al.*, 1997). Assim, a primeira coluna da tabela de vida representa a idade exata x dos indivíduos, e a segunda representa o número de indivíduos vivos à exata idade x , l_x , dos l_0 indivíduos iniciais à idade 0.

Seja X a variável aleatória idade à morte, a função distribuição do tempo de vida futuro de um indivíduo a partir de uma idade exata x , $T(X)$, pode ser estimada com base em relações de l_x (BOWERS *et al.*, 1997). Para cada idade $y \leq x$, a probabilidade de morrer entre as idades exatas x e $x + 1$, é dada por

${}_xq_y = P(x < X < x+1 / X > y) = \frac{l_x - l_{x+1}}{l_y}$. Portanto, para qualquer idade x tem-se, pela

razão dos l_x a partir daquela idade, toda a distribuição de probabilidade de morte em tempo futuro (BOWERS *et al.*, 1997).

Tendo-se essa distribuição de probabilidades pode-se, então, estimar o tempo de vida futuro para cada indivíduo de forma aleatória. Para implementar este algoritmo recorreu-se ao R, um ambiente livre e aberto que permite programação (R DEVELOPMENT CORE TEAM, 2013). A função EstimaTx é a função desenvolvida para estimar o tempo de vida futuro dos indivíduos com idade x assumindo um vetor de valores da função l_x da tabela de vida a partir da idade y desses indivíduos. A função “sample” do R é a responsável por fazer os sorteios aleatórios baseados na distribuição de probabilidade de morte a cada idade dada por ${}_xq_y$. Seja ω a maior idade possível de sobreviver, para cada idade y os possíveis valores de tempo de vida futuro vão de 0 até $\omega - y$.

```
EstimaTx=function(n,lx){ #Estima Tx
  lxMais1=c((lx[2:length(lx)]),0)
  qy=(lx-lxMais1)/lx[1]; qy=qy[1:(length(lx))]
  TxPossivel=0:(length(lx)-1)
  Tx=sample(TxPossivel,n,replace = T,qx
  Tx
}
```

Contudo, cada indivíduo da população pode ter uma idade atual y diferente, de modo que para cada indivíduo deve-se utilizar um conjunto de valores de l_y e de l_x diferentes, já que $y \leq x$. Resta, portanto, organizar um algoritmo de forma tal que essa função de estimação do tempo de vida futuro estime tempos de vida futuros aleatórios para indivíduos com idades e sexos iguais dentre um conjunto de indivíduos com idades e sexos variados. Esse processo será realizado em quatro passos. Primeiro identifica-se quais e quantos são os valores de idade e sexo que aparecem na matriz inicial. Depois se identifica as posições na matriz com idade e sexo iguais pela função **procNdif**. Em seguida, estima-se o número de tempos de vidas aleatórios futuros necessários para cada grupo de idade e sexo iguais. Por fim, atribuiu-se esses tempos de vida futuros às posições correspondentes na matriz. A função que executa esses passos é a função GeraTx.

```
GeraTx=function(MatIdade,MatSexo,lxM,lxF){
  #MatIdad = matriz com idades
  #MatSexo = Matriz com sexos (1 para Mulheres e 2 para homens)
  #lxM = tabela de vida para homens
  #lxF = tabela de vida para mulheres
  x=as.vector(MatIdade) #lê como vetor
  Sexo=as.vector(MatSexo) #lê como vetor
  VetorTx=rep(NA,length(x)) ##Vetor onde serão armazenados os resultados
  IndSexoF = which(Sexo %in% 1) #Identifica as posições na matriz em que há Mulheres
  IndSexoM = which(Sexo %in% 2) #Identifica as posições na matriz em que há Homens

  #Para sexo masculino
```

```

xM=x[IndSexoM]# Idade dos homens Novo vetor
xDiferentesM=procNdif(xM)  ##Identifica valores diferentes de
idades - função elaborada pelo autor
mM=length(xDiferentesM)      ##Identifica quantos são os
valores diferentes
xresM=as.vector(array(NA,dim=c(length(xM),1))) ##Cria vetor
para guardar resultados
for(i in 1:mM){
inds = which(xM %in% xDiferentesM[i])  ##identifica posições
do vetor que têm a mesma idade
nInds=length(inds)      ##Identifica o número de posições
do vetor que têm a mesma idade
if(xDiferentesM[i]>=(length(lxM)-1)){  ##Só estima para idades
abaixo da idade limite da tabela de vida. Diminui em 1
porque a tabela de vida começa na idade 0.
xresM[inds]=rep(0,nInds)  ##Se for maior que o limite
da tábu de vida, atribui 0.
}
else{
##Se for menor, calcula.
lxNovo=lxM[(xDiferentesM[i]+1):(length(lxM))]  ##Cria
novo vetor de lx apenas com valores acima da idade i
xresM[inds]=EstimaTx(nInds,lxNovo)
##função elaborada pelo autor
}
}
}
#Para sexo feminino
xF=x[IndSexoF]# Idade das mulheres Novo vetor
xDiferentesF=procNdif(xF)  ##Identifica valores diferentes de
idades
mF=length(xDiferentesF) ##Identifica quantos são os valores diferentes
xresF=as.vector(array(NA,dim=c(length(xF),1))) ##Cria vetor para
guardar resultados
for(i in 1:mF){
inds = which(xF %in% xDiferentesF[i])  ##identifica posições
do vetor que têm a mesma idade
nInds=length(inds)      ##Identifica o número de posições
do vetor que têm a mesma idade
if(xDiferentesF[i]>=(length(lxF)-1)){  ##Só estima para idades
abaixo da idade limite da tabela de vida. Diminui em 1
porque a tabela de vida começa na idade 0.
xresF[inds]=rep(0,nInds)  ##Se for maior que o limite
da tábu de vida, atribui 0.
}else{
##Se for menor, calcula.
lxNovo=lxF[(xDiferentesF[i]+1):(length(lxF))]  ##Cria novo
vetor de lx apenas com valores acima da idade i
xresF[inds]=EstimaTx(nInds,lxNovo) ##função elaborada
pelo autor
}
}
}

```

```

VetorTx[IndSexoM]=xresM
  VetorTx[IndSexoF]=xresF
  MatrizTx=array(VetorTx,dim=c(nrow(MatIdade),ncol(MatIdade)))
##Transforma vetor em matriz
  MatrizTx
}
procNdif=function(vet){  ##Procura valores diferentes em um vetor

  n=length(vet)
  indres=order(vet)
  res=vet[indres]
  aux=res[1]

  for(i in 2:n){
    if(res[i-1]!=res[i]){
      temp=res[i]
      aux=c(aux,temp)
    }
  }

  aux
}

```

4. Resultados

Para comprovar a eficácia e eficiência da função desenvolvidas simulou-se o tempo de vida futuro para matrizes de idade e sexo para 10 milhões de indivíduos, metade de cada sexo, tendo as funções de mortalidade definidas pelas tábuas de vida do IBGE 2010 para cada sexo. A Figura 1 apresenta a comparação entre essas duas distribuições para indivíduos inicialmente com 0, 40 e 80 anos de vida, para homens e para mulheres. Como pode-se observar, os resultados teóricos e simulados são praticamente idênticos, mostrando a eficácia do função proposta, que também é eficiente, já que os resultados são gerados quase instantaneamente pelo R.

5. Conclusão

Este trabalho teve como objetivo apresentar uma função de estimação do tempo de vida aleatório a partir de matrizes de idade e sexo dos indivíduos e da tabela de vida diferenciada por sexo. As simulações realizadas confirmam que a função desenvolvida estima coerentemente os valores desejados em um tempo mínimo, de forma que os resultados têm distribuição quase idêntica à distribuição teórica. Espera-se, com este trabalho, contribuir para o avanço das projeções demográficas probabilísticas, assim como para o desenvolvimento de outras áreas que demandem microssimulação estocástica.

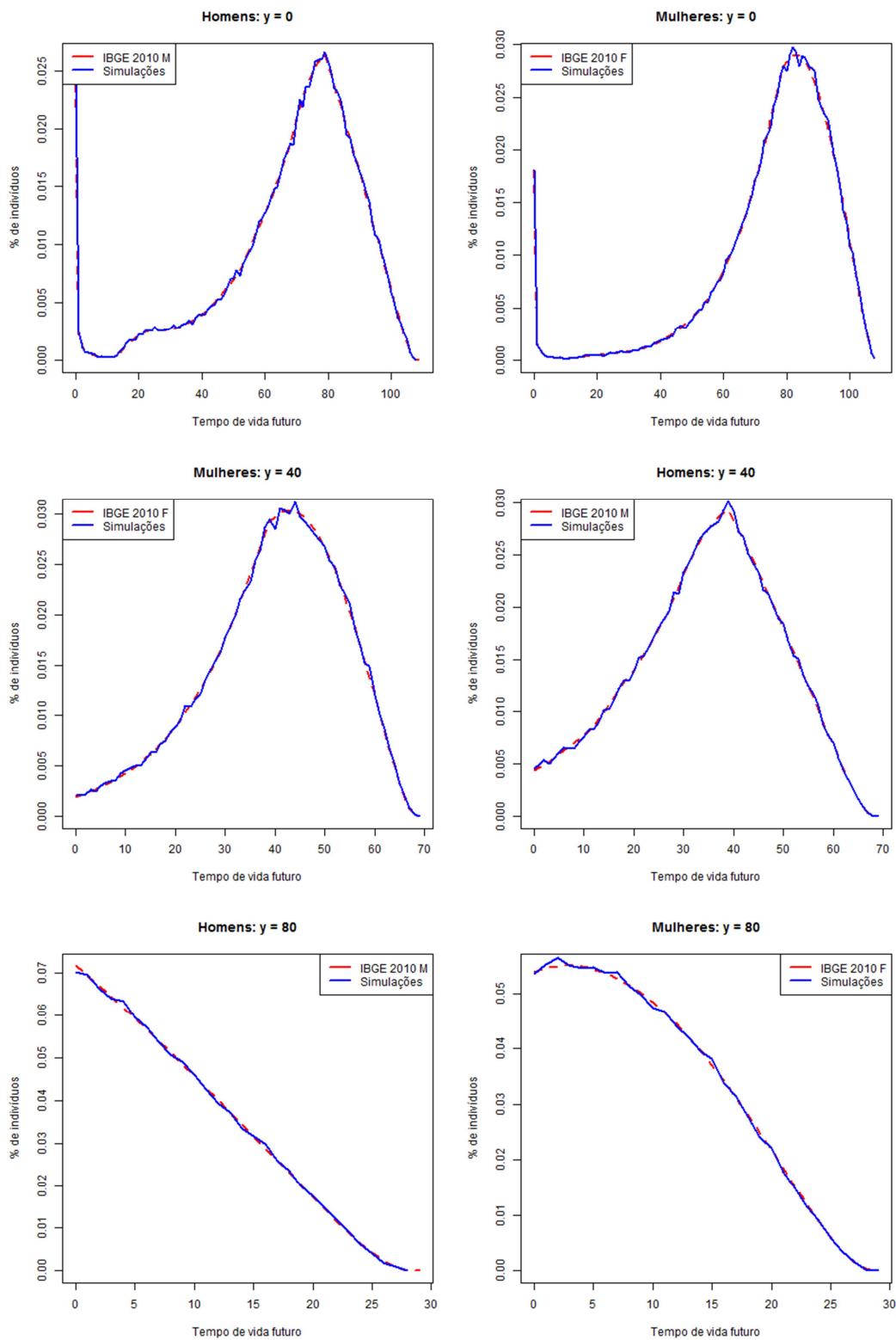


Figura 1 - Distribuição do tempo de vida futuro por 100 mil simulações e pela tábua IBGE 2010 para homens e mulheres com idades iniciais de 0, 40 e 80 anos.

Referências

- [1] BOWERS, N. L. et al. Actuarial Mathematics. [S.l.]: Society of Actuaries, 1997.
- [2] CORRÊA, Cristiane Silva. Tamanho populacional e aleatoriedade de eventos demográficos na solvência de RPPS municipais capitalizados. 2014. 252 f. Tese (Doutorado em Demografia) – CEDEPLAR/UFMG, Belo Horizonte, MG, 2014.
- [3] HYNDMAN, Rob J. Package “demography”: Forecasting mortality, fertility, migration and population data. [S.l: s.n.], 2014. Disponível em: <<https://cran.r-project.org/web/packages/demography/demography.pdf>>.
- [4] PRESTON, Samuel H; HEUVELINE, Patrick; GUILLOT, Michel. Demography: measuring and modeling population processes. Malden, MA: Blackwell Publishers, 2001.
- [5] R DEVELOPMENT CORE TEAM. R: A language and environment for statistical computing. Vienna, Austria: [s.n.], 2013. Disponível em: <<http://www.R-project.org/>>.
- [6] SHARROW, David J.; SEVCIKOVA, Hana. Package “LifeTables”: A package to implement HMD model life table system. [S.l: s.n.], 2012. . Acesso em: 8 abr. 2015.
- [7] SPEDICATO, Giorgio Alfredo; KAINHOFER, Reinhold; OWENS, Kevin J. Package “lifecontingencies”: A package to perform actuarial mathematics for life contingencies insurances. [S.l: s.n.], 2014. Disponível em: <<http://cran.r-project.org/web/packages/lifecontingencies/lifecontingencies.pdf>>.
- [8] VOS, Susan De; PALLONI, Alberto. Formal Models and Methods for the Analysis of Kinship and Household Organization. Population Index. Office of Population Research, 2. v. 55, p. 174–198, Summer 1989.
- [9] ZHAO, Zhongwei. Computer microsimulation and historical study of social structure: A comparative review of SOCSIM and CAMSIM. Revista de Demografia Histórica, XXIV. v. 2, p. 59–88, 2006.