SP_DemTech2_Problem Set 3

Sara Peters

2023-10-04

#set up

```
#Set working directory
setwd("C:/Users/saraa/OneDrive - UW-Madison/SOC 756- Demography Techniques II/Problem Sets/DemTe
ch2/Problem Set 3")

#load libraries
# install.packages("dplyr")
# install.packages("tidyverse")
# install.packages("ggplot2")
# install.packages("HMDHFDplus")

library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following object is masked from 'package:kableExtra':
##
##     group_rows
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(tidyverse)
```

```
## ── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0 ──
## ✓ forcats   1.0.0     ✓ readr     2.1.4
## ✓ ggplot2   3.4.3     ✓ stringr   1.5.0
## ✓ lubridate 1.9.2     ✓ tibble    3.2.1
## ✓ purrr     1.0.1     ✓ tidyr     1.3.0
```

```
## ── Conflicts ────────────────────────────────────────── tidyverse_conflicts() ──
## ✖ dplyr::filter()     masks stats::filter()
## ✖ dplyr::group_rows() masks kableExtra::group_rows()
## ✖ dplyr::lag()        masks stats::lag()
## ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to becom
e errors
```

```
library(ggplot2)
library (HMDHFDplus)

#HMD login info:
#username: speters27@wisc.edu
#password: @DemTech2

# Set the 'scipen' option to a large value to prevent scientific notation
options(scipen = 999)

# Set the 'digits' option to control the number of decimal places
options(digits = 6)  # Change the number to the desired decimal places
```

#1. Approximately 85,000 adolescents turn 16 each year in Wisconsin. Data from Fohr et al., 2005 suggest that the probability of being involved in a non-fatal motor vehicle accident among Wisconsin 16-year-olds is roughly 0.0486. The authors find that the probability declines dramatically with age, reaching 0.0145 by age 30.

#Assume that, in each year of life, the probability of experiencing a non-fatal motor vehicle accident is equal to 0.062 - 0.000053*(age2), where age is defined in discrete one-year intervals.

#Acquire an account with the Human Mortality Database and the Human Fertility Database. Use the HMDHFDplus package in R to obtain the 2005 single year age-specific death probabilities from the Human Mortality Database.

#Get Data from Human Mortality Database and Construct a standard life table

#Install the R package HMDHFDplus and get the lifetable values directly: https://cran.r-project.org/web/packages/HMDHFDplus/HMDHFDplus.pdf (https://cran.r-project.org/web/packages/HMDHFDplus/HMDHFDplus.pdf) The commands ask you to supply your user name and password, you'll still need to sign up at HMD first.

```
#get 2005 single year age-specific death probabilities from the Human Mortality Database
US_lt <-readHMDweb(CNTRY = "USA", item = "bltper_1x1", username = "speters27@wisc.edu" , passwor
d = "@DemTech2" )

#Filter age-specific death probabilities for only values where year = 2005
US_qx05 <- US_lt %>%
  filter(Year == 2005) %>%
  subset (Age  == 16:31)
```

```
## Warning in Age == 16:31: longer object length is not a multiple of shorter
## object length
```

```r
#convert dataframe to all numeric types
columns_to_process <- c("Age", "lx", "dx", "qx", "Lx")

for (col_name in columns_to_process) {
  US_qx05 [, col_name] <- as.numeric(US_qx05 [, col_name])
}

str(US_qx05)
```

```
## 'data.frame':    16 obs. of  11 variables:
##  $ Year        : int  2005 2005 2005 2005 2005 2005 2005 2005 2005 2005 ...
##  $ Age         : num  16 17 18 19 20 21 22 23 24 25 ...
##  $ mx          : num  0.00049 0.00065 0.00082 0.00091 0.00091 0.001 0.001 0.001 0.00098 0.000
99 ...
##  $ qx          : num  0.00049 0.00065 0.00081 0.00091 0.00091 0.001 0.001 0.001 0.00098 0.000
99 ...
##  $ ax          : num  0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 ...
##  $ lx          : num  98984 98936 98871 98791 98701 ...
##  $ dx          : num  48 64 81 90 90 99 99 98 97 97 ...
##  $ Lx          : num  98960 98903 98831 98746 98656 ...
##  $ Tx          : int  6169668 6070709 5971805 5872974 5774229 5675573 5577012 5478549 5380185
5281918 ...
##  $ ex          : num  62.3 61.4 60.4 59.5 58.5 ...
##  $ OpenInterval: logi  FALSE FALSE FALSE FALSE FALSE FALSE ...
```

```r
#Construct multiple decrements table

US_qx05_md = US_qx05 %>%
  select(Age, lx, dx, qx, Lx) %>%
  rename( x = Age) %>%
  mutate( nqx_accident = 0.062 - 0.000053 * (x^2) ) %>%
  mutate (nqx_death = qx) %>%
  mutate (nqx_all = nqx_death + nqx_accident)


#set radix at 85,000

US_qx05_md$lx_all[1] = 85000
    for(i in 2:nrow(US_qx05_md)) {
      US_qx05_md$lx_all[i] <- US_qx05_md$lx_all[i-1]*(1-US_qx05_md$nqx_all[i-1])
    }

#calculate dx for new radix
#ndx for Car
for(i in 1:nrow(US_qx05_md)) {
  US_qx05_md$ndx_accident[i] <- US_qx05_md$lx_all[i]*US_qx05_md$nqx_accident[i]
}


#ndx for death
for(i in 1:nrow(US_qx05_md)) {
  US_qx05_md$ndx_death[i] <- US_qx05_md$lx_all[i]*US_qx05_md$nqx_death[i]
}

##ndx all
US_qx05_md$ndx_all <- US_qx05_md$lx_all[1] - US_qx05_md$lx_all[2]
for(i in 1:nrow(US_qx05_md)) {
  US_qx05_md$ndx_all[i] <- US_qx05_md$lx_all[i]-US_qx05_md$lx_all[i+1]
}

US_qx05_md$ndx_all[16] = US_qx05_md$lx_all[16]




#create lx columns for death and accidents

# Initialize an empty vector
US_qx05_md$lx_accident <- numeric(length(US_qx05_md$lx_all))
US_qx05_md$lx_death <- numeric(length(US_qx05_md$lx_all))

#create lx accident
for (i in 1:nrow(US_qx05_md)) {

  US_qx05_md$lx_accident[i] <- sum(US_qx05_md$ndx_accident[i:16])

}
```

```
#create lx death
for (i in 1:nrow(US_qx05_md)) {

  US_qx05_md$lx_death[i] <- sum(US_qx05_md$ndx_death[i:16])

}


#Print life table
print (US_qx05_md)
```

```
##     x    lx  dx      qx    Lx nqx_accident nqx_death  nqx_all  lx_all
## 17 16 98984  48 0.00049 98960     0.048432   0.00049 0.048922 85000.0
## 18 17 98936  64 0.00065 98903     0.046683   0.00065 0.047333 80841.6
## 19 18 98871  81 0.00081 98831     0.044828   0.00081 0.045638 77015.2
## 20 19 98791  90 0.00091 98746     0.042867   0.00091 0.043777 73500.3
## 21 20 98701  90 0.00091 98656     0.040800   0.00091 0.041710 70282.7
## 22 21 98611  99 0.00100 98561     0.038627   0.00100 0.039627 67351.2
## 23 22 98512  99 0.00100 98462     0.036348   0.00100 0.037348 64682.3
## 24 23 98413  98 0.00100 98364     0.033963   0.00100 0.034963 62266.5
## 25 24 98315  97 0.00098 98267     0.031472   0.00098 0.032452 60089.5
## 26 25 98218  97 0.00099 98170     0.028875   0.00099 0.029865 58139.5
## 27 26 98121  98 0.00100 98072     0.026172   0.00100 0.027172 56403.2
## 28 27 98023  99 0.00101 97974     0.023363   0.00101 0.024373 54870.6
## 29 28 97924  99 0.00101 97875     0.020448   0.00101 0.021458 53533.2
## 30 29 97825  97 0.00099 97777     0.017427   0.00099 0.018417 52384.5
## 31 30 97728 101 0.00103 97678     0.014300   0.00103 0.015330 51419.7
## 32 31 97628 106 0.00109 97575     0.011067   0.00109 0.012157 50631.5
##    ndx_accident ndx_death   ndx_all lx_accident lx_death
## 17     4116.720   41.6500  4158.370   34059.996 924.0700
## 18     3773.930   52.5471  3826.477   29943.276 882.4200
## 19     3452.435   62.3823  3514.818   26169.346 829.8729
## 20     3150.739   66.8853  3217.624   22716.911 767.4907
## 21     2867.535   63.9573  2931.492   19566.172 700.6054
## 22     2601.576   67.3512  2668.927   16698.637 636.6481
## 23     2351.072   64.6823  2415.754   14097.062 569.2969
## 24     2114.758   62.2665  2177.025   11745.990 504.6146
## 25     1891.137   58.8877  1950.025    9631.231 442.3480
## 26     1678.778   57.5581  1736.336    7740.094 383.4603
## 27     1476.183   56.4032  1532.586    6061.316 325.9022
## 28     1281.941   55.4193  1337.360    4585.133 269.4991
## 29     1094.647   54.0685  1148.716    3303.192 214.0798
## 30      912.905   51.8606   964.765    2208.545 160.0113
## 31      735.302   52.9623   788.264    1295.640 108.1506
## 32      560.338   55.1883 50631.461     560.338  55.1883
```

```
#qxi = given probability
#dxi= nqxi * lx
#dxd = dx- dxi
#qxd = qx * dxd/dx
#lxi = sumdix above that age
```

##Answer the following questions:

A.What proportion of Wisconsinites who live to age 16 will live to age 31 without experiencing a motor vehicle accident?

```
#npx = lx + n/ lx, lx31//x16

31-16
```

```
## [1] 15
```

```
#15p16 =

US_qx05_md$lx_all[16]/US_qx05_md$lx_all[1]
```

```
## [1] 0.595664
```

```
#0.595664 or 60%
```

B.Among those who live to age 25 accident-free, what is the probability of experiencing an accident by age 31? sumndx_cause/lx, = sumndx between 25 and 31/l25

```
sum(US_qx05_md$ndx_accident[10:15])/US_qx05_md$lx_all[10]
```

```
## [1] 0.123492
```

C.Among those who survive to age 16, what is the probability of dying without experiencing an accident by age 31?

sumdx/lx = sumdx between 16 and 31/l16

```
# sumdx/lx = sumdx between

sum(US_qx05_md$ndx_death[1:15])/US_qx05_md$lx_all[1]
```

```
## [1] 0.0102221
```

D.If the experience of accidents and the probability of dying are process-dependent, is your estimate for C an overestimate or an underestimate of the true probability?

This life table process assumes that every person in the population has an equal risk of experiencing the event under consideration. But in reality, not everyone has the same amount of risk for experience an event like a non-fatal motor vehicle accident. So the our calculation in C is an overestimate of the probability of dying.

E.Push your code to GitHub and share the link with someone from class. Answer here the name of the person(s) to whom you shared the link.

Johanna. https://github.com/saraapeters/DemTech2 (https://github.com/saraapeters/DemTech2)

##2. A cohort of never-married individuals are subject to two forces of decrement assumed to be constant within each interval x to x+n. The following age-specific rates were calculated for this cohort: # - age-specific mortality rates for never-married individuals:$MnxD$ # - age-specific first marriage rages: $MnxM$ These are assumed to be zero above the age of 50.

A.Write an expression in terms of these age-specific transition rates ($MnxD$ and $MnxM$) for the probability of being never married at age 50 for a newborn.

B.Write an expression in terms of these age-specific transition rates for the probability of being never married at age 50 for a newborn, net of mortality.

C.A second cohort is subject to the same first marriage rates described above but experiences mortality rates that are 20% lower at each age. For this second cohort, write expressions for the probabilities described in (1) and (2). Compare these probabilities with those of the first cohort; which are larger?

D.You study the population of Sulawesi and observe that the proportion of newborns that are never married at age 50 has stayed constant over time. Yet mortality conditions of individuals have improved at every age during the same period. What can be concluded about trends in first marriage rates in Sulawesi? Why?

The size of the population at age 50 is increasing as a result of declining mortality rates which means that the denominator for determining the proportion of the population getting married is increasing. If the proporation of newborns that are never married at age 50 has stayed constant, this indicates that the number of people getting married before age 50 is also increasing.