Demographic Methods - Practical 4 (Life Tables II)

2025-10-20

The heading of the R script

The R Script starts by clearing all generated data if any. Packages and libraries are not required this time.

```
rm(list = ls())
```

Reading the data

In this exercise, we are given the number of survivors l_x from two life tables corresponding to the male and female populations in Athens in 1981. Additionally, we are provided with the life expectancy at age 10 e_{10}, for each population. As shown in the command lines below, the l_x values and the exact ages x are combined into a vector using the R function "c()".

```
x                    = c(0,1,5,10,15,20,25,30,35,40,45,50,55,60,65,70,75,80,85)
lx_m                 = c(100000, 97731, 97203, 96925, 96723, 96364, 95894, 95445,
                         94958, 94343, 93436, 91883, 89212, 84660, 77483, 66830,
                         52309, 34968, 18173)
lx_f                 = c(100000, 98179, 97884, 97752, 97658, 97512, 97313, 97068,
                         96761, 96343, 95721, 94669, 92945, 90131, 85492, 77598,
                         65075, 47402, 27701)
e10_m                = 63.885
e10_f                = 66.385

print(data.frame(x,lx_m,lx_f))
```

```
##     x   lx_m   lx_f
## 1   0 100000 100000
## 2   1  97731  98179
## 3   5  97203  97884
## 4  10  96925  97752
## 5  15  96723  97658
## 6  20  96364  97512
## 7  25  95894  97313
## 8  30  95445  97068
## 9  35  94958  96761
## 10 40  94343  96343
## 11 45  93436  95721
## 12 50  91883  94669
## 13 55  89212  92945
## 14 60  84660  90131
## 15 65  77483  85492
## 16 70  66830  77598
## 17 75  52309  65075
## 18 80  34968  47402
## 19 85  18173  27701
```

Data preparation

Using the R function "cbind($name_1$ = $vector_1$, ..., $name_n$ = $vector_n$)", we can create a matrix with n columns—one column of l_x values per population—and one row per age interval $[x, x+n)$. Similarly, we can create a vector with two elements, one e_{10} value for each population. Naming the elements of vectors or matrices is optional, but it helps with identification later. There is also a complementary R function, "rbind()", which performs the same operation by rows.

```
lx                   = cbind(male = lx_m, female = lx_f)
e10                  = cbind(male = e10_m, female = e10_f)
```

As is standard in life table calculations, we can define the length of age intervals based on the values of x, make assumptions about the age distribution of deaths within each interval using ${}_na_x$, and identify the open-ended age interval using logical operators.

```
n                    = c(diff(x,1),NA)
nax                  = c(0.1,0.4,rep(0.5, length(x) - 3),NA)
sEL                  = !is.na(n)
```

Note that the value of ${}_na_x$ may also vary by sex, but no adjustments have been made to account for this.

Life table formulae

$$ {}_nM_x = \frac{{}_nD_x}{{}_nN_x} \tag{1} $$

$$ {}_nq_x = \frac{n \cdot_n M_x}{1 + n \cdot (1 -_n a_x) \cdot_n M_x} \tag{2} $$

$$ {}_\infty q_x = 1.00 \tag{3} $$

$$ {}_np_x = 1 -_n q_x \tag{4} $$

$$ l_{x+n} =_n p_x \cdot l_x \tag{5} $$

$$ {}_nd_x =_n q_x \cdot l_x \tag{6} $$

$$ {}_nd_x = l_x - l_{x+n} \tag{7} $$

$$ {}_nL_x = n \cdot (l_x -_n d_x) + n \cdot_n a_x \cdot_n d_x \tag{8} $$

$$ {}_\infty L_x = \frac{l_x}{{}_\infty M_x} \tag{9} $$

$$ T_x = \sum_{a=x}^{\infty} {}_nL_a \tag{10} $$

$$ e_x = \frac{T_x}{l_x} \tag{11} $$

Exercise 1: life table revision question (Compulsory)

Calculate the following values and comment on your results. While you may choose to reconstruct the entire life table starting from l_x, this is not required to solve questions a, b, c, d, and e.

a. ${}_1p_0$, ${}_4p_1$, ${}_5p_{10}$, ${}_{10}p_{75}$.
Hint: All these quantities depend on l_x.

a.1
${}_1p_0 = \frac{l_1}{l_0}$

```r
lx[x == 1,]/lx[x == 0,]
```

```
##    male  female
## 0.97731 0.98179
```

a.2

$_4p_1 = \frac{l_5}{l_1}$

```r
lx[x == 5,]/lx[x == 1,]
```

```
##      male    female
## 0.9945974 0.9969953
```

a.3

$_5p_{10} = \frac{l_{15}}{l_{10}}$

```r
lx[x == 15,]/lx[x == 10,]
```

```
##      male    female
## 0.9979159 0.9990384
```

a.4

$_{10}p_{75} = \frac{l_{85}}{l_{75}}$

```r
lx[x == 85,]/lx[x == 75,]
```

```
##      male    female
## 0.3474163 0.4256781
```

b. $_1d_0$, $_4d_1$, $_{15}d_{50}$.
Hint: Decumulate l_x to calculate the number of deaths in a life table.

b.1

$_1d_0 = l_0 - l_1$

```r
lx[x == 0,] - lx[x == 1,]
```

```
##   male female
##   2269   1821
```

b.2

$_4d_1 = l_1 - l_5$

```r
lx[x == 1,] - lx[x == 5,]
```

```
##   male female
##    528    295
```

b.3

$_{15}d_{50} = l_{50} - l_{65}$

```r
lx[x == 50,] - lx[x == 65,]
```

```
##   male female
##  14400   9177
```

c. $_4q_1$, $_5q_5$, $_{15}q_{50}$.
Hint: These quantities can be calculated directly from l_x.

c.1

$_4q_1 = \frac{l_1 - l_5}{l_1}$

```
(lx[x == 1,] - lx[x == 5,])/lx[x == 1,]
```

```
##        male      female
## 0.005402585 0.003004716
```

c.2
$_5q_5 = \frac{l_5 - l_{10}}{l_5}$

```
(lx[x == 5,] - lx[x == 10,])/lx[x == 5,]
```

```
##        male      female
## 0.002859994 0.001348535
```

c.3
$_{15}q_{50} = \frac{l_{50} - l_{65}}{l_{50}}$

```
(lx[x == 50,] - lx[x == 65,])/lx[x == 50,]
```

```
##       male     female
## 0.15672105 0.09693775
```

d. $_1L_0$, $_4L_1$, $_5L_5$, $_5L_{45}$.
Hint: Use the general formula for $_nL_x$ in closed age intervals.

d.1
$_1L_0 = 1 \cdot l_1 + 1 \cdot_1 a_0 \cdot (l_0 - l_1)$

```
n[x == 0]*lx[x == 1,] + n[x == 0]*nax[x == 0]*(lx[x == 0,] - lx[x == 1,])
```

```
##    male  female
## 97957.9 98361.1
```

d.2
$_4L_1 = 4 \cdot l_5 + 4 \cdot_4 a_1 \cdot (l_1 - l_5)$

```
n[x == 1]*lx[x == 5,] + n[x == 1]*nax[x == 1]*(lx[x == 1,] - lx[x == 5,])
```

```
##     male   female
## 389656.8 392008.0
```

d.3
$_5L_5 = 5 \cdot l_{10} + 5 \cdot_5 a_5 \cdot (l_5 - l_{10})$

```
n[x == 5]*lx[x == 10,] + n[x == 5]*nax[x == 5]*(lx[x == 5,] - lx[x == 10,])
```

```
##   male female
## 485320 489090
```

d.4
$_5L_{45} = 5 \cdot l_{50} + 5 \cdot_5 a_{45} \cdot (l_{45} - l_{50})$

```
n[x == 45]*lx[x == 50,] + n[x == 45]*nax[x == 45]*(lx[x == 45,] - lx[x == 50,])
```

```
##     male   female
## 463297.5 475975.0
```

e. $_1M_0$, $_4M_1$, $_{10}M_{60}$.
Hint: All these quantities depend on $_nd_x$ and $_nL_x$.

e.1
$_1M_0 = \frac{_1d_0}{_1L_0} = \frac{l_0 - l_1}{1 \cdot l_1 + 1 \cdot_1 a_0 \cdot (l_0 - l_1)}$

```r
(lx[x == 0,] - lx[x == 1,])/(n[x == 0]*lx[x == 1,] + n[x == 0]*nax[x == 0]*(lx[x == 0,] - lx[x == 1,]))
```

```
##       male     female
## 0.02316301 0.01851342
```

e.2

$_4M_1 = \frac{_4d_1}{_4L_1} = \frac{l_1 - l_5}{4 \cdot l_5 + 4 \cdot {_4a_1} \cdot (l_1 - l_5)}$

```r
(lx[x == 1,] - lx[x == 5,])/(n[x == 1]*lx[x == 5,] + n[x == 1]*nax[x == 1]*(lx[x == 1,] - lx[x == 5,]))
```

```
##         male       female
## 0.0013550386 0.0007525357
```

e.3

$_{10}M_{60} = \frac{_5d_{65} + _5d_{60}}{_5L_{65} + _5L_{60}} = \frac{(l_{60} - l_{65}) + (l_{65} - l_{70})}{(5 \cdot l_{65} + 5 \cdot {_5a_{60}} \cdot (l_{60} - l_{65})) + (5 \cdot l_{70} + 5 \cdot {_5a_{65}} \cdot (l_{65} - l_{70}))}$

```r
(lx[x==60,] - lx[x==70,])/(n[x==60]*lx[x==65,] + n[x==60]*nax[x==60]*(lx[x==60,] - lx[x==65,])
  + n[x==65]*lx[x==70,] + n[x==65]*nax[x==65]*(lx[x==65,] - lx[x==70,]))
```

```
##       male     female
## 0.02327251 0.01480073
```

f. e_0 and e_1.

Hint: Start by calculating T_{10}, then find T_{85} using $_{75}L_{10}$. Once $_nL_x$ has been estimated for all ages, including the open-ended age interval, proceed to compute T_x and e_x following standard life table equations.

Life table analysis

```r
ndx                     = rbind(-diff(lx,1),lx[!sEL,])
nqx                     = ndx/lx
nLx                     = n*(lx - ndx) + n*nax*ndx
nLx[!sEL,]              = e10*lx[x == 10,] - colSums(nLx[sEL & x >= 10,])
nMx                     = ndx/nLx
Tx                      = matrix(1, length(x), 1)%*%colSums(nLx) - (apply(nLx, 2, cumsum) - nLx)
Tx                      = apply(apply(apply(nLx, 2, rev), 2, cumsum), 2, rev)
ex                      = Tx/lx
```

R function "apply(input, dimension, function)" repeats the same function for each row (dimension = 1) or each column (dimension = 2). Matrix or vector multiplication should use the notation "%*%".

f.1

e_0

```r
ex[x == 0,]
```

```
##     male   female
## 71.64988 74.68726
```

f.2

e_1

```r
ex[x == 1,]
```

```
##     male   female
## 72.31104 75.07068
```

g. Ask Copilot to estimate the number of years that a newborn female is expecting to live between exact ages 20 and 65.

h. Ask Copilot to estimate the number of years that a 20-year female is expecting to live between exact ages 20 and 65.

Exercise 2: application of the stationary population model (Compulsory)

A job training program has 150 training positions that are always filled. The program admits 30 new candidates each month. Every month, 10 trainees quit the program while 20 more find a placement. These conditions have prevailed as long as anyone can remember.

a. What is the monthly rate of leaving the training program?
$CDR = CBR = \frac{l_x}{T_x} = \frac{30}{150} = 200$ exits per 1,000

b. What is the probability that a trainee will leave the program by finding a placement?
$\frac{placements}{all\, exists} = \frac{20}{30} = 0.6666667$

c. How long on average does a trainee stay in the program?
$e_0 = \frac{1}{CDR} = \frac{1}{0.2} = 5$ months

d. Ask Copilot to solve and explain this exercise by including the problem and questions a, b, and c in the prompt.

Exercise 3: additional life table questions based on the data of Exercise 1 (Optional)

Hint: Bear in mind that both life tables use the same radix of 100,000 individuals, which does not reflect the slight excess of male births. To aggregate the two populations into a single life table, or to compare quantities that depend on population size—such as the number of survivors or person-years lived—it would be necessary to adjust for the Sex Ratio at Birth. You can assume 105 male births per 100 female births.
a. Calculate the Infant Mortality Rate.

```
SRB                     = 1.05
sum(ndx[x == 0,]*c(1.05,1))/sum(lx[x == 0,]*c(SRB,1))*1000
```

```
## [1] 20.50463
```

b. If you were told that there had been 35,000 males aged 5-9 in Athens in mid-1981, would you be able to say how many there would be aged 10-14 exactly 5 years later? What assumptions would you have to make?
Hint: The command "nLx[x == 5,"male"]" returns the function ${}_nL_x$ at x equal to 5 for the male population.
${}_5N_{10} = \frac{{}_5L_{10}}{{}_5L_5} \cdot {}_5N_5 = \frac{{}_5L_{10}}{{}_5L_5} \cdot 35000$

```
nLx[x == 10,"male"]/nLx[x == 5,"male"]*35000
```

```
##     male
## 34913.46
```

Assuming a closed population and keeping constant the mortality rates of 1981.

c. What proportion of male population is aged 5-9 in the stationary population represented by the life table in Table 1?
$\frac{{}_5L_5}{T_0}$

```
nLx[x == 5,"male"]/Tx[x == 0,"male"]
```

```
##       male
## 0.06773493
```

d. What would be the sex ratio in the 25-29 age group in the stationary populations represented by the life tables in Table 1?

Hint: you may need to make an assumption about the Sex Ratio at Birth.
$\frac{_5L_{25,male}}{_5L_{25,female}} \cdot SRB$

```r
sprintf("%.5f",nLx[x == 25,"male"]/nLx[x == 25,"female"]*SRB)
```

```
## [1] "1.03357"
```

e. What would be the Crude Birth Rate in the stationary populations represented by the life table of Athens in 1981? and the Crude Death Rate? $CDR = CBR = \frac{l_0}{T_0}$

```r
sum(lx[x == 0,]*c(SRB,1))/sum(Tx[x == 0,]*c(SRB,1))*1000
```

```
## [1] 13.67399
```

f. The two questions below pertain to girl twins were born to a woman on her 20th birthday. Her husband was exactly 5 years older than herself. You can assume that the mortality regimes of Athens in 1981 apply to each member of the family. 1. What is the probability that both mother and children are alive when the twins celebrate their 10^{th} anniversary, but that the father had died?

$P(event) = P(father\,died) \cdot P(mother\,survived) \cdot P(one\,twin\,survived) \cdot P(other\,twin\,survived)$
$P(event) = [1 - \frac{l_{35,male}}{l_{25,male}}] \cdot [\frac{l_{30,female}}{l_{20,female}}] \cdot [\frac{l_{10,female}}{l_{0,female}}]^2$

```r
sprintf("%.7f",(1 - lx[x == 35,"male"]/lx[x == 25,"male"])
        *(lx[x == 30,"female"]/lx[x == 20,"female"])
        *(lx[x == 10,"female"]/lx[x == 0,"female"])^2)
```

```
## [1] "0.0092844"
```

2. What is the probability that at least one child survived but only one parent is alive 10 years after the birth of the twins?
Hint: in order to obtain the probability that two events jointly occur, we have to multiply probabilities (AND = multiplication, assuming event independence); in other to obtain the probability that either one of two events occur we add the two probabilities (OR = addition).

$P(at\,least\,one\,twin\,survived) = 1 - P(both\,twins\,died)$

```r
P1 = 1 - (1 - lx[x == 10,"female"]/lx[x == 0,"female"])^2
sprintf("%.10f",P1)
```

```
## [1] "0.9994946496"
```

$P(only\,one\,parent\,survived) = 1 - P(both\,parents\,survived) - P(both\,parents\,died)$

```r
P2 = 1 - (lx[x == 30,"female"]/lx[x == 20,"female"])*(lx[x == 35,"male"]/lx[x == 25,"male"]) -
  (1 - lx[x == 30,"female"]/lx[x == 20,"female"])*(1 - lx[x == 35,"male"]/lx[x == 25,"male"])
sprintf("%.10f",P2)
```

```
## [1] "0.0142251761"
```

$P(event) = P(at\,least\,one\,twin\,survived) * P(only\,one\,parent\,survived)$

```r
sprintf("%.10f",P1*P2)
```

```
## [1] "0.0142179874"
```

Exercise 4: life tables and stationary population theory (Optional)

```r
rm(list = ls())
x                     = c(0, 10, 25, 65, 85)
lx                    = c(1000, 950, 750, 600, 300)
ex                    = c(60.0, NA, 50.0, 17.5, 5.0)
```

```r
LT                      = data.frame(x,lx,ex)
print(LT)
```

```
##    x   lx   ex
## 1  0 1000 60.0
## 2 10  950   NA
## 3 25  750 50.0
## 4 65  600 17.5
## 5 85  300  5.0
```

Given the stationary population described in the table and the fact that 16.25% of the population is between exact ages 0 and 10, answer the questions:

Life table analysis

As is standard in life table calculations, we can define the length of age intervals based on the values of x, and identify the open-ended age interval.

```r
n                       = c(diff(x),NA)
sEL                     = !is.na(n)
```

The number of deaths is calculated as the first difference of the number of survivors. An incomplete set of T_x is calculated as the product of the life expectancy and the number of survivors. An incomplete set of ${}_nL_x$ is calculated as the first difference of T_x.

```r
ndx                     = c(-diff(lx,1),lx[!sEL])
Tx                      = ex*lx
nLx                     = c(-diff(Tx,1),Tx[!sEL])
```

${}_{10}L_0$ accounts for 16.25% of T_0, as incorporated by the following line.

```r
nLx[x == 0]             = Tx[x == 0]*16.25/100
```

${}_{15}L_{10}$ is still missing. The variable s is created to identify the missing value. The missing value should be equal to the difference between T_0 and the sum of all ${}_nL_x$.

```r
s                       = is.na(nLx)
nLx[s]                  = Tx[x == 0] - sum(nLx[!s])
```

Once the function ${}_nL_x$ has been identified for all ages, T_x and e_x are recalculated to populate their missing values. Then, ${}_nM_x$ is calculated as the ratio of the observed number of events ${}_nd_x$, to the exposure to the risk of dying, measured by the number of person-years lived within the age interval ${}_nL_x$.

```r
Tx                      = rev(cumsum(rev(nLx)))
ex                      = Tx/lx
nMx                     = ndx/nLx

LT                      = data.frame(x,n,nMx,lx,ndx,nLx,Tx,ex)
print(LT)
```

```
##    x  n         nMx   lx ndx   nLx    Tx       ex
## 1  0 10 0.005128205 1000  50  9750 60000 60.00000
## 2 10 15 0.015686275  950 200 12750 50250 52.89474
## 3 25 40 0.005555556  750 150 27000 37500 50.00000
## 4 65 20 0.033333333  600 300  9000 10500 17.50000
## 5 85 NA 0.200000000  300 300  1500  1500  5.00000
```

a. What is the death rate in the age interval [65, 85) ${}_{20}M_{65}$?

```r
nMx[x == 65]
```

```
## [1] 0.03333333
```

b. What is the value of e_{10}?

```r
ex[x == 10]
```

```
## [1] 52.89474
```

c. Ask Copilot to find the value of e_{10}, analysing the data as an expert demographer.
d. Ask Copilot to calculate the life expectancy at birth for a theoretical population with a constant force of mortality equal to 0.02.
e. Ask Copilot to calculate e_{10} for the same theoretical population.