SOC 756: Problem Set 1

Mateo Frumholtz

September 8, 2025

1. Table 1 (next page) contains deaths by age for French males in 1985. These data also include mid-year population estimates and a set of nax values for French males for 1985. Table 1 is on our webpage in .csv format.

a. Use these data to construct a life table for the male population. Do this by performing operations on the vectors. You will need to calculate and fill in the following life table columns: nqx; lx; ndx; nLx; nmx; Tx; and ex.

```
# nNx = midyear population
# nDx = deaths between ages x and x + n
# nmx ~ nDx/nNx

ps1 |>
  arrange(x) |>
  mutate(
    time_diff = lead(x) - x,
    nmx = nDx / nNx,
    nqx = (time_diff * nmx) / (1 + ((time_diff - nax) * nmx)),
    nqx = case_when(
      is.na(nqx) ~ 1,
      TRUE ~ nqx
    ),
    npx = 1 - nqx,
    lx = accumulate(npx, `*`, .init = 100000)[-1],
    lx = lag(lx, default = 100000),
    ndx1 = lx - lead(lx),
    # Other option for ndx
    # ndx1 = case_when(
    #     is.na(ndx1) ~ lx,
    #     TRUE ~ ndx1
    # ),
```

```
  ndx = nqx * lx,
  nLx = (time_diff * lead(lx)) + (nax * ndx),
  nLx = case_when(
    is.na(nLx) ~ lx / nmx,
    TRUE ~ nLx
  ),
  Tx = rev(cumsum(rev(nLx))),
  ex = Tx / lx
) |>
select(x, nqx, lx, ndx, nLx, nmx, Tx, ex) |>
gt() |>
fmt_number(
  columns = c(lx, ndx, nLx, Tx, ex),
  decimals = 1
) |>
fmt_number(
  columns = c(nqx, nmx),
  decimals = 5
) |>
tab_header(
  title = "Life Table for French Males in 1985"
)
```

Life Table for French Males in 1985

x	nqx	lx	ndx	nLx	nmx	Tx	ex
0	0.00976	100,000.0	975.7	99,109.1	0.00985	7,131,028.1	71.3
1	0.00197	99,024.3	195.3	395,608.8	0.00049	7,031,918.9	71.0
5	0.00140	98,829.0	138.5	493,798.4	0.00028	6,636,310.1	67.1
10	0.00156	98,690.4	153.6	493,139.6	0.00031	6,142,511.7	62.2
15	0.00489	98,536.8	482.2	491,608.2	0.00098	5,649,372.1	57.3
20	0.00792	98,054.6	776.2	488,390.1	0.00159	5,157,763.9	52.6
25	0.00778	97,278.5	756.9	484,509.2	0.00156	4,669,373.8	48.0
30	0.00848	96,521.6	818.3	480,632.5	0.00170	4,184,864.6	43.4
35	0.01138	95,703.3	1,089.4	475,963.9	0.00229	3,704,232.1	38.7
40	0.01723	94,613.9	1,630.1	469,315.3	0.00347	3,228,268.2	34.1
45	0.02849	92,983.8	2,649.5	458,811.8	0.00577	2,758,952.9	29.7
50	0.04566	90,334.3	4,125.0	442,031.5	0.00933	2,300,141.1	25.5
55	0.06824	86,209.3	5,883.2	417,074.2	0.01411	1,858,109.6	21.6
60	0.09473	80,326.2	7,609.4	383,375.8	0.01985	1,441,035.4	17.9
65	0.13127	72,716.7	9,545.3	340,818.3	0.02801	1,057,659.6	14.5

70	0.20498	63,171.5	12,948.6	284,754.8	0.04547	716,841.3	11.3
75	0.31172	50,222.8	15,655.5	212,570.3	0.07365	432,086.5	8.6
80	0.45848	34,567.3	15,848.4	132,676.7	0.11945	219,516.2	6.4
85	1.00000	18,718.9	18,718.9	86,839.4	0.21556	86,839.4	4.6

b. Graph the following life table functions using either plot() or ggplot(): lx; ndx; and nmx. What do you observe?

c. What was life expectancy at age 40? How would you interpret this number?

d. What was the probability of surviving from birth to age 30?

e. What was the probability of surviving to age 65 for those who survived to age 30?

f. What was the probability that a newborn would die between 50 and 55?

g. How many years could a newborn expect to live in the interval 15-65?

h. If you only had the fourth column of Table 1, would you be able to distinguish this population as one with high mortality or low mortality? (What nax value in particular might help distinguish between the two?)

i. If the French population were stationary, what would be the crude death rate?

The CDR would be...

j. Extra credit part 1: push your code to your Github page and list the URL in your submitted answers.

The github link for this class can be found here.

~~**k. Extra credit part 2: install the Lifetables package in R. With nmx in hand, use lt.mx() to populate the other functions. Check your work in 1(a), noting discrepancies if you set nax=NULL.**~~

2. Think about the social phenomena / processes that most interest you. Might any of these processes be measured in the form of a lifetable? If yes:

a. What events would constitute "births" and "deaths"?

b. What could you learn from using a lifetable?

c. Where might you start looking for data to identify the size of the population at risk and the age-specific "death" rates or probabilities?

d. What issues might limit how the information produced in your lifetable can be interpreted?

~~If no, describe the issues that would make the lifetable an inappropriate analytical tool for the social processes that you study.~~