Precept 2: Life Tables

Kristin E. Bietsch

Office of Population Research, Princeton University

1 Equations

Note: The information in this section comes from Preston chapter 3

1.1 Period Life Table Construction

- Observed data:
  1. ${}_nN_x$ = mid-year population in age interval x to $x+n$
  2. ${}_nD_x$ = deaths between ages x and $x+n$ during the year
- Steps for period life table construction:
  1. ${}_nm_x \simeq {}_nM_x = \frac{{}_nD_x}{{}_nN_x}$
  2. ${}_na_x$: Refer to Section 1.3
  3. ${}_nq_x = \frac{n \cdot {}_nm_x}{1+(n-{}_na_x)\cdot {}_nm_x}$
     - Remember: ${}_\infty q_x = 1$
  4. ${}_np_x = 1 - {}_nq_x$
  5. l_0=100,000
  6. $l_{x+n} = l_x \cdot {}_np_x$
  7. ${}_nd_x = l_x - l_{x+n}$
  8. ${}_nL_x = n \cdot l_{x+n} + {}_na_x \cdot {}_nd_x$
     - Open-ended interval: ${}_\infty L_x = \frac{l_x}{{}_\infty m_x}$
  9. $T_x = \sum_{a=x}^{\infty} {}_nL_a$
  10. $e_x^0 = \frac{T_x}{l_x}$

1.2 Interpretating the Life Table

1.2.1 The basic columns

x	Exact age x
l_x	Number of people left alive at age x
${}_nd_x$	Number of people dying between ages x and $x+n$
${}_nq_x$	Probability of dying between ages x and $x+n$
${}_np_x$	Probability of surviving between ages x and $x+n$
${}_nL_x$	Person-years lived between ages x and $x+n$
T_x	Person-years lived above age x
e_x^0	Expectation of life at age x
${}_nm_x$	Age-specific death rate between ages x and $x+n$
${}_na_x$	Average person-years lived between ages x and $x+n$ for persons dying in the interval

1.2.2 Additional information

$\frac{l_y}{l_x} = {}_{y-x}p_x$	Probability of surviving from age x to age y
$1 - \frac{l_y}{l_x} = {}_{y-x}q_x$	Probability of dying from ages x and y
$l_x - l_y = {}_{y-x}d_x$	Number of people dying between ages x and y
$T_x - T_y = {}_{y-x}L_x$	Number of person-years lived between ages x and y
$\frac{{}_nd_x}{l_0}$	Probability that a newborn will die between ages x and y
$\frac{l_x - l_y}{l_0}$	Probability that a newborn will experience his death between ages x and y
$\frac{T_x - T_y}{l_0}$	Number of years that a newborn can expect to live between ages x and y

1.3 Methods for Constructing ${}_na_x$

1. Direct observation
2. Borrow from a similar population
3. **Assume even distribution**
4. **Keyfitz gradation**
5. **Assume constant hazard**

Note: Special assumptions are needed for ages < 5 (Table 3.3 in Preston)

	Males	Females
Values of ${}_1a_0$		
If ${}_1m_0 \geq .107$	.330	.350
If ${}_1m_0 < .107$	$.045 + 2.684 \cdot {}_1m_0$	$.053 + 2.800 \cdot {}_1m_0$
Values of ${}_4a_1$		
If ${}_1m_0 \geq .107$	1.352	1.361
If ${}_1m_0 < .107$	$1.651 - 2.816 \cdot {}_1m_0$	$1.522 - 1.518 \cdot {}_1m_0$

For ${}_na_x$ between 5 and the open ended inerval:

Assume even distribution	${}_na_x = \frac{n}{2}$
Keyfitz gradation	${}_na_x = \frac{-\frac{n}{24}{}_nd_{x-n} + \frac{n}{2}{}_nd_x + \frac{n}{24}{}_nd_{x+n}}{{}_nd_x}$
Assume constant hazard	${}_na_x = n + \frac{1}{{}_nm_x} - \frac{n}{1-e^{-n_nm_x}}$ ${}_np_x = 1 - {}_nq_x = e^{-n_nm_x}$

2 Stata

Note: The information in this section comes from Stata help files and http://data.princeton.edu/stata/programming.html#s41

2.1 in

Description: *in* at the end of a command means that the command is to use only the observations specified. *in* is allowed with most Stata commands.

Syntax: command in range

#
#/#
#/l
f/#

f stands for the first observation, and l for the last.

Negative numbers may be used to specify distance from the end of the data.

Some examples:

list price in 10	displays the 10th observation
list prince in 10/20	displays the 10th through 20th observation
list price in 20/l	displays the 20th through last observation
list prince in 1/10	displays the 1st through 10th observation
list price in f/10	displays the 1st through 10th observation
list price in -10/l	displays the last 10 observations

2.2 [_n+1]

Refers to the neighboring observations:

```
gen dprice = price - price[_n-1]
```

In the above line, the difference in price is constructed by subtracting the price from the line above from the price of the current line. You can also use [_n+1] which refers to the line below the current observation (or any number!). For example, here is a line of code from a current project of mine:

```
replace mar_interval=mar_interval[_n-5] in 6/1050
```

2.3 Local

- Local macros have names of up to 31 characters and are known only in the current context (the console, a do file, or a program).
- You *define* a local macro using

  ```
  local name [=] text
  ```

  and you *evaluate* it using

  ```
  `name'
  ```

  (Note the use of a backtick or left quote.)

Example:

```
local controls age agesq education income
regress outcome treatment `controls'
```

2.4 Loops

Loops are used to do repetitive tasks. Stata has commands that allow looping over sequences of numbers and various types of lists, including lists of variables.

Today, we will be using the *while* loop.

```
while condition {
        ... do something ...
}
```

- Where condition is an expression
- The loop executes as long as the condition is true (nonzero)
- Usually something happens inside the loop to make the condition false, otherwise the code would run forever

A typical use of *while* is in **iterative estimation procedures**, where you may loop while the difference in successive estimates exceeds a predefined tolerance. Usually an iteration count is used to detect lack of convergence.

2.5 Additionall Stata Commands

drop
replace
egen
local
log()
total()
abs()
format —%8.4f
list