Working with Demographic Life Table Data in Stata

Daniel C. Schneider
Max Planck Institute for Demographic Research

Stata Conference, July 27, 2017, Baltimore
Contents

Demographic Life Tables

The Human Mortality Database

hmddata
  text to .dta conversion
  data usage and examples
    mortality rates

lifetable
The Life Table

- displays death-related statistics of a cohort/population
- columns: age and age-related functions pertaining to mortality
- cohort life table vs. period life table: "synthetic cohort"
- based on triangles from a Lexis diagram
- calculation of life expectancy
- related: `ltab`le of official Stata
Life Table of the US, 2014

Demographic Life Table Data in Stata    D.C.Schneider   Stata Conference 2017

```
.hmddata use lifetables bothsexes, clear grid(5x1) popfilter(usa)
.list age mx-ex if year==2014, noobs sep(0)
```

<table>
<thead>
<tr>
<th>age</th>
<th>mx</th>
<th>qx</th>
<th>ax</th>
<th>lx</th>
<th>dx</th>
<th>Lx</th>
<th>Tx</th>
<th>ex</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.0060</td>
<td>0.0059</td>
<td>0.06</td>
<td>100,000</td>
<td>592</td>
<td>99,447</td>
<td>7,897,283</td>
<td>79.0</td>
</tr>
<tr>
<td>1</td>
<td>0.0002</td>
<td>0.0010</td>
<td>1.64</td>
<td>99,408</td>
<td>96</td>
<td>397,407</td>
<td>7,797,837</td>
<td>78.4</td>
</tr>
<tr>
<td>5</td>
<td>0.0001</td>
<td>0.0006</td>
<td>2.41</td>
<td>99,312</td>
<td>57</td>
<td>496,414</td>
<td>7,400,429</td>
<td>74.5</td>
</tr>
<tr>
<td>10</td>
<td>0.0001</td>
<td>0.0007</td>
<td>2.82</td>
<td>99,255</td>
<td>69</td>
<td>496,125</td>
<td>6,904,015</td>
<td>69.6</td>
</tr>
<tr>
<td>15</td>
<td>0.0004</td>
<td>0.0023</td>
<td>2.98</td>
<td>99,186</td>
<td>224</td>
<td>495,476</td>
<td>6,407,890</td>
<td>64.6</td>
</tr>
<tr>
<td>20</td>
<td>0.0008</td>
<td>0.0042</td>
<td>2.60</td>
<td>98,962</td>
<td>415</td>
<td>493,810</td>
<td>5,912,414</td>
<td>59.7</td>
</tr>
<tr>
<td>65</td>
<td>0.0147</td>
<td>0.0710</td>
<td>2.62</td>
<td>84,222</td>
<td>5,983</td>
<td>406,867</td>
<td>1,644,162</td>
<td>19.5</td>
</tr>
<tr>
<td>70</td>
<td>0.0228</td>
<td>0.1080</td>
<td>2.62</td>
<td>78,239</td>
<td>8,450</td>
<td>371,069</td>
<td>1,237,295</td>
<td>15.8</td>
</tr>
<tr>
<td>75</td>
<td>0.0361</td>
<td>0.1662</td>
<td>2.62</td>
<td>69,789</td>
<td>11,596</td>
<td>321,393</td>
<td>866,226</td>
<td>12.4</td>
</tr>
<tr>
<td>80</td>
<td>0.0600</td>
<td>0.2621</td>
<td>2.59</td>
<td>58,193</td>
<td>15,252</td>
<td>254,268</td>
<td>544,833</td>
<td>9.4</td>
</tr>
<tr>
<td>85</td>
<td>0.1023</td>
<td>0.4080</td>
<td>2.52</td>
<td>42,941</td>
<td>17,522</td>
<td>171,309</td>
<td>290,565</td>
<td>6.8</td>
</tr>
<tr>
<td>90</td>
<td>0.1785</td>
<td>0.6061</td>
<td>2.35</td>
<td>25,420</td>
<td>15,408</td>
<td>86,325</td>
<td>119,256</td>
<td>4.7</td>
</tr>
<tr>
<td>95</td>
<td>0.2801</td>
<td>0.7737</td>
<td>2.11</td>
<td>10,011</td>
<td>7,746</td>
<td>27,655</td>
<td>32,931</td>
<td>3.3</td>
</tr>
<tr>
<td>100</td>
<td>0.4170</td>
<td>0.8948</td>
<td>1.81</td>
<td>2,266</td>
<td>2,027</td>
<td>4,861</td>
<td>5,276</td>
<td>2.3</td>
</tr>
<tr>
<td>105</td>
<td>0.5695</td>
<td>0.9554</td>
<td>1.52</td>
<td>238</td>
<td>228</td>
<td>400</td>
<td>415</td>
<td>1.7</td>
</tr>
<tr>
<td>110</td>
<td>0.6923</td>
<td>1.0000</td>
<td>1.44</td>
<td>11</td>
<td>11</td>
<td>15</td>
<td>15</td>
<td>1.4</td>
</tr>
</tbody>
</table>
Life Expectancy

• life expectancy (LE, $e_x$):
  • $e_x$: average years ahead of a population member aged X
  • $e_0$ (LE at birth):
    • average years lived
      = mean age at death

• $e_0$ of period life table: average years lived under current (period) mortality conditions

• All statements are made with respect to members of a hypothetical cohort.
The Human Mortality Database (HMD)

- compiled by: UC Berkeley, Max Planck Institute for Demographic Research
- high-quality data
- variables: see next slide
- geographic coverage:
  - currently 39 countries / populations
  - many European countries, plus: US, Canada; Japan, Taiwan; Australia; Chile; Israel; Russia
- time coverage: Sweden 1750-, France 1816-, 10 other countries start before 1900
- www.mortality.org
- companion / similar databases: HLD, HFD, HFC
HMD: Data Contents

.`hmddata info concepts`

HMD full concepts:

- births
- deaths
- deathsbylexistriangles
- populationsize
- exposuretorisk
- deathrates
- lifetables
- lifeexpectancyatbirth

- period|raw
- period|raw
- period
- period|raw
- period|cohort
- period|cohort
- period|cohort
- period|cohort
HMD: Data Acquisition

- consent to user agreement and registration required, but free of charge
- Data are distributed in text files.
- zipped text files
  (http://www.mortality.org/cgi-bin/hmd/hmd_download.php)
  - available
    - by statistic / concept
    - by country
    - all data
  - **hmddata** can process any one and one or more of the zipped text files.
development goal was a data exploration tool for researchers:

- easy data access
- quick generation of working-quality tables and graphs

net install hmddata,  
from(http://user.demogr.mpg.de/schneider/stata)
**Syntax**

Set and query `hmddata` user settings

\[
\text{hmddata settings} \ [ \text{parameter} ] , [ \text{value(valstring)} ]
\]

Convert source data text files to `hmddata` files

\[
\text{hmddata convert fullconcepts} , \text{sourcesdir(dirstring)} [ \text{grid(gridlist)} \text{replace} ]
\]

Load HMD data

\[
\text{hmddata use fullconcepts} , [ \text{popfilter(poplistspec)} \text{grid(gridspec)} \text{long clear} ]
\]
hmddata: data handling

Generate age and year interval variables

\texttt{hmddata intervals [intvalvars], [noorder]}

Filter data set according to a subset of populations

\texttt{hmddata popfilter poplist, [iso noerror}
\texttt{droplist dummy(varname)]}

Generate graphs based on hmd data sets

\texttt{hmddata graph plottype plotvars xvar [if] [in], [at1(atspec) at2(atspec) by(varlist[, byopts]) plotopts(cline_options) twoway_options]}

Original graph from paper: Female life expectancy (LE) for selected countries and trend in record LE.
Graph replication using `hmddata`:

```
. hmddata use lifeexp, clear
. hmddata graph line female year
   if inrange(year, 1840, 2000), all(popname chile japan
   newzealand non nor usa) [...]
```
Original graph from paper: Male (blue) and female (red) LE in the record-holding country.
• replication of graph plus additional history
• More complicated graphs: use `graph twoway` instead of `hmddata graph`.

```stata
hmddata use lifeexp, clear
hmddata popfilter francecivil [...] [...] // generate LE rank variables
    // for males/females

twoway //
(scat fem year if rank_f==1, [...] || ///
(scat mal year if rank_m==1, [...] || ///
(lfit fem year if rank_f==1 [...] || ///
(lfit mal year if rank_m==1 [...] ))
```

---

Demographic Life Table Data in Stata  D.C.Schneider  Stata Conference 2017
Mortality Declines 1900-1949, 1950-2000

. hmddata use deathrates, grid (5x10) clear
. hmddata popfilter swe francetotal neth ita, dummy(d1)
. replace total = log10(total)
. hmddata graph line total age if d1, at1(year 1900 1950 2000) by(popname) [...]

Graphs by Country / Population name

France: Total population

Italy

Netherlands

Sweden
Mortality Declines 1900-1949, 1950-2000

Demographic Life Table Data in Stata    D.C.Schneider    Stata Conference 2017
lifetable

- not yet released
- development goal: versatile tool to generate and manipulate demographic life tables

principles / features:
- operation on multiple yet selected life tables at once
- standardized/prescribed variable names
- calculations using any valid minimum starting information

- to be added before release: CIs, methods for approximating $nax$, …
Thank you!

Questions? Comments?

contact: schneider@demogr.mpg.de
Human Mortality Database. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Available at www.mortality.org and www.humanmortality.de.