Università di Padova, Facoltà di Scienze Statistiche
Laurea Magistrale in Scienze Statistiche
Corso: **Teoria e modelli demografici**

Period life table

Maria Letizia Tanturri

Dipartimento di Scienze Statistiche

tanturri@stat.unipd.it

Lezione 10

Riferimenti

- Preston et al. (2001), ch. 3, p. 42-58

Università di Padova, Facoltà di Scienze Statistiche
Laurea Magistrale in Scienze Statistiche
Corso: **Teoria e modelli demografici**

From the cohort to the period lifetable

Lezione 10

From Cohort to Period Life Table

- Cohort life tables are easy to calculate, but they are of limited utility since they can only be computed once all members are all dead
- Moreover cohort data not very common
- Need ability to use "period" data that describe age-specific mortality in a given year or period

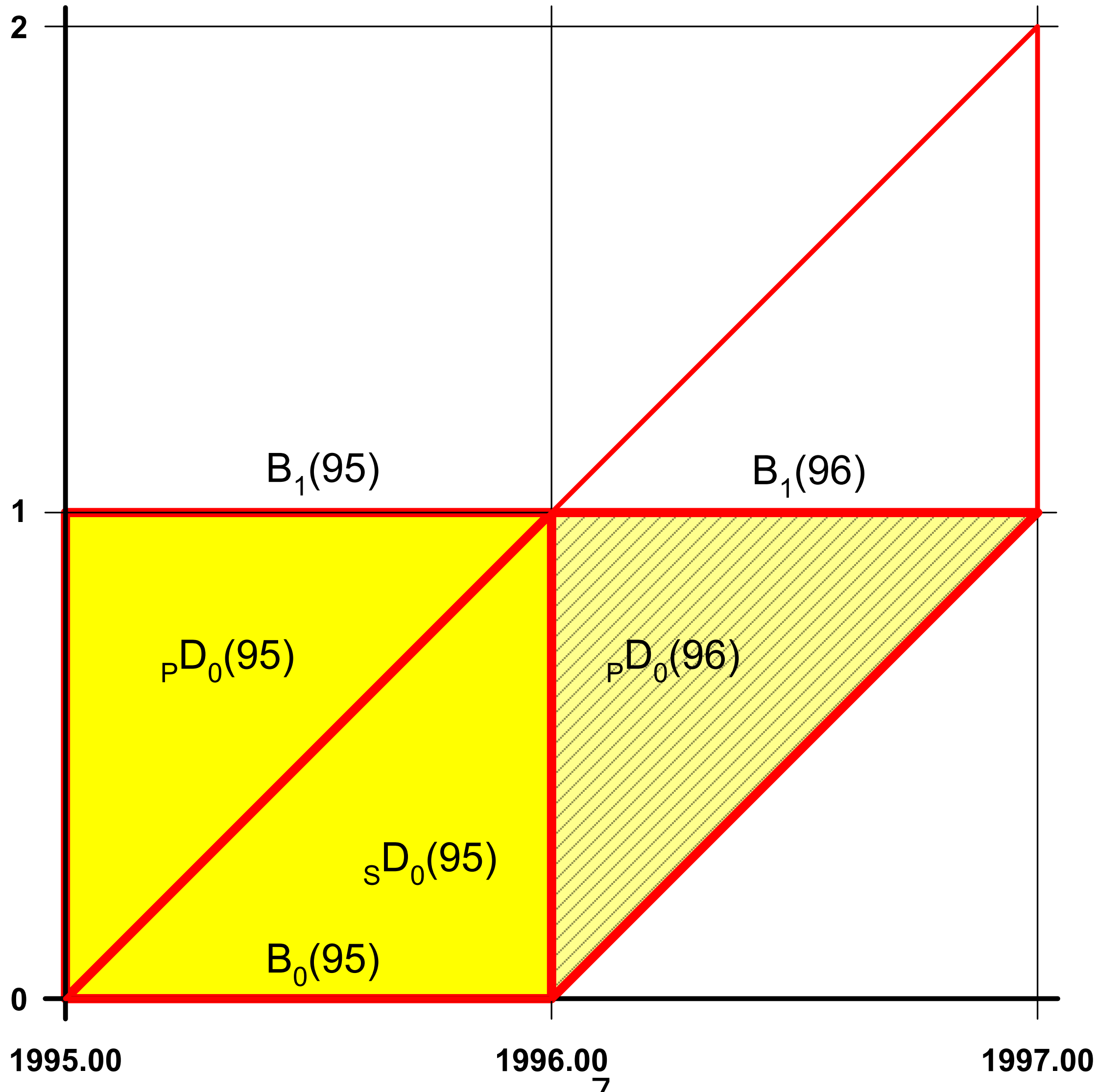
Period Life Table

- A “**period life table**” is exactly the same as a cohort life table except it describes the dying out of a “synthetic cohort” that experiences at each age the age-specific mortality associated with a given period
- A hypothetical group of people survives through the age-specific risk of dying associated with a period

Creating a Period Life Table

- The key to this is the fact that the hypothetical cohort experiences the age-specific probabilities of dying associated with the period
- The data available are usually observed age-specific mortality rates, ${}_nM_x$
- The trick then is to convert these observed age-specific mortality rates into one of the columns of a life table

Diagram of Synthetic $_nq_x$



Critical point

- Remember that to construct the life table we need first the summary information about how many people survive to different ages, or the equivalent, probability to survive from one of these ages to the next one.

- As for data, we usually have a tabulation of deaths by age in a given period

Creating a Period Life Table 2

- The most convenient choice is to convert ${}_nM_x$ to ${}_nq_x$
- ${}_nM_x$ to ${}_nq_x$ conversion →
- Critical assumption is that ${}_nM_x \sim {}_nm_x$
  - Mortality rates observed in the population during the period, nMx, apply to the synthetic cohort through its life course, on each age interval.

- The basic operation in calculating a period life table is to convert the mortality rates into underlying probabilities of death.
- Using life table notations, the two are related because:

$$ {}_nm_x = \frac{{}_nd_x}{{}_nL_x} = \frac{\text{Number of deaths in the cohort between ages } x \text{ and } x+n}{\text{Number of person-years lived in the cohort between ages } x \text{ and } x+n} $$

$$ {}_nq_x = \frac{{}_nd_x}{1_x} = \frac{\text{Number of deaths in the cohort between ages } x \text{ and } x+n}{\text{Number of survivors to age } x \text{ in the cohort}} $$

- The number of person-years lived between age x and x+n is n for every survivor to age x+n, but it is a fraction of n for people who survived to age x but didn’t make it to age x+n.
  - The fraction would be 0 if everybody dying between age x and x+n died exactly at age x, in general it will depend on the average age at death of people in the interval.

More formally:

$${}_nL_x = n \cdot l_{x+n} + {}_nA_x$$

${}_nL_x$: Number of person-years lived by the cohort between ages x and x+n

$n \cdot l_{x+n}$: Number of person-years lived in the interval by members of the cohort who survive the interval

${}_nA_x$: Number of person-years lived in the interval by members of the cohort who die in the interval

or:

$${}_nL_x = n \cdot l_{x+n} + {}_na_x \cdot {}_nd_x$$

${}_na_x$: Mean number of person-years lived in the interval by those dying in the interval

${}_nd_x$: Number of members of the cohort dying in the interval

Hence:

$${}_nL_x = n\,(l_x - {}_nd_x) + {}_na_x \cdot {}_nd_x$$

$$n \cdot l_x = {}_nL_x + n \cdot {}_nd_x - {}_na_x \cdot {}_nd_x$$

$$\boxed{l_x} = \frac{1}{n}\left[{}_nL_x + (n - {}_na_x) \cdot {}_nd_x \right]$$

- Now we substitute in q

$$ {}_nq_x = \frac{{}_nd_x}{l_x} = \frac{n \cdot {}_nd_x}{{}_nL_x + (n - {}_na_x)\ {}_nd_x} $$

And finally, we divide by L_x both numerator and denominator:

$$ {}_nq_x = \frac{n \cdot \frac{{}_nd_x}{{}_nL_x}}{\frac{{}_nL_x}{{}_nL_x} + (n - {}_na_x)\frac{{}_nd_x}{{}_nL_x}} = \frac{n \cdot {}_nm_x}{1 + (n - {}_na_x)\ {}_nm_x} $$

- If persons dying in the interval do so, on average half-way through the interval, then the equation

$$ {}_nq_x = \frac{n \cdot \frac{{}_nd_x}{{}_nL_x}}{\frac{{}_nL_x}{{}_nL_x} + (n - {}_na_x)\frac{{}_nd_x}{{}_nL_x}} = \frac{n \cdot {}_nm_x}{1 + (n - {}_na_x)\ {}_nm_x} $$

become:

$$ {}_nq_x = \frac{n \cdot {}_nm_x}{1 + \frac{n}{2}{}_nm_x} = \frac{2n \cdot {}_nm_x}{2 + n \cdot {}_nm_x} $$

Consequences:

- We can thus "extract" the life table probabilities from the rates as long as we can choose a set of $_na_x$. What are the options?

- $_nm_x \rightarrow {}_nq_x$ requires $_na_x$ … where do we get $_na_x$?

Strategies for choosing ${}_{n}a_{x}$

- 1) From calculating it directly
  - But the are: rare
  - They refer to population and not to cohort
  - They are not necessarily appropriate as they should depend only on the mortality function, not age distribution
    - Es. Extreme case<. If population aged 60 is 100 times as people aged 61-64, ${}_{5}a_{60}$ would be necessarily between ages 60 and 61 regardeless of mortality conditions

Strategies for choosing $_na_x$

- 2) From smoothing (graduating) the death distribution within each age interval
  - (a) the faster mortality increases in the interval, the more people will die toward the end of the interval (higher risk), but
  - (b) the higher mortality in the age interval the more people will die toward the beginning of the interval (fewer people at risk when age increases).
  - Gompertz model (death rates log-linearly related to age) or Keyfitz (distribution of deaths as second degree polinomial f)

e.g. age specific death rates

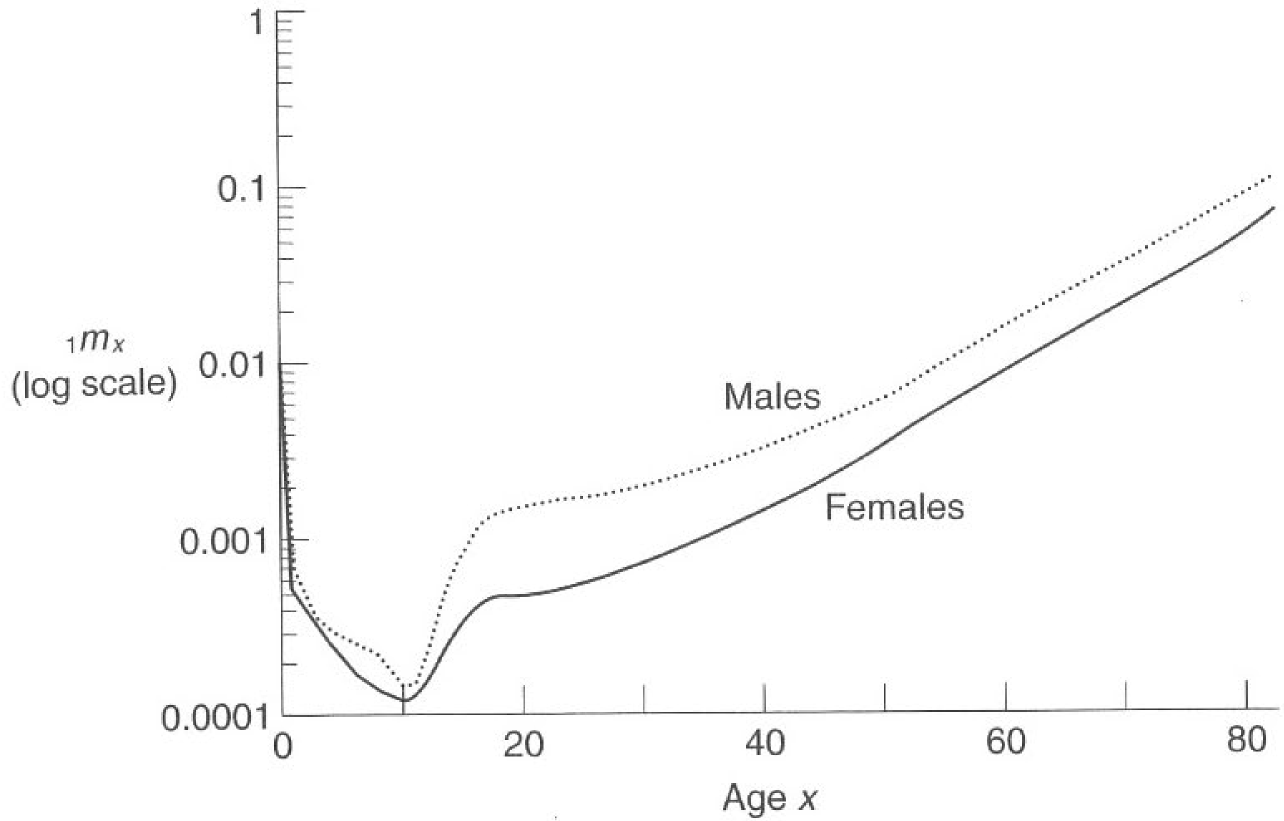


Figure 3.2 Age-specific death rates ($_1m_x$) by age, US, 1992, males and females
Data source: National Center for Health Statistics, 1996.

Strategies for choosing ${}_na_x$

- 3) Borrowing values from another population

Table 3.2: *Average person-years lived between ages x and x + n for persons dying in the interval* (${}_na_x$)

e_0^o	Sweden, 1900 Males	Sweden, 1900 Females	Sweden, 1985 Males	Sweden, 1985 Females	United States, 1985 Males	United States, 1985 Females	Guatemala, 1985 Males	Guatemala, 1985 Females
	51.528	54.257	73.789	79.830	71.266	78.422	60.582	64.415
Age x	Average person-years lived for people dying in the interval x to $x+n$							
0	0.358	0.375	0.083	0.081	0.090	0.086	0.165	0.150
1	1.235	1.270	1.500	1.500	1.500	1.500	1.500	1.500
5	2.500	2.500	2.500	2.500	2.500	2.500	2.500	2.500
10	2.456	2.469	3.006	2.773	3.014	2.757	2.469	2.390
15	2.639	2.565	2.749	2.617	2.734	2.644	2.711	2.665
20	2.549	2.536	2.569	2.578	2.564	2.552	2.628	2.601
25	2.481	2.514	2.561	2.665	2.527	2.588	2.573	2.563
30	2.505	2.509	2.600	2.649	2.571	2.632	2.593	2.627
35	2.544	2.521	2.638	2.625	2.622	2.678	2.545	2.566
40	2.563	2.522	2.695	2.662	2.666	2.706	2.541	2.543
45	2.572	2.561	2.705	2.722	2.688	2.702	2.604	2.592
50	2.574	2.578	2.706	2.694	2.684	2.683	2.596	2.627
55	2.602	2.609	2.687	2.670	2.657	2.671	2.623	2.661
60	2.602	2.633	2.673	2.689	2.626	2.650	2.635	2.623
65	2.591	2.628	2.643	2.697	2.608	2.642	2.616	2.676
70	2.561	2.585	2.607	2.706	2.571	2.631	2.557	2.607
75	2.500	2.517	2.547	2.650	2.519	2.614	2.486	2.532
80	2.415	2.465	2.471	2.607	2.460	2.596	2.409	2.447
85+	3.488	3.888	4.607	5.897	5.455	6.969	4.611	4.836

Source: Keyfitz and Flieger, 1968: 491; and 1990: 310, 348 and 528.

Strategies for choosing ${}_na_x$

- 4) Making one of two assumptions (thumb rules):
  - ${}_na_x$ is half the length of the age interval (n/2), or
  - ${}_nm_x$ is constant in the interval which negates the necessity of using ${}_na_x$ because there is a direct formula to calculate ${}_np_x$:

$$ {}_np_x = 1 - {}_nq_x = e^{-n \cdot {}_nm_x} $$

${}_na_x$ in Practice

- Usually ${}_na_x$ approximated by se n/2 for all age groups except the first
- Mortality rate between ages 0 and 5 changes very rapidly, falling very quickly at first and then flattening out
- Consequently most deaths early in life occur closer to 0 than to 5 and hence ${}_na_x$ is significantly less than n/2 in the first two age groups (0, 1-4)

${}_{n}a_{x}$ in Practice

- In general in other age groups where mortality is changing less rapidly, the overall life table is very insensitive to the exact choice of ${}_{n}a_{x}$

http://papp.iussp.org/sessions/papp101_s07/PAPP101_s07_090_010.html

${}_na_x$ for Very Young Ages

Coale –Demeny equation

	Males	Females
Value of ${}_1a_0$		
If ${}_1m_0 >= 0.107$	0.330	0.350
If ${}_1m_0 < 0.107$	$0.045+2.684({}_1m_0)$	$0.053+2.800({}_1m_0)$
Value of ${}_4a_1$		
If ${}_1m_0 >= 0.107$	1.352	1.361
If ${}_1m_0 < 0.107$	$1.651-2.816({}_1m_0)$	$1.522-1.518({}_1m_0)$

The Open-ended Age Interval

- Because n is effectively infinite for the open (last) age interval, we cannot calculate ${}_nL_x$ given the formulas we have

$${}_nm_x = \frac{{}_nd_x}{{}_nL_x}$$

$${}_\infty m_x = \frac{{}_\infty d_x}{{}_\infty L_x}$$

rearranging

$${}_\infty L_x = \frac{{}_\infty d_x}{{}_\infty m_x}$$

and for the open interval

$$l_x = {}_\infty d_x$$

so: $${}_\infty L_x = \frac{l_x}{{}_\infty m_x}$$

Università di Padova, Facoltà di Scienze Statistiche
Laurea Magistrale in Scienze Statistiche
Corso: **Demografia progredito**

Step by step

Lezione 10

Period Life Table Construction

- Step 1: Solving for age-specific death rates

$$ {}_n m_x \approx {}_n M_x = {}_n D_x \,/\, {}_n N_x $$

Period Life Table Construction

- Step 2: Decide on method for estimating ${}_{n}a_{x}$ (average person-years lived in the interval by those dying in the interval)
    - Direct observation
    - Graduation of the ${}_{n}m_{x}$ function
    - Borrowing values
    - Rules of thumb

Period Life Table Construction

- Step 3: Solve for ${}_nq_x$ (probability of dying in age interval)

$${}_nq_x = n * {}_nm_x \; / \; 1 + (n - {}_na_x) * {}_nm_x$$

$${}_nq_{85} = 1.00$$

Period Life Table Construction

- Step 4: Solve for ${}_np_x$ (probability of surviving an age interval)

$${}_np_x = 1 - {}_nq_x$$

$${}_np_{85} = 0.00$$

Period Life Table Construction

- Step 5: Set radix (l_0) and solve for number alive at age x (l_x).

$$l_{x+n} = l_x * {}_np_x$$

Period Life Table Construction

- Step 6: Solve for deaths experienced in each age interval ($_nd_x$)

$$_nd_x = l_x - l_{x+n}$$

Period Life Table Construction

- Step 7: Solve for person years lived between x and x+n (${}_nL_x$)

$${}_nL_x = n * l_{x+n} + {}_na_x * {}_nd_x$$

(open-ended interval: ${}_nL_x = l_x / {}_nm_x$)

Period Life Table Construction

- Step 8: Solve for person years lived above age x ($_nT_x$)

$$T_x = \Sigma {}_nL_a$$

Period Life Table Construction

- Step 9: Solve for life expectancy at age x (e^0_x)

$$e^0_x = T_x / l_x$$