Life tables: exercise

11. 1. 2023

ukolovae@natur.cuni.cz

Outline

Comments to exercise on standardization

Exercise on complete life tables

Exercise on abridged life tables

Exercise on abridged life tables from complete ones

Median length of life (probable life time)

Modal age at death

Life table notation

x	The age at start of the interval
n	The width of the age interval
${}_nq_x$	The probability of dying between exact ages x and x+n
${}_na_x$	The average proportion of the age interval x to x+n lived by those dying in the interval
${}_nk_x$	The average proportion of the age interval x to x+n lived by those surviving within the interval.
${}_np_x$	Probability of surviving between exact ages x and x+n
l_x	The number of survivors at exact age x
${}_nd_x$	The number of deaths between ages x and x+n
${}_nL_x$	The number of person-years lived between ages x and x+n
T_x	The total number of person-years lived after exact age x
e_x	The life expectancy et exact age x

Exercise 22

Compute **complete life tables for Spain, males and females in 2000 and 2019**. Apply indirect exponential method for the estimation of age specific probability of death. Consider open interval as 100+ and share of death in lower elementary set of events at age 0 equal to 0,88. Plot life table functions q_x, l_x, da and e_x.

[Data source: Eurostat, Human Mortality Database]

Exercise 23

Compute **abridged life tables for Spain, males and females in 2000 and 2019**. Apply indirect exponential method for the estimation of age specific probability of death. Consider open interval as 95+ and share of death in lower elementary set of events at age 0 equal to 0,88.

[Data source: Eurostat, Human Mortality Database]

Exercise 24

Compute **abridged life tables for Spain, males and females in 2000 and 2019 from complete life tables you computed in exercise 22.**

Median length of life (probable length of life)

The age by which half of an original cohort of births has died according to a particular set of age-specific death rates. Corresponds to the median age at death in a life table.

It is the age, at which l_x equals to 50 000.

$$\tilde{x} = x + n * \frac{l_y / 2 - l_x}{l_{x+1} - l_x}$$

Where y is exact age, for which we want to find out the median, x and x+1 are exact ages between which lives half of the persons surviving until age y

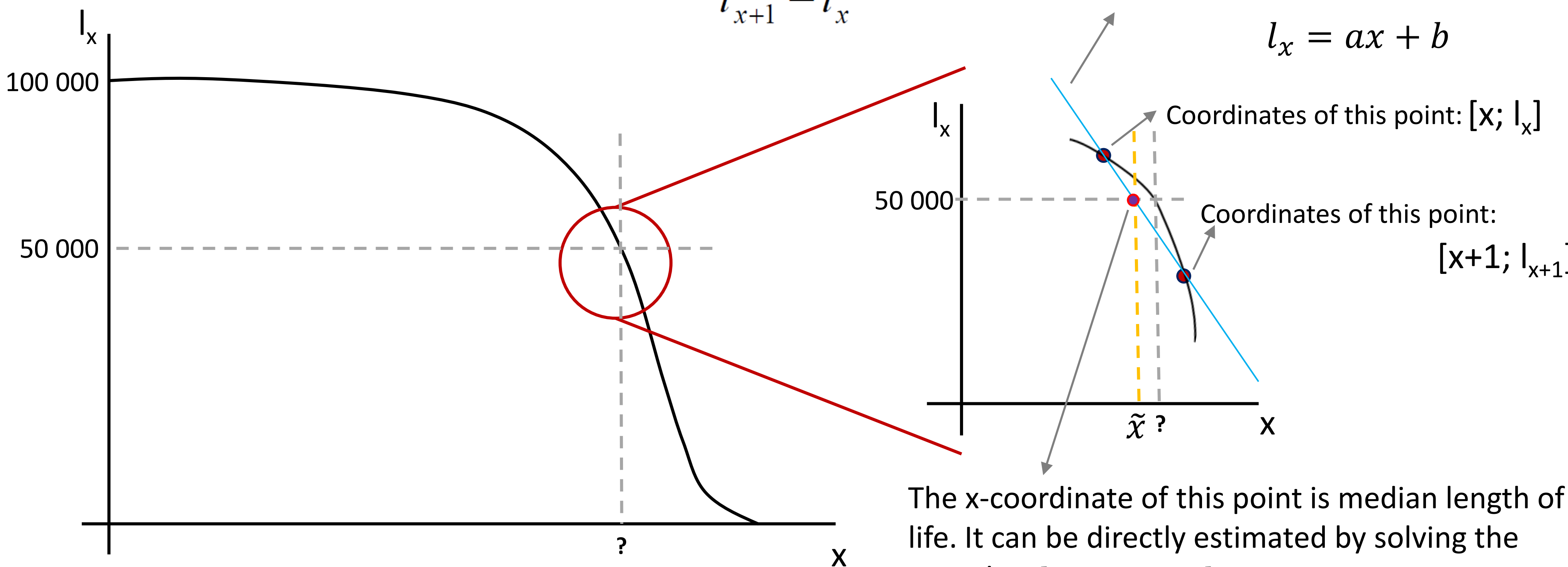
Median length of life (probable length of life)

What is behind the formula $\widetilde{x} = x + n * \frac{l_y / 2 - l_x}{l_{x+1} - l_x}$?

The problem can be understood as finding the x coordinate of fixed l_x equal to 50 000 in a two-dimensional space. This can be done by performing linear interpolation between two datapoints which delimit the interval, in which the x, which belongs to l_x= 50 000 lies.

Median length of life (probable length of life)

What is behind the formula $\tilde{x} = x + n * \dfrac{l_y/2 - l_x}{l_{x+1} - l_x}$?



The x-coordinate of this point is median length of life. It can be directly estimated by solving the equation $l_x = ax + b$ for lx = 50 000.

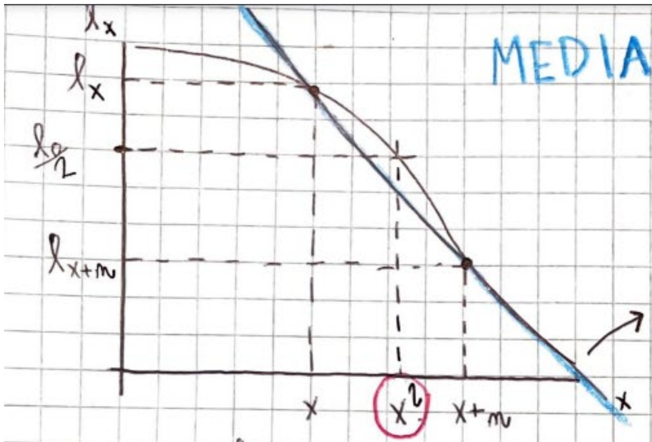
However, the parameters *a* nad *b* are so far unknown. Their estimation can be performed by solving the system of two equations with known datapoints around the median. Let us have the following data, which delimit the l_x=50 000:

x	l_x
80	66880
85	48676

See next slide for the solution.

Median length of life (probable length of life)

What is behind the formula $\tilde{x} = x + n * \dfrac{l_y/2 - l_x}{l_{x+1} - l_x}$?



x	l_x
80	66 880
85	48 676

x	l_x
80	66880
85	48676

estimate parameters of this function and solve x' value for y = 50 000 ($l_{x'}$)

1. slope = $\dfrac{\text{change in } y}{\text{change in } x} = \dfrac{48\,676 - 66\,880}{85-80} = -3\,640{,}81 = a$

2. model: $l_x = ax + b$

3. solve for one of the points: $66\,880 = -3640{,}81 \cdot 80 + b$
$66\,880 = -291\,265 + b$
$358\,145{,}1 = b$

4. solve for $l_x = 50\,000$: $50\,000 = -3640{,}81\,x + 358\,145{,}1$
$-308\,145{,}1 = -3640{,}81\,x$
$x = 84{,}63638 \doteq 84{,}64$

This formula

$$\tilde{x} = x + n * \frac{l_y/2 - l_x}{l_{x+1} - l_x}$$

yields exactly the same result:

$$\tilde{x} = 80 + 5 \cdot \frac{50\,000 - 66\,880}{48\,676 - 66\,880} = 80 + 5 \cdot \frac{-16\,880}{-18\,204}$$
$$= 80 + 5 \cdot 0{,}92727 = 84{,}63638 \doteq 84{,}64$$

Remark: *Obviously, this method of estimation (linear interpolation) is inaccurate for populations with steep decline in survival curve, because the space between the line and the curve is non-negligible in these cases.*

Modal age at death (normal age at death)

It is the age when density of life table death reaches its maximum. In other words, it is the age, when d_x is maximal.

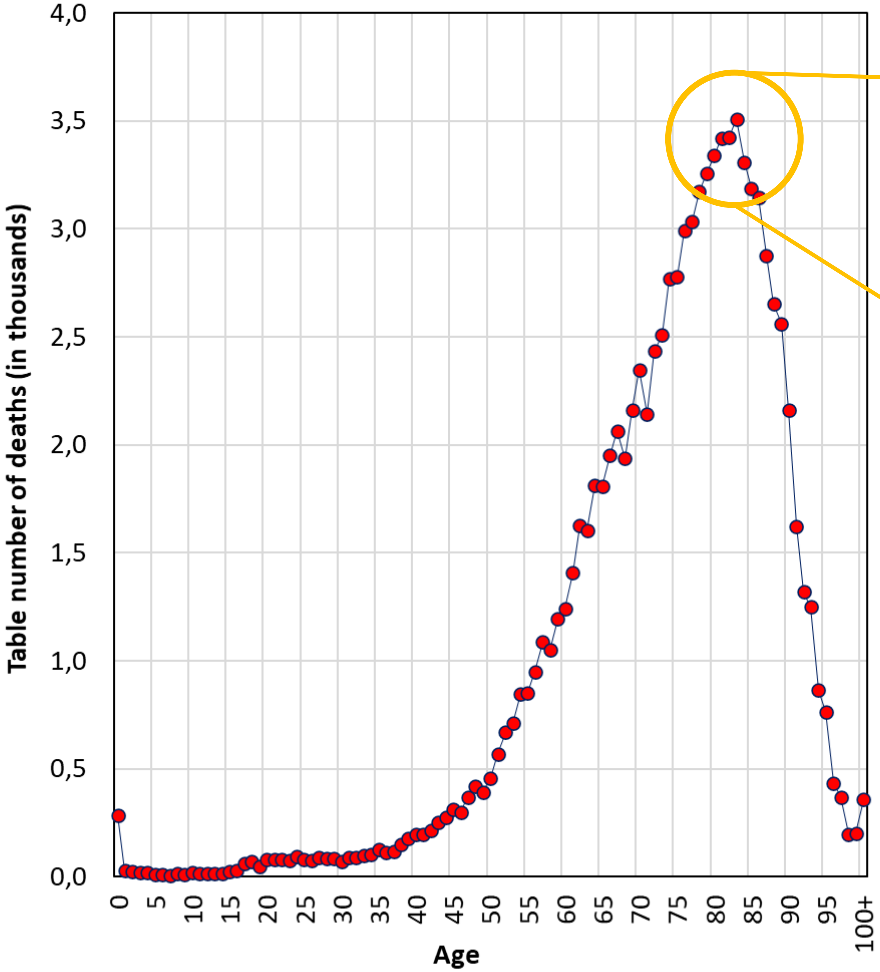
$$\hat{x} = x_c - \frac{n}{2} * \frac{d_{x+1} - d_{x-1}}{d_{x+1} + d_{x-1} - 2 * d_x}$$

Where x stands for completed age, at which d_x is maximal.

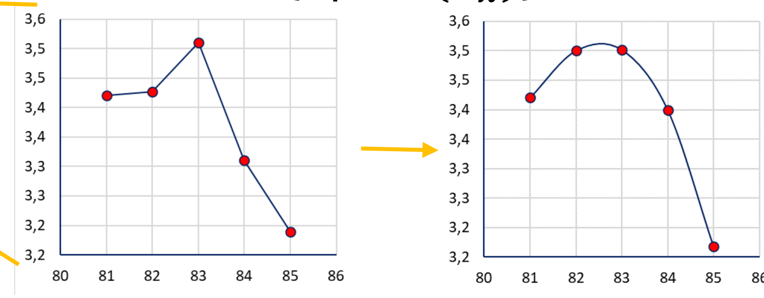
Modal age at death (normal age at death)

What is behid the formula $\hat{x} = x_c - \frac{n}{2} * \frac{d_{x+1} - d_{x-1}}{d_{x+1} + d_{x-1} - 2 * d_x}$?

Distribution of table number of deaths (function d_x)



We are looking for age, at which following condition is satisfied: $\hat{x} = \{x | \max(d_x)\}$.



The goal is to find maximum density of distribution of deaths. This maximum density can be located at age, at which we do not know the exact number of death (f. e. 80,28 or 81,74 etc.). Therefore, we fit polynomial function to three data points: the one, at which discrete d_x is maximal and to points around this value. Our polynomial function is parabola given by the equation:

$$d_x = \alpha x^2 + \beta x + \gamma$$

Maximum density of distribution of death is located at the maximum of this function. The function reaches its maximum, where its first derivative equals to zero. The first derivative of parabolic equation is equal to:

$$\frac{\partial d_x}{\partial x} = \frac{\partial(\alpha x^2 + \beta x + \gamma)}{\partial x} = 2\alpha x + \beta$$

Thus, the corresponding age of the maximum equals to: $x = -\frac{\beta}{2\alpha}$

All we have to do, is estimate α and β

Modal age at death (normal age at death)

What is behid the formula $\hat{x} = x_c - \frac{n}{2} * \frac{d_{x+1} - d_{x-1}}{d_{x+1} + d_{x-1} - 2 * d_x}$?

Application to the data from a fictive life table, where max d_x was found at age 11. Selected datapoints, which will be used to fit polynomial model are:

x	d_x
10	80
11	100
12	90

The parameters α, β, and γ will be found by solving the system of three equations with three unknown parameters. **BUT!** Table function d_x refers to completed age. We need to determine number of death exactly at the middle of an age interval. Under the assumption of uniformity of distribution of deaths, our datapoints shall be modified:

x	d_x
10,5	40
11,5	50
12,5	45

We use these datapoints to estimate the parameters α, β, and γ. And once we estimate them, we solve polynomials first derivative equal to zero for *x*.

See the solution on the next slide.

Modal age at death (normal age at death)

What is behid the formula $\hat{x} = x_c - \frac{n}{2} * \frac{d_{x+1} - d_{x-1}}{d_{x+1} + d_{x-1} - 2 * d_x}$?

x	d_x
10,5	40
11,5	50
12,5	45

1. $40 = \alpha \cdot 10{,}5^2 + \beta \cdot 10{,}5 + \gamma$
2. $50 = \alpha \cdot 11{,}5^2 + \beta \cdot 11{,}5 + \gamma$
3. $45 = \alpha \cdot 12{,}5^2 + \beta \cdot 12{,}5 + \gamma$

$40 = 110{,}25\alpha + 10{,}5\beta + \gamma$
$50 = 132{,}25\alpha + 11{,}5\beta + \gamma$
$45 = 156{,}25\alpha + 12{,}5\beta + \gamma \Rightarrow \gamma = -156{,}25\alpha - 12{,}5\beta + 45$

$40 = 110{,}25\alpha + 10{,}5\beta + (-156{,}25\alpha - 12{,}5\beta + 45)$
$50 = 132{,}25\alpha + 11{,}5\beta + (-156{,}25\alpha - 12{,}5\beta + 45)$

solve for β

$40 = 110{,}25\alpha + 10{,}5\beta - 156{,}25\alpha - 12{,}5\beta + 45$
$-5 = -46\alpha - 2\beta$
$2\beta = -46\alpha + 5$
$\beta = -23\alpha + \frac{5}{2} \Rightarrow \beta = -23(-7{,}5) + \frac{5}{2} = 175$

put to 2nd eq.

$50 = -24\alpha - \beta + 45$
$5 = -24\alpha - (-23\alpha + \frac{5}{2})$
$5 = -24\alpha + 23\alpha - \frac{5}{2}$
$\alpha = -5 - \frac{5}{2}$
$\alpha = -7{,}5$

find the exact age: $x = -\frac{\beta}{2\alpha}$

$x = -\frac{175}{-2 \cdot 7{,}5}$

$x = -\left(-\frac{175}{15}\right)$

$\underline{\underline{x = 11{,}6667}}$

All these computations lead exactly to the same results, as if we apply this formula:

$$\hat{x} = x_c - \frac{n}{2} * \frac{d_{x+1} - d_{x-1}}{d_{x+1} + d_{x-1} - 2 * d_x}$$

to the former data directly from the life table:

x	d_x
10	80
11	100
12	90

Thus:

$\hat{x} = 11{,}5 - \frac{1}{2} \cdot \frac{90 - 80}{90 + 80 - 2 \cdot 100} =$

$= 11{,}5 - 0{,}5 \cdot \frac{10}{-30} = 11{,}5 - \frac{1}{2} \cdot \left(-\frac{1}{3}\right) =$

$= 11{,}5 + \frac{1}{6} = \underline{\underline{11{,}6667}}$

The same logic works for abridged life tables too.