Basic methodological notes

The complete life tables for the Czech Republic and its NUTS-2 cohesion regions and NUTS-3 regions are calculated from the third main group of demographic events (age-period specific). Input death probabilities are computed indirectly, it means they are derived from real (observed) age-specific mortality rates (see below).The life tables are detailed i.e. by single year of age with an open age interval for 105+. They are computed separately for men and women.

To exclude random fluctuations, the life tables for regions (NUTS 3) were calculated for a two-year period. The input data always refers to the regional breakdown valid as of 1 January of the given year of the event.

Life tables indicators

*The **number of deaths** (D_x) states the absolute number of deaths by age x during the reference period.*

*The **number of inhabitants** (P_x) states the mid-year population by age x. In life tables for the Czech Republic, cohesion regions (NUTS 2), and regions (NUTS 3) the mid-year population is defined as the balanced population as of 1 July of a given year, or as the sum of mid-year populations in the case of multiple-year periods.*

*The **death probability** (q_x) expresses the probability that an individual at the exact age of x years will die in a given period, i.e. before the exact age of $x + 1$ years:*

$$q_x = \frac{m_x}{1 + (1 - a_x) \cdot m_x} \; .$$

where m_x is the mortality rate at a given age and the parameter a_x is the average number of years lived within the age interval $[x, x + 1)$ for people dying at that age (see below).

*The **table number of survivors** (l_x) is a hypothetical number of individuals alive at the exact age of x years out of 100,000 live births (table radix $l_0 = 100{,}000$), given the mortality conditions of the reference period:*

$$l_{x+1} = l_x \cdot (1 - q_x) \; .$$

*The **table number of deaths** (d_x) is a hypothetical number of individuals who die at the complete age of x years; it is computed as the difference between two subsequent table numbers of survivors:*

$$d_x = l_x - l_{x+1} \; .$$

*The **table number of person-years** (L_x) is a hypothetical number of person-years lived by the life-table population in the age interval x and $x + 1$:*

$$L_x = l_x - (1 - a_x) \cdot d_x \; .$$

*The **auxiliary indicator** (T_x) expresses the number of years of life to be lived by the life-table population (not by an individual) at a given age x. It is the accumulation of L_x from the age of x to the highest age of the table.*

$$T_x = \sum_x^{105+} L_x \,.$$

The **life expectancy** or the **expectation of life** (e_x) shows the average number of years an x-year-old individual can expect to live, given the mortality conditions of the reference period. It is a summary indicator reflecting mortality conditions in all age groups in a given year.

$$e_x = \frac{T_x}{l_x} \,.$$

Detailed description of a life table processing

A. Input data processing

1. Calculation of real (observed) mortality rates

The so-called real (observed) mortality rate (m_x) is computed from empirical data as the ratio of the number of deaths and the mid-year population at a given age, sex, and area:

$$_t m_x^{CR,cohesion\ region} = \frac{_t D_x}{_{1.7.(t)} P_x},$$

$$_{t-1,t} m_x^{region} = \frac{_{t-1} D_x + _t D_x}{_{1.7.(t-1)} P_x + _{1.7.(t)} P_x} \,.$$

2. Smoothing of real (observed) mortality rates

To exclude possible random fluctuations in real (observed) mortality rates, which are common, particularly at ages with low mortality or a low number of survivors, the mortality rates are smoothed by the generalised additive model (the GAM method) combined with the so-called P-splines; this method is sometimes called the P-GAM method (Eilers and Marx, 1996). The age is the predictor, the number of deaths is the predicted figure, and the population is used as an offset. Poisson distribution of the number of deaths is assumed. The level of smoothing is adapted to the number of deaths at a given age; it means that a lower number of deaths is smoothed more than a higher number of deaths; the significance of the original phenomenon/trend is thus maintained.

Mortality modelling at high ages

Regarding low numbers of deaths at high ages and lower reliability of data on the mid-period oldest population, the smoothed observed mortality rates are substituted by model figures at oldest-old ages. Generally, mortality models work with a continuous mortality rate function, the so-called force of mortality, denoted by μ_x. In practice, it is assumed, based on the monotonic increase of mortality with age that the force of mortality and the mortality rate at a given age are in this connection approximately (Thatcher et al., 1998):

$$m_x \cong \mu_{x+0.5} \,.$$

The model selected here is based on a logistic curve, which takes into consideration the deceleration in mortality increase with age. Specifically, the model named by Kannisto is used (Thatcher et al., 1998). So the programme models the age-specific mortality by the following function:

$$\mu_x = \frac{\alpha e^{\beta x}}{1 + \alpha e^{\beta x}} + c \ .$$

where x is the age, α, β are parameters, and c is a constant. The parameters are estimated by the maximum likelihood estimation method (MLE). The default parameters estimate is based on an assumed Poisson distribution of the number of deaths (D) at the age x. Based on these assumptions, the logarithm of the likelihood function (L) is maximized as:

$$ln\,L = \sum_x D_x\,ln(\mu_x) - E_x\mu_x.$$

During the optimization process (of the negative logarithm) of the likelihood function, the programme works with its default values of parameters. For estimation of the parameters, the age range of 70–90 years is set.

Regarding a transition from smoothed to model data, it is assumed that the age, from which model data characterize mortality better than smoothed data is higher than 75. For smooth progress of the substitution of smoothed data with model ones, the age denoted by y (higher than 75 years) is sought first, where the difference between the smoothed and the model figure is minimal:

$$|m_x^{smooth} - m_x^{model}| = min, \quad y \geq 75 \ .$$

A step further, a gradual change from the smoothed figures to model ones is made for the ages $x \in [(y - 4), (y + 4)]$ according to:

$$m_x^{grad} = \left(1 - \frac{x - y + 5}{10}\right) \cdot m_x^{smooth} + \left(\frac{x - y + 5}{10}\right) \cdot m_x^{model} \ .$$

3. *For the computation of life tables the following data on age-specific mortality rates are used as the input:*

at the age of 0	*real (observed) figures,*
at the age of 1 to $(y - 5)$	*smoothed figures,*
at the age of $(y - 4)$ to $(y + 4)$	*graduation from smoothed to model figures,*
at the age of $(y + 5)$ and higher	*model figures.*

B. Conversion of real smoothed data to table figures

The basis for the calculation of life tables in an indirect way is to convert age-specific mortality rates (m_x) to death probabilities (q_x). Based on the assumption that the real mortality rates and the table ones are equal and on the assumption that mortality rates in a hypothetical table cohort are constant, the following conversion formula is used (Chiang, 1984):

$$q_x = \frac{m_x}{1 + (1 - a_x) \cdot m_x} \quad \ .$$

ČSÚ

The parameter a_x quantifies the average number of person-years lived in the given age interval by those who die in the interval. The figure a_x can be observed directly from the real data (when the exact ages of deceased persons are available the same as their average), or be deduced from observed mortality rates or taken from a population with similar mortality schedule (Preston et al., 2000). With regards to the generally accepted assumption of a uniform distribution of an exact age of deceased persons, the value is set to 0.5 for all ages except 0 and the last (open) age interval.

In case of the estimation of the average age of deceased infants (a_0) the most exact way of calculation based directly on the observed data is used (Andreev and Kingkade, 2015); it is computed from the real (observed) individual data as the observed proportion of the mean age at death for children who died in their first year of life (expressed in days) in the length of the whole calendar year (365 days):

$$a_0^{poz} = \left(\frac{\sum X_{D_{i,0}}^{(days)}}{\sum D_{i,0}} \right) / 365 \; .$$

where $D_{i,0}$ is an individual death in the first year of life (infant period) and $X_{D_{i,0}}^{(days)}$ is an individual age of the deceased infant in days. The a_0 value is set as one of the parameters (the default value is 0.1). In case of lower regional breakdowns, the value is always the one calculated from data for the whole Czech Republic in the given period.

Concerning the last open age interval (105+ here) the a_x parameter is calculated as (see for example Preston et al., 2000):

$$a_{105+} = \frac{1}{m_{105+}} \; .$$

The next steps in life tables processing follow the computed q_x values (see Life tables indicators).

Bibliography

Andreev, E. and Kingkade, W. W. (2015). Average age at death in infancy and infant mortality level: Reconsidering the Coale-Demeny formulas at current levels of low mortality. *Demographic Research*. 33(13): 363−390.

Eilers, P. H. C. and Marx B. D. (1996). Flexible Smoothing with B-splines and Penalties. *Statistical Science*. 11(2):89-121.

Chiang, C. L. (1984). *The Life Table and Its Applications*. Robert E. Krieger Publishing Company, Malabar, FL, USA.

Preston, S. H., Heuveline, P., Guillot, M. (2001). *Demography: Measuring and modeling population processes*. Oxford: Blackwell Publishers.

Thatcher, A., Kannisto, V., and Vaupel, J. W. (1998). *The Force of Mortality at Ages 80 to 120*. Odense, Denmark: Odense University Press.