Estimation of the force of mortality -independently of the initial health state- from cross-longitudinal surveys using IMaCh version 0.97

The starting point (origin of time) of the duration of survival of each individual is the date of entry in the study, i.e. its age at the date of the first wave. The time of survival is measured until the date of the death if the subject died before the last interview or until the age at the last interview if the subject is still alive. The models classically used in analysis of the biographies consider only the duration of survival and suppose that all the individuals are interviewed at the same time. Because of the great disparities of the ages at the first wave, it is mandatory to take into account the age in the model of analysis of survival. The estimated parameters are calculated with the method of the maximum of probability.

Let be x_i the age at the first interview of individual i, x_i^d is the age at death, x_i^c is the age at the last interview and δ_i a dummy variable indicating the status (δ_i=0 if the individual is dead and 1 otherwise).

If the subject is dead, its contribution to the likelihood is the product of the survival probability between age x_i and x_i^d by the probability of dying between age x_i^d and $x_i^d + 1$. This contribution is

$$\mu(x_i^d) \exp\left(-\int_{x_i}^{x_i^d} \mu(u)du\right). \tag{1}$$

The contribution of a surviving suject to the date of the last wave is the survival probability between age x_i and x_i^c, i.e.

$$\exp\left(-\int_{x_i}^{x_i^c} \mu(u)du\right). \tag{2}$$

The total likelihood L of n independant sujects, indexed by i, is the product of the contributions of each individuals :

$$L = \Pi_{i=1}^{n} \left[\mu(x_i^d) \exp\left(-\int_{x_i}^{x_i^d} \mu(u)du\right)\right]^{(1-\delta_i)} \left[\exp\left(-\int_{x_i}^{x_i^c} \mu(u)du\right)\right]^{(\delta_i)} \tag{3}$$

where $\mu(x)$ is the force of mortality at age x. By definition, $\mu(x)dx$ is the probability for an individual aged x to die between ages x and $x + dx$.

The log-likelihood is then

$$l = \sum_{i=1}^{n}(1-\delta_i)\left[\left(-\int_{x_i}^{x_i^d} \mu(u)du\right) + \log(\mu(x_i^d))\right] + \delta_i\left[-\int_{x_i}^{x_i^c} \mu(u)du\right] \tag{4}$$

Suppose that the force of mortality is modelled by a Gompertz law where the two parameters are μ_{100} and θ_1. The force of mortality is $\mu(x) = \mu_{100} \exp(\theta_1(x - 100))$. The parameter μ_{100} is the force at age 100 ans and θ_1 is the slope.

Then the log-likelihood is

$$
\begin{aligned}
l(\mu_{100}, \theta_1) \quad = \quad & \sum (1-\delta_i)\left[-\frac{\mu_{100}}{\theta 1}\left(\exp(\theta_1 x_i^d) - \exp(\theta_1 x_i)\right) + \log(\mu_{100}) + \theta_1 (x_i^d)\right] \\
& +\delta_i \left[-\frac{\mu_{100}}{\theta 1}\left(\exp(\theta_1 x_i^c) - \exp(\theta_1 x_i)\right)\right] \qquad (5)
\end{aligned}
$$

The usual software of statistics cannot be employed to implement this parametric model because their procedures making it possible to carry out biographical analyses do not take into account the age. All the estimates and the construction of the confidence intervals were carried out with a program written in language C. We used a function of maximization based on the algorithm of Powell describes in the book *Numerical Recipes in C* (1992). The matrix of covariance is calculated like the reverse of the matrix hessienne to the optimum.