HMD-LQ Model Life Table Fitting TEMPLATE.xls – README

1. Introduction

`HMD-LQ Model Life Table Fitting Template.xls` is an Excel spreadsheet for fitting the log-quadratic model life table family proposed by Wilmoth *et al* (2011) to observed age-specific death rates. The usual free software disclaimer applies. THIS SOFTWARE COMES WITH ABSOLUTELY NO WARRANTY. If you do not understand both the model and its implementation in this spreadsheet, do not use the spreadsheet.

"HMD-LQ Model Life Table" refers to the log-quadratic model life family with the coefficients shown in Table 3 of Wilmoth *et al* 2011 (females and males) and in the file `HMD-719.csv` available at URL `demog.berkeley.edu/~jrw/LogQuad` given on page 14 of the reference (both sexes).

2. The HMD-LQ Model

HMD-LQ model ${}_nm_x$ values are specified by two parameters, $q(5) = {}_1q_0$ and k. Age-specific death rates corresponding to values of these two parameters are obtained as follows.

1. Calculate $h = \log(q(5))$.
2. Calculate age-specific death rates ${}_1m_0$, ${}_5m_5$, ..., ${}_5m_{105}$, ${}_\infty m_{110}$ for ages 0, 5-9, ..., 105-109, 110+ from equation (1), page 8, $\exp(a_x + b_x h + c_x h^2 + v_x k)$.
3. Calculate the ${}_1q_0$ corresponding to ${}_1m_0$.
4. Calculate the ${}_4q_1$ implied by ${}_1q_0$ and the input ${}_5q_0$,
5. Calculate the ${}_1m_0$ implied by ${}_4q_1$.

Steps 3 requires a formula for calculating ${}_nq_x$ from ${}_nm_x$ The formula used here is

$$ {}_nq_x = \frac{n\,{}_nm_x}{1+{}_nb_x n\,{}_nm_x} \tag{1}$$

where ${}_nb_x$ is the separation factor. This is equivalent to the formula

$$ {}_nq_x = \frac{n \times {}_nm_x}{1+(n-{}_na_x)\times {}_nm_x} \tag{2}$$

given in Preston *et al* (19*xx*) with ${}_nb_x = (n - {}_na_x)/n$. ${}_1b_0$ and ${}_4b_1$ are calculated from ${}_1a_0$ and ${}_4a_1$ calculated from the equations in the R functions `coale.demeny.a0` and `coale.demeny.4a1`, respectively, both available at `demog.berkeley.edu/~jrw/LogQuad`.

The ${}_4q_1$ implied by ${}_1q_0$ and the input ${}_5q_0$ (Step 4) is given by

$$ {}_4q_1 = 1 - \frac{(1-{}_1q_0)}{(1-{}_5q_0)} \tag{3}$$

For step 5, inverting the ${}_nq_x \sim {}_nm_x$ formula above gives

$$ {}_nm_x = \frac{{}_nq_x}{n(1-{}_nb_x\,{}_nq_x)}. \tag{4}$$

3. Using the Spreadsheet

The worksheet HMD-LQ Fit implements all fitting calculations. To use it to fit an observed set of age-specific death rates, proceed as follows.

1. Enter the country name, period and sex into cells B3, B4 and B5, respectively. Sex must be entered accurately as "Female", "Male" or "Both" to ensure that the correct coefficients are used (see the formulas in cells K4:M4). "Code" is not required (it is included for use in fitting the HMD tables, for which country codes are provided).
2. Paste age-specific death rates ${}_1m_0$, ${}_4m_1$, ${}_5m_5$, ..., into cells C15:C38. It is not necessary to have rates to 110+. Note that input cells are colored blue.
3. Enter initial values of h and k in cells G2 and G3, respectively. Use $k = 0$ as a default value if fitting one parameter only.
4. Edit the sum of squared differences formula in cell G11 to specify the range of ${}_nm_x$ to be fit. The template shows values for one of the HMD life tables, for which a full set of age-specific death rates is available. If the observed ${}_nm_x$ values end with an earlier open ended age group, reduce the range of the sum of squared differences formula to end with the ${}_nm_x$ preceding the open ended m value. Other selections may be used for specific purposes.
5. Use SOLVER to minimize the goodness of fit value in cell G11 varying either (a) h in cell G3 with $k = 0$ for a one parameter fit or (b) h and k in cells G3 and G2 for a two parameter fit. Note that h, not $q(5)$, is varied, this to avoid errors when using SOLVER.
6. Scrutinize the plots for goodness of fit.

The fitted age-specific death rates are given in cells D15:D38. These cells show all age-specific death rates to the 110+ age group, whatever observed rates are supplied as input.

4. Spreadsheet Implementation

Cells D15 and D17:D38 contain the model age-specific death rates ${}_nm_x$ calculated from the model formula $\exp(a_x + b_xh + c_xh^2 + v_xk)$. The values h and k are in cells G3 and G2, respectively. The coefficients (a_x b_x c_x v_x) used are in cells J15:M38 (less J16:M16). The formulas in these cells link them to the coefficients for females, males or both sexes shown at right. The linking uses the simple device of expressing each coefficient to be used as a weighted sum of the three possible coefficients, where the weights are (1,0,0) for females, (0,1,0) for males, or (0,0,1) for both sexes. The weights, shown in cells K4:M4, are calculated from the input cell B5 (sex) using the IF and MID functions. The same device is used to produce ${}_1b_0$ and ${}_4b_1$ values in cells I6 and I8.

Calculation of the fitted ${}_4m_1$ is done in cells G4:G10.

1. Cell G5 reproduces the fitted ${}_1m_0$ value in cell D15.
2. Cell G6 reproduces the ${}_1b_0$ value in cell I5.
3. Cell G7 calculates ${}_1q_0$ from ${}_1m_0$ and ${}_1b_0$ using formula (1).
4. Cell G8 calculates ${}_4q_1$ from ${}_4q_1$ and ${}_5q_0$ (cell G4) using formula (3).
5. Cell G9 reproduces the ${}_4b_1$ value in cell I8.
6. Cell G10 calculates ${}_4m_1$ from ${}_4q_1$ and ${}_4b_1$ using formula (4).

Cells D15 and G5, both of which contain the fitted ${}_1m_0$ are colored orange to highlight the linkage between them. Cell D16 reproduces ${}_4m_1$value in cell G10 resulting from the calculation. These cells are are colored pink to highlight the linkage between them.

Cells O2:T10 and U2:AB6 show the key spreadsheet formulas used, the spreadsheet expressions for formulas (1), (3) and (4) above as well as the formula expressing ${}_nb_x$ in terms of ${}_na_x$.

5. Plots

The Obs+Fit worksheet plots the observed and fitted Log(${}_nm_x$) values. Log(${}_nm_x$) rather than ${}_nm_x$ values are plotted to better show the age pattern. When ${}_nm_x$ are plotted, the high values for the youngest and oldest ages force a vertical scale that obscures in the complementary ages. Plotting Log(${}_nm_x$) has the additional advantage of showing the extent to which death rates at older ages conform to the Gompertz model.

The Residual sheet plots the residuals (Log(${}_nm_x$ Obs) - Log(${}_nm_x$ Fit)).

The RelRes mx sheet plots the relative residuals of the fitted ${}_nm_x$ — not Log(${}_nm_x$) — values, that is, the difference (${}_nm_x$ Observed - ${}_nm_x$ Fitted) divided by ${}_nm_x$ Observed. This plot provides an indication of how close the fitted values come to the observed values.

Plots may require manual adjustment of the vertical scale and the aspect ratio.

6. Life Tables

The Fitted Life Table sheet shows a life table calculated from the fitted ${}_nm_x$ values using formula (1) and the ${}_nb_x$ values in cells G9:G32, which are calculated from the ${}_na_x$ values in cells F9:F32. The ${}_na_x$ and ${}_nb_x$ for the first two age groups are calculated in cells H2:L6 as described above. For older age groups ${}_na_x$ is taken to be 2.5 through age 75-79 and to decline linearly to 1.25 for age 110+. This is a convenient expedient only.

The Observed Life Table sheet shows a life table calculated in the same way from the observed ${}_nm_x$ values. Observed values in the template go out to the 110+ age group. If observed values end with a younger open-ended age group, as they often will, some decision needs to be taken on how to proceed.

A simple expedient when life expectancy at birth is not very high is to use the fitted values for the age groups for which observed values are not available. A more complicated but perhaps preferable approach, again for cases in which life expectancy at birth is not high, is to fit the observed rates for older age groups by a Gompertz function and use the fit to extrapolate deaths rates for the oldest age groups. A third approach is to use an open-ended age group formula.

References

Preston, Samuel H., ... Heuveline, and Michel Guillot. 19*xx*. *Demography: Measuring and Modeling Population Models and Processes …. Place: Publisher.*

Wilmoth, John, Sarah Zureick, Vladimir Canudas-Romo, Mie Inoue, and Cheryl Sawyer. 2011. A flexible two-dimensional mortality model for use in indirect estimation. *Population Studies* **V**(N):1-28