



Max-Planck-Institut für demografische Forschung  
Max Planck Institute for Demographic Research  
Konrad-Zuse-Strasse 1 · D-18057 Rostock · GERMANY  
Tel +49 (0) 3 81 20 81 - 0; Fax +49 (0) 3 81 20 81 - 202;  
<http://www.demogr.mpg.de>

---

MPIDR TECHNICAL REPORT 2010-002  
MAY 2010

**Age-decomposition of a difference  
between two populations for any  
life-table quantity in Excel**

Vladimir M. Shkolnikov ([shkolnikov@demogr.mpg.de](mailto:shkolnikov@demogr.mpg.de))  
Evgeny M. Andreev ([andreev@demogr.mpg.de](mailto:andreev@demogr.mpg.de))

---

This technical report has been approved for release by: Vladimir Shkolnikov ([shkolnikov@demogr.mpg.de](mailto:shkolnikov@demogr.mpg.de)),  
Head of the Laboratory of Demographic Data.

© Copyright is held by the authors.

Technical reports of the Max Planck Institute for Demographic Research receive only limited review.  
Views or opinions expressed in technical reports are attributable to the authors and do not necessarily  
reflect those of the Institute.

# Age-decomposition of a difference between two populations for any life-table quantity in Excel

by Vladimir M. Shkolnikov and Evgeny M. Andreev

## **Abstract**

We provide a simple VBA/Excel program that decomposes by age a difference between two values of a life-table based quantity. For example, one might want to know what are contributions of different ages into the total difference between two populations in: values of life expectancy at birth or of temporary life expectancy between exact ages 20 and 65 or of standard deviation of ages at death over age 10 and older ages or of survival from age 0 to age 15 or of any other life-table-based measure. The program uses the general algorithm of stepwise replacement for estimating the age-components.

**Keywords:** decomposition, stepwise replacement algorithm, age components, life expectancy, life table, life table disparity

## **Background**

Comparative analyses often look at differences between values of aggregate demographic measures. Once such difference is known, the researcher may be interested in decomposing it according to impacts of underlying factors, such as (for example) age groups, sexes, and causes of death. The general decomposition problem is to estimate the additive contributions of differences between values of the factors to the overall difference between values of the aggregate measure. The aggregate measure of interest is to be considered as a dependent function or functional of the factors.

If the aggregate measure is a sum or a linear function of the factors, then the decomposition problem is trivial. E. Kitagava (1955) was the first to solve a non-trivial decomposition problem. She applied the *standardization* method to decomposing a difference between two crude mortality rates by splitting it into additive components related to differences between age-specific death rates and differences between population age structures. In the 1980s, demographers developed methods that permitted them to decompose differences between two life expectancies by ages and causes of death (Pollard, 1982, Andreev, 1982, Arriaga, 1984, Pollard, 1988), and to decompose differences between two values of completed cohort fertility by specific parity progressions (Pullum, Tedrow, Herting, 1989). Notably, each of these methods was able to solve the decomposition problem for one particular type of aggregate measure.

Later on, a more universal decomposition methodology was developed. Das Gupta was the first to attempt the decomposition of differences between two values of *any* aggregate measure. The measure has to be based on a multi-dimensional array of elementary rates (Das Gupta, 1991, Das

Gupta, 1999). His decomposition method includes all possible replacements of the elementary rates of the first population with the corresponding rates of the second population.

Building on this approach, Andreev, Shkolnikov, and Begun (2002) elaborated a general algorithm of stepwise replacement and applied it to several previously unresolved decomposition problems. In particular, the algorithm was used for numerical and analytical decomposition of a difference between health expectancies by age, and also by mortality and health within each age group. A numerical decomposition of a difference between total fertility rates and between parity progression ratios by age and birth order was performed as well. In a later study, the same algorithm was applied for an unusual decomposition of a difference between two life expectancies by age and educational group (Shkolnikov, Andreev, Begun, 2003).

Finally Horiuchi, Wilmoth, and Pletcher (2008) introduced a different type of universal decomposition method. While the prior approaches were based on discrete changes in the value of each factor, the former method is based on the assumption that values of the factors change continuously. This assumption can be considered to be a natural choice for time-trend analyses because many variables change gradually with time. In the paper, the proposed decomposition method was applied to the age-decomposition of changes in mean and median lengths of life, and in the standard deviation of ages at death, as well as to the decomposition of a difference in reported health between two U.S. states, in respect to ethnic structures and contextual effects of state.

### ***Age decomposition with the general stepwise replacement algorithm***

The general stepwise replacement algorithm (Andreev, Shkolnikov, Begun, 2002) allows us to execute age decompositions for any types of life-table-based quantities.

Consider two values of an aggregate mortality measure  $E$  calculated from life tables for two populations. The two quantities are values of functional of respective vectors of age-specific death rates  $E^1 = E(M^1)$  and  $E^2 = E(M^2)$ . Values  $E^1$  and  $E^2$  are to be returned by conventional computations of two life tables from vectors of age-specific mortality rates  $M^1 = \parallel m_x^1 \parallel$  and  $M^2 = \parallel m_x^2 \parallel, x=0, 1, 2, \dots, \omega$ . Transition from quantity  $E^1$  to the quantity  $E^2$  corresponds to a transformation of vector  $M^1$  into vector  $M^2$ . Such a transformation can be completed by a stepwise replacement in an age-by-age mode of the elements of the first vector by respective elements of the second vector.

Let  $M_{[x]}^1$  be the vector, consisting of elementary mortality rates  $m_y^2$  at ages  $y < x$  and  $m_y^1$  at ages  $y \geq x$  and  $M_{[x+1]}^1$  be the vector, consisting of elementary mortality rates  $m_y^2$  at ages  $y < x+1$  and  $m_y^1$  at ages  $y \geq x+1$ . Then the contribution to the overall difference  $E^2 - E^1$  produced by the difference between the death rates  $m_x^2$  and  $m_x^1$  at age  $[x, x+1)$  is

$$\delta_{x|x+1}^{2 \rightarrow 1} = E(M_{[x+1]}^1) - E(M_{[x]}^1). \quad (1a)$$

Formula (1) suggests that all age-specific components can be computed by stepwise replacement of age-specific death rates of population 1 by corresponding age-specific death rates of population 2 progressing from age 0 to the last age  $\omega$ . It was shown (Andreev, Shkolnikov,

Begun, 2002) that if  $E = e_0$ , then the age components returned by (1a) are exactly equal to the components produced by known methods for decomposition of differences between life expectancies by Andreev (1982) and Arriaga (1984). For reasons of symmetry, replacement 2→1 realized by (1a) should be complemented by replacement 1→2:

$$\delta_{x|x+1}^{1 \rightarrow 2} = E(M_{[x+1]}^2) - E(M_{[x]}^2). \quad (1b)$$

Numerically, components (1a) and (1b) have opposite signs and are often not exactly the same by absolute value (Shkolnikov, Andreev, Valkonen, 2001). The final age components are obtained by averaging of (1a) and (1b)

$$\delta_{x|x+1} = (\delta_{x|x+1}^{2 \rightarrow 1} - \delta_{x|x+1}^{1 \rightarrow 2}) / 2. \quad (2)$$

The resulting additive age-decomposition of the total difference between quantities  $E^2$  and  $E^1$  is

$$E^2 - E^1 = \sum_x \delta_{x|x+1}. \quad (3)$$

### **How it works**

The spreadsheet “Age-decomposition-complete-LT.xls” consists of three worksheets: “Life Table1,” “Life Table2,” and “Decomposition.”

The user must enter two input vectors of age-specific death rates for populations 1 and 2 in columns B (under column titles ‘mx) of worksheets “Life Table1” and “Life Table2,” respectively. Information about sex for the two populations (‘Males or ‘Females) should be entered in cells B2 in the same worksheets. The standard functions of the complete life table for populations 1 and 2 will be instantly computed in the worksheets “Life Table1” and “Life Table2.” Computation of the life table is being performed according to the Human Mortality Database method. In addition, three life-table dispersion measures will be computed in columns N, O, and P under column titles ‘e<sup>†</sup>x, ‘AIDx, and ‘STDx (see Shkolnikov and Andreev (2010) for more details).

Note that a user is free to calculate any other quantity in columns R, S, and other free columns in the worksheets “Life Table1” and “Life Table2.” For example, the user may be interested in the median length of life rather than the mean length of life or in the coefficient of variation rather than the standard deviation, or in years of life lost (YLL) rather than e-dagger.

Then it is necessary to specify quantities for the age decomposition in cells S2 in worksheets “Life Table1” and “Life Table2.” For example, if both cells S2=I2, then the difference between life expectancies at birth will be decomposed; if S2=(H19-H69)/G19, then the difference between temporary life expectancies between exact ages 15 and 65 will be decomposed; if S2=N14, then the difference between e-dagger values at age 10 will be decomposed.

Finally, the user should run the program for the age decomposition of the difference between quantities in cells S2 in the worksheets “Life Table1” and “Life Table2.” This can be done by clicking on button “Run macro” in the worksheet “Decomposition.” The age-specific components  $\delta_{x|x+1}^{2 \rightarrow 1}$ ,  $\delta_{x|x+1}^{1 \rightarrow 2}$ , and  $\delta_{x|x+1}$  will appear within one second in columns B, D, and F, respectively. The final age-components from (3) that have to be used in further analysis should

be taken from column F. Once these components are obtained, they are depicted in the figure in the worksheet "Decomposition."

## **References**

- Andreev E.M. 1982. Metod komponent v analize prodoljitelnosti zjizni. [The method of components in the analysis of length of life]", *Vestnik Statistiki*, 9,42-47.
- Andreev, E. M., Shkolnikov, V. M., Begun, A. Z. 2002. Algorithm for decomposition of differences between aggregate demographic measures and its application to life expectancies, healthy life expectancies, parity-progression ratios and total fertility rates. *Demographic Research*, 7:14, 499-522.
- Arriaga E. 1984. Measuring and explaining the change in life expectancies, *Demography*, 21(1), 83-96.
- Das Gupta, P. 1991. Decomposition of the Deference between Two Rates and Its Consistency When More e than Two Populations are Involved. *Mathematical Population Studies* 3(2): 105-125.
- Das Gupta, P. 1999. Decomposing the difference between rates when the rate is a function of factors that are not cross-classified. *Genus*, LV(1-2), 9-26.
- Horiuchi S., Wilmoth J.R., Pletcher S. 2008. A decomposition method based on a model of continuous change. *Demography*, Volume 45-Number 4, November: 785–801
- Kitagawa, E. 1955. Component of a difference between two rates. *Journal of American Statistical Association*. Vol. 50 No. 272. 1168-1194.
- Pollard J.H. 1982. The expectation of life and its relationship to mortality, *Journal of the Institute of Actuaries*, 109(442), Part II, 225-240.
- Pollard, J.H. 1988. On the decomposition of changes in expectation of life and differentials in life expectancy. *Demography*, 25, 265-276.
- Pullum, T. W., Tedrow, L. M., and Herting, J. R. 1989. Measuring change and continuity in parity distributions. *Demography* 26: 485-498.
- Shkolnikov, V. M., Andreev, E. M.; Begun, A. Z. 2003. Gini coefficient as a life table function: computation from discrete data, decomposition of differences and empirical examples/ *Demographic Research*, 8:11, 305-358
- Shkolnikov, V.M. and E.M.Andreev. 2010. Spreadsheet for calculation of life-table dispersion measures. Max Planck Institute for Demographic Research, TR-2010-001.