Taller 3 - Life Expectancy Decomposition
LEdecomp R-package

D. Atance[1]

[1] Universidad de Alcalá, Madrid, Spain.

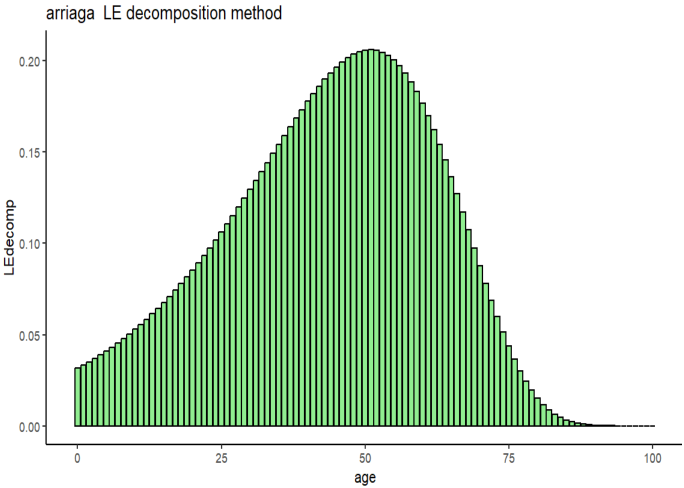
IV Congreso Usuarios de R 2025, Valencia, Octuber 2025

Index

1. Resources:

1. Arriaga, E.E. (1984). Measuring and explaining the change in life expectancies. Demography 21: 83–96. https://link.springer.com/article/10.2307/2061029
2. Andreev, E.M., Shkolnikov, V.M., and 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: 499–522. https://www.jstor.org/stable/26348070
3. Horiuchi, S., Wilmoth, J.R., and Pletcher, S.D. (2008). A decomposition method based on a model of continuous change. Demography 45: 785–801. doi:10.1353/dem.0.0033
4. **LEdecomp** R-package https://github.com/davidAtance/LEdecomp

2. Life expectancy decomposition methods



2. Life expectancy decomposition methods

- Life expectancy differences serve as an indicator of inequality between groups, while improvements in life expectancy reflect societal progress.
- Explaining these differences and changes by examining the age (and cause of death) structure of mortality provides essential insights.
- Several methods have been proposed to decompose differences in life expectancy and quantify the impact of changes in mortality rate on life expectancy at birth.
- The decomposition methods have been used to elucidate differences in length of life between genders, between racial groups, between geographical areas, among socioeconomic groups, in years lived with or without disabilities or over time.

2. Life expectancy decomposition methods

Different types of decomposition methods

- **Discrete-time life table functions** $\Rightarrow$ Andreev et al. (1982); Arriaga (1984); Lopez and Ruzicka (1977); Ponnapalli (2005).
- **Continuous time life table functions** $\Rightarrow$ Keyfitz (1968); Pollard (1982, 1988); Vaupel (1986).
- **Generalized algorithms** $\Rightarrow$ Andreev et al. (1982); Horiuchi et al. (2008).
- **Lifetable response experiment** $\Rightarrow$ Pollard (1988); Caswell (2019).
- **Extending the Lee-Carter model** (Lee and Carter, 1992) for cause-specific mortality data $\Rightarrow$ Villegas et al. (2025).

2. Life expectancy decomposition methods

Hypothesis:

The total effect, ${}_n\Delta_x^{2-1}$, of a difference in mortality rates between ages x and $x+n$ on life expectancy at birth between populations 2 and 1 can be expressed as follows:

$$\sum_x {}_n\Delta_x^{2\rightarrow -1} = e_0^2 - e_0^1, \tag{1}$$

where e_0^i denotes life expectancy (LE) at birth in population i with $i = 1, 2$, and refers to the average number of years a newborn will live after birth (Preston et al., 2000).

2. Life expectancy decomposition methods

Decomposition method included in the package:

- **Arriaga** $\Rightarrow \boldsymbol{\Delta}^{2\rightarrow 1} = \mathcal{A}(\mathbf{m}^1, \mathbf{m}^2)$ (Arriaga, 1984).
- **Arriaga symmetrical** $\Rightarrow {}_n\Delta_x = \left({}_n\Delta_x^{2\rightarrow 1} - {}_n\Delta_x^{1\rightarrow 2}\right)/2$ (Arriaga, 1984).
- **Stepwise** is a computational algorithm to decompose differences in a demographic measure $\Rightarrow {}_n\Delta_x = 1/2\left\{{}_n\Delta_x^{2\rightarrow 1}\left(m_x^1, m_x^2\right) -_n \Delta_x^{1\rightarrow 2}\left(m_x^2, m_x^1\right)\right\}$ (Andreev et al., 1982).
- **Horiuchi** quantifies the contribution of individual covariates to differences in a dependent demographic variable, such as LE, between two populations $\Rightarrow$ ${}_n\Delta_x = LE_0^1\left(m_x^1\right) - LE_0^2\left(m_x^2\right)$ (Horiuchi et al., 2008).

3. Sensitivity LE decomposition methods

- The sensitivity method is a perturbation analysis applied to mortality data designed to assess the adaptability of a demographic measure, typically LE, to changes in age-specific mortality rates (Van Raalte and Caswell, 2013).
- Sensitivity analysis quantifies how a small change in a parameter (e.g., a mortality rate at a given age) affects an outcome variable (e.g., LE), allowing for exploration of different scenarios and hence obtaining less obvious results (Keyfitz and Caswell, 2005).
- In contrast to decomposition methods, which provide a retrospective account of differences between two populations, sensitivity analysis provides how each age is affected by the increase/decrease in LE.
- Providing a simple way to respond directly to questions such as: What age-specific mortality improvements would be needed to increase LE by two years over the next decade?

3. 1. Sensitivity analysis for decomposition methods based on discrete time life table functions

A direct way to calculate the sensitivity of LE to changes in age-specific mortality rates can be estimated as follows:

$$s_x = \frac{{}_n\Delta_x}{m_x^2 - m_x^1}, \tag{2}$$

where, m_x^i for $i = 1, 2$ denotes the mortality rates at age x for each respective population i, and ${}_n\Delta_x$ is the age-specific contribution derived from the chosen decomposition method. From equation (2), it follows that:

$${}_n\Delta_x = s_x \cdot \left(m_x^2 - m_x^1\right), \tag{3}$$

where the decomposition of LE will depend on the way ro obtain the sensitivity function s_x. Therefore, depending on the estimation approach for s_x, this expression may yield exact or approximate results $\left({}_n\Delta_x \approx s_x \cdot \left(m_x^2 - m_x^1\right)\right)$

3. 2. Instantaneous sensitivity analysis for decomposition methods based on discrete life table

We propose a refined sensitivity analysis based on symmetric perturbations of average mortality rates. Two perturbed mortality vectors are constructed by applying symmetric adjustments to an average of mortality rates from the two populations considered. Providing two sensitivities of LE estimates, which can be presented as:

$$s_x^{1\to 2} = \frac{{}_n\overline{\Delta}_x^{1\to 2}}{\overline{m}_x^1 - \overline{m}_x^2}, \qquad \text{and} \qquad s_x^{2\to 1} = \frac{{}_n\overline{\Delta}_x^{2\to 1}}{\overline{m}_x^2 - \overline{m}_x^1}, \tag{4}$$

where $\overline{\Delta}_x^{i\to j}$ denotes the decomposition output based on the perturbed mortality vectors for each comparison direction, with i and j equal to one or two in opposite values. An instantaneous sensitivity is computed by averaging both directional results:

$$s_x^{\text{Instantaneous}} = \frac{s_x^{1\to 2} + s_x^{2\to 1}}{2}. \tag{5}$$

3. 3. Direct life-table sensitivity analysis

A direct analytical approach to estimate the sensitivity of LE to changes in age-specific mortality. We adopt the following expression:

$$s_x^{\mathrm{DLF}} = -L_x \cdot [e_x \cdot (1 - a_x) + e_{x+1} \cdot a_x] \,, \tag{6}$$

where: L_x represents the person-years lived of individuals with age x, e_x represents the LE at age x which corresponds to the expectation years of lived, and a_x represents the average person-years of life at age x for those individuals who die at that specific age x

3. 4. Numerical sensitivity analysis optimizing the weights

We propose a numerical sensitivity framework that optimizes the weighting of age-specific mortality rates across two populations, rather than relying on a simple arithmetic mean.

4. Decomposition analysis for all cause-of-death analyis

`LEdecomp()` R functions is designed to apply decomposition or sensitivity analysis. Its basic syntax is as follows:

```
library(LEdecomp)
LEdecomp(mx1, mx2, age, nx, sex1, sex2, method, ···)
```

- **mx1** and **mx2**: Age-structured mortality rates for population 1 and 2.
- **age**: Lower bound of each age group.
- **sex1** and **sex2**: `'m'`,`'f'`, or `'t'`, affects a_0 treatment.
- ...: optional arguments passed to `'numDeriv::grad()'` or other internals.

4. Decomposition analysis for all cause-of-death analyis

`LEdecomp()` R functions is designed to apply decomposition or sensitivity analysis. Its basic syntax is as follows:

LEdecomp(mx1, mx2, age, nx, sex1, sex2, method, $\cdots$)

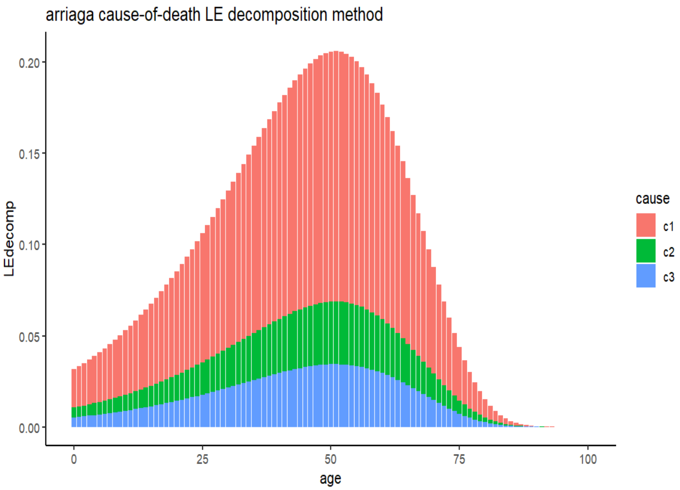
- **method**: `c('lifetable', 'arriaga', 'arriaga_sym', 'sen_arriaga', 'sen_arriaga_sym', 'sen_arriaga_inst', 'sen_arriaga_inst2', 'sen_arriaga_sym_inst', 'sen_arriaga_sym_inst2', 'chandrasekaran_ii', 'sen_chandrasekaran_ii', 'sen_chandrasekaran_ii_inst', 'sen_chandrasekaran_ii_inst2', 'chandrasekaran_iii', 'sen_chandrasekaran_iii', 'sen_chandrasekaran_iii_inst', 'sen_chandrasekaran_iii_inst2', 'lopez_ruzicka', 'lopez_ruzicka_sym', 'sen_lopez_ruzicka', 'sen_lopez_ruzicka_sym', 'sen_lopez_ruzicka_inst', 'sen_lopez_ruzicka_inst2', 'horiuchi', 'stepwise', 'numerical')`

4. Decomposition analysis for all cause-of-death analyis

Practical application in R.
Necessary files:

- 1.LEdecomp_AllCauses.R
- Spain.RData

5. Cause-of-death life expectancy decomposition



5. Cause-of-death life expectancy decomposition

- Search to estimate the contribution of every cause-of-death specific mortality rate to the differences between two life expectancies.
- We present different approaches to estimate the cause-of-death contribution to life expectancy differences.

5. Cause-of-death life expectancy decomposition

General framework for cause-of-death analysis (Arriaga, 1989; Preston et al., 2000) consider that the contribution of cause i to the difference in life expectancy between populations 2 and 1 at age group x to $x+n$, ${}_n\Delta_x^{i,2-1}$, can be written as:

$$ {}_n\Delta_x^{i,2\to 1} = {}_n\Delta_x^{2\to 1} \cdot \frac{{}_nm_x^{i,2} - {}_nm_x^{i,1}}{{}_nm_x^{2} - {}_nm_x^{1}} = s_x^{2\to 1} \cdot \left({}_nm_x^{i,2} - {}_nm_x^{i,1} \right), \tag{7} $$

where ${}_nm_x^{i,j}$ denotes the cause-specific mortality rate for cause i and population j with $j = 1, 2$ between age x and $x+n$.

5. Cause-of-death life expectancy decomposition

Instantaneous approach. We propose incorporating instantaneous sensitivity, into the general framework for decomposing causes of deaths:

$$ {}_n\Delta_x^{i,\text{Arriaga-Inst}} = s_x^{\text{Instantaneous}} \cdot \left({}_n m_x^{i,2} - {}_n m_x^{i,1} \right), \tag{8} $$

where $s_x^{\text{Instantaneous}}$ represents the instantaneous Arriaga sensitivity, which is derived from (5) using the two perturbed populations.

5. Cause-of-death life expectancy decomposition

When conducting a cause-of-death decomposition analysis, the primary distinction in the use of the `LEdecomp()` function is based on the input format of the mortality rates provided by the user. In this example, we provide cause of death mortality rates for US males and females in a matrix format (although a vector format can also be used), where each column corresponds to mortality rates by cause of death.

LEdecomp(mx1, mx2, age, nx, sex1, sex2, method, $\cdots$)

5. Cause-of-death life expectancy decomposition

Practical application in R.
Necessary files:

- 2.LEdecomp_Cause-of-Death.R

6. Conclusions

- The **LEdecomp** R-package seeks to standardize and facilitate the implementation of methods for decomposing life expectancy between two populations, offering practical tools.
- It focuses on classical methods (Arriaga, Stepwise, Horiuchi) and sensitivity analysis, including applications for all causes or cause of death analysis.
- Symmetric Arriaga is the recommended method for comparing sexes or periods within a country: it is accurate, additive, and computationally efficient.
- For analysis by cause of death, the symmetric Arriaga is recommended for its robustness and ability to avoid numerical errors.

7. References

Andreev, E. M. et al. (1982). Metod komponent v analize prodoljitelnosty zjizni.[The method of components in the analysis of length of life]. *Vestnik Statistiki*, 9(3):42–47.

Arriaga, E. E. (1984). Measuring and explaining the change in life expectancies. *Demography*, 21:83–96.

Arriaga, E. E. (1989). *Changing trends in mortality decline during the last decades*. In: Ruzicka L, Wunsch G, Kane P, eds. Differential mortality: Methodological issues and biosocial factors. Clarendon Press.

Caswell, H. (2019). *Sensitivity Analysis: Matrix Methods in Demography and Ecology*, chapter Sensitivity Analysis of Longevity and Life Disparity, pages 45–65. Springer.

Horiuchi, S., Wilmoth, J. R., and Pletcher, S. D. (2008). A decomposition method based on a model of continuous change. *Demography*, 45:785–801.

Keyfitz, N. (1968). Introduction to the Mathematics of Population : With Revisions(Addison-Wesley Series in Behavioral Science: Quantitative Methods, Addison-Wesley). *(No Title)*.

Keyfitz, N. and Caswell, H. (2005). *Applied mathematical demography*, volume 3rd ed. New York: Springer-Verlag.

Lee, R. D. and Carter, L. R. (1992). Modeling and forecasting US sex differentials in mortality. *International Journal of forecasting*, 8(3):393–411.

Lopez, A. D. and Ruzicka, L. T. (1977). The differential mortality of the sexes in australia. in n.d. mcglashan, (ed.) studies in australian mortality, Tasmania: University of Tasmania, environmental studies, occasional paper no. 4 Pollard, j.h. (1982) The expectation of life and its relationship to mortality. *Journal of the Institute of Actuaries*, 109, Part 2(442):225–240.

Pollard, J. H. (1982). The expectation of life and its relationship to mortality. *Journal of the Institute of Actuaries*, 109(2):225–240.

Pollard, J. H. (1988). On the decomposition of changes in expectation of life and differentials in life expectancy. *Demography*, 25(2):265–276.

Ponnapalli, K. M. (2005). A comparison of different methods for decomposition of changes in expectation of life at birth and differentials in life expectancy at birth. *Demographic Research*, 12:141–172.

Preston, S., Heuveline, P., and Guillot, M. (2000). *Demography: measuring and modeling population processes*. Wiley-Blackwell.

Van Raalte, A. A. and Caswell, H. (2013). Perturbation analysis of indices of lifespan variability. *Demography*, 50(5):1615–1640.

Vaupel, J. W. (1986). How change in age–specific mortality affects life expectancy. *Population Studies*, 40(1):147–157.

Villegas, A. M., Bajekal, M., and Haberman, S. (2025). Modelling mortality by cause of death and socio–economic stratification: an analysis of mortality differentials in England. *Decisions in Economics and Finance*, pages 1–39.

Thanks for your attention

Appendix-A: Equation used for Arriaga decomposition method

$$
{}_n\Delta_x^{2\to 1} = \begin{cases} \text{Direct} + \text{Indirect} = \ell_x^1 \cdot \left(\frac{{}_nL_x^2}{\ell_x^2} - \frac{{}_nL_x^1}{\ell_x^1} \right) + T_{x+n}^2 \cdot \left(\frac{\ell_x^1}{\ell_x^2} - \frac{\ell_{x+n}^1}{\ell_{x+n}^2} \right) & \forall x < \omega, \\ \text{Direct} = \ell_x^1 \cdot \left(\frac{{}_nL_x^2}{\ell_x^2} - \frac{{}_nL_x^1}{\ell_x^1} \right) & \forall x = \omega, \end{cases} \tag{9}
$$

where, l_x represents the number of people aged x who last birthday on January 1 of period t, ${}_nL_x$ represents person-years lived between ages x and $x+n$, T_x denotes the total number of expected years to live from age x to the maximum age achievable (ω) of the population in the life table in the year t and the subscript denotes the population. The superscripts 1 and 2 correspond to the specific population considered.