LEDecomp. Life Expectancy Decomposition R package

Tim Riffe[a,b], **David Atance**[c,*], **Josep Lledó**[d]
[a]University of the Basque Country, Spain
[c]Universidad de Alcalá. Spain
[d]Universitat de València, Spain
[*]Corresponding author, david.atance@uah.es.

Insurance Data Science Conference, Hanover, June 2026

Insurance
Data
Science

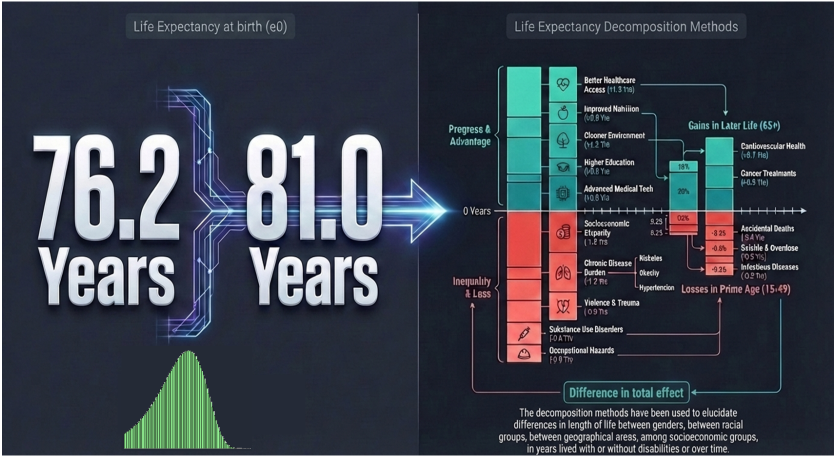
Universidad de Alcalá

Index

1. Introduction.

- Life expectancy (LE) at birth indicates the average years a cohort of newborns is expected to live. This indicator captures the cumulative social, economic, medical, and technological achievements of society, among others.
- LE differs among genders or even varies between periods in the same population.
- **LE differences represent** inequality between group of populations.
- Explaining these differences and changes by examining the age (and cause of death) structure of mortality gives a more detailed understanding of aggregate differences ⇒ **LE Decomposition Methods**

1. Introduction.



1. Introduction.

- There are many different decomposition methods that are similar, but not always identical.
- These methods are often self-made, shared among colleagues, unmaintained, and may differ in how they handle details such as **age intervals**, **or the open age group**, **symmetric or not**.
- We found it opportune to collect and standardize the main decomposition options, proposing several new methods, **36** in all, in the **LEdecomp R package**, and a complementary Excel tool.

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:

$$\sum_x {}_n\Delta_x^{2\rightarrow -1} = e_0^2 - e_0^1, \quad (1)$$

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

- **Direct lifetable methods**: Direct analytic lifetable functions (Andreev, 1982; Arriaga, 1984; Lopez and Ruzicka, 1977; Ponnapalli, 2005).
- **Sensitivity-based approximations**: direct lifetable-based sensitivity or numerical gradient functions (Keyfitz, 1968; Pollard, 1982; Mitra, 1978; Demetrius, 1979; Goldman and Lord, 1986; Vaupel, 1986; Pollard, 1988).
- **Generalized algorithms**: algorithms applied to decompose differences in LE, recalculating LE a large number of times (Vaupel, 1986; Pollard, 1988).

`LEdecomp` R package includes all these kind of methods.

3. Cause-of-death life expectancy decomposition

- To estimate the contribution of every cause-of-death specific mortality rate to the differences between two LE. The case of deaths is classified by a single underlying cause, c which implies: $m_x = \sum_c m_{x_c}$
- We present different approaches to estimate the cause-of-death contribution to life expectancy differences:
  1. **Split all-cause decomposition results by proportional differences.**
  2. **Multiply cause-specific differences into the all-cause sensitivity function.**
  3. **Generalized decomposition methods.**

4. LEdecomp example usage

`LEdecomp()` R functions is designed to apply decomposition or sensitivity analysis for all-causes or cause-of-death analysis. The syntax is as follows:

```
install.packages('LEdecomp')
library(LEdecomp)
LEdecomp(mx1, mx2, age, nx, sex1, sex2, method, ···)
```

- **mx1** and **mx2**: age-specific mortality rates for populations 1 and 2, respectively. These can be vectors, or data frames (with a column labeled mx)..
- **age**: a numeric vector of the unique age groups corresponding to mx1 and mx2.
- **nx**: a numeric vector specifying the width of each age interval, inferred from age if not given.
- **sex1** and **sex2**: `'m'`,`'f'`, or `'t'`, affects a_0 treatment.
- ...: optional arguments passed to `'numDeriv::grad()'` or other internals.

4. LEdecomp example usage

`LEdecomp()` R functions is designed to apply decomposition or sensitivity analysis for all-causes or cause-of-death analysis. The syntax is as follows:

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

- **available_methods()**

```
> available_methods()
 [1] "lifetable"                    "arriaga"                    "arriaga_sym"
 [4] "sen_arriaga"                  "sen_arriaga_sym"            "sen_arriaga_inst"
 [7] "sen_arriaga_inst2"            "sen_arriaga_sym_inst"       "sen_arriaga_sym_inst2"
[10] "andreev"                      "andreev_sym"                "sen_andreev"
[13] "sen_andreev_sym"              "sen_andreev_inst"           "sen_andreev_inst2"
[16] "sen_andreev_sym_inst"         "sen_andreev_sym_inst2"      "chandrasekaran_ii"
[19] "sen_chandrasekaran_ii"        "sen_chandrasekaran_ii_inst" "sen_chandrasekaran_ii_inst2"
[22] "chandrasekaran_iii"           "sen_chandrasekaran_iii"     "sen_chandrasekaran_iii_inst"
[25] "sen_chandrasekaran_iii_inst2" "lopez_ruzicka"              "lopez_ruzicka_sym"
[28] "sen_lopez_ruzicka"            "sen_lopez_ruzicka_sym"      "sen_lopez_ruzicka_inst"
[31] "sen_lopez_ruzicka_inst2"      "sen_lopez_ruzicka_sym_inst" "sen_lopez_ruzicka_sym_inst2"
[34] "numerical"                    "stepwise"                   "horiuchi"
```

5.1 Data

Data

- US mortality data from Human Mortality Database (2026).
- Data consists on mortality rates by year **2010, 2015 and 2020**, and age range **0-100**.
- Also, we use cause-of-death mortality data from Centers for Disease Control and Prevention (2024) which are categorized in **18 causes** based on International Classification of Diseases (ICD10).

5.2 All-cause life expectancy decomposition

```
LE_2010 <- LEdecomp(mx1 = mx_male_US2010, mx2 = mx_female_US2010,
            age = c(0:100), nx = 1, sex1 = "m", sex2 = "f",
            method = "arriaga_sym")
```

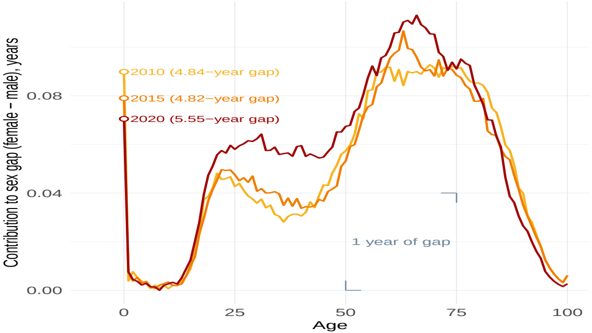
The output of the decomposition analysis returns an object (LE_2010) of class LEdecomp including:

- mx1, mx2, age, nx, sex1, sex2, and method. Input vlaues.
- LE1 and LE2: the total LE calculated from mx1 and mx2, respectively.
- LEdecomp: a vector or matrix of age-specific contributions to the difference in LE, structured consistently with mx1 and mx2.

5.2 All-cause life expectancy decomposition

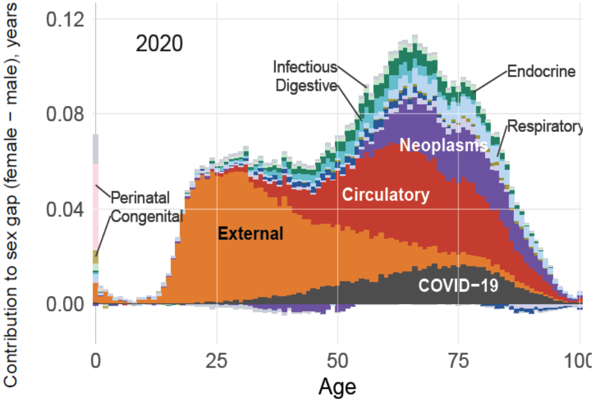
$e_{0,2010}^{\text{male}} = 76.2427$ and $e_{0,2010}^{\text{female}} = 81.0796.$
$e_{0,2015}^{\text{male}} = 76.3266$ and $e_{0,2015}^{\text{female}} = 81.1461.$
$e_{0,2020}^{\text{male}} = 74.3574$ and $e_{0,2020}^{\text{female}} = 79.9053.$

Figure: Symmetrical Arriaga decomposition of life expectancy sex gap (females - males) in the USA in 2010, 2015, and 2020.



5.3 Decomposition analysis by causes of death

Figure: Arriaga symmetric (instantaneous sensitivity) cause-of-death decomposition of LE between males and females in the US in 2020.



5.3 Decomposition analysis by causes of death

Cause (Name)	Cause ID	2020
Infectious	A00-B99	0.0869
Neoplasms	C00-D48	0.5384
Blood & immune	D50-D89	0.0065
Endocrine & metabolic	E00-E88	0.2414
Mental	F01-F99	0.0799
Nervous system	G00-G98	0.0641
Circulatory	I00-I99	1.4675
Respiratory	J00-J98	0.2216
Digestive	K00-K92	0.2038
Skin	L00-L98	0.0013
Musculoskeletal	M00-M99	-0.0102
Genitourinary	N00-N98	0.0632
Maternal	O00-O99	-0.0257
Perinatal	P00-P96	0.0364
Congenital	Q00-Q99	0.0117
Symptoms & ill-defined conditions	R00-R99	0.0623
Special codes	U00-U99	0.5978
External causes	V01-Y89	1.8870

Table: Cause-specific contributions for 2020.

5. Conclusions

- The **LEdecomp** R-package seeks to standardize and facilitate the implementation of methods for decomposing life expectancy between two populations, offering practical tools.
- All cause of death or causes-of-death analysis are included in the R-package.
- Symmetric Arriaga is the recommended method for all cause of death analysis due to its properties.
- For analysis by cause of death, the numerial or lifetable methods are recommended for its robustness and ability to avoid numerical errors.

6. References

Andreev, E. M. (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.

Centers for Disease Control and Prevention (2024). CDC WONDER Online Database. Accessed via CDC WONDER. Originally released April 1995. Accessed 2024-XX-XX.

Demetrius, L. (1979). Relations between demographic parameters. *Demography*, 16(2):329–338.

Goldman, N. and Lord, G. (1986). A new look at entropy and the life table. *Demography*, 23(2):275–282.

Human Mortality Database (2026). University of california, berkeley (usa), and max planck institute for demographic research (germany). Avaliable at www.mortality.org and www.humanmortality.de (accesed on 05/02/2026).

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

Lopez, A. D. and Ruzicka, L. T. (1977). The differential mortality of the sexes in australia. In McGlashan, N. D., editor, *Studies in Australian Mortality*, volume 4 of *Environmental Studies Occasional Paper*, pages 89–94. University of Tasmania, Hobart, Tasmania. Occasional Paper No. 4.

Mitra, S. (1978). A short note on the Taeuber paradox. *Demography*, 15(4):621–623.

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.

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

Thanks for your attention

We acknowledge the support of *Generalitat Valenciana* through projects CIGE/2023/7 (*Conselleria de Educación, Universidades y Empleo*), and PIUAH25/CSJ-050 funded by Universidad de Alcalá.

SCAN ME

Appendix-A: Methods summary

Figure: Empirically distinct decomposition classes implemented in LEdecomp. Although the package gives 36 variants, many formulations yield identical age-specific contributions and are therefore grouped

Empirical decomposition	Description	Symm.	Add.	LEdecomp method(s)
Asymmetrical lifetable decomp.	Direction-dependent direct lifetable decomposition	No	Yes	Includes "arriaga", "andreev", "lopez_ruzicka", and their indirect sensitivities.
Symmetrical lifetable decomp.	Average of directional decompositions	Yes	Yes	Includes "chandrasekaran_ii", "chandrasekaran_iii", all "*_sym" variants, and their indirect sensitivities.
Instantaneous implied sensitivities	Indirect sensitivities evaluated around optimized $\overline{m}$	Yes	Yes*	Includes e.g., "sen_andreev_inst", and all "*_inst" and "*_inst2" variants.
Discrete lifetable sensitivity	Derivative of e_0 with respect to m_x evaluated at an intermediate schedule	Yes	Yes*	"lifetable" and "numerical".
Stepwise replacement	Ordered cumulative replacement of rates between mortality schedules	Yes†	Yes	"stepwise".
Pseudocontinuous (Horiuchi)	Path integral along a linear interpolation between schedules	Yes	Asymp.	"horiuchi".

* Additivity is guaranteed if optimized with default arg opt = TRUE.
† Symmetry is guaranteed if default arg symmetry = TRUE, and results match those of symmetrical lifetable decomposition (row 2) if direction = "both" as well.

Appendix-B: Cause-of-death LE decomposition equations

- **Split all-cause decomposition results by proportional differences.**

$$\Delta_x = \Delta_{xc}\frac{\delta_{xc}}{\delta_x} = \Delta_x \frac{m_{xc}^2 - m_{xc}^1}{m_x^2 - m_x^1} \tag{2}$$

- **Multiply cause-specific differences into the all-cause sensitivity function.**

$$\Delta_{xc} = s_x \delta_{xc} = s_x \left(m_{xc}^2 - m_{xc}^1\right) \tag{3}$$

- **Generalized decomposition methods:** tackle cause-of death decomposition using brute force, i.e. by increasing the number of parameters to handle.