Arriaga meets Kitagawa. life expectancy decompositions including population subgroups

Dr. Timothy Riffe[1,2,3*]

[1*]Department of Sociology and Social Work, University of the Basque Country (UPV/EHU), Leioa, Spain, 48940, Bizkaia, Barrio Sarriena s/n.
[2]Ikerbasque (Basque Foundation for Science).
[3]Laboratory of Population Health, Max Planck Institute for Demographic Research.

Corresponding author(s). E-mail(s): tim.riffe@ehu.eus;

Abstract

Background: An Arriaga (1984) decomposition allows us to decompose differences in life expectancy into contributions due to mortality rate differences in each age. A Kitawaga (1955) decomposition allows us to decompose differences in a weighted mean into an effect from differences in structure and effects from differences in each element of the weighted value.
Objective: Sometimes we would like to decompose a difference between two populations that are each composed of like-defined subpopulations. Said decomposition would produce effects for each age of each subpopulation, as well as a marginal effect of composition differences. I propose a straightforward analytic method to do this.
Methods: In short, within-subpopulation life expectancy differences can be handled by the Arriaga method. The case of weighting together the life expectancies of subpopulations by way of a composed radix can be handled using the well-known Kitagawa method. The elements of the value component of the Kitagawa method tell us how to rescale the Arriaga results specific to each subpopulation.
Results: I show the proposed analytic method to be equivalent to a Horiuchi et al (2008) reframing of the same problem. Some mentionable properties: (i) There is no limit to the number of subpopulations, (ii) it is straightforward to incorporate cause-of-death information, (iii) composition is here only considered in the radix age. I currently have results for simulated mortality rates, but I promise to wrangle up an empirical application to demonstrate the method.
Conclusions: The analytic decomposition method I propose is advantageous compared to a Horiuchi method for this problem purely for reasons of computational efficiency. This method could help further disentangle the effects of

mortality and composition differences in explaining or clarifying paradoxes or secular change. I promise to think further about subpopulation weighting that might occur over all ages (using prevalence information).}

Keywords: Decomposition, Mortality, Cause of death, Population Structure, Mortality Inequalities

1 Introduction

An Arriaga (1984) decomposition allows us to decompose changes in life expectancy into contributions due to mortality changes in different ages. The method was designed to be practical, and framed in terms of age lifetable columns expressed in discrete age groups. A well-known property of the method is that mortality changes in different ages need not be proportional. Derived contributions also sum exactly to the observed life expectancy difference. A not-well-known property of the method is that it is asymmetrical, in the sense that the absolute values of age-specific contributions depend on whether we compare population 1 with 2, or population 2 with 1. A further property of the method is that it is designed to work with homogeneous populations, meaning that populations 1 and 2 are each homogeneous, in the sense that each is expressed with only one lifetable.

A Kitagawa (1955) decomposition allows us to decompose differences in a weighted mean due to differences in weights (structure) and differences in the value being weighted (often rates, but in our case life expectancies). This widely-used decomposition method is well-known to be exact in that the resulting structure and value components sum exactly to the observed difference in weighted means. The individual elements (ages, or life expectancies for us) of the value being weighted have identifiable effects. It is not well-known that the individual elements of the structure component do not have identifiable effects. Rather, the structure effects should be summed to a a marginal effect due to differences in structure.

Sometimes we have a situation where we would like to decompose a difference between two populations that are each composed of like-defined subpopulations. For example, life expectancy in France versus Spain, each with education-specific subpopulations. In such situations, a decomposition should tell us the contribution to the difference in overall life expectancy due to rate differences in each age in each subpopulation, and also separate an effect due to overall compositional change. In this paper, I would like to propose a straightforward analytic method to decompose in this way. In short, within-subpopulation life expectancy differences can be handled by the Arriaga method. The case of weighting together the life expectancies of subpopulations by way of a composed radix can be handled using the well-known Kitagawa method. The elements of the value component of the Kitagawa method tell us how to rescale the Arriaga results specific to each subpopulation so as to isolate the age-subpopulation-specific effects on the overall life expectancy difference. I will justify this rescaling, and show that the results of this procedure are fully consistent with the results of a Horiuchi, Wilmoth, and Pletcher (2008) reframing of the problem.

2 Method

2.1 Notation

Arriaga decompositions can be implemented in a number of relevant ways, per Riffe et. al. (2024), from which we here consider only *symmetrical* Arriaga decomposition. This implies following the original formulas twice, once in each direction, and averaging the sign-adjusted results. For this, we use the following lifetable columns:

- $\ell(a)$ lifetable survivorship at exact age a.
- ${}_nL_a$ lifetable person years lived in the interval $[a, a+n)$.
- T_a total lifetable person years lived beyond age a.
- e_a remaining life expectancy at exact age a.

We also use the superscript s to index subpopulations comprising the total population, and the superscript t to index time points, or some other way of differentiating the total population. The index $e_0^{s,t}$ reads as "life expectancy at birth for subpopulation s at time t", e.g. $e_0^{low,1990}$ could be the life expectancy in 1990 of a low education group. The first stage of our exercise is to decompose subgroup-specific differences in life expectancy (hence the ^{s} on everything):

$$\begin{aligned}\Delta^s &= e_0^{s,2} - e_0^{s,1} \\ &= \sum_x^{\omega} {}_n\Delta_x^s \quad ,\end{aligned} \tag{1}$$

where ω is the highest (open) age group, and Δ^s is the subgroup-specific (s) difference in life expectancy being decomposed, and it is composed of age-specific contributions, ${}_n\Delta_x^s$, which can be calculated following Arriaga's decomposition method following eq (2). I'll use a lifetable radix of 1, meaning $\ell_0 = 1$ to reduce the formula a bit.

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

Equation (2) is the first pass of our symmetrical decomposition, where we move *forward* in time from $t = 1$ to $t = 2$, so please permit . It

Supplementary information. If your article has accompanying supplementary file/s please state so here.

Acknowledgments. Acknowledgments are not compulsory. Where included they should be brief. Grant or contribution numbers may be acknowledged.

Please refer to Journal-level guidance for any specific requirements.

Declarations

Some journals require declarations to be submitted in a standardised format. Please check the Instructions for Authors of the journal to which you are submitting to see if you need to complete this section. If yes, your manuscript must contain the following sections under the heading 'Declarations':

- Funding
- Conflict of interest/Competing interests (check journal-specific guidelines for which heading to use)
- Ethics approval
- Consent to participate
- Consent for publication
- Availability of data and materials
- Code availability
- Authors' contributions

If any of the sections are not relevant to your manuscript, please include the heading and write 'Not applicable' for that section.

Appendix A Section title of first appendix

An appendix contains supplementary information that is not an essential part of the text itself but which may be helpful in providing a more comprehensive understanding of the research problem or it is information that is too cumbersome to be included in the body of the paper.

For submissions to Nature Portfolio Journals please use the heading "Extended Data".

References

Arriaga, E.E. (1984). Measuring and explaining the change in life expectancies. *Demography*, *21*, 83–96, https://doi.org/10.2307/2061029

Horiuchi, S., Wilmoth, J.R., Pletcher, S.D. (2008). A decomposition method based on a model of continuous change. *Demography*, *45*, 785–801, https://doi.org/10.1353/dem.0.0033

Kitagawa, E.M. (1955). Components of a difference between two rates. *Journal of the american statistical association*, *50*(272), 1168–1194, https://doi.org/10.1080/01621459.1955.10501299