

Life Expectancy Decompositions

ECO 572/ Soc 532 • SPRING 2017

Section 3.10 in the textbook considers the decomposition of a change in life expectancy in terms of the contribution of each age group. At issue are questions such as “How much of the gain in life expectancy in a recent period can be attributed to reductions in infant and child mortality?” We can answer the question in continuous or discrete time. The textbook focuses on the discrete method and gives a formula and example. We’ll briefly review both, focusing on expectation of life at birth.

Pollard

Pollard (1982) proposed a continuous-time decomposition. Recall that expectation of life at age a is the ratio of time lived after age a to the number of survivors to that age. Specifically, at birth we have

$$e(0) = \frac{1}{l(0)} \int_0^{\infty} l(x) dx = \int_0^{\infty} e^{-M(x)} dx$$

where $M(x) = \int_0^x \mu(a) da$ is the cumulative force of mortality up to age x . The difference in life expectancy between two time periods (or two countries) may be written as

$$e_2(0) - e_1(0) = \int_0^{\infty} (e^{-M_2(x)} - e^{-M_1(x)}) dx$$

Let us factor out the initial survival $e^{-M_1(x)} = l_1(x)/l_1(0)$ to obtain

$$e_2(0) - e_1(0) = \frac{1}{l_1(0)} \int_0^{\infty} (e^{M_1(x)-M_2(x)} - 1) l_1(x) dx$$

At this point we can integrate by parts using the fact that $l_1(x)$ is the derivative of $-\int_x^{\infty} l_1(a) da = -l_1(x)e_1(x)$, to obtain the main result

$$e_2(0) - e_1(0) = \frac{1}{l_1(0)} \int_0^{\infty} (\mu_1(x) - \mu_2(x)) e^{M_1(x)-M_2(x)} l_1(x) e_1(x) dx$$

Pollard notes that expanding the exponential in a Taylor series and taking just the leading term (which is 1) leads to the “well-known” approximation $\Delta e(0) \approx \Delta \mu(x) e_1(x) l_1(x) / l_1(0)$. In words, a small change in death rates at age x changes life expectancy at birth by the expectation of life remaining at x times the probability of surviving to that age.

Noting that $e^{-M_1(x)} = l_1(x)/l_1(0)$ we can do some cancellation to obtain the simpler formula

$$e_2(0) - e_1(0) = \int_0^{\infty} (\mu_1(x) - \mu_2(x)) \frac{l_2(x)}{l_2(0)} e_1(x) dx$$

Which combines survival probabilities in the new regime ($l^2(x)$) with life expectancy in the old ($e_1(x)$). Reversing the labels leads to an equivalent expression in terms of $l_1(x)$ and $e_2(x)$. Both exact.

This method is largely of theoretical interest because to apply it we need to evaluate the integrals, which requires approximations. (See Pollard's paper if you are interested.)

Arriaga

Arriaga (1988) proposed a discrete-time decomposition that is much easier to apply to conventional abridged life tables. We consider the contribution of a change in mortality rates at ages x to $x + n$ on life expectancy at age $a < x$. We focus here on life expectancy at birth, so $a = 0$. For consistency with the textbook I'll use superscripts for the two time periods or countries.

I find that it helps follow the argument to consider the average person-years lived at ages x to $x + n$, which Arriaga calls a "temporary" life expectancy and denotes ${}_n e_x L_x / l_x$. Changing mortality at ages x to $x + n$ has an affect at those ages and as we'll see, also an effect at later ages.

The first component, sometimes called the *direct* effect, reflects the fact that in the new regime people spend on average ${}_n e_x^2$ years at those ages instead of ${}_n e_x^1$, provided of course they make it to age x , so this first component is

$$\frac{l_x^1}{l_0^1} ({}_n e_x^2 - {}_n e_x^1) = \frac{l_x^1}{l_0^1} \left(\frac{{}_n L_x^2}{l_x^2} - \frac{{}_n L_x^1}{l_x^1} \right)$$

This is the first part of equation (3.11) in the textbook.

The second component reflects the fact that we now have more people coming out of the age group x to $x + n$. In the first regime we had l_{x+n}^1 exiting, but we now have $l_x^1 l_{x+n}^2 / l_x^2$ exiting. (It may help to think of the last ratio as the conditional probability of surviving from x to $x + n$ in the second regime.) The additional survivors represent more person-years at later ages even if the rates at those ages don't change and still average ${}_{\infty} e_{x+n}^1$ years. But of course the rates themselves have changed, and they will average ${}_{\infty} e_{x+n}^2$ years instead. The second component is then

$$\frac{1}{l_0^1} \left(l_x^1 \frac{l_{x+n}^2}{l_x^2} - l_{x+n}^1 \right) {}_{\infty} e_{x+n}^2 = \frac{T_{x+n}^2}{l_0^1} \left(\frac{l_x^1}{l_x^2} - \frac{l_{x+n}^1}{l_{x+n}^2} \right)$$

This is the second part of (3.11) in the textbook.

Arriaga further splits this term into an *indirect* effect attributable to the additional survivors at old rates, and an *interaction* effect due to the fact that those survivors face new rates. We will not distinguish these, but if you are interested the indirect effect is easily

computed using ${}_{\infty}e_{x+n}^1$ instead of ${}_{\infty}e_{x+n}^2$ in the above formula and the interaction is the difference between the two.

For the last open-ended age group there is only a direct effect, computed as

$$\frac{l_x^1}{l_0^1} \left(\frac{T_x^2}{l_x^2} - \frac{T_x^1}{l_x^1} \right)$$

This is expression (3.12) in the textbook, but it is really just a special case of the first formula above because ${}_{\infty}L_x = T_x$ for the open-ended age group.

If you apply these formulas to the example in the textbook make sure you use ${}_nL_x^2$ and T_x as printed, because accumulating person-years by age gives slightly different results, probably because of rounding. The data needed are available in file `box34.dat` in my website.

Pollard (1988) shows that his continuous-time formulation and Arriaga's discrete-time analysis are exactly equivalent in the limit when one uses finer and finer age intervals, with Pollard's equation corresponding to the sum of Arriaga's direct, indirect and interaction effects.

Keyfitz

Keyfitz considered the effects of absolute and relative changes in mortality at every age, and you'll find a nice writeup in Keyfitz and Caswell (2005, Section 4.3).

They show that if the rates change from $\mu(x)$ to $\mu(x) + \delta$, then the derivative of life expectancy w.r.t. δ evaluated at zero is $-\bar{x}e_0$ where \bar{x} is the mean age in the stationary population. For example if life expectancy is 70 and mean age is 35, reducing all rates by 0.001 would increase life expectancy by 2.45 years.

An alternative scenario posits a proportionate change in age-specific rates, where the force of mortality goes from $\mu(x)$ to $\mu(x)(1 + \delta)$. In this case the survival probability is raised to the power $(1 + \delta)$, and the resulting integral is hard to evaluate except in special cases. They show, however, that the derivative of the log of life expectancy w.r.t. δ can be written approximately as the product $-H\delta$, where

$$H = - \frac{\int \log\left(\frac{l(x)}{l(0)}\right) l(x) dx}{\int l(x) dx}$$

is a measure known as *entropy*, defined as the negative weighted average of $\log l(x)/l(0)$ using $l(x)$ as the weights. If everyone lived to age ω and then died $l(x)/l(0)$ would be one and its log zero, so $H = 0$. At the other extreme, if the force of mortality is constant we have an exponential distribution with $l(x)/l(0) = e^{-\mu x}$, for which $H = 1$.

The product $-H\delta$ estimates the relative change in life expectancy after a proportionate change in age-specific mortality. For U.S. females in 2013 entropy was around 0.134, so a ten percent decline in mortality at all ages would increase life expectancy by about 1.34%

or 1.09 years. A quick calculation using a single year life table and reducing rates by 10% yields an actual increase in life expectancy of 1.41 years.

These results are useful because they provide some insight into the relationship between death rates and life expectancy, but they are not terribly realistic because they apply to small absolute or relative changes at all ages. However, the decomposition procedures discussed above have very wide applicability and happen to be exact.

Revised 2/27/2017