

A note on independant time measures

Tim Riffe^{*1} and Joel Cohen²

¹Max-Planck-Institute for Demographic Research

²The Rockefeller University

March 2, 2018

Abstract

Background There are countless Lexis-like relationships between linearly dependant time measures, and these can be combined into higher-order Lexis identities with dense linear interdependencies. One such relationship is the demographic time identity between age, period, cohort, time to death, length of life, and time of death. Certain subsets of time measures in this and other higher order identities consist in time measures that are independant of one another.

Objective We aim to describe the relationship between independent time measures and the identities within which they are nested, with special attention to the demographic time identity and its three independant time dyads. We aim to determine whether data structured on such time dyads might be useful in demographic research.

Data and Methods We illustrate concepts visually based on data from the Colonial Québec Population Register and from the US Health and Retirement study.

Results We show that the Cartesian mappings of independant time dyads form a redundant mapping of the same demographic timespace. These add three previously undescribed demographic perspectives to the four Lexis-like perspectives, four a total of seven proposed standard perspectives.

Conclusions We show independant sets of time measures to be consistent with the temporal identities within which they are nested. We demonstrate that these perspectives can be useful for pattern detection and characterization in demographic research.

This isn't even a first draft! Literally it contains copied and pasted email conversations at this point. It needs a deep revision before even becoming a manuscript stub. So please read w caution.

1 Data

We illustrate concepts using two data sources: the *le registre de la population du Québec ancien* (RPQA) (Desjardins 1999), and RAND version P of the United States Health and

^{*}riffe@demogr.mpg.de

Retirement Study (HRS) (HRS 2017, RAND 2018). The RPQA data are used to illustrate perspectives on population structure, whereas the HRS data are used to illustrate how independent time perspectives relate can be used to take cross-sections of phenomena with multiple time dimensions.

1.1 RPQA

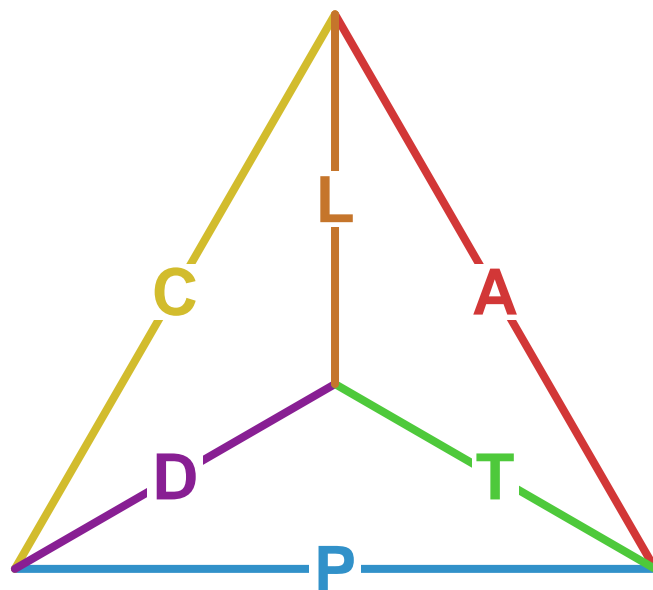
1.2 HRS

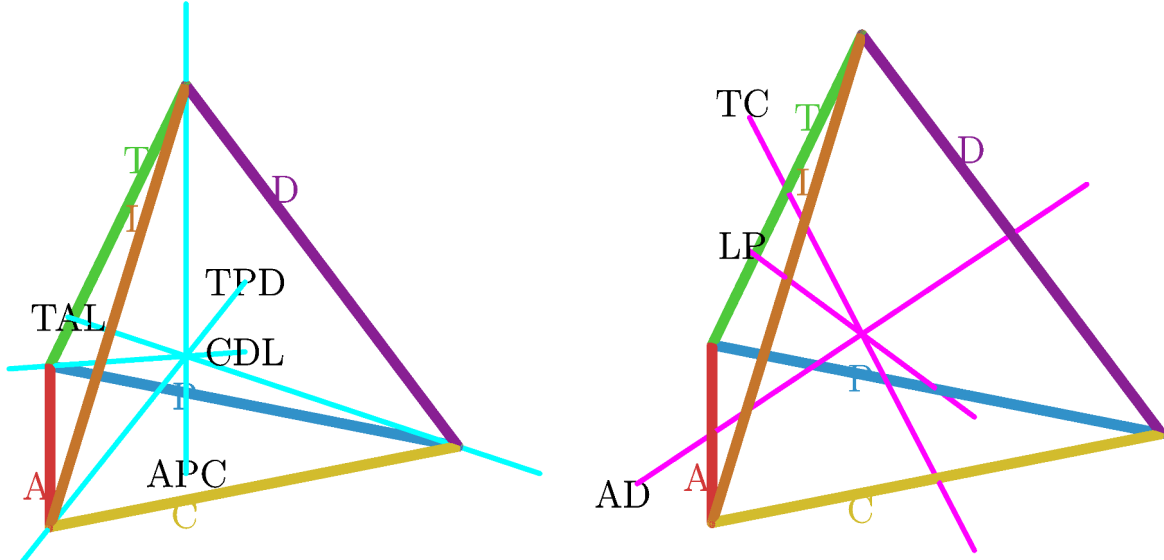
We select a single phenomenon to visualize based on information in the HRS. We follow the data smoothing steps described by Riffe et al. (2017) to prepare the data, and summarize these here.

2 Geometric relationships

The demographic time identity maps to the edges of a tetrahedron. When extended into three dimensions, the tetrahedron tessellates with octahedra in a 2:1 ratio. In its original description (Riffe et al. 2017), neither independent pairs of time measures nor the significance of octahedra in this geometric construct were considered in any depth. Figure 1 shows the tetrahedral graph of the demographic time identity between age (A), period (P), cohort (C), time to death (T), death cohort (D), and length of life (L). Independent dyads are those that do not share any vertex in the graph, and these appear already as perpendicular but disjoint in the present graph representation: CT, LP, and AD.

Figure 1: Tetrahedral graph of demographic time hexad identity, with edges labelled by the six time indices. Reproduced from Riffe et al. (2017).





(a) The four view axes of the triad dependencies map to the medians of the tetrahedron. (b) The three view axes of independent dyads map to the bimedians of the tetrahedron.

Figure 2: The seven principal view axes of the demographic time identity. Figures are interactive if viewed in a recent version of Adobe Reader. Click each figure to activate. A mouse wheel (or right click and drag) can be used to zoom. Left click and drag to rotate.

As with any arbitrary pair of quantitative variables, independent temporal dyads can define the axes of a standard Cartesian plane. Such planes are already an isometric projection, and these planes can also represent cross-sections of the same 3d demographic timespace previously defined (Riffe et al. 2017). The independent dyads can also be conceived of as axes of view on the demographic time space.

2.1 View axes

An angle of view directly at any of the four faces of the demographic time tetrahedron reveals the plane-perspective of a Lexis-like identities. In other words, the viewing axis is set as *normal* to one of the faces of the tetrahedron. Interactive figure 2 may help visualize the notion of viewing axes—the view on a plane is orthogonal if an axis line is reduced to appear as a point in the centroid of the tetrahedron. The view axes of the four triad identities map to the four medians (axes of 3-fold symmetry) of the tetrahedron, depicted in Figure 2a with cyan lines. Independent dyad planes are revealed when the viewing angle is such that the midpoints of opposite tetrahedral edges are aligned. There are three pairs of opposite edges (LP, TC, AD), and therefore three viewing angles of this kind, which map to the bimedians of the tetrahedron (axes of 2-fold symmetry), depicted in Figure 2b with magenta lines.

This is all to say, choosing to view the timespace using the coordinates of some independent dyad simply implies a different transection of the space. The three independent symmetry lines of Figure 2b are also orthogonal to the faces of the cube dual of any octahedron in the tetrahedral-octahedral timespace. Said cubes also form a space-filling tessellation, and are therefore sufficiently and redundantly define the time space.

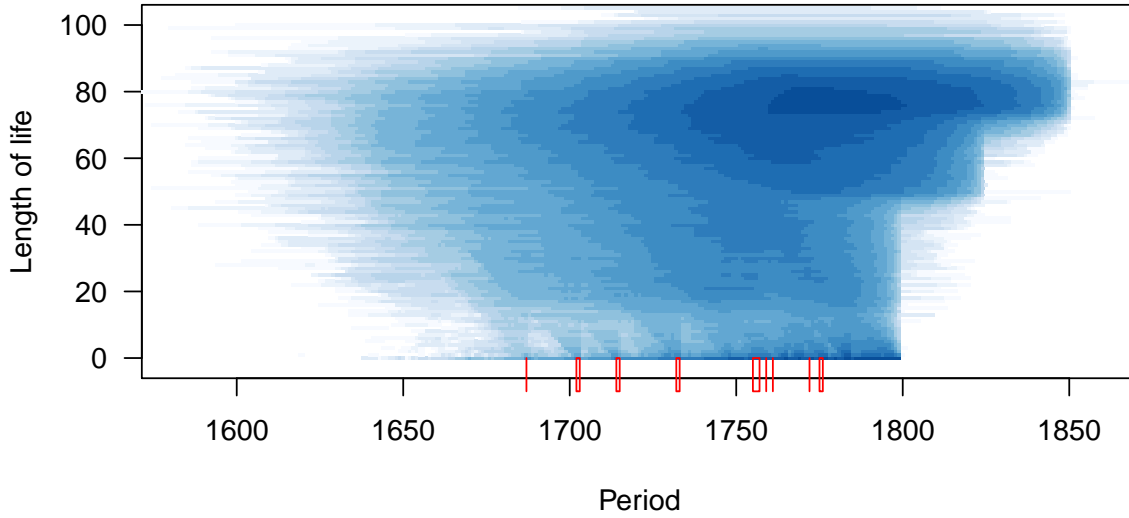


Figure 3: Population structure by length of life and year. Logged counts

2.2 Life lines

The planes defined by any of the triad identities have a clear relationship with life course, either looking back to birth (APC), forward to death (TPD), looking in both directions symmetrically (TAL), or compressed to a whole (LCD). Unambiguous orientation of this kind is a key ingredient to make any sort of inference about variation over the life course. Time planes defined by independent time measures also have a clear relationship with lifelines, but

3 Independant time surfaces with heterogeneity

Here use example of Quebec surfaces with interpretation.

3.1 LP: Lifespan and Period

Subsection structure: orientation of lifelines, definition of space. Expectations a priori about how mortality and fertility should influence shape of surface. show picture of Quebec exposure, discuss peculiarities,

Here for each P , we ask how many people are alive that will one day die at age L . The people with $L=100$ in 2016 have a distribution of calendar years of birth (birth cohorts) along the P axis, and an identically shaped distribution, located L years later along the P axis, of years of death. If mortality is improving, this distribution should have more probability mass toward the right side of the distribution (later years of birth, later years of death) and less toward the left. Conversely, the number of people with $L=100$ is a function of period. This number should be increasing with period if mortality is improving. The distribution of the prevalence of $L=100$ as a function of period could be an interesting indicator of mortality trends, and likewise for other values of L .

In this orientation, an individual born in 1700 that will eventually die at age 80 contributes 1 to the count at age 80 from 1700 all the way to 1780 (or so), so the individual increments a horizontal stripe to the surface. Infant deaths are very short segments in the bottom, whereas long lives are maximally overlapped. It would appear births were not collected after 1800, and some other data collection milestone happens in 1825. The triangles at the bottom look like artifacts at first, but actually they are mortality shocks (wow!). Take a given triangle: each is right-aligned on moment of death (otherwise a given lifeline would be higher up in y and extend further to the right in x). These shocks likely extend up into higher ages, but they get blurred out by higher baseline mortality in higher ages. I guess growth happens in the contours for the most part because the population was growing exponentially.

What are the boundaries on the extreme right, and why do some lifelines extend beyond the main boundaries? In addition, if you look at the triangle with right vertical axis in 1705 or so, and left oblique edge of slope -1 extending upward and to the left at 45° from horizontal axis, that oblique image seems to extend very far up, detectably as far up as life length 70. There is a parallel shift in density to the left of that oblique line space by about 10-15 years. I am wondering if we are looking at an artifact of some kind of grouping of the data or if there really was a crisis in mortality around 1690, 1705, 1720, 1740, 1750. If you have independent estimates of death rates in these years, it would probably be easy to compare those estimates with these apparent events.

Death info here comes from burial records, which had some age constraints in data collection: (1800-1825: deaths of 50+ only; 1826-1850: deaths of 75+ only). That explains the block edges on the right of LP, and the angle edges in TC. I'm also not very sure what accounts for the exceptions to these edges, but my guess would be that it owes to there being multiple potential sources for date of death. Here it is again, with known epidemics (i.e. main disease can be guessed) marked with red rectangles (if spanning more than 1 year) (cf Figure 1.1 Mazan 2011) Typhus, measles, smallpox, etc) were cross checked with other historical documents. These match the right edge of most of the triangles, and the others may have just been lesser or unstudied epidemics, or ones that I just missed. The left oblique edge of each triangle is in general more visible than the vertical right edge, at least in the early part of the series, because growth in the early phase (pre 1663 if I recall) was largely from migration.

Missingness in LP: Data collection began in 1621, but lifelines in the LP figure begin well left of that. This is because a person dying at some old age right at the start of the series will have of course been born well before the series itself began. Anything to the left of 1621 is incomplete in a sense, because we're missing the lifelines that would overlap the pre 1621 lives we already have (e.g. a life extinguished at age 50 in 1620). So, we could of course crop the surface at 1621, no problem. However, we have the same problem on the right side of the series. Ignoring the block shapes for now. A person dying in 1850 at age 60 would clearly overlap with part of the blue we have, but would here be excluded. The right-side zone of still-incomplete observation in LP is the upper-right triangle of the whole surface; so, draw a line inward from the lower right corner at slope -1. Anything above that is still missing the early part of some lives. Those funny observation blocks offset said missingness to some degree, a happy accident. As a rule, missingness is higher in LP as we go up the y axis on the right side.

Now imagine we have an LP surface of the prevalence $\pi(l,p)$ of some health condition from HRS. The observation window is much smaller, but we are located high up in the lifespan range. Most of the people whose prevalence should contribute to $\pi(90,2015)$ have not died yet, because many of them are still toddlers. Others not yet included in $\pi(90,2015)$ are those age 80 in 2015 that will not die until 2025. None of this will be a problem if we condition these planes on some other bounding time measure, I guess it's just something to be careful with. My first guess as to conditions that might vary in a regular way over LP would be prevalence of having ever smoked. Here's a screenshot of the TAL plane of the 1915 female birth cohort for ever-smoking prevalence:

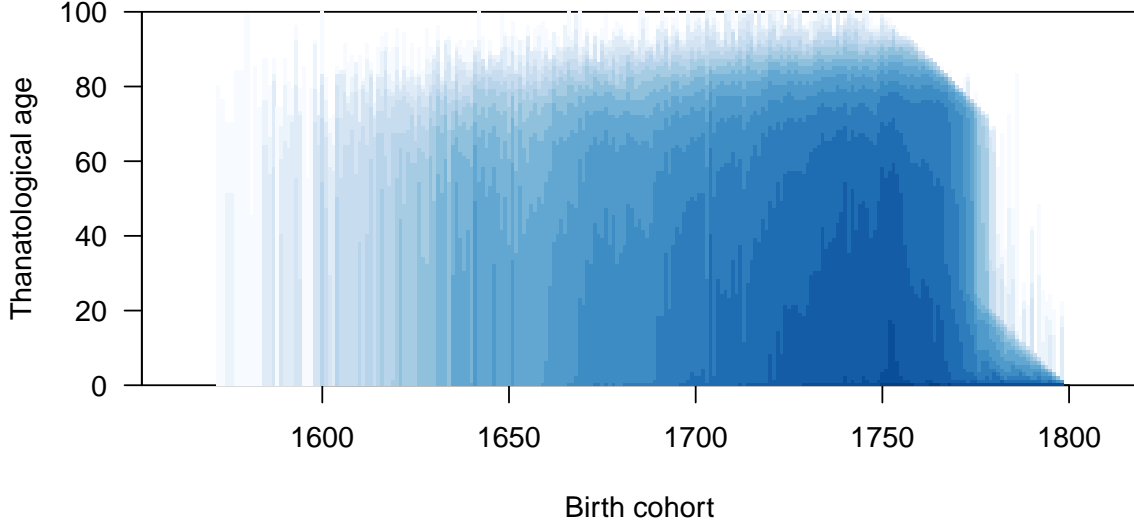


Figure 4: Population structure by time to death and birth cohort.

I used this as an example in another recent paper ([link](#)). I found this striking. Lifespan contours are those with slope -1. One explanation would be that smoking determines length of life before age 70 and things simply don't change thereafter (you can't undo being an ever smoker). As I recall dental visits had contours with this slope, but the surface was not monotonic. It's just a guess that strong patterns like these would carry over into LP.

The Quebec data has an incredible observation window, but no health except fertility (and extended family events). Lifeline surfaces like those produced for Quebec can be inferred from any Lexis surface of death counts.

3.2 TC

The thanatological age distribution of a birth cohort specifies that birth cohort's life table (distribution of T given C). Conversely, the distribution of birth cohorts of people of a given thanatological age reveals mortality history and inequalities in survival. If more of the people who will die in 5 years are from recent cohorts than from cohorts long ago, this population has high childhood mortality and low survival to old ages.

A lifeline on the CT plane extends from 0 upward until the age of death. Ergo, someone born in 1700 dying at age 80 contributes 1 to the count in each thanatological age from 0 to 80, the entire countdown. $T = 0$ is maximally heterogenous with respect to A (or L or D or P), and lifeline overlap decreases in higher T (long-lived people while they're young, i.e. in years closer to C itself). Here's that tabulation of this same data (also logged):

The first thing to pop out at me are birth dearths for particular years, possibly the same years that were the right edges of mortality shocks in the LP surface. Also a wider boom or two, which might just be natural growth. Here each vertical strip monotonically decreases as we go up to higher thanatological ages. I agree these are the stationary populations belonging to each birth cohort, each with a radix equal to the actual birth

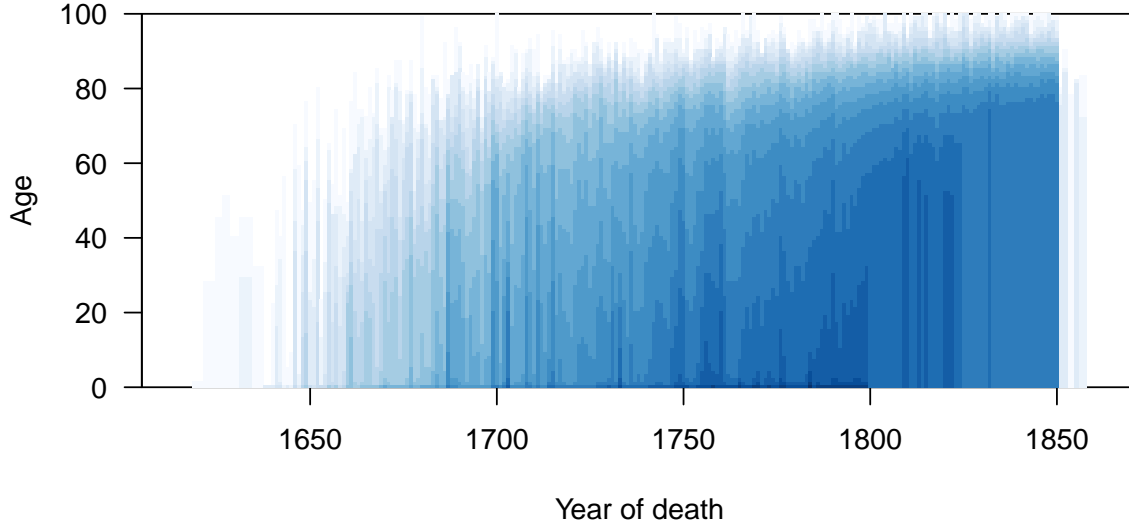


Figure 5: Population structure by age and year of death.

cohort size, so this is a bit more straightforward to wrap one's mind around than LP.

It seems that the upper envelope (that is, the thanatological age at which the lifelines terminate, as a function of cohort) is rising with increasing cohort, reflecting improvements in cohort life expectancy. These improvements are visible also in contour lines that are lower than the upper envelope. You suggested that growth happens in the contours mainly because the population was growing exponentially. If you wanted to remove the effect of population growth, you could plot a curve of births or a curve of population size or a curve of number of women in the population below the x axis with an inverted y axis and then normalize the counts by births or population size or number of women. That might reveal additional patterns. This kind of normalization by some measure of population size might also be applied to the next figure.

I now detect the affects of the right-side block observation of deaths. Had burial records for younger ages been captured in this data, we would have increasing contours as we go down towards the x axis on the right side. Here we just get the tips of the oldest deaths, each having needed to pass through young ages first, so they trace down to age 0.

3.3 AD

For a given age A , the distribution of the death cohort D of individuals gives the conditional life table, conditional on reaching the given age A , and hence gives, for example, the remaining life expectancy and the variability of remaining life, among other summaries of the conditional life table. Conversely, for a given death cohort D , the distribution of age A of individuals specifies the distribution of their birth cohorts and life lengths.

In the AD plane, we have the same sort of vertical overlap. Just as everyone experiences $T = 0$, everyone is also included in $A = 0$ here. If this were a stationary population then a vertical strip here would be equal to one from the CT plane, but we have lots

of non-stationary things happening. Vertical striation here is due to mortality shocks in particular years, and these are strictly positive. Anything dearthy-looking is either due to a 'harvest' or data collection oddities. At least I've never heard of a survival shock. So, to repeat, an individual born in 1700 destined to die at age 80 is located at 1780 in the x axes, and ages from the bottom up. Ergo, the bottom of the surface is young people, as we're used to, but they come from different birth cohorts. As with CT, the population is maximally mixed with respect to C (or L or P or T) at $A = 0$. An 80 year old and an infant dying in 1780 both contribute to the count at age 0 in 1780.

We see here the same high density in death year 1705 that we detected in the LP plane, so maybe it is not a data artifact. Whether there are also high densities of deaths in the other years we identified in the LP plane is not so clear to me; perhaps yes. There appear to be little white rectangles between 1760 and 1800 at ages zero through perhaps 3 or 4. Every lifeline must extend to age zero, so something is wrong here, I think. Again we see the upper envelope increasing with the year of death and rising contour lines.

4 Independant time surfaces as cross-sections

Here use example of HRS surfaces produced at novel *cut* angles.

4.1 LP

4.2 TC

4.3 AD

References

- Desjardins, B. (1999). Le registre de la population du qubec ancien. *Annales de Dmographie Historique* 1998(2), 215–226.
- HRS (2017). Health and retirement study public use dataset. Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant number NIA U01AG009740).
- Mazan, R. (2011). *Analyzing Epidemics in New France: The Measles Epidemic of 1714-15*. Ph. D. thesis, University of Western Ontario.
- RAND (2018). RAND HRS Data, Version P. Produced by the RAND Center for the Study of Aging, with funding from the National Institute on Aging and the Social Security Administration.
- Riffe, T., J. Schöley, and F. Villavicencio (2017). A unified framework of demographic time. *Genus* 73(7).
- Riffe, T., A. Van Raalte, and M. J. Bijlsma (2017). Healthy life expectancy, mortality, and age prevalence of morbidity. *MPIDR Working Papers*. WP-2017-015.