Time spent and left of transient states in stationary populations

Tim Riffe · Francisco Villavicencio

Abstract

Background The Brouard-Carey equality establishes equality between the distributions of years-lived and years-left in stationary populations.

Objective This equality can be generalized to account for time spent and time left in states within multistate stationary populations. We provide an intuitive proof that the distribution of time spent and left in transient states is equal in multistate stationary populations.

Conclusions We speculate that this equality may be helpful under certain constraints to estimate the distribution of otherwise unobserved onset timing for health or other states.

Keywords Brouard-Carey equality · Stationary population · Age structure · Symmetric patterns · Multistate models

Mathematics Subject Classification (2000) 91D20 · 60K05 · 92D25

1 Background

The Brouard-Carey equality establishes that the distributions of years-lived and years-left are identical in perfectly stationary populations (Brouard 1989, Vaupel 2009, Rao and Carey 2015, Villavicencio and Riffe 2016). A perfectly stationary population is either perfectly stationary because it is of infinite size or because it is finite and deterministically repeating, and in either case vital rate schedules must be fixed and the intrinsic growth rate constant at null.

Let's say that individuals in the stationary population can obtain different states over the life course. If birth and death rates do not vary by states, then it does not matter whether state transition rates are fixed or not, for in the aggregate the population remains stationary in the traditional sense. However, if vital rates depend on one's state, transition rates must also be fixed in order for stationarity to hold in the aggregate— This is the situation that we entertain in the following. Under fixed vital and state-transition rates, where vital rate schedules differ between sates, the standard set of aggregate invariant quantities of course remains: birth cohorts and death cohorts are of fixed and equal size. Further, the Brouard-Carey equality still holds in the aggregate. Under these conditions, once stationarity is acheived, the age-state structure of the stationary population is also invariant over time. As a consequence of fixed transition and vital rate schedules, the distribution of state-specific tenures for individuals entering a state at a given age is also a fixed attribute.

This setup is not entirely contrived, for it corresponds with the assumptions of common multistate markov models, often used to calculate state expectancies (cite caswell and tulja). In the present, however,

Tim Riffe

Max Planck Institute for Demographic Research. Konrad-Zuse-Str. 1. 18057 Rostock, Germany

E-mail: riffe@demogr.mpg.de

Tel.: +49 176 232 858 45

Fax: +49 381 2081 - 280

Francisco Villavicencio

Department of Public Health & Max-Planck Odense Center on the Biodemography of Aging. University of Southern Denmark, J.B. Winslws Vej 9B, 2nd floor DK-5000 Odense C, Denmark

E-mail: fvillavicencio@health.sdu.dk

we are not bound to memoryless transition probabilities. Instead, we simply require the consequence of invariant age-state structure and invariant conditional age-state tenure distributions, which may result from either memoryless transitions, or from arbitrarily linked dependencies. It is the fact of a fixed age-state structure on which we base the following observations. A fixed age-state structure implies a fixed distribution of lifelines, each consisting in a sequence of state-specific durations.

Under fixed vital and transition rates, and zero growth, we introduce a new theorem, which is a more general version of the Brouard-Carey equality. Namely, if there is some state s in the population, a randomly drawn individual from state s has equal probabilities of having entered s x years ago and exiting s in x years. This proposition is probably not intuitive at first glance, so we prove it, and then speculate as to how this equality may be put to good use.

2 An intuitive proof of transient symmetry

Theorem 1 *Given a stationary population and fixed state-transition rates, the probability that a randomly selected individual is in state s and entered s x years ago is equal to the probability of being in state s and exiting in x years.*

Proof The proof of this statement follows five intuitive steps, built up from lifelines.

1. A duration can be represented as a line segment, potentially a subset of a life-line. Points along a single within-person duration can be sampled over time in arbitrarily fine time steps, δ. Each time step can then be used to bisect the segment, collecting two sets of values: 1) time spent in the state prior to the sample point, and 2) time left until exiting the state. If observations are evenly spaced, and values are truncated to the nearest δ, by way of complements these two sets will consist of the same values.
   If the duration of the ith individual is called d^i, the age of entry is a_L^i and the age of exit is a_R^i, such that $d^i = a_R^i - a_L^i$. The set of time-spent values, A^i is defined as:

$$A^i = \left\{ \delta \cdot k \mid k \in \mathbb{Z} , \ 0 \le k \le \left\lfloor \frac{d_i}{\delta} \right\rfloor \right\} \qquad , \tag{1}$$

   where $\lfloor \dots \rfloor$ denotes truncation to the lower integer bound. The set of time left values, T^i is also:

$$T^i = \left\{ \delta \cdot k \mid k \in \mathbb{Z} , \ \left\lfloor \frac{d_i}{\delta} \right\rfloor \ge k \ge 0 \right\} \qquad . \tag{2}$$

   Figure 1 illustrates the construction of A^i and T^i in expressions (1) and (2). The central cut-point moves along the duration from the time of entry until the time of exit, creating two lists of values.

Fig. 1: A lifeline of individual i showing the construction of the sets A^i and T^i.
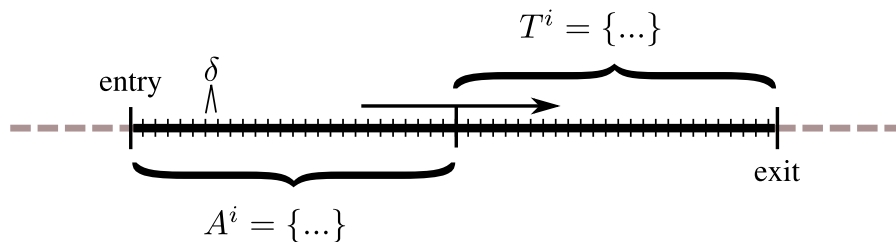
2. If this individual is reborn every δ time step, destined to relive the exact same life course as the first, we end up with a set of identically aligned and identically long segments placed side-by-side, spaced apart by δ. This potentially large set of segments may be imagined as a plane, although this would only hold in the limit. One could in this setting take a census at a single point in time, collecting a set of time spent and time left values, each from a unique individual in sequence. It is clear that the two sets observed at a single time point but drawn from the population of perfect clones, will be identical to the first two sets that were observed of a single duration over its entire length, if the same δ-truncation operation is followed.

Figure 2 illustrates this notion with uniformly-spaced lifelines in a Lexis configuration. The vertical line indicates a hypothetical census at time t of the population of this cloned individual. At time t, the blue-highlighted segments indicate the set of time-spent values in A^i, and red-highlighted segments are the time-left elements of T^i.

Fig. 2: The life of individual i repeated in δ time steps. A census with followup now constructs the sets A^i and T^i with values identical to the within-individual sets.
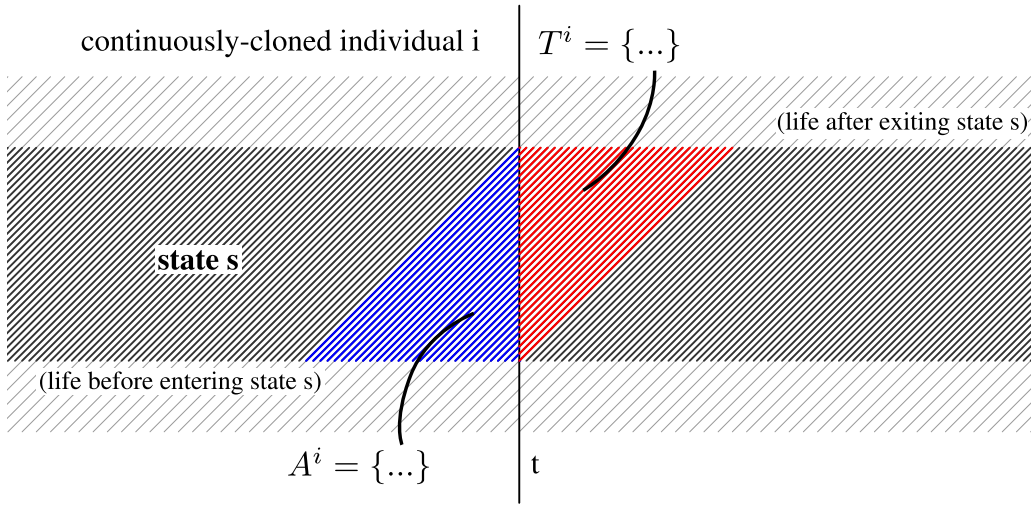
Formally, sets A^i and T^i consist of the same values as the previous, but coming from individuals born in a uniform series from between $t - a_R^i$ years ago until as recently as $t - a_L^i$ years ago. This demonstrates both period-cohort equality and time spent-left equality. The blue and red triangles in Figure 2 are simple rotations of one another.

3. Assume we have a second individual from the same birth cohort as individual i that enters the same state as the first, but at a different time and for a different total duration. We could demonstrate time spent and time left equality in the same way for this individual, by sampling in δ time steps. Since this individual has a different life course timing than the first individual, their sets of time-spent, A^2, and time left T^2 will be distinct. A^1 and T^1 range from 0 to d^1, whereas A^2 and T^2 range from 0 to d^2. However, their concatenations are identical:

$$\{A^1, A^2\} = \{T^1, T^2\} \tag{3}$$

4. If the second individual is also perfectly cloned in δ time steps as in step 2, then our census at a single point in time also yields identical sets of time spent and time left values, A^2 and T^2. Also from this census, the concatenation of the first and second time-spent sets and the concatenation of the first and second time-left sets are guaranteed to be identical, as in equation (3).

5. By induction we can keep adding durations in this state, infinitely if we please, and the collection of all resultant time-spent sets and the collection of time-left sets will continue being identical. Therefore the probability of selecting a particular value from the time-spent set is identical to the probability of selecting the same value from the time-left set.

Our point of departure was that of a perfectly stationary population with identically distributed lifelines that aggregate to a fixed age-state structure, and this is also the result that we arrive at by induction in the final step of the proof. In this strict environment it is therefore the case that our

theorem holds. A more rigorous proof would allow δ to decrease to 0 in the limit. The present approach does not allow this because probability statements would no longer hold. We may require a population of infinite size in order to make the jump from a fixed age-state structure to an invariant set of lifelines, where each lifeline is itself an invariant sequence of states.

3 Discussion

The approach from this proof is equally valid to prove the original statement of the Brouard-Carey equality (where the state is "alive"), but it is more general. The statement and proof is flexible enough to hold for irreversible and reversible states. It also applies to repeatable states, whether time spent in the state is kept in cumulative fashion over spells, or whether the clock resets to zero on each entry into the state. The equality may also be conditionable in curious ways: for example, the distribution of time-spent in a state conditional on having entered at age a must also be equal to the distribution of time-left in the state, conditional on having entered at age a. Likewise, one may condition statements on exit age. One may also arbitrarily merge states, and the equality still holds within the newly merged state.

At first glance, this equality is probably less intuitive than the original Brouard-Carey equality, because state entry is not necessarily aligned on age zero. It is less visible in commonly-produced plots because plots of stacked sequences look chaotic, often even if clustered or sorted. It might be tempting to think that due to state-varying vital rates, the equality simply ought not hold. However, the basis of the proof is the observation that if each individual duration is symmetrical by complements, then so are aggregations of durations, irrespective of alignment. Since each cohort in a perfectly stationary population of infinite size is an identical copy of the previous, census-like cross-sections are also equally-composed.

4 Potential applications

Empirical applications of the presently-described transient tenure equality may be easy to conjure up. For example, imagine a hypothetical health state that shows no noticeable symptoms, but that is medically measurable. One may take a census with regular follow-ups, until eventually the state is exited by each individual, whether by absorption into death or entry into another state. Then, if the assumption of stationarity is acceptable, one may be able to say something about onset timing in the aggregate, itself unobserved. We think that the present equality will come as good news to researchers in similar settings.

5 Simulation

The reader may wish to have a sense of how well this equality holds up under various violations of our assumptions. The assumption of invariant transitions is trickier to handle than that of finite populations with stochasticity. We conduct a simple markov chain simulation to produce random populations of different sizes derived from the same set of stationary transition probabilities. Our transition rates come from a published study of working life expectancy in older ages in the USA (Dudel and Myrskylä 2017), and refer to US black females around the year 1994. The transition matrix includes single ages 50-100, with lifetable closeout at age 100. Transient states include employed, unemployed, inactive, and retired. We generate random sequences of trajectories using the `rmarkovchain` function of the `markovchain R` package (Spedicato 2017).

For simplicity, all trajectories begin in a state of employment at age 50. We generate 1000, 10000, and 100000 individual trajectories. Out of these we sample observations of inactivity a total of 100, 1000, and 10000 times, respectively with replacement. We assume that the time of observation is half way through the year, but that spells begin on January 1 and end on December 31. For each observation of inactivity we measure the time spent in the spell up to the point of observation, and the time left until exiting the spell. These values are then tabulated for each simulation to produce distributions to compare. The theorem states that the distributions of time spent and left in this procedure should be asymptotically identical, but we have induced noise with the simulation, and we should be able to see distributions converge as simulated population and draw sizes increase.

6 Conclusion

Stationary populations may be more symmetrical with respect to time lived and spent than has been previously supposed. We have given an intuitive proof that the distribution of time spent and left is equal even within states in stationary populations. A more rigorous proof than the present will surely be forthcoming, but we hope to have left the reader assured of the truth of this statement so that it may be exploited in empirical research. We do not presume that this relationship will be immediately able to answer pressing questions, but we hope that it may inspire new approaches in empirical measurements. We think that that researchers working with left-censored or truncated data or any kind of multistate models should be generally aware of this equality in case it may come in handy as a heuristic.

The relationship between years lived and left established by the Brouard-Carey equality is itself amenable to non-zero growth rates (Riffe 2015). We speculate that the transient tenure equality we describe here may also be tractable in stable populations, but we have not investigated this possibility in detail. Future research, potentially more sophisticated simulations, should establish the impact of departures from stationarity and establis the fuzzy bounds of usability for this relationship.

References

Brouard N (1989) Mouvements et modèles de population. Institut de formation et de recherche démographiques

Dudel C, Myrskylä M (2017) Working life expectancy at age 50 in the united states and the impact of the great recession. Demography 10.1007/s13524-017-0619-6, URL https://doi.org/10.1007/s13524-017-0619-6

Rao ASS, Carey JR (2015) Generalization of Carey's equality and a theorem on stationary population. Journal of Mathematical Biology 71(3):583–594

Riffe T (2015) Renewal and stability in populations structured by remaining years of life. MPIDR Working Papers (WP-2015-007)

Spedicato GA (2017) Discrete time markov chains with R. The R Journal URL https://journal.r-project.org/archive/2017/RJ-2017-036/index.html, R package version 0.6.9.7

Vaupel JW (2009) Life lived and left: Carey's equality. Demographic Research 20:7–10, 10.4054/DemRes.2009.20.3

Villavicencio F, Riffe T (2016) Symmetries between life lived and left in finite stationary populations. Demographic Research 35(14):381–398