Skewed distributions of LRS: beyond mean and variance

Shripad Tuljapurkar*[1], Wenyun Zuo[1], Tim Coulson[2], Carol Horvitz[3], and Jean-Michel Gaillard[4]

[1]Department of Biology, Stanford University, Stanford, CA 94305-5020, USA
[2]Department of Zoology, University of Oxford, Oxford, OX1 3SZ, UK
[3]Department of Biology, University of Miami, Coral Gables, FL 33124-0421 USA.
[4]Laboratoire de Biométrie et Biologie Evolutive, Université Lyon 1, CNRS, UMR 5558, F-69622 Villeurbanne, France

*corresponding author: tulja@stanford.edu;

Abstract

Many field studies find that lifetime reproductive success (LRS) is highly skewed and often multimodal among individuals. Field biologists generate invaluable data on survival and reproductive rates, as a function of age and stage, that are used to parameterise structured models. These models often perform well at predicting population growth and mean LRS, but we do not know whether they accurately predict observed distributions of individual LRS. If the models fail to recreate these distributions, their use may be limited because the LRS is central to understanding life history evolution. We present powerful tools to generate distributions of LRS from age and/or stage structured models. Our methods reveal that structured models do perform well at generating distributions that agree with observations. Our approach also reveals why such skewed distributions arise, and helps resolve a debate about detecting signatures of selection in skewed distributions of LRS.

Introduction

Lifetime measures of reproductive performance describe the number of offspring an individual produces over its lifespan (Clutton-Brock 1988). Many empirical biologists consider such lifetime measures as individual fitness (e.g. Kruuk et al. 1999; Brommer et al. 2004), although, with the exception of annual species (Fisher 1930), these measures are correlates of fitness (Grafen 1988). Nonetheless, lifetime reproductive performance can provide generational estimators of the opportunity for natural and sexual selection and phenotypic evolution. For instance, many empirical studies have assessed the opportunity for selection using the standardized variance (I metric, Crow 1958) of reproductive performance and revealed high variation (e.g. in females from 0.177 in lions, *Panthera leo* to 3.62 in elephant seals, *Mirounga angustirostris* (Cabana and Kramer 1991); and in males from 0.59 in Atlantic salmon *Salmo salar* to 4.52 in bighorn sheep, *Ovis canadensis*, (Tatarenkov et al. 2008)). Comparison of standardized variances of lifetime reproductive performance between sexes within a population is commonly used for assessing the strength of sexual selection (e.g. Dubuc et al. (2014) on Rhesus macaques, *Macaca mulatta*). These empirical studies consistently reveal that distributions of lifetime reproductive performance measures are often non-normal, zero-inflated and highly skewed, and this begs the question: what are the demographic causes of these unusually-shaped distributions?

Lifetime Breeding Success (LBS) censuses offspring at birth (e.g. Slate et al. 2000), while Lifetime Reproductive Success (LRS) censuses them at some

later stage (e.g. weaning/fledging (van Noordwijk and van Schaik 1999), one year-old (Newton 1988), two years-old (Clutton-Brock 1988)). We analyze discrete-time population models that are formulated in terms of species-appropriate modes of breeding and counting, but for simplicity use just the term LRS. So the question is: do such models predict LRS distributions that mimic those observed in nature, and why?

In the models here, individuals are distinguished by age and/or stage, and life cycles are described by probabilities of survival, stage transitions, and reproduction. Simulations for many species yield a distribution of LRS that is dramatically skewed or multimodal (Tuljapurkar et al. 2009b; Steiner and Tuljapurkar 2012), in accord with the empirical results mentioned above. Hence in any cohort, many individuals have an LRS of 0 so they leave no offspring; a lucky few have an LRS much higher than the average value. This variation is driven purely by chance since all the individuals are governed by the same probability rules (Steiner and Tuljapurkar 2012; van Daalen and Caswell 2017) (hereafter the latter reference is abbreviated vDC), meaning that "luck" can be more important than "pluck" (Snyder and Ellner 2018).

To quantify such "neutral" variation, Steiner and Tuljapurkar (2012) obtained formulas for the mean and variance (and higher moments) using a fixed fertility at every age/stage. But, as emphasized by vDC, offspring only come in integers (0, 1, 2, ...), so reproduction at every age/stage should properly be described by probabilities that an individual has $0, 1, 2, \ldots$ offspring. Using the latter description, vDC derived formulas for the moments – mean, variance, and so on – of LRS for each birth stage (extending Caswell

(2011) and van Daalen and Caswell (2015)). The analyses by Snyder and Ellner (2018) also examined variance computed in this way as a measure of dispersion in the LRS. For populations structured only by age, the variance in LRS has previously been computed by Waples et al. (2011) and Charlesworth and Williamson (1975).

Yet, the moments do not provide the same information as the probability distribution of the LRS. Why? After all, the formulas in vDC easily yield the first three moments (also higher order moments but these don't reveal such interesting patterns as multimodality). For one thing, our intuitive understanding of these moments is often based on a roughly normal distribution and so can be of limited value. For another, a general approach to approximating a distribution by the first few moments, e.g., the Gram-Charlier expansion (Stuart et al. 1994), doesn't work for many (all the ones we've tried) LRS distributions, as we illustrate here. Thus the complete set of probabilities for LRS is often essential, certainly when the distribution is highly skewed or multimodal.

We present here a new method for computing the entire probability distribution of LRS for most species. We show below by examples that the distribution of LRS contains new and useful biological insights. Our methods highlight important questions: what factors determine the probability that individuals produce zero offspring during their lifetimes; how does a skewed LRS distribution affect natural selection, effective population size, and life history evolution? We discuss these questions, and point towards future work needed in this area.

The main text explains our method, and then focuses on examples. The first two examples are age-only models for human populations: the Hadza, a Tanzanian group of hunter-foragers, and a North American group called the Hutterites. The next example is a stage-only model for the large evergreen tree *Tsuga canadensis*. For these examples, the sources are cited in vDC, who also kindly provided matrices. The last example, and perhaps most general, is an age+stage model for Roe deer, *Capreolus capreolus*, based on the data in Plard et al. (2015). For each example we compute and display the full distribution, and also the first 3 moments (the same as in vDC).

Methods

Lifetime Reproductive Success (LRS) is a random quantity, and we find the probability for each possible value of LRS. We model demography in discrete time; vital rates may be structured only by age, only by stage, or by a combination of age and stage (age+stage). Mathematical details are in the online Appendix; we use and explain discrete convolutions and discrete Fourier transforms, unfamiliar to some biologists, but essential here. Code (based on R (R Core Team 2019) and the R package FFTW (Fastest Fourier Transform in the West) (Mersmann 2019)) is available from the authors. Our results should apply to continuous stages using established methods (Ellner et al. 2016; Snyder and Ellner 2018). Notation is always a nuisance; where possible, we use notation similar to Caswell (2011) and Caswell et al. (2018).

Our logic uses three points. First, every individual has a particular age

of death; second, at each age before death, there is a probability distribution for the number of offspring produced at that age; and third, convolution of these age-specific distributions (weighted appropriately by mortality) yields the distribution of LRS. We make the assumption that the probability distribution at each age is independent of the distribution at all other ages.

General Framework

Age is counted in discrete intervals $a = 1, 2, \ldots, \omega$ and newborns are always in the first age interval (parameters are defined in Table 1). Stage is defined in discrete categories, $s = 1, 2, \ldots, S$ (indexed as i or j) for stage-only and age+stage models. Stage transition probabilities (conditional on remaining alive) over a single time step are defined in a matrix $\boldsymbol{G}$ for stage-only models, or in a set of matrices $\boldsymbol{G}_a$ for age+stage models. One-period survival probabilities (from one interval of time to the next) are p_a for age-only models, $p(j)$ for stage-only models and $p_a(j)$ for age+stage models (Table 1). Multiplying stage transitions by stage survivals, we obtain an $S \times S$ matrix of unconditional stage transition probabilities, $\boldsymbol{U}$ for stage-only models and a set $\boldsymbol{U}_a$ for age+stage models (Table 1).

Age-at-death is a random variable (well-defined even for populations where one-period survival depends only upon stage). The probability that an individual dies at each age a is the product of the probability of living to the beginning of age interval a with the conditional probability of dying during that age interval.

Reproduction for age-only and age+stage models is only possible between some specified youngest age α and an oldest age β, respectively. Stage-only models may have non-reproductive stages, but such stages may occur at any age. The number of offspring produced during a single time interval is a random variable that can take on integer values $k = 0, 1, 2, \ldots, k_m$. The probability distribution of reproduction (in a single interval) depends on age ($\boldsymbol{\kappa}_a$), stage ($\boldsymbol{\kappa}_j$), or age+stage ($\boldsymbol{\kappa}_{a,j}$), respectively for the age-only, stage-only and age+stage models (Table 1).

Finally, the number of offspring produced by individuals during a lifetime, the LRS, is a random variable that can take on integer values $k = 0, 1, 2, \ldots, k_u$. The final goal of our analyses is always the probability distribution Γ whose elements are

$$\gamma(k) = \Pr[\text{LRS } = k]. \tag{1}$$

Our calculations assume that, within a single time interval, reproduction precedes death. Thus an individual alive at the beginning of a given time interval always reproduces according to its age and or stage at the beginning of the interval. Our formulas are easily changed to suit other assumptions.

The practical flow of decisions about which of the following methods to use is shown in the online Appendix (Fig. A.4).

Age-only Model

Here, survival and reproduction during a single time interval depend only upon the age of individuals. We want the probability distribution Γ of LRS (see equation (1)). The distribution of offspring $\boldsymbol{\kappa}_a$ at age a is taken to be binomial for age-structured human populations (so $k_m = 1$ and there are 0 or 1 offspring at each age); but our method below applies to any form of probability distribution. We break life into segments: newborns (age 1) who die before reaching age 2, those who live to age 2 but die before 3, and so on.

First, consider individuals who die before reaching age $a = 2$. The probability of such a death is $\phi_1 = l_1 [1 - p_1]$, and reproductive success at age 1 follows the distribution $\boldsymbol{\kappa}_1$, so the LRS for such individuals has the probability distribution

$$\boldsymbol{\gamma}_1 = \phi_1 \, \boldsymbol{\kappa}_1. \tag{2}$$

Next, consider individuals who reach age $a = 2$, reproduce and die before age $a = 3$. Such individuals have had two opportunities to reproduce. If these individuals have LRS = 2, they could have: i) produced 0 offspring at age 1 and 2 offspring at age 2, or ii) produced 1 offspring at age 1 and 1 offspring at age 2, or iii) produced 2 offspring at age 1 and 0 offspring at age 2. In other words, at ages $(1, 2)$, they could have $(2, 0)$, $(1, 1)$ or $(0, 2)$ offspring. Hence the probability of LRS = 2 is

$$\kappa_1(2)\kappa_2(0) + \kappa_1(1)\kappa_2(1) + \kappa_1(0)\kappa_2(2).$$

Each LRS $= k$ is analyzed in this way, and the corresponding probabilities are elements of the discrete convolution

$$\boldsymbol{\delta} = \boldsymbol{\kappa}_1 \star \boldsymbol{\kappa}_2. \tag{3}$$

For more detail on convolutions, see the online Appendix (Table A.1 and section A.1). There are $(2k_m + 1) = 3$ possible outcomes resulting from this convolution. Since the proportion of a cohort that dies at age 2 is ϕ_2, the LRS for individuals in a cohort who die before reaching age $a = 3$ is distributed according to the probabilities

$$\boldsymbol{\gamma}_2 = \phi_2 \, \boldsymbol{\kappa}_1 \star \boldsymbol{\kappa}_2. \tag{4}$$

In general then, an individual who lives to age $2 \leq a < \beta$ but dies before reaching $a + 1$, has LRS distributed as

$$\boldsymbol{\gamma}_a = \phi_a \, \boldsymbol{\kappa}_1 \star \boldsymbol{\kappa}_2 \cdots \star \boldsymbol{\kappa}_a. \tag{5}$$

Reproduction ends after age β, so survival to later ages does not change the LRS. Consequently, for an individual who lives to age $a \geq \beta$, the LRS has distribution

$$\boldsymbol{\gamma}_\beta = l_\beta \, \boldsymbol{\kappa}_1 \star \boldsymbol{\kappa}_2 \cdots \star \boldsymbol{\kappa}_\beta. \tag{6}$$

Adding, we have

$$\Gamma = \sum_{a=1}^{a=\beta} \boldsymbol{\gamma}_a. \tag{7}$$

Stage-only Model

For many species, stage (such as size or developmental stage) is predictive of vital rates rather than age (Lefkovitch 1965). Survival, stage-transition, and reproduction during a single time interval depend only on the stage s of individuals at the beginning of that interval.

The number of offspring (k) produced during a time interval depends only on stage j at the beginning of the time interval and is specified by probability distributions $\boldsymbol{\kappa}_j = \{\kappa_j(k)\}$ of producing $k = 0, 1, 2, \ldots$ offspring (Table 1). These distributions are often taken to be Poisson (e.g., some cases studied in vDC), but that is not a requirement of our analysis.

Stage-structure requires multiple convolutions that make the age-only procedure computationally unwieldy. This is because here, in contrast to an age-structured model, individuals may visit a stage any number of times before dying. They may also move back and forth between non-reproductive and reproductive stages. So we need a fast way of (a) doing repeated convolutions, and (b) summing over repeated convolutions.

To solve (a) we use the mathematical fact that the (discrete) Fourier transform of a (discrete) convolution is the product of Fourier transforms, and the numerical fact that the FFTW package (Mersmann 2019) does a fast discrete Fourier transform. To solve (b), we exploit our earlier work (Steiner and Tuljapurkar 2012) and combine a closed-form generating function with Fourier transforms, then use an inverse FFT. For the details (or if this brief description makes you curious) see the online Appendix (section A.5).

Thus we can compute the exact distribution of LRS for any stage-structured model. We can separately, and quickly, compute the probability that the LRS = 0. Both analyses are given in the online Appendix (sections A.5 and A.8).

Age+Stage Model

Here survival, stage transitions and reproduction during a single time interval depend on both age and stage (e.g. Coulson (2012), Table 1 and the online Appendix). In principle, here we can proceed as with the age-structured model. But the computation depends on the "size" of the life cycle, and determines a choice between two options. If all individuals reach a terminal age and die, and if the number of stages is modest, use the age-structured approach. If there are many ages and/or stages, and/or if all individuals reach an "old" age with stages in which they reproduce and remain until death, the age-structured approach is too slow. In that case, use a general stage-structured approach with block matrices (see online Appendix section A.6 for details).

Results

Age-Only Model: Two examples

We contrast the Hadza, a hunter-forager group thought to be representative of early humans, to the Hutterites who in the 1950s were famous for their high fertility (vDC). We consider only females with no migration. One-

period survival probabilities by age, p_a, yield survivorship l_a (the probability of surviving until age a). Reproduction ceases at menopause $\beta \approx 50$ yr, so we do not need later ages. The commonly used fertility rate f_a is the mean number of offspring at each age. Fertility at age a can be 0 or 1 for human populations (vDC), with

$$\begin{aligned} \kappa_a(0) &= \Pr[0 \text{ offspring}|\text{ parent age } a] = \big(1 - f_a\big), \\ \kappa_a(1) &= \Pr[1 \text{ offspring}|\text{ parent age } a] = f_a. \end{aligned} \tag{8}$$

For each age, these numbers constitute a distribution we denote $\boldsymbol{\kappa}_a$. We ignore twinning, which is rare in these populations (but can be included by adding the value of $\kappa_a(2)$). For other species (e.g., if there are litters with many offspring) we need additional information to create a distribution of offspring at each age.

We used our methods to compute the exact distributions of lifetime reproductive success (LRS), Hadza in Fig. 1(a), Hutterites in Fig. 1(b); the distribution is given by the solid points in the figure. Notable features:

(1) The distribution for the Hadza is dramatically bimodal, with a very high mode at 0 and another at LRS ≈ 6. These modes are very different from the mean LRS of 3.13; the spread in LRS is not described by the standard deviation ($= 3.36$). The third central moment ($= 21.7$) does imply the distribution is skewed.

(2) The Hutterites have a low mode at 0 but a high mode at LRS close to

the mean of 7.53. Excluding the mode at 0, the distribution appears to concentrate nicely around the mean value, with a standard deviation ($= 2.93$) that captures the spread of nonzero values. However the third central moment is negative ($= -14.49$), presumably because of the mode at 0. If we didn't have the full distribution on view, this value of the third moment would be hard to understand.

Mode at LRS $= 0$

Is this likely to be a common feature? A female may have 0 offspring if she dies before reproduction begins, or lives on but has 0 offspring at every age. For the Hadza and the Hutterite, reproduction can begin at age 13 yrs. The probabilities of dying by that age are 0.415 for Hadza, 0.043 for Hutterites. The probability of living on but childless is 0.035 for the Hadza and only 0.008 for the high-fertility Hutterites. So the mode at LRS $= 0$ is largely determined by high infant mortality in the Hadza and the much lower infant mortality among Hutterites. Thus in general we should see a mode at LRS $= 0$ unless early mortality is extremely low.

Do 3 moments predict the distribution?

We can compute the LRS distribution from our methods, but can we recover a reasonable approximation from just the first 3 moments? For the Hadza, the answer is emphatically no, as shown by the Gram-Charlier approximation in Fig. 1(a) (the dotted line). For the Hutterites, the approximation works rather well for nonzero LRS but entirely misses the mode at 0, as shown in

Fig. 1(b) (the dotted line); so the answer is still no. We've tried a different method, starting with a Poisson and then using a numerical search for the "best" modification, but the answer is still negative. We have not tried using higher moments, because they do not seem likely to change our results.

Stage-Only Model

An example: An evergreen tree *Tsuga canadensis*

We use a stage-only model for *Tsuga canadensis*, eastern hemlock, an evergreen tree that grows to be 40-75 feet tall in the wild (as cited in vDC). Populations of this tree, as noted by Lamar and McGraw (2005), usually include many saplings, small individuals that may remain in good condition beneath a canopy for hundreds of years. So individuals exhibit considerable stasis.

The LRS for *T. canadensis* has an extremely pronounced mode at 0, and the skew of the distribution is so dramatic that we had to plot the logarithms of the probabilities of LRS, as shown in Fig. 2. The average LRS is just 1.42, but the standard deviation of 37.57 is not useful as a measure of spread at either lower or higher values. The third moment is huge, $\approx 2.6 \times 10^6$, which certainly means great skew, but does not predict the actual distribution in the figure. We expect a similar skewed distribution for many species (including plants and marine organisms) that produce large numbers of offspring such as seeds or larvae, but few recruits, and even fewer adults. The skew seen in Fig. 2 is relevant to current paradigms about life history evolution and

selection, as discussed in the final section.

Mode at LRS $= 0$

Individuals may produce 0 offspring even when they are potentially reproductive, so our analysis identifies two related but distinct probabilities:

(i) One is $\Pr[\text{LRS} = 0]$, which is a high 0.9883 for *T. canadensis*! This number is the sum of the probability of surviving while in non-reproductive stages plus the probability of producing 0 offspring while in reproductive stages. We can always compute this probability exactly (see online Appendix section A.8).

(ii) So does it always make sense to ask (as in the age-structured case) for the conditional probability

$$\Pr[\text{LRS} = 0 \mid \text{die without reaching a stage capable of reproduction}]?$$

The answer depends on the life cycle. Suppose we start in a non-reproductive stage. In one type of life cycle, every individual that survives eventually makes an irreversible transition to one or more stages that are capable of reproducing (we call these reproductive stages). For this case, it makes sense to treat the reproductive stages as "absorbing," so we can compute the corresponding absorption probability. Then [1 − absorption probability] is just the conditional probability defined above (see online Appendix section A.8). The value of this con-

ditional probability is 0.9878 for *T. canadensis*; comparing this number with 0.9883, we see that most individuals who have an LRS of zero must die without ever having entered a stage that is capable of reproduction.

But, in life cycles of stage-structured populations that are not as described above, this probability doesn't make sense. In the general stage-structured life cycle, individuals can make any number of transitions back and forth between non-reproductive and reproductive stages and death can occur at any point.

Age+Stage Model

An example: Roe deer *Capreolus capreolus*

In general, an individual is classified by both age $a = 1, 2, \ldots, \omega$ and stage $s = 1, 2, \ldots, S$. We illustrate with an estimated age+stage model for Roe deer *Capreolus capreolus* (Plard et al. 2015). Age is in years, and size in 200 equal body mass intervals from 1 Kg to 40 Kg. There are 4 age groups; in each group, there is one set of transition and survival rates and reproduction probabilities (the age groups are yearlings, adults age 2-7 yr, adults age 8-11 yr, adults over age 11). Age advances in single years.

The "initial stage" is a combination of age (here, yearlings in age class 1) and size. For initial age $a = 1$ there are potentially 200 initial sizes. In practice, yearling sizes range from size class 40 to 100, but we analyze the full potential size range (see the online Appendix section A.10). Fig. 3 plots the LRS distribution for 3 different initial size classes, 1, 75, 200. The

distribution shifts from unimodal at LRS = 0 for the smallest yearlings, to unimodal at a nonzero LRS for the largest yearlings, and can be multimodal at intermediate sizes.

Very small offspring have a high mode at 0, and thus the highest probability of leaving no offspring – no surprise. Neither the average LRS nor the standard deviation seem to be useful descriptors of the LRS distribution for the smallest yearlings (plotted in the online Appendix Section A.10). Somewhat bigger yearlings (birth size class 70-84, see Fig. 3) have a bimodal distribution of LRS. Larger juveniles (size class 85 upward) have a single nonzero mode but still high odds of leaving 0 offspring. The nonzero mode shifts from an LRS ≈ 3 (for size class 80) to an LRS ≈ 5 (for size class 200) with increasing yearling size, and the distribution of LRS concentrates towards the mode. Very large offspring have a high nonzero mode at LRS ≈ 5, with a normal-like LRS distribution around the mode.

Mode at LRS = 0

The height of a mode at an LRS of 0 can be computed using a age+stage method described above (see online Appendix Section A.10). We found that Pr[LRS = 0] declines with increasing birth mass. This is expected, since in this population it pays to be born larger. However the gain from increasing birth weight is strikingly nonlinear, so that the advantage of being born 1 Kg heavier is much greater for small than for large yearlings.

To dissect this effect, we analyzed yearling death probability and the probability of having no offspring if you live to later ages (see online Ap-

pendix). There is a drop in both as birth mass increases, but the gain in yearling survival reduces faster than the gain in later survival/reproduction.

Discussion

We have presented an analytical method that computes the probabilities of getting every possible LRS, $0, 1, 2, 3, \ldots$. We illustrated our results for two human populations for age-structure, eastern hemlock for stage-structure, and Roe deer for age+stage structure. We found that the LRS distribution often has one mode at LRS $= 0$, and a second at nonzero LRS. For several of our examples, the mean, variance and third moment are poor descriptors of the LRS distribution. We hasten to add that these moments are useful descriptors if offspring (often censused as newborns or recruits) have high survival, probably uncommon in nature but common in modern human populations. Even in the latter case, though, we recommend using our methods to analyze the probability that LRS $= 0$. We find an even greater variability of the LRS than expected from previous work (Tuljapurkar et al. 2009b; Steiner and Tuljapurkar 2012; van Daalen and Caswell 2015; Snyder and Ellner 2018) and vDC. Here we discuss extensions and implications of our results, and directions for future work.

Descriptors and Sensitivity analyses

How best to describe the variability that we have found in the LRS? And given a descriptor, what are its sensitivities to vital rates? In ecology, sensi-

tivities determine how vital rates (or underlying parameters) affect the LRS distribution, whereas in evolution sensitivities capture selection pressures that depend on LRS Caswell et al. (2018). For future analytical work, note that our stage-structured analysis applies to any age+stage combinations.

One descriptor of the inequality of LRS is a Gini index (Weiner and Solbrig 1984; Shkolnikov et al. 2003), which corresponds to the mean of absolute differences in individual LRS relative to the average LRS, and is bounded between 0 and 1. The Gini index is commonly used in economics and social science to quantify inequality, but less frequently in ecology and evolution. It has been used to assess variation in the distribution of reproductive success of parasites (Dobson 1986), in the mass and fecundity distributions of a cestode infesting pikes (Shostak and Dick 1987), in the strength of forest edge effects (Matlack 1994) and in size hierarchy studies of competitive processes in plants (Weiner and Solbrig 1984). It has also been recently used to describe the shape of mortality within a population (Archer et al. 2018) and to quantify species differences in the distribution of reproductive ages across vertebrates (Healy et al. 2019). We propose that it will be valuable in analyzing variation between sexes or across species in the LRS distribution. To compute the Gini index, find the cumulative distribution

$$\xi(x) = \sum_{k=1}^{x} \Pr[\mathrm{LRS} = x] = \sum_{k=1}^{x} \Gamma(x)$$

and plot $\xi(x)$ versus values (x/k_U) of the LRS (remember k_U is the largest possible LRS). This is the Lorenz plot shown for our examples in the panels

of Fig.4. Equality means that every individual has the same odds of any LRS and is indicated by the dotted lines. The actual distributions are shown by the solid lines and indicates that a few individuals do most of the reproducing. The Gini index (shown on the plots as GI) is

$$GI = 2 \times \text{Area between the two curves shown}$$

Equality means that $GI = 0$ while complete inequality means that $GI = 1$. As we expect from Fig. 1, the Gini is much smaller for Hutteries than the Hadza. And the extreme inequality of LRS for *T. canadensis* is reflected in a very high Gini. Surprisingly (to us) the LRS for a fairly large Roe deer yearling (size class 100) is also quite unequal, with a Gini value halfway between Hazda and Hutterite.

A general approach to the sensitivity of inequality in LRS should be feasible using our results in conjunction with the methods in van Raalte and Caswell (2013) and Caswell et al. (2018). More generally, numerical perturbations can be used to examine changes in Γ produced by changes in vital rates, and we can compare distributions using, e.g., the Kullback-Leibler distance (Cover and Thomas 2012).

A useful measure highlighted by our examples is the probability of having no reproductive success:

$$\Pr[\text{LRS} = 0] = \Gamma(0)$$

A general expression and computational method for this probability is given in the online Appendix (section A.8). Derivatives of this probability with respect to these parameters should be computable using known approaches, e.g., Caswell et al. (2018) or Steiner et al. (2012).

Another useful measure is the nonzero modal value, call it k^*, of the LRS – if it exists. We found a nonzero mode in several examples, but there is no nonzero mode for the tree (*T. canadensis*). Assuming its existence, a nonzero mode means that

$$\Gamma(k^* - 1) < \Gamma(k^*) > \Gamma(k^* + 1)$$

This modal value may be more meaningful in some populations than the commonly used average LRS. We can use our result to find k^* and the methods in Caswell et al. (2018) should be extendable to compute sensitivities.

Implications

In our age-structured applications each individual is governed by the same set of demographic rates (probabilities). In other words, all individuals have the same chances of surviving and reproducing at each age. Yet, the distribution of LRS is multimodal; this reveals how much individual heterogeneity can be generated by chance alone. This variation caused Steiner and Tuljapurkar (2012) to conclude that entirely neutral processes could generate the empirically-observed multimodal distributions of LRS so frequently reported.

In our stage and age+stage structured models individuals that start life

with different phenotypic trait values can have contrasting distributions of LRS. For each phenotype there is still a large amount of individual heterogeneity in LRS similar to that observed for our age-only models, but for individuals with different phenotypes we will find differing LRS distributions. These differences will generate opportunities for selection. Whether such selection would result in evolution will depend upon whether there is a genetic basis for individuals entering the population with different phenotypes.

Stage- and age+stage structured models are powerful. They can be constructed for stochastic, density-dependent, and frequency-dependent environments to predict population (Tuljapurkar 1990), life history (Tuljapurkar et al. 2009a; Caswell 2001), and evolutionary dynamics (Coulson et al. 2011, 2017; Childs et al. 2016). A large number of ecological and evolutionarily important statistics can be calculated from these models, with analysis of these models revealing how different values of these statistics arise. By extending the toolkit of structured models to predict entire distributions of LRS we open the door for a number of powerful, new analyses. For example, we can ask how different growth trajectories will impact distributions of LRS, we can gain insight into why weak associations between birth weight and LRS are so frequently observed, and the conditions when lifetime measures of reproduction act as a good, or bad, surrogates of fitness. In addition, we could use our approach to generate distributions of other per-generation life-history statistics such as generation times, or future life expectancies at a given age.

Previous analyses of LRS distributions in vertebrates relied on the standardized variance and have supported the existence of a link between mating

systems and sex differences in LRS distributions, with similar distributions in males and females of monogamous species and much more widespread and skewed distributions in males than in females in highly polygynous species (Andersson 1994). However, our findings demonstrate that only using the first few moments such as mean and variance does not capture important characteristics of an LRS distribution. More comprehensive metric such as the Gini index are much more appropriate and should be used in future comparisons of LRS distributions. Our approach should thus allow accurate comparisons among LRS distributions that are currently accumulating thanks to long-term individual-based monitoring of populations in a broad range of species across the tree of life. Moreover, in addition to sex and species, our approach provides a suitable way to assess how ecological factors (e.g. environmental harshness), life history tactics (e.g. slow vs. fast life cycle, capital vs. income breeding), and population structure (e.g. in terms of age, size) influence LRS distributions within and among populations.

The ability of the discrete-time structured model to gain ecological, life history, and evolutionary understanding is remarkable, and with each year new uses for them are developed. Such utility is not surprising, as these models (Coulson et al. 2011) include the fundamental biological processes of development, inheritance (although not considered in the current paper), survival and reproduction, and these processes can be specified to vary across individuals as a function of age, sex, phenotype, genotype (although not considered in the current paper) and environment. It is consequently unsurprising that analysis of such models has generated so much understanding

in recent years. The methods we develop here open up opportunities for yet more biological insight.

Acknowledgements

We thank Ni Lao, Hal Caswell, and Maarten Broekmann for useful comments.

$a = 1, \ldots, \omega$	Age, discrete intervals. Newborns are in the first interval $a = 1$. ω is maximum age.
$s = 1, \ldots, S$	Stage, discrete categories. S is the total number of stages.
$\mathbf{G}$, $0 \leq \mathbf{G}(i,j) \leq 1$, $i, j = 1, 2, \ldots, S$,	Stage only transition matrix, probability individual in stage j will be in stage i in the next time interval, conditional on survival. $S \times S$ matrix, column stochastic.
$\mathbf{G}_a$, $0 \leq \mathbf{G}_a(i,j) \leq 1$, $i, j = 1, 2, \ldots, S$	Age+stage transition matrix, individual age a in stage j will be age $a+1$ in stage i, conditional on survival. $S \times S$ matrix, column stochastic.
p_a, $0 \leq p_a \leq 1$	One-period survival (age only) from age a to age $a+1$. $p_\omega = 0$.
$p(j)$, $0 \leq p(j) \leq 1$, $j = 1, 2, \ldots, S$	One-period survival (stage only) from stage j to any other stage.
$p_a(j)$, $0 \leq p_a(j) \leq 1$, $j = 1, 2, \ldots, S$	One-period survival (age+stage) from age a and stage j to the next time interval. $p_\omega(j) = 0$ for all j.
l_a, $0 \leq l_a \leq 1$	Survivorship to age a, probability of being alive at age a for a. $l_1 = 1$, $l_2 = p_1\ l_1$, $l_3 = p_2\ l_2$, $\ldots$, $l_{\omega+1} = 0$.
$\boldsymbol{L}_a$, $0 \leq \boldsymbol{L}_a(i,j) \leq 1$, $i, j = 1, 2, \ldots, S$, $a = 1, 2, \ldots, \omega$	Survivorship to age a by stage, probability of being alive in stage i at age a for those born into stage j. $S \times S$ matrix, changes with age (see text).

$\phi_a,\ 0 \leq \phi_a \leq 1,\ a = 1, 2, \ldots, \omega,\ \sum_a \phi_a = 1$	Probability of age at death, the proportion of a cohort that dies at each age a. $\phi_a = l_a\ (1 - p_a)$ for age-only model (for stage-only and age+stage models see text).
$\mathbf{U},\ 0 \leq \mathbf{U}(i,j) \leq p(j),\ i,j = 1, 2, \ldots, S,$ $\mathbf{U}(i,j) = \mathbf{G}(i,j)\ p(j)$	Cohort projection matrix (stage only), probability an individual in stage j is alive in stage i at the next time interval. $S \times S$ matrix, does not change with age.
$\mathbf{U}_a,\ 0 \leq \mathbf{U}_a(i,j) \leq p_a(j), i,j = 1, 2, \ldots, S,$ $\mathbf{U}_a(i,j) = \mathbf{G}_a(i,j)\, p_a(j)$	Cohort projection matrix (age + stage), probability an individual age a in stage j is alive at age $a+1$ in stage i. $S \times S$ matrix, may change with age.
$\boldsymbol{\kappa}_a$	Vector of probabilities that an individual age a produces k offspring, $\kappa_a(k)$, where $0 \leq \kappa_a(k) \leq 1$, $\sum_k \kappa_a(k) = 1$, and $a = 1, 2, \ldots, \omega$.
$\boldsymbol{\kappa}_j$	Vector of probabilities that an individual in stage j produces k offspring at a single time interval, $\kappa_j(k)$, where $0 \leq \kappa_j(k) \leq 1$, $\sum_k \kappa_j(k) = 1$, and $j = 1, 2, \ldots, S$.
$\boldsymbol{\kappa}_{a,j}$	Vector of probabilities an individual at age a and stage j produces k offspring, $\kappa_{a,j}(k)$, where $0 \leq \kappa_{a,j}(k) \leq 1$, $\sum_k \kappa_{a,j}(k) = 1$, and $a = 1, 2, \ldots, \omega$. For each a, an $k_m \times S$ matrix(each column is for a stage j).
$\Gamma(k),\ 0 \leq \Gamma(k) \leq 1,$ $\sum_k \Gamma(k) = 1$	Probability an individual produces k offspring during its lifetime.

Γ	Vector of probabilities, $\Gamma(0)$, $\Gamma(1)$, $\Gamma(2)$, $\ldots$, $\Gamma(k_U)$. k_U = maximum possible number of offspring.
$\boldsymbol{\gamma}_a = \phi_a \ \boldsymbol{\kappa}_1 \star \boldsymbol{\kappa}_2 \star \boldsymbol{\kappa}_3, \cdots \star \boldsymbol{\kappa}_a$	Vector of probabilities of LRS for individuals who live to age a for $a < \beta$.
$\boldsymbol{\gamma}_\beta = l(\beta) \ \boldsymbol{\kappa}_1 \star \boldsymbol{\kappa}_2 \star \boldsymbol{\kappa}_3, \cdots \star \boldsymbol{\kappa}_\beta$	Vector of probabilities of LRS for individuals who live to age β and beyond.
$\Gamma = \sum_{a=1}^{a=\beta} \boldsymbol{\gamma}_a$	Vector of probabilities of LRS, from summing across contributions by each age, up to and including the last reproductive age. (same as Γ above)

Table 1: Definitions

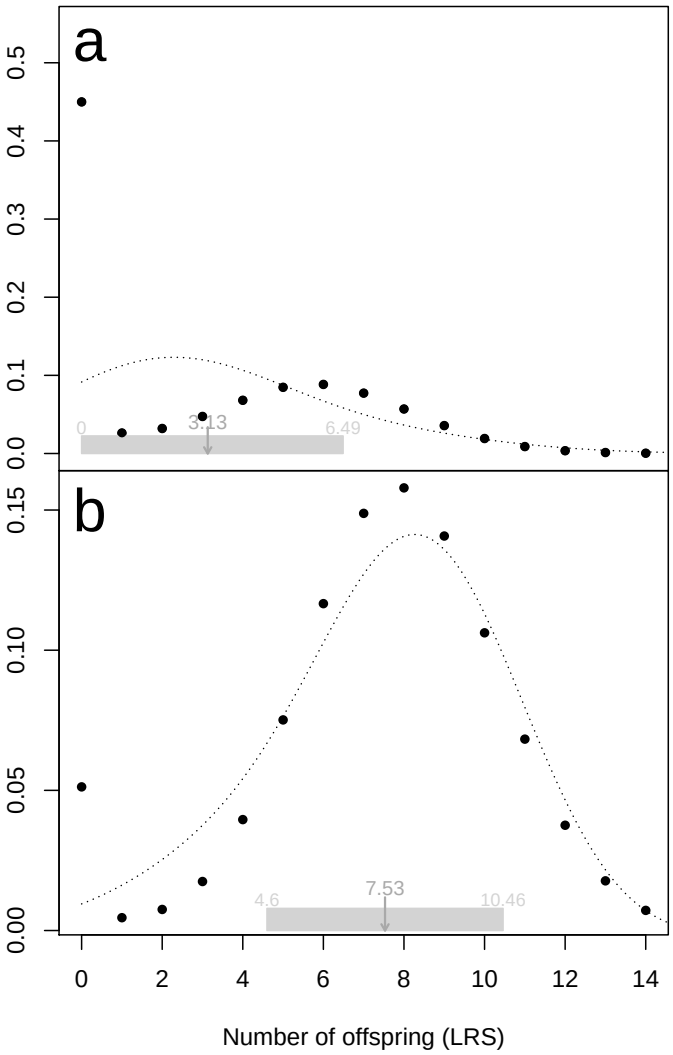

Figure 1: The LRS distribution of Hadza (a) and Hutterites (b) [solid points], and a Gram-Charlier approximation based on the first three moments [dotted lines].

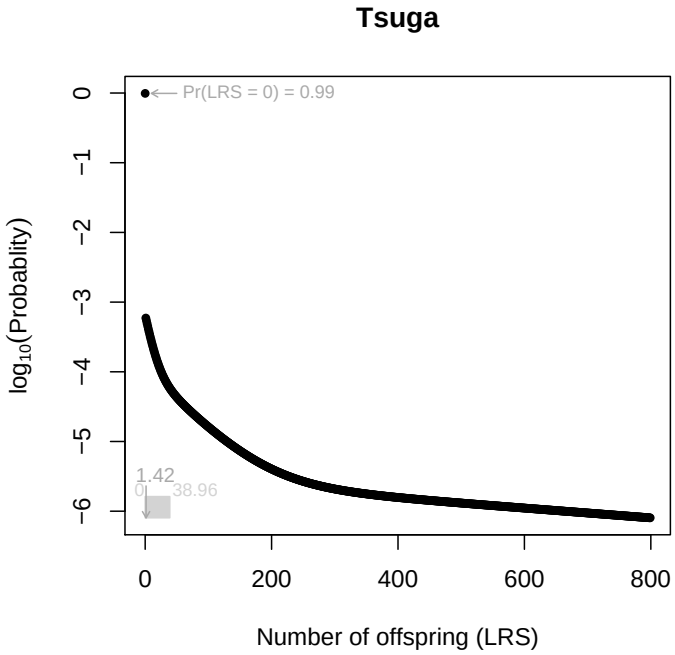

Figure 2: The LRS distribution of *Tsuga canadensis*, eastern hemlock.

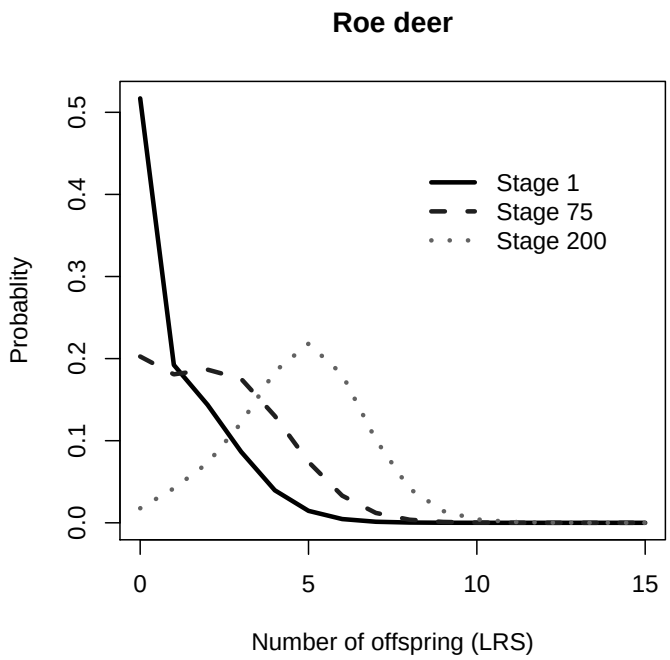

Figure 3: The LRS distribution of yearlings of Roe deer *Capreolus capreolus* for different initial size classes of 1, 75 or 200.

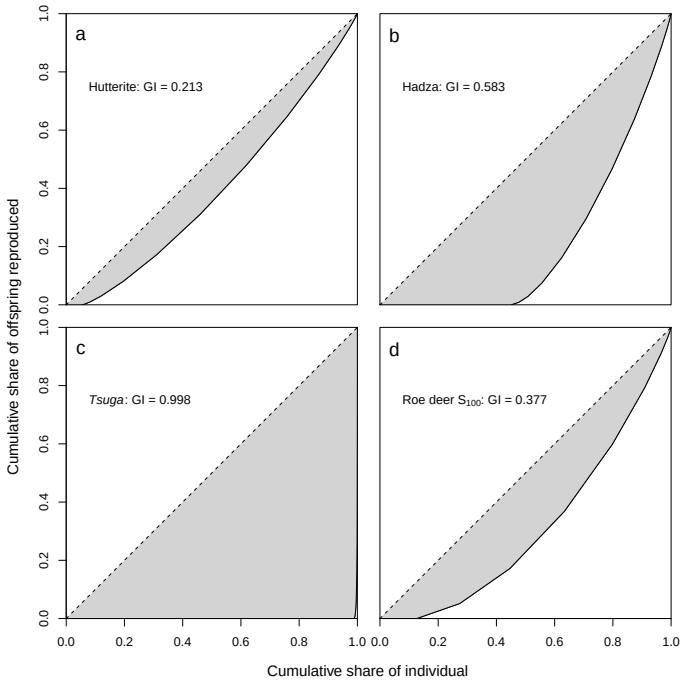

Figure 4: The Lorenz plots for Hutterite, Hadza, *T. canadensis*, and Roe deer (with initial stage 100).

References

Andersson, Malte B (1994). *Sexual Selection*. Princeton: Princeton University Press.

Archer, C. Ruth et al. (2018). "Diet has independent effects on the pace and shape of aging in Drosophila melanogaster". In: *Biogerontology* 19.1, pp. 1–12. ISSN: 1573-6768. DOI: 10.1007/s10522-017-9729-1. URL: https://doi.org/10.1007/s10522-017-9729-1.

Brommer, Jon E. et al. (2004). "Single-Generation Estimates of Individual Fitness as Proxies for Long-Term Genetic Contribution". In: *The American Naturalist* 163.4, pp. 505–517. ISSN: 00030147, 15375323. URL: http://www.jstor.org/stable/10.1086/382547.

Cabana, Gilbert and Donald L. Kramer (1991). "Random offspring Mortality and Variation in Parental Fitness". In: *Evolution* 45.1, pp. 228–234. ISSN: 00143820, 15585646. URL: http://www.jstor.org/stable/2409498.

Caswell, Hal (2001). *Matrix Population Models: Construction, Analysis, and Interpretation.* 2nd. Sinauer Associates Inc., Sunderland, MA.

— (2011). "Beyond R0: demographic models for variability of lifetime reproductive output". In: *PloS one* 6.6, e20809.

Caswell, Hal et al. (2018). "Age× stage-classified demographic analysis: a comprehensive approach". In: *Ecological monographs* 88.4, pp. 560–584.

Charlesworth, Brian and John A Williamson (1975). "The probability of survival of a mutant gene in an age-structured population and implications for the evolution of life-histories". In: *Genetics Research* 26.1, pp. 1–10.

Childs, Dylan Z., Ben C. Sheldon, and Mark Rees (2016). "The evolution of labile traits in sex- and age-structured populations". In: *Journal of Animal Ecology* 85.2, pp. 329–342. DOI: 10.1111/1365-2656.12483. eprint: https://besjournals.onlinelibrary.wiley.com/doi/pdf/10.1111/1365-2656.12483. URL: https://besjournals.onlinelibrary.wiley.com/doi/abs/10.1111/1365-2656.12483.

Clutton-Brock, T.H. (ed.) (1988). *Reproductive success: studies of individual variation in contrasting breeding systems*. Chicago: University of Chicago press.

Coulson, Tim (2012). "Integral projections models, their construction and use in posing hypotheses in ecology". In: *Oikos* 121.9, pp. 1337–1350.

Coulson, Tim et al. (2011). "Modeling Effects of Environmental Change on Wolf Population Dynamics, Trait Evolution, and Life History". In: *Science* 334.6060, pp. 1275–1278. ISSN: 0036-8075. DOI: 10.1126/science.1209441. eprint: https://science.sciencemag.org/content/334/6060/1275.full.pdf. URL: https://science.sciencemag.org/content/334/6060/1275.

Coulson, Tim et al. (2017). "Modeling Adaptive and Nonadaptive Responses of Populations to Environmental Change". In: *The American Naturalist* 190.3. PMID: 28829647, pp. 313–336. DOI: 10.1086/692542. eprint: https://doi.org/10.1086/692542. URL: https://doi.org/10.1086/692542.

Cover, Thomas M and Joy A Thomas (2012). *Elements of information theory*. John Wiley & Sons.

Crow, J F (1958). "Some possibilities for measuring selection intensities in man." In: *Human biology; an international record of research* 30.1, pp. 1–13.

Dobson, A. P. (1986). "Inequalities in the individual reproductive success of parasites". In: *Parasitology* 92.3, pp. 675–682. DOI: 10.1017/S0031182000065537.

Dubuc, Constance, William L. Allen, Dario Maestripieri, and James P. Higham (2014). "Is male rhesus macaque red color ornamentation attractive to females?" In: *Behavioral Ecology and Sociobiology* 68.7, pp. 1215–1224. ISSN: 1432-0762. DOI: 10.1007/s00265-014-1732-9. URL: https://doi.org/10.1007/s00265-014-1732-9.

Ellner, Stephen, Dylan Childs, and Mark Rees (2016). *A practical guide to the Integral Projection Model*. Springer.

Fisher, RA (1930). *The genetical theory of natural selection*. Oxford University Press.

Grafen, Alan (1988). "On the uses of data on lifetime reproductive success". In: *Reproductive success: studies of individual variation in contrasting breeding systems*. Ed. by T. H. Clutton-Brock. Chicago: The University of Chicago Press, pp. 454–471.

Healy, Kevin et al. (2019). "Beyond the fast-slow continuum in animal life-histories". In: *Nature Ecology and Evolution* In print.

Kruuk, L. E. B., T. H. CluttonBrock, K. E. Rose, and F. E. Guinness (1999). "Early determinants of lifetime reproductive success differ between the sexes in red deer". In: *Proceedings of the Royal Society of London Series B-Biological Sciences* 266.1429, pp. 1655–1661.

Lamar, W Robert and James B McGraw (2005). "Evaluating the use of remotely sensed data in matrix population modeling for eastern hemlock (Tsuga canadensis L.)" In: *Forest Ecology and Management* 212.1-3, pp. 50–64.

Lefkovitch, LP (1965). "The study of population growth in organisms grouped by stages". In: *Biometrics*, pp. 1–18.

Matlack, Glenn R. (1994). "Vegetation Dynamics of the Forest Edge – Trends in Space and Successional Time". In: *Journal of Ecology* 82.1, pp. 113–123. ISSN: 00220477, 13652745. URL: `http://www.jstor.org/stable/2261391`.

Mersmann, Olaf (2019). *fftw: Fast FFT and DCT Based on the FFTW Library*. R package version 1.0-5. URL: `https://CRAN.R-project.org/package=fftw`.

Newton, I. (1988). "Age and reproduction in the Sparrowhawk". In: *Reproductive success: studies of individual variation in contrasting breeding systems*. Ed. by T. H. Clutton-Brock. Chicago: The University of Chicago Press, pp. 201–219.

Plard, Floriane et al. (2015). "Quantifying the influence of measured and unmeasured individual differences on demography". In: *Journal of Animal Ecology* 84.5, pp. 1434–1445.

R Core Team (2019). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. URL: `https://www.R-project.org/`.

Shkolnikov, Vladimir M., Evgeny M. Andreev, and Alexander Begun (2003). "Gini coefficient as a life table function: Computation from discrete data, decomposition of differences and empirical examples". In: *Demographic Research* 8.11, pp. 305–358. DOI: 10.4054/DemRes.2003.8.11. eprint: https://www.demographic-research.org/volumes/vol8/11/8-11.pdf. URL: https://www.demographic-research.org/volumes/vol8/11/.

Shostak, Allen W. and Terry A. Dick (1987). "Individual variability in reproductive success of Triaenophorus crassus Forel (Cestoda: Pseudophyllidea), with comments on use of the Lorenz curve and Gini coefficient". In: *Canadian Journal of Zoology* 65.12, pp. 2878–2885. DOI: 10.1139/z87-437. eprint: https://doi.org/10.1139/z87-437. URL: https://doi.org/10.1139/z87-437.

Slate, J. et al. (2000). "Inbreeding Depression Influences Lifetime Breeding Success in a Wild Population of Red Deer (Cervus elaphus)". In: *Proceedings: Biological Sciences* 267.1453, pp. 1657–1662. ISSN: 09628452. URL: http://www.jstor.org/stable/2665912.

Snyder, Robin and Stephen Ellner (2018). "Pluck or Luck: Does Trait Variation or Chance Drive Variation in Lifetime Reproductive Success?" In: *The American Naturalist* 191.4, E90–E107.

Steiner, U.K. and S. Tuljapurkar (2012). "Neutral theory for life histories and individual variability in fitness components". In: *Proceedings of the National Academy of Sciences* 109, pp. 4684–4689.

Steiner, U.K., S. Tuljapurkar, T. Coulson, and C. Horvitz (2012). "Trading Stages: Life Expectancies in Structured Populations". In: *Experimental Gerontology*.

Stuart, Alan, J Keith Ord, and Sir Maurice Kendall (1994). *Distribution theory*. Edward Arnold; New York.

Tatarenkov, Andrey, Christiane I. M. Healey, Gregory F. Grether, and John C. Avise (2008). "Pronounced reproductive skew in a natural population of green swordtails, Xiphophorus helleri". In: *Molecular Ecology* 17.20, pp. 4522–4534. DOI: 10.1111/j.1365-294X.2008.03936.x. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-294X.2008.03936.x. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-294X.2008.03936.x.

Tuljapurkar, S. (1990). *Population dynamics in variable environments*. New York: Springer-Verlag, Berlin.

Tuljapurkar, S., JM Gaillard, and T. Coulson (2009a). "From stochastic environments to life histories and back." In: *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 364.1523, pp. 1499–1509.

Tuljapurkar, S., U.K. Steiner, and S.H. Orzack (2009b). "Dynamic heterogeneity in life histories". In: *Ecology Letters* 12.1, pp. 93–106.

van Daalen, Silke F. and Hal Caswell (2017). "Lifetime reproductive output: individual stochasticity, variance, and sensitivity analysis". In: *Theoretical Ecology* 10.3, pp. 355–374.

Van Daalen, Silke and Hal Caswell (2015). "Lifetime reproduction and the second demographic transition: Stochasticity and individual variation". In: *Demographic Research* 33, pp. 561–588.

Van Noordwijk, Maria A. and Carcl P. van Schaik (1999). "The effects of dominance rank and group size on female lifetime reproductive success in wild long-tailed macaques,Macaca fascicularis". In: *Primates* 40.1, pp. 105–130. ISSN: 1610-7365. DOI: 10.1007/BF02557705. URL: https://doi.org/10.1007/BF02557705.

Van Raalte, Alyson A. and Hal Caswell (2013). "Perturbation Analysis of Indices of Lifespan Variability". In: *Demography* 50.5, pp. 1615–1640. ISSN: 1533-7790. DOI: 10.1007/s13524-013-0223-3. URL: https://doi.org/10.1007/s13524-013-0223-3.

Waples, Robin S, Chi Do, and Julien Chopelet (2011). "Calculating Ne and Ne/N in age-structured populations: a hybrid Felsenstein-Hill approach". In: *Ecology* 92.7, pp. 1513–1522.

Weiner, Jacob and Otto T. Solbrig (1984). "The meaning and measurement of size hierarchies in plant populations". In: *Oecologia* 61.3, pp. 334–336. ISSN: 1432-1939. DOI: 10.1007/BF00379630. URL: https://doi.org/10.1007/BF00379630.

Appendix

Shripad Tuljapurkar*[1], Wenyun Zuo[1], Tim Coulson[2], Carol Horvitz[3], and Jean-Michel Gaillard[4]

[1]Department of Biology, Stanford University, Stanford, CA 94305-5020, USA
[2]Department of Zoology, University of Oxford, Oxford, OX1 3SZ, UK
[3]Department of Biology, University of Miami, Coral Gables, FL 33124-0421 USA.
[4]Laboratoire de Biométrie et Biologie Evolutive, Université Lyon 1, CNRS, UMR 5558, F-69622 Villeurbanne, France

A.1 Discrete Convolutions and Age structure

Consider an age-structured population with only two age classes. Say that an individual of age class 1 has some number of offspring $i_1 = 0, 1, \ldots, k_m$ with probabilities $\boldsymbol{\kappa}_1 = \{\kappa_1(i_1)\}$, and an individual of age class 2 has $i_2 = 0, 1, \ldots, k_m$ offspring with probabilities $\boldsymbol{\kappa}_2 = \{\kappa_2(i_2)\}$. Now consider an individual who may produce offspring at both ages 1 and 2, and call $\delta(n)$ the probability that such an individual produces n offspring, $n = 0, 1, 2, \ldots, 2k_m$. Given any n, some m_1 offspring are produced at age 1 and the rest, $k_2 = (n - m_1)$, must be produced at age 2. The number of ways to obtain a given n varies with n. For example, there is only one way to obtain $n = 0$ (both

*corresponding author: tulja@stanford.edu

$m_1 = 0$ and $m_2 = 0$). Similarly, there is only one way for $n = 2k_m$ (both $m_1 = k_m$ and $m_2 = k_m$). There are only two ways for $n = 1$ (either $m_1 = 1$ and $m_2 = 0$ or $m_1 = 0$ and $m_2 = 1$). In general, we sum over the probabilities of the different ways we could get a particular n

$$\delta(n) = \sum \kappa_1(m_1)\, \kappa_2(n - m_1), \tag{A.1}$$

where the number of terms in the summation varies from 1 to a maximum of $k_m + 1$. The probabilities $\{\delta(n)\}$ are components of a distribution that we write as $\boldsymbol{\delta}$. Also, the fact that the first argument in equation (A.1) is m_1 and the second is $(n - m_1)$ means that $\boldsymbol{\delta}$ is what is called a discrete convolution of the distributions $\boldsymbol{\kappa}_1$ and $\boldsymbol{\kappa}_2$, written as

$$\boldsymbol{\delta} = \boldsymbol{\kappa}_1 \star \boldsymbol{\kappa}_2. \tag{A.2}$$

Alternatively we can describe $\boldsymbol{\kappa}_1$ and $\boldsymbol{\kappa}_2$ by the probability generating functions

$$w_1(x) = \sum_{i_1} \kappa_1(i_1) x^{i_1},\ w_2(x) = \sum_{i_2} \kappa_2(i_2) x^{i_2}, \tag{A.3}$$

where x is a dummy variable. Let the probability generating function for $\boldsymbol{\delta}$ be

$$u(x) = \sum_n \delta(n) x^n. \tag{A.4}$$

Setting

$$u(x) = w_1(x)\, w_2(x), \tag{A.5}$$

also (algebra!) yields the probabilities in equation (A.1). Thus there is a correspondence between convolutions of distributions as in equation (A.2)) and products of generating functions (as in equation (A.5)).

Using equation (A.3), we see that the distribution $\boldsymbol{\kappa}_1 \star \boldsymbol{\kappa}_1$ has a generating function $[w_1(x)]^2$. As a result, we write $\boldsymbol{\kappa}_1 \star \boldsymbol{\kappa}_1$ as $\boldsymbol{\kappa}^2$. Also, $\boldsymbol{\kappa}_1 \star \boldsymbol{\kappa}_2 = \boldsymbol{\kappa}_2 \star \boldsymbol{\kappa}_1$. Finally, note that convolutions can be rapidly computed (for discrete and continuous functions).

For age structured populations, we can use repeated convolutions to compute the LRS distribution as in the main text. But this method doesn't work for stage structure so we need a different approach that starts with the property discussed next.

A.2 Fourier transforms

Fourier transforms require complex numbers. Denote

$$\mathbf{i} = \sqrt{-1}. \tag{A.6}$$

For an integer N, define the frequencies

$$\theta(z) = \exp\Big[\frac{-2\pi z \mathbf{i}}{N}\Big], z = 0, 1, \ldots, N-1. \tag{A.7}$$

Then for any integer q, define the discrete function $\boldsymbol{\kappa} = \{\kappa(0), \kappa(1), \kappa(2), \ldots, \kappa(q)\}$ (this could be a probability distribution). Then we define the Discrete Fourier Transform (DFT) of $\boldsymbol{\kappa}$ as a function $\widehat{\kappa}$, whose values are given for each frequency in equation (A.7) as,

$$\widehat{\kappa}(\theta(z)) = \sum_{b=0}^{q} \kappa(b)\, \theta(z)^b. \tag{A.8}$$

Now suppose that $\boldsymbol{\psi}$ is another series of numbers, $\boldsymbol{\psi} = \{\psi(0), \psi(1), \ldots, \psi(q)\}$. For frequencies $\theta(z)$, we can use equation (A.8) to find the Fourier transform $\widehat{\psi}(\theta(z))$. It is well known that (James 2002)

$$\boxed{\text{the Fourier transform of } \{\boldsymbol{\kappa} \star \boldsymbol{\psi}\} \text{ is the product } \widehat{\kappa}\,\widehat{\psi}}$$

In consequence, for any frequency $\theta(z)$ in equation (A.7): for any integer $m \geq 0$,

$$\widehat{\big(\kappa^m\big)}(\theta(z)) = \big(\widehat{\kappa}(\theta(z))\big)^m, \tag{A.9}$$

and for any non-negative integers m_a, m_b

$$\big(\widehat{\kappa^{m_a} \star \psi^{m_b}}\big)(\theta(z)) = \widehat{\kappa}(\theta(z))^{m_a}\, \widehat{\psi}(\theta(z))^{m_b}. \tag{A.10}$$

A.3 Digression: A Generating Function

We consider a general age+stage model following Steiner and Tuljapurkar (2012). Age and stage are used together: given A ages and S discrete stages, there are $(A \times S) = S_1$ unique combinations, so that we can proceed as if we had a stage-only model with S_1 stages. Let there be S_1 living stages, with transitions between them with probabilities $\mathbf{U}$ (Table 1 in main text). Form

the matrix $\mathbf{U}^T = \mathbf{Q}$. The element q_{ij} of matrix $\mathbf{Q}$ is the probability of being in stage i in one time interval and then surviving to reach stage j in the next time interval. The stages that are rows (or columns) of $\mathbf{Q}$ are transient, meaning that an individual starts in any stage then makes transitions to other stages until an eventual but certain death. Starting in a particular stage i let τ_j be the (random) time spent in stage j before death. (In the main text we use matrix $\mathbf{U}$, but here, we are using the transposed matrix $\mathbf{Q}$ and so reverse rows and columns).

The death probabilities for stages 1 to S are components d_i of a vector $\mathbf{d}$. These death rates are of course simply probabilities of not reaching another transient stage, so e.g.,

$$d_i = 1 - \sum_j q_{ij}.$$

We summarize by defining

$$\mathbf{e} = \text{ a column vector whose every entry is 1,} \tag{A.11}$$

and writing $\mathbf{I}$ for the identity matrix, so that

$$\mathbf{d} = \big(\mathbf{I} - \mathbf{Q}\big)\,\mathbf{e}. \tag{A.12}$$

Consider the possible outcome in which an individual spends a time m_j in stage j. In our notation, this outcome means that $\tau_j = m_j$, conditional on starting in stage i, and has probability,

$$P_i[m_1, m_2, \ldots, m_S] = \Pr[\tau_1 = m_1, \tau_2 = m_2, \ldots, \tau_S = m_S| \text{ starting stage } i]. \tag{A.13}$$

Let $w_i, i = 1, \ldots, S$ be a set of (dummy) variables, and make a generating function for the above probabilities,

$$V_i = \sum_{m_1, m_2, \ldots, m_S} P_i[m_1, m_2, \ldots, m_S]\, w_1^{m_1}\, w_2^{m_2} \ldots w_S^{m_S}. \tag{A.14}$$

Say we start in stage 1, so $i = 1$. Then $V_i = V_1$ is a polynomial, whose first term is $w_1 d_1$ – meaning an individual starts in 1 but dies before making a transition. The second term of V_1 is

$$w_1 q_{11} w_1 d_1 + w_1 q_{12} w_2 d_2 + \cdots + w_1 q_{1S} w_S d_S,$$

meaning an individual starts in stage 1, lives one period and reaches some stage j, but then dies before making a second transition. And so on.

Following Steiner and Tuljapurkar (2012), form the diagonal matrix

$$\mathbf{W} = \text{diag}\{w_i\}, i = 1, \ldots, S. \tag{A.15}$$

Define the vector function

$$\mathbf{V} = \mathbf{W}\left(\mathbf{I} - \mathbf{Q}\,\mathbf{W}\right)^{-1}\mathbf{d}, \tag{A.16}$$

where $\mathbf{I}$ is the identity matrix and $\mathbf{d}$ is defined in equation (A.12).

Let the components of $\mathbf{V}$ in equation (A.16) be $\{V_i, i = 1, \ldots, S\}$. Then this component V_i is indeed the generating function defined earlier in equation (A.14). The reader can derive this by expanding the inverse in equation (A.16) and seeing, e.g., that the resulting polynomial V_1 has the same terms as we describe above.

A.4 To the LRS

We now combine convolutions, the generating function, and Fourier transforms.

Start by defining frequencies. Choose a number N that is the largest possible number of offspring. This may be approximated by a product like $k_U \leq (\omega \times k_m)$ with ω being an estimated maximum age, and k_m being the maximum number of offspring per time period. Or a biological upper limit can be used. Either way, we can test the adequacy of our choice by computing Γ and then its sum, which should be very close to 1.

Given N, for every stage i we pad each offspring probability distribution $\boldsymbol{\kappa}_i$ by adding zeros so that the padded distribution has N elements. For each frequency $\theta(z)$, as in equation (A.7), let the Fourier transforms of the padded distributions $\boldsymbol{\kappa}_i$ be $\widehat{\boldsymbol{\kappa}}_i$, so for example

$$\widehat{\boldsymbol{\kappa}}_i(\theta(z)) = \sum_{k=0}^{N-1} \kappa_i(k)\,\theta(z)^k. \tag{A.17}$$

For any non-reproducing stage i we have $\kappa_i = \{1, 0, \ldots\}$, and so

$$\widehat{\boldsymbol{\kappa}}_i(\theta(z)) = 1, \text{ for all } \theta(z) \text{ for non-reproducing stages.} \tag{A.18}$$

Now as in the previous subsection, say that, starting in stage i an individual spends time m_j in each stage j before death (for stages $j = 1, 2, \ldots$). The total offspring numbers that result have the distribution

$$\boldsymbol{\kappa}_1^{m_1} \star \boldsymbol{\kappa}_2^{m_2} \star \cdots \star \boldsymbol{\kappa}_S^{m_S} = \boldsymbol{\delta}. \tag{A.19}$$

So in this case

$$\Pr[\text{number of offspring is } k|\text{ starting stage } i]$$

is given by the kth element of the convolution δ in equation (A.19)

The lifetime probability of having k offspring is the kth element $\gamma(k)$ of

$$\Gamma = \sum_{m_1,m_2,\ldots,m_S} P[m_1, m_2, \ldots, m_S|\text{ starting stage } i]\, \boldsymbol{\kappa}_1^{m_1} \star \boldsymbol{\kappa}_2^{m_2} \star \cdots \star \boldsymbol{\kappa}_S^{m_S}. \tag{A.20}$$

Here Γ can be thought of as a function of the discrete argument $k = 0, 1, \ldots$ with values $\gamma(0), \gamma(1), \ldots$. This is reminiscent of equation (A.14), but unfortunately here we have convolutions and not products of numbers. This is where Fourier transforms come in.

We denote the Fourier transform of this function by $\widehat{\Gamma}$. Note that this latter transformed function is defined for each discrete frequency; so if the frequencies are $\theta(0), \theta(1), \ldots$, the Fourier transform is the set of (scalar) values $\widehat{\Gamma}(\theta(0)), \widehat{\Gamma}(\theta(1)), \ldots$.

Using the facts about Fourier transforms and equation (A.20) we conclude that

$$\widehat{\Gamma}(\theta) = \sum_{m_1,m_2,\ldots,m_S} P[m_1, m_2, \ldots, m_S|\text{ starting stage } i]\, \widehat{\boldsymbol{\kappa}_1}(\theta)^{m_1} \ldots \widehat{\boldsymbol{\kappa}_S}(\theta)^{m_S}. \tag{A.21}$$

Now this is indeed similar to equation (A.14), so the sum here can be exactly computed using the generating function equation (A.16).

A.5 Method for Stages

Start with choosing a large number N, fix the starting stage i, then find the frequencies as in equation (A.7). For each frequency $\theta(z)$ and every stage j, compute the Fourier transforms $\widehat{\boldsymbol{\kappa}_j}(\theta(z))$, and then make the matrix

$$\widetilde{\mathbf{W}} = \begin{pmatrix} \widehat{\boldsymbol{\kappa}_1}(\theta(z)) & 0 & 0 & \ldots \\ 0 & \widehat{\boldsymbol{\kappa}_2}(\theta(z)) & 0 & \ldots \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix}. \tag{A.22}$$

Remembering the generating function, use equations (A.21) and (A.22) to compute the quantity

$$\mathbf{V} = \widetilde{\mathbf{W}}\Big(\mathbf{I} - \mathbf{Q}\,\widetilde{\mathbf{W}}\Big)^{-1}\mathbf{d} \text{ with components } \{V_i\}.$$

Then use equation (A.21) to conclude that

$$\widehat{\Gamma}(\theta(z)) = V_i. \tag{A.23}$$

Repeat for all frequencies. Thus one computes the set of (scalar) values $\widehat{\Gamma}(\theta(0)), \widehat{\Gamma}(\theta(1)), \ldots$ and hence the function $\widehat{\Gamma}$.

Final step: use the Inverse FFT to find Γ.

Note: since every age+stage model can be cast as a stage-only model, this method can also be used for age-only or age+stage models. But the age-structured analysis in the main text explains the basic logic that the general method does not. The age-based method is often computationally useless for stage based or age+stage models when we have a large number of stages, or when individuals can stay in a stage for a long time. In such cases, the method here is essential.

A.6 Block method for age+stage

An age+stage model has ages $a = 1, 2, \ldots, \omega$, stages $s = 1, 2, \ldots, S$. A unique age+stage combination is written a, s, and there are $A \times S$ such combinations. In some cases, the general method of the preceding section may be computationally lengthy and the method described below is faster.

A.6.1 Inverse

To make it efficient to find the generating function, start by computing the inverse of the matrix

$$\mathbf{H} = \left(\mathbf{I} - \mathbf{Q}\,\widetilde{\mathbf{W}}\right), \tag{A.24}$$

in which $\mathbf{Q} = \mathbf{U}^T$. Here $\widetilde{\mathbf{W}}$ is a diagonal matrix as in equation (A.22). Suppose that we have S stages at each age, and that ages $a = 1, 2, \ldots, A-1$ have corresponding and distinct $S \times S$ transition matrices $\mathbf{U}_a = \mathbf{Q}_a^T$. The transition matrix $\mathbf{U}_A = \mathbf{Q}_A^T$ for age A is repeating, and applies at age $A + 1, A+2, \ldots$ until death. Note that for this case, there are total $A * S$ stages. Write the diagonal elements of $\widetilde{\mathbf{W}}$ for a given frequency, $\theta(z)$,as

$$\widetilde{\mathbf{W}} = (\widetilde{\mathbf{W}}_1, \widetilde{\mathbf{W}}_2, \ldots, \widetilde{\mathbf{W}}_A), \tag{A.25}$$

$$\widetilde{\mathbf{W}}_a = (\widehat{\boldsymbol{\kappa}_{a,1}}(\theta(z)), \widehat{\boldsymbol{\kappa}_{a,2}}(\theta(z)), \ldots, \widehat{\boldsymbol{\kappa}_{a,S}}(\theta(z))), \tag{A.26}$$

and the matrix of transition probabilities as

$$\mathbf{Q} = \begin{pmatrix} \mathbf{0} & \mathbf{Q}_1 & \mathbf{0} & \dots \\ \mathbf{0} & \mathbf{0} & \mathbf{Q}_2 & \dots \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{Q}_{A-1} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{Q}_A \end{pmatrix}. \tag{A.27}$$

We can write the matrix of equation (A.24) in block form as

$$\mathbf{H} = \begin{pmatrix} \mathbf{I} & \mathbf{B}_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{B}_2 & \dots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{I} & \dots & \mathbf{B}_{A-1} \\ \mathbf{0} & \mathbf{0} & \dots & & \mathbf{C} \end{pmatrix}, \tag{A.28}$$

where

$$\mathbf{B}_a = -\mathbf{Q}_a \widetilde{\mathbf{W}}_{a+1}, (a < A), \tag{A.29}$$

$$\mathbf{C} = \mathbf{I} - \mathbf{Q}_A \widetilde{\mathbf{W}}_A. \tag{A.30}$$

Now use the result in Singh (1979) applied to block matrices. It is best to write this first in cases and then in general. First, say that $A = 5$ so that

$$\mathbf{H} = \begin{pmatrix} \mathbf{I} & \mathbf{B}_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{B}_2 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{B}_3 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{B}_4 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{C} \end{pmatrix}, \tag{A.31}$$

where

$$\mathbf{C} = \mathbf{I} - \mathbf{Q}_5 \widetilde{\mathbf{W}}_5. \tag{A.32}$$

Then

$$\mathbf{H}^{-1} = \begin{pmatrix} \mathbf{I} & -\mathbf{B}_1 & \mathbf{B}_1\mathbf{B}_2 & -\mathbf{B}_1\mathbf{B}_2\mathbf{B}_3 & \mathbf{B}_1\mathbf{B}_2\mathbf{B}_3\mathbf{B}_4\mathbf{C}^{-1} \\ \mathbf{0} & \mathbf{I} & -\mathbf{B}_2 & \mathbf{B}_2\mathbf{B}_3 & -\mathbf{B}_2\mathbf{B}_3\mathbf{B}_4\mathbf{C}^{-1} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} & -\mathbf{B}_3 & \mathbf{B}_3\mathbf{B}_4\mathbf{C}^{-1} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} & -\mathbf{B}_4\mathbf{C}^{-1} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{C}^{-1} \end{pmatrix}, \tag{A.33}$$

In general, the inverse matrix for equation (A.28) has the form

$$\mathbf{H}^{-1} = \begin{pmatrix} \mathbf{I} & \mathbf{J}_{1,2} & \mathbf{J}_{1,3} & \dots & \mathbf{J}_{1,A} \\ \mathbf{0} & \mathbf{I} & \mathbf{J}_{2,3} & \dots & \mathbf{J}_{2,A} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{I} & \mathbf{J}_{A-1,A} \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{C}^{-1} \end{pmatrix}, \tag{A.34}$$

Here,

$$\mathbf{J}_{i,j} = \left(-1\right)^{i+j} \left\{ \prod_{k=i}^{j-1} \mathbf{B}_k \right\}, \text{ for } i < j \leq (A-1), \tag{A.35}$$

$$\mathbf{J}_{i,A} = \left(-1\right)^{i+j} \left\{ \prod_{k=i}^{A-1} \mathbf{B}_k \right\} \mathbf{C}^{-1}, \text{ for } j = A. \tag{A.36}$$

A.6.2 The Generating Function

Fix a starting stage i out of all $A \times S$ combinations, where S is the number of stages in a given age and A is the age where the transition matrix $\mathbf{U}_A$ starts repeating for following age. Death rates by age+stage are elements (there are $A \times S$ elements) of the vector

$$\begin{aligned} \mathbf{d} &= \begin{pmatrix} \mathbf{d}_1 \\ \mathbf{d}_2 \\ \vdots \\ \mathbf{d}_A \end{pmatrix}, \\ \mathbf{d}_a &= \begin{pmatrix} \mathbf{d}_{a,1} \\ \mathbf{d}_{a,2} \\ \vdots \\ \mathbf{d}_{a,S} \end{pmatrix}. \end{aligned} \tag{A.37}$$

The generating function is

$$\mathbf{V} = \widetilde{\mathbf{W}}\, \mathbf{H}^{-1}\, \mathbf{d}, \tag{A.38}$$

We start with the product

$$\mathbf{H}^{-1}\, \mathbf{d}, \tag{A.39}$$

which is a vector of A blocks, each S-long.

Using the pattern in eqations (A.33 – A.36), we define a sequence:

$$\widehat{\mathbf{d}}_A = \mathbf{C}^{-1}\mathbf{d}_A, \tag{A.40}$$

$$\widehat{\mathbf{d}}_{A-1} = \mathbf{d}_{A-1} - \mathbf{B}_{A-1}\widehat{\mathbf{d}}_A, \tag{A.41}$$

$$\widehat{\mathbf{d}}_{A-2} = \mathbf{d}_{A-2} - \mathbf{B}_{A-2}\widehat{\mathbf{d}}_{A-1}, \tag{A.42}$$

$$\widehat{\mathbf{d}}_a = \mathbf{d}_a - \mathbf{B}_a\widehat{\mathbf{d}}_{a+1},\ 1 \leq a \leq (A-1). \tag{A.43}$$

Then set

$$\widehat{\mathbf{d}} = (\widehat{\mathbf{d}}_1, \widehat{\mathbf{d}}_2, \ldots, \widehat{\mathbf{d}}_A)^T, \tag{A.44}$$

so finally

$$\mathbf{H}^{-1}\,\mathbf{d} = \widehat{\mathbf{d}}. \tag{A.45}$$

Therefore the generating function is

$$\mathbf{V} = \{V_1, V_2, \ldots, V_{A\times S}\} = \widetilde{\mathbf{W}}\widehat{\mathbf{d}}. \tag{A.46}$$

This quantity gives us the Fourier Transform $\widehat{\Gamma}$ of the LRS distribution Γ – which is the distribution we want. So at the given frequency $\theta(k)$ we have found that

$$\widehat{\Gamma}(\theta(k)) = V_i. \tag{A.47}$$

Repeat for all frequencies. Thus one computes the set of (scalar) values $\widehat{\Gamma}(\theta(0)), \widehat{\Gamma}(\theta(1)), \ldots$ and hence the function $\widehat{\Gamma}$. The inverse Fourier transform then yields Γ.

A.7 Special probabilities: Ages only

To compute Pr[LRS = 0], observe that an individual has 0 offspring if it either (a) dies before reaching reproductive age α or (b) survives through the pre-reproductive period, but has 0 offspring thereafter. At reproductive ages a the probability of having 0 offspring is $\kappa_a(0)$. Assuming survival, the cumulative probability of having had no offspring by age a is

$$p'_a = \kappa_1(0)\,\kappa_2(0)\ldots\kappa_a(0). \tag{A.48}$$

Thus,

$$\begin{aligned}\Pr[\mathrm{LRS}=0] = \big(1-l_\alpha\big) &+ \\ &+ \sum_{a\geq\alpha}^{\beta-1} \phi_a\, p'_a + \\ &+ l_\beta\, p'_\beta. \end{aligned} \tag{A.49}$$

The first line gives us (a) above, the second line adds in reproduction at intermediate ages weighted by the odds of death, and the third line gives us the final contribution to childlessness.

A.8 Special probabilities: Stages

To compute $\Pr[\mathrm{LRS}=0]$ here, we need a different method. Each stage i has a corresponding probability that an individual in that stage produces 0 offspring. This probability is 1 for non-reproducing stages, and $\kappa_i(0)$ for all other stages i. Define $\mathbf{W}_{\mathrm{zero}}$ to be a diagonal matrix whose elements are

$$\begin{aligned}\mathbf{W}_{\mathrm{zero}}(i,i) &= 1,\ \text{non-reproductive stage } i, \\ \mathbf{W}_{\mathrm{zero}}(i,i) &= \kappa_i(0),\ \text{reproductive stage } i. \end{aligned} \tag{A.50}$$

Use the definitions of the preceding section and compute

$$\mathbf{V}_{\mathrm{zero}} = \mathbf{W}_{\mathrm{zero}} \Big(\mathbf{I} - \mathbf{Q}\,\mathbf{W}_{\mathrm{zero}}\Big)^{-1} \mathbf{d}. \tag{A.51}$$

Then the ith component of $\mathbf{V}_{\mathrm{zero}}$ is

$$\Pr[\mathrm{LRS}=0|\ \text{initial stage } i].$$

Since every age+stage model can be cast as a stage-only model, this answer can also be used for any model.

For some life cycles we can also compute

$$\Pr[\mathrm{LRS}=0|\ \text{die without reaching a stage capable of reproduction}]$$

Suppose we start in a non-reproductive stage. In suitable lifecycles, every individual eventually makes an irreversible transition to one or more reproducing stages, or dies. Here it makes sense to treat the reproducing stages as "absorbing" and then (1 - the absorption probability) is just what we want. See Caswell (2001) or Kemeny and Snell (1976) for details on how to do this.

A.9 Special distributions

A.9.1 Poisson

Here the average fertility is f_i for stage i, and $\boldsymbol{\kappa}_i$ is a Poisson distribution. So $\boldsymbol{\kappa}_i$ has the probability generating function

$$\kappa_i(x) = \exp\big[f_i\,(x-1)\big], \tag{A.52}$$

for x a dummy variable.

So for stage j, and any frequency $\theta(k)$ as defined in equation (A.7), the Fourier transform is

$$\widehat{\kappa_j}(\theta(z)) = \exp\big[f_j\,(\theta(z)-1)\big]. \tag{A.53}$$

A.9.2 Binomial

Here the average fertility is f_i for stage i, and $\boldsymbol{\kappa}_i$ is a Binomial (with 1 trial, also known as a Bernoulli) distribution. So $\boldsymbol{\kappa}_i$ has the probability generating function

$$\kappa_i(x) = \kappa_i(0) + (1-\kappa_i(0))x, \tag{A.54}$$

where $\kappa_i(0)$ is the probability that an individual in stage i has 0 offspring and x is a dummy variable.

So for stage i, and any frequency $\theta(k)$ as defined in equation (A.7), the Fourier transform is

$$\widehat{\kappa_i}(\theta(z)) = \kappa_i(0) + (1-\kappa_i(0))\theta(z). \tag{A.55}$$

A.10 Roe deer plots

Small offspring have a high mode at 0, and thus the highest probability of leaving no offspring. Neither the average LRS nor the standard deviation seem to be useful descriptors of the LRS distribution for the smallest yearlings (plotted in the Appendix, Fig. A.1). For the full spectrum of size class, Fig. A.2 shows 41 initial stages spread over size class 1 to 200. Between size class 5 to 200, there are 5 size class interval between each curve. Fig. A.3 shows the Pr[LRS = 0] declines with increasing birth mass.

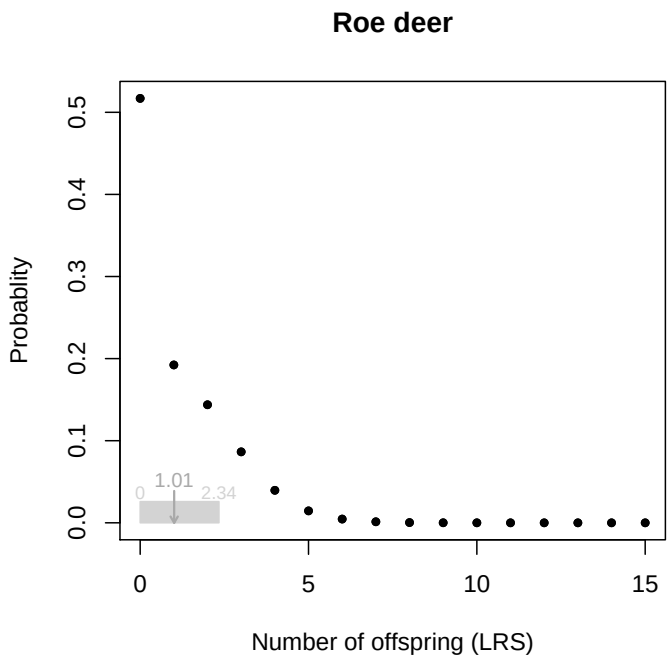

Figure A.1: The LRS distribution of Roe deer *Capreolus capreolus*. Mean LRS = 1.01 and the standard deviation is 1.34.

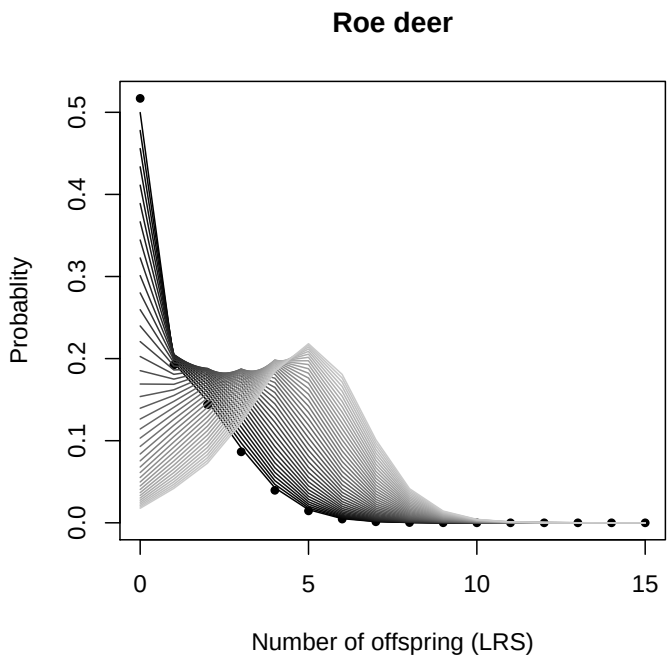

Figure A.2: The LRS distribution of Roe deer *Capreolus capreolus*. There are 41 initial stages. Solid points are for yearlings in size class 1. The gradient gray lines are the distribution spread over size class 5 to 200 with interval 5.

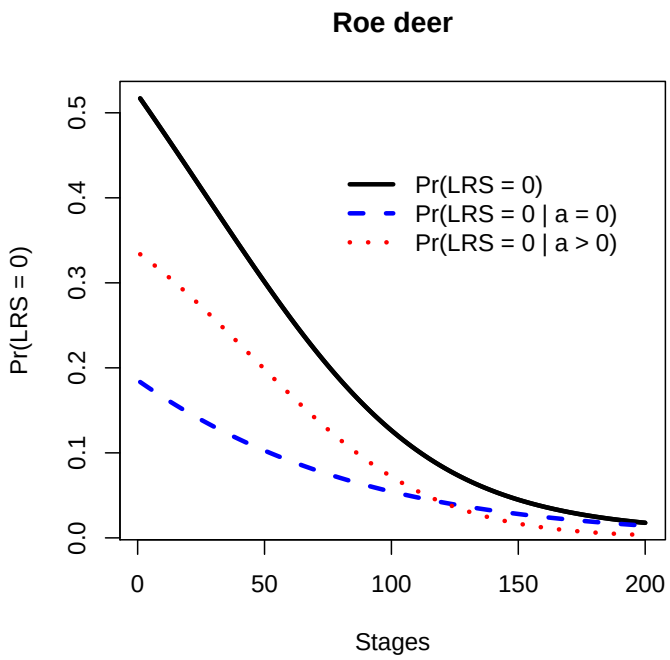

Figure A.3: The probability that LRS = 0 for different initial size classes for yearlings of Roe deer *Capreolus capreolus*. The blue dashed line shows the yearling death probability. The red dotted line shows the probability of having no offspring if you live to later ages.

A.11 The flow chart of decisions

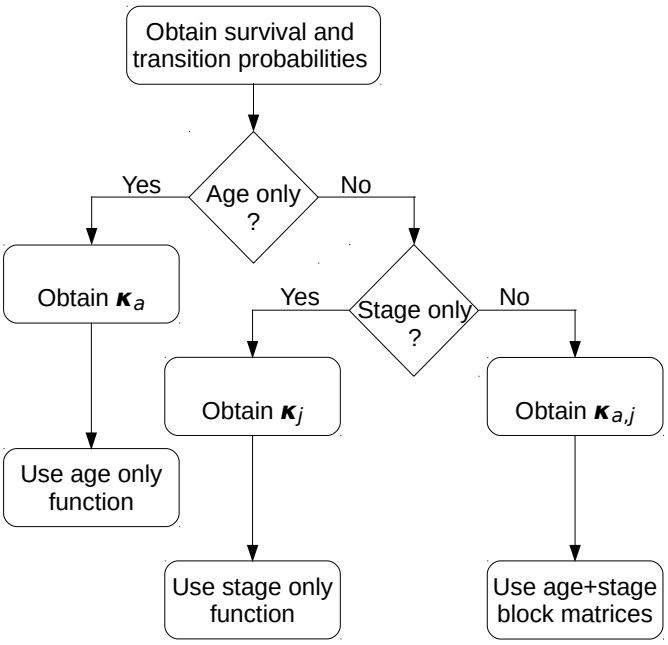

Figure A.4: The flow chart of decisions. Survival probabilities (Table 1) and/or survival-transition matrices (Table 1); Reproduction distributions $\boldsymbol{\kappa_a}$, $\boldsymbol{\kappa_j}$, $\boldsymbol{\kappa_{a,j}}$ (Table 1).

f_a, $0 \leq f_a \leq k_m$	Mean number offspring produced by individuals in age interval a. $f_a = \sum_k k\, \boldsymbol{\kappa}_a(k)$, $f_a = 0$ for $a < \alpha$ and $a > \beta$.
f_j, $0 \leq f_j \leq k_m$	Mean number offspring produced during one time interval by individuals in stage j. $f_a = \sum_k k\, \boldsymbol{\kappa}_j(k)$, $f_j = 0$ for non-reproductive stages.
$R_0 = \sum_a l_a\, f_a$, $a = 1, 2, \ldots, \omega$	Net Reproductive Rate, mean life time reproductive success.
k_m	Maximum number of offspring that can be produced by an individual during one time interval.
$k_m + 1$	Number of elements in a vector where each element references one possible number of offspring produced by an individual during one time interval.
$2k_m + 1$	Number of elements in a vector representing number of offspring produced cumulatively during two time intervals.
$\boldsymbol{\delta}$	Vector of probabilities that an individual produces a total of k offspring after reproducing at 2 different ages, $\delta(k)$, conditional on being alive at each of these ages. $\boldsymbol{\delta} = \boldsymbol{\kappa}_1 \star \boldsymbol{\kappa}_2$, Convolution of $\boldsymbol{\kappa}_1$ and $\boldsymbol{\kappa}_2$ (see main text Method:Age-only Model for computation)

Table A.1: Definitions

References

Caswell, Hal (2001). *Matrix Population Models: Construction, Analysis, and Interpretation.* 2nd. Sinauer Associates Inc., Sunderland, MA.

James, J. F. (2002). *A Student's Guide to Fourier Transforms : With Applications in Physics and Engineering.* Vol. 2nd ed. Cambridge University Press. ISBN: 9780521808262. URL: `https://stanford.idm.oclc.org/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=125073&site=ehost-live&scope=site`.

Kemeny, John G. and J. Laurie Snell (1976). *Finite Markov Chains.* Springer Verlag New York.

Steiner, U.K. and S. Tuljapurkar (2012). "Neutral theory for life histories and individual variability in fitness components". In: *Proceedings of the National Academy of Sciences* 109, pp. 4684–4689.