The equivocal mean age of parents in a cohort

François Bienvenu[1,2]

[1] *Center for Interdisciplinary Research in Biology (CIRB), CNRS UMR 7241, Collège de France, PSL Research University, Paris, France*
[2] *Laboratoire de Probabilités, Statistique et Modélisation (LPSM), CNRS UMR 8001, Sorbonne Université, Paris, France*

September 10, 2018

Abstract

The mean age at childbirth is an important notion in demography, ecology and evolution, where it is used as a measure of generation time. A standard way to quantify it is to compute the mean age of the parents of all offspring produced by a cohort, and the resulting measure is thought to represent the mean age at which a typical parent produces offspring. In this note, I explain why this interpretation is problematic. I also introduce a new measure of the mean age at reproduction and show that it can be very different from the mean age of parents of offspring of a cohort. In particular, the mean age of parents of offspring of a cohort systematically overestimates the mean age at reproduction, and can even be greater than the expected lifespan of parents.

1 Introduction

The mean age at reproduction is a central notion in the study of the evolution of reproductive timing and of the slow-fast continuum, and it also plays an important role in demography. However, as with many descriptors of populations, it is not clear how it should be defined – let alone quantified in practice. A standard way to measure it is to use the *mean age of parents of offspring produced by a cohort*: consider all offspring produced by a cohort of newborns over its lifetime; for each of these offspring, record the age that their parents (mother, in the case of a female-based model) had when the offspring was born, and take the average of these ages.

This quantity, which is is also frequently referred to as the *cohort generation time*, is straightforward to compute from complete census data. In practice however, it is usually estimated from life-tables using the following formula:

$$\mu_1 \;=\; \frac{\int_0^{+\infty} t\, m(t)\, \ell(t)\, dt}{\int_0^{+\infty} m(t)\, \ell(t)\, dt}\,. \tag{1}$$

E-mail address: francois.bienvenu@normalesup.org

In this expression, the *survivorship function* ℓ gives the probability that an individual of the chosen cohort reaches age t, and the *age-specific fertility* m represents its rate of offspring production in such a way that, assuming the individual remains alive between ages a and b, the expected number of offspring it will produce in that interval of time is $\int_a^b m(t)\,dt$. There is also a discrete-time version of formula (1):

$$\mu_1 \;=\; \frac{\sum_{t=1}^{+\infty} t\,\ell_t\,m_t}{\sum_{t=1}^{+\infty} \ell_t\,m_t}\,, \tag{2}$$

where ℓ_t is the probability that an individual survives to age t and m_t is the expected number of offspring produced at age t by individuals who reach that age.

Formulas (1) and (2) go back a long way, and are ubiquitous in the literature. They have been popularized by classic references such as Keyfitz (1968) and Coale (1972) in demography, and Charlesworth (1994) and Caswell (2001) in biology. They can also be found in more recent works of reference – including but not limited to Jørgensen and Fath (2008), Rockwood (2015) and Kliman (2016).

A consensual interpretation of μ_1 is that it represents the mean age at which a typical parent produces offspring. The aim of this note is to show that this interpretation is inaccurate and can be problematic in practice. To do so, I introduce a more direct measure of the mean age at reproduction of a typical parent: consider a typical parent, and compute the average of the ages at which it gives birth to its offspring. The expected value of this average is what we term the *mean age at reproduction*. Under standard assumptions, it is given by

$$\tau \;=\; \frac{1}{c}\int_0^{+\infty} \frac{\int_0^t s\,m(s)\,ds}{\int_0^t m(s)\,ds}\left(1-e^{-\int_0^t m(s)\,ds}\right) f(t)\,dt\,, \tag{3}$$

where f denotes the probability density function of the lifespan of an individual and the constant

$$c \;=\; \int_0^{+\infty}\left(1-e^{-\int_0^t m(s)\,ds}\right) f(t)\,dt \tag{4}$$

is the fraction of individuals that produce offspring during their lifetime. As with μ_1, there is a discrete-time formula for τ:

$$\tau \;=\; \frac{1}{c}\sum_{t\geqslant 1} \frac{\sum_{s=1}^t s\,m_s}{\sum_{s=1}^t m_s}\left(1-\prod_{s=1}^t e^{-m_s}\right) p_t\,, \tag{5}$$

where $p_t = \mathbb{P}(T_i = t) = \ell_t - \ell_{t+1}$ and

$$c \;=\; \sum_{t\geqslant 1}\left(1-\prod_{s=1}^t e^{-m_s}\right) p_t\,. \tag{6}$$

Using the expressions of μ_1 and of τ, we show by means of concrete examples that these two quantities can differ greatly, both under simple modelling hypotheses and in real-world applications. We also investigate the relation between them mathematically. In particular, we show that μ_1 is always greater than τ, and that the difference between the two can be arbitrarily large.

2 Interpretation of the expressions of μ_1 and τ

The detailed derivations of the expressions of μ_1 and τ are somewhat technical. They can be found in the Appendix, and a brief overview of some of the mathematical notions on which they rely is provided in the Online Supplements. Here we simply present the assumptions behind the formulas and explain what the quantities μ_1 and τ correspond to.

Let us start with our working hypotheses. Although seldom made explicit in sources presenting μ_1, the following mathematical assumptions are essential to the expressions of μ_1 given in the introduction.

In the continuous-time setting, we assume (1) that the lifetimes of individuals are independent copies of a random variable T such that $\mathbb{P}(T \geqslant t) = \ell(t)$; and (2) that births are punctual random events that occur while individuals are alive (but are independent of everything else), and that there exists a function m such that the expected number of offspring produced by any individual alive between ages a and b is $\int_a^b m(t)\,dt$ – in other words, that the birth events are the points of a point process with intensity m. Such models are known a Crump-Mode-Jagers processes (Crump and Mode, 1968, 1969; Jagers, 1969) – also sometimes referred to as generalized branching processes.

In the discrete-time setting, (2) is replaced by the assumption that at each age $t = 1, 2, \ldots$ at which it is alive, an individual produces a random number of offspring that is independent of everything else and has mean m_t.

Under these hypotheses, it can be shown (see Section B of the Appendix) that, if we let N be the random variable corresponding to the number of offspring produced by a typical individual over its lifetime and S be the sum of the ages at which it produces them, then the quantity μ_1 given in formulas (1) and (2) can rigorously be interpreted as

$$\mu_1 \;=\; \frac{\mathbb{E}(S)}{\mathbb{E}(N)}\,, \tag{7}$$

where $\mathbb{E}(\,\cdot\,)$ denotes the expected value. This however is neither the average age at which the individual produces offspring (as often claimed) nor the mean age of the parents of the offspring produced by a cohort. Nevertheless, using the assumption that individuals are independent, by a simple application of the law of large numbers we can show (see again Section B of the appendix) that the average of the ages of the parents of the offspring produced by a cohort tends to μ_1 as the size of the cohort goes to infinity.

Since the average age at which a typical individual produces offspring is the random variable S/N, a more natural measure for the mean age at reproduction than the ratio of expected values $\mathbb{E}(S)/\mathbb{E}(N)$ would be the expected value of S/N. A minor difficulty however is that this average is well-defined only when the individual produces some offspring, i.e. when $N > 0$. Thus, we define τ to be the conditional expectation

$$\tau \;=\; \mathbb{E}(S/N \mid N > 0)\,. \tag{8}$$

Equivalently, τ can be defined as follows: consider a typical parent (i.e. sample an individual uniformly at random among all individuals that produce some off-

spring), and denote by $\tilde{N}$ its number of offspring and by $\tilde{S}$ the sum of the ages at which it produces them. Then,

$$\tau \;=\; \mathbb{E}\Big(\tilde{S}/\tilde{N}\Big)\,. \tag{9}$$

Assuming that the point processes describing birth events are Poisson point processes (or, in the discrete-time setting, that the number of offspring produced each year by each individual follows a Poisson distribution), we prove in Section C of the Appendix that τ is given by formula (3) (resp. (5) in discrete time). Although these formulas rely on an extra assumption compared to the expressions of μ_1, observe that while the interpretation of μ_1 as an average on a large cohort depends crucially on the independence of individuals, this hypothesis is not required when working with τ, which is truly a characteristic of individuals.

When studying evolution, it is not uncommon for one to be interested in the average of a function z of the ages at which a parent produces offspring rather than in the average of the ages themselves [1]. One can then use that, letting A denote a uniformly chosen age at which a typical parent produces offspring, for every function z,

$$\mathbb{E}\Big(z(A)\Big) \;=\; \frac{1}{c}\int_0^{+\infty} \frac{\int_0^t z(s)\,m(s)\,ds}{\int_0^t m(s)\,ds}\left(1-e^{-\int_0^t m(s)\,ds}\right) f(t)\,dt\,, \tag{10}$$

with the constant c given in equation (4).

Finally, let us mention that expressions of μ_1 are also available for more general population structures. For instance, in matrix population models, if we let $\mathbf{S}$ be the survival matrix and $\mathbf{F}$ be the fertility matrix (i.e. if we decompose the projection matrix $\mathbf{A}$ into $\mathbf{A} = \mathbf{S} + \mathbf{F}$ to separate survival probabilities from fertilities) and denote by $\mathbf{w}$ the stable distribution of the population (the dominant right-eigenvector of $\mathbf{A}$) and $\mathbf{e} = (1, \ldots, 1)$ the row vector consisting only of ones, then we can use the following modern version of the classic formula of Cochran and Ellner (1992), which can be found in Ellner (2018):

$$\mu_1 \;=\; \frac{\mathbf{e}\mathbf{F}(\mathbf{I}-\mathbf{S})^{-2}\mathbf{F}\mathbf{w}}{\mathbf{e}\mathbf{F}(\mathbf{I}-\mathbf{S})^{-1}\mathbf{F}\mathbf{w}}\,. \tag{11}$$

Note that $(\mathbf{I}-\mathbf{S})^{-1} = \sum_{t\geqslant 1}\mathbf{S}^{t-1}$ and that $(\mathbf{I}-\mathbf{S})^{-2} = \sum_{t\geqslant 1} t\,\mathbf{S}^{t-1}$, so that this expression closely parallels (2). Also, the entries of $\mathbf{e}$ represent the weight given to each type of offspring when computing the average age of the parents. Taking $\mathbf{e} = (1, \ldots, 1)$ thus corresponds the natural choice to weight all types of offspring equally. But, should we wish to give more importance to some offspring type, any vector with positive entries could be used in place of $\mathbf{e}$ – and in fact Cochran and Ellner (1992) suggest using the reproductive values as weights. See Steiner et al. (2014) and Ellner (2018) for more on this.

Unfortunately, as explained at the end of Section C of the Appendix, we have not been able to obtain an analogue of (11) for τ. Nevertheless, its definition as the mean age at which a typical parent produces offspring still applies in the context of matrix population models, and it is straightforward to estimate it numerically via individual-based simulations.

[1] This was pointed out by Mauricio González-Forero.

3 Examples

Let us start with a simple but fundamental example, where individuals reproduce at constant rate m. In that case,

$$\mu_1 \;=\; \frac{\mathbb{E}\Big(\int_0^T ms\,ds\Big)}{\mathbb{E}\Big(\int_0^T m\,ds\Big)} \;=\; \frac{1}{2}\frac{\mathbb{E}(T^2)}{\mathbb{E}(T)} \tag{12}$$

and

$$\tau \;=\; \mathbb{E}\left(\frac{\int_0^{\tilde{T}} ms\,ds}{\int_0^{\tilde{T}} m\,ds}\right) \;=\; \frac{1}{2}\,\mathbb{E}\big(\tilde{T}\big)\,. \tag{13}$$

In particular, the expression of τ was foreseeable, because birth events are uniformly distributed on the lifetime of individuals, so on average they occur in the middle of them. Also, since

$$\mathbb{E}\big(\tilde{T}\big) \;=\; \frac{\mathbb{E}\Big(T\big(1-e^{-mT}\big)\Big)}{\mathbb{E}(1-e^{-mT})}\,, \tag{14}$$

and that for all $t > 0$, $1 - e^{-mt}$ increases to 1 as m goes to infinity, it follows by monotone convergence that

$$\mathbb{E}\big(\tilde{T}\big) \to \mathbb{E}(T) \quad \text{as } m \to +\infty\,. \tag{15}$$

By a similar argument that can be found in Section S3 of the Online Supplements, we also have

$$\mathbb{E}\big(\tilde{T}\big) \to \frac{\mathbb{E}(T^2)}{\mathbb{E}(T)} \quad \text{as } m \to 0\,. \tag{16}$$

Furthermore, since $\mathbb{E}(\tilde{T})$ can be shown to be a decreasing function of m (e.g., by a standard coupling argument), we conclude that when individual reproduce at a constant rate,

$$\tfrac{1}{2}\mathbb{E}(T) \;\leqslant\; \tau \;\leqslant\; \mu_1\,. \tag{17}$$

In fact, the second inequality holds for general age-specific fertility functions, as shown in Section S4 of the Online Supplements.

To make this example more concrete, let us assume that individuals die at constant rate η, so that T is an exponential variable and that $\ell(t) = e^{-\eta t}$. In that case, after simplifications we get

$$\mu_1 = \frac{1}{\eta} \quad \text{and} \quad \tau = \frac{1}{2\,\eta}\left(1 + \frac{1}{1+m/\eta}\right). \tag{18}$$

Note that here μ_1 is also equal to the expected lifespan in the population. Interpreting it as the mean age at which parents reproduce would therefore lead to a contradiction, because – in the case where the fertility m is large enough, so that most individuals get to reproduce during their lifetime and that the lifespan of a typical parent is not very different from that of a typical individual – this would imply that, on average, the age at which an individual reproduces is the same

as the age at which it dies. This is absurd, because unless individuals reproduce exactly when they die, the former has to be smaller than the latter.

We also see that for m/η large enough, $\mu_1 \approx 2\tau$. For $m = \eta$, which corresponds to the minimum ratio m/η to have a population that does not go extinct, the difference is already 33% of the value of τ.

Now consider the closely related discrete-time model where individual survive from one year to the other with probability p and produce Poisson(m) offspring at each age $t \geqslant 1$, so that

$$p_t = (1-p)\,p^t \quad \text{and} \quad \ell_t = p^t\,, \tag{19}$$

After straightforward calculations, we find that the numerator in formula (2), which corresponds to the mean sum of the ages at childbirth, is $mp/(1-p)^2$ and that the denominator is $R_0 = mp/(1-p)$. As a result,

$$\mu_1 = \frac{1}{1-p}\,. \tag{20}$$

Note that this model can also be seen as a 1×1 matrix population model with survival matrix $\mathbf{S} = (p)$ and fertility matrix $\mathbf{F} = (m)$, so formula (11) can also be used to get this result.

Because $\mathbb{E}(T) = p/(1-p)$, we see that

$$\mu_1 \;=\; \mathbb{E}(T) + 1\,, \tag{21}$$

which also corresponds to the expected lifespan of individuals that reach age 1. For the same reason as before, this implies that μ_1 is not credible as an estimate for the mean age at which a typical parent produces offspring.

After standard calculations that have been relegated to Section S5 of the Online Supplements, we find that

$$\tau = \frac{1}{2}\left(\frac{1}{1-p} + \frac{1}{1-p\,e^{-m}}\right). \tag{22}$$

As previously, $\frac{1}{2}\mu_1 \leqslant \tau \leqslant \mu_1$, but the difference between μ_1 and τ can be quite high, even for very reasonable values of p and m: for instance, with $p = 0.5$ and $m = 2$ the discrepancy is 30% of the value of τ; for $p = 0.9$ and $m = 2$, which yields a model that is similar to many real-world matrix populations models, it is already 80%.

What about real-world situations? Consider the matrix population models for the tropical palm *Astrocaryum mexicanum* (Pinero et al., 1984; Cochran and Ellner, 1992), whose projection matrices can be found in Section S6 of the Online Supplements. These models have frequently been used as examples of matrix population models, as illustrated by the fact that one of them is shipped with the ULM software for studying population dynamics (Legendre and Clobert, 1995).

Table 1 lists some relevant descriptors of these models. As already mentioned, we do not know how to compute τ for general matrix population models but it is straightforward to estimate it from individual-based simulations, which is what we have done here.

	High density	Medium density	Low density
T_{R_0}	197.5	169.2	153.6
$\bar{A}$	152.5	122.8	105.1
μ_1	275.2	261.8	275.5
$\hat{\tau}$	152.0	131.3	184.8
$\hat{L}$	232.2	207.6	296.0

Table 1: Comparison of several measures of reproductive timing for three real-world models for the demography of the tropical palm *Astrocaryum mexicanum*, in part taken from Table 4 of Cochran and Ellner (1992): T_{R_0} denotes the R_0 generation time, which corresponds to the time it takes for the population to grow by a factor of its net reproductive rate; $\bar{A}$ is the mean age of parents of offspring in a population that has reached its stable distribution; μ_1 is computed as in formula (11); $\hat{\tau}$ and $\hat{L}$ are estimates of the mean age at reproduction and of the expected lifespan conditional on producing offspring, respectively, that were obtained using individual-based simulations (see Section S6 of the Online Supplements). All values are expressed in years.

The difference between μ_1 and τ is significant in all three scenarios – a factor 2 in the *Medium density* model. Moreover, μ_1 is also consistently much greater than $\bar{A}$, the mean age of the parents of the offspring produced by the population when it has reached its stable distribution, and than T_{R_0}, the R_0 generation time. This is intriguing, because these three quantities are all supposed to give a measure of generation time. Overall, τ seems to be in much better agreement with T_{R_0} and $\bar{A}$ than μ_1. Finally, and most surprisingly, we see that in the *High density* case, μ_1 is greater than the expected lifespan conditional on reproduction. This counter-intuitive fact casts serious doubts on the relevance of μ_1 as a measure of reproductive timing in the life-cycle.

4 Discussion

The expressions of μ_1 and τ and the examples of the previous section make it clear that the mean age of the parents of the offspring produced by a cohort and the mean age at reproduction are two genuinely different notions. It is thus natural to wonder why they have not been recognized as such before. Part of the answer probably comes from the fact that precise definitions of these quantities are seldom given. For instance, in the references given above – which are or have been among the most influential in the field – μ_1 is variously described as the "mean age at childbearing in the stationary population[2]" by Keyfitz (1968); as the "mean age of childbearing in a cohort" by Coale (1972, eq. (2.10) p. 19); as the "mean age at reproduction of a cohort of females" by Charlesworth (1994, eq. (1.47a) p. 30); and as the "mean age of the mothers at the time of their daughter's birth" by Rockwood (2015, eq. (4.12) p. 98). Yet these four definitions fail to mention explicitly the crucial fact that each offspring produced by the cohort should be considered when computing the average. In particular, they are all compatible with the idea of

[2] What Keyfitz calls the *stationary population* is actually a cohort.

sampling an individual uniformly at random among the parents of a cohort and computing the mean age at which it produces offspring – and thus apply just as well to τ.

It is not obvious from the definitions of μ_1 and τ how these two quantities are related – or indeed why they should differ at all. One helpful way to think about it is the following: μ_1 can be seen as an *offspring-centric* measure of the mean age of parents, whereas τ is a *parent-centric* measure of it. Indeed, to compute μ_1 we ask each newborn produced by a cohort "how old is your parent?", while for τ we ask a parent "how old are you going to be when you have children?" These questions have distinct answers because they correspond to two different ways to sample a parent.

Among other things, this explains why μ_1 is greater than τ: indeed, parents that live longer tend to have more offspring, and thus have a higher probability of being sampled via their offspring than when the sampling is done uniformly at random. As a result, they contribute more to μ_1 than to τ. Since these parents with longer lifespans are also those that tend to have a higher mean age at reproduction, this biases μ_1 upward compared to τ.

This also explains why the difference $\mu_1 - \tau$ goes to zero as the fertility becomes vanishingly small: in that case, the proportion of parents that give birth to more than one offspring during their lifetime goes to zero, and as the result the two parent-sampling schemes become equivalent.

Now that we better understand the link between μ_1 and τ, two important questions arise: (1) does the difference matter in practice? (2) which descriptor should be used in which context?

There are strong arguments in favor of a positive answer to the first question. First, observe that, from a purely mathematical point of view, the difference between μ_1 and τ can be made arbitrarily large. Indeed, recall that, when individuals reproduce at a constant rate m, $\mu_1 = \mathbb{E}(T^2)/\mathbb{E}(T)$ and $\tau \to \frac{1}{2}\mathbb{E}(T)$ as $m \to +\infty$. Thus, by choosing an appropriate distribution for the lifespan T and taking m large enough, we can make μ_1 arbitrarily large and τ arbitrarily small. Although this requires using biologically unrealistic lifespans with large variances, this is informative because it shows conclusively that μ_1 and τ are not just two slightly different formalizations of the same concept. Second, the examples of Section 3 show that, both under very standard modelling hypotheses and in real-world applications, μ_1 can easily be roughly twice as large as τ. Of course, this is does not mean that the difference is going to be significant in typical real-world situations – and, in fact, since $\mu_1 = \tau$ when the lifespan of individuals is deterministic, it seems likely that μ_1 and τ will be similar in situations where there is little variation in the lifespan of individuals – but this at least indicates that special caution should be taken when interpreting μ_1.

Regarding the second question, is still unclear which of μ_1 or τ should be favored in which context. From a practical point of view, the expressions of τ are, admittedly, more complex than those of μ_1. This of course is not a problem for real-world applications, where they are going to be evaluated numerically; but for theoretical applications this does make exact calculations harder, if possible at all.

Another thing to take into account is the slightly different domain of validity of both measures. While the interpretation of μ_1 hinges on the assumption that there are no interactions between individuals, the expression of τ relies on that of Poissonian births. One might cynically argue that this is hardly a problem, because both hypotheses are often used jointly in theoretical models – and never met in real-world applications. Nevertheless, there is a real difference here that should be taken into account when deciding which measure to choose.

Lastly, τ has the advantage of having a more direct interpretation than μ_1 – and judging from the phrasing used by several authors, it seems that it is sometimes τ they have in mind, even when working with μ_1. Moreover, the interpretation of μ_1 might not be as intuitive as we usually assume; notably, the fact that it can be greater not only than the expected lifespan but also than the expected lifespan conditional on reproduction (as illustrated by the *Medium density* scenario for *Astrocaryum mexicanum* in Table 1) is likely to come as a surprise to many researchers.

Acknowledgments

Mauricio González-Forero and Stéphane Legendre encouraged me to write this note and provided helpful comments on a first version of it. Jean-Jil Duchamps did a careful cross-checking of the math and helped me clarify a few points. Florence Débarre helped me increase the accessibility of the manuscript. Finally, my deepest thanks go to Amaury Lambert and Stephen P. Ellner who, in addition to being attentive to details, pointed me to the correct interpretation of the classic formulas for the mean age at childbirth in a cohort. Stephen Ellner should also be credited for the simple proof of Proposition 1.

References

Caswell, H. 2001. Matrix Population Models: Construction, Analysis, and Interpretation. 2nd ed. Sinauer Associates.

Charlesworth, B. 1994. Evolution in Age-Structured Populations. Cambridge Studies in Mathematical Biology, 2nd ed. Cambridge University Press.

Coale, A. J. 1972. The Growth and Structure of Human Populations. Princeton University Press.

Cochran, M. E., and S. P. Ellner. 1992. Simple methods for calculating age-based life history parameters for stage-structured populations. Ecological Monographs 62:345–364.

Crump, K., and C. J. Mode. 1969. A general age-dependent branching process II. Journal of Mathematical Analysis and Applications 25:8–17.

Crump, K. S., and C. J. Mode. 1968. A general age-dependent branching process I. Journal of Mathematical Analysis and Applications 24:494–508.

Daley, D. J., and D. Vere-Jones. 2003. An Introduction to the Theory of Point Processes. Probability and its Applications, 2nd ed. Springer.

Ellner, S. P. 2018. Generation time in structured populations. The American Naturalist, 192:105–110.

Ellner, S. P., D. Z. Childs, and M. Rees. 2016. Data-driven modelling of structured populations. Springer International Publishing.

Jagers, P. 1969. A general stochastic model for population development. Scandinavian Actuarial Journal 1969:84–103.

Jørgensen, S. E., and B. Fath. 2008. Encyclopedia of Ecology. 1st ed. Elsevier Science.

Keyfitz, N. 1968. Introduction to the Mathematics of Population. Addison-Wesley.

Kingman, J. F. C. 1992. Poisson Processes. Oxford Studies in Probability. Clarendon Press.

Kliman, R. M. 2016. Encyclopedia of Evolutionary Biology. Academic Press.

Legendre, S., and J. Clobert. 1995. ULM, a software for conservation and evolutionary biologists. Journal of Applied Statistics 22:817–834. The ULM software can be downloaded at http://www.biologie.ens.fr/~legendre/ulm/ulm.html.

Pinero, D., M. Martinez-Ramos, and J. Sarukhan. 1984. A population model of astrocaryum mexicanum and a sensitivity analysis of its finite rate of increase. The Journal of Ecology 72:977–991.

Rockwood, L. L. 2015. Introduction to Population Ecology. 2nd ed. Wiley–Blackwell.

Steiner, U. K., S. Tuljapurkar, and T. Coulson. 2014. Generation time, net reproductive rate, and growth in stage-age-structured populations. The American Naturalist 183:771–783.

A An explicit model for the population

Here we recall and further detail the assumptions on which the expressions of μ_1 and τ and their interpretations rely, and introduce some notation.

The setting that we use is that of a Crump-Mode-Jagers process (Crump and Mode, 1968, 1969; Jagers, 1969), where the population consists of a discrete set of individuals such that:

(i) Each individual i has a random lifespan T_i with distribution ν and which is independent of everything else.

(ii) Individual i produces a new offspring at age t for every point of P_i at t such that $t \leqslant T_i$, where P_i is a point process with intensity m on $[0, +\infty[$ that is independent of everything else.

Note that the point processes P_i are not homogeneous (m is a function of the age of individuals) and that they do not have to be simple (an individual can give birth to several offspring simultaneously). For mathematical tractability however, it is often convenient to work with *Poisson point processes*. As explained in Section S1 of the Online Supplements, where a few useful results about Poisson point processes can also be found, these allow to formalize the familiar idea that events "occur at rate m". While the assumption that P_i are Poisson point processes is not needed in the study of μ_1, it will be required to derive explicit formulas for τ.

In this setting, the definition and interpretation of the survivorship function and of the age-specific fertility are straightforward. The survivorship is defined by [3]

$$\ell(t) \;=\; \mathbb{P}(T_i \geqslant t) \;=\; \nu\Big([t, +\infty[\Big)\,. \tag{A.1}$$

Working with the measure ν is convenient because it makes it possible to treat the case where T_i is a continuous random variable and the case where it is a discrete random variable simultaneously. However, in many applications T_i will have a density f. Thus, we will do most of our calculations with ν but express our final results in terms of f or ℓ, as in formulas (1) and (3). Note that this essentially consists in replacing $d\nu(t)$ by $f(t)\,dt$ in integrals, and that either of f and ℓ can be deduced from the other, since $\ell(t) = 1 - \int_0^t f(s)\,ds$ and $f(t) = -\ell'(t)$.

The age-specific fertility is the function m. If we denote by $M_i(a, b)$ the integer-valued random variable corresponding to the number of offspring produced by i between ages a and b, then assuming that $b \leqslant T_i$ we have, as expected,

$$\mathbb{E}(M_i(a, b)) = \int_a^b m(t)\,dt\,, \tag{A.2}$$

Obviously, the framework of Crump-Mode-Jagers processes is not meant to take into account all phenomena that shape the structure and dynamics of real-world populations. For instance, it assumes that individuals are independent and thus excludes any kind of density dependence. Similarly, the (optional) assumption that

[3] In probability theory and statistics, the *survival function* almost invariably refers to the complementary cumulative distribution function of T_i, $t \mapsto \mathbb{P}(T_i > t)$. Here, however, we will stick to the convention used in biology.

individuals reproduce at rate m is constraining, and in particular implies that they cannot produce several offspring simultaneously. Nevertheless, this framework is close to the minimal setting containing all the ingredients needed to define most descriptors of populations, whilst being simple enough to remain tractable and make it possible to derive explicit formulas for these descriptors. Moreover, the hypotheses above correspond quite well to the assumptions that are made, typically implicitly, to obtain the classic expressions of many of descriptors of populations.

Finally, to obtain discrete-time equivalents of formulas (1) and (3) we will need to consider the following version of the model, which allows simultaneous births: we keep assumption (i) under the extra hypothesis that the lifespan T_i is an integer-valued random variable, and we replace (ii) by the assumption that at each age $t = 1, \dots, T_i$, individual i gives birth to $M_t^{(i)}$ new individuals. Again, this corresponds quite well to the usual hypotheses on which many classic formulas rely.

B The mean age of the parents of the offspring produced by a cohort

We now give a rigorous interpretation of the quantity μ_1 given by formulas (1) and (2). As we will see, this interpretation is more subtle than what is usually assumed. This is because μ_1 does *not* correspond to the expected value of the average of the ages of the parents of the offspring produced by a cohort, but only to the limit of this average when the size of this cohort goes to infinity.

Let $\mathcal{C}$ denote a cohort, that is, a set of n individuals considered from the time of their birth to the time of their death. Let T_i be the lifespan of individual i, and P_i be the set of ages at which it produces offspring. Note that in our setting, conditional on T_i, P_i is a point process with intensity m on $[0, T_i]$.

The average of the ages of the parents of the offspring produced by the cohort over its lifetime is

$$Z_{\mathcal{C}} = \frac{\sum_{i\in\mathcal{C}} \sum_{t\in P_i} t}{\sum_{i\in\mathcal{C}} \sum_{t\in P_i} 1} = \frac{\sum_{i\in\mathcal{C}} S_i}{\sum_{i\in\mathcal{C}} N_i}, \tag{B.1}$$

where $N_i = \sum_{t\in P_i} 1$ is the number of offspring produced by individual i, and $S_i = \sum_{t\in P_i} t$ is the sum of the ages at which it produces them. Note that $Z_{\mathcal{C}}$ is well-defined only when $\sum_{i\in\mathcal{C}} N_i > 0$, but that this happens with probability arbitrarily close to one for a large enough cohort.

As we have already seen, the expected number of offspring produced by an individual i whose lifespan is $T_i = t$ is

$$\int_0^t m(s)\, ds\,. \tag{B.2}$$

This quantity can be thought of as "$\mathbb{E}(N_i \mid T_i = t)$", even though this interpretation is subject to some caution. At any rate, it follows that

$$\mathbb{E}(N_i) = \int_0^{+\infty} \Big(\int_0^t m(s)\, ds \Big) d\nu(t)\,. \tag{B.3}$$

Moreover, using Fubini's theorem,

$$\int_0^{+\infty}\Big(\int_0^t m(s)\,ds\Big)\,d\nu(t) \;=\; \int_0^{+\infty} m(s)\Big(\int_s^{+\infty} d\nu(t)\Big)ds\,. \tag{B.4}$$

Using that $\int_s^{+\infty} d\nu(t) = \ell(s)$, we get the well-known expression for R_0, the *mean number of offspring produced by an individual during its lifetime*:

$$R_0 \;=\; \mathbb{E}(N_i) \;=\; \int_0^{+\infty} m(t)\,\ell(t)\,dt \tag{B.5}$$

Using Campbell's formula (equation (S1.11) in the Online Supplements) and the exact same reasoning, we can express the *mean sum of the ages at which an individual produces offspring* as

$$\mathbb{E}(S_i) \;=\; \int_0^{+\infty} t\,m(t)\,\ell(t)\,dt \tag{B.6}$$

Let N (resp. S) denote a random variable that has the common distribution of the variables N_i (resp. S_i). Then, as pointed out in most sources presenting the measure μ_1, we have

$$\mu_1 \;=\; \frac{\mathbb{E}(S)}{\mathbb{E}(N)}\,. \tag{B.7}$$

This however does not establish a link between μ_1 and $Z_{\mathcal{C}}$, the average age of the parents of offspring produced by the cohort. To see how these two quantities are related, observe that since the variables N_i (resp. S_i) are independent, if we denote by $n = \mathrm{Card}(\mathcal{C})$ the size of the cohort then by the law of large numbers, as $n \to +\infty$,

$$\frac{1}{n}\sum_{i\in\mathcal{C}} N_i \to \mathbb{E}(N) \quad \text{and} \quad \frac{1}{n}\sum_{i\in\mathcal{C}} S_i \to \mathbb{E}(S)\,. \tag{B.8}$$

As a result,

$$Z_{\mathcal{C}} \;=\; \frac{\frac{1}{n}\sum_{i\in\mathcal{C}} S_i}{\frac{1}{n}\sum_{i\in\mathcal{C}} N_i} \;\xrightarrow[n\to+\infty]{}\; \mu_1\,, \tag{B.9}$$

where the convergence is almost sure, i.e. happens with probability 1.

Importantly, note that μ_1 is *not* the expected value of S_i/N_i or of $Z_{\mathcal{C}}$. In fact, the expected value of S_i/N_i (conditional on this variable being well-defined) is precisely what we have termed the mean age at reproduction. We explain how to compute it in the next section.

C The mean age at reproduction τ

Recall that we have defined the mean age at reproduction to be the expected value of the average of the ages at which a typical parent produces offspring. Formally, assuming that individual i has some offspring, the average age at which it produces them is

$$\bar{X}_i = \frac{1}{N_i}\sum_{t\in P_i} t\,, \tag{C.1}$$

where, as before, N_i is the total number of offspring produced by i and P_i is the set of ages at which it produces them. The mean age at reproduction is thus

$$\tau = \mathbb{E}\Big(\bar{X}_i \,\Big|\, N_i > 0\Big)\,, \tag{C.2}$$

which, given our assumptions, does not depend on i or on the composition of the population.

To compute τ, let I be a "typical parent", i.e. be uniformly sampled among the individuals that produce offspring during their lifetime. We then have

$$\mathbb{E}\Big(\bar{X}_i \,\Big|\, N_i > 0\Big) = \mathbb{E}\Big(\bar{X}_I\Big)\,. \tag{C.3}$$

Moreover, letting $\tilde{T}$ denote the lifespan of I, $\bar{X}_I$ is the average of a point process with intensity m on $[0, \tilde{T}]$. As explained in Section S1 of the Online Supplements, in the case of a Poisson point process, the expected value of this average is simply the expected value of a random point of $[0, \tilde{T}]$ with density $t \mapsto m(t)/\int_0^{\tilde{T}} m(s)\,ds$. The remarkable fact that it does not depend on the value of N_i is a consequence of the absence of internal structure of Poisson point processes. From this, we get

$$\mathbb{E}\Big(\bar{X}_I \,\Big|\, \tilde{T}\Big) \;=\; \frac{\int_0^{\tilde{T}} s\, m(s)\,ds}{\int_0^{\tilde{T}} m(s)\,ds}\,. \tag{C.4}$$

As a result,

$$\tau \;=\; \int_0^{+\infty} \frac{\int_0^t s\, m(s)\,ds}{\int_0^t m(s)\,ds}\, d\tilde{\nu}(t)\,, \tag{C.5}$$

where $\tilde{\nu}$ is the law of the lifespan $\tilde{T}$ of I. Note that it is different from ν, the lifespan of a fixed individual, because conditioning on the fact that an individual produces offspring biases its lifespan; for instance, if – as frequently the case in real applications – there exists an age α such that $m(t) = 0$ for $t < \alpha$, then individuals that produce offspring all live longer than α, whereas it is not necessarily the case for other individuals.

The last thing that we need to do in order to get an explicit formula for τ is thus to determine $\tilde{\nu}$. For this, note that

$$\begin{aligned} \mathbb{P}\Big(\tilde{T} \leqslant t\Big) &= \mathbb{P}(T_i \leqslant t \mid N_i > 0) \\ &= \frac{\mathbb{P}(T_i \leqslant t, N_i > 0)}{\mathbb{P}(N_i > 0)} \end{aligned} \tag{C.6}$$

Conditioning on T_i, using the void probabilities of Poisson point processes (see equation (S1.1) in the Online Supplements) for the probability that an individual with lifetime s produces some offspring, and finally integrating against ν, we get

$$\mathbb{P}(T_i \leqslant t, N_i > 0) = \int_0^t \Big(1 - e^{-\int_0^s m(r)\,dr}\Big)\, d\nu(s)\,. \tag{C.7}$$

As a result,

$$d\tilde{\nu}(t) \;=\; \frac{1}{c}\Big(1 - e^{-\int_0^t m(s)\,ds}\Big)\, d\nu(t)\,, \tag{C.8}$$

where the constant $c = \mathbb{P}(N_i > 0)$ is given by

$$c \;=\; \int_0^{+\infty} \left(1 - e^{-\int_0^t m(s)\,ds}\right) d\nu(t)\,. \tag{C.9}$$

Note that, by integrating by parts and using that $\ell(t) \to 0$ as $t \to +\infty$, we can also express c directly in terms of ℓ and m as

$$c \;=\; \int_0^{+\infty} e^{-\int_0^t m(s)\,ds}\, m(t)\,\ell(t)\,dt\,. \tag{C.10}$$

Putting the pieces together in the case where T_i has a density f, we get formula (1):

$$\tau \;=\; \frac{1}{c}\int_0^{+\infty} \frac{\int_0^t s\,m(s)\,ds}{\int_0^t m(s)\,ds} \left(1 - e^{-\int_0^t m(s)\,ds}\right) f(t)\,dt\,. \tag{C.11}$$

Note that neither the biological interpretation of τ nor the derivation of its expression depend on the assumption that individuals are independent.

Formula (10) for the average of a function z of the ages at which a parent produces offspring is obtained similarly, except that we have to work with

$$\bar{W}_i \;=\; \frac{1}{N_i}\sum_{t\in P_i} z(t) \tag{C.12}$$

instead of $\bar{X}_i$, and use equation (S1.10) instead of equation (S1.6) to get

$$\mathbb{E}\left(\bar{W}_I \,\middle|\, \tilde{T}\right) \;=\; \frac{\int_0^{\tilde{T}} z(s)\,m(s)\,ds}{\int_0^{\tilde{T}} m(s)\,ds}\,. \tag{C.13}$$

The justification of the expression of τ for discrete age structures can be found in Section S2 of the Online Supplements. It essentially consists in approaching the discrete-time model with the continuous-time one by choosing appropriate age-specific fertilities, and relies on the assumption that the number of offspring produced each year by each individual follows a Poisson distribution. It should also be pointed out that, because in the discrete-time setting individuals can produce several offspring simultaneously, there are two possibilities to define the average age at offspring production: counting all births equally, or weighting them by the number of offspring produced. Formula (2) is obtained by weighting the ages by the number of offspring produced when averaging them.

Finally, to obtain an equivalent of formula (5) for more general population structures, such as those allowed by matrix population models, one would need to (1) find the law of the conditional trajectory of an individual in the life cycle given that it produces offspring and (2) integrate the average of the ages at which it produces offspring against this law. While the first of these steps is feasible [4], it is unclear whether the resulting expression – if it can be obtained – would be simple enough to be useful.

[4] This was explained to me by Stephen Ellner– see e.g. Chapter 3 of Ellner et al. (2016).

Online Supplements

S1 Basic facts about Poisson point processes

In this section we recall basic results about Poisson point processes, focusing on the properties on which our calculations rely. Thus no attempt is made to state the results in full generality, and we do not preoccupy ourselves with technical conditions such as measurability. For a detailed presentation of Poisson point processes, see e.g. Kingman (1992) or Daley and Vere-Jones (2003).

It is common in modelling to assume that an event *occurs at rate* $r(t)$ *at time* t. Loosely speaking, this means that the probability that the event happens between t and $t + dt$ is independent of its previous occurrences, and is approximately $r(t)\,dt$. The rigorous way to formalize this is to say that the events are distributed according to an (inhomogeneous) Poisson point process with intensity r. Such a process can be seen as a random set of points characterized by the following properties: writing $N(I)$ for the number of points that fall in a fixed set $I \subset \mathbb{R}$,

(i) $N(I)$ is a Poisson random variable with mean $\int_I r(t)\,dt$.

(ii) $N(I)$ and $N(J)$ are independent whenever I and J are disjoint.

Note that the following useful fact is an immediate consequence of (i):

$$\mathbb{P}(N(I) = 0) = \exp\Big(-\int_I r(t)\,dt\Big)\,. \tag{S1.1}$$

Property (ii), often known as the *independent scattering* property, essentially says that Poisson point processes have a "completely random" structure.

From now on, we consider a fixed set $I \subset \mathbb{R}$ such that $\int_I r(t)\,dt < +\infty$. We let P be a Poisson point process with intensity r on I and denote by $N = \mathrm{Card}(P)$ its number of points. Let X be a random point of I with density $t \mapsto r(t)/\int_I r(t)\,dt$, i.e. whose distribution is characterized by

$$\forall A \subset I, \quad \mathbb{P}(X \in A) = \frac{\int_A r(t)\,dt}{\int_I r(t)\,dt}\,, \tag{S1.2}$$

and note in passing that

$$\mathbb{E}(X) = \frac{\int_I t\,r(t)\,dt}{\int_I r(t)\,dt}\,. \tag{S1.3}$$

Then, conditional on $N = n$, P consists of n independent copies of X – that is, for every function φ,

$$\mathbb{E}(\varphi(P) \mid N = n) = \mathbb{E}\Big(\varphi(\{X_i : i = 1, \dots, n\})\Big)\,, \tag{S1.4}$$

where X_i, $i = 1, \dots, n$, are independent copies of X.

A consequence of this is that the expected value of the average of the points in P is $\mathbb{E}(X)$. Formally, if $N > 0$ then we can define a random variable $\bar{X}$ by

$$\bar{X} = \frac{1}{N}\sum_{t \in P} t\,. \tag{S1.5}$$

We then have

$$\mathbb{E}\Big(\bar{X} \,\Big|\, N > 0\Big) = \mathbb{E}(X)\,. \tag{S1.6}$$

Indeed,

$$\mathbb{E}\Big(\bar{X} \,\Big|\, N > 0\Big) \;=\; \frac{1}{\mathbb{P}(N>0)} \sum_{n \geqslant 1} \mathbb{E}\Big(\bar{X} \,\Big|\, N = n\Big)\mathbb{P}(N = n)\,, \tag{S1.7}$$

and, for every $n \geqslant 1$,

$$\mathbb{E}\Big(\bar{X} \,\Big|\, N = n\Big) \;=\; \mathbb{E}\Big(\frac{X_1 + \cdots + X_n}{n}\Big) \;=\; \mathbb{E}(X)\,. \tag{S1.8}$$

In fact, given a function f, the exact same reasoning can be applied to

$$\bar{W} \;=\; \frac{1}{N} \sum_{t \in P} f(t) \tag{S1.9}$$

to show that

$$\mathbb{E}\Big(\bar{W} \,\Big|\, N > 0\Big) = \mathbb{E}(f(X))\,. \tag{S1.10}$$

We close this short overview with a fundamental result known as Campbell's formula. This formula states that, for every function f,

$$\mathbb{E}\Big(\sum_{t \in P} f(t)\Big) \;=\; \int_I f(t)\, r(t)\, dt\,. \tag{S1.11}$$

In contrast to (S1.6) and (S1.10), which are consequences of the independent scattering property, Campbell's formula is not specific to Poisson point processes.

S2 Expression of τ for discrete age structures

In discrete time, individual i has an integer-valued lifespan T_i and, at each age $t = 1, \ldots, T_i$, produces $M_t^{(i)}$ new individuals, where the variables $M_t^{(i)}$ are integer-valued and independent of everything else. Here we will also need to assume that each variable $M_t^{(i)}$ is a Poisson random variable with mean m_t.

In that setting, the average $\bar{X}_i$ of the ages at which individual i produces offspring can be defined as

$$\bar{X}_i = \frac{1}{N_i} \sum_{t=1}^{T_i} t\, M_t^{(i)}\,, \tag{S2.1}$$

this definition being valid only when $N_i = \sum_{t=1}^{T_i} M_t^{(i)} > 0$. Note that, in this expression, each age at which i produces offspring is weighted by the number of offspring produced. This is similar to what is done for μ_1, where each offspring contributes to the average age of the parents. But another possibility would be to weight all ages equally, that is, use the variable

$$\bar{Y}_i = \frac{\sum_{t=1}^{T_i} t\, I_t^{(i)}}{\sum_{t=1}^{T_i} I_t^{(i)}}\,, \tag{S2.2}$$

where $I_t^{(i)} = 1$ if $M_t^{(i)} > 0$ and 0 otherwise.

Since $\bar{X}_i = \bar{Y}_i$ when individuals cannot give birth to several offspring simultaneously (or, more generally, when the number of offspring produced is either 0 or some constant m), the two definitions were equivalent in the continuous-time setting. But now, $\bar{X}_i$ and $\bar{Y}_i$ are two different and legitimate candidates for the "average age at which i produces offspring". However, $\bar{Y}_i$ does not lend itself to analysis as easily as $\bar{X}_i$ and to obtain formula (5) – which is arguably the natural discrete-time equivalent of formula (3) – it is $\bar{X}_i$ that should be used. Therefore, we define τ to be $\mathbb{E}\Big(\bar{X}_i \,\Big|\, N_i > 0\Big)$.

The reasoning that lead to (3) could be adapted to obtain an expression for τ. However, it is also possible to deduce this expression directly from our results in continuous time. Indeed, the calculations of Section C of the Appendix are valid for general lifespans, including discrete ones: when ν is discrete, we simply have for any function φ

$$\int_0^{+\infty} \varphi(t)\, d\nu(t) \;=\; \sum_{t\geqslant 1} \varphi(t)\, p_t\,, \tag{S2.3}$$

where $p_t = \mathbb{P}(T_i = t)$.

Moreover, observe that if we let the age-specific fertility m be the piecewise constant function defined by

$$m(t) = \sum_{s\geqslant 1} m_s \mathbb{1}_{]s-1,s]}(t)\,, \tag{S2.4}$$

where $\mathbb{1}_{]s-1,s]}$ is the function that evaluates to 1 if $t \in\,]s-1, s]$ and 0 otherwise, then the number of offspring produced by an individual between ages $(t-1)$ and t is a Poisson variable with parameter m_t. Thus, the only difference with the discrete setting is that the ages at which these offspring are produced are uniformly distributed in $]t-1, t]$ instead of all equal to t.

Now, if we take the age-specific fertility to be the function $m^{(\varepsilon)}$ defined by

$$m^{(\varepsilon)}(t) = \sum_{s\geqslant 1} \frac{m_s}{\varepsilon} \mathbb{1}_{]s-\varepsilon,s]}(t)\,, \tag{S2.5}$$

then the number of offspring produced between ages $(t-1)$ and t is still a Poisson variable with parameter m_t, but this time the ages at which these offspring are produced are uniformly distributed in $]t-\varepsilon, t]$. Taking ε to zero, the mean age at childbirth will therefore tend to that of the discrete-time model. We spare the reader the straightforward but somewhat technical argument by which this can be made rigorous. Noting that, for continuous functions g,

$$\int_0^t g(s)\, m^{(\varepsilon)}(s)\, ds \;\xrightarrow[\varepsilon\to 0]{}\; \sum_{s=1}^{t} g(s)\, m_s\,, \tag{S2.6}$$

we obtain the following discrete-time equivalent of (3):

$$\mu_1 \;=\; \frac{1}{c} \sum_{t\geqslant 1} \frac{\sum_{s=1}^t s m_s}{\sum_{s=1}^t m_s} \left(1 - \prod_{s=1}^{t} e^{-m_s}\right) p_t\,, \tag{S2.7}$$

where $p_t = \mathbb{P}(T_i = t) = \ell_t - \ell_{t+1}$, and

$$c = \sum_{t\geqslant 1} \left(1 - \prod_{s=1}^{t} e^{-m_s}\right) p_t\,. \tag{S2.8}$$

S3 Proof of $\mathbb{E}(\tilde{T}) \to \mathbb{E}(T^2)/\mathbb{E}(T)$ as $m \to 0$

In this section we prove that, when the lifespan T of a fixed individual has a second moment – a condition that is always met in practice – and the age-specific fertility is constant and equal to m, then the expected lifespan of individuals that produce offspring during their lifetime converges to $\mathbb{E}(T^2)/\mathbb{E}(T)$ as $m \to 0$. As seen in the main text, it follows immediately that $\tau \to \mu_1$, both in the continuous setting where offspring production occurs at a constant rate m during the lifetime of individuals and in the discrete setting where individuals produce a Poisson(m) number of offspring at each integer-valued age $t \geqslant 1$.

Proposition 1. *Let T denote the lifespan of a fixed individual, and let $\tilde{T}$ have the distribution of T conditional on reproduction in the model where reproduction happens at constant rate (or in the model where individuals produce* Poisson(m) *offspring at each integer age $t \geqslant 1$ at which they are alive), i.e.*

$$\mathbb{E}\big(\tilde{T}\big) \;=\; \frac{\mathbb{E}\big(T(1-e^{-mT})\big)}{\mathbb{E}(1-e^{-mT})}\,. \tag{S3.1}$$

Then, if $\mathbb{E}(T^2) < +\infty$,

$$\mathbb{E}\big(\tilde{T}\big) \to \frac{\mathbb{E}(T^2)}{\mathbb{E}(T)} \quad \text{as } m \to 0\,. \tag{S3.2}$$

Proof. The following proof is due to Stephen P. Ellner and is a welcome simplification of my original proof.

Let $g(m,t) = (1-e^{-mt})/m$, so that

$$\mathbb{E}\big(\tilde{T}\big) \;=\; \frac{\mathbb{E}(Tg(m,T))}{\mathbb{E}(g(m,T))}\,. \tag{S3.3}$$

Since

$$\frac{\partial}{\partial m} g(m,t) \;=\; \big(1+mt-e^{mt}\big)\frac{e^{-mt}}{m^2} \tag{S3.4}$$

and that $1+x \leqslant e^x$ for all x, we see that $g(m,t)$ increases to t as m decreases to 0. By the monotone convergence theorem, it follows that $\mathbb{E}(g(m,T)) \uparrow \mathbb{E}(T)$ and $\mathbb{E}(Tg(m,T)) \uparrow \mathbb{E}(T^2)$ as $m \downarrow 0$. This terminates the proof. □

S4 Proof of $\tau \leqslant \mu_1$

In this section we prove that μ_1, as defined by formulas (1) and (2), is always greater than or equal to τ, as defined by formulas (3) and (5). This will be a simple consequence of the following lemma.

Lemma 1. *Let X be a positive random variable, and let g and h be positive functions such that $x \mapsto g(x)/x$ is nondecreasing and $x \mapsto h(x)/x$ is nonincreasing. Then,*

$$\mathbb{E}\left(\frac{g(X)h(X)}{X}\right)\mathbb{E}(X) \;\leqslant\; \mathbb{E}(g(X))\,\mathbb{E}(h(X)) \tag{S4.1}$$

Proof. Let Y be a random variable with the same distribution as X and that is independent of X. We have to show

$$\begin{aligned}
& \mathbb{E}(g(X))\,\mathbb{E}(h(Y)) - \mathbb{E}\left(\frac{g(Y)h(Y)}{Y}\right)\mathbb{E}(X) \;\geqslant\; 0 \\
\iff \quad & \mathbb{E}\left(Xh(Y)\left(\frac{g(X)}{X} - \frac{g(Y)}{Y}\right)\right) \;\geqslant\; 0 \\
\iff \quad & \mathbb{E}\left(Xh(Y)\left(\frac{g(X)}{X} - \frac{g(Y)}{Y}\right)\mathbb{1}_{\{X>Y\}}\right) \\
& \quad + \;\; \mathbb{E}\left(Xh(Y)\left(\frac{g(X)}{X} - \frac{g(Y)}{Y}\right)\mathbb{1}_{\{X<Y\}}\right) \;\geqslant\; 0
\end{aligned} \tag{S4.2}$$

Since $x \mapsto g(x)/x$ is nondecreasing,

$$\left(\frac{g(X)}{X} - \frac{g(Y)}{Y}\right)\mathbb{1}_{\{X<Y\}} \;\leqslant\; 0 \tag{S4.3}$$

and since $x \mapsto h(x)/x$ is nonincreasing,

$$0 \;\leqslant\; Xh(Y)\mathbb{1}_{\{X<Y\}} \;\leqslant\; Yh(X)\mathbb{1}_{\{X<Y\}}\,. \tag{S4.4}$$

As a result,

$$\begin{aligned}
& \mathbb{E}\left(Xh(Y)\left(\frac{g(X)}{X} - \frac{g(Y)}{Y}\right)\mathbb{1}_{\{X<Y\}}\right) \\
\geqslant\; & \mathbb{E}\left(Yh(X)\left(\frac{g(X)}{X} - \frac{g(Y)}{Y}\right)\mathbb{1}_{\{X<Y\}}\right) \\
=\; & -\mathbb{E}\left(Xh(Y)\left(\frac{g(X)}{X} - \frac{g(Y)}{Y}\right)\mathbb{1}_{\{X>Y\}}\right)
\end{aligned} \tag{S4.5}$$

Plugging this into (S4.2) finishes the proof. □

Proposition 2. *Let T denote the lifespan of a fixed individual. Define the random variables M and M^* by*

$$M = \int_0^T m(s)\,ds \quad \textit{and} \quad M^* = \int_0^T s\,m(s)\,ds \tag{S4.6}$$

in the case where reproduction occurs at a constant rate, and by

$$M = \sum_{s=1}^{T} m_s \quad \textit{and} \quad M^* = \sum_{s=1}^{T} s\,m_s \tag{S4.7}$$

in the case where it takes place at integer-valued ages $t \geqslant 1$, so that, in both cases

$$\mu_1 \;=\; \frac{\mathbb{E}(M^*)}{\mathbb{E}(M)} \quad \textit{and} \quad \tau \;=\; \frac{\mathbb{E}\left(\frac{M^*}{M}(1-e^{-M})\right)}{\mathbb{E}(1-e^{-M})}\,. \tag{S4.8}$$

Then, $\tau \leqslant \mu_1$.

Proof. First, observe that M^* is actually a deterministic function of M. Indeed, let ψ (resp. ψ^*) denote the function such that $M = \psi(T)$ (resp. $M^* = \psi^*(T)$). Since ψ is nondecreasing, if we define θ by

$$\theta(x) = \inf\{t \geqslant 0 : \psi(t) \geqslant x\}\,, \tag{S4.9}$$

then we have $M^* = \psi^*(\theta(M))$. To see this, note that $\theta(M) \leqslant T$ by construction and that $\theta(M) < T$ implies $\int_{\theta(M)}^T m(s)\,ds = 0$ (resp. $\sum_{s=\theta(M)}^T m_s = 0$), which in turn implies $\int_{\theta(M)}^T s\,m(s)\,ds = 0$ (resp. $\sum_{s=\theta(M)}^T s\,m_s = 0$). Thus, writing

$$g(x) = \psi^*(\theta(x)) \quad \text{and} \quad h(x) = 1 - e^{-x}\,, \tag{S4.10}$$

we have to prove

$$\mathbb{E}(g(M))\,\mathbb{E}(h(M)) \;\geqslant\; \mathbb{E}\left(\frac{g(M)h(M)}{M}\right)\mathbb{E}(M)\,. \tag{S4.11}$$

Clearly, M is a positive random variable, and the functions g and h are positive. Therefore, all we have to do to finish the proof is to show that $x \mapsto h(x)/x$ is nonincreasing and that $x \mapsto g(x)/x$ is nondecreasing, so that we can apply Lemma 1. First,

$$\frac{d}{dx}\left(\frac{h(x)}{x}\right) \;=\; \frac{e^{-x}(1+x)-1}{x^2} \;\leqslant\; 0\,, \tag{S4.12}$$

since $1 + x \leqslant e^x$. Second,

$$\frac{g(x)}{x} \;=\; \frac{\psi^*(\theta(x))}{\psi(\theta(x))} \;=\; F(\theta(x)) \tag{S4.13}$$

where $F\colon t \mapsto \psi^*(t)/\psi(t)$. The function θ is nondecreasing by construction. The fact that F is nondecreasing can be shown by straightforward calculations, e.g., in the continuous case,

$$\frac{d}{dt}F(t) \;=\; \frac{m(t)\Big(\int_0^t (t-s)m(s)\,ds\Big)}{\Big(\int_0^t m(s)\,ds\Big)^2} \;\geqslant\; 0\,. \tag{S4.14}$$

However, it is more satisfying to see that $F(t)$ can be interpreted as the expectation of a random variable X_t with density $f_t(s) = m(s)\mathbb{1}_{[0,t]}(s)/\psi(t)$ in the continuous case, and with probability mass function $p_s^{(t)} = m_s/\psi(t)$ for $s = 1, \ldots, t$ in the discrete case. It then is easy to see that X_t is stochastically dominated by $X_{t'}$ for $t < t'$, and so if follows immediately that $F(t) = \mathbb{E}(X_t) \leqslant \mathbb{E}(X_{t'}) = F(t')$. □

S5 Example of calculation of τ

In this section we detail the calculations of τ in the case where the lifespan T of a fixed individual is a geometric variable such that for all $t \geqslant 0$, $\mathbb{P}(T_i = t) = (1-p)p^t$, and the age-specific fertility is constant and equal to m for all ages $t \geqslant 1$.

First, since

$$\frac{\sum_{s=1}^{t} s\, m}{\sum_{s=1}^{t} m} = \frac{t+1}{2} \tag{S5.1}$$

we see that, writing $\tilde{T}$ for the lifespan conditional on reproduction,

$$\tau \;=\; \frac{1}{2}\left(\mathbb{E}\left(\tilde{T}\right)+1\right), \tag{S5.2}$$

where

$$\mathbb{E}\left(\tilde{T}\right) \;=\; \frac{\mathbb{E}\left(T(1-e^{-mT})\right)}{\mathbb{E}(1-e^{-mT})}. \tag{S5.3}$$

Now,

$$\begin{aligned}\mathbb{E}\left(1-e^{-mT}\right) &= 1-\sum_{t\geqslant 0} e^{-mt}\,(1-p)\,p^t \\ &= 1-\frac{1-p}{1-p\,e^{-m}} \\ &= \frac{p\,(1-e^{-m})}{(1-pe^{-m})}\,.\end{aligned} \tag{S5.4}$$

Similarly,

$$\begin{aligned}\mathbb{E}\left(T(1-e^{-mT})\right) &= \sum_{t\geqslant 0} t\,(1-p)\,p^t - \sum_{t\geqslant 0} t\,e^{-mt}\,(1-p)\,p^t \\ &= \frac{p}{1-p} - \frac{(1-p)\,p\,e^{-m}}{(1-p\,e^{-m})^2} \\ &= \frac{p\,(1-e^{-m})(1-p^2e^{-m})}{(1-p)(1-p\,e^{-m})^2}\,.\end{aligned} \tag{S5.5}$$

As a result,

$$\mathbb{E}\left(\tilde{T}\right) = \frac{1-p^2e^{-m}}{(1-p)(1-p\,e^{-m})} \tag{S5.6}$$

and finally,

$$\tau = \frac{1}{2}\left(\frac{1}{1-p}+\frac{1}{1-p\,e^{-m}}\right). \tag{S5.7}$$

S6 Projection matrices for *A. mexicanum*

In this section we give the projection matrices for *Astrocaryum mexicanum* that are behind Table 1. These matrices are taken from Appendix 6 of Cochran and Ellner (1992), who averaged them from several projection matrices of Pinero et al. (1984). Note that there is a small typo in the projection matrix for the *Low density* model given by Cochran and Ellner (1992): the entry $(9, 8)$ of the projection matrix given is 0.8775, when it should be 0.08775. Correcting this, we find the same descriptors as in their Table 4.

The model is mostly size-based. Stage 1 corresponds to seedlings; stages 2–4 to non-reproducing juveniles and stages 5–10 to full-grown adults. In the matrices below, entries in bold correspond to reproductive transitions.

$$\mathbf{A}^{\text{high}} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & \mathbf{1.4792} & \mathbf{8.1560} & \mathbf{9.9513} & \mathbf{14.259} & \mathbf{23.594} \\ 0.037349 & 0.83093 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.015881 & 0.89666 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.048969 & 0.95944 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.029778 & 0.90496 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.082074 & 0.91348 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.086520 & 0.90553 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.094467 & 0.87733 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.088200 & 0.88642 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.11358 & 0.9950 \end{pmatrix}$$

$$\mathbf{A}^{\text{med.}} = \begin{pmatrix} 0 & 0 & 0 & 0 & \mathbf{0.18385} & \mathbf{4.222} & \mathbf{8.41} & \mathbf{8.8405} & \mathbf{16.676} & \mathbf{19.904} \\ 0.03629 & 0.84127 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.014582 & 0.91636 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.058131 & 0.93735 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.051565 & 0.91462 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.065923 & 0.8468 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.1424 & 0.8725 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.1200 & 0.84332 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.14800 & 0.913030 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.086966 & 0.9950 \end{pmatrix}$$

$$\mathbf{A}^{\text{low}} = \begin{pmatrix} 0 & 0 & 0 & 0 & \mathbf{0.33} & \mathbf{0.918} & \mathbf{8.0875} & \mathbf{16.606} & \mathbf{13.068} & \mathbf{16.875} \\ 0.030332 & 0.850010 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.026738 & 0.93928 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.04966 & 0.94548 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.04804 & 0.9185 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.0815 & 0.9313 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.0687 & 0.86362 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.13637 & 0.91225 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.08775 & 0.87867 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.12133 & 0.9950 \end{pmatrix}$$

The estimates of τ and L (the expected lifespan conditional on reproduction) presented in Table 1 were obtained from these matrices using straightforward individual-based simulations. The code is available as an Online Enhancement to the manuscript.