

Demographic Window to Aging in the Wild: Constructing Life Tables from Marked Individuals of Unknown Age

Hans-Georg Müller¹, Jane-Ling Wang¹, James R. Carey^{2,3}, Edward P. Caswell-Chen⁴, Carl Chen⁴, Nikos Papadopoulos² and Fang Yao¹

¹Department of Statistics, University of California, Davis, One Shields Ave., Davis, CA 95616, USA

²Department of Entomology, University of California, Davis, One Shields Ave., Davis, CA 95616, USA

³Center for the Economics and Demography of Aging, University of California, Berkeley, CA 94720, USA

⁴Department of Nematology, University of California, Davis, One Shields Ave., Davis, CA 95616, USA

March 17, 2004

Correspondence

James R. Carey, Department of Entomology, University of California, Davis, One Shields Ave., Davis, CA 95616, USA. Tel: +1 530 752 6217; fax: +1 530 752 1537; e-mail: jrcarey@ucdavis.edu

Summary

We address the problem of establishing a survival schedule for wild populations. A demographic key identity is established that leads to a method whereby age-specific survival and mortality can be deduced from a marked cohort lifetable that is established for individuals that are randomly sampled at unknown age and marked, with subsequent recording of time-to-death. This identity permits to construct lifetables from data where the birthdate of subjects is unknown. The model is explored for lifetables and continuous lifetime data and is implemented with statistical methods using nonparametric curve estimation methods to obtain estimates for the unknown survival in the wild. The analytical model provided here serves as a starting point to develop more complex models for estimating the survival properties of wild populations based on more realistic assumptions; and furthermore it is a first step towards a broad new concept of 'expressed demographic information content of marked or captured individuals'.

Keywords: capture, demographic identity, information content, lifetable, nonparametric estimation, remaining lifetime, survival.

Introduction

The life table is one of the most important tools in demographic and gerontological research because it is used to characterize the mortality and survival properties of cohorts and to quantify the actuarial rate of aging. The historical application of classical life table methods in aging science has been largely restricted to the use of mortality data from either humans or experimental animals maintained in the laboratory with minimal application of life table methods to wild populations. This limited use for analysis of field populations is partly due to the need for marking individuals of known age which generally requires capturing and marking young individuals for monitoring throughout their lives until they die. Because of the importance of the life table in aging research and the growing interest in understanding aging in the wild (Austad 1993, Congdon *et al.* 1994, Finch 2001, Reznick *et al.* 2001, Tatar and Yin 2001) we describe a life table identity that, by making certain key assumptions, enables us to estimate the age-specific life table rates from data based on the mark, release, and monitoring of randomly-captured individuals of unknown age from the time of their entry into the study (*i.e.* marking) to their death. The purpose of this paper is to establish the concept, formalize the model, apply it to an idealized case using medfly laboratory data, and discuss the importance of the technique as a baseline from which to build more realistic models for studying aging in the wild.

Consider a population that is assumed to be stable, stationary and closed. Individuals are captured with equal probability at an unknown age and marked, then their time-to-death is recorded. The question we address is this: Can the information on time-to-death for this randomly-captured marked subgroup provide the necessary information to construct a life table for the population at large? We will demonstrate that the answer to this question is affirmative because of a life table identity that reveals a mathematical relationship between the distribution of deaths in the marked cohort and the age structure of the original population. Individuals in the captured and marked sample are assumed to have remaining lifetimes as in the wild. We note that this model may be particularly adequate for some human populations.

Survival and longevity of animals living in the wild is of interest for research in various fields such as ecology, conservation biology, evolution and biodemography (Morris 1959, 1965; Gaillard *et al.* 1994; Southwood 2000). Methods that have been used such as capture-recapture, tele-tracking and others are fraught with difficulties (Lincoln 1930; Caughley 1977; Krebs 1999). In contrast, it is easier to monitor age-at-death and lifespan for a marked sample cohort, consisting of subjects that are randomly sampled from a wild population, then marked, their remaining lifetime between marking and death being recorded subsequently.

The problem of constructing a survival schedule from incomplete data has been studied in anthropology (Müller *et al.* 2002) and has applications to human populations such as the !Kung and the Ache for which only incomplete demographic data are available (Hawkes *et al.* 1998; Howell 1979; Hill & Hurtado 1996; Jones *et al.* 2002). An anthropologist may encounter a group of people whose ages are unknown but whose remaining lifetime can be recorded. The key identity, on which the reconstruction of the survival schedule that we propose is based, asserts that for such situations a lifetable may be ob-

tained, under certain assumptions. Application of a key identity provided in this paper then establishes a new way to construct lifetables and estimate survival functions.

We derive this key identity for both discrete lifetables and situations which are modeled by continuous survival times. This identity is a consequence of a close relationship between the density of the remaining lifetimes in a cohort of randomly sampled subjects and the survival schedule of the population from which the subjects were sampled. We provide statistical implementations of this identity by applying suitably adapted nonparametric density estimation methods. The proposed model is developed for a stable population but possesses sufficient flexibility to allow for modifications. This model will be useful for a variety of ecological studies (Promislow, 1991; Sinclair, 1977; Lehane 1985). The proposed methodology is also quite general in as much as it can be applied to virtually all populations from humans and large ungulates to insects and other invertebrates.

A Key Demographic Identity

The data on remaining lifetime after capture and marking that are obtained from the marked sample are assembled in a “marked sample lifetable”. Assuming that the process of capture and marking does not alter an individual’s remaining lifespan, the corresponding “marked sample” and “wild” lifetables are compared for a hypothetical situation in Table 1.

That it is possible to obtain the survival schedule in the wild, as summarized by the wild lifetable, from the marked sample lifetable is due to a basic relationship between these two lifetables. Assuming that the population is stationary and stable, i.e., is neither increasing nor decreasing, and without immigration or emigration, the number of subjects of age x is $c_x = l_x / \sum l_y = c_0 l_x$ (Caswell, 2001, see Table 1 for definitions). The death rates in the marked sample lifetable at age x' are by definition $d_{x'}^* = l_{x'}^* - l_{x'+1}^*$. These death rates are generated by subjects that enter the marked sample lifetable at various (unknown) ages, survive to “marked age” (i.e., age counted in days after capture and marking) x' and do not survive to “marked age” $x' + 1$. For all subjects that enter the marked sample cohort at age z , the contribution to $d_{x'}^*$ is therefore

$$c_z \frac{l_{z+1}}{l_z} \frac{l_{z+2}}{l_{z+1}} \dots \frac{l_{z+x'}}{l_{z+x'-1}} \left(1 - \frac{l_{z+x'+1}}{l_{z+x'}}\right) = c_z \left(\frac{l_{z+x'}}{l_z} - \frac{l_{z+x'+1}}{l_z}\right) = c_0 (l_{z+x'} - l_{z+x'+1}),$$

where l_z refers to the survival function or survival schedule of the wild population at age z .

The contributions of subjects entering the marked sample lifetable at various ages are additive, therefore, adding the contributions over all ages of entry z ,

$$d_{x'}^* = \sum_z c_0 (l_{z+x'} - l_{z+x'+1}) = c(0) l_{x'} = c_{x'},$$

and this relationship implies that the columns c_x indicating the age distribution in the wild lifetable and d_x^* indicating the distribution of deaths in the marked sample lifetable are identical. We can see from Table 1 that this is indeed the case for the hypothetical case considered there. As $l_x = \frac{c_x}{c_0}$, this relationship between the two lifetables leads to

$$l_x = \frac{d_x^*}{d_0^*},$$

thus enabling the reconstruction of the survival schedule l_x in the wild lifetable from the survival schedule l_x^* of the marked sample lifetable. Statistical estimates implementing this probabilistic relationship can be easily found, for example by plugging in empirical observed frequencies for l_x^* and d_x^* , thus replacing expected population values as they appear in Table 1 with their corresponding sample estimates.

Continuous Lifetimes

These considerations can be extended to the case where age-at-death is considered to be continuous and survival distributions are smooth rather than discrete. Denoting by X the age-at-death (lifetime) for an individual in the wild, by $\bar{F}(x) = P(X > x)$ the survival function in the wild, where x is a continuous age variable and P denotes probability, we find for the density of the age-distribution in the wild $c(x) = \bar{F}'(x) / \int_0^\infty \bar{F}(x) dx$, and consequently $\bar{F}(x) = \frac{c(x)}{c(0)}$.

The unknown age A at the time of capture and marking and the unknown age-at-death X are related with the known remaining lifetime X^* of an individual by $X^* = X - A$. Denote the densities of the distributions of X, X^* by f_X, f_{X^*} , and consider the conditional density $f_X(\cdot | X \geq x)$ of lifetime conditional on the event that the individual survives to age x . Then one obtains for the density $f_{X^*}(a)$ of X^* , evaluated at the age-at-death a ,

$$\begin{aligned} f_{X^*}(a) &= \int_0^\infty c(x) f_X(x+a | X \geq x) dx = \int_0^\infty \frac{\bar{F}(x)}{\int \bar{F}(a) da} \frac{f_X(x+a)}{\bar{F}(x)} dx \\ &= \frac{\bar{F}(a)}{\int \bar{F}(a) da} = c(a). \end{aligned}$$

This implies the key identity for the continuous case,

$$\bar{F}(x) = \frac{f_{X^*}(x)}{f_{X^*}(0)},$$

providing the relationship between the marked cohort mortality and survival in the wild. This type of relationship has been noted previously in the literature on renewal processes (Doob, 1948; Feller, 1968; Winter, 1989).

Estimating the Survival Schedule of the Wild Population

Given a sample of continuous lifetimes X_1^*, \dots, X_n^* that are observed in the marked sample cohort and measured in terms of relative age counted from the time of marking, we may substitute nonparametric kernel density estimators (compare for example Müller 1997) for $f_{X^*}(z)$, given by

$$\hat{f}_{X^*}(z) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{z - X_i^*}{h}\right).$$

Here $h = h(n)$ is a sequence of bandwidths and K is a kernel function. We then obtain asymptotically consistent estimates of the survival function of the wild population,

$$\hat{\bar{F}}(x) = \frac{\hat{f}_{X^*}(x)}{\hat{f}_{X^*}(0)}.$$

The implementation is less straightforward than it may seem. One difficulty is that the estimates $\hat{f}_{X^*}(0)$ that appear in the denominator are density estimates at a boundary point of the support of the data and therefore are subject to higher variability than density estimates in the interior of the support (compare Müller & Stadtmüller, 2001). A second difficulty is that the above estimate is not necessarily a survival function, which would be monotone declining from 1 to 0. This can be ensured by adding a monotonization step through the pool adjacent violators algorithm (PAVA, see Barlow et al. 1977; Robinson & Dykstra, 1988). We note that using analogous kernel density estimators $\hat{f}'_{X^*}(x)$ for the derivative of f_{X^*} , we may obtain estimates for the density f_X of the survival schedule of the wild cohort, $\hat{f}(x) = -\hat{f}'_{X^*}(x)/\hat{f}_{X^*}(0)$. Analogously, estimates for the hazard rate $h_X(x) = f_X(x)/\bar{F}(x)$ are obtained as $\hat{h}(x) = -\hat{f}'_{X^*}(x)/\hat{f}_{X^*}(x)$.

To assess the age at capture for a subject for which an additional lifetime x was observed in the marked cohort lifetable, we use the conditional density $f_{A|X^*}(a|x) = f_X(x+a)/\bar{F}(x)$ to infer the conditional expectation

$$E(A|X^* = x) = \frac{1}{\bar{F}(x)} \int_x^\infty (z-x)f_X(z) dz.$$

Plugging the above estimates into the right hand side of this equation then leads to consistent estimates of conditional mean age at capture. We note that monotonized density estimates similar to those above were proposed by Watelet & Winter (1991) in a reliability context.

Illustration

We illustrate the reconstruction of the survival schedule of the wild population from the observations made on the marked sample in a simulation study. The underlying survival schedule of the wild population is modeled as the survival function of a real cohort. The starting point is a cohort consisting of 1000 female Mediterranean fruit flies, *Ceratitis capitata*, commonly known as the medfly, whose survival has been described and analyzed in Carey et al. (1998).

Using acceptance-rejection sampling based on the graph of the survival function for these 1000 flies, we randomly sample N flies (with replacement) to create one simulated marked sample. Each of the flies that is selected for the marked sample has a random age, following the age distribution of the flies in the entire “wild population”, and also an associated remaining lifetime that is recorded as “marked lifespan”. Kernel density estimation as described above is implemented by local linear smoothing after an initial prebinning step (see Müller, 1997) and combined with the PAVA method.

The resulting survival function estimates along with the target survival function for six generated marked sample cohorts of sizes $N = 1000$ and $N = 50$ can be seen in Figure 1. We find that the method

of reconstructing the survival schedule of the wild population works very well for the larger sample and reasonably well for the smaller sample. The infant survival estimates show a higher degree of variability than the survival estimates for the mid-age period since not very many early deaths will be recorded in the marked cohort.

Discussion: Window on Aging in the Wild and a Generalization

In this paper we demonstrated that age-specific life tables can be constructed from mortality data derived from randomly-captured individuals of unknown age in stable, stationary and closed populations. The importance of our model is that it provides a starting point to develop more complex models whose purpose is to estimate the life table properties of populations based on more realistic assumptions (non-stable, non-stationary populations). However, we believe that the significance of the general approach extends beyond the life table and applies to the concept of expressed information content of marked (or captured) individuals. For the current case the expressed information is the remaining post-capture life span of marked individuals that is used to estimate the life table of the population at large. But this idea of expressed information content is general if it is assumed that: (1) the experiences of individuals early in life influence the expression and pattern of their life history traits (mortality, reproduction, behavior) later in life; and (2) these patterns expressed in later life can be traced to early-life experience. The concept of extracting knowledge of both an individual's age and its early-life experience to gain insights into the demographic and gerontological characteristics of the field population then can be used as the conceptual foundation for a new sampling concept for understanding aging in the wild. Examples of the types of information that can be extracted from wild-caught (or marked) individuals at the individual level include remaining life span, age-specific reproduction (relative to time of capture), details of reproduction including birth interval, clutch size, post-reproductive period, overall patterns of individual reproduction, total reproduction, and time from capture to first egg, timing and magnitude of peak reproduction (Carey *et al.* 1998, Müller *et al.* 2001), mating status and number of mating, behavioral measures such as calling (males, see Papadopoulos *et al.* 2002), mating, oviposition, and overall activity, and physiological measures such as metabolic rate.

We believe that this new concept for extracting information about aging in the wild is important for several reasons. First, life course analysis will both encourage and require a deep understanding of the interdependencies of various components of an individual's life course including reproduction, behavior, and death. In particular the approach will require an understanding of the relationship between reproduction at young ages and mortality risk at older ages, the age patterns of reproduction that are unique to different stages in the adult life course, and the linkages between different behavioral patterns and death. Second, the approach will encourage a greater integration of laboratory and field studies. Specifically, the method will require the creation of reference 'libraries' consisting of the life history patterns of individuals maintained under different conditions in the laboratory. These 'libraries' will be used for comparing the observed life history patterns (birth and death) of wild-caught flies maintained in the laboratory. Third, the results of studies using the methods we propose to develop will shed new

light on both aging and aging structure of wild populations. This includes aging data on populations of invertebrate species such as *C. elegans* which are difficult to study under natural conditions in the wild but which are extraordinarily important model organisms in aging science (Gershon and Gershon, 2002; Reznick, 1993). The combination of laboratory and field studies will provide the means for testing various theories about aging in the wild and also for testing models used in both forecasting and backcasting.

Acknowledgments

This research was supported by NIH grant P01-AG08761 and NSF grant DMS-02-04869. We thank J. Cardenas for technical assistance, L. Harshman, L. Partridge, and A. Yashin for discussion and J. Vaupel and K. Wachter for comments on a previous draft.

Appendix: Asymptotic Confidence Intervals and Variances

Based on the estimation of the survival schedule of the wild population, one can derive asymptotic confidence intervals for important characteristics of the survival schedule, such as the survival function $\bar{F}(x)$ for discrete and continuous lifetimes, the density $f_{X^*}(x)$, and the hazard rate $h_{X^*}(x)$ for continuous lifetimes. These confidence intervals can be used for inference. Another option is to employ the bootstrap.

We first investigate confidence intervals for the survival function $\bar{F}(x)$ for discrete lifetimes, i.e., the survival schedule at age x , i.e., $l_x = d_x^*/d_0^*$, where x is an arbitrary nonnegative integer. Let \hat{d}_0 and \hat{d}_x denote the estimates obtained by plugging in empirical observed frequencies for d_0^* and d_x^* . Let $W_n(x)$ denote the number of deaths in $(x, x + 1]$. It is easily seen that $W_n(x) \sim \mathbf{B}(n, d_x^*)$, for $x = 0, 1, \dots$, and $\hat{d}_x/\hat{d}_0 = W_n(x)/W_n(0)$, where $\mathbf{B}(n, d_x^*)$ denotes the binomial distribution with n trials and probability of success d_x^* , and n is the total number of subjects. Then from the Central Limit Theorem, one can obtain the asymptotic joint distribution of multinomial random variable $(W_n(x), W_n(0))^T$ which is $\mathbf{N}_2((d_x^*, d_0^*)^T, \Sigma)$, where \mathbf{N}_2 denotes the bivariate normal distribution, and Σ is 2×2 matrix with $(\Sigma)_{11} = d_x^*(1 - d_x^*)$, $(\Sigma)_{22} = d_0^*(1 - d_0^*)$, and $(\Sigma)_{12} = (\Sigma)_{21} = -d_x^*d_0^*$. Applying Cramér-Wold device leads to the asymptotic distribution of $\hat{F}(x)$ as follows,

$$(\hat{l}_x - l_x)/\sqrt{n} \xrightarrow{\mathcal{D}} \mathbf{N}(0, \frac{d_x^*(1 - d_x^*)}{d_0^{*2}} + \frac{d_x^{*2}(1 - d_0^*)}{d_0^{*3}} - \frac{2d_x^{*2}}{d_0^{*2}}),$$

Then the $100(1 - \alpha)\%$ confidence interval of l_x is obtained by substituting the empirical estimates of l_x as follows, based on Slutsky Theorem,

$$\hat{l}_x \pm \Phi(1 - \alpha/2) \sqrt{\hat{d}_x^*(1 - \hat{d}_x^*)/\hat{d}_0^{*2} + \hat{d}_x^{*2}(1 - \hat{d}_0^*)/\hat{d}_0^{*3} - 2\hat{d}_x^{*2}/\hat{d}_0^{*2}}$$

where $\Phi(\cdot)$ is the the cumulative distribution function of the standard normal random variable.

For continuous lifetimes, the survival function is estimated by $\hat{F}(x) = \hat{f}_{X^*}(x)/\hat{f}_{X^*}(0)$. Assume that kernel function K in $\hat{f}_{X^*}(x)$ and the boundary kernel function K_0 in $\hat{f}_{X^*}(0)$ (K_0 is used for eliminating the boundary effects) are both compactly supported on $[-1, 1]$. The bandwidth h used in kernel density estimates $\hat{f}_{X^*}(x)$ and $\hat{f}_{X^*}(0)$ is assumed to satisfy $h \rightarrow 0$ and $nh \rightarrow \infty$, as $n \rightarrow \infty$. For any fixed x ,

when n is sufficiently large, one has $h < x - h$, i.e., no X_i^* 's are included in both $[0, h]$ and $[x - h, x + h]$, whence the estimates $\hat{f}_{X^*}(x)$ and $\hat{f}_{X^*}(0)$ are asymptotically independent. From standard results for kernel density estimation (see Müller 1997 for references), one can easily obtain the asymptotic joint distribution of $[\hat{f}_{X^*}(x), \hat{f}_{X^*}(0)]$ as follows,

$$\sqrt{nh}\{\hat{f}_{X^*}(x) - E[\hat{f}_{X^*}(x)], \hat{f}_{X^*}(0) - E[\hat{f}_{X^*}(0)]\} \xrightarrow{\mathcal{D}} \mathbf{N}_2(\mathbf{0}_2, \begin{pmatrix} f_{X^*}(x)\|K\|^2, & 0 \\ 0, & f_{X^*}(0)\|K_0\|^2 \end{pmatrix}),$$

where $\mathbf{0}_2 = (0, 0)^T$, $\|K\|^2 = \int K^2(u)du$ and $\|K_0\|^2 = \int K_0^2(u)du$. Since the bias $E[\hat{f}_{X^*}(x)] - f_{X^*}(x) = O(h^2)$ for $x = 0$ and $x > 0$, we can ignore biases for small values of h . Assuming this is the case and applying the delta method, we obtain the asymptotic normal approximation to the distribution of $\hat{F}(x) = \hat{f}_{X^*}(x)/\hat{f}_{X^*}(0)$,

$$\hat{F}(x) - \bar{F}(x) \approx \mathbf{N}(0, \frac{1}{nh} \left[\frac{f_{X^*}(x)\|K\|^2}{f_{X^*}^2(0)} + \frac{f_{X^*}^2(x)(\|K_0\|^2)}{f_{X^*}^3(0)} \right]).$$

Then the $100(1 - \alpha)\%$ confidence interval for $\bar{F}(x)$ is obtained by substituting the consistent kernel estimates $\hat{f}_{X^*}(x)$ and $\hat{f}_{X^*}(0)$ for the true values, applying Slutsky's theorem, i.e., the $100(1 - \alpha)\%$ confidence interval for $\bar{F}(x)$ is

$$\hat{F}(x) \pm \Phi(1 - \alpha/2) \sqrt{[\hat{f}_{X^*}(x)\|K\|^2/\hat{f}_{X^*}^2(0) + \hat{f}_{X^*}^2(x)\|K_0\|^2/\hat{f}_{X^*}^3(0)]/(nh)}.$$

To construct the confidence interval for the density estimate $\hat{f}(x) = \hat{f}'_{X^*}(x)/\hat{f}_{X^*}(0)$, we note that $\hat{f}'_{X^*}(x)$ has slower convergence rate than $\hat{f}_{X^*}(0)$. Slutsky's theorem implies that $\hat{f}'_{X^*}(x)/\hat{f}_{X^*}(0)$ is asymptotically equivalent to $\hat{f}'_{X^*}(x)/f_{X^*}(0)$. From the asymptotic distribution of the kernel estimator for the derivative $f'_{X^*}(x)$, and ignoring the bias terms as argued earlier, one has $\hat{f}'_{X^*}(x) - f'_{X^*}(x) \approx \mathbf{N}(0, f_{X^*}(x)\|K_1\|^2/(nh^3))$, where K_1 is the kernel function used in $\hat{f}'_{X^*}(x)$. (Usually we will choose $K_1 = K^{(1)}$). Thus the asymptotic distribution of the density estimate $\hat{f}'(x)$ is approximately $\mathbf{N}(0, f_{X^*}(x)\|K_1\|^2/[nh^3 f_{X^*}^2(0)])$, and the $100(1 - \alpha)\%$ confidence intervals can be obtained by substituting the kernel estimates for $f'_{X^*}(x)$ and $f_{X^*}(0)$, whence one obtains the intervals

$$\hat{f}'(x) \pm \Phi(1 - \alpha/2) \sqrt{\hat{f}_{X^*}(x)\|K_1\|^2/[nh^3 \hat{f}_{X^*}^2(0)]}.$$

Similarly, the $100(1 - \alpha)\%$ confidence interval of the hazard rate $h(x)$, estimated by $\hat{h}(x) = \hat{f}'_{X^*}(x)/\hat{f}_{X^*}(x)$, is obtained by

$$\hat{h}(x) \pm \Phi(1 - \alpha/2) \sqrt{\hat{f}_{X^*}(x)\|K_1\|^2/[nh^3 \hat{f}_{X^*}^2(x)]}.$$

We note that common choices for kernels K, K_0, K_1 are $K(x) = 0.75(1 - x^2)$ on $[-1, 1]$, $K_0(x) = 12(x + 1)(x + 1/2)$ on $[-1, 0]$, and $K_1(x) = -(3/2)x$.

References

- Austad SN (1993) Retarded senescence in an insular population of Virginia opossums (*Didelphis virginiana*). *Journal of Zoology* **229**, 695-708.
- Carey JR (2002) Development of a model for assessing the demography of a captive cohort. *Unpublished manuscript*.
- Carey JR, Liedo P, Müller HG, Wang JL, Chiou JM (1998) Relationship of age patterns of fecundity to mortality, longevity, and lifetime reproduction in a large cohort of Mediterranean fruit fly females. *Journal of Gerontology – Biol. Sci.* **53**, 245-251.
- Caswell H (2001) *Matrix population models: construction, analysis, and interpretation* (2nd ed). Sunderland, Mass.: Sinauer Associates.
- Caughley G (1977) *Analysis of Vertebrate Populations*. Chichester: John Wiley & Sons.
- Congdon JD, Dunham AE, VanLobenSels RC (1994) Demographics of common snapping turtles (*Chelydra serpentina*): Implications for conservation and management of long-lived organisms. *American Zoologist* **34**, 397-408.
- Doob JL (1948) Renewal theory from the point of view of the theory of probability. *Transactions of the American Mathematical Society* **63**, 429.
- Feller W (1968) *Introduction to Probability Theory, vol. I*. New York: Wiley.
- Finch, CE (2001) History and prospects: symposium on organisms with slow aging. *Experimental Gerontology* **36**, 593-597.
- Gaillard JM, Allaine D, Pontier D, Yoccoz NG, Promislow DEL (1994) Senescence in natural populations of mammals: A reanalysis. *Evolution* **48**, 509-516.
- Gershon, H. and D. Gershon (2002) *Caenorhabditis elegans*—a paradigm for aging research: advantages and limitations. *Mechanisms in Ageing and Development* **123**, 261-274.
- Hawkes K, O’Connell JF, Jones NGB, Alvarez H, Charnov EL (1998) Grandmothering, menopause, and the evolution of human life histories. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 1336-1339.
- Hill K, A.M. Hurtado AM (1996) *Ache Life History*. New York: A. de Gruyter.
- Howell N (1979) *Demography of the Dobe !Kung*. New York: Academic Press.
- Jones NGB, Hawkes K, O’Connell JF (2002) Antiquity of postreproductive life: Are there modern impacts on hunter-gatherer postreproductive life spans? *American Journal of Human Biology* **14**, 184-205.
- Krebs CJ (1999) *Ecological Methodology* (2nd ed). Menlo Park: Addison Wesley.

- Lehane MJ (1985) Determining the age of an insect. *Parasitology Today* **1**, 81-85.
- Lincoln FC (1930) Calculating waterfowl abundance on the basis of banding returns. *U.S. Department of Agriculture Circular* **118**.
- Morris RF (1959) Single-factor analysis in population dynamics. *Ecology* **40**, 580-588.
- Morris RF (1965) Contemporaneous mortality factors in population dynamics. *The Canadian Entomologist* **97**, 1173-1184.
- Müller HG (1997) Density estimation. In *Encyclopedia of Statistical Science* (Kotz S, Read CB, Banks DL, ed). New York: Wiley, pp. 185-200.
- Müller HG, Carey JR, Wu D, Liedo P, Vaupel JW (2001) Reproductive potential predicts longevity of female Mediterranean fruit flies. *Proceedings of the Royal Society B* **268**, 445-450.
- Müller HG, Love B, Hoppa R (2002) A semiparametric method for estimating paleodemographic profiles from age indicator data. *American Journal of Physical Anthropology* **116**, 1-14.
- Müller HG, Wang JL (1994) Hazard rate estimation under random censoring with varying kernels and bandwidths. *Biometrics* **50**, 61-76.
- Papadopoulos NT, Carey JR, Katsoyannos BI, Kouloussis NA, Müller HG, Liu X (2002) Supine behaviour predicts time to death in Mediterranean fruit flies (*Ceratitis capitata*). *Proceedings of the Royal Society B* **269**, 1633-1637.
- Papadopoulos NT, Katsoyannos BI, Kouloussis NA, Carey JR, Müller H-G, Zhang Y. (2004) High sexual calling rates predict extended life span in male Mediterranean fruit flies. *Oecologia* (in press).
- Promislow DEL (1991) Senescence in natural populations of mammals: a comparative study. *Evolution* **45**, 1869-1887.
- Reznick, D. (1993) New model systems for studying the evolutionary biology of aging: crustacea. *Genetica* **91**, 79-88.
- Reznick D, Buckwalter G, Groff J, Elder D (2001) The evolution of senescence in natural populations of guppies (*Poecilia reticulata*): a comparative approach. *Experimental Gerontology* **36**, 791-812.
- Sinclair ARE (1977) *The African buffalo: A study of resource limitation or populations*. Chicago: University of Chicago Press.
- Southwood TRE, Henderson PA (2000) *Ecological Methods* (3rd ed). Malden, Massachusetts: Blackwell Science,
- Tatar M, Yin CM (2001) Slow aging during insect reproductive diapause: why butterflies, grasshoppers and flies are like worms. *Experimental Gerontology* **36**, 723-738.

Watelet L, Winter BB (1991) Nonparametric estimation of nonincreasing densities and use of data from renewal processes. *Communications in Statistics – Theory and Methods* **20**, 2073-2094.

Winter BB (1989) Joint simulation and forward recurrence times in a renewal process. *Journal of Applied Probability* **26**, 404-407.

Table 1: Illustration of the relationship between hypothetical 'wild' and 'marked sample' lifetables in the stationary case (from Carey, 2002). The 'wild' cohort consists of n individuals at each age x with corresponding schedules of survival l_x and age structure $c_x = l_x / \sum l_y$, with lifetable in the leftmost subtable. The 'marked sample' cohort consists of initially 20 'marked' individuals with the same age structure as the 'wild' cohort, all simultaneously entering the marked sample cohort at the age of capture and marking $x^* = 0$. Remaining lifetimes are recorded for the marked sample, N_{x^*} is the number of animals that remain alive at age x^* after marking, and l_{x^*} is the survival schedule of the marked sample cohort, with death rates $d_{x^*} = l_{x^*+1} - l_{x^*}$, as listed in the rightmost subtable. The survival schedules separately for age cohorts $x = 0, x = 1, x = 2$ and $x = 3$ in dependency on marked sample cohort age x^* are listed in the corresponding columns of the sub-table in the middle. In this hypothetical example, the initial marked sample cohort at marked sample cohort age $x^* = 0$ has an age structure identical to c_x (bolded row in middle subtable is identical to bolded c_x column of leftmost sub-table). The key identity is revealed by the equality of bolded columns c_x and d_{x^*} in leftmost and rightmost sub-tables. This key relationship allows to deduce the wild survival schedule from the marked sample survival schedule.

Wild Cohort				Age Distribution in Marked Sample Cohort				Marked Sample			
x	N_x	l_x	c_x	$x = 0$	$x = 1$	$x = 2$	$x = 3$	x^*	N_{x^*}	l_{x^*}	d_{x^*}
0	40	1.000	0.40	0.40	0.30	0.25	0.05	0	20	1.00	0.40
1	30	0.750	0.30	0.30	0.25	0.05		1	12	0.60	0.30
2	25	0.625	0.25	0.25	0.05			2	6	0.30	0.25
3	5	0.125	0.05	0.05				3	1	0.05	0.05
4	0	0.000	0.00	0.00				4	0	0.00	0.00
	100	2.5								1.95	1.00

Legend

Figure 1: Reconstruction of survival function of wild population from six simulated marked sample cohorts of sizes $N=50$ (upper panel) and $N=1000$ (lower panel). This reconstruction is based on a key demographic identity and corresponding nonparametric estimation methods as described in text. The solid curve is the target survival function that corresponds to the observed survival schedule of a cohort of medflies.

