

动态生命表寿命数据收集 与处理过程中常见问题的探讨*

宛新荣 钟文勤 王梦军 王广和 刘 伟

(中国科学院动物研究所农业虫鼠害综合治理研究国家重点实验室, 北京, 100080)

摘要: 种群动态生命表是研究野生动物存活过程的有用工具, 而其可靠性则依赖于生命数据收集与处理过程中的正确性。文章就野生动物动态生命表收集和处理过程中常见的右删失数据处理问题、个体死亡时间估计、同生群初始值 N_0 估计以及早期死亡数据的估算问题提出了一点建议和讨论。

关键词: 野生动物; 动态生命表; 寿命数据; 右删失数据

中图分类号: Q141 文献标识码: A 文章编号: 1000 - 1050 (2001) 02 - 0137

- 06

生命表 (life table) 是描述种群死亡过程的有用工具, 在种群生态学研究已显示出广泛的应用前景。它最初应用于人口统计学, 尤其是在人寿保险事业^[1, 2]。近年来, 已有不少学者着手动植物生命表的研究^{3~12]}。

生命表分为两大类, 即动态生命表和静态生命表。前者是根据观察一群同时出生的生物个体的死亡或存活过程而获得的数据编制生命表; 后者则是在特定时间里通过调查生物种群的年龄结构而编制。动态生命表的个体经受的环境条件是相同的, 因而在实际应用中, 动态生命表能提供更可靠, 更有价值的信息。一般来说, 动态生命表的数据收集过程中存在着相当大的困难^{13~16]}, 具体体现在如下几个方面: (1) 动物个体的标识与年龄鉴定; (2) 个体死亡时间估计; (3) 同生群 (cohort) 初始值 N_0 的确定; (4) 动物早期死亡数据的估计。因此, 如何收集野生动物种群的寿命数据、克服数据处理过程中的难点, 成为编制动态生命表的关键问题。目前, 有关兽类动态生命表的研究国内报道不多, 国外涉及此类研究的也少见^[7]。在自然条件下, 确定每个动物个体的实际年龄是寿命数据收集的首要问题^{16]}, 一般可用标志重捕结合观测或者部分采用无线电遥测法测定, 其具体操作过程因动物类群、种类特点而异, 这里不再赘述。本文仅就野生动物寿命数据收集与处理过程中其他几个经常遇到的问题如个体死亡时间估计、同生群初始值 N_0 的估计、早期死亡数据的估计以及目前容易忽略的右删失数据处理问题提出解决途径的设想和建议, 谨供商讨。

* 基金项目: 国家自然科学基金 (39870136) 以及自然科学基金重点资助项目 (39730090)

作者简介: 宛新荣 (1969 -), 男, 博士, 助理研究员, 主要从事动物种群生态学研究。

收稿日期: 1999 - 11 - 24; 修回日期: 2000 - 08 - 29

1 右删失数据的处理

右删失数据 (right-censored data) 只出现在动态生命表的寿命数据收集过程中。目前, 涉及右删失数据处理方法的文献很少, 因而它一直没有引起研究人员的足够重视, 往往为研究人员所忽略。然而, 妥善处理好右删失数据是保证分析结果可靠性的重要条件, 下文首先讨论这个问题。

按照 Nelson 的定义^[17], 常见寿命数据主要为寿终数据 (complete data)、右删失数据、左删失数据 (left-censored data) 和区间性数据 (interval data)。寿终数据又叫完全寿命数据, 即一个个体的确切寿命已知 (顾名思义: 寿终正寝)。这是寿命数据最普通、最重要的一种数据类型。而右删失数据的定义为一个个体的确切寿命不知道, 只知道其寿命大于某个值 (该数值即为右删失时间或右删失年龄)。如生命数据收集过程中由于人为因素或其它不可避免的因素造成的数据收集中断将直接导致右删失数据的产生。例如: 由于偷猎原因造成调查动物的伤亡; 取样中有意去除的一部分个体 (removals); 重捕操作中的捕捉丧失 (losses on traps) 或追踪个体迁出样地而中断追踪, 这些个体的实际寿命要大于上述事件发生时刻该动物的年龄, 但其确切值无从知晓。由于野外条件下不可控因素很多, 右删失数据在野生动物寿命收集过程中极为常见, 虽然比例不高, 但必须将其与普通的寿终数据区别对待, 否则将使生存函数的估计值偏低, 从而影响估算结果的正确性。如果收集到的寿命数据中包含右删失数据, 那么, 最好采用具有删失机制的寿命表法 (life table method) 或者乘积限估计 (Product-Limit Estimation) 估算其生存函数^[18]。

2 个体死亡时间估计

对于右删失数据, 一般能够准确推算其右删失日龄。但对于正常个体死亡时间的判别, 在野外条件下就存在一定的困难^[19], 具体地说, 在野外直接观测到动物的死亡时间近乎不可能。某些大型动物, 一旦死亡, 其尸体很快便被其它动物吃掉; 对小型穴居动物而言, 多死于地下洞内或被天敌捕食, 仅少量个体死于地面, 而地面上的尸体也很快被食尸动物 (如步行虫类) 吃掉。诚然, 生命表法允许某些个体的死亡时间以区间形式来表示^[18], 但该方法在野生动物上 (尤其对出生时间不一致的种类) 使用的局限性也是很明显的: 假如收集 1999 年 5 月份出生的雄性布氏田鼠同生群的生命数据, 由于不能保证同生群中的所有个体都出生于同一日期, 所以同一调查时刻 (不妨假定在 6 月 5 日), 不同个体的实际年龄 (或日龄) 仍有一定的差异如 10, 15, 17, 23, 26, 34, ... 等, 那么, 假定在某个固定时间间隔 (如 10 天, 即在 6 月 15 日再次调查), 如果部分个体已经有死亡, 就可能得到如下情形的寿命数据: [10, 20), [15, 25), [17, 27), [23, 33), [26, 36), [34, 44), ... 等。至此, 虽然获得了区间型的生命数据, 但由于每个个体的寿命区间分别与其他多个个体的寿命区间交叉迭合 (此外, 由于在野外条件下, 取样间隔多半不固定, 取样期长度往往还受天气条件的影响, 因此不同个体寿命区间的实际迭合混杂程度远远高于此例), 这类数据实际上根本不能按照传统的生命表法来分析处理。因此, 动物个体死亡时间估计问题事实上是研究人员无法回避的。

至于如何^{万方数据}解决个体死亡时间估计问题, 现有文献却鲜有详尽的论述, 一些生态学教

科书和工具手册也完全忽略了这个问题。为此，作者结合自身工作的体验，尝试以中点值估计法推算特定个体的死亡时间——供大家探讨。其具体规程为：假设在一系列连续取样（重捕或观测中）， t_i 为最后一次已知某个个体还“活着”的取样（重捕或观测）时间，而且在其后的任何一次取样中均未见到该鼠的活动，那么即假定该个体确已死亡（其前提是该个体未迁出样地），其死亡时间为 $(t_i + t_{i+1})/2$ ，其中 t_{i+1} 为下一次取样的时间。中点值估计法的可靠性主要依据两个方面：（1）每次取样的捕获率或观测率大小，这个值越大，估算值可靠性越高（否则无法确认个体在 t_{i+1} 时刻已经死亡）；（2）两次观测的间隔，显然，间隔越小，偏差越小。如果在实验设计中同时保证较高的捕获率（或观测率）并同时缩短取样间隔，那么，中点值死亡时间估计值则接近于真实值，从而大大降低估计偏差。

3 同生群初始值 N_0 的确定

对同生群初始值 N_0 的估计可以说是动态生命表的基石， N_0 估计值出现偏差将导致整个生命表估计值出现偏差。一般来说，实验室饲养条件下很容易得到 N_0 ^[9]，而自然条件下，野生动物的同生群初始值 N_0 在多数场合是很难确定的。例如调查小鸟带环以后的存活状况虽然比较容易，但要知道某一年该鸟的总产卵数或孵化数便十分困难^[20]。有些研究有蹄类的学者干脆放弃了对 0~1 龄个体的存活追踪而直接从 1 龄个体开始计算。Connell 追踪了两种藤壶 (*Balanus balanoides*, *B. glandula*) 的存活过程，也是从藤壶在岩石上固着后而不是从自由游泳幼体初态就开始追踪^[5,6]。国内有些学者则采用啮齿动物的初次出窝时间反推其出生日期^[10]。上述处理方法都同样回避了对同生群初始值 N_0 的估计，这样，估算动物的生命期望时只能从其开始追踪时刻算起而不能从出生时刻算起，否则就很可能由于忽略种群早期死亡数据而导致对种群期望寿命的估计值偏高。

同生群的初始值 N_0 的估计十分重要，其估计偏差将直接影响到生存函数的估计，从而影响生命期望以及存活曲线类型的判断。对野生动物动态生命表编制者而言，对它的估计则是很难回避的。在不同的动物类群，估计 N_0 的方法与技巧应各有差异。在现有生态学教科书中，也很少有这方面的论述，为此，作者结合自身的工作经验，摸索一种针对啮齿动物的 N_0 估计方法。其具体操作程序如下：（1）全面跟踪特定时间调查样地内的雌鼠繁殖记录得到其产仔窝数。可借助标志重捕技术以及染色观测（染色观测只对昼行性鼠有效）来获得。（2）确定特定时间（年份/月份）雌鼠的平均胎仔数。这可借助夹捕解剖资料或者对野外捕获鼠的直接饲养繁殖获得。必要的话还可以区分一定的年龄组（如当年鼠或者越冬鼠），分别统计各年龄组的产仔窝数以及胎仔数。通过（1）与（2）的乘积便可近似估计到同生群初始值 N_0 （含雌雄两性）。

在做生命表分析处理时，一般需要将两性分开估算。因此，还须分别估算出两性的初始值 N_0 ，这样就必须首先确定其出生性比。估算其出生性比，有两种备择方案：（1）通过捕获野外临近分娩鼠的饲养而调查幼鼠的出生性比，但需要保证一定的临近分娩鼠的数量，其操作程序略显繁琐；（2）利用假设：新生鼠从出生到出窝（或初捕日龄）这段时间内存活率与性别无关。这项假设的依据是：幼鼠在此之前远远未达到性成熟，因此不考虑此期间存活率的性别差异可能是合理的。张知彬等在研究大仓鼠种群时发现，可以用成体性比作为其胚胎性别的估计值^[21]。第 2 种方案无疑更为简洁，但需要验证相应的假设。按照方案（2），雌雄两性个体的初生数目估计方法如下（方案 1 的步骤与

其相似):

N_0 : 初生个体数 (两性); N_{0m} : 初生雄体数; N_{0f} : 初生雌体数; N_m : 出窝时 (或者最初进入重捕年龄) 已知存活的雄体数; N_f : 出窝时 (或者最初进入重捕年龄) 已知存活的雌体数

通过野外调查可直接得到 N_0 , N_m 以及 N_f 。那么依据假设就有:

$$N_{0m} = \frac{N_0 N_m}{N_m + N_f}; \quad N_{0f} = \frac{N_0 N_f}{N_m + N_f}$$

这样两性个体的初始数目 N_{0m} 、 N_{0f} 就可以通过上述方法确定。上式的计算得到的 N_{0m} 与 N_{0f} 若带有小数部分, 则采用四舍五入法取整, 或者通过随机函数表决定其取舍 (若随机数小于分数部分则进位, 随机数大于分数部分则舍去)。

4 早期死亡个体寿命数据的处理方法

静态生命表的一个严重不足在于其早期死亡数据的缺乏, 由于幼年动物死亡率相对较高, 而根据动物种群年龄结构编制的生命表很容易低估其早期死亡率。而在动态生命表的编制过程中, 一旦解决了同生群初始值 N_0 的估计问题, 结合动物的出窝数量, 就很容易计算出早期死亡个体的数目 ($N_{0m} - N_m$)。但这里也出现一个新的问题: 对早期死亡个体寿命数据的处理问题。对于进入初捕年龄之后的个体, 其死亡时间估计可参照前文的中点值估计法, 但若以此法估计进入初捕年龄之前 (或者出窝前) 的早期死亡个体的死亡时间, 则显得很不当。因为早期死亡个体数很多, 按中点值估计就假定了所有这些个体均在同一日龄 (从其出生时间到初捕平均日龄的中点时间) 死亡, 这样势必会造成存活曲线的大起大落。那么究竟如何处理这些早期死亡个体的死亡时间呢? 在此作者也提出两种途径: (1) 将出生到出窝这段时间视为一个整体, 统一从整体上估算存活率, 不再分别考虑每个个体的死亡时间。这种方案的优点是简单, 又能基本保证编制生命表的可靠性, 其缺点是不利于进行不同同生群间的生存函数的统计比较。(2) 为了便于生存模式间的比较分析, 需要估算出每个早期死亡数据的具体数值, 这样就必须了解从出生到出窝这段时间上存活函数的分布特征。最简单的方法为假定在这段时间内存活率处处相等, 这样存活函数可用指数函数关系来表示:

$$S(t) = \exp(-kt) \quad (4)$$

其中 $S(t)$ 为生存函数, k 为正常数, 定义其平均出窝年龄为 t' , 那么, 当 $t = t'$ 时,

有: $S(t) = \frac{N_m}{N_{0m}}$, 上述关系代入 (4) 式, 得:

$$S(t) = \frac{N_m}{N_{0m}} = \exp(-kt') \quad (5)$$

由于 N_m , N_{0m} 和 t' 已知, 这样可以先确定参数 k 的数值:

$$k = \frac{1}{t'} \cdot (\ln \frac{N_{0m}}{N_m}) \quad (6)$$

那么，每个早期死亡时间 t_i 估算如下：

$$\frac{N_{0m} + 0.5 - i}{N_{0m}} = \exp(-kt') \quad (7)$$

(7) 式中，依次以 $i = 1, i = 2, \dots$ ，和 $i = (N_{0m} - N_m)$ ，代入即可解出各个早期死亡数据的数值 t_i ，其中绝大多数含有小数位数，可根据具体情况取整。

为便于读者理解这一步骤，下面举一应用实例：设同生群数量为 30 ($N_{0m} = 30$)，出窝数量为 20 ($N_m = 20$)，那么，早期死亡个体数量为 10 ($N_{0m} - N_m = 10$)，假定平均出窝年龄为 30d ($t' = 30$)，根据假设给出的条件，估算 10 个早期死亡年龄数据的步骤如下：

I. 根据公式 (6) 确定参数 k 的数值 (为保证精确度，宜保留多位小数)

$$k = \frac{1}{t'} \cdot \left(\ln \frac{N_{0m}}{N_m} \right) = \frac{1}{30} \cdot \ln \frac{30}{20} = 0.013 52$$

II. 根据公式 (7) 依次以 $i = 1, i = 2, \dots$ ，和 $i = 10$ 代入解出 t_i 的近似数值

$$\frac{30 + 0.5 - 1}{30} = \exp(-0.013 52 \cdot t_1), \text{ 得到: } t_1 = 1.2 \text{ (d)}$$

$$\frac{30 + 0.5 - 2}{30} = \exp(-0.013 52 \cdot t_2), \text{ 得到: } t_2 = 3.8 \text{ (d)}$$

.....

$$\frac{30 + 0.5 - 10}{30} = \exp(-0.013 52 \cdot t_{10}), \text{ 得到 } t_{10} = 28 \text{ (d)}$$

这样，就获得了 10 个早期死亡的动物的寿命估计值：1.2 d, 3.8 d, ..., 28 d 等。

5 结语

动态生命表的编制是研究野生动物种群的一个重要研究内容，寿命数据收集和处理过程是否得当则决定了所编制的动态生命表的可靠性，由于多数生态学教科书和工具手册类书籍很少涉及相关问题的讨论，目前对于寿命数据收集与处理过程中遇到的问题仍然缺乏足够的重视。本文所列举的 4 个问题，从事野生动物动态生命表研究的工作人员可能都会碰到，但如何妥善地解决这些问题，确有必要进行专门的讨论。本文仅就野生动物动态生命表寿命数据收集与处理过程中几个常见问题，结合作者的工作体验，提出一点建议和设想，目的在于抛砖引玉，以期在广大同行的踊跃参与下，最终能够妥善地处理好动态生命表数据收集与处理过程中的难点。

参考文献：

- [1] Krebs C J. Ecology [M]. New York: Harper and Row Publishers, 1996. 168 - 195.
- [2] 孙儒泳. 动物生态学原理 [M]. 北京: 北京师范大学出版社, 1992. 130 - 146.
- [3] Crisp M D, Lange R T. Age structure distribution and survival under grazing of arid zone shrub *Acacia burkittii* [J]. *Oikos*, 1976, 27: 86 - 92.
- [4] Leverich W J, Levin D A. Age-specific survivorship and reproduction in *Phlox drummondii* [J]. *American Naturalist*.

- 1979, 113 : 881 - 903.
- [5] Connell J H. The effects of competition predation by *Thais lapillus* and other on natural population of the barnacle, *Balanus balanoides* [J]. *Ecological Monographs*, 1961, 31 : 61 - 104.
- [6] Connell J H. A predator-prey system in the marine intertidal region, *Balanus glandula* and several predatory species of *Thais* [J]. *Ecological Monographs*, 1970, 40 : 49 - 78.
- [7] Lowe V P W. Population dynamics of the red deer (*Cervus elaphus*) on Rhum [J]. *Journal of Animal Ecology*, 1969, 38 : 425 - 457.
- [8] Sinclair A R E. The African Buffalo [M]. Chicago : University of Chicago Press, 1977.
- [9] 梁杰荣, 孙儒泳. 根田鼠生命表和繁殖的研究 [J]. 动物学报, 1985, 31 (2): 170 - 177.
- [10] 王学高, 戴克华. 高原鼠兔自然寿命研究 [J]. 兽类学报, 1989, 9 (1): 56 - 62.
- [11] 魏辅文, 胡锦涛, 许光瓚, 江明道, 邓启涛, 钟肇敏. 野生大熊猫生命表初编 [J]. 兽类学报, 1989, 9 (2): 81 - 86.
- [12] 李进华, 王岐山, 李明. 短尾猴种群生态学研究 III. 年龄结构和生命表 [J]. 兽类学报, 1995, 15 (1): 31 - 35.
- [13] Sherman P W, Morton M L. Demography of Belding's ground squirrels [J]. *Ecology*, 1984, 65 : 1617 - 1628.
- [14] Sherman P W, Zammuto R M. Time-specific and cohort life tables for Belding's ground squirrels [J]. *Ecology*, 1993, 74 : 2168 - 2169.
- [15] Menkens G E Jr, Boyce M S. Comments on the use of time-specific and cohort life tables [J]. *Ecology*, 1993, 74 : 2164 - 2168.
- [16] Hartman J D. Age determination, age structure, and longevity in the mole, *Scalopus aquaticus* (Mammalia : Insectivora) [J]. *Journal of Zoology*, 1995, 237 : 107 - 122.
- [17] Nelson W. Applied life data analysis [M]. New York : John Wiley & Sons, 1982.
- [18] 陈家鼎. 生存分析与可靠性引论 [M]. 合肥 : 安徽教育出版社. 1982. 1 - 43.
- [19] Clinton W L, Le Boeuf B J. Sexual selections effects on male life history and the pattern of male Mortality [J]. *Ecology*, 1993, 74 : 1884 - 1892.
- [20] 伊藤嘉昭. 动物生态学研究法 [M]. 北京 : 科学出版社, 1986.
- [21] 张知彬, 朱靖, 杨荷芳. 大仓鼠种群年龄组存活率的估算 [J]. 动物学报, 1993, 39 (1): 56 - 63.

DISCUSSION ON SOME COMMON PROBLEMS IN THE LIFETIME DATA COLLECTING AND DISPOSING PROCESSES FOR CONSTRUCTING COHORT LIFE TABLE

WAN Xinrong ZHONG Wenqin WANG Mengjun WANG Guanghe LIU Wei

(Chinese State Key Lab of Integrated Management of Pest Insects and
Rodents, Institute of Zoology, the Chinese Academy of Sciences, Beijing, 100080)

Abstract : The cohort life table is a convenient format to describe survival course of wildlife population. Its accuracy and reliability depend on the lifetime data collecting and disposing processes. Some commonly encountered problems, such as the right-censored data disposing, the failure time estimating, the initial size of cohort estimating, and the early failure data disposing are proposed and discussed here.

Key words : Wild animal ; Cohort life table ; Lifetime data ; Right-censored data