

# 关于Logistic增长模型参数估计方法的再探讨

杨益民<sup>1</sup>, 付必谦<sup>2</sup>

(1. 北京工商大学 理学院, 北京, 100084; 2. 首都师范大学 生命科学学院, 北京 100084)

**摘 要:** 对于 Logistic 增长模型  $N=K/(1+e^{a-rt})$  中的三个常数  $K$ 、 $r$  和  $a$ , 绝大多数拟合方法均将其作为三个独立参数来估计, 这样不仅使拟合得到的理论初值  $\hat{N}_0$  与实测初值  $N_0$  不等, 而且可能明显影响  $K$  和  $r$  的估计结果。文章对利用线性化方法拟合 Logistic 模型进行了再探讨, 给出了一种估计参数  $K$ 、 $r$  和  $a$  的新方法并证明了其合理性。

**关键词:** Logistic 增长模型; 参数估计; 线性回归; 黄金分割法; “最佳”观测点数定向搜索法

**中图分类号:** O212

**文献标识码:** A

**文章编号:** 1002-6487(2015)13-0028-05

## 0 引言

$$\text{Logistic 增长模型 } N = \frac{K}{1 + e^{a - rt}} \quad (1)$$

是描述有限环境条件下生物种群 S-型增长的最基本和最常用的模型, 在种群数量动态研究中占有十分重要的地位。

在利用式(1)定量研究种群动态时, 必须依据实验观测数据估计式中的  $K$ 、 $r$  和  $a$ 。较早提出且目前仍然十分常用的方法是将式(1)移项取对数后, 转化为线性方程

$$y = \ln \frac{K - N}{N} = a - rt \quad (2)$$

先根据原始数据和经验预先给定  $K$ , 然后再利用一元线性回归估计  $r$  和  $a$ <sup>[1-3]</sup>。后来一些学者认为线性化方法主观色彩较浓或遗漏“最优”参数估计值的可能性较大, 提出采用一些非线性最优化方法如麦夸法和单纯形加速法等<sup>[4,5]</sup>, 也有作者对利用数值方法拟合进行了一些探讨<sup>[6]</sup>。但是, 无论是利用线性化还是非线性化方法(包括直接调用统计软件包), 在 Logistic 模型参数估计中均存在一些共同的需要特别重视的问题。此外, 对现有线性化方法进行改进, 探讨一种既简便易行又客观准确的线性化方法, 对于缺乏较深数学基础的实际工作者仍不失一种较好的选择。

本文对目前 Logistic 模型参数估计中的问题进行了分析, 并对线性化方法进行了再探讨, 提出了一种新的线性化方法——“最佳”观测点数定向搜索法。将该方法应用于实例的分析, 获得了满意的拟合效果。

## 1 Logistic 增长模型中 $a$ 的估计方法

对于 Logistic 增长模型的拟合, 绝大多数作者均将式(1)中的  $a$ 、 $K$  和  $r$  作为 3 个独立的参数进行估计。在统计软件包相关教程中, Verhulst 模型的表达式也为

$y = b_1 / (1 + b_3 \exp(-b_2 x))$ , 其中  $b_i (i=1, 2, 3)$  为待估计的参数<sup>[7]</sup>。

利用这些方法估计参数, 计算得到的理论初值常与实测初值不等, 甚至差异很大。虽然唐启义等<sup>[8]</sup>曾对上述作法提出过质疑, 但似乎并未引起重视。

因此, Logistic 模型参数估计所必须解决的一个重要问题是: 能将  $a$  作为与  $K$  和  $r$  一样的独立的参数估计吗? 经过拟合得到的理论初值是否应与实测初值  $N_0$  相同?

为探讨上述问题, 将 Logistic 增长模型:

$$\begin{cases} \frac{dN}{dt} = rN(1 - \frac{1}{K}N) \\ N(0) = N_0 \end{cases} \quad (3)$$

解得:

$$N = \frac{K}{1 + (\frac{K - N_0}{N_0})e^{-rt}} \quad (4)$$

与(1)是比较, 易得:

$$a = \ln(\frac{K - N_0}{N_0}) \quad (5)$$

由上可见, Logistic 增长模型只有  $K$  和  $r$  两个独立参数; 而  $a$  是依赖于  $K$  与种群初值  $N_0$  的常数, 而且  $N_0$  应在解微分方程之前就已确定而非获得通解后再行估计。此外, 从生态学角度考虑, 初始种群大小对于种群增长非常重要, 并常常是种群增长实验中唯一一个可以准确测定而非通过取样估计的值。因此, 笔者赞成唐启义等<sup>[8]</sup>的意见, 即  $a$  应据式(4)直接由参数  $K$  和种群实测初值  $N_0$  解出而非作为一个独立的参数进行估计。这样得到的  $a$  值可以保证理论初值与实测初值相同, 使  $K$  和  $r$  的估计更为客观正确。前述大部分作者<sup>[1-5]</sup>的拟合方法均不能保证理论初值与实测初值相同, 虽然其中一些方法理论值与实测值的残差平方和  $SS_e$  ( $SS_e = \sum_{i=0}^n (N_i - \hat{N}_i)^2$ ) 较小, 但其拟合结果可能并不能真实反映实验观测数据本身的特点。当利用线性回归进行拟合时, 由式(4)移项取对数得,

**基金项目:** 北京市第十届教学名师专项资助项目(PXM2015.014213.00032)

$$Y = \ln\left(\frac{K - N_0}{N_0}\right) - \ln\left|\frac{K - N}{N}\right| = rt \quad (6)$$

则待拟合的方程应为式(6)而非式(2)。

## 1 利用线性回归分析估计参数K和r的新方法:“最佳”观测点数定向搜索法

关于参数K和r的估计,许多线性化方法是首先估计一个K值,然后将其代入式(2)估计r,因此r的估计在很大程度上依赖于K的估计精度并受回归分析采用的观测点数的影响。若K值估计精度不够高或不确定性较大或选择的观测点数目不恰当,均可能严重影响最终拟合结果。万昌秀和梁中宇<sup>[9]</sup>提出的枚举选优法将种群达到平衡后的最大和最小观测值作为K的上、下界,分别将该范围内的每个整数作为K值,并对每个K值均分别从全部观测值中选取前m个观测值( $m=4,5,6,\dots,n$ ;n为观测值总数)进行回归分析估计可能的r值,然后逐一比较各组参数估计值对应的 $SS_e$ ,以 $SS_e$ 最小为标准确定最终参数估计值。该方法考虑全面、稳妥可靠,明显提高了拟合优度;但因计算量大(种群数量很大和观测数据很多的实验数据尤其如此)而需编制专门的计算机程序,并难以处理非整数型数据或种群未达平衡状态便终止实验的不完整的观测数据。

针对上述问题,在提出利用式(6)作为拟合方程的基础上,本文提出一种新的线性化方法——“最佳”观测点数定向搜索法。该方法的要点为:(1)首先确定适当的K值初始搜索区间;(2)利用该区间端点和中值确定一个在整个K值搜索过程中进行线性回归分析时均可采用的适宜的观测点数目,从而不必对每个K值均选取前m个观测值进行分析;(3)利用黄金分割法定向搜索K的最优估计值。

### 1.1 K值初始搜索区间 $[K_a, K_b]$ 的确定

#### 1.1.1 完整实验数据初始搜索区间的确定

对于实验数据相对完整、有多个观测值围绕平衡数量波动的情况,一般来讲,参数K的最优估计值 $K^*$ 应在这些观测值的平均值( $K_m$ )附近,因此可直接将 $K_m$ 作为初始搜索区间的中值 $K_{mid}$ 来确定初始搜索区间 $[K_a, K_b]$ ( $K_{mid} = (K_a + K_b)/2$ )。若在这些数据中存在个别观测值明显偏大或偏小,可去除这些观测值后再计算 $K_m$ 。初始搜索区间应稍大一些,最好包括围绕平衡数量波动的所有观测值(异常值除外)。

#### 1.1.2 不完整实验数据初始搜索区间的确定

若种群未达平衡状态便终止了实验,则无法依据上述方法确定 $K_{mid}$ 和初始搜索区间,而需预先估计一个搜索区间,然后再行调整。

(1)初始搜索区间的预估。对于增长曲线上端已经明显变缓的实验数据,可直接通过目测预估初始搜索区间 $[K'_a, K'_b]$ ,计算中值 $K'_{mid}$ 。若实验数据明显不完整,难以估计上渐近线的大致位置,则可随机选取3组等时间间距观测数据依据Pearl和Reed的三点法十分方便地估计3个K值,计算其均值 $K_m$ ,以 $K_m$ 作为区间中值 $K'_{mid}$ 初步估计初

始搜索区间 $[K'_a, K'_b]$ 。初始搜索区间应包括上述3个K值并向上、下有一定延伸;对于增长曲线上端已出现转折的实验数据,还应包括转折段的观测值,以减小遗漏 $K^*$ 的可能性。

(2)初始搜索区间的调整。分别将 $K'_a$ 、 $K'_b$ 和 $K'_{mid}$ 代入式(6),采用前 $m_i$ 个观测数据进行回归分析估计 $r(m_i$ 的确定方法见2.2),然后将3个K值和相应的r值代入式(4)估计 $\hat{N}_i$ ,计算各个K值对应的 $SS_e$ 。由式(4)和式(6)可知,对于特定的一组实验观测数据, $SS_e$ 仅随K和m的变化而变化,因此可通过比较 $SS_e$ 的大小决定K值的取舍或确定适宜的m值。若某端点的 $SS_e$ 最低,则以该端点作为新区间的中值,将搜索区间平移。重复上述调整过程,直至区间中值对应的 $SS_e$ 最低。将最终的搜索区间确定为 $[K_a, K_b]$ 。

### 1.2 回归分析采用的“最佳”观测点数( $M^*$ )的确定

在K值搜索过程中,若初始搜索区间 $[K_a, K_b]$ 设置比较合理, $K^*$ 一般应位于区间中值 $K_{mid}$ 附近,搜索过程中(尤其后期)大多数探测点的K值往往也距 $K_{mid}$ 相对较近,因此可利用区间端点和中值探测一个在整个K值搜索过程中进行回归分析时均可采用的适宜观测点数(“最佳”观测点数 $M^*$ )。具体方法如下:

(1)对于初始搜索区间 $[K_a, K_b]$ ,分别根据式(6)计算端点 $K_a$ 、 $K_b$ 及中值 $K_{mid}$ 对应的Y值,在同一图中作3个K值对应的Y-t折线,然后在图中找到符合下列条件的一点:在该点和该点之前,不仅3条Y-t折线的线性关系均较强且比较类似,而且包括的观测点数目尽可能多,使依据这些观测数据估计r时能够利用更多实验数据的信息和获得更高的拟合优度。可以看到,该点一般对应于3条Y-t折线呈明显“Y”形分离的分叉点(将其对应的观测点数记为 $m_i$ ),但有时也可能为分叉点之前或之后的一点,因此比较适宜的观测点数分别为 $m_i-1$ 、 $m_i$ 和 $m_i+1$ 。

(2)将 $K_{mid}$ 代入式(6),分别以前m个观测数据(一般选取 $m=m_i-1, m_i, m_i+1$ 即可,同时应大致满足 $N < K_a$ )估计r,然后将 $K_{mid}$ 和这些r值代入式(4)估计 $\hat{N}_i$ 。计算各m值对应的 $SS_e$ ,以 $SS_e$ 最小的m作为 $M^*$ 。

本方法的关键在于以 $K_a$ 、 $K_b$ 及 $K_{mid}$ 对应的3条Y-t折线确定 $m_i$ ,因此可称为“三线定点法”。

“三线定点法”的合理性证明:

在初始搜索区间 $[K_a, K_b]$ 中,设 $K_c = K_a + \Delta K$ ( $0 \leq \Delta K \leq K_b - K_a$ ),则对于任一 $K_c$ ,均有

$$Y_c = \ln\left(\frac{K_c - N_0}{N_0}\right) - \ln\left|\frac{K_c - N}{N}\right| = \ln\left(\frac{K_a + \Delta K - N_0}{N_0}\right) - \ln\left|\frac{K_a + \Delta K - N}{N}\right|$$

则:

$$\frac{\partial Y_c}{\partial N} = \frac{K_c}{N(K_c - N)}, \quad \frac{\partial Y_c}{\partial \Delta K} = \frac{N_0 - N}{(K_c - N_0)(K_c - N)}$$

从而,若 $N < K_c$ ,则 $\frac{\partial Y_c}{\partial N} > 0$ ,  $\frac{\partial Y_c}{\partial \Delta K} < 0$ ,  $Y_c$ 随N的增加而增加,随 $\Delta K$ 的增加而下降。反之,若 $N > K_c$ ,则

$\frac{\partial Y_c}{\partial N} < 0$ ,  $\frac{\partial Y_c}{\partial \Delta K} > 0$ ,  $Y_c$  随  $N$  的增加而下降, 随  $\Delta K$  的增加而增加。

因此, 在  $Y-t$  图中, 随着观测点数目变化 (反映  $N$  的变化), 不仅特定  $K$  值对应的  $Y_c$  值的变化趋势, 而且不同  $K$  值对应的  $Y_c$  值彼此之间的大小关系均可能发生明显改变, 由此必然影响参数  $r$  的估计结果。随着  $N$  的增加, 各个  $K_c = K_a + \Delta K$  对应的  $Y-t$  折线将按照  $\Delta K$  的大小依次向下发生转折, 直至  $K_c$  大于全部  $N$  值。在区间  $[K_a, K_b]$  中,  $K_a$  最小, 因此在  $N < K_a$  的数据范围内, 各个  $K$  值对应的  $Y-t$  折线不仅均处于上升段而且位于两端点对应的  $Y-t$  折线之间。若以区间中值  $K_{mid}$  对应的  $Y-t$  折线作为参照线, 则在  $K_a, K_b$  和  $K_{mid}$  对应的 3 条  $Y-t$  折线均呈较强线性关系的观测点范围内, 其他任一  $K_c$  值对应的  $Y-t$  折线也应保持类似关系, 因此可在整个  $K$  值搜索过程中以同样的观测点数目进行回归分析估计  $r$ 。

以 Gause (1934) 双小核草履虫种群增长实验 (添加一环细菌)<sup>[9]</sup> 为例。该实验数据完整, 有多个观测值围绕平衡数量上下波动, 因此可直接以这些观测值的平均值 ( $K_m = 448$ ) 作为中值  $K_{mid}$  确定初始搜索区间  $[418, 478]$ 。为了更加清晰地反映不同  $K$  值对应的  $Y-t$  折线之间的关系, 避免实测值在平衡数量附近的波动对图形的影响, 首先以唐启义等<sup>[10]</sup> 的预测值代替实测值进行分析 (图 1 左图)。在预测值中,  $N_7 = 409 < K_a$ ,  $N_8 = 429 > K_a$ ,  $N_{max} = N_{16} = 439$ 。由图可见, 在前 8 个观测点范围内, 所有  $K_c = K_a + \Delta K (0 < \Delta K \leq K_b - K_a)$  对应的  $Y_c-t$  折线均位于  $Y_a-t$  折线之下且与  $Y_a$  的偏离程度随  $\Delta K$  的增大而增大。在该范围之后, 对于所有  $K_c < N_{max} = 439$ , 其对应的  $Y_c-t$  折线均随  $N$  的增加依次向下转折, 最终的  $Y_c$  值随  $\Delta K$  的增大而增大。而对于其他  $K_c$  值, 因  $K_c > N_{max}$ , 其对应的  $Y_c$  值单调增加直至稳定, 且随  $\Delta K$  的增大而减小。各条  $Y-t$  折线明显“Y”形分叉点对应的观测点数  $m_s = 7$ 。

若直接以实验观测值计算  $Y_c$ , 则得图 1 右图。实验观测值  $N_8 = 396 < K_a$ ,  $N_9 = 443 > K_a$ 。由图可见, 尽管该图右部各  $Y-t$  折线波动剧烈, 但在前 9 个观测点范围内,  $Y-t$  折线的排列却十分有序,  $Y_c$  与  $Y_a$  的偏离程度随  $\Delta K$  的增大而增大。各条  $Y-t$  折线明显“Y”形分叉点对应的观测点数也为  $m_s = 7$ 。

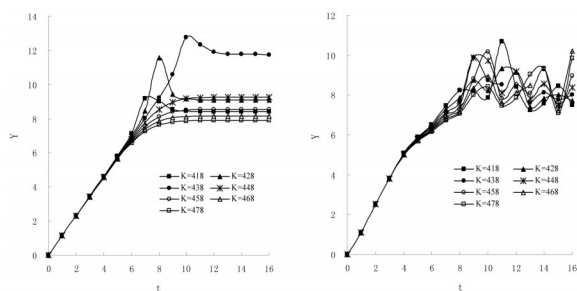


图 1 “三线定点法”示意图。依据 Gause (1934) 双小核草履虫种群增长数据 (添加一环细菌)<sup>[9]</sup> 作图。左: 利用唐启义等预测值<sup>[10]</sup> 计算  $Y_c$ ; 右: 利用 Gause 实测值计算  $Y_c$ 。

### 1.3 $K$ 的“最优”估计值 ( $K^*$ ) 的定向搜索

黄金分割法是运筹学中单变量非线性最优化的重要方法<sup>[18]</sup>, 该法具有搜索方向明确、收敛速度快、计算简单且计算量小的优点, 而且适用于整数型或非整数型观测数据。利用该法搜索  $K^*$  的具体作法如下:

(1) 对于初始搜索区间  $[K_a, K_b]$ , 搜索区间长度  $L = K_b - K_a$ 。在搜索区间内距两个端点各  $0.382 L$  处对称插入两个内插点  $K_c$  和  $K_d$ , 下内插点  $K_c = K_a + 0.382 L$ , 上内插点  $K_d = K_b - 0.382 L = K_a + 0.618 L$ 。分别将  $K_c$  和  $K_d$  代入式 (6), 选取前  $M^*$  个观测数据进行回归分析估计  $r$ , 然后计算对应的  $SS_c$ 。

(2) 比较两内插点对应的  $SS_c$ , 若下内插点  $K_c$  的  $SS_c$  较大, 则舍弃下端点  $K_a$ , 将  $K_c$  作为新区间的端点, 并在上内插点  $K_d$  与上端点  $K_b$  之间插入一个新的上内插点; 反之, 若上内插点  $K_d$  的  $SS_c$  较大, 则舍弃上端点  $K_b$ , 将  $K_d$  作为新区间端点, 并在  $K_c$  与  $K_a$  之间插入新的下内插点。仍按照 (i) 中的方法计算新内插点的  $K$  值 (此时  $L$  应为缩小后的搜索区间长度)。

(3) 多次重复步骤 (ii), 通过不断舍弃端点来逐渐缩小搜索范围 (第  $n$  次搜索后, 搜索长度将缩小为初始长度  $L$  的  $0.618^n$  倍), 同时不断插入新内插点以加大对重点区域的搜索密度。当  $SS_c$  下降十分缓慢或达到实验要求的精度时, 即可将此时的  $K$  值作为“最优”估计值  $K^*$ 。

### 1.4 方法小结

综上所述, “最佳”观测点数定向搜索法的主要分析步骤如下:

(1) 完整实验数据的拟合。

① 实验观测数据的预处理。若原始实验数据的初始时间  $t_0 \neq 0$ , 则以  $t' = t - t_0$  代替  $t$ 。

② 选取围绕平衡数量波动的观测值计算其平均值  $K_m$ , 以  $K_m$  作为中值  $K_{mid}$ , 确定初始搜索区间  $[K_a, K_b]$ 。

③ 利用“三线定点法”确定一个在整个  $K$  值搜索过程中进行线性回归时均可采用的“最佳”观测点数  $M^*$ 。

④ 利用黄金分割法定向搜索参数  $K$  的“最优”估计值  $K^*$  和相应的  $r$  值。

⑤ 此外, 搜索到  $K^*$  值之后, 还可对其进一步调整回归分析采用的观测点数目, 以提高参数  $r$  的估计精度、获得更好的拟合效果。

⑥ 将  $K^*$  和最终获得的  $r$  估计值  $r^*$  代入式 (6) 或式 (1) (积分常数  $a$  据  $K^*$  和实测初值  $N_0$  由式 (5) 解出), 即可获得 Logistic 增长模型。

(2) 不完整实验数据的拟合。

若实验数据不完整, 则需首先预估一个  $K$  值搜索区间 (具体方法见 1.1.2 步骤 (i)), 作该区间端点和中值对应的  $Y-t$  图, 确定  $m_s$ ; 然后分别将端点和中值代入式 (6) 估计  $r$  (因两个端点的  $K$  值往往相差较大, 一般直接由前  $m_s$  个观测数据进行回归分析即可), 计算 3 个  $K$  值对应的  $SS_c$ 。若一侧端点的  $SS_c$  最低, 则以该端点作为新区间的中值, 将搜



索区间向该侧平移。重复上述调整过程,直至中值的 $SS_e$ 最低,最终确定初始搜索区间 $[K_a, K_b]$ 。

其他步骤与上述完整实验数据的分析相同。

## 2 实例分析及结果

利用“最佳”观测点数定向搜索法对7个实例进行了分析。其中,Gause(1934)草履虫种群增长实验<sup>[9]</sup>是种群生态学中最为经典和重要的实验。

### 2.1 完整实验数据的分析过程

仍以Gause(1934)双小核草履虫种群增长实验(添加一环细菌)<sup>[9]</sup>为例。由前面分析可知,本例初始搜索区间 $[418, 478]$ ,  $K_{mid} = 448$ ,  $m_s = 7$ 。取 $K_{mid} = 448$ ,分别以前6、7、8个观测数据估计 $r$ 。3个 $r$ 值对应的 $SS_e$ 分别为17436、14504和14970。 $m_s = 7$ 对应的 $SS_e$ 最小,因此确定 $M^* = 7$ 。以前7个观测数据进行回归分析估计 $r$ ,经8次搜索(表1),得 $K^* = 438$ ,  $r^* = 1.1532$ 。

表1 Gause(1934)双小核草履虫实验数据(添加一环细菌)<sup>[9]</sup>K值搜索过程

搜索次数	搜索点和搜索范围				两个内插点对应的 $SS_e$		应舍弃的端点
	下端点	下内插点	上内插点	上端点	下内插点	上内插点	
1	418	441	455	478	13697	16225	478
2	418	432	441	455	14002	13697	418
3	432	441	446	455	13697	14181	455
4	432	437	441	446	13645	13697	446
5	432	435	437	441	13732	13645	432
6	435	437	439	441	13645	13633	435
7	437	439	440	441	13633	13655	441
8	437	438	439	440	13630	13633	

### 2.2 不完整实验数据初始搜索区间的确定

以王振中和林孔勋花生锈病显症动态试验(I)前19天的实验数据<sup>[12]</sup>为例。该例观测时间为第9~19天,相应的观测值(孢子堆数目)依次为79、126、257、625、1856、4000、7389、12133、16057、18898和22361。首先将原观测时间转换为 $t=0\sim10$ ;然后分别选取 $t=(0,5,10)$ 、 $(1,5,9)$ 和 $(2,5,8)$ 对应的观测值,按照三点法<sup>[1]</sup>计算 $K$ 值,得 $K=21564$ 、17843和19204。将3个 $K$ 值的均值 $K_m=19537$ 作为 $K'_{mid}$ ,分别向上、下各延伸4000预估初始搜索区间,得 $[15537, 23537]$ 。分别作 $K=15537$ 、19537和23537对应的 $Y-t$ 图,3条折线的明显“Y”形分叉点对应于 $m_s=7$ 。选取前7个观测数据进行回归分析估计3个 $K$ 值对应的 $r$ 值,计算 $SS_e$ ,分别为107790143、33024912和4105099。由于上端点 $K=23537$ 对应的 $SS_e$ 最低,因此将其作为新区间的中值,将初始搜索区间平移至 $[19537, 27537]$ 。对新区间进行类似分析,得 $m_s=8$ ;选取前8个观测数据估计 $r$ ,以中值对应的 $SS_e$ 最小,因此确定初始搜索区间为 $[19537, 27537]$ 。

### 2.3 个实例的分析结果

表2给出了7个实例的分析结果,相应的 $Y-t$ 折线图分别见图1~图7。获得“最优”估计值 $K^*$ 后,均对其进行进一步调整 $m$ 值,以期提高参数 $r$ 的估计精度,但最终均以 $m=M^*$ 对应的 $SS_e$ 最低。

表2

“最佳”观测点数定向搜索法应用实例分析

实验数据来源	初始搜索区间 $[K_a, K_b]$	$K_{mid}$	$m_s$	$M^*$	搜索次数	$K^*$	$r^*$
双小核草履虫实验(一环细菌) <sup>[9]</sup>	[418, 478]	448	7	7	8	438	1.1532
双小核草履虫实验(半环细菌) <sup>[9]</sup>	[232, 264]	248	5	5	5	243	1.3890
大草履虫(P. caudatum)实验(一环细菌) <sup>[9]</sup>	[122, 150]	136	5	5	6	130	1.1933
大草履虫实验(半环细菌) <sup>[9]</sup>	[42, 74]	58	4	4	6	59	0.9796
酵母种群实验 <sup>[11]</sup>	[650, 670]	660	13	12	5	664	0.5405
稻瘿蚊黄柄黑蜂实验 <sup>[13]</sup>	[70, 90]	80	5	6	11	77.8	0.5090
花生锈病显症动态试验(前19天数据) <sup>[12]</sup>	[19537, 27537]	23537	8	9	18	23994	0.8014

周赛花<sup>[13]</sup>误将邬祥光<sup>[4]</sup>第17天的实测值23.50看作28.50,因此拟合结果存在一定偏差。但与周赛花提出的逐次加密

搜索法进行比较,本文仍按其给出的数据进行拟合。

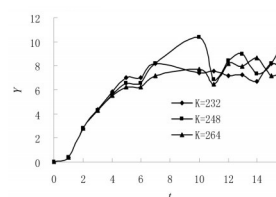


图2 Gause(1934)双小核草履虫(添加半环细菌)<sup>[17]</sup>Y-t图

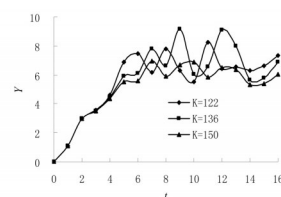


图3 Gause(1934)大草履虫(添加一环细菌)<sup>[17]</sup>Y-t图

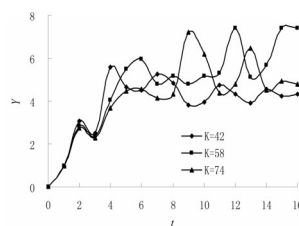


图4 Gause(1934)大草履虫(添加半环细菌)<sup>[17]</sup>Y-t图

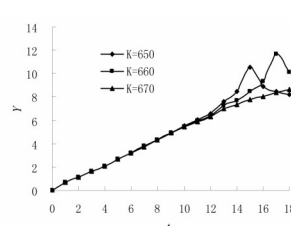


图5 Carlson(1913)酵母种群<sup>[19]</sup>Y-t图

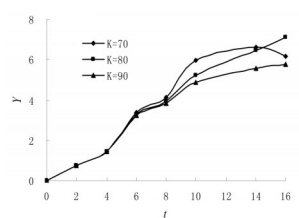


图6 稻瘿蚊黄柄黑蜂<sup>[8]</sup>Y-t图

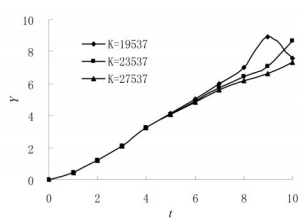


图7 花生锈病显症动态试验 I (前19天数据)<sup>[12]</sup>Y-t图

## 3 与其他作者拟合结果的比较

将上述7个实例的拟合结果与其他作者的拟合结果以及笔者调用SPSS非线性回归的分析结果进行比较,得表3~表7。其中,SPSS非线性回归以式(4)作为待拟合模型,以实验观测值中的最大值作为参数 $K$ 的初始值,以 $t=1$ 时的观测值据式(6)估计参数 $r$ 的初始值:

$$r = \ln\left(\frac{K - N_0}{N_0}\right) - \ln\left|\frac{K - N_1}{N_1}\right|$$

其他作者的理论初值( $N_0$ )和残差平方和( $SS_e$ )数据为笔者依据各作者的参数估计值计算得到。

表3 Gause(1934)双小核草履虫种群增长实验数据<sup>[17]</sup>拟合结果的比较

作者	拟合方法	每日添加一环细菌					每日添加半环细菌				
		K	r	a	N <sub>0</sub>	SS <sub>e</sub>	K	r	a	N <sub>0</sub>	SS <sub>e</sub>
Andrewartha (1954) <sup>[1]</sup>	平均值法	450	1.022	5.041	2.89	13132	240	1.460	5.101	1.45	2811
万昌秀等(1981) <sup>[3]</sup>	枚举选优法	443	1.0451	5.0683	2.77	12489	241	1.5284	5.2021	1.32	2692
王振中等(1987) <sup>[12]</sup>	四点式平均值法	452	0.9576	4.9098	3.31	14926	248	1.1978	4.7058	2.22	4974
王莽莽等(1986) <sup>[4]</sup>	麦夸法	448	0.8048	3.8171	9.64	8640	244	1.2378	4.2512	3.43	2367
马占山(1991) <sup>[5]</sup>	单纯形加速法	448	0.8040	3.8143	9.67	8641	245	1.2345	4.2407	3.47	2367
唐启义等(1996) <sup>[8]</sup>	麦夸法(修正后)	438.75	1.1446	5.3862	2.00	13598	243.3	1.3976	4.7929	2.00	2477
吴新元(1990) <sup>[6]</sup>	一种数值方法	449	0.8611	5.4094	2.00	71377	246	1.1854	4.8040	2.00	6631
笔者(方法1)	SPSS非线性回归分析	438.74	1.1447	5.3862	2.00	13598	243.29	1.3979	4.7929	2.00	2477
笔者(方法2)	“最佳”观测点数定向搜索法	438	1.1532	5.3845	2.00	13630	243	1.3890	4.7916	2.00	2485

\* 实测初值 N<sub>0</sub> = 2。表4 Gause(1934)大草履虫种群增长实验数据<sup>[17]</sup>拟合结果的比较

作者	每日添加一环细菌					每日添加半环细菌				
	K	r	a	N <sub>0</sub>	SS <sub>e</sub>	K	r	a	N <sub>0</sub>	SS <sub>e</sub>
Andrewartha (1954) <sup>[1]</sup>	140	1.015	3.894	2.79	3117	65	0.855	3.185	2.58	1233
万昌秀等(1981) <sup>[3]</sup>	130	1.1414	4.0101	2.32	2245	59	0.9379	3.2086	2.29	971
王振中等(1987) <sup>[12]</sup>	131	1.0939	3.9223	2.54	2243	69	0.5706	2.6858	4.40	1697
王莽莽等(1986) <sup>[4]</sup>	131	1.05797	3.713	3.12	2222	60	0.821	2.819	3.38	956
马占山(1991) <sup>[5]</sup>	131	1.0554	3.6977	3.17	2222	60	0.8200	2.8597	3.25	957
唐启义等(1996) <sup>[8]</sup>	130.46	1.1766	4.1779	1.97	2266	59.46	0.9674	3.3579	2.00	976
吴新元(1990) <sup>[6]</sup>	134	1.0756	4.1897	2.00	2653	62	0.8393	3.4012	2.00	1129
笔者(方法1)	130.46	1.1767	4.1625	2.00	2263	59.45	0.9677	3.3578	2.00	976
笔者(方法2)	130	1.1933	4.1589	2.00	2271	59	0.9796	3.3499	2.00	978

\* 各作者拟合方法及实测初值大小同表3。

表5 Carlson(1913)酵母种群增长实验数据<sup>[18]</sup>拟合结果的比较

作者	拟合方法	K	r	a	N <sub>0</sub>	SS <sub>e</sub>
Pearl(1927) <sup>[11]</sup>	三点法	665	0.5355	4.1896	9.93	246
王莽莽等(1986) <sup>[4]</sup>	麦夸法	662	0.5475	4.2758	9.08	204
笔者	SPSS非线性回归分析	664	0.5406	4.2220	9.60	211
笔者	“最佳”观测点数定向搜索法	664	0.5405	4.2220	9.60	212

\* 实测初值 N<sub>0</sub> = 9.6。表6 稻螟蚊黄柄黑蜂种群增长数据<sup>[8]</sup>拟合结果的比较

作者	拟合方法	K	r	a	N <sub>0</sub>	SS <sub>e</sub>
周赛花(1992) <sup>[13]</sup>	逐次加密搜索法	77.4	0.5566	10.2853	1.19	46
笔者	SPSS非线性回归分析	77.7	0.5169	3.8494	1.62	49
笔者	“最佳”观测点数定向搜索法	77.8	0.5090	3.8507	1.62	51

\* 将实测时间转换为 t<sub>0</sub>=0 后, 实测初值 N<sub>0</sub> = 1.62。表8 花生锈病显症动态试验( I )前 19 天实验数据<sup>[9]</sup>拟合结果的比较

作者	拟合方法	K	r	a	N <sub>0</sub>	SS <sub>e</sub>
王振中等(1987) <sup>[12]</sup>	四点式平均值法	25137	0.8100	13.2194	66.69	4153444
笔者	SPSS非线性回归分析	23598	0.8128	5.6961	79.00	1423560
笔者	“最佳”观测点数定向搜索法	23994	0.8014	5.7128	79.00	1781790

\* 将实测时间转换为 t<sub>0</sub>=0 后, 实测初值 N<sub>0</sub> = 79。

由上可见, 与其他拟合方法比较, 本文提出的“最佳”观测点数定向搜索法具有如下优点:

(1) 不仅可以满足 Logistic 增长模型的初值条件, 保证

理论初值与实测初值相等, 使参数 K 和 r 的估计更为客观准确, 而且能够获得理想的拟合效果, 拟合优度与唐启义等<sup>[16]</sup>以及 SPSS 非线性回归模块(依据式(4))采用麦夸法的结果十分接近。

(2) 每一步的分析结果都比较明确, 最终拟合结果的不确定性小。

(3) 可用于完整或不完整、整数型或非整数型实验数据的拟合, 适用范围较之前的很多线性化方法明显拓广。

(4) 原理简单, 计算量小, 仅靠 Excel 即可轻松完成全部计算, 很适于不熟悉统计软件包的研究者利用。

#### 参考文献:

- [1] Andrewartha H G, Birch L G. The Distribution and Abundance of Animals[M]. Chicago: University of Chicago Press, 1954.
- [2] 考克斯. 普通生态学实验手册[M]. 蒋有绪译. 北京: 科学出版社, 1979.
- [3] 万昌秀, 梁中宇. 逻辑斯谛曲线的一种拟合方法[J]. 生态学报, 1983, 3(3).
- [4] 王莽莽, 李典谟. 用麦夸方法最优拟合逻辑斯谛曲线[J]. 生态学报, 1986, 6(2).
- [5] 马占山. 单纯形加速法拟合生态学中的非线性模型[J]. 生物数学学报, 1992, 7(2).
- [6] 吴新元. 逻辑斯谛曲线拟合的一种数值方法[J]. 生物数学学报, 1990, 5(1).
- [7] 陆纹岱. SPSS for Windows 统计分析[M]. (第3版). 北京: 电子工业出版社, 2006.
- [8] 唐启义, 胡国文, 冯明光, 胡阳. Logistic 方程参数估计中的错误与修正[J]. 生物数学学报, 1996, 11(4).
- [9] Gause G F. The struggle for existence[M]. Baltimore: Williams and Wilkins, 1934.
- [10] 瓦格纳. 运筹学原理与应用(第2版)[M]. 邓三瑞等译. 北京: 国防工业出版社, 1992.
- [11] Pearl R. The Growth of Populations[J]. Quarterly Review of Biology, 1927.
- [12] 王振中, 林孔勋. 逻辑斯谛曲线 K 值的四点式平均值估计法[J]. 生态学报, 1987, 7(3).
- [13] 周赛花. 逻辑斯谛方程中参数的估计[J]. 数理统计与管理, 1992, 11(5).
- [14] 祁祥光. 昆虫生态学的常用数学分析方法(修订版)[M]. 北京: 农业出版社, 1985.

(责任编辑/亦 民)