

离散分支过程模型应用及其随机模拟验证

518030910319 张嘉乐

摘要

为探究生物种群数量演化的规律,本文用离散分支过程对此进行建模,首先以马尔科夫链理论分析其短期表现。为了研究生物种群一段时间后消亡的概率,在第二部分用概率母函数方法证明了灭绝定理,由此计算出种群初始容量和后代数量分布确定时种群灭绝的概率和相应参数。

为了验证理论的正确性,本文引入了一个简单实例进行理论计算并进行大量轮次的随机模拟,最终实验结果和理论计算结果高度符合。

第三部分以此模型对现今我国人口演化作出分析。

1 离散分支过程

1.1 问题的提出

建国以来我国的人口数量发生了很大变化,为了适应时代需要,人口政策也相应有过多次调整。一个很重要的问题是采取不同的人口政策到底会对人口变化起到多大的影响。那么为此我们需要建立模型分析人口增长的规律。

由于人类社会现象的复杂性,为了简化模型,我们将问题抽象:

某一种生物每代产生不定数量的后代,建模分析给定初始条件下种群个体数的变化规律。

1.1 离散分支过程模型的建立

• 模型假设

- 1. 种群数量在时间上变化是离散的,用 $\{X_n, n \geq 0\}$ 表示种群第 n 代的个体总数

假设说明:实际问题中生物由于繁殖年龄不完全相同,不可能严格的划分代数,但这一假设使得问题可以用离散参数链(函数值显然是整数)来建模,使后续的分析能够进行,同时对于人类来说两代人的年龄差范围变化不大。

- 2. 新一代个体数为原先一代个体的后代总数,即计算 X_{n+1} 时仅考虑 X_n 的后代而不考虑其本身
假设说明:如果将这一假设理解为新一代个体产生时原先一代个体全部消亡,那么这看似仅对低等生物成立,对于高等生物如人类不太合理。但要注意 $\{X_n, n \geq 0\}$ 并非 $t = n$ 时存活的所有个体,而是 $t = n$ 时种群中处于某一年龄段的个体数(比如我们日常所说的70后,00后指的是出生于这一段时间的所有人一样)

由于生育年龄的有限性(如人大约在20至40岁), X_{n+1} 几乎就是 X_n 产生的后代数,与 X_{n-1} 无关,这也就是为何在 X_n 中不能加上 X_{n-1} 的原因,否则一部分下一代的个体是上两代

个体产生的，这对于人口模型是不正确的。这一步假设实际说明了种群数量这一离散参数链具有马尔可夫性，为一马尔可夫链。

- 3. 每一代的每一个个体的后代数满足相同的分布律

假设说明：这一假设说明我们的模型仅针对局部地区的局部时间，忽略了不同地域后代数量分布的差异和人口政策的调整，当需要考虑这些因素时只需拼接两个随机过程即可。这一假设实际上说明了 X_n 为为一时间齐次马尔可夫链

• 模型建立

设 $\{X_n, n \geq 0\}$ 表示该生物种群第 n 代的个体总数， $\xi_{n,i}$ 表示第 n 代第 i 个个体产生的后代。假设每一代所有个体繁殖年龄相同，下一代数量不包括上一代的个体；且 $\xi_{n,i}$ 满足相同的分布。有转移方程

$$X_{n+1} = \sum_{i=1}^{X_n} \xi_{n,i}$$

则 $\{X_n, n \geq 0\}$ 称为离散分支过程。下面的定理给出了当 $X_0 = 1$ 时种群灭绝的概率 π_1

1.2 灭绝定理

• 定理

$X_0 = 1$ 时，离散分支过程的灭绝概率 π_1 是方程 $s = A_1(s)$ 的最小正根。

其中 $A_n(s)$ 是 X_n 的概率母函数， $A_n(s) = E(s^{X_n})$ 。 $A_1(s)$ 即为个体后代数的概率母函数 $E(s^{\xi_{n,i}})$

• 证明：

- 引理： $A_{n+1}(s) = A_1(A_n(s))$
- 引理证明：根据概率母函数的定义有

$$\begin{aligned} A_{n+1}(s) &= E(X_{n+1}) = \sum_{j=1}^{\infty} E(s^{X_{n+1}} | X_n = j) P\{X_n = j\} = \sum_{j=1}^{\infty} E(s^{\sum_{i=1}^j \xi_{n,i}}) P\{X_n = j\} \\ &= \sum_{j=1}^{\infty} E(s^{\xi_{n,i}})^j P\{X_n = j\} = \sum_{j=1}^{\infty} A_1(s)^j P\{X_n = j\} = A_n(A_1(s)) \\ &= A_{n-1}(A_1(A_1(s))) = A_{n-1}(A_2(s)) = A_{n-2}(A_3(s)) = A_{n-3}(A_4(s)) = \cdots \\ &= A_1(A_n(s)) \end{aligned}$$

- 定理证明：

①令 $p_{i,j} = P(X_i = j)$ ，则 $\pi_1 = \lim_{n \rightarrow \infty} p_{n,0}$ 。当 $n \rightarrow \infty$ 时， $p_{n,0} = p_{n+1,0}$ 。

注意到 $p_{n,0} = A_n(0)$ ，代入引理即得 $A_n(0) = A_{n+1}(0) = A_1(A_n(0))$ 。

所以灭绝概率 π_1 为方程 $s = A_1(s)$ 的根。

②由概率分布的归一性有 $A_1(1) = 1$ ，所以方程 $s = A_1(s)$ 存在正根 s 。

由后代数和概率的非负性， $A_n(x)$ 在 $(0, +\infty)$ 上单调递增，有 $p_{1,0} = A_1(0) < A_1(s) = s$

假设 $n \leq k$ 时， $p_{n,0} < s$ ，则 $p_{k+1,0} = A_{k+1}(0) = A_1(A_k(0)) = A_1(p_{k,0}) < A_1(s) = s$

由数学归纳法得 $p_{n,0} < s$ 对 $\forall n \in N^*$ 成立, 所以 $\pi_1 = \lim_{n \rightarrow \infty} p_{n,0} \leq s$.

所以 π_1 为 $s = A_1(s)$ 的最小正根

• 补充说明

若 $\mu = E(\xi_{i,j})$ 为平均后代数, 若 $\mu \leq 1$, 则 $\pi_1 = 1$ 。这是容易理解的, 由于期望后代数小于当前种群个体数, 足够长时间后一定会灭绝。严格证明 $s = 1$ 为 $s = A_1(s)$ 最小正根过程在此略去。

1.3 举例

- 模型: 设 $\{X_n, n \geq 0\}$ 为离散分支过程, $X_0 = k$ 。 $\xi_{i,j}$ 分布为

$$p_0 = \frac{1}{4}, p_1 = \frac{1}{2}, p_2 = \frac{3}{16}, p_3 = \frac{1}{16}$$

- 解: $A_1(s) = \frac{s^3}{16} + \frac{3s^2}{16} + \frac{s}{16} + \frac{1}{4}$, $A_1(s) = s$ 最小正根 $s = 2\sqrt{2} - 2$
所以 $k = 1$ 时灭绝概率 $\pi_1 = 2\sqrt{2} - 2 \approx 0.828$, $\pi_k = \pi_1^k$

k	1	2	3	4	5	10	15	20	30	40
π_k	0.828	0.686	0.568	0.470	0.389	0.151	0.059	0.023	0.003	0.001

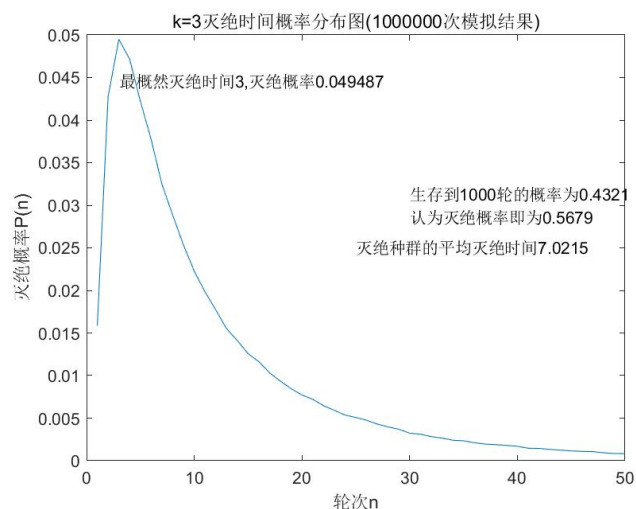
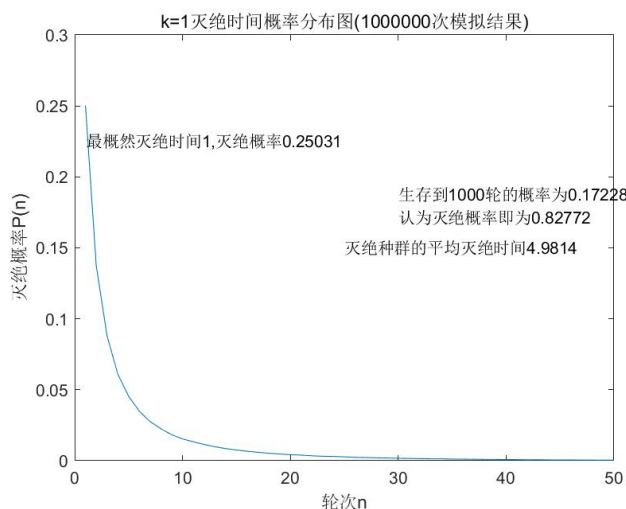
• 随机模拟实验

◦ 实验方法

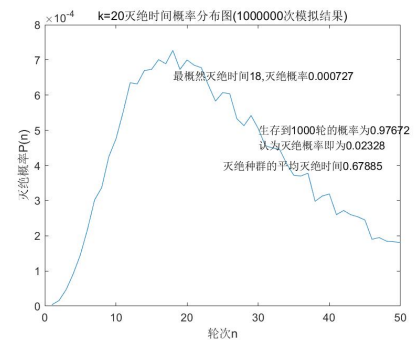
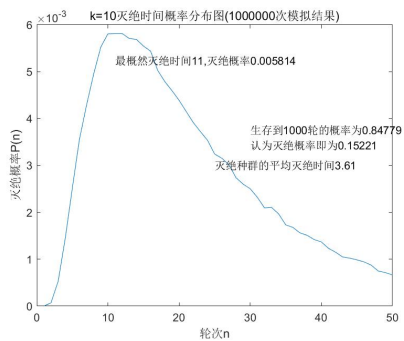
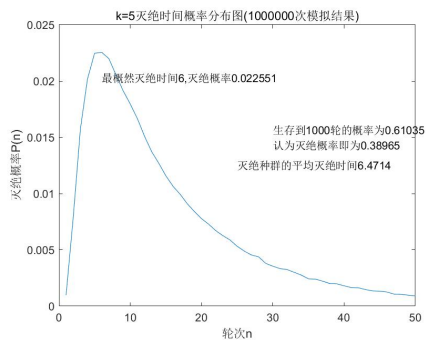
用Matlab进行1000000次随机模拟获得每次模拟灭绝时间, 以频率代替概率来表现灭绝时间的概率分布, 并获得最终不会灭绝的概率。

在模拟中通过生成 $[0, 1]$ 上均匀分布的随机数, 取其不同范围来模拟 $\xi_{n,i}$ 的概率分布。对于时间 $n = +\infty$ 的模拟困难, 为了控制运算时间, 当种群容量 $X_n > 1000$ 或生存轮数 $n > 1000$ 时即认为不会灭绝, 由此作出灭绝时间概率分布图。由理论计算和模拟结果都可知 $X_n > 100$ 或 $n > 1000$ 时灭绝概率在0.01%以下, 该近似不影响实验精度。

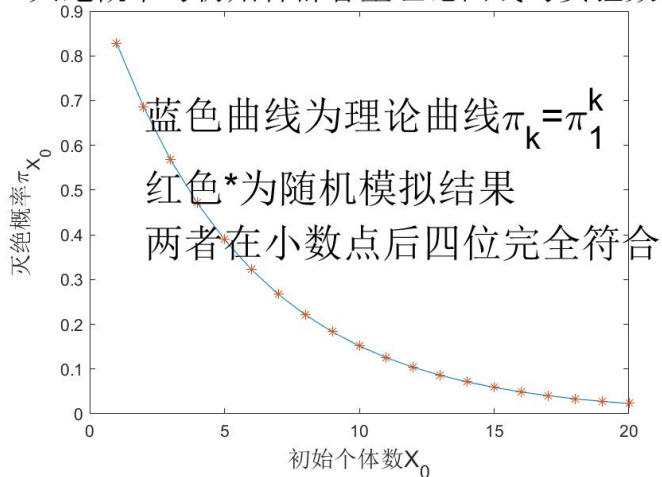
◦ 实验结果



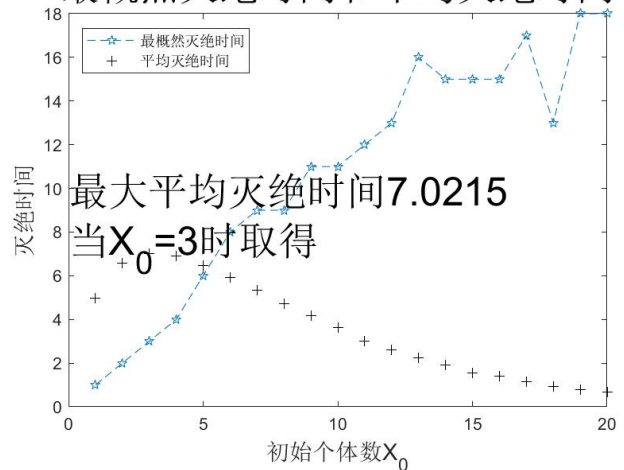
1.4 平均灭绝时间的理论计算



灭绝概率与初始种群容量理论曲线与实验数据



最概然灭绝时间和平均灭绝时间



在前述样例中我们通过模拟实验不仅以很高的精度验证了灭绝定理，还得到了不同初始条件下灭绝时间的概率分布情况，并且从中抽取了最概然灭绝时间和平均灭绝时间这两个特征，下面我们希望从马尔科夫链理论出发作一些推导以得到这些结论

2 排队系统

2.1 排队系统的描述

2.2 M/M/1 排队系统的建模与求解

2.3