复杂网络和人类行为动力学演化模型

郭进利 著

科学出版社

复杂网络和人类行为动力学演化模型

郭进利　著

科学出版社
北京

内 容 简 介

本书致力于系统地介绍复杂网络和人类行为动力学演化模型的理论分析方法，涉及的模型几乎均给出解析过程和模拟，大部分模型给出实证分析或实际背景. 本书力图取材广泛新颖、理论严谨有深度、模型与实证相结合.

复杂网络部分在模型中引入随机过程，介绍多种分析和实证模型的方法；以供应链为背景，建立连续时间动态网络模型；将加权网络统一在竞争网络下处理，体现出竞争网络的普适性；基于随机服务系统建立节点退出机制的网络模型；提出变速增长网络模型，讨论加速增长网络；论述幂律分布和几何级数增长网络度分布的计算，给出计算确定性网络分布的方法. 人类行为动力学部分介绍基于排队系统的人类行为动力学模型，利用非齐次 Poisson 过程研究人类行为动力学，对供应链采购行为动力学特性进行分析.

本书几乎是自封闭的，尽量使读者阅读时不用查找其他参考文献. 本书可作为理工科、管理类硕士研究生和博士研究生的教材，也可供自然科学和工程技术领域的研究人员参考阅读.

图书在版编目(CIP)数据

复杂网络和人类行为动力学演化模型/郭进利著. —北京：科学出版社，2013.11

ISBN 978-7-03-038907-7

Ⅰ. ①复… Ⅱ. ①郭… Ⅲ. ①复杂性－网络系统－动力学模型－研究 ②人类－行为－动力学模型－研究 Ⅳ. ①TN94 ②C912.68

中国版本图书馆 CIP 数据核字（2013）第 253765 号

责任编辑：赵彦超 李静科／责任校对：赵桂芬
责任印制：赵德静／封面设计：陈 敬

科学出版社 出版
北京东黄城根北街 16 号
邮政编码：100717
http://www.sciencep.com

文林印刷厂 印刷

科学出版社发行 各地新华书店经销

*

2013 年 11 月第 一 版 开本：720×1000 1/16
2013 年 11 月第一次印刷 印张：15 1/4
字数：292 000

定价：68.00 元

（如有印装质量问题，我社负责调换）

前　　言

十年前, 在新疆飞上海的途中, 我看到脚下大片沙荒戈壁和天空翻滚的乌云, 想象有一天能实现南云北调, 大西北就成了绿洲. 科学技术的飞速发展, 将不可能变为可能. 今天的南云北调幻想明天可能变为设想, 明天的设想后天可能变为现实. 我们很荣幸处在科技加速发展时代. 如果三国时期诸葛亮发明的不是木牛流马, 是飞机大炮或手机电话, 那孟德和仲谋可能就不战而降. 几个世纪前, 人们要收集人类行为的数据是不可想象的, 但今天已成为可能. 随着互联网、通信、IT 等科技的发展, 任何人和物都打上了电子烙印, 产生了海量的数据. 用句时髦的话来说, 我们已经进入了大数据时代. 近年来, 虽然公安机关利用监控录像和通信数据推测犯罪分子的行踪, 但没有引起复杂科学研究者的足够重视. 如果一般人说人类行为大部分是可以预测的, 估计没有人相信, 然而, Barabási 的人类行为 93%可以预测的推断引起了复杂科学研究者的高度重视. 这是因为 Barabási 有名人效应优势、基金资助优势、团队人才优势、数据获取优势. 这就是“马太效应”, Price 称之为“累积优势”, 也是 Barabási 等说的 “择优选择”. 增长和择优选择是形成无标度网络的重要机制之一. 20 世纪 90 年代末, Watts 和 Strogatz 的小世界网络模型、Barabási 和 Albert 等的无标度网络模型的提出, 开辟了复杂网络研究的新纪元.

在复杂网络和人类行为动力学研究领域, 根据实际背景建立模型和模拟指标并不难, 但解析模型比较困难. 作者从 2011 年下半年开始动笔, 在写作过程中发现, 想准确反映复杂网络和人类行为动力学演化模型, 其难度远远超出了作者的想象. 其原因有三：首先, 复杂网络的研究文献浩瀚; 其次, 要从繁多的文献中选出既有实际背景也有比较严谨解析过程的模型并不是一件容易的事情; 最后, 这个领域的交叉性太强, 要写一本自封闭的教材很困难. 到目前为止, 作者还没有看到一本在数学上有严格论证的复杂网络和人类行为动力学理论专著. 我们希望本书内容不仅能使复杂科学方面的学者感兴趣, 而且能使数学专业背景的学者接受. 复杂网络的发展必然涉及数学理论的研究, 它也必将成为数学的一个研究领域. 自从无标度网络的提出, 许多复杂网络中的理论问题就饱受争议, 特别是, 在权威期刊上对阿波罗网络度分布仍没有令人信服的结论, 复杂网络领域的严谨性是亟待解决的问题. 这将为你提供机会, 也许下一个对此学科有重要贡献的就是你.

本书包含 12 章内容. 第 1 章是预备知识, 介绍概率论、Poisson 过程、非齐次 Poisson 过程、更新过程和随机服务系统的基础知识, 并给出大部分定理的详细证明. 第 2 章以图、随机图和复杂网络为主线, 拓展随机图模型, 将小世界网络的

WS(Watts-Strogatz) 模型纳入随机图的框架, 介绍评价复杂网络的指标体系和复杂网络数学定义. 第 3 章介绍 Barabási –Albert 模型及其三种常用的拓扑结构分析方法, 给出计算集聚系数与平均路径的计算方法. 第 4 章分析 Barabási –Albert 方法的缺陷, 在网络模型中引入随机过程, 给出分析模型的 Poisson 过程方法、差分方程方法和近似估计方法. 第 5 章将加权网络统一在竞争网络模型下处理, 体现出竞争网络的普适性. 第 6 章以供应链为背景, 建立连续时间动态有向网络模型. 第 7 章介绍内部边加速增长模型、边随机增长模型, 讨论新节点的边对网络无标度性影响、无标度网络的幂律加速不变性和加速增长网络的量变到质变问题. 第 8 章介绍节点老化模型, 基于随机服务系统对节点具有寿命的网络分类, 建立节点退出机制的网络模型, 给出基于齐次 Markov 链的度分布计算方法. 第 9 章介绍几类确定性的无标度网络, 给出几何级数增长随机变量幂律分布的指数与累积分布的指数的关系, 它是用累积分布计算几何增长无标度网络的度分布和度指数的依据, 并对阿波罗网络度分布进行论证. 第 10 章介绍 Barabási 人类行为动力学模型, 即基于排队系统的人类行为动力学. 第 11 章基于非齐次 Poisson 过程的人类行为动力学, 在博客实证数据的基础上, 研究兴趣衰减的人类行为动力学和兴趣衰稳的人类行为动力学模型. 第 12 章分析供应链中采购行为动力学特性, 总结复杂网络和人类行为动力学中的实证统计方法, 它们是计数数据频数法、连续型数据频数法和对数装箱方法, 特别是详细介绍对数装箱方法; 基于更新理论分析供应链中采购行为动力学模型, 试图为人类行为动力学的研究提供新思路, 为供应链中采购订单预测提供参考.

本书从开始撰写到形成讲义, 经过上海理工大学 2011 级和 2012 级研究生试用, 并几经修改终于与读者见面. 作者借此机会, 衷心感谢本领域的合作者周涛、韩筱璞、胡海波、杨涵新等教授和博士. 感谢上海理工大学徐福缘教授、王恒山教授、车宏安教授和钱省三教授等, 他们给了作者许多支持和鼓励. 虽然本书主要总结作者的研究工作, 但也引用了国内外同行的许多重要成果, 在此也向他们表示感谢. 本书得到了国家自然科学基金项目 (项目编号: 70871082) 和上海市一流学科项目 (项目编号: S1201YLXK) 资助, 作者深表谢意. 最后, 感谢赵彦超、李静科及对全书做了辛苦编辑工作的科学出版社的同志.

由于复杂网络和人类行为动力学是交叉学科, 理论体系并不成熟, 限于作者水平, 书中难免有不妥之处, 敬请读者不吝批评指正.

郭进利

2013 年 5 月于上海

目　　录

第 1 章 预 备 知 识

1.1 随机变量与随机过程简介

概率论中的随机试验是指: 其结果在事先不能确定的试验. 所有可能的试验结果的集合称为试验的样本空间, 记为 S. 一个事件是样本空间的子集, 如果试验的结果是这个子集中的元素, 则称该事件发生. 假设对样本空间 S 的每个事件 E 定义一个实数 $P(E)$, 满足如下三条公理:

公理 (1) $0 \leqslant P(E) \leqslant 1$.

公理 (2) $P(S) = 1$.

公理 (3) 对于任何一列互不相交的事件 $E_1, E_2, \cdots$, 有 $P\left(\bigcup\limits_{i=1}^{\infty} E_i\right) = \sum\limits_{i=1}^{\infty} P(E)$.

我们称 $P(E)$ 为事件 E 的概率.

频率是指事件发生的频繁度, 在相同条件下, 进行 n 次试验, 事件 E 发生的次数 n_E 称为事件 E 发生的频数, 频数与试验次数的比值 n_E/n 称为事件 E 发生的频率. 概率的统计定义可以表示为: 当不断重复同一试验时, 事件 E 发生的频率的极限称为事件 E 的概率.

随机变量 X 是一个函数, 它给 S 中的每个结果指定一个实数. 对于任何实数集 A, X 取的值含于集 A 中的概率等于试验结果包含在 $X^{-1}(A)$ 中的概率, 即

$$P\{X \in A\} = P(X^{-1}(A)),$$

其中, $X^{-1}(A)$ 是由一切满足 $X(\omega) \in A$ 的点 $\omega \in S$ 构成的事件.

随机变量 X 的 (累积) 分布函数 F 由下式定义: 对于任意一实数 x, 有

$$F(x) = P\{X \leqslant x\} = P\{X \in (-\infty, x]\},$$

$\overline{F}(x) = 1 - F(x)$ 称为 (累积) 分布函数 F 的补, 则

$$\overline{F}(x) = P\{X > x\}.$$

在复杂网络研究中也称 $\overline{F}(x)$ 为累积 (计) 分布, 因此, 本书中不区分累积分布和 (累积) 分布函数的补 (complementary cumulative distribution function, CCDF 或 ccdf).

一个随机变量 X 称为离散的, 如果它可能取的值集合是有限个或可数值 x_1, $x_2, \cdots, x_n, \cdots$. 若记 $P(X = x_k) = p_k(k = 1, 2, \cdots)$, 则 X 取值的概率分布由 $\{p_k\}$ 完全确定. 称 $\{p_k\}$ 为 X 的概率分布列 (复杂网络中称为概率分布). $\{p_k\}$ 有以下性质: (1) $p_k \geqslant 0$;(2) $\sum_k p_k = 1$.

离散随机变量 X 的分布函数为

$$F(x) = \sum_{y \leqslant x} P\{X = y\}.$$

这是一个阶梯函数. 阶梯函数永远也不可能是幂函数, 而在复杂网络研究中主要考察离散随机变量分布函数是否是幂函数, 因此, 下面涉及的离散随机变量 X 的累积分布函数的定义域均指与 X 的值域相同.

一个随机变量 X 称为连续型随机变量, 如果存在函数 $f(x)$, 使得

$$F(x) = \int_{-\infty}^{x} f(x)\mathrm{d}x,$$

其中, $f(x)$ 称为概率密度函数.

随机变量 X 的均值 (期望) 为

$$EX = \int_{-\infty}^{+\infty} x\mathrm{d}F(x) = \begin{cases} \displaystyle\int_{-\infty}^{+\infty} xf(x)\mathrm{d}x, & X \text{ 是连续的}, \\ \displaystyle\sum_{x} xP\{X = x\}, & X \text{ 是离散的}. \end{cases}$$

对于任意正整数 n, 随机变量 X 的 n 阶矩, 记作 μ_n, 定义为

$$\mu_n' = E(X^n).$$

X 的 n 阶中心矩, 记为

$$\mu_n = E(X - EX)^n.$$

对于随机变量 X 来说, 除期望以外, 最重要的矩莫过于二阶中心矩, 这个矩被称作方差

$$DX = E(X - EX)^2.$$

方差的算术平方根称为标准差, 即

$$\sigma(X) = \sqrt{D(X)} = \sqrt{E(X - EX)^2}.$$

方差反映了随机变量的分散程度, 方差越大, 表明该随机变量取值越分散.

设 X 和 Y 是两个随机变量, 如果

$$\operatorname{cov}(X,Y)=E(X-EX)(Y-EY)$$

存在, 则上式称为 X 和 Y 的协方差. 协方差的绝对值反映了 X 和 Y 相关的密切程度.

随机变量 X 和 Y 之间的相关系数定义为

$$\rho(X,Y)=\frac{\operatorname{cov}(X,Y)}{\sqrt{DX}\sqrt{DY}}.$$

显然, $0\leqslant|\rho(X,Y)|\leqslant 1$, 相关系数绝对值的大小直接反映了随机变量 X 和 Y 之间线性相关的程度. 当 $|\rho(X,Y)|=1$ 时, X 和 Y 之间有完全线性关系. 当 $|\rho(X,Y)|<1$ 时, 表明 X 和 Y 之间存在某种程度的线性关系.

一个随机过程 $\{N(t),t\in T\}$ 是一族随机变量, 即对指标集 T 中的每个 t, $N(t)$ 是一个随机变量. 我们常把 t 解释为时间. 如果指标集是可数的, 则称 $\{N(t),t\in T\}$ 为离散时间随机过程; 如果指标集是一个连续统, 则称 $\{N(t),t\in T\}$ 为连续时间随机过程.

连续时间随机过程 $\{N(t),t\in T\}$ 称为有独立增量, 若对一切 $t_0<t_1<\cdots<t_n$, 随机变量 $N(t_1)-N(t_0), N(t_2)-N(t_1),\cdots,N(t_n)-N(t_{n-1})$ 相互独立. 称 $\{N(t),t\in T\}$ 有平稳增量, 如果 $N(t+s)-N(t)$ 对一切 t 有相同的分布.

为了方便, 本书将随机过程 $\{N(t),t\in T\}$ 简记为 $N(t)$.

1.2 Poisson 过程

随机过程 $\{N(t),\ t\geqslant 0\}$ 称为一个计数过程, 若 $N(t)$ 表示到时刻 t 为止已发生的“事件”的总数, 因此, 一个计数过程 $N(t)$ 满足

(Ⅰ) $N(t)\geqslant 0$,

(Ⅱ) $N(t)$ 是整数值,

(Ⅲ) 若 $s<t$, 则 $N(s)\leqslant N(t)$,

(Ⅳ) 若 $s<t$ 时, $N(t)-N(s)$ 等于区间 $(s,\ t]$ 中发生的事件个数.

如果在不相交的时间区间中发生的事件个数是独立的, 则称计数过程有独立增量. 若在任意时间区间中发生的事件个数的分布只依赖于时间的区间长度, 则称计数过程有平稳增量.

假设 T_n 表示计数过程 $\{N(t),\ t\geqslant 0\}$ 第 n 个事件与第 $n-1$ 个事件之间的间隔时间 (称为点间间距), 则确定了过程的点发生时间序列

$$S_0=0,\quad S_n=\sum_{i=1}^{n}T_i,\quad n=1,2,\cdots.$$

计数过程 $\{N(t),\ t \geqslant 0\}$ 的强度 (若定义的极限存在)

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{P\{N(t+\Delta t) - N(t) > 0\}}{\Delta t}$$

是一个能很好地刻画过程计数性质的量, 可以通过它来描述一个计数过程. $\lambda(t)\mathrm{d}t$ 是在时间 t 和 $t+\mathrm{d}t$ 之间事件出现的概率. 强度 $\lambda(t)$ 可能是一个常数, 也可能是一个随着时间 t 变化的函数, 根据强度 $\lambda(t)$ 的不同类型出现了各种计数过程.

定义 1.1 计数过程 $\{N(t),\ t \geqslant 0\}$ 称为 Poisson 过程 (齐次 Poisson 过程), 具有强度 $\lambda, \lambda > 0$, 如果

(Ⅰ) $N(0) = 0$,

(Ⅱ) 过程有独立增量,

(III) 在任一长度为 t 的区间中事件的个数服从 Poisson 分布, 即对一切 $s,\ t \geqslant 0$,

$$P\{N(t+s) - N(s) = n\} = \mathrm{e}^{-\lambda t} \frac{(\lambda t)^n}{n!}, \quad n = 0, 1, 2, \cdots. \tag{1.1}$$

从条件 (III) 可知 Poisson 过程有平稳增量, 且 $E[N(t)] = \lambda t$, 这正是称 λ 为速率或强度的原因 (单位时间内发生事件的平均数).

为了判断一个计数过程是 Poisson 过程, 必须验证它满足条件 (Ⅰ)、(Ⅱ) 和 (III). 条件 (Ⅰ) 只是说明事件的计数是从时刻 $t = 0$ 开始的. 条件 (Ⅱ) 可以从我们对过程了解的情况去直接验证. 然而全然不知如何确定条件 (III). 为此, 我们引入 Poisson 过程的另一个等价定义.

定义 1.2 计数过程 $\{N(t),\ t \geqslant 0\}$ 称为 Poisson 过程, 具有强度 $\lambda, \lambda > 0$, 如果

(Ⅰ) $N(0) = 0$,

(Ⅱ) 过程有平稳与独立增量,

(III) $P\{N(h) = 1\} = \lambda h + o(h)$,

(Ⅳ) $P\{N(h) \geqslant 2\} = o(h)$.

现在来确定 Poisson 过程的间隔时间 T_n 的分布. 首先注意到事件 $\{T_1 > t\}$ 发生当且仅当 Poisson 过程在区间 $[0, t]$ 内没有发生事件, 因而

$$P\{T_1 > t\} = P\{N(t) = 0\} = \mathrm{e}^{-\lambda t}.$$

因此, T_1 的分布函数为

$$F(t) = 1 - \mathrm{e}^{-\lambda t}, \quad t \geqslant 0, \lambda > 0. \tag{1.2}$$

这个分布函数称为参数为 λ 的指数分布, 其密度函数为

$$f(t) = \lambda \mathrm{e}^{-\lambda t}, \quad t \geqslant 0, \lambda > 0.$$

由独立增量性和平稳性, 有

$$\begin{aligned}P\{T_2>t\big|T_1=s\}=&P\{\text{在}(s,s+t)\text{内没有事件}\big|T_1=s\}\\=&P\{\text{在}(s,s+t)\text{内没有事件}\}\\=&\mathrm{e}^{-\lambda t}.\end{aligned}$$

所以,

$$P\{T_2>t\}=\int_0^\infty P\{T_2>t\big|T_1=s\}\mathrm{d}F(s)=\int_0^\infty \mathrm{e}^{-\lambda t}\mathrm{d}F(s)=\mathrm{e}^{-\lambda t}.$$

因此, T_2 独立于 T_1, 且服从参数为 λ 的指数分布.

不断重复上述同样的推导得到下列命题: $T_n\ (n=1,2,\cdots)$ 为独立同分布的参数为 λ 的指数随机变量. 这个命题的逆命题也成立, 我们有如下定理.

定理 1.1 计数过程 $\{N(t),\ t\geqslant 0\}$ 为强度 λ 的 Poisson 过程当且仅当它的间隔时间 $\{T_n\}$ 是相互独立同分布参数为 λ 的指数随机变量序列.

定理 1.2 强度为 λ 的 Poisson 过程的点发生时间序列 $S_n(n=0,1,2,\cdots)$ 服从参数为 n 与 λ 的 Gamma 分布 $\varGamma(n,\lambda)$, 其密度函数为

$$f(t)=\lambda\mathrm{e}^{-\lambda t}\frac{(\lambda t)^{n-1}}{(n-1)!},\quad t\geqslant 0. \tag{1.3}$$

证明 注意到第 n 个事件在时刻 t 或之前发生当且仅当到时刻 t 已发生的事件数目至少是 n, 即

$$N(t)\geqslant n\Leftrightarrow S_n\leqslant t.$$

因此,

$$P\{S_n\leqslant t\}=P\{N(t)\geqslant n\}=\sum_{j=n}^\infty \mathrm{e}^{-\lambda t}\frac{(\lambda t)^j}{j!}.$$

对上式求导, 可获得 S_n 的密度函数为

$$f(t)=-\lambda\sum_{j=n}^\infty \mathrm{e}^{-\lambda t}\frac{(\lambda t)^j}{j!}+\lambda\sum_{j=n}^\infty \mathrm{e}^{-\lambda t}\frac{(\lambda t)^{j-1}}{(j-1)!}=\lambda\mathrm{e}^{-\lambda t}\frac{(\lambda t)^{n-1}}{(n-1)!},\quad t\geqslant 0.$$ 证毕

定理 1.3 在已知 $N(t)=n$ 的条件下, n 个来到时刻 $S_1,\cdots,S_n$ 与相应于 n 个 $(0,t)$ 上均匀分布的独立随机变量的顺序统计量有相同的分布.

证明 计算给定 $N(t)=n$ 时, $S_1,\cdots,S_n$ 的条件概率密度函数, 设 $0<t_1<t_2<\cdots<t_{n+1}=t$, 且取 h_i 充分小, 使得, $t_i+h_i<t_{i+1}, i=1,2,\cdots,n$. 现在

$$P\{t_i\leqslant S_i\leqslant t_i+h_i, i=1,2,\cdots,n\big|N(t)=n\}$$

$$
\begin{aligned}
&=\frac{P\{\text{在 } [t_i,t_i+h_i]\text{中恰有一个事件}, i=1,2,\cdots,n, \text{在}[0,t]\text{的别处无任何事件}\}}{P\{N(t)=n\}}\\
&=\frac{\lambda h_1 \mathrm{e}^{-\lambda h_1}\cdots\lambda h_n \mathrm{e}^{-\lambda(t-h_1-h_2\cdots-h_n)}}{\mathrm{e}^{-\lambda t}(\lambda t)^n/n!}\\
&=\frac{n!}{t^n}h_1h_2\cdots h_n.
\end{aligned}
$$

因此,

$$\frac{P\{t_i \leqslant S_i \leqslant t_i+h_i, i=1,2,\cdots,n|N(t)=n\}}{h_1h_2\cdots h_n}=\frac{n!}{t^n}.$$

令 $h_i \to 0$, 得到 $S_1,\cdots,S_n$ 在已知 $N(t)=n$ 的条件下的条件密度为

$$f(t_1,\cdots,t_n)=\frac{n!}{t^n},\quad 0<t_1<\cdots<t_n.$$

证毕

直观地, 常说在已知 $(0,t]$ 内发生 n 个事件的条件下, 各事件发生的时刻 $S_1,\cdots,S_n$, 看作不排序的随机变量, 是相互独立的且服从 $(0,t)$ 上的均匀分布.

定理 1.4　假设参数为 λ 的 Poisson 过程的各个事件被分成 1-类或 2-类, 一事件在时刻 s 发生, 以概率 $P(s)$ 被归为 1-类, 否则被归为 2-类, 且与其他事件归作什么类相互独立. 若 $N_i(t)$ 表示到时刻 t, i-类事件发生的个数, $i=1,2$, 则 $N_1(t)$ 与 $N_2(t)$ 是独立的 Poisson 随机变量, 分别有均值 λpt 及 $\lambda(1-p)t$, 其中

$$p=\frac{1}{t}\int_0^t P(s)\mathrm{d}s.$$

证明　对 $N(t)$ 取条件, 计算 $N_1(t)$ 和 $N_2(t)$ 的联合分布:

$$
\begin{aligned}
&P\{N_1(t)=n,N_2(t)=m\}\\
=&\sum_{k=0}^{\infty}P\{N_1(t)=n,N_2(t)=m|N(t)=k\}P\{N(t)=k\}\\
=&P\{N_1(t)=n,N_2(t)=m|N(t)=n+m\}P\{N(t)=n+m\}.
\end{aligned}
$$

现考虑在区间 $[0,t]$ 中发生的任何一事件. 如果它在时刻 s 发生, 则它是 1-类事件的概率为 $P(s)$. 因为已知 $N(t)=n+m$, 由定理 1.3 可知此事件发生的时刻在 $(0,t)$ 上均匀分布, 所以, 它是一个 1-类事件的概率为

$$p=\frac{1}{t}\int_0^t P(s)\mathrm{d}s,$$

与其他事件为什么类型相互独立. 因此,

$$P\{N_1(t)=n,N_2(t)=m|N(t)=n+m\},$$

刚好等于 $n+m$ 次试验中 n 成功 m 次失败的概率, 而 p 是各次试验成功的概率, 即

$$P\{N_1(t)=n, N_2(t)=m \big| N(t)=n+m\} = \begin{pmatrix} n+m \\ n \end{pmatrix} p^n(1-p)^m.$$

从而

$$\begin{aligned} &P\{N_1(t)=n, N_2(t)=m\} \\ =&\frac{(n+m)!}{n!m!}p^n(1-p)^m\frac{(\lambda t)^{n+m}}{(n+m)!}\mathrm{e}^{-\lambda t} \\ =&\mathrm{e}^{-\lambda t}\frac{(\lambda t p)^n}{n!}\mathrm{e}^{-\lambda t(1-p)}\frac{(\lambda t(1-p))^m}{m!}. \end{aligned}$$

证毕

上述 Poisson 过程具有常数强度 λ, 即齐次 Poisson 过程, 如果去掉过程的平稳性, 我们得到非齐次 Poisson 过程.

定义 1.3 计数过程 $\{N(t),\ t \geqslant 0\}$ 称为非齐次 Poisson 过程, 具有强度函数 $\lambda(t)$, 如果

(Ⅰ) $N(0)=0$,

(Ⅱ) 过程有独立增量,

(Ⅲ) 在任一长度为 t 的区间中事件的个数服从 Poisson 分布, 即对一切 $s,\ t \geqslant 0$,

$$P\{N(t+s)-N(s)=n\} = \mathrm{e}^{-\int_s^{t+s}\lambda(s)\mathrm{d}s}\frac{\left(\displaystyle\int_s^{t+s}\lambda(s)\mathrm{d}s\right)^n}{n!}, \quad n=0,1,2,\cdots, \tag{1.4}$$

这里 $\lambda(t)$ 是 $[0,+\infty)$ 上的非负函数, 它在任意有限区间是可积的.

我们引入非齐次 Poisson 过程的另一个等价定义.

定义 1.4 计数过程 $\{N(t),\ t \geqslant 0\}$ 称为非齐次 Poisson 过程, 具有强度函数 $\lambda(t) \geqslant 0, t \geqslant 0$, 如果

(Ⅰ) $N(0)=0$,

(Ⅱ) 过程有独立增量,

(Ⅲ) $P\{N(t+h)-N(h)=1\} = \lambda(t)h + o(h)$,

(Ⅳ) $P\{N(t+h)-N(h) \geqslant 2\} = o(h)$,

这里 $\lambda(t)$ 在任意有限区间是可积的.

定理 1.5 设 $\{N(t),\ t \geqslant 0\}$ 为具有强度函数 $\lambda(t)$ 的非齐次 Poisson 过程, 对于任意给定的正整数 n, 过程的前 n 个点发生时间 $S_1, S_2, \cdots, S_n$ 的联合分布密度函数是

$$f_{S_1,S_2,\cdots,S_n}(x_1,x_2,\cdots,x_n) = \begin{cases} \mathrm{e}^{-\int_0^{x_n}\lambda(x)\mathrm{d}x}\displaystyle\prod_{i=1}^n \lambda(x_i), & 0 < x_1 \leqslant \cdots \leqslant x_n, \\ 0, & \text{其他}. \end{cases} \tag{1.5}$$

证明

$$
\begin{aligned}
&P\{x_i-\Delta x_i<S_i\leqslant x_i, 1\leqslant i\leqslant n\}\\
=&P\{N(x_1-\Delta x_1)=0; N(x_i)-N(x_i-\Delta x_i)=1,\ 1\leqslant i\leqslant n;\\
&N(x_{i+1}-\Delta x_{i+1})-N(x_i)=0,\ 1\leqslant i\leqslant n-1\}\\
=&\mathrm{e}^{-\int_0^{x_1-\Delta x_1}\lambda(x)\mathrm{d}x}\prod_{i=1}^{n}\mathrm{e}^{-\int_{x_i-\Delta x_i}^{x_i}\lambda(x)\mathrm{d}x}\int_{x_i-\Delta x_i}^{x_i}\lambda(x)\mathrm{d}x\prod_{i=1}^{n-1}\mathrm{e}^{-\int_{x_i}^{x_{i+1}-\Delta x_{i+1}}\lambda(x)\mathrm{d}x}\\
=&\prod_{i=1}^{n}\left(\int_{x_i-\Delta x_i}^{x_i}\lambda(x)\mathrm{d}x\right)\mathrm{e}^{-\int_0^{x_n}\lambda(x)\mathrm{d}x}.
\end{aligned}
$$

于是, 按照密度函数的定义,

$$
\begin{aligned}
&f_{S_1,S_2,\cdots,S_n}(x_1,x_2,\cdots,x_n)\\
=&\lim_{\max\Delta x_i\to 0}\frac{1}{\prod_{i=1}^{n}\Delta x_i}P\{x_i-\Delta x_i<S_i\leqslant x_i, 1\leqslant i\leqslant n\}\\
=&\mathrm{e}^{-\int_0^{x_n}\lambda(x)\mathrm{d}x}\prod_{i=1}^{n}\lambda(x_i).
\end{aligned}
$$

证毕

定理 1.6　设 $\{N(t), t\geqslant 0\}$ 为具有强度函数 $\lambda(t)$ 的非齐次 Poisson 过程, 对于任意给定的正整数 n, 过程的 n 个点发生时间为 $S_1, S_2, \cdots, S_n$, 第 $n+1$ 个事件与第 n 个事件之间的间隔时间 T_{n+1} 的条件分布函数是

$$
P\{T_{n+1}\leqslant t|S_n=x_n,\cdots,S_1=x_1\}=1-\exp\left\{-\int_{x_n}^{x_n+t}\lambda(u)\mathrm{d}u\right\}. \tag{1.6}
$$

证明　由非齐次 Poisson 过程的性质知道, 给定 $S_1, S_2, \cdots, S_n$ 时, T_{n+1} 的条件分布函数为

$$
\begin{aligned}
&P\{T_{n+1}\leqslant t|S_n=x_n,\cdots,S_1=x_1\}\\
=&\int_0^t f_{T_{n+1}|S_n,S_{n-1},\cdots,S_1}(x|x_n,x_{n-1},\cdots,x_1)\mathrm{d}x\\
=&\int_0^t f_{S_{n+1}|S_n,S_{n-1},\cdots,S_1}(x_n+x|x_n,x_{n-1},\cdots,x_1)\mathrm{d}x\\
=&\int_0^t \frac{f_{S_{n+1},S_n,S_{n-1},\cdots,S_1}(x_n+x,x_n,x_{n-1},\cdots,x_1)}{f_{S_n,S_{n-1},\cdots,S_1}(x_n,x_{n-1},\cdots,x_1)}\mathrm{d}x\\
=&\int_0^t \frac{\mathrm{e}^{-\int_0^{x_n+x}\lambda(x)\mathrm{d}x}\prod_{i=1}^{n}\lambda(x_i)\lambda(x_n+x)}{\mathrm{e}^{-\int_0^{x_n}\lambda(x)\mathrm{d}x}\prod_{i=1}^{n}\lambda(x_i)}\mathrm{d}x
\end{aligned}
$$

$$
\begin{aligned}
&=\int_0^t \mathrm{e}^{-\int_{x_n}^{x_n+x}\lambda(x)\mathrm{d}x}\lambda(x_n+x)\mathrm{d}x\\
&=1-\mathrm{e}^{-\int_{x_n}^{x_n+t}\lambda(x)\mathrm{d}x}.
\end{aligned}
$$

证毕

1.3 更新过程

Poisson 过程一个最大的特性是它的点间间距相互独立且具有相同指数分布. 如果只保留对点间间距相互独立同分布的要求, 允许它们有任意的分布, 就自然得到 Poisson 过程另一个推广 —— 更新过程.

设 $\{X_n,\ n\geqslant 1\}$ 为一列非负的随机变量. 令

$$S_0=0,\quad S_n=X_1+\cdots+X_n,\quad n\geqslant 1,$$

对 $t\geqslant 0$, 定义

$$N(t)=\max\{n\geqslant 0: S_n\leqslant t\}=\sum_{n=1}^{\infty}\chi(S_n\leqslant t), \tag{1.7}$$

其中 $\chi(S_n\leqslant t)$ 为事件$\{S_n\leqslant t\}$的示性随机变量, 即

$$\chi(S_n\leqslant t)=\begin{cases}1, & \{S_n\leqslant t\}\text{发生},\\ 0, & \{S_n\leqslant t\}\text{不发生}.\end{cases}$$

由定义式 (1.7) 可知, 对一切 $n\geqslant 1$, $t\geqslant 0$,

$$\{N(t)\geqslant n\}=\{S_n\leqslant t\}, \tag{1.8}$$

$$N(t)+1=\inf\{n\geqslant 1: S_n>t\}. \tag{1.9}$$

若 $\{X_n,n\geqslant 1\}$ 独立同分布 (简记为 i.i.d), 其共同分布为 F:

$$F(t)=P\{X_1\leqslant t\}, \tag{1.10}$$

满足条件

$$F(0)<1, \tag{1.11}$$

则称$\{N(t),t\geqslant 0\}$为更新过程. 称 X_n 为第 n 个更新间隔.

由强大数定律, $n\to\infty$ 时

$$\frac{S_n}{n}\to\mu=EX_1,\quad \text{a.s(几乎处处)}.$$