arXiv:2304.06916v1 [stat.ME] 14 Apr 2023

非齐次计数过程：边缘化泊松驱动的 Cox 过程

Shuying Wang*德克萨斯大学奥斯汀分校统计与数据科学系和
Stephen G. Walker†德克萨斯大学奥斯汀分校数学系

摘要

本文研究一种 Cox 过程，其中泊松数据模型的随机强度函数本身是一个非齐次泊松过程。我们证明，可以得到边缘数据过程，即一种呈现过度离散的非齐次计数过程。虽然强度函数是非递减的，但通过对数据进行简单变换，即可使非递减的强度函数适用。我们重点研究一个过程到达时间的时间序列，尤其是能够为观测数据的边缘概率求得精确形式，从而可以通过直接计算似然函数，轻松实现估计算法。

关键词：泊松过程，强度函数，边缘过程。

1 引言

Cox 过程族是一类具有随机强度函数的非齐次泊松过程。它们之间的差异在于随机函数的构造方式不同。最初的应用是对异质空间和/或时空点过程进行建模，例如 Wolpert 和 Ickstadt（1998）的研究。近期的发展主要依赖于对数高斯 Cox 过程，参见 Brix 和 Diggle（2001）、Diggle 等人（2013）、Teng 等人（2017）以及 Bayisa 等人（2020）。双随机过程背后的基本原理是解决过度离散问题。众所周知，泊松随机变量的均值与方差相等，这与许多观测到的现象相反，在这些现象中，点过程的方差超过了其均值。

*通信请联系：shuying.wang@utexas.edu
†s.g.walker@math.utexas.edu

在空间场景中，甚至在一维时间序列的情况下，也需要在强度函数的建模中引入协方差结构。因此，高斯过程是主要工具，不过，由于需要保证强度函数为正数，实际采用的是对数高斯过程。事实上，任何能够将高斯过程转换为正函数的方法都是可行的。其他类型的强度模型还包括散粒噪声过程和其他马尔可夫点过程。关于随机强度函数的早期研究包括 Thomas（1949）以及 Neyman 和 Scott（1958）的工作，他们采用了简化的散粒噪声型过程。

虽然 Cox 过程的应用主要与空间建模有关，但本文重点关注时间序列数据，因此我们只考虑定义在 $(0,\infty)$ 上的随机强度函数。Wu 等（2013）的研究也采用了这一设定，不过这些作者假设观测到的是重复点过程，并使用函数型数据分析技术，而我们则假设只观测到一个过程。

为建立模型并说明符号，我们将聚焦于一维情形，以时间 t, X 为索引；在给定随机过程 $Y(\cdot)$ 的条件下，$\cdot$ 是强度函数为 $Y(t)$ 的泊松过程。Cox 过程存在的条件是，对所有 $T<\infty$， $\int_0^T Y(s)\,ds<\infty$。

强度过程的一种常见且经典的形式如下：

$$Y(t)=\sum_i w_i\, f(t;s_i),$$

对于某组时间点 $S=(s_i)$，且 f 是概率密度函数，其中 w_i 是点 s_i 的权重。这就是散粒噪声 Cox 过程的一般形式。平均强度函数为：

$$\rho(t)=\int\int w\, f\,(t;s)\; ds\,\chi(dw)$$

对于定义在 $(0,\infty)$ 上的某个密度 χ。Neyman 和 Scott（1958）类过程具有 $w_i=w$，而当 f 基于高斯密度时，则为 Thomas（1949）过程。

当 $f(t;\ s)\ =f(t-s)$ 时，该过程是平稳的，在这种情况下，均值函数为 $\rho(t)\ =\rho=\int w\,\chi(w)\,dw.$。有关散粒噪声 Cox 过程的更多内容，请参见 Moller（2003）；Møller 和 Torrisi（2005）；Dassios 等（2015）；Jalilian 等（2015）。有关包括行列式散粒 Cox 过程在内的最新进展，请参见 Moller 和 Vihrs（2022）；有关 Cox 过程及其相关过程的全面综述，请参见 Jang 和 Oh（2021）。

另一方面，对数高斯过程（参见 Møller 等（1998））是通过令 $Y(t)\ =\exp\{Z(t)\}$，其中 Z 为高斯过程而得到的；也就是说，对于任意点集

$(t_1, \ldots, t_m)$，也就是说，$(Z(t_1) \ldots, Z(t_m))$ 服从多元高斯分布。目前已有许多论文专门研究对数高斯考克斯过程，涵盖其理论研究以及用于估计该模型的算法。后者的最新进展必然涉及马尔可夫链蒙特卡洛（MCMC）技术。此时需要构建网格，因此通常采用高斯过程在一组点上的离散化近似；例如，参见 Shirota 和 Gelfand（2016）。Schnoerr 等人（2016）还研究了适用于 Y 的其他类型的过程。

本文以一种新颖的方式对 $Y(t)$ 进行建模，从而使我们能够求得 X 的边际非齐次计数过程。我们要求 Y 为非递减函数，并说明如何利用这样的随机强度函数对任意计数数据进行建模。我们假设 $Y(t)$ 由一个非齐次泊松过程驱动；也就是说

$$Y(t) = \sum_i 1(s_i \leq t),$$

其中，(s_i) 来自一个强度函数为$\gamma(t)$的非齐次泊松过程。因此， $X(t)$的平均强度函数为 $\rho(t) = \int_0^t \gamma(u)\, du$. 本文研究的一项重要成果是，可以推导出 X 过程的边际联合概率。也就是说，如果 $[0, T]$ 上的 X 在时间位置$t_1 < \cdots < t_m$处发生了 m 次变化，那么我们就能够给出 $p(t_1, \ldots, t_m, m)$。这使得推断算法易于实现，无需基于离散化进行近似，也无需对潜在随机过程进行采样。

鉴于 Y 过程取值于特定状态，即非负整数，该模型也可以视为连续时间隐马尔可夫模型，其中模型的隐马尔可夫部分是非齐次的。这类模型以难以估计而闻名，即使隐连续时间过程是齐次的，且取值于有限个状态也是如此。通常的应对方法是将时间离散化，和／或使用 EM算法最大化似然。即使采用这些方法，实现起来也很困难。尽管如此，对于隐非齐次泊松过程，我们能够给出一个易于计算的似然函数表达式。虽然我们可以将该模型视为非齐次连续时间隐马尔可夫模型，但我们将其归为 Cox 过程这一类。

本文其余部分安排如下。第 2 节将以无穷小概率的形式给出模型，利用期望 $\mathrm{E}_Y\left[p_{X|Y}(x \mid Y)\right]$定义 X 过程的边缘概率，并介绍全文将使用的符号。第 3 节将 $x^{(n)}$ 和 $Y^{(n)}$ 描述为 x 和Y 的 n-网格离散化，分别构成离散时间计数路径和离散时间计数过程。第 4 节通过递归推导期望 $\mathrm{E}_{Y^{(n)}}\left[p_{X|Y}\left(x^{(n)} \mid Y_{(n)}\right)\right]$。第 5 节证明，当 n 时，$\mathrm{E}_{Y^{(n)}}\left[p_{X|Y}\left(x^{(n)} \mid Y_{(n)}\right)\right]$ 收敛于 $\mathrm{E}_Y\left[p_{X|Y}(x \mid Y)\right]$。

趋于无穷大，因此取该期望的极限，以获得关于 X 的边缘概率的解析表达式，进而得到似然。

2 模型与背景

考虑 Cox 过程 $\{X(t)\}_{0\le t\le T}$，它是一个非齐次泊松过程，其随机强度函数为 $\beta(t,\ Y(t)) = b(t) + wY(t)$，其中 $\{Y(t)\}_{0\le t\le T}$是一个强度函数为 $\gamma(t)$ 的非齐次泊松过程。本文中，我们取 $b(t) = \beta_0$ 为非负常数，但也可以将其推广为任意实可积函数。

Y 条件下 X 的无穷小概率为

$$\begin{cases} \mathrm{P}\left(X(t+h) = x+1 \mid X(t) = x,\ Y(t) = y\right) = \beta(t,\ y)h + o(h) \\ \mathrm{P}\left(X(t+h) = x \mid X(t) = x,\ Y(t) = y\right) = 1 - \beta(t,\ y)h + o(h), \end{cases}$$

以及 Y 的无穷小概率为

$$\begin{cases} \mathrm{P}\left(Y(t+h) = y+1 \mid Y(t) = y\right) = \gamma(t)h + o(h) \\ \mathrm{P}\left(Y(t+h) = y \mid Y(t) = y\right) = 1 - \gamma(t)h + o(h). \end{cases}$$

我们的目标是通过对 Y 过程进行积分消去，求得边缘过程 $\{X(t)\}_{0<t\le T}$。为了说明这并非一个简单的问题，考虑该模型最简单的版本；即给定 Y， $X(t)$ 服从均值为 $\int_0^t Y(s)\,ds$ 的泊松分布，而 $Y(t)$ 是均值为 t 的泊松过程。 $X(t)$ 的拉普拉斯变换由下式给出。

$$-\log Ee^{-\theta X(t)} = \phi^{-1}(1 - e^{-\phi t}),$$

其中 $\phi = 1 - e^{-\theta}$。这无法识别出 $X(t)$ 的分布，更不用说该过程了。

$\{X(t)\}_{0\le t\le T}$和 $\{Y(t)\}_{0\le t\le T}$的右连续样本路径，分别记为$x\colon [0,T] \to \mathbb{N}$和 $y\colon [0,T] \to \mathbb{N}$，可以通过跳跃时刻来刻画。令 $\{t_i\}$ 为 x 中的跳跃时刻， $\{s_i\}$ 为 y 中的跳跃时刻，则

$$x(t) = \sum_{t_i} \mathbf{1}(t_i \le t), \quad y(t) = \sum_{s_i} \mathbf{1}(s_i \le t), \quad \forall\, t \in [0,T].$$

给定 Y 时， X 的条件分布为

$$p_{X|Y}(x \mid y) = \prod_{t_i} \beta(t_i-,\ y(t_i-)) \exp\left\{-\int_0^T \beta(t,\ y(t))\ dt\right\}, \tag{1}$$

且 Y 的边缘分布由下式给出。

$$p_Y(y) = \prod_{s_i} \gamma(s_i-) \, \exp\left\{ - \int_0^T \gamma(t) \, dt \right\}.$$

边际 al likelihood of X 可表示为关于 Y 的期望,

$$p_X(x) = \int p_Y(y) \, p_{X|Y}(x \mid y) \, dy = \mathrm{E}_Y \left[p_{X|Y}(x \mid Y) \right],$$

这是关于 y 的期望值。由于 Y 是一个连续时间过程，因此很难直接求出其期望值。唯一简单的情况是当 $m = 0$ 时，此时.....

$$p_X(x) = \exp\left\{ - \int_0^T (1 - e^{-(T-s)}) \gamma(s) \, ds \right\},$$

其中，我们还使用了 $w = 1$ 和 $\beta_0 = 0$ 来得到这个表达式。

我们的方法是在 n 网格离散化上对 x 和 Y 进行积分，并推导出期望 $\mathrm{E}_{Y^{(n)}} \left[p_{X|Y} \left(x^{(n)} \mid Y_{(n)} \right) \right]$，其中包含离散时间计数路径 $x^{(n)}$ 和离散时间计数过程 $Y^{(n)}$。随后，我们令 $n \to \infty$，取该期望的极限，以恢复所需的极限期望。期望与极限的交换（因此也包括期望收敛）可通过 $x^{(n)} \to x$ 在 Skorokhod拓扑 下收敛以及 $Y^{(n)} \to Y$ 在分布意义下收敛来证明。

求出具有相应事件时间的 m 个事件的边缘概率，步骤如下：首先求出离散化过程的 $m = 0$。然后利用归纳论证，求出一般 m 的似然函数；完成后，再将离散化过程还原为连续过程，并求出边缘概率的相应极限。

为更清楚地说明离散化过程，假设 Y_n 是一列随机变量，且 Y_n 依分布收敛于随机变量 Y。因此，Y_n 可以是对 Y（结果）的离散化，也可以是对它的其他近似。众所周知，对于所有 g 连续且有界的函数，都有

$$\lim_{n\to\infty} E(g(Y_n)) = E(g(Y)).$$

简而言之，我们使用 $E(g(Y_n))$ 的极限来估计 $E(g(Y))$，并在 Ys 为某些随机过程时应用这一思路。因此，我们先描述相关过程的（时间）离散化／近似。然后求 n 近似的期望，最后求 $n \to \infty$ 时的极限。

3 过程的离散化

期望 $\mathrm{E}_{Y^{(n)}}\left[p_{X|Y}\left(x^{(n)} \mid Y_{(n)}\right)\right]$ 是针对离散时间点过程 $Y^{(n)}$ 和离散时间计数路径 $x^{(n)}$ 求得的，但条件概率函数 $p_{X|Y}(\cdot \mid \cdot)$ 尚未离散化。因此，本节将定义 $x^{(n)}$ 和 $Y^{(n)}$，分别作为计数路径 x 和非齐次泊松过程 Y 的 n- 网格离散化。

令 $D[0,T]$ 为定义在 $[0,T]$ 上且右连续且有左极限的实值函数空间。此外，令 $D_c[0,T]$ 为 $D[0,T]$ 中计数路径的集合。若 D 中的函数“单调不减、取整数值，且在其每个不连续点处恰有一个大小为 1 的跳跃”（Billingsley，1999），则称其为计数路径，因此计数路径是点过程的样本路径。

取第 2 节中定义的样本路径 x，因此 $x \in D_c[0,T]$。对每个 x，将其 n-网格离散化记为 $x^{(n)}$。令 $x^{(n)}$ 在 $\mathcal{T}$ 中的离散时间点 x 处与其一致 $_{(n)}$ ：$= \{kT/n:\ k=1,\ \ldots\ldots,\ n\}$，并在每个时间区间 $[\,(k-1)\,T/n,\ kT/n)$ 内保持常数，其中 $k=1,\ \ldots\ldots,\ n$。因此，$x^{(n)}$ 也是 $D_c[0,T]$ 中右连续的计数路径，但只会在 $t \in \mathcal{T}_{(n)}$ 处发生跳跃。这里我们不考虑 $x^{(n)}$ 在同一时刻发生两次跳跃的情况，因为只要 n-网格足够精细，对于较大的 n，这就不会构成问题。将 $x^{(n)}$ 中的跳跃时刻记为 $\left\{t_i^{(n)}\right\}$，因此 $t_i^{(n)} = \min\left\{s \in \mathcal{T}_{(n)}:\ s \geq t_i\right\}$。

取第 2 节中定义的非齐次泊松过程 Y，并将其 n-网格离散化记为 $Y_{(n)}$。设 $Y^{(n)}$ 是一个只能在离散时间点 $t \in \mathcal{T}_{(n)}$ 发生跳跃的点过程，概率为 $\gamma(t)\ T/n$，因此我们有

$$\mathrm{P}\left(Y^{(n)}\left(t+T/n\right) = Y^{(n)}(t)+1\right) = \gamma(t)\,T/n \qquad \forall\ t \in \mathcal{T}^{(n)},\ t \neq T.$$

现在设 $y^{(n)}$ 为 $Y_{(n)}$ 的样本路径。为简化记号，设 $h = T/n$ 为时间网格的大小。令 $x_k = x^{(n)}(kh)$ 和 $y_k = y^{(n)}(kh)$ 对于 $k=0,\ \ldots\ldots,\ n$，则样本路径 $x^{(n)}$ 和 $y^{(n)}$ 可分别表示为 $(x_0,\ \ldots x_n)$ 和 $(y_0,\ \ldots y_n)$。我们还将强度函数的记号简化为 $\gamma_k = \gamma(kh)$ 和 $\beta_k = \beta(kh,\ \ y_k)$，其中 $k=0,\ \ldots\ldots,\ n$。

$Y^{(n)}$ 的概率函数为

$$\mathrm{P}_{Y^{(n)}}\left(y^{(n)}\right) = \prod_{k=0}^{n-1} \mathrm{P}\left(y_{k+1} \mid y_k\right),$$

其中

$$\mathrm{P}\left(y_{k+1} \mid y_k\right) = \begin{cases} \gamma_k h & \text{for}\ \ y_{k+1} = y_k + 1 \\ 1 - \gamma_k h & \text{for}\ \ y_{k+1} = y_k. \end{cases}$$

将公式 (1) 中的 x 和 y 替换为 $x^{(n)}$ 和 $y^{(n)}$后，函数 $p_{X|Y}\left(x^{(n)} \mid y^{(n)}\right)$可写成如下乘积形式:

$$\begin{aligned} p_{X|Y}\left(x^{(n)} \mid y^{(n)}\right) &= \prod_{t_i^{(n)}} \beta\left(t_i^{(n)} - h,\ y^{(n)}\left(t_i^{(n)} - h\right)\right) \exp\left(-\sum_{k=0}^{n-1} \beta_k h\right) \\ &= \prod_{k=0}^{n-1} p(x_{k+1} \mid x_k,\ y_k), \end{aligned}$$

其中

$$p(x_{k+1} \mid x_k,\ y_k) = \begin{cases} \beta_k \exp(-\beta_k h) & \text{for } \ x_{k+1} = x_k + 1 \\ \exp(-\beta_k h) & \text{for } \ x_{k+1} = x_k. \end{cases}$$

设 $\mathcal{Y}^{(n)}$ 表示 $Y^{(n)}$的所有可能样本路径构成的集合。即使 $n \to \infty$ ， $\mathcal{Y}^{(n)}$ 仍然是可数的，因为 $Y^{(n)}$ 只能发生有限次跳跃。因此，在假设已知 $y_0 = 0$ 的情况下，我们可以通过对所有 $y^{(n)} \in \mathcal{Y}^{(n)}$求和，推导出期望 $\mathrm{E}_{Y^{(n)}}\left[p_{X|Y}\left(x^{(n)} \mid Y_{(n)}\right)\right]$。

$$\begin{aligned} &\mathrm{E}_{Y^{(n)}}\left[p_{X|Y}\left(x^{(n)} \mid Y^{(n)}\right)\right] = \sum_{y^{(n)} \in \mathcal{Y}^{(n)}} p_{X|Y}\left(x^{(n)} \mid y^{(n)}\right) \mathrm{P}_{Y^{(n)}}\left(y^{(n)}\right) \\ &= \sum_{y^{(n)} \in \mathcal{Y}^{(n)}} \prod_{k=0}^{n-1} p(x_{k+1} \mid x_k,\ y_k) \prod_{k=0}^{n-1} \mathrm{P}(y_{k+1} \mid y_k) \\ &= p(x_1 \mid x_0,\ y_0) \sum_{y_1} p(x_2 \mid x_1,\ y_1)\ \mathrm{P}(y_1 \mid y_0) \sum_{y_2} p(x_3 \mid x_2,\ y_2)\ \mathrm{P}(y_2 \mid y_1) \cdots\cdots \\ &\sum_{y_{n-1}} p(x_n \mid x_{n-1},\ y_{n-1})\ \mathrm{P}(y_{n-1} \mid y_{n-2}) \sum_{y_n} \mathrm{P}(y_n \mid y_{n-1}). \end{aligned}$$

为了进一步简化记号，我们逐步定义求和，

$$\begin{aligned} s_1(y_{n-1}) &= \sum_{y_n} \mathrm{P}(y_n \mid y_{n-1}) \\ s_2(y_{n-2}) &= \sum_{y_{n-1}} p(x_n \mid x_{n-1},\ y_{n-1})\ \mathrm{P}(y_{n-1} \mid y_{n-2})\ s_1(y_{n-1}) \\ &\cdots\cdots \\ s_k(y_{n-k}) &= \sum_{y_{n-k+1}} p(x_{n-k+2} \mid x_{n-k+1},\ y_{n-k+1})\ \mathrm{P}(y_{n-k+1} \mid y_{n-k})\ s_{k-1}(y_{n-k+1}) \\ &\circ\ \cdot\circ\ \cdot\circ\ \cdot\circ\ \cdot\circ \\ s_n(y_0) &= \sum_{y_1} p(x_2 \mid x_1,\ y_1)\ \mathrm{P}(y_1 \mid y_0)\ s_{n-1}(y_1). \end{aligned}$$

这里对 $y^{(n)} \in \mathcal{Y}^{(n)}$ 的求和是沿着时间线逆向进行的，因此我们首先对 y_n 求和，得到 $s_1(y_{n-1})$，然后对 y_{n-1} 求和，得到 $s_2(y_{n-2})$，依此类推。当我们推导出截至 $s_k(y_{n-k})$ 的求和时，已经对 y_n、y_{n-1}、……、y_{n-k+1} 求过和，但尚未处理到 y_{n-k-1}，因此 $s_k(y_{n-k})$ 仅依赖于 y_{n-k}。它也依赖于 x_n、x_{n-1}、……、x_{n-k+1}，因此 $s_k(y_{n-k})$ 实际上应写作 $s_k\left(y_{n-k}、\ x^{(n)}\right)$，但这里我们省略了每个 $s_k(y_{n-k})$ 对 $x^{(n)}$ 的依赖关系，只明确指出其对 $y^{(n)}$ 的依赖关系。我们可以尝试找出 $s_k(y_{n-k})$ 与 $s_{k+1}(y_{n-k-1})$ 之间的归纳模式，以推导出截至 $s_n(y_0)$ 的求和，从而得到由……给出的期望。

$$\mathrm{E}_{Y^{(n)}}\left[p_{X|Y}\left(x^{(n)} \mid Y^{(n)}\right)\right] = p(x_1 \mid x_0,\ y_0)\ s_n(y_0).$$

在下一节中，我们将求出离散化过程的边缘概率。

4 边缘概率

这里我们针对一个简单情形推导 $\mathrm{E}_{Y^{(n)}}\left[p_{X|Y}\left(x^{(n)} \mid Y_{(n)}\right)\right]$，假设 $x_0 = y_0 = 0$，且$\beta(t,\ y) = \beta_0 + wy$，其中 $\beta_0 \geq 0$ 和 $w > 0$，因此 $\beta_k = \beta(kh,\ y_k) = \beta_0 + wy_k$。我们对 $x^{(n)}$中的跳跃次数采用归纳论证。

在推导每一步的求和时，我们需要对 $y^{(n)}$ 求和，并考虑在相应的时间点 $x^{(n)}$ 是否发生跳跃。例如，给出的求和步骤为

$$s_{n-k}(y_k) = \sum_{y_{k+1}} p(x_{k+2} \mid x_{k+1},\ y_{k+1})\,\mathrm{P}(y_{k+1} \mid y_k)\,s_{n-k-1}(y_{k+1})$$

是对 y_{k+1} 求和，并且我们需要考虑 x_{k+2}，因此，当 x_{k+2}处没有跳跃时，即 $x_{k+2} = x_{k+1}$时，我们有

$$p(x_{k+2} \mid x_{k+1},\ y_{k+1})\,\mathrm{P}(y_{k+1} \mid y_k) = \begin{cases} \exp(-\beta_k h)\,(1-\gamma_k h) & \text{for } y_{k+1} = y_k \\ \exp\{-(\beta_k + w)h\}\,\gamma_k h & \text{for } y_{k+1} = y_k + 1, \end{cases}$$

并且当在 x_{k+2} 处发生跳跃时，即 $x_{k+2} = x_{k+1} + 1$，我们有

$$p(x_{k+2}|x_{k+1},\,y_{k+1})\mathrm{P}(y_{k+1}|y_k) = \begin{cases} \beta_k \exp(-\beta_k h)(1-\gamma_k h) & \text{for } y_{k+1} = y_k \\ (\beta_k + w)\exp\{-(\beta_k + w)h\}\gamma_k h & \text{for } y_{k+1} = y_k + 1. \end{cases}$$

4.1 无跳跃

首先，假设 $x^{(n)}$ 是一个在时间区间 $[0, T]$ 内没有跳跃的样本路径，因此对于所有 $k = 1, \ldots, n$，均有 $x_k = x_{k-1}$。

求和的推导从 $s_1(y_{n-1}) = \sum_{y_n} \mathrm{P}(y_n \mid y_{n-1}) = 1$ 开始。然后我们对 y_{n-1} 求和，得到

$$\begin{aligned} s_2(y_{n-2}) &= \sum_{y_{n-1}} p(x_n \mid x_{n-1},\ y_{n-1})\ \mathrm{P}(y_{n-1} \mid y_{n-2}) \\ &= \exp(-\beta_{n-2}h)(1-\gamma_{n-2}h) + \exp\{-(\beta_{n-2}+w)h\}\gamma_{n-2}h \\ &= \exp(-\beta_{n-2}h)\ \left\{1 - \left(1 - e^{-wh}\right)\gamma_{n-2}h\right\}. \end{aligned}$$

接下来，我们对 y_{n-2} 求和以得到

$$\begin{aligned} s_3(y_{n-3}) &= \sum_{y_{n-2}} p(x_{n-1} \mid x_{n-2},\ y_{n-2})\ \mathrm{P}(y_{n-2} \mid y_{n-3})\ s_2(y_{n-2}) \\ &= \left[\exp(-2\beta_{n-3}h)(1-\gamma_{n-3}h) + \exp\{-2(\beta_{n-3}+w)h\}\gamma_{n-3}h\right]\ (1-\lambda_{n-2}h) \\ &= \exp(-2\beta_{n-3}h)\ (1-\lambda_{n-2}h)\ (1-\lambda_{n-3}h), \end{aligned}$$

其中 $\lambda_i := \left(1 - e^{-(n-i-1)wh}\right)\gamma_i$. 不断重复这个过程，我们可以得到，对于 $2 \le k \le n$,

$$s_k(y_{n-k}) = \exp\{-(k-1)\beta_{n-k}h\} \prod_{i=n-k}^{n-2} (1-\lambda_i h). \tag{2}$$

上面的等式可以通过归纳法轻松证明。我们知道它对$s_2(y_{n-2})$成立，而且也很容易证明，当它对 $s_k(y_{n-k})$ 成立时，对 $s_{k+1}(y_{n-k-1})$ 也成立。

因此，在求和结束时，我们将得到

$$\mathrm{E}_{Y^{(n)}}\left[p_{X|Y}\left(x^{(n)} \mid Y^{(n)}\right)\right] = p(x_1 \mid x_0,\ y_0)\ s_n(y_0) = \exp(-n\beta_0 h) \prod_{i=0}^{n-2}(1-\lambda_i h).$$

以此为出发点，我们现在通过归纳法，求出对应的 m 跳跃形式。

4.2 归纳论证

在本节中，我们将通过归纳法和递归，推导出对于时间区间 $[0,T]$ 内具有任意数量跳跃的样本路径 $x^{(n)}$，$\mathrm{E}_{Y^{(n)}}[p_{X|Y}\left(x^{(n)} \mid Y_{(n)}\right)]$ 的一般表达式。

我们将跳跃时刻记为 $t_i^{(n)} = (r_i + 2)h$，并按降序排列，其中 $t_i^{(n)} > t_j^{(n)}$ 对应于 $i < j$。首先，我们通过归纳法推导 $s_k(y_{n-k})$ 的一般表达式。

定理 1。给定一个样本路径 $x^{(n)}$，令 m 表示在 $x^{(n)}$ 中截至 s_k（y_{n-k}）已计入求和的跳跃次数，即 $m = \sum_{r_i} \mathbf{1}$（$r_i \ge n-k$）。

那么，存在一组系数 $\left\{c_j^{(m,n)}{}_{,\ k}\right\}_{j=0}^{m}$，它们仅取决于 $x^{(n)}$，而不取决于 $y^{(n)}$，使得 s_k（y_{n-k}）在 $2 \le k \le n$ 下的一般表达式为：

$$s_k(y_{n-k}) = \left(\sum_{j=0}^{m} c_{j,\ k}^{(m,n)}\, w^j\, \beta_{n-k}^{m-j}\right) \exp\{-(k-1)\beta_{n-k}h\} \prod_{i=n-k}^{n-2}(1-\lambda_i h). \qquad (3)$$

证明见附录。

现在，我们已经得到了适用于任意 $2 \le k \le n$ 的 $s_k(y_{n-k})$ 的一般表达式，但仍需推导系数 $\left\{c_j^{(m,n)}{}_{,\ k}\right\}_{j=0}^{m}$ 的具体表达式。如定理 1 所示，系数 $\left\{c_j^{(m,n)}{}_{,\ k}\right\}_{j=0}^{m}$ 取决于 $x^{(n)}$，其中 m 是迄今为止在求和中已计入的 $x^{(n)}$ 中的跳跃次数，即 $m = \sum_{r_i} \mathbf{1}(r_i \ge n-k)$。现在，我们需要将 $\left\{c_j^{(m,n)}{}_{,\ k}\right\}_{j=0}^{m}$ 的定义展开为 $m \le \sum_{r_i} \mathbf{1}(r_i \ge n-k)$。

对于样本路径 $x^{(n)}$，其跳跃时刻 $t_i^{(n)} = (r_i+2)h$ 按降序排列，定义 $\left\{c_j^{(m,n)}{}_{,\ k}\right\}_{j=0}^{m}$，其中 $m \le \sum_{r_i} \mathbf{1}(r_i \ge n-k)$ 是样本路径 $x^{(m,n)}$（其跳跃时刻为 $\left\{t_i^{(n)}\right\}_{i=1}^{m}$）的 $s_k\left(y_{n-k},\, x^{(m,n)}\right)$ 系数。这意味着，沿时间线向后计数时，$x^{(m,n)}$ 与 $x^{(n)}$ 有相同的前 m 次跳跃，但之后不再发生其他跳跃。例如，$c_0^{(0,n)}{}_{,\ k} = 1$ 是 $s_k\left(y_{n-k},\, x^{(0,n)}\right)$ 的系数，其中 $x^{(0,n)}$ 是一条没有跳跃的样本路径，而 $\{c_{0,\ k}^{(1,n)},\, c_{1,\ k}^{(1,n)}\}$ 是 $s_k\left(y_{n-k},\, x^{(1,n)}\right)$ 的系数，其中 $x^{(1,n)}$ 是一条仅在 $t_1^{(n)}$ 处发生一次跳跃的样本路径。

在定理 1 的证明中，我们已经推导出一个递推方程，该方程在 x_{n-k+1} 时没有发生跳跃的情况下，根据方程（6），用 $\left\{c_j^{(m,n)}{}_{,\ k}\right\}_{j=0}^{m}$ 表示 $\left\{c_j^{(m,n)}{}_{,\ k+1}\right\}_{j=0}^{m}$；同时，我们还推导出了另一个递推方程，在 x_{n-k+1} 时发生跳跃的情况下，根据方程 (7)，用 $\left\{c_j^{(m,n)}{}_{,\ k}\right\}_{j=0}^{m}$ 表示 $\left\{c_j^{(m+1,n)}{}_{,\ k+1}\right\}_{j=0}^{m+1}$。

基于上述两个递推方程，我们将能够表示 $\left\{c_j^{(m,n)}{}_{\ k}\right\}_{j=0}^{m}$ 关于 $\left\{c_j^{(m-1,n)}{}_{,\ k}\right\}_{j=0}^{m-1}$ 对于任意 $m > 0$ 和 $n - r_m \le k \le n$。这将是我们推导最终结果时使用的关键递推关系，因为它为我们提供了一种递归地从 $c_0^{(0,n)}{}_{,\ n}$ 得到 $\left\{c_j^{(m,n)}{}_{,\ n}\right\}_{j=0}^{m}$ 的方法，该方法适用于任意 $m > 0$，从而可以递归地得到具有任意跳跃次数的 $s_n(y_0)$。这个递推方程将在下面的定理中给出，并通过归纳法证明。

定理 2。定理 1 中定义的系数具有递归关系。

从 $m-1$ 到 m，定义为

$$c_{j,\ k}^{(m,n)} = \begin{cases} c_{0,\ k}^{(m-1,n)} & \text{for } j=0 \\ \sum_{i=0}^{j-1} c_{i,\ k}^{(m-1,n)} \binom{m-i-1}{j-i-1} \sum_{i=n-k}^{r_m} \alpha_i h + c_{j,\ k}^{(m-1,n)} & \text{for } j=1,2,\dots m-1 \\ \sum_{i=0}^{m-1} c_{i,\ k}^{(m-1,n)} \sum_{i=n-k}^{r_m} \alpha_i h & \text{for } j=m, \end{cases} \quad (4)$$

其中 $\alpha_i := e^{-(n-i-1)wh}\,\gamma_i$，对于任意 $m>0$ 和 $n-r_m \le k \le n$。

证明见附录。

根据第 4.1 节，当 $x^{(n)}$ 中没有跳跃时，系数为$c_{0\ ,\ k}^{(0,n)} = 1$，对于 $2 \le k \le n$。
根据定理 2，当 $x^{(n)}$ 中只有一个跳跃时，系数为

$$c_{0,\ k}^{(1,n)} = c_{0,\ k}^{(0,n)} = 1 \quad \text{and} \quad c_{1,\ k}^{(1,n)} = c_{0,\ k}^{(0,n)} \sum_{i=n-k}^{r_1} \alpha_i h = \sum_{i=n-k}^{r_1} \alpha_i h, \quad \text{for } n-r_1 \le k \le n.$$

当 $x^{(n)}$ 中有两个跳跃时，系数为

$$\begin{aligned} c_{0,\ k}^{(2,n)} &= c_{0,\ k}^{(1,n)} = 1 \\ c_{1,\ k}^{(2,n)} &= c_{0,\ k}^{(1,n)} \sum_{i=n-k}^{r_2} \alpha_i h + c_{1,\ k}^{(1,n)} = \sum_{i=n-k}^{r_2} \alpha_i h + \sum_{i=n-k}^{r_1} \alpha_i h \\ c_{2,\ k}^{(2,n)} &= (c_{0,\ k}^{(1,n)} + c_{1,\ k}^{(1,n)}) \sum_{i=n-k}^{r_2} \alpha_i h = \left(\sum_{i=n-k}^{r_1} \alpha_i h + 1 \right) \sum_{i=n-k}^{r_2} \alpha_i h, \end{aligned}$$

对于 $n-r_2 \le k \le n$。

因此，根据定理 1，对于 m 个跳跃的情况，我们得到了 $\mathrm{E}_{Y^{(n)}}\left[p_{X|Y}\left(x^{(n)} \mid Y_{(n)}\right)\right]$ 的一般表达式，如下所示

$$\begin{aligned} \mathrm{E}_{Y^{(n)}}\left[p_{X|Y}\left(x^{(n)} \mid Y^{(n)}\right)\right] &= p(x_1 \mid x_0,\ y_0) s_n(y_0) \\ &= \left(\sum_{j=0}^{m} c_{j,\ n}^{(m,n)}\ w^j \beta_0^{m-j} \right) \exp(-n\beta_0 h) \prod_{i=0}^{n-2} (1-\lambda_i h), \end{aligned}$$

其中，系数 $\left\{c_{j\ ,\ n}^{(m,n)}\right\}_{j=0}^{m}$ 可通过在 $k=n$ 处应用定理 2，从 $c_{0\ ,\ n}^{(0,n)} = 1$ 递归推导得出。

5 收敛结果

本节将证明期望的收敛性

$$\mathrm{E}_{Y^{(n)}}\left[p_{X|Y}\left(x^{(n)} \mid Y^{(n)}\right)\right] \to \mathrm{E}_{Y}\left[p_{X|Y}(x \mid Y)\right],$$

通过证明 $Y^{(n)} \to Y$ 在分布上收敛，以及 $x^{(n)} \to x$ 在Skorokhod拓扑下收敛。

5.1 弱收敛

取第 2 节中定义的非齐次泊松过程 Y 及其在第 3 节中定义的 n-网格离散化 $Y_{(n)}$。这里，我们通过考察有限维分布的收敛来证明 $Y^{(n)} \to Y$在分布上收敛；参见 Billingsley（1999）。

引理 1。 $Y^{(n)}$ 的所有有限维分布弱收敛于 Y的相应有限维分布。

证明。对于给定的时间点 $t \in [0,T]$，令 S_t和 $S_t^{(n)}$分别表示 Y和 $Y^{(n)}$发生下一次跳跃前的等待时间，即。

$$S_t = \inf\{s \in [0, T-t] : Y(t+s) = Y(t) + 1\}$$

and

$$S_t^{(n)} = \inf\{s \in [0, T-t] : Y^{(n)}(t+s) = Y^{(n)}(t) + 1\}.$$

只需证明，对于任意 $t \in [0,T]$， $S_t^{(n)}$依分布收敛至 S_t，因为 Y 和 $Y^{(n)}$ 都是马尔可夫过程。

S_t 和 $S_t^{(n)}$ 的累积分布函数为，

$$F_t(s) = \mathrm{P}(S_t \le s) \quad \text{and} \quad F_t^{(n)}(s) = \mathrm{P}\left(S_t^{(n)} \le s\right),$$

因此，我们想要证明，对于任意 $t \in [0,T]$ 和 $s \in [0, T-t]$， $F_{(n)\,t}(s) \to F_t(s)$。

对于 Y，等待时间的累积分布函数为，

$$F_t(s) = \mathrm{P}(S_t \le s) = 1 - \exp\left\{-\int_t^{t+s} \gamma(u)\, du\right\}.$$

For $Y_{(n)}$，我们可以考虑在 $[t, t+s]$中没有发生跳跃的概率 ，

$$F_t^{(n)}(s) = 1 - \mathrm{P}\left(S_t^{(n)} > s\right) = 1 - \prod_{u \in \mathcal{T}^{(n)} \cap [t, t+s]} \left[1 - \gamma(u)\frac{1}{n}\right].$$

当 $n \to \infty$，根据乘积积分（Dollard and Friedman，1984），其极限为 b y

$$\lim_{n\to\infty} F_t^{(n)}(s) = 1 - \exp\left\{-\int_t^{t+s} \gamma(u) du\right\} = F_t(s).$$

因此，在任意给定的时间点 $t \in [0,T]$， $Y^{(n)}$ 下一次跳跃的等待时间依分布收敛于 Y 的等待时间，即 $S_t^{(n)} \to S_t$ 依分布收敛，由此完成证明。□

根据 Billingsley（1999）定理 12.6， D_c 中有限维分布的收敛意味着弱收敛，因此我们有 $Y_{(n)} \xrightarrow{d} Y$。

现在，固定 x，并将条件概率函数 $p_{X|Y}(x \mid y)$ 视为关于 y 的函数。根据公式 (1)，该函数将从 D_c 到来的计数路径 y 映射为一个实数。对于任意固定的计数路径 $x \in D_c$，该函数

(i) 有界；因为 x 只能有有限多个跳跃，即 $\{t_i\}$ 是一个有限集合，并且强度函数 $\beta(t,\ Y(t))$ 有界；(ii) 几乎处处连续；因为不连续点只会在 y 于 $t \in \{t_i\}$ 发生跳跃时出现。

现在，利用连续映射定理（参见 Billingsley（1999）中的定理 2.7），对于有界且几乎处处连续的函数 $p_{X|Y}(x \mid y)$，由 $Y_{(n)} \xrightarrow{d} Y$，我们得到期望收敛：$\mathrm{E}_{Y^{(n)}}\left[p_{X|Y}\left(x \mid Y_{(n)}\right)\right] \to \mathrm{E}_Y\left[p_{X|Y}(x \mid Y)\right]$。

5.2 Skorokhod拓扑中的收敛

取第 2 节中定义的计数路径 x 及第 3 节中定义的其 n-网格离散化 $x^{(n)}$。

离散化所引起的跳跃时刻误差不会大于时间网格的大小，即 $\left|t_i^{(n)} - t_i\right| < T/n$，因此 $t_i^{(n)} \to t_i$ 当 $n \to \infty$。因此，存在一个时间尺度变形，记为 $\tau^{(n)}$：$[0,T] \to [0,T]$，它是连续、递增且满射的，使得对于任意 $t \in [0,T]$，$\lim_{n\to\infty} x^{(n)}\left(\tau^{(n)}(t)\right) = x(t)$。此外，$\left|\tau_{(n)}(t) - t\right| < T/n$，因此 $\tau_{(n)}(t) \to t$ 一致地，且 $x^{(n)}(t) \to x(t)$ 对于 x 的所有连续点 t。注意，$x \in D_c$ 只有有限个不连续点，因此根据 Skorokhod拓扑可得 $x^{(n)} \to x$ （参见 Billingsley（1999）第124页）。

考虑 $\mathrm{E}_Y\left[p_{X|Y}(x \mid Y)\right]$ 作为一个将 $D_c[0,T]$ 映射到 $\mathbb{R}$的 x 函数，则该函数关于定义 Skorokhod 拓扑的度量是连续的（见 Billingsley (1999) 中的 (12.13)）。因此，$x^{(n)} \to x$ 在 Skorokhod 拓扑下收敛，会导致如下期望收敛：$\mathrm{E}_Y\left[p_{X|Y}\left(x^{(n)} \mid Y\right)\right] \to \mathrm{E}_Y\left[p_{X|Y}(x \mid Y)\right]$.

5.3 期望收敛

定理3。我们有如下期望收敛

$$\mathrm{E}_{Y^{(n)}}\left[p_{X|Y}\left(x^{(n)} \mid Y^{(n)}\right)\right] \to \mathrm{E}_Y\left[p_{X|Y}(x \mid Y)\right],$$

也就是说，对于任意 $\epsilon > 0$，都存在一个整数 N，使得

$$\left|\ \mathrm{E}_{Y^{(n)}}\left[p_{X|Y}\left(x^{(n)} \mid Y^{(n)}\right)\right] \ - \ \mathrm{E}_Y\left[p_{X|Y}(x \mid Y)\right]\ \right| < \epsilon, \quad \forall\ n > N.$$

证明。根据三角不等式，$\left|\mathrm{E}_{Y^{(n)}}\left[p_{X|Y}\left(x^{(n)}|Y_{(n)}\right)\right]-\mathrm{E}_{Y}\left[p_{X|Y}(x\mid Y)\right]\right|$的上界为

$$\left|\mathrm{E}_{Y^{(n)}}\left[p_{X|Y}\left(x^{(n)}|Y^{(n)}\right)\right]-\mathrm{E}_{Y}\left[p_{X|Y}\left(x^{(n)}|Y\right)\right]\right|+\left|\mathrm{E}_{Y}\left[p_{X|Y}\left(x^{(n)}|Y\right)\right]-\mathrm{E}_{Y}\left[p_{X|Y}(x\mid Y)\right]\right|.$$

根据第 5.1 节，我们有 $\mathrm{E}_{Y^{(n)}}\left[p_{X|Y}\left(x\mid Y_{(n)}\right)\right]\to\mathrm{E}_{Y}\left[p_{X|Y}(x\mid Y)\right]$，因此对于任意 $\epsilon>0$，存在一个整数 N_1，使得

$$\left|\mathrm{E}_{Y^{(n)}}\left[p_{X|Y}\left(x^{(n)}\mid Y^{(n)}\right)\right]-\mathrm{E}_{Y}\left[p_{X|Y}\left(x^{(n)}\mid Y\right)\right]\right|<\epsilon/2,\quad\forall\, n>N_1.$$

根据第 5.2 节，我们有 $\mathrm{E}_{Y}\left[p_{X|Y}\left(x^{(n)}|Y\right)\right]\to\mathrm{E}_{Y}\left[p_{X|Y}(x\mid Y)\right]$，因此对于任意 $\epsilon>0$，都存在一个 N_2，使得

$$\left|\mathrm{E}_{Y}\left[p_{X|Y}\left(x^{(n)}\mid Y\right)\right]-\mathrm{E}_{Y}\left[p_{X|Y}(x\mid Y)\right]\right|<\epsilon/2,\quad\forall\, n>N_2.$$

因此，结合第 5.1 节和第 5.2 节，我们得到的是

$$\left|\mathrm{E}_{Y^{(n)}}\left[p_{X|Y}\left(x^{(n)}\mid Y^{(n)}\right)\right]-\mathrm{E}_{Y}\left[p_{X|Y}(x\mid Y)\right]\right|<\epsilon,\quad\forall\, n>\max\left(N_1,N_2\right),$$

因此，当 $n\to\infty$ 时，$\mathrm{E}_{Y^{(n)}}\left[p_{X|Y}\left(x^{(n)}\mid Y_{(n)}\right)\right]\to\mathrm{E}_{Y}\left[p_{X|Y}(x\mid Y)\right]$。

□

5.4 边缘似然

现在，我们可以取出在第 4 节中推导的 $\mathrm{E}_{Y^{(n)}}\left[p_{X|Y}\left(x^{(n)}\mid Y_{(n)}\right)\right]$ 的一般表达式，并令 $n\to\infty$，以恢复所需的期望 $\mathrm{E}_{Y}\left[p_{X|Y}(x\mid Y)\right]$，从而得到第 2 节中所述 Cox 过程 X 的边缘似然。

定理 4。取第 2 节中定义的 Cox 过程 $\{X(t)\}_{0\le t\le T}$ 及其样本路径 x。假设样本路径 x 由按降序排列的跳跃时刻 $\{t_i\}_{i=1}^{M}$刻画，即对于 $i<j$，有 $t_i>t_j$。则 X 的边缘似然为

$$p_X(x)=\left(\sum_{j=0}^{M}c_j^{(M)}w^j\beta_0^{M-j}\right)\exp\left\{-\beta_0 T-\int_0^T\lambda(t)dt\right\},$$

其中，$\lambda(t):=\left(1-e^{-w(T-t)}\right)\gamma(t)$，并且系数 $\left\{c_j^{(M)}\right\}_{j=0}^{M}$ 由以下递推方程给出，对于 $m=1$，2、.....、M，

$$\begin{aligned}
c_0^{(0)}&=1\\
c_j^{(m)}&=\begin{cases}
c_0^{(m-1)} & \text{for } j=0\\
\sum_{i=0}^{j-1}c_i^{(m-1)}\binom{m-i-1}{j-i-1}\int_0^{t_m}\alpha(t)dt+c_j^{(m-1)} & \text{for } j=1,2,\dots m-1\\
\sum_{i=0}^{m-1}c_i^{(m-1)}\int_0^{t_m}\alpha(t)dt & \text{for } j=m,
\end{cases}
\end{aligned}\tag{5}$$

其中 $\alpha(t)=e^{-w(T-t)}\gamma(t)$。

证明见附录。

6 图示

首先，在第6.1节中，我们详细介绍模拟 Cox 过程和计算边缘似然的算法。然后，在第6.2节中，我们展示模拟数据的示例。所有后验抽样均采用 Metropolis-Hastings算法完成，该算法会对定义强度函数$\gamma(t)$的参数提出候选值。候选值来自相互独立的正态分布，其方差经过调整，以使链达到特定的混合水平。

我们始终将强度设为多项式函数，因此参数就是该多项式的系数。在模拟数据时，我们假设w 已知。另一方面，当我们考察真实数据集（见第6.3节）并将数据转换为递增的强度函数时，需要以特定方式指定 w ，具体内容将在第6.3节中说明。

本节所有插图的代码均可在该网页中找到（ https://github.com/ShuyingWang/marginal-cox-process）。

6.1 算法

为了模拟第 2 节中定义的 Cox 过程，我们采用模拟非齐次泊松过程的标准方法；参见 Kingman（1992）。首先，我们需要通过标准泊松过程的转换，模拟强度函数为 $\gamma(t)$ 的潜在非齐次泊松过程 $\{Y(t)\}_{0\leq t\leq T}$。然后，给定 Y 的一个样本路径，我们将 X 模拟为强度为 $(\beta_0 + wY(t))$ 的非齐次泊松过程；参见算法1。利用定理 4 中的递归表达式计算边缘似然的过程详见算法2。

6.2 模拟数据

首先，我们使用模拟数据，对强度 $\gamma(t)$ 为多项式函数的情况进行一些示例说明。具体而言，我们假设强度函数分为 4 种类型，即常数、线性、二次和三次。每种情况下的数据均使用算法1生成。对于所有情况，均采用方差较大的独立正态分布作为弱先验。此外，每种情况下都使用 Metropolis抽样器，从多项式函数系数的后验分布中抽取样本。我们对独立正态提议分布的方差进行调节，使提出整个新函数时的接受率达到 0.2 至 0.3。也就是说，要么所有参数都发生变化，要么函数保持在当前状态。

图1展示了每种情况下的样本函数以及真实函数。

算法1： 模拟 Cox 过程

输入： T, $\gamma(t)$ β_0, w

输出： $\{t_i\}$

定义函数： $\Gamma(t) \leftarrow \int_0^t \gamma(t)dt$;

绘制 $u \sim$ Exponential(1);

$t^{(y)} \leftarrow \Gamma^{-1}(u)$;

$t^{(x)} \leftarrow 0$;

$y \leftarrow 0$;

当 $t^{(x)} < T$ **时做**

  绘制 $r \sim$ 指数分布$(\beta_0 + wy)$;

  若 $t^{(x)} + r <$ min$(t^{(y)}, T)$ **然后**

    $t^{(x)} \leftarrow t^{(x)} + r$;

    将 $t^{(x)}$ 添加到 $\{t_i\}$ 中

  end

  否则如果 $t^{(y)} < T$ **那么**

    抽取 $r \sim$ Exponential(1);

    $u \leftarrow u + r$;

    $t^{(y)} \leftarrow \Gamma^{-1}(u)$;

    $y \leftarrow y + 1$;

    $t^{(x)} \leftarrow t^{(y)}$

  end

  else

    $t^{(x)} \leftarrow T$

  end

end

算法2： 计算边缘似然

输入： $T,\ \gamma(t),\ \beta_0,\ w,\ \{t_i\}$

输出： 边缘似然

按降序排列 $\{t_i\}$ ；

$M \leftarrow$ $\{t_i\}$的计数;

$c_0^{(0)} \leftarrow 1$;

$c_0^{(1)} \leftarrow 0$;

对于 $m = 1:\quad M$ **执行**

  $c_0^{(m)} \leftarrow 1$;

  对于 $j = 1:\quad m$ **做**

    $c_j^{(m)} \leftarrow \sum_{i=0}^{j-1} c_i^{(m-1)} \binom{m-i-1}{j-i-1} \int_0^{t_m} e^{-w(T-t)} \gamma(t)\ dt + c_j^{(m-1)}$;

  end

  $c_{m+1}^{(m)} \leftarrow 0$;

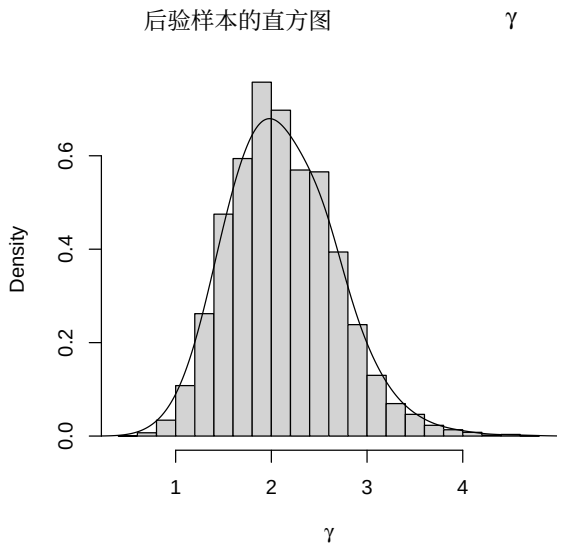
end

边缘似然 $\leftarrow \left(\sum_{j=0}^{M} c_j^{(M)} w^j \beta_0^{M-j} \right) \exp\ \left\{ -\beta_0 T - \int_0^T \left(1 - e^{-w(T-t)}\right) \gamma(t)\ dt \right\}$;

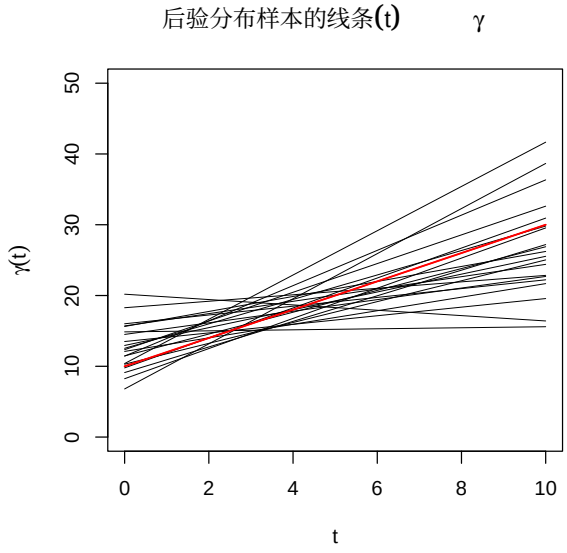
在图1a中，该函数由一个常数表示，后验样本以直方图呈现，并叠加了经验密度曲线。 γ 的真实值为 2，后验均值为 2.12，后验众数为 1.98。其他所有情况（即图1b、图1c 和图1d）中，都绘制了多个后验函数，并同时展示以红色标出的 $\gamma(t)$ 真实函数。

6.3 真实数据

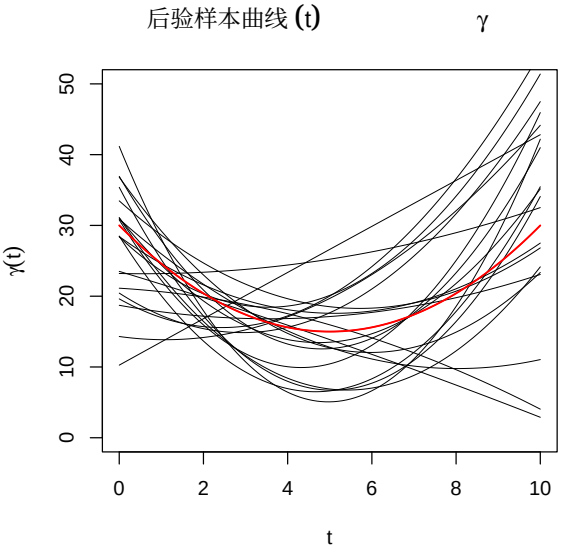
对于真实数据示例，我们使用 AEGISS 数据集，该数据集最初来源于 Diggle 等人（2005 年），并由 Taylor 等人（2013 年）进行了详细介绍。这是一个时空疾病监测数据集，记录了疾病事件发生的时间和地点。本文仅关注时间序列数据，因此，我们从该数据集的空间范围中选取一个矩形区域，并将该区域内发生的事件的时间点视为一个计数过程。为了拟合我们的 Cox 过程模型，我们将原始计数过程转换为一个具有非递减强度的适应性计数过程。数据适应方法如下。假设原始计数过程在时间区间 $[0, T]$ 内有 $M^\star$ 次跳跃。令 $\{t_i^\star\}_{i=1}^{M^\star}$ 表示跳跃时刻， $x^\star(t) = \sum_{i=1}^{M^\star} \mathbf{1}(t_i \le t)$ 表示原始计数过程的样本路径。首先，我们对 $x^\star(t)$ 进行积分，并通过常数 w 进行缩放，得到一个由 $\tilde{x}(t) = w \int_0^t x^\star(s) ds$ 给出的积分函数。其次，我们取 $\tilde{x}(t)$ 首次达到各个整数值的时间点，并将其记为 $\{\tilde{t}_i\}_{i=1}^{M}$，因此 $\tilde{t}_i = \inf\{t : \tilde{x}(t) = i\}$。第三，我们将 $\tilde{x}(t)$ 转换为一个计数路径 $x(t)$，其具有



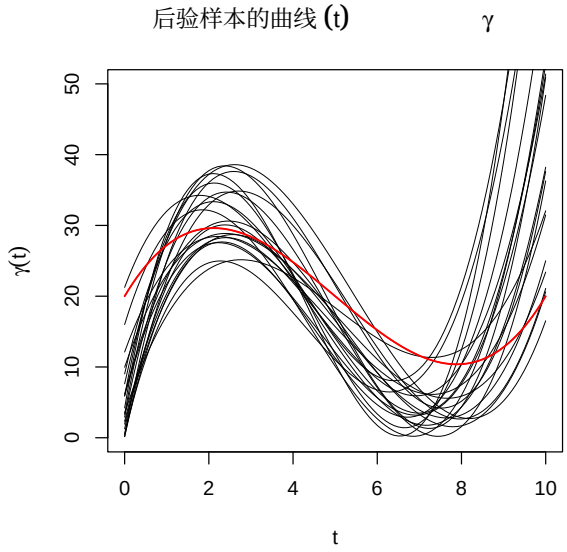
(a) 常数 γ



(b) 线性 $\gamma(t)$



(c) 二次 $\gamma(t)$



(d) 三次 $\gamma(t)$

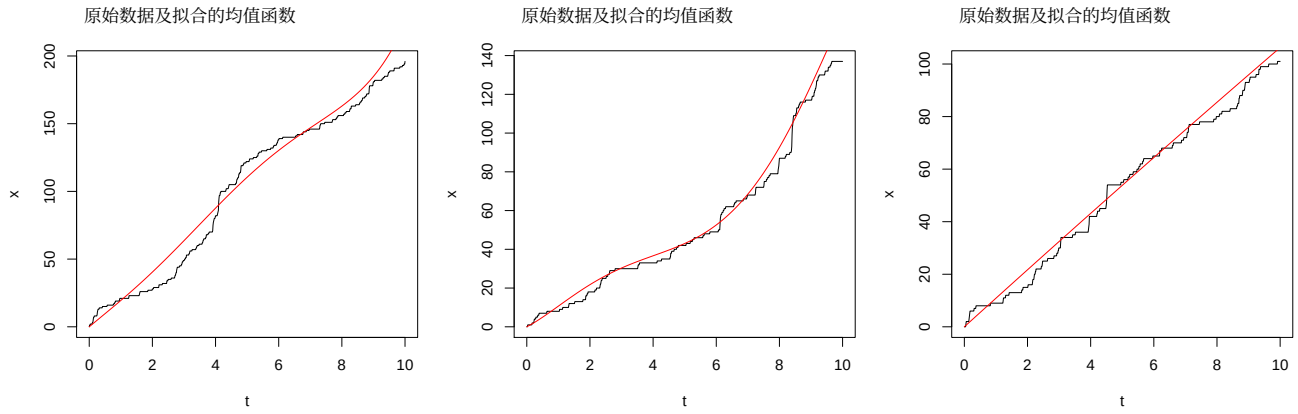
图1：模拟数据示例：这四幅图分别展示了常数、线性、二次和三次多项式的后验分布样本。第一种情形中的后验样本是常数值；对于后三种情形，我们展示了从后验分布中随机抽取的多项式曲线，其中 $\gamma(t)$ 的真实曲线以红色表示。

M 次跳跃，使得在每个 $\tilde{t}_i$ 处， $x(t)$ 与 $\tilde{x}(t)$ 重合，并且截至每个 $\tilde{t}_i$， $x(t)$ 下方的面积等于 $\tilde{x}(t)$ 下方的面积，即 $x\left(\tilde{t}_i\right) = \tilde{x}\left(\tilde{t}_i\right)$ 且 $\int_0^{t_i} x\left(t\right) dt = \int_0^{t_i} \tilde{x}(t) dt$ 对于$i = 1, \ldots, M$。此外，我们可以调整 w 的值以获得 $M = M^{\star}$，从而

调整后的计数过程 $x(t)$ 与原始计数过程 $x^{\star}(t)$ 的跳跃次数相同。

调整时间序列数据后，我们可以使用 $\beta_0 = 0$ 拟合 Cox 过程模型，并采用与缩放积分时相同的 w 值。我们使用多项式 $\gamma(t)$ 函数，并为系数设定弱独立正态先验；然后使用独立正态提议分布的 Metropolis 抽样器，获取系数的后验样本。最后，为了说明起见，我们使用每个系数的后验均值来估计 $\gamma(t)$ 函数。

我们对从原始数据集中选取的三个不同矩形区域重复上述步骤，并获得每个时间序列的拟合多项式 $\gamma(t)$，将其作为平均强度函数；见图 2。



(a) 4 阶多项式 $\gamma(t)$　　(b) 4 阶多项式 $\gamma(t)$　　(c) 线性 $\gamma(t)$

图 2：真实数据示例：这三幅图分别是从原始时空数据集中选取的三个不同矩形区域的时间序列数据。红色曲线是拟合的 $\gamma(t)$ 函数的积分。

7 讨论

在本文中，我们构建了一种非齐次计数过程。具体而言，我们采用了一种具有充分理论依据的方法：首先为非齐次泊松过程建立一个潜在强度过程，然后通过对其积分将其消去，从而获得边缘过程。简而言之，这就是一种边缘化的 Cox 过程。因此，无需进行近似，也无需对潜在过程进行采样，即可直接计算似然。尽管本文采用低维多项式来展示这一方法，但将其推广到更一般类型的函数是相当.....

这是可行且简单的；只需确保函数族在 $[0, T]$ 上可积。未来的工作将探讨如何将边缘化扩展到多维情形。

参考文献

Bayisa, F. L., ̊Adahl, M., Ryd´en, P. 和 Cronie, O.（2020）。利用时空对数高斯 Cox 过程对瑞典北部救护车呼叫进行大规模建模与预测。《空间统计》，39：100471。

Billingsley, P.（1999）。《概率测度的收敛》。John Wiley & Sons, Inc，纽约。

Brix, A. 和 Diggle, P. J. (2001). 对数高斯 Cox 过程的时空预测。英国皇家统计学会杂志：*B*辑（统计方法学），63(4):823–841.

Dassios, A.、Jang, J. 和 Zhao, H.（2015）。一种采用更新型噪声 Cox 过程的风险模型。《保险：数学与经济学》，65：55–65。Diggle, P.、Moraga, P.、Rowlingson, B. 和 Taylor, B.（2013）。空间与时空对数高斯 Cox 过程：拓展地统计学范式。《统计科学》，4：542–563。

Diggle, P., Rowlingson, B., and Su, T.-l. (2005). 用于在线时空疾病监测的点过程方法。《环境计量学》：国际环境计量学会官方期刊，16(5):423–434。

Dollard, J. D. 和 Friedman, C. N.（1984）。《乘积积分及其在微分方程中的应用》。剑桥大学出版社。Jalilian, A.、Guan, Y.、Mateu, J. 和 Waagepetersen, R.（2015）。多元乘积噪声 Cox 点过程模型。《生物统计学》，71(4):1022–1033。
Jang, J. 和 Oh, R.（2021）。关于泊松过程、Cox 过程、Hawkes 过程、散粒噪声泊松过程、动态传染过程及其复合过程的综述。《精算科学年刊》，15:623–644。
Kingman, J.（1992）。《泊松过程》。《牛津概率学研究》。Clarendon 出版社。
Moller, J.（2003）。噪声过程。《应用概率进展》，35:614–640。

Møller, J. 和 Torrisi, G. L.（2005）。广义散粒噪声 Cox 过程。《应用概率进展》，37(1):48–74。Moller, J. 和 Vihrs, N.（2022）。行列式散粒噪声 Cox 过程。*Stat*，11:e502。Møller, J.、Waagepetersen, R. 和 Syversveen, A.（1998）。对数高斯 Cox 过程。《斯堪的纳维亚统计学杂志》，25:451–482。Neyman, J. 和 Scott, E.（1958）。宇宙学问题的统计方法。《英国皇家统计学会杂志》*B*辑，20:1–29。Schnoerr, D.、Grima, R. 和 Sanguinetti, G.（2016）。随机反应–扩散过程的 Cox 过程表示与推断。《自然通讯》，7(1):11729。Shirota, S. 和 Gelfand, A.（2016）。利用近似边际后验对数高斯 Cox 过程进行推断。*ArXiv:1611.10359v1*。Taylor, B. M.、Davies, T. M.、Rowlingson, B. S. 和 Diggle, P. J.（2013）。lgcp：用于空间和时空对数高斯 Cox 过程推断的 R 软件包。《统计软件杂志》，52:1–40。Teng, M.、Nathoo, F. 和 Johnson, T. D.（2017）。对数高斯 Cox 过程的贝叶斯计算：方法比较分析。《统计计算与模拟杂志》，87(11):2227–2252。Thomas, M.（1949）。泊松二项极限的推广。《生物计量学》，36:18–25。Wolpert, R. L. 和 Ickstadt, K.（1998）。用于空间统计的泊松/伽马随机场模型。《生物计量学》，85(2):251–267。Wu, S.、Muller, H. 和 Zhang, Z.（2013）。稀有事件点过程的函数型数据分析。《中华统计学志》，23:1–23。

附录

定理 1 的证明。为了通过归纳法证明该定理，我们首先证明它在 $k = 2$ 处的基础情形成立，然后证明如果它在 k 处成立，那么它在 $k + 1$ 处也成立。

首先，对于 $k = 2$ 处的基础情形，我们已在第 4.1 节中推导出其结果，即

$$s_2(y_{n-2}) = \exp(-\beta_{n-2}h)\ (1 - \lambda_{n-2}h),$$

它符合公式（3）给出的一般表达式，其中 $m = 0$ 和 $c_0^{(0,\,n)}{}_{,\,2} = 1$。

接下来，为了证明归纳步骤，假设存在一个 $2 \le k < n$，使得$s_k(y_{n-k})$ 满足一般表达式。然后，我们需要根据 $s_{k+1}(y_{n-k-1})$ 处是否发生跳跃，讨论两种不同的情况。

当在 x_{n-k+1} 处没有发生跳跃时，也就是说，当 $t = (n - k + 1)h$ 不是跳跃时刻，即 $n - k + 1 \neq r_{m+1} + 2$ 时，我们有 $x_{n-k+1} = x_{n-k}$ 并且

$$\begin{aligned} s_{k+1}(y_{n-k-1}) &= \sum_{y_{n-k}} p(x_{n-k+1} \mid x_{n-k},\, y_{n-k})\, \mathrm{P}(y_{n-k} \mid y_{n-k-1})\, s_k(y_{n-k}) \\ &= \sum_{j=0}^{m} c_{j,\,k}^{(m,n)}\, w^j \left\{\beta_{n-k-1}^{m-j}(1 - \gamma_{n-k-1}h) + (\beta_{n-k-1} + w)^{m-j} e^{-kwh}\gamma_{n-k-1}h\right\} \\ &\quad \times \exp(-k\beta_{n-k-1}h) \prod_{i=n-k}^{n-2} (1 - \lambda_i h) \end{aligned}$$

当在 x_{n-k+1} 处发生一次跳跃时，也就是说，当 $t = (n - k + 1)h$ 是一个跳跃时间点，即 $n - k + 1 = r_{m+1} + 2$ 时，我们有 $x_{n-k+1} = x_{n-k} + 1$ 和

$$\begin{aligned} &s_{k+1}(y_{n-k-1}) \\ &= \sum_{j=0}^{m} c_{j,\,k}^{(m,n)}\, w^j \left\{\beta_{n-k-1}^{m+1-j}(1 - \gamma_{n-k-1}h) + (\beta_{n-k-1} + w)^{m+1-j} e^{-kwh}\gamma_{n-k-1}h\right\} \\ &\quad \times \exp(-k\beta_{n-k-1}h) \prod_{i=n-k}^{n-2} (1 - \lambda_i h). \end{aligned}$$

为更清楚地展示以下代数步骤，这里省略下标，

$$\begin{aligned} &\beta^i\ (1 - \gamma h)\ +\ (\beta + w)^i\ e^{-kwh}\ \gamma h \\ =&\beta^i \left\{1 - \left(1 - e^{-kwh}\right)\gamma h\right\} + \left\{(\beta + w)^i - \beta^i\right\}\ e^{-kwh}\ \gamma h \\ =&\left[\beta^i + \left\{(\beta + w)^i - \beta^i\right\} e^{-kwh}\ \gamma h\right] \left\{1 - \left(1 - e^{-kwh}\right)\gamma h\right\} + o(h) \\ =&\left[\beta^i + \left\{(\beta + w)^i - \beta^i\right\}\ e^{-kwh}\ \gamma h\right] (1 - \lambda h). \end{aligned}$$

h^2 项可以视为一个 $o(h)$，当我们设置 $n \to \infty$时它将变为 0，因此本节其余部分可以省略这类 $o(h)$项。定义 $\alpha_i := e^{-(n-i-1)wh}\ \gamma_i$，并将各项合并起来，我们可以得到

$$s_{k+1}(y_{n-k-1}) = \begin{cases} \sum_{j=0}^{m} c_{j,\ k}^{(m,n)}\ w^j \left[\beta^{m-j} + \{(\beta+w)^{m-j} - \beta^{m-j}\}\alpha_{n-k-1}h\right] \\ \times \exp(-k\beta h) \prod_{i=n-k-1}^{n-2}(1-\lambda_i h) \qquad \text{for } x_{n-k+1} = x_{n-k} \\ \\ \sum_{j=0}^{m} c_{j,\ k}^{(m,n)}\ w^j \left[\beta^{m+1-j} + \{(\beta+w)^{m+1-j} - \beta^{m+1-j}\}\alpha_{n-k-1}h\right] \\ \times \exp(-k\beta h) \prod_{i=n-k-1}^{n-2}(1-\lambda_i h) \qquad \text{for } x_{n-k+1} = x_{n-k}+1, \end{cases}$$

其中，β 的下标（由 $n-k-1$ 给出）被省略了。我们可以看到，确实存在“$e^{-k\beta h}\prod(1-\lambda_i h)$”这一部分，它遵循公式（3）给出的一般表达式，因此现在我们只关注 $s_{k+1}(y_{n-k-1})$ 的多项式部分。

When 在 x_{n-k+1} 处没有跳跃， $s_{k+1}(y_{n-k-1})$ 的多项式部分由 b 给出 y

$$\begin{aligned} &\sum_{i=0}^{m} c_{i,\ k}^{(m,n)}\ w^i \left[\beta^{m-i} + \{(\beta+w)^{m-i} - \beta^{m-i}\}\alpha h\right] \\ =&\sum_{i=0}^{m} c_{i,\ k}^{(m,n)}\ w^i \left[\beta^{m-i} + \left\{\sum_{l=1}^{m-i} \binom{m-i}{l} w^l \beta^{m-i-l}\right\}\alpha h\right] \\ =&\sum_{j=0}^{m} c_{j,\ k+1}^{(m,n)} w^j \beta^{m-j}. \end{aligned}$$

通过匹配系数，我们可以得到

$$c_{j,\ k+1}^{(m,n)} = \begin{cases} c_{0,\ k}^{(m,n)} & \text{for } j = 0 \\ \sum_{i=0}^{j-1} c_{i,\ k}^{(m,n)} \binom{m-i}{j-i} \alpha_{n-k-1}h + c_{j,\ k}^{(m,n)} & \text{for } j = 1, 2, \ldots m, \end{cases} \tag{6}$$

适用于 $x_{n-k+1} = x_{n-k}$。

当在 x_{n-k+1}处发生一次跳跃时， $s_{k+1}(y_{n-k-1})$的多项式部分由下式给出。

$$\begin{aligned} &\sum_{i=0}^{m} c_{i,\ k}^{(m,n)}\ w^i \left[\beta^{m+1-i} + \{(\beta+w)^{m+1-i} - \beta^{m+1-i}\}\alpha h\right] \\ =&\sum_{j=0}^{m+1} c_{j,\ k+1}^{(m+1,n)}\ w^j\ \beta^{m+1-j}. \end{aligned}$$

通过匹配系数，我们可以得到

$$c_{j,\ k+1}^{(m+1,n)} = \begin{cases} c_{0,\ k}^{(m,n)} & \text{for } j = 0 \\ \sum_{i=0}^{j-1} c_{i,\ k}^{(m,n)} \binom{m+1-i}{j-i} \alpha_{n-k-1}h + c_{j,\ k}^{(m,n)} & \text{for } j = 1, 2, \ldots m \\ \sum_{i=0}^{m} c_{i,\ k}^{(m,n)}\ \alpha_{n-k-1}h & \text{for } j = m+1, \end{cases} \tag{7}$$

对于 $x_{n-k+1} = x_{n-k} + 1$。因此，在这两种情况下， $s_{k+1}(y_{n-k-1})$ 均遵循定理 1 给出的一般表达式，从而完成证明。□

定理 2 的证明。为了通过归纳法证明该定理，我们首先证明它在 $k = n - r_m$ 处的基础情形成立，然后证明如果它在 k 处成立，那么它在 $k+1$ 处也成立。

在此证明中，我们将省略 $j = 0$ 的情形，因为该情形显然易于证明，并定义 $c^{(m-1,n)}_{m\ k} := 0$ 以合并 $j = 1, \ldots, m-1$ 和 $j = m$ 的情形。

首先，我们需要证明在 $k = n - r_m$ 处的基础情形。这里在x_{r_m+2} 处有一个跳跃，这是求和中计入的第 m 个跳跃，因此我们可以使用方程 (7) 得到，

$$\begin{aligned} c^{(m,n)}_{j,\ n-r_m} &= \sum_{i=0}^{j-1} c^{(m-1,n)}_{i,\ n-r_m-1} \binom{m-i}{j-i} \alpha_{r_m} h + c^{(m-1,n)}_{j,\ n-r_m-1} \\ &= \sum_{i=0}^{j-1} c^{(m-1,n)}_{i,\ n-r_m-1} \left\{ \binom{m-i-1}{j-i-1} + \binom{m-i-1}{j-i} \right\} \alpha_{r_m} h + c^{(m-1,n)}_{j,\ n-r_m-1}, \end{aligned}$$

对于 $j = 1, \ldots m$。

然后，根据方程（6），我们可以将上述方程中的 $c^{(m-1,n)}_{j\ ,\ n-r_m-1}$ 替换为

$$c^{(m-1,n)}_{j,\ n-r_m-1} = c^{(m-1,n)}_{j,\ n-r_m} - \sum_{i=0}^{j-1} c^{(m-1,n)}_{i,\ n-r_m-1} \binom{m-1-i}{j-i} \alpha_{r_m} h.$$

$\left\{ c^{(m-1,n)}_{i\ ,\ n-r_m-1} \right\}_{i=0}^{j-1}$ 可以通过添加一项 $o(h)$，直接替换为 $\left\{ c^{(m-1,n)}_{i\ ,\ n-r_m} \right\}_{i=0}^{j-1}$。因此，我们可以得到

$$c^{(m,n)}_{j,\ n-r_m} = \sum_{i=0}^{j-1} c^{(m-1,n)}_{i,\ n-r_m} \binom{m-i-1}{j-i-1} \alpha_{r_m} h + c^{(m-1,n)}_{j,\ n-r_m}, \quad \text{for } j = 1, 2, \ldots m,$$

这就完成了 $k = n - r_m$ 处基础情形的证明。

接下来，我们需要证明归纳步骤：当方程 (4) 对 $\left\{ c^{(m,n)}_{j\ ,\ k} \right\}_{j=0}^{m}$ 成立时，它对 $\left\{ c^{(m,n)}_{j\ ,\ k+1} \right\}_{j=0}^{m}$ 也成立。

现在，假设我们有一个 $n - r_m \le k \le n$，使得方程 (4) 成立。根据方程（6），我们可以得到

$$c^{(m,n)}_{j,\ k+1} = \sum_{i=0}^{j-1} c^{(m,n)}_{i,\ k} \binom{m-i}{j-i} \alpha_{n-k-1} h + c^{(m,n)}_{j,\ k}, \quad \text{for } j = 1, 2, \ldots, m.$$

然后，由于我们假设方程 (4) 对于 k成立，因此可以利用方程 (4) 将上述方程中的 $\left\{c_i^{(m,n)}{}_{,\ k}\right\}_{i=0}^{j}$ 代入，

$$\begin{aligned}
c_{j,\ k+1}^{(m,n)} =& \sum_{i=1}^{j-1}\left\{\sum_{l=0}^{i-1} c_{l,\ k}^{(m-1,n)}\binom{m-l-1}{i-l-1}\sum_{l=n-k}^{r_m}\alpha_l h + c_{i,\ k}^{(m-1,n)}\right\}\binom{m-i}{j-i}\alpha_{n-k-1}h \\
&+ c_{0,\ k}^{(m-1,n)}\binom{m}{j}\alpha_{n-k-1}h + \sum_{i=0}^{j-1} c_{i,\ k}^{(m-1,n)}\binom{m-i-1}{j-i-1}\sum_{i=n-k}^{r_m}\alpha_i h + c_{j,\ k}^{(m-1,n)} \\
=& \sum_{i=1}^{j-1}\left\{\sum_{l=0}^{i-1} c_{l,\ k}^{(m-1,n)}\binom{m-l-1}{i-l-1}\right\}\binom{m-i}{j-i}\alpha_{n-k-1}h\sum_{i=n-k-1}^{r_m}\alpha_i h + o(h) \\
&+ \sum_{i=0}^{j-1} c_{i,\ k}^{(m-1,n)}\left\{\binom{m-i-1}{j-i-1} + \binom{m-i-1}{j-i}\right\}\alpha_{n-k-1}h \\
&+ \sum_{i=0}^{j-1} c_{i,\ k}^{(m-1,n)}\binom{m-i-1}{j-i-1}\sum_{i=n-k}^{r_m}\alpha_i h + c_{j,\ k}^{(m-1,n)}.
\end{aligned}$$

然后，我们使用方程（6）将上述方程中的 $c_{j\ \ k}^{(m-1,n)}$，替换为，

$$c_{j,\ k}^{(m-1,n)} = c_{j,\ k+1}^{(m-1,n)} - \sum_{i=0}^{j-1} c_{i,\ k}^{(m-1,n)}\binom{m-1-i}{j-i}\alpha_{n-k-1}h,$$

我们将得到

$$\begin{aligned}
c_{j,\ k+1}^{(m,n)} =& \sum_{i=1}^{j-1}\left\{\sum_{l=0}^{i-1} c_{l,\ k}^{(m-1,n)}\binom{m-l-1}{i-l-1}\right\}\binom{m-i}{j-i}\alpha_{n-k-1}h\sum_{i=n-k}^{r_m}\alpha_i h \\
&+ \sum_{i=0}^{j-1} c_{i,\ k}^{(m-1,n)}\binom{m-i-1}{j-i-1}\sum_{i=n-k-1}^{r_m}\alpha_i h + c_{j,\ k+1}^{(m-1,n)}.
\end{aligned}$$

然后，我们再次使用方程（6）代入上述方程中的 $\left\{c_i^{(m-1,n)}{}_{,\ k}\right\}_{i=0}^{j-1}$

并获取

$$
\begin{aligned}
c_{j,\ k+1}^{(m,n)} =& \sum_{i=1}^{j-1}\left\{\sum_{l=0}^{i-1} c_{l,\ k}^{(m-1,n)}\binom{m-l-1}{i-l-1}\right\}\binom{m-i}{j-i}\alpha_{n-k-1}h\sum_{i=n-k-1}^{r_m}\alpha_i h \\
&+\sum_{i=1}^{j-1}\left\{c_{i,\ k+1}^{(m-1,n)}-\sum_{l=0}^{i-1} c_{l,\ k}^{(m-1,n)}\binom{m-l-1}{i-l}\alpha_{n-k-1}h\right\}\binom{m-i-1}{j-i-1}\sum_{i=n-k-1}^{r_m}\alpha_i h \\
&+c_{0,\ k+1}^{(m-1,n)}\binom{m-1}{j-1}\sum_{i=n-k-1}^{r_m}\alpha_i h + c_{j,\ k+1}^{(m-1,n)} \\
=& \sum_{l=0}^{j-2} c_{l,\ k}^{(m-1,n)}\left\{\sum_{i=l+1}^{j-1}\binom{m-l-1}{i-l-1}\binom{m-i}{j-i}\right\}\alpha_{n-k-1}h\sum_{i=n-k-1}^{r_m}\alpha_i h \\
&-\sum_{l=0}^{j-2} c_{l,\ k}^{(m-1,n)}\left\{\sum_{i=l+2}^{j}\binom{m-l-1}{i-l-1}\binom{m-i}{j-i}\right\}\alpha_{n-k-1}h\sum_{i=n-k-1}^{r_m}\alpha_i h \\
&+\sum_{i=0}^{j-1} c_{i,\ k+1}^{(m-1,n)}\binom{m-i-1}{j-i-1}\sum_{i=n-k-1}^{r_m}\alpha_i h + c_{j,\ k+1}^{(m-1,n)} \\
=& \sum_{i=0}^{j-1} c_{i,\ k+1}^{(m-1,n)}\binom{m-i-1}{j-i-1}\sum_{i=n-k-1}^{r_m}\alpha_i h + c_{j,\ k+1}^{(m-1,n)},
\end{aligned}
$$

至此，证明完毕。 □

定理 4 的证明. 根据第 4.2 节和第 5.3 节，我们有

$$
\begin{aligned}
p_X(x) =& \mathrm{E}_Y\left[p_{X|Y}(x\mid Y)\right] = \lim_{n\to\infty}\mathrm{E}_{Y^{(n)}}\left[p_{X|Y}\left(x^{(n)}\mid Y^{(n)}\right)\right] \\
&= \lim_{n\to\infty}\left(\sum_{j=0}^{M} c_{j,\ n}^{(M,n)}\ w^j\beta_0^{M-j}\right)\exp(-n\beta_0 h)\prod_{i=0}^{n-2}(1-\lambda_i h) \\
&= \left(\sum_{j=0}^{M}\lim_{n\to\infty} c_{j,\ n}^{(M,n)}\ w^j\beta_0^{M-j}\right)\exp(-\beta_0 T)\lim_{n\to\infty}\prod_{i=0}^{n-2}(1-\lambda_i h).
\end{aligned}
$$

根据乘积积分（Dollard and Friedman，1984），我们有

$$
\begin{aligned}
\lim_{n\to\infty}\prod_{i=0}^{n-2}(1-\lambda_i h) &= \lim_{n\to\infty}\prod_{i=0}^{n-2}\left\{1-\left(1-e^{-(n-i-1)wh}\right)\gamma_i h\right\} \\
&= \lim_{h\to 0}\prod_{t=0}^{T-2h}\left\{1-\left(1-e^{-w(T-t-h)}\right)\gamma(t)h\right\} \\
&= -\int_0^T\left(1-e^{-w(T-t)}\right)\gamma(t)\,dt
\end{aligned}
$$

现在，我们需要证明，当 $n \to \infty$ 时，对于任意 $m \geq 0$ 和 $j = 0, \ldots, m$，方程 (4) 中给出的 $c_{j\ n}^{(m,n)}$，收敛于方程 (5) 中给出的 $c_j^{(m)}$；这可以通过归纳法证明。对于 $m = 0$ 处的基础情形，显然 $\lim_{n\to\infty} c_{0\ ,\ n}^{(0,n)} = c_0^{(0)} = 1$。

对于归纳步骤，我们需要证明 $\lim_{n\to\infty} c_{j\ ,\ n}^{(m-1,n)} = c_j^{(m-1,n)}$ 蕴含 $\lim_{n\to\infty} c_{j\ ,\ n}^{(m,n)} = c_j^{(m)}$。同样，这里我们将省略 $j = 0$ 的情况，因为该情况很容易证明，并定义 $c_m^{(m-1)} := 0$ 以合并 $j = 1, \ldots, m-1$ 和 $j = m$ 的情况。根据方程 (4)，我们有

$$\begin{aligned}\lim_{n\to\infty} c_{j,\ n}^{(m,n)} &= \lim_{n\to\infty} \left\{ \sum_{i=0}^{j-1} c_{i,\ n}^{(m-1,n)} \binom{m-i-1}{j-i-1} \sum_{i=0}^{r_m} \alpha_i h + c_{j,\ n}^{(m-1,n)} \right\} \\ &= \sum_{i=0}^{j-1} \lim_{n\to\infty} c_{i,\ n}^{(m-1,n)} \binom{m-i-1}{j-i-1} \left\{ \lim_{n\to\infty} \sum_{i=0}^{r_m} \alpha_i h \right\} + \lim_{n\to\infty} c_{j,\ n}^{(m-1,n)}.\end{aligned}$$

根据第 5.2 节，我们有 $t_i^{(n)} \to t_i$ as $n \to \infty$，因此根据黎曼积分，我们可以得到

$$\lim_{n\to\infty} \sum_{i=0}^{r_m} \alpha_i h = \lim_{h\to 0} \sum_{t=0}^{t_m - 2h} e^{-(T-t-h)w}\, \gamma(t)\, h = \int_0^{t_m} e^{-(T-t)w}\, \gamma(t)\, dt.$$

因此，如果我们有 $\lim_{n\to\infty} c_{j\ ,\ n}^{(m-1,n)} = c_j^{(m-1)}$ 对于 $j = 0, \ldots, m-1$，那么我们也将有 $\lim_{n\to\infty} c_{j\ ,\ n}^{(m,n)} = c_j^{(m)}$ 对于 $j = 0, \ldots, m$，从而完成证明。□