

带有多终止事件的复发事件数据可加可乘比率模型

张铭鸿¹, 韦程东^{1*}, 岑泰林²

¹南宁师范大学, 数学与统计学院, 广西 南宁

²贺州学院, 教师教育学院(师范学院), 广西 贺州

收稿日期: 2025年3月25日; 录用日期: 2025年4月25日; 发布日期: 2025年5月9日

摘要

在已有的联合建模方法基础上, 深入分析了具有两类终止事件的复发事件数据。研究中采用含脆弱变量的可加可乘比率模型来描述复发事件过程, 同时运用含脆弱变量的Cox型模型刻画终止事件过程。两类事件通过共同的脆弱变量建立联系, 从而揭示终止事件与复发事件之间的关联性。进一步提出了联合模型中边缘参数和关联参数的估计方程, 并探讨了估计量的渐进性质。

关键词

复发事件, Cox模型, 脆弱模型, 偏边际比率模型, 可加可乘模型

Additive and Multiplicative Ratio Modeling of Recurrent Event Data with Multiple Termination Events

Minghong Zhang¹, Chengdong Wei^{1*}, Tailin Cen²

¹School of Mathematical and Statistics, Nanning Normal University, Nanning Guangxi

²School of Teacher Education (Normal College), Hezhou University, Hezhou Guangxi

Received: Mar. 25th, 2025; accepted: Apr. 25th, 2025; published: May 9th, 2025

Abstract

Recurrent event data with two types of termination events were analyzed in depth on the basis of existing joint modeling approaches. In the study, an additive and multiplicative ratio model with

*通讯作者。

文章引用: 张铭鸿, 韦程东, 岑泰林. 带有多终止事件的复发事件数据可加可乘比率模型[J]. 理论数学, 2025, 15(5): 1-13. DOI: 10.12677/pm.2025.155148

fragile variables is used to describe the recurrent event process, while a Cox-type model with fragile variables is applied to portray the termination event process. The two types of events are linked through a common vulnerability variable, thus revealing the correlation between termination and recurrence events. The estimation equations for the edge and association parameters in the joint model are further proposed and the asymptotic nature of the estimates is explored.

Keywords

Recurrent Events, Cox Model, Fragility Model, Skewed Marginal Ratio Model, Additive and Multiplicative Model

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

复发事件数据在多个学科领域,如医学、经济学和生物学等,是一种常见的特殊数据类型。这种数据类型下,研究中的个体可能在特定时期内多次经历同一事件,具体表现包括:肿瘤再次出现、多次入院治疗、重复感染、连续服药以及周期性经济下滑等现象[1]。基于复发事件的强度或速率函数,已有多方法用于分析复发事件数据[2]-[4],Cook 和 Lawless (2007) [1]等人对复发事件方法进行了出色的综述。

在许多应用中,复发事件个体可能会出现死亡、治愈、接受其他治疗等终止事件,从而停止随访。通常终止事件与研究复发事件往往存在着紧密联系。分析带有复发事件的终止事件的方法可以分为强度模型、边际模型和偏边际模型[9]。强度模型通过引入一个共同的脆弱性变量,描述复发事件与终止事件之间的相互依赖。这些模型假设事件强度由可观测的协变量和未观测的脆弱变量共同决定,从而提供了理解事件关系的新视角[6]。作为强度模型的替代方案,一些作者提出了边际模型[7]-[9],边际模型通过存活者与死亡个体比例的平均值来定义复发事件的发生概率,具有假设灵活、对泊松分布偏差稳健和解释直观等优点,适用于复杂关系情境。然而,其简化计算也导致模型参数的解释力减弱[18]。

偏边际模型主要研究存活个体的事件复发率,并可通过脆弱变量来描述复发和终止事件间的关联系[2][5][6][11]-[13]。例如, Cook 和 Lawless (1997) [11]提出了个体在特定时间点的复发事件平均值和复发率概念。Ye 等(2007) [12]构建了一个基于脆弱变量的联合半参数建模框架,在保持边际模型特性的同时,有效刻画了终止事件与复发事件的相依关系,即仅以协变量和脆弱变量为条件,而不考虑过程的先前历史。

Kalbfleisch 等(2013) [10]提出了一种基于估计方程的新方法,用于估计联合脆弱性模型中的边际参数和关联参数。这种方法具有三个主要优点:无需泊松过程假设即可进行参数估计,提高了鲁棒性;直接估计共同脆弱性分布;能够估计事件间的依赖程度。然而实际应用中,加法模型作为乘法模型的重要替代选择,也具有其独特价值。Pan 和 Schaubel (2009) [5]采用加性模型来估计条件复发事件率,并使用比例风险模型来估计终止事件风险。这种方法可以评估治疗对幸存者生存率和复发事件率的影响,有助于解释研究结果间的差异。Zeng 和 Cai (2010) [2]以及 Sun 和 Kang (2013) [13]分别研究了加法率模型和加法、乘法率模型,其乘法率模型体现在使用比例风险率模型对终止事件进行建模,由于基线率函数依赖于非参数的脆弱变量,因此复发和终止事件之间关联的脆弱变量被视为一干扰项。Chen 等(2016) [14]考虑了部分 Aalen 加性模型,该模型对复发事件条件比率和终止事件风险率均采用了乘性脆弱变量。但他

们并没有建立所提估计值的大样本性质, Sun 等(2017) [15]提出了一种新模型, 利用共享脆弱变量同时分析复发和终止事件, 阐释二者的关联性, 并给出了大样本估计特性. Lin 和 Ying(1995) [16]则指出, 实际应用中, 协变量可能同时呈现加性和乘性效应. 基于这一观点, Sun 和 Kang (2013) [13]开发了一个复发事件的可加可乘模型, 引入非参数形式的脆弱变量. 然而, 这种处理方式使得模型在解释复发事件和终止事件相关性方面存在冗余. 孙琴等(2019) [17]构建了一个联合模型: 复发事件采用含脆弱变量的可加和可乘结构, 而终止事件则使用带脆弱变量的乘法模型. 但因其复发和终止事件是唯一的, 不能应用于更复杂的情况, 在实际情况中, 一个复发事件往往伴随着多个终止事件, 就如上文所述, 而研究者也十分关心它们之间的联系.

在孙琴等(2019) [17]这篇文章的展望提出的新的复发事件与终止事件的联合分析模型的基础上, 继续对带有两种终止事件相关的脆弱变量的复发事件加法乘法比率模型进行分析, 进一步提出了联合模型中边缘参数和关联参数的估计方程, 最后再给出估计量的参数估计及其证明.

2. 模型假设和参数估计

定义 $\tilde{N}^R(t)$ 为 $(0, t]$ 时间窗内观察到的复发事件累积计数过程, $J(t) = (Z(t)^T, W(t)^T)^T$ 为 $p+q$ 维外生协变量 [18], 其中 $Z(t)$ 为 $p \times 1$ 维协变量, $W(t)$ 为 $q \times 1$ 维协变量. 定义 D_i 和 C_i 分别代表第 i 个观察对象的终止事件时间和删失时间. 其中终止事件会停止未来复发事件的发生, 即当 $t > D$ 时, $d\tilde{N}^R(t) = 0$. Y_i 表示终止事件类型, 个体历经 k 种终止事件, 则 $Y_i = k$, k 取值为 1, 2. 记 $T = (C \wedge D)$, $\Delta(t) = I(T \geq t)$, 其中 $I(\cdot)$ 是示性函数, 由于数据删失, $\tilde{N}^R(t)$ 未能完全观测, 令 $N^R(t) = \tilde{N}^R(t \wedge T)$ 表示复发事件在区间 $(0, t]$ 上实际数量, 同理, 令 $N^{Dk}(t) = \tilde{N}^{Dk}(t \wedge T)$ 表示终止事件在区间 $(0, t]$ 上实际数量, 其中 $\tilde{N}^{Dk}(t) = I(D \leq t)$, $k = 1, 2$. 考虑包含 n 个独立同分布个体的样本, 其观测数据形式为 $\{N_i^R(t), N_i^{D1}(t), N_i^{D2}(t), T_i, \Delta_i(t), J_i(t), 0 \leq t \leq T_i, i = 1, \dots, n\}$.

令 v_1 和 v_2 是两个非负不可观测的脆弱变量且分别独立于 $J(t)$. 根据 Ye [12] 和 Kalbfleisch [10]. 我们考虑一个给定 $J(t)$, $D = s$ 及脆弱变量 v 的复发事件的偏边际比率模型

$$d\Lambda_R(t|v) = P\{d\tilde{N}^R(t) = 1 | J(t), D = s, v\}, s \geq t \quad (1)$$

特别是, $d\Lambda_R(t|v)$ 可能会受协变量 $J(t)$ 和脆弱变量 v 影响, 不过与终止事件 $D = s \geq t$ 无关. 说明当协变量给定时, v 对复发事件与终止事件进行联合建模.

则表达式(1)可以表示成

$$\Lambda_R(t|v) = P\{\tilde{N}^R(t) = 1 | J(t), D \geq t, v\} \quad (2)$$

这表明给定 $J(t)$ 和 v , $d\Lambda_R(t|v)$ 指定了存活到时间 t 的受试者中复发事件的边际比率. 为了方便分析, 假设存在两种终止事件, 因此考虑如下的可加可乘比率模型

$$d\Lambda_R(t|v) = v_1 v_2 \left[\exp\{\gamma^T Z(t)\} d\Lambda_0^R(t) + \beta^T W(t) dt \right] \quad (3)$$

其中回归参数 γ 和 β 分别是 $p \times 1$ 维向量和 $q \times 1$ 维向量, $d\Lambda_0^R(t)$ 是一个未指定的基线比率函数.

令 $d\Lambda_{Dk}(t|v) = P\{d\tilde{N}^{Dk}(t) = 1 | J(t), D \geq t, v\}$ 为给定协变量 $J(t)$ 和脆弱变量 v 下的风险函数. 我们指定以下对于终止事件的乘法风险率模型:

$$d\Lambda_{Dk}(t|v) = v_k \exp\{\alpha^T J(t)\} d\Lambda_{0k}^D(t), k = 1, 2 \quad (4)$$

其中回归参数 α 是 $(p+q) \times 1$ 维度向量, $d\Lambda_{0k}^D(t)$ 是一个未指定的基线风险函数, 因而两种终止事件的风

险函数可以用(4)式表达, 为了便于表示, 模型(3)和(4)的 $J(t)$ 假定为同一组协变量, 所提出的估计方法可进一步推广, 适用于处理两个模型中存在差异的协变量集合。

此外, 类似 Ye (2007) [12], 假设脆弱变量 v_1 和 v_2 分别服从伽马分布 $G(1, \theta_1)$ 和 $G(1, \theta_2)$, 为了可识别这个模型, 固定 $E(v_1)=1$, $E(v_2)=1$, 且脆弱变量分布假设的稳健性已在孙琴[17]和 Qu [15]中得到验证。我们在下文假设给定 $J(\cdot)$, 删失时间 C 与 $\{\tilde{N}^R(\cdot), \tilde{N}^{D1}(\cdot), \tilde{N}^{D2}(\cdot), D, v\}$ 独立。

3. 估计方法

令 $\eta = (\gamma^T, \beta^T)^T$ 。注意到, 如果脆弱变量 v_1 和 v_2 已知, 可以利用 Liu [19]以及 Gill [3]的方法分别应用到模型(3)和(4), 从而得到 η 、 α_1 和 α_2 的估计方程。但在实际情况中, v_1 和 v_2 通常是未知的, 这使得直接应用上述方法变得不可行。对此, 通过考虑一个包含 η 、 α_1 和 α_2 的边际模型, 其通过给定 $D \geq t$ 和 $X(t)$ 的条件下, 取模型(3)和(4)的条件期望。假设脆弱变量 v_1 和 v_2 服从伽马分布情况下有:

$$d\Lambda_R(t) = P\{d\tilde{N}^R(t)=1 | J(t), D \geq t\} \quad (5)$$

$$= \psi_1(t)\psi_2(t) \left[\exp\{\gamma^T Z(t) d\Lambda_0^R(t) + \beta^T W(t) dt\} \right]$$

$$P\{d\tilde{N}^{Dk}(t)=1 | D \geq t, J(t)\} = \psi_k(t) \exp\{\alpha_k^T J(t)\} d\Lambda_{0k}^D(t) \quad (6)$$

其中

$$\psi_k(t) = E[v_k | D \geq t, Z(t)] = \left[1 + \theta_k \exp\{\alpha_k^T J(t)\} d\Lambda_{0k}^D(t) \right]^{-1}, k=1, 2$$

$$\psi_3(t) = E[v_1 v_2 | D \geq t, X(t)] = \psi_1(t)\psi_2(t)$$

给其定义

$$dM_i^R(t) = \psi_{3i}(t) dN_i^R(t) - \Delta_i(t) \left[\exp\{\gamma^T Z_i(t)\} + \beta^T W_i(t) dt \right]$$

$$dM_i^{Dk}(t) = \psi_{ki}(t) dN_i^{Dk}(t) - \Delta_i(t) \exp\{\alpha_k^T J_i(t)\} d\Lambda_{0k}^D(t), k=1, 2.$$

其中

$$\psi_{ki}(t) = \left[1 + \theta_k \int_0^t \exp\{\alpha_k^T J_i(t)\} d\Lambda_{0k}^D(t) \right]^{-1}, k=1, 2.$$

在假设下, 模型(3)和(4)可以得出 $M_i^R(t)$ 和 $M_i^{Dk}(t)$ 是零均值随机过程。

当 $\psi_1(t), \psi_2(t)$ 已知时, 相当于知道 $\psi_3(t)$, 可以用以下方程估计 $\Lambda_0^R(t), \Lambda_{0k}^D(t), \eta, \alpha_k$, 其中 $k=1, 2$

$$\sum_{i=1}^n \left[\psi_{3i}(t) dN_i^R(t) - \Delta_i(t) \left\{ \exp\{\gamma^T Z_i(t)\} d\Lambda_0^R(t) + \beta^T W_i(t) dt \right\} \right] = 0, 0 \leq t \leq \tau,$$

$$\sum_{i=1}^n \left[\psi_{ki}(t) dN_i^{Dk}(t) - \Delta_i(t) \exp\{\alpha_k^T J_i(t)\} d\Lambda_{0k}^D(t) \right] = 0, 0 \leq t \leq \tau, k=1, 2,$$

$$\sum_{i=1}^n \int_0^\tau [D_i(t; \eta) - \bar{D}(t; \eta)] \left[\psi_{3i}(t) dN_i^R(t) - \Delta_i(t) \beta^T W_i(t) dt \right] = 0, k=1, 2$$

$$\sum_{i=1}^n \int_0^\tau \{J_i(t) - \bar{J}_k(t)\} \psi_{ki}(t) dN_i^{Dk}(t) = 0, k=1, 2$$

其中 τ 是常数, 使得 $P(T_i \geq \tau) > 0$, 并且

$$\bar{J}_k(t) = \sum_{i=1}^n \Delta_i(t) \exp\{\alpha_k^T J_i(t)\} J_i(t) / \sum_{i=1}^n \Delta_i(t) \exp\{\alpha_k^T J_i(t)\}$$

根据 Lin 和 Ying [16], 我们选取

$$D_i(t; \eta) = \left(Z_i^T(t), W_i^T(t) / \exp\{(\alpha_1 + \alpha_2)^T J_i(t)\} \right)^T$$

定义 $\bar{D}(t; \eta) = \left(\bar{Z}^T(t; \eta), \bar{W}^T(t; \eta) \right)^T$, 其中

$$\bar{Z}(t; \eta) = \sum_{i=1}^n \Delta_i(t) \exp\{\gamma^T Z_i(t)\} Z_i(t) / \sum_{i=1}^n \Delta_i(t) \exp\{\gamma^T Z_i(t)\}$$

$$\bar{W}(t; \eta) = \sum_{i=1}^n \Delta_i(t) W_i(t) / \sum_{i=1}^n \Delta_i(t) \exp\{\gamma^T Z_i(t)\}$$

而 $\psi_k(t)$ 是未知的, 为了估计 θ_k , 令

$$w_{1i}^k(t) = E[N_i^R(t) | dN_i^{Dk}(t) = 1, J_i(t)]$$

和

$$w_{2i}(t) = E[N_i^R(t) | \Delta_i(t) = 1, J_i(t)]$$

在模型假设下, 可得

$$w_{1i}^k(t) = (\theta_k + 1) \psi_{3i}(t)^{-1} \int_0^t \left[\exp\{\gamma^T Z_i(t)\} d\Lambda_0^R(u) + \beta^T W_i(u) du \right]$$

和

$$w_{2i}(t) = \psi_{3i}(t)^{-1} \int_0^t \left[\exp\{\gamma^T Z_i(t)\} d\Lambda_0^R(u) + \beta^T W_i(u) du \right]$$

因此有

$$w_{1i}^k(t) / w_{2i}(t) = \theta_k + 1 \quad (7)$$

当 $k=1, 2$, θ_1 和 θ_2 描述了复发事件和各个终止事件的关联性。

如 Kalbfleisch [10] 所讨论的, 为 θ_k 构造了以下估计方程:

$$\sum_{i=1}^n \int_0^\tau \{N_i^R(t) - (\theta_k + 1) G_k(t) w_{2i}(t)\} dN_i^{Dk}(t) = 0, k=1, 2$$

其中

$$G_k(t) = \sum_{i=1}^n w_{2i}(t)^{-1} \Delta_{ki}^*(t) N_i^R(t) / \sum_{i=1}^n \Delta_{ki}^*(t)$$

$\Delta_{ki}^*(t) = \Delta_i(t) \{1 - N_i^{Dk}(t)\}$ 代表个体 i 在 t 时刻仍存活且未经历终止事件, 并且在之后的某个时间点发生了第 k 类终止事件。

令 $\gamma = (\eta^T, \alpha_1^T, \alpha_2^T, \theta_1, \theta_2, \Lambda_0^R, \Lambda_{01}^D, \Lambda_{02}^D)^T$, 通过解估计方程 $U(\gamma) = (U_1^T, U_{21}^T, U_{22}^T, U_{31}, U_{32}, U_4, U_{51}, U_{52})^T = 0$ 来获得 γ 的估计值 $\hat{\gamma} = (\hat{\eta}^T, \hat{\alpha}_1^T, \hat{\alpha}_2^T, \hat{\theta}_1, \hat{\theta}_2, \hat{\Lambda}_0^R, \hat{\Lambda}_{01}^D, \hat{\Lambda}_{02}^D)^T$, 其中

$$U_1 = \sum_{i=1}^n \int_0^\tau [D_i(t; \eta) - \bar{D}(t; \eta)] [\psi_{1i}(t) \psi_{2i}(t) dN_i^R(t) - \Delta_i(t) \beta^T W_i(t) dt],$$

$$U_{2k} = \sum_{i=1}^n \int_0^\tau \{J_i(t) - \bar{J}_k(t)\} \psi_{ki}(t) dN_i^{Dk}(t), k=1, 2$$

$$U_{3k} = \sum_{i=1}^n \int_0^\tau \{N_i^R(t) - (\theta_k + 1) G_k(t) w_{2i}(t)\} dN_i^{Dk}(t), k=1, 2$$

$$U_4 = \sum_{i=1}^n \left[\psi_{3i}(t) dN_i^R(t) - \Delta_i(t) \left\{ \exp \left\{ \gamma^T Z_i(t) \right\} d\Lambda_0^R(t) + \beta^T W_i(t) dt \right\} \right], 0 \leq t \leq \tau$$

$$U_{5k} = \sum_{i=1}^n \left[\psi_{ki}(t) dN_i^{Dk}(t) - \Delta_i(t) \exp \left\{ \alpha^T J_i(t) \right\} d\Lambda_{0k}^D(t) \right], 0 \leq t \leq \tau, k=1,2$$

由于每个参数的估计式取决于其他参数的估计式，因此可以通过递归过程的方获得估计方程的解。因此，提出以下的迭代算法来求解 $U(\gamma) = 0$ 。

步骤 0 选取初值 $\theta_1^{(0)}, \theta_2^{(0)}, \alpha_1^{(0)}, \alpha_2^{(0)}$ 和 $\Lambda_{01}^{D(0)}(t), \Lambda_{02}^{D(0)}(t)$ 。

步骤 1 令 $\psi_{ki}^{(0)}(t) = \psi_{ki}(t; \theta_k^{(0)}, \alpha_k^{(0)}, \Lambda_{0k}^{D(0)})$ ，由上述方程得

$\psi_{3i}^{(0)}(t) = \psi_{3i}(t; \theta_1^{(0)}, \theta_2^{(0)}, \alpha_1^{(0)}, \alpha_2^{(0)}, \Lambda_{01}^{D(0)}(t), \Lambda_{01}^{D(0)}(t))$ ，把 $\psi_{ki}^{(0)}(t)$ 、 $\psi_{3i}^{(0)}(t)$ 代入 $U_1 = 0, U_{2k} = 0, U_4 = 0$ 和 $U_{5k} = 0$ ，其中 $k=1,2$ ，则求解得到更新估计 $\eta^{(1)}, \alpha_1^{(1)}, \alpha_2^{(1)}, \Lambda_0^{R(1)}(t)$ 和 $\Lambda_{01}^{D(1)}(t), \Lambda_{02}^{D(1)}(t)$ 。

步骤 2 给定 $\eta^{(1)}, \Lambda_0^{R(1)}(t)$ ， $\psi_{3i}^{(0)}(t) = \psi_{3i}(t; \theta_1^{(0)}, \theta_2^{(0)}, \alpha_1^{(1)}, \alpha_2^{(1)}, \Lambda_{01}^{D(1)}(t), \Lambda_{02}^{D(1)}(t))$

解出 $U_{31} = 0, U_{32} = 0$ ，得到估计 $\theta_1^{(1)}$ 和 $\theta_2^{(1)}$ 。

步骤 3 重复步骤 1 和 2，更新估计值，直到收敛。

注意，初始的估计值 $\theta_k^{(0)}, \alpha_k^{(0)}$ 和 $\Lambda_{0k}^{D(0)}(t)$ ， $k=1,2$ 可以有多种的选取方法，为了简单起见，因而选择 $\theta_1^{(0)}=1, \theta_2^{(0)}=1, \alpha_1^{(0)}=0, \alpha_2^{(0)}=0, \Lambda_{01}^{D(0)}(t)$ 和 $\Lambda_{02}^{D(0)}(t)$ 为累计基线风险函数的 Nelson-Aalen 类型估计。对于收敛性，该算法大多数时候会收敛，但偶尔会发生非收敛，具体取决于初始值设置。

令 $\xi_1 = (\eta^T, \alpha_1^T, \theta_1^T)^T$ ， $\xi_2 = (\eta^T, \alpha_2^T, \theta_2^T)^T$ ， $\xi = \xi_1 \cup \xi_2 = (\eta^T, \alpha_1^T, \alpha_2^T, \theta_1^T, \theta_2^T)^T$ ，则 $\hat{\xi}_1 = (\hat{\eta}^T, \hat{\alpha}_1^T, \hat{\theta}_1^T)^T$ ， $\hat{\xi}_2 = (\hat{\eta}^T, \hat{\alpha}_2^T, \hat{\theta}_2^T)^T$ ， $\hat{\xi} = \hat{\xi}_1 \cup \hat{\xi}_2 = (\hat{\eta}^T, \hat{\alpha}_1^T, \hat{\alpha}_2^T, \hat{\theta}_1^T, \hat{\theta}_2^T)^T$ ， ξ_1, ξ_2 和 ξ 的真值分别为 $\xi_{10} = (\eta_0^T, \alpha_{10}^T, \theta_{10}^T)^T$ ， $\xi_{20} = (\eta_0^T, \alpha_{20}^T, \theta_{20}^T)^T$ ， $\xi_0 = \xi_{10} \cup \xi_{20} = (\eta_0^T, \alpha_{10}^T, \alpha_{20}^T, \theta_{10}^T, \theta_{20}^T)^T$ 。下面描述了所提出的估计器的渐近属性。首先考虑 $\hat{\xi}, \hat{\Lambda}_0^R(t), \hat{\Lambda}_{01}^D(t)$ 和 $\hat{\Lambda}_{02}^D(t)$ 的存在性、唯一性和强相合性。

4. 估计量的渐进性质

为了研究所提出的估计的渐近性质，需要以下正则性条件：

(C1) $\{N_i^R(\cdot), N_i^{D1}(\cdot), N_i^{D2}(\cdot), T_i, J_i(\cdot)\}, i=1, \dots, n$ 是独立同分布的。

(C2) $E\{N_i^R(\tau)\}$ 是有界的， $\Pr(T \geq \tau) > 0$ 。

(C3) $Z_i(t)$ 在 $[0, \tau]$ 上几乎处处为有界变差随机过程。

(C4) 存在 η_0 的一个紧集 $\mathcal{B}$ ，满足 $\eta \in \mathcal{B}$ ， $\Gamma(\eta)$ 是非奇异的，其中 $\Gamma(\eta)$ 是 $-\partial \tilde{U}(\eta) / \partial \eta^T$ 的极限值，其中 $\tilde{U}(\eta)$ 的定义在 (A.4) 中给出。

定理 1 $\hat{\xi}, \hat{\Lambda}_0^R(t), \hat{\Lambda}_{01}^D(t)$ 和 $\hat{\Lambda}_{02}^D(t)$ 几乎处处存在且唯一。

证明 令 $\psi_{ki}(t; \Lambda_{0k}^D) = \psi_{ki}(t; \theta_k, \alpha_k, \Lambda_{0k}^D)$ ，其中 $k=1,2$ ，

$\Lambda_0^R(t; \Lambda_0^D) = \Lambda_0^R(t; \alpha_1, \alpha_2, \theta_1, \theta_2, \Lambda_{01}^D, \Lambda_{02}^D)$ 。对任意向量 γ^* 和变量 Z^* ，定义

$$S_{kZ^*}(t; \gamma^*) = \frac{1}{n} \sum_{i=1}^n \Delta_i(t) \exp \left\{ \gamma^{*T} Z_i^*(t) \right\} Z_i^{*\otimes k}(t), k=0,1,$$

其中对于任意向量 a ，有 $a^{\otimes 0} = 1, a^{\otimes 1} = a$ 。令

$$\bar{N}^{Dl}(t) = n^{-1} \sum_{i=1}^n \psi_{li}(t; \tilde{\Lambda}_{0l}^D) N_i^{Dl}(t), \bar{N}^R(t) = n^{-1} \sum_{i=1}^n \psi_{li}(t; \tilde{\Lambda}_{0l}^D) \psi_{2i}(t; \tilde{\Lambda}_{02}^D) N_i^R(t)$$

及 $s_{kz}^*(t; \gamma^*)$ 为 $S_{kz}^*(t; \gamma^*)$ 的极限 ($k=0,1, l=1,2$)。

对给定的 ξ , 令 $\tilde{\Lambda}_0^R(t; \xi)$ 及 $\tilde{\Lambda}_{01}^D(t; \xi_1)$ 和 $\tilde{\Lambda}_{02}^D(t; \xi_2)$ 分别为 $U_4=0$ 及 $U_{51}=0$ 和 $U_{52}=0$ 的解。则

$$\tilde{\Lambda}_0^R(t; \xi) = \frac{1}{n} \sum_{i=1}^n \int_0^t \frac{\psi_{1i}(t; \tilde{\Lambda}_{01}^D) \psi_{2i}(t; \tilde{\Lambda}_{02}^D) dN_i^R(u) - \Delta_i(u) \beta^T W_i(u) du}{S_{0Z}(u; \gamma)}$$

且 $\tilde{\Lambda}_{01}^D(t; \xi_1)$ 和 $\tilde{\Lambda}_{02}^D(t; \xi_2)$ 均满足方程

$$\tilde{\Lambda}_{0k}^D(t; \xi_k) = \frac{1}{n} \sum_{i=1}^n \int_0^t \frac{\psi_{ki}(t; \tilde{\Lambda}_{0k}^D) dN_i^{Dk}(u)}{S_{0X}(u; \alpha_k)} \quad (8)$$

其中 $k=1,2$, (8) 是 Volterra 积分方程, 且存在唯一解。

固定 ξ , 令

$$\Lambda_k^D(t; \xi_k) = \int_0^t \frac{E\{d\bar{N}^{Dk}(u)\}}{S_{0X}(u; \alpha_k)} \quad (9)$$

的解为 $\Lambda_k^D(t; \xi_k)$, 其中 $k=1,2$ 。

这是 Volterra 积分方程且有唯一解, 满足 $\Lambda_1^D(t; \xi_{10}) \equiv \Lambda_{01}^D(t)$, $\Lambda_2^D(t; \xi_{20}) \equiv \Lambda_{02}^D(t)$, 令

$$K_l(t; \xi_l) = \frac{1}{n} \sum_{i=1}^n \int_0^t \exp\{\alpha_l^T J_i(u)\} \frac{\theta_l dN_i^{Dl}(u)}{S_{0X}(u; \alpha_l)}$$

利用(8)和(9)可得

$$\tilde{\Lambda}_{0l}^D(t; \xi_l) - \Lambda_l^D(t; \xi_l) = \int_0^t [\tilde{\Lambda}_{0l}^D(t; \xi_l) - \Lambda_l^D(t; \xi_l)] dK_l(t; \xi_l) + \mathcal{E}_{nl}(t; \xi_l)$$

其中

$$\mathcal{E}_{nl}(t; \xi_l) = \frac{1}{n} \sum_{i=1}^n \int_0^t \frac{dM_i^{Dl}(u)}{S_{0X}(u; \alpha_{l0})}$$

$l=1,2$, 由强大数定律可知, 在 $t \in [0, \tau]$ 和 $\xi \in \mathcal{B}$ 条件下, 几乎必然一致地有 $\mathcal{E}_{nl}(t; \xi_l) \rightarrow 0$ 。

因为上式也是一个 Volterra 积分方程, 求解可得

$$\tilde{\Lambda}_{0l}^D(t; \xi_l) - \Lambda_l^D(t; \xi_l) = \frac{1}{P_{nl}(t)} \int_0^L P_{nl}(u-) d\mathcal{E}_{nl}(t; \xi_l) \quad (10)$$

其中

$$P_{nl}(t) = \prod_{s \leq t} \{1 - dK_l(t; \xi_l)\}$$

$l=1,2$, 是 $\mathcal{E}_{nl}(t; \xi)$ 在 $[0, t]$ 上的乘积积分[1]及 $P_n(u-)$ 表示 $P_n(u)$ 的左连续版本。利用乘积积分的渐进性质[20], 一致强大数定律[21]及 Lin 和 Ying (2011) [9] 得出在 $t \in [0, \tau]$ 和 $\xi \in \mathcal{B}$ 时, $\tilde{\Lambda}_{01}^D(t; \xi_1)$ 和 $\tilde{\Lambda}_{02}^D(t; \xi_2)$ 分别几乎处处收敛到 $\Lambda_1^D(t; \xi_1)$ 和 $\Lambda_2^D(t; \xi_2)$, 设

$$\Lambda^R(t; \xi) = \int_0^t \frac{E\{d\bar{N}^R(u)\} - \beta^T s_{1W}^*(u) du}{s_{0Z}(u; \gamma)}$$

令 $\Lambda^R(t; \eta_0) \equiv \Lambda_0^R(t)$, 其中, $s_{1W}^*(u)$ 为 $s_{1W}(t; 0)$ 的极限值, 在 $t \in [0, \tau]$ 和 $\xi \in \mathcal{B}$ 上, $\tilde{\Lambda}_0^R(t; \xi)$ 几乎处处一致收敛到 $\Lambda^R(t; \xi)$ 。因此证明 $\hat{\xi}$ 、 $\hat{\Lambda}_{01}^D(t)$ 、 $\hat{\Lambda}_{02}^D(t)$ 及 $\hat{\Lambda}_0^R(t)$ 的存在性和唯一性, 只需证明:

$$\tilde{U}(\xi) = (\tilde{U}_1(\xi)^T, \tilde{U}_{21}(\xi_1)^T, \tilde{U}_{22}(\xi_2)^T, \tilde{U}_{31}(\xi), \tilde{U}_{32}(\xi))^T = 0 \quad (11)$$

有唯一解, 则

$$\begin{aligned}\tilde{U}_1(\xi) &= \sum_{i=1}^n \int_0^\tau \{D_i(t; \eta) - \bar{D}(t; \eta)\} \{\psi_{1i}(t; \tilde{\Lambda}_{01}^D) \psi_{2i}(t; \tilde{\Lambda}_{02}^D) dN_i^R(t) - \Delta_i(t) \beta^T W_i(t) dt\} \\ \tilde{U}_{2k}(\xi_k) &= \sum_{i=1}^n \int_0^\tau \{J_i(t) - \bar{J}(t)\} \psi_{ki}(t; \tilde{\Lambda}_{0k}^D) dN_i^{Dk}(t) \\ \tilde{U}_{3k}(\xi) &= \sum_{i=1}^n \int_0^\tau \{N_i^R(t) - (\theta_k + 1) \tilde{G}_k(t) \tilde{\omega}_{2i}(t)\} dN_i^{Dk}(t)\end{aligned}$$

其中 $k=1, 2$, 则

$$\tilde{\omega}_{2i}(t) = \psi_{1i}^{-1}(t; \tilde{\Lambda}_{01}^D) \psi_{2i}^{-1}(t; \tilde{\Lambda}_{02}^D) \int_0^t \left[\exp\{\gamma^T Z_i(u)\} d\tilde{\Lambda}_0^R(u) + \beta^T W_i(u) du \right]$$

$\tilde{G}_k(t)$ 为 $G_k(t)$ 的 $\omega_{2i}(t)$ 替换为 $\tilde{\omega}_{2i}(t)$ 可得。

设

$$\hat{\Gamma}(\xi) = -n^{-1} \partial \tilde{U}(\xi) / \partial \xi^T$$

根据强大数定律以及 $\tilde{\Lambda}_{01}^D(t; \hat{\xi}_1)$, $\tilde{\Lambda}_{02}^D(t; \hat{\xi}_2)$ 和 $\tilde{\Lambda}_0^R(t; \hat{\xi})$ 的一致收敛性, 则对于 $\xi \in \mathcal{B}$ 的 $\hat{\Gamma}(\xi)$ 一致收敛到非随机过程 $\Gamma(\xi)$ 。可证:

$$n^{-1} \tilde{U}(\xi_0) \rightarrow 0$$

所以, $\hat{\Gamma}(\xi)$ 的一致相合性及(C4)得: 当 n 足够大, $\xi \in \mathcal{B}$ 时, $\hat{\Gamma}(\xi)$ 非奇异。根据逆函数定理[18], 则 $\tilde{U}(\xi) = 0$ 存在唯一解 $\hat{\xi}$, 即存在唯一估计

$$\hat{\xi}, \quad \hat{\Lambda}_{01}^D(t) \equiv \tilde{\Lambda}_{01}^D(t; \hat{\xi}_1), \quad \hat{\Lambda}_{02}^D(t) \equiv \tilde{\Lambda}_{02}^D(t; \hat{\xi}_2) \text{ 和 } \hat{\Lambda}_0^R(t) \equiv \tilde{\Lambda}_0^R(t; \hat{\xi}) (0 \leq t \leq \tau)$$

定理 1 证毕。

定理 2 $\hat{\xi}$ 强相合于 ξ_0 , 且在 $t \in [0, \tau]$ 上, $\hat{\Lambda}_0^R(t)$ 几乎处处收敛到 $\Lambda_0^R(t)$, $\hat{\Lambda}_{01}^D(t)$ 和 $\hat{\Lambda}_{02}^D(t)$ 几乎处处分别收敛到 $\Lambda_{01}^D(t)$ 和 $\Lambda_{02}^D(t)$ 。

证明 对(11)进行一阶泰勒展开, 可得

$$n^{-1} \tilde{U}(\hat{\xi}) - n^{-1} \tilde{U}(\xi_0) = -\hat{\Gamma}(\xi_0) (\hat{\xi} - \xi_0) + o(\|\hat{\xi} - \xi_0\|)$$

因此几乎处处有

$$\hat{\Gamma}(\xi_0) (\hat{\xi} - \xi_0) + o(\|\hat{\xi} - \xi_0\|) = o(1)$$

由于 $\xi \in \mathcal{B}$, $\hat{\Gamma}(\xi)$ 非奇异, 上述等式可知 $\hat{\xi}$ 是强相合的。由 $\tilde{\Lambda}_{01}^D(t; \hat{\xi})$, $\tilde{\Lambda}_{02}^D(t; \hat{\xi})$ 和 $\tilde{\Lambda}_0^R(t; \hat{\xi})$ 的一致收敛性表明: 对于 $t \in [0, \tau]$, $\hat{\Lambda}_{01}^D(t)$, $\hat{\Lambda}_{02}^D(t)$ 和 $\hat{\Lambda}_0^R(t)$ 几乎处处一致地分别收敛到

$$\Lambda_1^D(t; \xi_{10}) \equiv \Lambda_{01}^D(t), \quad \Lambda_2^D(t; \xi_{20}) \equiv \Lambda_{02}^D(t) \text{ 和 } \Lambda^R(t; \xi_0) \equiv \Lambda_0^R(t)。$$

定理 2 证毕。

定理 3 $\sqrt{n}(\hat{\xi} - \xi_0)$ 按分布收敛到具有均值为 0, 协方差矩阵为 $\Gamma^{-1} \Sigma (\Gamma^T)^{-1}$ 的正态随机变量。

证明 令 $\tilde{\Lambda}_{01}^D(t) = \hat{\Lambda}_{01}^D(t; \xi_{10})$, $\tilde{\Lambda}_{02}^D(t) = \hat{\Lambda}_{02}^D(t; \xi_{20})$, 由定理 1 得

$$\tilde{\Lambda}_{0l}^D(t) - \Lambda_{0l}^D(t) = \sum_{i=1}^n \int_0^t [\tilde{\Lambda}_{0l}^D(t) - \Lambda_{0l}^D(t)] dK_l(u; \xi_{l0}) + \frac{1}{n} \sum_{i=1}^n \int_0^t \frac{dM_i^{Dl}(u)}{S_{0X}(u; \alpha_{l0})}$$

解得

$$\tilde{\Lambda}_{0l}^D(t) - \Lambda_{0l}^D(t) = \frac{1}{\hat{P}_{nl}(t; \xi_{l0})} \int_0^t \hat{P}_{nl}(u-; \xi_{l0}) \frac{\sum_{i=1}^n dM_i^{Dl}(u)}{nS_{0X}(u; \alpha_{l0})},$$

其中 $l=1, 2$,

$$\hat{P}_{nl}(t; \xi_{l0}) = \prod_{s \leq t} \{1 - dK_l(s; \xi_{l0})\}$$

表示关于 $K_l(s; \xi_0)$ 在 $[0, t]$ 的乘积积分。根据 $\tilde{\Lambda}_{01}^D(t)$ 和 $\tilde{\Lambda}_{02}^D(t)$ 的一致收敛性, 结合一致强大数定律[21] 及 Lin 和 Ying (2011) [9] 研究, 则 $t \in [0, \tau]$ 时, 有

$$\tilde{\Lambda}_{0l}^D(t) - \Lambda_{0l}^D(t) = \frac{1}{n} \sum_{i=1}^n \phi_{li}^l(t) + o_p(n^{-1/2}) \quad (12)$$

其中

$$\phi_{li}^l(t) = \frac{1}{P_l(t; \xi_{l0})} \int_0^t P_l(u-; \xi_{l0}) \frac{dM_i^{Dl}(u)}{S_{0X}(u; \alpha_{l0})}$$

$P_l(t; \xi_0)$ 是 $\hat{P}_{nl}(t; \xi_0)$ 极限, $l=1, 2$ 。

令

$$K^*(t; \xi_0) = \frac{1}{n} \sum_{i=1}^n \int_0^t \exp\left\{(\alpha_{10} + \alpha_{20})^T J_i(u)\right\} \frac{\theta_{10} \theta_{20} dN_i^R(u)}{S_{0Z}(u; \gamma_0)}.$$

由(12)得

$$\tilde{\Lambda}_0^R(t) - \Lambda_0^R(t) = \int_0^t \left\{ \tilde{\Lambda}_{01}^D(u) \tilde{\Lambda}_{02}^D(u) - \Lambda_{01}^D(u) \Lambda_{02}^D(u) \right\} dK^*(u; \xi_0) + \frac{1}{n} \sum_{i=1}^n \int_0^t \frac{dM_i^R(u)}{S_{0Z}(u; \gamma_0)} = \frac{1}{n} \sum_{i=1}^n \phi_{2i}(t) + o_p\left(n^{-\frac{1}{2}}\right) \quad (13)$$

其中

$$\phi_{2i}(t) = \int_0^t \phi_{1i}^1(u) \phi_{1i}^2(u) dK^*(u; \xi_0) + \int_0^t \frac{dM_i^R(u)}{S_{0Z}(u)}$$

在(11)中, 对 $\tilde{U}_1(\xi_0)$ 可分解为

$$\begin{aligned} \tilde{U}_1(\xi_0) &= \sum_{i=1}^n \int_0^\tau \left\{ D_i(t; \eta_0) - \bar{D}(t; \eta_0) \right\} \left\{ \psi_{1i}(t; \tilde{\Lambda}_{01}^D) \psi_{2i}(t; \tilde{\Lambda}_{02}^D) dN_i^R(t) - \Delta_i(t) \beta^T W_i(t) dt \right\} \\ &= \sum_{i=1}^n \int_0^\tau \Delta_i(t) \exp\left\{(\alpha_{10} + \alpha_{20})^T J_i(t)\right\} \left\{ D_i(t; \eta_0) - \bar{D}(t; \eta_0) \right\} \left\{ \tilde{\Lambda}_{01}^D(u) \tilde{\Lambda}_{02}^D(u) - \Lambda_{01}^D(u) \Lambda_{02}^D(u) \right\} dN_i^R(t) \\ &\quad + \sum_{i=1}^n \int_0^\tau \left\{ D_i(t; \eta_0) - \bar{D}(t; \eta_0) \right\} dM_i^R(t) \end{aligned}$$

令

$$\tilde{H}_1(t) = \frac{1}{n} \sum_{i=1}^n \int_0^t \theta_{10} \theta_{20} \Delta_i(u) \exp\left\{(\alpha_{10} + \alpha_{20})^T J_i(t)\right\} \left\{ D_i(u; \eta_0) - \bar{D}(u; \eta_0) \right\} dN_i^R(u)$$

且 $\tilde{d}(t; \eta_0)$ 和 $H_1(t)$ 分别是 $\bar{D}(t; \eta_0)$ 和 $\tilde{H}_1(t)$ 的极限。可以证

$$\tilde{U}_1(\xi_0) = \sum_{i=1}^n \mathcal{A}_{1i} + o_p\left(n^{-\frac{1}{2}}\right) \quad (14)$$

其中

$$\mathcal{G}_{1i} = \int_0^\tau \{D_i(t; \eta_0) - \bar{d}(t; \eta_0)\} dM_i^R(t) + \int_0^t \phi_{1i}^1(t) \phi_{1i}^2(t) dH_1(t)$$

类似地

$$\tilde{U}_2(\xi_0) = \sum_{i=1}^n \mathcal{G}_{2i}^l + o_p\left(n^{\frac{1}{2}}\right) \quad (15)$$

其中

$$\mathcal{G}_{2i}^l = \int_0^\tau \{J_i(t) - \bar{J}(t)\} dM_i^{Dl}(t) + \int_0^t \phi_{1i}^l(t) dH_{2l}(t)$$

$\bar{J}(t)$ 和 $H_{2l}(t)$ 分别是 $\bar{J}(t)$ 和 $\tilde{H}_{2l}(t)$ 的极限, $l=1,2$, 且

$$\tilde{H}_{2l}(t) = \frac{1}{n} \sum_{i=1}^n \int_0^t \theta_{l0} \Delta_i(u) \exp\{\alpha_{l0}^T J_i(u)\} \{J_i(u) - \bar{J}(u)\} dN_i^{Dl}(u)$$

设

$$\begin{aligned} \omega_{2i}^*(t) &= \psi_{1i}^{-1}(t; \Lambda_{01}^D) \psi_{2i}^{-1}(t; \Lambda_{02}^D) \int_0^t \left[\exp\{\gamma^T Z_i(u)\} d\Lambda_0^R(u) + \beta^T W_i(u) du \right] \\ \tilde{Q}(t) &= \frac{\sum_{i=1}^n \Delta_i^*(t) \tilde{\omega}_{2i}^{-1}(t) N_i^R(t)}{\sum_{i=1}^n \Delta_i^*(t)} \end{aligned}$$

且 $q(t)$ 为 $\hat{Q}(t)$ 的极限。经计算可得

$$\begin{aligned} \tilde{U}_{3l}(\xi_0) &= \sum_{i=1}^n \int_0^\tau \{N_i^R(t) - (\theta_{l0} + 1) \omega_{2i}^*(t) q(t)\} dN_i^{Dl}(t) \\ &\quad - (\theta_{l0} + 1) \sum_{i=1}^n \int_0^\tau \tilde{Q}(t) \{\tilde{\omega}_{2i}(t) - \omega_{2i}^*(t)\} dN_i^{Dl}(t) \\ &\quad - (\theta_{l0} + 1) \sum_{i=1}^n \int_0^\tau \{\tilde{Q}(t) - q(t)\} \omega_{2i}^*(t) dN_i^{Dl}(t) \end{aligned} \quad (16)$$

其中 $l=1,2$, 和(A.7)的证明一样: (A.9)右边的第二项化简得

$$(\theta_{l0} + 1) \sum_{i=1}^n \int_0^\tau [\phi_{1i}^l(t) dH_{3l}(t) - \phi_{2i}(t) dH_{4l}(t)] + o_p\left(n^{\frac{1}{2}}\right) \quad (17)$$

其中 $H_{3l}(t)$ 和 $H_{4l}(t)$ 分别是 $\tilde{H}_{3l}(t)$ 和 $\tilde{H}_{4l}(t)$ 的极限, 且

$$\tilde{H}_{3l}(t) = \frac{1}{n} \sum_{i=1}^n \int_0^t \theta_{l0} \left[\exp\{\alpha_{l0}^T J_i(s)\} \tilde{Q}(s) \psi_{1i}^{-1}(t; \tilde{\Lambda}_{01}^D) \psi_{2i}^{-1}(t; \tilde{\Lambda}_{02}^D) \omega_{2i}^*(s) \right] dN_i^{Dl}(s)$$

和

$$\tilde{H}_{4l}(t) = \frac{1}{n} \sum_{i=1}^n \int_0^t \exp\{\gamma_0^T Z_i(s)\} \tilde{Q}(s) \psi_{1i}^{-1}(t; \tilde{\Lambda}_{01}^D) \psi_{2i}^{-1}(t; \tilde{\Lambda}_{02}^D) dN_i^{Dl}(s)$$

设 $d\Phi_i(t) = E\{\omega_{2i}^*(t)\} dN_i^{Dl}(t)$, 同样的, (A.9)右边的第三项等于

$$\begin{aligned} &(\theta_{l0} + 1) \sum_{i=1}^n \int_0^\tau [\phi_{1i}^l(t) H_{5l}(t) + \phi_{2i}(t) H_{6l}(t)] d\Phi(t) \\ &- (\theta_{l0} + 1) \sum_{i=1}^n \int_0^\tau \left[\frac{\omega_{2i}^{*-1}(t) \Delta_i^*(t) N_i^R(t)}{E\{\Delta_i^*(t)\}} - \frac{q(t)}{E\{\Delta_i^*(t)\}} \Delta_i^*(t) \right] d\Phi(t) + o_p(n^{1/2}) \end{aligned} \quad (18)$$

其中 $H_{5l}(t)$ 和 $H_6(t)$ 分别是 $\tilde{H}_{5l}(t)$ 和 $\tilde{H}_6(t)$ 的极限, 且

$$\begin{aligned}\tilde{H}_{5l}(t) &= \frac{\theta_{l0} \sum_{i=1}^n \exp\{\alpha_{l0}^T J_i(t)\} \psi_{1i}^{-1}(t; \tilde{\Lambda}_{01}^D) \psi_{2i}^{-1}(t; \tilde{\Lambda}_{02}^D) \tilde{\omega}_{2i}^{-1}(t) \Delta_i^*(t) N_i^R(t)}{\sum_{i=1}^n \Delta_i^*(t)} \\ \tilde{H}_6(t) &= \frac{\sum_{i=1}^n \exp\{\gamma_0^T Z_i(t)\} \psi_{1i}^{-1}(t; \tilde{\Lambda}_{01}^D) \psi_{2i}^{-1}(t; \tilde{\Lambda}_{02}^D) \tilde{\omega}_{2i}^{-1}(t) \omega_{2i}^{*-1}(t) \Delta_i^*(t) N_i^R(t)}{\sum_{i=1}^n \Delta_i^*(t)}\end{aligned}$$

由上面计算, 则

$$\tilde{U}_{3l}(\xi_0) = \sum_{i=1}^n \mathcal{G}_{3i}^l + o_p(n^{1/2}) \quad (19)$$

$l=1, 2$, 其中

$$\begin{aligned}\mathcal{G}_{3i}^l &= \sum_{i=1}^n \int_0^\tau \{N_i^R(t) - (\theta_{l0} + 1) \omega_{2i}^*(t) q(t)\} dN_i^{Dl}(t) \\ &\quad - (\theta_{l0} + 1) \sum_{i=1}^n \int_0^\tau [\phi_{1i}^l(t) dH_{3l}(t) - \phi_{2i}(t) dH_{4l}(t)] \\ &\quad + (\theta_{l0} + 1) \sum_{i=1}^n \int_0^\tau [\phi_{1i}^l(t) H_{5l}(t) + \phi_{2i}(t) H_6(t)] d\Phi(t) \\ &\quad - (\theta_{l0} + 1) \sum_{i=1}^n \int_0^\tau \left[\frac{\omega_{2i}^{*-1}(t) \Delta_i^*(t) N_i^R(t)}{E\{\Delta_i^*(t)\}} - \frac{q(t)}{E\{\Delta_i^*(t)\}} \Delta_i^*(t) \right] d\Phi\end{aligned} \quad (20)$$

令

$$\mathcal{G}_i = (\mathcal{G}_{1i}^T, \mathcal{G}_{2i}^{1T}, \mathcal{G}_{2i}^{2T}, \mathcal{G}_{3i}^1, \mathcal{G}_{3i}^2)^T$$

及 $\Gamma = \Gamma(\xi_0)$, 根据泰勒一阶展开可得

$$n^{1/2}(\hat{\xi} - \xi_0) = n^{-1/2} \Gamma^{-1} \sum_{i=1}^n \mathcal{G}_i + o_p(1)$$

通过多元中心极限理论, $n^{1/2}(\hat{\xi} - \xi_0)$ 是零均值和协方差矩阵为 $\Gamma^{-1} \Sigma (\Gamma^T)^{-1}$ 的渐进正态分布, 其中 $\Sigma = E\{\mathcal{G}_i \mathcal{G}_i^T\}$ 。定理 3 证毕。

定理 4 当 $t \in [0, \tau]$, $\{\sqrt{n}(\hat{\Lambda}_0^R(t) - \Lambda_0^R(t))\}$, $\{\sqrt{n}(\hat{\Lambda}_{01}^D(t) - \Lambda_{01}^D(t))\}$ 和 $\{\sqrt{n}(\hat{\Lambda}_{02}^D(t) - \Lambda_{02}^D(t))\}$ 均弱收敛到零均值随机过程。

证明 由于证明 $n^{1/2}\{\hat{\Lambda}_{01}^D(t) - \Lambda_{01}^D(t)\}$ 和 $n^{1/2}\{\hat{\Lambda}_{02}^D(t) - \Lambda_{02}^D(t)\}$ 的弱收敛性类似, 为方便起见, 只需证明其一即可, 不妨证明 $n^{1/2}\{\hat{\Lambda}_{01}^D(t) - \Lambda_{01}^D(t)\}$ 和 $n^{1/2}\{\hat{\Lambda}_0^R(t) - \Lambda_0^R(t)\}$ 的弱收敛性, 注意到

$$\hat{\Lambda}_{01}^D(t) - \Lambda_{01}^D(t) = \{\hat{\Lambda}_{01}^D(t; \hat{\xi}_1) - \hat{\Lambda}_{01}^D(t; \xi_{10})\} + \{\hat{\Lambda}_0^D(t; \xi_{10}) - \Lambda_0^D(t)\} \quad (21)$$

令

$$\hat{Y}_{11}(t; \xi_1) = \partial \hat{\Lambda}_{01}^D(t, \xi_1) / \partial \xi_1$$

根据一致强大数定律, 当 $t \in [0, \tau]$ 和 $\xi \in \mathcal{B}$ 时, 可证明 $\hat{Y}_{11}(t; \xi_1)$ 几乎处处收敛到非随机函数 $Y_{11}(t; \xi_1)$, 由泰勒一阶展开和(12), (20), (21)得

$$n^{1/2}\{\hat{\Lambda}_{01}^D(t) - \Lambda_{01}^D(t)\} = n^{-1/2} \sum_{i=1}^n \Psi_{1i}(t) + o_p(1)$$

其中

$$\Psi_{1i}^1(t) = -Y_{11}(t; \xi_{10}) \Gamma^{-1} \mathcal{G}_i + \phi_{1i}^1(t)$$

类似的, 由(13)和(20)得

$$n^{1/2} \{ \hat{\Lambda}_0^R(t) - \Lambda_0^R(t) \} = n^{-1/2} \sum_{i=1}^n \Psi_{2i}(t) + o_p(1)$$

其中

$$\Psi_{2i}(t) = -Y_2(t; \xi_0) \Gamma^{-1} \mathcal{G}_i + \phi_{2i}(t)$$

及 $Y_2(t; \xi_0)$ 为 $\partial \hat{\Lambda}_0^R(t; \xi) / \partial \xi$ 极限, 给定

$$\Psi_i(t) = (\Psi_{1i}^1(t), \Psi_{1i}^2(t), \Psi_{2i}^2(t))^T$$

在时刻 t 的条件下, $\Psi_i(t) (i=1, \dots, n)$ 独立同分布且具有零均值。由多元中心极限定理, $n^{1/2} \{ \hat{\Lambda}_{01}^D(t) - \Lambda_{01}^D(t) \}$, $n^{1/2} \{ \hat{\Lambda}_{02}^D(t) - \Lambda_{02}^D(t) \}$ 和 $n^{1/2} \{ \hat{\Lambda}_0^R(t) - \Lambda_0^R(t) \}$ 的有限维分别弱收敛于零均值过程。此外, $\Psi_i(t)$ 可表示为 t 的单调函数的累加或乘积形式, 故此过程是紧密的[22]。则 $n^{1/2} \{ \hat{\Lambda}_{01}^D(t) - \Lambda_{01}^D(t) \}$, $n^{1/2} \{ \hat{\Lambda}_{02}^D(t) - \Lambda_{02}^D(t) \}$ 和 $n^{1/2} \{ \hat{\Lambda}_0^R(t) - \Lambda_0^R(t) \}$ 是紧的, 并在 (s, t) 处的协方差函数为 $E \{ \Psi_i(s) \Psi_i(t)^T \}$ 。定理 4 证毕。

5. 结束语

针对已存在双重终止事件的复发事件联合模型, 继续对其进行研究, 通过建立估计方程, 推导出方程参数的相合性与渐近正态性。

因为模型(1)和(2)允许脆弱变量 v 与复发事件与终止事件之间相关性是正向, 也可以是负向。对于这种情况, 可以采用以下模型:

$$d\Lambda_R(t|v) = v^{-1} \exp \{ \gamma^T Z(t) d\Lambda_0^R(t) + \beta^T W(t) dt \},$$

其中, 模型(2)不变, v 为单位均值, 方差小于 1 的伽马分布。可得

$$\psi^*(t) = E[v^{-1} | X(t), D \geq t] = \psi(t) / (1 - \theta)$$

和

$$\omega_{1i}(t) / \omega_{2i}(t) = 1 - \theta.$$

这表明复发事件的发生率与终止事件呈反向关联, 因为在时间 t 有一个终止事件的个体预计, 会比在时间 t 之后有一个终止事件的个体有更少的复发事件。此外, 通过在 U_1 , U_3 和 U_4 中分别用 $\psi^*(t)$ 和 $1 - \theta$ 代替 $\psi(t)$ 和 $1 + \theta$, 可以构造与前几节相同的估计方程。

然而, 在分析终止事件和复发事件数据的联合模型时, 协变量的选择标准与确定方法应当遵循以下原则: 首先需要通过 Wald 统计检验和效应量评估, 即标准化系数绝对值大于 1.96 来识别协变量的显著性; 对于显著协变量, 应进一步通过似然比检验比较其在加法模型和乘法模型中的拟合优度, 以确定其更适合纳入可加部分 $W(t)$ 还是可乘部分 $Z(t)$ 。当协变量维度较高时, 则采用 LASSO 惩罚回归方法进行变量筛选; 对于存在高度删失的数据, 应当使用逆概率加权法或多重插补技术来校正选择偏差。此外, 需要通过 Bootstrap 重抽样来验证最终模型的稳定性, 要求核心参数的变化率不超过 15%。这一系统化的协变量选择流程既保证了统计严谨性, 又能适应不同类型的数据特征。

采用广义估计方程法进行参数估计, 在半参数模型框架下可能存在效率提升空间。未来研究可着眼

于开发更简洁高效的推断方法。另外,为评估模型(1)和(2)对数据的拟合程度,可考虑基于残差 $d\hat{M}_i^R(t)$ 和 $d\hat{M}_i^D(t)$ 的图形和数值分析方法[13]。其中, $d\hat{M}_i^R(t)$ 和 $d\hat{M}_i^D(t)$ 通过将 $dM_i^R(t)$ 和 $dM_i^D(t)$ 中的参数替换为相应估计值获得。这一问题的理论深化和数值模拟值得进一步探索。

基金项目

国家自然科学基金项目(11561010), 广西高校中青年教师科研基础能力提升项目(2022KY0707)资助。

参考文献

- [1] Cook, R.J. and Lawless, J.F. (2007) The Statistical Analysis of Recurrent Events. Springer.
- [2] Zeng, D. and Cai, J. (2010) A Semiparametric Additive Rate Model for Recurrent Events with an Informative Terminal Event. *Biometrika*, **97**, 699-712. <https://doi.org/10.1093/biomet/asq039>
- [3] Andersen, P.K. and Gill, R.D. (1982) Cox's Regression Model for Counting Processes: A Large Sample Study. *The Annals of Statistics*, **10**, 1100-1120. <https://doi.org/10.1214/aos/1176345976>
- [4] Sun, L., Zhao, X. and Zhou, J. (2011) A Class of Mixed Models for Recurrent Event Data. *Canadian Journal of Statistics*, **39**, 578-590. <https://doi.org/10.1002/cjs.10132>
- [5] Pan, Q. and Schaubel, D.E. (2009) Flexible Estimation of Differences in Treatment-Specific Recurrent Event Means in the Presence of a Terminating Event. *Biometrics*, **65**, 753-761. <https://doi.org/10.1111/j.1541-0420.2008.01157.x>
- [6] Liu, L., Wolfe, R.A. and Huang, X. (2004) Shared Frailty Models for Recurrent Events and a Terminal Event. *Biometrics*, **60**, 747-756. <https://doi.org/10.1111/j.0006-341x.2004.00225.x>
- [7] Ghosh, D. and Lin, D.Y. (2002) Marginal Regression Models for Recurrent and Terminal Events. *Statistica Sinica*, **12**, 663-688.
- [8] Schaubel, D.E. and Zhang, M. (2010) Estimating Treatment Effects on the Marginal Recurrent Event Mean in the Presence of a Terminating Event. *Lifetime Data Analysis*, **16**, 451-477. <https://doi.org/10.1007/s10985-009-9149-x>
- [9] Zhao, X., Zhou, J. and Sun, L. (2010) Semiparametric Transformation Models with Time-Varying Coefficients for Recurrent and Terminal Events. *Biometrics*, **67**, 404-414. <https://doi.org/10.1111/j.1541-0420.2010.01458.x>
- [10] Kalbfleisch, J.D., Schaubel, D.E., Ye, Y. and Gong, Q. (2013) An Estimating Function Approach to the Analysis of Recurrent and Terminal Events. *Biometrics*, **69**, 366-374. <https://doi.org/10.1111/biom.12025>
- [11] Cook, R.J. and Lawless, J.F. (1997) Marginal Analysis of Recurrent Events and A Terminating Event. *Statistics in Medicine*, **16**, 911-924. [https://doi.org/10.1002/\(sici\)1097-0258\(19970430\)16:8<911::aid-sim544>3.0.co;2-j](https://doi.org/10.1002/(sici)1097-0258(19970430)16:8<911::aid-sim544>3.0.co;2-j)
- [12] Ye, Y., Kalbfleisch, J.D. and Schaubel, D.E. (2007) Semiparametric Analysis of Correlated Recurrent and Terminal Events. *Biometrics*, **63**, 78-87. <https://doi.org/10.1111/j.1541-0420.2006.00677.x>
- [13] Sun, L. and Kang, F. (2013) An Additive-Multiplicative Rates Model for Recurrent Event Data with Informative Terminal Event. *Lifetime Data Analysis*, **19**, 117-137. <https://doi.org/10.1007/s10985-012-9228-2>
- [14] Chen, C., Shen, P. and Chuang, Y. (2015) The Partly Aalen's Model for Recurrent Event Data with a Dependent Terminal Event. *Statistics in Medicine*, **35**, 268-281. <https://doi.org/10.1002/sim.6625>
- [15] Qu, L., Sun, L. and Liu, L. (2017) Joint Modeling of Recurrent and Terminal Events Using Additive Models. *Statistics and Its Interface*, **10**, 699-710. <https://doi.org/10.4310/sii.2017.v10.n4.a12>
- [16] Lin, D.Y. and Ying, Z. (1995) Semiparametric Analysis of General Additive-Multiplicative Hazard Models for Counting Processes. *The Annals of Statistics*, **23**, 1712-1734. <https://doi.org/10.1214/aos/1176324320>
- [17] 孙琴, 曲连强. 带相依终止事件的复发事件数据的可加可乘比率模型[J]. 数学学报(中文版), 2019, 62(1): 87-102.
- [18] Prentice, K., John, D. and Ross, L. (2011) The Statistical Analysis of Failure Time Data. John Wiley & Sons.
- [19] Liu, Y., Wu, Y., Cai, J. and Zhou, H. (2010) Additive-Multiplicative Rates Model for Recurrent Events. *Lifetime Data Analysis*, **16**, 353-373. <https://doi.org/10.1007/s10985-010-9160-2>
- [20] Gill, R.D. and Johansen, S. (1990) A Survey of Product-Integration with a View toward Application in Survival Analysis. *The Annals of Statistics*, **18**, 1501-1555. <https://doi.org/10.1214/aos/1176347865>
- [21] Pollard, D. (1990) Empirical Processes: Theory and Applications. Institute of Mathematical Statistics.
- [22] Van Der Vaart, A.W. and Wellner, J.A. (1996) Weak Convergence and Empirical Processes. Springer.