arXiv:2303.04915v1 [stat.ME] 8 Mar 2023

离散随机效应分布在单变量和多变量生存模型中诱导的相对脆弱性方差的形态

Maximilian Bardo[1*]和 Steffen Unkel[1,2]

[1]医学统计系，哥廷根大学医学中心，德国。[2]V学院：生命科学学院，锡根大学，德国。

*maximilian.bardo@med.uni-goettingen.de

摘要

在用于分析事件发生时间数据的统计模型中，通常通过一个或多个随机效应（也称为脆弱性）来考虑个体异质性。绝大多数文献都假设随机效应服从连续概率分布。然而，在某些应用领域，离散脆弱性分布可能更为合适。我们考察并比较了现有的各种离散单变量和共享脆弱性模型族，重点关注存活者中相对脆弱性分布的方差。存活者中的相对脆弱性方差（RFV）提供了一种易于解释的度量，用于描述由脆弱性模型表征的人群异质性如何随时间演变。我们探讨 RFV 的变化形态，以用于模型选择，并回顾了这一背景下可用的离散随机效应分布。我们发现，离散单变量和共享脆弱性模型的 RFV 轨迹可能呈非单调变化，这种性质并不常见。此外，我们证明，对于具有（不具有）零点原子的离散时不变单变量和共享脆弱性模型，如果离散分布的支撑集可以按升序排列，则 RFV 的极限趋于无穷大（零）。在共享脆弱性模型中，RFV 与交叉比函数存在一一对应关系；我们将这一关系推广到高元情形，因此，我们的结果也适用于簇内关联模式。文中还讨论了与离散时变脆弱性模型的扩展和比较，以及与相关离散脆弱性模型的比较。

关键词： 关联、离散分布、脆弱性、异质性、相对脆弱性方差、生存时间模型

1 引言

在事件发生时间数据的建模中，通常通过随机效应（也称为脆弱性）来考虑观测单位之间未观测到的异质性（例如，参见 Aalen 等人，2008年；Duchateau 和 Janssen，2008年；Hougaard，2000；Wienke，2010）。脆弱性模型为量化聚类事件时间之间的关联，以及表征由难以或不可能测量的因素所导致的观测单位间异质性，提供了一种概念上颇具吸引力的方法。绝大多数文献都将随机效应视为连续变量；伽马分布、逆高斯分布和对数正态分布是脆弱性建模中常见的选择（例如，参见上述著作及其中的参考文献）。然而，在某些应用领域（其中一些将在下文提及），离散脆弱性分布可能更为合适。

Pickles 和 Crouchley（1994）、Begun 等人（2000）、Caroni 等人（2010）、Wienke（2010）第3.2节及其中的参考文献、Farrington 等人（2012）、Ata 和 Ozel（2013）、Bijwaard（2014）、Palloni 和 Beltr´án- S´ánchez（2017）、de Souza 等人（2017）、Cancho 等人（2018、2020）、Gasperoni 等人（2020）、Mohseni 等人（2020）、Cancho 等人（2021）、Molina 等人（2021）以及 Cancho 等人（2022）均研究过适用于单变量和多变量数据的离散脆弱性分布。此类研究所采用的分布包括几何分布（Caroni 等人，2010；Cancho 等人，2021）、负二项分布、泊松分布（Caroni 等人，2010）、零膨胀幂级数分布（上述分布均为其特例）（Cancho 等人，2018、2020）、零修正幂级数分布（零膨胀幂级数分布为其特例）（Molina 等人，2021），以及超泊松分布¨（de Souza 等人，2017；Mohseni 等人，2020）。Ata 和 Ozel（2013）提出了一种离散复合泊松过程，其中泊松分布决定随机变量的个数，这些随机变量服从二项分布、负二项分布或泊松分布，并且它们的总和构成脆弱性。Farrington 等人（2012）提出了Addams族离散脆弱性分布，其中包括缩放的负二项分布、二项分布、泊松分布，以及作为连续型特例的伽马分布。此外，还提出了 k点分布，即脆弱性的有限支撑集服从具有 k 个类别的多项分布；例如，可参见 Wienke（2010）第3.2节及其中的参考文献、Begun 等人（2000）以及 Pickles 和 Crouchley（1994）。关于 k点分布背景下的单变量和多变量生存数据、复发事件、竞争风险数据，以及共享脆弱性和相关脆弱性的讨论，见 Bijwaard（2014）。Gasperoni 等人（2020）提出了 k点脆弱性分布的非参数版本，其中 k 也需要估计。对于未观测异质性源自遗传因素的情形，有研究提出使用 k点分布（参见 Pickles 和 Crouchley，1994；Begun 等人，2000；Wienke，2010）。其他应用包括心力衰竭患者首次再入院的时间，其中未观测异质性

被认为是由医疗服务提供者的表现所引起（Gasperoni 等人，2020）。Bijwaard (2014）提出了劳动力市场方面的应用。Palloni 和 Beltr´an-S´anchez(2017）提出了一种延迟二元脆弱性模型，即一种超额风险模型，其中脆弱性加权的超额风险从随访开始后某个时间点起发挥作用。由连续随机变量控制的延迟脆弱性效应，旨在表示影响成年期健康和死亡率的早期生活条件。Cancho 等人 (2022还提出了一种具有相关脆弱性而非共享脆弱性的离散脆弱性模型，并将泊松分布和伽马分布用作潜在结构的组成部分。

除了 k点分布外，前述离散分布的支撑集均为包含零的自然数。若脆弱性以乘法形式作用于基线风险率，则零处的正概率会导致观测单位不会面临发生目标事件的风险，也就是说，这些单位已被治愈（Aalen 等人，2008年，第244页）。因此，离散脆弱性模型的一个潜在应用领域，是适用治愈率的数据。关于离散治愈率模型在黑色素瘤数据中的死亡时间或复发时间分析上的应用，参见de Souza 等人 (2017)、Cancho 等人(2018年、2020以及 Molina 等人 (2021)。关于这些模型在霍奇金淋巴瘤死亡时间分析中的应用，参见Mohseni 等人 (2020)。

Caroni 等人 (2010)提出了一种不含治愈率的单变量离散脆弱性模型，方法是在脆弱性加权的基线风险率上增加一个额外的风险率。其在可靠性分析中的应用可见Caroni 等人(2010)。Caroni 等人 (2010)和Molina 等人 (2021)还提出了非治愈率模型，方法是从离散脆弱性分布的支撑集中截去零。

离散脆弱性模型中，聚类事件时间的关联形态以及观测单位间异质性的模式，迄今尚未得到充分研究。只有Farrington 等人 (2012年研究了Addams族离散脆弱性分布所诱发的异质性，而Cancho 等人 (2022年考察了其提出的相关脆弱性模型所诱发的簇内关联。因此，本文分别研究离散时不变单变量和共享脆弱性模型的关联形态与异质性模式。我们考察上述若干模型，进而从这一角度对离散时不变单变量和共享脆弱性模型展开一般性研究。具体而言，我们研究存活者相对脆弱性分布的方差（Hougaard，1984），即相对脆弱性方差（RFV）；它是随时间变化的观测单位间异质性的局部度量。对于双变量共享脆弱性模型，RFV 与交叉比函数（CRF）存在一一对应关系（Anderson 等人，1992），后者是随时间变化的簇内关联局部度量。我们对 CRF 的定义稍作调整，使其在每个观测单位具有两个以上事件发生时间变量时仍与 RFV 保持这一关系。因此，在共享脆弱性模型中，关于 RFV 的结果同样适用于（经调整的）CRF。与 CRF 相比，使用 RFV 的优势在于，它还可用于

单变量情形，即每个观测单位只有一个事件发生时间变量的情况，并可作为异质性度量。我们证明，对于具有（不具有）零点原子的离散时不变单变量和共享脆弱性模型，RFV 的极限分别为无穷大（零）。我们还将这一发现推广至分段常数的离散时变单变量和共享脆弱性模型。作为对比，我们表明，Caroni 等人 (2010)提出的模型不符合我们对时不变单变量脆弱性模型的定义，因此不具有这一性质。在多变量情境下，我们将关于离散共享脆弱性模型的发现与Cancho 等人 (2022年提出的离散相关脆弱性模型类别进行比较。

本文其余部分的组织如下。第 2节定义了本文始终考虑的时间不变单变量和共享脆弱性模型。此外，本文重新审视了 RFV 的概念及其与 CRF 的关系。第 3 节研究了选定的离散单变量和共享脆弱性模型的 RFV。随后对离散单变量和共享脆弱性模型的 RFV 进行一般性研究。本节最后将前述结果扩展并比较至时间变化的单变量和共享脆弱性模型，以及相关脆弱性模型。第 4节给出总结性说明。进一步的技术细节和证明见附录。

2 风险率的单变量和共享脆弱性模型

考虑事件发生时间随机向量（RVe） $\boldsymbol{T} = [T(1) \ldots, T^{(J)}]^T$，其中$J \geq 1$，且 $T^{(j)} \in \mathbb{R}_{>0}\ \forall j$。我们将来自 $\boldsymbol{T}$的特定 T（$_j$）称为目标变量。令 $\boldsymbol{x}^{(j)}$（t），其中 $t \in \mathbb{R}_{\geq 0}$，且 $j = 1, \ldots, J$表示可观测、可能随时间变化的协变量向量的实现值，这些向量允许在 j、 $j = 1, \ldots, J$之间变化。给定协变量，RVe $\boldsymbol{T}$服从总体（多元）事件发生时间分布 $\pi_{\boldsymbol{T}}$，其分布参数为$\boldsymbol{\theta}_{\boldsymbol{T}}$。协变量可能影响事件发生时间分布 $\pi_{\boldsymbol{T}}$ 本身和/或分布参数 $\boldsymbol{\theta}_{\boldsymbol{T}}$。对于不同的观测单位 i 和 i'，在给定协变量的条件下，假设 $\boldsymbol{T}_i$与 $\boldsymbol{T}_{i'}$相互独立。在多元情形下，即对于 $J \geq 2$、 $T^{(j)}$和 T（$_{j'}$），其中 $j \neq j'$，假设在给定不可观测的簇特异性特征 z 和协变量时，它们在同一观测单位（簇）内相互独立。假设单位特异性潜在特征 z是随机变量（RVa） $Z \geq 0$的一个实现值，该随机变量服从分布 π_Z，其参数向量为 $\boldsymbol{\theta}_Z$。分布 π_Z 和/或分布参数 $\boldsymbol{\theta}_Z$ 可能取决于随访开始时（即时间 $t = 0$）的一部分单位特异性协变量，这些协变量汇集于向量 $\tilde{\boldsymbol{x}}$中。假设从 Z 抽取的值在给定 $\tilde{\boldsymbol{x}}$的条件下于各观测单位之间相互独立。请注意，例如，这包括时间和目标均不变的协变量的随机斜率，即 $z = z_1 \exp\{\ln\{\boldsymbol{z}_2^T\}\tilde{\boldsymbol{x}}\}$，其中$[z_1, \boldsymbol{z}_2^T]$ 是一个不可观测的、单位特异性的非负RVe实现值，并且该变量在观测单位之间独立同分布。

假设决定条件事件发生时间分布的条件目标特异风险率具有以下形式

$$\begin{aligned}\lambda^{(j)}(t \mid Z=z) &= \lim_{\Delta \to 0} \frac{P(t \leq T^{(j)} < t+\Delta \mid T^{(j)} \geq t, Z=z, \boldsymbol{x}^{(j)}(t))}{\Delta} \\ &= z\lambda_x^{(j)}(\boldsymbol{x}^{(j)}(t), t),\end{aligned} \tag{1}$$

含协变量和目标特异性风险率 $\lambda_x^{(j)}(\boldsymbol{x}^{(j)}(t)\ t) > 0$，以及 $t \in \mathbb{R}_{>0}$。我们称（1）为时间不变脆弱性模型，因为 z 不会随 t变化。如果 $J > 1$，则（1）是时间不变共享脆弱性模型，即 z 在以下各项之间没有差异： j =1、.....、 J。

方程（1）中的单位特异性潜在特征 z 此后也称为随机效应（RE）或脆弱性。脆弱性会导致簇内事件发生时间目标之间的关联，以及簇之间或观测单位之间的异质性。本文将全篇同义使用（共享）脆弱性模型、时间不变（共享）脆弱性模型或RE模型这些术语，并在此情形下指代方程（1）。除非另有说明，我们将仅讨论时间不变（共享）脆弱性模型。

以下我们假设，所有量均以潜在协变量为条件，但不在记号中明确说明。因此，我们将在风险率的记号中省略协变量，并令 $\lambda_x^{(j)}(\boldsymbol{x}^{(j)}(t)\ t)$ 等于基线风险率 $\lambda_0^{(j)}\ (t)$。

对于多元事件发生时间数据，即 $J \geq 2$，CRF（Clayton，1978年）是适用于簇内的关联度量。CRF 最初被引入作为双变量关联的度量，可表示如下：

$$\mathrm{CRF}^{(j,j')}(\boldsymbol{t}) = \frac{h^{(j)}(t_j \mid T^{(j')} = t_{j'})}{h^{(j)}(t_j \mid T^{(j')} > t_{j'})}, \tag{2}$$

其中， $\boldsymbol{t} = [t_j, t_{j'}]^T$，且 $h^{(j)}(t \mid A)$ 表示在可观测条件 A 下的总体风险率，即 $\lim_{\Delta \to 0} {}^{P}(t \leq T^{(j)} < t+\Delta | T(_j) \geq t, A)\ \Delta$。~~在时不变共享脆弱性模型~~中，CRF 的上标具有对称性，即 $\mathrm{CRF}^{(j,j')}(\boldsymbol{t}) = \mathrm{CRF}^{(j',j)}(\boldsymbol{t})$，因为 CRF 可以重新表示为 $\Lambda^{(j,j')}\ 0\ (\boldsymbol{t}) = \Lambda_0^{(j)}\ (t_j) + \Lambda^{(j')}\ 0\ (t_{j'}) = \Lambda^{(j',j)}\ 0\ (\boldsymbol{t})$ 的函数，其中 $\Lambda_0^{(j)}\ (t) = \int_0^t \lambda_0^{(j)}\ (u)du$(Oakes，1989)。然而，对于 $\boldsymbol{t} = [t_j, t]^T$，如果 $j' \neq j''$ and $\Lambda^{(j')}\ 0\ (t) \neq \Lambda_0^{(j'')}\ (t)$，则 $\mathrm{CRF}^{(j,j')}(\boldsymbol{t}) \neq \mathrm{CRF}^{(j,j'')}(\boldsymbol{t})$，因为 CRF 的隐含自变量从 $\Lambda^{(j,j')}\ 0\ (\boldsymbol{t})$ 变为 $\Lambda^{(j,j'')}\ 0\ (\boldsymbol{t})$。前两个 CRF 不相等的原因并不是两对 (j, j') 和 (j, j'') 之间的关联模式不同 —— 对于这两个元组，该模式均由簇特异性脆弱性 z 决定——而是由决定脆弱性选择过程的累积基线风险所定义的通用时间（Farrington 等，2012）随 t 的推进方式不同。因此，我们重新定义 CRF，以避免前述不等式

从而获得一种同样适用于高维情形的关联度量。令

$$\mathrm{CRF}^{(j,j')}(\boldsymbol{t}) = \frac{h^{(j)}(t_j \mid T^{(j')} = t_{j'}, T^{(l)} > t_l \; \forall \; l \neq j, j')}{h^{(j)}(t_j \mid T^{(l)} > t_l \; \forall \; l \neq j)},$$

此时总体风险率与公式 (2) 相同，只是可观测条件 A in $h^{(j)}(t \mid A)$ 扩展为任意数量的事件发生时间脆弱性随机变量，并相应地扩展 $\boldsymbol{t} = [t_1, \ldots, t_J]^T$。如果将 CRF 视为累积基线风险之和的函数，即 $\mathrm{CRF}^{(j,j')}(\boldsymbol{t}) = \mathrm{CRF}^{(j,j')}(\Lambda_0(\boldsymbol{t}))$，其中 $\Lambda_0(\boldsymbol{t}) = \sum_{j=1}^{J} \Lambda_0^{(j)}(t_j)$，那么 $\mathrm{CRF}^{(j,j')}(\boldsymbol{t}) = \mathrm{CRF}^{(j,j'')}(\boldsymbol{t})$。即使对于部分或全部 t, $j' \neq j''$ and $\Lambda^{(j')}(t) \neq \Lambda^{(j'')}(t)$，这仍然成立，因为参数 $\Lambda_0(\boldsymbol{t})$ 不受元组 (j, j') 和 (j, j'') 选择的影响。由于 CRF 不依赖于 j andj' 的选择，以下将省略上标。这一调整使 CRF 成为通用时间尺度 $\Lambda \in \{\Lambda_0(\boldsymbol{t}) : \boldsymbol{t} \in \mathbb{R}_{\geq 0}\}$ 上簇内的多变量关联度量。注意，对于所有 j, $\Lambda_0^{(j)}(0)$ 定义为零。

CRF 的这个定义保留了$\boldsymbol{t}$，即 $= 1 +$RFV$\boldsymbol{t}$，其中相对脆弱性方差 RFV$\boldsymbol{t}$为 $= \mathrm{Var}_{Z|\boldsymbol{T}>\boldsymbol{t}}$除以 $\mathrm{E}_{Z|\boldsymbol{T}>\boldsymbol{t}^2}$，这一结果已在双变量情形中得知（Anderson 等人，1992），并适用于每个簇中包含两个以上事件发生时间 RVa 的情况。推导过程与双变量情形基本相同，见附录 A。与 CRF 相比，RFV 的优势在于，它也适用于单变量生存时间数据（$J = 1$）。在单变量和多变量情形下，RFV 都可以解释为：在多变量情形下，给定簇截至 $\boldsymbol{t}$ 时仍无事件的条件下，观测单位特异性条件风险率相对于总体风险率的离散度。更准确地说，

$$\begin{aligned}\mathrm{RFV}(\boldsymbol{t}) &= \frac{\lambda_0^{(j)}(t_j)^2 \, \mathrm{Var}(Z \mid \boldsymbol{T} > \boldsymbol{t})}{\lambda_0^{(j)}(t_j)^2 \, \mathrm{E}(Z \mid \boldsymbol{T} > \boldsymbol{t})^2} \\ &= \frac{\mathrm{Var}(\lambda^{(j)}(t_j \mid Z) \mid T^{(l)} > t_l \; \forall l \neq j)}{h^{(j)}(t_j \mid T^{(l)} > t_l \; \forall l \neq j)^2},\end{aligned}$$

在时不变共享脆弱性模型中，这一结果不依赖于 j 的选择。CRF 与 RFV 之间的一一对应关系表明，首先，RFV 也可以表示为通用时间尺度 Λ 的函数。其次，对于共享脆弱性模型，根据 RFV 衡量的异质性所作的陈述，与根据 CRF 衡量的关联所作的陈述等价。

RFV 唯一确定了 Z、 $\mathcal{L}(s) = \int_0^\infty \exp\{-zs\} g(z) dz$ 的拉普拉斯变换，因此也确定了 Z 的密度 $g(z)$，以及（多变量）总体事件发生时间生存函数 $S(\boldsymbol{t}) = P(\boldsymbol{T} > \boldsymbol{t}) = \mathcal{L}(\Lambda(\boldsymbol{t}))$，均值 $\mathrm{E}(Z)$ 除外（Farrington 等人，2012年）。这凸显出，RFV 是模型构建过程中应予考虑的重要量

，因为错误设定的 RFV 会导致模型错误设定，并可能造成无效推断。据我们所知，迄今尚未对离散脆弱性模型中的 RFV 进行充分研究。因此，在下一节中，我们将研究众多离散脆弱性模型的 RFV 轨迹，以及离散脆弱性模型通常能够诱导出的 RFV 的长期形状。

3 相对脆弱性的形态

3.1 离散脆弱性分布族及其相对脆弱性

本节考虑的离散脆弱性分布包括负二项分布（$\mathcal{NB}$）、二项分布（$\mathcal{B}$）、泊松分布（$\mathcal{P}$）以及 k点分布。我们还考虑零修正泊松分布（$\mathcal{ZMP}$），其中 $Z = 0$ 处的概率由一个收缩/膨胀附加参数建模。我们进一步区分 $\mathcal{NB}$ 和 $\mathcal{NB}_{>0}$：前者的支撑集包含零，即其底层脆弱性随机变量为“观察到 ν 次成功之前的失败次数”；后者的支撑集不包含零，即其底层脆弱性随机变量为“观察到 ν 次成功之前的试验次数”。

文献中的离散脆弱性模型通常用作治愈率模型（¨），即 $P(Z = 0) > 0$（参见 Caroni 等人，2010；Ata 和 Ozel，2013年；de Souza 等人，2017；Cancho 等人，2018年、2020；Mohseni 等人，2020；Cancho 等人，2021；Molina 等人，2021）。Caroni 等人（2010）在单变量情形下提出了不含治愈率的离散脆弱性模型，该模型可按如下方式推广至多变量情形：$\lambda^{(j)}(t \mid Z = z) = z\lambda_0^{(j)}(t) + \lambda_+^{(j)}(t)$，其中 $\lambda_+^{(j)}(t) \in \mathbb{R}_{\geq 0}$对所有 t均成立。用乘法风险率表示时，该模型可写为 $\lambda^{(j)}(t \mid Z = z) = (z + p^{(j)}(t))\lambda_0^{(j)}(t)$，其中 $p^{(j)}(t) = \lambda_+^{(j)}(t)\lambda_0^{(j)}(t)$。一般而言，该模型不对应于时间不变（共享）脆弱性模型的方程（1），因为脆弱性取决于时间（以及 j）。施加比例性约束 $p = \lambda_+^{(j)}(t)\lambda_0^{(j)}(t)$，对 $j = 1, ..., J$及所有 t均成立时，该模型便是符合方程（1）的（共享）脆弱性模型，即脆弱性既不取决于时间 t，也不取决于 j。此时，脆弱性分布是某种支撑集平移了 $p \in \mathbb{R}_{\geq 0}$的（离散）分布。本文研究的这类离散（共享）脆弱性模型包括平移负二项分布（$\mathcal{NB} + p$）、平移二项分布（$\mathcal{B} + p$）和平移泊松分布（$\mathcal{P} + p$）。我们将支撑集由附加参数 p平移的脆弱性分布称为平移模型。

Caroni 等人 (2010) 还建议从离散脆弱性分布的支撑集中截去零。这一思路被 Molina 等人(2021) 采用，其中通过零修正脆弱性分布的一个特例隐式地考虑了截断，因此，我们通过该$\mathcal{ZMP}$ 分布来考察这类模型。我们将从离散脆弱性分布的支撑集中截去零的模型称为截断模型。我们将其中

$P(Z=0)$ 的概率通过附加参数进行削减或增加的模型称为零修正模型。

使用平移模型和零修正模型存在一个缺点。一方面，如果平移模型允许零值，那么参数 p必须位于参数空间的边界上，才能允许 $P(Z=0)>0$。另一方面，在零修正模型中，缩减/膨胀参数必须位于参数空间的边界上，才能对应截断模型，即满足 $P(Z=0)=0$。如果其中一种情况是真实的数据生成过程，或最符合当前数据，这可能会导致估计过程出现困难。我们将展示，在平移模型中，即使 p 的值非常小，或者在零修正模型中 $P(Z=0)$ 的值非常小，也会对相对脆弱性产生巨大影响。

Farrington 等人（2012年）提出了离散脆弱性分布的 Addams 族（$\mathcal{AF}$）。$\mathcal{AF}$ 避免了上述问题，即 $P(Z=0)>0$ 或 $P(Z=0)=0$ 位于参数空间边界上的问题。对于这一族，有 $\text{RFV}(t)=\gamma\exp\{\alpha\Lambda_0(t)\}$，其中 $\gamma>0$，且 $\alpha\in\mathbb{R}$。若 $\alpha<0$，则 Z服从缩放后的 $\mathcal{NB}_{>0}$。若 $\alpha>0$，则 Z要么服从缩放后的 $\mathcal{NB}$，要么服从 $\mathcal{B}$ 或 $\mathcal{P}$，具体取决于 γ。情形 $\alpha=0$ 是一个连续例外，此时 Z 服从伽马分布（$\mathcal{G}$）。因此， $\mathcal{AF}$ 可用作探索性工具，以判断零是否应属于离散脆弱性分布的支撑集。然而，Farrington 等人（2012年）仅提出了这一类分布（$\mathcal{AF}$）；目前尚无可用的估计方法。本文将重新介绍 $\mathcal{AF}$ 的相对脆弱性的形状。

在开始考察 RFV 的轨迹之前，我们将讨论平移模型和零修正模型的拉普拉斯变换及 RFV。我们先从平移模型开始，记作 $Z=Z_*+p$，其中 Z_* 引入了 Z 中的随机性，而 p 是支撑集平移参数。设 $\mathcal{L}_*(s)$ 为 Z_* 的拉普拉斯变换。那么， Z 的拉普拉斯变换为 $\mathcal{L}(s)=\mathcal{L}_*(s)\exp\{-ps\}$。平移模型的 RFV 为 $\text{RFV}(t)=\text{RFV}_*(t)\left[\frac{\mathcal{L}_*'(\Lambda_0(t))}{\mathcal{L}_*'(\Lambda_0(t))}\right]^2$，其中 $\text{RFV}_*(t)=\frac{\text{Var}(Z_*|T>t)\,\text{E}(Z_*|T>t)}{\mathcal{L}_*'(\Lambda_0(t))\mathcal{L}_*(\Lambda_0(t))}-1$，而 $\mathcal{L}'(s)$ $\mathcal{L}''(s)$ 分别表示拉普拉斯变换的一阶和二阶导数。注意，使用 Z 而非 Z_* 作为 RVa，可以保留 $\text{RFV}(t)=\text{CRF}(t)-1$ 这一关系；对于 $\text{RFV}_*(t)$，情况则并非如此，因为 $\mathcal{L}_*(s)$ 失去了与总体生存函数 $S(t)=P(T>t)=\mathcal{L}(\Lambda_0(t))$ 的联系，因而也失去了与总体风险率的联系。因此，$\text{RFV}(t)$ 是一种合适的关联度量和异质性度量，而 $\text{RFV}_*(t)$ 则不是。由此，我们认为，将 p 解释为脆弱性分布的支撑集移位参数，比将其视为模型中的某个加性常数更为合理。

零修正模型中的脆弱性概率质量函数（pmf）定义如下

$$g(z)=\begin{cases}\phi g_*(0) & \text{for } z=0\\ \frac{g_*(z)(1-\phi g_*(0))}{1-g_*(0)} & \text{for } z=1,2,\dots\end{cases},$$

其中， $g_*(z)$ 是某种离散参考分布，例如 $\mathcal{P}$，以及$\phi \in [0, \frac{1}{g_*(0)})$。参数 ϕ 是缩减/膨胀参数，用于调整 $Z = 0$处的概率； $\phi = 0$ 对应于截断模型。零修正模型的拉普拉斯变换为 $\mathcal{L}(s) = \frac{1-\phi g_*(0)}{1-g_*(0)} \mathcal{L}_*(s) + g_*(0)^{\frac{\phi-1}{1-g_*(0)}}$，其中 $\mathcal{L}_*$ 表示参考分布的拉普拉斯变换。零修正模型中的相对脆弱性（RFV）为 $\text{RFV}(t) = \text{RFV}_*(t) + \frac{\phi-1}{1-\phi g_*(0)} g_*(0)^{\text{RFV}_*(\Lambda_0 t + 1)}$。对于移位模型和零修正模型，如果指的是特定参考分布 $\mathcal{L}_*, \Lambda_0$，我们将分别用 $\mathcal{L}_\pi(s)$ 和 $\text{RFV}_\pi(\Lambda_0(t))$ 替换 $\mathcal{L}_*(s)$ 和 $\text{RFV}_*(\Lambda_0(t))$。

表B1 附录B 概述了拉普拉斯变换、$\mathcal{L}(s)$、相对脆弱性(Λ)、其导数相对脆弱性（详细信息）$'(\Lambda)$，以及 Z 在 $Z \sim \mathcal{NB}$、 $\mathcal{NB}_{>0}$、 $\mathcal{B}$、 $\mathcal{P}$、 $\mathcal{NB}+p$、 $\mathcal{B}+p$、 $\mathcal{P}+p$、 $\mathcal{AF}$以及 $\mathcal{ZMP}$分布的支撑集。

相对于通用时间尺度绘制的所得相对脆弱性图示 $\Lambda \in \{\Lambda_0(t): t \in \mathbb{R}^J_{\geq 0}\}$ 上述（共享）脆弱性模型的图示见图(a)至(f)的子图1。非平移模型， $\mathcal{B}$， $\mathcal{P}$ 在图例中以 $p = 0$ 标出，见图1(c) 和1(d)。请注意，相对脆弱性的截距、陡峭程度以及极小值和极大值的位置（若存在）取决于参数值的选择，但图1中相对脆弱性的总体轨迹足够清晰。下文中，我们将“尾部”一词用于表示相对脆弱性从时间轴上的该点 Λ 开始的轨迹；该点之后相对脆弱性的斜率最后一次改变符号，若斜率符号未发生变化，则“尾部”指整个轨迹。

$\mathcal{AF}$ 模型的相对脆弱性轨迹已在 Farrington 等人（2012年）中讨论，见图1(a)。相对脆弱性在任何情况下都是单调的，但可以严格递增、严格递减或保持不变。 $\mathcal{AF}$ 的未缩放成员会产生单调的相对脆弱性轨迹（ $\mathcal{NB}$ 和 $\mathcal{NB}_{>0}$见图1(b)， $\mathcal{B}$见图1(c)， $\mathcal{P}$见图1(d)）。由于 $\mathcal{AF}$ 包含所有这些分布，相对脆弱性的趋势并非预先确定，而是由 α决定。对于 $\alpha > 0$，相对脆弱性严格递增，即脆弱性分布为缩放后的 $\mathcal{NB}$、 $\mathcal{P}$或 $\mathcal{B}$。此时， $Z = 0$ 属于离散脆弱性分布的支撑集。若 $\alpha < 0$，相对脆弱性则严格递减，且相对脆弱性是缩放后的$\mathcal{NB}_{>0}$。此时， $Z = 0$ 不属于脆弱性分布的支撑集。若缩放参数 α 趋近于零，脆弱性分布将趋近于 $\mathcal{AF}$的连续例外情况，即 $\mathcal{G}$ 分布。由此，恒定的相对脆弱性可能是 $\mathcal{AF}$的某个成员所产生的一种形态；离散共享脆弱性模型不可能产生这种形态，因为 $\mathcal{G}$ 是唯一具有恒定相对脆弱性的分布。

平移模型（图 1(b)、1(c)、1(d)），以及 $\mathcal{ZMP}$ 模型（图 1(e)、1(f)）具有一个有趣的性质：相对脆弱性的轨迹不受限于单调变化。对于平移模型，当 $p > 0$ 时，相对脆弱性的尾部单调递减；也就是说，只要 $Z = 0$ 不属于脆弱性分布的支撑集，就会如此。只有在 $p = 0$ 的边界情形下，相对脆弱性的尾部才会单调递增；也就是说，如果 $Z = 0$ 属于

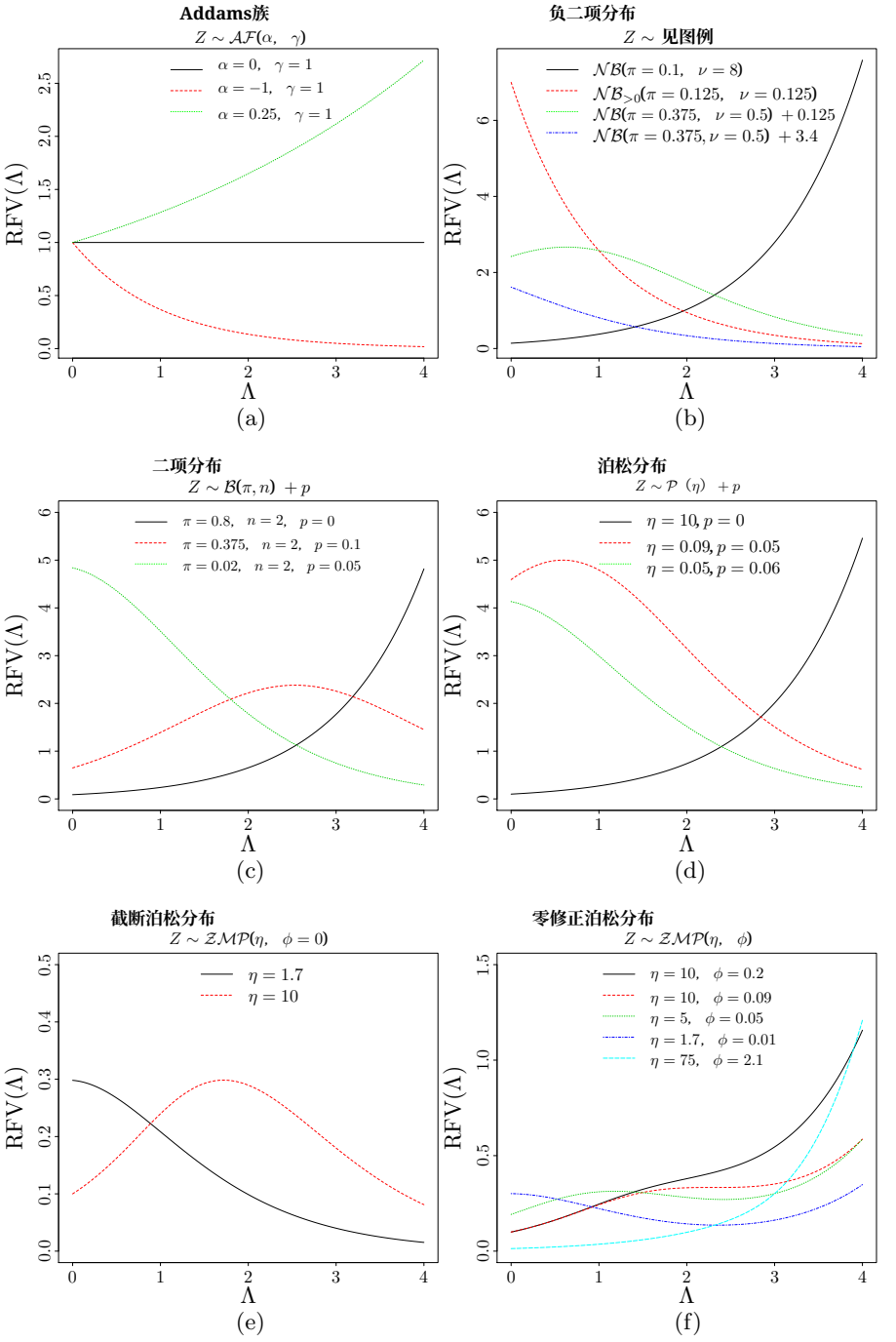

图1：离散（共享）脆弱性模型中相对脆弱性的形状。

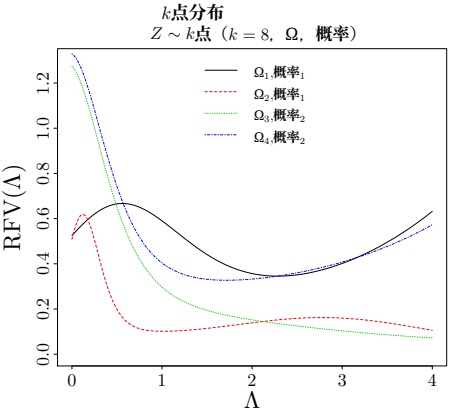

图2： k点分布脆弱性的相对脆弱性形状示例。参数为：$\Omega_1 = \{0.99, 2.02, 2.22, 2.41, 2.51, 2.52, 3.96, 10.44\}$、$\Omega_2 = \Omega_4$，但 $z_{(1)} = 0$、$\Omega_3 = \{0.35, 0.41, 0.49, 0.60, 1.09, 1.39, 3.75, 5.63\}$、$\Omega_4 = \Omega_3$，但 $z_{(1)} = 0$、**概率**$_1 = [0.03, 0.22, 0.01, 0.03, 0.18, 0.03, 0.16, 0.34]$，以及**概率**$_2 = [0.04, 0.14, 0.25, 0.21, 0.17, 0.08, 0.07, 0.04]$。

脆弱性分布的支撑集。所述移位模型的相对脆弱性可能只有一个驻点，因此对于 $p > 0$，其轨迹可能是非单调的。

模型的相对脆弱性轨迹见图 $\mathcal{ZMP}$ 1(e)和1(f)。对于 $\mathcal{ZMP}$ 模型，当 $\phi > 0$时，相对脆弱性的尾部单调递增；也就是说，当 $Z = 0$ 属于脆弱性分布的支撑集时（图1(f)）。在 $\phi = 0$的边界情形下，即 $Z = 0$ 不属于脆弱性分布的支撑集时，相对脆弱性的尾部递减（图1(e)）。 $\mathcal{ZMP}$ 模型的相对脆弱性最多可有两个驻点。此情形仅在 $\phi > 0$时出现，此时局部最大值之后是局部最小值。如果相对脆弱性只有一个驻点，那么对于 $\phi > 0$，该驻点要么是全局最小值，要么是鞍点。如果 $\phi = 0$，则驻点为全局最大值。

图2展示了 k点分布的相对脆弱性（RFV）的一些示例。 k点分布的参数包括潜在组的数量$k \in \mathbb{N}$、支撑集 $\Omega = \{z_{(1)}, \ldots, z_{(k)}\} \in \mathbb{R}^k_{\geq 0}$（其中 $z_{(1)} < \cdots < z_{(k)}$），以及概率 $\mathrm{pr}_{(m)} = P(Z = z_{(m)})$（ $m = 1, \ldots, k$），这些参数汇总在向量 pr 中。 k点分布的相对脆弱性（RFV）等于$\sum_{m'=1}^{k} \sum_{m=1}^{k} z^2$ ($\sum_{m'=1}^{k} \sum_{m=1}^{k} \frac{m}{z_{(m)} z_{(m')}} \{-z_{(m)} + z_{(m')}\Lambda\}$ $\frac{(m)}{(m)}$ $\frac{(m')}{(m')}$) exp$\{-(z_{(m)} + z_{(m')})\Lambda\}$pr pr exp ( ) pr pr$_{(m')}$ 1。请注意，与图1不同，图2中相对脆弱性（RFV）的轨迹（其中 k 取值为8）仅为示例，并未对 k 点分布族中相对脆弱性（RFV）的形状进行全面研究。此外，请注意，图1(c)表示两点分布的一种特殊情况。我们观察到了单调和非单调的相对脆弱性（RFV）轨迹。图2展示的轨迹具有

最多三个驻点。对于图2中的示例，当 $z_{(1)} > 0$ 时，尾部严格递减；当 $z_{(1)} = 0$时，尾部严格递增。我们模拟了几个 k点分布的示例，从未发现当 Z 时，相对脆弱性（RFV）存在超过4个驻点的分布，即使 k非常大也是如此。

3.2 RFV的渐近行为

请注意，对于图 1 中展示的所有模型以及图 2中的示例，只有当离散脆弱性分布在 $Z = 0$ 上具有正概率时，RFV 的尾部才会递增，CRF 的尾部也因此递增；而只有当离散分布在$Z = 0$ 上的概率为零时，尾部才会递减。这并非巧合，下面的命题将予以说明。

命题 1 *Let $Z \geq 0$ be a discrete RVa for which the support can be arranged in ascending order i.e. $= \{z_{(1)}, z_{(2)}, \dots\}$. The support is bounded from below, $\Omega\,\Omega$, , by minimum $z_{(1)}$ but not necessarily from above. Then in a time-invariant (shared) frailty model according to equation 1 where the frailties z are realizations of Z and ()*

a $t_j \in \boldsymbol{t}$ exists such that 极限$_{t_j \to \infty} \Lambda_0(\boldsymbol{t}) \to \infty$,

$$\lim_{t_j \to \infty} \mathrm{RFV}(\Lambda_0(\boldsymbol{t})) = \begin{cases} \infty & \text{if } z_{(1)} = 0 \text{ (i)} \\ 0 & \text{if } z_{(1)} > 0 \text{ (ii)}. \end{cases}$$

Proof 命题1 的证明见附录C。 □

因此，若命题1的假设成立，且相对脆弱性（CRF）具有尾部，则在离散时不变（共享）脆弱性模型中，当且仅当最小脆弱性为零时，相对脆弱性（CRF）的尾部严格递增；当且仅当最小脆弱性大于零时，其尾部严格递减。

命题1的结果也可推广至某一类离散时变（共享）脆弱性模型。设 $Z(t)$在互不相交的时间区间上为分段常数，即 $Z(t) = \sum_{q=1}^{Q} \mathcal{I}(t \in [t_{q-1}, t_q)) Z_q$，其中 $t_Q = \infty > t_{Q-1} > \cdots > t_1 > 0 = t_0$，$\mathcal{I}$是示性函数：当括号中的表达式为真时取值为1，否则取值为0；并且每个 Z_q，其中$q = 1, \dots, Q$，都是离散脆弱性随机变量。于是，根据命题 $Z(t_Q$，的支撑集决定了相对脆弱性（CRF）的长期行为。相应的证明可视为从时间 t_Q开始进行。注意，我们并未声称这是使命题1的结果能够推广至离散时变脆弱性模型的限制最少的一组假设；它只是最便于将证明调整至时变情形的一组假设。证明的调整过程也见附录C。

命题 1的结果可以解释如下。我们先从聚类视角出发，关注 CRF，然后再从异质性视角出发，即关注 RFV。对这两种视角的解释都依赖于这样一个事实：$P(Z = z_{(1)} \mid \boldsymbol{T} > \boldsymbol{t})$ 随着 $\boldsymbol{t}$ 增加，并最终趋近于一。

在 $z_{(1)} = 0$ 的情况下，随着时间增加，可以越来越准确地将簇分为非脆弱簇和脆弱簇。从这个意义上说，无事件簇未来发生任何事件的可能性越来越低，因为这些簇不脆弱的可能性越来越高。然而，脆弱簇在经过一段时间后会发生事件。这意味着，对于足够大的 t，可以几乎完美地将簇中的无事件簇和部分无事件簇区分为永远不会发生任何事件的簇，以及最终会发生所有事件的簇。因此，以 CRF 衡量的簇内关联趋于无穷大。

在 $z_{(1)} > 0$ 的情况下，所有簇都具有脆弱性。因此，所有簇最终都会经历事件。由于具有 $z_{(1)}$ 的簇最终会在总体中占据主导地位，同时也会经历事件，因此，对于 CRF 定义中的较大 t，这些簇的风险会主导总体风险率。因此，对于较大的 t，有关事件是否发生的信息不会为该簇的风险率提供相关信息。由此，簇内的关联（如 CRF 所表示的）会随着时间推移而减弱。

在异质性的背景下，即聚焦于 RFV，对于无事件观测单位，$\lambda^{(j)}(t \mid Z)$ 的期望 $h^{(j)}(t_j \mid T(l) > t_l, l \neq j)$ 在较大的 t 下趋近于 $z_{(1)}\lambda_0^{(j)}(t)$。在没有发生事件的条件下，观测单位特异性风险率 $Z\lambda_0^{(j)}(t)$ 的变异趋近于零，因为当 t 较大时，具有脆弱性 $z_{(1)}$ 的簇会在总体中占据主导地位。因此，如果 $z_{(1)} > 0$，存活者的观测单位特异性风险率相对于总体风险率平方的变异趋近于零。然而，在 $z_{(1)} = 0$ 的情况下，总体风险率 $h^{(j)}(t_j \mid T(l) > t_l, l \neq j)$ 趋近于零，而相对于总体风险率平方的变异趋近于无穷大。

命题 1 还表明，离散脆弱性分布的支撑集具有灵活性十分重要。对支撑集的限制决定了相对脆弱性尾部的轨迹。在许多情况下，最小可能脆弱性值的选择决定了相对脆弱性的整个轨迹，因为大多数标准离散脆弱性分布的相对脆弱性都是单调的。也可以利用支撑集方面更加灵活的模型来验证治愈率假设。从这个意义上说，估计得到的递减相对脆弱性可视为治愈率不存在的反证，而递增的相对脆弱性则可视为存在治愈率的迹象。

命题 1还意味着，在离散（共享）脆弱性设定下，相对脆弱性（因而 CRF）永远不会收敛到正渐近值（大于 1），这或许可视为一种缺点。因此，在一个总体为离散的分布族中纳入连续的特殊情形就显得尤为重要，因为在这些情形下，相对脆弱性可以收敛到正渐近值，例如 $\mathcal{G}$ 中的 $\mathcal{AF}$，其相对脆弱性为常数。

或者，命题1 的基本假设可能不成立，因此相对脆弱性的极限行为不同于命题 1中所述的（共享）脆弱性模型的两种极端情形。我们讨论两个例子：首先是相关脆弱性模型，其次是具有时变支撑集的模型。

在多变量情境（$J > 1$）中，可以使用离散相关脆弱性模型来构建一种模型，使 CRF 收敛到大于 1 的正渐近值。为说明这一点，我们简要考察 Cancho等人（2022）提出的模型。假设条件目标特异性风险率为 $\lambda^{(j)}(t \mid Z^{(j)} = z^{(j)}) = z^{(j)}\lambda_0^{(j)}(t)$，其中 $j = 1, \ldots, J$。给定$W = w$，脆弱性随机变量 $Z^{(j)}, j = 1, \ldots, J$服从 $\mathcal{P}(\eta_j w)$，并且对于任意一对 $Z^{(j)}$ 和 $Z^{(j')}$及 $j \neq j'$，条件独立。RVa $W \geq 0$是该模型中的关联诱导特征，其分布为 π_W。 $Z^{(j)}$与 $Z^{(j')}$的相关性等于 Var(W $\sqrt{\eta_j \eta_{j'}}$)) E( ))(Var$(w)\eta_{j'}$+E(w))，其中 $j \neq j'$的取值范围为区间 $(0, 1]$，当 Var(W) η_j时；$\eta_{j'} > 0$。该模型在数学上较为简便，因为$\mathcal{P}$ 的拉普拉斯变换是指数函数（见附录 B 中的表 B1），因此，$S(\boldsymbol{t} \mid W = w) = \exp\{-wd(\boldsymbol{t})\}$，其中 $d(\boldsymbol{t}) = \sum_{j=1}^{J} \eta_j(1 - \exp\{-\Lambda_0^{(j)}(t_j)\}$。由此，$S(\boldsymbol{t}) = \mathcal{L}_{\pi_W}(d(\boldsymbol{t}))$，即生存函数由 W 的拉普拉斯变换决定；这与相对效应分布为 W的共享脆弱性模型相同，不同之处在于其隐式自变量由$\mathcal{P}$ 分布通过 $d \in \{d(\boldsymbol{t}) : \boldsymbol{t} \in \mathbb{R}_{\geq 0}\}$确定。注意，$\lim_{t_j \to \infty} d(\boldsymbol{t}) = \eta_j + \sum_{j' \neq j} \eta_{j'}(1 - \exp\{-\Lambda_0^{(j')}(t_{j'})\}$，且 $\lim_{\boldsymbol{t} \to \infty} d(\boldsymbol{t}) = \sum_{j=1}^{J} \eta_j$。

Cancho 等人 (2022年) 在其提出的相关脆弱性模型中研究了 $W \sim \mathcal{G}$ 对应的 CRF。然而，我们假设 π_W的取值任意。由于 $d(\boldsymbol{t}$关于 t_j 和 $t_{j'}$的交叉偏导数，在 $j \neq j'$时等于零，因此 CRF 与元组 (j, j') 的取值无关。此外，CRF 可表示为 d 的函数，且完全由 W的拉普拉斯变换决定，即

$$\begin{aligned} \mathrm{CRF}(\boldsymbol{t}) &= \mathrm{CRF}(d(\boldsymbol{t})) \\ &= \frac{\mathcal{L}''_{\pi_W}(d(\boldsymbol{t}))\mathcal{L}_{\pi_W}(d(\boldsymbol{t}))}{\mathcal{L}'_{\pi_W}(d(\boldsymbol{t}))^2}. \end{aligned}$$

因此， W决定了 CRF 在 $d \in [0, \sum_{j=1}^{J} \eta_j]$上的形状，就如同它是共享脆弱性模型中的相对效应一样。然而， $\mathcal{P}$ 分布决定了这种形状随时间变化的速度 $\boldsymbol{t}$ ，该变化通过 $d(\boldsymbol{t})$ 实现。最终， $\mathcal{P}$ 分布也决定了 CRF 的极限，$\lim_{\boldsymbol{t} \to \infty} \mathrm{CRF}(d(\boldsymbol{t})) = \mathrm{CRF}(\sum_{j=1}^{J} \eta_j)$，通常是一个大于一的正数。图 1 展示了在相应的 π_W 取值下 CRF 的一些变化轨迹，其中只需将横坐标轴截断于 $\sum_{j=1}^{J} \eta_j$ ，并在纵坐标轴上加上一。由此可见，在 Cancho 等人 (2022年) 提出的离散相关脆弱性模型类别中，CRF 通常必然趋近于一个大于一的正数，这与离散共享脆弱性模型形成对比。请注意，在相关脆弱性模型中，CRF 与相对脆弱性之间不再存在关联。

到目前为止，我们只考虑了Caroni 等人(2010)提出的模型中的一部分，即平移模型。平移模型属于命题1的适用范围。不过，如果允许平移参数

依赖于时间，这些模型便不符合我们根据公式 (1) 对时间不变（共享）脆弱性模型的定义，因此违反了命题1的假设。此类模型的离散潜在随机变量为$Z(t) = Z_* + p(t)$，其中 $Z_* \geq 0$ 是某个离散随机变量，而 $p(t) \geq 0$ 是依赖于时间的平移参数。风险率等于 $\lambda^{(j)}(t \mid Z_* = z_*) = z(t)\lambda_0^{(j)}(t)$，其中 $z(t) = z_* + p(t)$。该模型的相对脆弱性等于 $\mathrm{RFV}(t) = \mathrm{RFV}_*(\Lambda_0(t))\left[\mathcal{L}_{*}{}'_{\Lambda_0}{}_{t}^{\mathcal{L}_{*}{}'}(\Lambda_0(t))\,(\,(\,))-p(t)\mathcal{L}_*(\Lambda_0(t))\right]^2$，与~~平移模型的情况类似~~。若 Z_* 选为 $\mathcal{P}$ 分布，且 $p(t) = \exp\{-\frac{\Lambda(t)}{2}\}\eta$，则 $\mathrm{RFV}(t) = \frac{1}{\eta}\left[\exp\{-\frac{\Lambda(t)}{2}\} + 1\right]^{-2}$，其收敛于 $\frac{1}{\eta}$。请注意，对于这类模型而言， $p(t)$在长期内趋近于零，是相对脆弱性趋近于正渐近值的必要条件，但并非充分条件（无论 $\mathcal{P}$ 分布如何选择）。若改为 $p(t) = \exp\{-\frac{\Lambda(t)}{2}\}\eta(\sin(\Lambda(t)) + 2)$，则 $\mathrm{RFV}(t) = \frac{1}{\eta}\left[\exp\{-\frac{\Lambda(t)}{2}\} + \sin(\Lambda(t)) + 2\right]^{-2}$，既不收敛，也不趋于无穷大。若 $p(t) = \exp\{-\Lambda(t)\}\eta$ ，相对脆弱性将趋于无穷大。若对于某个 $\epsilon \in \mathbb{R}_{>0}$，所有大于某个时间阈值的 t 都满足 $p(t) > \epsilon$，则无论为 Z_*选择何种 $\mathcal{P}$ 分布，相对脆弱性都将趋于零。

4 讨论

对于时间不变的离散单变量模型和共享脆弱性模型，我们研究了 RFV（CRF）的轨迹。我们证明，最小脆弱性类别的选择决定了这两个函数的长期行为。如果脆弱性分布的支撑集可按从小到大的顺序排列，且其中包含零点原子（例如基于经典的 $\mathcal{B}$、 $\mathcal{NB}$ 和 $\mathcal{P}$ 分布的模型），那么 RFV（CRF）将趋于无穷大。如果离散脆弱性分布不包含零点原子，RFV（CRF）将趋于零（一）。

在 RFV 函数形式方面表现出更大灵活性的离散分布族，也就是说，包含既有零点原子又无零点原子的分布类型的离散分布族，似乎更适合作为一种通用建模方法，并可作为诊断工具，用于寻找现有数据中存在治愈率的证据。Addams 族、零修正幂级数分布以及平移分布均符合这一标准。

本文讨论的分布中，由 Farrington 等人(2012年) 提出的 Addams族具有一项优势：治愈率的存在与否都不处于参数空间的边界。然而，目前尚无针对 Addams族的估计方法。另一种选择是采用零修正幂级数分布，通过附加参数调节治愈率。然而，当不存在治愈率时，该情形处于缩减与膨胀参数参数空间的边界，这可能导致对这类数据进行估计时出现困难

对于治愈率似乎不合适的数据。如第 3 节所示，要灵活调整相对脆弱性（RFV）的轨迹，一种直接的方法是借助附加参数移动脆弱性分布的支撑集，该参数的取值介于零与无穷大之间。然而在这种情况下，治愈率处于平移参数参数空间的边界。

这类平移模型和零修正模型有一个有趣的性质：RFV 不受单调性的限制。对于时间不变的脆弱性分布而言，这种性质十分罕见；例如，幂方差族、扩展广义伽马族和库默尔族都会产生单调变化的 RFV 轨迹（CRF）。关于这三个族及其 RFV 的研究，Farrington 等人(2012年) 中有所讨论。

我们已经证明，对于时间不变的离散单变量脆弱性分布和共享脆弱性分布，RFV（CRF）无法收敛到正渐近值（大于 1）。这与现有的连续分布族（例如库默尔族）形成对比。离散脆弱性分布的 Addams 族通过将伽马分布纳入其中作为连续特例，解决了这一问题；该特例具有恒定的 RFV（CRF），因而具有正渐近值（大于 1），除非脆弱性方差为零。如果长期非零但有限的 RFV 对当前数据的建模很重要，那么估计过程可能会趋近于这种情况。

如果我们脱离时间不变单变量或共享脆弱性模型的框架，则RFV可能趋近于一个正渐近值。如果允许Caroni 等人(2010)所引入的离散脆弱性模型中的支撑集移位参数依赖于时间，并且该支撑集移位参数从长期来看趋近于零，则RFV可能趋近于一个正渐近值。在具有离散脆弱性的多变量情形下，对于Cancho 等人(2022年)所引入的离散相关脆弱性模型类，CRF趋近于一个大于1的正常数。在这种情况下，RFV与CRF失去了联系。

我们的研究重点是随机效应不随时间变化的离散脆弱性模型。不过，我们也将研究结果推广到了分段常数离散时变共享脆弱性模型。对于这一类脆弱性模型，RFV（CRF）会趋于无穷大或零（一），具体取决于最终随机效应的支撑集中是否包含零。受Caroni 等人(2010) 在第2 节讨论的平移模型启发，我们可以通过脱离时不变的单变量共享脆弱性模型设定，并令脆弱性分布的支撑集移位参数在互不相交的时间区间上呈分段常数，从而利用这些模型生成更灵活、非单调的RFV（CRF）轨迹。这既能保持模型结构相对简单，也可能使RFV（CRF）在递减和递增阶段之间多次切换。此外，该模型是上述分段常数时变脆弱性模型的特例，因此，RFV（CRF）会趋于零（一）或无穷大，分别取决于最后一次移位是否大于零或等于零。

与现有标准连续脆弱性分布相比，离散脆弱性模型的RFV呈现出更多样的变化形态。这一特性或可用于开发具有灵活RFV（CRF）轨迹的时变相关离散脆弱性模型。

致谢。 本研究获得了德国研究联合会（DFG；德国研究基金会）的资助，项目编号为 UN400/2-1。感谢 Fabian K¨uck 博士对本文提出的意见。

附录 A 重新定义的 CRF 及其与 RFV 的联系

我们证明，对于 $J \geq 2$，CRF($\boldsymbol{t}$) $= 1+$RFV($\boldsymbol{t}$)。令 $S(\boldsymbol{t})$ 表示总体生存函数，即 $S(\boldsymbol{t}) = P(\boldsymbol{T} > \boldsymbol{t})$（给定协变量）。请注意，对于单个自变量的函数（例如 $\mathcal{L}$），我们使用拉格朗日记号表示导数；对于多个自变量的函数（例如 S，Λ），则使用莱布尼茨记号。给定 $T^{(j')} = t_{j'}, j' \neq j$ 时，$T_{(j)}$ 的总体风险率为

$$h^{(j)}(t_j \mid T^{(j')} = t_{j'}, T^{(l)} > t_l \ \forall l \neq j, j') = -\frac{\partial^2 S(\boldsymbol{t})/\partial t_j \partial t_{j'}}{\partial S(\boldsymbol{t})/\partial t_{j'}}.$$

将其改写为拉普拉斯变换的形式，可得

$$\begin{aligned} h^{(j)}(t_j \mid T^{(j')} = t_{j'}, T^{(l)} > t_l \ \forall l \neq j, j') &= -\frac{\mathcal{L}''(\Lambda_0(\boldsymbol{t}))}{\mathcal{L}'(\Lambda_0(\boldsymbol{t}))} \frac{\partial \Lambda_0(\boldsymbol{t})/\partial t_j \ \partial \Lambda_0(\boldsymbol{t})/\partial t_{j'}}{\partial \Lambda_0(\boldsymbol{t})/\partial t_{j'}} \\ &= -\frac{\mathcal{L}''(\Lambda_0(\boldsymbol{t}))}{\mathcal{L}'(\Lambda_0(\boldsymbol{t}))} \partial \Lambda_0(\boldsymbol{t})/\partial t_j. \end{aligned}$$

对 CRF 的分母进行类似推导可得

$$\begin{aligned} h^{(j)}(t_j \mid T^{(l)} > t_l \ \forall l \neq j) &= -\frac{\partial S(\boldsymbol{t})/\partial t_j}{S(\boldsymbol{t})} \\ &= -\frac{\mathcal{L}'(\Lambda_0(\boldsymbol{t}))}{\mathcal{L}(\Lambda_0(\boldsymbol{t}))} \partial \Lambda_0(\boldsymbol{t})/\partial t_j \end{aligned}$$

这得出 CRF$^{(j,j')}(\boldsymbol{t}) = \frac{\mathcal{L}''}{(\Lambda_0(\boldsymbol{t}))\mathcal{L}(\Lambda_0(\boldsymbol{t}))\ \mathcal{L}'(\Lambda_0(\boldsymbol{t}))^2}$。注意，基线风险率会抵消，因此，在时间不变的共享脆弱性模型中，j 和 j' 的选择与 CRF 无关，相应地可以省略上标。此外，CRF 可视为累积基线风险之和的函数，即 CRF($\boldsymbol{t}$) $=$ CRF($\Lambda_0(\boldsymbol{t})$)。

RFV 可以通过条件矩母函数表示，因此也可以通过条件拉普拉斯变换表示，即存活者的拉普拉斯变换。存活者的拉普拉斯变换 $\mathcal{L}_{\boldsymbol{t}}(s) = \int_0^\infty \exp\{-zs\} g(z \mid \boldsymbol{T} > \boldsymbol{t}) dz = \frac{\mathcal{L}(s+\Lambda_0(\boldsymbol{t}))}{\mathcal{L}(\Lambda_0(\boldsymbol{t}))}$ 可通过 $-\mathcal{L}'_{\boldsymbol{t}}(0)$ 求得 E($Z \mid \boldsymbol{T} > \boldsymbol{t}$)，以及 E($Z^2 \mid$

$\boldsymbol{T} > \boldsymbol{t}$) 可通过 $\mathcal{L}_t''(0)$ 求得。因此,

$$\begin{aligned}\mathrm{RFV}(\boldsymbol{t}) &= \frac{\mathrm{Var}(Z \mid \boldsymbol{T} > \boldsymbol{t})}{\mathrm{E}(Z \mid \boldsymbol{T} > \boldsymbol{t})^2} \\ &= \frac{\mathcal{L}''(\Lambda_0(\boldsymbol{t}))\mathcal{L}(\Lambda_0(\boldsymbol{t})) - \mathcal{L}'(\Lambda_0(\boldsymbol{t}))^2}{\mathcal{L}'(\Lambda_0(\boldsymbol{t}))^2} \\ &= \mathrm{CRF}(\boldsymbol{t}) - 1.\end{aligned}$$

与 CRF 类似，RFV 可以视为累积基线风险之和的函数。

附录B 所选模型的相对脆弱性形状

在 $\mathcal{ZMP}$ 模型中，RFV 关于一般时间 Λ 的导数为

$$\mathrm{RFV}'(\Lambda) = \mathrm{RFV}_{\mathcal{P}}(\Lambda)\left(1 + \frac{b}{r(\Lambda)}\right),$$

其中 $b = \frac{\phi^{-1}}{1-\phi}\exp\{-\eta\}$ 且 $r(\Lambda) = \mathrm{RFV}_{\mathcal{P}}(\Lambda)^2(1+\mathrm{RFV}_{\mathcal{P}}(\Lambda))^2 - \mathrm{RFV}_{\mathcal{P}}(\Lambda)\exp\left\{\frac{1}{\mathrm{RFV}_{\mathcal{P}}(\Lambda)}\right\}$。当$\phi \geq 1$ （膨胀）时，RFV 单调递增；而当 $\phi < 1$ （收缩）时，则可能存在驻点。驻点满足 $-b = r(\Lambda)$。具体而言， $r(0) = \exp\frac{\{\eta\}}{\eta^2+\eta+1}$ 且 $\lim_{\Lambda\to\infty} r(\Lambda) = 1$，并且 $r(\Lambda)$ 在 $\eta \leq 1$ 时严格递增，但在 $\eta > 1$ 时于区间 $(0,\infty)$ 内存在全局最小值，即在 Λ 处，$\mathrm{RFV}_{\mathcal{P}}(\Lambda) = 1$，且取值为 $\exp_3^{\{1\}}$。因此，对于 $\phi < 1$，RFV 中可能没有、存在一个或存在两个驻点。注意，在 $\phi < 1$的情况下， $0 < -b \leq 1$。

在驻点处计算的相对脆弱性的二阶导数 is

$$\frac{\mathrm{RFV}_{\mathcal{P}}(\Lambda) - 1}{(1 + \mathrm{RFV}_{\mathcal{P}}(\Lambda))^2 - \mathrm{RFV}_{\mathcal{P}}(\Lambda)}.$$

因此，如果驻点位于 Λ ，且相对脆弱性$_{\mathcal{P}}(\Lambda) = 1$，那么相对脆弱性只有一个驻点，该驻点为鞍点，并且相对脆弱性单调递增。在其他情况下，驻点要么是局部最小值，要么是局部最大值。如果相对脆弱性有两个驻点，则局部最大值之后是局部最小值。对于 $\phi > 0$，相对脆弱性在其尾部严格递增。当且仅当 $\phi = 0$ 时，相对脆弱性存在严格递减的尾部，即对于截断的 $\mathcal{P}$ 模型。

表 B1 列出了第 3 节讨论的相对效应模型的详细信息。对于表 B1 中列出的平移模型，我们假设 $p > 0$。关于 $p = 0$ 的相应信息可从非平移对应模型所在行获得。对于 $\mathcal{NB} + p$ 模型， $c(\Lambda)$ 定义为

$\nu(1-\pi)+p(\exp\{\Lambda\}-(1-\pi))$。常数 c^* 对于 $\nu>p>0$ 等于 $\ln\{\binom{(1-\pi)(\nu-p)}{p}\}$，对于 $\nu\le p$ 等于 $-\infty$。对于 $\mathcal{B}+p$ 模型，$c(\Lambda)$ 定义为 $\pi\exp\{-\Lambda\}(p+n)+p(1-\pi)$，且常数 c^* 等于 $-\ln\{\frac{p}{(1-\pi)\,\pi(p+n)}\}$。对于 $\mathcal{P}+p$ 模型，$c(\Lambda)$ 定义为 $\eta\exp\{-\Lambda\}+p$，常数 c^* 等于 $\ln\{\frac{\eta}{p}\}$。

附录C 命题1的证明

设 $\Lambda\in\{\sum_{j=1}^{J}\Lambda_0^{(j)}(t_j):\boldsymbol{t}\in\mathbb{R}_{\ge 0}^{J}\}$。证明将基于令 Λ趋于无穷大。注意，在这种情况下，我们隐含地指的是 $t_j\to\infty$。下面的证明基于以下假设：$\mathrm{E}(Z)=\mu_1$ 和 $\mathrm{E}(Z^2)=\mu_2$ 存在。然而，情况不一定如此。关于在第一个二阶矩不存在（在 $\boldsymbol{t}=0$ 处）时如何调整证明的说明，可在本节证明之后找到。

为了证明命题 1，我们需要坦纳里定理。

坦纳里定理
设 $\lim_{\Lambda^*\to\infty}v_k(\Lambda^*)=w_k$ 和 $F(\Lambda^*)=\sum_{k=1}^{\infty}v_k(\Lambda^*)$ 是收敛级数。若 $|v_k(\Lambda^*)|\le M_k$，其中 M_k 与 Λ^* 无关，且 $\sum_{k=1}^{\infty}M_k<\infty$，则 $\lim_{\Lambda^*\to\infty}F(\Lambda^*)=\sum_{k=1}^{\infty}w_k$。

Tannery 定理可见于Hofbauer（2002）。随后，Λ 将取代 Λ^*。函数 F 将是概率质量函数、加权生存函数或期望，因此根据 Ω 的不同，它将是收敛级数或有限和。命题1 的证明如下。

Proof Farrington 等人（2012年）证明，如果脆弱性是一个在 $Z=0$ 上具有正概率的混合分布，则相对脆弱性趋于无穷大（第685页）。这也包括 $Z>0$ 的分布为离散的情形，因此（i）成立。我们证明（ii）。

某个 $z\in\Omega$ 的条件概率质量函数，

$$
\begin{aligned}
g(z,\Lambda_0(\boldsymbol{t})) &\equiv P(Z=z\mid \boldsymbol{T}\ge\boldsymbol{t})\\
&=\frac{P(\boldsymbol{T}\ge\boldsymbol{t}\mid Z=z)g(z)}{P(\boldsymbol{T}\ge\boldsymbol{t})}\\
&=\frac{\exp\{-z\Lambda_0(\boldsymbol{t})\}g(z)}{\sum_k\exp\{-z_{(k)}\Lambda_0(\boldsymbol{t})\}g(z_{(k)})} \qquad \text{(C1)}\\
&=\frac{g(z)}{\sum_k\exp\{-[z_{(k)}-z]\Lambda_0(\boldsymbol{t})\}g(z_{(k)})}, \qquad \text{(C2)}
\end{aligned}
$$

其中求和遍历 Ω（各元素的下标）。

证明基于这样的思路：$\lim_{\Lambda\to\infty}g(z_{(1)},\Lambda)=1$ 和 $\lim_{\Lambda\to\infty}\mu_q(\Lambda)=\mathrm{E}(Z^q\mid \boldsymbol{T}\ge\boldsymbol{t})$ $q=1,2$。从技术上说，这是由 Tannery 定理推导得出的，该定理给出了在类似 $\lim_{\Lambda\to\infty}\sum_{k=1}^{\infty}$ 的表达式中交换极限次序的条件。

我们将两次使用坦纳里定理。首先，$\lim_{\Lambda\to\infty}g(z_{(1)},\Lambda)=1$ 可由分母中的求和式的 (C2) 和 $M_k=g(z_{(k)})$ 立即得出。因此，$\lim_{\Lambda\to\infty}g(z_{(k)},\Lambda)=0,\ \forall k>1$。其次，该定理应用于 $\mathrm{E}(z^q\mid T\ge t)$ $q=1,2$，这需要多做一些工作。

$Z \sim$	$\mathcal{L}(s)$	$\text{RFV}(\Lambda)$	$\text{RFV}'(\Lambda)$	Support
$\mathcal{NB}(\pi,\nu)$	$\left(\frac{\pi}{1-(1-\pi)\exp\{-s\}}\right)^{\nu}$	$\exp\{\Lambda\}/(1-\pi)\nu$	$>0 \forall \Lambda$	$\{0,1,\dots\}$
$\mathcal{NB}_{>0}(\pi,\nu)$	$\left(\frac{\pi}{\exp\{s\}-(1-\pi)}\right)^{\nu}$	$(1-\pi)\exp\{-\Lambda\}/\nu$	$<0 \forall \Lambda$	$\{\nu,\nu+1,\dots\}$
$\mathcal{B}(\pi,n)$	$((1-\pi)+\pi\exp\{-s\})^n$	$(1-\pi)\exp\{\Lambda\}/n\pi$	$>0 \forall \Lambda$	$\{0,\dots,n\}$
$\mathcal{P}(\eta)$	$\exp\{\eta(\exp\{-s\}-1)\}$	$\exp\{\Lambda\}/\eta$	$>0 \forall \Lambda$	$\{0,1,\dots\}$
$\mathcal{NB}(\pi,\nu)+p$	$\mathcal{L}_{\mathcal{NB}}(s)$ $\times\exp\{-sp\}$	$\exp\{\Lambda\}\nu(1-\pi)$ $\times c(\Lambda)^{-2}$	$>0\ \forall\Lambda:\Lambda<c^*$, $=0\ \forall\Lambda:\Lambda=c^*$, <0 else	$\{p,p+1,\dots\}$
$\mathcal{B}(\pi,n)+p$	$\mathcal{L}_{\mathcal{B}}(s)$ $\times\exp\{-sp\}$	$(1-\pi)\pi n\exp\{-\Lambda\}$ $\times c(\Lambda)^{-2}$	$>0\ \forall\Lambda:\Lambda<c^*$, $=0\ \forall\Lambda:\Lambda=c^*$, <0 else	$\{p,p+1,\dots\ p+n\}$
$\mathcal{P}(\eta)+p$	$\mathcal{L}_{\mathcal{P}}(s)$ $\times\exp\{-sp\}$	$\eta\exp\{-\Lambda\}/c(\Lambda)^2$	$>0\ \forall\Lambda:\Lambda<c^*$, $=0\ \forall\Lambda:\Lambda=c^*$, <0 else	$\{p,p+1,\dots\}$
$\mathcal{AF}(\alpha,\gamma)$	See Farrington et al (2012)	$\gamma\exp\{\alpha\Lambda\}$	For $\alpha<0$: $<0\forall\Lambda$	$\{-\alpha\nu,-\alpha(1+\nu),\dots\}$
			For $\alpha=0$: $=0\forall\Lambda$	$\mathbb{R}_{>0}$
			For $\alpha>\gamma$: $>0\forall\Lambda$	$\{0,\alpha,2\alpha,\dots\}$
			For $0<\alpha<\gamma, \frac{\gamma}{\gamma-\alpha}\in\mathbb{N}$: $>0\forall\Lambda$	$\{0,\alpha,2\alpha,\dots,b\alpha\}$
$\mathcal{ZMP}(\eta,\phi)$	$\frac{1-\phi\exp\{-\eta\}}{1-\exp\{-\eta\}}\mathcal{L}_{\mathcal{P}}(s)$ $+\frac{\phi-1}{\exp\{\eta\}-1}$	$\exp\{\frac{-1}{\text{RFV}_{\mathcal{P}}(\Lambda)}\}$ $\times\frac{\phi-1}{1-\phi\exp\{-\eta\}}$ $\times(1+\text{RFV}_{\mathcal{P}}(\Lambda))$ $+\text{RFV}_{\mathcal{P}}(\Lambda)$	For $\phi=0$: RFV can be strictly decreasing or can have inner global maximum.	$\{1,2,\dots\}$
			For $\phi>0$: RFV can be strictly increasing or can have up to two stationary points.	$\{0,1,\dots\}$

Table B1: Overview of some characteristics of discrete frailty distributions

为此，将证明存在一个 k_Λ^*，使得 $\frac{\partial g(z_{(k)},\Lambda)}{\partial\Lambda}=g'(z_{(k)},\Lambda)<0\ \forall k\geq k_\Lambda^*$，并且一旦某个导数在某处变为负值，那么对所有 $\Lambda>\Lambda^*$，它都将始终为负。接着，对于 $k\geq k_0^*$，有 $M_k=z^q(k)g(z_{(k)})$，其中 k_0^* 是满足 $g'(z_{(k)},\Lambda)|_{\Lambda=0}<0$ 的首个 k。请注意，在后一种情况下，我们所指的是右侧导数，即 $\Lambda\in[0,\infty)$。

关于 Λ 的条件概率质量函数导数，其中 $z\in\Omega$

$$g'(z,\Lambda)=\frac{g(z,\Lambda)^2}{g(z)}\left(\sum_k[z_{(k)}-z]\exp\{-[z_{(k)}-z]\Lambda\}g(z_{(k)})\right). \tag{C3}$$

请注意，这还要求在 (C1) 的分母中交换无穷求和与求导算子。可以通过证明级数 $\sum_k -z_{(k)}\exp\{-z_{(k)}\Lambda\}g(z_{(k)})$ 一致收敛来说明这一点；利用 $z_{(k)}\exp\{-z_{(k)}\Lambda\}g(z_{(k)})\leq z_{(k)}g(z_{(k)})\forall\Lambda$ 以及一阶矩 μ_1 的存在性即可确立这一点。对于 $z_{(1)}$，导数只包含非负项，且至少有一项为正，因为根据假设，$\mathrm{Var}(Z)>0$。因此，$g'(z_{(1)},\Lambda)>0\ \forall\Lambda$。令 k_Λ^* 表示首个导数为负的指标。这样的 k_Λ^* 必定存在，因为一阶导数为正（对所有 Λ 均如此），且关于 Ω 求和时所有概率之和为 1（给定 Λ）；也就是说，至少存在一个 k，其概率会随时间增加而降低。我们省略 k_Λ^* 中的指标，但强调它取决于（一般的）时间。通过重新表达导数，使 $z_{(k^*)}$ 成为参照点，即

$$g'(z_{(\tilde{k})},\Lambda)=\frac{g(z_{(\tilde{k})},\Lambda)^2}{g(z_{(\tilde{k})})}\times\left(\exp\{[z_{(\tilde{k})}-z_{(k^*)}]\Lambda\}\sum_k[z_{(k)}-z_{(\tilde{k})}]\exp\{-[z_{(k)}-z_{(k^*)}]\Lambda\}g(z_{(k)})\right),$$

可以看出，对于所有 k，$g'(z_{(\tilde{k})},\Lambda)<0,\ \forall\tilde{k}>k^*$。这是因为求和中决定符号的贡献项 $z_{(k)}-z_{(\tilde{k})}$ 小于其对应项 $g'(z_{(k^*)}\Lambda)$，$z_{(k)}-z_{(k^*)}$，而求和中的权重相对于参考值 k^* 保持不变。

由 (C3) 可知，对于所有 $\Lambda_+>\Lambda$ 和 $k>1$，$\frac{g'(z_{(k)},\Lambda_+)}{g(z_{(k)},\Lambda_+)^2}<\frac{g'(z_{(k)},\Lambda)}{g(z_{(k)},\Lambda)^2}$，因为导数求和中每个负贡献项的权重增加，而每个正贡献项的权重则减小。这意味着，一旦导数变为负值，随着时间增加，它将始终为负，即对于所有 $\Lambda>0$，$k_0^*\geq k_\Lambda^*$。因此，对于所有 $k\geq k_0^*$ 和 $\Lambda>0$：$g(z_{(k)})>g(z_{(k)},\Lambda)$。于是，$\mu_q(\boldsymbol{t})\leq\mu_q$，且 $\mu_q(\boldsymbol{t})$ 存在（因为单调递增级数 $\mathrm{E}(Z^q\mid\boldsymbol{T}\geq\boldsymbol{t})$ 有界于零与 μ_q 之间），其中 $q=1$ 为 2。现在，对于所有 $k\geq k_0^*$，$M_k=z^q(k)g(z_{(k)})\geq z^q(k)g(z_{(k)},\Lambda)$，且这与 Λ 无关。设对于所有 $k<k_0^*$，$M_k=z^q(k)\geq z^q(k)g(z_{(k)},\Lambda)$。那么，$\sum_{k=1}^\infty M_k=\mathrm{E}(Z^q)+\sum_{k=1}^{k_0^*-1}(1-g(z_{(k)}))z^q(k)$ 收敛，因为后半部分是有限和，最终可得出在 $\lim_{\Lambda\to\infty}\mathrm{E}(Z^q\mid T\geq t)$ 中交换极限的充分条件。

通过后一次应用坦纳里定理，$\lim_{\Lambda\to\infty}\mathrm{E}(Z^q\mid\boldsymbol{T}\geq\boldsymbol{t})=\frac{z^q_1}{g(z_{(1)})|_{()}}$ 结合前一次对上文 $\boldsymbol{T}\geq\boldsymbol{t}$ $z^q>0$（）应用坦纳里定理即可得出。由于根据假设(1)，

$$\lim_{\Lambda\to\infty}\mathrm{RFV}(\Lambda)=\lim_{\Lambda\to\infty}\frac{\mathrm{E}(Z^2\mid\boldsymbol{T}\geq t)-\mathrm{E}(Z\mid\boldsymbol{T}\geq t)^2}{\mathrm{E}(Z\mid\boldsymbol{T}\geq t)^2}$$

$$= \frac{z_{(1)}^2 - z_{(1)}^2}{z_{(1)}^2} = 0$$

证明完毕。 □

关于 $\boldsymbol{E}(Z)$,$\mathrm{V}(Z)$ 不存在时如何调整证明的说明：
若 $\mathrm{E}(Z) = \infty$ 和/或 $\mathrm{V}(Z) = \infty$ 不存在，可以认为证明从 $\boldsymbol{t} = 0$ 之后的某个时刻开始，此时 $\mathrm{E}(Z \mid \boldsymbol{T} \geq \boldsymbol{t}) \equiv \mu_1(\boldsymbol{t})$ 和 $\mathrm{E}(Z^2 \mid \boldsymbol{T} \geq \boldsymbol{t}) \equiv \mu_2(\boldsymbol{t})$ 存在。下文将证明，存在一个 $\boldsymbol{t}^*$，使得条件一阶和二阶矩存在。可以认为上述证明从时刻 $\boldsymbol{t}^*$ 开始，且证明中的所有 $\boldsymbol{t}$ 都大于 $\boldsymbol{t}^*$。阈值 k_0^* 必须重新定义为$k^*_{\Lambda(\boldsymbol{t}^*)}$。无条件矩 μ_1 和μ_2 必须分别替换为 $\mu_1(\boldsymbol{t}^*)$和 $\mu_2(\boldsymbol{t}^*)$。此外，时间自变量Λ 必须重新定义为 $\Lambda(\boldsymbol{t}^*) + \Lambda$和$\Lambda \in \{\sum_{j=1}^{J} \Lambda^{(j)}(t_j) - \Lambda^{(j)}(t_j^*) \colon \ \boldsymbol{t} \geq \boldsymbol{t}^*\}$。注意，若 $g(\boldsymbol{0}) > 0$，则 $g(0 \mid \boldsymbol{T} \geq \boldsymbol{t}) > 0$，且 $g(\boldsymbol{0}) = 0$也遵循相同规则。因此，与矩存在的情形一样，$z_{(1)}$ 决定了 RFV 的极限行为。

如果 $Z \mid \boldsymbol{T} \geq \boldsymbol{t}$ 的矩存在，则可以通过对 s 求导并令 $s = 0$，利用存活者的拉普拉斯变换 $\mathcal{L}_{\boldsymbol{t}}(s) = \mathcal{L}_{(s+\Lambda(\boldsymbol{t}))}\ \mathcal{L}(\Lambda(\boldsymbol{t}))$ 来推导这些矩。从 $\boldsymbol{t}^*$ 最新起，我们通过分子来证明这些矩存在。

拉普拉斯变换 $\mathcal{L}(s + \Lambda(\boldsymbol{t}))$ 在 $s \in (-\Lambda(\boldsymbol{t}) \ \infty)$ 上收敛。然而，我们只关心 $s = 0$ 处的导数（以及可能需要其在 $s > 0$ 处的值来计算存活者的生存函数）。因此，对于某个固定的 $r \in (0, 1 - \ ^2 \exp\{1\})$，我们将定义域限制为 $s \in (-r, \infty)$。我们选择 $\boldsymbol{t}^*$，使得 $\Lambda(\boldsymbol{t}^*) = 1$。由此，我们证明部分和序列 $\{S_n'(s \mid \boldsymbol{t})\}_{n=1}^{\infty} = \sum_{k=1}^{n} (^{-z_{(k)}})^q \exp\{-z_{(k)}(s + \Lambda(\boldsymbol{t}))\} g(^{z_{(k)}})$ 一致收敛，即 $\mathcal{L}(s + \Lambda(\boldsymbol{t}))$ 中每一项关于 s 的一阶（$q = 1$）和二阶（$q = 2$）导数之和一致收敛，从而证明对于每个 $s \in (-r, \infty)$，给定每个$\boldsymbol{t} \geq \boldsymbol{t}^*$，该级数的极限也是 $\mathcal{L}(s + \Lambda(\boldsymbol{t}))$ 的导数。注意，对于所有 k，且给定每个 $\boldsymbol{t} \geq \boldsymbol{t}^*$，$z^q(k) \exp\{-z_{(k)}(s + \Lambda(\boldsymbol{t}))\} g(^{z_{(k)}}) \leq z^q(k) \exp\{-z_{(k)}(1 - r)\} g(^{z_{(k)}})$。当 $z^q(k) < \exp\{z_{(k)}(1 - r)\}$ 时，对于所有 k，序列 $\{S_n\}_{n=1}^{\infty} = \{\sum_{k=n}^{\infty} (^{-z_{(k)}})^q \exp\{-z_{(k)}(1 - r)\} g(^{z_{(k)}})\}_{n=1}^{\infty}$ 收敛到 0（并且 $q \in \{1, 2\}$ (r 和 $\Lambda(\boldsymbol{t}^*)$ 的选取保证这些不等式成立）。因此，对于每个 $\epsilon > 0$，存在一个 $\tilde{n} \in \mathbb{N}$，使得 $|S_n| < \epsilon, \forall n > \tilde{n}$。这实质上证明了，当 $|S_n'(s \mid \boldsymbol{t}) - \sum_{k=1}^{\infty} (^{-z_{(k)}})^q \exp\{-z_{(k)}(s + \Lambda(\boldsymbol{t}))\} g(^{z_{(k)}})| \trianglelefteq |S_{n+1}| < \epsilon$ 时，对于所有 $n > \tilde{n}$ 以及所有 $s \in (-r, \infty)$，级数 $S'(s \mid \boldsymbol{t})$ 一致收敛。对所有 $\boldsymbol{t} \geq \boldsymbol{t}^*$，这一结论均成立。因此，$\mathcal{L}_{\boldsymbol{t}}(s)$ 在 $s = 0$ 处的一阶和二阶导数存在，因而 $\mu_1(\boldsymbol{t})$ $\mu_2(\boldsymbol{t})$ 以及 RFV（$\boldsymbol{t}$）也对所有 $\boldsymbol{t} \geq \boldsymbol{t}^*$ 存在。注意，这里并未声称 $\boldsymbol{t}^*$ 是两个条件矩都存在时的最低阈值。

将证明调整为适用于离散分段常数时变脆弱性的说明：

如第3节末尾所述，该证明可调整后用于某些随时间变化的（共享）脆弱性模型。考虑一个分段常数的时变（共享）脆弱性模型。若 $t \in [t_{q-1}, t_q)$，则脆弱性 $Z(t) = Z_q$。脆弱性随机变量 Z_1、……、Z_Q 是离散的。Z_q 的有序支撑集为 $\{z_{q,(1)}, z_{q,(2)},\ldots\}$。$Z_q$ 的支撑集大小为 K_q，其可能为无穷大。Z_Q 的概率质量函数记为 g_Q。令 $\Lambda_q = \sum_{j=1}^{J} \int_{t_{q-1}}^{t_q} \lambda_0^{(j)}(t)dt$, $\Lambda_Q(\boldsymbol{t}) = \sum_{j=1}^{J} \int_{t_{Q-1}}^{t_j} \lambda_0^{(j)}(u)du$，其中 $\boldsymbol{t} \geq t_{Q-1}$，且 $\Lambda_Q \in \{\Lambda_Q(\boldsymbol{t}) \mid \boldsymbol{t} \geq t_{Q-1}\}$。此外，我们记 $c(z) = \sum_{k_1=1}^{K_1} \cdots \sum_{k_{Q-1}=1}^{K_{Q-1}} \exp\{-\sum_{q=1}^{Q-1} z_{q,(k)}\Lambda_q\} P(Z_1 = z_{1,(k_1)} \cdots, Z_{Q-1} = z_{Q-1,(k_{Q-1})} \mid Z_Q = z)$。注意，$c(z) = P(\boldsymbol{T} > t_{Q-1} \mid Z_Q = z)$，因此其值介于零和一之间。对于 $\boldsymbol{t} \geq t_{Q-1}$，$P(\boldsymbol{T} > \boldsymbol{t} \mid Z_Q = z)$ 可简化为 $\exp\{-z\Lambda_Q(\boldsymbol{t})\}c(z)$。此外，$P(\boldsymbol{T} \geq \boldsymbol{t}) = \sum_{k_1}^{K_1} \cdots \sum_{k_Q}^{K_Q} P(\boldsymbol{T} \geq \boldsymbol{t} \mid Z_1 = z_{1,(k_1)} \cdots, Z_Q = z_{Q,(k_Q)}) P(Z_1 = z_{1,(k_1)} \cdots, Z_{Q-1} = z_{Q-1,(k_{Q-1})} \mid Z_Q = z_{Q,(k_Q)}) g_Q(Z_Q = z_{Q,(k_Q)})$。借助前面的定义，可将其简化为 $\sum_{k=1}^{K_Q} \exp\{-z_{Q,(k)}\Lambda_Q(t)\} g_Q(z_{Q,(k)}) c(z_{Q,(k)})$，其中 $\boldsymbol{t} \geq t_{Q-1}$。

为将证明推广到给定的时变（共享）脆弱性情形，需要进行两处调整。首先，证明假定从 t_{Q-1} 开始，即假定向量 $\boldsymbol{t}$ 的每个元素都大于 t_{Q-1}。其次，求导时应对 Λ_Q 求导，而不是对 Λ 求导。“M_k”仅在记号上作相应调整，分别记为 $z_{Q,\,k}$ 或 $z_{Q,\,k} g_Q$（$z_{Q,\,k}$）。为说明这一点，我们重新考察 P（$Z_Q = z \mid \boldsymbol{T} \geq \boldsymbol{t}$），并将其记号调整为适用于时变情形的形式，即令其为 g_Q（z，Λ_Q（$\boldsymbol{t}$））。

给定生存至 $\boldsymbol{t} \geq t_Q$，$Z_Q = z$ 的概率可表示为 s

$$
\begin{aligned}
g_Q(z, \Lambda_Q(\boldsymbol{t})) &= \frac{P(\boldsymbol{T} \geq \boldsymbol{t} \mid Z_Q = z) g_Q(z)}{P(\boldsymbol{T} \geq \boldsymbol{t})} \\
&= \frac{\exp\{-z\Lambda_Q(\boldsymbol{t})\} c(z) g_Q(z)}{\sum_{k=1}^{K_Q} \exp\{-z_{Q,(k)}\Lambda_Q(\boldsymbol{t})\} g_Q(z_{Q,(k)}) c(z_{Q,(k)})}.
\end{aligned} \tag{C4}
$$

现在，c（z）g_Q（z）可以重新定义为 g（z），而（C4）是（C1）的时变对应形式。由于 $0 \leq c$（z）≤ 1，证明的其余部分不受影响。因此，当 $z_{Q,\,(1)} > 0$ 时，RFV 趋近于零；而当 $z_{Q,\,(1)} = 0$时，RFV 趋近于无穷大。

参考文献

Aalen OO、Borgan Ø、Gjessing HK（2008年）《事件史分析：过程点视角》。施普林格，多德雷赫特

Anderson JE, Louis TA, Holm NV 等（1992）二元生存分布的时间依赖关联度量。《美国统计学会杂志》87(419):641–650。https://doi.org/10.2307/2290199

¨Ata N, Ozel G（2013）基于离散复合泊松过程的脆弱性模型的生存函数。《统计计算与模拟杂志》83(11):2105–2116。
https://doi.org/10.1080/00949655.2012.679943

Begun AZ、Iachine IA、Yashin AI（2000）个体脆弱性的遗传本质：两种方法的比较。《双胞胎研究与人类遗传学》3(1)：51–57。
https://doi.org/10.1375/twin.3.1.51

Bijwaard G（2014）具有脆弱性的多状态事件史分析。《人口学研究》30：1591–1620。https://doi.org/10.4054/DemRes.2014.30.58

Cancho VG、Zavaleta KEC、Macera MAC 等（2018）一种由离散脆弱性引起离散度的贝叶斯治愈率模型。《统计应用与方法通讯》25(5)：471–488。https://doi.org/10.29220/CSAM.2018.25.5.471

Cancho VG、Macera MAC、Suzuki AK 等（2020）一种由离散脆弱性引起离散度的新型长期生存模型。《寿命数据分析》26(2)：221–244。
https://doi.org/10.1007/s10985-019-09472-2

Cancho VG, Barriga G, Le˜ao J 等（2021）基于离散脆弱性构建的生存模型：用于对具有长期存活者和变化点的寿命数据进行建模。《统计学通讯：理论与方法》50(5):1161–1172。https://doi.org/10.1080/03610926.2019.1648826

Cancho VG, Suzuki AK, Barriga GDC 等（2022年）基于离散脆弱性构建的多变量生存模型。《统计学通讯：模拟与计算》51(11):6572–6590。
https://doi.org/10.1080/03610918.2020.1806323

CaroniC, Crowder M, Kimber A（2010）含离散脆弱性的比例风险模型。《寿命数据分析》16(3):374–384。https://doi.org/10.1007/s10985-010-9151-3

Clayton DG (1978年) 双变量生命表中的关联模型及其在慢性病发病率家族倾向流行病学研究中的应用。《生物统计学》 65(1):141–151.
https://doi.org/10.2307/2335289

Duchateau L, Janssen P (2008年) 《脆弱性模型》。施普林格，多德雷赫特

Farrington CP, Unkel S, Anaya-Izquierdo K(2012年) 相对脆弱性方差与共享脆弱性模型。《英国皇家统计学会杂志 B辑》 74(4):673–696.
https://doi.org/10.1111/j.1467-9868.2011.01021.x

Gasperoni F、Ieva F、Paganoni AM 等（2020）用于分层生存时间数据的非参数脆弱性Cox模型。Biostatistics 21(3)：531–544。

https://doi.org/10.1093/biostatistics/kxy071

Hofbauer J (2002) 关于相关恒等式的一个简单证明。《美国数学月刊》109(2):196–200. https://doi.org/10.1080/00029890.2002.11919855

Hougaard P (1984) 异质人群的生命表方法：描述异质性的分布。《生物统计学》71(1):75–83.https://doi.org/10.2307/2336399

Hougaard P (2000)《多变量生存数据分析》。施普林格，纽约，https://doi.org/10.1007/978-1-4612-1304-8

Mohseni N, Maboudi AAK, Baghestani A 等 (2020) 霍奇金淋巴瘤患者确诊后的离散脆弱性治愈率模型。《生物科学进展档案》11(4):15–22.https://doi.org/10.22037/aab.v11i4.32576

Molina KC, Calsavara VF,Tomazella VD, 等（2021）零修正幂级数离散脆弱性所诱导的生存模型：在一组黑色素瘤数据中的应用。《医学研究统计方法》30(8):1874–1889。 https://doi.org/10.1177/09622802211011187

Oakes D（1989）脆弱性所诱导的双变量生存模型。《美国统计学会杂志》84(406):487–493。 https://doi.org/10.2307/2289934

Palloni A, Beltr´an-S´anchez H（2017）离散巴克脆弱性与高龄时扭曲的死亡率动态。《人口学》54(2):655–671。 https://doi.org/10.1007/s13524-017-0548-4

Pickles A, Crouchley R（1994）生存和事件数据脆弱性模型的推广与应用。《医学研究统计方法》3(3):263–278。 https://doi.org/10.1177/096228029400300305

de Souza D, Cancho VG,Rodrigues J, 等（2017）生存分析中由脆弱性诱导的贝叶斯治愈率模型。《医学研究统计方法》26(5):2011–2028。https://doi.org/10.1177/0962280217708671

Wienke A（2010）生存分析中的脆弱性模型。Chapman & Hall/CRC Biostatistics 级数，泰勒-弗朗西斯，Boca Raton，https://doi.org/10.1201/9781420073911