扩展*saemix* 软件包，使*R* 能够拟合非高斯结局

预印本

Emmanuelle Comets*
巴黎西岱大学 &
巴黎北索邦大学
法国国家健康与医学研究院、IAME，法国巴黎75018区
雷恩大学，法国国家健康与医学研究院，法国高等公共卫生学院
Irset UMR _S 1085
F-35000 雷恩 ，法国
emmanuelle.comets@inserm.fr

Maud Delattre
巴黎萨克雷大学
法国国家农业、食品与环境研究院，MaIAGE
法国茹伊昂若萨78350
maud.delattre@inrae.fr

Belhal Karimi
巴黎综合理工学院
萨克雷路，帕莱索
法国
belhal.karimi@polytechnique.edu

2026年2月27日

通讯作者：
Emmanuelle Comets
法国国家健康与医学研究院 - UMR1137
比沙医学培训与研究单位
亨利·于沙尔街16号
巴黎75018区, 法国
电话：(33) 6 25 82 49 50
emmanuelle.comets@inserm.fr

arXiv:2603.03154v1 [stat.ME] 3 Mar 2026

*通讯作者

摘要

背景与目的：临床试验中越来越多地收集纵向数据，以增进对治疗作用和疾病进展的了解。随后，可以对连续生物标志物（如目标激素浓度或病毒载量）的变化轨迹进行建模，并将其与康复或住院等事件的发生联系起来。然而，还可以收集许多其他类型的纵向数据，例如疼痛评分（分类或二元）、发作次数（计数）或事件发生情况（生存）。纵向数据首选的统计分析方法是非线性混合效应模型（NLMEM），因为这类模型在刻画总体潜在变化趋势的同时，能够自然地处理个体间轨迹差异。尽管已有若干商业软件包可用于估计此类模型的参数，例如 NONMEM 和 Monolix，但 R 的 saemix 软件包是少数开源解决方案之一，也是最灵活的解决方案之一。本文介绍该软件包的扩展版本，它支持多种非高斯数据模型。

方法：saemix 软件包通过随机逼近期望最大化（SAEM）算法估计参数。在该软件包中，非高斯模型通过其对数似然函数进行指定，使用户能够最大程度地控制模型构建。我们还扩展了基于视觉预测检验的探索性图和诊断图，并采用自助法估计参数不确定性。我们向软件包中添加了多个数据集，展示不同类型的数据和分析，并提供了更详尽的笔记本，说明代码及其输出。我们还基于趾甲数据集开展了一项模拟研究，以评估 saemix 估计二元模型参数的性能。

结果：saemix 在模拟研究中表现良好，能够恢复真实参数值，且在不同参数初始值下均保持稳定。不同示例的结果与此前对这些数据的分析一致。诊断图和统计准则有助于选择合适的模型。我们还实现了一种算法，可联合搜索协变量模型和个体间变异性模型，以构建协变量模型；该算法应用于一个包含肺癌患者生存数据的数据集。

结论： saemix 是一款功能强大且完全开源的非线性混合效应模型参数估计工具。该软件包已扩展为支持非高斯模型，并提供灵活的接口，允许用户指定自己的似然函数。

关键词 非线性混合效应模型 · SAEM 算法 · saemix · R 软件包 · 参数估计 · 纵向数据 · 事件发生时间 · 有序数据 · 计数数据 · 二元数据

1 引言

纵向数据在临床试验期间常规采集，用于监测生物标志物、药物暴露或临床状况的变化。对这些响应变量进行建模，既可用于常规监测，也可用于跟踪或预测疾病轨迹 [Hendrickx 等人， 2024], 并且生物标志物轨迹可通过联合模型与生存相关联 [Desmée 等人， 2017]。非线性混合效应模型（NLMEM）用于描述纵向数据，因为它们适用于来自同一总体的受试者中重复测量的层次结构。无论置于最大似然还是贝叶斯框架中，都需要特定算法来规避技术挑战，例如计算似然，而在非线性混合效应模型中似然没有闭式形式。在最大似然估计（MLE）框架中，已提出多种算法，可分为两大类别：最小化（负）对数似然近似的基于梯度的算法 [Lindstrom 和 Bates， 1990] 以及将个体参数视为缺失数据并使用完全对数似然的 EM 算法。贝叶斯框架对感兴趣的参数指定先验分布，以计算后验分布。实现这些方法的软件包括 NONMEM [Beal 等人，1989年-2022],Monolix [Lavielle， 2014] 或较新的 Pumas [Rackauckas 等人， 2020], 等商业软件，并包含专门例程，用于处理药代动力学/药效学数据的特殊性，如给药方案或重复场合。在更通用的软件（如 R [R Development Core Team， 2006] ）中拟合非线性混合效应模型也是可能的。R 的基本安装附带 nlme 软件包，作为十五个推荐软件包之一 [Pinheiro 等人， 2023] ，实现了一阶条件近似（FOCE）算法 [Pinheiro 和 Bates， 2000]。还开发了其他软件包，如 lme4 [Bates 等人， 2015], ，其中似然要么基于非线性模型函数的线性化，要么使用自适应高斯求积进行计算。对于贝叶斯用户，rstan 软件包 [Stan Development Team， 2023] 提供了使用哈密顿蒙特卡罗连接 Stan 软件的接口 [Carpenter 等人，2017]。saemix（Comets 等人，2011年）是首个实现随机近似 EM（SAEM）算法的 R 软件包 [Kuhn 和 Lavielle， 2005], ，该算法在理论和实践上均已被证明在非线性混合效应模型中表现极为出色。

所研究的结局类型日趋多样，不再局限于药物浓度或生物标志物水平等连续测量。R 中有一些专用软件包可用于拟合某一类模型，例如用于广义线性混合效应模型的 glmm [Hadfield，2010], ，可处理计数或分类数据；以及用于生存数据参数化建模的 flexsurv [。SAEM 算法已扩展至非高斯结局，Savic 等人对此进行了研究，并表明对于计数和分类数据，与基于对数似然近似的算法相比，该算法性能更优 [Savic and Lavielle, 2009, Savic 等人，2010]。本文介绍 saemix 软件包的 3.2 版 [Comets 等人，2017], ，该版本扩展后可用于离散型和生存型结局模型 [Karimi 等人，2020]。这个新版本可从 CRAN（综合 R 档案网络）获取，是一款全面且灵活的工具，可用于在各种情境下拟合非线性混合效应模型。其他值得注意的新增内容包括：利用新版 R 的 npde 软件包改进连续响应模型的诊断图 [Comets 等人, 2008, Comets 和 Mentré，2021], ，以及针对新型结局的探索性图形和基于模拟的诊断。

本文旨在介绍 saemix 软件包近期针对非线性混合效应模型中非高斯数据分析所作的扩展。我们通过二元、分类、计数和事件发生时间（TTE）结局的案例研究，展示这些新功能。除此之外，我们还通过一项基于二元模型的模拟研究评估扩展后软件包的性能，并在附录中提供计数数据的其他结果。讨论部分将探讨该软件包的灵活性所支持的进一步扩展。

2 统计方法

2.1 模型

关于非线性混合效应模型的详细介绍，可参见 [Davidian 和 Giltinan，1995；Pinheiro 和 Bates，2000；Lavielle， 2014]。

连续结局：我们假设观测值服从高斯分布，并可写出以下一般非线性混合效应模型，用于描述受试者 i 的第 j 个观测值， $y_{ij} \in \mathrm{R}$:

$$\begin{aligned} y_{ij} &= f(x_{ij}, \psi_i) + g(x_{ij}, \psi_i, \sigma)\epsilon_{ij} \ , \ 1 \le i \le N \ , \ 1 \le j \le n_i \\ \psi_i &= H(\mu, c_i, \eta_i) = H(\phi_i), \quad \eta_i \sim_{i.i.d.} \mathcal{N}(0, \Omega) \\ \epsilon_{ij} &\sim_{i.i.d.} \mathcal{N}(0,1) \end{aligned} \tag{1}$$

其中:

- N 是受试者的数量,
- n_i 是受试者 i的观测值数量,
- f 是定义纵向结局行为的结构模型，取值于R,
- 回归变量，或称设计变量，（x_{ij}）假定为已知，个体协变量 c_i同样假定为已知,

- 对于受试者 i，向量 $\psi_i = (\psi_{i,l};\ 1 \le \ell \le n_\psi) \in \mathrm{R}^{n_\psi}$ 是模型 f尺度上的 n_ψ 个体参数向量，而 ϕ_i 是正态分布尺度上的 n_ψ 个体参数向量,
- μ 是一个大小为 n_μ 的未知固定效应向量,
- η_i 是一个大小为 n_η 的未知正态分布随机效应向量。

我们假设 ϵ_{ij} 和 η_i 相互独立。进一步假设，经过逆变换后，参数 ϕ_i 可以表示为固定效应 μ、协变量 c_i 和随机效应 η_i的线性组合。连接函数 H 用于将 ϕ_i 变换为进入函数 f的参数。残差方差由实值函数 g 和一些未知参数 σ定义。saemix 中可以使用不同的误差模型，包括常数型（$g = \sigma = a$）、比例型（ $g = b\,f$和 $\sigma = b$）、组合型（ $g = \sqrt{a^2 + b^2\,f^2}$ 和 $\sigma = (a, b)$ ）以及指数分布型；对于指数分布型，我们将观测值变换为 $y = f\exp(a\epsilon)$，并使用常数型误差模型。

离散结局： 离散模型和生存时间模型通过观测值的分布进行定义，而用于定义个体参数的第二层模型保持不变。对于任意个体 i，将观测值（$y_{ij},\ \ 1 \le j \le n_i$）视为一系列条件独立随机变量，其取值属于固定且有限的名义类别集合 $\{M_1, M_2, \ldots, M_K\}$，则该模型完全由概率质量函数 $\mathrm{P}(y_{ij} = M_k|\psi_i)$（适用于 $k = 1$、... ...、 K和 $1 \le j \le n_i$）定义。对于给定的（$i,\ \ j$）， K个概率之和为 1，因此实际上只需定义其中的 $K - 1$ 个。以最一般的方式，只要模型定义了概率分布，任何模型都可以被考虑；也就是说，对于每个 k, $\mathrm{P}(y_{ij} = M_k|\psi_i) \in [0,1]$，并且 $\sum_{k=1}^{K} \mathrm{P}(y_{ij} = M_k|\psi_i) = 1$。有序数据进一步假设类别具有顺序，即存在一种顺序 $\preceq$ 使得

$$M_1 \preceq M_2 \preceq \ldots \preceq M_K$$

例如，我们可以考虑疼痛等级（轻度 $<$ 中度 $<$ 重度），或者离散量表上的评分，比如从 1 到 10。与其定义每个类别的概率，不如定义累积概率 $\mathrm{P}(y_{ij} \preceq M_k|\psi_i)$（适用于 $k = 1$、.....、 $K - 1$），或者反过来定义 $\mathrm{P}(y_{ij} \succeq M_k|\psi_i)$（适用于 $k = 2$、.....、 K），可能更方便。只要模型定义了概率分布，即满足

$$0 \le \mathbb{P}(y_{ij} \preceq M_1|\psi_i) \le \mathbb{P}(y_{ij} \preceq M_2|\psi_i) \le \cdots \le \mathbb{P}(y_{ij} \preceq M_K|\psi_i) = 1 \quad \text{and} \sum_{k=1}^{K} \mathbb{P}(y_{ij} = M_k|\psi_i) = 1$$

这个定义很容易扩展到计数数据。计数数据可以取多个可能值，甚至可能有无穷多个，因此通常用少量参数定义一个概率分布，例如泊松分布（示例见第 4.3 节的案例研究）。

事件发生时间结局： 设 T_i 表示事件发生时间，δ_i 表示个体 i 的指示变量；若观察到事件，则 $\delta_i = 1$；若 T_i 为删失时间，则 $\delta_i = 0$。个体 i 发生事件的瞬时风险由风险函数描述:

$$h(t, \psi_i) = \lim_{dt \to 0^+} \frac{Pr(t \leq T_i < t + dt | T_i \geq t; \psi_i)}{dt} \tag{2}$$

参数风险函数通过对风险函数积分至时间 t，可确定在时间 t 之前未发生事件的概率 S:

$$S(t, \psi_i) = \exp\left(-\int_0^t h(x, \psi_i) dx\right) \tag{3}$$

对于个体 i，在给定个体参数 ψ_i的条件下，恰好在时间 t观测到事件的似然，是生存函数（=即在 t之前**未**发生事件的概率）与风险函数（在 t观测到事件的概率）的乘积:

$$l(t, \delta_i, \psi_i) = S(t, \psi_i)\ h(t, \psi_i)^{\mathbb{1}_{\delta_i = 1}} = S(t, \psi_i)\ h(t, \psi_i)^{\delta_i}.$$

2.2 SAEM算法

SAEM 算法最初由 [Kuhn 和 Lavielle 于 2005]提出。下面简要介绍该算法；更多详情请参阅 [Lavielle, 2014，Comets 等人 2017]，或 saemix 软件包的文档（`https://github.com/iame-researchCenter/saemix/blob/fcf6231d6ee8f89af01cc1a81642472560c7eb0e/docsaem.pdf`）。

群体参数的估计

SAEM 算法计算未知参数集合 $\theta = (\mu,\ \Omega\ \sigma)$ 的最大似然估计量，即使对数似然最大化的参数 θ：

$$LL(\theta | y) = \sum_{i=1}^{N} \log\left(p(y_i | \theta)\right) = \sum_{i=1}^{N} \log\left(\int p(y_i | \theta, \psi_i) p(\psi_i | \theta) d\psi_i\right). \tag{4}$$

由于当 f 为非线性时积分难以求解，EM算法会在 E步和 M步之间迭代：E步在第 k 次迭代时计算完全对数似然的条件期望

$$Q_k(\theta) = \sum_{i=1}^{N} \mathbb{E}[\log p(y_i, \psi_i | \theta) | y_i, \theta_{k-1}] \tag{5}$$

而 M步则使用使 $Q_k(\theta)$ 最大化的值 θ_k 更新对 θ 的估计。根据 [Dempster 等人（1977）和 Wu, 1983], 在一般正则条件下，EM序列 (θ_k) 收敛到观测似然的一个驻点（另见附录A）。当回归函数 f 不与随机效应线性相关时，E步无法以闭式形式执行，因此由随机过程取代。在 SAEM 的第 k 次迭代中:

- 模拟步骤：从条件分布 $p(\cdot | y_i;\ \theta_k)$ 中抽取 $\psi_{(k)\,i}$，$1 \leq i \leq N$。
- 随机逼近：按照以下方式更新 $Q_k(\theta)$，使其逼近式 (5)

$$\tilde{Q}_k(\theta) = \tilde{Q}_{k-1}(\theta) + \gamma_k(\log p(y_i, \psi_i^{(k)} | \theta) - \tilde{Q}_{k-1}(\theta)) \tag{6}$$

  其中 $(\gamma_k)_k$ 是一列步长，在 $[0, 1]$, 中递减并趋近于 0，且满足 $\sum_k \gamma_k = \infty$ 和$\sum_k \gamma_k^2 < \infty$。

- 最大化步骤：根据 θ_k 进行更新

$$\theta_{k+1} = \operatorname{Arg}\max_{\theta} \tilde{Q}_k(\theta)$$

Delyon 等人 [、 1999] 证明，在非常一般的条件下，SAEM 会收敛到观测值似然的一个极大值（局部或全局）。
在模拟步骤中，无法直接从其条件分布中抽取个体参数 $\psi^{(k)} = \{\psi_i^{(k)}, i = 1, \ldots, N\}$。相反，使用马尔可夫链蒙特卡罗（MCMC）方法对其进行模拟，最常用的是 Metropolis–Hastings（MH）算法。可以使用连续且不同的转移核及其相关提议分布（详情请参见 [Lavielle， 2014]）。

在最大化步骤中，根据模型的充分统计量更新群体参数。对于某些模型结构（组合误差模型、无个体间变异的参数等），可能还会执行额外的最小化步骤，具体取决于模型是否属于曲指数族 [Lavielle， 2014]。

迭代阶段： 为了有效探索参数空间，算法分为两个阶段。在前 K_1 次迭代（探索阶段），我们通过一个包含 5 次迭代的预热序列来初始化算法，并在每次迭代后丢弃样本；随后通过设置 $\gamma_k = 1$，让算法探索参数空间。在平滑阶段（K_2 次迭代），我们设置 $\gamma_k = 1/（k - K_1 + 1）$，以确保算法几乎必然收敛至最大似然估计。默认情况下，在探索阶段的前半段实施模拟退火，以增大提议分布的方差。

为了提高整体收敛性，可以运行多个 L 马尔可夫链，而不是只运行一条；这意味着在模拟步骤中抽取序列 $\psi^{(k,1)}$,..., $\psi^{(k,L)}$，然后结合随机近似和蒙特卡罗近似：

$$\tilde{Q}_k(\theta) = \tilde{Q}_{k-1}(\theta) + \gamma_k \left(\frac{1}{L} \sum_{\ell=1}^{L} \log p(y_i, \psi_i^{(k,\ell)} | \theta) - \tilde{Q}_{k-1}(\theta) \right)$$

f-SAEM： 快速 SAEM 算法于 3.0 版加入 saemix，以加快最大似然估计的收敛速度 [Karimi 等人， 2020]。模拟步骤中新增了第四个核，使用一个提议分布来近似个体参数的条件分布。对于连续数据模型，提议分布的协方差通过对模型进行一阶线性化得到；对于非高斯模型，则通过拉普拉斯近似得到。

似然估计

式（4）中定义的非线性混合效应模型的似然无法以闭式形式计算。一种方法是将该似然近似为高斯模型的似然；这个高斯模型是对函数 f 在个体参数的预测值处进行线性化后，从非线性混合效应模型推导得到的（ϕ_i, $1 \leq i \leq N$）。
或者，也可以通过蒙特卡洛方法或求积方法，在不对模型作任何近似的情况下估计似然。在 saemix 中，随机积分通过重要性抽样方法实现，而数值积分则采用高斯-埃尔米特求积方法。对于这两种方法，计算质量取决于个体分布条件均值和方差的估计值。

不确定性的估计

对于连续结局模型，saemix 默认通过基于期望 Fisher 信息矩阵（FIM）的渐近近似来估计参数估计值的不确定性。FIM 使用模型函数的线性近似计算 f （也称为 FOCE 方法）。由于这种近似对离散结局模型的效果不佳，当前版本的 saemix 不会为这些模型计算线性化 FIM。已有研究为非线性混合效应模型（NLMEM）提出了自助法 [Comets 等人， 2021]，且该方法已在软件包中实现。基于 MH 算法的替代方法正在研究中 [Guhl 等，2022，Delattre 和 Kuhn，2023，Lavalley-Morelle 等人， 2023]。

个体参数的估计

模型参数估计完成后，通过个体参数的条件分布估计个体参数（ψ_i）。我们首先估计服从正态分布的个体参数（ϕ_i），再通过变换 $\psi_i = H(\phi_i)$推导出（ψ_i）的估计值。通过模拟步骤中使用的 MCMC 过程估计条件分布：以总体估计值作为先验分布进行迭代，直至达到预设容差。随后，可将条件分布的众数（最大后验）或均值作为个体参数的估计值，而条件标准差则用于估计其不确定性。

3 saemix 3

该saemix 软件包：该软件包的第一个版本于 2011 [Comets 等人 2011] 上传至 CRAN，是原始 Monolix Matlab 代码的 R 版本 [Lavielle， 2014]。此后，软件包定期更新，加入新功能以便于模型构建和评估；此外，一篇展示 saemix 与 nlme 和 nlmer 相比具有良好性能的论文也已发表 2017 [Comets 等人， 2017]。对非高斯结局的扩展已纳入 3.0 版；截至提交日期，当前版本为 3.5（2026 年 2 月）。非线性混合效应模型的诊断图是诊断模型缺陷的有用工具 [Nguyen 等人， 2017]；在 3.0 版转而使用 ggplot 时 2 [Wickham， 2016] ，我们更改了 saemix 生成的图，以便使用配套软件包 npde [Comets 和 Mentré， 2021]。此外，saemix 还实现了适用于非线性混合效应模型的贝叶斯信息准则（BIC） [Delattre 等人， 2014]，以及用于自动构建协变量模型和变异结构的逐步算法 [Delattre 和 Poursat， 2020] 。最后，我们新增了选项，可在估计过程的前几次迭代中切换至 [Karimi 等人提出的快速 SAEM 算法 2020] ，以加快收敛。

软件包主要函数的详细操作指南可在文档中找到，文档可从 GitHub 上的公共代码库下载（ `https://github.com/iame-researchCenter/saemix`）。为简化本文内容，代码本身请参阅随论文提供的笔记本。软件包的开发版本可在开发版 GitHub 代码库（`https://github.com/saemixdevelopment/saemixextension`）中获取；所有代码以及其他分析也都通过针对每种结局类型的专用笔记本提供，位于文件夹（`https://github.com/saemixdevelopment/saemixextension/blob/master/paperSaemix3/`）中。

软件包扩展：为了在 saemix 中拟合离散结局模型，我们编写模型函数来计算观测值的对数似然。与纵向结局建模的一个重要区别是，在进行这项计算时，我们需要将观测值作为预测变量传入，因此必须将其纳入 saemixData 对象。

可通过 plotDiscreteData() 函数绘制图形，以便根据结局类型探索数据；指定结局类型为二元、分类、计数或事件即可。包中的许多函数用法相似，例如拟合后的结果提取或 summaryfunctions，而其他函数则会根据结局类型进行调整。例如，预测函数会根据拟合模型预测观测值的对数似然。与纵向结局的另一个区别是，由于用户可以直接编写对数似然，因此默认情况下没有与之关联的模拟函数。若要在模型下进行模拟，用户可以在模型对象中添加一个额外元素，指定一个基于观测值所依据的概率模型的模拟函数。这样一来，saemix 就能生成模型下的模拟数据集，并可通过 discreteVPC() 函数进行一些诊断。

最后，由于 saemix 中目前采用的近似是 FIM 的一阶近似，而众所周知这种近似在离散数据结局中表现不佳，因此无法再像目前实现的那样通过费 Fisher 信息矩阵获得估计标准误。取而代之的是，saemix 中已实现自助法，可应用于这类模型 [Comets 等人， 2021], 并指出，基于残差的自助法可能需要定义模拟函数，与诊断时相同。

使用示例：本文通过多个示例展示典型的建模工作流程，包括数据探索与格式整理、模型定义、参数估计、估计标准误、基于统计准则的模型比较以及模型诊断。现在可通过 npde 软件包 [Comets 和 Mentré, 2021], 使用 saemix 中拟合模型的模拟结果，对纵向结局进行模型诊断。本次扩展新增了离散结局的示例，并在软件包中提供了相应的数据集及其说明：二元数据（toenail.saemix）、分类数据（knee.saemix）、计数数据（rapi.saemix）和生存时间数据（lung.saemix）。

4 使用 SAEM 算法的非高斯数据模型

4.1 二元数据

4.1.1 案例研究：thetoenaildata

saemix 软件包中的 toenail.saemix 数据集包含一项随机临床试验的二元数据，该试验比较了两种真菌性趾甲感染治疗方法；在 R 中，该数据作为 prLogistic R 软件包中的 Toenail 数据集提供 [Amorim 和 Ospina, 2021]。数据来自 [De Backer 等人, 1998], 开展的一项多中心随机比较，比较了两种口服治疗（A 和 B）对趾甲感染的疗效。294 名患者在七次访视中接受测量（基线=第 0 周、第 4、8、12、24、36 和 48 周），共计 1908 次测量。主要终点是是否存在趾甲感染，关注的结局是二元变量甲剥离，该变量表示甲板与甲床分离的程度（无或轻度，与中度或重度）。

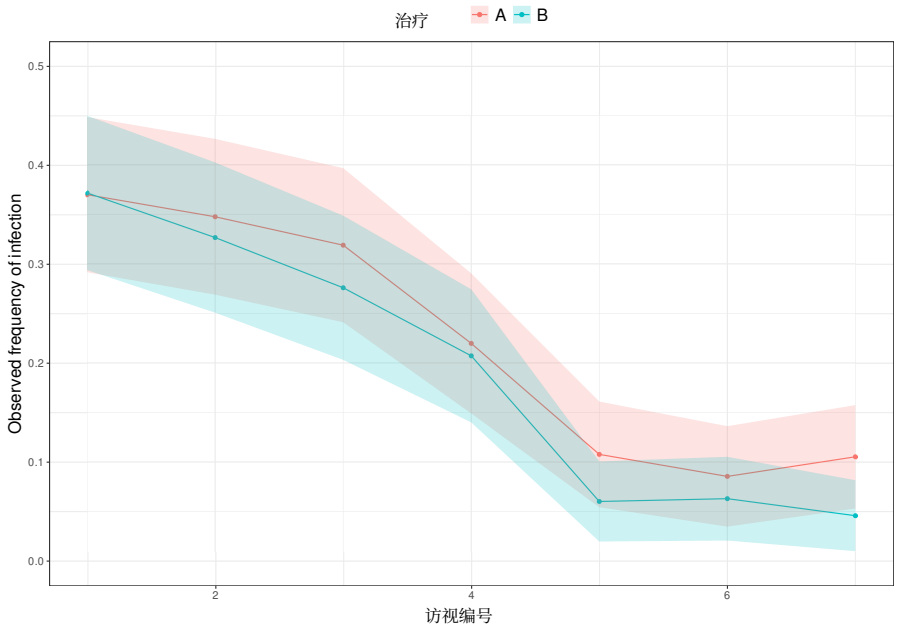

图 1：各次访视时趾甲感染受试者比例的时间进程，按治疗分层。

已有多项分析，例如 [Lesaffre 和 Spiessens, 2001, Lin 和 Chen, 2011]；本研究考虑了 [Lesaffre 和 Spiessens, 2001] 提出的逻辑斯蒂随机效应模型。该模型包含一个随机截距（$\theta_{1,\,i}$）、时间效应（θ_2）以及作为影响斜率的协变量的治疗（A 或 B）（β） θ_2。我们考虑了时间与治疗之间的交互项，但未考虑治疗对截距的影响，因为随机化过程意味着各组的截距不应有所不同。由于在检验不同模型时未发现显著的个体间变异，因此假设不同受试者的时间变化趋势相似。我们使用 saemix 在 logit 标度上拟合线性模型（ $\text{logit}(p) = \ln\left(\frac{p}{1-p}\right)$）：

$$
\begin{aligned}
y_{ij} \mid \theta_{1,i} &\sim \text{Bernoulli}(p_{ij}),\\
\text{logit}(p_{ij}) &= \theta_{1,i} + (\theta_2 + \beta \mathbb{1}(\text{trt}_i = B))\, t_{ij},\\
\theta_{1,i} &\sim \mathcal{N}(\theta_1, \omega_1^2).
\end{aligned}
\tag{7}
$$

表1展示了使用该模型估计的参数，以及采用两种自助法估计的估计标准误（SE）：经典的个案自助法（在个体层面进行重采样）和针对非线性混合效应模型提出的非参数条件自助法 [Comets 等人， 2021]。除 θ_1 外，使用条件自助法得到的所有参数的标准误均较低。因此，使用个案自助法估计的标准误时，治疗效应并不显著；而采用条件自助法时，该效应与 0 的差异变得（略微）显著。在本例中，我们无法确定哪种方法更准确，但下一节的模拟研究表明，在此情况下，条件自助法得到的标准误可能更接近经验标准误。

参数	估计值	个案自助法		条件自助法	
		RSE%	CI	RSE%	CI
θ_1 (-)	-1.71	21	[-2.44, -1.08]	18	[-2.20, -0.98]
θ_2 (d^{-1})	-0.39	18	[-0.56, -0.28]	10	[-0.49, -0.32]
β (-)	-0.15	87	[-0.44, 0.07]	47	[-0.29, -0.02]
ω_1 (-)	4.02	13	[3.22， 5.31]	9	[3.16， 4.64]

表1：使用 saemix 对趾甲数据进行拟合的逻辑回归模型参数估计值。标准误和置信区间通过个案自助法和条件自助法估计。

4.1.2 saemix的性能

本节评估该软件包在二元模型中估计群体参数的性能，并以趾甲数据集作为模拟研究的模板。

模拟设置：模拟时间设定为方案中包含的原始7个时间点，即由周换算为月（四舍五入）后的基线、1、2、3、5.5、8和11月。设置包括274名受试者，每个治疗组137名。

模拟了两种设计方案。第一种模拟趾甲数据集，使用表1中报告的参数（四舍五入至2位）。第二种假设两个参数均存在变异性，即$\theta_{1,i} \sim \mathcal{N}(\theta_1, \omega_1^2)$，且(7)中的 θ_2 替换为$\theta_{2,i} \sim \mathcal{N}(\theta_2, \omega_2^2)$，其中 $\omega_1 = 1, \omega_2 = 0.2$。通过使用 [Ueckert和Mentre 2017]提出的MC/AGQ 方法计算预测费 Fisher 信息矩阵（FIM），检验了设计方案的可识别性。在两种情景中，我们模拟了200个数据集，并使用saemix估计群体参数。为评估估计值的稳定性，我们采用了3种设置：从参数真值开始；从基于基线和末次访视时趾甲感染频率的粗略总体估计开始（无治疗效应，且变异性初值较小）（$\theta_{1,0} = -0.5, \theta_{2,0} = -0.19, \beta_0 = 0, \omega_{1,0} = 1$）；或从参数取0、变异性取4开始。所有情况下均使用10条链，并在前 $K_1/2$ 次迭代中启用模拟退火（算法细节见文档）。

评估：我们将相对偏倚计算为相对估计误差的均值，将相对均方根误差（RMSE）计算为其标准差，结果以百分比表示：

$$
\begin{aligned}
RB(\hat{\theta}) &= \frac{1}{S}\sum_{s=1}^{S} \frac{\hat{\theta}_s - \theta^0}{\theta^0} \times 100\\
RRMSE(\hat{\theta}) &= \sqrt{\frac{1}{S}\sum_{s=1}^{S} \left(\frac{\hat{\theta}_s - \theta^0}{\theta^0}\right)^2} \times 100
\end{aligned}
\tag{8}
$$

其中， θ^0 表示参数的真实值，$\hat{\theta}_s$ 表示第 s 次模拟中的估计值， S 表示模拟数据集的数量。

结果： 图 2 展示了两种设计方案中各参数相对估计误差（REE）的小提琴图。所有设置均得到了相同结果，表明使用 10 条链进行的运行结果稳定。因此，为了清晰起见，我们仅展示以真实参数作为初始值的设置下的图（完整结果见附录）。如附录中的表7所示，在原始设计中，所有参数的偏倚均小于 5%，且除 β 外，所有参数的 RRMSE 均在 10% 至 20% 之间； β 的估计值变异性更大（50%）。当我们为模型中的两个参数都设定变异性时，观察到 β 的偏倚约为 12%， θ_2 的偏倚为 8%；不同重复样本的估计值之间的变异性也有所增加，尤其是 β 和斜率 ω_2，因此这两个参数的 RRMSE 最高。估计得到的 RRMSE 与根据设计方案和参数通过 FIM 预测的值接近，但第二种情景中的 β 除外，其 RRMSE（56%）小于预测值（112%）。最后，我们可以将附录表7中给出的第一种情景的相对 RMSE，与真实数据示例中针对同一模型通过自助法得到的估计值（表1）进行比较，发现条件自助法估计值与 200 次模拟所得的经验估计值非常接近，而且在此情形下比个案自助法估计值更准确。

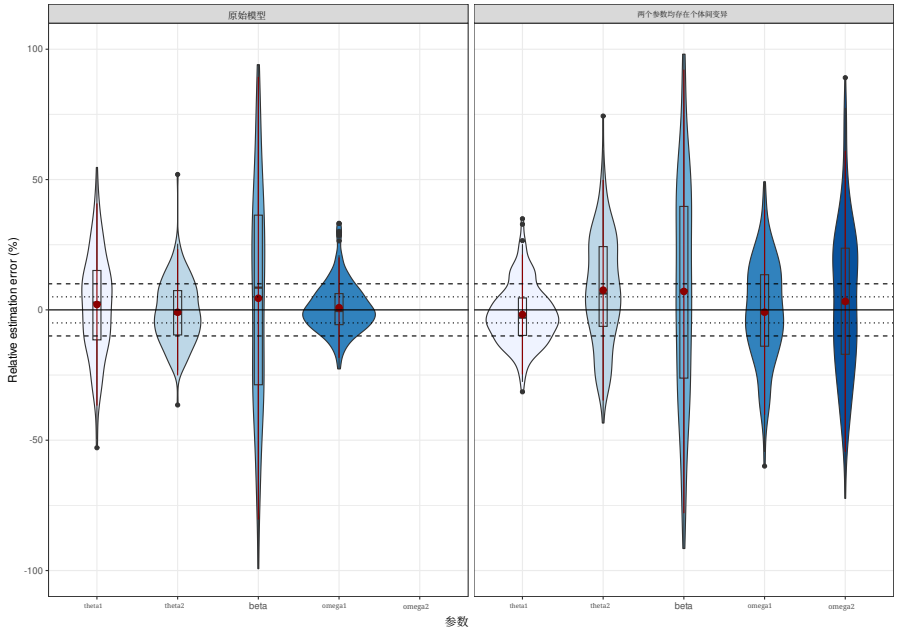

图 2：原始设计（左）以及两个参数均具有变异性的设计（右）中各参数的相对估计误差。虚线标示绝对相对偏差在 10% 以内的范围，点线表示 5% 以内的偏倚。红色插图突出显示对应于均值（红点）加减 2 个标准差的区间。该图被裁剪至 +/-100，省略了 β 的 26 个值。

4.2 分类数据

数据： 接下来，我们展示如何使用 saemix 对分类数据进行建模。knee.saemix 数据集记录了 127 名因运动相关损伤而接受两种不同治疗方法的患者在一项临床研究中的疼痛评分。患者在治疗 3、7 和 10 天后，膝关节运动时出现的疼痛均有记录。该数据集取自 R 的 catdata 软件包 [Schauberger 和 Tutz, 2020]（数据集 knee），并经过重新整理以用于 saemix，同时添加了表示测量日期的时间列（0 表示基线值），因此每次观测都对应数据集中的一行。治疗重新编码为 0/1（安慰剂/治疗），性别重新编码为 0/1（男性/女性），Age2 表示以年龄均值为中心化后的年龄平方。

我们可以使用 plotDiscreteData() 函数观察各评分所占比例随时间的变化，并将 outcome 设为分类变量，按治疗分层（见附录中的图8）。在两个治疗组中，较低评分（1 或 2）的概率均随时间增加，而较高疼痛评分（4 或 5）的概率则下降，这表明患者正在康复；治疗组本身似乎对中位数的影响大于对极端疼痛评分的影响。

模型： knee 数据集是 [Tutz 分析的数据集之一， 2011]采用了多种方法进行分析，这些方法在 catdata 软件包文档的插图中有所描述，但主要是对 10 天后的响应变量进行逻辑回归，或是在将响应变量二分化后进行混合二元回归。在此，我们对完整数据拟合比例优势模型。与时间 t_{ij} 时事件 Y_{ij} 相关的概率 $p_{ij} = P(Y_{ij} = 1|\psi_i)$ 的 logit 由以下方程给出：

$$\begin{aligned}
\text{logit}(P(Y_{ij} = 1|\psi_i)) &= \theta_{1,i} + \alpha_i t_{ij}, \\
\text{logit}(P(Y_{ij} = 2|\psi_i)) &= \theta_{1,i} + \theta_{2,i}, \\
\text{logit}(P(Y_{ij} = 3|\psi_i)) &= \theta_{1,i} + \theta_{2,i} + \theta_{3,i}, \\
\text{logit}(P(Y_{ij} = 4|\psi_i)) &= \theta_{1,i} + \theta_{2,i} + \theta_{3,i} + \theta_{4,i}, \\
P(Y_{ij} = 4|\psi_i) &= 1 - \sum_{k=1}^{4} P(Y_{ij} = k|\psi_i),
\end{aligned} \tag{9}$$

其中，$\theta_{1,\,i}$ 是时间 0 时疼痛评分为 1 的概率经过 logit 转换后的结果，其他 $\theta_{k,\,i}$ 参数表示进入更高疼痛类别的增量风险，而 α_i 表示个体 i的时间效应。

我们使用 SAEM 算法运行了 10 条链，在探索阶段和平滑阶段分别进行了 600 次和 100 次迭代。首先，我们拟合了一个不含协变量的模型，仅对 $\theta_{1,\,i}$ 和 α_i 设置变异性，以下称为基础模型。假设 $\theta_{1,\,i}$ 服从正态分布，$\theta_{1,\,i} \sim \mathcal{N}(\theta_1, \omega^2_{\theta_1})$，而 α_i 服从对数正态分布 $\alpha_i \sim \text{LN}(\alpha, \omega^2_\alpha)$，以反映随着时间推移疼痛评分较低的概率增加。其余参数被视为固定参数（无变异性），并被约束为正值，指定为 $\theta_{k,\,i} = \exp(\theta_k)$，适用于所有 $i = 1, \ldots, N$ 和 $k = 2, 3, 4$，以保持数据的有序结构。第二步中，我们使用了由 [Delattre 和 Poursat 实现的逐步算法， 2020]该算法基于非线性混合效应模型 [Delattre 等人的 BICc， 2014] 同时探索变异结构和协变量模型。在这一步中，假设参数 $\theta_{k,\,i}$，$k = 2, 3, 4$ 服从对数正态分布，$\theta_{k,\,i} \sim \text{LN}(\theta_k, \omega^2_{\theta_k})$。通过该逐步方法得到的最终模型称为协变量模型。我们对基础模型和协变量模型进行了诊断。

结果： 表 2 报告了基线模型和协变量模型的参数估计值。与之前一样，我们使用自助法评估参数的不确定性，并报告采用条件自助法得到的自助法置信区间，因为在二元响应变量模拟中，该方法的表现更好。

图3展示了按治疗分层的视觉预测检验（VPC）。图中显示模型存在一定的设定错误，尤其是在中间的疼痛评分上；此外，模型还倾向于高估改善程度，这主要是因为最低疼痛评分的增加。这表明，对于最低的三个疼痛类别以及整个时间段内的治疗效应，模型均未能妥善考虑。虽然性别未保留在协变量模型中，但按性别分层的 VPC 似乎显示，男性和女性的疼痛评分分布有所不同，尤其是在第一和第三类别（无痛或中度疼痛水平；附录 C 中的图9）。我们还可以查看各治疗组中位数评分的 VPC，发现模型倾向于低估疼痛评分，尤其是在接受治疗1的组中（附录 C 中的图10）。综合来看，这些结果表明，比例优势模型可能不足以描述这些数据，因此可以考虑研究其他模型，例如差异优势模型或有界整数模型。

参数	基础模型		协变量模型	
	估计值	自助法置信区间	估计值	自助法置信区间
θ_1	-15.2	[-19.0, -10.7]	-23.3	[-28.1, -13.6]
β_{Age^2,θ_1}	-	-	0.06	[0.03, 0.09]
θ_2	6.5	[4.6, 7.7]	3.5	[1.3, 4.8]
β_{Trt,θ_2}	-	-	1.0	[0.43, 1.6]
θ_3	8.5	[6.6, 10.6]	8.5	[5.8, 11.2]
θ_4	12.5	[9.7, 17.0]	16.4	[9.4, 20.7]
α	0.87	[0.6, 1.0]	0.88	[0.55, 1.0]
$\beta_{Trt,\alpha}$	-	-	0.56	[0.21, 0.83]
ω_{θ_1}	13.8	[10.5, 17.0]	15.6	[9.3, 17.8]
ω_{θ_2}	-	-	0.43	[0.01, 0.32]
ω_{θ_3}	-	-	0.74	[0.25, 0.89]
ω_{θ_4}	-	-	0.69	[0.18, 1.1]
ω_α	0.74	[0.55, 0.95]	-	-

表 2：无协变量和含协变量的比例优势模型的参数估计值。相对标准误和置信区间通过条件自助法获得。

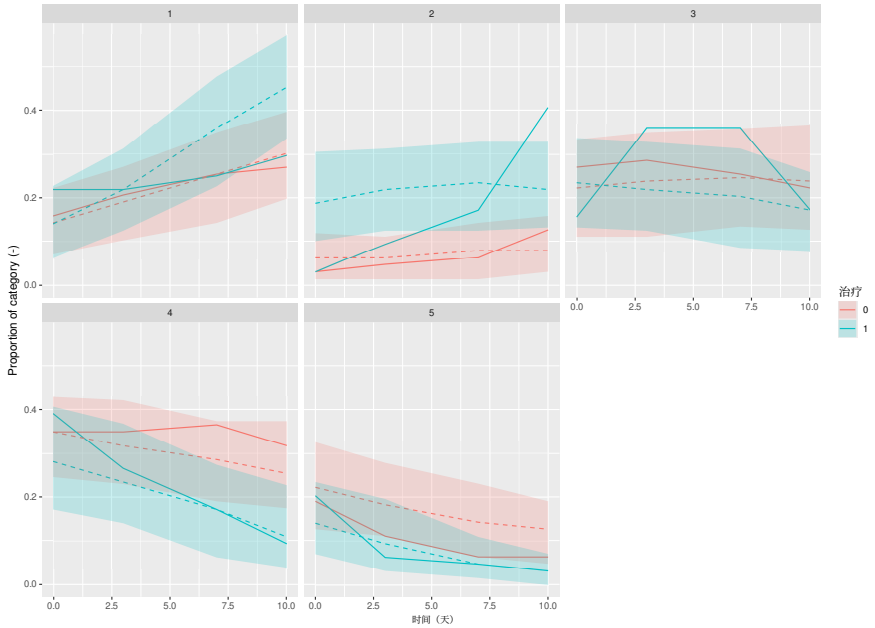

图3：含协变量的比例优势模型中各评分值的视觉预测检验，按治疗分层。视觉预测检验通过在原始数据集的模型和设计下模拟描述性统计量，并将其与观测值进行比较得出。此处，我们按照相同的设计模拟了 1000 个数据集的评分，然后计算了各治疗组中各评分随时间变化的比例的 95% 预测区间。随后叠加了观测比例。

4.3 计数数据

数据： saemix 软件包中的 rapi.saemix 数据集包含计数数据，该数据由 David Atkins（华盛顿大学）慷慨提供，出自其关于计数数据建模的教程 [Atkins 等人， 2013]。这些数据来自一项随机对照试验，该试验评估了基于网络的个性化规范反馈干预的有效性

关于酒精摄入 [Neighbors 等，2010b,a]。rapi.saemix 数据集记录了酒精相关问题，这些问题通过罗格斯酒精问题指数（RAPI）进行测量 [White 和 Labouvie， 1989], 研究对象为有过量饮酒风险的大一新生。学生每六个月报告一次酒精相关问题的数量，该数据集包含 818 名受试者的 3500 多条此类计数的重复测量。这个数据集有两个值得关注的特点：首先，其纵向特性使我们能够评估随时间的变化；其次，计数的分布形态。数据直方图显示，大量数据点为零，表明个体或某些时段没有出现问题。图 11（附录 D）展示了研究两年期间各计数所占比例的变化。Atkins 等人将该数据集用于教程，介绍如何使用 lme4 软件包中的 glmer() 函数进行混合效应计数回归[Atkins 等， 2013] [Bates 等， 2015]。这里，我们将展示如何使用 saemix 拟合计数数据分析中使用的主要模型。

模型：考虑计数数据时，首先想到的模型是**泊松模型**，它将事件数的均值的对数（λ_i）视为时间的线性函数：

$$P(Y_{ij}=n|\alpha_{0,i},\alpha_{1,i})=\frac{\lambda_i(t_{ij})^n\ e^{-\lambda_i(t_{ij})}}{n!}$$

其中

$$\log(\lambda_i(t))=\alpha_{0,i}+\alpha_{1,i}t \tag{10}$$

按照 [Atkins 等人 2013], 的做法，我们在两个参数中都纳入性别效应：

$$\alpha_{k,i}=\alpha_k+\beta_{male,\alpha_k}\mathbb{1}_{(\text{gender}_i=\text{male})}+\eta_{k,i}\ ,\ \eta_{k,i}\sim\mathcal{N}(0,\omega^2_{\alpha_k})\ ,\ k=0,1 \tag{11}$$

该模型的视觉预测检验见附录 D（图 13），结果表明，泊松模型的灵活性不足以应对数据中大量的零计数。

另外两种可用模型是零膨胀泊松（ZIP）模型和障碍模型。**ZIP 模型**使用混合分布来处理零计数的比例：

$$P(Y_{ij}=0|\alpha_{0,i},\alpha_{1,i})=p_0+(1-p_0)e^{-\lambda_i(t_{ij})}$$
$$P(Y_{ij}=n|Y_{ij}>0,\alpha_{0,i},\alpha_{1,i})=\frac{\lambda_i(t_{ij})^n\ e^{-\lambda_i(t_{ij})}}{n!}$$

其中，$\lambda_i(t)$ 的定义与式（10）相同。p_0 不设置变异性，它表示总体层面的比例，其取值被限制在 $[0,1]$内。个体参数的定义与式（11）相同。另一方面，**障碍模型** 是由两个模型组合而成，分别拟合于不同的数据集：二元逻辑回归用于确定零计数的概率（在整个数据集中拟合，将数据建模为发生一个或多个事件与未发生事件两类），截断泊松模型则用于描述非零计数（仅在严格为正的计数上拟合）。在此模型中，我们拟合：

Hurdle 0: binary logistic regression

$$P(Y_{ij}>0|\alpha_{0,i},\alpha_{1,i})=e^{\alpha_{0,i}+\alpha_{1,i}\ t_{ij}}/(1+e^{\alpha_{0,i}+\alpha_{1,i}\ t_{ij}})$$

Hurdle 1: truncated Poisson model

$$P(Y_{ij}=n|Y_{ij}>0,\lambda_i)=\frac{\lambda_i(t_{ij})^n\ e^{-\lambda_i(t_{ij})}}{(1-e^{-\lambda_i(t_{ij})})\ n!},$$

其中，$\lambda_i(t)$ 的定义见式 (10)–(11)。这两个模型的主要区别在于，障碍模型将零计数与非零计数分开处理，而在零膨胀泊松模型中，非零计数的数量由两个子模型共同决定。

我们运行 SAEM 算法，在 S 步骤中使用 10 条链，并进行 600 次预热迭代，随后进行 100 次细化迭代。对于泊松模型中的模拟函数，我们将计数截断至观测计数的最大值，以避免模拟出异常计数。

结果：表3比较了 saemix 中三种模型得到的估计值。我们可以比较在同一数据集上拟合的泊松模型（ 21508.6）和零膨胀泊松模型（20512.9）的贝叶斯信息准则，两者相差约 100

个点。这一改进也体现在观测零计数比例与模型预测值的一致性更好：模型预测值为 p_0（泊松模型的预测值为 0.15），而观测值为 0.21。不过，这两个模型估计的参数非常接近。最后，障碍模型对零计数比例的预测也非常准确（0.21）；同样，描述非零计数的截断泊松模型部分的所有参数估计值也十分相近。此外，我们的参数估计值与 [Atkins 等人此前在这两个模型的表1和表2中报告的结果非常接近。 2013]

	泊松	ZIP (截断泊松分布)	栏模型
α_0 (-)	1.49	1.57	1.61
β_{male,α_0} (-)	0.20	0.19	0.21
α_1 (mo$^{(-1)}$)	-0.04	-0.04	-0.02
β_{male,α_1} (-)	0.02	0.02	0.01
p_0 (-)*	-	0.08	-
ω_{α_0} (-)	0.96	0.89	0.85
ω_{α_1} (-)	0.06	0.06	0.05
$\rho(\alpha_0, \alpha_1)$ (-)	-0.14	-0.09	-0.29

* 仅出现在零膨胀泊松模型中表3：泊松、零膨胀泊松以及障碍模型中截断泊松分布部分的参数估计值。包括不确定性估计值在内的完整结果，以及针对零计数的逻辑回归模型，见附录（8）。

根据这些估计值，我们可以推导性别效应的相对风险。例如，与女性相比，男性的基线RAPI评分高出22%（$1.22 = \exp(0.20)$）。同样，在研究期间每过一个月，前两个模型的RAPI计数预计下降4%（$0.96 = \exp(-0.04)$），栏模型则下降2%。需要注意的是，栏模型在分别建模零计数与非零计数的演变方面具有更大的灵活性，这可以从两个子成分上性别效应的不同估计值看出（见附录中的表 8）。图4展示了零膨胀泊松模型的诊断图；视觉预测检验通过对拟合模型进行模拟获得，并将每个计数（或高计数的计数类别）的经验比例叠加到预期区间上，并按性别分层。这些图显示所有计数的拟合良好，并表明该模型能够很好地解释总体趋势以及男性与女性之间的差异。附录中给出了栏模型的类似诊断图。

与计数数据相关的笔记本（参见 GitHub）展示了如何在 saemix 中执行这些分析。其中还包括其他可用于处理过度离散的模型，例如广义泊松模型（该模型引入了过度离散参数），以及负二项模型（使用伽马分布对该参数进行建模） λ。

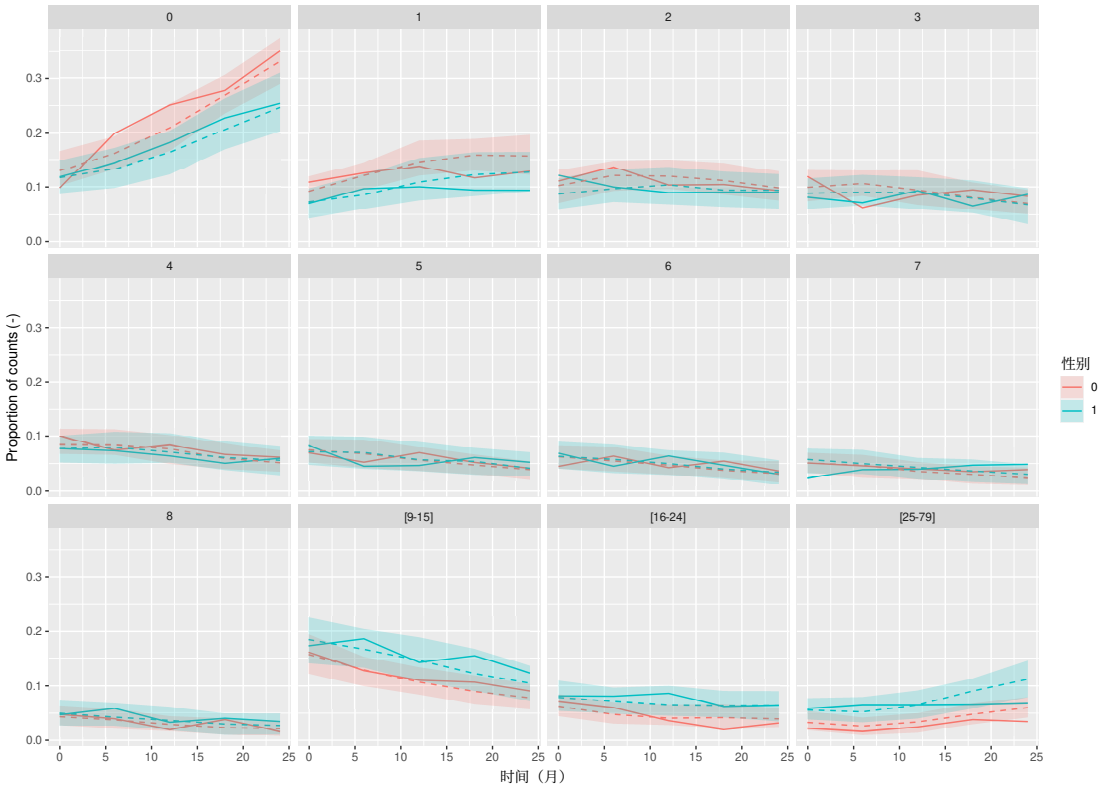

图4：零膨胀泊松模型的视觉预测检验图，将高计数类别合并，并按性别分层。

4.4 生存时间数据

数据： 用于说明如何在 saemix 中分析事件发生时间（TTE）数据的示例，是 NCCTG 肺癌数据集。该数据集描述了北方中部癌症治疗组中晚期肺癌患者的生存情况 [Loprinzi 等人， 1994]。研究中测量的协变量包括用于评定患者执行日常活动能力的体能评分。我们将 R 的 survival 软件包中提供的癌症数据集重新整理为 saemix 格式：从数据集中删除了年龄、性别、机构或医生评估信息缺失的患者，最终数据集包含 225 名患者，原始数据则包含 228 名患者。将状态重新编码为：死亡为 1，删失事件为 0；同时添加了删失列，用于表示患者在最后一次观察时是死亡还是存活。为所有受试者添加了时间=0 处的一行，因为观察期的开始是该算法的必要条件。每位患者记录的协变量包括性别（0：男性，1：女性）、年龄（岁）、由医生评估的 ECOG（东部合作肿瘤学组）体能状态（0–5），以及由医生和患者共同评估的 Karnofsky 体能状态（0–100）。我们可以通过绘制生存曲线来探索数据，如图 5 所示，其中我们绘制了按性别或基线 ECOG 评估分层的生存曲线。

模型： 使用 saemix，我们可以拟合似然可写出的任意参数模型。这里，我们将拟合几种经典参数生存模型：指数分布、对数逻辑斯蒂分布、伽马分布、威布尔分布和戈珀茨模型，其方程见表4。所有参数均假定服从对数正态分布，因为它们应保持为正值。对于单个事件，个体间变异难以识别，因此我们任意将其设定在参数 T_e 上。结构模型的模型选择基于贝叶斯信息准则。

若要使用自动诊断，必须在模型对象中添加模拟函数。与其他非高斯模型相比，从 TTE 模型进行模拟稍微复杂一些。当风险函数存在反函数时，我们可以使用逆累积分布函数法（或逆变换算法）从 TTE

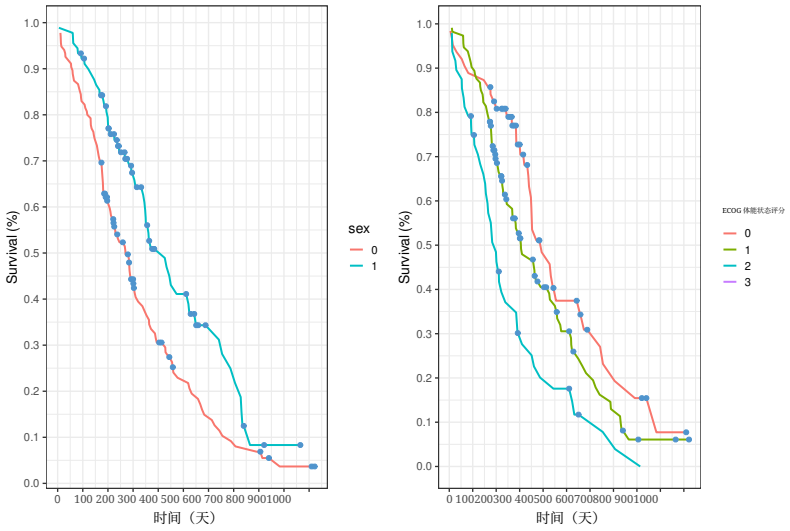

图 5：肺癌数据中的生存情况，按性别（左）或 ECOG 评估（右）分层。事先转换数据后，也可以绘制以月而非天为单位的相同图表；但由于数据集以天为单位，因此整个分析中我们都使用这一单位。

模型	Hazard 函数 $h(t)$	累积风险 $H(t)$	逆生存函数 $S^{-1}(p)$	参数
指数分布	$\frac{1}{T_e}$	$\frac{t}{T_e}$	$-T_e \ln(p)$	T_e
威布尔分布	$\frac{\gamma}{T_e}\left(\frac{t}{T_e}\right)^{\gamma-1}$	$\left(\frac{t}{T_e}\right)^{\gamma}$	$T_e(-\ln(p))^{1/\gamma}$	T_e, γ
冈珀茨分布	$\frac{\gamma}{T'_e}\, e^{\frac{t}{T'_e}}$	$\gamma(e^{\frac{t}{T'_e}}-1)$	$T'_e \ln\left(1-\frac{\ln(p)}{\gamma}\right)$	T_e, γ
伽马分布	$\frac{t^{\gamma-1}e^{-\frac{t}{T_e}}}{T_e^{\gamma}\Gamma(\gamma)}$	$\frac{\Gamma(\gamma,\, t/T_e)}{\Gamma(\gamma)}$	无闭式解	T_e, γ
对数逻辑斯蒂分布	$\frac{\frac{\gamma}{T_e}\left(\frac{T}{T_e}\right)^{\gamma-1}}{\left(\frac{T}{T_e}\right)^{\gamma}}$	$\ln\left(1+\left(\frac{T}{T_e}\right)^{\gamma}\right)$	$T_e\left(e^{-\ln(p)}-1\right)^{1/\gamma}$	T_e, γ

表4：在肺癌数据上测试的生存模型。 T_e 为尺度参数， γ 控制模型的形状。在戈珀茨模型中，我们将参数 T'_e 定义为 $T'_e = \frac{T_e}{ln\ (1+\frac{ln(2)}{\gamma})}$。在伽马模型中， $\Gamma(x)$表示伽马分布，$\Gamma(x,t)$表示不完全伽马分布。

模型生成随机样本。该方法利用了这样一个事实：连续累积密度函数 F 是从累积分布函数的定义域到区间 (0,1) 的一一映射。因此，如果 U是区间 (0,1) 上的均匀随机变量，那么 $X = F^{-1}(U)$ 服从分布 F。

对于单事件威布尔模型：

$$F(T) = 1 - e^{-\int_0^T h(u)du} = 1 - e^{-\left(\frac{T}{\gamma}\right)^{\beta}} \sim \mathcal{U}(0,1) \tag{12}$$

而反函数为：

$$F^{-1}(U) = \gamma\left(-\ln(1-U)\right)^{1/\beta} \tag{13}$$

假设我们从 $\mathcal{U}(0,1)$ 中模拟 $V = 1 - U$，即可从威布尔参数模型中获得样本，如 $F^{-1}(V)$ 所示。除伽马分布外，所有模型的逆生存函数均列于表4；伽马分布的反函数没有闭式表达式。随后将使用这些模拟结果生成 Kaplan-Meier 视觉预测检验 [Björnsson 和 Simonsson， 2011]。

结果：五种模型得出的贝叶斯信息准则见表 5。

模型	BIC
指数分布	2303.05
威布尔分布	2291.02
冈珀茨分布	2291.78
伽马分布	2378.34
对数逻辑斯蒂分布	2306.24

表5：在肺癌数据上比较生存模型。

威布尔模型和戈珀茨模型的贝叶斯信息准则非常接近，其诊断图见图 6。两种模型在第二年都略微低估了生存率；威布尔模型似乎对曲线末端的拟合稍好一些，但此时研究中剩余的受试者已经很少。

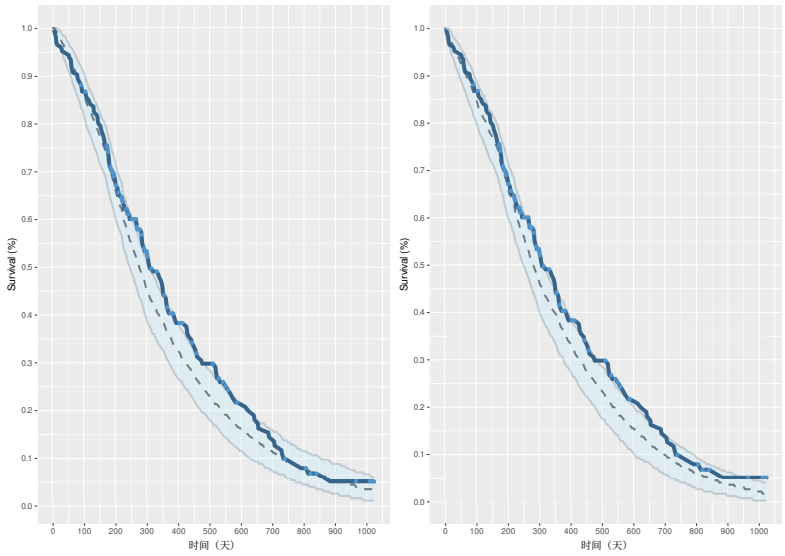

图6：比较威布尔分布（左）和冈珀茨分布（右）模型的期望生存。浅蓝色阴影区域表示Kaplan-Meier VPC，预测生存率以虚线显示，观察到的生存率以深蓝色线显示。

随后，使用威布尔模型并应用自动化算法 [Delattre 和 Poursat, 2020], 将ECOG评分合并为3个类别（0=参照组、1和2-3），再合并为两类（ECOG评分 ≥ 2 与较低评分）。最终模型纳入了性别和 T_e上的ECOG状态。最终模型的参数估计值见表6。若要将该模型拟合为混合效应模型，就需要在一个参数上纳入个体间变异性，但由于每位受试者仅提供一次观测，因此无法识别该变异性 [Kuhn 和 Lavielle,2005]。出于同样的原因，我们采用个案自助法计算置信区间，因为在条件自助法中抽样随机效应没有意义。模型预测女性的进展速度较慢， T_e增加至619天，而男性为432天（中位生存时间分别为470天（16个月）和328天（11个月））；男性的ECOG评分达到2分或以上时， T_e 则进一步降至264天。

参数	估计值	自助法置信区间
T_e (天)	405.4	[335.4-464.3]
$\beta_{gender,Te}$ (-)	0.36	[0.17-0.64]
β_{ECOG23,T_e} (-)	-0.49	[-0.89–0.24]
γ (-)	1.47	[1.21-1.73]
ω_γ (-)	0.085	[0.08-0.56]

表6：使用威布尔模型并纳入协变量对肺癌数据进行分析所得的参数估计值。采用个案自助法计算95% 置信区间。

5 讨论

本文介绍了如何扩展用于 R 的 saemix 软件包，以便使用 SAEM 算法拟合非高斯数据的非线性混合效应模型 [Kuhn 和 Lavielle， 2005]。saemix 于 2011年 首次在 CRAN 上发布。与同样用于连续结局建模的 nlme 和 lme4 软件包相比 [Comets 等人， 2017] 发现 saemix 的表现优于这些 FOCE 算法，估计失败较少，统计性能也更好。用于非高斯数据的模型已整合到 2022年2月 发布的第3版中。本文将介绍若干案例研究，这些研究使用了软件包在线帮助中提供的数据集。

第一类非高斯结局是分类数据，其中有限数量的类别对应于离散概率分布；二元数据是其特例，仅包含两个类别，例如“是/否”。本文展示了两项应用：一项比较两种治疗真菌性趾甲感染的随机试验，以及一项临床研究，其中膝痛评分以 5 种形式记录。在二元回归示例中得到的参数估计值与使用 Monolix 得到的估计值几乎完全相同，也与 [Molenberghs 和 Verbeke， 2009], 报告的结果相似，不过在后者的分析中，截距中加入了一个（不显著的）治疗效应。Monolix 使用随机方法给出的标准误估计值与条件自助法得到的估计值非常接近；后者也得到了模拟研究所得经验标准误的验证，这表明在此情形下，个案自助法高估了标准误。另一方面，在 [Molenberghs 和 Verbeke， 2009] 中，采用数值求积法拟合广义混合效应模型时，治疗效应未达到显著，但在截距中加入的（不显著）治疗效应存在差异。随后，将趾甲数据作为模板开展模拟研究，以检验算法的性能。我们发现，在原始情景中估计值没有偏倚；而在具有个体间变异的第二种情景中，最难识别的参数出现了偏倚。对于非高斯数据，我们使用多条链来稳定算法，并检查了采用不同参数初始值时结果是否相近。随后，在采用比例优势模型对膝部数据集建模时也使用了类似设置。诊断图显示，模型对两个较高的疼痛评分拟合良好；第一治疗组的疼痛评分下降幅度较大，且膝痛随治疗过程有所改善。不过，诊断图也突出显示，模型对较低疼痛评分的拟合明显不足。有序数据模型还面临一项挑战：模型对参数初始估计值更加敏感。我们可以研究多项模型，以允许每个评分值具有不同的时间趋势；也可以考虑马尔可夫模型等其他类型的模型，但这些目前尚不属于 saemix 的范围。

第二类非高斯结局是计数数据，其可能取值的数量可以无限增加。在 saemix 中，定义计数所服从的概率分布非常简单，我们可以拟合 [Atkins 等人 2013] 提出的不同模型，且结果非常相似。为保持一致，所有模型均使用相同的个体间变异，并在两个参数上都设置方差；而 [Atkins 等人 2013]提出的障碍模型中仅使用了一个方差项。诊断图显示，使用泊松模型时出现的过度离散，可以通过零膨胀泊松模型或障碍模型加以纠正。与分类数据示例一样，诊断图可用于考察模型在不同评分和协变量下的拟合情况。我们还利用此示例探究了快速 SAEM 算法的性能。该算法此前已在连续数据和生存时间数据中进行评估 [Karimi 等人 2020], , 结果发现它在此处的表现也相当不错（数据未展示）；不过，对于其底层条件分布拉普拉斯近似的性能，以及该近似对错误初始参数的敏感性，还需要进行更深入的评估。

最后，我们可以使用 saemix 处理单次和重复事件发生时间模型，并使用参数模型描述风险函数。在这些模型中，计算概率需要求出事件发生前的生存函数，这可能需要对风险函数进行积分，从而耗费大量计算时间。saemix 在事件发生时间分析中的性能已在另一项研究中进行了考察 [Lavalley-Morelle 等人， 2024]。

saemix 是一个协作项目，拥有专门的开发 GitHub (`https://github.com/saemixdevelopment/saemixextension`)。我们欢迎新的贡献者，尤其欢迎针对软件包当前局限性的改进。代码结构足够简单，便于实现新功能并

与软件包对接。具体而言，该算法已纳入 mkin 软件包，用于拟合毒代动力学数据模型 [Ranke, 2020]。

由于非线性混合效应模型的各种特性可能变得相当复杂，saemix 在指定模型时不采用传统的 R 公式。这样可以提高灵活性：对于连续结局，观测模型可以通过结构模型定义；对于非高斯结局，则可以通过概率分布定义。数据也需要通过特定函数进行结构化，以定义其层级特性，这与 nlme 软件包中的 groupedData 结构类似 [Pinheiro 等， 2023]。saemix 为软件包中定义的每类对象提供了许多便于用户使用的功能，帮助用户探索数据集、检查模型、汇总结果并诊断拟合情况。与连续结局的选项相比，非高斯结局目前尚不支持某些功能。可以生成视觉预测检验诊断图，也可以按协变量分层；但更复杂的诊断图需要用户自行编写代码。基于模拟的诊断要求在模型对象中添加模拟函数，因此需要由用户负责定义。调用 saemix 及软件包中的其他函数所生成的对象，除参数估计外，还包含个体估计值和预测值，R 用户可据此绘制各种图表并计算统计量。此外，软件包还提供了补充在线帮助的笔记本，其中包含脚本和使用示例。这些内容展示了如何获取统计准则、诊断图和评估工具，可用于比较模型并评估拟合优度。

saemix 应用于非高斯结局时的一个当前问题是如何估计不确定性。该软件包中的估计标准误，是通过 FO 近似计算 Fisher 信息矩阵获得的；该近似使用个体参数的条件估计值，对于非高斯响应变量的表现很差 [Ueckert 和 Mentre，2017，Riviere 等人， 2016]。目前已有多种方法用于估计精确的 FIM，通常通过 Louis 方法分解海森矩阵期望 [Louis， 1982]。如 Monolix 中所采用的 FIM 随机计算方法 [Lavielle，2014，Savic 和 Lavielle， 2009]目前仍在开发中，SAEM 算法运行期间对 FIM 进行随机逼近的方法也是如此[Kuhn 和 Lavielle， 2005]。近期有人提出另一种计算方法，利用对数似然海森矩阵的期望与其梯度乘积的负期望等价这一性质 [Delattre 和 Kuhn， 2023], ；我们已在一项关于联合模型的研究中实现了该方法 [Lavalley-Morelle 等人， 2023], ，相关代码可在 GitHub 上获取。计算 FIM 是最优设计方法的重要组成部分。Rivière 等人提出了非高斯结局的精确 FIM [Riviere 等人， 2016], ，他们采用哈密顿蒙特卡罗方法进行随机计算；Ueckert 等人 [Ueckert 和 Mentre， 2017]则使用自适应高斯求积进行数值计算。他们用于计算期望 FIM 的代码也可用于计算观测 FIM，不过计算并求和各个个体的 FIM 可能会耗费大量计算资源。或者，也可以通过自助法或抽样重要性重抽样（SIR）等非参数方法来计算不确定性 [Dosne 等人， 2016]。saemix 中实现了多个适用于非线性混合效应模型的自助法版本 [Comets 等人， 2021] ，并已用于本文的示例；SIR 代码则已上传至 GitHub，等待在未来版本中纳入该软件包。

缺少微分方程是 saemix 的主要局限之一。虽然在该软件包的模型函数中加入微分方程相对容易，但这样做会损失 R 中向量化的效率，使估计过程变得极其缓慢。为此，需要采用更好的解决方案，例如使用 R 中高效的求解器，或将计算函数外包给 C。mkin 软件包便是一个此类方案的实例，它使用 saemix 拟合毒代动力学数据，并在 C 中对微分方程进行积分 [Ranke，2020]。目前正在开发一种基于网格近似的扩展，以便与复杂的 ODE 系统对接；这些系统用于定义由 PKSim [Willmann et al.，2003] 软件求解的基于药理学的药代动力学模型，并已在伊曲康唑研究中进行了概念验证应用 [Teutonico et al.，2026]。

目前，saemix 仅支持单响应模型。不过，我们已开发出联合模型扩展，并通过多个示例进行了评估。这些示例涉及线性和非线性纵向标志物与生存时间数据的组合型模型；生存时间数据采用标准参数风险函数以及竞争风险模型进行建模 [Lavalley-

Morelle 等人， 2024]。目前正在开发面向多个响应变量的通用扩展，并将在该软件包的未来版本中实现。

总之，saemix 的这一扩展提供了一个能够拟合描述高斯结局和非高斯结局的模型的软件包。该软件包易于使用，并在多种模型中表现良好；在线帮助及其 GitHub 上的附加笔记本中均提供了许多详细的使用示例。它会定期更新，未来展望包括增强建模能力，例如多响应模型、处理删失数据或变异的层级的算法，以及实现更便于用户使用的功能，例如分类协变量的自动变换、改进的诊断，以及更便于检查初始固定效应的函数。

致谢

作者衷心感谢 David Atkins 教授，他慷慨允许将 RAPI 数据纳入 saemix。作者还要感谢 Marc Lavielle 教授创建 Monolix，并为 saemix 的首个版本做出贡献。saemix 是一项协作成果，我们衷心感谢各位贡献者，尤其是 Johannes Ranke 在其用于毒代动力学的 mkin 软件包中使用 saemix（`https://www.jrwb.de/`），Alexandra Lavalley-Morelle 实现了可用于拟合联合模型的代码扩展（可在 GitHub 上获取（`https://github.com/saemixdevelopment/saemixextension/tree/master/joint`），Lucie Fayette 将 saemix 与 Monolix 及较早版本的 Pumas 进行了比较（`https://pumas.ai/`），以及 Lucas Bodelle 目前正在开展将 saemix 扩展至多个响应变量的工作。

伦理审批声明：不适用

资金：无

利益冲突：无

Code

saemix 软件包可在 CRAN 上获取（当前版本：3.5），后续版本存档于 IAME researchcenter 的 GitHub 仓库（`https://github.com/iame-researchCenter/saemix`），并可在此下载 saemix 的用户指南（`https://github.com/iame-researchCenter/saemix/blob/main/docsaem.pdf`）。
saemix 的开发版本可在 GitHub 仓库中找到（`https://github.com/saemixdevelopment/saemixextension`）。本文各部分对应的笔记本和 R 脚本已分别放置在各自的文件夹中（`https://github.com/saemixdevelopment/saemixextension/tree/master/paperSaemix3`）。

运行附录 B 至 E 中给出的代码之前，需要执行以下命令加载 saemix 软件包：

```
1 程序包载入( saemix )
```

这也应该会加载所需的软件包（npde、ggplot2）。如果没有，请同时加载这些库。

参考文献

Leila Amorim 和 Raydonal Ospina。通过逻辑回归估计患病率比：R 中的一种工具。巴西科学院院刊，93(4):e20190316，2021。doi:10.1590/0001-3765202120190316。David Atkins、S Baldwin、C Zheng、R Gallop 和 C Neighbors。纵向物质使用数据的计数回归和零值改变计数模型教程。成瘾行为心理学，27:166–77，2013。Douglas Bates、Martin Mächler、Ben Bolker 和 Steve Walker。使用 lme4 拟合线性混合效应模型。《统计软件杂志》，67:1–48，2015。Stuart Beal、Lewis Sheiner、Alison Boeckmann 和 Robert Bauer。Nonmem 7.5.1 用户指南。1989—2022年。ICON plc，盖瑟斯堡，https://nonmem.iconplc.com/nonmem751。Marcus A Björnsson 和 Ulrika SH Simonsson。萘普西诺、萘普生和安慰剂给药后牙痛模型中疼痛强度及信息性脱落的建模。英国临床药理学杂志，71(6): 899–906，2011。Bob Carpenter、Andrew Gelman、Matthew D. Hoffman、Daniel Lee、Ben Goodrich、Michael Betancourt、Marcus Brubaker、Jiqiang Guo、Peter Li 和 Allen Riddell。Stan：一种概率编程语言。《统计软件杂志》，76(1):132，2017。doi:10.18637/jss.v076.i01。URL https://www.jstatsoft.org/index.php/jss/article/view/v076i01。Emmanuelle Comets 和 France Mentré。开发用于评估非线性混合效应模型的工具：npde 探索之旅 20 年。*AAPS*期刊，23(4):75，2021年5月。Emmanuelle Comets、Karl Brendel 和 France Mentré。计算标准化预测分布误差以评估非线性混合效应模型：R 的 npde附加软件包。《生物医学计算机方法与程序》，90(2):154–166，2008。Emmanuelle Comets、Audrey Lavenu 和 Marc Lavielle。SAEMIX：SAEM 算法的 R 语言版本。*PAGE20*，第 Abstr 2173 页，2011。Emmanuelle Comets、Audrey Lavenu 和 Marc Lavielle。使用 saemix（SAEM 算法的 R 语言实现）进行非线性混合效应模型的参数估计。《统计软件杂志》，80:1–41，2017。Emmanuelle Comets、Christelle Rodrigues、Vincent Jullien 和 Moreno Ursino。非线性混合效应模型的条件非参数自助法。药学研究，38:1057–66，2021。Marie Davidian 和 David M. Giltinan。重复测量数据的非线性模型。CRC，伦敦，1995。M De Backer、C De Vroey、Emmanuel Lesaffre、I Scheys 和 P De Keyser。针对皮肤癣菌引起的趾甲甲真菌病，连续 12 周口服疗法：特比萘芬 250 mg/天与伊曲康唑 200 mg/天的双盲对照试验。美国皮肤病学会杂志，38(5):S57–S63，1998。Maud Delattre 和 Estelle Kuhn。通过随机逼近，在潜变量模型中计算经验费希尔信息矩阵估计值。*Computo*，2023。Maud Delattre 和 Marie-Anne Poursat。混合效应模型中协变量与随机效应联合选择的迭代算法。国际生物统计学杂志，16(2):20190082，2020。doi:doi:10.1515/ijb-2019-0082。URL https://doi.org/10.1515/ijb-2019-0082。Maud Delattre、Marc Lavielle 和 Marie-Anne Poursat。关于混合效应模型中贝叶斯信息准则的说明。电子统计学杂志，8:456–475，2014。Bernard Delyon、Marc Lavielle 和 Eric Moulines。EM算法随机逼近版本的收敛性。统计年鉴，27:94–128，1999。Arthur P Dempster、Nan M Laird 和 Donald B Rubin。通过 EM算法从不完整数据中求得最大似然。英国皇家统计学会杂志*B*辑，39(1):1–38，1977。附讨论。

Solène Desmée、F Mentré'、Christine Veyrat-Follet、Bernard Sébastien 和 Jérémie Guedj。使用 SAEM 算法构建机制联合模型，刻画前列腺癌患者非线性 PSA 动力学与生存之间的关系。《生物统计学》，73:305–12，2017。Anne-Gaelle Dosne、Martin Bergstrand、Kajsa Harling 和 Mats O Karlsson。使用抽样重要性重抽样改进非线性混合效应模型中参数不确定性分布的估计。《药代动力学与药效动力学杂志》，43(6):583–596，2016。doi:10.1007/s10928-016-9487-8。Mélanie Guhl、Julie Bertrand 和 Emmanuelle Comets。非线性混合效应模型中有限距离不确定性的计算。*PAGE30*，2022年。 www.page-meeting.org/?abstract=10152。Jarrod D. Hadfield。多响应广义线性混合模型的 MCMC 方法：MCMCglmm R 软件包。《统计软件杂志》，33(2):1–22，2010。URL https://www.jstatsoft.org/v33/i02/。Niels Hendrickx、France Mentré、Andreas Traschütz、Cynthia Gagnon、Rebecca Schüle、ARCA Study group、EVIDENCE-RND consortium、Matthis Synofzik 和 Emmanuelle Comets。罕见神经退行性疾病中纳入参数不确定性的个体疾病进展预测：以常染色体隐性痉挛性共济失调 Charlevoix-Saguenay 型（ARSACS）为例。*AAPS*期刊，26:57，2024。Christopher Jackson。flexsurv：用于 R 中参数生存建模的平台。《统计软件杂志》，70 (8):1–33，2016。doi:10.18637/jss.v070.i08。Belhal Karimi、Marc Lavielle 和 Eric Moulines。f-SAEM：用于非线性混合效应模型的快速 EM 算法随机逼近。《计算统计与数据分析》，141:123–38，2020年。Estelle Kuhn 和 Marc Lavielle。非线性混合效应模型中的最大似然估计。《计算统计与数据分析》，49:1020–1038，2005年。Alexandra Lavalley-Morelle、France Mentré、Emmanuelle Comets 和 Jimmy Mullaert。联合建模生物标志物动态与竞争风险下的生存，以预测因重症感染性疾病住院患者的预后。*PAGE31*，2023年。 www.page-meeting.org/?abstract=10394。Alexandra Lavalley-Morelle、France Mentré、Emmanuelle Comets 和 Jimmy Mullaert。扩展开源 saemix 软件包中的代码以拟合联合模型。《生物医学计算机方法与程序》，247:108095，2024年。Marc Lavielle。群体方法的混合效应模型：模型、任务、方法与工具。Chapman & Hall CRC 生物统计学丛书，佛罗里达州博卡拉顿，2014年。Emmanuel Lesaffre 和 Bart Spiessens。逻辑斯蒂随机效应模型中求积点数量的影响：一个示例。《英国皇家统计学会杂志：*C*辑（应用统计）》，50(3):325–335，2001年。Kuo-Chin Lin 和 Yi-Ju Chen。使用非参数平滑方法对逻辑斯蒂-正态模型进行拟合优度检验。《统计规划与推断杂志》，141(2):1069–1076，2011年。Mary J. Lindstrom 和 Douglas M. Bates。重复测量数据的非线性混合效应模型。《生物统计学》，46(3):673687，1990年。ISSN 0006341X，15410420。URL http://www.jstor.org/stable/2532087。Charles Lawrence Loprinzi、John A Laurie、H Sam Wieand、James E Krook、Paul J Novotny、John W Kugler、Joan Bartel、Marlys Law、Marilyn Bateman 和 Nancy E Klatt。基于患者填写的问卷对预后变量进行前瞻性评估。北方中部癌症治疗组。《临床肿瘤学杂志：美国临床肿瘤学会官方期刊》，12:601–7，1994年。Thomas A. Louis。使用 EM 算法时如何求得观测信息矩阵。《英国皇家统计学会杂志：*B*辑（方法论）》，44(2):226–233，1982年。ISSN 2517-6161。Geert Molenberghs 和 Geert Verbeke。纵向与不完整临床研究。《日本计算统计学会杂志》，22:1–32，2009年。Charles J Neighbors、Nancy P Barnett、Damaris J Rohsenow、Suzanne M Colby 和 Peter M Monti。医院急诊科涉酒青少年动机干预的成本效益。《酒精与药物研究杂志》，71(3):384–394，2010a。

Clayton Neighbors、Melissa A. Lewis、David C. Atkins、Megan M. Jensen、Theresa Walter、Nicole Fossos、Christine M. Lee 和 Mary E. Larimer。基于网络的个性化规范反馈的效果：一项为期两年的随机对照试验。《咨询与临床心理学杂志》，78:898–911，2010b。THT Nguyen、Samer Mouksassi、Nicholas Holford、N Al-Huniti、Imanuel Freedman、Andrew C Hooker、J John、Mats O Karlsson、Diane R Mould、JJ Pérez Ruixo 等。连续型数据药物计量学模型的模型评估：指标与图形。《*CPT*：药物计量学与系统药理学》，6(2):87–109，2017年。ISSN 2163- 8306，2163-8306。doi:10.1002/psp4.12161。José Pinheiro、Douglas Bates 和 R Core Team。*nlme*：线性与非线性混合效应模型，2023年。URLhttps://CRAN.R-project.org/package=nlme。R 软件包版本 3.1-162。José C. Pinheiro 和 Douglas M. Bates。*S* 和 *S-PLUS* 中的混合效应模型。施普林格，纽约，2000年。doi:10.1007/b98882。R 开发核心团队。*R*：一种用于统计计算的语言与环境。R统计计算基金会，维也纳，奥地利，2006年。URL http://www.R-project.org。ISBN 3-900051-07-0。Chris Rackauckas、Yingbo Ma、Andreas Noack、Vaibhav Dixit、Patrick Kofod Mogensen、Simon Byrne、Shubham Maddhashiya、José Bayoán Santiago Calderón、Joakim Nyberg、Jogarao VS Gobburu 和 Vijay Ivaturi。借助 Pumas 加速预测性医疗保健分析：一个高性能药物建模与模拟平台。*BioRxiv*，2020年。URL https://doi.org/10.1101/2020.11.28.402297。Johannes Ranke。*mkin*：化学降解数据的动力学评估，2020年。URL https://CRAN.R-project.org/package=mkin。R 软件包版本 0.9.50.3。Marie-Karelle Riviere、Sebastian Ueckert 和 France Mentré。用于评估非线性混合效应模型费 Fisher 信息矩阵的 MCMC 方法。《生物统计学》，17:kxw020，2016年5月。Radojka M Savic 和 Marc Lavielle。计数数据群体模型的性能，第二部分：一种新的 SAEM 算法。《药代动力学与药效动力学杂志》，36:367–79，2009年。Radojka M. Savic、France Mentré 和 Marc Lavielle。SAEM 算法在纵向有序分类数据中的实现与评估，并以药代动力学/药效学为例。*AAPS*期刊，13(1):44–53，2010年11月。ISSN 1550-7416。doi:10.1208/s12248-010-9238-5。Gunther Schauberger 和 Gerhard Tutz。*catdata*：分类数据，2020年。URL https://CRAN.R-project.org/package=catdata。R 软件包版本 1.2.2。Stan 开发团队。RStan：Stan 的 R 接口，2023年。URL https://mc-stan.org/。R 软件包版本 2.21.8。Donato Teutonico、David Marchionni、Marc Lavielle 和 Laurent Nguyen。将群体方法与生理药代动力学模型相结合：一种参数估计的新框架。《*CPT*：药物计量学与系统药理学》，15(1):e70186，2026年。doi:https://doi.org/10.1002/psp4.70186。Gerhard Tutz。分类数据回归。《剑桥统计与概率数学丛书》。剑桥大学出版社，2011年。ISBN 9781139499583。URL https://books.google.fr/books?id=hvxuqoxD00kC。Sebastian Ueckert 和 France Mentre。利用蒙特卡洛抽样和自适应高斯求积评估离散混合效应模型费 Fisher 信息矩阵的新方法。《计算统计与数据分析》，111:203–219，2017年。Helene R White 和 Erich W Labouvie。探讨青少年问题饮酒的评估。《酒精研究杂志》，50:30–7，1989年。Hadley Wickham。*ggplot2*：优雅的数据分析图形。施普林格出版社纽约，2016年。ISBN 978-3-319-24277-4。URL https://ggplot2.tidyverse.org。Stefan Willmann、Jorg Lippert、Michael Sevestre、Juri Solodenko、Franco Fois 和 Walter Schmitt。PK-Sim：生理药代动力学全身模型。*Biosilico*，1:121–4，2003年。

C.-F. Jeff Wu。关于EM算法的收敛性质。《统计年刊》，11(1):95–103，1983年。

附录A - SAEM 算法

我们的目标是最大化方程 4 中 θ 的对数似然。由于需要对未知随机效应进行积分，这个表达式在非线性混合效应模型中没有闭式形式。然而，与完整数据（y, ψ_i）相关的完全对数似然可以明确写出：

$$L(\theta|y,\psi) = \sum_{i=1}^{N} \log\left(p(y_i,\psi_i|\theta)\right) = \sum_{i=1}^{N} \log\left(p(\psi_i|y_i,\theta)\right) + \log\left(p(y_i|\theta)\right)$$

因此，我们有：

$$L(\theta|y) = L(\theta|y,\psi) - \sum_{i=1}^{N} \log\left(p(y_i|\theta)\right)$$

由于 $L(\theta|y)$ 不依赖于 ψ，在给定 $\theta^{(k)}$ 的条件下，对 ψ 取期望可将前面的等式写为：

$$L(\theta|y) = \mathbb{E}_{\psi}\left[L(\theta|y,\psi)|\theta^{(k)}\right] - \mathbb{E}_{\psi}\left[\sum_{i=1}^{N} \log\left(p(y_i|\theta)\right)|\theta^{(k)}\right]$$

接下来可以证明，取 $\theta^{(k+1)} = \arg\max_{\theta} \mathrm{E}_{\psi}\left[L(\theta|y,\ \psi)|\theta^{(k)}\right]$ 可使 $L(\theta^{(k)}|y)$ 收敛到一个局部最大值。

此外，

$$\begin{aligned}
\mathbb{E}_{\psi}\left[L(\theta|y,\psi)|\theta^{(k)}\right] &= \int L(\theta|y,\psi_i)p(\psi_i|\theta^{(k)})d\psi_i \\
&= \int \sum_{i=1}^{N} \log\left(p(y_i,\psi_i|\theta)\right) p(\psi_i|\theta^{(k)})d\psi_i \\
&= \mathbb{E}_{\psi}\left[\log\left(p(y,\psi|\theta)\right)|y,\theta^{(k)}\right]
\end{aligned}$$

因此，在EM算法的第 k 次迭代中，首先计算完全对数似然的条件期望

$$Q_k(\theta) = \mathbb{E}\left[\log\left(p(y,\psi;\theta)\right)|y,\theta^{(k-1)}\right]$$

，然后在M步中计算下一个 $\theta^{(k)}$，以最大化 $Q_k(\theta)$。

附录中的所有代码都需要加载 saemix 软件包。请在运行附录 B 至 E 中给出的代码之前执行以下命令。

```
程序包载入( saemix )
```

Thi还应加载所需的软件包（npde、ggplot2）。如果尚未加载，请也加载这些软件包 s.

附录 B - 二元数据分析

趾甲示例的 saemix 代码

```
数据(toenail.saemix)
saemix.数据<-saemixData(name.数据=toenail.saemix,name.group=c("id"),
    名称.预测变量=c("time","y"),          名称.响应="y",名称.协变量=c("治疗"),
    单位=列表(x="d",y="-"),       详尽=FALSE)
# 模型和        模拟          函数
二元。模型<-函数(psi,id,xidep){
    时间<-xidep[,1]
    y<-xidep[,2]
    inter<-psi[id,1]
    斜率<-心理参数[个体编号, 2]
    logit<-inter+slope*tim
    事件概率<-exp(logit)/(1+exp(logit))
    logpdf<-rep(0,length(tim))
    P.obs = (y==0)*(1-pevent)+(y==1)*pevent
    概率密度函数<-对数(P.obs)
    返回(对数概率密度函数)
}
simulBinary<-函数(psi,id,xidep){
    tim<-xidep[,1]
    y<-xidep[,2]
    截距<-psi[id, 1]
    斜率<-psi[id, 2]
    逻辑<-截距+斜率*时间
    事件<-1/(1+exp(-逻辑))
    模拟<-rbinom(长度(时间),大小=1,概率=事件)
    返回(ysim)
}
# saemix模型
saemix.模型<-saemixModel(模型=二元.模型,描述="二元                                    模型",
    模拟.函数=模拟二元,                    模型类型="似然",
    psi0=matrix(c(-0.5,-.15,0,0),ncol=2,byrow=TRUE,
    维度名称=列表(NULL,c("theta1","theta2"))),
    transform.par=c(0,0), covariate.模型=c(0,1),
    covariance.模型=matrix(c(1,0,0,0),ncol=2),
    omega.init=诊断(c(0.5,0.3)),       冗余信息=FALSE)

# saemix      拟合
saemix.选项<-列表(随机种子=1234567,保存=FALSE,保存.图形=FALSE,
    显示进度=FALSE,          链数=10,          fim=FALSE, 打印=FALSE)
binary.fit<-saemix(saemix.模型,saemix.data,saemix.options)
摘要(binary.fit)

# 诊断信息
nsim<-1000
binary.fit <- simulateDiscreteSaemix(binary.fit, nsim=nsim)
离散VPC(binary.fit,          结局="binary",        其中.cov="treatment")
```

二分类结局模拟的其他结果

偏倚和相对均方根误差

表7显示了两个情景中所有估计参数的相对偏倚（%）和相对均方根误差（%，括号内）。每个情景均测试了3种初始条件，但无论设置如何，结果都非常接近。

参数	原始模型			两个参数均存在个体间变异		
	True	Pop	Far	True	Pop	Far
θ_1	2.06 (19)	2.07 (19)	2.07 (19)	-1.76 (11)	-0.60 (11)	-1.65 (11)
θ_2	-0.92 (12)	-0.91 (12)	-0.92 (12)	7.57 (21)	7.49 (20)	9.22 (21)
β	3.94 (47)	3.91 (47)	3.96 (47)	11.57 (56)	12.38 (56)	12.50 (57)
ω_1	0.85 (10)	0.84 (10)	0.85 (10)	-0.89 (20)	0.39 (20)	-0.19 (20)
ω_2				3.29 (29)	5.19 (25)	6.47 (26)

表7：两种情景下，从不同初始值集合出发时各参数的偏倚（%）（相对均方根误差，%）（真实：用于数据模拟的参数；Pop：按真实数据集初始化的参数（θ_1 =−0.5, $\theta_2 = -0.19, \beta = 0, \omega_1 = 1, \omega_2 = 1$）；Far：固定效应初始化为0，变异性初始化为 $\omega_1 = 2, \omega_2 = 0.7$）。

小提琴图

图7展示了三种不同设置下图 2 的小提琴图。与前一个表格一样，该图表明初始条件对参数估计值没有重大影响。

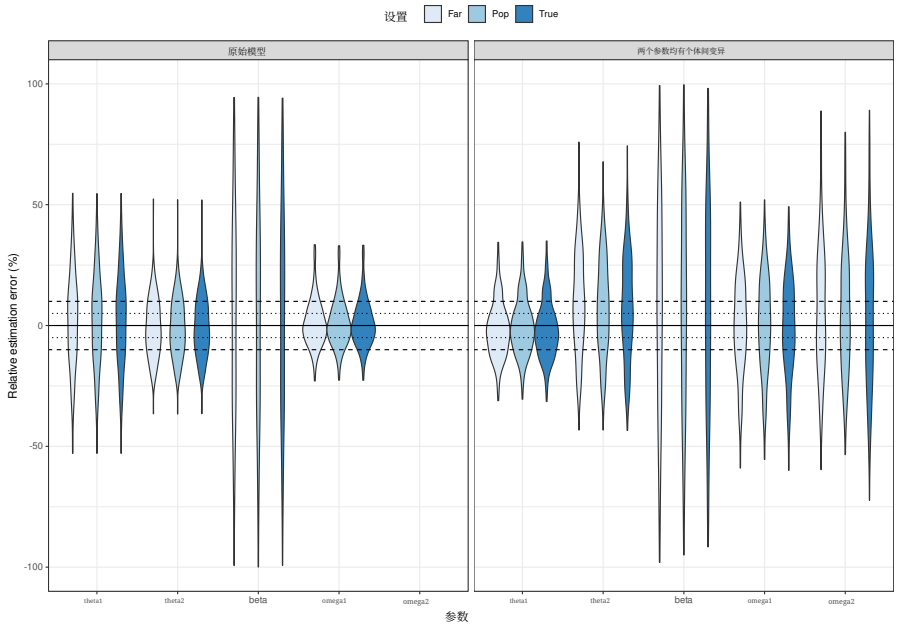

图7：在不同初始条件下，原始设计（左）和两个参数均存在变异性的设计（右）中各参数的相对估计误差。虚线标示绝对相对偏倚在 10% 以内的范围，点线标示偏倚在 5% 以内的范围。图表截取至 +/-100，3 次分析中共省略了 β 个值。

附录 C - 分类数据

用于分析膝痛评分的 saemix 代码

```
数据(knee.saemix)

ordknee.data <-saemixData(name.data =knee.saemix,name.group =c("id"),name.predict
                                                                                    "治疗")
plotDiscreteData(ordknee.data, 结局="分类", 其中。协变量=                          "性别")
11 # 序数响应模型 12 有序.model<- 函数(psi, id, xidep) { 13 y<- xidep [ ,1 ]14 时间<- xidep [ ,2 ]15 alp 1 <- psi [
id,
alp
ex
28
pg
(lo
alp
45
    logit1<-alp1 + beta*time
46  logit2<-logit1+alp2
47  logit3<-logit2+alp3
48  logit4<-logit3+alp4
49  pge1<-exp(logit1)/(1+exp(logit1))
50  pge2<-exp(logit2)/(1+exp(logit2))
```

```
1 pge3<-exp(logit3)/(1+exp(logit3))2 pge4<-exp(logit4)/(1+exp(logit4))3 x<-runif(l
5 e
5 p
5 v
5 p
5 n
5 s
5 n
6 0
6 0
6 o
6 B
6 s
6 a
6 n
6 o
6 n
7 s
7 n
7 1
7 o
7 e
7 c
7 o
7 f
7 t
7 o
8 o
8
8
8
8
8
8
87
8
8
90
9
9
9
9
9
9
9
```

膝关节数据的补充结果

图 图8显示了 knee 数据中按治疗组分层的疼痛评分比例随时间的变化 a.

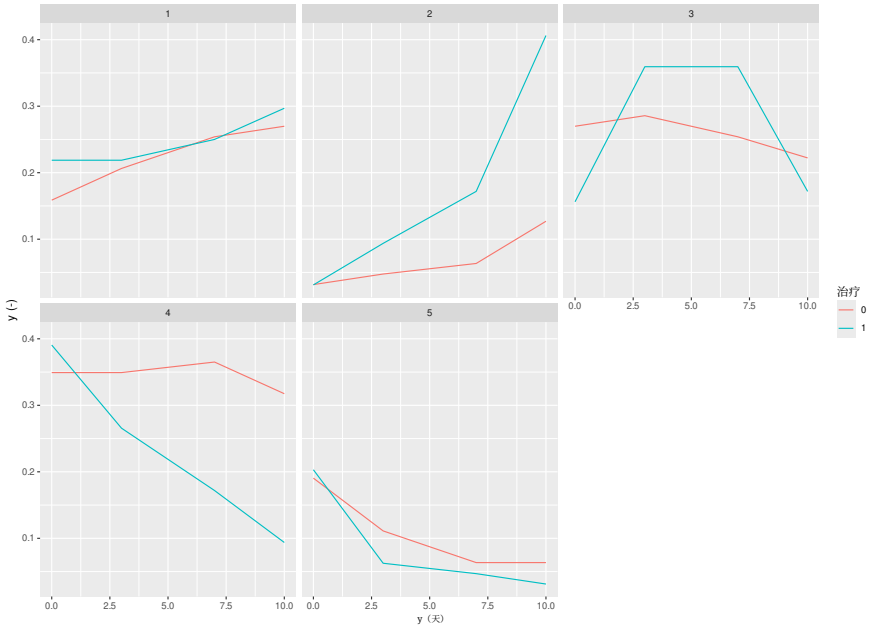

图8：按治疗组分层的各评分观察比例随时间的变化。

图9给出了用于模型评估的其他诊断结果，其中疼痛评分按性别分层；图10显示了每种治疗的平均疼痛评分及其预测区间。

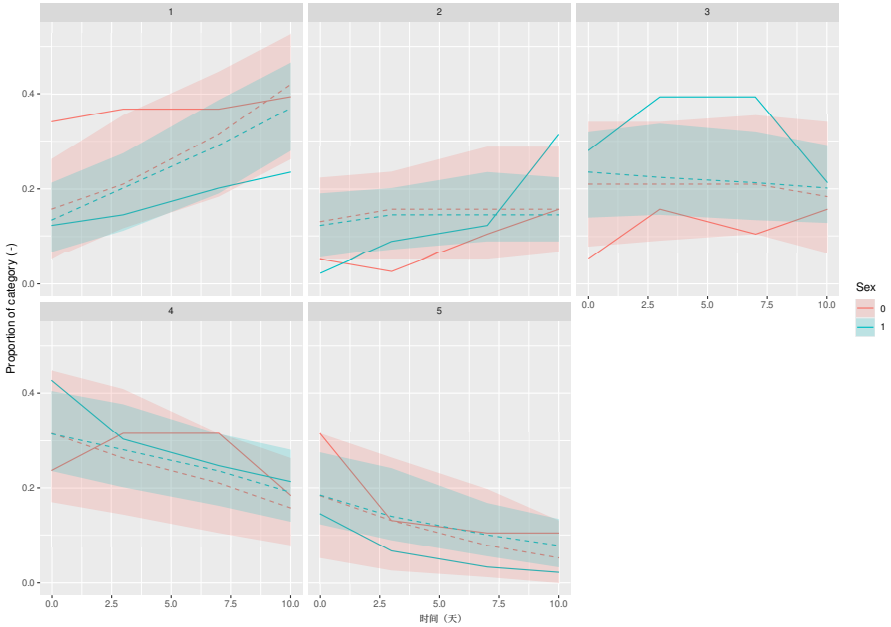

图9：按性别分层的比例优势模型中各评分值的视觉预测检验。

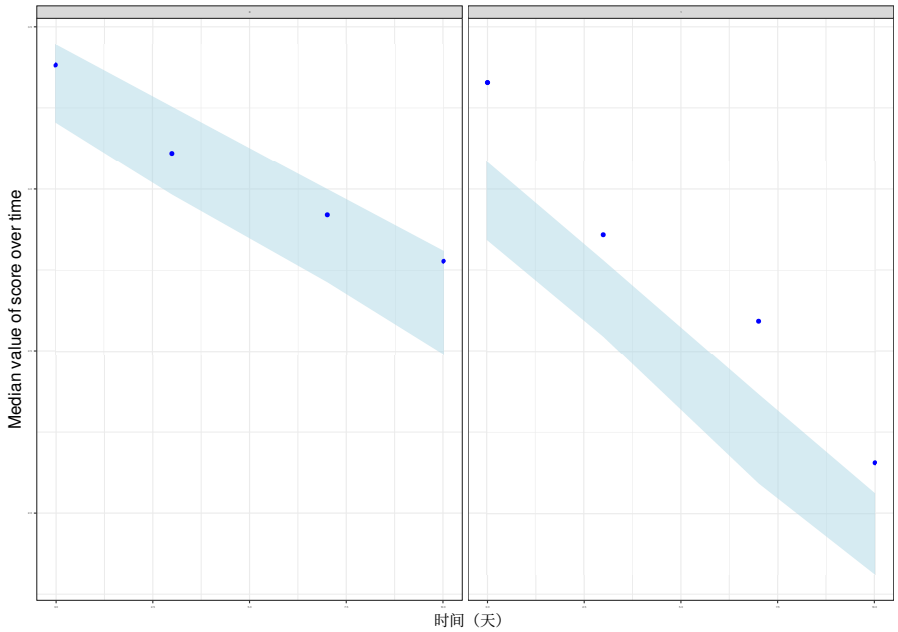

图 10：各时间点平均得分的视觉预测检验。

附录 D - 计数数据模型

saemix 代码

```
1 data(rapi.saemix)2 rapi.saemix$性别  <-ifelse(rapi.saemix$性别 =="男性",1,0)#
```

```
y<-xidep[,2]
ymax<-max(y)
intercept <-psi[id,1]
slope <-psi[id,2]
p0<-ps
```

RAPI 数据集的其他结果

图 11 显示了按性别分层的、呈现各评分的受试者比例随时间的变化。我们观察到，随着时间推移，报告无酒精相关问题的受试者比例有所增加，这源于所报告的暴饮事件数量缓慢但持续下降。我们还注意到，在趋势和所报告事件数量上均存在明显的性别差异。

表 8 给出了三个模型的参数估计及其使用条件自助法估计的 95% 置信区间。对于障碍模型，我们展示了两个子模型的参数估计：针对零计数的逻辑回归和针对非零计数的截断泊松分布。

图 12 比较了按性别分层的零计数比例随时间的演变。泊松模型的视觉预测检验显示，在整个研究期间，对报告无饮酒事件的受试者比例存在低估，在女性中尤为明显，而另外两个模型表现良好。

泊松模型的诊断图见图 13，而障碍模型非零计数的相应图（图 14）表明其在各计数类别上具有更好的拟合度。

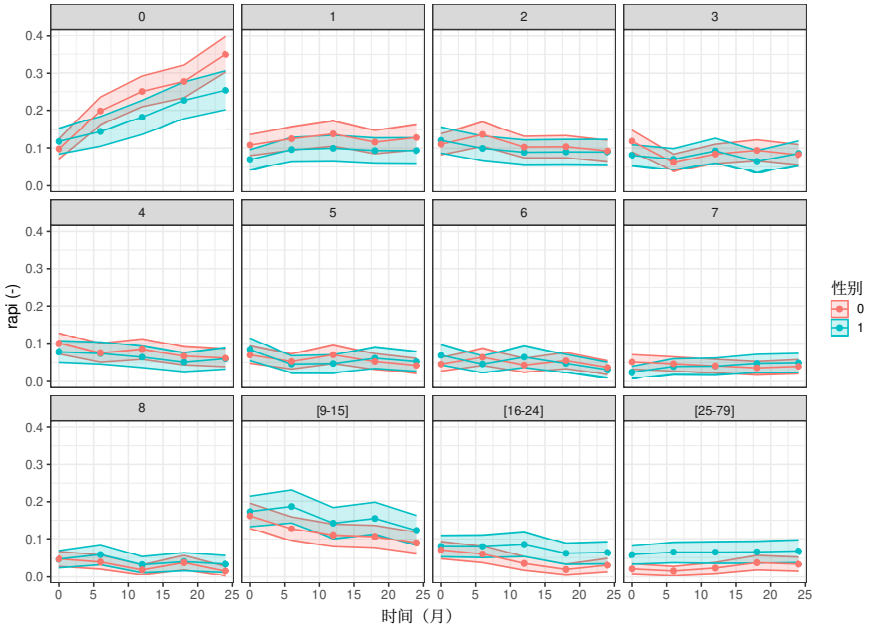

图 11：按性别分层的各计数观测比例随时间的变化。较大的计数已合并（见分面标题）。

	泊松	ZIP	栏模型	
			逻辑回归	截断泊松分布
α_0 (-)	1.49 (0.05)	1.57 (0.045)	2.49 (0.165)	1.61 (0.05)
β_{male,α_0} (-)	0.20 (0.07)	0.19 (0.07)	0.01 (0.19)	0.21 (0.07)
α_1 ($mo^{(-1)}$)	-0.039 (0.0035)	-0.035 (0.0035)	-0.054 (0.011)	-0.018 (0.003)
β_{male,α_1} (-)	0.017 (0.005)	0.016 (0.005)	0.038 (0.015)	0.01 (0.004)
p_0 (-)	-	0.076 (0.005)	-	-
ω_{α_0} (-)	0.92 (0.06)	0.80 (0.05)	1.54 (0.47)	0.72 (0.05)
ω_{α_1} (-)	0.004 (0.0003)	0.003 (0.0002)	0.007 (0.002)	0.002 (0.0002)
$\rho(\alpha_0, \alpha_1)$ (-)	-0.14 (0.047)	-0.085 (0.05)	0.38 (0.20)	-0.29 (0.055)

表 8：泊松、零膨胀泊松和障碍模型的参数估计及其由条件自助法估计的标准误。

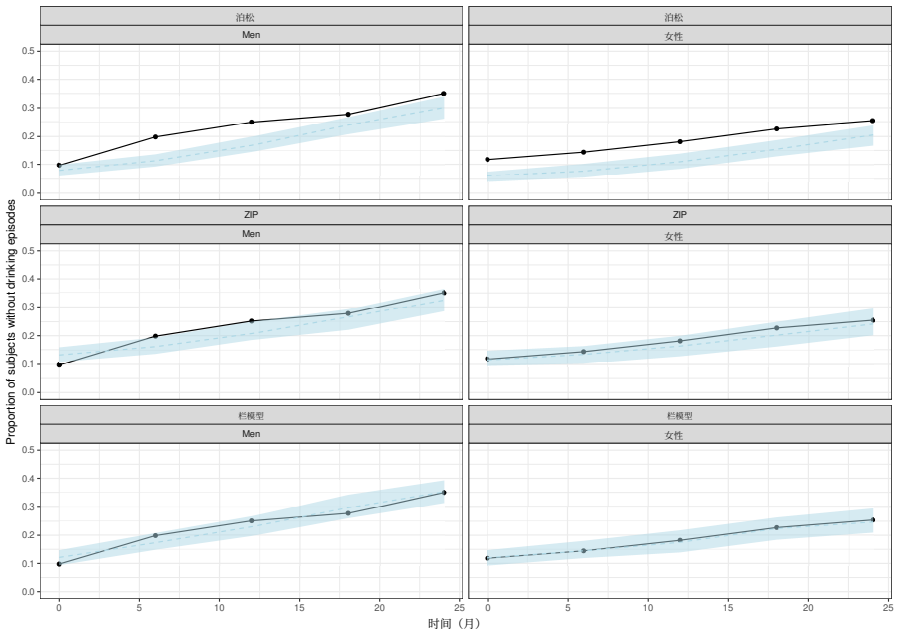

图 12：按性别分层的各模型（泊松、零膨胀泊松和栏模型）中未报告任何饮酒事件的受试者比例的视觉预测检验。

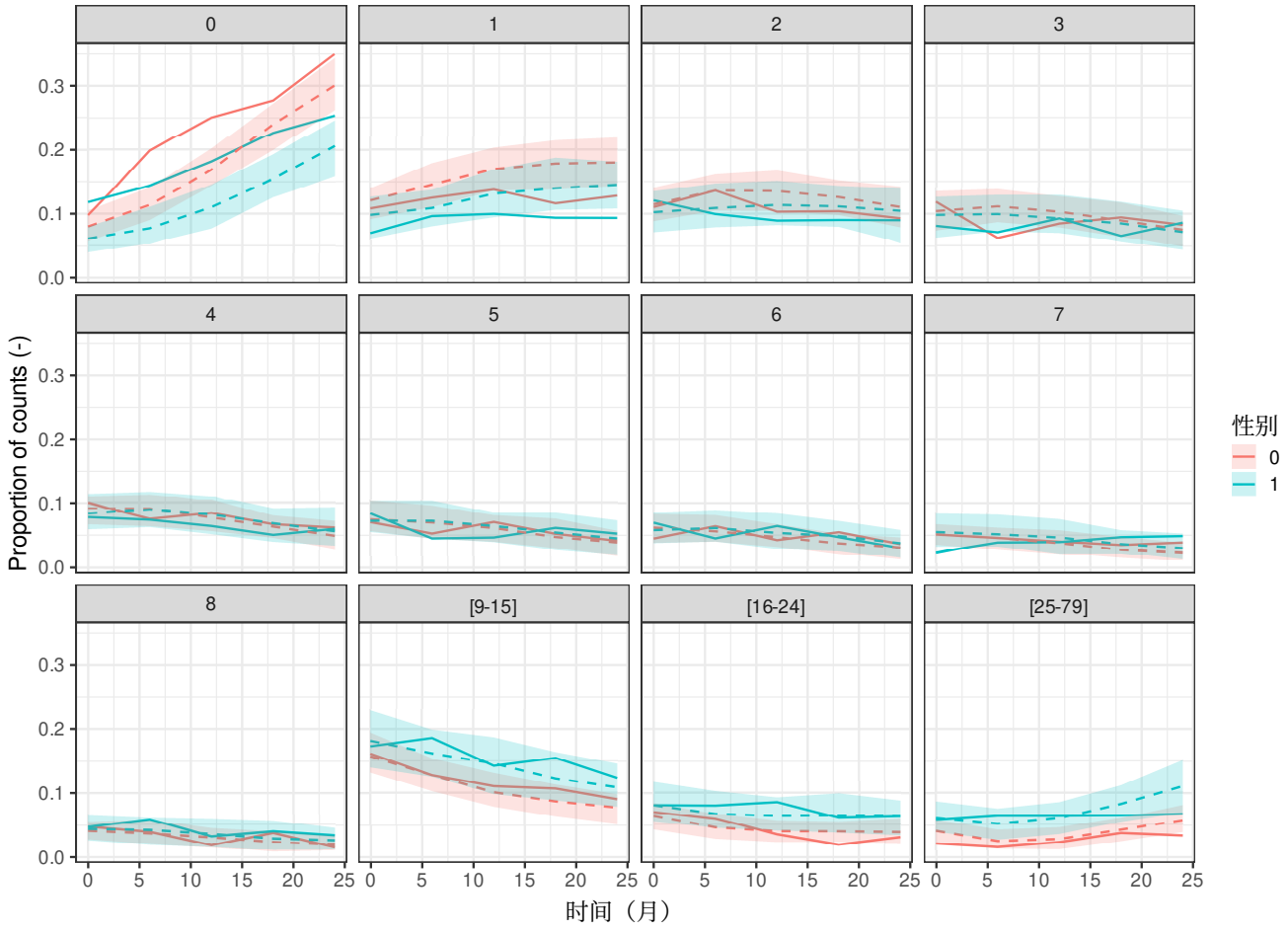

图 13：按性别分层的泊松模型视觉预测检验图。

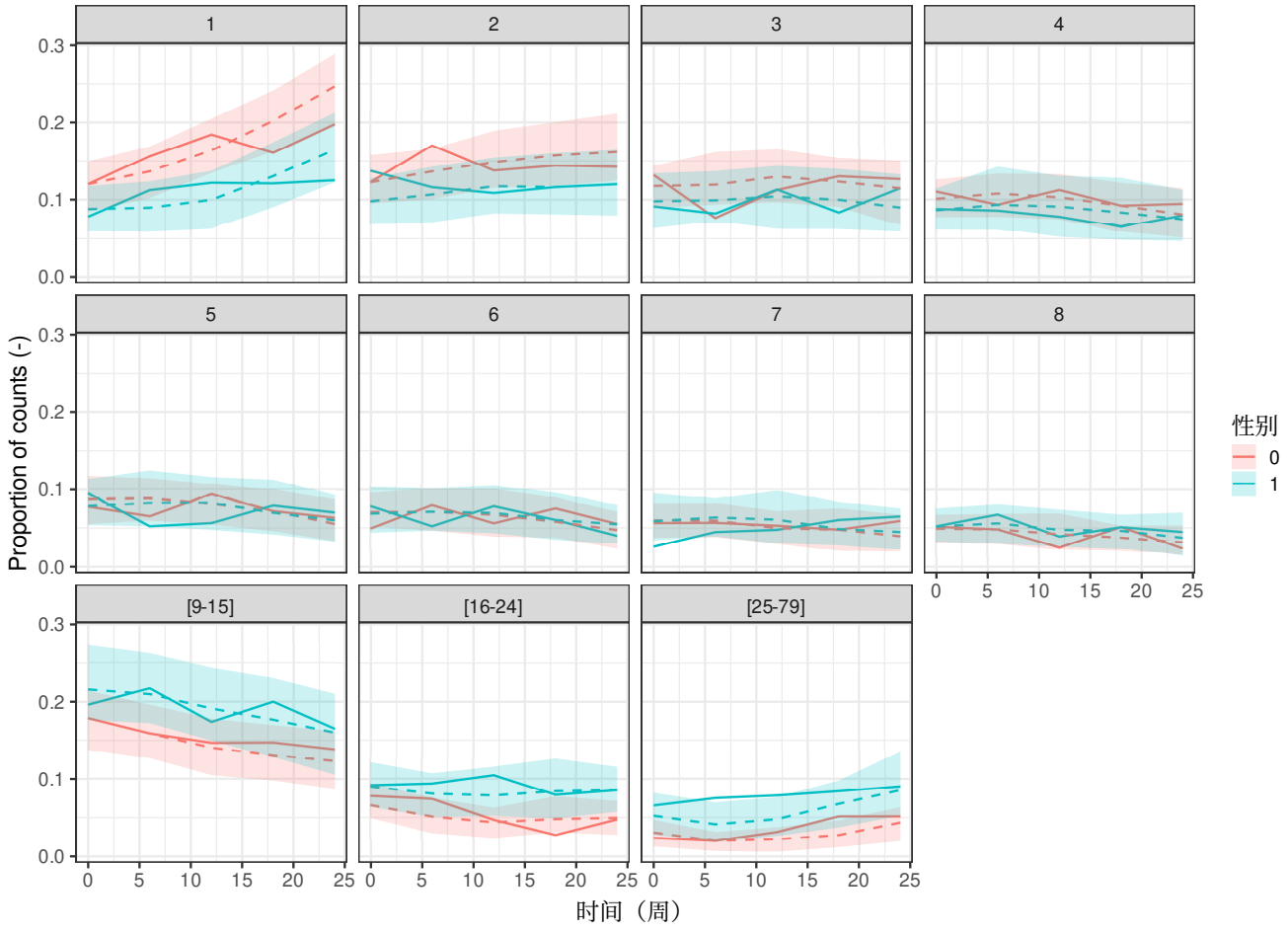

图 14：使用栏模型、按性别分层的非零类别 VPC图。

附录 E - 生存时间模型

saemix 代码

```
data(lung.saemix)2lung2<-lung.saemix3lung2$ecog1<-ifelse(lung2$ph.ecog==1,
```

```
covmodel  <-cbind(c(0,1,0,1,0,0),rep(0,6))9weibull.model.cov2<-saemixModel(
n
p
r
T
=
w
s
y
g
b
g
n
```