采用康威–麦克斯韦–泊松设定的贝叶斯死亡率预测

Jackie Siaw Tze Wong [a]和 Emiliano A. Valdez, [b]

[a]数学、统计与精算科学学院，埃塞克斯大学，Wivenhoe Park，科尔切斯特，CO4 3SQ，英国。电子邮箱：jw19203@essex.ac.uk。通讯作者。[b]数学系，康涅狄格大学，斯托斯，CT 06269-1009，美国。电子邮箱：emiliano.valdez@uconn.edu。

arXiv:2601.01686v1 [stat.ME] 4 Jan 2026

2026年1月6日

摘要

本文提出一种随机死亡率建模的新方法，使用康威–麦克斯韦–泊松（CMP）分布对死亡人数建模。与标准泊松分布或负二项分布不同，CMP 是一种适应性更强的选择，能够处理数据中不同程度的变异性，这一特征称为离散程度。具体而言，它可以处理欠离散（变异性低于预期）、等离散（变异性与预期相当）和过度离散（变异性高于预期）的数据。我们构建了一种贝叶斯建模，将离散程度视为未知参数，并采用伽马先验，以稳健且一致的方式综合处理参数、过程和分布不确定性。模型采用马尔可夫链蒙特卡罗（MCMC）方法进行校准，并使用残差分析和评分规则等标准统计准则评估模型表现。基于英格兰和威尔士男性死亡率数据的实证研究表明，与传统泊松模型和负二项模型相比，我们基于 CMP 的模型对现有数据和未来预测的拟合效果更好，尤其是在数据呈现过度离散时。最后，我们针对先验设定进行了敏感性分析，以评估模型的稳健性。

关键词：贝叶斯估计；Conway–Maxwell–Poisson (CMP)；李-卡特；死亡率预测；过度离散；不确定性校准

1 引言

死亡率预测对人口统计学家、精算师和政策制定者至关重要，因为它能够帮助他们了解未来的人口动态、退休制度和社会保障制度的可持续性，以及保险公司的偿付能力。Lee–Carter (LC) 模型（Lee 和 Carter，1992年）的提出是该领域的一项重大研究进展，该模型引入了一种简单而强大的随机框架，用于分解并外推年龄别死亡率。Basellini 等人（2023）以及 Booth 和 Tickle（2008）对后续发展进行了全面综述。虽然 LC 等基础模型通常假设死亡人数服从简单分布（如泊松分布），但现代研究的一个重要方向是恰当地对离散程度建模，即刻画死亡数据中模型的主要结构通常无法捕捉的意外噪声。考虑过度离散至关重要；若未能准确地对这种额外变异性建模，模型的预测区间就会过窄。这进而会导致长寿风险被低估，并可能造成重大的财务后果。

本文提出了三项主要贡献。首先，我们引入 Conway–Maxwell–Poisson（CMP）分布，将其作为一种灵活的方法，用于对存在离散程度的死亡人数进行建模。CMP 分布最初由 Conway 和 Maxwell（1962）提出，据我们所知，此前尚未应用于死亡率建模领域。该模型能够捕捉死亡数据中的欠离散、等离散和过度离散现象，并为展示其在人口统计学和精算应用中的多功能性及实际相关性提供了机会。

其次，我们对所提出的模型进行了贝叶斯建模，将离散程度视为一个未知但可估计的参数，并为其指定伽马先验。该设定能够对离散程度的性质和水平进行明确推断，同时在统一的概率框架内一致地整合多个不确定性来源。贝叶斯方法的优势已在相关文献中得到充分论述，例如 Cairns 等人（2011）和 Wong 等人（2018）。此外，近期将贝叶斯方法应用于死亡率和长寿建模的研究也日益增多，例如 Diana 等人（2025）、Goes 等人（2025）和 Shi 等人（2024）。

最后，我们使用英格兰和威尔士的男性死亡数据展示所提出的方法。实证结果令人鼓舞，并表明相较于传统的泊松和负二项设定，该方法有明显改进。

本文其余部分安排如下。第 1 节余下内容介绍数据、符号表示，以及用于死亡人数建模的泊松设定。第 2 节介绍 CMP 分布，将其作为处理过度离散的灵活替代方案。第 3 节概述贝叶斯实现，包括先验设定及相应的估计算法。第 4 节报告基于英格兰和威尔士男性死亡率数据的实证结果。最后，我们对先验设定进行敏感性分析，并作出总结。

1.1 数据与符号

令 d_{xt} 表示年龄组 x 在年份 t的观测死亡人数，其中 $x = \{x_1, \ldots, x_A\}$和 $t = \{t_1, \ldots, t_T\}$ 分别表示一组包含 A 个不同年龄组和 T个不同年份的集合。进一步假设， e_{xt} 和 μ_{xt} 分别表示相应的中心风险暴露量和中心死亡率。

用于说明的数据是从人类死亡率数据库（HMD）[†]中提取的英格兰和威尔士男性死亡数据及其相应的暴露量。数据按单岁年龄（0至99岁）和年份（1961年至2002年）分类。因此，这里有 $\{x_1, \ldots, x_A\} = \{0, \ldots, 99\}$ 和 $\{t_1, \ldots, t_T\} = \{1961, \ldots, 2002\}$，分别对应 $A = 100$ 和

[†]参见 http://www.mortality.org。

$T = 42$。2003—2021 年的数据被保留为验证集（见第 4 节 ）.

1.2 用于死亡人数建模的泊松分布

泊松分布通常用于对死亡人数建模（例如，参见 McCullagh，2019 年或 Selvin，2004 年）。较新的应用包括 Hayat 和 Higgins（2014 年）、Tawiah 等人（2020 年）、Zhang 等人（2023 年）等。设 D_{xt} 为表示 t年 x 岁死亡人数的随机变量，则

$$D_{xt}|\mu_{xt} \sim \text{Poisson}(e_{xt}\mu_{xt}), \tag{1}$$

其对应的概率质量函数（PMF）为

$$\mathbb{P}[D_{xt} = d_{xt}|\mu_{xt}] = \frac{(e_{xt}\mu_{xt})^{d_{xt}} \exp(-e_{xt}\mu_{xt})}{d_{xt}!} \propto \mu_{xt}^{d_{xt}} \exp(-e_{xt}\mu_{xt}). \tag{2}$$

μ_{xt} 有多种结构假设，每种假设都定义了文献中的不同死亡率模型。例如，指定广受欢迎的李-卡特模型（Lee 和 Carter，1992年），可得:

$$\begin{cases} D_{xt}|\mu_{xt} \sim \text{Poisson}(e_{xt}\mu_{xt}) \ , \\ \log \mu_{xt} = \alpha_x + \beta_x \kappa_t \ , \end{cases} \tag{3}$$

为确保模型的可识别性，采用以下约束:

$$\sum_x \beta_x = 1 \quad \text{and} \quad \sum_t \kappa_t = 0.$$

由此得到 Brouhns 等人（2002）提出的**泊松李-卡特模型（泊松–LC）**。参数可解释如下:

α_x： 表示一段时间内死亡率对数的平均值（即 $\alpha_x = {}^{\sum_t \log_T}{}^{\mu_{xt}}$）。$\beta_x$： 表示死亡率改善的年龄特异性模式，用于量化各年龄的死亡率对整体变化的敏感性。κ_t： 表示死亡率的总体时间趋势，并由 β_x 进行适当调节。

在本文中，我们还考虑纳入队列效应以进行比较。遵循 Renshaw和Haberman（2005年）的表述,

$$\begin{cases} D_{xt}|\mu_{xt} \sim \text{Poisson}(e_{xt}\mu_{xt}) \ , \\ \log \mu_{xt} = \alpha_x + \beta_x \kappa_t + \gamma_c \ , \end{cases} \tag{4}$$

其中， γ_c为队列参数， $c = t - x \in \{1, \ldots, C\}$ 表示队列索引，对应于出生于 $\{1861, \ldots, 2001\}$ 的队列，且 $C = A + T - 1$。我们将遵循 Wong 等人（ 2023 ）的建议，施加以下约束条件 †:

$$\sum_t \kappa_t = \sum_t t\kappa_t = \sum_c \gamma_c = \sum_c c\gamma_c = \sum_c c^2\gamma_c = 0, \tag{5}$$

我们将这一模型称为**含队列效应的泊松李-卡特模型（泊松–LCC）**模型。

†我们注意到，这些约束不同于 Renshaw和Haberman（2005年）以及 Hunt和Villegas（2015年）所采用的约束，但这一选择能够实现稳定的计算，并具有良好的收敛性。

1.3 过度离散

泊松模型假定均值-方差相等（$\mathbb{E}[D_{xt}]$ =Var$[D_{xt}] = e_{xt}\mu_{xt}$），这一限制性假设在实践中很少成立。它意味着每个年龄-时期单元内部具有同质性，即假设相同年龄、相同年份的个体具有相同的死亡风险。这并不现实，因为死亡率还会受到吸烟、收入、族裔和遗传背景等因素的影响 Brown（2003），从而导致个体之间出现额外差异，这种现象通常称为过度离散。

多篇文章都指出，泊松分布之外的变异会导致过度离散，例如 Brillinger（1986）、Breslow（1984）、Hinde 和 Dem´etrio（1998）、Hougaard 等（1997）、Dean 和 and（1989）以及 Dean（1992）。这里，我们用皮尔逊残差的平方来说明这一现象，如下所示:

$$r_{xt}^2 = \left.\frac{(d_{xt} - \mathbb{E}[D_{xt}])^2}{\text{Var}[D_{xt}]}\right|_{\mu_{xt}=\hat{\mu}_{xt}} = \frac{(d_{xt} - e_{xt}\hat{\mu}_{xt})^2}{e_{xt}\hat{\mu}_{xt}}, \tag{6}$$

其中，$\hat{}\mu_{xt}$ 是潜在死亡率的最大似然估计（MLE）。具体而言，

$$\hat{\mu}_{xt} = \begin{cases} \exp\left(\hat{\alpha}_x + \hat{\beta}_x\hat{\kappa}_t\right), & \text{for Poisson–LC,} \\ \exp\left(\hat{\alpha}_x + \hat{\beta}_x\hat{\kappa}_t + \hat{\gamma}_c\right), & \text{for Poisson–LCC.} \end{cases}$$

其中，$\hat{}\alpha_x$、$\hat{}\beta_x$、$\hat{}\kappa_t$、$\hat{}\gamma_c$ 是参数的最大似然估计。随后可构建 r_{xt}^2 的颜色编码热图，以直观展示 Poisson–LC 和 Poisson–LCC 模型对死亡数据的拟合不足情况，如图1所示。

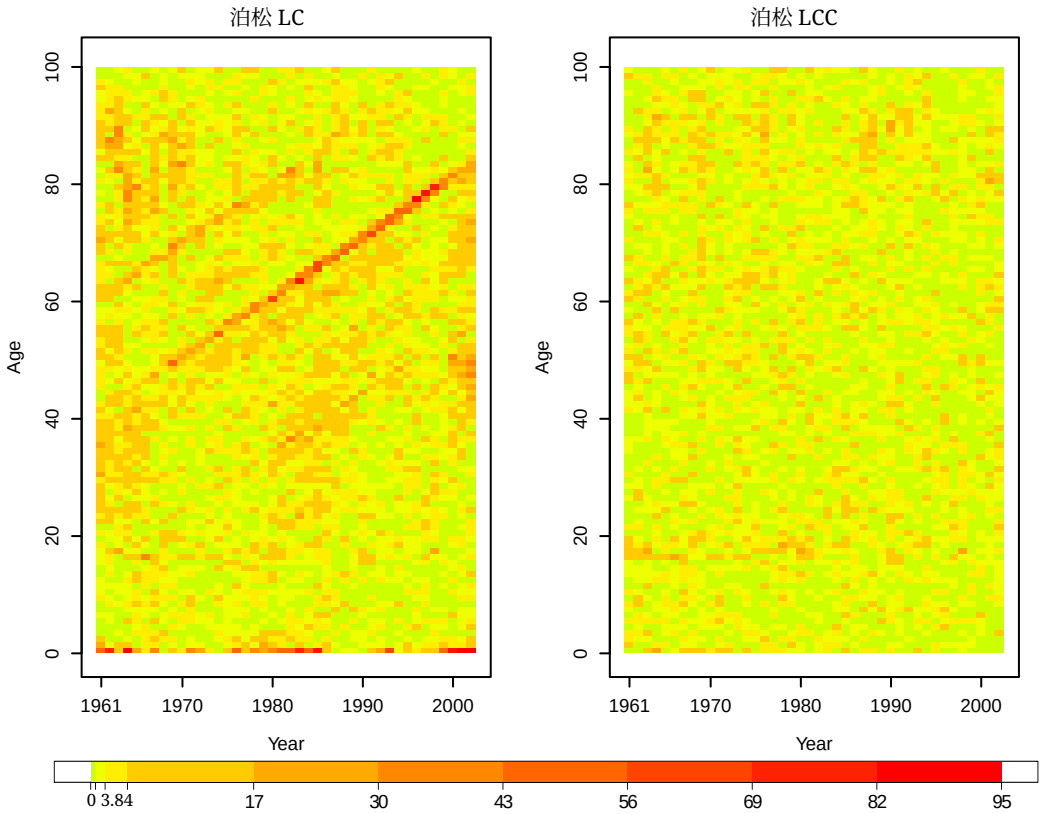


图1：泊松–LC（左）和泊松–LCC（右）模型的 r_{xt}^2 热图。绿色/黄色矩形单元格表示拟合良好的区域；橙色/红色单元格表示拟合明显不佳的区域。

在 Poisson–LC 模型为真实底层模型的零假设下（并满足一些温和条件），每个 r_{xt}^2 渐近地近似服从自由度为一（χ_1^2）的卡方分布。理想情况下，矩形单元格（$AT \times 0.05 = 210$）中应只有约 5% 拟合不佳（我们将其定义为 $r_{xt}^2 > 3.84$，其中 3.84 是 χ_1^2 的第 95^{th} 百分位数）。然而，从图1可以明显看出，Poisson–LC 模型的热图中散布着超过预期数量的橙色/红色单元格（1128 个，约占 27%）；这一现象在婴儿以及约于 1920 年出生的队列中尤为明显（橙色/红色单元格呈对角线分布），这表明该模型无法充分解释数据中的额外变异。即使像 Poisson–LCC 模型那样纳入队列效应，热图中仍散布着橙色/红色单元格（474 个，约占 11%），说明整体拟合不佳。

此外，我们还可以进行Pearson卡方总体拟合优度检验。具体而言，模型偏差计算为 r_{xt}^2之和，

$$r^2 = \sum_{x,t} r_{xt}^2 = \sum_{x,t} \frac{(d_{xt} - e_{xt}\hat{\mu}_{xt})^2}{e_{xt}\hat{\mu}_{xt}},$$

对于泊松–LC和Poisson–LCC模型，所得数值分别为16709.85和6628.97。再次假设该模型与数据具有良好拟合这一零假设，r^2 应近似服从自由度（df）为 $df = (A-1)(T-2)$ =3960的卡方分布（参见Renshaw和Haberman，2005年）。由于16709.85和6628.97的模型偏差明显大于传统卡方统计量的临界值（即 χ_{df}^2 的95^{th} 百分位数为4107.51），这清楚地表明，泊松–LC和Poisson–LCC模型均无法对数据提供令人满意的拟合。

正如 Wong 等人（2018年）所指出的，忽略过度离散会导致方差被低估，从而造成过拟合，并产生过窄的预测。考虑过度离散可以得到更灵活的拟合，更好地捕捉异质性，并生成更可靠的预测区间（见第 4 节.3）。此外，纳入队列效应可进一步改善数据与不确定性之间的校准，在保持区间宽度适当的同时减少投影偏差。在接下来的章节中，我们将展示 CMP分布如何处理死亡人数中的超泊松变异。

2 Conway–Maxwell–Poisson（CMP）分布

CMP分布最初由 Conway 和 Maxwell（1962年）提出，作为泊松分布的一种推广。服从CMP分布的随机变量 Y的概率质量函数为：

$$\mathbb{P}[Y = y|\lambda, \nu] = \frac{1}{Z(\lambda, \nu)} \frac{\lambda^y}{(y!)^\nu}, \qquad \text{for } y \in \{0, 1, 2, \ldots\},$$

其中

$$Z(\lambda, \nu) = \sum_{j=0}^{\infty} \frac{\lambda^j}{(j!)^\nu}$$

是确保概率质量函数总和为 1 的归一化常数。模型参数满足$\lambda > 0$ 和 $\nu \geq 0$。这通常记作 $Y \sim$ CMP (λ, ν)。

2.1 一些性质

该分布的主要优势在于，它能够通过离散参数 ν 对呈现过度离散和欠离散的计数数据进行建模。我们注意到，

$Y \sim$ CMP (λ, ν) 的矩为

$$\mathbb{E}[Y] = \sum_{j=0}^{\infty} \frac{j\lambda^j}{(j!)^\nu Z(\lambda,\nu)}, \qquad \text{and} \qquad \text{Var}[Y] = \sum_{j=0}^{\infty} \frac{j^2\lambda^j}{(j!)^\nu Z(\lambda,\nu)} - (\mathbb{E}[Y])^2.$$

更具体地说，当 ν 和 λ不太小时，我们有

$$\mathbb{E}[Y] \approx \lambda^{1/\nu} + \frac{1}{2\nu} - \frac{1}{2} \qquad \text{and} \qquad \text{Var}[Y] \approx \frac{\lambda^{1/\nu}}{\nu}, \tag{7}$$

近似成立（例如参见 Chanialidis 等人，2018年）。这使得 $\text{Var}[Y] > \mathbb{E}[Y]$ 可以过度离散（当 $\nu < 1$时），$\text{Var}[Y] < \mathbb{E}[Y]$ 可以欠离散（当 $\nu > 1$时），或 $\text{Var}[Y] = \mathbb{E}[Y]$ 可以等离散（当 $\nu = 1$时），具体取决于 ν的值。

另请注意，后续概率之比

$$\frac{\mathbb{P}[Y = y-1|\lambda,\nu]}{\mathbb{P}[Y = y|\lambda,\nu]} = \frac{y^\nu}{\lambda},$$

这表明衰减率符合以下情况之一：

- $\nu < 1$：衰减速度慢于线性下降，这意味着该分布具有更重/更长的尾部（过度离散）。
- $\nu = 1$：衰减速度呈线性，因此该分布成为普通的泊松分布（等离散）。
- $\nu > 1$：衰减速度快于线性下降，这意味着该分布的尾部更轻、更短（欠离散）。
- $\nu = 0$：这是一个特殊情形，此时 CMP分布退化为几何分布，概率为 $1-\lambda$，其中 $0 < \lambda < 1$。
- $\nu \to \infty$：这是一个特殊情形，此时 CMP分布退化为伯努利分布，成功概率为 $\frac{\lambda}{1+\lambda}$。

涉及 CMP分布的计算可使用 R 软件包 *COM-PoissonReg* （Sellers 等人，2023）。关于 CMP分布的更多细节，参见 Daly 和 Gaunt（2016）。

2.2 用于死亡人数建模的 CMP分布

Wong 等人（2018）在死亡率建模/预测的背景下探讨了泊松假设中的过度离散问题，并针对上述设定提出了两种扩展。在本研究中，我们提出使用 CMP分布作为一种替代方案，它引入了一个通用的离散参数，以放宽泊松模型的严格假设。也就是说，我们提出

$$D_{xt}|\mu_{xt},\nu \sim \text{CMP}\left((e_{xt}\mu_{xt} + \frac{1}{2} - \frac{1}{2\nu})^\nu, \nu\right), \tag{8}$$

其中，ν是待估计的离散参数。为简便起见，假设记作 $\lambda_{xt} = (e_{xt}\mu_{xt} + \frac{1}{2} - \frac{1}{2\nu})^\nu$，相应的概率质量函数为

$$\mathbb{P}[D_{xt} = d_{xt}|\mu_{xt},\nu] = \frac{1}{Z(\lambda_{xt},\nu)} \frac{\lambda_{xt}^{d_{xt}}}{(d_{xt}!)^\nu},$$

其中

$$Z(\lambda_{xt},\nu)=\sum_{j=0}^{\infty}\frac{\lambda_{xt}^{j}}{(j!)^{\nu}}\ .$$

请注意，这种（略显非常规的）表述旨在确保预期死亡人数等于暴露量与死亡率的乘积，即

$$\begin{aligned}\mathbb{E}[D_{xt}] &\approx ((e_{xt}\mu_{xt}+\frac{1}{2}-\frac{1}{2\nu})^{\nu})^{1/\nu}+\frac{1}{2\nu}-\frac{1}{2}\\ &= e_{xt}\mu_{xt}+\frac{1}{2}-\frac{1}{2\nu}+\frac{1}{2\nu}-\frac{1}{2}\\ &= e_{xt}\mu_{xt},\end{aligned}\tag{9}$$

如（7）所示。

我们指出，为实现 $\mathbb{E}[D_{xt}]\approx e_{xt}\mu_{xt}$，指定 $D_{xt}|\mu_{xt},\nu\sim$CMP $((e_{xt}\mu_{xt})^{\nu},\nu)$或许就足够了，尤其是因为 ν 通常不会太小。不过，我们仍采用（8）中的表述，因为它能在不显著增加复杂性的情况下进行更准确的调整。

该设定还意味着，以均值表示的死亡人数方差为

$$\begin{aligned}\mathrm{Var}[D_{xt}] &\approx \frac{((e_{xt}\mu_{xt}+\frac{1}{2}-\frac{1}{2\nu})^{\nu})^{1/\nu}}{\nu}\\ &= \frac{e_{xt}\mu_{xt}+\frac{1}{2}-\frac{1}{2\nu}}{\nu}\\ &= \frac{\mathbb{E}[D_{xt}]+\frac{1}{2}-\frac{1}{2\nu}}{\nu}.\end{aligned}\tag{10}$$

因此， ν 决定了过度离散的程度，其中术语 $1/\nu$ 也可以（近似地）解释为 D_{xt} 的方差相对于其均值的倍数增加。例如，如果 $\nu<1$，则有 $\mathrm{Var}[D_{xt}|\mu_{xt},\nu]>\mathbb{E}[D_{xt}|\mu_{xt},\nu]$，这属于过度离散的情况，反之亦然。而当 $\nu=1$时，$\mathrm{Var}[D_{xt}|\mu_{xt},\nu]=\mathbb{E}[D_{xt}|\mu_{xt},\nu]$，表示等离散的情况。

与之前类似，我们考虑两个关于 μ_{xt}的结构模型，

$$\log\mu_{xt}=\begin{cases}\alpha_x+\beta_x\kappa_t,\\ \alpha_x+\beta_x\kappa_t+\gamma_c,\end{cases}\tag{11}$$

并结合使用（8）。我们将前者称为**CMP–LC** 模型，后者称为**CMP–LCC** 模型。

利用独立性和上述概率质量函数，我们还可以得到完整数据似然为

$$f(\boldsymbol{d}|\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\kappa},\nu)=\prod_{x,t}\mathbb{P}[D_{xt}=d_{xt}|\mu_{xt},\nu]=\prod_{x,t}\frac{1}{Z(\lambda_{xt},\nu)}\frac{(\lambda_{xt})^{d_{xt}}}{(d_{xt}!)^{\nu}}.\tag{12}$$

3 贝叶斯估计

采用贝叶斯框架的动机已在文献中得到广泛讨论（例如，Wong等人，2018）。在我们的设定中，这种方法允许将离散参数ν 视为未知随机变量，并直接、自然地从数据中进行估计。这种数据驱动的设定无需预先指定 ν的固定值，即可对离散程度进行推断。这也为不确定性校准提供了一个一致的框架，其中包括过度离散所导致的不确定性，以及通过先验分布体现的参数不确定性等其他不确定性来源。在以下小节中，我们将介绍先验设定，推导联合后验分布，并概述我们提出的MCMC抽样过程。

3.1 先验设定

原则上，所设定的先验分布应反映我们对模型参数的先验知识。但在大多数应用中，我们并不具备领域先验知识，因此贝叶斯方法允许我们从数据中学习所有参数（包括 ν）。在本文中，我们将主要遵循Wong 等人（2023）提出的设定，但对于 ν，我们将探讨各种设定。

首先，我们设定

$$\begin{aligned}\boldsymbol{\alpha} \sim &\quad N(-5 \cdot \mathbf{1}_A, 4 \cdot I_A),\\ \boldsymbol{\beta} \sim &\quad N\left(\frac{1}{A} \cdot \mathbf{1}_{A-1}, 0.005 \cdot \left(I_{A-1} - \frac{1}{A} \cdot J_{A-1}\right)\right),\\ \beta_1 = 1 - &\sum_{x=2}^{A} \beta_x,\end{aligned}$$

其中，$\boldsymbol{\alpha}=(\alpha_1、\ldots\ldots、\alpha_A)^\top$、$\boldsymbol{\beta}=(\beta_2、\ldots\ldots、\beta_A)^\top$，$\mathbf{1}_A$是一个长度为$-A$的全 1 向量，$I_A$和 J_A 分别是单位矩阵和维度为 $A \times A$的全 1 矩阵。

遵循 Czado 等人（2005年）的表述，我们对 κ_t 施加

$$\left.\begin{aligned}&\kappa_t - \eta_t = \rho(\kappa_{t-1} - \eta_{t-1}) + \epsilon_t, \quad \text{for } t = 2, 3, \ldots, T,\\ &\kappa_1 = \eta_1 + \epsilon_1,\end{aligned}\right\} \tag{13}$$

其中，$\eta_t = \psi_1 + \psi_2 t$表示线性漂移，$\epsilon_t \overset{\text{ind}}{\sim} N(0,\ \sigma_\kappa^2)$。然后，根据$\rho$的取值，该式对于 κ_t构成一阶自回归（AR(1)）模型（$|\rho| < 1$）或随机游走模型（$\rho = 1$）。式（13）可将其（连同约束）表示为多元形式:

$$\left.\begin{aligned}&\boldsymbol{\kappa}|\rho, \boldsymbol{\psi}, \sigma_\kappa^2 \sim N\left(\boldsymbol{\mu}_\kappa, \sigma_\kappa^2 V\right),\\ &\kappa_1 = -\textstyle\sum_{t=2}^{T} \kappa_t,\end{aligned}\right\} \tag{14}$$

其中，$\boldsymbol{\kappa}=(\kappa_2、\ldots\ldots、\kappa_T)^\top$、$\boldsymbol{\mu}_\kappa=(I_{T-1} - B_{21}B_{11}^{-1}\mathbf{1}_{T-1}^\top) \cdot X\boldsymbol{\psi}$、$X = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 2 & 3 & \cdots & T \end{pmatrix}^\top$，$\boldsymbol{\psi}=(\psi_1,\ \psi_2)^\top$，关于 B 和 V 的表达式见附录 A。对于其他超参数,

$$\begin{aligned}\frac{\rho+1}{2} \sim &\quad \text{Beta}(3,2),\ \text{where } \rho \in (-1,1),\\ \sigma_\kappa^{-2} \sim &\quad \text{Gamma}(1, 0.0001),\\ \boldsymbol{\psi} \sim &\quad N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2000 & 0 \\ 0 & 2 \end{pmatrix}\right),\end{aligned}$$

对于队列参数 γ_c，我们采用 Villegas 等人（2018年）使用的 ARIMA(1,1,0)，即

$$\left.\begin{aligned}&(\gamma_c - \gamma_{c-1}) = \rho_\gamma(\gamma_{c-1} - \gamma_{c-2}) + \epsilon_c^\gamma, \quad \text{for } c = 3, \ldots, C,\\ &\gamma_2 - \gamma_1 = \tfrac{1}{\sqrt{1-\rho_\gamma^2}}\epsilon_2^\gamma,\\ &\gamma_1 = 100\epsilon_1^\gamma,\end{aligned}\right\} \tag{15}$$

其中，$\epsilon_c^\gamma \overset{\text{ind}}{\sim} N(0, \sigma_\gamma^2)$ 对于 $c = 1, \ldots, C$ 成立，且 ρ_γ和 σ_γ^2为超参数。对（15）应用约束条件 $\sum_c \gamma_c = \sum_c c\gamma_c = \sum_c c^2\gamma_c = 0$后，可写为

$$\boldsymbol{\gamma} \sim N(\mathbf{0}, \sigma_\gamma^2 W_\gamma), \tag{16}$$

其中， $\boldsymbol{\gamma} = (\gamma_2, \ldots, \gamma_{71}, \gamma_{73}, \ldots, \gamma_{C-1})^\top$，而 W_γ 见附录 A。其余队列分量 $\{\gamma_1, \gamma_{72}, \gamma_C\}$ 随后可按如下方式推导

$$\begin{aligned}
\gamma_1 &= \frac{1}{71 \times (C-1)} \sum_{c \neq 1,72,C} (c-72)(C-c)\gamma_c, \\
\gamma_{72} &= -\frac{1}{69 \times 71} \sum_{c \neq 1,72,C} (C-c)(c-1)\gamma_c, \\
\gamma_C &= \frac{1}{69 \times (C-1)} \sum_{c \neq 1,72,C} (c-1)(72-c)\gamma_c.
\end{aligned}$$

按照 Wong 等人（2023）的建议，我们设定先验

$$\rho_\gamma \sim N(0,1) \qquad \text{and} \qquad \sigma_\gamma \sim \text{Uniform}(0.1).$$

对于 离散参数 ν，我们首先采用（非正规）平坦先验，即

$$\pi(\nu) \propto 1 \ , \ \text{for } \nu > 0 \tag{17}$$

这样，我们的贝叶斯分析所得结果便可直接与频率学派框架下的结果进行比较。这体现了数据驱动推断的理念，即让数据“自己说话”。这也有助于判断数据中是否包含足够的信息来估计 ν。我们指出，使用非正规平坦先验十分常见且方便，并且通常会得到正规的后验分布，正如 Gelfand 和 Sahu（1999）所指出的那样。

先验设定对结果的影响也值得探究。我们可以考察 $\pi(\nu)$的不同取值。我们重点关注以下形式的先验：

$$\pi(\nu) \sim \text{Gamma}(a, b),$$

其中， a 和 b 是待按如下方式设定的超参数：

- $a = 1$ 和 $b = 100$。
- $a = 1$ 和 $b = 10$。
- $a = 1$ 和 $b = 1$。
- $a = 1$ 和 $b = 0.1$。
- $a = 1$ 和 $b = 0.01$。

从上到下， ν 的先验方差从 1×10^{-4} 增至 1×10^4，这表明先验逐渐变得模糊或缺乏信息。稍后将在第 4 节中对 a 和 b的这些取值进行敏感性分析。

3.2 后验分布

首先，记 $\boldsymbol{\theta} = \{\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\kappa}, \boldsymbol{\psi}, \sigma_\kappa^2, \rho, \nu\}$ 为 CMP–LC 模型的完整参数集，则联合后验分布为

$$\begin{aligned}
\pi(\boldsymbol{\theta}|\boldsymbol{d}) &\propto f(\boldsymbol{d}|\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\kappa}, \nu) \cdot \pi(\boldsymbol{\theta}) \\
&= f(\boldsymbol{d}|\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\kappa}, \nu) \cdot \pi(\boldsymbol{\alpha}) \cdot \pi(\boldsymbol{\beta}) \cdot \pi(\boldsymbol{\kappa}|\boldsymbol{\psi}, \sigma_\kappa^2, \rho) \cdot \pi(\boldsymbol{\psi}) \cdot \pi(\sigma_\kappa^2) \cdot \pi(\rho) \cdot \pi(\nu),
\end{aligned}$$

其中， $f(\boldsymbol{d}|\boldsymbol{\alpha}$、$\boldsymbol{\beta}$、$\boldsymbol{\kappa}$、$\nu)$ 是式（12）给出的数据似然，其余部分则是第3.1节所述的所有其他参数的先验分布。类似地，对于 CMP–LCC模型， $\boldsymbol{\theta} = \{\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\kappa}, \boldsymbol{\psi}, \sigma_\kappa^2, \rho, \boldsymbol{\gamma}, \sigma_\gamma^2, \rho_\gamma, \nu\}$，得到联合后验分布

$$\begin{aligned}
\pi(\boldsymbol{\theta}|\boldsymbol{d}) \propto\ & f(\boldsymbol{d}|\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\kappa}, \nu) \cdot \pi(\boldsymbol{\alpha}) \cdot \pi(\boldsymbol{\beta}) \cdot \pi(\boldsymbol{\kappa}|\boldsymbol{\psi}, \sigma_\kappa^2, \rho) \cdot \pi(\boldsymbol{\psi}) \cdot \pi(\sigma_\kappa^2) \\
& \cdot \pi(\rho) \cdot \pi(\boldsymbol{\gamma}|\sigma_\gamma^2, \rho_\gamma) \cdot \pi(\sigma_\gamma^2) \cdot \pi(\rho_\gamma) \cdot \pi(\nu).
\end{aligned}$$

3.3 MCMC更新方案

我们采用 O'Hagan 和 Forster（2004）所述的逐变量随机游走 Metropolis-Hastings（MH）算法。有关所有其他参数的 MCMC更新方案，请参阅补充材料中的附录B。下面我们详细介绍离散参数 ν 的 MCMC更新步骤。

注意，ν 的条件后验分布正比于似然函数与其先验分布的乘积，即

$$\pi(\nu|\boldsymbol{d},\boldsymbol{\theta}_{-\nu}) \propto f(\boldsymbol{d}|\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\kappa},\nu) \times \pi(\nu), \tag{18}$$

其中，$\boldsymbol{\theta}_{-\nu} = \boldsymbol{\theta}\backslash\{\nu\} = \{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\kappa},\boldsymbol{\psi},\sigma_\kappa^2,\rho\}$ 是除 ν之外的全部参数。随后迭代应用MCMC更新方案，从式（18）中抽样；这些样本可与其他后验样本结合，形成联合后验分布 $\pi(\boldsymbol{\theta}|\boldsymbol{d})$的样本。

为避免提出负值，我们在对数尺度上执行更新方案，即

$$\log(\nu^*) \sim N(\log(\nu^{(i)}), \sigma_\nu^2),$$

其中，$\nu^{(i)}$ 是当前迭代值，σ_ν^2 是提议方差，其取值应使接受率介于0.15–0.45之间（Roberts 和Rosenthal，2001）。根据我们的试运行结果，设置$\sigma_\nu^2 = 0.005$ 可使接受率约为0.40。

ν^* 的接受率为

$$\alpha(\nu^*|\boldsymbol{\theta}_{-\nu}^{(i)},\nu^{(i)}) = \min\left\{1, \frac{f(\boldsymbol{d}|\boldsymbol{\alpha}^{(i)},\boldsymbol{\beta}^{(i)},\boldsymbol{\kappa}^{(i)},\nu^*) \times \pi(\nu^*)}{f(\boldsymbol{d}|\boldsymbol{\alpha}^{(i)},\boldsymbol{\beta}^{(i)},\boldsymbol{\kappa}^{(i)},\nu^{(i)}) \times \pi(\nu^{(i)})}\right\}. \tag{19}$$

如果使用式（17）中的平坦先验，那么式（19）就简化为似然函数之比。

3.4 初始化与收敛性诊断

对于 $\boldsymbol{\alpha}$、$\boldsymbol{\beta}$、$\boldsymbol{\kappa}$ 和 $\boldsymbol{\gamma}$的初始化，我们使用通过在 R 中迭代拟合泊松–LCC模型（调用glm 函数）得到的最大似然估计。通过对 $\boldsymbol{\kappa}$ 拟合带线性漂移的AR(1)模型，得到 σ_κ^2 和 ρ 的初始值（在 R 中使用$forecast$包中的$arima$ 函数）。通过对 $\boldsymbol{\gamma}$的最大似然估计拟合 ARIMA (1,1,0)模型，初始化队列相关参数σ_γ 和ρ_γ。最后，将 $\boldsymbol{\psi}$ 初始化为（0,0）$^\top$，而离散参数的初始值设为 $\nu = 0.5$。请注意，为加快收敛性，初始化时采用了接近最大似然估计的取值，但这对估计过程的影响应该很小。

我们采用 1，000 次迭代的预热阶段，以减轻初始化的影响。此外，对每个参数应用 50^{th} 后验样本稀疏抽取（每 50 次迭代收集一个实现值），以降低这些序列的自相关。丢弃预热阶段的迭代并进行稀疏抽取后，在 CMP–LC 和 CMP–LCC 模型下，每个参数均获得大小为 10，000 的样本。

在进行任何推断比较之前，可以使用轨迹图和自相关图（例如，参见 Lunn 等，2013，第 3 章）作为诊断工具，以检测马尔可夫链蒙特卡罗生成的后验样本中的异常。参见附录 C，部分随机选取参数的轨迹图显示收敛似乎已经达到，混合良好，且没有明显异常。稀疏抽取后，样本自相关似乎也衰减得相当快，但 κ_t 可能除外，其相关性相对较强。补充材料中介绍了几种常用的收敛诊断方法，例如 Gelman 和 Rubin（1992）以及 Geweke（1991）提出的方法。具体而言，根据 CMP–LC 和 CMP–LCC 模型下的后验样本计算得到的 Gelman统计量，即潜在尺度缩减因子（PSRF），均非常接近 1，表明不存在收敛问题。采用 Geweke（1991）的方法计算的 Z统计量也满足收敛判据。

总之，MCMC生成的后验样本似乎表现良好，因此可用于后续计算，以得出准确的推断。

4 结果

在呈现结果时，我们还会在适当情况下列出采用贝叶斯方法的泊松设定和负二项设定下得到的结果 D_{xt} ，以供比较（分别参见 Czado 等人，2005年和 Wong 等人，2018年）。泊松设定如第 1.2 节所述，负二项分布（NB）设定如下：

$$D_{xt} \sim \text{Neg-Bin}\left(\phi, \frac{\phi}{e_{xt}\mu_{xt} + \phi}\right), \tag{20}$$

其中 ϕ 是过度离散参数。在负二项分布设定下， $\mathbb{E}[D_{xt}] = e_{xt}\mu_{xt}$ 和 $\text{Var}[D_{xt}] = \mathbb{E}[D_{xt}] \times (1 + \mathbb{E}[D_{xt}]/\phi) > \mathbb{E}[D_{xt}]$。本比较旨在首先凸显考虑过度离散的重要性，其次展示两种纳入过度离散的方法之间的差异（即 CMP 分布与负二项分布之间的差异）。

4.1 估计参数

图2和图3展示了 $\boldsymbol{\alpha}$、 $\boldsymbol{\beta}$ 和 $\boldsymbol{\kappa}$ 的拟合值（后验中位数），以及分别在 LC 模型和 LCC 模型下（采用 CMP、泊松和负二项分布设定）相应的 95%可信区间。显然，这些模型的拟合值相当相似，而过度离散模型得出的可信区间略宽。这是考虑过度离散的模型普遍具有的特征：由于模型似然提供了额外的灵活性，响应值（D_{xt}）允许具有更大的变异性，使参数更易波动，因此可信区间也更宽。 $\boldsymbol{\beta}$ 在 CMP–LC 模型中存在一些例外，CMP 设定得出的可信区间明显更宽。负二项–LC 模型中的 $\boldsymbol{\kappa}$ 似乎也比 CMP–LC 模型呈现出更强的线性。

LC 模型和 LCC 模型中其他参数的核后验密度如图4和图5所示。CMP 与泊松设定的后验分布大多相似。CMP 与负二项分布设定的后验分布也相当相似（从视觉上看），但 σ_κ^2、ψ_1 和 ρ 除外。具体而言，负二项分布设定下估计的 σ_κ^2 较小，而估计的 ψ_1 和 ρ 较大。有趣的是，ρ 的后验分布呈单峰，且取值接近 1；这与 Wong 等人（2018）对英格兰和威尔士女性死亡数据的研究结果相反，在该研究中， ρ 呈现出双峰后验分布。

CMP–LC 模型和 CMP–LCC 模型下估计的平均离散参数 ν 分别约为 0.237 和 0.578。回顾公式（10）， $1/\nu$ 衡量 D_{xt} 的方差相对于其均值的相对（乘法）增幅。将后验均值代入该表达式可得 $\frac{1}{\nu} \approx 4.22$ [†] 和 $\frac{1}{\nu} \approx 1.73$，表明 CMP–LCC 所隐含的过度离散程度低于 CMP–LC 模型。此结果与 Wong 等人（2023）的发现一致：纳入队列效应有助于更好地校准数据信号与不确定性，从而避免将未捕捉到的数据信号误判为过度离散。

4.2 模型评估

图 6 和图 7 展示了按年龄和时间划分的残差热图，以说明样本内拟合优度。总体而言，CMP 和负二项分布设定都产生了表现良好的残差，表明误差分布校准良好，因此模型 adequacy 良好。相比之下，泊松设定明显导致残差分布较差，凸显了其不足。比较 CMP 和负二项分布设定，负二项分布似乎能更有效地捕捉 60–80 岁年龄段的变化，但在较年轻年龄段表现欠佳，而 CMP 的表现更好

[†]这意味着 $\text{Var}[D_{xt}]$ 约为 $\mathbb{E}[D_{xt}]$的 4.22 倍。

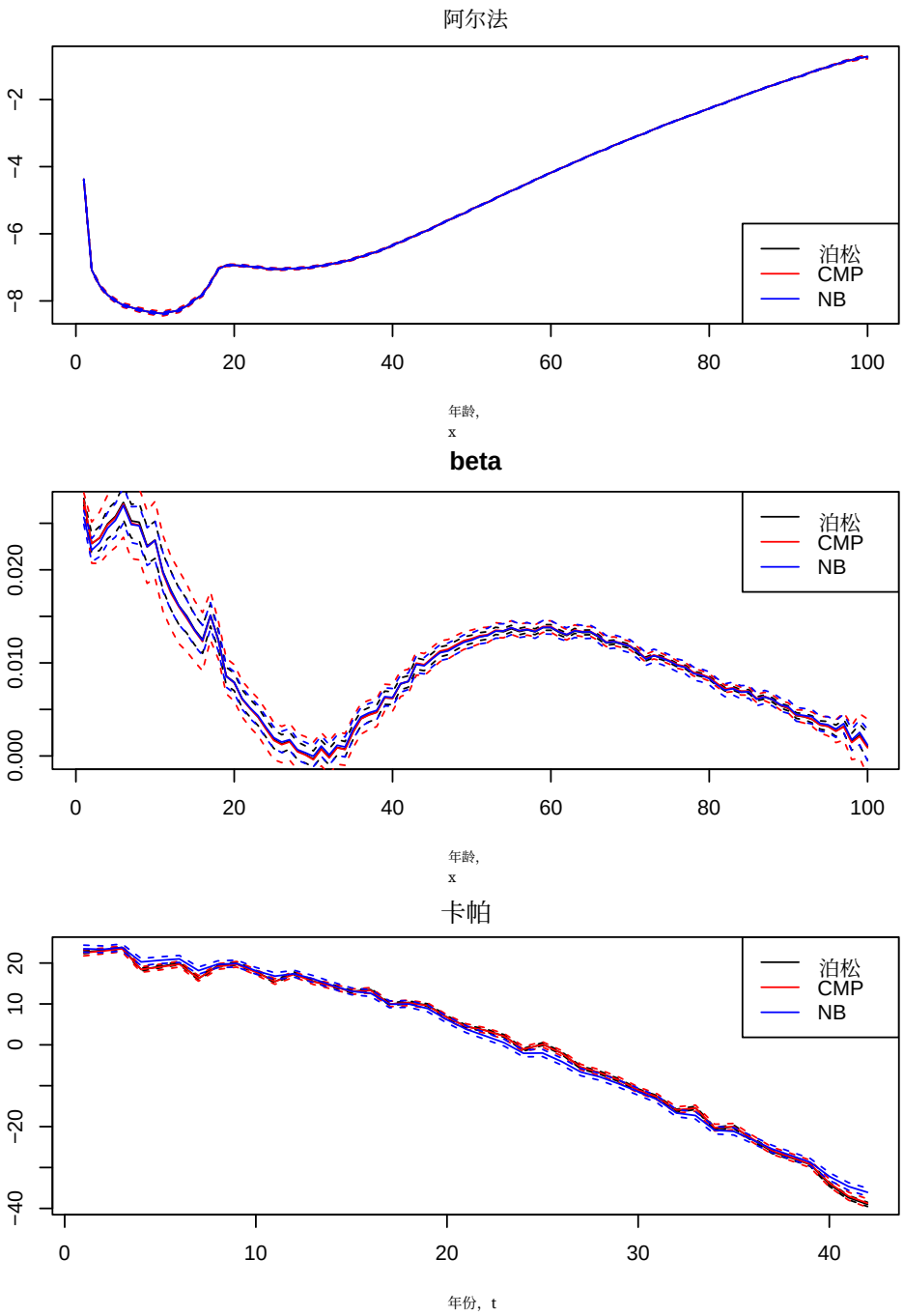


图 2：泊松–LC、CMP–LC 和负二项–LC 模型下估计的模型参数中位数（实线）和 95%可信区间（虚线）图。

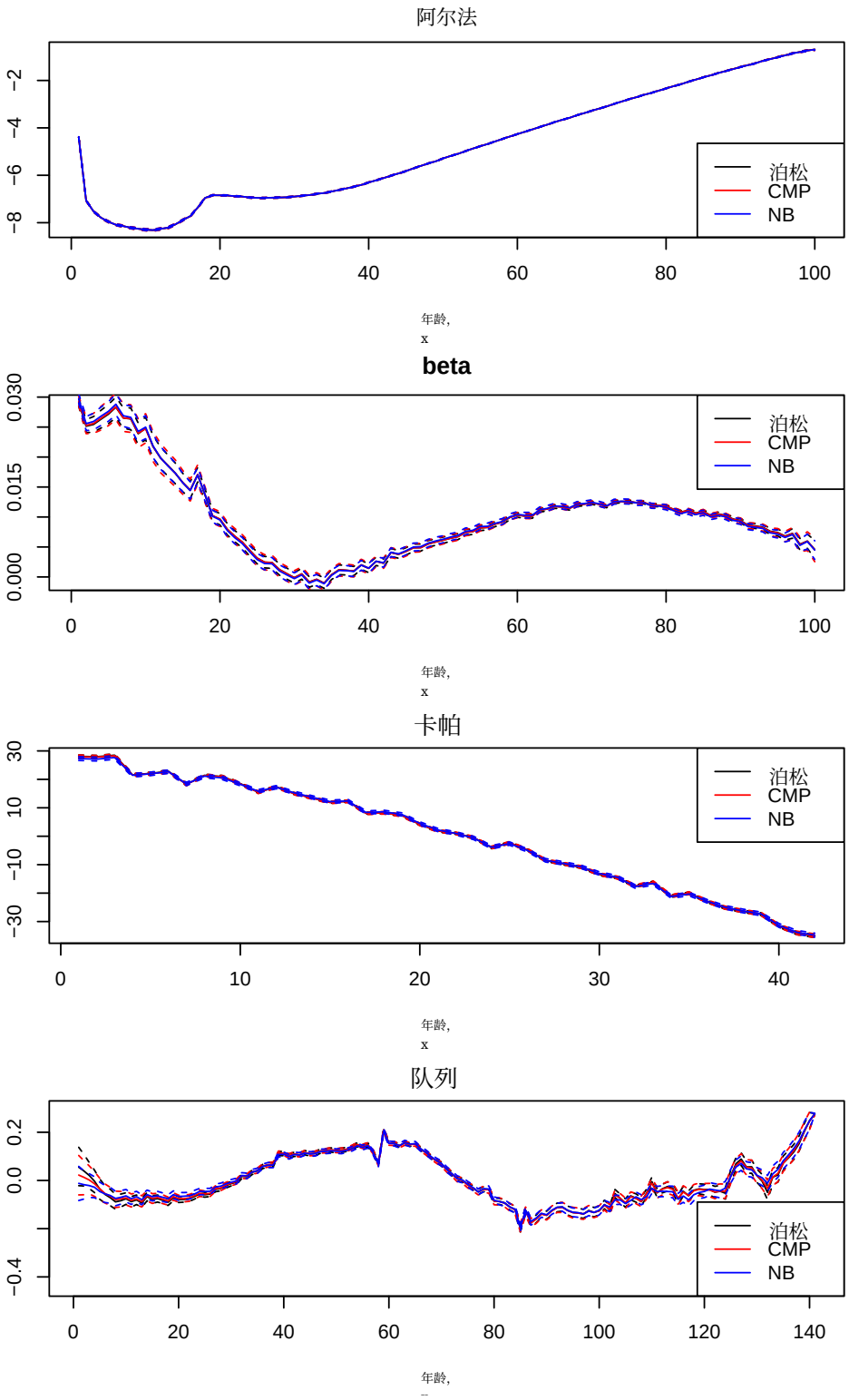


图3：泊松–LCC、CMP–LCC 和负二项–LCC模型下估计的模型参数的中位数（实线）和 95%可信区间（虚线）图。

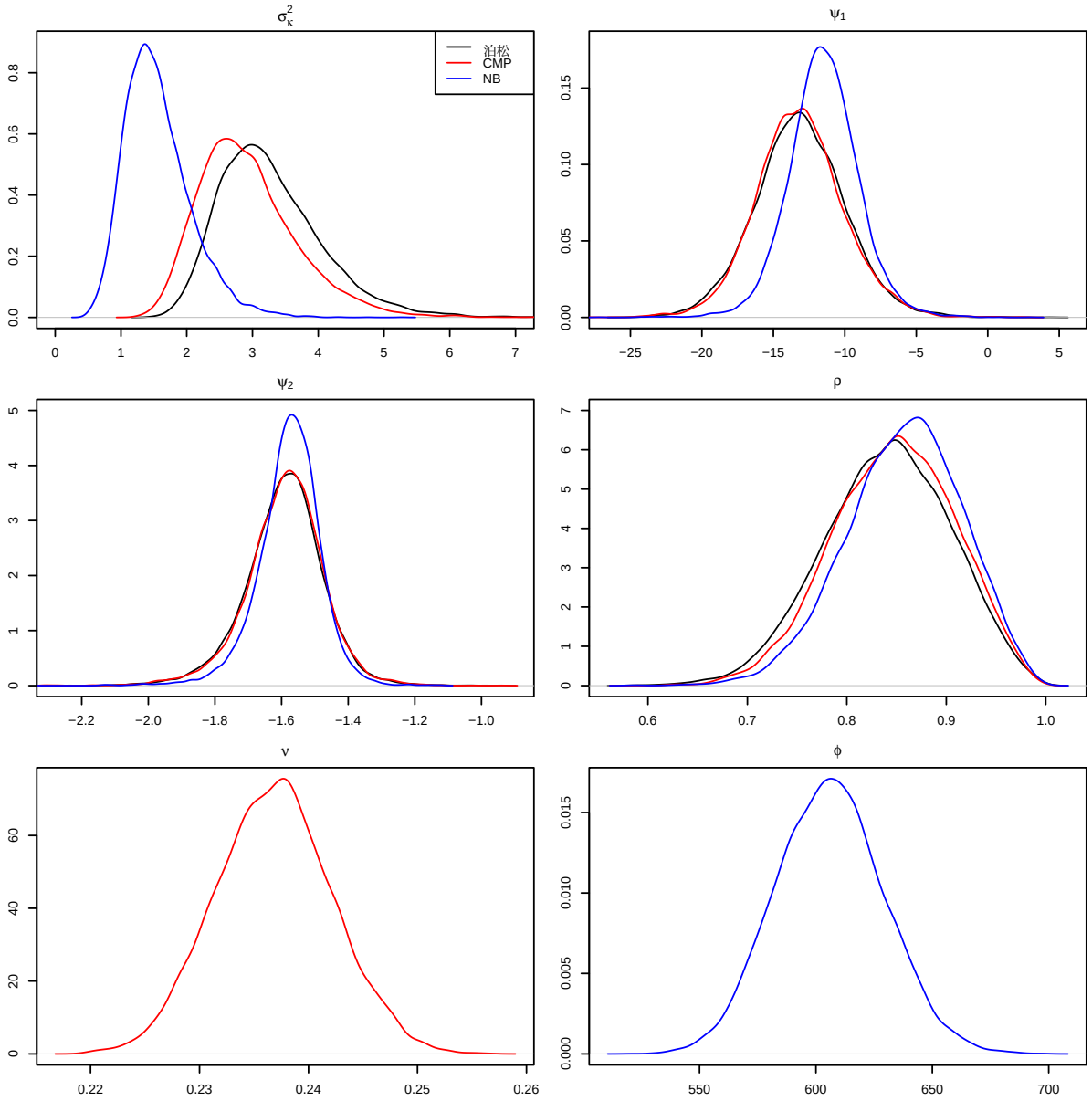


图4：LC 模型估计参数的密度图。

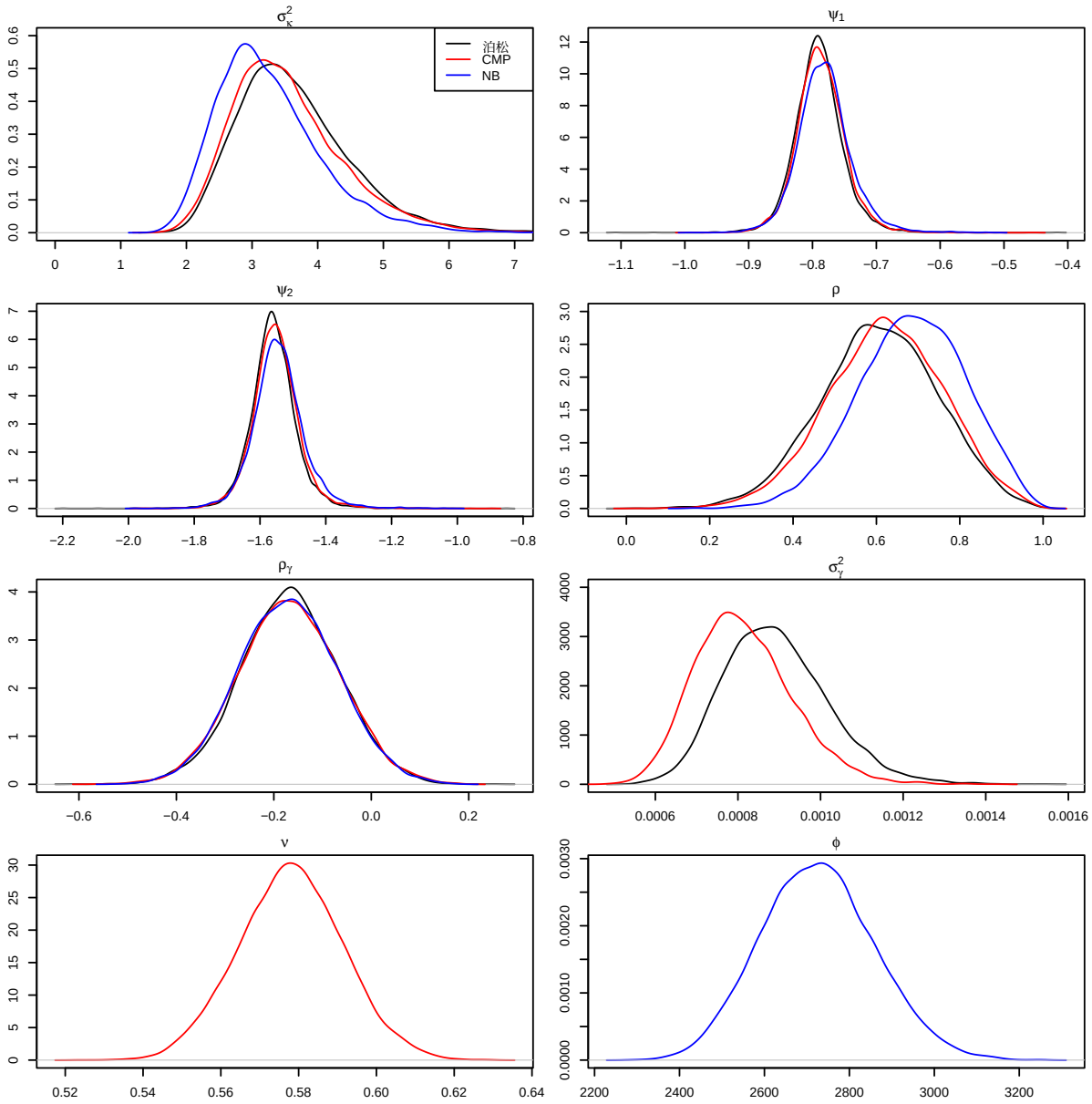


图5：LCC 模型的估计参数密度图。

在较年轻的年龄段。总体而言，残差模式表明 CMP 设定的拟合效果更佳。

表 1 报告了模型残差总量 r^2，其定义见第 1.3 节，用于评估整体拟合优度。显然，CMP 设定的整体偏差更小，优于其他模型。

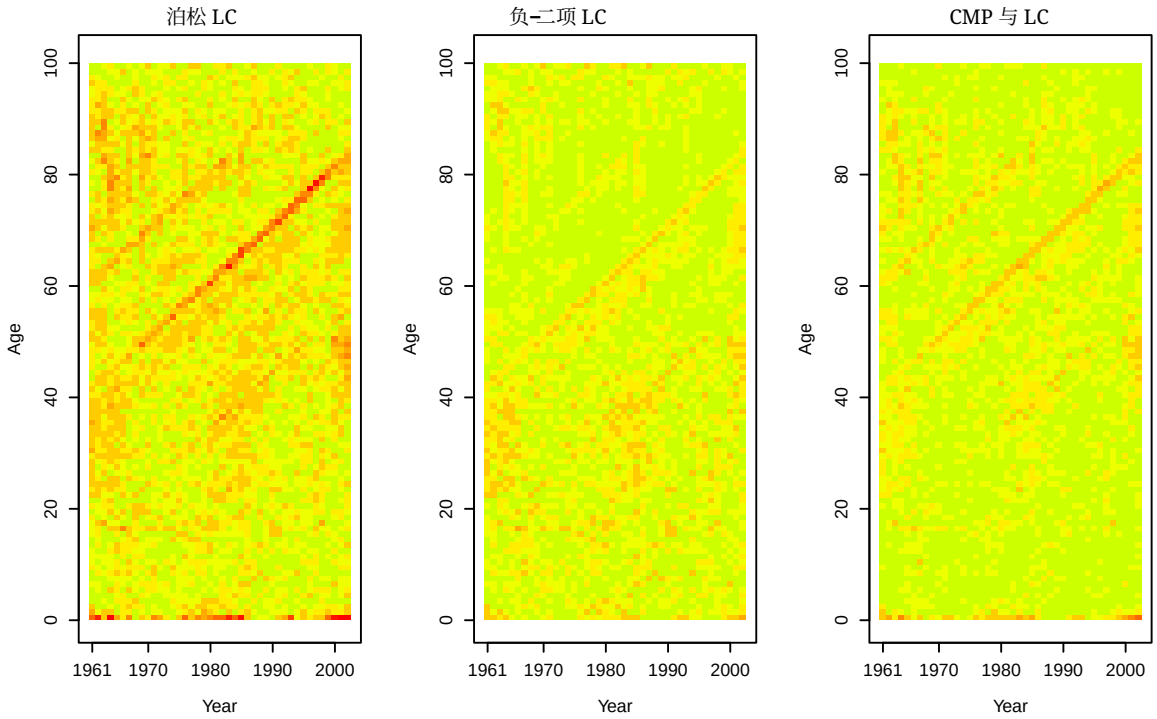


图 6：LC 模型的 r^2_{xt} 热力图。

表 1：各模型设定下估计的模型残差总量 r^2。括号中还列出了拟合不佳的单元格所占百分比，即 $r^2_{xt} > 3.84$。拟合不佳单元格占比低于 5% 的模型表现良好。

模型	泊松	NB	CMP
LC	16709.85	4392.80	**3970.60**
	(26。86%)	(5.62%)	**(4.40%)**
LCC	6628.97	4006.45	**3850.75**
	(11.29%)	(4.71%)	**(4.33%)**

通过评估模型的对数评分（LOGS）、连续排序概率评分（CRPS）和 Dawid-Sebastiani 评分（DSS），比较模型的样本外预测性能，例如参见 Kr¨uger 等人（2021）。这些指标可根据 D_{xt} 的马尔可夫链蒙特卡罗后验样本推导得出（对于 $x = 1, \ldots, A$ 和 $t = 2003, \ldots, 2021$），并使用 R 软件包 $scoringRules$ 计算（Jordan 等人，2019），结果见表 2。结果表明，针对死亡人数的 CMP 设定比相应的其他设定产生了更好的样本外预测。

4.3 拟合粗死亡率与预测粗死亡率

我们还使用粗死亡率（可观测）而非潜在死亡率 μ_{xt}（不可观测）来评估模型的性能。要获得拟合粗死亡率，首先通过方程（8）生成拟合死亡人数 D^F_{xt}，其中联合

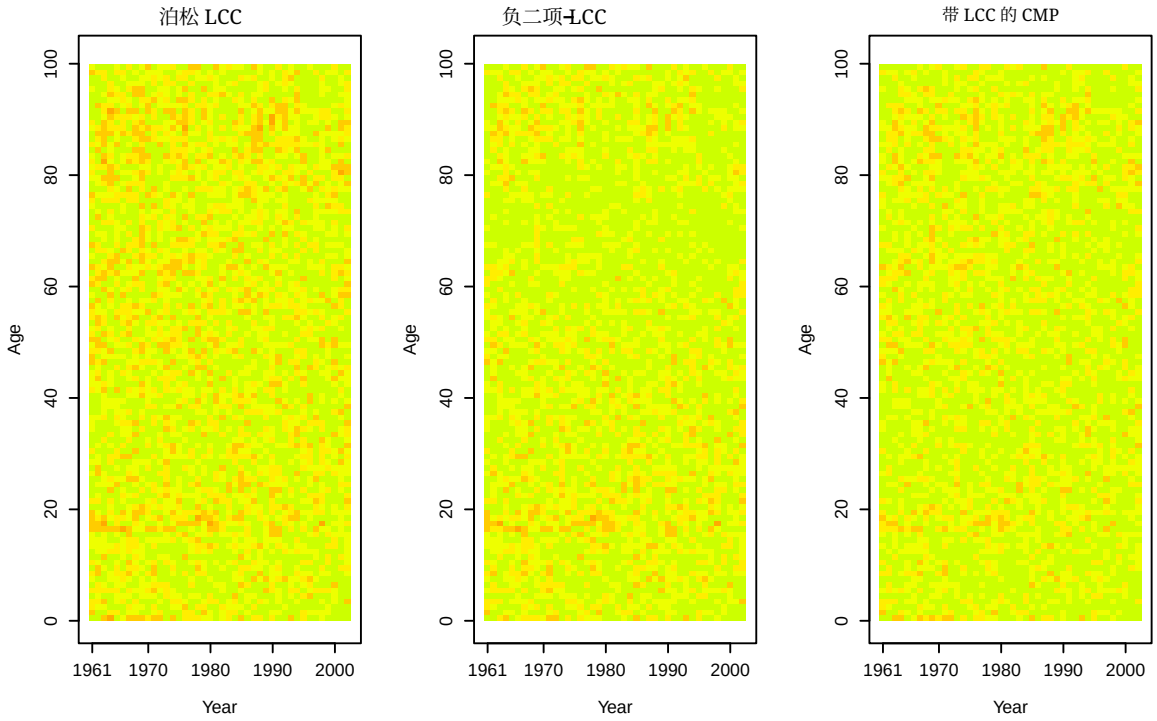


图 7：LCC 模型的 r_{xt}^2 热力图。

表 2：用于评估各模型设定下样本外预测性能的各种评分（越小越好）。

模型	指标	泊松	NB	CMP
LC	LOGS	10.82	9.05	**7.20**
	CRPS	235.62	254.97	**227.60**
	DSS	15.37	13.91	**12.71**
LCC	LOGS	15.57	13.86	**8.95**
	CRPS	153.95	153.32	**151.85**
	DSS	15.34	14.32	**13.29**

相应地代入 α_x、 β_x、 κ_t、 γ_c 和 ν 的后验样本。随后计算每个 $x = 1, \ldots, A$ 和 $t = 1, \ldots, T$ 对应的粗拟合死亡率，结果为 $\hat{}\mu_{xt} = \frac{D^F_{xt}}{e_{xt}}$。

对于死亡率预测，我们注意到，例如在 LCC 模型下（年龄参数保持固定），每个年龄的一年期死亡人数后验预测分布为

$$\begin{aligned} & f(D_{x\,T+1}|\boldsymbol{d}) \\ = & \int f(D_{x\,T+1}|\alpha_x, \beta_x, \kappa_{T+1}, \gamma_{T+1-x}, \nu) \times f(\kappa_{T+1}|\kappa_T, \boldsymbol{\psi}, \rho, \sigma^2_\kappa) \times f(\gamma_{T+1-x}|\gamma_{T-x}, \gamma_{T-x-1}, \rho_\gamma, \sigma^2_\gamma) \\ & \times \pi(\alpha_x, \beta_x, \kappa_T, \boldsymbol{\psi}, \rho, \sigma^2_\kappa, \gamma_{T+1-x}, \rho_\gamma, \sigma^2_\gamma, \nu|\boldsymbol{d}) \mathrm{d}\alpha_x \ldots \mathrm{d}\phi, \end{aligned} \tag{21}$$

其中， $\pi(\alpha_x, \beta_x, \kappa_T, \boldsymbol{\psi}, \rho, \sigma^2_\kappa, \gamma_{T+1-x}, \rho_\gamma, \sigma^2_\gamma, \nu|\boldsymbol{d})$ 是联合后验分布。这表明可采用以下生成程序来预测粗死亡率:

(i) 根据方程（13）生成 κ_{T+1}，其中适当代入 $(\boldsymbol{\psi}, \rho, \kappa_T, \sigma^2_\kappa)$ 的联合后验样本。

(ii) 根据方程（15）生成 γ_T，其中代入 $(\gamma_{T-1}, \gamma_{T-2}, \rho_\gamma, \sigma^2_\gamma)$ 的联合后验样本。其他“观测值”队列分量 $(\gamma_{T-99}, \ldots, \gamma_{T-3})$ 也可从后验样本中获得。

(iii) 使用（8）为每个 x 生成预测死亡人数，即

$$D^F_{x\,T+1} \sim \mathrm{CMP}\left((e_{x\,T+1} \exp(\alpha_x + \beta_x \kappa_{T+1} + \gamma_{T+1-x}) + \frac{1}{2} - \frac{1}{2\nu})^\nu, \nu\right),$$

其中， $e_{x\,T+1}$ 为留出期中心风险暴露量，代入（α_x， β_x， ν）的联合后验样本， κ_{T+1} 和 γ_{T+1-x} 分别来自步骤（i）和（ii）。

(iv) 对每个 x 计算预测粗死亡率，即 $\hat{}\mu_{x\,T+1} = \frac{D^F_{x\,T+1}}{e_{x\,T+1}}$。

依此类推，可通过递归实施上述算法获得 h年期预测。生成 $x = 1$、.....、 A 和 $t = 1$、.....、 $T + h$对应的 $\hat{}\mu_{xt}$ 样本后，可计算样本百分位数，构建中位数和95%区间。

在泊松和负二项分布设定下，我们使用相关参数的后验样本，从 $D^F_{xt} \sim$ Poisson($e_{xt}\mu_{xt}$)和 $D^F_{xt} \sim$ Neg-Bin(ϕ, $\frac{\phi}{e_{xt}\mu_{xt}+\phi}$)中生成样本，然后将拟合粗死亡率计算为 $\hat{}\mu_{xt} = \frac{D^F_{xt}}{e_{xt}}$。

请注意，上述拟合和预测死亡率的生成程序纳入了随机变异（泊松、CMP 和负二项分布）、参数以及预测不确定性所带来的不确定性。这使我们能够在考虑各种不确定性来源的基础上，更好地直观比较各分布设定在 D_{xt} 下的总体表现。

图8展示了$h = 19$时若干选定年龄的拟合和预测粗死亡率，其中也包括观测值和留出期粗死亡率。总体而言，CMP 和负二项分布设定均能产生性能良好的死亡率预测，与留出期粗死亡率十分接近，且 95%可信区间具有较高的覆盖率。相比之下，泊松设定得出的可信区间过窄，凸显了考虑过度离散的重要性。比较 CMP 和负二项分布设定可见，对于较低年龄，CMP 设定产生的 95%可信区间更宽；而在负二项分布情形下则恰好相反，其为较高年龄产生的区间更宽。

纳入队列效应通常能更好地描述拟合率和预测率背后的死亡率趋势。尤其是，队列模型（所有设定）的95%预测区间尽管宽度较窄，仍能以合理的覆盖率涵盖留出率。未纳入队列效应的模型通常无法捕捉（

并因此无法预测）死亡率趋势，导致中位数预测不切实际，且区间宽度不必要地过大。同时纳入过度离散和队列分量，有助于校准信号与误差，从而降低死亡率预测中的偏差，并以足够宽的区间刻画不确定性。这与 Wong 等人（2023）的发现一致。比较各种设定可见，泊松设定因区间过窄而覆盖率偏低，而 CMP 和负二项分布设定则能产生足够宽且覆盖率较高的区间。请注意，由于新冠肺炎的影响，观测到最近两个死亡率异常偏高；本文不对此进行讨论。

4.4 敏感性分析

最后，我们考察先验设定对 ν 估计的影响，并评估结果对超参数 a和 b的敏感性。为便于说明，本分析仅针对 CMP–LC模型进行。图9展示了在第 3.1 节所述的不同伽马先验设定下， ν 的后验分布。

首先可以看到，在采用 Gamma(1，0。01)、Gamma(1，0。1)、Gamma(1，1) 和 Gamma(1，10) 先验时， ν 的后验分布与采用平坦先验时的后验分布十分相似。这并不出人意料，因为这些先验的方差远大于 ν的后验方差（约为 3×10^{-5}），因此对后验估计的影响极小。在这些情况下， ν 的估计主要由数据决定。

例外情况是指定先验 $\nu \sim$ Gamma(1，100)。该先验的均值为 0。01，方差为 1×10^{-4}，与数据所隐含的方差处于相近量级。因此，该先验包含足够的信息，能够影响贝叶斯学习过程，使 ν 的后验分布略微向先验均值 0。01 偏移。这反映出，相较于其他设定，该先验被赋予了更高的确定性，因而对贝叶斯估计产生了更显著的影响。

总之，随着伽马先验的方差减小（即 b 的值增大），先验在 ν估计中的作用更加显著。

5 结论

本文介绍了 CMP分布，将其作为死亡人数建模的灵活框架。该分布被嵌入 LC 和 LCC死亡率模型中，并采用贝叶斯方法进行估计。所提出的框架突出了 CMP分布通过显式离散参数 ν 处理欠离散、等离散和过度离散的能力，其中 ν 的值可以直接从数据中推断得到。通过这种方式，该方法能够一致地纳入所有相关的不确定性来源，尤其适用于存在过度离散的情形。为便于进行贝叶斯推断，我们针对离散参数采用非信息伽马先验，并开发了相应的 MCMC算法以生成后验样本。

所提出的方法使用英格兰和威尔士整个年龄范围内的男性死亡率数据进行了说明。实证结果表明，与泊松和负二项模型等常用的替代分布相比，基于 CMP 的设定在样本内和样本外表现更优；这一结论是通过残差诊断和一系列适当评分规则评估得出的。这些发现凸显了将灵活的分布假设与适当的结构组成部分相结合在死亡率建模中的重要性。特别是，在纳入队列效应的同时采用能够处理离散程度的更灵活分布，可以改善系统性死亡率趋势与随机变异之间的校准。此外，还进行了敏感性分析，以评估先验选择对贝叶斯估计的影响；结果表明，在所提出的设定下，模型表现似乎具有稳健性。

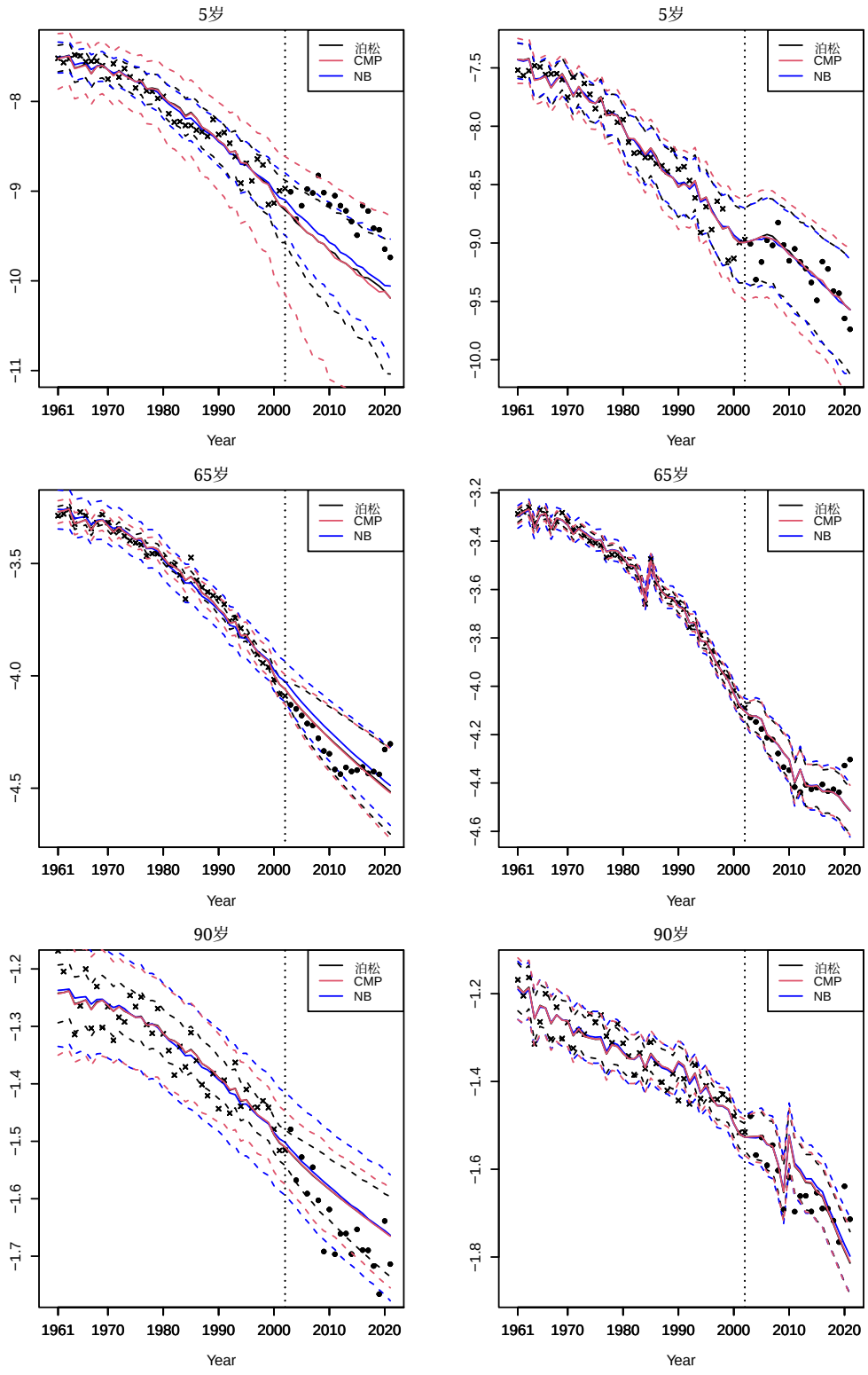


图8：在 LC（左侧面板）和 LCC（右侧面板）率模型下，若干选定年龄的观测值和留出对数粗死亡率、拟合对数粗死亡率，以及粗对数死亡率未来19年预测的图示，并以虚线标示相应的95%可信区间。

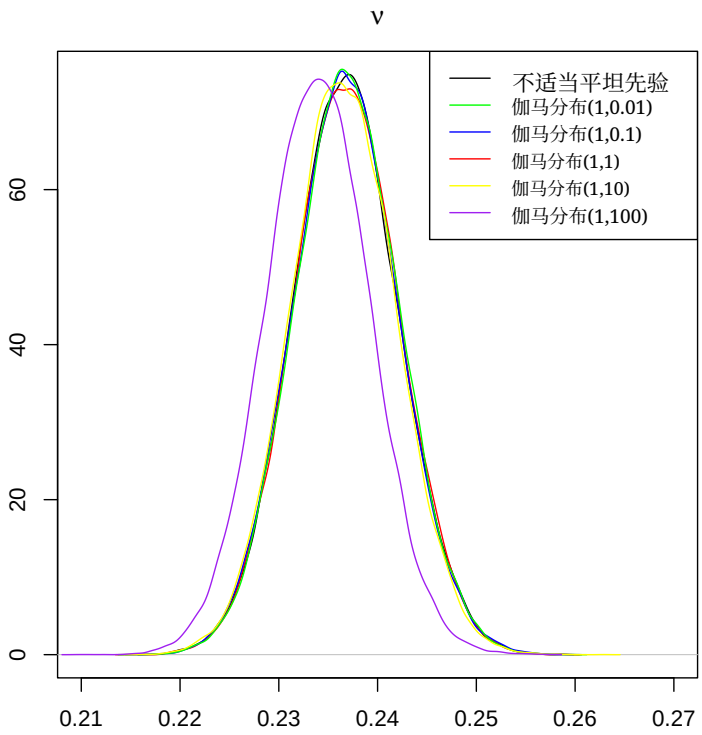


图9：不同先验设定下，CMP–LC模型的 ν 后验分布核密度图。

我们的研究还提出了若干值得未来探讨的方向。首先，时期效应和队列效应（ κ_t 和 γ_c）的时间序列动态采用了文献中的标准形式。将这一框架扩展为可在不同时间序列设定之间进行正式模型选择，或许能进一步提高模型稳定性，尤其是在应用于其他死亡率数据集时；不过，对于本文所考察的数据，改进幅度可能有限。其次，我们假设离散参数 ν 在不同年龄和时期之间保持恒定，因此仅捕捉总体离散程度。这一假设可能过于严格，因为离散程度可能会在死亡率曲面上呈现系统性变化。允许离散程度随年龄或时间变化，例如采用 ν_x 或 ν_t，是一个颇有前景的未来研究方向。

总之，我们的研究表明，将CMP分布融入成熟的死亡率模型，为分析死亡人数提供了一种强大而灵活的方法。通过在统一的贝叶斯框架中明确对离散程度进行建模，所提出的方法改善了模型拟合，提供了更丰富的推断结果，并为死亡率建模与预测领域进一步开展方法和应用研究奠定了坚实基础。

致谢

作者衷心感谢编辑、副编辑以及所有参与审稿的审稿人提出的意见，这些意见有助于提升论文质量。

基金资助

本研究未获得公共、商业或非营利部门任何基金资助机构的专项资助。

利益冲突

作者声明，他们在本文发表方面不存在利益冲突。作者未披露任何可能影响本研究的财务或个人关系。

数据可用性声明

支持本研究结果的数据和代码可在以下 GitHub 仓库中获取 [录用后添加]。

附录：补充材料

与本研究相关的补充材料可在出版商网站上的文章在线版本中找到。

参考文献

Basellini, U., C. G. Camarda, and H. Booth (2023). 三十年后：李-卡特方法在死亡率预测中的回顾。《国际预测学杂志》*39*(3), 1033–1049. Booth, H. and L. Tickle (2008). 死亡率建模与预测：方法综述。《精算科学年刊》*3*(1–2), 3–43. Breslow, N. E. (1984). 对数线性模型中的超泊松变异。《英国皇家统计学会杂志》*C*辑（应用统计学）*33*(1), 38–44. Brillinger, D. R. (1986). 一篇附讨论的《生物统计学》特邀论文：生命率及相关统计量的自然变异性。生物统计学 *42*(4), 693–734. Brouhns, N., M. Denuit, and J. K. Vermunt (2002). 采用泊松对数双线性回归构建预测生命表。《保险：数学与经济学》**31**(3), 373–393. Brown, J. R. (2003). 再分配与保险：具有死亡率异质性的强制年金化。《风险与保险杂志》**70**(1), 17–41. Cairns, A. J., D. Blake, K. Dowd, G. Coughlan, and M. Khalaf-Allah (2011). 针对两个群体的贝叶斯随机死亡率建模。*ASTIN Bulletin*：*IAA*期刊 *41*(1), 29–59. Chanialidis, C., L. Evers, T. Neocleous, and A. Nobile (2018). COM-泊松回归模型的高效贝叶斯推断。《统计学与计算》*28*, 595–608. Conway, R. and W. Maxwell (1962). 一种服务率依赖于状态的排队模型。《工业工程杂志》*12*(2), 132–136. Czado, C., A. Delwarde, and M. Denuit (2005). 贝叶斯泊松对数双线性死亡率预测。《保险：数学与经济学》**36**(3), 260–284. Daly, F. and R. E. Gaunt (2016). 康威–麦克斯韦–泊松分布：分布理论与近似。Dean, C. and J. F. L. and (1989). 泊松回归模型中过度离散的检验。《美国统计学会杂志》*84*(406), 467–472.

Dean, C. B. (1992). 泊松回归模型和二项回归模型中过度离散的检验。《美国统计学会杂志》*87*(418), 451–457. Diana, A., J. S. T. Wong, and A. Pittea (2025). 用于死亡率模型选择的统一贝叶斯框架。《精算科学年刊》*19*(3), 462–481. Gelfand, A. E. and S. K. Sahu (1999). 广义线性模型的可识别性、非正规先验与吉布斯抽样。《美国统计学会杂志》*94*(445), 247–253. Gelman, A. and D. B. Rubin (1992). 使用多条序列从迭代模拟中进行推断。统计科学**7**(4), 457–511. Geweke, J. (1991). 评估基于抽样的方法在计算后验矩时的准确性。技术报告，明尼阿波利斯联邦储备银行。Goes, J., K. Barigou, and A. Leucht (2025). 存在大流行病时的贝叶斯死亡率建模：消失跳跃方法。《英国皇家统计学会杂志》*C*辑：应用统计学 *74*(4), 1150–1182. Hayat, M. J. and M. Higgins (2014). 理解泊松回归。护理教育杂志 *53*(4), 207–215. Hinde, J. and C. G. Dem´etrio (1998). 过度离散：模型与估计。计算统计与数据分析 *27*(2), 151–170. Hougaard, P., M.-L. T. Lee, and G. A. Whitmore (1997). 利用泊松变量混合与泊松过程分析过度离散计数数据。生物统计学 *53*(4), 1225–1238. Hunt, A. and A. M. Villegas (2015). 存在队列效应时李-卡特模型的稳健性与收敛性。《保险：数学与经济学》*64*, 186–202. Jordan, A., F. Kr¨uger, and S. Lerch (2019). 使用scoringrules评估概率预测。《统计软件杂志》*90*(12), 1—37. Kr¨uger, F., S. Lerch, T. Thorarinsdottir, and T. Gneiting (2021). 基于马尔可夫链蒙特卡罗输出的预测推断。国际统计评论 *89*(2), 274–301. Lee, R. D. and L. R. Carter (1992). 美国死亡率的建模与预测。《美国统计学会杂志》**87**(419), 659–671. Lunn, D., C. Jackson, N. Best, A. Thomas, and D. Spiegelhalter (2013). 《*BUGS*手册：贝叶斯分析实用入门》。Chapman and Hall/CRC。McCullagh, P. (2019). 广义线性模型。Routledge。O’Hagan, A. and J. Forster (2004). *Kendall*《高级统计学理论》（第2版），第2B卷。Kendall统计学丛书。Renshaw, A. E. and H. Haberman (2005). 李-卡特死亡率降低因子的基于队列的扩展。《保险：数学与经济学》**38**(3), 556–570. Roberts, G. O. and J. S. Rosenthal (2001). 各类Metropolis-Hastings算法的最优缩放。统计科学**16**(4), 351–367. Sellers, K. F., T. Lotze, and A. Raim (2023). *COMPoissonReg*：康威–麦克斯韦–泊松（*COM*-泊松）回归。CRAN。R 软件包版本0.8.1。Selvin, S. (2004). 流行病学数据的统计分析，第35卷。牛津大学出版社。

Shi, J., Y. Shi, P. Wang, 和 D. Zhu（2024）。多人口死亡率建模：贝叶斯分层方法。*ASTIN Bulletin*：*IAA*期刊 54(1), 46–74。Tawiah, K., S. Iddi, 和 A. Lotsi（2020）。零膨胀分层泊松模型及其在孕产妇死亡率数据中的应用。《国际数学与数学科学杂志》*2020*(1), 1407320。Villegas, A. M., V. K. Kaishev, 和 P. Millossovich（2018）。StMoMo：用于随机死亡率建模的R软件包。《统计软件杂志》**84**(3), 1–38。Wong, J. S. T., J. J. Forster, 和 P. W. F. Smith（2018）。考虑过度离散的贝叶斯死亡率预测。《保险：数学与经济学》**83**(2018), 206–221。Wong, J. S. T., J. J. Forster, 和 P. W. F. Smith（2023）。死亡率预测的贝叶斯模型比较。《皇家统计学会期刊*C*辑：应用统计》*72*(3), 566–586。Zhang, X. 和 R. M. and（2023）。使用分段泊松模型与ARIMA模型的新型组合预测波动的每日新冠肺炎死亡人数序列。《应用统计学杂志》*50*(11-12), 2561–2574。