

【统计理论与方法】

DOI:10.20207/j.cnki.1007-3116.20250703.001

泊松分布下基于鞍点逼近的相对风险置信区间构造

侯 健^{a,b},刘 硕^{a,b},田茂再^{a,b}

(中国人民大学 a. 应用统计科学研究中心; b. 统计学院,北京 100872)

摘要: 相对风险是流行学随机对照试验中的一种常用指标,用于比较两个群体的发病率或死亡率。对其进行区间估计是一种有效的统计推断方法,可以反映疾病诊断的不确定性和可信度。然而,传统的区间估计方法,如 Wald 区间、Wilson 区间、Fisher 区间等,都存在一定的局限性:误差大、覆盖率低、区间长度长、区间不合理等。为了克服这些缺点,提出了基于鞍点逼近的区间估计方法:在泊松分布的假设下,利用泰勒展开和拉普拉斯积分等推导了相对风险的逼近公式,并给出了相应置信区间的构造方法。通过蒙特卡洛模拟和实证分析,将鞍点逼近方法与部分常用的区间估计方法进行了比较,结果表明,鞍点逼近方法在小样本或大样本情况下都能获得最短的置信区间,同时区间覆盖率更加接近名义水平,为相关风险提供了一种优越的区间估计方法。最后对鞍点逼近方法的优点和局限性,以及在流行病学研究中的应用价值和前景进行了讨论。

关键词: 泊松分布; 鞍点逼近; 区间估计; 相对风险; 随机对照试验

中图分类号:O212 文献标识码:A 文章编号:1007-3116(2025)08-0003-11

引用格式:侯健,刘硕,田茂再:泊松分布下基于鞍点逼近的相对风险置信区间构造[J].统计与信息论坛,2025,40(8):3-13.

Citation Form: HOU Jian, LIU Shuo, TIAN Maozai. Relative risk confidence interval construction based on saddle point approximation under Poisson distribution[J]. Journal of statistics and information, 2025, 40(8): 3-13.

一、引言

随着现代人生活习惯和环境的不断变化,细胞突变现象变得愈发普遍,不受控制的细胞分裂可能在体内形成恶性肿瘤。在恶性肿瘤研究中,细胞突变被视为稀有事件,因为其不可预知性和聚集性,泊松(Poisson)分布被广泛用于描述这类小概率事件的发生规律。假设某疾病在特定周期内的平均发病水平为 λ ,则该疾病在某一周期内发生 x 例病例的概率可以通过这一平均水平计算。在流行病学研究中,随机对照试验的重要性日益增加,各种指标的应用推动了医学研究中的统计推断和流行趋势分析。常用的随机对照试验指标包括相对风险(Relative Risk, RR)、风险差(Risk Difference, RD)、比值比(Odds Ratio, OR)等。这些指标在评估疾病发生的概率和风险因素的影响方面起着至关重要的作用^[1]。

相对风险(RR)是流行病学中的重要指标,其定义为暴露组中事件的发生率 λ_E 与对照组中事件的发生率 λ_C 之比^[2],公式为:

$$RR = \frac{\lambda_E}{\lambda_C} = \frac{\frac{a}{t_E}}{\frac{b}{t_C}} = \frac{a \cdot t_C}{b \cdot t_E}$$

收稿日期:2024-05-30

基金项目:国家自然科学基金项目“基于分位回归的当代统计学逆问题重大基础理论和方法及其应用”(11861042);中国
人民大学科学研究基金(中央高校基本科研业务费专项资金资助)项目成果“当代复杂大数据的动态稳健建模
及应用研究”(22XNL016)

作者简介:侯 健,男,上海人,博士生,研究方向:复杂数据分析与建模;

刘 硕,男,北京人,博士生,研究方向:张量数据分析与建模;

田茂再(通讯作者),男,湖南凤凰人,博士,教授,博士生导师,研究方向:大数据建模与分位回归。

其中, t_E 与 t_C 分别表示暴露组与对照组的观察时间。当 $RR=1$ 时, 暴露对结果没有影响; $RR>1$ 表示暴露增加结果发生的概率, 属于“风险因子”; $RR<1$ 表示暴露降低结果发生的概率, 为“保护因子”。然而, RR 只反映相关性而不代表因果关系。例如, 在医院和家中患癌症的相对风险可能显著大于 1, 但这是因为患癌症的人更可能去医院。因此, 相对风险常用于衡量暴露与结果之间的相关性, 通过区间估计可以为评估疾病风险提供有利的信息和辅助。另外, 更加精确和更短的区间长度能够降低误判度, 使诊断结果更加准确^[3]。

目前, 已有研究针对流行病学指标的置信区间构造进行了较为广泛的探索。例如, Tang 等采用负二项抽样方法构造了风险差的得分检验型置信区间^[4]; 夏丽丽等在二项分布下构造总体成数的各类置信区间^[5]; 孟令宾等研究了优势比的渐进置信区间^[6]; 白永昕等采用 Poisson 抽样方法, 通过枢轴量方法构造置信区间, 结果显示其区间长度与覆盖率表现优良^[7]。王维贤等使用鞍点逼近方法对相关差指标进行 Poisson 抽样构造置信区间, 证明了鞍点逼近在小样本情况下的拟合性较好^[8]。Cannon-Albright 等利用包含 19 世纪犹他州先驱者家谱的资源, 并与犹他州死亡证数据相关联, 评估了基于特定家族史的阿尔茨海默症相对风险的 Wald 型置信区间, 结果表明任何一代患病的亲属都会显著增加后代患病风险^[9]。Byerley 等探讨了疫苗接种的相对风险与感染 COVID-19 的风险之间的对比。通过提供科学证据和数据, 结果表明疫苗接种带来的风险与未接种而感染 COVID-19 的风险相比是相对较小的^[10]。

然而, 针对泊松分布下相对风险指标的置信区间构造问题研究甚少。泊松分布在处理稀有事件和小概率事件中具有独特的优势, 尤其是在医学研究和公共卫生领域, 能够有效描述某些疾病的发病率情况。因此, 深入研究泊松分布下相对风险的置信区间构造具有重要的理论和实际意义。本文在构造传统置信区间的基础上, 引入鞍点逼近方法, 通过蒙特卡洛模拟和实证研究, 证明了鞍点逼近方法在相对风险置信区间中的适用性。研究结果显示, 无论在小样本还是大样本情况下, 鞍点逼近置信区间的精度更高, 覆盖率更接近名义水平。此外, 本文还探讨了不同样本规模和参数设定下鞍点逼近方法的性能表现, 为未来的研究提供了新的思路和方法。

综上所述, 本文不仅在理论上丰富了泊松分布下相对风险置信区间的构造方法, 而且在实际应用中也为医学研究和公共卫生领域提供了更为可靠的统计工具。通过鞍点逼近方法, 本文为小样本条件下的统计分析提供了一种新的解决方案, 具有重要的现实意义和应用价值。

二、相对风险置信区间的构造

在一次对照实验中设定暴露组为 $i=1$, 对照组为 $i=0$, 对第 i 组中的 N_i 个研究对象展开跟踪研究并得到 Y_i 例发病数。此时 Y_i 服从参数为 $\lambda_i N_i^*$ 的泊松分布, 其中 N_i^* 表示单位时间内研究对象数目的平均值, λ_0 和 λ_1 分别表示暴露组和对照组的发病率且大于零, 则随机变量的概率质量函数为:

$$P(Y_i=y_i|\lambda_i,N_i)=\frac{(\lambda_i N_i^*)^{y_i}}{y_i!}\exp(-\lambda_i N_i^*), y_i=0,1,\dots \quad (1)$$

在泊松分布下, 相对风险是发病率的比值, 由于通常无法知晓发病率的真实值, 因此所得到的置信区间往往是近似的而不是精确的^[11]。若采用其极大似然估计值, 此时似然函数是:

$$L(\lambda_0,\lambda_1;y_0,y_1)=\frac{(\lambda_0 N_0^*)^{y_0}}{y_0!}\exp(-\lambda_0 N_0^*)\frac{(\lambda_1 N_1^*)^{y_1}}{y_1!}\exp(-\lambda_1 N_1^*) \quad (2)$$

则发病率对应的极大似然估计分别为 $\hat{\lambda}_0=\frac{y_0}{N_0^*}$ 和 $\hat{\lambda}_1=\frac{y_1}{N_1^*}$, 那么 $\widehat{RR}=\hat{\lambda}_0/\hat{\lambda}_1$ 简记为 \hat{R} 。下面将给出关于 \hat{R} 置信区间构造的几种常用方法。

(一) 反转假设检验统计量

考察假设检验中所构造统计量的显著性时可以很自然与其区间估计联系在一起, 所以我们可以在给定显著性水平 α 的条件下, 通过反转统计量来获得一个 $100(1-\alpha)\%$ 的置信区间。考虑假设 $H_0: R=\hat{R}$ 对 $H_1: R\neq\hat{R}$, 此时有如下检验:

(1) Wald 检验

在大样本条件下, 反转 Wald 检验与正态近似法所得到的置信区间等价, 由中心极限定理可知, 其统计

量是^[12-13]:

$$T_{\text{Wald}} = \frac{\hat{R} - R}{\sqrt{\text{Var}(\hat{R})}} \sim N(0, 1) \quad (3)$$

反转统计量 T_{Wald} 后得到 R 的置信区间为:

$$(\max(\hat{R} - Z_{\alpha/2} \sqrt{\text{Var}(\hat{R})}, 0), \hat{R} + Z_{\alpha/2} \sqrt{\text{Var}(\hat{R})}) \quad (4)$$

其中, $Z_{\alpha/2}$ 是标准正态分布的 $\alpha/2$ 分位点。实际上 $\text{Var}(\hat{R})$ 是难以计算的, 但可以通过 Δ 方法来计算其近似替代值 $\widehat{\text{Var}}(\hat{R})$ 。 Δ 方法是统计应用性非常广的一种方法, 利用了数学方法中的泰勒展开式原理, 取随机变量函数的一阶泰勒展开式作为其逼近。在已知单个随机变量的分布下, 可得到随机变量函数的数字特征如均值或方差等。根据 Papanicolaou 的研究有 Δ 定理^[14]:

定理 1: 设随机变量序列 X_n 满足 $\sqrt{n}(X_n - \mu) \xrightarrow{d} N(0, \sigma^2)$, 存在光滑函数 g 在给定的 μ 处满足: $g'(\mu)$ 存在且不为 0, 则有:

$$\frac{\sqrt{n}(g(X_n) - g(\mu))}{|g'(\mu)|\sigma} \approx N(0, 1) \quad (5)$$

根据上述定理有如下结论:

$$\begin{aligned} \text{Var}(\hat{R}) &= \text{Var}\left(\frac{\hat{\lambda}_0}{\hat{\lambda}_1}\right) \approx \frac{1}{\hat{\lambda}_1^2} \text{Var}(\lambda_0) + \frac{\mu_{\lambda_0}^2}{\hat{\lambda}_1^4} \text{Var}(\lambda_1) - 2 \frac{\mu_{\lambda_0}}{\hat{\lambda}_1^3} \text{Cov}(\lambda_1, \lambda_2) \\ &= \left(\frac{E\hat{\lambda}_0}{E\hat{\lambda}_1}\right)^2 \left[\frac{\text{Var}(\hat{\lambda}_0)}{E\hat{\lambda}_0^2} + \frac{\text{Var}(\hat{\lambda}_1)}{E\hat{\lambda}_1^2} - \frac{\text{Cov}(\hat{\lambda}_0, \hat{\lambda}_1)}{E\hat{\lambda}_0 E\hat{\lambda}_1} \right] \end{aligned} \quad (6)$$

由于暴露组与对照组间的实验结果需保持独立, 所以计算上式即可得到 $\widehat{\text{Var}}(\hat{R})$ 。

(2) 对数化 Wald 检验

Katz 等认为在样本量相对较小时, 直接使用正态近似法可能会存在问题^[13]。而对数化的随机变量 $\log(\hat{R})$ 才会更加接近正态分布, 此时的 Wald 检验统计量为^[15]:

$$\frac{\log(\hat{R}) - \log(R)}{\sqrt{\text{Var}(\log(\hat{R}))}} \sim N(0, 1) \quad (7)$$

同理使用 Delta 方法对计算 $\log(\hat{R})$ 的方差, 指数化反转的统计量可以得到 R 的置信区间为:

$$(\max(\hat{R} \exp\{-Z_{\alpha/2} \sqrt{\text{Var}(\log(\hat{R}))}\}, 0), \hat{R} \exp\{Z_{\alpha/2} \sqrt{\text{Var}(\log(\hat{R}))}\}) \quad (8)$$

(3) 似然比检验

假设参数空间为 Θ , 原假设下 R 是来自参数空间某个子集 Θ_0 的参数, 备择假设则是来自其补集 Θ_0^c 。大样本下原假设成立时似然比统计量为^[16]:

$$\Lambda(Y_0, Y_1) = -2 \ln \left[\frac{\sup_{(\lambda_0, \lambda_1) \in \Theta_0} L(\lambda_0, \lambda_1; Y_0, Y_1)}{\sup_{(\lambda_0, \lambda_1) \in \Theta} L(\lambda_0, \lambda_1; Y_0, Y_1)} \right] \rightarrow \chi^2(1) \quad (9)$$

因此, 反转似然比统计量可以得到 R 的渐近置信区间为:

$$\{R : \Lambda(Y_0, Y_1) \leq \chi^2_{1-\alpha}(1)\} \quad (10)$$

(4) Fieller 检验

Fieller 定理是一种统计方法, 用于构建两个正态随机变量比值的置信区间。它由英国统计学家 Fieller 提出, 特别适用于在估计比率时的情况^[17]。令 $R' = \hat{\lambda}_0 - R\hat{\lambda}_1$ 后有:

$$\begin{aligned} E(R') &= E(\hat{\lambda}_0 - R\hat{\lambda}_1) = \lambda_0 - R\lambda_1 = 0 \\ \text{Var}(R') &= \text{Var}(\hat{\lambda}_0 - R\hat{\lambda}_1) = \text{Var}(\hat{\lambda}_0) + R^2 \text{Var}(\hat{\lambda}_1) \end{aligned} \quad (11)$$

则在大样本下有如下结果:

$$\frac{R' - E(R')}{\sqrt{\text{Var}(R')}} \sim N(0, 1) \quad (12)$$

反转该统计量后可以得到置信区间为:

$$(\max\{(B - \sqrt{B^2 - AC})/A, 0\}, (B + \sqrt{B^2 - AC})/A) \quad (13)$$

其中 $A = \hat{\lambda}_1^2 - Z_{\alpha/2}^2 \hat{\lambda}_1 (1 - \hat{\lambda}_1) N_1$, $B = \hat{\lambda}_0 \hat{\lambda}_1$, $C = \hat{\lambda}_0^2 - Z_{\alpha/2}^2 \hat{\lambda}_0 (1 - \hat{\lambda}_0) / N_0$ 。

(5) 精确条件检验

相对风险可以表示为多参数指数分布族的形式：

$$P(Y_0, Y_1; \lambda_0, \lambda_1) = e^{-\lambda_0 + \lambda_1} \exp(\theta u + rt) (Y_0 ! Y_1 !)^{-1} \quad (14)$$

其中, $\theta = \ln(\lambda_1 / \lambda_0)$, $r = \ln(\lambda_0)$, $u = Y_1$, $t = Y_0 + Y_1$, 则双边假设检验问题的水平为 α 的 UMPUT 为:

$$\begin{cases} 0, & c_1(t) < u < c_2(t) \\ r_i(t), & u = c_i(t), i = 1, 2 \\ 1, & u < c_1(t) \text{ or } u > c_2(t) \end{cases} \quad (15)$$

其中, $r_i(t)$ ($0 \leq r_i(t) \leq 1$) 和 $c_i(t)$ 由下式确定:

$$\begin{aligned} E_R[\phi(U, T) | T=t] &= \alpha \\ E_R[U\phi(U, T) | T=t] &= \alpha E_R[U | T=t] \end{aligned} \quad (16)$$

也就是说, 只要在直线 $Y_0 + Y_1 = t$ 的整数点上寻找最优势的条件检验。则在给定 $Y_0 + Y_1 = t$ 时, Y_1 的条件分布是:

$$P(Y_1 | Y_0 + Y_1 = t) = \binom{t}{Y_1} p^{Y_1} (1-p)^{t-Y_1}, Y_1 = 0, 1, \dots, t \quad (17)$$

其中, $p = \lambda_1 / (\lambda_0 + \lambda_1) = e^\theta / (1 + e^\theta)$ 。这样原有的假设检验问题转化为关于二项分布参数 p 的假设检验问题, 此时的置信区间为反转二项检验统计量^[18]:

$$\begin{aligned} \left\{ p \mid P[\text{Bin}(N_0 + N_1; p) \leq y_1] = \frac{\alpha}{2}, P[\text{Bin}(N_0 + N_1; p) \geq y_1] = \frac{\alpha}{2} \right\} \\ = \text{Beta}\left(\frac{\alpha}{2}; y_1, y_0 + 1\right) < p < \text{Beta}\left(1 - \frac{\alpha}{2}; y_1 + 1, y_0\right) \end{aligned} \quad (18)$$

其中, $\text{Bin}(N_0 + N_1; p)$ 表示试验次数为 $N_0 + N_1$ 成功概率为 p 的二项分布。

(二) Woolf 计算法

Woolf 方法是估计 R 置信区间的快速且最常用的方法。假设 \hat{R} 是 R 的极大似然估计, 根据其具有同变性原理 $\log \hat{R}$ 仍是 $\log R$ 的极大似然估计^[19], 在大样本条件下 $\log \hat{R}$ 近似服从均值为 $E(\log \hat{R})$ 、方差为:

$$\text{Var}(\log \hat{R}) = \left(\frac{1}{Y_0} + \frac{1}{N_0 - Y_0} + \frac{1}{Y_1} + \frac{1}{N_1 - Y_1} \right) \quad (19)$$

的正态分布。由枢轴量法可以得到一个置信水平为 $1 - \alpha$ 的渐近置信区间:

$$\log \hat{R} - Z_{\alpha/2} \cdot \sqrt{\text{Var}(\log \hat{R})} = C_L \leq \log R \leq C_U = \log \hat{R} + Z_{\alpha/2} \cdot \sqrt{\text{Var}(\log \hat{R})} \quad (20)$$

将区间指数化后有:

$$(\max(\exp C_L, 0), \exp C_U]$$

是相对风险的一个区间估计。

三、基于鞍点逼近的置信区间构造

鞍点逼近方法摒弃了传统的使用正态分布进行近似的思维, 通过利用各阶矩来更加完整精确地描述分布的行为来产生更加合理的区间估计。本文将首先给出鞍点逼近的基本原理, 在此基础上给出泊松分布下相对风险的鞍点逼近型置信区间。

(一) 拉普拉斯逼近

存在 $f(x)$ 为一个密度函数, 令 $h(x) = \log f(x)$, 则 $f(x) = \exp h(x)$ 。在 x_0 点处对 $f(x)$ 进行泰勒展开

$$f(x) = \exp h(x) \approx \exp \left\{ h(x_0) + (x - x_0) h'(x_0) + \frac{(x - x_0)^2}{2!} h''(x_0) \right\} \quad (21)$$

令 $x_0 = \hat{x}$ 上式得到简化, 其中 $h'(\hat{x}) = 0$ 时有:

$$f(x) \approx \exp \left\{ h(\hat{x}) + \frac{(x - \hat{x})^2}{2!} h''(\hat{x}) \right\} \quad (22)$$

对上式两边积分得到:

$$\int_{-\infty}^{+\infty} f(x) dx \approx \int_{-\infty}^{+\infty} \exp \left\{ h(\hat{x}) + \frac{(x - \hat{x})^2}{2!} h''(\hat{x}) \right\} dx \approx \exp \{ h(\hat{x}) \} \left(\frac{2\pi}{-h''(\hat{x})} \right)^{1/2} \quad (23)$$

进一步将 $f(x)$ 看成为一个边际密度, 即:

$$f(x) = \int_{-\infty}^{+\infty} m(x, t) dt$$

令 $\kappa(x, t) = \log m(x, t)$, 考虑 $\exp \{\kappa(x, t)\}$ 关于 t 积分的 Laplace 逼近, 对于任意固定的 x 有:

$$f(x) = \int_{-\infty}^{+\infty} \exp \left\{ \kappa(x, \hat{t}(x)) + \frac{(t - \hat{t}(x))^2}{2!} \frac{\partial^2 \kappa(x, t)}{\partial t^2} \Big|_{\hat{t}(x)} \right\} dt = \exp \{ \kappa(x, \hat{t}(x)) \} \left(-\frac{2\pi}{\left(\frac{\partial^2 \kappa(x, t)}{\partial t^2} \right) \Big|_{\hat{t}(x)}} \right)^{1/2}$$

其中, 对每个 $x, \hat{t}(x)$ 满足 $\partial \kappa(x, t)/\partial t = 0$ 和 $\partial^2 \kappa(x, t)/\partial t^2 < 0$ 。对于密度函数 $f(x)$ 其矩生成函数为:

$$\varphi_X(t) = E \exp(tx) = \int_{-\infty}^{+\infty} \exp(tx) f(x) dx$$

运用逆转公式^[20], 可得:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \varphi_X(it) \exp(-itx) dx = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \exp(K_X(it) - itx) dt$$

其中, $i = \sqrt{-1}$, $K_X(t) = \log \varphi_X(t)$ 是累积生成函数, $\varphi_X(it)$ 是特征函数。做变换 $\hat{t} = it$, 由复变函数积分理论, 对于 0 附近的 t 有:

$$\begin{aligned} f(x) &= \frac{1}{2i\pi} \int_{r-\infty}^{r+\infty} \exp \{ K_X(\hat{t}) - \hat{t}x \} dt \approx \frac{1}{2i\pi} \int_{r-\infty}^{r+\infty} \exp \left\{ K_X(\hat{t}(x)) - \hat{t}(x)x + \frac{t - \hat{t}(x)^2}{2} K''_X(\hat{t}(x)) \right\} dt \\ &\approx \left(\frac{1}{2\pi K''_X(\hat{t}(x))} \right)^{1/2} \exp \{ K_X(\hat{t}(x)) - \hat{t}(x)x \} \end{aligned} \quad (24)$$

其中, $\hat{t}(x)$ 满足 $K'(t(x)) = x$ 。在复平面上, $\hat{t}(x)$ 既不是最大值也不是最小值, 而是 $K_X(t) - tx$ 上的鞍点^[21-22]。

(二) 鞍点逼近型置信区间

同样为简化表达, 将相对风险 RR 简记为 R 。在 Poisson 分布下, 相对风险的极大似然估计是 $\hat{R} = \hat{\lambda}_0 / \hat{\lambda}_1$, 考虑 \hat{R} 的累积分布函数如下:

$$P(\hat{R} \leq r) = P \left(\frac{\hat{\lambda}_0}{\hat{\lambda}_1} \leq r \right) = P(Y_0 N_1 - r N_0 Y_1 \leq 0) \quad (25)$$

令一个变量 $Z = N_1 Y_0 - r N_0 Y_1$, 则 Z 的矩生成函数是:

$$\varphi_Z(\theta) = E(e^{\theta Z}) = E(e^{\theta Y_0 N_1}) E(e^{-r N_0 Y_1}) = e^{N_0 \lambda_0 (e^{\theta N_1} - 1)} e^{N_1 \lambda_1 (e^{-\theta N_0 r} - 1)} \quad (26)$$

由此可得 Z 的累积分布函数如下:

$$K_Z(\theta) = \ln \varphi_Z(\theta) = N_0 \lambda_0 (e^{\theta N_1} - 1) + N_1 \lambda_1 (e^{-\theta N_0 r} - 1) \quad (27)$$

进一步有其各阶导函数:

$$\mu_Z(\theta) = K'_Z(\theta) = N_1 N_0 \lambda_0 e^{\theta N_1} - N_1 N_0 \lambda_1 r e^{-\theta N_0 r}$$

$$\sigma_Z^2(\theta) = K''_Z(\theta) = N_1^2 N_0 \lambda_0 e^{\theta N_1} + N_1 N_0^2 \lambda_1 r^2 e^{-\theta N_0 r}$$

$$K_Z^{(3)}(\theta) = N_1^3 N_0 \lambda_0 e^{\theta N_1} - N_1 N_0^3 \lambda_1 r^3 e^{-\theta N_0 r}$$

$$K_Z^{(4)}(\theta) = N_1^4 N_0 \lambda_0 e^{\theta N_1} + N_1 N_0^4 \lambda_1 r^4 e^{-\theta N_0 r}$$

根据上文中鞍点逼近原理, 令 $K'_Z(\theta) = 0$, 得到如下鞍点方程:

$$N_1 N_0 \lambda_0 e^{\theta N_1} - N_1 N_0 \lambda_1 r e^{-\theta N_0 r} = 0$$

求解方程可得到鞍点 $\hat{\theta}$, 鞍点方程通常没有显式解, 可以通过数值方法求解^[23]。运用 Jensen 中的逼近公式 (2.2.10)^[24]:

$$\begin{aligned} P(\hat{R}' \geq r) &= \frac{\varphi_Z(\theta) \exp(-\theta r)}{\sigma_Z(\theta) \{1 - \exp(-|\theta|)\}} \left\{ B_0(\lambda) + \left[\frac{1}{\sigma_Z(\theta)} \left(\frac{1}{|\theta|} - \gamma_\theta \right) B_1(\lambda) + \frac{\zeta_3(\theta)}{6} \operatorname{sgn}(\theta) B_3(\lambda) \right] \right. \\ &\quad \left. + \left[\gamma_\theta \left(\frac{1}{2} - \frac{1}{|\theta|} - \gamma_\theta^2 \right) \frac{B_2(\lambda)}{\sigma_Z^2(\theta)} + \frac{1}{\sigma_Z(\theta)} \left(\frac{1}{|\theta|} - \gamma_\theta \right) \frac{\zeta_3(\theta)}{6} \operatorname{sgn}(\theta) B_4(\lambda) + \frac{\zeta_4(\theta)}{24} B_5(\lambda) + \frac{\zeta_3(\theta)^2}{72} B_6(\lambda) \right] \right\} \end{aligned} \quad (28)$$

其中,

$$\lambda = \sqrt{n} |\theta| \sigma_Z(\theta) = \sqrt{K''_Z(\theta)}, \gamma_\theta = \exp(-|\theta|) / \{1 - \exp(-|\theta|)\}$$

$$B_0(\lambda) = \lambda \exp(\lambda^2/2 \{1 - \Phi(\lambda)\})$$

$$B_1(\lambda) = -\lambda \{B_0(\lambda) - (2\pi)^{-1/2}\}$$

$$B_2(\lambda) = \lambda^2 \{B_0(\lambda) - (2\pi)^{-1/2}\}$$

$$B_3(\lambda) = -\{\lambda^3 B_0(\lambda) - (\lambda^3 - \lambda)(2\pi)^{-1/2}\}$$

$$B_4(\lambda) = \lambda^4 B_0(\lambda) - (\lambda^4 - \lambda^2)(2\pi)^{-1/2}$$

$$B_5(\lambda) = -\{\lambda^5 B_0(\lambda) - (\lambda^5 - \lambda^3 + 3\lambda)(2\pi)^{-1/2}\}$$

$$B_6(\lambda) = \lambda^6 B_0(\lambda) - (\lambda^6 - \lambda^4 + 3\lambda^2)(2\pi)^{-1/2}$$

$$\zeta_3(\theta) = \frac{K_Z^{(3)}(\theta)}{K_Z^{(2)}(\theta)^{3/2}}, \zeta_4(\theta) = \frac{K_Z^{(4)}(\theta)}{K_Z^{(2)}(\theta)^2}$$

此外, $\Phi(\cdot)$ 是标准正态分布的分布函数, sgn 是符号函数。将鞍点 $\hat{\theta}$ 带入 $P(\hat{R}' \geq r)$, 那么 \hat{R} 的近似累积分布函数为 $F_R(r) \approx 1 - P(\hat{R}' \geq r)$ 。对于任意的 r , 置信区间上下限 \hat{R}_U 和 \hat{R}_L 都能定义为:

$$F_R(r | \hat{R}_U(r)) = 1 - \frac{\alpha}{2}, F_R(r | \hat{R}_L(r)) = \frac{\alpha}{2}$$

求解 r 则能得到 \hat{R} 一个置信水平为 $1 - \alpha$ 的置信区间 $(\max\{\hat{R}_L(r), 0\}, \hat{R}_U(r)]^{[25-26]}$ 。

四、模拟研究

采用蒙特卡洛模拟法, 对所提出的区间估计方法进行性能比较和验证。在实际应用中, 如果所构建的置信区间具有接近显著性水平的覆盖概率、较短的区间长度, 并且置信区间的端点符合实际情况, 则说明该估计方法表现更佳。具体而言可以使用覆盖率、平均置信区间长度、区间范围是否合理三个指标进行评价。

选择基于 Poisson 分布的对照实验, 并设计了三种可能的情形: $R > 1, R = 1, R < 1$ 。其中, 暴露组和对照组的样本量分别设为 5、15、25、100、150 和 200, 以便同时考察在小样本和大样本情形(本文中将样本量大于 30 定义为大样本)的不同估计方法的性能, 名义水平设定为 0.05。为尽可能避免随机性带来的误差, 每组实验将重复 2 000 次, 在每组实验结果中报告各估计方法得到的平均上限值、平均下限值、平均长度、区间是否合理。

鞍点逼近方法被记为 SPA, 精确条件检验法被记为 Exact, 似然比检验法被记为 LRT, 对数转换法被记为 log。为更加直观地汇报实验结果, 将按照 R 的设定形式进行分组汇总。其中表 1~3 汇报了 $R > 1$ 时不同样本量的区间估计, 另外表 4~6 为 $R = 1$, 表 7~9 为 $R < 1$ 。其中, 黑底数字突出显示了在当前评价指标下取得的最优数值。

(一) 相对风险大于 1 的情形

在该组模拟实验中, 设定相对风险大于 1 即暴露组中的事件发生风险高于对照组的情况。

在所有实验条件下, SPA 法均表现出较高的覆盖率和合理的区间长度。特别是在表 1 和表 2 中, SPA 法的覆盖率分别为 95.05% 和 94.2%, 在所有方法中均表现最佳。此外, SPA 法的区间长度也较为紧凑, 表明其在保证覆盖率的同时, 能够有效缩短区间长度, 提供更精确的估计。在表 1 和表 2 中, Wald 法的表现与 SPA 法相当。然而在表 3 中, Wald 法的覆盖率为 95.45%, 比 SPA 更加接近名义水平, 同时区间长度更短, 这表明在大样本下当发病率差距较大时 SPA 法可能没有 Wald 法稳健。Woolf 法和 LRT 法这两种方法在不

表 1 $R=2, N_0=100, N_1=100, \lambda_0=0.8, \lambda_1=0.4$

方法	平均下限	平均上限	平均区间长度	覆盖率	区间是否合理	$E(\hat{R})$
SPA	1.335 7	2.911 9	1.576 2	0.950 5	Y	2.044 4
Wald	1.274 0	2.872 2	1.598 2	0.951 5	Y	
Woolf	1.038 5	4.066 7	3.028 2	0.979 0	Y	
LRT	1.155 5	2.977 1	1.821 6	0.970 0	Y	
log	1.573 0	2.672 2	1.099 2	0.832 5	Y	
Fieller	1.578 9	2.682 5	1.103 6	0.835 0	Y	
Exact	1.372 7	3.078 7	1.706 0	0.965 0	Y	

表 2 $R=1.67, N_0=10, N_1=15, \lambda_0=3, \lambda_1=1.8$

方法	平均下限	平均上限	平均区间长度	覆盖率	区间是否合理	$E(\hat{R})$
SPA	1.070 4	2.545 0	1.484 5	0.942 0	Y	1.709 4
Wald	0.968 8	2.521 9	1.553 0	0.934 0	Y	
Woolf	0.674 5	12.402 1	11.727 6	0.002 0	Y	
LRT	0.852 5	2.621 0	1.768 4	0.962 0	Y	
log	0	3.256 2	4.575 4	1	N	
Fieller	0	3.987 1	4.772 1	1	N	
Exact	1.060 1	2.744 2	1.674 1	0.962 0	Y	

同条件下的表现波动较大。尤其是在表 3 中, Woolf 法的覆盖率明显偏低(51.05%),说明其在高风险比条件下的置信区间估计不够稳定。虽然在部分情况下 Exact 法表现不错,但在样本量较大时其区间长度显著增加。

表 3 $R=6, N_0=100, N_1=100, \lambda_0=6, \lambda_1=1$

方法	平均下限	平均上限	平均区间长度	覆盖率	区间是否合理	$E(\hat{R})$
SPA	5.112 4	7.247 1	2.134 7	0.933 0	Y	6.054 9
Wald	5.010 6	7.121 2	2.110 6	0.954 5	Y	
Woolf	3.087 0	15.366 1	12.279 1	0.510 0	Y	
LRT	4.864 1	7.282 3	2.418 2	0.973 0	Y	
log	5.001 2	7.429 1	2.489 0	0.894 0	Y	
Fieller	4.979 8	7.601 7	2.621 7	1	Y	
Exact	5.077 6	7.245 0	2.167 4	0.958 5	Y	

(二) 相对风险等于 1 的情形

在该组模拟实验中,设定相对风险等于 1 即暴露组中的事件发生风险与对照组相同的情况。

通过表 4~6 中的实验结果可以看出,SPA 法在大部分条件下均表现出色,无论在覆盖率还是区间长度方面均取得了较好的平衡。在表 4 中,SPA 法的覆盖率为 95.1%,且区间长度较短(0.262 6),表明其在较大样本条件下提供了稳定且精确的估计。在表 5 中,虽然样本量较小,但 SPA 法仍然维持了较高的覆盖率(95.6%),显示了其在小样本情况下的优势。而在表 6 中,SPA 法的表现依然优异,覆盖率达到 94.7%,且区间长度相对其他方法较短,进一步证实了 SPA 法在不同样本量和条件下的鲁棒性。相比之下,Wald 法在部分情况下表现接近 SPA 法,尤其在表 6 中的覆盖率达到 94.55%,但其区间长度通常略长。其他方法如

表 4 $R=1, N_0=150, N_1=200, \lambda_0=\lambda_1=2$

方法	平均下限	平均上限	平均区间长度	覆盖率	区间是否合理	$E(\hat{R})$
SPA	0.824 3	1.086 0	0.262 6	0.951 3	Y	1.000 8
Wald	0.863 3	1.141 4	0.278 1	0.945 0	Y	
Woolf	0.404 9	3.445 2	3.040 3	0.970 2	Y	
LRT	0.843 7	1.162 0	0.318 2	0.974 0	Y	
log	0.423 5	3.041 6	2.618 1	0.946 0	Y	
Fieller	0.426 9	3.065 3	2.638 4	0.934 9	Y	
Exact	0.868 5	1.152 0	0.283 5	0.946 5	Y	

表 5 $R=1, N_0=15, N_1=20, \lambda_0=\lambda_1=2$

方法	平均下限	平均上限	平均区间长度	覆盖率	区间是否合理	$E(\hat{R})$
SPA	0.643 6	1.397 8	0.754 1	0.956 0	Y	1.021 0
Wald	0.631 4	1.433 1	0.801 7	0.943 0	Y	
Woolf	0.072 6	4.113 8	4.036 0	0.910 5	Y	
LRT	0.471 8	1.604 9	1.330 9	0.965 0	Y	
log	0.148 5	2.672 0	2.523 5	0.886 0	Y	
Fieller	0.170 1	2.673 5	2.503 4	0.862 0	Y	
Exact	0.610 3	1.685 1	1.074 7	0.972 0	Y	

Woolf 和 LRT 法在某些情况下表现较差,特别是在表 5 中的覆盖率偏低,这表明这些方法在某些高风险比和小样本条件下可能不够可靠。

表 6 $R=1, N_0=200, N_1=200, \lambda_0=\lambda_1=0.75$

方法	平均下限	平均上限	平均区间长度	覆盖率	区间是否合理	$E(\hat{R})$
SPA	0.751 8	1.186 5	0.434 7	0.947 0	Y	1.006 2
Wald	0.782 2	1.241 3	0.459 0	0.945 5	Y	
Woolf	0.632 3	1.603 6	0.971 3	0.999 5	Y	
LRT	0.748 1	1.272 4	0.524 2	0.971 5	Y	
log	0.894 9	1.124 1	0.229 2	0.662 0	Y	
Fieller	0.894 4	1.123 5	0.229 1	0.659 0	Y	
Exact	0.794 9	1.268 3	0.473 4	0.954 5	Y	

(三) 相对风险小于 1 的情形

在该组模拟实验中,设定相对风险大于 1 即暴露组中的事件发生风险低于对照组的情况。

实验结果表明,SPA 法在所有测试中保持较高的覆盖率和合理的区间长度,特别是在表 7~9 中,覆盖率达 94.6%、94.9% 和 94.6%。相比之下,Wald 法在部分条件下接近 SPA 法,但区间长度明显较长(如表 9 中为 1.701 0,远高于 SPA 的 0.684 8);Woolf 和 LRT 法在部分情形下覆盖率较高,但区间长度通常不及 SPA。总体来看,SPA 法在小样本和低风险比条件下表现尤佳,是构建相对风险量置信区间的优选方法。

表 7 $R=1/3, N_0=100, N_1=100, \lambda_0=0.1, \lambda_1=0.3$

方法	平均下限	平均上限	平均区间长度	覆盖率	区间是否合理	$E(\hat{R})$
SPA	0.190 1	0.534 2	0.368 1	0.946 0	Y	0.342 3
Wald	0.166 1	0.569 1	0.379 0	0.940 5	Y	
Woolf	0.190 3	0.608 1	0.417 8	0.976 0	Y	
LRT	0.139 2	0.558 8	0.419 6	0.959 5	Y	
log	0.208 2	0.548 5	0.340 4	0.934 0	Y	
Fieller	0.205 3	0.543 3	0.338 0	0.928 5	Y	
Exact	0.190 2	0.590 1	0.399 8	0.966 5	Y	

表 8 $R=1/3, N_0=15, N_1=5, \lambda_0=1, \lambda_1=3$

方法	平均下限	平均上限	平均区间长度	覆盖率	区间是否合理	$E(\hat{R})$
SPA	0.186 9	0.605 3	0.408 4	0.948 5	Y	0.342 2
Wald	0.151 1	0.553 4	0.412 2	0.936 5	Y	
Woolf	0.081 1	1.162 6	1.081 5	0.475 0	Y	
LRT	0.121 5	0.579 7	0.458 2	0.953 0	Y	
log	0.259 3	0.641 4	0.558 9	1	Y	
Fieller	0.271 8	0.690 3	0.570 4	0.998 9	Y	
Exact	0.182 6	0.627 8	0.445 2	0.964 0	Y	

表 9 $R=5/9, N_0=15, N_1=5, \lambda_0=0.05, \lambda_1=0.09$

方法	平均下限	平均上限	平均区间长度	覆盖率	区间是否合理	$E(\hat{R})$
SPA	0.302 8	0.987 4	0.684 8	0.946 0	Y	0.558 5
Wald	0	1.701 0	1.701 0	0.984 0	N	
Woolf	0.027 1	6.656 0	6.628 9	1	Y	
LRT	0	1.865 6	1.865 6	0.999 8	N	
log	0.048 8	1.379 1	1.330 3	0.847 5	Y	
Fieller	0.042 4	1.379 1	1.336 7	1	Y	
Exact	0.016 9	1.027 2	0.990 3	0.964 0	Y	

模拟研究中,本文分别考察了高/低发病率与样本量的组合情况。由于 R 的极大似然估计在样本量小或 $R_1=0, R=\infty$ 情况下不稳定,我们采用类似逆抽样策略,确保每次实验中 Y_0, Y_1 至少有一项非零。表 1~9 的结果显示,鞍点逼近法不仅使用极大似然信息,还利用 R 分布的前四阶矩,从而获得更短、更合理的区

间,覆盖率接近名义水平。

相比之下,log 和 LRT 等法易得负下限;Exact 法虽准确,但依赖大样本。其余大样本方法虽可得合理区间,但区间偏宽。综上,鞍点逼近法在各种样本量与发病率条件下均表现优异,是令人满意的置信区间估计方法。

五、真实数据分析

(一)过食快餐对女性甲状腺癌的影响

甲状腺癌已成为发展中国家的常见病,许多引发甲状腺癌的因素已被确认,如年龄、性别、碘缺乏、经常食用快餐等。数据来源于 Asif 的研究^[27],该研究旨在通过调查不同的危险因素来了解该疾病在城乡地区的病因和影响。为此,从拉合尔核医学和肿瘤研究所与拉合尔谢赫扎耶德医院收集病例数据,共纳入 88 名农村患者和 232 名城市患者,通过问卷调查的方式调查甲状腺癌的可能危险因素,其中是否过多食用快餐是城市女性患甲状腺癌最重要的因素。

从表 10 可知 $\hat{R}=1.32$, $\hat{\lambda}_0=0.63$, $\hat{\lambda}_1=0.48$,同时使用上述七种方法估计 \hat{R} 的 95% 置信区间,结果如表 11 所示。

不难看出,所有估计方法都正确覆盖了 \hat{R} ,其中鞍点逼近方法取得了最短的区间长度,对数 Wald 检验和 Fieller 检验的结果其次,Woolf 方法获得的置信区间最长。另一方面,由 $\hat{R}>1$ 可以认为暴露组的风险高于对照组,即过食快餐对女性患甲状腺癌而言是危险因子。

(二)锻炼频率对 CHF 的影响

Belardinelli 等将 99 名稳定型充血性心力衰竭(CHF)患者随机分配至锻炼计划($N_0=50$)组或不锻炼($N_1=49$)组,并每周对患者进行两次随访,为期一年^[28]。感兴趣的结果是全因死亡率。治疗组的受试者每周锻炼 3 次,持续 8 周,然后每周锻炼两次持续 1 年。研究结果表明,与没有训练的人相比,进行运动训练的人死亡率较低。从表 12 可知 $\hat{R}=0.441$, $\hat{\lambda}_0=0.18$, $\hat{\lambda}_1=0.41$,同时使用上述七种方法估计 \hat{R} 的 95% 置信区间。从表 13 的结果可知,鞍点逼近方法置信区间的长度显著短于其他六种估计方法,且正确覆盖,Wald 检验和 Fieller 检验的效果其次,Woolf 方法置信区间长度依然最长。由于 $\hat{R}<1$ 暴露组风险小于对照组风险,即进行锻炼对于 CHF 患者的生存是保护因素。

(三)吸烟对患肺癌的影响

Doll 等报告了肺癌病因学回顾性研究的结果^[29-30]。将 20 家医院的 709 名肺癌患者与对照组 709 名无肺癌患者进行比较。对于每名肺癌患者,从同一家医院、同一性别、同一年龄组中选择一名对照患者。每组中的每位患者都接受了一名社会工作者的访谈,所有访谈者都使用相同的问卷,以保证病例组和对照组之间唯一的实质性差异在于他们报告的吸烟习惯。该项研究的随访时间长达 40 年,与许多其他类似的研究一样,他们都强烈表明吸烟与患肺癌的风险之间存在关联。

表 10 城市女性甲状腺癌患者的随机对照实验结果

过食快餐	患甲状腺癌		
	是	否	总计
是	68	40	108
否	59	65	124
总计	127	105	232

表 11 甲状腺癌数据集置信区间估计结果

方法	下限	上限	区间长度	\hat{R}
SPA	1.312 7	1.334 0	0.021 3	1.323 3
Wald	0.861 8	1.784 7	0.922 9	
Woolf	0.782 0	2.239 3	1.457 4	
LRT	0.795 6	1.851 0	1.055 4	
log	1.046 5	1.673 2	0.626 7	
Fieller	1.048 5	1.677 4	0.628 9	
Exact	0.919 9	1.908 3	0.988 4	

表 12 锻炼对 CHF 患者生存影响的随机对照实验结果

治疗	状态		
	死亡	存活	总计
锻炼	9	41	50
不锻炼	20	29	49
总计	29	70	99

表 13 CHF 数据集 \hat{R} 置信区间估计结果

方法	下限	上限	区间长度	\hat{R}
SPA	0.421 6	0.461 0	0.039 4	0.441 0
Wald	0.094 1	0.787 9	0.693 9	
Woolf	0.175 9	1.105 8	0.929 9	
LRT	0.044 2	0.837 8	0.793 5	
log	0.223 2	0.871 3	0.648 1	
Fieller	0.212 5	0.844 0	0.631 5	
Exact	0.176 8	1.013 2	0.836 4	

表 14 吸烟对患肺癌风险影响的对照实验结果

是否吸烟	实验组		
	肺癌组	对照组	总计
吸烟者	83	72	155
非吸烟者	3	14	17
总计	86	86	172

从表 14 可知 $\hat{R} = 3.03$, $\hat{\lambda}_0 = 0.54$, $\hat{\lambda}_1 = 0.18$, 同时使用上述七种方法估计 \hat{R} 的 95% 置信区间, 结果如表 15 所示。

该实例中, 由于对照组的发病率显著小于暴露组, 所以导致 \hat{R} 的数值变大。在这种情况下 Wald 检验、似然比检验得到小于 0 的错误区间下限, 而精确条件检验获得的区间长度因低发病率而变得异常长。反观鞍点逼近方法并未受到影响, 这与模拟实验中的结果一致。可见, 鞍点逼近方法在分布复杂或非标准的情况下也能获得最优的区间估计。

表 15 肺癌数据集置信区间估计结果

方法	下限	上限	区间长度	\hat{R}
SPA	2.9218	3.1563	0.2345	3.0344
Wald	0	6.5296	6.5296	
Woolf	0.8384	10.9824	10.1440	
LRT	0	7.0315	7.0315	
log	1.0754	8.5620	7.4865	
Fieller	1.2372	9.1032	7.8660	
Exact	1.0029	15.0131	14.0102	

六、结论与启示

通过鞍点逼近方法, 本文构造了泊松分布下相对风险更加精确的置信区间, 模拟和实证研究均表明, 在常见和特殊的对照实验设定下, 鞍点逼近相较于现有方法具有更加稳定的估计效果且区间长度更短。这也意味着鞍点逼近方法可以更好地为流行病学研究问题进行指导。需要指出的是: 从方法上, 尽管鞍点逼近能够得到更短的置信区间和更加准确的覆盖率, 但对于一些复杂的检验问题往往无法显式地构造矩生成函数, 使得无法方便地使用鞍点逼近方法; 从应用上, 较短的置信区间可能也会导致对暴露因素影响程度的高估, 从而导致错误的诊断, 所以在实际应用中仍需要考虑其他辅助信息^[31-33]。因此, 如何进一步简化鞍点逼近的计算和提高适用度是需要进一步讨论的内容。

参考文献:

- [1] BRESLOW N E, DAY N E, HESELTINE E. Statistical methods in cancer research — The analysis of case-control studies [M]. Lyon: IARC Scientific Publications, 1980: 85-115.
- [2] SEDGWICK P, MARSTON L. Relative risks[J]. BMJ, 2010, 341: c3983.
- [3] WANG W, LU S, XIE T. Optimal confidence intervals for the relative risk and odds ratio[J]. Statistics in medicine, 2023, 42(3): 281-296.
- [4] TANG M L, TIAN M. Approximate confidence interval construction for risk difference under inverse sampling[J]. Statistics and computing, 2010, 20: 87-98.
- [5] 夏丽丽, 田茂再, 朱钰. 二项分布下基于鞍点逼近的总体成数置信区间的构造[J]. 统计与信息论坛, 2019, 34(9): 3-9.
- [6] 孟令宾, 李二倩, 田茂再. 基于鞍点逼近的二项抽样下优势比的置信区间构造[J]. 数理统计与管理, 2017, 36(1): 85-102.
- [7] 白永昕, 田茂再. Poisson 分布下基于鞍点逼近的慢性病风险差的置信区间构造[J]. 高校应用数学学报 A 编, 2017, 32(3): 253-266.
- [8] 王维贤, 田茂再. 鞍点逼近下相关差的置信区间构造[J]. 高校应用数学学报 A 编, 2019, 34(3): 264-272.
- [9] CANNON-ALBRIGHT L A, FOSTER N L, SCHLIEP K, et al. Relative risk for Alzheimer disease based on complete family history[J]. Neurology, 2019, 92(15): e1745-e1753.
- [10] BYERLEY C O N, HORNE D, GONG M, et al. An effective COVID-19 vaccine hesitancy intervention focused on the relative risks of vaccination and infection[J]. Scientific reports, 2024, 14(1): 7419.
- [11] EDERER F, MANTEL N. Confidence limits on the ratio of two Poisson variables[J]. American journal of epidemiology, 1974, 100(3): 165-167.
- [12] CROW E L. Confidence intervals for a proportion[J]. Biometrika, 1956, 43(3/4): 423-435.
- [13] KATZ D, BAPTISTA J, AZEN S P, et al. Obtaining confidence intervals for the risk ratio in cohort studies[J]. Biometrics, 1978, 34(3): 469-474.
- [14] PAPANICOLAOU A. Taylor approximation and the delta method [R/OL]. (2009-04-28) [2024-03-21]. https://www.stat.rice.edu/~dobelman/notes_papers/math/TaylorAppDeltaMethod.pdf.
- [15] BARKER L, CADWELL B L. An analysis of eight 95 percent confidence intervals for a ratio of Poisson parameters when events are rare[J]. Statistics in medicine, 2008, 27(20): 4030-4037.
- [16] MURPHY S A. Likelihood ratio-based confidence intervals in survival analysis[J]. Journal of the American Statistical Association, 1995, 90(432): 1399-1405.

- [17] FIELLER E C. Some problems in interval estimation[J]. Journal of the royal statistical society series B: statistical methodology, 1954, 16(2): 175-185.
- [18] FAY M P. Confidence intervals that match Fisher's exact or Blaker's exact tests[J]. Biostatistics, 2010, 11(2): 373-374.
- [19] WOOLF B. On estimating the relation between blood group and disease[J]. Annals human genetics, 1955, 19(4): 251-253.
- [20] FELLER W. An introduction to probability theory and its applications, Volume 2[M]. New York: John Wiley & Sons, 1991: 500-510.
- [21] DANIELS H E. Saddlepoint approximations in statistics[J]. The annals of mathematical statistics, 1954, 25(4): 631-650.
- [22] 田茂再,吴喜之,李远,等.逆抽样下流行病发病率的逼近与渐近置信区间[J].系统科学与数学,2008(5):513-523.
- [23] ZHANG W, BRAVINGTON M V, FEWSTER R M. Fast likelihood-based inference for latent count models using the saddlepoint approximation[J]. Biometrics, 2019, 75(3): 723-733.
- [24] JENSEN J L. Saddlepoint approximations[M]. Oxford: Clarendon Press, 1995: 32.
- [25] 钱政超,张晨阳,孟令宾,等.二项抽样下基于鞍点逼近方法的流行病相对风险置信区间构造[J].数学的实践与认识,2014,44(21):204-217.
- [26] 闵素芹,张宁.鞍点逼近法在NNT区间估计中的应用与模拟研究[J].统计与信息论坛,2014,29(9):11-14.
- [27] ASIF F, AHMAD M R, MAJID A. Risk factors for thyroid cancer in females using a logit model in Lahore, Pakistan[J]. Asian Pacific journal of cancer prevention, 2015, 16(15): 6243-6247.
- [28] BELARDINELLI R, GEORGIOU D, CIANCI G, et al. Randomized, controlled trial of long-term moderate exercise training in chronic heart failure: effects on functional capacity, quality of life, and clinical outcome[J]. Circulation, 1999, 99(9): 1173-1182.
- [29] DOLL R, HILL A B. Smoking and carcinoma of the lung[J]. British medical journal, 1950, 2: 739-748.
- [30] ARMITAGE P, BERRY G, MATTHEWS J N S. Statistical methods in medical research[M]. New York: John Wiley & Sons, 2008: 648-716.
- [31] 郭冰清,姚强.新时代中国居民健康保障获得感变化趋势及影响因素研究[J].统计与信息论坛,2023,38(9):95-111.
- [32] 田梓璇,谢小月.大数据序贯检验方法及其应用[J].统计与信息论坛,2024,39(9):13-22.
- [33] 熊巍,王娟,潘晗,等.多重稳健的高维缺失数据插补研究[J].统计与信息论坛,2023,38(2):3-15.

Relative Risk Confidence Interval Construction Based on Saddle Point Approximation under Poisson Distribution

HOU Jian^{a,b}, LIU Shuo^{a,b}, TIAN Maozai^{a,b}

(a. Center for Applied Statistics; b. School of Statistics, Renmin University of China, Beijing 100782, China)

Abstract: Relative risk (RR) is a commonly used indicator in epidemiological randomized controlled trials to compare morbidity or mortality between two groups. Interval estimation of RR is an effective statistical inference method that reflects uncertainty and confidence. However, traditional interval estimation methods, such as Wald's interval, Wilson interval, Fisher interval, etc., have certain limitations: large error, low coverage, long interval length, and unreasonable interval. In order to overcome these shortcomings, an interval estimation method based on saddle point approximation is proposed: in the context of Poisson distribution, the approximation formula of RR is derived by using Taylor expansion and Laplace integral, and the corresponding confidence interval construction method is given. Through Monte Carlo simulation and empirical analysis, the saddle point approximation method is compared with some existing interval estimation methods, and the results show that the saddle point approximation method can obtain the shortest confidence interval in the case of small or large samples, and the interval coverage is closer to the nominal level, which provides a superior interval estimation method for RR. Finally, the advantages and limitations of the saddle point approximation method, as well as its application value and prospect in epidemiological research are discussed.

Key words: Poisson distribution; saddle point approximation; interval estimation; relative risk; randomized controlled trial

(责任编辑:崔国平)