

极端频率情形下二项分布比例置信区间的比较*

张学新

(湖北工程学院 数学与统计学院, 湖北 孝感 432000)

摘要:为在二项分布比例参数 p 的点估计 \hat{p} 值接近于 0 或 1 的极端情形下选择 p 的置信区间,对 Wald, Plus 4, Wilson, CP, Jeffreys, Hall, Kott-Liu, Cai 和 T-approximation 等 17 种二项分布比例参数 p 的置信区间的平均覆盖率与期望长度进行模拟比较,分析各置信区间覆盖率随比例参数 p 变动时的表现。研究结果表明:Wilson Score 检验的置信区间总体最优,在二项分布比例的点估计 \hat{p} 值很小或很大时,Cai, Kott-Liu 置信区间的覆盖率较好。

关键词:二项分布比例 p ; 点估计 \hat{p} 值极端; 置信区间; 覆盖率; 平均区间长度

中图分类号: O212.1

文献标志码: A

文章编号: 1672-6693(2016)03-0091-07

某类样本出现很小,甚至为 0 频数时,如何估计该类占总体的比例? 如果运用传统的二项分布比例参数的置信区间,例如利用 Wald 检验方法,得出关于 p 的 95% 置信区间等于 $\hat{p} \pm u_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = (0, 0)$, 几乎没有意义。国内外已有一些对二项分布比例参数的置信区间修正研究。刘镇瑜、王军等人^[1]根据二项分布离散的特点提出的一种暴力算法,指出 Wald 方法在各方面的性能都很差,建议进行修改。姚励^[2]、李丹^[3]比较了 Wilson, Agresti-Coull 和 Jeffreys 等 3 个区间的优劣。Zhou 等人^[4]基于 p 的 Logit 变换的 Edgeworth 展开,提出一种二项比例的置信区间,比较了 Wilson, Agresti-Coull 及 Jeffreys 区间的 2 个性能。Alan Agresti 等人^[5-11]对分类数据分析的最优置信区间进行了探讨,对 Wald 检验、Score 检验、P(Plus) 4 检验和 Clopper-Pearson 检验进行了分析对比。

本文模拟比较多达 17 种方法的置信区间,特别考虑了极端频率情形,即点估计 \hat{p} 值接近于 0 或接近于 1 时,有关置信区间的表现。

1 沃尔德(Wald)检验及 p 的置信区间

设二项分布比例参数 p 的点估计为 \hat{p} , SE 为 \hat{p} 的标准误不为零。沃尔德统计量为 $u_w = \frac{\hat{p} - p_0}{SE} = \frac{\hat{p} - p_0}{\sqrt{\hat{p}(1-\hat{p})/n}}$, 在 $H_0: p = p_0$ 时近似服从标准正态分布。解关于 p_0 不等式组 $|u_w| < u_{1-\frac{\alpha}{2}}$, 得:

$$\hat{p} \mp u_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \quad (1)$$

它是一种传统置信区间,依据中心极限定理,除非 n 非常大,否则该置信区间很不准确。本文的模拟表明,当 p 的值接近于 0 或 1 时,它的概率通常会低于名义置信水平 $100(1-\alpha)\%$ 。

2 改进的置信区间

2.1 Wilson 置信区间^[5]

原假设 $H_0: p = p_0$ 。为利用对数似然函数 $L(p)$ 的斜率和曲率。令 $u(p_0) = \frac{m}{p_0} - \frac{n-m}{1-p_0}$, $l(p_0) = \frac{n}{p_0(1-p_0)}$,

* 收稿日期:2015-07-28 修回日期:2015-11-19 网络出版时间:2016-04-29 18:34

作者简介:张学新,男,副教授,博士,研究方向为概率论与数理统计方法应用,E-mail:zxxhss@sina.com

网络出版地址:http://www.cnki.net/kcms/detail/50.1165.n.20160429.1834.020.html

则统计量 $u_s = \frac{u(p_0)}{[l(p_0)]^{\frac{1}{2}}}$, 即 $u_s = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \sim N(0, 1)$ 。与沃尔德统计量 u_w 使用 \hat{p} 所对应的标准误不同, 得分

统计量 u_s 使用 p_0 对应的标准误。相应置信区间可表示为:

$$\frac{n\hat{p} + \frac{1}{2}u_{1-\frac{\alpha}{2}}^2}{n+u_{1-\frac{\alpha}{2}}^2} \mp \frac{u_{1-\frac{\alpha}{2}}\sqrt{n}}{n+u_{1-\frac{\alpha}{2}}^2} \sqrt{\hat{p}(1-\hat{p}) + \frac{u_{1-\frac{\alpha}{2}}^2}{4n}} = \hat{p} \left(\frac{n}{n+u_{1-\frac{\alpha}{2}}^2} \right) + \frac{1}{2} \left(\frac{u_{1-\frac{\alpha}{2}}^2}{n+u_{1-\frac{\alpha}{2}}^2} \right) \mp u_{1-\frac{\alpha}{2}} \times \sqrt{\frac{\left[\hat{p}(1-\hat{p}) \left(\frac{n}{n+u_{1-\frac{\alpha}{2}}^2} \right) + \left(\frac{1}{2} \right) \left(\frac{1}{2} \right) \left(\frac{u_{1-\frac{\alpha}{2}}^2}{n+u_{1-\frac{\alpha}{2}}^2} \right) \right]}{n+u_{1-\frac{\alpha}{2}}^2}}}. \quad (2)$$

可以这样理解: 置信区间的中点 $\tilde{p} = \hat{p} \left(\frac{n}{n+u_{1-\frac{\alpha}{2}}^2} \right) + \frac{1}{2} \left(\frac{u_{1-\frac{\alpha}{2}}^2}{n+u_{1-\frac{\alpha}{2}}^2} \right)$ 是关于 \hat{p} 和 $\frac{1}{2}$ 的加权平均数, 进一步可得 $\tilde{p} = (m+u_{1-\frac{\alpha}{2}}^2)/(n+u_{1-\frac{\alpha}{2}}^2)$, 等同于在发生与不发生两类观测值中各增加了 $u_{1-\frac{\alpha}{2}}^2$ 个观测值后所得到的新样本中的点估计, 样本大小是 $n+u_{1-\frac{\alpha}{2}}^2$ 而不是 n 。

Wilson 置信区间也称 Wilson Score 置信区间、Rao's Score、Score^[11] 置信区间。

当 $\hat{p} \rightarrow 0$ 时, 令 λ 是 $\sum_{k=0}^{\hat{p}-1} \frac{\lambda^k}{k!} e^{-\lambda} = 1 - \alpha$ 的解, 把下限用 λ/n 取代得到修正的 Wilson p 置信区间。

为了使最小的覆盖率与名义值匹配, 可以做 Yates 修正, 得置信区间限^[12]为:

$$\left\{ \begin{array}{l} w^- = \max \left\{ 0, \frac{2n\hat{p} + u_{1-\frac{\alpha}{2}}^2 - \left[u_{1-\frac{\alpha}{2}} \sqrt{u_{1-\frac{\alpha}{2}}^2 - \frac{1}{n} + 4n\hat{p}(1-\hat{p}) + (4\hat{p}-2) + 1} \right]}{2(n+u_{1-\frac{\alpha}{2}}^2)} \right\}, \\ w^+ = \min \left\{ 1, \frac{2n\hat{p} + u_{1-\frac{\alpha}{2}}^2 + \left[u_{1-\frac{\alpha}{2}} \sqrt{u_{1-\frac{\alpha}{2}}^2 - \frac{1}{n} + 4n\hat{p}(1-\hat{p}) - (4\hat{p}-2) + 1} \right]}{2(n+u_{1-\frac{\alpha}{2}}^2)} \right\}. \end{array} \right. \quad (3)$$

2.2 P(Plus) 4 检验方法及 p 的置信区间^[5]

P(Plus) 4 检验是在 $\frac{|\hat{p} - p_0|}{\sqrt{\frac{p_0(1-p_0)}{n}}} < u_{1-\frac{0.5}{2}}$ 中, 把 $u_{1-\frac{0.5}{2}} = 1.96$ 以 2 代替, 得 P(Plus) 4 检验比例参数的 95% 置信区间为:

$$\frac{n\hat{p} + 2}{n+4} \mp 2 \sqrt{\frac{(n\hat{p} + 2) \left(1 - \frac{n\hat{p} + 2}{n+4} \right)}{n+4}}. \quad (4)$$

不难看出, 它实际为沃尔德(Wald)检验的置信区间的一种调整, 也称为 Agresti-Coull 置信区间。

2.3 Jeffreys 检验的方法及 p 的置信区间^[12]

这是一种等尾 Bayes 置信区间: 设 $B(1-\alpha; a, b)$ 表示 Beta 分布的 $1-\alpha$ 分位数, 则 Jeffreys 置信区间上下限为:

$$L_J(x) = B\left(\frac{\alpha}{2}; n\hat{p} + 0.5, n(1-\hat{p}) + 0.5\right), U_J(x) = B\left(1 - \frac{\alpha}{2}; n\hat{p} + 0.5, n(1-\hat{p}) + 0.5\right), \quad (5)$$

规定 $L_J(0) = 0, U_J(n) = 1$ 。该区间的端点值需要借助软件做数值计算。

修正的 Jeffreys p 置信区间, 当 $\hat{p} \rightarrow 0$ 或 $\hat{p} \rightarrow 1$ 时, 取

$$L_J(x) = 1 - \sqrt{\frac{\alpha}{2}}, U_J(x) = \sqrt{\frac{\alpha}{2}}. \quad (6)$$

2.4 近似检验的方法及 p 的置信区间^[13]

令 $c = \frac{\text{var}(V(\hat{p}, n))}{2E(V(\hat{p}, n))}$, $v = \frac{2[E(V(\hat{p}, n))]^2}{\text{var}(V(\hat{p}, n))}$, 把 $\text{var}(V(\hat{p}, n))$ 近似展开: $\text{var}(V(\hat{p}, n)) \approx \text{var}(n\hat{p})/n^4 +$

$\text{var}((n\hat{p})^2)/n^6 - 2\text{Cov}(n\hat{p}, (n\hat{p})^2)/n^5 = (p-p^2)/n^3 + [p + (6n-7)p^2 + 4(n-1)(n-3)p^3 - 2(n-1)(2n-3)p^4]/n^5 - 2[p + (2n-3)p^2 - 2(n-1)p^3]/n^4$, 结果记之为 $\Omega(\hat{p}, n)$ 。

假设 $\hat{p}, V(\hat{p}, n)$ 是近似独立的, 依据 \hat{p} 的渐进正态性, 则可得 $\frac{\hat{p}-p}{\sqrt{V(\hat{p}, n)}} = \frac{(\hat{p}-p)\sqrt{E(V(\hat{p}, n))}}{\sqrt{\frac{V(\hat{p}, n)}{cv}}}$, 这就是近似

T 检验, 自由度 $v \approx \frac{2V(\hat{p}, n)^2}{\Omega(\hat{p}, n)}$ 。相应 p 的置信区间为:

$$\hat{p} \mp t_{v, \frac{\alpha}{2}} \sqrt{V(\hat{p}, n)}. \quad (7)$$

2.5 Clopper-Pearson 检验方法及 p 的置信区间^[10]

Clopper-Pearson 检验方法计算分布的每个尾部超过 $\frac{\alpha}{2}$ 的 p 值, 区间的上限和下限通过求解 p_0 的方程组得

$$\begin{cases} \sum_{k=n\hat{p}}^n \binom{n}{k} p_0^k (1-p_0)^{n-k} = \frac{\alpha}{2} \\ \sum_{k=0}^{n\hat{p}} \binom{n}{k} p_0^k (1-p_0)^{n-k} = \frac{\alpha}{2} \end{cases}, \text{ 关于 Clopper-Pearson 检验比例参数的置信区间, 当 } \hat{p}=0 \text{ 时, 下限为 } 0; \text{ 当 } \hat{p}=1 \text{ 时,}$$

上限为 1。当 $n\hat{p}=1, 2, \dots, n-1$ 时, 根据二项分布之和与 beta 函数之间的联系以及相应的 beta 累积分布函数和 F 分布, 相应置信区间等于

$$\left[1 + \frac{n(1-\hat{p})+1}{n\hat{p}F_{2n\hat{p}, 2(n(1-\hat{p})+1)}\left(1-\frac{\alpha}{2}\right)} \right]^{-1} < p < \left[1 + \frac{n(1-\hat{p})}{(n\hat{p}+1)F_{2(n\hat{p}+1), 2(n(1-\hat{p}))}\left(\frac{\alpha}{2}\right)} \right]^{-1}, \quad (8)$$

其中 $F_{a,b}(c)$ 表示自由度为 a 和 b 的 F 分布的第 $1-c$ 个分位数。

理论上讲, Clopper-Pearson 方法近乎完美, 故也称之为精确 (CP) 方法。然而, 由于离散性的存在, 置信区间对于任何 p 的实际涵盖概率都至少不小于名义上的置信水平, 而且可能会比后者大很多。在本文模拟中 will 看到, Clopper-Pearson 检验 $100(1-\alpha)\%$ 的置信区间覆盖率比名义上设定的置信水平要高。

2.6 其他方法的置信区间

Mid-P^[14] 置信区间可以视为 Jeffreys p 的置信区间的改进, 上下限为:

$$\begin{cases} L_{mid-p}(x) = 0.5B\left(1-\frac{\alpha}{2}; n\hat{p}, n(1-\hat{p})+1\right) + 0.5B\left(1-\frac{\alpha}{2}; n\hat{p}+1, n(1-\hat{p})\right), \\ U_{mid-p}(x) = 0.5B\left(\frac{\alpha}{2}; n\hat{p}, n(1-\hat{p})+1\right) + 0.5B\left(\frac{\alpha}{2}; n\hat{p}+1, n(1-\hat{p})\right). \end{cases} \quad (9)$$

$$\text{Hall}^{[14]} \text{ 置信区间} \quad \hat{p} + \left(\frac{u_{1-\frac{\alpha}{2}}^2 + 1}{3} + \frac{1}{6}\right) \frac{1-2\hat{p}}{n} \mp u_{1-\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n-1}}. \quad (10)$$

$$\text{Kott-Liu}^{[15]} \text{ 置信区间} \quad \hat{p} + \left(\frac{u_{1-\frac{\alpha}{2}}^2 + 1}{3} + \frac{1}{6}\right) \frac{1-2\hat{p}}{n} \mp \sqrt{u_{1-\frac{\alpha}{2}}^2 \frac{\hat{p}(1-\hat{p})}{n-1} + \left(\frac{u_{1-\frac{\alpha}{2}}^2 + 1}{3} + \frac{1}{6}\right)^2}. \quad (11)$$

令 $\eta = \frac{u_{1-\frac{\alpha}{2}}^2 + 1}{3} + \frac{1}{6}$, $\bar{p} = \frac{\hat{p} + \frac{\eta}{n}}{1 + 2\frac{\eta}{n}}$, $\gamma_1 = -\frac{13}{18}u_{1-\frac{\alpha}{2}}^2 - \frac{17}{18}$, $\gamma_2 = \frac{1}{18}u_{1-\frac{\alpha}{2}}^2 + \frac{7}{36}$, 则 Cai^[16] 置信区间为:

$$\bar{p} \mp \frac{u_{1-\frac{\alpha}{2}}}{\sqrt{n}} \sqrt{\hat{p}(1-\hat{p}) + \frac{\gamma_1 \hat{p}(1-\hat{p}) + \gamma_2}{n}}. \quad (12)$$

反正弦变换的置信区间^[10] 为:

$$\left(\sin^2 \left(\arcsin(\sqrt{\hat{p}}) - \frac{u_{1-\frac{\alpha}{2}}}{2\sqrt{n}} \right), \sin^2 \left(\arcsin(\sqrt{\hat{p}}) + \frac{u_{1-\frac{\alpha}{2}}}{2\sqrt{n}} \right) \right), \quad (13)$$

该方法的 $\text{Var}(\arcsin(\sqrt{\hat{p}})) \approx \frac{1}{4n}$, 比较稳定。

记 $\hat{\lambda} = \ln\left(\frac{\hat{p}}{1-\hat{p}}\right), \hat{V} = \frac{1}{n\hat{p}(1-\hat{p})}$, 则 Logit 置信区间^[10]为:

$$\left(\frac{e^{\hat{\lambda}-u_{1-\frac{\alpha}{2}}\sqrt{\hat{V}}}}{1+e^{\hat{\lambda}-u_{1-\frac{\alpha}{2}}\sqrt{\hat{V}}}}, \frac{e^{\hat{\lambda}+u_{1-\frac{\alpha}{2}}\sqrt{\hat{V}}}}{1+e^{\hat{\lambda}+u_{1-\frac{\alpha}{2}}\sqrt{\hat{V}}}}\right) \quad (14)$$

Poisson 置信区间^[15], 当 n 很大而 p 很小时, 用 Poisson 分布近似二项分布, 置信区间为:

$$\left(\chi_{\frac{\alpha}{2}}^2(2n\hat{p})/(2n), \chi_{1-\frac{\alpha}{2}}^2(2(n\hat{p}+1))/(2n)\right) \quad (15)$$

似然比 p 置信区间^[17], 在原假设 $H_0: p=p_0$ 下, $-2\ln \frac{p_0^{n\hat{p}}(1-p_0)^{n(1-\hat{p})}}{\hat{p}^{n\hat{p}}(1-\hat{p})^{n(1-\hat{p})}} \leq u_{1-\frac{\alpha}{2}}^2$ 。极限分布为:

$$2\left[n\hat{p}\ln\left(\frac{\hat{p}}{p_0}\right) + n(1-\hat{p})\ln\left(\frac{1-\hat{p}}{1-p_0}\right)\right] \sim \chi^2(1), \quad (16)$$

求分位数获得置信区间 (L_{LR}, U_{LR}) 。

3 各种置信区间的模拟研究

统计上, 一个置信区间的覆盖率是它包含感兴趣真值的次数占总次数的比例。通常 $100(1-\alpha)\%$ 的置信区间的置信水平 $1-\alpha$ 只是名义覆盖率, 覆盖率是区间包含感兴趣真值的概率, 可以通过试验估计。

模拟研究基于如下 3 个命题, 证明从略。

命题 1 若随机数 X 服从均匀分布, 记作 $X \sim U(0, 1)$, 则 $P(X \leq p) = p$ 。

命题 2 若 $X_i (i=1, 2, 3, \dots, n)$ 相互独立且服从 0-1 分布, 即 $X_i \sim \begin{pmatrix} 0 & 1 \\ 1-p & p \end{pmatrix}$, 则 $X = \sum_{i=1}^n X_i$ 服从二项分布。

命题 3 设 $CI = [L(n\hat{p}, n), U(n\hat{p}, n)]$ 是二项分布比例参数 p 的一个置信区间, 则其覆盖率与平均宽度分别是 $C_n(p) = \sum_{n\hat{p}=0}^n I(n\hat{p}, p) C_n^{n\hat{p}} p^{n\hat{p}} (1-p)^{n(1-\hat{p})}, EW_n(p) = \sum_{n\hat{p}=0}^n (U(n\hat{p}, n) - L(n\hat{p}, n)) C_n^{n\hat{p}} p^{n\hat{p}} (1-p)^{n(1-\hat{p})}$, 其中 $I(n\hat{p}, p) = \begin{cases} 1, & p \in CI \\ 0, & p \notin CI \end{cases}$ 。

试验方案 1 (n, p 都变动), $X \sim B(n, p)$ 。取真实的 p 值从 0.01~0.99, 步长 0.01; 参数 n 分别取值 5, 25, 100。每个置信区间重复 10 000 次, 比较各种方法的性能, 从中剔除在参数 $n=5, 25, 100$ 时平均覆盖率都显著低于名义水平的 Arc sine, Mid-P, CP, Modified Jeffreys 和 logit 置信区间, 及显著低于名义水平的 Wilson correction, Poisson 置信区间后, 汇总在表 1~3。

试验方案 2 (仅 p 变动), $X \sim B(50, p)$ 。取真实的 p 值从 0~0.3, 步长 0.05, 做出覆盖率函数随 p 变化的图像, 从表中选择出性能较好的 4 种方法的置信区间的表现见图 1 所示。

表 1 小样本情形各比例参数置信区间表现

Tab. 1 Performance of small sample CI for proportion parameter

样本数量	方法	95% 置信区间平均覆盖率	95% 置信区间平均宽度
5	Wald	0.647 3	0.248 2
5	Plus4/ Agresti-Coull	0.895 9	0.075 7
5	Score/Wilson	0.955 3	0.028 1
5	Score correction	0.967 0	0.139 9
5	CP	0.667 5	0.265 0
5	Jeffreys	0.956 1	0.026 0
5	Modified Wilson	0.972 8	0.025 2
5	Hall	0.671 9	0.262 5
5	Kott-Liu	0.997 4	0.003 2
5	Cai	0.961 1	0.024 0
5	T-approximation	0.971 6	0.022 4

表 2 中等大小样本情形各比例参数置信区间表现

Tab. 2 Performance of medium size sample CI for proportion parameter

样本数量	方法	95%置信区间平均覆盖率	95%置信区间平均宽度
25	Wald	0.869 9	0.141 2
25	Plus4/Agresti-Coull	0.939 0	0.033 4
25	Score/Wilson	0.951 8	0.015 1
25	Score correction	0.974 1	0.008 5
25	CP	0.909 4	0.156 6
25	Jeffreys	0.951 2	0.018 1
25	Modified Wilson	0.967 8	0.012 0
25	Hall	0.893 3	0.181 7
25	Kott-Liu	0.967 7	0.012 8
25	Cai	0.953 9	0.016 7
25	T-approximation	0.964 7	0.012 9

表 3 大样本情形各比例参数置信区间表现

Tab. 3 Performance of larger sample CI for proportion parameter

样本数量	方法	95%置信区间平均覆盖率	95%置信区间平均宽度
100	Wald	0.927 5	0.050 1
100	Plus4/Agresti-Coull	0.947 3	0.014 2
100	Score/Wilson	0.949 5	0.009 8
100	Score correction	0.965 9	0.008 2
100	CP	0.954 0	0.052 4
100	Jeffreys	0.950 0	0.010 0
100	Modified Wilson	0.964 3	0.010 3
100	Hall	0.939 9	0.059 8
100	Kott-Liu	0.952 6	0.010 1
100	Cai	0.950 2	0.010 0
100	T-approximation	0.958 0	0.012 5
100	Log-likelihood-ratio	0.948 0	0.214 0

从图 1 可看出,当 p 在 0.05 附近时,比较好的是 Score 与 P4 置信区间,而 T -近似置信区间的覆盖率对较小的 p 值偏低。

试验方案 3(仅 p 变动), $X \sim B(5, p)$ 。模拟 1 000 次,比较 p 靠近两端时,几种修正方法置信区间的覆盖率,见表 4 所示。

下面再对 Plus4/Agresti-Coull, Score/Wilson 与精确(CP)方法的置信区间平均区间长度随参数 p 的变化做个比较,见图 2 所示。

各置信区间 CI 的期望宽度在 $p=0.5$ 最宽,在 $p \rightarrow 0$ 或 $p \rightarrow 1$ 时,退化为一点了。由于离散性的存在,尽管管理

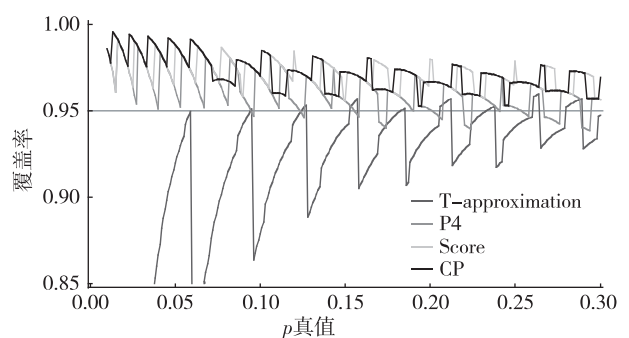


图 1 4 种较优方法置信区间的覆盖率曲线

Fig. 1 Coverage curves of CIs obtained by 4 optimal methods

论上 Clopper-Pearson 方法近乎完美,但是它的 CI 的期望宽度比 Plus4/Agresti-Coull 置信区间、Score/Wilson 置信区间的长。

表 4 几种修正置信区间在比例参数 P 为极端情形的表现

Tab. 4 Performances of modified CIs for extreme proportion parameter P

方法	小 P 值				大 P 值			
	$P=0$	$P=0.01$	$P=0.02$	$P=0.05$	$P=0.8$	$P=0.9$	$P=0.99$	$P=1$
Mid-P	0	0	0	0.214 0	0.606 0	0.422 0	0	0
Hall	0	0.053 0	0.096 0	0.233 0	0.664 0	0.423 0	0.050 0	0
Kott-Liu	0	1.000 0	1.000 0	0.998 0	1.000 0	1.000 0	1.000 0	1.000 0
Cai	1.000 0	1.000 0	0.996 0	0.979 0	0.939 0	0.906 0	0.999 0	1.000 0

4 讨论

从模拟结果看,总体上 Score/Wilson 置信区间最好。WilsonScore/Wilson, Agresti-Coull 置信区间覆盖率较好,但是在 $p \rightarrow 0$ 或 $p \rightarrow 1$ 时差,Jeffreys 区间期望长度最短。Score/Wilson 置信区间不仅长度最短,而且经常对准了正确的覆盖率。精确(CP)方法区间则是过度覆盖。而且精确(CP)方法区间长度较长除非真实比例接近 1 或 0,此时它最短,但几乎是退化的一个点。精确(CP)方法区间并非理论上期望的那样好。

在样本量 $n \leq 40$ 时,Score/Wilson、Jeffreys 置信区间较好。对大的 n , Score/Wilson、Jeffreys 和 Agresti-Coull 区间仍具可比性,不过 Agresti-Coull 区间的表达最为简单。

Wald, Arcsine 置信区间在覆盖率上有大的偏差与振荡 Poisson 有很大的覆盖率。

小样本情形,当 $p \rightarrow 0$, Cai 和 Kott-Liu 置信区间方法表现较好。当 $p \rightarrow 1$, Cai 和 Kott-Liu 置信区间较好, Mid-P 和 Hal 置信区间很差。

当点估计 $\hat{p}=0$ 时,许多置信区间的下限估计值为负值。对此,张学新等人^[18]在假定 n 次试验中都不成功的概率不低于 $\alpha\%$ 下,给出比例参数 $100(1-\alpha)\%$ 的置信区间 $\left(0, -\frac{1}{n} \ln \alpha\right)$ 。

参考文献:

- [1] 刘镇瑜,王军,王凌艳,等. 二项分布参数 p 置信区间评估算法分析[J]. 无线电工程,2013,43(12):13-16.
Liu Z Y, Wang J, Wang L Y, et al. Analysis of a binomial parameter p confidence interval evaluation algorithm[J]. Radio Engineering,2013,43(12):13-16.
- [2] 姚励. 二项分布的区间估计分析[D]. 合肥:安徽大学,2008.
Yao L. Interval estimation of the binomial distribution analysis[D]. Hefei: Anhui University,2008.
- [3] 李丹. 二项分布参数置信区间的比较[D]. 长春:东北师范大学,2008.
Li D. A comparison of confidence intervals for the binomial parameter[D]. Changchun: Northeast Normal University, 2008.
- [4] Zhou X H, Li C M, Yang Z. Improving interval estimation of binomial proportions[J]. Philos Transact A: Math Phys Eng Sci,2008,366(1874):2405-2418.
- [5] Agresti A. 分类数据分析[M]. 齐亚强,译. 重庆:重庆大学出版社,2012.
Agresti A. Categorical data analysis[M]. Qi Y Q, translation. Chongqing: Chongqing University Press,2012.
- [6] Agresti A, Coull B. Approximate better than 'exact' for CIs for binomial parameters[J]. American Statistician, 2008,52:119-126.
- [7] Agresti A, Min Y. On small-sample confidence intervals for parameters in discrete distributions[J]. Biometrics, 2001,

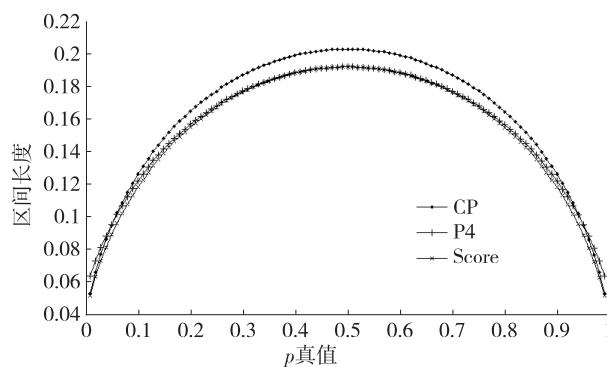


图 2 3 种比例置信区间的平均长度 ($n=100$)

Fig. 2 Average lengths of 3 CIs for proportion parameter ($n = 100$)

- 57:963-971.
- [8] Agresti A, Gottard A. Nonconservative exact small-sample inference for discrete data[J]. *Comput Statist & Data Anal*, 2007, 51:6447-6458.
- [9] Agresti A, Ryu E. Pseudo-score inference for parameters in discrete statistical models[J]. *Biometrika*, 2010, 97: 215-222.
- [10] Brown L D, Cai T T, Dasgupta A. Interval estimation for a binomial proportion[J]. *Statistical Science*, 2001, 16(2): 101-133.
- [11] Bera A K, Biliyas Y. Rao's score, Neyman's $C(\alpha)$ and Silvey's LM tests[J]. *Journal of Statistical Planning and Inference*, 2001, 97:9-44.
- [12] Cai T T. One-sided confidence intervals in discrete distributions[J]. *Journal of Statistical Planning and Inference*, 2005, 131:63-88.
- [13] Pan W. Approximate confidence intervals for one proportion and difference of two proportions[J]. *Computational Statistics and Data Analysis*, 2001, 40(1):143-157.
- [14] Hall P. Improving the normal approximation when constructing one-sided confidence intervals for binomial or Poisson parameters[J]. *Biometrika*, 1982, 69(3):647-652.
- [15] Liu Y, Kott P. Evaluating alternative one-sided coverage intervals for an extreme binomial proportion[R]. Zürich: Socioeconomic Institute, University of Zurich, 2007.
- [16] Cai T. One-sided confidence intervals in discrete distributions[J]. *Journal of Statistical Planning and Inference*, 2004, 131(1):63-88.
- [17] Wilks S S. The large-sample distribution of the likelihood ratio for testing composite hypotheses[J]. *Ann Math Statistics*, 1938, 9(1):60-62.
- [18] 张学新, 林先安. 关于二项分布比例参数置信区间教学的一点注记[J]. *湖北工程学院学报:自然科学版*, 2015, 35(3):94-97.
- Zhang X X, Lin X A. A note of teaching binomial proportion confidence interval[J]. *Journal of Hubei Engineering University; Natural Science Edition*, 2015, 35(3):94-97.

Comparisons of Confidence Intervals for the Binomial Proportion with Extremely Frequency

ZHANG Xuexin

(School of Mathematics and Statistics, Hubei Engineering University, Xiaogan Hubei 432000, China)

Abstract: To select confidence intervals for the binomial proportion in extreme cases that its parameter point estimation value is close to 0 or 1, this paper carried on some simulation comparisons of coverage probability and average length of confidence intervals amongst Wald, Plus 4, Wilson, CP, Jeffreys, Hall, Kott-Liu, Cai, T-approximation, etc. more than a dozen methods. Also performances of their coverage probabilities with the proportion parameter p in Binomial distribution changing were analyzed. The results showed that Wilson Score-test-based confidence interval is optimal as a whole, when the point estimator of p , is small, or large, Cai, Kott-Liu confidence intervals are better in terms of coverage probabilities of their confidence intervals.

Key words: binomial proportion p ; point estimator of p being extreme value; confidence interval; coverage probability; expected length

(责任编辑 游中胜)