多项分布结果的精确最小体积置信集相交问题

Heguang Lin[1] Binhao Chen[2] MengzeLi[3] Daniel Pimentel-Alarcón[3] MatthewL.Malloy[3]heguangl@scripps.edu binhao_chen@brown.edu mli562@wisc.edu pimentelalar@wisc.edu matthew.malloy@wisc.edu

arXiv:2601.18145v1 [stat.ML] 26 Jan 2026

摘要

置信集的计算是数据科学和机器学习的核心，为 A/B 测试提供了基础，也是强化学习算法运行与分析的基石 [1]。在所有有效的多项分布参数置信集中，最小体积置信集（MVCs）是最优的，因为它们能使平均体积最小化 [2]，但它们被定义为精确 p值的水平集，而该值不连续且难以计算。本文不试图直接刻画 MVC 的几何形状，而是研究一个基础且具有实际意义的判定问题：给定两个观测到的多项分布结果，能否认证它们对应的 MVC 是否相交？我们针对这一相交问题提出了一种可认证且考虑容差的算法。该方法利用似然排序在对数优势坐标中诱导半空间约束这一事实，从而能够对参数空间进行自适应几何划分，并计算每个单元上 p值的下界和上界。对于三个类别，这一方法构成了一种高效且有严格可靠性保证的算法，可认证相交、认证不相交，或在判定结果处于预设容差范围内时返回不确定结果。我们还进一步展示了该方法如何扩展至高维。结果表明，尽管 MVC 的几何形状不规则，仍可为 A/B 测试中的核心任务提供可靠、可认证的判定程序。

I. 引言

置信集、置信区域和置信区间是数据科学、统计推断和机器学习中的基础工具，用于刻画模型参数的一系列合理取值。为简化计算和分析，大多数构造置信集的方法都依赖于近似或界限，而这些近似或界限在小样本情形下较为宽松 [2]–[4]。尽管这些方法通常在渐近意义下是最优的，但在小样本情形下使用更紧的置信集，可以降低 A/B 测试、强化学习算法以及应用数据科学中的其他问题的样本复杂度 [5]–[9]。

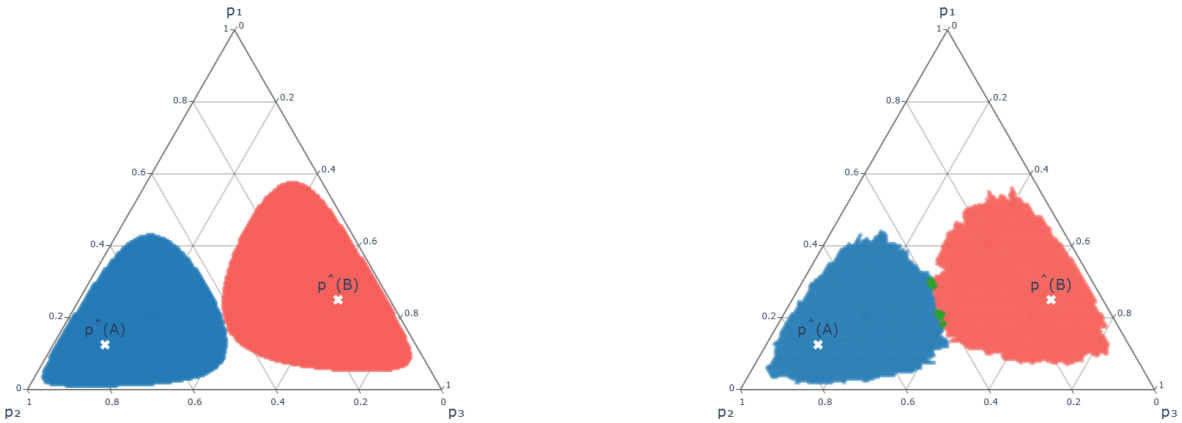
对于分类数据，构造多项分布参数的紧置信集是一个长期受到研究的问题。近期研究 [2], 基于假设检验与置信集之间的经典对偶关系 [10]–[12], 表明，将精确 p值的水平集定义为置信集，能够在对经验结果取平均时达到最小体积最优性。这些最小体积置信集 （MVC）因此在严格意义上具有统计最优性。

尽管具有最优性，最小体积置信集在计算上仍面临重大挑战。对单个参数值进行成员资格检验，需要计算一个精确的 p值，该值定义为对多项分布结果求和；这一任务的计算开销会随着样本量和维度 [13] 的增加而迅速攀升。此外，由于 p值是参数的一个不连续函数，最小体积置信集呈现出高度不规则的几何形态：它们通常是非凸的，可能不连通，也无法用简单的几何特征加以刻画 [14]。这些特点使得即便是基本的推断任务，也难以直接使用最小体积置信集。

本文聚焦于一项此类任务，它是 A/B 测试中的核心问题：给定两个观测到的多项分布结果，判断其对应的最小体积置信集是否相交。若两者相交，则可确认存在一个参数值，能在给定置信水平下同时与两个结果相容；若两者不相交，则意味着底层分布之间存在统计显著差异。基于网格化或连续松弛的朴素方法，由于最小体积置信集具有不连续且碎片化的结构，无法提供保证。图 1 通过一个包含三个类别的简单多项分布示例说明了这一难题。图 1a 所示的渐近卡方（似然比/Wilks）置信区域 [15], 彼此不相交，因此会提示两个经验结果之间存在统计显著差异。相比之下，图 1b 所示的、通过 p 值判定定义的精确最小体积置信集却彼此相交，从而正确确认存在一个参数值，能在规定的置信水平下同时与两个观测值相容。这种差异凸显了渐近近似在小样本情形下的局限性，并说明了为何需要精确且算法上可处理的方法来判断最小体积置信集是否相交。

我们的贡献是针对这一相交问题提出一种可认证决策过程。我们不尝试显式计算最小体积置信集，而是利用一个关键的结构性质：多项分布结果之间的似然比较在对数优势坐标中会转化为线性不等式。该观察使我们能够将参数空间划分为几何单元，并在这些单元上计算精确 p 值的紧致下界和上界。通过自适应细化，我们获得了一种算法，可以认证相交、认证不相交，或在决策处于预设容差范围内时报告不确定。我们首先详细介绍三个类别情形下的方法，然后将其扩展到一般的 k 维度。

[1]斯克里普斯研究所，加利福尼亚州 拉霍亚，加利福尼亚州，美国。 [2]布朗大学，美国罗得岛州普罗维登斯。 [3]威斯康星大学麦迪逊分校，美国威斯康星州麦迪逊 A.



(a) 卡方（似然比检验/Wilks）置信集。　　(b) 最小体积置信集。

图 1：单纯形 Δ_3 上渐近置信集与精确置信集的比较。卡方近似（左图）产生互不相交的区域，而最小体积置信集（右图）则存在交集。设定：$n = 8$，$k = 3$ 和 $\alpha = 0.17$，两个观测结果分别为 $A = [1, 6, 1]$ 和 $B = [2, 1, 5]$.

II. 符号与基本定义

设 $X_1, \ldots, X_n$ 为某个分类随机变量的独立同分布样本，该变量从有限个类别 $\mathcal{X} = \{x_1, \ldots, x_k\}$中取 k 个可能值之一。经验分布 $\widehat{\boldsymbol{p}}$ 是 $X_1, \ldots, X_n$中 $\mathcal{X}$每个元素出现的相对比例，即 $\widehat{\boldsymbol{p}} = [n_1/n, \ldots, n_k/n]$，其中 $n_i = \sum_{j=1}^{n} \mathbb{1}_{\{X_j = x_i\}}$。令 $\Delta_{k,n}$ 表示基于 n 个样本、涵盖 k 个类别的离散单纯形：

$$\Delta_{k,n} := \left\{ \widehat{\boldsymbol{p}} \in \{0,\ 1/n,\ 2/n,\ \ldots,\ 1\}^k \ :\ \sum_{i=1}^{k} \widehat{p}_i = 1 \right\},$$

并定义 $m = |\Delta_{k,\,n}| = \binom{n+k-1}{k-1}$。将连续单纯形记为 $\Delta_k = \left\{ \boldsymbol{p} \in [0,1]^k : \ \sum_i p_i = 1 \right\}$。我们用 $\mathcal{P}(\Delta_{k,\,n})$ 表示 $\Delta_{k,\,n}$ 的幂集，并用 $\mathcal{P}(\Delta_k)$ 表示 Δ_k 的勒贝格可测子集集合。对于任意 $\mathcal{S} \subset \Delta_{k,\,n}$，我们将 $\mathbb{P}_{\boldsymbol{p}}(\mathcal{S})$ 简记为 $\mathbb{P}_{\boldsymbol{p}}(\{X \in \mathcal{X}^n : \ \widehat{\boldsymbol{p}}(X) \in \mathcal{S}\})$，其中 $\mathbb{P}_{\boldsymbol{p}}(\cdot)$ 表示多项分布参数 $\boldsymbol{p} \in \Delta_k$ 下的概率测度。最后，令 $\mathbb{R}^{k+} := \{\boldsymbol{x} \in \mathbb{R}^k : \ x_i > 0,\ i = 1、\ldots、k\}$。

定义 1.（置信集）设 $\mathcal{C}_\alpha(\widehat{\boldsymbol{p}}) : \Delta_{k,n} \to \mathcal{P}(\Delta_k)$ 为一个集合值函数，它将观测到的经验分布$\widehat{\boldsymbol{p}}$ 映射到 k-单纯形的一个子集。$\mathcal{C}_\alpha(\widehat{\boldsymbol{p}})$ 是置信水平为 $1 - \alpha$ 的置信集，当且仅当满足以下条件：

$$\sup_{\boldsymbol{p} \in \Delta_k} \mathbb{P}_{\boldsymbol{p}}\left(\boldsymbol{p} \notin \mathcal{C}_\alpha(\widehat{\boldsymbol{p}})\right) \ \le \ \alpha. \tag{1}$$

定义 2.（p-值）固定一个结果 $\widehat{\boldsymbol{p}}$。原假设 $\boldsymbol{p}$的函数 p-值定义为：

$$\rho_{\widehat{\boldsymbol{p}}}(\boldsymbol{p}) \ = \sum_{\widehat{\boldsymbol{q}} \in \Delta_{k,n} : \mathbb{P}_{\boldsymbol{p}}(\widehat{\boldsymbol{q}}) \le \mathbb{P}_{\boldsymbol{p}}(\widehat{\boldsymbol{p}})} \mathbb{P}_{\boldsymbol{p}}\left(\widehat{\boldsymbol{q}}\right). \tag{2}$$

对于固定的结果 $\widehat{\boldsymbol{p}}$，为简便起见，我们记作 $\rho(\boldsymbol{p})$。

命题 1.（最小体积置信集（*MVCs*）*[2]*）。*MVCs* 定义为

$$\mathcal{C}^\star_\alpha(\widehat{\boldsymbol{p}}) \ := \ \left\{\boldsymbol{p} \in \Delta_k \ : \ \rho_{\widehat{\boldsymbol{p}}}(\boldsymbol{p}) \ge \alpha\right\},$$

并满足

$$\sum_{\widehat{\boldsymbol{p}} \in \Delta_{k,n}} \mathrm{vol}\left(\mathcal{C}^\star_\alpha(\widehat{\boldsymbol{p}})\right) \le \sum_{\widehat{\boldsymbol{p}} \in \Delta_{k,n}} \mathrm{vol}\left(\mathcal{C}_\alpha(\widehat{\boldsymbol{p}})\right)$$

对于任意 c 置信集 $\mathcal{C}_\alpha(\cdot)$；此处 $\mathrm{vol}(\cdot)$ 表示勒贝格测度。证明见 *[2]*　　.

III. A/B 测试相交问题

本文专门关注一项由 A/B 测试引出的计算任务：给定两个观测到的多项分布结果，判断它们的最小体积置信集（MVC）是否相交。

设 $\widehat{\boldsymbol{p}}^{(A)},\ \widehat{\boldsymbol{p}}^{(B)} \in \Delta_k,\ n$ 表示来自 n 个独立同分布分类样本、涵盖 k 个类别的两个观测经验分布。对于置信水平 $1-\alpha$，回顾第 II 节，与 $\widehat{\boldsymbol{p}}$ 相关联的最小体积置信集为

$$\mathcal{C}_\alpha^\star(\widehat{\boldsymbol{p}}) \;=\; \{\boldsymbol{p} \in \Delta_k :\ \rho_{\widehat{\boldsymbol{p}}}(\boldsymbol{p}) \geq \alpha\}.$$

我们的目标是判断

$$\mathcal{C}_\alpha^\star(\widehat{\boldsymbol{p}}^{(A)}) \;\cap\; \mathcal{C}_\alpha^\star(\widehat{\boldsymbol{p}}^{(B)}) \;\neq\; \emptyset. \tag{3}$$

等价地，定义 $\rho_A(\boldsymbol{p}) := \rho_{\widehat{\boldsymbol{p}}^{(A)}}(\boldsymbol{p})$ 和 $\rho_B(\boldsymbol{p}) := \rho_{\widehat{\boldsymbol{p}}^{(B)}}(\boldsymbol{p})$。那么，当且仅当以下条件成立时，式 (3) 成立

$$\max_{\boldsymbol{p}\in\Delta_k} \;\min\{\rho_A(\boldsymbol{p}), \rho_B(\boldsymbol{p})\} \;\geq\; \alpha.$$

我们强调，并不旨在计算或可视化类似 [14] 的最小体积置信集边界；我们只寻求对式 (3) 作出可靠的“是/否”判断。

在有限精度计算中，当最优值接近阈值 α 时，决策在数值上可能较为微妙。因此，我们引入决策裕度 $\tau > 0$，并以在 α 附近留有缓冲区的经认证决策为目标。

定义 3（鲁棒相交 决策）。 固定 $\alpha \in (0,1)$ 和 $\tau > 0$。如果存在……，我们称最小体积置信集稳健相交</*style*】【。$\boldsymbol{p} \in \Delta_k$ 使得

$$\rho_A(\boldsymbol{p}) \geq \alpha + \tau \qquad \textit{and} \qquad \rho_B(\boldsymbol{p}) \geq \alpha + \tau.$$

我们称这些最小体积置信集是鲁棒不相交 的，如果对于所有 $\boldsymbol{p} \in \Delta_k$,

$$\min\{\rho_A(\boldsymbol{p}), \rho_B(\boldsymbol{p})\} \leq \alpha - \tau.$$

我们的算法将返回以下结果之一

- 相交（认证鲁棒相交）
- 不相交（证明鲁棒不相交），或
- 不确定（当无法在所选容差范围内判定实例时）。

除 τ 外，算法还接受一个几何细化容差 $\varepsilon > 0$，用于控制自适应划分中单元直径的最大值。

IV. Certified Intersection Test for $k = 3$

我们首先针对 $k = 3$ 清晰地介绍该方法，此时单纯形是二维的，算法尤其有效。随后，第 V 节将此方法推广到任意 k。

A. 对数优势坐标与结果概率

对于 $k = 3$，我们将 $\boldsymbol{p} = (p_1, p_2, p_3) \in \Delta_3^\circ$，所有坐标均严格为正，因此对数优势坐标是有限的。将对数优势坐标 $(u, v) \in \mathbb{R}^2$ 定义为

$$u = \log\frac{p_1}{p_3}, \qquad v = \log\frac{p_2}{p_3},$$

具有逆映射

$$p_1(u,v) = \frac{e^u}{1+e^u+e^v}, \quad p_2(u,v) = \frac{e^v}{1+e^u+e^v}, \quad p_3(u,v) = \frac{1}{1+e^u+e^v}. \tag{4}$$

对于任意的 c ount 向量 $r = (r_1, r_2, r_3) \in \mathbb{N}^3$，其中 $r_1 + r_2 + r_3 = n$，定义多项式系数

$$\kappa(r) = \frac{n!}{r_1!\, r_2!\, r_3!}.$$

在参数 $\boldsymbol{p}(u, v)$ 下观察到 r 的概率为

$$\mathbb{P}_{\boldsymbol{p}(u,v)}(r) \;=\; \kappa(r)\, p_1(u,v)^{r_1}\, p_2(u,v)^{r_2}\, p_3(u,v)^{r_3}.$$

取对数并利用 (4) 可得

$$\log \mathbb{P}_{\boldsymbol{p}(u,v)}(r) = \log\kappa(r) \;+\; r_1 u + r_2 v \;-\; n\log(1+e^u+e^v). \tag{5}$$

存在零计数时，相交判定理解为允许 $\boldsymbol{p}$ 位于 Δ_k 的边界上。如第五节 E 所示，这可以归结为单纯形各个面上的有限个低维多项分布。

B. 似然排序半空间

固定一个 观测结果 $\hat{r} = (\hat{r}_1,\ \hat{r}_2,\ \hat{r}_3)$（等价地，$\widehat{\boldsymbol{p}} = \hat{r}/n \in \Delta_{3,\ n}$）。回想一下，p 值 i s

$$\rho_{\widehat{\boldsymbol{p}}}(\boldsymbol{p}) = \sum_{r:\ \mathbb{P}_{\boldsymbol{p}}(r) \le \mathbb{P}_{\boldsymbol{p}}(\hat{r})} \mathbb{P}_{\boldsymbol{p}}(r).$$

一个关键事实是，似然排序比较 $\mathbb{P}_{\boldsymbol{p}(u,v)}(r) \le \mathbb{P}_{\boldsymbol{p}(u,v)}(\hat{r})$ 对 $(u,\ v)$ 是线性的。

引理 1（(u, v)）**的半空间形式。**对于任意 r 以及固定的 $\hat{r}$，不等式 $\mathbb{P}_{\boldsymbol{p}(u,v)}(r) \le \mathbb{P}_{\boldsymbol{p}(u,v)}(\hat{r})$ 等价于

$$(r_1 - \hat{r}_1)u + (r_2 - \hat{r}_2)v \ \le\ \log\kappa(\hat{r}) - \log\kappa(r). \tag{6}$$

证明。将 r 对应的式子 (5) 减去 $\hat{r}$ 对应的式子。共有项 $-n\log(1 + e^u + e^v)$ 抵消。两边取指数即可得到 (6)。 □

因此，对于固定的 $\hat{r}$，每个 r 都在 $(u,\ v)$ 中定义一个半空间，而使 r 落在 p 值尾部（即其概率不大于 $\hat{r}$）的 $(u,\ v)$ 构成一个半空间边界。

C. 紧致搜索域

坐标 $(u,\ v)$ 的取值范围为 $\mathbb{R}^2$，因此我们将搜索限制在一个紧致区域内，而不会遗漏任何稳健决策。以下这个粗略界已足够。

引理 2（通用 p 值上界）。设 $m = |\Delta_{3,\ n}|$。对于任意观测到的 $\hat{r}$ 以及任意 $\boldsymbol{p} \in \Delta_{3}$，

$$\rho_{\widehat{\boldsymbol{p}}}(\boldsymbol{p}) \ \le\ m\,\mathbb{P}_{\boldsymbol{p}}(\hat{r}).$$

证明。根据定义，p 值中的尾部集合至多包含 m 个结果；尾部和中的每一项根据定义均为 $\le \mathbb{P}_{\boldsymbol{p}}(\hat{r})$。求和即可得证。详细证明见附录 VI-A。 □

根据引理 2，如果 $\mathbb{P}_{\boldsymbol{p}}(\hat{r}) < (\alpha - \tau)/m$，那么 $\rho_{\widehat{\boldsymbol{p}}}(\boldsymbol{p}) < \alpha - \tau$。因此，任何可能作为鲁棒相交见证的 $\boldsymbol{p}$（定义 3）都必须同时满足两者

$$\mathbb{P}_{\boldsymbol{p}}(\hat{r}^{(A)}) \ge (\alpha - \tau)/m \qquad \text{and} \qquad \mathbb{P}_{\boldsymbol{p}}(\hat{r}^{(B)}) \ge (\alpha - \tau)/m.$$

在 $(u,\ v)$ 坐标中，每个约束都是凹函数 $(u,\ v) \mapsto \log \mathbb{P}_{\boldsymbol{p}(u,v)}(\cdot)$ 的超水平集，因此是凸且有界的。令 $\mathcal{W} \subset \mathbb{R}^2$ 为任意一个包含这两个超水平集交集的紧凸集。实际应用中，只需取该交集周围一个简单的包围盒；下文的算法只要求 $\mathcal{W}$ 包含所有可能的鲁棒见证点。

D. 多项式概率的单元下界

我们将 $\mathcal{W}$ 划分为三角形。令 $T \subset \mathbb{R}^2$ 为一个顶点为 $w^{(1)},\ w^{(2)},\ w^{(3)} \in \mathcal{W}$ 的三角形，其中 $w^{(j)} = (u^{(j)},\ v^{(j)})$。

对于任意结果 r，定义 $f_r(u,\ v) := \log \mathbb{P}_{\boldsymbol{p}(u,v)}(r)$。根据 (5)，$f_r$ 关于 $(u,\ v)$ 是凹的，因为它是仿射函数减去 $n\log(1 + e^u + e^v)$（而 $\log(1 + e^u + e^v)$ 是凸的）。凹性意味着可以在 T 上为 $\mathbb{P}_{\boldsymbol{p}(u,v)}(r)$ 给出一个简单且可计算的下界。

引理 3 (凹对数概率的顶点最小值). 对于任意结果 r 以及任意顶点为 $\{w^{(j)}\}_{j=1}^{3}$ 的三角形 T，

$$\min_{(u,v)\in T} f_r(u,v) \ =\ \min_{j\in\{1,2,3\}} f_r(w^{(j)}).$$

因此，该量

$$\underline{P}_T(r) \ :=\ \min_{j\in\{1,2,3\}} \mathbb{P}_{\boldsymbol{p}(w^{(j)})}(r)$$

对于所有 $(u, v) \in T$，均满足 $\mathbb{P}_{\boldsymbol{p}(u,v)}(r) \ge \underline{P}_T(r)$。

证明。任何 $(u, v) \in T$ 都是顶点的凸组合。由凹性可得 $f_r(u, v) \ge \sum_j \lambda_j f_r(w^{(j)}) \ge \min_j f_r(w^{(j)})$。详情请参阅附录 VI-C。 □

E. 三角形上的经认证的 p 值界

固定一个观测结果 $\hat{r}$。对每个 r，定义线性函数

$$g_{r;\hat{r}}(u,v) := (r_1 - \hat{r}_1)u + (r_2 - \hat{r}_2)v - \big(\log\kappa(\hat{r}) - \log\kappa(r)\big).$$

根据引理 1，r 在 (u,v) 处属于 p 值尾部，当且仅当 $g_{r;\hat{r}}(u,v) \le 0$。由于 $g_{r;\hat{r}}$ 是线性的，只需检查顶点，即可验证结果是否属于三角形内的集合。

定义肯定属于尾部的 $\mathcal{I}_T(\hat{r})$ 和肯定不属于尾部的 $\mathcal{O}_T(\hat{r})$ 索引集，对应三角形 T：

$$\mathcal{I}_T(\hat{r}) := \{r:\ g_{r;\hat{r}}(w^{(j)}) \le 0 \text{ 对所有 } j=1,2,3\}, \qquad \mathcal{O}_T(\hat{r}) := \{r:\ g_{r;\hat{r}}(w^{(j)}) > 0 \text{ 对所有 } j=1,2,3\}.$$

所有其余结果在 T 上均存在歧义。

命题 2（三角形上的认证 p 值区间）。令 $\hat{r}$ 为固定值，并令 T 为一个三角形。定义

$$\underline{\rho}_T(\hat{r}) := \sum_{r\in\mathcal{I}_T(\hat{r})} \underline{P}_T(r), \qquad \overline{\rho}_T(\hat{r}) := 1 - \sum_{r\in\mathcal{O}_T(\hat{r})} \underline{P}_T(r), \tag{7}$$

其中 $P_{\underline{T}}(r)$ 来自引理 3。于是，对所有 $(u,v) \in T$，

$$\underline{\rho}_T(\hat{r}) \ \le\ \rho_{\widehat{\boldsymbol{p}}}(\boldsymbol{p}(u,v)) \ \le\ \overline{\rho}_T(\hat{r}).$$

证明。若 $r \in \mathcal{I}_T(\hat{r})$，则在所有顶点处都有 $g_{r;\hat{r}} \le 0$，因而在 T 上处处成立，所以对于所有 $(u,v) \in T$，r 都属于尾部。因此，p 值尾部和包含所有 $r \in \mathcal{I}_T(\hat{r})$，而引理 3 给出的下界表明，每个概率都至少为 $P_T(r)$，从而得到下界。

若 $r \in \mathcal{O}_T(\hat{r})$，则在所有顶点处都有 $g_{r;\hat{r}} > 0$，因而在 T 上处处成立，所以对于所有 $(u,v) \in T$，r 都不属于尾部。因此，p 值至多为不属于 $\mathcal{O}_T(\hat{r})$ 的结果的总概率质量。利用对于每个 $r \in \mathcal{O}_T(\hat{r})$，$\mathbb{P}_{\boldsymbol{p}(u,v)}(r) \ge P_T(r)$，并且 $\sum_r \mathbb{P}_{\boldsymbol{p}(u,v)}(r) = 1$，即可得到式（7）中的上界。 □

F. 三角形细分与经认证的相交检验

对于 A/B 测试，我们要求两个 p 值都大于 α。在三角形 T 上，定义

$$\underline{\rho}_{A,T} := \underline{\rho}_T(\hat{r}^{(A)}), \quad \overline{\rho}_{A,T} := \overline{\rho}_T(\hat{r}^{(A)}), \qquad \underline{\rho}_{B,T} := \underline{\rho}_T(\hat{r}^{(B)}), \quad \overline{\rho}_{B,T} := \overline{\rho}_T(\hat{r}^{(B)}),$$

以及最小值的相应界：

$$\underline{M}_T := \min\{\underline{\rho}_{A,T}, \underline{\rho}_{B,T}\}, \qquad \overline{M}_T := \min\{\overline{\rho}_{A,T}, \overline{\rho}_{B,T}\}.$$

那么：(i) 如果 $\underline{M}_T \ge \alpha + \tau$，$T$，则可认证鲁棒相交；(ii) 如果 $\overline{M_T} < \alpha - \tau$，$T$，则可将其舍弃（$T$ 中不存在鲁棒见证）；(iii) 否则，细分 T，除非它已经足够小。

我们用三角形在 $(u,\ v)$ 中的直径来衡量其大小：

$$\operatorname{diam}(T) := \max_{i,j\in\{1,2,3\}} \|w^{(i)} - w^{(j)}\|_2.$$

我们采用最长边二分法进行细分（任何形状正则细化方法均可）。到目前为止的讨论使我们能够为 $k=3$ 设计经认证的 MVC 相交检验算法 1。

G. 正确性保证

定理 1（算法 1 的可靠性）。令 $m = |\Delta_{3,\ n}|$ 并定义

$$S_A := \{(u,v) \in \mathbb{R}^2 : \mathbb{P}_{\boldsymbol{p}(u,v)}(\hat{r}^{(A)}) \ge (\alpha-\tau)/m\}, \quad S_B := \{(u,v) \in \mathbb{R}^2 : \mathbb{P}_{\boldsymbol{p}(u,v)}(\hat{r}^{(B)}) \ge (\alpha-\tau)/m\}.$$

假设紧致搜索域 $\mathcal{W}$ 满足 $\mathcal{W} \supseteq S_A \cap S_B$。则算法 1 在以下意义下是可靠的：

- 如果返回相交，则最小体积置信集鲁棒相交（定义 3）。
- 如果返回不相交，则最小体积置信集鲁棒不相交（定义 3）。

如果返回不确定，则真实决策在至少一个直径不超过 ε 的三角形上，距离阈值 τ 邻域以内。

证明。对于任意三角形 T，命题 2 对 T 中的所有点给出每个 p 值的界，因此 $M_T \le \min\{\rho_A, \rho_B\} \le \underline{M}_T$ 在 T 上。如果对于某个已处理的 T 有 $\underline{M}_T \ge \alpha + \tau$，那么对于所有 $\boldsymbol{p} \in T$，$\min\{\rho_A(\boldsymbol{p})\ \rho_B(\boldsymbol{p})\} \ge \alpha + \tau$，因此存在鲁棒见证，且最小体积置信集鲁棒相交。如果队列为空且没有未解决的三角形，那么对于 $\mathcal{W}$ 的（细化后）三角剖分中的每个三角形，我们有 $M_T < \alpha - \tau$，因此对于所有 $\boldsymbol{p} \in \mathcal{W}$，$\min\{\rho_A(\boldsymbol{p})\ \rho_B(\boldsymbol{p})\} < \alpha - \tau$。现在考虑任意坐标为（$u,\ v$）的 $\boldsymbol{p} \in \Delta_3 \notin \mathcal{W}$。由于 $\mathcal{W} \supseteq S_A \cap S_B$，我们有（$u,\ v$）$\notin S_A$ 或（$u,\ v$）$\notin S_B$。如果（$u,\ v$）$\notin S_A$，那么 $\mathbb{P}_{\boldsymbol{p}}(\hat{r}^{(A)}) < (\alpha-\tau)/m$，并且根据引理 2，$\rho_A(\boldsymbol{p}) < \alpha - \tau$。对于 B，同理可得。无论哪种情况，对于所有 $\boldsymbol{p} \notin \mathcal{W}$，$\min\{\rho_A(\boldsymbol{p})\ \rho_B(\boldsymbol{p})\} \le \alpha - \tau$ 均成立。结合关于 $\mathcal{W}$ 的界，即可证明在所有 Δ_3 上均具有鲁棒不相交性。

□

算法 1 针对 $k = 3$ 的经认证的 MVC 相交检验

要求： 观测结果 $\hat{r}^{(A)}, \hat{r}^{(B)} \in \mathbb{N}^3$，总和为 n，显著性水平为 $\alpha \in (0, 1)$，决策裕度 $\tau > 0$，细化容差 $\varepsilon > 0$，紧致搜索域 $\mathcal{W} \subset \mathbb{R}^2$。

确保： 相交 ，不相交 ，或不确定。

初始化 $\mathcal{T}$ 的三角剖分 $\mathcal{W}$（例如，将一个包围盒分成两个三角形）。
初始化一个队列 $\mathcal{Q} \leftarrow \mathcal{T}$。
当 $\mathcal{Q}$ 不为空**时**
    从 $\mathcal{Q}$ 中取出一个三角形 T。
    计算（$\rho_{-A,T}$ $\overline{\rho}_{A,T}$）和（$\rho_{-B,T}$ $\overline{\rho}_{B,T}$）根据命题 2。
    集合 $\underline{M}_T \leftarrow \min\{\underline{\rho}_{A,T}\ \underline{\rho}_{B,T}\}$ 和 $\overline{M}_T \leftarrow \min\{\overline{\rho}_{A,}{}^{T,}\ {}^{\rho_B}{}_{,T}\}$.
    if $\underline{M}_T > \alpha + \tau$ **然后**
        返回 相交。
    否则如果 $M_T < \alpha - \tau$ **则**
        继续剪枝 T
    否则如果 直径$(T) \le \varepsilon$ **则**
        将 T 标记为未解决。
    else
        将 T 二分为两个三角形 T_1, T_2（例如，最长边二分）。
        将 T_1, T_2 推入 $\mathcal{Q}$。
    结束 if
结束 while
如果不存在未解决的三角形**则**
    返回 不相交。
否则
    返回 不确定。
结束 if

H. 计算复杂度

朴素的逐三角形成本是 $O(m)$ 个结果乘以一个较小的常数，其中 $m = |\Delta_{3,\ n}| = \binom{n+2}{2} = O(n^2)$。在实践中，可以通过以下方式获得显著的加速：(i) 一次性预计算所有 r 的 $\log \kappa(r)$；(ii) 在共享顶点的三角形之间缓存 $\mathbb{P}_{\boldsymbol{p}(w^{(j)})}(r)$；以及 (iii) 使用 $M_T < \alpha - \tau$ 进行提前剪枝。

V. Extension to general k

现在我们将其推广 到任意的 k，并强调哪些内容发生了变化，哪些内容保持不变 l.

A. $k-1$ 维度中的对数优势坐标

对于 $\boldsymbol{p} \in \Delta_k^\circ$，定义 $u \in \mathbb{R}^{k-1}$ 为

$$u_i = \log \frac{p_i}{p_k}, \qquad i = 1, \dots, k-1,$$

带有逆映射

$$p_i(u) = \frac{e^{u_i}}{1 + \sum_{j=1}^{k-1} e^{u_j}},\ i \le k-1, \qquad p_k(u) = \frac{1}{1 + \sum_{j=1}^{k-1} e^{u_j}}.$$

对于任意结果 $r = (r_1, \dots, r_k)$，其中 $\sum_i r_i = n$，令 $\kappa(r) = \frac{n!}{\prod_{i=1}^k r_i!}$。
则

$$\log \mathbb{P}_{\boldsymbol{p}(u)}(r) = \log \kappa(r) \ + \ \sum_{i=1}^{k-1} r_i u_i \ - \ n \log\Big(1 + \sum_{j=1}^{k-1} e^{u_j}\Big).$$

B. 半空间结构与单纯形单元

固定一个观测到的 $\hat{r}$。与引理 1 完全相同，对于任意 r,

$$\mathbb{P}_{\boldsymbol{p}(u)}(r) \le \mathbb{P}_{\boldsymbol{p}(u)}(\hat{r}) \quad \Longleftrightarrow \quad \sum_{i=1}^{k-1} (r_i - \hat{r}_i) u_i \ \le \ \log \kappa(\hat{r}) - \log \kappa(r),$$

这是 $\mathbb{R}^{k-1}$ 中的一个半空间。

为降低高维情况下顶点检查的成本，我们建议将紧致搜索区域 $\mathcal{W} \subset \mathbb{R}^{k-1}$划分为单纯形。维度为 $d = k-1$ 的单纯形只有 $d+1 = k$ 个顶点（而超立方体有 2^d 个顶点）。

C. 单纯形单元上的经认证的 p 值界

设 $S \subset \mathbb{R}^{k-1}$ 是顶点为 $\{w^{(j)}\}_{j=1}^{k}$的单纯形。定义线性函数

$$g_{r;\hat{r}}(u) := \sum_{i=1}^{k-1}(r_i - \hat{r}_i)u_i - \big(\log\kappa(\hat{r}) - \log\kappa(r)\big).$$

与前面一样，定义

$$\mathcal{I}_S(\hat{r}) := \{r :\ g_{r;\hat{r}}(w^{(j)}) \le 0 \text{ for all } j\}, \qquad \mathcal{O}_S(\hat{r}) := \{r :\ g_{r;\hat{r}}(w^{(j)}) > 0 \text{ for all } j\}.$$

函数 $u \mapsto \log \mathbb{P}_{\boldsymbol{p}(u)}(r)$ 关于 $\mathbb{R}^{k-1}$ 是凹函数，因此顶点最小值下界可直接推广：

$$\underline{P}_S(r) := \min_{j\in\{1,\dots,k\}} \mathbb{P}_{\boldsymbol{p}(w^{(j)})}(r) \quad \Rightarrow \quad \mathbb{P}_{\boldsymbol{p}(u)}(r) \ge \underline{P}_S(r) \ \ \forall u \in S.$$

因此，命题 2 中的 p 值界在将 T 替换为 S 后仍然成立：

$$\underline{\rho}_S(\hat{r}) = \sum_{r\in\mathcal{I}_S(\hat{r})} \underline{P}_S(r), \qquad \overline{\rho}_S(\hat{r}) = 1 - \sum_{r\in\mathcal{O}_S(\hat{r})} \underline{P}_S(r).$$

D. 算法与扩展性讨论

算法 1 可通过将三角形替换为 $(k-1)$-单纯形，并采用形状正则细化规则（例如最新顶点二分法）来推广。相交/不相交判定的可靠性保持不变。

对于 $k > 3$，主要区别在于计算规模：$m = |\Delta_{k,\ n}| = \binom{n+k-1}{k-1}$ 在 k 固定时随 n^{k-1} 增长；在最坏情况下，自适应细化可能需要维度为 $k-1$ 时的大量单元。尽管如此，对于适中的 k（分类 A/B 测试中很常见），该方法仍然实用，因为：(i) 单纯形单元只使用 k 个顶点；(ii) 可以积极缓存顶点概率评估；(iii) 许多单元可通过 $M_S < \alpha - \tau$及早剪枝。

E. 通过面分解处理计数为零的类别

到目前为止的推导隐含地假设 $\boldsymbol{p} \in \Delta_k^\circ$，即对所有 i，$p_i > 0$，因此对数优势坐标均为有限值。在实践中，观测到的多项分布结果往往包含零计数，而任何正确的相交检验都必须考虑单纯形的边界面。我们展示了如何通过对低维面进行有限分解来简洁地处理这一问题。

令 $\hat{r}^{(A)}$，$\hat{r}^{(B)} \in \mathbb{N}^k$ 为观测结果。定义索引集

$$S_A := \{i : \hat{r}_i^{(A)} > 0\}, \qquad S_B := \{i : \hat{r}_i^{(B)} > 0\}, \qquad S_0 := S_A \cup S_B.$$

引理 4 (零计数排除**).** 若对某个结果 $\hat{r}$ 和 $p_i = 0$，有 $\hat{r}_i > 0$，则 $\rho_{\hat{r}}(\boldsymbol{p}) = 0.$

证明。若 $p_i = 0$ 且 $\hat{r}_i > 0$，则 $\mathbb{P}_{\boldsymbol{p}}(\hat{r}) = 0$。p 值尾部 $\{r : \mathbb{P}_{\boldsymbol{p}}(r) \le \mathbb{P}_{\boldsymbol{p}}(\hat{r})\}$ 仅包含概率为零的结果，因此 p 值总和等于零。 □

引理 4 表明，任何满足 $\rho_A(\boldsymbol{p}) \ge \alpha$ 且 $\rho_B(\boldsymbol{p}) \ge \alpha$ 的参数 $\boldsymbol{p}$，对于所有 $i \in S_0$都必须满足 $p_i > 0$。因此，在潜在的相交见证中，只有两个结果中计数均为零的类别才可设为零。我们在附录VI-E中给出此类情况的示例。

Let

$$Z := \{1, \dots, k\} \setminus S_0$$

表示共同零计数类别的集合。对于任意子集 $T \subseteq Z$，定义面

$$\Delta_k^{(T)} := \{\boldsymbol{p} \in \Delta_k :\ p_i = 0 \ \forall i \in T\}.$$

在此面上，多项分布约化为关于剩余坐标的（$k - |T|$）类别多项分布，样本量保持不变 n。相应的约化结果可通过从 $\hat{r}^{(A)}$ 和 $\hat{r}^{(B)}$ 中删除 T 中的坐标得到。

因此，判定最小体积置信集是否相交，可归结为在约化维度下检查每个面 $\Delta_k^{(T)}$、 $T \subseteq Z$ 上的鲁棒相交。我们在附录VI-F中给出了一个此类案例的示例。算法 2中给出了一种可扩展至 k 维并处理零计数的通用算法。

VI. 总结

本文研究了与多项分布参数的最小体积置信集相关的一个基本计算问题：判定对应于两个观测结果的置信集是否相交。我们没有尝试直接计算或可视化最小体积置信集，而是构建了一种经认证的决策过程，利用了对数优势坐标中由似然排序所诱导的几何结构。

对于三个类别的情形，我们提出了一种自适应划分算法，可计算每个单元上精确 p 值的严格上下界，从而在用户指定的容差范围内，给出可证明可靠的相交或不相交证书。该方法通过在似然排序固定的区域上进行推理，自然地处理了最小体积置信集的不连续性和非凸性。我们还展示了该方法如何扩展到高维，以及如何通过将单纯形分解为有限个低维面的集合，妥善处理由计数为零的类别引起的边界情形。

参考文献

[1] K. Jamieson、M. Malloy、R. Nowak 和 S. Bubeck，“lil'ucb：多臂老虎机的最优探索算法”，载于学习理论会议。PMLR，2014 年，第 423–439 页。[2] M. L. Malloy、A. Tripathy 和 R. D. Nowak，“多项分布参数的最优置信集”，载于*2021*年*IEEE*国际信息论研讨会 *(ISIT)*，2021 年，第 2173–2178 页。[3] G. Casella 和 R. L. Berger，统计推断。圣智学习，2021 年。[4] D. Chafai 和 D. Concordet，“小样本量下多项分布参数的置信区域”，美国统计学会会刊，第 104 卷，第 487 期，第 1071–1079 页，2009 年。[5] M. L. Malloy、A. Tripathy 和 R. D. Nowak，“多项分布参数的最优置信区域”，*arXiv* 预印本 *arXiv:2002.01044*，2020 年。[6] K. Jamieson、M. Malloy、R. Nowak 和 S. Bubeck，“如何在众多均值中找到最大值”，*arXiv* 预印本 *arXiv:1306.3917*，2013 年。[7] M. L. Malloy、S. Alfeld 和 P. Barford，“通过凸松弛进行污染估计”，载于*2015*年*IEEE*国际信息论研讨会 *(ISIT)*。IEEE，2015 年，第 1189–1193 页。[8] M. L.Malloy、G.Tang 和 R. D.Nowak，“稀有分布的快速搜索”，载于*2012*年第*46*届信息科学与系统年会 *(CISS)*。IEEE，2012 年，第 1–6 页。[9] ——，“多个总体搜索的样本复杂度”，*IEEE*信息论汇刊，第 59 卷，第 8 期，第 5039–5050 页，2013 年。[10] T. E. Sterne，“关于置信限或保真限的一些评论”，*Biometrika*，第 41 卷，第 1/2 期，第 275–278 页，1954 年。[11] E. L. Crow，“比例的置信区间”，*Biometrika*，第 43 卷，第 3/4 期，第 423–435 页，1956 年。[12] L. D. Brown、G. Casella 和 J. Gene Hwang，“最优置信集、生物等效性与帕斯卡蜗线”，美国统计学会会刊，第 90 卷，第 431 期，第 880–889 页，1995 年。[13] J. Resin，“精确多项检验的简单算法”，2020 年。[14] H. Lin、M. Li、D. Pimentel-Alarcón 和 M. L. Malloy，“最小体积置信集的几何”，载于*2022*年*IEEE*国际信息论研讨会 *(ISIT)*。IEEE，2022 年，第 3180–3185 页。[15] S. S. Wilks，“检验复合假设时似然比的大样本分布”，数学统计年刊，第 9 卷，第 1 期，第 60–62 页，1938 年。

附录

A. 引理 *2* 的详细证明

引理 2 的重述。设 $k \geq 2$ 和 $n \geq 1$，并设 $\Delta_{k,n}$ 表示由 n 个来自 k 个类别的独立同分布样本所产生的所有经验分布（等价地，计数向量除以 n）的集合。设 $m := |\Delta_{k,n}| = \binom{n+k-1}{k-1}$。任取一个观测结果 $\hat{r}$（或等价地，$\widehat{\boldsymbol{p}} = \hat{r}/n \in \Delta_{k,n}$）。那么对每个多项分布参数 $\boldsymbol{p} \in \Delta_k$，

$$\rho_{\widehat{\boldsymbol{p}}}(\boldsymbol{p}) \;\leq\; m\,\mathbb{P}_{\boldsymbol{p}}(\hat{r}).$$

等价地，将 $\rho_{\hat{r}}(\boldsymbol{p})$ 记作相同的 p 值，

$$\rho_{\hat{r}}(\boldsymbol{p}) \;\leq\; m\,\mathbb{P}_{\boldsymbol{p}}(\hat{r}).$$

证明。我们先展开本文中始终使用的（精确）p 值的定义。任取 $\hat{r}$ 和 $\boldsymbol{p} \in \Delta_k$。回顾一下，在原假设 $\boldsymbol{p}$ 下，与 $\hat{r}$ 相关联的 p 值为

$$\rho_{\hat{r}}(\boldsymbol{p}) \;:=\; \sum_{r \in \mathcal{T}_{\hat{r}}(\boldsymbol{p})} \mathbb{P}_{\boldsymbol{p}}(r), \qquad \text{where} \quad \mathcal{T}_{\hat{r}}(\boldsymbol{p}) := \Big\{ r \in \Delta_{k,n} :\; \mathbb{P}_{\boldsymbol{p}}(r) \leq \mathbb{P}_{\boldsymbol{p}}(\hat{r}) \Big\}. \tag{8}$$

这里，$\Delta_{k,n}$ 索引了来自 n 次抽样的所有可能经验结果（等价地，所有计数向量 $r = (r_1, \ldots, r_k) \in \mathbb{N}^k$ 满足 $\sum_i r_i = n$），而 $\mathbb{P}_{\boldsymbol{p}}(r)$ 是在 $\boldsymbol{p}$ 下观测到该结果的多项分布概率。Th
详细证f is a di明 i交 d-boundi叉论证。
我们首先证明，尾部集合至多包含 m 个结果。根据定义，$\mathcal{T}_{\hat{r}}(\boldsymbol{p}) \subseteq \Delta_{k,\,n}$。由于 $|\Delta_{k,\,n}| = m$，我们得到平凡的基数上界

$$|\mathcal{T}_{\hat{r}}(\boldsymbol{p})| \;\leq\; |\Delta_{k,n}| \;=\; m. \tag{9}$$

然后我们证明，p 值中的每一项都被 $\mathbb{P}_{\boldsymbol{p}}(\hat{r})$ 所界定。对于任意 $r \in \mathcal{T}_{\hat{r}}(\boldsymbol{p})$，$\mathcal{T}_{\hat{r}}(\boldsymbol{p})$ 的定义性质给出

$$\mathbb{P}_{\boldsymbol{p}}(r) \;\le\; \mathbb{P}_{\boldsymbol{p}}(\hat{r}). \tag{10}$$

这对 p 值求和式 (8) 中包含的每个 r 都成立。

现在我们来界定总尾部概率质量。利用 (8)、(10)，以及求和中共有 $|\mathcal{T}_{\hat{r}}(\boldsymbol{p})|$ 项这一事实，可得

$$\rho_{\hat{r}}(\boldsymbol{p}) \;=\; \sum_{r\in\mathcal{T}_{\hat{r}}(\boldsymbol{p})} \mathbb{P}_{\boldsymbol{p}}(r) \;\le\; \sum_{r\in\mathcal{T}_{\hat{r}}(\boldsymbol{p})} \mathbb{P}_{\boldsymbol{p}}(\hat{r}) \;=\; |\mathcal{T}_{\hat{r}}(\boldsymbol{p})|\,\mathbb{P}_{\boldsymbol{p}}(\hat{r}). \tag{11}$$

最后，将基数上界 (9) 应用于 (11)：

$$\rho_{\hat{r}}(\boldsymbol{p}) \;\le\; m\,\mathbb{P}_{\boldsymbol{p}}(\hat{r}),$$

这正是所需的不等式。 □

B. 超水平集限制的几何解释

现在我们来证明引理 2 后面用到的几何论述，即该约束 t

$$\mathbb{P}_{\boldsymbol{p}}(\hat{r}) \;\ge\; \frac{\alpha-\tau}{m}$$

在对数优势坐标中定义了一个凸的有界区域。

我们首先将约束表示为一个超水平集。固定一个观测结果 $\hat{r}$，并考虑函数

$$f(u,v) \;:=\; \log \mathbb{P}_{\boldsymbol{p}(u,v)}(\hat{r}),$$

其中 $(u,v) \in \mathbb{R}^2$ 是由下式定义的对数优势坐标

$$p_1 = \frac{e^u}{1+e^u+e^v}, \qquad p_2 = \frac{e^v}{1+e^u+e^v}, \qquad p_3 = \frac{1}{1+e^u+e^v}.$$

不等式

$$\mathbb{P}_{\boldsymbol{p}(u,v)}(\hat{r}) \;\ge\; \frac{\alpha-\tau}{m}$$

对两边取对数后，等价于

$$f(u,v) \;\ge\; \log(\alpha-\tau) - \log m. \tag{12}$$

因此，满足 (12) 的点集 (u,v) 是函数 f 的超水平集，即

$$\mathcal{S} \;:=\; \{(u,v)\in\mathbb{R}^2 :\; f(u,v)\ge c\}, \qquad c := \log(\alpha-\tau) - \log m.$$

现在我们证明对数似然函数的凹性。回顾正文中的 (5)：

$$f(u,v) = \log\kappa(\hat{r}) + \hat{r}_1 u + \hat{r}_2 v - n\log(1+e^u+e^v).$$

前三项是 $(u,\ v)$ 的仿射函数。最后一项 $-\log(1+e^u+e^v)$ 是凸函数的相反数，因为 $\log(1+e^u+e^v)$ 在 $\mathbb{R}^2$ 上是凸函数。因此，$f(u,\ v)$ 在 $\mathbb{R}^2$ 上是凹函数。凸分析中的一个标准结果指出，若 $f:\ \mathbb{R}^d \to \mathbb{R}$ 是凹函数，则对于任意常数 $c\in\mathbb{R}$，超水平集

$$\{x :\; f(x)\ge c\}$$

是 $\mathbb{R}^d$ 的凸子集。将该结果应用于 $f(u,\ v)$，可知上文定义的集合 $\mathcal{S}$ 是凸的。

现在我们证明 $\mathcal{S}$ 有界。由于 $\|(u,\ v)\|\to\infty$，u 或 v 至少有一个趋于 $+\infty$ 或 $-\infty$。在所有这些情况下，项 $-n\log(1+e^u+e^v)$ 占主导并趋于 $-\infty$，而线性项 $\hat{r}_1 u + \hat{r}_2 v$ 的增长至多为线性。因此，

$$\lim_{\|(u,v)\|\to\infty} f(u,v) = -\infty.$$

这意味着，对于任意有限常数 c，不等式 $f(u,\ v)\ge c$ 只能在 $\mathbb{R}^2$ 的有界区域内成立。因此，超水平集 $\mathcal{S}$ 是有界的。

因此，约束 $\mathbb{P}_{\boldsymbol{p}}(\hat{r})\ge(\alpha-\tau)/m$ 在 (u,v)-坐标下定义了 $\mathbb{R}^2$ 的一个有界凸子集。分别将这一论证应用于 $\hat{r}^{(A)}$ 和 $\hat{r}^{(B)}$，我们可以得出，两个相应超水平集的交集同样是凸且有界的。因此，任何包含该交集的紧凸集 $\mathcal{W}$ 都可以用作算法 1 中的搜索域。

C. 引理 *3* 的详细证明

引理 3 的重述。 固定 $k = 3$ 和 $n \geq 1$。对于任意结果 $r = (r_1, r_2, r_3) \in \mathbb{N}^3$，满足 $r_1 + r_2 + r_3 = n$，定义 e

$$f_r(u, v) \;:=\; \log \mathbb{P}_{\boldsymbol{p}(u,v)}(r),$$

其中，$\boldsymbol{p}(u,\ v)$ 由 softmax 映射 (4) 给出。设 $T \subset \mathbb{R}^2$ 为任意一个顶点为 $w^{(1)}$、$w^{(2)}$、$w^{(3)} \in \mathbb{R}^2$ 的（闭）三角形。则

$$\min_{(u,v)\in T} f_r(u, v) \;=\; \min_{j\in\{1,2,3\}} f_r\Big(w^{(j)}\Big).$$

因此，量

$$\underline{P}_T(r) \;:=\; \min_{j\in\{1,2,3\}} \mathbb{P}_{\boldsymbol{p}(w^{(j)})}(r)$$

对所有 $(u,\ v) \in T$，均满足 $\mathbb{P}_{\boldsymbol{p}(u,v)}(r) \geq P_T(r)$。

证明。我们分三步进行。首先，我们验证 f_r 在 $\mathbb{R}^2$ 上的凹性。其次，我们利用关于凸多面体上凹函数的一个一般事实：多面体上的最小值在某个极点（顶点）处取得。第三，我们将其专门应用于三角形，并得出概率下界。

我们首先证明 $f_r(u,\ v)$ 关于 $(u,\ v)$ 是凹函数。回顾对数优势坐标中的多项分布概率（参见 (5)）：

$$f_r(u, v) = \log \kappa(r) \;+\; r_1 u + r_2 v \;-\; n \log(1 + e^u + e^v), \tag{13}$$

其中，$\kappa(r) = \frac{n!}{r_1!\ r_2!\ r_3!}$ 相对于 $(u,\ v)$ 为常数。(13) 右侧的前三项是 $(u,\ v)$ 的仿射函数，因此既是凸函数，也是凹函数。剩余项

$$(u, v) \;\mapsto\; \log(1 + e^u + e^v)$$

在 $\mathbb{R}^2$ 上是凸函数（这是标准的“log-sum-exp”函数）。凸函数乘以 $-n < 0$ 会得到凹函数。因此，仿射项与凹项 $-n \log(1 + e^u + e^v)$ 之和是凹函数，由此证明 f_r 在 $\mathbb{R}^2$ 上是凹函数。

接下来，我们证明凹函数在三角形上的最小值出现在顶点处。三角形 $T = \mathrm{conv}\{w^{(1)}, w^{(2)}, w^{(3)}\}$ 是一个紧凸集。我们将证明，对于任意 $(u,\ v) \in T$，

$$f_r(u, v) \;\geq\; \min_{j\in\{1,2,3\}} f_r\Big(w^{(j)}\Big). \tag{14}$$

由于 T 是紧集，且 f_r 连续（因为它是连续函数之和），根据极值定理，f_r 在 T 上取得最小值。因此，不等式 (14) 表明，该最小值不可能小于顶点上的最小值，故最小值必定在至少一个顶点处取得。

为证明 (14)，任取 $(u,\ v) \in T$。根据凸包的定义，存在系数 $\lambda_1,\ \lambda_2,\ \lambda_3 \geq 0$，满足 $\lambda_1 + \lambda_2 + \lambda_3 = 1$，使得

$$(u, v) \;=\; \lambda_1 w^{(1)} + \lambda_2 w^{(2)} + \lambda_3 w^{(3)}.$$

f_r 的凹性恰恰意味着，对于任意这样的凸组合，

$$f_r\Big(\lambda_1 w^{(1)} + \lambda_2 w^{(2)} + \lambda_3 w^{(3)}\Big) \;\geq\; \lambda_1 f_r(w^{(1)}) + \lambda_2 f_r(w^{(2)}) + \lambda_3 f_r(w^{(3)}). \tag{15}$$

利用 $(u,\ v)$ 的表示形式并应用 (15)，得到

$$f_r(u, v) \;\geq\; \lambda_1 f_r(w^{(1)}) + \lambda_2 f_r(w^{(2)}) + \lambda_3 f_r(w^{(3)}).$$

最后，由于 $\lambda_1,\ \lambda_2,\ \lambda_3 \geq 0$ 且它们的和为 1，右侧是以下三个数的凸组合 $\{f_r(w^{(1)})\ f_r(w^{(2)})\ f_r(w^{(3)})\}$ an 因此不能小于它们的最小值。形式上，

$$\lambda_1 f_r(w^{(1)}) + \lambda_2 f_r(w^{(2)}) + \lambda_3 f_r(w^{(3)}) \;\geq\; \min_{j\in\{1,2,3\}} f_r(w^{(j)}).$$

将两个不等式结合起来即可得到（14）。

现在我们来证明最小值相等。令 $j^\star \in \arg\min_{j\in\{1,2,3\}} f_r(w^{(j)})$。由于 $w^{(j^\star)} \in T$，我们有

$$\min_{(u,v)\in T} f_r(u, v) \;\leq\; f_r\Big(w^{(j^\star)}\Big) \;=\; \min_j f_r(w^{(j)}).$$

另一方面，不等式 (14) 对所有 $(u,v) \in T$ 都成立，因此对两边取关于 $(u,v) \in T$ 的最小值可得

$$\min_{(u,v)\in T} f_r(u,v) \;\geq\; \min_j f_r(w^{(j)}).$$

这两个不等式结合起来可推出

$$\min_{(u,v)\in T} f_r(u,v) \;=\; \min_{j\in\{1,2,3\}} f_r(w^{(j)}),$$

从而证明了第一个结论。

步骤 4：将结果转换为概　　率。指数运算保序，因此由 $f_r(u,\ v) \geq \min_j f_r(w^{(j)})$ 可得

$$\mathbb{P}_{\boldsymbol{p}(u,v)}(r) \;=\; e^{f_r(u,v)} \;\geq\; \exp\Big(\min_j f_r(w^{(j)})\Big) \;=\; \min_j \exp(f_r(w^{(j)})) \;=\; \min_{j\in\{1,2,3\}} \mathbb{P}_{\boldsymbol{p}(w^{(j)})}(r).$$

这正是对所有 $(u,\ v) \in T$ 都成立的 $\mathbb{P}_{\boldsymbol{p}(u,v)}(r) \geq P_T(r)$，证明完毕。 □

推广到 $k > 3$：当维度为 $d = k-1$ 、以具有顶点 $\{w^{(j)}\}_{j=1}^k$ 的单纯形 $S \subset \mathbb{R}^{k-1}$ 替换 T 时，同样的论证可原样成立，因为单纯形中的每个点都是其顶点的凸组合，且 $\log \mathbb{P}_{\boldsymbol{p}(u)}(r)$ 关于 u 仍然是凹函数。

D. 命题 *2* 的详细证明

命题 2 的重述。固定一个观测结果 $\hat{}r \in \mathbb{N}^3$，其中 $\hat{}r_1 + \hat{}r_2 + \hat{}r_3 = n$，并令 $\widehat{\boldsymbol{p}} = \hat{}r/n \in \Delta_{3,\ n}$。令 $T \subset \mathbb{R}^2$ 为一个三角形，其顶点为 $w^{(1)}$、$w^{(2)}$、$w^{(3)}$。对于每个结果 $r \in \mathbb{N}^3$（其总和为 n），定义

$$g_{r;\hat{r}}(u,v) := (r_1 - \hat{r}_1)u + (r_2 - \hat{r}_2)v - \big(\log\kappa(\hat{r}) - \log\kappa(r)\big),$$

以及索引集

$$\mathcal{I}_T(\hat{r}) := \{r:\ g_{r;\hat{r}}(w^{(j)}) \leq 0 \text{ for all } j = 1,2,3\}, \qquad \mathcal{O}_T(\hat{r}) := \{r:\ g_{r;\hat{r}}(w^{(j)}) > 0 \text{ for all } j = 1,2,3\}.$$

Let

$$\underline{P}_T(r) := \min_{j\in\{1,2,3\}} \mathbb{P}_{\boldsymbol{p}(w^{(j)})}(r),$$

并定义

$$\underline{\rho}_T(\hat{r}) := \sum_{r\in\mathcal{I}_T(\hat{r})} \underline{P}_T(r), \qquad \overline{\rho}_T(\hat{r}) := 1 - \sum_{r\in\mathcal{O}_T(\hat{r})} \underline{P}_T(r).$$

那么，对于每个 $(u,\ v) \in T$,

$$\underline{\rho}_T(\hat{r}) \;\leq\; \rho_{\widehat{\boldsymbol{p}}}(\boldsymbol{p}(u,v)) \;\leq\; \overline{\rho}_T(\hat{r}),$$

其中

$$\rho_{\widehat{\boldsymbol{p}}}(\boldsymbol{p}(u,v)) = \sum_{r:\ \mathbb{P}_{\boldsymbol{p}(u,v)}(r) \leq \mathbb{P}_{\boldsymbol{p}(u,v)}(\hat{r})} \mathbb{P}_{\boldsymbol{p}(u,v)}(r)$$

即式 (2) 中定义的精确 p 值。

*Pr*嗷。 设三角形 T的顶点为 $w^{(1)}$、$w^{(2)}$、$w^{(3)}$。在全文中，令 $(u,v) \in T$为任意值。

我们首先刻画半空间尾部归属。根据引理 1，对于任意结果 r，有如下等价关系

$$\mathbb{P}_{\boldsymbol{p}(u,v)}(r) \leq \mathbb{P}_{\boldsymbol{p}(u,v)}(\hat{r}) \quad\Longleftrightarrow\quad g_{r;\hat{r}}(u,v) \leq 0. \tag{16}$$

此外，$g_{r;\hat{r}}(u,v)$ 是 (u,v) 的仿射函数（实际上是线性函数加常数）。

接下来，我们可以明确“确定属于尾部”的结果对应的顶点。考虑任意 $r \in \mathcal{I}_T(\hat{}r)$。根据定义，对于所有顶点 $j = 1$、2、3，$g_{r;\hat{r}}(w^{(j)}) \leq 0$。由于 $g_{r;\hat{r}}$是仿射函数，且 T是由其顶点构成的凸包，因此每个点 $(u,v) \in T$都可表示为凸组合

$$(u,v) = \lambda_1 w^{(1)} + \lambda_2 w^{(2)} + \lambda_3 w^{(3)}, \qquad \lambda_j \geq 0,\ \sum_{j=1}^3 \lambda_j = 1.$$

仿射函数保持凸组合，因此

$$g_{r;\hat{r}}(u,v) = \lambda_1 g_{r;\hat{r}}(w^{(1)}) + \lambda_2 g_{r;\hat{r}}(w^{(2)}) + \lambda_3 g_{r;\hat{r}}(w^{(3)}) \;\leq\; 0.$$

利用 (16)，我们得出，对于每个 $(u,\ v) \in T$,

$$r \in \mathcal{I}_T(\hat{r}) \quad\Longrightarrow\quad \mathbb{P}_{\boldsymbol{p}(u,v)}(r) \leq \mathbb{P}_{\boldsymbol{p}(u,v)}(\hat{r}), \tag{17}$$

即， r 属于三角形中所有参数的 p 值尾部。

现在考虑“确定不属于尾部”的结果对应的顶点。现在考虑任意 $r \in \mathcal{O}_T(\hat{r})$。根据定义，对于所有顶点， $g_{r;\hat{r}}(w^{(j)}) > 0$ 。同样的凸组合论证意味着，对于每个 $(u,\ v) \in T$,

$$g_{r;\hat{r}}(u,v) = \lambda_1 g_{r;\hat{r}}(w^{(1)}) + \lambda_2 g_{r;\hat{r}}(w^{(2)}) + \lambda_3 g_{r;\hat{r}}(w^{(3)}) \ > \ 0,$$

因此，根据式 (16),

$$r \in \mathcal{O}_T(\hat{r}) \quad \Longrightarrow \quad \mathbb{P}_{\boldsymbol{p}(u,v)}(r) > \mathbb{P}_{\boldsymbol{p}(u,v)}(\hat{r}) \tag{18}$$

对于所有 $(u,\ v) \in T$。特别是，这样的 r在 T 上的任何位置都被排除在 p 值尾部之外。

概率的统一下　　界。根据引理 3，对于每个结果 r 以及所有 $(u,\ v) \in T$,

$$\mathbb{P}_{\boldsymbol{p}(u,v)}(r) \ \geq \ \underline{P}_T(r) \ := \ \min_{j\in\{1,2,3\}} \mathbb{P}_{\boldsymbol{p}(w^{(j)})}(r). \tag{19}$$

p 值的下界如下：令 $\mathcal{T}(u,\ v)$ 表示 $(u,\ v)$ 处的 p 值尾部集合：

$$\mathcal{T}(u,v) := \{r :\ \mathbb{P}_{\boldsymbol{p}(u,v)}(r) \leq \mathbb{P}_{\boldsymbol{p}(u,v)}(\hat{r})\}.$$

根据 (17)，对于所有 $(u,\ v) \in T$，每个 $r \in \mathcal{I}_T(\hat{r})$ 都属于 $\mathcal{T}(u,\ v)$，因此

$$\rho_{\widehat{\boldsymbol{p}}}(\boldsymbol{p}(u,v)) = \sum_{r\in\mathcal{T}(u,v)} \mathbb{P}_{\boldsymbol{p}(u,v)}(r) \ \geq \ \sum_{r\in\mathcal{I}_T(\hat{r})} \mathbb{P}_{\boldsymbol{p}(u,v)}(r).$$

逐项应用一致概率下界 (19)，得到

$$\rho_{\widehat{\boldsymbol{p}}}(\boldsymbol{p}(u,v)) \ \geq \ \sum_{r\in\mathcal{I}_T(\hat{r})} \underline{P}_T(r) \ = \ \underline{\rho}_T(\hat{r}).$$

由于 $(u,\ v) \in T$ 是任意的，这就证明了该下界在整个三角形上一致成立：

$$\underline{\rho}_T(\hat{r}) \ \leq \ \rho_{\widehat{\boldsymbol{p}}}(\boldsymbol{p}(u,v)) \qquad \forall (u,v) \in T.$$

类似地，我们可以给 p 值一个上界。根据 (18)，对于任意 $(u,\ v) \in T$，每个 $r \in \mathcal{O}_T(\hat{r})$ 都不在尾部 $\mathcal{T}(u,\ v)$ 中。因此，p 值尾部包含于 $\mathcal{O}_T(\hat{r})$ 的补集中：

$$\mathcal{T}(u,v) \ \subseteq \ \Delta_{3,n} \setminus \mathcal{O}_T(\hat{r}).$$

因此

$$\rho_{\widehat{\boldsymbol{p}}}(\boldsymbol{p}(u,v)) = \sum_{r\in\mathcal{T}(u,v)} \mathbb{P}_{\boldsymbol{p}(u,v)}(r) \ \leq \ \sum_{r\in\Delta_{3,n}\setminus\mathcal{O}_T(\hat{r})} \mathbb{P}_{\boldsymbol{p}(u,v)}(r) = 1 - \sum_{r\in\mathcal{O}_T(\hat{r})} \mathbb{P}_{\boldsymbol{p}(u,v)}(r). \tag{20}$$

对每个 $r \in \mathcal{O}_T(\hat{r})$ 应用一致下界 (19)，得到

$$\sum_{r\in\mathcal{O}_T(\hat{r})} \mathbb{P}_{\boldsymbol{p}(u,v)}(r) \ \geq \ \sum_{r\in\mathcal{O}_T(\hat{r})} \underline{P}_T(r),$$

代入（20）得

$$\rho_{\widehat{\boldsymbol{p}}}(\boldsymbol{p}(u,v)) \ \leq \ 1 - \sum_{r\in\mathcal{O}_T(\hat{r})} \underline{P}_T(r) \ = \ \overline{\rho}_T(\hat{r}).$$

同样，由于（u,v） $\in T$ 是任意的，这对三角形中的所有点都一致成立。

结合这两个界可得，对所有（u,v） $\in T$,

$$\underline{\rho}_T(\hat{r}) \ \leq \ \rho_{\widehat{\boldsymbol{p}}}(\boldsymbol{p}(u,v)) \ \leq \ \overline{\rho}_T(\hat{r}),$$

证明完毕。 □

$\mathcal{I}_T(\hat{r})$ 中的结果保证在 T 上的任何位置都会计入 p 值尾部，而 $P_T(r)$ 对其贡献给出了统一下界，从而得到 $\rho_T(\hat{r})$。 $\mathcal{O}_T(\hat{r})$ 中的结果保证在 T 上绝不会计入；通过对其概率质量给出下界，我们可以用 1 减去该下界，得到剩余尾概率的上界，从而得到 $\overline{\rho}_T(\hat{r})$。

E. 零计数排除示例。

考虑 $k=3$ 个类别和 $n=5$ 个样本。假设观测结果为 $\hat{r}=(2,\ 2,\ 1)$，因此类别 3 至少出现一次。现在考虑参数 $\boldsymbol{p}=(1/2,\ 1/2,\ 0)$，它给类别 3 分配了零概率。那么

$$\mathbb{P}_{\boldsymbol{p}}(\hat{r})=\frac{5!}{2!\,2!\,1!}(1/2)^2(1/2)^2(0)^1=0.$$

由于 $\mathbb{P}_{\boldsymbol{p}}(\hat{r})=0$，p 值尾部仅包含概率为零的结果，因此 $\rho_{\hat{r}}(\boldsymbol{p})=0$。特别地，对于任意置信水平 $\alpha>0$，$\boldsymbol{p}$ 都不可能属于最小体积置信集。

相反，如果 $\hat{r}=(2,\ 3,\ 0)$ 且 $\boldsymbol{p}=(1/2,\ 1/2,\ 0)$，那么 $\mathbb{P}_{\boldsymbol{p}}(\hat{r})>0$，且 p 值不一定为零。这说明，在寻找交集见证点时，只有在两个结果中计数均为零的类别才能设为零概率。如果某个类别 i 在两个观测结果中的计数都为零，那么将 $p_i=0$ 设为零不会使任一似然失效。在这种情况下，多项式模型退化到单纯形的一个低维面，对应于剩余类别上的约化多项式分布。重要的是，这种退化是**精确的**，而非近似的：给定$p_i=0$ 时结果的分布恰好是一个类别数少一、样本量仍为 n 的多项式分布。

从 p 值的角度来看，这意味着将 $\boldsymbol{p}$ 限制到面 $\Delta_k^{(T)}$ 上（对于 $T\subseteq Z$），会产生与从参数和观测结果中同时删除计数为零的坐标所得排序相同的 p 值排序。因此，在面 $\Delta_k^{(T)}$ 上检验最小体积置信集的交集，等价于对维度为 $k-|T|$ 的约化结果运行相同的经认证的交集检验。

算法 2 利用这一结构，枚举所有对应于共同零计数类别的面，并在每个约化问题上应用内部算法。这样就能精确处理边界行为，无需近似或特殊情况逻辑，同时保持与内部区域相同的可靠性保证。

F. 映射到单纯形低维面上的示例

考虑一个有 $k=4$ 个类别、样本量为 $n=5$ 的 A/B 检验问题。假设观测结果为

$$\hat{r}^{(A)}=(3,2,0,0),\qquad \hat{r}^{(B)}=(4,1,0,0).$$

类别 1 和 2 在两组结果中都至少出现一次，而类别 3 和 4 从未被观测到。因此，我们定义

$$S_0=\{1,2\},\qquad Z=\{3,4\}.$$

假设我们考虑参数 $\boldsymbol{p}$，其中 $p_1=0$。由于 $\hat{r}_1^{(A)}>0$ 和 $\hat{r}_1^{(B)}>0$，在 $\boldsymbol{p}$ 下两个观测结果的多项式似然均为零。根据引理 4，这意味着

$$\rho_A(\boldsymbol{p})=\rho_B(\boldsymbol{p})=0.$$

因此，$\boldsymbol{p}$ 不可能属于任一最小体积置信集，并且任何相交见证点都不能位于满足 $p_1=0$ 的任何面上。对 $p_2=0$ 也适用同样的推理。

因此，S_0 中的类别在任何候选交点处都必须严格为正。

相比之下，类别 3 和 4 在两个观测结果中的计数均为零。令 $p_3=0$ 或 $p_4=0$ 不会导致任一观测结果的似然变为零。因此，p 值 $\rho_A(\boldsymbol{p})$ 和 $\rho_B(\boldsymbol{p})$ 可能仍严格为正，而此类参数无法依据先验信息排除。

每个子集 $T\subseteq Z$ 都对应于单纯形的一个面，在该面上

$$p_i=0\quad\forall i\in T.$$

在此示例中，相关的面为：

- $T=\emptyset$：p_1、p_2、p_3、$p_4>0$ 的完整内部；
- $T=\{3\}$：面 $p_3=0$；
- $T=\{4\}$：面 $p_4=0$；
- $T=\{3,4\}$：面 $p_3=p_4=0$，将问题简化为类别 1 和 2 上的双类别（二项）模型。

每个参数 $\boldsymbol{p}\in\Delta_4$ 都恰好位于这些面中的一个面的相对内部。算法 2 使用相应约化维度下经认证的相交检验，对每个面进行穷尽检查。

算法 2 带零计数处理的经认证 MVC 相交检验 要求：观测结果 $\hat{r}^{(A)}$、$\hat{r}^{(B)} \in \mathbb{N}^k$，显著性水平 $\alpha \in (0,1)$，决策裕度 $\tau > 0$，细化容差 $\varepsilon > 0$。确保：相交、不相交或不确定。 1: 计算 $S_0 = \{i : \hat{r}_i^{(A)} > 0\} \cup \{i : \hat{r}_i^{(B)} > 0\}$ 和 $Z = \{1, \dots, k\} \setminus S_0$。 2: 对每个子集 $T \subseteq Z$ 执行 3: 删除 T 中的坐标，得到约化结果 $\hat{r}_{-T}^{(A)}$、$\hat{r}_{-T}^{(B)}$。 4: 设置 $k' \leftarrow k - |T|$。 5: 构造一个紧致搜索域 $\mathcal{W}_T \subset \mathbb{R}^{k'-1}$。 6: 对 $(\hat{r}_{-T}^{(A)}, \hat{r}_{-T}^{(B)})$ 运行算法 1（或其 $(k'-1)$ 维推广）。 7: 如果输出为相交，则 8: 返回相交。 9: 结束如果 10: 结束循环 11: 如果所有面子问题均返回不相交，则 12: 返回不相交。 13: 否则 14: 返回不确定。 15: 结束如果

G. W 在 (u, v) 坐标中的包围盒构造

我们提供两种推导 W 界限的方法，一种是闭式方法，另一种是解析式方法。我们始终从带有对数优势坐标的 $k = 3$ 开始

$$u = \log \frac{p_1}{p_3}, \qquad v = \log \frac{p_2}{p_3},$$

因此

$$p_1(u,v) = \frac{e^u}{1 + e^u + e^v}, \quad p_2(u,v) = \frac{e^v}{1 + e^u + e^v}, \quad p_3(u,v) = \frac{1}{1 + e^u + e^v}.$$

对于观测计数向量 $\hat{r} = (\hat{r}_1, \hat{r}_2, \hat{r}_3)$，其中 $\sum_i \hat{r}_i = n$，定义

$$f_{\hat{r}}(u,v) \triangleq \log P_{p(u,v)}(\hat{r}) = \log \kappa(\hat{r}) + \hat{r}_1 u + \hat{r}_2 v - n \log\bigl(1 + e^u + e^v\bigr),$$

其中 $\kappa(\hat{r}) = \frac{n!}{\hat{r}_1! \hat{r}_2! \hat{r}_3!}$。对于固定阈值 $t > 0$，超水平集

$$S_{\hat{r}}(t) \triangleq \{(u,v) \in \mathbb{R}^2 : \ P_{p(u,v)}(\hat{r}) \geq t\} = \{(u,v) : \ f_{\hat{r}}(u,v) \geq \log t\}$$

是凸的（事实上是紧的），因为 $f_{\hat{r}}$ 在 (u, v) 中是凹的。

在经过认证的相交过程 中，我们需要一个包含所有可能见证点的紧致搜索域。一个方便的 t 充分条件是

$$W \subseteq S_{\hat{r}^{(A)}}(t) \cap S_{\hat{r}^{(B)}}(t), \qquad t = \frac{\alpha - \tau}{m}, \quad m = |\Delta_{3,n}|.$$

一个轴对齐包围盒 对于 $S_{\hat{r}}(t)$（或两个此类集合的相交），需要计算坐标极值 例如

$$u_{\max} = \sup\{u : \exists v \text{ s.t. } f_{\hat{r}}(u,v) \geq \log t\}, \qquad v_{\max} = \sup\{v : \exists u \text{ s.t. } f_{\hat{r}}(u,v) \geq \log t\},$$

并且类似地，$u_{\min}$，$v_{\min}$。边界点满足等值集方程

$$\log \kappa(\hat{r}) + \hat{r}_1 u + \hat{r}_2 v - n \log(1 + e^u + e^v) = \log t.$$

重新整理得

$$1 + e^u + e^v = \exp\left(\frac{\log \kappa(\hat{r}) + \hat{r}_1 u + \hat{r}_2 v - \log t}{n} \right),$$

其中混合了 e^u、e^v，以及关于 (u, v) 的仿射形式的指数函数。给定其中一个变量求解另一个变量会导出一个超越方程；一般不存在代数闭式。当对相交部分进行界定时，同样的障碍依然存在

$$S_{\hat{r}^{(A)}}(t) \cap S_{\hat{r}^{(B)}}(t),$$

其极值来自具有相同形式边界条件的凸优化程序。

尽管边界本身没有闭式表达式，但每个单变量切片都存在闭式最大化解。固定 u 并考虑$v \mapsto f_{\hat{r}}(u,v)$。微分可得

$$\frac{\partial f_{\hat{r}}}{\partial v}(u,v) = \hat{r}_2 - n\frac{e^v}{1+e^u+e^v}.$$

令该导数等于零，得到

$$\frac{e^v}{1+e^u+e^v} = \frac{\hat{r}_2}{n} \quad \Longrightarrow \quad e^v = \frac{\hat{r}_2}{n-\hat{r}_2}(1+e^u),$$

因此，唯一的切片最大化解为

$$v^*(u) = \log\left(\frac{\hat{r}_2(1+e^u)}{n-\hat{r}_2}\right).$$

由凹性可知，对于固定的 u，超水平集

$$\{v:\ f_{\hat{r}}(u,v) \geq \log t\}$$

要么为空，要么是一个区间

$$[v_-(u), v_+(u)],$$

其端点满足 $f_{\hat{r}}(u,v)$ = 对数 t。这些端点没有闭式解，但可以通过一维区间夹逼和二分法稳健地求得。由对称性可得，对于 $u^*(v)$，类似结论同样成立。

若需要一个纯闭式（但非常宽松）的界，可以使用以下不等式

$$1+e^u+e^v \leq 3e^{\max\{0,u,v\}},$$

这意味着

$$\log(1+e^u+e^v) \leq \log 3 + \max\{0,u,v\}.$$

代入 $f_{\hat{r}}$ 可得

$$f_{\hat{r}}(u,v) \geq \log\kappa(\hat{r}) + \hat{r}_1 u + \hat{r}_2 v - n\log 3 - n\max\{0,u,v\}.$$

因此，任何满足以下条件的 $(u,\ v)$

$$\log\kappa(\hat{r}) + \hat{r}_1 u + \hat{r}_2 v - n\log 3 - n\max\{0,u,v\} \geq \log t$$

都必定属于 $S_{\hat{r}}(t)$。该条件是分段线性的，通过分别考虑 $\max\{0,u,v\} = 0$、u 或 v 的情形，可以将其展开为显式的线性不等式。

反过来，利用下界

$$\log(1+e^u+e^v) \geq \max\{0,u,v\},$$

我们得到如下推论

$$f_{\hat{r}}(u,v) \leq \log\kappa(\hat{r}) + \hat{r}_1 u + \hat{r}_2 v - n\max\{0,u,v\}.$$

只要右侧严格小于 $\log\ t$，点 $(u,\ v)$ 就肯定在 $S_{\hat{r}}(t)$ 之外。由此可得到一个闭式（但保守的）排除区域，并可从中推导出一个显式的轴对齐外包盒。