arXiv:2304.03056v1 [math.PR] 6 Apr 2023

狄利克雷加权和的尖锐偏差界及其在贝叶斯算法分析中的应用

Denis Belomestny,PierreM´enard, Alexey Naumov, Daniil Tiapkin, Michal Valko o

2023年1月

摘要

在本文中，我们推导出了狄利克雷随机变量加权和的尖锐非渐近偏差界。这些界基于狄利克雷加权和密度的一种新型积分表示。利用几何学和复分析方法，该表示使我们能够获得和的分布的类高斯近似。我们的结果推广了开创性论文 Alfers and Dinges [1984]中针对贝塔分布得到的类似界。此外，我们的结果还可以被视为 Ganesh and O'Connell[1999] 在贝叶斯框架下研究的萨诺夫定理逆定理的一个尖锐非渐近版本。基于这些结果，我们推导出了狄利克雷过程后验均值的新偏差界，并将其应用于贝叶斯自助法。最后，我们将所得估计应用于多臂老虎机中的多项式汤普森采样（TS）算法分析，通过使现有遗憾界与臂分布的支撑集大小无关，显著改进了这些遗憾界。

1 引言

作为一种主要的多元分布，狄利克雷分布受限于多维空间中的单纯形。该分布有许多应用。组成数据的建模 [Hijazi and Jernigan, 2009], 参数和非参数贝叶斯统计 [Congdon，2014年、Ghosal and Van der Vaart, 2017], 主题建模 [Blei 等人，2003年、Teh 等人, 2006], 强化学习 [Osband 等人，2013年、Osband and Van Roy, 2017], 统计遗传学 [Lange, 1995], 可靠性 [Somerville 等人, 1997], 概率约束编程模型 [Dentcheva, 2006] 都只是它的一些应用实例。它具有多种概率刻画性质，包括条件分布、零回归和独立性特征。在这些特征中，我们希望特别指出，狄利克雷分布是贝叶斯统计中多项分布参数的共轭先验。

本文关注狄利克雷随机变量加权和的偏差界。这类加权和自然出现在贝叶斯自助法的分析中 [Rubin, 1981] 并被用作后验均值的近似。我们旨在推导这类加权和穿越概率的紧致界，使其能够最优地依赖于底层狄利克雷分布的维数和加数个数。一般而言，多维空间中的概率计算通常需要求解多重积分，是概率论中一项极具挑战性的课题。维数效应有时也称为“维数灾难”，是造成复杂性的一项关键因素。它指出，维数每增加一维，计算上都会面临显著挑战。鉴于这一现象，对多元概率作出直接的表示和界定，能够为理论研究和实际应用提供有价值的信息。

本文的贡献有三点。首先，我们推导出一种新的积分表示，用于表示狄利克雷分布随机变量加权和的密度。这一表示推广并改进了现有表示；参见 Ng 等人 [2011]的第 2 章节。其次，基于这一表示，我们得到了均值偏差的双侧高斯型界，其对狄利克雷参数之和

（“样本量”）及狄利克雷分布维度的依赖关系达到最优。若该维度等于 2，则得到贝塔分布。在这方面，我们的结果将 Alfers and Dinges [1984] 的非渐近界推广到了狄利克雷分布情形。请注意，Alfers and Dinges [1984] 中的界远比基于中心极限定理近似得到的界精确；比较研究见 Zubkov 和 Serov [2013] 。将我们的结果应用于有限支撑集上测度的贝叶斯推断，可得到 Ganesh and O'Connell [1999]所提出的所谓萨诺夫定理逆定理的非渐近版本。第三，我们应用推导出的估计来分析老虎机中的多项式汤普森采样算法。所得的依赖于问题实例的遗憾界比文献中此前已知的所有结果都紧得多；参见 Riou 和 Honda [2020]。具体而言，其首项达到最优，剩余项则与状态空间维度无关。

关于狄利克雷加权和穿越概率（即偏离均值固定幅度的概率）的现有研究成果，在文献中还不够丰富。首先，我们提及 Cho 和 Cho [2001] 针对参数相同的狄利克雷分布给出的穿越概率精确公式。遗憾的是，该公式提供的信息并不充分，而且似乎难以据此推导出恰当的界。在 Baudry 等人 [2021], 的研究中，作者给出了狄利克雷穿越概率的一些界。请注意，下界（推论 C.5.1）包含一个额外的指数因子（关于维数），使得该界在维数较大时较为粗糙。这个额外的指数因子源自对狄利克雷随机向量密度的类乘积估计。应用已知的概率论结果（例如大偏差界或集中不等式）无法得到所需的界，原因有几个。首先，狄利克雷分布的各分量具有很强的依赖性，因此所有针对独立同分布随机变量之和的概率界都不适用。其次，文献中关于相关随机变量的已知界很少，而且无法用于获得我们所需的尖锐概率穿越界。例如，Marchal 和 Arbel [2017] 证明了狄利克雷分布的次高斯性。然而，相应次高斯界中的代理方差取决于底层狄利克雷分布参数的最大值，因此对穿越概率的估计较为粗糙。

记号 设 $(\mathbf{X}\ \mathcal{X})$ 为一个可测空间， $\mathcal{P}(\mathbf{X})$为该空间上所有概率测度的集合。对于 $p \in \mathcal{P}(\mathbf{X})$，我们用 $\mathbb{E}_p$ 表示关于 p 的期望。对于随机变量 $\xi : \mathbf{X} \to \mathbb{R}$ 记号 $\xi \sim p$ 表示$\mathrm{Law}(\xi) = p$。我们也写 $\mathbb{E}_{\xi\sim p}$ 代替 $\mathbb{E}_{p}$。对于相互独立（分别地，同分布独立）的随机变量 $\xi_\ell \overset{\text{ind}}{\sim} p_\ell$（分别地，$\xi_\ell \overset{\text{i.i.d}}{\sim} p$），$\ell = 1, \ldots, d$，我们将写作 $\mathbb{E}_{\xi_\ell \overset{\text{ind}}{\sim} p_\ell}$（分别地，$\mathbb{E}_{\xi_\ell \overset{\text{i.i.d}}{\sim} p}$），以表示关于 $(\mathbf{X}^d, \mathcal{X}^{\otimes d})$ 上乘积测度的期望。对于任意$p, q \in \mathcal{P}(\mathbf{X})$，Kullback–莱布勒散度 $\mathrm{KL}(p, q)$ 定义为

$$\mathrm{KL}(p,q) = \begin{cases} \mathbb{E}_p[\log \frac{\mathrm{d}p}{\mathrm{d}q}], & p \ll q, \\ +\infty, & \text{otherwise.} \end{cases}$$

对于任意 $p \in \mathcal{P}$ $f \to \mathbb{R}$ $pf = \mathbb{E}_p[f]$ $p \in \Delta_d$ f $\{0, \ldots, d\} \to (\mathbf{X})$和: $\mathbf{X}$, 。特别地, 对于任意和: $\mathbb{R}$, $pf = \sum_{\ell=0}^{d} \ell\, f(\ell) p(\ell)$。定义 $\mathrm{Var}_p(f) = \mathbb{E}_{s'\sim p}[(f(s') - pf)^2] = p[f^2] - (pf)$。

对于 $d \in \mathbb{N}_{++}$，对于任意 $\alpha = (\alpha_0, \ldots, \alpha_d) \in \mathbb{R}_{++}^{d+1}$，我们用 $\mathrm{Dir}(\alpha)$ 表示定义在单纯形 Δ_d 上的狄利克雷分布，其前 d 个分量的密度为 $p_\alpha(x_0, \ldots, x_{d-1}) = \prod_{i=0}^{d} x_i^{\alpha_i - 1}$，并约定 $x_d = 1 - \sum_{i=0}^{d-1} x_i$。特别地，对于 $d = 1$，狄利克雷随机向量具有 $(\xi, 1-\xi)$ 的形式，对于 ξ，其服从贝塔分布，记为 $\mathrm{Beta}(\alpha_0, \alpha_1)$。

2 主要结果

设 $\mathcal{P}[0, b]$ 为所有支撑于线段 $[0, b]$ 上的概率测度构成的空间。然后，我们定义测度 $\nu \in \mathcal{P}[0, b]$ 与实数 $\mu \in [0, b]$ 的最小 Kullback– 莱布勒散度,

$$\mathcal{K}_{\inf}(\nu, \mu) = \inf\{\mathrm{KL}(\nu, \eta) : \eta \in \mathcal{P}[0, b],\, \mathbb{E}_{X\sim\eta}[X] \geq \mu\}\,. \tag{1}$$

这个量可以解释为测度 ν （或其投影）到所有期望至少为 μ 的测度集合之间的距离，其中距离由 KL散度衡量。满足优化问题（1）的测度 η 称为矩投影（M-投影）或反向信息投影（rI-投影），例如参见 [Csiszar 和 Matus，2003年，Bishop 和 Nasrabadi，2006] 以及 [Murphy，2022]。由于 KL散度并不对称，因此自然会将此类投影与更常见的信息投影（I-投影）进行比较

$$I(\nu,\mu) = \inf\{\mathrm{KL}(\eta,\nu) : \eta \in \mathcal{P}[0,b],\, \mathbb{E}_{X\sim\eta}[X] \geq \mu\}\,,$$

例如，它会出现在 Sanov型偏差界中 [Sanov，1961]。这些 I-投影具有出色的几何解释，因为 KL散度可以视为 Bregman散度。这些 M-投影并非 Bregman散度，因而缺乏几何解释。不过，当测度 ν 是样本的经验测度 [Csisz´ar and Shields，2004，引理 3。1] 时，它们与最大似然估计有着深刻联系。此外，M-投影自然地作为贝叶斯框架中大偏差原理的速率函数出现 [Ganesh and O'Connell，1999]，并且也出现在多臂老虎机的下界中，参见 [Lai and Robbins，1985，Burnetas and Katehakis，1996]。关于多臂老虎机的补充说明，请参见第 4章。

为简化后续记号，我们定义有限支撑测度的最小Kullback-Leibler距离版本。固定函数 f: $\{0,\dots,m\} \mapsto [0,b]$，并对 $p \in \Delta_m$ 和 $\mu \in \mathbb{R}$ 定义，

$$\mathcal{K}_{\inf}(p,\mu,f) = \inf\{\mathrm{KL}(p,q) : q \in \Delta_m, qf \geq \mu\}\,.$$

接下来，我们给出关于穿越概率的主要结果。这些结果总结在以下定理中。

定理 1 *For any* $\alpha = \alpha_0\ \alpha_1\ \dots\ \alpha_m \in \mathbb{R}^{m+1}_{++}$ *define* $\overline{p} \in \Delta_m$ *with* $\overline{p}\ \ell\ =\alpha_\ell/\overline{\alpha}\ \ell =$。 (,,,) 0,,,, 0 … m 0,1 *and assume that* $\alpha_0\ m =\ \alpha_0\ \alpha_m$ *and* $\alpha = \sum_{j=0}^m \alpha_j$. *Let* $\varepsilon \in$ $\geq c_0 \cdot \varepsilon^{-2}$() ,min(,) (): *for an absolute constant* $c_0 > 0$ *defined in* 7 . *Let* f $\{0,\dots,m\} \to [0,b]$ *be any mapping such* (), () *that* f 0 $= 0$ $f\ m\ = b$.

Lo下界 *Define* $\alpha^+ = (\alpha_0, \alpha_1, \dots, \alpha_{m-1}, \alpha_m + 1)$ *then for any* $\mu \in (\overline{p}f, b)$ *it hol* *ds*

$$\mathbb{P}_{w\sim\mathrm{Dir}(\alpha^+)}(wf \geq \mu) \geq (1-\varepsilon)\mathbb{P}_{\zeta\sim\mathcal{N}(0,1)}\left(\zeta \geq \sqrt{2\overline{\alpha}\,\mathcal{K}_{inf}(\overline{p},\mu,f)}\right). \tag{2}$$

Upper 上界 *Set* $\alpha^- = (\alpha_0 + 1, \alpha_1,\ \dots\ \alpha_{m-1}, \alpha_m)$ *then for any* $\mu \in (\overline{p}f, b)$ *it hol* *ds*

$$\mathbb{P}_{w\sim\mathrm{Dir}(\alpha^-)}(wf \geq \mu) \leq (1+\varepsilon)\mathbb{P}_{\zeta\sim\mathcal{N}(0,1)}\left(\zeta \geq \sqrt{2\overline{\alpha}\,\mathcal{K}_{inf}(\overline{p},\mu,f)}\right). \tag{3}$$

请注意，上述定理仅在 $\mu \in (\overline{p}f, b)$ 时成立。我们可以通过引入一个类似于 Alfers and Dinges [1984] 在贝塔分布情形下所用的额外定义，将此定理推广到任意 $\mu \in (0, b)$。首先，我们定义如下量

$$\mathcal{K}^\star_{\inf}(p,\mu,f) = \inf\{\mathrm{KL}(p,q) : q \in \Delta_m, qf = \mu\}$$

它度量了将测度 p 移动到期望为 μ 的一族测度所需的反向 $\mathcal{K}_{\inf}\ p\ b - \mu\ b - f\quad \mu < pf$ $p\ \mu\quad p\ \mu\quad p\ \mu\ =$ KL 输运代价。事实上，对于 $\mu \geq pf$，有 $\mathcal{K}^\star_{\inf}(,,f) = \mathcal{K}_{\inf}(,,f)$ 和 $\mathcal{K}^\star_{\inf}(,,f)$ (,,) for 。然后我们定义

$$A(p,\mu,f) = \mathrm{sgn}(\mu - pf)\cdot\sqrt{2\,\mathcal{K}^\star_{\inf}(p,\mu,f)}.$$

对于任意非退化的 $p \in \Delta_m$，此函数构成 $[0,b]$ 和 $\mathbb{R}$ 之间的双射。

推论 1 *For any* $\alpha = \alpha_0\ \alpha_1\ \dots\ \alpha_m \in \mathbb{R}^{m+1}_{++}$ *define* $\overline{p} \in \Delta_m$ *with* $\overline{p}\ \ell\ =\alpha_\ell/\overline{\alpha}\ \ell =$。 (,,,) (),,,, 0 … m 0,1 *and assume that* $\alpha_0\ m =\ \alpha_0\ \alpha_m$ *and* $\alpha = \sum_{j=0}^m \alpha_j$. *Let* $\varepsilon \in$ $\geq c_0 \cdot \varepsilon^{-2}$() ,min(,) () : *for an absolute constant* $c_0 > 0$ *defined in* 7 . *Let* f $\{0,\dots,m\} \to [0,b]$ *be any mapping such* () ,() *that* f 0 $= 0$ $f\ m\ = b$.

Lo下界 *Define* $\alpha^+ = (\alpha_0, \alpha_1, \ldots, \alpha_{m-1}, \alpha_m + 1)$ *then for any* $\mu \in (0, b)$ *it holds*

$$\mathbb{P}_{w\sim \mathrm{Dir}(\alpha^+)}(wf \geq \mu) \geq (1-\varepsilon)\mathbb{P}_{\zeta\sim\mathcal{N}(0,1)}\Big(\zeta \geq \sqrt{\overline{\alpha}} \cdot A(\overline{p}, \mu, f)\Big). \tag{4}$$

上界 *Set* $\alpha^- = (\alpha_0 + 1, \alpha_1, \ldots, \alpha_{m-1}, \alpha_m)$ *then for any* $\mu \in (0, b)$ *it holds*

$$\mathbb{P}_{w\sim \mathrm{Dir}(\alpha^-)}(wf \geq \mu) \leq (1+\varepsilon)\mathbb{P}_{\zeta\sim\mathcal{N}(0,1)}\Big(\zeta \geq \sqrt{\overline{\alpha}} \cdot A(\overline{p}, \mu, f)\Big). \tag{5}$$

推论 **2** *For any* $\alpha = (\alpha_0 \ldots \alpha_m) \in \mathbb{R}^{m+1}_{++}$ *define* $\overline{\alpha} = \sum_{i=0}^m \alpha_i$ *and two distributions* $\overline{p}^+ \in \Delta_m$ $\overline{p}^- \in \Delta_m$ *with*

$$\begin{aligned} \overline{p}^+(\ell) &= \alpha_\ell/(\overline{\alpha}-1), \quad \ell < m, \quad \overline{p}^+(m) = (\alpha_m - 1)/(\overline{\alpha}-1), \\ \overline{p}^-(\ell) &= \alpha_\ell/(\overline{\alpha}-1), \quad \ell > 0, \quad \overline{p}^-(0) = (\alpha_0 - 1)/(\overline{\alpha}-1). \end{aligned}$$

Assume that $\alpha_{0,m} = \min(\alpha_0, \alpha_m) \geq c_0 \cdot \varepsilon^{-2} + 1$ *for an absolute constant* $c_0 > 0$. *Then the following two-sided bound holds for any* $\mu \in (0, b)$

$$(1-\varepsilon)\mathbb{P}\Big(\zeta \geq \sqrt{\overline{\alpha}-1} \cdot A(\overline{p}^+, \mu, f)\Big) \leq \mathbb{P}(wf \geq \mu) \leq (1+\varepsilon)\mathbb{P}\Big(\zeta \geq \sqrt{\overline{\alpha}-1} \cdot A(\overline{p}^-, \mu, f)\Big),$$

where $w \sim$狄利克雷分布(α) *and* $\zeta \sim \mathcal{N}(0,1)$。

讨论 让我们讨论该定理的条件。

- 关于 $\alpha_{0,\ m}$ 是为了在 wf的密度的积分表示的渐近展开中界定余项所必需的，参见章节 A.2。在贝叶斯框架中，采用足够大的均匀先验分布即可自动满足这一条件。

- 关于权重函数 f 的条件可以通过将 u 和 f 同时适当地平移相同的常数来实现；

- .....的使用 α^+ 和 α^- 等价于Alfersand Dinges [1984]所采用的参数平移。在推论2 中，我们有上界和下界彼此接近，因为$\overline{\alpha} \to +\infty$ 因为 $\|\overline{p}^+ - \overline{p}^-\|_1 = 2/\overline{\alpha}$ 以及Hondaand Takemura [2010] 的定理 7 成立。

让我们将定理 1 的界与文献中已有的界进行比较。首先，注意到狄利克雷分布向量 $\xi \sim$狄利克雷分布(α) 的分量具有很强的依赖性。特别地，

$$\mathrm{Cov}[\xi_i, \xi_j] = -\frac{\alpha_i \alpha_j}{\overline{\alpha}^2(1+\overline{\alpha})} < 0, \quad i \neq j.$$

这意味着，对于 ξf 的方差，

$$\mathrm{Var}(\xi f) = \sum_{i\neq j} \frac{f_i f_j \alpha_i \alpha_j}{\overline{\alpha}^2(1+\overline{\alpha})} + \sum_i \frac{(\overline{\alpha} - \alpha_i)\alpha_i f_i^2}{\overline{\alpha}^2(1+\overline{\alpha})}.$$

如果 ξ 的各个分量相互独立，我们就可以将高斯逼近结果（参见例如 Fang 和 Koike [2021] 及其中的参考文献）应用于 $(\xi f - \overline{p} f)/\sqrt{\mathrm{Var}(\xi f)}$，并得到以下界

$$\mathbb{P}_{\xi\sim\mathrm{Dir}(\alpha)}(\xi f \geq \mu) = \mathbb{P}_{\xi\sim\mathrm{Dir}(\alpha)}\Big((\xi f - \overline{p} f)/\sqrt{\mathrm{Var}(\xi f)} \geq (\mu - \overline{p} f)/\sqrt{\mathrm{Var}(\xi f)}\Big).$$

然而，即使在独立情形下，如何获得涉及 $\mathcal{K}_{\inf}$ 作为偏差度量的乘法形式双侧界，仍不清楚。获得形如（2）和（3）的界的另一种可能方法是采用大偏差技术（Sanov 型界）。除了文献中关于依赖情形的结果寥寥无几这一事实外，这种技术只能得到渐近结果，因此不能直接用于获得我们所需的非渐近

界。例如，所谓的逆 Sanov 型大偏差原理由 Ganesh and O'Connell [1999]建立。 Ganesh and O'Connell [1999]在文献中考虑了一个贝叶斯推断问题，其对象是在大小为 $m+1$的空间上的有限支撑分布 $p \in \Delta_m$。取先验分布 ρ^0 为狄利克雷分布 Dir(α^0)，则后验分布 ρ^n 也为狄利克雷分布 Dir(α^n)，其中 $\sum_{j=0}^{m} \alpha_j^n \to \infty$。因此，文献 Ganesh and O'Connell [1999] 中的定理 1 表明，对于形如 $B = \{w$ 的集合 $B: \quad wf \geq \mu\}$ 其中$\mu \geq pf$， 以下渐近界成立

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}_{w\sim\rho^n}[wf \geq \mu] = \inf_{q: qf \geq \mu} \mathrm{KL}(p, q) = \mathcal{K}_{\inf}(p, \mu, f).$$

因此，我们可以将我们的定理 1 视为贝叶斯推断中这一大偏差原理的非渐近版本，其对先验分布的条件较为宽松。最后，我们提及其与加权自助法偏差界限文献之间的联系，参见例如 Broniatowski [2017] 及其中的参考文献。这些研究针对自助法经验测度（由点 x_1，……， x_n 构成）给出了 Sanov 型大偏差结果。

$$P_n^W = \sum_{i=1}^{n} W_i^n \delta_{x_i}, \quad W_i^n = \frac{Y_i}{\sum_{i=1}^{n} Y_i}$$

其中 Y_1，…， Y_n 表示一列相互独立、取非负实数值且期望为 1 的随机变量。如果我们为 Y 选择伽马分布，那么这些权重就服从狄利克雷分布。这种方法的主要问题在于，只能通过这种方式得到渐近结果，而且余项难以量化。

3 应用：狄利克雷过程后验的偏差 bounds r

考虑一个空间X 及其 σ-代数 $\mathcal{X}$，并定义 ν为(X $\mathcal{X}$)上的有限非零测度。随机过程 G由 $\mathcal{X}$ 的元素 B 索引，若其参数为ν ($G \sim$ 狄利克雷过程 (ν))，则称其为狄利克雷过程，当

$$G(B_1), \ldots, G(B_d) \sim \mathrm{Dir}(\nu(B_1), \ldots, \nu(B_d)),$$

对于任意可测划分（B_1，……， B_d），参见X，见Ferguson [1973]。

LetP$_n = n^{-1} \sum_{i=1}^{n} \delta_{Z_i}$是独立同分布样本的经验测度 $Z_1,\ldots, Z_n$，其分布为 P，定义在 (X $\mathcal{X}$) 上；给定 $Z_1, \ldots, Z_n$，令 $\widetilde{P}_n$ 服从基测度为 $\nu + n\widehat{P}_n$ 的狄利克雷过程。这里 ν 是样本空间上的有限测度（不一定是概率测度），且对所有 $\widetilde{P}_n | Z_1, \ldots, Z_n \sim$，均有 DP($\nu + n\widehat{P}_n$)，这是在为观测值的分布 $Z_1, Z_2, \ldots, Z_n$ 赋予基测度为 ν 的狄利克雷过程先验分布后得到的后验分布。完整定义和性质请参阅 Ghosal and Van der Vaart [2017] 第4章中的综述。

~我们将关注过程 $P_n g$ 对 X 上某个有界函数 g 的偏差界。值得注意的是，以下表示成立。

$$\widetilde{P}_n g = V_n Q g + (1 - V_n) \frac{\sum_{i=1}^{n} W_i g(Z_i)}{\sum_{i=1}^{n} W_i} \tag{6}$$

与 $V_n \sim$贝塔分布($|\nu|, n$) $Q \sim$狄利克雷过程(ν)，并且 $W_1, W_2, \ldots$是均值为 1 的独立同分布指数随机变量。所有变量 $V_n, Q, W_1, W_2, \ldots$均相互独立。证明参见，例如， Ghosal and Van der Vaart [2017] 中的 $\sigma = 0$。一些渐近形式（大偏差原理）的偏差界可见于Ganesh and O'Connell [2000]。

令 g为一个取值于 $[0, 1]$ 的有界函数，并假设我们知道 x_0、 x_1，使得 $g(x_0) = 0$ 以及 $g(x_1) = 1$。特别地，可以通过将这两个点添加到集合X 中来保证这一点。然后引入基测度 $\nu_\gamma = \gamma\delta_{x_0} + \gamma\delta_{x_1}$，其中 $\gamma > 0$ 是一个固定数。于是， $Q \sim$ 狄利克雷过程(ν_γ)对应于支撑集为 x_0 和 x_1上的狄利克雷分布：

$$Q = w_0 \cdot \delta_{x_0} + w_1 \cdot \delta_{x_1}, \qquad (w_0, w_1) \sim \mathrm{Dir}(\gamma, \gamma),$$

并且，特别地， g 的后验均值由下式给出

$$\widetilde{P}_n g = \sum_{i=0}^{n+1} w_i \cdot g(Z_i), \qquad w \sim \mathrm{Dir}(\gamma, 1, \ldots, 1, \gamma),$$

其中 $Z_0 = x_0$ 且 $Z_{n+1} = x_1$ 按惯例。值得注意的是，以有限支撑集的基测度作为先验分布的狄利克雷过程会导出贝叶斯自助法[Rubin, 1981]。根据推论2，对于任意 $\varepsilon > 0$，在选择 $\gamma = c_0 \cdot \varepsilon^{-2} + 1$ 时，对任意 $t \in (0, 1 - Pg)$在给定 $Z = (Z_1, \ldots, Z_n)$的条件下，以下偏差不等式成立：

$$\mathbb{P}(\widetilde{P}_n g - \widehat{P}_n g \geq t | Z) \leq (1+\varepsilon)\mathbb{P}\left(\zeta \geq \sqrt{2(n+2\gamma-1)\,\mathcal{K}_{\inf}\left(\overline{\nu}_{n,g}, \widehat{P}_n g + t\right)}\right),$$

其中 $\overline{\nu}_{n,\,g}$ 是定义在 $[0,1]$ 上的后验测度，定义如下：

$$\overline{\nu}_{n,g} = \frac{\gamma - 1}{n + 2\gamma - 1}\delta_0 + \frac{1}{n + 2\gamma - 1}\sum_{i=1}^{n} \delta_{g(Z_i)} + \frac{\gamma}{n + 2\gamma - 1}\delta_1,$$

并且我们将 $\mathcal{K}_{\inf}$ 视为所有支撑于$[0,1]$上的测度之间的最小 Kullback–莱布勒散度。根据Garivier 等人 [2022], 的引理 11，我们有

$$\mathcal{K}_{\inf}\left(\overline{\nu}_{n,g}, \widehat{P}_n g + t\right) \geq 2\left(\frac{2\gamma - 1}{n + 2\gamma - 1}\left[\widehat{P}_n g - \frac{\gamma}{2\gamma - 1}\right] + t\right)_+^2$$

这意味着后验狄利克雷过程满足 Hoeffding 型不等式。

命题 1（狄利克雷过程的 Hoeffding 型不等式）。设 $Z_1, \ldots, Z_n$ 是来自某可测空间 $(\mathbb{X}\ \mathcal{X})$ 上分布 P的样本，且设 $g\colon \mathbb{X} \to [0,1]$ 是满足 $g(x_0) = 0, g(x_1) = 1$ 的函数，其中 $x_0, x_1 \in \mathbb{X}$。此外，设 $\widetilde{P}_n$ 取自带有基测度 $\nu_\gamma = \gamma\delta_{x_0} + \gamma\delta_{x_1}$的后验分布 $\mathrm{DP}(\nu_\gamma + n\widehat{P}_n)$。那么，对于任意$\delta \in (0,1)$ $\gamma \geq c_0\varepsilon^{-2} + 1$ 以及固定的 $\varepsilon \in (0,1)$，我们有

$$\mathbb{P}\left(\widetilde{P}_n g - \widehat{P}_n g \geq \sqrt{\frac{\log((1+\varepsilon)/\delta)}{2(n+2\gamma-1)}} + \frac{\gamma}{n+2\gamma-1} \;\middle|\; Z\right) \leq \delta.$$

此外，利用一种稍微复杂一些的技术，我们可以得到一个伯恩斯坦型不等式。

~命题 2（狄利克雷过程的伯恩斯坦型不等式）。令 P_n 服从基测度为 $\nu_\gamma = \gamma\delta_{x_0} + \gamma\delta_{x_1}$的后验分布狄利克雷过程（$\nu_\gamma + n\widehat{P}_n$）。那么，对于任意$\delta \in (0,1)$ $\gamma \geq c_0\varepsilon^{-2} + 1$ 以及固定的 $\varepsilon \in (0,1)$，我们有

$$\mathbb{P}\left(\widetilde{P}_n g - \widehat{P}_n g \geq \sqrt{\frac{4\mathrm{Var}_{\widehat{P}_n}[g] \cdot \log((1+\varepsilon)/\delta)}{n+2\gamma-1}} + \frac{4\log((1+\varepsilon)/\delta) + 5\gamma}{n+2\gamma-1} \;\middle|\; Z\right) \leq \delta\,,$$

其中 $\mathrm{Var}_{\widehat{P}_n}[g] = \widehat{P}_n[g^2] - [\widehat{P}_n g]^2$ 是经验测度 $\widehat{P}_n$ 的方差。

Proof. 由于 $\mathcal{K}_{\inf}(\overline{\nu}_{n,g}, \mu)$ 关于第二个参数连续，因此对于任意 $\delta \in (0,1)$，存在 μ_δ 使得

$$\mathcal{K}_{\inf}(\overline{\nu}_{n,g}, \mu_\delta) = \frac{\log((1+\varepsilon)/\delta)}{n+2\gamma-1}.$$

令 η 为满足 $\mathcal{K}_{\inf}(\overline{\nu}_{n,\,g}、\ \mu_\delta) = \mathrm{KL}(\overline{\nu}_{n,\,g}、\ \eta)$ 的测度。根据 Honda 和 Takemura[2010], 的引理 9，该测度具有有限支撑集，并且与测度$\overline{\nu}_{n}$的支撑集$_g$ 相同，因为$\overline{\nu}_{n,\,g}\{1\} > 0$。因此， Talebi 和 Maillard [2018] 的推论 11 适用，从而得到

$$\mathbb{E}_{X\sim\eta}[X] - \mathbb{E}_{X\sim\overline{\nu}_{n,g}}[X] \leq \sqrt{2\mathrm{Var}_{X\sim\eta}[X]\,\mathrm{KL}(\overline{\nu}_{n,g}, \eta)}.$$

通过方差变化论证（例如参见Tiapkin 等人[2022]），

$$\mathrm{Var}_{X\sim\eta}[X] \le 2\mathrm{Var}_{X\sim\overline{\nu}_{n,g}}[X] + 4\,\mathrm{KL}(\overline{\nu}_{n,g},\eta) \le 2\mathrm{Var}_{\widehat{P}_n}[g] + \frac{3(2\gamma-1)}{n+2\gamma-1} + 4\,\mathrm{KL}(\overline{\nu}_{n,g},\eta)$$

我们推导出

$$\mathbb{E}_{X\sim\eta}[X] \le \widehat{P}_n + \sqrt{\frac{4\mathrm{Var}_{X\sim\overline{\nu}_{n,g}}[X]\cdot\log((1+\varepsilon)/\delta)}{n+2\gamma-1}} + \frac{(\sqrt{8}+1)\log((1+\varepsilon)/\delta)+4\gamma}{n+2\gamma-1},$$

因此，

$$\mathbb{P}\left(\widetilde{P}_n g - \widehat{P}_n g \ge \sqrt{\frac{4\mathrm{Var}_{\widehat{P}_n}[g]\cdot\log((1+\varepsilon)/\delta)}{n+2\gamma-1}} + \frac{4\log((1+\varepsilon)/\delta)+4\gamma}{n+2\gamma-1} \;\middle|\; Z\right) \le \delta.$$

□

需要强调的是，我们的界也可以用类似的方法，直接估计相应后验均值与真实均值之间的差异。

4 应用：对多项式汤普森采样的精细分析

本章节中，我们将应用定理1 来分析一种用于随机 K臂老虎机问题的算法。在这个序贯问题中，智能体在每一轮从 K 个相互独立的臂中抽取一个臂（或分布），并在长度为 T 轮的过程中获得相应样本作为奖励。遗憾衡量的是，事先知道每个臂均值的预言机所收集的累积奖励与智能体所收集的累积奖励之间的差异，是该问题中研究最广泛的性能指标之一。

在其开创性工作中，Lai 和 Robbins [1985] 为参数老虎机推导出了一个渐近的问题依赖型下界，其中各臂属于同一参数化分布族。该下界指出，任何优秀的算法都至少会遭受 $\kappa\log(T)$ 的遗憾；当 T趋于无穷大时，其中 κ是衡量问题信息复杂度的某种指标。此后，许多算法被证明具有渐近最优性，达到了 Lai 和 Robbins [1985]给出的下界。值得注意的是，著名的上置信界（KL-UCB）算法 [Agrawal, 1995, Auer et al., 2002] 对于多种指数族分布而言具有渐近最优性 [Agrawal, 1995, Burnetas and Katehakis, 1996, Garivier and Capp´e, 2011]。此外，汤普森抽样（TS，Thompson，1933）算法在实证中表现良好 [Chapelle and Li, 2011], ，也已被证明具有渐近最优性 [Korda et al., 2013, Agrawal and Goyal, 2013]。

本文研究了更具挑战性的非参数随机老虎机设定，其中我们仅假设各臂的支撑集位于单位区间内，即奖励是有界的。该设定最早由 Honda 和 Takemura [2010]考虑。特别地，他们证明了 DMED 算法达到了由Lai 和 Robbins [1985] 推广的、由 Burnetas 和 Katehakis [1996] 推广到有界奖励的非参数设定下的问题依赖型下界（另见 Honda 和 Takemura [2015]）。随后，Capp´e 等人 [2013], Garivier 等人 [2022], 利用经验似然方法，提出了KL-UCB 算法的一种扩展，即 KL-UCB，用于具有有界奖励的老虎机问题，并且该算法同样具有渐近最优性。

在最近的研究中，Riou 和 Honda [2020] 提出了一种面向TS 的扩展，适用于有界奖励，称为非参数汤普森抽样（NPTS）。该算法在每轮中为所有臂随机分配权重，计算观测奖励的加权平均值，并选择平均值最高的臂。关于该算法在有界奖励情形之外的扩展，请参见Baudry 等人 [2021] 。虽然NPTS 本身并不依赖后验抽样机制，Riou 和 Honda [2020] 声称，NPTS 仍然是一种随机化算法，但并非严格意义上的

TS 算法。不过，狄利克雷过程在6中的表示，结合$\mathsf{X} = [0,1]$ 和 $g(x) = x$ 表明，NPTS 确实是一种TS型算法，其贝叶斯模型的先验分布为基测度是狄拉克分布的狄利克雷过程 DP(δ_1)。

然而，处理非参数设定需要付出显著的时间复杂度和空间复杂度代价。事实上，所介绍的所有算法都需要存储全部观测奖励；NPTS 需要存储整个历史记录的样本权重，或者KL-UCB和 DMED 需要在每一轮求解一个凸规划，最终在第 t轮的空间复杂度至少为 $\mathcal{O}(t)$，时间复杂度为 $\mathcal{O}(t)$。这一事实与参数设定形成鲜明对比，在参数设定中，所有轮次的时间复杂度均为 $\mathcal{O}(1)$，空间复杂度均为 $\mathcal{O}(K)$。

有趣的是，论文 Riou和 Honda [2020] 也考虑了多项式奖励设置，其中臂的支撑集是同一个大小为 m的有限集合。我们可以将这一设置视为处于非参数老虎机问题与参数老虎机问题之间的边界。注意到，这一设置此前已由Capp´e 等人考虑过。对于这一设置，Riou 和 Honda[2020] 提出了多项式汤普森采样（MTS）算法，事实证明，该算法是一个在臂上采用狄利克雷先验/后验的 TS 算法。

本文为 MTS 推导了一个更精细的、依赖于问题实例的遗憾界。值得注意的是，我们的遗憾界与臂的有限支撑集大小无关。这一性质与此前针对 MTS 的遗憾界形成鲜明对比，后者随支撑集大小呈指数增长。此外，这一改进表明，基于随机舍入技术的 MTS 扩展是最优的。然而，更重要的是，这种扩展的空间复杂度为 $\mathcal{O}(K\log(T))$，每轮的时间复杂度为 $\mathcal{O}(\log(T))$。

先从问题设定的描述开始。我们考虑一个有 K 个臂的老虎机 $\bar{\nu} = (\nu_1, \dots, \nu_K)$，这些臂属于某个概率分布集合 $\mathcal{V}$，$\nu_k \in \mathcal{V}, \forall k \in [K]$。在每一轮 $t \in [T]$，智能体选择一个臂 $A^t \in [K]$，并从分布 ν_{A^t} 中独立随机抽取一个随机奖励 Y_t。衡量智能体表现的一个典型指标是期望遗憾。

$$\mathfrak{R}^T = \mathbb{E}\left[\sum_{t=1}^{T}(\mu^\star - \mu_{A^t})\right] = \sum_{a=1}^{K}(\mu^\star - \mu_a)\mathbb{E}[N_a^t],$$

其中，$\mu_a = \mathbb{E}_{Y\sim\nu_a}[Y]$ 是臂 a 的平均奖励，$\mu^\star = \max_{a\in[K]}\mu_a$ 是最优臂的平均奖励，$N_a^t = \sum_{s=1}^{t}\mathbb{1}\{A^s = a\}$ 是在第 t 轮结束时抽取臂 a 的次数。为简化起见，我们假设最优臂 $a^\star$ 唯一：$\mu^\star = \mu_{a^\star}$。

我们现在描述Lai 和 Robbins 提出的 [1985], 另请参见Burnetas 和 Katehakis [1996] 以及 Garivier 等人 [2018]。这个下界中的一个关键量是第2节中引入的所谓最小 Kullback–莱布勒散度；对于臂 a，且$\mu \in \mathbb{R}$

$$\mathcal{K}_{\inf}{}^{\mathcal{V}}(\nu_a, \mu) = \inf\{\mathrm{KL}(\nu, \eta) : \eta \in \mathcal{V},\, \mathbb{E}_{X\sim\eta}[X] \geq \mu\}\,.$$

那么，对于任何“合理的”智能体[1]，对于任何老虎机问题 $\bar{\nu}$，均有

$$\liminf_{T\to\infty}\frac{\mathfrak{R}^T}{\log T} \geq \sum_{a:\Delta_a>0}\frac{\Delta_a}{\mathcal{K}_{\inf}{}^{\mathcal{V}}(\nu_a, \mu^\star)}$$

其中，$\Delta_a = \mu^\star - \mu_a$ 是臂 a的次优性差距。在下文中，我们关注问题相关的最优算法，对于这类算法，反向不等式使用上极限而非下极限成立。我们重点关注两种特定情形：多项式奖励情形，其中每个臂都有一个具有有限支撑集的分布；以及有界奖励情形。

4.1 多项式奖励

在本章节中，我们假设臂的支撑集为集合 $\{0, 1/m, \dots, 1\}$，其中 m 为固定值。特别地，我们有 $\mathcal{V} = \mathcal{P}(\{0, 1/m, \dots, 1\})$，并将最小 Kullback 散度记为

[1]参见 Garivier 等人 [2018] 以获取正式定义。

$\mathcal{K}_{\inf}{}^{(m)}(\nu_a,\mu) = \mathcal{K}_{\inf}{}^{\mathcal{V}}(\nu_a,\mu) = \mathcal{K}_{\inf}(p_a,\mu,f)$ 其中 $p_a \in \Delta_m$ 定义为 $p_a(i) = \nu_a\{i/m\}$ 对于 $i \in \{0,\dots,m\}$，且 $f(x) = x/m$。

现在我们介绍MTS 算法，该算法由Riou 和 Honda [2020] 但我们首先需要介绍相关的贝叶斯模型。由于奖励是从多项分布中抽取的，因此适合使用狄利克雷分布作为先验分布，因为它是多项分布的共轭先验。因此，设 $\rho_a^0 = \mathrm{Dir}(\alpha_a^0)$ 是参数为 $\alpha_a^0 \in \mathbb{R}_+^{m+1}$ 、定义在单纯形 Δ_m 上的臂 a的狄利克雷先验分布。经过 t 轮后，后验分布仍为狄利克雷分布，记作

$$\rho_a^t = \qquad \alpha_a^t \qquad\qquad\qquad \alpha_a^t\ i\ = \alpha_a^0\ i\ + \sum_{s=1}^t \mathbb{1}\{A_s = \mathrm{Dir}(\ ),\ \text{其}$$

$a, Y_s\ =\ i/m\}$ $\qquad\qquad i\ =\ 0\ \dots\ m \qquad\qquad t$

参数为 ( ) ( )，对所有类别,,均如此。在第 轮，MTS 为所有臂生成后验样本$w_a^t \sim \rho_a^{t-1}$，并选择 $a \in [K]$ 臂 $A^t = \mathrm{argmax}_{a\in[K]}\, w_a^t f$，即后验样本均值最高的臂。完整描述请参见附录1 中的算法C 。

Riou 和 Honda [2020] 证明，对于足够大的 T，该算法具有最优的问题相关遗憾。然而，他们的分析实际上是渐近性的，因为遗憾界的二阶项会随支撑集的大小呈指数增长 m。这意味着 T应为 e 的量级$^{e^m}$，才能正确地刻画 T的（非渐近）依赖关系。通过在定理 1 中谨慎选择先验分布，我们在附录D 中证明了以下定理，其对 m的依赖关系锐化了许多。

定理 **2** *For MTS with Dirichlet prior parameter given by $\alpha_0^0 = \alpha_m^0 = 4c_0 + 1$ where c_0。 is a constant defined in Theorem 1 and $\alpha_i^0 = 1/\ m-2$ for all $i \in \{1,\dots,m-1\}$ for any , () ,sub-optimal arm a*

$$\mathbb{E}[N_a^T] \le \frac{\log T}{\mathcal{K}_{inf}{}^{(m)}(\nu_a,\mu^\star)} + \mathcal{O}\Big(\log^{9/10}(T)\Big)$$

and consequently

$$\mathfrak{R}^T \le \sum_{a:\Delta_a>0} \frac{\Delta_a \log(T)}{\mathcal{K}_{inf}{}^{(m)}(\nu_a,\mu^\star)} + \mathcal{O}\Big(\log^{9/10}(T)\Big).$$

In particular, the regret bound does not depend on the size of support m.

注释：Riou 和 Honda 之前的分析 [2020] （另见 Baudry 等人，2023年）依赖于一种关于狄利克雷随机变量的加权和的偏差不等式，其形式为$\mathbb{P}(w_a^t f \ge \mu) \gtrsim n^{-m}\exp(-n\,\mathcal{K}_{\inf}{}^{(m)}(\nu_a^n,\mu))$，其中 $\mu > \mu_a^n$， $n = N_a^t$表示臂 a的拉取次数， μ_a^n 表示在时间 t从臂 a 收集的 n 奖励的经验分布 ν_a^n 的均值。支撑集大小 m 的指数依赖性直接转化为关于 m呈指数增长的遗憾界。相反，利用定理 1，我们得到了形式为 $\mathbb{P}(w_a^t f \ge \mu) \gtrsim n^{-1/2}\exp(-n\,\mathcal{K}_{\inf}{}^{(m)}(\nu_a^n,\mu))$的更精确偏差不等式，从而可以得到与支撑集大小无关的遗憾界。还应注意，我们的偏差不等式与 Tiapkin 等人提供的不等式 [2022]非常接近。不过，我们要指出两点主要差异。首先，为了证明其不等式，Tiapkin 等人的作者 [2022] 需要为大小为 2 的人工奖励额外引入一个狄利克雷先验参数 α_{2m}^0。此外，该参数必须取为 $\mathcal{O}(\log(T))$的量级，而不是像我们当前设定中那样取常数。后一项改进对于推导定理 2 的界至关重要。

4.2 有界奖励

在本章节中，我们假设 $\mathcal{V} = \mathcal{P}[0,1]$ 是支撑在线段 $[0,1]$上的所有有限概率测度所构成的空间。请注意，对于这种设定， $\mathcal{K}_{\inf}{}^{\mathcal{V}}(\nu_a,\mu) = \mathcal{K}_{\inf}(\nu_a,\mu)$ 按照第 2 章中针对 $b = 1$的定义。接下来，我们通过 Riou 和 Honda [2020], 讨论的随机舍入方法，将 MTS 算法扩展到有界奖励设定[2013]（另请参阅 Agrawal 和 Goyal）。固定大小 m，并考虑有限网格 $\{0, 1/m, \dots, 1\}$。如果对某个 $i \in \{0,\dots,m-1\}$， $Y \in [i/m, (i+1)/m)$，我们就抽取新的奖励 $\widetilde{Y} = (i+B)/m$，使其服从 Ber($mY - i$)，其中 $B \sim$，并且该奖励与 Y具有相同的期望，且服从多项分布。然后，我们只需将变换后的奖励 $\widetilde{Y}$输入 MTS 算法即可。我们将该算法称为舍入多项式汤普森采样（RMTS），并在附录 C 的算法2中给出完整描述。

我们用 $\nu_a^{(m)}$表示来自臂 a的变换后奖励 $\widetilde{Y}$的分布。根据Riou 和 Honda 提出的引理 6 [2020], , 我们可以量化舍入过程对最小 Kullback–莱布勒散度的影响

$$\mathcal{K}_{\inf}(\nu_a, \mu^\star) - \frac{1}{m(1-\mu^\star)-1} \leq \mathcal{K}_{\inf}{}^{(m)}(\nu_a^{(m)}, \mu^\star) \leq \mathcal{K}_{\inf}(\nu_a, \mu^\star)\,.$$

因此，如果我们选择一个阶为 $m = \mathcal{O}(\log\ T)$ 的网格大小，结合前述不等式和 MTS 的遗憾界，我们便证明了在有界奖励的情形下，RMTS 达到了问题相关的最优性。我们强调，只有在 MTS 的遗憾界不依赖于支撑集大小 m 时，才能得到这样的结果。

定理 3. *For* ***RMTS****, with the same prior as in Theorem 2, it holds for all sub-optimal armsa,*

$$\mathbb{E}[N_a^T] \leq \frac{\log T}{\mathcal{K}_{inf}(\nu_a, \mu^\star)} + \mathcal{O}\Big(\log^{9/10}(T)\Big)$$

and consequently

$$\mathfrak{R}^T \leq \sum_{a:\Delta_a>0} \frac{\Delta_a \log(T)}{\mathcal{K}_{inf}(\nu_a, \mu^\star)} + \mathcal{O}\Big(\log^{9/10}(T)\Big).$$

有趣的是，RMTS 只需要 $\mathcal{O}(K\log(T))$ 量级的空间复杂度，每轮的时间复杂度为 $\mathcal{O}(K\log(T))$ 量级。相比之下，NPTS 的空间复杂度为 $\mathcal{O}(T)$ 量级，时间复杂度为 $\mathcal{O}(T)$ 量级；在保持问题相关最优性的同时，我们观察到了显著的改进。

注 1。RMTS 算法假设时间范围 T 已提前知晓。然而，在 T 未知的情况下，我们可以使用倍增技巧，同时保持相同的时间复杂度。修改后的算法如下：我们记录从真实模型中获得的所有奖励，但在达到 2^k 回合后，通过对每个 2^k 样本再次应用舍入过程，将支撑集的大小从 k 增加到 $k+1$。在 T回合后，舍入过程的总成本不超过 $\sum_{k=0}^{\lfloor \log_2 T \rfloor} 2^k \leq 2T$。这意味着重新舍入步骤的摊销每轮时间复杂度仅为 $\mathcal{O}(1)$。尽管这一修改将空间复杂度增加到 $\mathcal{O}(T)$，但在时间范围 T未知的情况下，仍保持原有的 $\mathcal{O}(K\log(T))$时间复杂度。

5 主要结果的证明

在本章节中，我们给出关于狄利克雷投影的偏差界主要结果的证明：定理 1 和推论 1。

5.1 定理的证明1

首先基于狄利克雷加权和的密度的积分表示（命题 3）及其非渐近展开（命题 4），我们证明以下引理。

引理 **1** *Consider a function f* $0 = 0$ *and f* $m = b$ *and.: a fixed positive vector* $\alpha = \alpha_0\ \alpha_1\ \ldots\ \alpha_m \in \mathbb{R}_+^{m+1}$. *Define* $\overline{p} \in \Delta_m$ *such that* $\overline{p}\ \ell$ $\{0,\ldots,m\} \to [0,b]$ *such that* $f(\,)\,(\,),(\,,\,,)\,0$ $= \alpha_\ell/\alpha$ *for* $\overline{\alpha} = \sum_{j=0}^m \alpha_j$. *Let* $\varepsilon > 0$ *and assume that* $\alpha_{0\ m} =$ $\{\alpha_0, \alpha_n\} \geq c_0 \cdot \varepsilon^{-2}$ *where* ,min ,

$$c_0 = \frac{2}{\pi}\left(8 + \frac{49\sqrt{6}}{9}\right)^2. \tag{7}$$

- *Let* $w^+ \sim$狄利克雷分布(α^+) *for* $\alpha^+ = (\alpha_0, \alpha_1, \ldots, \alpha_{m-1}, \alpha_m + 1)$. *Then for any* $u \in (\overline{p}f, b)$ *the density of the random variable* $Z^+ = w^+ f$ *could be lower-bounded as follows*

$$p_{Z^+}(u) \geq (1-\varepsilon)\sqrt{\frac{\overline{\alpha}}{2\pi(1-\lambda^\star(b-u))^2\sigma^2}} \cdot \exp(-\overline{\alpha}\,\mathcal{K}_{inf}(\overline{p}, u, f)).$$

- *Let* $w^- \sim$狄利克雷分布(α^-) *for* $\alpha^- = (\alpha_0 + 1, \alpha_1, \ldots, \alpha_{m-1}, \alpha_m)$. *Then for any* $u \in (\overline{p}f, b)$ *the density of a random variable* $Z^- = w^- f$ *could be upper-bounded as follows*

$$p_{Z^-}(u) \leq (1+\varepsilon)\sqrt{\frac{\overline{\alpha}}{2\pi(1+\lambda^\star u)^2\sigma^2}} \cdot \exp(-\overline{\alpha}\mathcal{K}_{inf}(\overline{p}, u, f)).$$

Proof. 我们结合命题 3 和命题 4 来证明 $\ell = m$

$$p_{Z^+}(u) \geq \frac{\overline{\alpha}}{\sqrt{2\pi}}\left(1 - \frac{c_1 \cdot \mathrm{e}^{-c_\kappa \alpha_{0,m}}}{\sqrt{2\pi\alpha_{0,m}}} - \frac{c_2}{\sqrt{2\pi\alpha_{0,m}}} - \frac{c_3 \cdot \mathrm{e}^{-c_\kappa \alpha_{0,m}}}{\sqrt{2\pi}}\right)\frac{\exp(-\overline{\alpha}\mathcal{K}_{\inf}(\overline{p}, u, f))}{(1-\lambda^\star(b-u))\sqrt{\overline{\alpha}\cdot\sigma^2}}.$$

目标　　是为 $\alpha_{0,m}$ 找到一个下界，使得以下不等式成立。

$$\frac{c_1 \cdot \exp(-c_\kappa \alpha_{0,m})}{\sqrt{2\pi c_\kappa \alpha_{0,m}}} + \frac{c_2}{\sqrt{2\pi\alpha_{0,m}}} + \frac{c_3 \cdot \exp(-c_\kappa \alpha_{0,m})}{\sqrt{2\pi}} \leq \varepsilon. \tag{8}$$

注意，对于较大的 $\alpha_{0,m}$，第二项占主导地位。首先，我们需要保证

$$\frac{c_2}{\sqrt{2\pi\alpha_{0,m}}} \leq \varepsilon/2,$$

这对 $\alpha_{0,m} \geq \frac{4c_2^2}{2\pi\varepsilon^2}$ 成立。要满足式 (8)，只需保证

$$\left(\frac{c_1}{2c_2\sqrt{c_\kappa}} \cdot \varepsilon + \frac{c_3}{\sqrt{2\pi}}\right)\exp(-c_\kappa\alpha_{0,m}) \leq \varepsilon/2 \iff \alpha_{0,m} \geq \frac{1}{c_\kappa}\log\left(\frac{c_1}{c_2\sqrt{c_\kappa}} + \frac{2c_3}{\sqrt{2\pi}} \cdot \frac{1}{\varepsilon}\right).$$

接下来，注意到对于任意 $\varepsilon \in (0, 1)$，由于 c_1、c_2、c_3、c_κ 的特定取值，以下不等式成立

$$\frac{2c_2^2}{\pi\varepsilon^2} \geq \frac{1}{c_\kappa}\log\left(\frac{c_1}{c_2\sqrt{c_\kappa}} + \frac{2c_3}{\sqrt{2\pi}} \cdot \frac{1}{\varepsilon}\right).$$

因此，对于 $\alpha_{0,\ m} \geq 2c_2^2/(\pi\varepsilon^2)$，引理 1 的第一个结论成立。对于第二个结论，只需对 $\ell = 0$ 使用命题 3 和命题 4 的相同组合。

□

在给出最终证明之前，我们先推导一个重要的技术结果。

引理2. *For any* $u \in (\overline{p}f, b)$ *it holds*

$$\frac{1}{2}(\lambda^\star)^2\sigma^2\big(1-\lambda^\star(b-u)\big)^2 \leq \mathcal{K}_{inf}(\overline{p}, u, f) \leq \frac{1}{2}(\lambda^\star)^2\sigma^2\big(1+\lambda^\star u\big)^2.$$

Proof. 定义函数 $\varphi_u(\lambda) = \mathbb{E}[\log(1-\lambda(f(X)-u))]$ 以及 $\lambda_u = \lambda^\star$。注意到 $\sigma^2 = -\varphi_u''(\lambda_u)$。根据 φ_u 的泰勒展开以及 λ_u 的定义，有

$$0 = \varphi_u(0) = \varphi_u(\lambda_u) + 0 + \frac{\lambda_u^2}{2}\varphi_u''(y\lambda_u) \iff \varphi_u(\lambda_u) = \frac{\lambda_u^2}{2}(-\varphi_u''(y\lambda_u))$$

对于某个 $y \in (0, 1)$。我们将从两侧界定上文出现的负二阶导数。首先，注意到

$$-\varphi_u''(y\lambda_u) = \mathbb{E}\left[\frac{(f(X)-u)^2}{\big(1-\lambda_u(f(X)-u)\big)^2}\left(\frac{1-\lambda_u(f(X)-u)}{1-y\lambda_u(f(X)-u)}\right)^2\right].$$

我们分两种情况讨论。

1) $f(X) \leq u$。利用 $y \in (0,1)$ 这一事实，我们可以推导出以下界限。

$$1 + \lambda_u u \geq \frac{1-\lambda_u(f(X)-u)}{1-y\lambda_u(f(X)-u)} \geq 1 \geq 1 - \lambda_u(b-u).$$

2) $f(X) \geq u$。通过对称论证

$$1+\lambda_u u \geq 1 \geq \frac{1-\lambda_u(f(X)-u)}{1-y\lambda_u(f(X)-u)} \geq 1-\lambda_u(b-u)\,.$$

特别是，利用 $\varphi''(\lambda_u) = -\sigma^2$ 的定义，可得

$$\sigma^2\big(1-\lambda_u(b-u)\big)^2 \leq -\varphi_u''(y\lambda_u) \leq \sigma^2\big(1+\lambda_u u\big)^2\,.$$

将此不等式代入 φ_u 的积分表示中，即可得出该结论。

□

利用这一引理，我们可以继续证明最终结果。

Proof of Theorem 1. 定义 $Z^+ = wf$ 对于 $w \sim$狄利克雷分布(α^+)。由引理 1，

$$\mathbb{P}_{w\sim\mathrm{Dir}(\alpha^+)}[wf \geq \mu] = \int_\mu^b p_{Z^+}(u)\mathrm{d}u \geq (1-\varepsilon)\sqrt{\frac{\overline{\alpha}}{2\pi}}\cdot\int_\mu^b \frac{\exp(-\overline{\alpha}\,\mathcal{K}_{\inf}(\overline{p},u,f))}{\sqrt{\sigma^2(1-\lambda^\star(b-u))^2}}\,\mathrm{d}u.$$

根据定理 6（由Honda 和 Takemura 提出） [2010]，

$$\frac{\partial}{\partial u}\,\mathcal{K}_{\inf}(\overline{p},u,f) = \lambda^\star.$$

因此，我们可以定义一个变量替换 $t^2/2 = \mathcal{K}_{\inf}(\overline{p},u,f)$ $t\mathrm{d}t = \lambda^\star \mathrm{d}u$ 并写出

$$\mathbb{P}\big[Z^+ \geq \mu\big] \geq (1-\varepsilon)\int_{\sqrt{2\,\mathcal{K}_{\inf}(\overline{p},\mu,f)}}^{+\infty} D^+(u)\sqrt{\frac{\overline{\alpha}}{2\pi}}\exp(-\overline{\alpha}t^2/2)\mathrm{d}t,$$

其中 $D^+(u)$ 定义如下

$$D^+(u) = \sqrt{\frac{2\,\mathcal{K}_{\inf}(\overline{p},u,f)}{(\lambda^\star)^2\sigma^2(1-\lambda^\star(b-u))^2}}.$$

根据引理 2， $D^+(u) \geq 1$ 因此，对于 $g \sim \mathcal{N}(0,1)$

$$\mathbb{P}\big[Z^+ \geq \mu\big] \geq (1-\varepsilon)\int_{\sqrt{2\,\mathcal{K}_{\inf}(\overline{p},\mu,f)}}^{+\infty}\sqrt{\frac{\overline{\alpha}}{2\pi}}\mathrm{e}^{-\overline{\alpha}t^2/2}\mathrm{d}t = (1-\varepsilon)\mathbb{P}\big(g \geq \sqrt{2\overline{\alpha}\,\mathcal{K}_{\inf}(\overline{p},\mu,f)}\big).$$

对于 陈述的第二部分，定义 $Z^- = wf$ 对于 $w \sim$狄利克雷分布(α^-)。根据引理 1 ，

$$\mathbb{P}_{w\sim\mathrm{Dir}(\alpha^-)}[wf \geq \mu] = \int_\mu^b p_{Z^-}(u)\mathrm{d}u \leq (1+\varepsilon)\sqrt{\frac{\overline{\alpha}}{2\pi}}\cdot\int_\mu^b \frac{\exp(-\overline{\alpha}\,\mathcal{K}_{\inf}(\overline{p},u,f))}{\sqrt{\sigma^2(1+\lambda^\star u)^2}}\,\mathrm{d}u.$$

通过变量替换 $t^2/2 = \mathcal{K}_{\inf}(\overline{p},u,f)$ $t\mathrm{d}t = \lambda^\star\mathrm{d}u$

$$\mathbb{P}\big[Z^- \geq \mu\big] \leq (1+\varepsilon)\int_{\sqrt{2\,\mathcal{K}_{\inf}(\overline{p},\mu,f)}}^{+\infty} D^-(u)\sqrt{\frac{\overline{\alpha}}{2\pi}}\exp(-\overline{\alpha}t^2/2)\mathrm{d}t,$$

其中 $D^-(u)$ 定义如下

$$D^-(u) = \sqrt{\frac{2\,\mathcal{K}_{\inf}(\overline{p},u,f)}{(\lambda^\star)^2\sigma^2(1+\lambda^\star u)^2}}.$$

由引理 2， $D^-(u) \leq 1$ 因此

$$\mathbb{P}\big[Z^- \geq \mu\big] \leq (1+\varepsilon)\mathbb{P}_{g\sim\mathcal{N}(0,1)}\big[g \geq \sqrt{2\overline{\alpha}\,\mathcal{K}_{\inf}(\overline{p},\mu,f)}\big].$$

□

5.2 推论的证明1

Proof. 由定理 1 两个结论对 $\mu \geq \overline{p}f$均成立。因此，我们可以假设 $\mu < \overline{p}f$从而

$$A(\overline{p}, \mu, f) = -\sqrt{2\,\mathcal{K}_{\inf}(\overline{p}, b-\mu, b-f)}.$$

让我们从下界开始。对于 $w \sim$狄利克雷分布(α^+)，我们有

$$\mathbb{P}(wf \geq \mu) = 1 - \mathbb{P}(wf \leq \mu) = 1 - \mathbb{P}(w(b-f) \geq b-\mu). \tag{9}$$

注意，对于函数 $b-f$，0 和 m 原子$b-f)(0) = b$ 以及 $(b-f)(m) = 0$的作用发生了变化。为了得到右侧概率的上界，我们可以直接应用定理 1中的上界。

对于任意向量 $x \in \mathbb{R}_+^{m+1}$，定义反转操作 rev(x)，使得 $y =$rev$(x) \iff y_i = x_{m-i}$ 对于任意 $i \in \{0, \ldots, m\}$。然后定义向量 $\beta =$rev(α)，对于该向量，我们可以通过将 β_0 和 β_m 的值分别增加 1 来定义 β^- 和 β^+。此外，定义测度 $\bar{q} \in \Delta_m$，使得 $\bar{q} =$ rev$(\overline{p})$，以及函数 $g =$ rev(f)。对于 $w \sim$ 狄利克雷分布(α^+)，我们有 $w' =$ rev(w) 服从狄利克雷分布(β^-)，并且，$w'(b-g) = w(b-f)$ 以及 $\bar{q}(b-g) \leq b-\mu$，因此定理 1 适用

$$\mathbb{P}_{w' \sim \mathrm{Dir}(\beta^-)}[w'(b-g) \geq b-\mu] \leq (1+\varepsilon)\mathbb{P}\Big[\zeta \geq \sqrt{2\overline{\alpha}\,\mathcal{K}_{\inf}(\bar{q}, b-g, b-\mu)}\Big]$$

对于 $\zeta \sim \mathcal{N}(0,1)$。注意到 $\mathcal{K}_{\inf}(\bar{q}, b-g, b-\mu) = \mathcal{K}_{\inf}(\overline{p}, b-f, b-\mu)$，我们得到以下上界

$$\mathbb{P}(w(b-f) \geq b-\mu) \leq (1+\varepsilon)\mathbb{P}\Big[\zeta \geq -\sqrt{\overline{\alpha}} \cdot A(\overline{p}, \mu, f)\Big].$$

因此，由于 $A(\overline{p}, \mu, f) < 0$，且 ζ 是对称的，我们有

$$\mathbb{P}(wf \geq \mu) \geq 1 - (1+\varepsilon)\mathbb{P}\Big[\zeta \geq -\sqrt{\overline{\alpha}} \cdot A(\overline{p}, \mu, f)\Big] \geq (1-\varepsilon)\mathbb{P}\Big[\zeta \geq \sqrt{\overline{\alpha}} \cdot A(\overline{p}, \mu, f)\Big].$$

上界也同样成立：从表示式（9）出发，并利用反转模型下 $\mathbb{P}[w(b-f) \geq b-\mu]$ 的概率下界。 □

附录

目录

引理证明1 15A.1 狄利克雷分布加权和的密度.15A.2 鞍点法.19B 线性代数基础 25C MTS算法 26D 遗憾界证明 27D.1 定理证明 2.27D.2 定理证明 3.28D.3 命题证明 5-7 .29E 技术引理 33

引理的证明1

在本章节中，我们将经常使用以下记号。令 F

$[0,b] \mid f\ 0\ =0\ f\ m\ =b\}$ $\alpha = \alpha_0\ \alpha_1\ \ldots\ \alpha_m\ \in \mathbb{R}_{++}^{m+1}$ $_m\ b\ =\{f\ \ \{0,\ldots,m\} \to (\,):(\,),(\,).$ 对任意 $(,,,)$，定义 $(\,)$，使得 $\overline{p} = \overline{p}\ \alpha\ \in \Delta_m$

$\overline{p}(\ell)\ = \alpha_\ell/\overline{\alpha}, \ell = 0,\ldots,m,$ 其中 $\overline{\alpha} = \sum_{j=0}^m \alpha_j$。

A.1 狄利克雷分布加权和的密度

本章节中，我们计算随机变量 $Z = wf$ 的密度，其中 $w \sim$ 狄利克雷(α) 且 $f \in \mathrm{F}_m(b)$。

命题 3。令 $f \in \mathrm{F}_m(b)$ 且 $\alpha = (\alpha_0, \alpha_1, \ldots, \alpha_m) \in \mathbb{R}_+^{m+1}$，使得 $\overline{\alpha} = \sum_{j=0}^m \alpha_j > 1$。假设 Z 不退化。那么对于任意 $0 \le u < b$

$$p_Z(u) = \frac{\overline{\alpha}-1}{2\pi} \int_{\mathbb{R}} \prod_{j=0}^m (1 + \mathrm{i}(f(j)-u)s)^{-\alpha_j}\,\mathrm{d}s.$$

命题的证明 3 将在本段末尾给出。

函数 $g\colon \mathbb{R}^{m+1} \to \mathbb{R}$ 被称为一个次数为

$t\qquad \gamma > 0\qquad x \in A\qquad g\ \gamma x\ = \gamma^t g\ x\qquad \widetilde{\Delta}_m =\qquad 0\ \Delta_m\ = \{w \in$ 的正齐次函数，

$\mathbb{R}_+^{m+1}\quad \sum_{j=0}^m w_j \le 1\}\qquad\qquad \Delta_m\qquad\qquad 0\qquad r > 0$

定义在锥 $A \subseteq \mathbb{R}^{m+1}$ 上，即对于任意，我们有 $(\,)(\,)$。将 Conv$(,)$: 定义为一个底面为、顶点位于 的棱锥。对于，我们记 $\Delta_m(r) = \{w \in \mathbb{R}_+^{m+1} : \sum_{\ell=0}^m w_\ell = r\}$。那么，显然 $\Delta_m = \Delta_m(1)$。对于任意 $a \in \mathbb{R}^{m+1}$ 定义 H

$_a = \{w \in \mathbb{R}^{m+1} :\ \langle a, w\rangle = 0\}$。此外，对于任意矩阵 $M \in \mathbb{R}^{m+1\times k}$ 定义 $\mathrm{H}_M = \{w \in \mathbb{R}^{m+1} : Mw = 0\}$。

对于任意维数为 $n < m+1$ 的可测集 $A \subseteq \mathbb{R}^{m+1}$ 和任意函数 $g\colon \mathbb{R}^{m+1} \to \mathbb{R}$，定义

$$I_n(g, A) = \int_A g(w)\mathcal{H}^n(\mathrm{d}w),$$

其中 $\mathcal{H}^n$ 是 n维豪斯多夫测度（定义见 Evans 和 Garzepy，2018年）。若 $A = \mathcal{L}(Y)$ 对于线性映射 $\mathcal{L}\colon \mathbb{R}^n \to \mathbb{R}^{m+1}$ 和 $Y \subseteq \mathbb{R}^n$成立，则可写为

$$I_n(g, \mathcal{L}(Y)) = [\mathcal{L}] \cdot \int_Y g(\mathcal{L}(y))\lambda_n(\mathrm{d}y),$$

其中 λ_n 是定义在 Y 上的 n 维勒贝格测度，$[\mathcal{L}]$ 是映射 $\mathcal{L}$ 的雅可比矩阵，可计算为 $[\mathcal{L}] = \sqrt{\det(\mathcal{L}\mathcal{L}^\intercal)}$。定义仿射映射 $\mathcal{L}_a^t\colon \mathbb{R}^m \to \mathbb{R}^{m+1}$，将 $\mathbb{R}^m$ 变换为 H_a^t，具体而言，将 $x_1, \ldots, x_m$ 映射到 $w_1, \ldots, w_m$，并且对 $a_0 > 0$ 有 $w_0 = \frac{t - \sum_{j=1}^m a_j x_j}{a_0}$（不失一般性）。该映射的线性部分的雅可比矩阵等于 $[\mathcal{L}_a^t] = \frac{\|a\|_2}{a_0}$（见引理 6）。此外，定义 $\mathcal{L}_a = \mathcal{L}_{a^\circ}^0$ 定义 $\mathcal{L}_M$ 为映射到 $\mathcal{H}_M$ 的映射，其中矩阵为 M；定义见附录 B。对于两个向量 $a, b \in \mathbb{R}^{m+1}$，我们定义 $\mathcal{H}^{t_1,t_2}_{[a,b]} = \{x \in \mathbb{R}^{m+1} :\ \langle a, x\rangle = t_1, \langle b, x\rangle = t_2\}$，以及映射到该集合的对应规范映射 $\mathcal{L}^{t_1,t_2}_{[a,b]}$。注意，该映射仅取决于空间 span$(a, b)$，而不取决于向量本身。

引理

3 *Let g be a positively homogeneous function of degree $t > -m$ on $\mathbb{R}_{++}^{m+1}$ and assume 。that $a_0 \ne a_1$. Then we have*

$$I_m(g, \widetilde{\Delta}_m \cap \mathrm{H}_a) = \frac{\|a\|_2}{|a_0 - a_1| \cdot [\mathcal{L}_{[a,\mathbf{1}]}]} \frac{I_{m-1}(g, \Delta_m \cap \mathrm{H}_a)}{m+t}.$$

Proof. 首先，我们通过映射来应用变量替换公式 [Evans 和 Garzepy，2018年，3.3。3] $\mathcal{L}_a$

$$I_m(g, \widetilde{\Delta}_m \cap \mathrm{H}_a) = [\mathcal{L}_a] \int_{\mathcal{L}_a(x)\ge 0,\ \mathbf{1}^\intercal \mathcal{L}_a(x) \le 1} g(\mathcal{L}_a(x))\lambda^m(\mathrm{d}x).$$

我们可以定义一个映射，使其与 $\mathcal{L}_a^\intercal \mathbf{1}$ 构成标量积，并应用余面积公式 [Evans 和 Garzepy，2018年，3.4。3]

$$I_m(g, \widetilde{\Delta}_m \cap \mathrm{H}_a) = \frac{[\mathcal{L}_a]}{[\mathcal{L}_a^\intercal \mathbf{1}]} \int_0^1 \mathrm{d}t \int_{\mathcal{L}_a(x)\geq 0,\ \mathbf{1}^\intercal \mathcal{L}_a(x)=t} g(\mathcal{L}_a(x)) \mathcal{H}^{m-1}(\mathrm{d}x).$$

接下来需要计算内层积分。为此，我们定义一个映射 $\mathcal{L}^t_{\mathcal{L}^\intercal a \mathbf{1}}$，它将 $\mathbb{R}^m$ 映射到集合 $\{x : \mathbf{1}^\intercal \mathcal{L}_a(x) = t\}$。注意，该映射的雅可比矩阵不依赖于参数 t。因此，根据变量替换公式，我们有

$$\begin{aligned}\int_{\mathcal{L}_a(x)\geq 0,\ \mathbf{1}^\intercal \mathcal{L}_a(x)=t} g(\mathcal{L}_a(x)) \mathcal{H}^{m-1}(\mathrm{d}x) &= [\mathcal{L}_{\mathcal{L}_a^\intercal \mathbf{1}}] \int_{\mathcal{L}^{0,t}_{[a,\mathbf{1}]}(y)\geq 0} g(\mathcal{L}^{0,t}_{[a,\mathbf{1}]}(y)) \lambda^{m-1}(\mathrm{d}y) \\ &= \frac{[\mathcal{L}_{\mathcal{L}_a^\intercal \mathbf{1}}]}{[\mathcal{L}_{[a,\mathbf{1}]}]} I_{m-1}(g, \Delta_m(t) \cap \mathrm{H}_a).\end{aligned}$$

这里我们利用引理7 来简化映射 $\mathcal{L}_a \circ \mathcal{L}^t_{\mathcal{L}_a \mathbf{1}} = \mathcal{L}^{0,\ t}_{[a,\mathbf{1}]}$。根据正齐次函数的定义以及豪斯多夫测度 $\mathcal{H}^{m-1}$ 的性质，我们得到

$$\begin{aligned}I_m(g, \widetilde{\Delta}_m \cap \mathrm{H}_a) &= \frac{[\mathcal{L}_a][\mathcal{L}_{\mathcal{L}_a^\intercal \mathbf{1}}]}{[\mathcal{L}_{[a,\mathbf{1}]}][\mathcal{L}_a^\intercal \mathbf{1}]} \int_0^1 I_{m-1}(g, \Delta_m(t) \cap \mathrm{H}_a) \mathrm{d}t \\ &= \frac{[\mathcal{L}_a][\mathcal{L}_{\mathcal{L}_a^\intercal \mathbf{1}}]}{[\mathcal{L}_{[a,\mathbf{1}]}][\mathcal{L}_a^\intercal \mathbf{1}]} \frac{I_{m-1}(g, \Delta_m \cap \mathrm{H}_a)}{m+t}.\end{aligned}$$

接下来我们简化该表达式，使用映射到超平面的雅可比矩阵的精确公式（引理6）

$$\frac{[\mathcal{L}_a][\mathcal{L}_{\mathcal{L}_a^\intercal \mathbf{1}}]}{[\mathcal{L}_a^\intercal \mathbf{1}]} = \frac{\|a\|_2 \|\mathcal{L}_a^\intercal \mathbf{1}\|_2}{|a_0| \cdot |(\mathcal{L}_a^\intercal \mathbf{1})_0| \cdot \|\mathcal{L}_a^\intercal \mathbf{1}\|_2} = \frac{\|a\|_2}{|(\mathcal{L}_a^\intercal \mathbf{1})_0| \cdot |a_0|}$$

其中，不失一般性，我们假设 $(\mathcal{L}_a^\intercal \mathbf{1})_0 \neq 0$ 和 $a_0 \neq 0$。要最终简化该表达式，只需注意 $(\mathcal{L}_a^\intercal \mathbf{1})_0 = (a_0 - a_1)/a_0$。 □

接下来，我们给出积分 $I_m(g, \widetilde{\Delta}_m \cap \mathrm{H}_a)$的另一种表示。我们遵循Lasserre [2020]，并采用同样基于拉普拉斯变换的技巧。

引理 4 *Let g be a positively homogeneous function of degree t on $\mathbb{R}^{m+1}_{++}$ such that $t > -(1+m)$ and $\int_{\widetilde{\Delta}_m} |g(w)| \mathrm{d}w < \infty$. Then*

$$I_m(g, \widetilde{\Delta}_m \cap \mathrm{H}_a) = \frac{1}{\Gamma(1+m+t)} \int_{\mathbb{R}^{m+1}_+ \cap \mathrm{H}_a} g(w) \exp\left(-\sum_{\ell=0}^m w_\ell\right) \mathcal{H}^m(\mathrm{d}w).$$

Proof. 考虑 $h(y) = \int_{w\geq 0,\ \langle \mathbf{1},\ w\rangle \leq y,\ \langle a,\ w\rangle = 0} g(w) \mathcal{H}^m(\mathrm{d}w)$。显然，$h(1) = I_m(g, \widetilde{\Delta}_m \cap \mathrm{H}_a)$。由于 g 是正齐次函数，我们得到 $h(y) = y^{m+t} h(1)$。这意味着，h 的拉普拉斯变换等于 $L(\lambda) = \int_0^\infty h(y) \mathrm{e}^{-\lambda y} \mathrm{d}y = h(1) \frac{\Gamma(m+t+1)}{\lambda^{m+t+1}}$。另一方面，拉普拉斯变换 $L(\lambda)$可以通过对子空间进行线性参数化来计算

$\langle a,\ w\rangle = 0$ 使用映射 $\mathcal{L}_a$ 和富比尼定理

$$\begin{aligned} L(\lambda) &= \int_0^\infty \mathrm{e}^{-\lambda y}\left[\int_{w\in\mathbb{R}_+^{m+1},\langle \mathbf{1},w\rangle\le y,\langle a,w\rangle=0} g(w)\mathcal{H}^m(\mathrm{d}w)\right]\mathrm{d}y \\ &= [\mathcal{L}_a]\int_{\mathcal{L}_a(x)\ge 0}\mathrm{d}x\cdot g(\mathcal{L}_a(x))\cdot\left[\int_{\langle \mathbf{1},\mathcal{L}_a(x)\rangle\le y}\mathrm{e}^{-\lambda y}\mathrm{d}y\right] \\ &= \frac{[\mathcal{L}_a]}{\lambda}\int_{\mathcal{L}_a(x)\ge 0} g(\mathcal{L}_a(x))\exp(-\lambda\langle \mathbf{1},\mathcal{L}_a(x)\rangle)\mathrm{d}x \\ &= \frac{[\mathcal{L}_a]}{\lambda^{m+t+1}}\int_{\mathcal{L}_a(x)\ge 0} g(\mathcal{L}_a(x))\exp(-\langle \mathbf{1},\mathcal{L}_a(x)\rangle)\mathrm{d}x \\ &= \frac{1}{\lambda^{m+t+1}}\int_{\mathbb{R}_+^{m+1}\cap \mathrm{H}_a} g(w)\exp\left(-\sum_{\ell=0}^m w_\ell\right)\mathcal{H}^m(\mathrm{d}w). \end{aligned}$$

通过确定拉普拉斯变换的两种计算方法，我们完成了证明。 □

现在我们计算引理4右侧的积分。我们将采用傅里叶变换方法，并沿用Dirksen [2015]的方法。

引理 5. *Let* $g(w) = w_0^{\alpha_0-1}\cdot\ldots\cdot w_m^{\alpha_m-1}$. *Then we have*

$$\int_{\mathbb{R}_+^{m+1}\cap \mathrm{H}_a} g(w)\exp\left(-\sum_{i=0}^m w_i\right)\mathcal{H}^m(\mathrm{d}w) = \frac{\|a\|_2\cdot\prod_{j=0}^m\Gamma(\alpha_j)}{2\pi}\int_{\mathbb{R}}\prod_{j=0}^m(1+\mathrm{i}a_j\tau)^{-\alpha_j}\mathrm{d}\tau.$$

Proof. 对任意 $t\in\mathbb{R}$，记

$$G(t) = \int_{\forall j:\langle a,w\rangle=t,w_j\ge 0} g(w)\exp\left(-\sum_{\ell=0}^m w_\ell\right)\mathcal{H}^m(\mathrm{d}w).$$

接下来，我们写出 G 关于 $\tau\in\mathbb{R}$ 的傅里叶变换

$$\mathcal{F}[G](\tau) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}\mathrm{e}^{-\mathrm{i}t\tau}\mathrm{d}t\int_{\langle a,w\rangle=t,w\ge 0} g(w)\exp\left(-\sum_{\ell=0}^m w_\ell\right)\mathcal{H}^m(\mathrm{d}w).$$

我们将 $\mathrm{e}^{-\mathrm{i}t\tau}$ 移到第二个积分号下，并将 t 替换为 $\langle a,w\rangle$。注意，这样做是正确的，因为函数 $g(w)\exp\left(-\sum_{\ell=0}^m w_\ell\right)\mathrm{e}^{-\mathrm{i}t\tau}$ 和 $g(w)\exp\left(-\sum_{\ell=0}^m w_\ell\right)\mathrm{e}^{-\mathrm{i}\tau\langle a,w\rangle}$ 在集合 $\langle a,w\rangle = t$ 上几乎处处相等。因此，我们有

$$\mathcal{F}[G](\tau) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}\mathrm{d}t\left[\int_{\langle a,w\rangle=t} g(w)\mathbb{1}\{w\ge 0\}\exp(-\langle w,\mathbf{1}+\mathrm{i}a\tau\rangle)\mathcal{H}^m(\mathrm{d}w)\right]$$

现在我们 可以对由与 a 的标量积定义的映射应用余面积公式。我们有

$$\mathcal{F}[G](\tau) = \frac{[a]}{\sqrt{2\pi}}\int_{\mathbb{R}^{m+1}} g(w)\mathbb{1}\{w\ge 0\}\exp(-\langle w,\mathbf{1}+\mathrm{i}a\tau\rangle)\lambda^{m+1}(\mathrm{d}w).$$

我们得到 m 个独立的 $\Gamma(\alpha_\ell,1)$ 分布随机变量的特征函数之积

$$\mathcal{F}[G](\tau) = \frac{[a]}{\sqrt{2\pi}}\cdot\prod_{j=0}^m\Gamma(\alpha_j)\cdot\frac{1}{(1+\mathrm{i}a_j\tau)^{\alpha_j}},$$

其中， $A^{(j)}$ 是矩阵 A 的第 j 列。最后，根据傅里叶逆变换，并注意到 $[a] = \|a\|_2$

$$G(0) = \frac{\|a\|_2}{\sqrt{2\pi}}\int_{\mathbb{R}}\mathcal{F}[G](\tau)\mathrm{d}\tau = \frac{\|a\|_2\cdot\prod_{j=0}^m\Gamma(\alpha_j)}{2\pi}\int_{\mathbb{R}}\prod_{j=0}^m(1+\mathrm{i}a_j\tau)^{-\alpha_j}\mathrm{d}\tau.$$

□

注 2。注意，积分右侧的值是实数，因为被积函数的实部是偶函数，虚部是奇函数。

推论 **3** *Let* $g\ w\ = \Gamma\ \overline{\alpha}\ \prod_{j=0}^{m} w_j^{\alpha_j - 1}/\Gamma\ \alpha_j$ *where* $\overline{\alpha} > 1$ *and* $a \in \mathbb{R}^{m+1}$ *such that*. () () () $a_0 \neq a_1$. *Then*

$$I_{m-1}(g, \Delta_m \cap \mathrm{H}_a) = (\overline{\alpha} - 1) \cdot [\mathcal{L}_{[a,\mathbf{1}]}] \cdot |a_0 - a_1| \cdot \frac{1}{2\pi} \int_{\mathbb{R}} \prod_{j=0}^{m} (1 + \mathrm{i} a_j s)^{-\alpha_j} \mathrm{d}s.$$

Proof. 注意， g是次数为正齐次函数 $\overline{\alpha} -\ (m+1)\ > -m$ 在 $\mathbb{R}_{++}^{m+1}$ 上。因此，我们可以应用引理3和引理4。我们得到

$$\begin{aligned} I_{m-1}(g, \Delta_m \cap \mathrm{H}_a) &= \frac{(\overline{\alpha} - 1) \cdot [\mathcal{L}_{[a,\mathbf{1}]}] \cdot |a_0 - a_1|}{\|a\|_2} \cdot I_m(g, \widetilde{\Delta}_m \cap \mathrm{H}_a) \\ &= \frac{(\overline{\alpha} - 1) \cdot [\mathcal{L}_{[a,\mathbf{1}]}] \cdot |a_0 - a_1|}{\|a\|_2 \cdot \Gamma(\overline{\alpha})} \int_{\mathbb{R}_+^{m+1} \cap \mathrm{H}_a} g(w) \exp\left(- \sum_{\ell=0}^{m} w_\ell \right) \mathcal{H}^m(\mathrm{d}w). \end{aligned}$$

最后一个积分可由引理 5 计算。我们有

$$I_{m-1}(g, \Delta_m \cap \mathrm{H}_a) = (\overline{\alpha} - 1) \cdot [\mathcal{L}_{[a,\mathbf{1}]}] \cdot |a_0 - a_1| \cdot \frac{1}{2\pi} \int_{\mathbb{R}} \prod_{j=0}^{m} (1 + \mathrm{i} a_j s)^{-\alpha_j} \mathrm{d}s.$$

□

现在我们准备好证明命题 3。

Proof of Proposition 3. 首先，我们给出一个关于 p_Z 的、用 I_{m-1}表示的公式。我们从将该概率改写为通常的积分开始

$$\mathbb{P}_{w \sim \mathrm{Dir}(\alpha)}[wf \leq \mu] = \int_{w \geq 0, \sum_{i=1}^{m} w_i \leq 1, wf \leq \mu} g\left(1 - \sum_{i=1}^{m} w_i, w_1, \ldots, w_m\right) \mathrm{d}w_1, \ldots, \mathrm{d}w_m$$

其中 $g(w)\ = \Gamma(\overline{\alpha}) \prod_{j=0}^{m} w_j^{\alpha_j - 1}/\Gamma(\alpha_j)$ 是狄利克雷分布的密度。我们注意到，此变换恰好定义了一个映射 $\mathcal{L}_{\mathbf{1}}^1$。然后我们应用变量变换公式 [Evans 和 Garzepy， 2018年，3.4。3] 将与向量 $c = (\mathcal{L}_{\mathbf{1}}^1)^\intercal f$ 的标量积作为映射。

$$\mathbb{P}_{w \sim \mathrm{Dir}(\alpha)}[wf \leq \mu] = \frac{1}{[c]} \int_0^\mu \left[\int_{\mathcal{L}_{\mathbf{1}}^1(x) \geq 0, f^\intercal \mathcal{L}_{\mathbf{1}}^1(w) = u} g(\mathcal{L}_{\mathbf{1}}^1(x)) \mathcal{H}^m(\mathrm{d}x) \right] \mathrm{d}u$$

然后，我们将变量替换公式 [Evans 和 Garzepy，2018年，3.3。3] 应用于内层积分，所用参数化由映射 $\mathcal{L}_c^u$

$$\mathbb{P}_{w \sim \mathrm{Dir}(\alpha)}[wf \leq \mu] = \frac{1}{[c]} \int_0^\mu \left[[\mathcal{L}_c^u] \int_{\mathcal{L}_{\mathbf{1}}^1(\mathcal{L}_c^u w) \geq 0} g(\mathcal{L}_{\mathbf{1}}^1(\mathcal{L}_c^u(z))) \mathrm{d}z \right] \mathrm{d}u$$

注意， $\mathcal{L}_c^u$的雅可比矩阵不依赖于平移参数 u，因此可以将其移出积分号。接下来，我们对映射 $\mathcal{L}_{\mathbf{1}}^1 \circ \mathcal{L}_c^u$应用变量替换公式

$$\mathbb{P}_{w \sim \mathrm{Dir}(\alpha)}[wf \leq \mu] = \frac{[\mathcal{L}_c]}{[c][\mathcal{L}_{\mathbf{1}} \circ \mathcal{L}_c]} \int_0^\mu \left[\int_{\Delta_m, wf = u} g(w) \mathcal{H}^{m-1}(\mathrm{d}w) \right] \mathrm{d}u.$$

根据 $c = \mathcal{L}_1^\intercal f$ 的表达式，我们可以应用引理7。因此， p_Z 可以表示为以下积分

$$p_Z(u) = \frac{[\mathcal{L}_{\mathcal{L}_{\mathbf{1}}^\intercal f}]}{[\mathcal{L}_{\mathbf{1}}^\intercal f][\mathcal{L}_{[\mathbf{1}, f]}]} I_{m-1}(g, \Delta_m \cap \mathrm{H}_f^u),$$

其中 $\mathrm{H}_f^u = \{w \in \mathbb{R}^{m+1} \colon wf = u\}$。遗憾的是，我们无法直接应用前面的结果，因为超平面 H_f^u 通常不与 0 相交。为解决这一问题，定义以下向量 $a(u)_j = f(j) - u$。注意，$\langle w,\ a(u)\rangle = 0$ 当且仅当 $\langle w,\ f - u\mathbf{1}\rangle = wf - u = 0$，这里我们使用了 $\langle w,\ \mathbf{1}\rangle = 1$。因此 $\mathrm{H}_f^u \cap \Delta_m = \mathrm{H}_{a(u)} \cap \Delta_m$。我们可以将推论 3 应用于子空间 $\mathrm{H}^{a(u)}$

$$
\begin{aligned}
I_{m-1}(g, \Delta_m \cap \mathrm{H}_f^u) &= \frac{\overline{\alpha} - 1}{2\pi} [\mathcal{L}_{[a(u),\mathbf{1}]}] \cdot |a_0 - a_1| \cdot \int_{\mathbb{R}} \prod_{j=0}^{m} (1 + \mathrm{i} a(u)_j s)^{-\alpha_j} \,\mathrm{d}s \\
&= \frac{\overline{\alpha} - 1}{2\pi} [\mathcal{L}_{[f,\mathbf{1}]}] \cdot |a_0 - a_1| \cdot \int_{\mathbb{R}} \prod_{j=0}^{m} (1 + \mathrm{i}(f(j) - u) s)^{-\alpha_j} \,\mathrm{d}s,
\end{aligned}
$$

其中我们用到了矩阵 $[a(u), \mathbf{1}]$ 和 $[f, \mathbf{1}]$ 可通过初等变换相互等价。综上，密度可表示为

$$
p_Z(u) = \frac{[\mathcal{L}_{\mathcal{L}_{\mathbf{1}}^{\mathsf{T}} f}] |f(0) - f(1)|}{[\mathcal{L}_{\mathbf{1}}^{\mathsf{T}} f]} \cdot \frac{\overline{\alpha} - 1}{2\pi} \int_{\mathbb{R}} \prod_{j=0}^{m} (1 + \mathrm{i}(f(j) - u) s)^{-\alpha_j} \,\mathrm{d}s.
$$

最后，由引理 6 可得

$$
p_Z(u) = \frac{|f(0) - f(1)|}{|(\mathcal{L}_{\mathbf{1}}^{\mathsf{T}} f)_0|} \cdot \frac{\overline{\alpha} - 1}{2\pi} \int_{\mathbb{R}} \prod_{j=0}^{m} (1 + \mathrm{i}(f(j) - u) s)^{-\alpha_j} \,\mathrm{d}s.
$$

By 直接计算可得 $|(\mathcal{L}_{\mathbf{1}}^{\mathsf{T}} f)_0| = |f(0) - f(1)|$，由此得出结论 nt. □

A.2 鞍点法

我们从加权狄利克雷分布之和的密度表达式中出现的积分的渐近分解开始。

命题 4. 设 $f \in \mathrm{F}_m(b)$，并设 $\alpha = (\alpha_0, \alpha_1, \ldots, \alpha_m) \in \mathbb{R}_{>0}^{m+1}$ 为一个固定的正向量，其中 $\alpha_{0,m}$，$\min\{\alpha_0, \alpha_m\} \geq 2$, $\overline{\alpha} = \sum_{i=0}^{m} \alpha_i$。那么，对于任意 $u \in (\overline{p} f, b)$ 以及任意固定的 $\ell \in \{0, \ldots, m\}$

$$
\begin{aligned}
\int_{\mathbb{R}} \frac{\prod_{j=0}^{m} (1 + \mathrm{i}(f(j) - u) s)^{-\alpha_j}}{(1 + \mathrm{i}(f(\ell) - u) s)} \mathrm{d}s &= \sqrt{\frac{2\pi}{\overline{\alpha}\,(1 - \lambda^\star (f(\ell) - u))^2\, \sigma^2}} \cdot \exp(-\overline{\alpha}\, \mathcal{K}_{\inf}(\overline{p}, u, f)) \\
&\quad + (R_2(\alpha) - R_1(\alpha)) \exp(-\overline{\alpha}\, \mathcal{K}_{\inf}(\overline{p}, u, f)) + R_3(\alpha),
\end{aligned}
$$

where

$$
\begin{aligned}
\sigma^2 &= \mathbb{E}_{X \sim \overline{p}} \left[\left(\frac{f(X) - u}{1 - \lambda^\star (f(X) - u)} \right)^2 \right], \\
|R_1(\alpha)| &\leq \frac{c_1}{(1 - \lambda^\star (f(\ell) - u)) \sqrt{\overline{\alpha}\, \sigma^2}} \cdot \frac{\exp(-c_\kappa \alpha_{0,m})}{(c_\kappa \alpha_{0,m})^{1/2}}, \\
|R_2(\alpha)| &\leq \frac{c_2}{(1 - \lambda^\star (f(\ell) - u)) \sqrt{\overline{\alpha}\, \sigma^2}} \cdot \frac{1}{\sqrt{\alpha_{0,m}}}, \\
|R_3(\alpha)| &\leq \frac{\exp(-\overline{\alpha}\, \mathcal{K}_{\inf}(\overline{p}, u, f))}{1 - \lambda^\star (f(\ell) - u)} \cdot \frac{c_3}{\sqrt{\overline{\alpha}\, \sigma^2}} \exp(-c_\kappa \alpha_{0,m}).
\end{aligned}
$$

其中 $c_1 = 2\sqrt{2}, c_2 = \left(8 + \frac{49\sqrt{6}}{9}\right) \mathrm{e}\pi, c_\kappa = \underline{1/2} \cdot \log(5/4)$ 且 $\lambda^\star$ 为优化问题的一个解

$$
\lambda^\star(\overline{p}, u, f) = \operatorname*{arg\,max}_{\lambda \in [0, 1/(b-u)]} \mathbb{E}_{X \sim \overline{p}} [\log(1 - \lambda (f(X) - u))].
$$

Proof. 我们首先将积分改写为以下形式

$$\begin{aligned} I &= \int_{\mathbb{R}} \frac{\prod_{j=0}^{m}(1+\mathrm{i}(f(j)-u)s)^{-\alpha_j}}{1+\mathrm{i}(f(\ell)-u)s}\,\mathrm{d}s \\ &= \int_{\mathbb{R}} \frac{\exp\Big(-\overline{\alpha}\sum_{j=0}^{m}\overline{p}_j\log(1+\mathrm{i}(f(j)-u)s)\Big)}{1+\mathrm{i}(f(\ell)-u)s}\,\mathrm{d}s \\ &= \int_{\mathbb{R}} \frac{\exp(-\overline{\alpha}\,\mathbb{E}_{X\sim\overline{p}}[\log(1+\mathrm{i}(f(X)-u)s)])}{1+\mathrm{i}(f(\ell)-u)s}\,\mathrm{d}s, \end{aligned} \tag{10}$$

其中我们选择复对数函数的主值分支。记 $S^{\mathbb{E}_{X\sim\overline{p}}[\quad 1+\ f\ X\ -u\ z\ ]}$ $\mathbb{E}$ $\mathbb{E}_{X\sim\overline{p}}$ $z\ =$) log(i(( )))。在下文中，为简便起见，我们将使用 代替 。

由于 $f(X)\le b$，该函数在 $|\operatorname{Im} z|<1/(b-u)$ 且 $\operatorname{Re} z\in\mathbb{R}$。最后的积分表示（10）允许使用鞍点法 Olver [1997]。接下来，我们将计算函数 S。为此，计算函数 S 在复数点 $z=x+\mathrm{i}y$

$$S'(z)=\mathrm{i}\mathbb{E}\left[\frac{f(X)-u}{1+\mathrm{i}(f(X)-u)z}\right]=\mathbb{E}\left[\frac{x(f(X)-u)^2+\mathrm{i}(f(X)-u)(1-y(f(X)-u))}{(1-y(f(X)-u))^2+x^2(f(X)-u)^2}\right]=0.$$

注意，只有当 $x=0$ 时，上述表达式的实部才为零。因此，鞍点只能位于虚轴 $\mathrm{i}\mathbb{R}$ 上。它们可由以下方程求得。

$$S'(\mathrm{i}y)=\mathrm{i}\mathbb{E}\left[\frac{f(X)-u}{1-y(f(X)-u)}\right]=0.$$

注意，对于 $y\ge 0$，该方程与 $\lambda^\star$ 在 $\mathcal{K}_{\inf}(\overline{p},u,f)$的定义中的最优性条件相一致。由于 $\overline{p}f<u<b$，函数 $y\mapsto S(\mathrm{i}y)=\mathbb{E}[\log(1-y(f(X)-u))]$ 在 y 上严格凹，因此，方程 $S'(\mathrm{i}y)=0$ 有唯一解 $y=\lambda^\star$。因此， S 的唯一鞍点等于 $z_0=\mathrm{i}\lambda^\star$。接下来，我们将积分路径更改为$\gamma^\star=\mathbb{R}+\mathrm{i}\lambda^\star$。为了证明该路径是合适的，我们来证明 S 的实部在 z_0 处、在所有 $z\in\gamma^\star$ 中取得最小值

$$\operatorname{Re} S(x+\mathrm{i}\lambda^\star)=\frac{1}{2}\mathbb{E}\big[\log\big((1-\lambda^\star(f(X)-u))^2+x^2(f(X)-u)^2\big)\big].$$

$\operatorname{Re} S(x+\mathrm{i}\lambda^\star)$ 的最小值在 $x=0$ 时取得，因此轮廓 $\gamma^\star$ 是合适的。于是，经过简单的坐标变换后，我们可以应用拉普拉斯方法

$$I=\int_{\mathbb{R}}\frac{\exp(-\overline{\alpha}\,\mathbb{E}[\log(1-\lambda^\star(f(X)-u)+\mathrm{i}s(f(X)-u))])}{1-\lambda^\star(f(\ell)-u)+\mathrm{i}(f(\ell)-u)s}\,\mathrm{d}s$$

记

$$T(s)=\mathbb{E}[\log(1-\lambda^\star(f(X)-u)+\mathrm{i}s(f(X)-u))],\quad P(s)=\frac{1}{1-\lambda^\star(f(\ell)-u)+\mathrm{i}s(f(\ell)-u)}.$$

固定截断参数 $K>0$，并定义 $\kappa_1=T(-K)-T(0)$, $\kappa_2=T(K)-T(0)$。接下来，类似于 Olver 的第4章（第6节）， [1997], 我们定义变量替换 $v_1=T(-s)-T(0)$ $v_2=T(s)-T(0)$ 以及隐函数 $q_1(v_1)=P(-s)\,T'(-s)$ $q_2(v_2)=P(s)\,T'(s)$。利用一阶泰勒展开，我们可以写出 $q_1(v_1)=P(0)\sqrt{2T''(0)\cdot v_1}+r_1(v_1)$ $q_2(v_2)=P(0)\sqrt{2T''(0)\cdot v_2}+r_2(v_2)$。于是得到如下分解。

$$I=\int_{-\infty}^{\infty}P(s)\exp(-\overline{\alpha}T(s))\,\mathrm{d}s=\left(P(0)\cdot\sqrt{\frac{2\pi}{\overline{\alpha}T''(0)}}-R_1(\alpha)+R_2(\alpha)\right)\exp(-\overline{\alpha}T(0))+R_3(\alpha),$$

其中

$$
\begin{aligned}
R_1(\alpha) &= \left(\Gamma\left(\frac{1}{2}, \kappa_1 \overline{\alpha}\right) + \Gamma\left(\frac{1}{2}, \kappa_2 \overline{\alpha}\right)\right) \frac{P(0)}{\sqrt{2T''(0)\,\overline{\alpha}}}, \\
R_2(\alpha) &= \int_0^{\kappa_1} e^{-\overline{\alpha} v_1} r_1(v_1)\,\mathrm{d}v_1 + \int_0^{\kappa_2} e^{-\overline{\alpha} v_2} r_2(v_2)\,\mathrm{d}v_2, \\
R_3(\alpha) &= \int_{\mathbb{R}\setminus[-K,K]} P(s)\exp(-\overline{\alpha}\,T(s))\,\mathrm{d}s,
\end{aligned}
$$

其中 $\Gamma(\alpha, x)$ 是上不完全伽马函数，并且关于 v_1, v_2 的积分分别沿连接点 0 与 κ_1、κ_2 的直线进行。定义 $\sigma^2 = T''(0)$。

项 R_2 我们将从对泰勒型展开中的余项求上界开始$r_2(v)$

$$
\begin{aligned}
|r_2(v)| &= \left|\frac{P(s)}{T'(s)} - \frac{P(0)}{\sqrt{2T''(0)(T(s)-T(0))}}\right| \\
&\le P(0)\left|\frac{1}{T'(s)} - \frac{1}{\sqrt{2T''(0)(T(s)-T(0))}}\right| + \frac{|P(s)-P(0)|}{|T'(s)|} \\
&= P(0)\cdot \bar{r}_2(v) + \tilde{r}_2(v).
\end{aligned}
$$

Upper bound for $\bar{r}_2(v)$

通过泰勒展开

$$
\begin{aligned}
T(s) &= T(0) + T'(0)\cdot s + \frac{T''(0)}{2}s^2 + \frac{T'''(\xi_1)}{6}s^3, && \xi_1 \in (0, s) \\
T'(s) &= T'(0) + T''(0)s + \frac{T'''(\xi_2)}{2}s^2, && \xi_2 \in (0, s).
\end{aligned}
$$

注意到 $T'(0) = 0$，因此

$$
\begin{aligned}
|\bar{r}_2(v)| &= \left|\frac{1}{T'(s)} - \frac{1}{\sqrt{2T''(0)(T(s)-T(0))}}\right| \\
&= \left|\frac{\sqrt{T''(0)^2 s^2 + T''(0)\frac{T'''(\xi_1)}{3}s^3} - T''(0)s - \frac{T'''(\xi_2)}{2}s^2}{[T''(0)s + \frac{T'''(\xi_2)}{2}s^2]\cdot\sqrt{T''(0)^2 s^2 + T''(0)\frac{T'''(\xi_1)}{2}s^3}}\right| \\
&= \left|T''(0)\frac{\sqrt{1 + \frac{T'''(\xi_1)}{3T''(0)}s} - 1 - \frac{T'''(\xi_2)}{2T''(0)}s}{s\cdot[T''(0) + \frac{T'''(\xi_2)}{2}s]\cdot\sqrt{T''(0)^2 + T''(0)\frac{T'''(\xi_1)}{3}s}}\right|.
\end{aligned}
$$

接下来由 应用平方根的泰勒展开式 $\sqrt{1+x} - 1 = \frac{x}{2} - \frac{x^2}{8(1+\xi_3)^{3/2}}$ 对于 ~~$|\xi_3| < x$,~~

$$
\begin{aligned}
|\bar{r}_2(v)| &= \left|T''(0)\frac{\frac{T'''(\xi_1)}{6T''(0)} - \frac{T'''(\xi_2)}{2T''(0)} - s\cdot\frac{(T'''(\xi_1)/T''(0))^2}{8\cdot 9\cdot(1+\xi_3)^{3/2}}}{[T''(0) + \frac{T'''(\xi_2)}{2}s]\cdot\sqrt{T''(0)^2 + T''(0)\frac{T'''(\xi_1)}{3}s}}\right| \\
&= \left|\frac{\frac{T'''(\xi_1)}{6} - \frac{T'''(\xi_2)}{2} - s\cdot\frac{(T'''(\xi_1))^2}{8\cdot 9\cdot(1+\xi_3)^{3/2}\cdot T''(0)}}{[T''(0) + \frac{T'''(\xi_2)}{2}s]\cdot\sqrt{T''(0)^2 + T''(0)\frac{T'''(\xi_1)}{3}s}}\right| \\
&\le \frac{\frac{2}{3}\sup_{u\in(0,s)}|T'''(u)| + s\cdot\frac{\sup_{u\in(0,s)}|T'''(u)|^2}{72\cdot|T''(0)|}}{|T''(0) + \frac{T'''(\xi_2)}{2}s|\cdot\sqrt{|T''(0)^2 + T''(0)\frac{T'''(\xi_1)}{3}s|}}.
\end{aligned}
$$

让我们分析 T 的二阶和三阶导数。

$$T''(s) = \mathbb{E}\left[\left(\frac{f(X)-u}{1-\lambda^\star(f(X)-u)+is(f(X)-u)}\right)^2\right],$$

$$T'''(s) = -2i\mathbb{E}\left[\left(\frac{f(X)-u}{1-\lambda^\star(f(X)-u)+is(f(X)-u)}\right)^3\right].$$

定义随机变量 $Y_s = \frac{f(X)-u}{1-\lambda^\star(f(X)-u)+is(f(X)-u)}$，然后 $T''(s) = \mathbb{E}[Y_s^2]$，$T'''(s) = -2i\mathbb{E}[Y_s^3]$。让我们计算 $T'''(s)$ 绝对值的上界

$$|T'''(s)| \le 2\mathbb{E}\left[\frac{|f(X)-u|^3}{((1-\lambda^\star(f(X)-u))^2+s^2(f(X)-u)^2)^{3/2}}\right] \le 2\mathbb{E}[|Y_0|^3].$$

通过选择 $1/(2K) = \max\left\{\frac{b-u}{1-\lambda^\star(b-u)}, \frac{u}{1+\lambda^\star u}\right\}$，我们确保 $\mathbb{E}[Y_0^2] - s\mathbb{E}[|Y_0|^3] \ge \frac{1}{2}\mathbb{E}[Y_0^2]$ 对所有 $0 \le s < K$ 均成立，因为

$$\mathbb{E}[|Y_0|^3] \le \text{最大值}_{j\in\{0,\dots,m\}} |f(j)-u|\,(f(j)-u)\,\mathbb{E}[Y_0^2] \le \text{最大值}\left\{\frac{b-u}{1-\lambda^\star(b-u)}, \frac{u}{1+\lambda^\star u}\right\}\mathbb{E}[Y_0^2] \le \frac{1}{2K}\mathbb{E}[Y_0^2].$$

因此

$$\begin{aligned}|\bar{r}_2(v)| &\le \frac{\frac{4}{3}\mathbb{E}[|Y_0|^3] + s\frac{\mathbb{E}[|Y_0|^3]^2}{18\cdot\mathbb{E}[Y_0^2]}}{(\mathbb{E}[Y_0^2]-\mathbb{E}[|Y_0|^3]s)\cdot\sqrt{\mathbb{E}[Y_0^2]}\cdot\sqrt{\mathbb{E}[Y_0^2]-\mathbb{E}[|Y_0|^3]\frac{2s}{3}}} \\ &\le \frac{4/3+1/36}{1/2\cdot\sqrt{2/3}}\cdot\frac{\mathbb{E}[|Y_0|^3]}{\mathbb{E}[Y_0^2]^2} \le \frac{49\sqrt{6}}{36\mathbb{E}[Y_0^2]}\cdot\max\left\{\frac{b-u}{1-\lambda^\star(b-u)}, \frac{u}{1+\lambda^\star u}\right\}.\end{aligned}$$

接下来，利用该界

$$\mathbb{E}[Y_0^2] = \sum_{j=0}^m \frac{\alpha_j}{\overline{\alpha}}\left(\frac{f(j)-u}{1-\lambda^\star(f(j)-u)}\right)^2 \ge \frac{\alpha_{0,m}}{\overline{\alpha}}\left(\max\left\{\frac{b-u}{1-\lambda^\star(b-u)}, \frac{u}{1+\lambda^\star u}\right\}\right)^2,$$

我们得到

$$|\bar{r}_2(v)| \le \frac{49\sqrt{6}}{36\sqrt{\sigma^2}}\cdot\sqrt{\frac{\overline{\alpha}}{\alpha_{0,m}}}.$$

Upper bound for $\tilde{r}_2(v)$ 我们接下来的目标是分析第二项 $\tilde{r}_2(v)$。我们应用如下形式的泰勒展开 $T'(s) = T''(0)s + T'''(\xi_2)s^2/2$ 以及 $P(s) = P(0) + P'(\eta)s$ 来推导

$$\tilde{r}_2(v) = \left|\frac{P(s)-P(0)}{T'(s)}\right| = \frac{|P'(\eta)\cdot s|}{|T''(0)s + T'''(\xi_2)s^2/2|} \le \frac{\sup_{\eta\in(0,s)}|P'(\eta)|}{|T''(0)+T'''(\xi_2)s/2|}.$$

首先注意到 $|P'(\eta)|$ 在 $\eta = 0$ 处取最大值，因为

$$|P'(\eta)| = \frac{|f(\ell)-u|}{(1-\lambda^\star(f(\ell)-u))^2+\eta^2(f(\ell)-u)^2}.$$

接下来，通过定义随机变量 $Y_s = \frac{f(X)-u}{1-\lambda^\star(f(X)-u)+is(f(X)-u)}$，并根据我们对 K 的选择，我们得出

$$|T''(0)+T'''(\xi_2)s/2| \ge \mathbb{E}[Y_0^2] - s\mathbb{E}[|Y_0|^3] \ge \mathbb{E}[Y_0^2]/2 = \sigma^2/2.$$

它给出

$$\tilde{r}_2(v) \le \frac{2|f(\ell)-u|}{(1-\lambda^\star(f(\ell)-u))^2\sigma^2} = \frac{2}{(1-\lambda^\star(f(\ell)-u))\sqrt{\sigma^2}}\sqrt{\frac{\frac{(f(\ell)-u)^2}{(1-\lambda^\star(f(\ell)-u))^2}}{\mathbb{E}[Y_0^2]}}.$$

由一个界

$$\mathbb{E}[Y_0^2] \geq \frac{\alpha_{0,m}}{\overline{\alpha}} \left(\max\left\{ \frac{b-u}{1-\lambda^\star(b-u)}, \frac{u}{1+\lambda^\star u} \right\} \right)^2 \geq \frac{\alpha_{0,m}}{\overline{\alpha}} \frac{(f(\ell)-u)^2}{(1-\lambda^\star(f(\ell)-u))^2}$$

我们得到

$$\tilde{r}_2(v) \leq \frac{2}{(1-\lambda^\star(f(\ell)-u))\sqrt{\sigma^2}} \sqrt{\frac{\overline{\alpha}}{\alpha_{0,m}}}.$$

统一 $\bar{r}_2$ 和 $\tilde{r}_2$ 的界，我们得到

$$|r_2(v)| \leq \frac{1}{(1-\lambda^\star(f(\ell)-u))\sqrt{\sigma^2}} \sqrt{\frac{\overline{\alpha}}{\alpha_{0,m}}} \left(2 + \frac{49\sqrt{6}}{36} \right).$$

根据对称性， $r_1(v)$ 也成立类似的界。令 $c_2' = 2 + \frac{49\sqrt{6}}{36}$，则

$$\begin{aligned} |R_2(\alpha)| &\leq \frac{c_2'}{\sqrt{(1-\lambda^\star(f(\ell)-u))^2\sigma^2}} \cdot \sqrt{\frac{\overline{\alpha}}{\alpha_{0,m}}} \cdot \left| \int_0^{\kappa_2} e^{-\overline{\alpha}v} \mathrm{d}v + \int_0^{\kappa_1} e^{-\overline{\alpha}v} \mathrm{d}v \right| \\ &\leq \frac{2c_2'}{\sqrt{\overline{\alpha}(1-\lambda^\star(f(\ell)-u))^2\sigma^2}} \cdot \frac{1+\exp(-\overline{\alpha}\kappa)}{\sqrt{\alpha_{0,m}}}, \end{aligned}$$

其中 $\kappa = \min\{\mathrm{Re}\ \kappa_1, \ \mathrm{Re}\,\kappa_2\}$. 使用单位矩阵

$$\mathrm{Re}\,\kappa_1 = \mathrm{Re}\,\kappa_2 = \frac{1}{2}\mathbb{E}\left[\log\left(\frac{(1-\lambda^\star(f(X)-u))^2 + K^2(f(X)-u)^2}{(1-\lambda^\star(f(X)-u))^2} \right) \right] = \frac{1}{2}\mathbb{E}\left[\log\left(1+K^2Y_0^2\right)\right]$$

以及不等式

$$\begin{aligned} \mathbb{E}\left[\log\left(1+K^2Y_0^2\right)\right] &= \sum_{j=0}^{m} \frac{\alpha_j}{\overline{\alpha}} \log\left(1 + K^2 \cdot \left(\frac{f(j)-u}{1-\lambda^\star(f(j)-u)} \right)^2 \right) \\ &\geq \frac{\alpha_{0,m}}{\overline{\alpha}} \log\left(1 + K^2 \cdot \left(\max\left\{ \frac{b-u}{1-\lambda^\star(b-u)}, \frac{u}{1+\lambda^\star u} \right\} \right)^2 \right) \geq \frac{\alpha_{0,m}}{\overline{\alpha}} \log(5/4) \end{aligned}$$

我们有 $\kappa = \mathrm{Re}\,\kappa_2 = \mathrm{Re}\ \kappa_1 \geq c_\kappa \cdot \frac{\alpha_{0,\ m}}{\overline{\alpha}}$ 其中 $c_\kappa = 1/2 \cdot \log(5/4)$。由于 $\alpha_{0,\ m} > 0$，我们还有 $\exp(-c_\kappa \alpha_{0,\ m}) \leq 1$。最后令 $c_2 = 4 \cdot c_2' = 8 + \frac{49\sqrt{6}}{9}$，我们推导出关于 R_2 的以下界大小规律

$$|R_2(\alpha)| \leq \frac{c_2}{\sqrt{\overline{\alpha}(1-\lambda^\star(b-u))\sigma^2}} \cdot \frac{1}{\sqrt{\alpha_{0,m}}}.$$

术语 R_1 根据 Borwein 和 Chan 的定理2.4 [2007] 我们对任意复数 z 满足 $\mathrm{Re}\,z > 0$

$$\left| \Gamma\left(\frac{1}{2}, z \right) \right| \leq \frac{2e^{-\mathrm{Re}\,z}}{|z|^{1/2}}.$$

因此，

$$|R_1(\alpha)| \leq \frac{4}{\sqrt{2T''(0)|\kappa|}} \cdot \frac{\exp(-\overline{\alpha}\,\kappa)}{\overline{\alpha}}.$$

注意到 $T''(0) = \sigma^2$，因此

$$|R_1(\alpha)| \leq \frac{2\sqrt{2}}{\sqrt{\sigma^2 \cdot \overline{\alpha}}} \cdot \frac{\exp(-c_\kappa \alpha_{0,m})}{(c_\kappa \alpha_{0,m})^{1/2}}.$$

项 R_3 我们从该界开始。

$$\left|\int_K^\infty P(s)\exp(-\overline{\alpha}T(s))\,\mathrm{d}s\right| \le \exp(-\overline{\alpha}T(0))\cdot\exp(-\overline{\alpha}(T(K)-T(0)))\cdot\sup_{s\in\mathbb{R}}|P(s)| \cdot\int_K^\infty \exp(-\overline{\alpha}\,\mathrm{Re}[T(s)-T(K)])\,\mathrm{d}s.$$

我们先从对与 $P(s)$ 相连的额外乘子的分析开始。

$$\sup_{s\in R}|P(s)| = \sup_s\sqrt{\frac{1}{(1-\lambda^\star(f(\ell)-u))^2+s^2(f(\ell)-u)}} = \frac{1}{1-\lambda^\star(f(\ell)-u)}$$

我们的目标是给最后一个积分找到界。让我们分析变量替换 $s\mapsto t+K$ 后指数下的函数。

$$q(t)\triangleq \mathrm{Re}[T(t+K)-T(K)] = \frac{1}{2}\mathbb{E}\left[\log\left(\frac{(1-\lambda^\star(f(X)-u))^2+(t+K)^2(f(X)-u)^2}{(1-\lambda^\star(f(X)-u))^2+K^2(f(X)-u)^2}\right)\right]$$
$$= \frac{1}{2}\mathbb{E}\left[\log\left(1+\frac{(t^2+2TK)(f(X)-u)^2}{(1-\lambda^\star(f(X)-u))^2+K^2(f(X)-u)^2}\right)\right]$$

定义一个函数 $g(j)=\frac{(f(j)-u)^2}{(1-\lambda^\star(f(j)-u))^2+K^2(f(j)-u)}\ge 0$。
然后

$$\exp(-\overline{\alpha}q(t)) = \prod_{j=0}^m\left(\frac{1}{1+(2tK+t^2)g(j)}\right)^{\alpha_i/2} \le \prod_{j=0}^m\left(\frac{1}{1+t^2g(j)}\right)^{\alpha_i/2}. \tag{11}$$

取一个序列 $q_i\in\mathbb{R}_+$ 使得 $\sum_{i=0}^m q_i^{-1}=1$。遵循 G¨otze 等人的思想, [2019], 应用广义霍¨尔德不等式

$$\int_K^\infty \exp(-\overline{\alpha}\,\mathrm{Re}[T(s)-T(K)])\,\mathrm{d}s = \int_0^\infty \exp(-\overline{\alpha}q(t))\mathrm{d}t$$
$$\le \int_0^\infty\prod_{j=0}^m\left(\frac{1}{1+t^2g(j)}\right)^{\alpha_j/2}\mathrm{d}t \le \prod_{j=0}^m\left(\int_0^\infty\left(\frac{1}{1+t^2g(j)}\right)^{\alpha_jq_j/2}\mathrm{d}t\right)^{1/q_j}.$$

接下来我们精确计算积分并得到

$$\int_0^\infty\left(\frac{1}{1+t^2g(j)}\right)^{\alpha_jq_j/2}\mathrm{d}t \le \frac{\sqrt{\pi}\Gamma((\alpha_jq_j-1)/2)}{2\Gamma(\alpha_jq_j/2)}\cdot\frac{1}{\sqrt{g(j)}}$$

根据Guo 等人 [2007] 和假设 $\alpha_jq_j>1$ 我们有

$$\frac{\Gamma((\alpha_jq_j-1)/2)}{\Gamma(\alpha_jq_j/2)} \le \frac{\sqrt{2\mathrm{e}}}{\sqrt{\alpha_jq_j-1}}.$$

最后，我们得到

$$\int_0^\infty\exp(-\overline{\alpha}q(t))\mathrm{d}t \le \frac{\sqrt{2\mathrm{e}\pi}}{2}\prod_{j=0}^m\left(\frac{1}{\sqrt{g(j)(\alpha_jq_j-1)}}\right)^{1/q_j}.$$

现在我们固定 q_j，使得 $g(j)(\alpha_jq_j-1)=\tau$，其中常数 τ 由 $\sum_{j=0}^m q_j^{-1}=1$ 确定。因此

$$\frac{1}{q_j}=\frac{\alpha_jg(j)}{\tau+g(j)},\quad \sum_{j=0}^m\frac{\alpha_jg(j)}{\tau+g(j)}=1.$$

特别地，由第二个等式可知 $\tau\le\sum_{j=0}^m\alpha_jg\ j$ $\tau+\ _kg\ k\ \ge()$ 和 $\max()()$。我们注意到，由于 $\alpha_{0\ m}\ge 2$ $\tau\ge\frac{1}{2}\sum_{j=0}^m\alpha_jg\ j$ $\sum_{j=0}^m\alpha_jg\ j$.我们有 $()$。接下来我们需要

保证 e，即 $\alpha_j q_j > 1$，但对于我们选择的 τ，这显然成立，因为 $\alpha_j q_j = (\tau + g(j))/g(j) > 1$. 因此 我们有

$$\int_K^\infty \exp(-\overline{\alpha}\,\mathrm{Re}[T(s) - T(K)])\,\mathrm{d}s \le \sqrt{\frac{\mathrm{e}\pi}{\overline{\alpha}\mathbb{E}[g(X)]}}.$$

接下来我们希望将 $\mathbb{E}[g(X)]$ 与 σ^2 联系起来。根据定义

$$\mathbb{E}[g(X)] = \mathbb{E}\left[\frac{(f(X)-u)^2}{(1-\lambda^\star(f(X))-u))^2 + K^2(f(X)-u)^2)}\right]$$

根据选择 $1/(2K) = \max\{1-\lambda^\star \frac{b-u}{-(b-u)}, \frac{u}{1+\lambda^\star u}\}$，从而 $K^2 \cdot (f(j) - u)^2 \le (1-\lambda^\star(f(j) - u))^2/4$对于任意 $j \in \{0, \ldots, m\}$。因此

$$\mathbb{E}[g(X)] \ge \mathbb{E}\left[\frac{(f(X)-u)^2}{(1-\lambda^\star(f(X))-u))^2 + (1-\lambda^\star(f(X))-u))^2/4}\right] \ge \frac{4}{5}\cdot\sigma^2.$$

最终，我们通过对称性得到余项的以下界

$$|R_3(\alpha)| \le \frac{\exp(-\overline{\alpha}\,\mathcal{K}_{\inf}(\overline{p}, u, f))}{1-\lambda^\star(f(\ell)-u)} \cdot \sqrt{\frac{5\mathrm{e}\pi}{\overline{\alpha}\sigma^2}}\exp(-c_\kappa \alpha_{0,m}).$$

□

B 线性代数基础概念

在本章节中，我们给出一些有用的引理，以简化雅可比矩阵的计算。首先回顾一下，每个满秩矩阵 $A \in \mathbb{R}^{k\times n}$ for $k < n$ 都有如下分解。

$$A = C\begin{bmatrix} I_k & \hat{A}\end{bmatrix} P, \tag{12}$$

其中， $C \in \mathbb{R}^{k\times k}$ 是非退化矩阵，$\hat{}A \in \mathbb{R}^{k\times(n-k)}$ 是自由变量前的系数矩阵，而 $P \in \mathbb{R}^{n\times n}$ 是置换矩阵。该分解直接由高斯–若尔当消元法得到，此分解中的中心矩阵对应于简化行阶梯形。众所周知，该分解的中心矩阵是唯一的。在此情况下，我们可以定义将 $\mathbb{R}^{n-k}$ 映射到$\mathrm{H}_A = \{x \in \mathbb{R}^n$ 的矩阵 $\mathcal{L}_A \in \mathbb{R}^{n\times(n-k)}$ 如下： $Ax = 0\}$

$$\mathcal{L}_A = P^\mathsf{T}\begin{bmatrix} -\hat{A} \\ I_{n-k}\end{bmatrix} \tag{13}$$

这对应于一种通过约化矩阵计算 H_A 的基的规范方式。事实上，可以验证

$$A\mathcal{L}_A = C\begin{bmatrix} I_k & \hat{A}\end{bmatrix} PP^\mathsf{T}\begin{bmatrix} -\hat{A} \\ I_{n-k}\end{bmatrix} = 0$$

且 $\mathcal{L}_A$ 是满秩的，因此它是到 H_A的真映射。注意， $\mathcal{L}_A$ 仅取决于矩阵 A 的简化形式，并且在对矩阵A执行任何行运算时都不会改变。接下来我们展示这些矩阵的若干性质。

引理 **6** *For any non-zero vector $v \in \mathbb{R}^n$ such that $v_0 \neq 0$ and $t \in \mathbb{R}$ a Jacobian of map $\mathcal{L}_v^t$ is equal to $\|v\|_2/|v_0|$.*

（线性参数化的雅可比矩阵）。

Proof. 注意，梯度向量不依赖于常数平移。因此，梯度矩阵等于线性映射 $\mathcal{L}_v$。定义一个向量 $\tilde{v} = [v_1, \ldots, v_n]$，则方形雅可比矩阵等于

$$[\mathcal{L}_v]^2 = \det(\mathcal{L}_v^\mathsf{T}\mathcal{L}_v) = \det\left(\begin{bmatrix}\tilde{v}/v_0 & I_{n-1}\end{bmatrix}\begin{bmatrix}\tilde{v}^\mathsf{T}/v_0 \\ I_{n-1}\end{bmatrix}\right) = \det\left(\frac{\tilde{v}\tilde{v}^\mathsf{T}}{v_0^2} + I_{n-1}\right).$$

这个矩阵是一个秩一矩阵加上单位矩阵。其特征值为 $\|\tilde{v}\|^2/v_0^2 + 1$ 和 $n-2$ 个 $\mathbf{1}$。因此 $[\mathcal{L}_v]^2 = \|\tilde{v}\|^2/v_0^2 + 1 = \|v\|^2/v_0^2$。 □

引理 7. *Let* $A \in \mathbb{R}^{k\times n}, c \in \mathbb{R}^n$ *and* $A_c = \begin{bmatrix} A \\ c^\intercal \end{bmatrix} \in \mathbb{R}^{(k+1)\times n}$ *for* $k < n$. *Then*

$$\mathcal{L}_A \mathcal{L}_{c^\intercal \mathcal{L}_A} = \mathcal{L}_{A_c}. \tag{14}$$

Proof. 首先，我们用矩阵 A 和向量 c来计算 $\mathcal{L}_{A_c}$。不失一般性，我们可以假设 $C = I$ 和 $P = I$。将向量 c 分成两部分 $c^\intercal = \begin{bmatrix} c^\intercal{}_{1()} & c^\intercal{}_{(2)} \end{bmatrix}$，其中 $c_{(1)} \in \mathbb{R}^k$ 和 $c_{(2)} \in \mathbb{R}^{n-k}$。然后，我们可以得到 A_c 的行消元过程如下

$$\begin{bmatrix} A \\ c^\intercal \end{bmatrix} \mapsto \begin{bmatrix} I_k & \hat{A} \\ c_{(1)}^\intercal & c_{(2)}^\intercal \end{bmatrix} \mapsto \begin{bmatrix} I_k & \hat{A} \\ 0_k & c_{(2)}^\intercal - c_{(1)}^\intercal \hat{A} \end{bmatrix}.$$

定义 $\hat{c} = c_{(2)} - \hat{A}^\intercal c_{(1)}$。不失一般性，我们可以假设第一个坐标 $\hat{c}_1 \neq 0$，因此可以将 $\hat{c} =$ 分解为 $\hat{c}_1 \begin{bmatrix} 1 & \tilde{c} \end{bmatrix}$，其中 $\tilde{c}$ 是 $\hat{c}$ 除第一个坐标外其余坐标经 $\hat{c}_1$缩放后的结果。接下来，我们将上三角形式化为简化形式，如下所示

$$\begin{bmatrix} I_k & \hat{A} \\ 0_k & \hat{c}^\intercal \end{bmatrix} \mapsto \begin{bmatrix} I_k & \hat{A}^{(1)} & \hat{A}^{(2:)} \\ 0_k & 1 & \tilde{c}^\intercal \end{bmatrix} \mapsto \begin{bmatrix} I_k & 0_{n-k} & \hat{A}^{(2:)} - \hat{A}^{(1)} \tilde{c}^\intercal \\ 0_k & 1 & \tilde{c}^\intercal \end{bmatrix},$$

其中 $\hat{A}^{(1)}$ 是矩阵 $\hat{A}$ 的第一列，$\hat{A}^{(2:)}$ 是矩阵 $\hat{A}$ 除第一列外的其余部分。由于我们将矩阵化为简化形式，因此得到

$$\mathcal{L}_{A_c} = \begin{bmatrix} \hat{A}^{(1)} \tilde{c}^\intercal - \hat{A}^{(2:)} \\ -\tilde{c}^\intercal \\ I_{n-k-1} \end{bmatrix}.$$

接下来，我们将证明方程（14）的左侧等于得到的表达式 $\mathcal{L}_{A_c}$。首先，我们注意到

$$\mathcal{L}_A^\intercal c = \begin{bmatrix} -\hat{A}^\intercal & I_{n-k} \end{bmatrix} \begin{bmatrix} c_{(1)} \\ c_{(2)} \end{bmatrix} = c_{(2)} - \hat{A}^\intercal c_{(1)} = \hat{c}.$$

接下来，对于只有一个方程的系统，通过显式的行化简可得

$$\mathcal{L}_{c^\intercal \mathcal{L}_A} = \begin{bmatrix} -\tilde{c}^\intercal \\ I_{n-k-1} \end{bmatrix},$$

因此

$$\mathcal{L}_A \mathcal{L}_{c^\intercal \mathcal{L}_A} = \begin{bmatrix} -\hat{A}^{(1)} & -\hat{A}^{(2:)} \\ 1 & 0_{n-k-1} \\ 0_{n-k-1} & I_{n-k-1} \end{bmatrix} \begin{bmatrix} -\tilde{c}^\intercal \\ I_{n-k-1} \end{bmatrix} = \begin{bmatrix} \hat{A}^{(1)} \tilde{c}^\intercal - \hat{A}^{(2:)} \\ -\tilde{c}^\intercal \\ I_{n-k-1} \end{bmatrix}.$$

□

C MTS 算法

本章节将详细介绍MTS 算法，见算法 1。我们还在算法 2中描述了RMTS ， 该算法在第 4章中介绍。

算法 1 MTS

1: 输入：所有臂的先验分布 ρ_a^0 $a \in [K]$。 2: 对于 $t \in [T]$ ， 执行
3: 为所有臂生成后验样本 $w_a^t \sim \rho_a^{t-1}$ $a \in [K]$。
4 $A_t \in$ $a\in[K] \sum_{i=0}^m w_a^t \frac{i}{m}$: 抽取臂 arg
5 Y_t
max: 获取奖励并更新后验分布。6: 结束循环。

算法2 RMTS

输入：网格大小 m。先验分布 ρ_a^0 在网格 $\{0, 1/m, \ldots, 1\}$ 上适用于所有臂 $a \in [K]$。

对 $t \in [T]$ 执行
  为所有臂 $a \in [K]$ 生成后验样本 $w_a^t \sim \rho_a^{t-1}$。
  抽取臂 $A_t \in \arg\max_{a\in[K]} \sum_{i=0}^{m} w_a^t \frac{i}{m}$
  获得奖励 Y_t 和类别 $i_t = \lfloor mY_t \rfloor$。
  样本虚拟奖励 $\widetilde{Y}_t = (i_t + b_t)/m$，其中 $b_t \sim \mathrm{Ber}(mY_t - i_t)$。
  使用 $\widetilde{Y}_t$ 更新后验分布。
结束 for

D 遗憾界的证明

在本章节中，我们给出定理 2 和定理 3。

D.1 定理证明 2

对于 $\alpha \in \mathbb{R}_+^{M+1}$，我们将 $\overline{p}(\alpha) \in \Delta_M$ 定义为如下概率分布：$\overline{p}(\alpha)_i = \alpha_i / (\sum_{j=0}^{m} \alpha_j)$。此外，定义以下两个概率分布。

$$\overline{p}^-(\alpha) = \overline{p}(\alpha_0 - 1, \alpha_1 \ldots, \alpha_{m-1}, \alpha_m), \qquad \overline{p}^+(\alpha) = \overline{p}(\alpha_0, \alpha_1, \ldots, \alpha_{m-1}, \alpha_m - 1).$$

令 $\varepsilon > 0$ 为稍后指定的值。令 $a^\star$ 为唯一的最优臂。定义以下事件

$$\mathcal{E}_\varepsilon^{\mathrm{conc},1}(t,a) = \left\{ (N_a^t + \overline{\alpha}^0 - 1)\, \mathcal{K}_{\inf}(\overline{p}^-(\alpha_a^t), \mu^\star - \varepsilon, f) \geq N_a^t\, \mathcal{K}_{\inf}(p_a, \mu^\star - \varepsilon, f) - \widetilde{\varepsilon}(N_a^t) \right\};$$

$$\mathcal{E}_\varepsilon^{\mathrm{conc},2}(t) = \left\{ (N_{a^\star}^t + \overline{\alpha}^0 - 1)\, \mathcal{K}_{\inf}(\overline{p}^-(\alpha_{a^\star}^t), 1 - \mu^\star + \varepsilon, 1 - f) \right.$$
$$\left. \geq N_{a^\star}^t\, \mathcal{K}_{\inf}(p_{a^\star}, 1 - \mu^\star + \varepsilon, 1 - f) - \widetilde{\varepsilon}(N_{a^\star}^t) \right\};$$

$$\mathcal{E}_\varepsilon^{\mathrm{conc},3}(t) = \left\{ (N_{a^\star}^t + \overline{\alpha}^0 - 1)\, \mathcal{K}_{\inf}(\overline{p}^+(\alpha_{a^\star}^t), \mu^\star, f) \leq 3\log(2N_a^t + 1) + (\overline{\alpha}^0 - 1) \log\left(\frac{1}{1-\mu^\star}\right) \right\};$$

$$\mathcal{E}_\varepsilon^{\mathrm{conc}}(t,a) = \mathcal{E}_\varepsilon^{\mathrm{conc},1}(t,a) \cap \mathcal{E}_\varepsilon^{\mathrm{conc},2}(t) \cap \mathcal{E}_\varepsilon^{\mathrm{conc},3}(t),$$

其中，N_a^t 表示截至时刻 t 臂 a 的拉取次数，$\overline{\alpha}^0$ 是先验分布的权重，并且

$$\widetilde{\varepsilon}(n) = \sqrt{\frac{4\log(n)\gamma}{n}} + \frac{3(\overline{\alpha}^0 - 1)}{n} \log\left(\frac{n + \overline{\alpha}^0 - 1}{1 - \mu^\star + \varepsilon}\right)$$

for

$$\gamma = \frac{1}{\sqrt{1 - \mu^\star + \varepsilon}} \left(16\mathrm{e}^{-2} + \log^2\left(\frac{1}{1 - \mu^\star + \varepsilon}\right)\right).$$

我们为任意次优臂 $a \neq a^\star$定义如下分解

$$\mathbb{E}[N_a^T] = \underbrace{\mathbb{E}\left[\sum_{t=1}^{T} \mathbb{1}\{A^t = a, w_a^t f \geq \mu^\star - \varepsilon, \mathcal{E}_\varepsilon^{\mathrm{conc}}(t-1,a)\}\right]}_{\mathrm{PostCV}(\varepsilon)}$$
$$+ \underbrace{\mathbb{E}\left[\sum_{t=1}^{T} \mathbb{1}\{A^t = a, w_a^t f \leq \mu^\star - \varepsilon, \mathcal{E}_\varepsilon^{\mathrm{conc}}(t-1,a)\}\right]}_{\mathrm{PreCV}(\varepsilon)}$$
$$+ \underbrace{\mathbb{E}\left[\sum_{t=1}^{T} \mathbb{1}\{A^t = a, \neg\mathcal{E}_\varepsilon^{\mathrm{conc}}(t-1,a)\}\right]}_{\mathrm{Conc}(\varepsilon)}.$$

接下来我们有以下三个命题，用于描述 T、ε 和 M 之间的依赖关系。命题 5. 对于多项式汤普森采样，任意 $\varepsilon > 0$ 和任意次优臂 a 均满足大发时时彩

$$\mathrm{Conc}(\varepsilon) = \mathcal{O}(1).$$

命题 6。对于多项式汤普森采样，对于任意 $\varepsilon > 0$ 和任意次优臂 a

$$\mathrm{PostCV}(\varepsilon) \leq \frac{\log T}{\mathcal{K}_{\inf}(p_a, \mu^\star, f)} + \mathcal{O}\Big(\varepsilon \cdot \log T + \sqrt{\log(T)} \cdot \log\log(T)\Big).$$

特别地，该项不依赖于 M。

P命题 7。对于多项式汤普森采样，对于任意 $\varepsilon > 0$ 和任意次优臂 a

$$\mathrm{PreCV}(\varepsilon) = \mathcal{O}\big(\varepsilon^{-9}\big).$$

这三个命题的证明大量运用了定理 1 ，并推迟到本章节末尾进行。它们意味着

$$\mathbb{E}[N_a^T] \leq \frac{\log T}{\mathcal{K}_{\inf}(p_a, \mu^\star, f)} + \mathcal{O}\Big(\varepsilon \cdot \log T + \sqrt{\log(T)} \cdot \log\log(T) + \varepsilon^{-9}\Big).$$

取 $\varepsilon = \mathcal{O}(\log^{-1/10}(T))$，我们得出该结论。

D.2 定理 3

Riou 和 Honda 给出的引理 6 [2020] 对任意 $m > 1/(1-\mu^\star)$ 均成立如下结论

$$\mathcal{K}_{\inf}(\nu_a, \mu^\star) - \frac{1}{m(1-\mu^\star) - 1} \leq \mathcal{K}_{\inf}{}^{(m)}(\nu_a^{(m)}, \mu^\star) \leq \mathcal{K}_{\inf}(\nu_a, \mu^\star).$$

此外，定理 2 表明，对于 RMTS 任意 $m > \max_{a \neq a^\star} \frac{1 + \mathcal{K}_{\inf}(\nu_a, \mu^\star)\,(1-\mu^\star) \cdot \mathcal{K}_{\inf}(\nu_a, \mu^\star)}{\mathcal{K}_{\inf}(\nu_a, \mu^\star) \cdot (1-\mu^\star)}$

$$\mathbb{E}[N_a^T] \leq \frac{\log T}{\mathcal{K}_{\inf}{}^{(m)}(\nu_a^{(m)}, \mu^\star)} + \mathcal{O}\Big(\log^{9/10}(T)\Big)$$
$$\leq \frac{\log T}{\mathcal{K}_{\inf}(\nu_a, \mu^\star)} \cdot \frac{1}{1 - [\mathcal{K}_{\inf}(\nu_a, \mu^\star) \cdot (m(1-\mu^\star) - 1)]^{-1}} + \mathcal{O}\Big(\log^{9/10}(T)\Big).$$

最后 y，对于 $x \in (0, 1/2)$ 我们有 $1/(1-x) \leq 1 + 2x$，因此对于 $m > \max_{a \neq a^\star} \frac{2 + \mathcal{K}_{\text{下确界}}(\nu_a, \mu^\star)\,(1-\mu^\star) \cdot \mathcal{K}_{\text{下确界}}(\nu_a, \mu^\star)}{\mathcal{K}_{\text{下确界}}(\nu_a, \mu^\star) \cdot (1-\mu^\star)}$

$$\mathbb{E}[N_a^T] \leq \frac{\log T}{\mathcal{K}_{\inf}(\nu_a, \mu^\star)} + \frac{2\log T}{(\mathcal{K}_{\inf}(\nu_a, \mu^\star))^2} \cdot \frac{1}{m \cdot (1-\mu^\star) - 1} + \mathcal{O}\Big(\log^{9/10}(T)\Big).$$

取 $m \geq \log(T)$，当 T 足够大时，即可得出该结论。值得注意的是，当 T 足够大时，关于 m 的所有条件均得到满足。

D.3 命题5–7

在本章节中，我们汇总了命题 5、命题 6 和命题 7。

D.3.1 命题5

由并集界可得

$$\mathrm{Conc}(\varepsilon) \le \sum_{t=1}^{T} \mathbb{P}[\neg \mathcal{E}_\varepsilon^{\mathrm{conc},1}(t-1,a)] + \sum_{t=1}^{T} \mathbb{P}[\neg \mathcal{E}_\varepsilon^{\mathrm{conc},2}(t-1)] + \sum_{t=1}^{T} \mathbb{P}[\neg \mathcal{E}_\varepsilon^{\mathrm{conc},3}(t-1)].$$

先从第一项的界开始。首先注意到， α_a^t 的分布仅取决于 $N^t(a)$ 和分布 p_a。因此，如果我们固定某个臂 a 的 $N^t(a) = n$，就可以在此条件下确定任意臂的 α_a^t 的分布。我们将相应的随机变量记为 $\alpha_a^{[n]}$。由并集界可得

$$\sum_{t=1}^{T} \mathbb{P}[\neg \mathcal{E}_\varepsilon^{\mathrm{conc},1}(t-1,a)] \le 1 + \sum_{n=1}^{T} \mathbb{E}\left[\sum_{t=0}^{T-1} \mathbb{1}\{\neg \widetilde{\mathcal{E}_\varepsilon^{\mathrm{conc}}}(N_a^t, a), N_a^t = n\}\right],$$

其中

$$\widetilde{\mathcal{E}_\varepsilon^{\mathrm{conc},1}}(n,a) = \left\{(n+\overline{\alpha}^0-1)\,\mathcal{K}_{\inf}(\overline{p}^-(\alpha_a^{[n]}), \mu^\star - \varepsilon, f) \ge n(\mathcal{K}_{\inf}(p_a, \mu^\star - \varepsilon, f) - \widetilde{\varepsilon}(n))\right\}.$$

根据引理 10 对于我们选择的 $\widetilde{\varepsilon}(n)$ 可得

$$\sum_{t=1}^{T} \mathbb{P}[\neg \mathcal{E}_\varepsilon^{\mathrm{conc},1}(t-1,a)] \le 1 + \sum_{n=1}^{T} \mathbb{P}[\neg \widetilde{\mathcal{E}_\varepsilon^{\mathrm{conc},1}}(n,a)] \le 1 + \sum_{n=1}^{T} \frac{1}{n^2} \le 1 + \frac{\pi^2}{6}.$$

对于第二项，我们有完全相同的论证，故略去。对于第三项，我们使用完全相同的技巧，将对 t 的求和改为对 n 的求和。

$$\sum_{t=1}^{T} \mathbb{P}[\neg \mathcal{E}_\varepsilon^{\mathrm{conc},3}(t-1)] \le 1 + \sum_{n=1}^{T} \mathbb{P}[\neg \widetilde{\mathcal{E}_\varepsilon^{\mathrm{conc},3}}(n)],$$

其中

$$\widetilde{\mathcal{E}_\varepsilon^{\mathrm{conc},3}}(n) = \left\{(n+\overline{\alpha}^0-1)\,\mathcal{K}_{\inf}(\overline{p}^+(\alpha_{a^\star}^{[n]}), \mu^\star, f) \le 3\log(2n+1) + (\overline{\alpha}^0-1)\log\left(\frac{1}{1-\mu^\star}\right)\right\}.$$

应用引理11 可得

$$\sum_{t=1}^{T} \mathbb{P}[\neg \mathcal{E}_\varepsilon^{\mathrm{conc},3}(t-1)] \le 1 + \mathrm{e}\sum_{n=1}^{T} \frac{1}{(2n+1)^2} \le 1 + \frac{\mathrm{e}\pi^2}{8}.$$

D.3.2 命题6 的证明

Proof. 定义一个常数 n_0，其具体取值将在后文给出。于是我们有如下结果。

$$\begin{aligned}\mathrm{PostCV}(\varepsilon) &\le n_0 + \mathbb{E}\left[\sum_{t=1}^{T} \mathbb{1}\{A^t = a, w_a^t f \ge \mu^\star - \varepsilon, \mathcal{E}_\varepsilon^{\mathrm{conc},1}(t-1,a), N_a^{t-1} \ge n_0\}\right] \\ &\le n_0 + \sum_{t=0}^{T-1}\sum_{n=n_0}^{T} \mathbb{P}\big[w_a^{t+1} f \ge \mu^\star - \varepsilon, N_a^t = n, \mathcal{E}_\varepsilon^{\mathrm{conc},1}(t,a)\big],\end{aligned}$$

接下来我们注意到， N_a^t 和 $\mathbb{1}\{\mathcal{E}_\varepsilon^{\mathrm{conc},1}(t,a)\}$ 都是 α_a^t 的函数，因此，根据条件期望的塔式性质

$$\mathbb{P}\big[w_a^{t+1} f \ge \mu^\star - \varepsilon, N_a^t = n, \mathcal{E}_\varepsilon^{\mathrm{conc},1}(t,a)\big] = \mathbb{E}\big[\mathbb{P}[w_a^{t+1} f \ge \mu^\star - \varepsilon | \alpha_a^t] \mathbb{1}\{N_a^t = n, \mathcal{E}_\varepsilon^{\mathrm{conc},1}(t,a)\}\big],$$

其中我们可以假设 α_a^t 满足 $\mathcal{E}^{\text{conc}}_{,\varepsilon}$ $^1(t,a)$ 和 $N_a^t = n$ 的定义。

接下来我们注意到

$$\mathbb{P}[w_a^{t+1} f \geq \mu^\star - \varepsilon | \alpha_a^t] = \mathbb{P}_{w\sim \text{Dir}(\alpha_a^t)}[wf \geq \mu^\star - \varepsilon].$$

并且我们自动得到 $\overline{\alpha} = n + \overline{\alpha}^0$。接下来，我们注意到，如果 $\overline{p}^{(-)}(\alpha_a^t) f > \mu^\star - \varepsilon$，则所需的概率上界为 1。否则，我们可以应用定理 1

$$\mathbb{P}_{w\sim \text{Dir}(\alpha_a^t)}[wf \geq \mu^\star - \varepsilon] \leq 2\mathbb{P}_{g\sim\mathcal{N}(0,1)}\left[g \geq \sqrt{2(n+\overline{\alpha}^0 - 1)\, \mathcal{K}_{\inf}(\overline{p}^-(\alpha_a^t), \mu^\star - \varepsilon, f)}\right].$$

由于 $\mathcal{K}_{\inf}$ 是非负的，该上界在 $\overline{p}^{(-)}(\alpha_a^t) f > \mu^\star - \varepsilon$时同样成立。因此，在事件 $\mathcal{E}^{\text{conc}}_{,\varepsilon}$ $^1(t,a)$ 对于 $N_a^t = n$ 我们有

$$\begin{aligned}
&\mathbb{P}_{g\sim\mathcal{N}(0,1)}\left[g \geq \sqrt{2(n+\overline{\alpha}^0 - 1)\, \mathcal{K}_{\inf}(\overline{p}^-(\alpha_a^t), \mu^\star - \varepsilon, f)}\right] \\
&\qquad \leq \mathbb{P}_{g\sim\mathcal{N}(0,1)}\left[g \geq \sqrt{2n(\mathcal{K}_{\inf}(p_a, \mu^\star - \varepsilon, f) - \widetilde{\varepsilon}(n))}\right].
\end{aligned}$$

总体而言，对于任意 $t = 0, \ldots, T-1$，我们得到

$$\begin{aligned}
&\mathbb{P}\left[w_a^{t+1} u \geq \mu^\star - \varepsilon | \alpha_a^t\right] \mathbb{1}\{N_a^t = n, \mathcal{E}_\varepsilon^{\text{conc},1}(t,a)\} \\
&\qquad \leq 2\mathbb{P}_{g\sim\mathcal{N}(0,1)}\left[g \geq \sqrt{2n(\mathcal{K}_{\inf}(p(a), \mu^\star - \varepsilon, f) - \widetilde{\varepsilon}(n))}\right] \mathbb{1}\{N_a^t = n\}.
\end{aligned}$$

概率的右侧关于 n 严格递增，因为 $n = \Omega(1)$，我们可以保证 $\mathcal{K}_{\inf}(p(a)\ \mu^\star - \varepsilon, f) - \widetilde{\varepsilon}(n) \geq 1/2\, \mathcal{K}_{\inf}(p(a)\ \mu^\star - \varepsilon, f)$。因此，我们可以获得对所有 $n \geq n_0$ 的一致上界，这意味着

$$\text{PostCV}(\varepsilon) \leq n_0 + 2T \cdot \mathbb{P}_{g\sim\mathcal{N}(0,1)}\left[g \geq \sqrt{2n_0(\mathcal{K}_{\inf}(p, \mu^\star - \varepsilon, f) - \widetilde{\varepsilon}(n_0))}\right].$$

接下来，对于任意 $x > 0$，以下正态分布尾部的上界成立（例如参见Vershynin [2018]）

$$\mathbb{P}_{g\sim\mathcal{N}(0,1)}\left[g \geq \sqrt{2x}\right] \leq \frac{\exp(-x)}{2\sqrt{\pi x}}.$$

因此，对于任意 $\varepsilon' > 0$，取 $n_0 = (1+\varepsilon')\log T / \mathcal{K}_{\inf}(p, \mu^\star - \varepsilon, u)$，我们有

$$\text{PostCV}(\varepsilon) \leq \frac{(1+\varepsilon')\log T}{\mathcal{K}_{\inf}(p_a, \mu^\star - \varepsilon, f)} + 2T \cdot \frac{\exp\left(-(1+\varepsilon')\log T + \mathcal{O}\left(\sqrt{\log\log(T) \cdot \log(T)}\right)\right)}{\sqrt{2\pi(1+\varepsilon')\log(T)}}$$

特别地，我们可以取 $\varepsilon' = \log\log(T)/\sqrt{\log(T)}$ 以使最后一项在 T 足够大时趋于零。最后，利用 Garivier等人 [2022] 的引理 11，我们有

$$\text{PostCV}(\varepsilon) \leq \frac{\log T}{\mathcal{K}_{\inf}(p_a, \mu^\star, f)} + \mathcal{O}\left(\varepsilon \cdot \log T + \sqrt{\log(T)} \cdot \log\log(T)\right).$$

□

D.3.3 命题7

Proof. 我们用 $\mathcal{F}^t$ 表示轮次 t结束时可获得的信息。首先注意到，引入具有最大索引 $\widetilde{A}^t = \max_{a' \neq a^\star} w_{a'}^t f$的次优臂，根据后验样本的独立性，有

$$\begin{aligned}
\mathbb{P}[A^t = a^\star, w_a^t f \leq \mu^\star - \varepsilon | \mathcal{F}^{t-1}] &\geq \mathbb{P}[\widetilde{A}^t = a, w_a^t f \leq \mu^\star - \varepsilon, w_{a^\star}^t f > \mu^\star - \varepsilon | \mathcal{F}^{t-1}] \\
&= \left(1 - \mathbb{P}[w_{a^\star}^t f > \mu^\star - \varepsilon | \mathcal{F}^{t-1}]\right) \mathbb{P}[\widetilde{A}^t = a, w_a^t f \leq \mu^\star - \varepsilon | \mathcal{F}^{t-1}]
\end{aligned}$$

and

$$\begin{aligned}\mathbb{P}[A^t = a, w_a^t f \le \mu^\star - \varepsilon | \mathcal{F}^{t-1}] &\le \mathbb{P}[\widetilde{A}^t = a, w_a^t f \le \mu^\star - \varepsilon, w_{a^\star}^t f \le \mu^\star - \varepsilon | \mathcal{F}^{t-1}] \\ &= \mathbb{P}[\widetilde{A}^t = a, w_a^t f \le \mu^\star - \varepsilon | \mathcal{F}_{t-1}] \mathbb{P}[w_{a^\star}^t f \le \mu^\star - \varepsilon | \mathcal{F}^{t-1}] .\end{aligned}$$

由这两个不等式，我们推得

$$\mathbb{P}[A^t = a, w_a^t f \le \mu^\star - \varepsilon | \mathcal{F}^{t-1}] \le \frac{\mathbb{P}[w_{a^\star}^t f \le \mu^\star - \varepsilon | \mathcal{F}^{t-1}]}{1 - \mathbb{P}[w_{a^\star}^t f \le \mu^\star - \varepsilon | \mathcal{F}_{t-1}]} \mathbb{P}[A^t = a^\star, w_a^t f \le \mu^\star - \varepsilon | \mathcal{F}^{t-1}] .$$

对 $\mathcal{F}^{t-1}$ 条件化并使用前一个不等式，然后从全局时间切换到计数，我们得到

$$\begin{aligned}\mathrm{PreCV}(\varepsilon) &= \mathbb{E}\left[\sum_{t=1}^{T} \mathbb{1}\{A^t = a, w_a^t f \le \mu^\star - \varepsilon, \mathcal{E}_\varepsilon^{\mathrm{conc}}(t-1, a)\}\right] \\ &\le \mathbb{E}\left[\sum_{t=1}^{T} \frac{\mathbb{P}[w_{a^\star}^t f \le \mu^\star - \varepsilon | \mathcal{F}^{t-1}]}{1 - \mathbb{P}[w_{a^\star}^t f \le \mu^\star - \varepsilon | \mathcal{F}_{t-1}]} \mathbb{1}\{A^t = a^\star, \mathcal{E}_\varepsilon^{\mathrm{conc},2}(t-1) \cap \mathcal{E}_\varepsilon^{\mathrm{conc},3}(t-1)\}\right] \\ &\le \mathbb{E}\left[\sum_{n=0}^{T-1} \frac{\mathbb{P}_{w \sim \mathrm{Dir}(\alpha_{a^\star}^{(n)})}[wf \le \mu^\star - \varepsilon]}{1 - \mathbb{P}_{w \sim \mathrm{Dir}(\alpha_{a^\star}^{(n)})}[wf \le \mu^\star - \varepsilon]} \mathbb{1}\{\widetilde{\mathcal{E}_\varepsilon^{\mathrm{conc},2}}(n) \cap \widetilde{\mathcal{E}_\varepsilon^{\mathrm{conc},3}}(n)\}\right]\end{aligned}$$

其中 $\alpha_{a^\star}^{(n)}$ 是观测 n 次后最优臂的后验分布参数，且$\widetilde{\quad}$ $\widetilde{\quad}$事件 $\mathcal{E}^{\mathrm{conc}}{}_{,\varepsilon}{}^{2}(n)$, $\mathcal{E}^{\mathrm{conc}}{}_{,\varepsilon}{}^{3}(n$在命题 5 的证明中定义，满足

$$\{\mathcal{E}_\varepsilon^{\mathrm{conc},2}(\tau_n) \cap \mathcal{E}_\varepsilon^{\mathrm{conc},3}(\tau_n)\} = \{\widetilde{\mathcal{E}_\varepsilon^{\mathrm{conc},2}}(n) \cap \widetilde{\mathcal{E}_\varepsilon^{\mathrm{conc},3}}(n)\},$$

其中 τ_n 是满足 $N_{a^\star}^{\tau_n} = n$ 的第一个时刻。

为了恰当地处理这个期望，我们将其分为两种情况。

$$(\mathbf{A}) = \sum_{n=0}^{T} \mathbb{E}\left[\frac{\mathbb{P}_{w \sim \mathrm{Dir}(\alpha_{a^\star}^{(n)})}[wf \le \mu^\star - \varepsilon]}{\mathbb{P}_{w \sim \mathrm{Dir}(\alpha_{a^\star}^{(n)})}[wf > \mu^\star - \varepsilon]} \cdot \mathbb{1}\{\widetilde{\mathcal{E}_\varepsilon^{\mathrm{conc},3}}(n), \overline{p}(\alpha_{a^\star}^{(n)}) \cdot f \le \mu^\star - \varepsilon\}\right],$$

and

$$(\mathbf{B}) = \sum_{n=0}^{T} \mathbb{E}\left[\frac{\mathbb{P}_{w \sim \mathrm{Dir}(\alpha_{a^\star}^{(n)})}[wf \le \mu^\star - \varepsilon]}{\mathbb{P}_{w \sim \mathrm{Dir}(\alpha_{a^\star}^{(n)})}[wf > \mu^\star - \varepsilon]} \cdot \mathbb{1}\{\widetilde{\mathcal{E}_\varepsilon^{\mathrm{conc},2}}(n), \overline{p}(\alpha_{a^\star}^{(n)}) \cdot f > \mu^\star - \varepsilon\}\right].$$

我们的目标是分别界定这两项。

项 (A) 对于这一项，我们有

$$\mathbb{P}_{w \sim \mathrm{Dir}(\alpha_{a^\star}^{(n)})}[wf \le \mu^\star - \varepsilon] \le 1$$

并由定理 1,

$$\begin{aligned}\mathbb{P}_{w \sim \mathrm{Dir}(\alpha_{a^\star}^{(n)})}[wf > \mu^\star - \varepsilon\}] &\ge \frac{1}{2} \mathbb{P}_{g \sim \mathcal{N}(0,1)}\left[g \ge \sqrt{2(n + \overline{\alpha}^0 - 1)\, \mathcal{K}_{\inf}(\overline{p}^+(\alpha_{a^\star}^{(n)}), \mu^\star - \varepsilon, f)}\right] \\ &\ge \frac{1}{2} \mathbb{P}_{g \sim \mathcal{N}(0,1)}\left[g \ge \sqrt{2(n + \overline{\alpha}^0 - 1)\, \mathcal{K}_{\inf}(\overline{p}^+(\alpha_{a^\star}^{(n)}), \mu^\star, f)}\right].\end{aligned}$$

根据事件的定义。 $\widetilde{\mathcal{E}^{\mathrm{conc}}{}_{,\varepsilon}}$ $^{3}(n)$ 我们有

$$(n + \overline{\alpha}^0 - 1)\, \mathcal{K}_{\inf}(\overline{p}^+(\alpha_{a^\star}^{(n)}), \mu^\star, f) \le 3 \log(2n+1) + (\overline{\alpha}^0 - 1) \log\left(\frac{1}{1 - \mu^\star}\right).$$

注意到　在 $n \geq 1$ 处，我们有 $3\log(2n+1) \geq 1$，因此高斯尾部的下界意味着

$$\mathbb{P}_{w\sim \mathrm{Dir}(\alpha_{a^\star}^{(n)})}[wf > \mu^\star - \varepsilon\}] \geq \frac{(2n+1)^{-3}\cdot(1-\mu^\star)^{\overline{\alpha}^0-1}}{4\sqrt{2}\sqrt{3\log(2n+1)+(\overline{\alpha}^0-1)\log\left(\frac{1}{1-\mu^\star}\right)}}.$$

因此，我们得到以下界

$$\begin{aligned}(\mathbf{A}) &\leq \sum_{n=0}^{T}\mathcal{O}(n^3\log^{1/2}(n))\mathbb{E}\Big[\mathbb{1}\{\overline{p}(\alpha_{a^\star}^{(n)})\cdot f \leq \mu^\star-\varepsilon\}\Big]\\ &\leq \sum_{n=0}^{T}\mathcal{O}(n^3\log^{1/2}(n))\mathbb{P}_{\alpha\sim\mathrm{Mult}(n,p_{a^\star})}\big[\overline{p}(\alpha+\alpha^0)\cdot f \leq \mu^\star-\varepsilon\big].\end{aligned}$$

为了给其他概率一个上界，我们使用霍夫丁不等式

$$\begin{aligned}&\mathbb{P}_{\alpha\sim\mathrm{Mult}(n,p_{a^\star})}\big[\overline{p}(\alpha+\alpha^0)\cdot u \leq \mu^\star-\varepsilon\big]\\ &\qquad\leq \mathbb{P}_{\alpha\sim\mathrm{Mult}(n,p_{a^\star})}\big[|\langle\alpha,u\rangle - n\mu^\star| \geq n\varepsilon - \overline{\alpha}^0(\mu^\star-\langle\alpha^0,u\rangle-\varepsilon)\big]\\ &\qquad\leq \exp\left(-2n\left(\varepsilon-\frac{\overline{\alpha}^0}{n}(\mu^\star-\langle\alpha^0,u\rangle-\varepsilon)\right)^2\right).\end{aligned}$$

定义截断点 $n_0' = 4\overline{\alpha}^0\mu^\star/\varepsilon$。对于 $n \geq n_0'$，我们有

$$\mathbb{P}_{\alpha\sim\mathrm{Mult}(n,p_{a^\star})}\big[\overline{p}(\alpha+\alpha^0)\cdot u \leq \mu^\star-\varepsilon\big] \leq \exp\left(\frac{-n\varepsilon^2}{2}\right).$$

对于 $n \leq n_0'$，由 1 可得一个平凡的上界。因此，我们得到

$$(\mathbf{A}) \leq \mathcal{O}\left(\frac{\log^{1/2}(1/\varepsilon)}{\varepsilon^4}\right) + \mathcal{O}\left(\sum_{n=n_0'}^{T} n^3\log^{1/2}(n)\mathrm{e}^{-n\varepsilon^2/2}\right).$$

接下来，我们希望选择下一个截断点 n_0''，使得

$$(n_0'')^{3.5} \leq \exp(n_0''\varepsilon^2/4) \iff 14\log(n_0'') \leq n_0''\cdot\varepsilon^2.$$

根据不等式 $\log(x) \leq \beta\cdot x^{1/\beta}$，对于任意 $x>0$ 和 $\beta>0$，我们可以得到 n_0'' 的值，该值确实满足上述不等式 $n_0'' = (14\beta)^{\beta/(\beta-1)}\cdot\varepsilon^{-2\beta/(\beta-1)}$，对于任意常数 $\beta>0$。因此，我们有

$$(\mathbf{A}) \leq \mathcal{O}\left(\frac{\log^{1/2}(1/\varepsilon)}{\varepsilon^4}\right) + \mathcal{O}\Big(\varepsilon^{-4\beta/(\beta-1)}\log^{1/2}(\varepsilon)\Big) + \mathcal{O}\left(\sum_{n=n_0''}^{T}\mathrm{e}^{-n\varepsilon^2/4}\right).$$

取 $\beta = 17/9$，我们可以看到第二项在 ε 中占主导，而且，利用不等式 $\log^{1/2}(1/\varepsilon) \leq 2\varepsilon^{1/2}$，我们有

$$(\mathbf{A}) = \mathcal{O}\big(\varepsilon^{-9}\big).$$

项 **(B)** 为了分析这一项，我们再次从期望下的概率入手。首先，根据定理 1，

$$\mathbb{P}_{w\sim\mathrm{Dir}(\alpha_{a^\star}^{(n)})}[wf > \mu^\star-\varepsilon] \geq \mathbb{P}_{w\sim\mathrm{Dir}(\alpha_{a^\star}^{(n)}))}\big[wf > \overline{p}^+(\alpha^\tau)\cdot f\big] \geq \frac{1}{2}$$

此外，根据定理1

$$\mathbb{P}_{w\sim\mathrm{Dir}(\alpha^{(n)}_{a^\star})}[wf \le \mu^\star - \varepsilon] = \mathbb{P}_{w\sim\mathrm{Dir}(\alpha^{(n)}_{a^\star})}[w(1-f) \ge 1-\mu^\star+\varepsilon]$$
$$\le \frac{3}{2}\mathbb{P}_{g\sim\mathcal{N}(0,1)}\left[g \ge \sqrt{2(n+\overline{\alpha}^0-1)\,\mathcal{K}_{\inf}(\overline{p}^-(\alpha^{(n)}_{a^\star})), 1-\mu^\star+\varepsilon, 1-f)}\right].$$

根据事件的定义 $\mathcal{E}^{\mathrm{conc}}$, ε ^{2} (n) 以及 Garivier 等人的引理 11Garivier等人 [2022]，我们有。

$$\mathbb{P}_{w\sim\mathrm{Dir}(\alpha^{(n)}_{a^\star})}[wf \le \mu^\star - \varepsilon] \le \frac{3}{2}\mathbb{P}_{g\sim\mathcal{N}(0,1)}\left[g \ge \sqrt{2n(2\varepsilon^2 - \widetilde{\varepsilon}(n))}\right].$$

定义截断点 $\tilde{n}_0$ 使得 $\varepsilon^2 \ge \widetilde{\varepsilon}(\tilde{n}_0)$，通过选择$n_0 = \mathcal{O}(\varepsilon^{-5})$ 可自动满足该条件。因此，根据正态分布尾部的上界，我们得到

$$(\mathbf{B}) \le \mathcal{O}(\varepsilon^{-5}) + \sum_{n=\tilde{n}_0}^{T} \frac{\exp(-n\varepsilon^2)}{\sqrt{n\varepsilon^2}} = \mathcal{O}(\varepsilon^{-5}).$$

总体而言，我们得到

$$\mathrm{PreCV}(\varepsilon) = \mathcal{O}(\varepsilon^{-9}).$$

□

E 技术引理

首先，我们先证明一个辅助结果，类似于Tiapkin 等人 [2022]

引理 8 *Let f $\{0,\dots,m\} \to [0,b]$ be a function such that f 0 $= 0$ and f m $= b$.。: Let α_0 ... α_m $\in \mathbb{R}^{m+1}_+$ be a vector such that $\alpha_m > 0$ and define $\overline{\alpha} = \sum_{i=0}^m \alpha_i$. Define a () ( ) (, ,) distribution $\overline{p} \in \Delta_m$ such that $\overline{p}(i)$ $= \alpha_i/\overline{\alpha}$ and consider a real number $\mu \in (\overline{p}f, b)$ Then*

$$1 - \lambda^\star(\overline{p},\mu,f)\cdot(b-\mu) \ge \overline{p}(m),$$

where $\lambda^\star(\overline{p},\mu,f)$ is defined as an optimal dual variables to the variational formula

$$\lambda^\star(p,\mu,f) = \underset{\lambda\in[0,1/(b-\mu)]}{\arg\max}\ \mathbb{E}_{X\sim p}[\log(1+\lambda\cdot(f(X)-\mu))].$$

Proof. 在条件 $\overline{p}f < \mu < b$ 和 $\overline{p}(m)$ > 0 下，值 $\lambda^\star = \lambda^\star(\overline{p},\mu,f)$ 满足以下方程

$$\mathbb{E}\left[\frac{f(X)-\mu}{1-\lambda^\star\cdot(f(X)-\mu)}\right] = \overline{p}(m)\frac{b-\mu}{1-\lambda^\star\cdot(b-\mu)} + \sum_{j=0}^{m-1}\overline{p}(j)\frac{f(j)-\mu}{1-\lambda^\star\cdot(f(j)-\mu)} = 0. \tag{15}$$

定义一个分布 $\widetilde{p}$其中 $\widetilde{p}(i) = \overline{p}(i)/(1-\overline{p}(m))$，当 $i \in \{0,\dots,m-1\}$ 且 $\widetilde{p}(m)$ $= 0$。则（15）中的期望可以写成

$$\overline{p}(m)\frac{b-\mu}{1-\lambda^\star\cdot(b-\mu)} + (1-\overline{p}(m))\mathbb{E}_{X\sim\widetilde{p}}\left[\frac{f(X)-\mu}{1-\lambda^\star\cdot(f(X)-\mu)}\right] = 0.$$

定义一个函数 $w(x,u) = \frac{x-u}{1-\lambda^\star\cdot(x-u)}$，其关于 x. 是凸函数。根据詹森不等式，

$$\mathbb{E}_{X\sim\widetilde{p}}\left[\frac{f(X)-\mu}{1-\lambda^\star\cdot(f(X)-\mu)}\right] \ge \frac{\widetilde{p}f-\mu}{1-\lambda^\star\cdot(\widetilde{p}f-\mu)}.$$

因此

$$\overline{p}(m)\frac{b-\mu}{1-\lambda^\star\cdot(b-\mu)} \le -(1-\overline{p}(m))\frac{\widetilde{p}f-\mu}{1-\lambda^\star\cdot(\widetilde{p}f-\mu)} = (1-\overline{p}(m))\frac{\mu-\widetilde{p}f}{1+\lambda^\star\cdot(\mu-\widetilde{p}f)}.$$

注意， $\mu \geq \overline{p}_n f \geq \widetilde{p}_n f$，因此我们可以如下重新排列各项

$$\begin{aligned}\frac{1}{\overline{p}(m)}\left(\frac{1}{b-\mu}-\lambda^\star\right) &\geq \frac{1}{1-\overline{p}(m)}\left(\frac{1}{\mu-\widetilde{p}f}+\lambda^\star\right)\\ &= \frac{1}{1-\overline{p}(m)}\left(\frac{1}{\mu-\widetilde{p}f}+\frac{1}{b-\mu}\right)-\frac{1}{1-\overline{p}(m)}\left(\frac{1}{b-\mu}-\lambda^\star\right).\end{aligned}$$

因此，

$$\left(\frac{1}{b-\mu}-\lambda^\star\right) \geq \overline{p}(m)\cdot\frac{b-\widehat{p}f}{(b-\mu)(\mu-\widehat{p}f)} \geq \overline{p}(m)\frac{1}{b-\mu}.$$

□

接下来，我们分析当第一个参数（分布）发生微小变化时， $\mathcal{K}_{\inf}$ 如何变化。对于任意非负向量 $s \in \mathbb{R}_+^{m+1}$，我们将 $\overline{p}_s$ 定义为 $\overline{p}_s(i) = s_i/(\sum_{i=0}^m s_i)$。

引理 **9** *Let $f : \{0,\dots,m\} \to [0,b]$ be a function such that $f(0) = 0$ and $f(m) = b$. Let $\alpha = (\alpha_0, \dots, \alpha_m) \in \mathbb{R}_+^{m+1}$ and $\beta = (\beta_0, \dots, \beta_m) \in \mathbb{R}_+^{m+1}$ be non-negative vectors such that $\beta_m + \alpha_m \geq 1$. Set $\overline{\alpha} = \sum_{i=0}^m \alpha_i$ and $\overline{\beta} = \sum_{i=0}^m \beta_i$ then the following bounds hold for any, $\mu \in (0,b)$,*

$$\mathcal{K}_{inf}(\overline{p}_\alpha,\mu,f) - \frac{3\overline{\beta}}{\overline{\alpha}}\log\left(\frac{b(\overline{\alpha}+\overline{\beta})}{b-\mu}\right) \leq \mathcal{K}_{inf}(\overline{p}_{\alpha+\beta},\mu,f) \leq \frac{\overline{\alpha}\cdot\mathcal{K}_{inf}(\overline{p}_\alpha,\mu,f)+\overline{\beta}\cdot\log\left(\frac{b}{b-\mu}\right)}{\overline{\alpha}+\overline{\beta}}.$$

Proof. 我们从下界开始。设 $q^\star$ 是一个概率测度，它在 $\mathcal{K}_{\inf}(\overline{p}_{\alpha+\beta},\mu,f)$ 的定义中取到极值，也就是说，

$$q^\star = \arg\min_{q\in\Delta_m}\{\mathrm{KL}(\overline{p}_{\alpha+\beta},q) : qf \geq \mu\}.$$

于是我们有

$$\begin{aligned}\mathcal{K}_{\inf}(\overline{p}_{\alpha+\beta},\mu,f) = \mathrm{KL}(\overline{p}_{\alpha+\beta},q^\star) = &\underbrace{\sum_{i:\overline{p}_{\alpha+\beta}(i)>0}(\overline{p}_{\alpha+\beta}(i)-\overline{p}_\alpha(i))\log\left(\frac{\overline{p}_{\alpha+\beta}(i)}{q^\star(i)}\right)}_{(\mathbf{A})}\\ &+\underbrace{\sum_{i:\overline{p}_{\alpha+\beta}(i)>0}\overline{p}_\alpha(i)\log\left(\frac{\overline{p}_{\alpha+\beta}(i)}{q^\star(i)}\right)}_{(\mathbf{B})}.\end{aligned}$$

项 **(A)** 让我们分析第一个求和。由于 $\overline{p}_{\alpha+\beta}(m+1) > 0$，根据 $\mathcal{K}_{\inf}$ 的对偶表示，可以得到最优 $q^\star$ 的公式

$$q^\star(i) = \frac{\overline{p}_{\alpha+\beta}(i)}{1-\lambda^\star\cdot(f(i)-\mu)},$$

其中， $\lambda^\star \in [0, 1/(b-\mu))$ 是 $\mathcal{K}_{\inf}(\overline{p}_{\alpha+\beta},\mu,f)$ 的变分问题的一个解。因此，我们有

$$\sum_{i:\overline{p}_{\alpha+\beta}(i)>0}(\overline{p}_{\alpha+\beta}(i)-\overline{p}_\alpha(i))\log\left(\frac{\overline{p}_{\alpha+\beta}(i)}{q^\star(i)}\right) \geq -\|\overline{p}_{\alpha+\beta}-\overline{p}_\alpha\|_1\cdot\|\log(1-\lambda^\star\cdot(f-\mu))\|_\infty.$$

对于第一个乘子，我们有以下界

$$\|\overline{p}_{\alpha+\beta}-\overline{p}_\alpha\|_1 = \left\|\frac{\overline{\alpha}}{\overline{\alpha}+\overline{\beta}}\overline{p}_\alpha+\frac{1}{\overline{\alpha}+\overline{\beta}}\beta-\overline{p}_\alpha\right\|_1 = \frac{1}{\overline{\alpha}+\overline{\beta}}\left\|\beta-\overline{\beta}\cdot\overline{p}_\alpha\right\|_1 \leq \frac{2\overline{\beta}}{\overline{\alpha}+\overline{\beta}}.$$

为了分析第二个乘子，我们注意到，由于对数是单调函数且 $\lambda^\star$ 固定，最大值只能在 $i=0$ 或 $i=m$ 处取得。因此

$$\max_{i\in\{0,\dots,m\}}|\log(1-\lambda^\star\cdot(f(i)-\mu))|=\max\left\{\log(1+\lambda^\star\mu),\log\left(\frac{1}{1-\lambda^\star(b-\mu)}\right)\right\}.$$

应用引理 8，我们得到

$$\max\left\{\log(1+\lambda^\star\mu),\log\left(\frac{1}{1-\lambda^\star(b-\mu)}\right)\right\}\le\max\left\{\log\left(\frac{b}{b-\mu}\right),\log\left(\frac{1}{\overline{p}_{\alpha+\beta}(m)}\right)\right\}.$$

因此，由于 $\alpha_m+\beta_m\ge 1$，我们有 $\overline{p}_n(m)\ge 1/(\overline{\alpha}+\overline{\beta})$，从而

$$(\mathbf{A})\ge-\frac{2\overline{\beta}}{\overline{\alpha}+\overline{\beta}}\cdot\max\left\{\log\left(\frac{b}{b-\mu}\right),\log(\overline{\alpha}+\overline{\beta})\right\}\ge-\frac{2\overline{\beta}}{\overline{\alpha}+\overline{\beta}}\log\left(\frac{b\cdot(\overline{\alpha}+\overline{\beta})}{b-\mu}\right).$$

项（B）接下来我们分析第二个和。由于 $\mathrm{supp}(\overline{p}_\alpha)\subseteq\mathrm{supp}(\overline{p}_{\alpha+\beta})=\mathrm{supp}(q^\star)$，我们有

$$\sum_{i:\overline{p}_{\alpha+\beta}(i)>0}\overline{p}_\alpha(i)\log\left(\frac{\overline{p}_{\alpha+\beta}(i)}{q^\star(i)}\right)=\sum_{i:\overline{p}_\alpha(i)>0}\overline{p}_\alpha(i)\log\left(\frac{\overline{p}_\alpha(i)}{q^\star(i)}\right)+\sum_{i:\overline{p}_\alpha(i)>0}\overline{p}_\alpha(i)\log\left(\frac{\overline{p}_{\alpha+\beta}(i)}{\overline{p}_\alpha(i)}\right)$$

第一项等于 KL($\overline{p}_\alpha\ q^\star$ $\mathcal{K}_{\inf}\ \overline{p}_\alpha\ \mu\ f$ $q^\star f\ge,),$
$\overline{p}_\alpha\ i\ >0\iff\alpha_i>0$
并且可以用 (,,) $^{\mu}$ 作为下界，因为 。对于第二项，我们分析条件（ ）下的对数。可得

$$\log\left(\frac{\overline{p}_{\alpha+\beta}(i)}{\overline{p}_\alpha(i)}\right)=\log\left(\frac{\alpha_i+\beta_i}{\alpha_i}\right)+\log\left(\frac{\overline{\alpha}}{\overline{\alpha}+\overline{\beta}}\right)\ge-\log\left(1+\frac{\overline{\beta}}{\overline{\alpha}}\right)\ge-\frac{\overline{\beta}}{\overline{\alpha}},$$

因此

$$(\mathbf{B})\ge\mathcal{K}_{\inf}(\overline{p}_\alpha,\mu,f)-\frac{\overline{\beta}}{\overline{\alpha}}.$$

最后，我们有

$$\mathcal{K}_{\inf}(\overline{p}_{\alpha+\beta},\mu,f)\ge\mathcal{K}_{\inf}(\overline{p}_\alpha,\mu,f)-\frac{3\overline{\beta}}{\overline{\alpha}}\log\cdot\left(\frac{b\cdot(\overline{\alpha}+\overline{\beta})}{b-\mu}\right).$$

对于上界，我们使用 $\mathcal{K}_{\inf}$ 的变分公式和一致上界（参见Honda和Takemura [2010]）以得到

$$\begin{aligned}(\overline{\alpha}+\overline{\beta})\,\mathcal{K}_{\inf}(\overline{p}_{\alpha+\beta},\mu,f)&=\max_{\lambda\in[0,1/(b-\mu)]}\sum_{i=0}^{m}(\alpha_i+\beta_i)\log(1-\lambda\cdot(f(i)-\mu))\\&\le\max_{\lambda\in[0,1/(b-\mu)]}\sum_{i=0}^{m}\alpha_i\log(1-\lambda\cdot(f(i)-\mu))\\&\quad+\max_{\lambda\in[0,1/(b-\mu)]}\sum_{i=0}^{m}\beta_i\log(1-\lambda\cdot(f(i)-\mu))\\&\le\overline{\alpha}\,\mathcal{K}_{\inf}(\overline{p}_\alpha,\mu,f)+\overline{\beta}\log\left(\frac{b}{b-\mu}\right).\end{aligned}$$

□

设 $X_1,\dots,X_n$ 是来自类别分布 $p\in\Delta_m$的独立同分布样本。我们将值 j的频数记为 ν_j：$\nu_j=\sum_{i=1}^n\mathbb{1}\{X_i=j\}$，并将经验概率分布记为 $\widehat{p}_n$，定义为 $\widehat{p}_n(j)=\nu_j/n$，对所有 $j\in\{0,\dots,m\}$均成立。此外，令 $\alpha^0\in\mathbb{R}^{m+1}$ 为先验分布计数器，使得$\sum_{j=0}^m\alpha_j^0=n_0$。然后我们定义平移后的经验概率分布为 $\overline{p}_n(j)=\alpha_j^n/(n+n_0)$，其中 $\alpha_j^n=\alpha_j^0+\nu_j$。在此设置下，我们给出关于平移分布的偏差与集中性的技术性结果。

引理10. *Let* $f\colon \{0,\dots,m\} \to [0,b]$ *be a function such that* $f(0) = 0$ *and* $f(m) = b$. *Consider a real number* $\mu \in (\overline{p}_n f, b)$ *and define*

$$\gamma = \sqrt{\frac{b}{b-\mu}\left(16\mathrm{e}^2 + \log^2\left(\frac{b}{b-\mu}\right)\right)}.$$

Assume that $\alpha_m^0 \geq 1$. *Then we have for any* $0 < \delta < \gamma/2$ *we have*

$$\mathbb{P}\left[n\,\mathcal{K}_{inf}(\overline{p}_n, \mu, f) \leq n(\mathcal{K}_{inf}(p, \mu, f) - \delta) - 3n_0 \log\left(\frac{b\cdot(n+n_0)}{b-\mu}\right)\right] \leq \exp(-n\delta^2/(2\gamma)).$$

Proof. 注意到 $\widehat{p}_n$ 根据引理 $\overline{p}_\nu$，等于 9，而且 $\overline{p}_n$ 等于 $\overline{p}_{\nu+\alpha^0}$。因此，我们可以结合引理 9 和 Garivier 等人的命题16 [2022]，从而得出该结论。 □

此外，我们陈述移位经验分布所需的偏差结果。

引理 11. *For any* $u \geq 0$ *we have*

$$\mathbb{P}\left[(n+n_0)\,\mathcal{K}_{inf}(\overline{p}_n, pf, f) \geq nu + n_0 \log\left(\frac{b}{b-\mu}\right)\right] \leq \mathrm{e}(2n+1)\mathrm{e}^{-nu}.$$

Proof. 由引理9

$$(n+n_0)\,\mathcal{K}_{\inf}(\overline{p}_n, \mu, f) \leq n\,\mathcal{K}_{\inf}(\widehat{p}_n, \mu, f) + n_0 \log\left(\frac{b}{b-\mu}\right)$$

然后我们可以应用 Garivier等人 [2022] 提出的命题14来处理 $\mathcal{K}_{\inf}(\widehat{p}_n,\ \mu,\ f)$。 □

参考文献

R. Agrawal. 用于多臂老虎机问题的基于样本均值的索引策略，遗憾为 $o(\log n)$。 *Advances in Applied Probability*，27(4):1054–1078，1995年。

Shipra Agrawal和 Navin Goyal。《汤普森抽样的进一步最优遗憾界》。载于 CarlosM. Carvalho 和Pradeep Ravikumar（编），*Proceedings of the Sixteenth International Conference on Artificial Intelligence and Statistics* 31 *Proceedings of*，第卷，第– 页，斯科茨 *Machine Learning Research* 99 107, 29 01 2013 代尔，亚利桑那州，美国，4 月–5 月。PMLR。网址 https://proceedings.mlr.press/v31/agrawal13a.html。

DuncanAlfers和Hermann Dinges。贝塔分布和伽马分布尾部概率的正态近似。 *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*，65:399–420，1984年。网址 https://link.springer.com/content/pdf/10.1007/BF00533744.pdf。

P. Auer、N. Cesa-Bianchi、Y. Freund和 R.E. Schapire。非随机多臂老虎机问题。 *SIAM Journal on Computing*，32(1):48–77，2002年。

Dorian Baudry、Patrick Saux和Odalric-Ambrym Maillard。从最优性到稳健性：随机老虎机中的自适应重采样策略。载于 M. Ranzato、A. Beygelzimer、Y. Dauphin、P.S. Liang 和J. Wortman Vaughan（编）， *Advances in Neural Information Processing Systems* 34, 14029 14041 2021 ，卷，页码 – 。Curran Associates, Inc.。网址https://proceedings.neurips.cc/paper_files/paper/2021/file/75429d136f65d2d6168b9b6c5f6ec。

Dorian Baudry、Kazuya Suzuki 和 Junya Honda。随机多臂老虎机算法分析的一般方法，2023年。

Christopher M Bishop 和 Nasser M Nasrabadi。 *Pattern recognition and machine learning*，第4卷。Springer，2006年。

David M Blei、Andrew Y Ng 和 Michael I Jordan . 潜在狄利克雷分配 *Journal of Machine Learning research*, 3(Jan):993–1022, 2003 , (1月) : – 。

Jonathan（Jon）Borwein 和 O-Yeat Chan。《互补不完全伽马函数的一致界》。*Mathematical Inequalities and Applications*，12，2007 年 01 月。doi：10.7153/ mia-12-10。

Michel Broniatowski . 用于散度最小化问题的加权自助法程序 . 在*Analytical Methods in Statistics: AMISTAT*、 *Prague*、 *November 2015*中，第1–22页。Springer，2017年。

A.N. Burnetas 和 M.N. Katehakis. 序贯分配问题的最优自适应策略. *Advances in Applied Mathematics*, 17(2):122–142, 1996.

奥利维耶·卡普 埃、奥雷 利安·加里维耶、Odalric-Ambrym Maillard、雷 米·穆诺和吉勒·斯托尔茨 . 最优序贯分配的 Kullback–Leibler 上置信界。 *Ann. Statist.*, 41 3 1516 1541, 2013年 0090 5364 101214 13 1119 ( ): – . ISSN - . doi: . / -AOS .

Olivier Chapelle 和 Lihong Li . 汤普森抽样的实证评估 . In J. Shawe- Taylor、R. Zemel、P. Bartlett、F. Pereira和 K.Q. Weinberger，编者，*Advances in Neural Information Processing Systems* 24 2011 , 第 卷。Curran Associates, Inc.， 。 URLhttps://proceedings.neurips.cc/paper_files/paper/2011/file/e53a0a2978c28872a4505bdb51db06d

Yunhi Cho 和 Eungchun Cho。被半空间截切的单纯形体积。*Applied mathematics letters* 14 6 731 735, 2001 , ( ): – 。

Peter Congdon . *Applied bayesian modelling*. John Wiley & Sons，2014年 .

I. Csiszar 和 F. Matus。信息投影再探。*IEEE Transactions on Information Theory* 49 6 1474 1490, 2003 10 1109 2003 810633 , ( ): – 。 doi: . /TIT. . .

I. Csiszár 和 P. C. Shields。信息论与统计学：教程。*Foundations and Trends® in Communications and Information Theory*，1(4):417–528，2004年。ISSN 1567-2190。doi: 10.1561/0100000004。网址 http://dx.doi.org/10.1561/0100000004。

Darinka Dentcheva。 *Optimization Models with Probabilistic Constraints*，第49–97页。Springer London，伦敦，2006年。ISBN978-1-84628-095-5。doi：10.1007/1-84628-095-8_2。网址 https://doi.org/10.1007/1-84628-095-8_2。 *Sections of simplices and cylinders. Volume. formulas and estimates* 2015 . 豪克·卡尔-埃尔温·迪尔克森 .

博士论文，. 网址 https://macau.uni-kiel.de/receive/diss_mods_00018308.

Lawrence C Evans 和 Ronald F Garzepy。 *Measure theory and fine properties of functions*。Routledge，2018年。Xiao Fang 和 Yuta Koike。通过 Stein 方法证明的高维中心极限定理。*The Annals of Applied Probability* 31 4 1660 1686, 2021 , ( ): –。

Thomas S Ferguson。对若干非参数问题的贝叶斯分析。*The annals of statistics* 209 230, 1973 , 页码–。Ayalvadi Ganesh 和 Neil O'Connell。Sanov 定理的逆定理。*Statistics & Probability Letters*，42(2):201–206，1999年。ISSN 0167-7152。doi: https://doi.org/10.1016/S0167-7152(98)00215-6。URL https://www.sciencedirect.com/science/article/pii/S0167715298002156。Ayalvadi J. Ganesh 和 Neil O'Connell。狄利克雷后验分布的大偏差原理。*Bernoulli*，6(6):1021 – 1034，2000年。

A. Garivier、P. M´enard 和 G. Stoltz。《先探索，再利用：老虎机问题中遗憾的真实形态》。 *Mathematics of Operations Research*，2018年。即将发表；同时参见 arXiv 预印本 arXiv:1602.07182。Aur´elien Garivier 和 Olivier Capp´e。《用于有界随机老虎机及更广泛场景的 kl-ucb 算法》。载于 *Proceedings of the 24th annual conference on learning theory* 359 376, , 2011

第 – 页。Aur´elien Garivier、H´edi Hadiji、Pierre M´enard 和 Gilles Stoltz。《Kl-ucb-switch：从分布相关和分布无关两个角度看随机老虎机的最优遗憾界》。 *Journal of Machine Learning Research*，23(179):1–66，2022。网址：http://jmlr.org/papers/v23/20-717.html。Subhashis Ghosal 和 Aad Van der Vaart。《 *Fundamentals of nonparametric Bayesian inference*》，第44卷。剑桥大学出版社，2017年。Friedrich G¨otze、Alexey Naumov、Vladimir Spokoiny 和 Vladimir Ulyanov。《大球概率、高斯比较与反集中》。 *Bernoulli*，25(4A):2538 – 2563，2019。doi: 10.3150/18-BEJ1062。网址：https://doi.org/10.3150/18-BEJ1062。Senlin Guo、Feng Qi 和 Hari Srivastava。《两类函数对数完全单调的充要条件》。 *Integral Transforms and Special Functions* 18 819 826, 11 2007 10 1080 10652460701528933 , :

–。doi: . / . Rafiq Hijazi 和 Robert Jernigan。《使用狄利克雷回归模型对组成数据建模》。 *Journal of Applied Probability and Statistics*，4:77–91，2009年1月。Junya Honda 和 Akimichi Takemura。《一种适用于有界支撑集模型的渐近最优老虎机算法》。载于 Adam Tauman Kalai 和 Mehryar Mohri 编，*COLT*，第67–79页。Omnipress，2010年。ISBN 978-0-9822529-2-5。网址：http://dblp.uni-trier.de/db/conf/colt/colt2010.html#HondaT10。Junya Honda 和 Akimichi Takemura。《一种适用于半有界奖励的新老虎机算法的非渐近分析》。 *Journal of Machine Learning Research*，16(113):3721–3756，2015年。网址：http://jmlr.org/papers/v16/honda15a.html。N. Korda、E. Kaufmann 和 R. Munos。《用于一维指数族老虎机的汤普森抽样》。载于 *Proceedings of the 26th International Conference on Neural Information Processing Systems* 13, 1448 1456, 2013 ,

NIPS' 第 – 页。T.L. Lai 和 H. Robbins。《渐近有效的自适应分配规则》。 *Advances in Applied Mathematics* 6 4 22, 1985 , :

–。K Lange。《狄利克雷分布在法证匹配概率中的应用》。 *Genetica*，96 1 2 107 117，1995年（-）：–。Jean-Bernard Lasserre。《单纯形上多项式积分的简明公式》。 *BIT Numerical Mathematics* 61 08 2020 10 1007 10543 020 00828 , ,。

doi: . /s - - -x。Olivier Marchal 和 Julyan Arbel。《贝塔分布和狄利克雷分布的次高斯性》。 *Electronic Communications in Probability*，22(论文编号54):1–14，2017年。Kevin P. Murphy。《 *Probabilistic Machine Learning: An introduction*》。MIT Press，2022。网址：probml.ai。K.W. Ng、G.L. Tian 和 M.L. Tang。《 *Dirichlet and Related Distributions: Theory Methods and Applications* 》，。2011 9781119998419 概率与统计威利丛书。Wiley，。ISBN 。网址：https://books.google.ru/books?id=k8GS868oyo4C。

Frank W, 马萨诸塞州韦尔斯利, 1997年。ISBN 1-56881-069-5。1979年首版重印[Peters, Ltd.
版社；MR0435697（55 #8655）]。

IanOsband和 Benjamin Van Roy。为什么后验抽样比乐观主义更适用于强化学习？载于 Doina Precup和 Yee WhyeTeh（编），*Proceedings of the 34th International Conference on Machine Learning*，第 70 卷，*Proceedings of Machine Learning Research*，2701–2710，06–11 2017年。PMLR，–月。网址 https://proceedings.mlr.press/v70/osband17a.html。

Ian Osband、Daniel Russo 和 Benjamin Van Roy。通过后验抽样实现（更）高效的强化学习。载于 C.J. Burges、L. Bottou、M. Welling、Z. Ghahramani 和 K.Q. Weinberger编，*Advances in Neural Information Processing Systems* 26，2013，第 卷。Curran Associates, Inc.，。网址https://proceedings.neurips.cc/paper_files/paper/2013/file/6a5889bb0190d0211a991f47bb19a77

CharlesRiou 和 Junya Honda。基于汤普森抽样的有界奖励分布多臂老虎机算法。在Aryeh Kontorovich和 Gergely Neu，编者，*Proceedings of the 31st International Conference on Algorithmic Learning Theory*，117 *Proceedings of Machine Learning Research*，卷，777–826，08–11 2020年，页码 –。PMLR，2月– 2月。网址 https://proceedings.mlr.press/v117/riou20a.html。Donald B Rubin。贝叶斯自助法。*The annals of statistics*，页码 130–134，1981年。Ivan Nicolaevich Sanov。关于随机变量发生大偏差的概率。*Selected Translations in Mathematical Statistics and Probability* 1 213–244，1961，：–。

Ian F. Somerville、Duane L. Dietrich 和Thomas A. Mazzuchi。使用狄利克雷先验分布进行贝叶斯可靠性分析，重点研究按随机顺序运行的加速寿命试验。*Nonlinear Analysis: Theory, Methods & Applications*，30(7):4415–4423，1997年。ISSN 0362-546X。doi: https://doi.org/10.1016/S0362-546X(96)00120-4。网址 https://www.sciencedirect.com/science/article/pii/S0362546X96001204。第二届世界非线性分析学家大会论文集。Mohammad Sadegh Talebi和 Odalric-AmbrymMaillard。马尔可夫决策过程中的无折扣强化学习的方差感知遗憾界。在 *Algorithmic Learning Theory*，页码 770–805，2018年。Yee Whye Teh、Michael I Jordan、Matthew J Beal 和David M Blei。分层狄利克雷过程。*Journal of the American Statistical Association*，101(476):1566–1581，2006年。doi: 10.1198/016214506000000302。网址https://doi.org/10.1198/016214506000000302。William R Thompson。根据两个样本的证据，一个未知概率超过另一个未知概率的可能性。*Biometrika*，25(3-4):285–294，12月1933年。ISSN0006-3444。doi: 10.1093/biomet/25.3-4.285。网址 https://doi.org/10.1093/biomet/25.3-4.285。Daniil Tiapkin、Denis Belomestny、Eric Moulines、Alexey Naumov、Sergey Samsonov、Yunhao Tang、Michal Valko 和 PierreMenard。从狄利克雷到 Rubin：强化学习中无奖励项的乐观探索。在 Kamalika Chaudhuri、Stefanie Jegelka、Le Song、Csaba Szepesvari、GangNiu 和 Sivan Sabato，编者，*Proceedings of the 39th International Conference on Machine Learning*，162 *Proceedings of Machine Learning Research*，卷，21380–21431，17–23 2022年，页码 –。PMLR，–7月。网址 https://proceedings.mlr.press/v162/tiapkin22a.html。Roman Vershynin。*High-Dimensional Probability: An Introduction with Applications in Data Science*。剑桥统计与概率数学丛书。剑桥大学出版社，2018年。doi: 10.1017/9781108231596。。

Andre M Zubkov 和 Aleksandr A Serov。二项分布分布函数普适不等式的完整证明。*Theory of Probability & Its Applications*，57(3)：539–544，2013年。