arXiv:2611.28811v1 [cond-mat.stat-mech] 21 Nov 2026

有限总体格兰诺维特阈值模型中的大偏差与临界标度

José F. Fontanari[a]

[a]圣卡洛斯物理研究所，圣保罗大学，*13566-590* 圣卡洛斯，圣保罗，巴西

摘要

我们对 Granovetter 阈值模型的精确有限总体解进行了系统的渐近分析。在该模型中，单个初始发起者会在采用阈值独立分布的个体中触发级联。通过将传播概率表示为一个次序统计量越过边界的事件，我们将其宏观大偏差率表述为一个受约束的相对熵最小化问题。对于 Beta 分布的阈值，我们获得了所有形状参数 α 和 β下的显式解，并确定了四种可能的最优经验阈值剖面：完全自由、完全钉扎、钉扎–自由和自由–钉扎。这些剖面决定了宏观级联的指数级概率，以及导致其发生的阈值构型。重要性抽样估计证实了解析速率函数，而序贯蒙特卡洛模拟则独立地再现了预测的最优剖面。我们还重新考察了 $\alpha_c = 1$ 附近的临界标度，并确认了主要的有限尺寸标度律：对于均匀阈值，平均活跃比例按 $N^{-1/2}$衰减；对于 $\alpha = 1$和 $0 < \beta < 1$，则按 N^{-1}衰减。在后一种状态下，我们进一步从微观分枝过程极限推导出了对数临界窗口的显式标度函数。我们的结果为确定阈值驱动的级联过程中罕见集体事件的概率及其特征构型提供了一个通用框架。

1. 引言

集体行动通常始于少数个体，而他们的决定会鼓励其他人跟随 [1, 2]。抗议活动可能会不断扩大，因为每一位新参与者都会说服更多人加入 [1, 3]；新技术的传播可能会随着其采用情况日益显著而加快 [4, 5]；金融恐慌则可能在少数投资者的行为促使其他人撤出资金时形成 [6, 7]。所有这些例子中的一个核心问题是：最初的小规模扰动会消失，还是会发展成为大规模集体事件 [8, 9, 10]。

阈值模型为研究这一问题提供了一个简单的框架。在格兰诺维特的模型中 [1], , 每个个体都有一个阈值，用来表示其
之前必须已经参与的总体最低比例

该个体愿意加入之前，总体中必须已有的最低参与比例。这些阈值因个体而异，并取自共同概率分布。随着参与度提高，可能会有更多阈值被超过，从而使级联得以在总体中传播。

对于规模非常大的总体，活跃比例的演化通常由涉及累积阈值分布 $F(x)$ [11] 的确定性方程描述。这类方程有助于识别稳定集体状态，以及宏观级联可能发生的条件。然而，它们无法完整描述有限总体中的情况。即使确定性描述预测微小扰动无法扩展，个体阈值的某种特定实现也可能特别有利于传播。此时，由单个个体发起的级联可能招募总体中相当大一部分成员，尽管此类事件发生的概率可能极低。

理解这些异常级联正是本研究的主题。我们不再询问宏观级联是否常见，而是探究当总体规模 N 变大时，这种情况发生的概率有多低。更确切地说，我们要研究单个永久煽动者引发涉及总体中有限比例 x 成员的级联的概率，即使这一结果需要阈值出现非典型实现才能实现。大偏差理论适用于描述此类概率；它刻画的是发生概率随系统规模 [12]呈指数级下降的事件。

近期，人们得到了该问题的一个精确有限总体表述 [13]。其核心观察是，成功传播会对被招募个体的阈值施加一系列条件。若将这些阈值按递增顺序排列为 $X_{(1)} \le X_{(2)} \le \cdots$，那么除初始发起者外，级联只有在以下条件成立时才能招募 $k-1$ 个体：

$$X_{(j)} \le \frac{j}{N}, \qquad j = 1, \ldots, k-1. \tag{1}$$

这些不等式的含义很直观。第一个招募对象必须愿意在仅有一个个体处于活跃状态时加入；第二个必须愿意在有两个个体处于活跃状态时加入；依此类推。只要其中任一条件不成立，级联就无法通过这组拟招募对象继续传播。

这种序列要求是该问题的主要难点。总体中有许多阈值相对较低的个体还不够；这些阈值还必须以适当的方式分布，使级联能够逐步推进并经过每一个中间阶段。因此，要估计大规模级联发生的概率，不能只统计阈值低于最终激活比例的个体数量。

对于均匀分布的阈值，满足公式（1）的概率与经典的停车函数问题密切相关 [14, 15, 16], ，而该问题已有精确的计数公式。对于一般的阈值分布，相应概率的计算要困难得多。特别是，传统的鞍点近似可能会失效，因为产生大规模级联的最可能阈值配置未必位于允许区域的

内部。这一困难在最近一篇关于有限总体表述的评论中得到了强调。

当我们考察参与罕见级联的个体所具有的阈值分布时，这一困难的物理根源便更加清晰。假设级联波及总体的比例为 $x = k/N$ 。在大-N 极限下，公式（1）中的序贯条件要求招募阈值的经验累积分布 $G(t)$满足

$$G(t) \geq \frac{t}{x}, \qquad 0 \leq t \leq x. \tag{2}$$

因此，经验分布不能低于一条直线边界。否则，在某个中间阶段，愿意参与的个体数量就会过少，传播便会停止。

我们方法的核心思想是找出能够维持宏观级联的最可能阈值配置。我们将这一问题表述为约束相对熵极小化，由此确定传播概率的主导指数衰减率，以及与罕见级联相关的典型阈值配置。重要的是，必须将连续变分问题与精确的有限总体约束联系起来。我们通过概率测度变换、分位数耦合以及均匀停车函数的精确计数，建立了这种联系。

我们将这一框架应用于贝塔阈值分布的双参数族 Beta(α，β)，该分布族涵盖递减、递增、钟形和 U 形密度。对于所有 α，$\beta > 0$，我们都得到了宏观传播率的显式表达式，并确定了相应的最优阈值剖面。我们使用重要性抽样和序贯蒙特卡洛模拟对分析预测进行了数值验证；两种方法均独立地重现了预测的阈值构型。

我们还重新考察了 $\alpha_c = 1$ 附近的临界标度；此前 Ref. [13] 曾对其进行研究。我们确认了平均最终活跃比例的主要有限尺寸标度律，并推导出 $0 < \beta < 1$ 时对数临界窗口的显式标度函数。这一结果将有限尺寸行为与微观分枝过程描述联系起来，并补充了宏观大偏差分析。

本文结构安排如下。第 2 节介绍精确的有限总体表述，第 3 节阐述约束熵方法并证明其渐近有效性。第4节推导贝塔分布族的传播率和最优剖面，第 5 节对其进行数值验证。第 6 节探讨临界有限尺寸标度，第 7 节总结全文。

2. 精确有限总体传播事件

我们考虑一个由 N 个个体组成的总体，其中一人被指定为永久发起者。该个体无条件参与并发起

级联。其余 $N-1$ 个体具有独立的激活阈值 X_1，……， X_{N-1}，这些阈值取自连续累积分布 $F(t)$，其概率密度为 $f(t)$，支持集为 $[0,1]$。当已激活个体所占比例达到或超过某个未激活个体的阈值时，该个体便会被激活。个体一旦被激活便保持激活状态，因此级联会不可逆地持续进行，直到无法再招募更多个体为止 [1]。

我们的第一个目标是确定级联恰好以包含初始发起者在内的 k 个活跃个体停止的概率 $P_N(k)$。如参考文献 [13], 所示，该概率可以根据最终参与的个体的阈值精确表示。

为了了解其运作方式，假设级联以 k 个活跃个体终止。除发起者外，必须有一个由 $k-1$ 个体组成的群体，其阈值允许它们变为活跃状态。这个群体有 $\binom{N-1}{k-1}$ 种可能的选择。

考虑这样一个群体，并将其阈值按递增顺序排列，
$X_{(1)} \le X_{(2)} \le \cdots \le X_{(k-1)}$。最初只有发起者处于活跃状态，因此第一个招募对象的阈值必须不超过 $1/N$。该个体加入后，活跃比例变为 $2/N$，因而如果第二个招募对象的阈值不超过 $2/N$，它也可以参与。此后的每个阶段都适用相同的推理。

因此，当且仅当该群体的有序阈值满足 $X_{(j)} \le j/N$对每个 $j = 1$、……、$k-1$均成立时，整个群体才能被完全招募，如公式（1）所述。这些不等式体现了级联的序贯性质：在每个阶段，至少必须有一个额外的个体愿意加入；否则，招募就会停止。

我们将一个指定的、包含 $k-1$ 个体的群体满足所有这些条件的概率记为

$$Q_{N,k} = \mathbb{P}\left\{ X_{(j)} \le \frac{j}{N}, \quad j = 1, \ldots, k-1 \right\}, \tag{3}$$

其中$Q_{N、1} = 1$。我们将$Q_{N、k}$ 称为传播概率。

选定群体的成功招募并不足以使级联在规模 k处停止。我们还必须要求其余 $N-k$个体中没有任何一个变为活跃。由于最终激活比例为 k/N，这恰好发生在它们的阈值全部超过 k/N时。该事件的概率为$[1-F(k/N)]^{N-k}$。

由于不同个体的阈值彼此独立，招募和停止的概率可以相乘。此外，两个不同的、包含$k-1$个招募对象的群体不可能同时代表最终的活跃总体：某一群体之外的个体的阈值大于 k/N ，因此不可能属于另一个成功招募的群体。

对招募对象的各种可能选择求和，可得精确的级联规模分布 [13],

$$P_N(k) = \binom{N-1}{k-1} Q_{N,k} \left[1 - F\left(\frac{k}{N} \right) \right]^{N-k}, \tag{4}$$

对于 $k = 1$、.....、 N。与通常情况一样，当 $k = N$ （ $[1 - F(1)]^0 \equiv 1$）时，与不存在非活跃代理相关的因子视为 1，因此$P_N(N) = Q_{N,\ N}$。

方程 (4) 将大小为 k 的级联概率分解为三项：可能被招募群体的数量、所选群体能够被激活的概率，以及该群体激活后招募停止的概率。唯一非平凡的量是传播因子$Q_{N,\ k}$，因为它涉及有序随机阈值的多个同时条件。

2.1. 传播概率的精确递推

对于任意有限的N [13]，都可以通过递归计算传播概率$Q_{N,\ k}$。为得到递推式，考虑所有 $k-1$ 个阈值均低于 $(k-1)/N$这一事件，并减去招募在更早的大小 $j < k$时停止的互斥事件。由此得到

$$Q_{N,k} = F\left(\frac{k-1}{N}\right)^{k-1} - \sum_{j=1}^{k-2} \binom{k-1}{j-1} Q_{N,j} \left[F\left(\frac{k-1}{N}\right) - F\left(\frac{j}{N}\right)\right]^{k-j}, \quad (5)$$

对于 $k \geq 2$，其中$Q_{N,\ 1} = 1$。该递推与参考文献 [13], 中推导的归一化恒等式等价，但避免了从 1 中减去较小的概率，因此在 N适中时数值上更优。不过，对于宏观级联，$Q_{N,\ k}$ 可能比递推式中的各项小得多，从而导致严重的消去误差。此外，该递推无法直接揭示大偏差率，这促使我们在下文中发展渐近方法。

2.2. 区分传播的两个要求

现在我们考虑宏观级联，即随着总体规模的增长，最终激活比例趋近于固定值 x:

$$\frac{k}{N} \longrightarrow x, \qquad 0 < x < 1. \quad (6)$$

我们的目标是确定传播概率随N降低的速率。

第一步是区分成功招募的两个不同要求。首先，所有 $k-1$ 个被招募个体的阈值都必须足够低。其次，这些阈值必须以适当的方式分布，才能使招募得以持续进行而不中断。

为明确区分这两个要求，定义

$$n = k - 1, \qquad a = \frac{n}{N}. \quad (7)$$

这里， n 表示被招募个体的数量，不包括永久煽动者；而a 是与其全部招募相容的最大阈值。

公式（1）中的最后一个不等式要求，最大的招募阈值至多为 a。因此，所有 n 阈值都必须位于区间 $(0, a)$ 内。由于这些阈值相互独立，它们全部落在该区间内的概率就是 $F(a)^n$。

现在假设我们已经知道所有 n 阈值都属于 $(0, a)$。其条件概率密度为

$$q_a(t) = \frac{f(t)}{F(a)} \mathbf{1}_{(0,a)}(t). \tag{8}$$

换句话说，q_a 是限制在 $(0, a)$ 上并归一化为单位概率的原始阈值分布。因此，传播概率可以精确写成

$$Q_{N,k} = F(a)^n R_{N,k}, \tag{9}$$

其中 $R_{N,\,k}$ 是这样的条件概率：现在独立地从 q_a 中抽取的 n阈值满足序贯招募条件。由于$a = n/N$，这些条件可以写成

$$E_n = \left\{ X_{(j)} \le \frac{aj}{n}, \quad j = 1, \dots, n \right\}, \tag{10}$$

因此

$$R_{N,k} = \mathbb{P}_{q_a}(E_n). \tag{11}$$

请注意，$R_{N,\,k}$ 由最终级联规模 k标记，而 E_n 只涉及 $n = k - 1$ 个已招募的个体。因此，这两个量从不同角度描述了同一传播过程。

因式分解（9）很重要，因为它将足够低阈值的总体稀缺性（由 $F(a)^n$ 表示）与以能够维持级联的方式排列这些阈值的额外困难（由 $R_{N,\,k}$ 表示）区分开来。正如我们将看到的，第二个因子可能与第一个同样重要。特别是，即使在将所有阈值限制在 a 以下之后，成功传播仍可能极不可能，并且其概率呈指数级小。

2.3. 均匀情形与停车函数

当条件密度 q_a 在 $(0, a)$ 上均匀时，缩放阈值$U_i = X_i/a$ 相互独立，且在 $[0, 1]$ 上均匀。传播事件 E_n 变为

$$U_{(j)} \le \frac{j}{n}, \qquad j = 1, \dots, n. \tag{12}$$

这些不等式等价于经典停车函数条件 [14,15, 16]。事实上，令 $p_i = \lceil nU_i \rceil$，可将该事件映射为具有偏好 $p_i \in \{1, \dots, n\}$ 的 n 辆汽车成功停车。由于 n^n 个偏好序列中恰有 $(n + 1)^{n-1}$ 个能够成功停车，我们得到

$$\mathbb{P}_{\text{unif}}(E_n) = \frac{(n+1)^{n-1}}{n^n} \sim \frac{e}{n}. \tag{13}$$

因此，在均匀抽样下，序贯传播约束只会引入代数级而非指数级的概率代价。

对于原本为均匀阈值分布的情形，$F(t) = t$，因此因式分解式 (9) 给出了精确结果

$$Q_{N,k} = \frac{k^{k-2}}{N^{k-1}}, \tag{14}$$

这与参考文献 [13] 一致。

然而，对于非均匀阈值分布，条件传播概率 $R_{N,\,k}$ 可能随 N 指数衰减。确定这一额外的指数成本，就是下文要讨论的大偏差问题。

3. 约束熵与有限样本速率判据

在前一节中，我们将传播概率分解为两个因子，$Q_{N,\,k} = F(a)^n R_{N,\,k}$，其中 $n = k-1$ 且 $a = n/N$。第一个因子是所有被招募个体的阈值均低于 a 的概率。第二个因子则是这些阈值的分布足够均匀，能够在每个阶段维持招募的概率。我们关注的是宏观级联，在这种情况下，$k/N \to x \in (0,\,1)$ 当 $N \to \infty$。由于 $n/N \to x$ 且 $a \to x$，第一个因子对指数衰减率的贡献可直接得出：

$$\lim_{N\to\infty} \frac{1}{N} \log F(a)^n = x \log F(x). \tag{15}$$

因此，主要困难在于确定$R_{N,k}$的大-N 行为。这是统计物理中的一个一般性问题：我们如何计算许多随机变量同时满足大量约束的概率，尤其是在所要求的事件为指数级不可能发生的情况下?

我们分两步处理这一问题。首先，通过最小化相对熵代价，确定满足这些约束的最可能配置。随后，我们展示如何通过适当的概率测度变换，将这一连续体预测与精确的有限样本问题联系起来。

3.1. 稀有事件的最可能实现

考虑从式（8）中引入的条件密度 $q_a(t)$抽取的 n 个独立阈值。对于较大的 n，它们的经验分布通常会接近 q_a。这是大数定律的结果：属于任意给定区间的抽样阈值比例，会趋近于 q_a赋予该区间的概率。

然而，稀有事件可能要求经验分布与通常观测到的分布大不相同。大偏差理论的核心思想是，这类偏离并非都同样不可能。在与稀有事件相容的各种分布中，所需统计偏离最小的分布决定了其主导指数级概率。

为了定量表述这一点，设 $g(t)$是定义在 $[0,x]$上的归一化概率密度，并将其累积分布记为

$$G(t)=\int_0^t g(s),ds, \tag{16}$$

其中 $G(0)=0$ 和 $G(x)=1$。

$G(t)$ 的含义很直观：它表示阈值不超过 t 的被招募个体所占的比例。由于被招募对象占整个总体的比例趋近于 x，在宏观极限下，这些个体在总体中所占的相应比例为 $xG(t)$。单个永久发起者的贡献为 $1/N$ 量级，在此极限下趋于零。

要使招募得以继续，在每个中间阶段，总体中都必须有足够多阈值较低的个体。尤其是，阈值不超过 t 的个体比例 $xG(t)$ 必须至少为t。否则，级联就无法达到活跃比例 t，继而无法招募阈值更高的个体。由此得到的连续传播约束为

$$G(t)\geq\frac{t}{x},\qquad 0\leq t\leq x. \tag{17}$$

考察这一条件如何从精确的有限种群要求中得出，是很有启发性的。回顾式（10），成功招募 $n=k-1$ 个体需要满足

$$X_{(j)}\leq\frac{aj}{n},\qquad j=1,\dots,n, \tag{18}$$

其中 $a=n/N$。引入招募阈值的经验累积分布，

$$\widehat{G}_n(t)=\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{\{X_i\leq t\}}. \tag{19}$$

顺序统计量不等式与下式完全等价

$$\widehat{G}_n\left(\frac{aj}{n}\right)\geq\frac{j}{n},\qquad j=1,\dots,n. \tag{20}$$

因此，在每个 n 离散招募阶段，经验累积分布都必须达到或超过对角线上对应的点。与 $n/N\to x$ 中的 $N\to\infty$ 一样，这些网格点在 $[0,x]$ 中变得稠密，从而得到连续统约束（17）。

从几何上看，这一约束要求累积分布 $G(t)$ 位于连接 $(0,\ 0)$ 与 $(x,\ 1)$ 的直线之上或直线上。该直线本身对应于均匀密度 $g(t)=1/x$，此时，阈值足够低的个体供给恰好达到连续体约束所要求的最低数量。严格位于该直线之上的分布，表示在相应的 t取值下，低阈值个体数量有余。相反，如果

$G(t^*) < t^*/x$ 在某个中间点 $t^* \in (0, x)$，阈值低于 t^* 的个体太少，无法维持招募直到该点。因此，级联必须在达到预期的宏观规模之前停止。边界 $G(t) = t/x$ 因而是临界传播条件。正如我们将要看到的，与成功传播相容的最可能分布，可能在某些区间与这一边界重合，而在其他区间则严格高于边界。

我们用 $\mathcal{A}_x$ 表示满足连续传播约束的归一化密度集合：

$$\mathcal{A}_x = \left\{ g \geq 0 : \int_0^x g(t)\,dt = 1, \quad \int_0^t g(s)\,ds \geq \frac{t}{x} \text{ for all } t \in [0,x] \right\}. \tag{21}$$

从 q_x 中抽取的独立样本，其经验分布接近指定剖面 g 的概率有多低?

Sanov定理给出了指数尺度上的答案 [18, 12]。当 n 很大时，观测到经验分布接近 g 的概率，其主导指数行为为

$$\mathbb{P}\{\text{empirical distribution near } g\} \asymp \exp[-nD_{\mathrm{KL}}(g\|q_x)], \tag{22}$$

其中

$$D_{\mathrm{KL}}(g\|q_x) = \int_0^x g(t) \log \frac{g(t)}{q_x(t)}\,dt \tag{23}$$

是 Kullback–Leibler散度，也称为相对熵。这里的 $\asymp$ 表示在主导指数尺度上的相等，而不指定代数前因子。相对熵衡量的是将原始分布 q_x 替换为非典型分布 g 的统计代价。它非负，且仅当两个分布几乎处处相同时才为零。因此，典型分布的代价为零，而非典型分布的代价为正。

对于可以由许多不同经验剖面实现的事件，主导贡献来自相对熵最小的可容许剖面。这是大偏差理论中与统计力学熟知的鞍点原理相对应的原理。因此，我们得到以下约束最小化问题

$$\mathcal{I}_F(x) = \inf_{g \in \mathcal{A}_x} D_{\mathrm{KL}}(g\|q_x). \tag{24}$$

记为 g_x^* 的最小化密度描述了在能够维持宏观级联的各个实现中，最可能出现的阈值分布。连续统论证表明

$$R_{N,k} \asymp \exp[-n\mathcal{I}_F(x)]. \tag{25}$$

由于 $n/N \to x$，将该表达式与因子 $F(a)^n$ 相结合，可得到候选传播率

$$\psi_Q(x) = x \log F(x) - x\mathcal{I}_F(x). \tag{26}$$

这个公式还有另一种有用的解释。由于在 $[0,x]$ 上，$q_x(t) = f(t)/F(x)$，我们有

$$D_{\mathrm{KL}}(g\|q_x) = \int_0^x g(t)\log\frac{g(t)}{f(t)}\,dt + \log F(x). \tag{27}$$

因此，

$$\psi_Q(x) = -x\inf_{g\in\mathcal{A}_x}\int_0^x g(t)\log\frac{g(t)}{f(t)}\,dt. \tag{28}$$

由于每个可容许密度 g 都在 $[0,x]$ 上归一化，并且在该区间之外为零，因此该积分恰好是它相对于原始阈值密度 f的相对熵。

因此，传播率取决于满足传播障碍（公式（17））的原始阈值分布所能进行的代价最低的统计重排。极小化剖面不一定与原始分布相似。特别是，在某个有限区间上，其累积分布可能与传播边界重合，这对应于约束达到饱和的区域。此类边界接触将在后续章节中发挥重要作用。

3.2. 为什么连续统论证并不充分

分

熵最小化为指数衰减率提供了自然的预测。然而，这并不能自动证明精确的有限总体概率具有这一速率。原因在于，原始事件 E_n 比连续统描述所暗示的限制更严格。它要求 n 个有序阈值中的每一个都满足相应的不等式，哪怕是微小的违反也不允许。三个困难尤其值得关注。

首先，标准的大偏差论证描述的是经验分布集合的概率。允许约束出现微小违反的宽松事件，其下界未必与精确事件相同。其次，最优累积分布可能在整个区间内与边界 $G(t) = t/x$重合。在这样的钉扎区间上，普通的样本波动会出现在边界两侧。要求经验分布在每一点都保持在允许的一侧，可能会带来额外的概率代价。第三，最初几步招募只涉及少数个体。无论总体变得多大，这些微观约束始终存在。仅凭连续统描述，并不能保证招募能够通过这些初始瓶颈。

关键问题在于，这些有限样本限制只会改变次指数因子，还是也会改变主导的指数衰减率。我们将证明，对于此处考虑的贝塔分布，这些限制不会改变熵最小化所预测的指数速率。证明结合了概率测度变换（与统计力学中的重要性抽样密切相关）以及精确的停车函数恒等式 (13)。

3.3. 改变抽样分布

估计稀有概率的难点通常在于生成样本所依据的分布。在条件密度 q_a 下，成功传播可能极不可能发生，且其概率会随规模呈指数级下降。我们可以改为从另一种分布中生成阈值，在该分布下成功传播更为常见，并对抽样概率的变化进行校正。这就是重要性抽样 [19]的原理。统计力学中的伞式抽样和直方图重加权也运用了密切相关的思想 [20, 21]。

令 $h_a(t)$ 为定义在 $(0,\ a)$ 上的辅助概率密度，其累积分布为

$$H_a(t) = \int_0^t h_a(s)\, ds. \tag{29}$$

最终，我们将选择 h_a 为连续问题中得到的熵最小化密度 g_a^*，并在有限端点 $a = n/N$ 处取值。为使测度变换有效，我们要求只要 $q_a(t) > 0$，就有 $h_a(t) > 0$，但零测集除外。定义对数似然比

$$\ell_a(t) = \log \frac{q_a(t)}{h_a(t)}. \tag{30}$$

由 h_a 生成的一组阈值配置 $Y_1, \ldots, Y_n$ 获得的统计权重为

$$W_n = \prod_{i=1}^{n} \frac{q_a(Y_i)}{h_a(Y_i)} = \exp\left[\sum_{i=1}^{n} \ell_a(Y_i)\right]. \tag{31}$$

该权重恰好补偿了使用辅助分布所带来的影响。因此，条件传播概率可表示为

$$R_{N,k} = \mathbb{E}_{h_a}\left[\mathbf{1}_{E_n} \exp\left(\sum_{i=1}^{n} \ell_a(Y_i)\right)\right]. \tag{32}$$

这里，$Y_1, \ldots, Y_n$ 是来自辅助密度 h_a 的独立样本。重要的是，传播事件 E_n 与原始密度 q_a 下的事件完全相同：

$$E_n = \left\{Y_{(j)} \le \frac{aj}{n}, \quad j = 1, \ldots, n\right\}. \tag{33}$$

因此，当所有这些不等式都成立时，指示函数 $\mathbf{1}_{E_n}$ 等于一；否则等于零。测度变换只会改变分配给每种阈值配置的概率，而不会改变定义成功传播的条件。似然权重恰好补偿了这一变化。方程（32）对每个有限的 N都精确成立。这里没有进行任何大N 近似。

这种表示法的优势取决于是否选择了合适的辅助密度。我们希望在h_a下成功传播足够常见，同时将似然权重控制在适当范围内，以确定其指数贡献。对于第4节构造的最优 Beta 提议分布，以下四个性质可以实现这一点。

性质*1*：障碍支配。

我们要求

$$H_a(t) \geq \frac{t}{a}, \qquad 0 \leq t \leq a. \tag{34}$$

这意味着，对于每个阈值水平，提议分布在该水平以下所分配的概率至少与 (0，a) 上的均匀分布一样多。

性质 *2*：单调似然权重。

我们要求 $\ell_a(t)$ 为非递减函数，并且其网格平均值收敛于其连续平均值：

$$\frac{1}{n}\sum_{j=1}^{n} \ell_a\left(\frac{aj}{n}\right) = \frac{1}{a}\int_0^a \ell_a(t)\,dt + o(1). \tag{35}$$

单调性将使传播不等式能够给出总似然权重的直接上界。

性质 *3*：空间平均与统计平均相等。

对于所关注的提议分布，

$$\frac{1}{a}\int_0^a \ell_a(t)\,dt = \int_0^a h_a(t)\ell_a(t)\,dt =: m_a. \tag{36}$$

这种等式是最优提议分布结构所特有的；对于任意测度变换，该等式并不成立。第二个积分是在 h_a 下对数权重的平均值。根据相对熵的定义，

$$m_a = -D_{\mathrm{KL}}(h_a \| q_a). \tag{37}$$

因此，对于最优提议分布， m_a 是满足传播约束所需的最小熵成本的负值。我们还要求 $m_a \to m_x$ ，其中 $N \to \infty$，且 $k/N \to x$。

性质 *4*：指数集中。

在从 h_a 独立抽样的情况下，对数权重的平均值预期将趋近于 m_a。我们要求固定幅度的向下偏差的概率按指数速度递减：

$$\mathbb{P}_{h_a}\left\{\frac{1}{n}\sum_{i=1}^{n} \ell_a(Y_i) < m_a - \varepsilon\right\} \leq e^{-c_\varepsilon n}, \tag{38}$$

对于每个 $\varepsilon > 0$，存在某个与 N 无关的 $c_\varepsilon > 0$ ，只要 N 足够大且 $k/N \to x \in$ (0,1)。这一条件强于通常的大数定律。它确保即使与概率随 n 按代数速度递减的事件相比，异常小的似然权重也可以忽略不计。

对于下一节构造的最优 Beta 提议分布，性质 1 和 3 可直接从其结构得出。性质1成立，因为熵极小化分布属于可容许集合 $\mathcal{A}_a$。为理解性质 3，请注意，

最优提议分布由自由区间（其中 $h_a = q_a$，因此 $\ell_a = 0$）和钉扎区间（其中 $h_a = 1/a$）组成。在这两种情况下，$h_a(t)\ell_a(t) = \ell_a(t)/a$，由此立即得到式 (36)。性质 2 和 4 需要额外的估计，尤其是在 Beta 密度具有端点奇异性时。其验证推迟至附录 A。

3.4. 停车函数与匹配的指数界

现在我们说明这四个性质如何确定指数传播率。论证包含两个要点：提议分布下成功传播概率的下界，以及相应似然权重的上界和下界。

提议分布下的成功传播..... 性质1意味着，可以将一个取自 h_a 的随机变量与一个均匀随机变量耦合，使得前者始终不大于后者的 a 倍。为说明这一点，令 U 在 $[0,1]$ 上均匀分布，并生成

$$Y = H_a^{-1}(U). \tag{39}$$

由于 $H_a(t) \geq t/a$，相应的分位数满足

$$H_a^{-1}(u) \leq au, \qquad 0 < u < 1. \tag{40}$$

将这一构造应用于 n 个独立的均匀随机变量，得到

$$Y_{(j)} \leq aU_{(j)}, \qquad j = 1, \ldots, n. \tag{41}$$

因此，只要这些均匀随机变量满足停车函数条件 $U_{(j)} \leq j/n$，提议变量就自动满足 $Y_{(j)} \leq aj/n$。于是，停车函数恒等式(13)给出了以下界

$$\mathbb{P}_{h_a}(E_n) \geq \frac{(n+1)^{n-1}}{n^n} \sim \frac{e}{n}. \tag{42}$$

这是决定性的观察。在辅助分布下，成功传播并非指数级罕见。其概率下界为一个仅随 n代数级递减的量。

剩下的任务是确定公式（32）中似然权重的指数贡献。为了用相匹配的渐近界夹住$R_{N,\,k}$ 我们必须排除两种相反的可能性：成功传播事件携带异常大的权重（上界），或携带异常小的权重（下界）。

上界。 为理解上界，回顾性质 2：对数似然比 $\ell_a(t) = \log[q_a(t)/h_a(t)]$ 是递增的。直观地说，传播约束会迫使有序阈值趋向更小的值，而在这些值处似然比不会更大。因此，满足传播屏障不会使似然权重超过将每个有序阈值置于其允许的最大值 $Y_{(j)} = aj/n$ 时得到的数值。

形式化地说，每当 E_n 出现时，对于每个 $j = 1$、.....、 n，都有 $Y_{(j)} \le aj/n$。由于 ℓ_a 是递增的，

$$\sum_{i=1}^{n} \ell_a(Y_i) = \sum_{j=1}^{n} \ell_a(Y_{(j)}) \le \sum_{j=1}^{n} \ell_a\left(\frac{aj}{n}\right) = nm_a + o(n), \tag{43}$$

其中最后一步由性质2和性质3得出。

因此，在每个成功传播事件上，似然权重都可以确定性地从上方界定为 $e^{nm_a+o\ (n)}$。将此上界代入公式（32）可得

$$R_{N,k} \le e^{nm_a+o(n)} \mathbb{P}_{h_a}(E_n) \le e^{nm_a+o(n)}. \tag{44}$$

取对数，除以 N，并利用 $n/N \to x$ 和 $m_a \to m_x$，可得

$$\limsup_{N\to\infty} \frac{1}{N} \log R_{N,k} \le x m_x. \tag{45}$$

下界。 对于下界，我们仅考虑似然权重并非异常小的实现。固定一个容差 $\varepsilon > 0$ 并定义$B_{n,\ \varepsilon}$为满足以下条件的阈值构型（Y_1，.....， Y_n）： 样本平均对数似然比不会比其期望 m_a低超过 ε：

$$B_{n,\varepsilon} = \left\{ \frac{1}{n} \sum_{i=1}^{n} \ell_a(Y_i) \ge m_a - \varepsilon \right\}. \tag{46}$$

因此，$B_{n,\ \varepsilon}$ 包含所有似然权重并非异常小的实现，而其补集 $B^c_{n,\ \varepsilon}$ 则由平均对数似然比出现异常大的向下偏差的构型组成。

在交集 $E_n \cap B_{n,\ \varepsilon}$上，传播成功（$\mathbf{1}_{E_n} = 1$），且似然权重至少为 $e^{n(m_a-\varepsilon)}$。由于公式（32）中的被积函数处处非负，将期望限制在 $E_n \cap B_{n,\ \varepsilon}$上可得

$$R_{N,k} \ge \mathbb{E}_{h_a}\left[\mathbf{1}_{E_n \cap B_{n,\varepsilon}} \exp\left(\sum_{i=1}^{n} \ell_a(Y_i) \right) \right] \ge e^{n(m_a-\varepsilon)} \mathbb{P}_{h_a}(E_n \cap B_{n,\varepsilon}). \tag{47}$$

为了估计剩余概率，我们使用基本不等式

$$\mathbb{P}_{h_a}(E_n \cap B_{n,\varepsilon}) \ge \mathbb{P}_{h_a}(E_n) - \mathbb{P}_{h_a}(B^c_{n,\varepsilon}). \tag{48}$$

重要的是，该不等式不要求 E_n 和 $B_{n,\ \varepsilon}$ 彼此独立。第一个概率由停车函数结果给出下界，而根据性质 4，第二个概率呈指数级微小。因此

$$\mathbb{P}_{h_a}(E_n \cap B_{n,\varepsilon}) \ge \frac{(n+1)^{n-1}}{n^n} - e^{-c_\varepsilon n}. \tag{49}$$

关键在于，在辅助测度下成功传播的概率最多呈代数级下降，而

似然权重异常小的概率则呈指数级下降。因此，在指数尺度上，后者不可能剔除前者中很大一部分。

事实上，当 n 足够大时，代数项占主导地位，超过指数级小的修正项，因此可得

$$\lim_{N\to\infty}\frac{1}{N}\log\mathbb{P}_{h_a}(E_n\cap B_{n,\varepsilon})=0. \tag{50}$$

将这一结果代入式（47）可得

$$\liminf_{N\to\infty}\frac{1}{N}\log R_{N,k}\geq x(m_x-\varepsilon). \tag{51}$$

由于 $\varepsilon>0$ 可以任意小，我们得出

$$\liminf_{N\to\infty}\frac{1}{N}\log R_{N,k}\geq xm_x. \tag{52}$$

上界和下界相等。因此，我们证明了

$$\lim_{N\to\infty}\frac{1}{N}\log R_{N,k_N}=xm_x. \tag{53}$$

对于最优提议分布 $h_x=g_x^*$，有 $m_x=-\mathcal{I}_F(x)$。将这一结果与因式分解 $Q_{N,k}=F(a)^nR_{N,k}$ 结合，最终可得

$$\lim_{N\to\infty}\frac{1}{N}\log Q_{N,k_N}=x\log F(x)-x\mathcal{I}_F(x). \tag{54}$$

只要提议密度满足四个性质，这就证明了连续体预测（26）。

这一方法在阈值模型之外同样具有重要意义。熵最小化剖面确定了稀有事件最可能的实现，而测度变换则使得在有限样本量下实现该剖面成为可能。停车函数耦合确保约束事件在新分布下不会变得指数稀有，而似然界则恢复了该事件原本的指数成本。

在下一节中，我们将针对 Beta 阈值分布显式构造最优提议分布，并验证其屏障性质和平均性质（性质 1 和 3）。其余技术要求（性质 2 和 4）将在附录 A 中建立。

4. Beta 密度与四个最优分布

现在，我们将第 3 节中的约束熵方法应用于 Beta 阈值分布族。我们的目标是确定与规模为 x 的宏观级联相容的最可能阈值剖面，并据此计算指数传播率 $\psi_Q(x)$。

贝塔分布族的主要优势在于，其密度可以呈递减、递增、钟形或 U 形。这四种情况会导致不同的

传播约束满足方式，从而产生不同的最优分布。

在本节中，最优剖面 g_x^* 使相对熵 $D_{\mathrm{KL}}(g\|q_x)$（式 (27)）最小，且满足

$$G(t) = \int_0^t g(s)\,ds \geq \frac{t}{x}, \qquad 0 \leq t \leq x. \tag{55}$$

一旦确定该极小化函数，传播率即可由式 (26) 得出：

$$\psi_Q(x) = x \log F(x) - x D_{\mathrm{KL}}(g_x^* \| q_x). \tag{56}$$

下面的计算将确定最优分布并评估其熵代价。其最优性在附录 附录 B 中得到证明，而匹配点的存在性和唯一性则在附录 附录 C 中讨论。

4.1. 密度形状的作用

Beta密度及其累积分布为

$$f(t) = \frac{t^{\alpha-1}(1-t)^{\beta-1}}{\mathrm{B}(\alpha,\beta)}, \qquad F(t) = \int_0^t f(s)\,ds, \qquad \alpha,\beta > 0. \tag{57}$$

f 的定性行为由其对数导数决定，

$$\frac{f'(t)}{f(t)} = \frac{\alpha-1}{t} - \frac{\beta-1}{1-t}. \tag{58}$$

根据 α 和 β的取值，密度可能偏向较低的阈值、较高的阈值、中间阈值，或两端的极端阈值。

这种形状直接决定了经过条件化的阈值分布 $q_x(t) = f(t)/F(x)$能否在无需进一步调整的情况下维持顺序招募。如果 q_x 已在较小阈值处配置了足够的概率质量，使其累积分布在整个 $[0,x]$上都满足 $F(t)/F(x) \geq t/x$，则无需重排，且 $g_x^* = q_x$。反之，当这一条件不成立时，要维持规模为 x 的级联，就需要将阈值非典型地重新分布至较小的取值，并付出相应的熵成本。

在归一化和传播约束 $G(t) \geq t/x$ 下最小化相对熵 $D_{\mathrm{KL}}(g\|q_x)$，可得到具有简单分段结构的最优剖面 g_x^* 。对于贝塔分布族，区间 $[0,x]$ 至多包含一个自由区间和一个钉扎区间：

(i) 在自由区间上，累积约束处于非激活状态（在区间内部，$G(t) > t/x$）。相对熵最小化的驻定条件要求比值 $g_x^*(t)/q_x(t)$ 为常数。因此，最优密度保留了原始分布的形状： $g_x^*(t) \propto f(t)$。

(ii) 在钉扎区间上，累积分布达到传播边界 $G(t) = t/x$，这使得密度必须均匀：$g_x^*(t) = 1/x$。

对于本文所考虑的 Beta剖面，比例常数和接触点由自由区间与钉扎区间边界处的归一化条件和匹配条件确定。所得最优密度在这些边界处连续。

这种自由/钉扎分解为全部四个参数区间提供了统一框架。与其在每种情形下分别求解变分问题，不如通过确定哪些区间被钉扎、哪些区间保持自由，以及它们的接触点如何随 x变化，来对大偏差行为进行分类。

为后文使用，定义

$$\begin{aligned} J(c) &= \int_0^c \log f(t)\, dt \\ &= (\alpha-1)(c\log c - c) - (\beta-1)\big[(1-c)\log(1-c)+c\big] - c\log \mathrm{B}(\alpha,\beta). \end{aligned} \tag{59}$$

在 $c = 0$处，表达式 $c\log\, c$按连续性理解。函数 $J(c)$使我们能够在确定最优剖面后，以闭式形式写出传播速率。

4.2. 递减密度：无需重排

首先考虑

$$\alpha \le 1, \qquad \beta \ge 1, \tag{60}$$

其中 Beta密度是非递增的。在这种情况下，低阈值至少与高阈值一样可能出现。因此，我们预计条件分布 q_x 中包含足够多的低阈值个体，无需任何额外调整即可满足传播约束。

这一期望可以直接验证。由于 f 是非递增的，只要满足 $t \le x$，它在 $[0,t]$ 上的平均值就至少与在 $[0,x]$ 上的平均值一样大：

$$\frac{F(t)}{t} \ge \frac{F(x)}{x}. \tag{61}$$

等价地，

$$\frac{F(t)}{F(x)} \ge \frac{t}{x}. \tag{62}$$

但是， $F(t)/F(x)$恰好是 q_x的累积分布。因此，条件密度已经满足传播屏障。

由于相对熵非负，且仅当两个分布一致时才为零，因此改变分布并无益处。最优剖面就是

$$g_x^*(t) = q_x(t) = \frac{f(t)}{F(x)}. \tag{63}$$

其熵成本为零，由公式（56）立即得到

$$\psi_Q(x) = x \log F(x). \tag{64}$$

这是变分问题最简单的可能结果。主要的指数成本完全来自要求所有招募阈值都低于 x。额外的序贯传播约束不会带来进一步的指数惩罚。

4.3. 递增密度：整个剖面都被钉定

现在考虑

$$\alpha \geq 1, \qquad \beta \leq 1, \tag{65}$$

其中 f 是非递减的。此时情况正好相反。较高的阈值更受青睐，因此条件分布在原点附近赋予的概率相对较小。除均匀情形外，这种分布无法维持序贯招募，除非将阈值重新分配到较小的取值。

最优重排尤其简单：累积分布在整个区间内都沿着传播边界变化，

$$G_x^*(t) = \frac{t}{x}. \tag{66}$$

求导可得

$$g_x^*(t) = \frac{1}{x}, \qquad 0 < t < x. \tag{67}$$

因此，整个区间都被钉定。最优剖面是均匀的，尽管原始密度通常是递增的。

均匀最优剖面的出现，体现了最小化熵成本与满足各阶段招募屏障之间的平衡。由于原始密度是非递减的，将概率质量移向较小的阈值会产生熵惩罚。传播约束要求低阈值个体的累积供给至少达到某一最低水平，该水平由边界 $G(t) = t/x$ 表示。任何严格高于这条边界的可容许剖面，都包含额外的低阈值个体。对于非递减的原始密度，这种过量供给无法降低相对熵代价。因此，代价最低的可容许剖面会在整个区间内使传播约束取等号。附录 附录 B 中给出的相对熵最优性论证，严格地推出了这一结论。

要计算速率，将 $g_x^* = 1/x$ 代入相对熵：

$$D_{\mathrm{KL}}(g_x^* \| q_x) = \frac{1}{x} \int_0^x \log \frac{F(x)}{x f(t)}\, dt. \tag{68}$$

利用 $J(x)$ 的定义，我们得到

$$D_{\mathrm{KL}}(g_x^* \| q_x) = \log F(x) - \log x - \frac{J(x)}{x}. \tag{69}$$

方程 (56) 随后给出

$$\psi_Q(x) = x\log x + J(x). \tag{70}$$

与递减密度情形不同，成功传播通常需要对阈值分布进行大规模重组。由此产生的熵代价会改变主导指数传播率。

两个极限情形与早期的精确结果建立了有益的联系。首先，对于 $\alpha = 1$和 $0 < \beta < 1$，Zeraoulia [17] 获得了微观级联的极限分布，其中最终级联规模 k 在 $N \to \infty$时保持不变。特别地，传播因子满足

$$Q_{N,k} \sim \frac{\beta^{k-1}k^{k-2}}{N^{k-1}}, \qquad k \geq 2. \tag{71}$$

相反，我们的结果描述的是宏观级联，其中 $k/N \to x \in (0,\ 1)$。然而，这两种渐近描述在主导指数依赖关系的层面上是一致的。事实上，对于 $\alpha = 1$ 和 $x \to 0$,

$$\psi_Q(x) = x\log(\beta x) + O(x^2). \tag{72}$$

形式上令 $x = k/N$，可得到因子 $(\beta k/N)^k$，其具有与微观表达式相同的主导指数依赖关系。仅凭宏观速率无法恢复其余的代数因子，因为极限$k = O(1)$和 $k/N \to x > 0$ 描述的是不同的渐近区域。

第二个基准情形是均匀分布， $\alpha = \beta = 1$，详见第2.3节。在这种情况下，停车函数恒等式给出有限总体的精确表达式

$$Q_{N,k} = \frac{k^{k-2}}{N^{k-1}}, \tag{73}$$

因此 $\psi_Q(x) = x\log\ x$，这与式（70）的预测完全一致。由此，均匀情形在有限总体组合学与宏观大偏差率之间建立了精确联系。

4.4. 钟形密度：先钉扎区域，后自由区域

接下来考虑

$$\alpha > 1, \qquad \beta > 1. \tag{74}$$

此时，Beta密度呈钟形，在其众数之前递增

$$m = \frac{\alpha - 1}{\alpha + \beta - 2} \tag{75}$$

此后则递减。单调性的这一变化表明，最优剖面可能结合上述两种机制。

对于级联规模 $x \leq m$，密度在相关区间 $[0, x]$ 内始终递增。因此，前一小节的论证仍然适用：整个最优剖面都受到约束，

$$g_x^*(t) = \frac{1}{x}, \qquad x \leq m, \tag{76}$$

and

$$\psi_Q(x) = x \log x + J(x). \tag{77}$$

当 $x > m$ 时，情况就更有意思了。此时，原始密度在区间的一部分递增，在其余部分递减。因此，我们预计传播约束在较小的阈值处会生效（此时原始密度递增），但沿区间继续前进后会失效。

由此得到一个包含钉扎区域 $0 < t \le c_x$ 和自由区域$c_x < t < x$的剖面：

$$g_x^*(t) = \begin{cases} \dfrac{1}{x}, & 0 < t \le c_x, \\ \dfrac{f(t)}{x f(c_x)}, & c_x < t < x. \end{cases} \tag{78}$$

匹配点 c_x 必须满足 $0 < c_x < m$。事实上，若要使 $G_x^*(t)$离开 $t = c_x$后立即上升到传播屏障 t/x 之上，则当 t 略大于 c_x时，自由区域的密度必须超过 $1/x$。这要求在刚越过 c_x后，$f(t) > f(c_x)$，而这只有在众数的递增一侧（$c_x < m$）才有可能。

自由区域中的比例常数由要求最优密度在 c_x处连续来确定：$g_x^*(c_x) = 1/x$。这解释了公式（78）第二个分支中的因子$1/[x f(c_x)]$ 。剩下的未知量——位置 c_x 本身——则由要求整个剖面归一化来确定。

归一化条件为

$$\int_0^{c_x} \frac{dt}{x} + \int_{c_x}^{x} \frac{f(t)}{x f(c_x)}\, dt = 1, \tag{79}$$

由此可得

$$F(x) - F(c_x) = (x - c_x) f(c_x). \tag{80}$$

这个方程具有简单的几何意义。原始密度在 c_x 和 x 之间曲线下的面积，必须等于一个宽度为 $x - c_x$、高度为 $f(c_x)$的矩形面积。匹配条件确保自由部分恰好包含完成归一化分布所需的概率质量。

对于每个 $x > m$，都存在唯一解 $c_x \in (0, m)$。相应的累积分布在 $t \le c_x$ 上位于传播边界，在 $c_x < t < x$ 上升至边界之上，并在 $t = x$ 处再次与边界相交。这些结论的证明见附录 附录 C。

现在我们可以计算传播率。在钉扎区间上，$g_x^* = 1/x$；而在自由区间上，$g_x^* = f(t)/[x f(c_x)]$。利用 ψ_Q的变分表达式，两个部分为

$$\psi_Q(x) = \int_0^{c_x} \log[x f(t)]\, dt + (x - c_x) \log[x f(c_x)]. \tag{81}$$

第一项是钉扎区域的贡献。第二项来自自由区域，在该区域中，原始密度与最优密度之比为

常数。合并各项可得

$$\psi_Q(x) = x \log x + J(c_x) + (x - c_x) \log f(c_x). \tag{82}$$

为方便起见，我们可以设定 $c_x = x$ when $x \le m$。按照这一约定，公式（82）也描述了完全钉扎区域。

*4.5. U*形密度：先是自由区域，随后是钉扎区域

最后，考虑

$$\alpha < 1, \qquad \beta < 1. \tag{83}$$

Beta密度呈U形。从原点开始递减，直至在以下位置达到最小值

$$m = \frac{1 - \alpha}{2 - \alpha - \beta} \tag{84}$$

随后逐渐递增至上端点。极小阈值的丰富分布起初有利于传播。然而，密度的递增部分最终会带来不同的情况：条件分布可能不再包含足够多的低阈值个体，无法在整个 $[0, x]$ 内满足传播屏障。因此，我们预计会出现两个区域，它们由一个特征值 d分隔。

为确定值 d，考虑条件密度的累积分布，

$$G(t) = \frac{F(t)}{F(x)}. \tag{85}$$

传播约束等价于

$$\frac{F(t)}{t} \ge \frac{F(x)}{x}, \qquad 0 < t \le x. \tag{86}$$

因此，相关量是平均密度 $F(t)/t$。其导数为

$$\frac{d}{dt}\frac{F(t)}{t} = \frac{t f(t) - F(t)}{t^2}. \tag{87}$$

对于 U 形 Beta密度，该平均值先减小后增大。其最小值出现在满足以下条件的唯一点 $d > m$

$$F(d) = d f(d). \tag{88}$$

从几何上看，密度 $f(d)$ 等于其在区间 $[0, d]$ 上的平均值。 d 的存在性和唯一性在附录附录 C 中得到证明。

对于 $x \le d$，平均密度 $F(t)/t$ 在整个 $[0, x]$ 上均为非递增。因此，条件分布已经满足传播屏障：

$$\frac{F(t)}{F(x)} \ge \frac{t}{x}. \tag{89}$$

无需进行重排，最优剖面为

$$g_x^*(t) = q_x(t), \qquad x \le d. \tag{90}$$

相应的速率为

$$\psi_Q(x) = x \log F(x). \tag{91}$$

然而，对于 $x > d$，条件密度不再满足可行性要求，必须对最优剖面进行修改。与钟形情形相反，此时自由区域先出现，随后是钉扎区域：

$$g_x^*(t) = \begin{cases} \dfrac{f(t)}{x f(d)}, & 0 < t \le d, \\ \dfrac{1}{x}, & d < t < x. \end{cases} \tag{92}$$

自由部分保留了原始密度的形状，包括其在原点附近的高度集中。钉扎部分则沿着剩余区间上的传播边界。匹配条件 $F(d) = d f(d)$ 保证了归一化：

$$\frac{F(d)}{x f(d)} + \frac{x - d}{x} = 1. \tag{93}$$

它还确保累积分布在 $t = d$ 处达到传播边界。在此之前，最优累积分布位于边界之上；此后则与边界重合。

这一情形的一个有趣特点是，匹配点 d 不依赖于最终级联规模 x。一旦 $x > d$，增大级联规模会延长钉扎区间，但自由区间保持不变。

速率由自由区域和钉扎区域的贡献相加得到：

$$\psi_Q(x) = d \log[x f(d)] + \int_d^x \log[x f(t)]\, dt. \tag{94}$$

根据 J的定义，我们得到

$$\psi_Q(x) = x \log x + d \log f(d) + J(x) - J(d), \qquad x > d. \tag{95}$$

因此，U形族在最优剖面中经历了不同的变化：对于 $x \le d$ ，原始条件分布已经是可容许的；而对于 $x > d$ ，区间的上端出现了钉扎区域。

图1展示了熵最小化阈值分布的四种可能几何形态。与条件密度 q_x 的比较表明，传播约束在哪些位置要求重新分配阈值，以及原始密度剖面在哪些位置可以保持其形状。重要的是，在自由区间上，最优密度与 q_x成正比，但不一定与其重合。这是因为比例常数由整个最优分布的归一化决定，其中包括任何钉扎区间。只有当整个区间都是自由区间时，才有 $g_x^* = q_x$，如子图（a）所示。

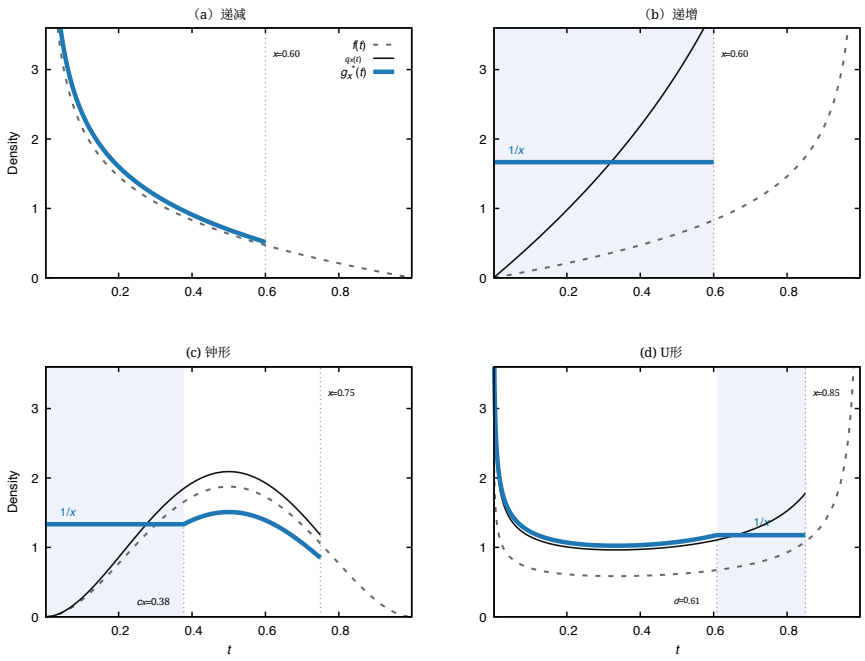


图1：贝塔分布阈值对应的熵最小化阈值密度 $g_x^*(t)$ 的四种情形。灰色虚线曲线表示原始密度 $f(t)$，黑色细曲线表示条件密度 $q_x(t) = f(t)/F(x)$，蓝色粗曲线表示最优剖面 $g_x^*(t)$。(a) 递减密度：$\alpha = 0.6$, $\beta = 2$, $x = 0.60$；整个区间均为自由区间。(b) 递增密度：$\alpha = 2$, $\beta = 0.6$, $x = 0.60$；整个区间均为钉扎区间。(c) 钟形密度：$\alpha = \beta = 3$, $x = 0.75$；先是钉扎区间，随后是自由区间，其中 $c_x \approx 0.38$。(d) U形密度：$\alpha = 0.7$, $\beta = 0.4$, $x = 0.85$；先是自由区间，随后是钉扎区间，其中 $d \approx 0.61$。浅色阴影表示钉扎区间，在这些区间内 $g_x^*(t) = 1/x$。条件密度和最优密度仅对 $0 < t < x$ 有定义。

这些显式解也完成了第 3 节的有限样本论证：它们满足传播屏障（性质1），并且其自由区间/固定区间结构确保了似然平均值相等（性质 3）。其余技术条件（性质 2 和 4）在附录 A 中得到验证。因此，上文推导出的速率正是有限总体传播概率的精确领先指数速率。

4.6. 最终级联规模速率

前述分析确定了传播因子 $Q_{N,\,k}$ 的主导指数速率。利用公式 (4) 和斯特林公式，我们得到最终级联规模分布相应的大偏差率，

$$I_P(x) = -\lim_{N\to\infty} \frac{1}{N} \log P_N(k) = x\log x + (1-x)\log(1-x) - \psi_Q(x) - (1-x)\log[1-F(x)], \tag{96}$$

其中 $k/N \to x \in (0,\ 1)$。因此，上一节推导出的显式传播速率也确定了宏观最终级联规模出现的主导指数概率。

第2.3节讨论的均匀情形 $\alpha = \beta = 1$ 提供了一个精确基准。停车函数恒等式给出 $\psi_Q(x) = x \log x$ and $I_P(x) = 0$，尽管任意特定宏观级联规模的概率都按$N^{-3/2}$衰减。这说明，大偏差率趋于零并不意味着极限概率非零。

5. 数值验证与最优经验分布

我们针对两种具有代表性的分布，检验第4节推导出的解析大偏差率和最优经验阈值剖面：Beta(2,2)（钟形）和 Beta$(1/2, 1/2)$（U形）。我们使用重要性抽样估计传播速率，并使用序贯蒙特卡洛考察以成功发生宏观级联为条件的阈值分布。

对于重要性抽样估计，我们固定 $k = \lfloor xN \rfloor$、 $n = k - 1$ 和$a = n/N$，并从第4节推导出的最优提议分布 $h_a = g_a^*$ 生成阈值。利用式（10）中定义的传播事件 E_n ，条件传播概率估计为

$$\widehat{R}_{N,k} = \frac{1}{M} \sum_{m=1}^{M} \mathbf{1}_{E_n}^{(m)} \prod_{i=1}^{n} \frac{q_a(Y_i^{(m)})}{h_a(Y_i^{(m)})}, \tag{97}$$

其中， M 是独立样本的数量。有限总体速率估计由精确因式分解（9）得出:

$$\widehat{\psi}_{Q,N}(x) = \frac{n}{N} \log F(a) + \frac{1}{N} \log \widehat{R}_{N,k}. \tag{98}$$

对于每个 x 和 $N = 100$，我们使用 $M = 20000$ 个样本， $N = 100$分别为 400 和 1600，并以对数形式累积似然比，以避免数值下溢。

为了对以传播为条件的经验阈值分布进行采样，我们采用了单独的序贯蒙特卡洛算法。在 q_a条件下依次生成有序阈值，并在每一步施加传播约束。给定 $Y_{(j-1)} = y$，下一个次序统计量满足其约束的概率为

$$p_j(y) = 1 - \left[\frac{F(a) - F(j/N)}{F(a) - F(y)} \right]^{n-j+1}, \tag{99}$$

其中， $y = 0$ 适用于 $j = 1$。粒子权重按 $p_j(y)$更新；当有效样本量低于预设阈值时，进行系统重采样。使用估计的传播概率作为权重，合并独立运行结果。由于此方法不使用解析最优解 g_a^*，因此可独立检验预测的条件分布。

图 2(a) 显示，随着 N 增加，有限总体速率估计逐渐趋近于解析预测；对于钟形分布，有限尺寸偏差更为显著。图 2(b) 验证了预测的四种

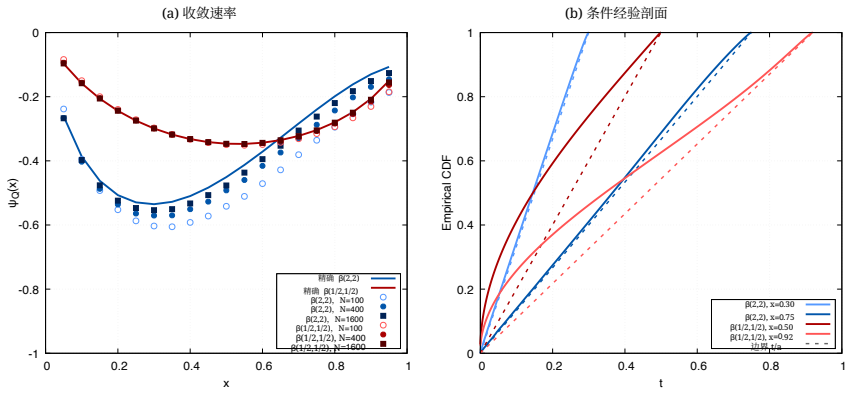


图 2：传播速率和条件经验阈值剖面的数值验证。(a) 对 Beta(2,2)（蓝色）和 Beta$(1/2, 1/2)$（红色）的有限总体 $N^{-1}\log Q_{N,k}$ 估计，其中 $k = \lfloor xN \rfloor$。空心圆、实心圆和方块分别对应 $N = 100$、400 和 1600。实线为解析渐近速率。每个估计使用 20000 个重要性采样样本。(b) 通过序贯蒙特卡洛得到的 $N = 1600$ 条件经验阈值累积分布函数。曲线对应于 $x = 0.30$ 为 0.75 时的 Beta(2,2)，以及 $x = 0.50$ 为 0.92 时的 Beta$(1/2, 1/2)$。虚线表示传播边界 $G(t) = t/a$，其中 $a = (k-1)/N$。

最优剖面几何形态：Beta(2,2) 对应完全钉扎和钉扎–自由，Beta$(1/2, 1/2)$ 对应完全自由和自由–钉扎。这种一致性为指数传播速率以及导致罕见宏观级联的阈值构型提供了相互补充的证据。

6. α = 1 附近的临界标度

我们首先回顾无限总体极限下的确定性动力学。从单个发起者开始，其初始比例随 $N \to \infty$ 消失，已激活个体的比例按照 $x_{t+1} = F(x_t)$ 演化。其不动点满足 $x^* = F(x^*)$，并且 $x^* = 0$ 始终是一个解。该解的稳定性由 $F'(0^+) = f(0^+)$ 决定：当 $f(0^+) < 1$ 时稳定，而当 $f(0^+) > 1$ 时不稳定。

对于 Beta(α, β) 族，$f(0^+)$ 在 $\alpha > 1$ 时 $= 0$，而在 $\alpha < 1$ 时发散；与此同时，$f(0^+)$ 在 $\alpha = 1$ 处 $= \beta$。因此，对于 $0 < \beta \leq 1$，$\alpha_c = 1$ 将无穷小的初始扰动会消退的区域与能够触发宏观传播的区域分隔开来。在 $\alpha = 1$ 处，不动点 $x^* = 0$ 对于 $\beta = 1$ 是边际稳定的，但对于 $0 < \beta < 1$ 仍保持线性稳定。

现在，我们来考察这一相变如何反映在平均最终级联比例的有限尺寸标度中，

$$\rho_\infty(N) = \frac{1}{N} \sum_{k=1}^{N} k P_N(k). \tag{100}$$

此前，文献 [13] 研究过幂律阈值（$\beta = 1$）以及满足 $0 < \beta < 1$ 的贝塔分布下的这一问题。然而，其中采用的大偏差分析依赖于近似 $\psi_Q(x) = x \log F(x)$，而当最优经验阈值剖面包含钉扎区间时，这一近似并非普遍成立。

此处得到的修正后传播率促使我们重新审视这些标度结果。具体而言，对于 $k/N \to x > 0$ 的宏观级联，其受大偏差率 $I_P(x)$ 控制；而对于 k 在 $N \to \infty$ 时仍保持有限的微观级联，则需要单独分析早期招募过程。这两种情形都可能对式（100）有所贡献，其相对重要性取决于阈值分布。

6.1. 幂律阈值：$\beta = 1$

对于 $\beta = 1$，累积阈值分布为 $F(x) = x^{\alpha}$。相变发生在 $\alpha_c = 1$，此时阈值分布为均匀分布。

我们首先考察相变附近的宏观大偏差率。记 $\alpha = 1 + \epsilon$，其中 $|\epsilon| \ll 1$，修正后的速率取决于 ϵ 的符号。

对于 $\epsilon < 0$，阈值密度递减，且最优剖面完全自由。因此，$\psi_Q(x) = x \log F(x)$，最终级联规模速率简化为二元相对熵

$$I_P(x) = x \log \frac{x}{F(x)} + (1 - x) \log \frac{1 - x}{1 - F(x)}. \tag{101}$$

在固定 $0 < x < 1$ 下展开，得到

$$I_P(x) = \frac{x(\log x)^2}{2(1 - x)} \epsilon^2 + O(|\epsilon|^3), \qquad \epsilon < 0. \tag{102}$$

对于 $\epsilon > 0$，密度递增，且最优剖面完全钉扎。传播率则为

$$\psi_Q(x) = x \log x + \int_0^x \log(\alpha t^{\alpha - 1})\, dt, \tag{103}$$

可将其显式计算为

$$\psi_Q(x) = x \log x + x \log \alpha + (\alpha - 1)(x \log x - x). \tag{104}$$

代入 $I_P(x)$ 的精确表达式，并在 $\alpha = 1$ 附近展开，得到

$$I_P(x) = \frac{x}{2} \left[1 + \frac{(\log x)^2}{1 - x} \right] \epsilon^2 + O(\epsilon^3), \qquad \epsilon > 0. \tag{105}$$

方程 (102) 和 (105) 表明，在相变两侧，宏观速率关于$\alpha - 1$ 均呈二次关系，但系数不同。特别是，钉扎区间会额外贡献一项 $x\epsilon^2/2$，而参考文献 [13] 所采用的近似中不存在该项。

对于固定的宏观 x，因此当 $N(\alpha-1)^2 = O(1)$ 时，指数抑制开始变得显著，从而确定尺度

$$|\alpha - 1| \sim N^{-1/2}. \tag{106}$$

这一结论涉及宏观大偏差包络；级联规模均值的标度也取决于其亚指数前因子。

在临界点 $\alpha = \beta = 1$，传播因子可精确求得，

$$Q_{N,k} = \frac{k^{k-2}}{N^{k-1}}. \tag{107}$$

由此得到的最终级联规模分布满足

$$kP_N(k) = \binom{N}{k} \left(\frac{k}{N}\right)^k \left(1 - \frac{k}{N}\right)^{N-k}. \tag{108}$$

因此，

$$\rho_\infty(N) \sim \sqrt{\frac{\pi}{2N}}. \tag{109}$$

这一精确结果不受宏观传播速率修正的影响。因此，指数 $1/2$ 可独立于大偏差近似而确定。修正后的理论保留了宏观尺度 $N^{-1/2}$，同时揭示了转变两侧的不对称性。

6.2. 0 < β < 1 时的 *Beta* 阈值

当 $0 < \beta < 1$ 时，情况在定性上有所不同。在这种情况下，$\alpha = 1$ 处的相变是连续的，非零的确定性不动点通过本性奇异性趋近原点。

事实上，对于较小的 x，Beta 累积分布的展开式为

$$F(x) = \frac{x^\alpha}{\alpha B(\alpha, \beta)}[1 + O(x)]. \tag{110}$$

令 $\alpha = 1 + \epsilon$，并求解 $F(\rho_\infty) = \rho_\infty$，对于 $\epsilon \to 0^-$，可得

$$\log \rho_\infty = \frac{\log \beta}{-\epsilon} + O(1). \tag{111}$$

因此，当 $\alpha \to 1^-$ 时，确定性序参量以本性奇异性的形式趋于零。

然而，在临界点 $\alpha = 1$，由单个发起者启动的级联行为受微观招募支配。由于

$$F(x) = 1 - (1 - x)^\beta = \beta x + O(x^2), \tag{112}$$

随着 $N \to \infty$，早期招募过程收敛到一个平均后代数为 β 的泊松伽尔顿–沃森分枝过程。

Zeraoulia [17]严格地确立了这种联系。具体而言，他证明了对于 $\alpha = 1$ 和 $0 < \beta < 1$，极限级联规模分布是 Borel 分布

$$\lim_{N\to\infty} P_N(k) = \frac{e^{-\beta k}(\beta k)^{k-1}}{k!}, \qquad k = 1, 2, \ldots. \tag{113}$$

由于 $\beta < 1$，该分枝过程是次临界的，其期望总后代数（包括初始发起者）为

$$\langle k\rangle \longrightarrow \frac{1}{1-\beta}. \tag{114}$$

因此，

$$\rho_\infty(N) \sim \frac{1}{(1-\beta)N}, \qquad 0 < \beta < 1. \tag{115}$$

因此，对于 $\beta = 1/2$，精确的极限振幅为 2，而不是文献 [13] 中蒙特卡罗模拟所估算的值。

现在，我们将这一微观描述扩展到有限规模临界区域，并令 α 随总体规模递增而同时趋近于 1。这一联合极限给出了一个显式标度函数，而仅凭临界Borel分布无法得到该函数。

沿用文献 [13]，中引入的对数标度变量，我们写作

$$\alpha = 1 + \frac{z}{\log N}, \tag{116}$$

其中，当 $N \to \infty$ 时， z 保持不变。利用方程 (110)，我们发现，对于每个固定的正整数 j，

$$NF(j/N) = \frac{N^{1-\alpha} j^\alpha}{\alpha B(\alpha, \beta)}[1 + O(N^{-1})] \longrightarrow \beta j e^{-z}. \tag{117}$$

因此，因子 $N^{1-\alpha} = e^{-z}$ 会改变微观级联规模下的有效招募率。

在联合缩放极限下，早期招募过程再次收敛到泊松 Galton–Watson 分枝过程，但此时其有效平均后代数为

$$\lambda(z) = \beta e^{-z}. \tag{118}$$

相应地，每个固定级联规模的极限分布为

$$\lim_{N\to\infty} P_N(k) = \frac{e^{-\lambda(z)k}[\lambda(z)k]^{k-1}}{k!}, \qquad k = 1, 2, \ldots. \tag{119}$$

当 $\lambda(z) < 1$ 时，这是一个均值有限的归一化次临界Borel分布

$$\langle k\rangle \longrightarrow \frac{1}{1-\lambda(z)} = \frac{1}{1-\beta e^{-z}}. \tag{120}$$

因此，亚临界区间内平均活跃比例的有限尺寸标度为

$$N\rho_\infty(N) \longrightarrow \mathcal{F}_\beta(z) = \frac{1}{1-\beta e^{-z}}, \qquad z = (\alpha-1)\log N > \log\beta. \tag{121}$$

方程 (121) 是一项明确的有限尺寸标度预测，它将微观分枝过程描述延伸至临界线 $\alpha = 1$ 之外。它还为此前从参考文献 [13] 的数值数据中发现的对数标度变量提供了理论解释。当 $z = 0$ 时，该标度函数再现了 Zeraoulia 的临界振幅，

$$\mathcal{F}_\beta(0) = \frac{1}{1-\beta}, \tag{122}$$

而对于 $z \to +\infty$， $\mathcal{F}_\beta(z) \to 1$，对应于孤立发起者。

限制条件 $z > \log\beta$ 具有直接的物理解释。当 $z = \log\beta$ 时，有效分枝过程达到临界状态，其总后代平均数发散。当 $z < \log\beta$ 时，微观分枝过程处于超临界状态，其有限后代分布不再能描述完整的级联规模分布。在此状态下，还必须考虑规模随 N 增长的级联所作出的贡献。

因此，对数标度变量描述了微观招募状态之间的交叉，而方程 (121) 给出了其次临界侧的明确标度函数。结合本研究推导出的修正后宏观大偏差速率，这一结果更完整地描述了 $\alpha_c = 1$ 附近的有限尺寸行为。

7. 讨论

我们对 Granovetter 阈值模型 [1]的精确有限总体解进行了系统的渐近分析。通过用个体阈值的经验分布表示传播概率，我们将大偏差问题重新表述为约束相对熵极小化问题。该方法避免了文献 [13]中所采用的鞍点近似的局限性，并为任意连续阈值分布给出了正确的宏观传播速率。

对于贝塔分布阈值，最小化问题有一个显式解，其特征是最优经验阈值剖面包含自由区间和钉扎区间。根据阈值密度的形状，可出现四种不同的几何类型：完全自由、完全钉扎、钉扎–自由和自由–钉扎。由此得到的解析速率函数与有限总体重要性抽样估计相符，而序贯蒙特卡洛模拟也独立地再现了预测的条件经验剖面。因此，除了确定罕见级联发生的指数级概率外，这一形式体系还揭示了导致其发生的阈值构型。

我们还重新考察了文献 [13] 中研究的 $\alpha_c = 1$ 附近的临界行为。尽管该分析采用了错误的宏观传播速率，但主要的有限尺寸标度律依然成立。具体而言，对于均匀阈值，平均最终活跃比例按 $N^{-1/2}$ 标度；对于 $\alpha = 1$ 和 $0 < \beta < 1$，则按 N^{-1} 标度；经修正的临界振幅 $1/(1-\beta)$ 由 Zeraoulia [17] 确立。我们还为亚临界微观区间中的对数临界窗口（$\alpha - 1$） $\log N$ 推导出了显式标度函数，从而将有限尺寸行为与底层的分枝过程描述联系起来。

本方法的主要优势在于其通用性。约束熵表述不依赖于贝塔分布族，可应用于其他阈值分布，包括那些难以通过传统鞍点方法进行论证的分布。更广泛地说，结合经验测度大偏差与序贯激活约束，为研究不可逆采纳、传染和级联过程中的罕见集体事件提供了一个框架。只要动力学能够通过对经验分布的约束来表达，同一策略就有望给出罕见结果的指数速率，并刻画产生这些结果的微观构型。

致谢

JFF 部分łձšu Conselho Nacional de Desenvolvimento Científico e Tecnológico 资助，项目编号为 300200/2026-9。

附录 A. 有限样本速率的技术论证

在第 3 节中，我们基于辅助概率密度h_a的四个性质，建立了指数传播率的有限样本判据。对于第4节中构造的最优提议分布 $h_a = g_a^*$，性质 1 和 3 可由其显式形式直接推出。

这里我们验证剩余的两个性质：似然网格平均值的收敛性（性质 2）以及似然权重的指数集中（性质 4）。之所以需要这些估计，是因为 Beta密度

$$f(t) = \frac{t^{\alpha-1}(1-t)^{\beta-1}}{B(\alpha,\beta)}, \qquad \alpha,\beta > 0, \tag{A.1}$$

可能在原点处趋于零或发散。

在本附录中，我们考虑序列 $a = n/N \to x$，其中 $x \in$ 固定在 (0,1) 内，并记为

$$\ell_a(t) = \log \frac{q_a(t)}{h_a(t)}, \qquad m_a = \int_0^a h_a(t)\ell_a(t)\,dt. \tag{A.2}$$

附录 *A.1.* 似然网格平均值的收敛性我们首先验证性质 2，
即

$$\frac{1}{n}\sum_{j=1}^{n}\ell_a\left(\frac{aj}{n}\right)=\frac{1}{a}\int_0^a \ell_a(t)\,dt+o(1). \tag{A.3}$$

对于最优 Beta 提议分布，对数似然比 $\ell_a(t)$ 在 $(0,\ a)$ 上单调不减。

当 $\alpha>1$ 时，Beta 密度在原点处按 $f(t)\sim Ct^{\alpha-1}$ 的方式趋于零。在钉扎区间上，其中 $h_a(t)=1/a$，因此我们有

$$\ell_a(t)=(\alpha-1)\log t+O(1),\qquad t\to 0. \tag{A.4}$$

尽管该表达式发散至负无穷，但其奇异性可积。此外，右端点网格 $a/n,\ 2a/n$、.....、a 不包含奇异点 $t=0$。

由单调性可得单侧估计

$$\frac{1}{n}\sum_{j=1}^{n-1}\ell_a\left(\frac{aj}{n}\right)\le\frac{1}{a}\int_0^a \ell_a(t)\,dt. \tag{A.5}$$

加上最后一个网格值，并利用性质 3，可得

$$\frac{1}{n}\sum_{j=1}^{n}\ell_a\left(\frac{aj}{n}\right)\le m_a+\frac{\ell_a(a)}{n}. \tag{A.6}$$

对于处于任意固定 $x\in(0,\ 1)$ 的紧邻域内的 a，$\ell_a(a)$ 保持有界。因此，公式 (A.6) 已给出了第 3.4 节上界所需的单侧估计。

为建立公式 (A.3) 中的完全收敛，我们将 ℓ_a 可能存在的对数奇异性与正则部分分离开来。正则贡献通过通常的黎曼和收敛来处理。对于对数贡献，相关的和可归结为如下形式的表达式

$$\sum_{j=1}^{m}\log\frac{j}{n}=\log(m!)-m\log n. \tag{A.7}$$

随后，斯特林公式表明，归一化和与其相应积分之间的差为 $O(\log n/n)$。

当 a 均匀地趋近于内部点 x 时，同样的估计也成立，包括固定区间和自由区间发生变化的接触过渡点。因此，

$$\frac{1}{n}\sum_{j=1}^{n}\ell_a\left(\frac{aj}{n}\right)=m_a+O\left(\frac{\log n}{n}\right), \tag{A.8}$$

这证明了性质 2。

附录 *A.2.* 似然权重的指数集中

现在验证性质 4。对于每个固定的 $\varepsilon > 0$，我们必须证明

$$\mathbb{P}_{h_a}\left\{\frac{1}{n}\sum_{i=1}^{n}\ell_a(Y_i) < m_a - \varepsilon\right\} \le e^{-c_\varepsilon n}, \tag{A.9}$$

对于某个 $c_\varepsilon > 0$，当 a 充分接近 x时一致成立。变量 $Y_1, \dots, Y_n$是来自 h_a 的独立样本。

对于 $\theta > 0$，定义

$$\Lambda_a(\theta) = \log \mathbb{E}_{h_a}\left[e^{-\theta \ell_a(Y)}\right] = \log \int_0^a h_a(t) e^{-\theta \ell_a(t)}\, dt. \tag{A.10}$$

对正随机变量 $\exp[-\theta \sum_i \ell_a(Y_i)]$应用马尔可夫不等式，得到

$$\mathbb{P}_{h_a}\left\{\frac{1}{n}\sum_{i=1}^{n}\ell_a(Y_i) < m_a - \varepsilon\right\} \quad \le \exp\left\{n\left[\theta(m_a - \varepsilon) + \Lambda_a(\theta)\right]\right\}. \tag{A.11}$$

我们必须证明，方程（A.10）中的矩生成函数对于充分小的正 θ 是有限的。

对于非增 Beta 密度（$\alpha \le 1,\ \ \beta \ge 1$），最优提议分布为$h_a = q_a$。因此，$\ell_a = 0$，而方程（A.9）中的概率恒等于零。

对于 U 形 Beta 密度 ($\alpha < 1, \beta < 1$)，似然比在原点附近为常数，在 $(0,\ a)$ 的其余部分有界。因此，对每个 $\theta > 0,\ \ \Lambda_a(\theta)$ 都是有限的。

对于递增且钟形的 Beta 密度 ($\alpha > 1$)，$\ell_a(t)$ 中唯一可能出现无界负贡献的地方，是原点附近的一个钉扎区间，在该区间内

$$\ell_a(t) = (\alpha - 1)\log t + O(1). \tag{A.12}$$

由于 $h_a(t) = 1/a$ 在该处成立，式 (A.10) 中的被积函数表现为

$$h_a(t) e^{-\theta \ell_a(t)} \propto t^{-\theta(\alpha-1)}. \tag{A.13}$$

只要满足以下条件，该函数在 $t = 0$ 附近可积

$$0 < \theta < \frac{1}{\alpha - 1}. \tag{A.14}$$

因此，在每一种 Beta 情形下，对于充分小的正 $\theta,\ \ \Lambda_a(\theta)$ 都是有限的；当 a 取遍任意固定 $x \in (0,\ 1)$ 的某个紧邻域时，这一结论一致成立。此外，

$$\Lambda_a(0) = 0, \qquad \Lambda_a'(0) = -\int_0^a h_a(t)\ell_a(t)\, dt = -m_a. \tag{A.15}$$

因此，局部展开为

$$\Lambda_a(\theta) = -\theta m_a + O(\theta^2), \tag{A.16}$$

其余项在 a 附近的 x范围内均匀有界。代入方程（A.11）可得

$$\mathbb{P}_{h_a}\left\{\frac{1}{n}\sum_{i=1}^{n}\ell_a(Y_i) < m_a - \varepsilon\right\} \le \exp\left\{-n\left[\theta\varepsilon - O(\theta^2)\right]\right\}. \tag{A.17}$$

当 $\theta > 0$ 足够小时，系数 $\theta\varepsilon - O(\theta^2)$ 严格为正。将其记为 c_ε，即可证明方程（A.9）。

至此完成了对性质 2 和 4 的验证。结合第 4 节中为最优 Beta 提议分布建立的性质 1 和 3，这些估计证明了第 3 节中推导的有限样本大偏差速率。

附录 B. Beta剖面的最优性

我们验证第 4 节中构造的剖面在累积约束下使相对熵最小。

设 h 表示所提出的最优密度之一，其累积分布为 $H(t) = \int_0^t h(s)\,ds$，并定义

$$\varphi(t) = \log\frac{h(t)}{q_x(t)}. \tag{B.1}$$

对于任意其他可容许密度 g，其累积分布为 G，相对熵恒等式

$$D_{\mathrm{KL}}(g\|q_x) - D_{\mathrm{KL}}(h\|q_x) = D_{\mathrm{KL}}(g\|h) + \int_0^x [g(t) - h(t)]\varphi(t)\,dt \tag{B.2}$$

展开对数即可得出。右侧第一项非负。因此，我们只需证明第二项也非负。

所提出的 Beta剖面包含两类区间。在自由区间上，h与 q_x成正比，因此 φ 为常数。在固定区间上，$h = 1/x$，因此

$$\varphi'(t) = -\frac{f'(t)}{f(t)} \le 0, \tag{B.3}$$

因为原始密度在每个固定区间上都是非递减的。

由于 g 和 h 都已归一化，$G(0) - H(0) = G(x) - H(x) = 0$。分部积分可得

$$\int_0^x (g-h)\varphi\,dt = -\int_0^x (G-H)\varphi'(t)\,dt. \tag{B.4}$$

内部连接点处的边界项相互抵消，因为 φ 是连续的。在自由区间上，$\varphi' = 0$。在固定区间上，$H(t) = t/x$，而可容许性

要求 $G(t) \geq t/x$。因此，在 φ' 可能非零的任何地方都有 $G - H \geq 0$。由于在这些区间上$\varphi' \leq 0$，我们得出

$$\int_0^x (g - h)\varphi \, dt \geq 0. \tag{B.5}$$

方程 (B.2) 现在表明

$$D_{\mathrm{KL}}(g \| q_x) \geq D_{\mathrm{KL}}(h \| q_x). \tag{B.6}$$

因此，所有提议的剖面都是全局熵最小化解。相对熵的严格凸性保证了除零测集外的唯一性。

附录 *B.1.* 似然平均值相等

我们还验证第 3 节的性质 3。在有限端点 a 处，令 h_a 为最优提议分布，并记

$$\ell_a(t) = \log \frac{q_a(t)}{h_a(t)}. \tag{B.7}$$

在固定区间上，$h_a = 1/a$，因此空间似然平均值与统计似然平均值逐点相等。在自由区间 $(u, \ v)$ 上，似然比 ℓ_a 为常数，记为 ℓ_0。此外，该区间的端点位于传播边界上：

$$H_a(u) = \frac{u}{a}, \qquad H_a(v) = \frac{v}{a}. \tag{B.8}$$

当 $u = 0$ 或 $v = a$ 时，这些恒等式也成立。因此，

$$\int_u^v h_a(t) \, dt = \frac{v - u}{a}, \tag{B.9}$$

因此

$$\int_u^v h_a(t)\ell_a(t) \, dt = \ell_0 \frac{v - u}{a} = \frac{1}{a} \int_u^v \ell_a(t) \, dt. \tag{B.10}$$

将所有自由区间和钉扎区间的贡献相加，得到

$$\int_0^a h_a(t)\ell_a(t) \, dt = \frac{1}{a} \int_0^a \ell_a(t) \, dt. \tag{B.11}$$

这就证明了性质 3。该等式依赖于最优提议分布的分段结构，对于一般的测度变换并不成立。

附录 C. Beta剖面的匹配点与可行性

我们确立了为钟形和 U 形 Beta 密度引入的匹配点的性质。

附录 *C.1*。钟形密度

假设 $\alpha > 1$ 和 $\beta > 1$，并令

$$m = \frac{\alpha - 1}{\alpha + \beta - 2} \tag{C.1}$$

为众数。对于 $x > m$，定义

$$\Phi_x(c) = F(x) - F(c) - (x - c)f(c). \tag{C.2}$$

匹配条件 (80) 等价于 $\Phi_x(c) = 0$。由于 $f(0) = 0$,

$$\Phi_x(0) = F(x) > 0. \tag{C.3}$$

此外,

$$\Phi_x(m) = \int_m^x [f(t) - f(m)]\, dt < 0, \tag{C.4}$$

因为 f 在其众数右侧严格递减。求导得到

$$\Phi_x'(c) = -(x - c)f'(c). \tag{C.5}$$

由于对于 $0 < c < m$, $f'(c) > 0$，函数 Φ_x 在此区间上严格递减。因此，恰好存在一个根 $c_x \in (0, m)$。

接下来，我们验证相应的最优剖面满足传播约束。

对于 $t \le c_x$，剖面是均匀的，因此

$$H(t) = \frac{t}{x}. \tag{C.6}$$

对于 $c_x < t < x$，定义

$$D(t) = H(t) - \frac{t}{x}. \tag{C.7}$$

匹配条件意味着

$$D(c_x) = D(x) = 0. \tag{C.8}$$

此外,

$$D'(t) = \frac{f(t) - f(c_x)}{x f(c_x)}. \tag{C.9}$$

由于 f 是钟形的且 $c_x < m$，差值 $f(t) - f(c_x)$ 起初为正；随着 t 从 c_x 增至 x，最终变为负值。因此， $D(t)$ 先增大，随后减小至零，但不会变为负值。由此可得,

$$H(t) \ge \frac{t}{x} \tag{C.10}$$

贯穿整个区间。

对于 $x \le m$，完全钉扎的剖面显然是可接受的。

附录 *C.2. U*形密度

设 $\alpha < 1$ 和 $\beta < 1$，并令

$$m = \frac{1-\alpha}{2-\alpha-\beta} \tag{C.11}$$

为密度的最小值。定义

$$\eta(t) = F(t) - tf(t). \tag{C.12}$$

其导数为

$$\eta'(t) = -tf'(t). \tag{C.13}$$

由于 f在 $(0, m)$ 上递减，在 $(m, 1)$ 上递增，因此 η在 $(0, m)$ 上递增，在 $(m, 1)$ 上递减。在原点附近，$f(t) \sim Ct^{\alpha-1}$，因此

$$\eta(t) \sim \left(\frac{1}{\alpha} - 1\right) tf(t) > 0. \tag{C.14}$$

当 $t \to 1$时，密度发散，而 $F(t) \to 1$，因此

$$\eta(t) \to -\infty. \tag{C.15}$$

因此存在唯一的 $d \in (m,1)$，满足

$$\eta(d) = 0, \qquad F(d) = df(d). \tag{C.16}$$

Now

$$\frac{d}{dt}\frac{F(t)}{t} = -\frac{\eta(t)}{t^2}. \tag{C.17}$$

因此，$F(t)/t$ 在 $(0, d)$ 上递减，并在 $(d, 1)$ 上递增。对于 $x \le d$，这意味着

$$\frac{F(t)}{F(x)} \ge \frac{t}{x}, \qquad 0 < t \le x. \tag{C.18}$$

因此，条件密度 q_x 是可容许的。

对于 $x > d$，考虑分段轮廓 (92)。在 $0 < t \le d$ 上，其累积分布为

$$H(t) = \frac{F(t)}{xf(d)}. \tag{C.19}$$

由于对于 $t \le d$，$F(t)/t \ge F(d)/d = f(d)$，我们得到

$$H(t) \ge \frac{t}{x}. \tag{C.20}$$

在 $t = d$ 处，等号成立，因为 $F(d) = df(d)$。对于 $d < t < x$，密度等于$1/x$，因此累积分布恰好沿着传播边界延续：

$$H(t) = \frac{t}{x}. \tag{C.21}$$

因此，对于每个 $x > d$，所提出的 U形轮廓都是可容许的。

结合附录 B 的最优性论证，这些结果确立了第4节中使用的四种剖面。

参考文献

[1] Granovetter, M.：《集体行为阈值模型》。Am. J. Sociol.**83**，1420–1443（1978）[2] Schelling, T.C.：《微观动机与宏观行为》。W. W. Norton，纽约（1978）[3] Kuran, T.：《火花与草原野火：意外的政治革命理论》。《公共选择》**61**，41–74（1989）[4] Bass, F.M.：耐用消费品新产品增长模型。Manage. Sci.**15**，215–227（1969）[5] Young, H.P.：异质群体中的创新扩散：传染、社会影响与社会学习。Am. Econ. Rev. **99**，1899–1924（2009）[6] Diamond, D.W., Dybvig, P.H.：银行挤兑、存款保险与流动性。J. Polit. Econ.**91**，401–419（1983）[7] Bikhchandani, S., Hirshleifer, D., Welch, I.：将时尚潮流、时尚、习俗和文化变迁视为信息级联的理论。J. Polit. Econ.**100**，992–1026（1992）[8] Watts, D.J.：随机网络上全球级联的简单模型。Proc. Natl. Acad. Sci. U.S.A. **99**，5766–5771（2002）[9] Castellano, C., Fortunato, S., Loreto, V.：社会动态的统计物理。Rev. Mod. Phys. **81**，591–646（2009）[10] Jusup, M., Holme, P., Kanazawa, K., Takayasu, M., Romić, I., Wang, Z., Geček, S., Lipić, T., Podobnik, B., Wang, L., Luo, W., Klanjšček, T., Fan, J., Boccaletti, S., Perc, M.：社会物理学。Phys. Rep.**948**，1–148（2022）[11] Granovetter, M., Soong, R.：扩散与集体行为的阈值模型。J. Math. Sociol. **9**，165–179（1983）[12] Touchette, H.：统计力学的大偏差方法。Phys. Rep.**478**，1–69（2009）[13] Fontanari, J.F.：有限总体中 Granovetter 阈值模型的精确解。《数学》**14**，2620（2026）[14] Konheim, A.G., Weiss, B.：一种占据规则及其应用。SIAM J. Appl. Math. **14**，1266–1274（1966）[15] Stanley, R.P.：《计数组合学》，第2卷。剑桥大学出版社，剑桥（1999）[16] Pitman, J.：《组合随机过程》。数学讲义，第1875卷。Springer，柏林（2006）

[17] Zeraoulia, R.：有限总体Granovetter模型中的临界振幅差异与边界-鞍点阻碍。预印本，ResearchGate，https://doi.org/10.13140/RG.2.2.12004.54406（2026）[18] Dembo, A.、Zeitouni, O.：《大偏差技术及其应用》，第2版。Springer，纽约（1998）[19] Hammersley, J.M.、Handscomb, D.C.：《蒙特卡罗方法》。Methuen，伦敦（1964）[20] Ferrenberg, A.M.、Swendsen, R.H.：研究相变的新蒙特卡罗技术。Physical Review Letters **61**，2635–2638（1988）。[21]Kumar, S.、Rosenberg, J.M.、Bouzida, D.、Swendsen, R.H.、Kollman, P.A.：生物分子自由能计算的加权直方图分析方法。I. 方法。Journal of Computational Chemistry **13**，1011–1021（1992）。