arXiv:2601.01642v1 [stat.ME] 4 Jan 2026

Wasserstein分布鲁棒稀有事件模拟

Dohyun Ahn*、Huiyi Chen

香港中文大学

Lewen Zheng

华为香港研究中心

2026年1月

摘要

标准的稀有事件模拟技术要求精确指定分布，这限制了它们在存在分布不确定性时的有效性。为此，我们提出了一种新框架，用于估计面临此类分布模型风险的稀有事件概率。具体而言，我们关注最坏情形稀有事件概率的计算，这些概率定义为针对以特定名义分布为中心的Wasserstein歧义集所给出的分布鲁棒界。通过利用这一界的对偶刻画，我们提出了分布鲁棒重要性采样（DRIS），这是一种计算上可处理的方法，旨在大幅降低估计对偶分量时的方差。所提出的方法易于实现，且采样成本低。最重要的是，它实现了 消失相对误差——这是稀有事件模拟中极难达到的最强效率保证。数值研究证实，DRIS 的表现优于现有基准方法。

1. 引言

从管理金融尾部风险到预测极端气候事件，量化稀有事件发生的可能性对于系统稳定性和安全性至关重要 [Glasserman, 2003, Asmussen and Glynn, 2007,

*通讯作者，电子邮箱：dohyun.ahn@cuhk.edu.hk

Rubino and Tuffin，2009]。其基本数学任务是估计随机向量落入关键稀有事件集的概率。由于标准蒙特卡罗方法在此类任务中的计算效率低下，人们针对各种模型和问题开发了复杂的方差缩减技术，例如重要性抽样、条件蒙特卡洛、分裂法和分层抽样；参见 Glasserman 等人 [2000, 2008], Juneja and Shahabuddin[2002], Bassamboo 等人 [2008], Blanchet and Lam [2014],Bai 等人 [2022],Ahn和Zheng [2025],Deo and Murthy [2025]及其中的参考文献。

然而，一个重要的理论缺口依然存在：这些经典方法假设能够精确了解底层概率分布，因此容易受到模型设定错误的影响。在现实场景中，这类细致信息鲜少可得，尤其是在数据稀缺或含噪时，因而会产生分布不确定性。为克服这一局限，我们采用分布鲁棒方法进行稀有事件模拟。具体而言，我们着重于高效计算一族合理分布上的最坏情形稀有事件概率，并将其数学形式化为以名义分布模型为中心的 Wasserstein 球。据我们所知，这是首项针对存在分布模型风险的情形，引入高效蒙特卡洛方法来估计稀有事件概率的研究。

在开发模型不确定性下最坏情况期望的仿真方法方面，我们的方法与 Glasserman 和 Xu [2014] 以及 Blanchet 等人 [2017]的方法密切相关。前者提出了所谓的鲁棒蒙特卡洛方法，用于估计由相对熵和 α-散度定义的分布模糊集上的风险度量；后者则着重计算边缘分布固定、但依赖结构未知的两个随机向量的最坏情况期望。尽管这些研究在方法上有所进展，但它们均未专门针对稀有事件模拟进行方差缩减，因此，这些方法在此类情形中的有效性尚未得到证实。

具体而言，关于稀有事件的分布鲁棒性，现有文献主要采用基于优化或基于极值理论的方法，而非仿真方法；例如，Lam 和 Mottet [2017], Blanchet 等人 [2020]和 Bai 等人[2023]。与此同时，Huang 等人 [2023] 近期的一项研究利用随机游走尾概率

分析稀有事件概率对尾部不确定性的脆弱性，并指出与轻尾情形相比，重尾情形对模型设定错误更为敏感。相比之下，我们着重于仿真，并提出一种方差缩减技术来估计最坏情形稀有事件概率，从而弥补这一空白。

具体而言，本文提出了一种新颖的<style id='1'>分布鲁棒重要性抽样（DRIS）方法，用于估计上述凸目标集的最坏情形稀有事件概率。借助 Wasserstein 分布鲁棒优化中的一般对偶性结果，所关注的概率可以重构为名义分布下目标集邻域的概率。从计算角度来看，这一对偶重构需要分两步进行：首先估计该邻域，然后将其纳入最终的概率计算。由于这两个步骤都涉及稀有事件模拟，我们的 DRIS方法旨在通过一种统一、高效且易于实施的算法来满足这些要求。

最重要的是，我们证明了 DRIS估计量满足中心极限定理，并表现出<style id='1'>消失相对误差 （定理2和定理3）。这些主要理论结果建立在以下两方面之上：（i）结合 Vapnik–Chervonenkis型论证的经验过程理论；以及（ii）仿真中的稀有事件分析。值得强调的是，消失相对误差这一性质确保了随着目标事件变得越来越稀有，相对误差趋于零；可以说，这是稀有事件模拟中最高层次的效率，而以往研究很少能够实现这一点。

本文其余部分安排如下。第2节对主要问题进行建模。第3节回顾在最坏情形概率背景下Wasserstein分布鲁棒优化的强对偶性，并给出初步理论结果。第4节介绍所提出的 DRIS 程序，并确立其在稀有事件情形下的理论性能保证。第5节通过数值实验验证该算法的有效性。第6节总结全文。所有证明均放在附录中。

2. 问题表述

令 $\mathcal{P}$ 表示支撑于 n 维欧几里得空间上的所有概率分布的集合。则 P_0,$\mathsf{P} \in \mathcal{P}$ 之间的 2-瓦瑟斯坦距离定义为

$$\mathcal{W}_2(\mathsf{P}_0, \mathsf{P}) = \inf_{\pi \in \Pi(\mathsf{P}_0, \mathsf{P})} \left(\mathsf{E}_{(\mathbf{X}_0, \mathbf{X}) \sim \pi}\left[\|\mathbf{X}_0 - \mathbf{X}\|^2\right]\right)^{1/2},$$

其中， $\Pi(\mathsf{P}_0,\mathsf{P})$ 表示 P_0 和 P 的所有耦合集合，也就是边缘分布分别为 P_0 和 P 的所有联合分布集合。因此，以名义分布 P_0 为中心、半径为 $\delta > 0$ 的 2-Wasserstein 球定义为

$$\mathcal{B}_\delta(\mathsf{P}_0) = \{\mathsf{P} \in \mathcal{P} : \mathcal{W}_2(\mathsf{P}_0, \mathsf{P}) \leq \delta\}.$$

在本文中，我们研究如下定义的最坏情况概率的估计：

$$p_* = \sup_{\mathsf{P} \in \mathcal{B}_\delta(\mathsf{P}_0)} \mathsf{P}(\mathbf{X} \in \mathcal{E}), \tag{1}$$

其中， $\delta \in (0, \infty)$ 是一个固定常数， $\mathcal{E}$ 是一个非空、满维、闭且凸的集合，且不包含原点； P_0 是 n 维标准正态分布。该量对应于 Wasserstein 分布鲁棒优化中的一种内层最坏情况问题，近期文献对此给予了广泛关注 [Zhang 等， 2025]。尽管我们聚焦于高斯名义分布，所提出的方法也可自然推广至其他多元椭圆族。鉴于高斯设定在运筹学与管理科学文献中十分常见，我们优先研究这一设定；在这些文献中，关键指标往往对应于由标准正态分布决定的稀有事件概率 [Bucklew, 2004，第 9] 章。以下是金融领域中的一个例子：

示例 1. 根据 *Glasserman* 等人 *[2000],* ，在时间区间 $[t, t + \mathrm{d}t]$ 内，欧式看涨/看跌期权投资组合的损失可近似表示为

$$L \coloneqq V(\mathbf{S}_t, t) - V(\mathbf{S}_t + \mathrm{d}\mathbf{S}, t + \mathrm{d}t) \approx -\frac{\partial V}{\partial t}\mathrm{d}t - \Delta^\top \mathrm{d}\mathbf{S} - \frac{1}{2}\mathrm{d}\mathbf{S}^\top \Gamma \mathrm{d}\mathbf{S} =: \tilde{L},$$

其中， $\mathbf{S}_t$ 和 $V(\mathbf{S}_t, t)$ 分别表示 n 风险因子和投资组合价值， $\mathrm{d}\mathbf{S} = \mathbf{S}_{t+\mathrm{d}t} - \mathbf{S}_t$, $\Delta = \nabla_{\mathbf{S}} V^\top$, $\Gamma = \nabla^2_{\mathbf{S}} V$，且“≈”由 *delta-gamma*近似成立。如果

$\tilde{\mathrm{d}}\mathbf{S} = \mathbf{S}_{t+\mathrm{d}t} - \mathbf{S}_t$ 服从多元正态分布，且该近似是精确的（即 $L = L$），

$$\mathsf{P}(L > \ell) = \mathsf{P}\left(a + \sum_{i=1}^{n} \left(b_i X_i + c_i X_i^2\right) > \ell\right),$$

对于损失阈值 $\ell > 0$、固定常数 $a, b_1, \ldots, b_n, c_1, \ldots, c_n$（其中 $c_1, \ldots, c_n \leq 0$，且 $X_1, \ldots, X_n$ 独立同分布$\sim \mathcal{N}(0,\ 1)$）。这一量通常用于定义投资组合风险度量；当 ℓ 较大时，它表示相互独立的标准正态变量属于某个凸稀有事件集的概率。

除了这个例子，许多连续时间随机模型（例如几何布朗运动和高斯马尔可夫过程）都可以通过欧拉格式模拟为标准正态变量的加权和。这不仅对金融建模至关重要，也有助于分析其他领域的系统稳定性：排队论中的重载近似依赖于由布朗运动驱动的扩散过程，而供应链管理中的需求过程通常建模为高斯随机游走。

值得注意的是，如果 $\mathbf{X}$ 服从 n 维非标准正态分布，则可以$\tilde{}$找到 $\boldsymbol{\mu} \in \mathbb{R}^n$ 和 $\boldsymbol{\Lambda} \in \mathbb{R}^{n \times m}$，使得 $n \geq m$，从而 $\mathbf{X}$ 与 $\boldsymbol{\mu} + \boldsymbol{\Lambda}\mathbf{X}$服从相同的分布，其中$\tilde{}\mathbf{X}$ 服从 m 维标准正态分布。因此，概率 $\mathsf{P}(\mathbf{X} \in \mathcal{E})$$\tilde{}$等于 $\mathbf{X}$ 属于另一个凸集的概率，该凸集由 $\{\mathbf{x} :\ \boldsymbol{\mu} + \boldsymbol{\Lambda}\mathbf{x}\ \in \mathcal{E}\}$ 给出。因此，对于所有高斯模型，只需将分析限制在标准正态分布即可。

不失一般性，我们假设 $\mathbf{x}^* := \arg\min_{\mathbf{x} \in \mathcal{E}} \|\mathbf{x}\|$ 位于 x_1 轴上。这可以通过适当的坐标旋转和变量重排来实现；由于标准正态分布在此类变换下保持不变，因此不会影响 (1)。此外，我们关注 $\{\mathbf{X} \in \mathcal{E}\}$ 是一个稀有事件的情形，即其发生概率接近于零。我们通过考虑一列由稀有度参数 $r > 0$ 索引的集合，从数学上研究这一情形:

$$\mathcal{E}_r = \left\{ \frac{r}{\|\mathbf{x}^*\|} \mathbf{x} : \mathbf{x} \in \mathcal{E} \right\}, \tag{2}$$

在这种情况下，$(r,\ 0,\ \ldots,\ 0) = \arg\min_{\mathbf{x} \in \mathcal{E}_r} \|\mathbf{x}\|$。因此，集合 $\mathcal{E}_r$ 随着$r \to \infty$远离原点，从而得到 $\lim_{r \to \infty} \mathsf{P}_0(\mathbf{X}\ \in \mathcal{E}_r) = 0$。

为了分析所提出估计量的效率，我们采用稀有事件模拟文献中广泛使用的以下性能准则 [，例如 Bassamboo 等人，2008，Nakayama 和 Tuffin， 2023]:

定义 1。令 q_r 表示一个关注量，满足当 $r \to \infty$时 $q_r \to 0$ 。假设无偏估计量 $Q_{N,\ r}$ 用于估计 q_r，由 N 个独立同分布样本构造，并且对任意 $r > 0$都满足渐近方差为 ξ_r^2的中心极限定理；也就是说，当 $N \to \infty$时，$\sqrt{N}(Q_{N,\ r} - q_r) \Rightarrow \mathcal{N}(0,\ \xi_r^2)$ 依分布收敛，其中 $\Rightarrow$表示依分布收敛，而 $\mathcal{N}(\gamma,\ \nu^2)$ 表示均值为γ 、方差为 ν^2的正态随机变量。于是，如果 $Q_{N,\ r}$ 满足以下条件，我们就称其具有消失相对误差：

$$\limsup_{r\to\infty} \frac{\xi_r}{q_r} = 0.$$

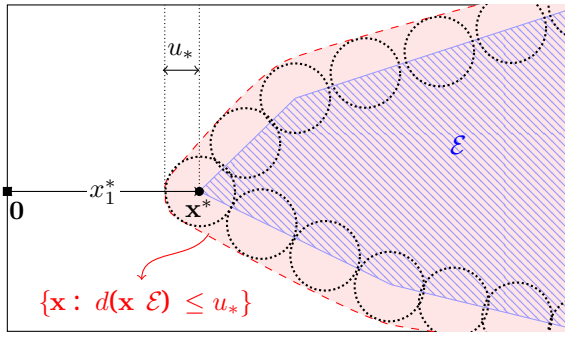
在稀有事件仿真中，消失相对误差通常被视为最高的效率标准。正如 Botev [2017], 所指出的，对于轻尾分布，蒙特卡洛估计量很少具有消失相对误差这一性质。该性质保证，在样本量固定且较大的情况下，随着目标事件变得越来越罕见，相应估计量的精度会提高。

3. 预备知识

本节回顾了针对式 (1) 中目标量的一个强对偶性结果，并介绍了初步的理论分析。这两者对于使问题易于处理以及推动第4节的主要分析都起着至关重要的作用。在深入讨论细节之前，我们先简要介绍全文使用的符号约定。我们用 E_0 表示在名义分布 P_0, 下的期望，并用 $d(\mathbf{x}\ \mathcal{S}) = \min_{\mathbf{y}\in\mathcal{S}} \|\mathbf{x} - \mathbf{y}\|$ 表示点 $\mathbf{x} \in \mathbb{R}^n$ 与集合 $\mathcal{S} \subset \mathbb{R}^n$ 之间的距离。此外，为简洁起见，对于任意函数 g 和任意事件 $\mathcal{A}$，我们记 $\mathrm{E}_0[g(\mathbf{X}); \mathcal{A}] \coloneqq \mathrm{E}_0[g(\mathbf{X})\mathbb{1}\{\mathcal{A}\}]$ 。

(1) 的强对偶性。(1) 中的优化问题是无限维的，因此无法直接求解。幸运的是，关于 Wasserstein 分布鲁棒优化的既有文献结果表明，(1) 的对偶形式在计算上可处理。我们在本研究框架下重述这些结果的一种形式，并讨论其对稀有事件模拟的影响。

图1： 基于对偶性结果，目标集与其膨胀版本之间关系的图示



引理*1*（*Blanchet* 和 *Murthy* 的定理 *2* [2019]）。设 $h(u) = E_0[d(X\ \mathcal{E})^2;\ d(X\ \mathcal{E}) \leq u]$ 且 $p(u) = P_0(d(X\ \mathcal{E}) \leq u)$。则，式 *(1)* 中的概率 p_* 等于 $p(u_*)$，其中 $u_* = h^{-1}(\delta^2)$。

这一对偶性结果的意义在于，它将最坏情况概率 p_* 表示为在名义分布 P_0, 下，目标事件的一个膨胀后的超集的概率，该超集由 $\{\mathbf{x}: d(\mathbf{x}\ \mathcal{E}) \leq u_*\}$给出。图1展示了目标集与其膨胀版本之间的联系：蓝色斜线区域表示目标集 $\mathcal{E}$，红色阴影区域则对应其膨胀版本 $\{\mathbf{x}: d(\mathbf{x}\ \mathcal{E}) \leq u_*\}$。虚线圆表示半径为 u_*的圆；以 $\mathcal{E}$ 中所有点为圆心的这类圆的并集刻画了膨胀后的超集。根据第2节中的假设，$\mathbf{x}^*$ 位于 x_1轴上，因此其到原点的距离为 x_1^*。

由于引理1对任意集合 $\mathcal{E}$ 均成立，因此该引理中的函数 $h(\cdot)$ 和数值 u_* 对式 (2) 中的集合序列 $\{\mathcal{E}_r\}_{r>0}$ 也有类似定义，具体如下：对于 $r > 0$ 和 δ，$u \geq 0$，令

$$h_r(u) = \mathsf{E}_0[d(\mathbf{X}, \mathcal{E}_r)^2; d(\mathbf{X}, \mathcal{E}_r) \leq u] \quad \text{and} \quad u_r = h_r^{-1}(\delta^2).$$

于是，由上述引理可得

$$p_r := \sup_{\mathsf{P} \in \mathcal{B}_\delta(\mathsf{P}_0)} \mathsf{P}(\mathbf{X} \in \mathcal{E}_r) = \mathsf{P}_0(d(\mathbf{X}, \mathcal{E}_r) \leq u_r). \tag{3}$$

尽管 u_r 和 p_r 取决于 2-Wasserstein球的半径 δ，但记号中省略了这种依赖关系。

初步理论结果。鉴于我们所处的稀有事件区域中，r 趋于 ∞，我们分析当 r 增大时，u_r 和 p_r 的行为。首先，以下引理描述了 u_r的行为：

引理 2 （ u_r的渐近行为）。 对于任意 δ、 $M > 0$，存在 $r_0 > 0$ ，使得对于所有$r \geq r_0$，

$$M < r - u_r < \bar{\Phi}^{-1}\left(\frac{\delta^2}{r^2}\right), \tag{4}$$

¯其中， $\Phi(\cdot)$ 表示标准正态互补累积分布函数。

注意到， $r - u_r$ 表示原点与 $\mathcal{E}_r$的膨胀版本之间的距离。因此，根据（4）中的第一个不等式，引理 2 确认，随着 r 增加，膨胀后的超集会远离原点，这表明 p_r 在（3）中的再次是一个稀有事件概率。这促使我们开发一种高效的稀有事件模拟算法来估计该概率。

此外，如附录 A 所示， $\bar{\Phi}^{-1}(\delta^2/r^2)$ 在（4）中随 $r \to \infty$以次线性速率增长。因此， （4）中的第二个不等式意味着，该距离以次线性速率发散。这表明，最坏情况概率 p_r 的衰减速度慢于名义概率 $\mathrm{P}_0(\mathbf{X} \in \mathcal{E}_r)$的指数速率。我们将在下面的定理中形式化这一观察结果，该定理刻画了 p_r 随 $r \to \infty$变化时的渐近下界。

定理 1 (p_r的渐近行为)。 对任意 $\delta > 0$，$\liminf_{r\to\infty} r^2 p_r \geq \delta^2$。

根据该定理，若要在p_r 的估计中达到消失相对误差（定义 1），则需要构造一个基于 N个样本的无偏估计量，其渐近方差以快于 r^{-4}的速率衰减。在下一节中，我们提出一种满足该条件的新型重要性抽样估计量。

4. 主要算法与结果

引理1允许我们通过两个步骤计算最坏情况概率 p_*： (a) 求解 $h(u) = \delta^2$以获得 u_* ， 以及 (b) 评估 $p(u_*)$。这两项任务都涉及在定义于形如 $\{\mathbf{x} : \ d(\mathbf{x}, \mathcal{E}) \leq u\}$ 的稀有事件集上的名义分布P_0 下估计期望（见第3节）。因此，在本节中，我们提出一种全面且易于处理的算法，用于解决这两个稀有事件估计步骤，并证明该算法能够实现消失相对误差。

4.1. DRIS算法

对于上述任务，在稀有事件集 $\{\mathbf{x}: \ d(\mathbf{x}\ \mathcal{E}) \leq u\}$附近对$\mathbf{X}$ 进行抽样，对于任何可行的 u而言都是必不可少的。我们将 X_1 确定为所述稀有事件的主要驱动因素，因为对于所有 u，$\{\mathbf{x}: \ d(\mathbf{x}\ \mathcal{E}) \leq u\} \subseteq \{\mathbf{x}: \ x_1 \geq x_1^* - u\}$ 均成立。此外，稀有事件集 $\{\mathbf{x}: \ d(\mathbf{x}\ \mathcal{E}) \leq u\}$是两个凸集 $\mathcal{E}$ 和 $\{\mathbf{x}: \ \|\mathbf{x}\| \leq u\}$的闵可夫斯基和，因此也为凸集。因此，受Ahn和Zheng [2023],中的条件重要性抽样方法启发，我们的重要性抽样方法包括：(a) 通过 $X_1 = x_1^* - u + Y/(x_1^* - u)$生成 X_1，其中 Y从标准指数分布中抽取；以及 (b) 从标准正态分布中对$(X_2, \ldots, X_n)$进行抽样。

接着，我们定义$\mathbf{Z} = (Y, \ X_2, \ \cdots, \ X_n)^\top$，并用 E 表示关于其分布的期望。我们还定义一个变换 $f_u: \mathbb{R}^n \to \mathbb{R}^n$，其定义如下：

$$f_u(\mathbf{z}) = \left(x_1^* - u + \frac{z_1}{x_1^* - u}, z_2, \cdots, z_d \right)^\top, \tag{5}$$

该变换将$\mathbf{Z}$ 映射到$\mathbf{X}$。最后，令

$$L_u(\mathbf{z}) := \frac{\exp(-z_1^2/(2(x_1^* - u)^2) - (x_1^* - u)^2/2)}{(x_1^* - u)\sqrt{2\pi}} \mathbb{1}\{z_1 \geq 0\} \tag{6}$$

表示与我们重要性抽样方法相关的似然比。在此设定下，很容易看出

$$\begin{cases} h(u) = \mathsf{E}[d(f_u(\mathbf{Z}), \mathcal{E})^2 \mathbb{1}\{d(f_u(\mathbf{Z}), \mathcal{E}) \leq u\} L_u(\mathbf{Z})]; \\ p(u) = \mathsf{E}[\mathbb{1}\{d(f_u(\mathbf{Z}), \mathcal{E}) \leq u\} L_u(\mathbf{Z})]. \end{cases}$$

这构成了 $h(u)$ 和 $p(u)$ 的无偏估计量，并使我们能够针对 p_* 制定以下估计步骤：

(i) 取$\mathbf{Z}$的 N 个独立同分布副本，记为 $\{\mathbf{Z}_i\}_{i=1}^n$;

(ii) 令 $H(\cdot, u) := d(f_u(\cdot)\ \mathcal{E})^2 \mathbb{1}\{d(f_u(\cdot)\ \mathcal{E}) \leq u\} L_u(\cdot)$，对于 $u \geq 0$，并将 $h(\cdot)$ 的估计定义为

$$\widehat{h}_N(u) = \frac{1}{N} \sum_{i=1}^N H(\mathbf{Z}_i, u) \ \text{ for } u \geq 0; \tag{7}$$

(iii) 计算估计值 $\widehat{u}_N := \inf\{u: \ \widehat{h}_N(u) \ > \delta^2\}$ for u_*;

算法 1： 分布鲁棒重要性采样（DRIS）

输入：N、x_1^* 和 δ 从标准指数分布中生成 N 个样本 $\{y_i\}_{i=1}^N$ of Y 从 $(n-1)$ 维标准正态分布中抽取 N 个样本 $\{(x_{2,i},\ldots,x_{n,i})\}_{i=1}^N$ of $(X_2,\ldots,X_n)$ 对于 $i=1,\ldots,N$，设 $\mathbf{z}_i=(z_{1,i},\ldots,z_{n,i})$，其中 $z_{1,i}=y_i$ and $z_{j,i}=x_{j,i}$ for $j=2,\ldots,n$ 对任意 $u\geq 0$，设 $h_N(u)$ $=N^{-1}\sum_{i=1}^N d(f_u(\mathbf{z}_i)\ \mathcal{E})^2\mathbb{1}\{d(f_u(\mathbf{z}_i)\ \mathcal{E})\ \leq u\}L_u(\mathbf{z}_i)$，其中 $f_u(\cdot)$ 和 $L_u(\cdot)$ 分别如 (5) 和 (6) 中所定义 通过（确定性）求根过程求得 $u_N=\inf\{u:\ h_N(u)\ >\delta^2\}$ 返回：$p_N=N^{-1}\sum_{i=1}^N\mathbb{1}\{d(f_{u_N}(\mathbf{z}_i)\ \mathcal{E})\ \leq u_N\}L_{u_N}(\mathbf{z}_i)$

(iv) 令 $P(\cdot,\ u):=\mathbb{1}\{d(f_u(\cdot)\ \mathcal{E})\ \leq u\}L_u(\cdot)$，并将 $p(u)$ 的估计量定义为

$$\widehat{p}_N(u)=\frac{1}{N}\sum_{i=1}^N P(\mathbf{Z}_i,u)\ \text{ for } u\geq 0; \tag{8}$$

(v) 通过计算 $\widehat{p}_N(\widehat{u}_N$ 来评估最坏情况概率 p_* 的估计量).

我们将这一方法及其估计量 $\widehat{p}_N(\widehat{u}_N)$ 分别称为分布鲁棒重要性采样（*DRIS*）和 DRIS估计量。我们在算法 1 中详细说明其过程。需要强调的是，尽管步骤 (iii) 涉及求根过程，但无需额外的采样成本，这与通常和重要性采样结合使用的求根过程不同 [He 等人，2024]。此外，DRIS方法的实现计算成本很低：尽管其中涉及求根过程，但该算法避免了其他环节中的高成本操作。具体而言，我们的抽样分布（即指数分布和正态分布）易于模拟，从而确保较低的采样成本。

4.2. DRIS 的效率

现在我们证明，所提出的方法具有强大的理论性能保证，满足定义 1 中的效率准则。为此，我们首先在以下结果中刻画 DRIS估计量 $\widehat{p}_N(\widehat{u}_N)$ 的中心极限定理：

定理 2 (中心极限定理)。 假设存在 $u_L, u_U \in (0, x_1^*)$，使得对所有充分大的 N， $u_*, \widehat{u}_N \in [u_L, u_U]$。那么，

$$\sqrt{N}(\widehat{p}_N(\widehat{u}_N) - p_*) \Rightarrow \mathcal{N}\left(0, \mathsf{Var}\left(P(\mathbf{Z}, u_*) - \frac{H(\mathbf{Z}, u_*)}{u_*^2}\right)\right) \quad as\ N \to \infty. \tag{9}$$

可以直接验证，在我们的渐近 regime 下，对于集合序列 $\{\mathcal{E}_r\}_{r>0}$，上述中心极限定理成立。具体而言，对所有 $r > 0$，式 (3) 中针对 p_r 的 DRIS估计量具有如下渐近方差

$$\sigma_r^2 := \mathsf{Var}\left(\mathbb{1}\{d(f_{u_r}(\mathbf{Z}), \mathcal{E}_r) \le u_r\} L_{u_r}(\mathbf{Z}) \left(1 - \frac{d(f_{u_r}(\mathbf{Z}), \mathcal{E}_r)^2}{u_r^2}\right)\right).$$

基于这一渐近方差，以下定理给出了本文的主要发现：对我们的 DRIS估计量的渐近效率进行刻画。该结果表明，使用固定的一组样本来估计 $h(\cdot)$ 和 $p(\cdot)$ 的有效性。

定理 3 （消失相对误差）。 对于任意 $\delta > 0$，$\limsup_{r\to\infty} r^2(r - u_r)^2\sigma_r^2/p_r^2 < \infty$。

由于 $r - u_r \to \infty$ 当 $r \to \infty$ （引理 2）时，前述定理表明，DRIS估计量的相对误差渐近变化速率至多为 $r^{-1}(r - u_r)^{-1}$ 当 $r \to \infty$，这意味着 DRIS估计量实现了消失相对误差。

5. 数值实验

在本节中，我们通过数值实验验证所提方法的性能。为了在数值上比较 DRIS方法与现有蒙特卡罗方法的应用效果，我们报告以下各项实验的两个性能指标：方差比（VR）和效率比（ER）。对于运行时间为 τ^{MC} 的粗蒙特卡罗估计量 Z^{MC} 和运行时间为 τ 的目标估计量 Z，我们定义 VR $:=$ $\mathsf{Var}(Z^{\mathsf{MC}})/\mathsf{Var}(Z)$ 和 ER $:=$ VR $\times\, \tau^{\mathsf{MC}}/\tau$。我们还报告估计量 Z 在 95% 置信水平下的相对误差，其定义为 $1.96\sqrt{\mathsf{Var}(Z)}/\mathrm{E}[Z]$。虽然 ER 通常被视为更全面的效率衡量指标，但计算时间容易受到硬件性能和实现细节的影响；因此，我们将 VR 作为一项重要的补充指标予以报告。

5.1. 实验设置

我们在数值实验中采用以下两个示例。

一个玩具示例。 我们首先考虑一个简单的二维设置，其中目标集由 $\mathcal{E}_r = \{\mathbf{x} \in \mathbb{R}^2 : x_1 - 5x_2 \geq r, x_1 + 5x_2 \geq r\}$ 给出，2-Wasserstein球的半径设为$\delta = 0.001$。我们使用样本量 10^7 估计 $\widehat{u}_N$ 和 $\widehat{p}_N(\widehat{u}_N)$，并将整个过程重复100次，以计算每种算法的平均运行时间和方差。据我们所知，目前还没有专门用于估计Wasserstein分布鲁棒稀有事件概率的仿真方法。因此，我们将DRIS方法的表现与粗蒙特卡洛（MC）和经典指数扭曲（ET）方案进行比较；后两者按照与DRIS方法中（7）和（8）类似的方式，用于估计 $h(\cdot)$和 $p(\cdot)$。

投资组合损失概率。接下来，我们重新考察第2节中的示例 1，以估计投资组合损失概率。我们构建一个由 $n = 5$ 个不相关基础资产组成的投资组合，并采用 Glasserman 等人 [2000] 中的参数设置。具体而言，我们假设每年有 250 个交易日、无风险利率为 5%，且 $\mathrm{d}t = 0.04$。每项基础资产的初始值为 $S_0 = 100$，波动率为 $\sigma = 0.3$。对于每项资产，投资组合均持有 10 份平值看涨期权和 5 份平值看跌期权的多头头寸。所有期权的期限均为半年期。所有情况下，损失阈值ℓ 均设为 120。为与我们的稀有事件设定保持一致，对于不同的 r 值，我们将风险因子 $\mathbf{X}$缩放 r^{-1}。最后，我们设置 $\delta = 0.01$ ，并采用与前述玩具示例相同的基准方法、样本量和宏观重复实验次数。

5.2. 数值结果总结

表1和表2报告了第5.1节所述两个示例中 u_r 和 p_r 的估计值以及各算法的运行时间，同时给出了相应的95%相对误差、方差比和效率比。在我们考虑的所有情形下，所提出的方法都完全优于两种基准方法，展现出更大的方差缩减和更高的效率。DRIS与另外两种方法之间显著的性能差距，且该差距随 r 增大而扩大，验证了我们的理论结果。尽管ET在我们的数值实验中表现出较强的竞争力，但它在这些问题上的表现缺乏

理论依据；更重要的是，DRIS始终能取得更优的结果。与蒙特卡洛相比，ET和DRIS的运行时间更长，这是因为这些算法中嵌入的求根过程需要转换样本，并求解依赖于所评估的 u的距离；相比之下，粗略蒙特卡洛算法中的样本保持不变。

表1： 第5.1节玩具示例的数值结果

方法	r	u_r （95%相对误差）	p_r （95%相对误差）	时间（秒）	VR	ER
MC	2	0.0027（1.62%）	2.40×10^{-3} (1.13%)	11	–	–
	3	0.0141 (3.29%)	2.39×10^{-4} (2.97%)	12	–	–
	4	0.0931（8.11%）	2.36×10^{-5} (7.80%)	12	–	–
	5	0.5245（12.43%）	3.10×10^{-6} (19.91%)	13	–	–
ET	2	0.0027 (0.42%)	2.40×10^{-3} (0。29%)	144	16	1.3
	3	0.0147 (0.30%)	2.40×10^{-4} (0.27%)	151	125	9.8
	4	0.0965 (0.18%)	2.31×10^{-5} (0.15%)	144	2，600	225
	5	0.5163 (0.08%)	3.08×10^{-6} (0.07%)	112	78，036	9，107
DRIS	2	0.0027 (0.24%)	2.41×10^{-3} (0.16%)	149	48	3.7
	3	0。0146（0。15%）	2.40×10^{-4} (0.13%)	148	559	45
	4	0.0965（0.08%）	2.31×10^{-5} (0.08%)	163	10，108	772
	5	0.5162 (0.04%)	3.08×10^{-6} (0.04%)	116	220，943	24，978

6. 总结

本文研究如何在分布模型风险下高效估计稀有事件概率。借助 Wasserstein 分布鲁棒优化中的强对偶性结果，我们构造了一种新颖且计算上可处理的重要性抽样方法，称为 DRIS，可显著降低估计上述概率时的方差。我们严格证明，所提出的 DRIS 估计量能够实现消失相对误差，这被视为稀有事件模拟领域中最强的效率概念。所有数值实验均支持这些理论结论。

表 2： 示例 1 中投资组合损失概率估计的数值结果

方法	r	u_r (95% 相对误差)	p_r (95% 相对误差)	时间（秒）	VR	ER
MC	2	1.42 (1.865%)	1.05×10^{-4} (2.499%)	7	–	–
	3	8.46 (2.278%)	1.37×10^{-5} (3.482%)	7	–	–
	4	24.50 (2.512%)	4.40×10^{-6} (3.362%)	7	–	–
ET	2	1.40 (0.056%)	1.05×10^{-4} (0.042%)	48	3，615	526
	3	8.60 (0.023%)	1.35×10^{-5} (0.016%)	50	48，120	6，620
	4	24。73 (0。013%)	4.39×10^{-6} (0.009%)	51	145，230	19，182
DRIS	2	1.40 (0.024%)	1.05×10^{-4} (0.034%)	53	5，269	691
	3	8.60（0.009%）	1.35×10^{-5} (0.013%)	54	71，806	9，212
	4	24.73 (0.004%)	4.39×10^{-6} (0.007%)	51	227，647	30，143

作为首个专门用于估计分布不确定性下稀有事件概率的方法框架，我们提出的方法依赖于特定的建模假设，也由此带来了若干值得探索的未来研究方向。首先，我们将凸集作为目标事件，这一选择受到相关文献中若干实例的启发。尽管如此，将我们的方法推广到非凸目标集虽颇具挑战，却能大幅拓展其实践适用性。其次，我们仅关注高斯名义分布的情形。如前所述，该框架可扩展至其他椭圆名义分布，但这些情形下的理论性能仍有待验证。探索非椭圆名义分布的情形也将很有意义。最后，为确保理论分析的可处理性，我们使用 2-Wasserstein球定义分布不确定集。放宽这一限制将是一个有前景的研究方向，因为引理1中的对偶性结果可推广至更广泛的不确定集，包括带有 $p \geq 1$ 的 p-Wasserstein 球。

A. 理论结果的证明

引理 *2* 的证明。 固定 $K > M > 0$。反设，对于某个 $r > M$， $u_r \geq r - M$。于是，我们观察到

$$\frac{\delta^2}{r^2} = \frac{h_r(u_r)}{r^2} \geq \mathsf{E}\left[\frac{d(\mathbf{X}, \mathcal{E}_r)^2}{r^2}; d(\mathbf{X}, \mathcal{E}_r) \leq r - M, \|\mathbf{X}\| \leq K\right]. \tag{10}$$

由于 $d(\cdot, \mathcal{E}_r)$ 是 1-利普希茨的，因此，对于任意满足 $\|\mathbf{x}\| \leq K$ 的$\mathbf{x}$，我们有 $d(\mathbf{x}, \mathcal{E}_r) \geq d(\mathbf{0}, \mathcal{E}_r) - \|\mathbf{x}\| \geq r - K$。这意味着 $\liminf_{r\to\infty} d(\mathbf{x}, \mathcal{E}_r)^2/r^2 = 1$。

固定$\mathbf{x} \in \mathbb{R}^n$，使得 $x_1 > M$。然后令 $t_r := (\|\mathbf{x}\|^2 - M^2)/(2rx_1 - 2rM) > 0$，经过直接计算可得 $\|\mathbf{x} - rt_r\mathbf{e}_1\| = rt_r - M$。因此，由于 $d(\cdot, \mathcal{E}_r)$ 是 1-利普希茨的，对于所有足够大的 r（满足 $t_r \in (0, 1)$），我们有 $d(\mathbf{x}, \mathcal{E}_r) \leq d(rt_r\mathbf{e}_1, \mathcal{E}_r) + \|\mathbf{x} - rt_r\mathbf{e}_1\| = r - rt_r + rt_r - M = r - M$。相应地，对（10）应用法图引理，我们得到

$$\liminf_{r\to\infty} \frac{\delta^2}{r^2} \geq \mathsf{E}_0\left[\liminf_{r\to\infty} \frac{d(\mathbf{X}, \mathcal{E}_r)^2}{r^2}; d(\mathbf{X}, \mathcal{E}_r) \leq r - M, \|\mathbf{X}\| \leq K\right] \geq \mathsf{P}(X_1 > M, \|\mathbf{X}\| \leq K) > 0.$$

这与 是一个常数这一事实矛盾。因此，对于所有充分大的 .

此外，显然有 $p_r = \mathsf{P}(d(\mathbf{X}, \mathcal{E}_r) \leq u_r) \leq \mathsf{P}(X_1 \geq r - u_r) = \Phi(r - u_r)$。因此，我们得到 $\delta^2 = h_r(u_r) = \mathsf{E}[d(\mathbf{X}, \mathcal{E}_r)^2$；对所有充分大的 r，均有 $d(\mathbf{X}, \mathcal{E}_r) \leq u_r] \leq u_r^2 p_r \leq r^2\bar{\Phi}(r - u_r)$。由此即得所需结果。

□

定理 *1* 的证明。根据标准正态分布米尔斯比的渐近行为，我们有 exp()，当.....时（例如，参见 Niewiadomska-Bugaj 和 Bartoszyński，）。这意味着 ()，当.....时。因此，令 ()，我们有 ()，当.....增大时。然后，将等式 (4) 两边除以，并令.....，得到 lim。此外，我们注意到

$$h_r(u_r) = \mathsf{E}_0[d(\mathbf{X}, \mathcal{E}_r)^2; d(\mathbf{X}, \mathcal{E}_r) \leq u_r] \leq u_r^2 \mathsf{P}_0(d(\mathbf{X}, \mathcal{E}_r) \leq u_r) = u_r^2 p_r.$$

因此，$\liminf_{r\to\infty} r^2 p_r \geq \delta^2/\lim_{r\to\infty}(u_r/r)^2 = \delta^2$。 □

定理 *2* 的证明。 我们分四步证明该结论。在此证明中，对于任意函数 $A\colon \mathbb{R}^n \to \mathbb{R}$，我们用 $\|\cdot\|_2$ theL^2 范数表示相对于抽样分布的范数，即 $\|A\|_2 = \sqrt{\mathsf{E}[A(\mathbf{Z})^2]}$ 。

$\widehat{\quad}$ $\widehat{\quad}$步骤 *1*： 的一致收敛 h_N。 在此步骤中，我们旨在证明式 (7) 中的 h_N 在 $\Theta := [u_L, u_U]$ 上一致收敛。由于每个 Donsker 类都满足一致大数定律 [（Van Der Vaart 和 Wellner，2023，第 130], 页），只需证明 $\mathcal{H} := \{H(\cdot,\ u) :\ u \in \Theta\}$ 是 Donsker 类。

我们定义两个函数类 $\mathcal{H}_1$和 $\mathcal{H}_2$如下： $\mathcal{H}_1 := \{\mathbf{z} \mapsto (d(f_u(\mathbf{z})\ \mathcal{E}) \wedge u_U)^2 L_u(\mathbf{z}) :\ u \in \Theta\}$和 $\mathcal{H}_2 := \{\mathbf{z} \mapsto \mathbb{1}\{d(f_u(\mathbf{z})\ \mathcal{E}) \le u\} :\ u \in \Theta\}$。我们注意到，对于任意 $u, v \in \Theta$,

$$
\begin{aligned}
&\left|(d(f_u(\mathbf{z}), \mathcal{E}) \wedge u_U)^2 L_u(\mathbf{z}) - (d(f_v(\mathbf{z}), \mathcal{E}) \wedge u_U)^2 L_v(\mathbf{z})\right| \\
&\le (d(f_u(\mathbf{z}), \mathcal{E}) \wedge u_U)^2 |L_u(\mathbf{z}) - L_v(\mathbf{z})| + |L_v(\mathbf{z})| \left|(d(f_u(\mathbf{z}), \mathcal{E}) \wedge u_U)^2 - (d(f_v(\mathbf{z}), \mathcal{E}) \wedge u_U)^2\right| \\
&\le u_U^2 |L_u(\mathbf{z}) - L_v(\mathbf{z})| + \bar{L} \left|(d(f_u(\mathbf{z}), \mathcal{E}) \wedge u_U)^2 - (d(f_v(\mathbf{z}), \mathcal{E}) \wedge u_U)^2\right| \\
&\le u_U^2 |L_u(\mathbf{z}) - L_v(\mathbf{z})| + 2u_U \bar{L} \|f_u(\mathbf{z}) - f_v(\mathbf{z})\|,
\end{aligned}
$$

其中，第一个不等式由三角不等式得出；第二个不等式成立，因为$\bar{L} := \sup_{\mathbf{z} \in \mathbb{R}^n, u \in \Theta} L_u(\mathbf{z}) < \infty$; 最后一个不等式显然成立，因为 $|a^2 - b^2| \le 2c|a - b|$for $a, b \in [0, c]$ and $c \ge 0$，且 $d(\cdot, \mathcal{E})$ 是 1-利普希茨函数。容易验证，存在一个多项式函数 G，满足 $u_U^2 |L_u(\mathbf{z}) - L_v(\mathbf{z})| + 2u_U \bar{L} \|f_u(\mathbf{z}) - f_v(\mathbf{z})\| \le G(\mathbf{z})|u - v|$，对所有$\mathbf{z} \in \mathbb{R}^n$ 和 $u, v \in \Theta$均成立。由于 $\|G\|_2 < \infty$ 和 Θ 是紧集，根据 Van Der Vaart 和 Wellner 的定理2.7.17和2.5.6， $\mathcal{H}_1$ 是 Donsker 类 [2023]。

给定一个集合族 $\mathcal{C}$，其 VC维数记为 $V(\mathcal{C})$，定义为满足 $|\{X \cap C :\ C \in \mathcal{C}\}| = 2^{|X|}$的最大集合 X 的基数。若函数类 $\mathcal{F}$ 的所有下图集 $\{\{(\mathbf{z}\ t) :\ t < f(\mathbf{z})\} :\ f \in \mathcal{F}\}$ 构成的集合族具有有限的 VC维数，则称该函数类为 VC类。假设对于某些 m 个点 $(\mathbf{z}1, t_1) \ldots, (\mathbf{z}_m, t_m)$，有$|\{\{(\mathbf{z}1, t_1) \ldots, (\mathbf{z}_m, t_m)\} \cap \{(\mathbf{z}\ t) :\ t < \mathbb{1}\{d(f_u(\mathbf{z})\ \mathcal{E}) \le u\}\} :\ u \in \Theta\}| = 2^m$，且$\in (0, \infty) \times \mathbb{R}^{n-1} \times \mathbb{R}$。由于条件 $t < \mathbb{1}\{d(f_u(\mathbf{z})\ \mathcal{E}) \le u\}$ 仅在 $t \in [0, 1)$时才是非平凡的，因此不失一般性，我们可以选取 $t_1 = \cdots = t_m = 0$。在这种情况下，下图集的打散条件等价于直接利用函数值打散点 $\mathbf{z}1, \ldots, \mathbf{z}_m$，即 $|\{(\mathbb{1}\{d(f_u(\mathbf{z}_1)\ \mathcal{E}) \le u\}, \ldots, \mathbb{1}\{d(f_u(\mathbf{z}_m)\ \mathcal{E}) \le u\}) :\ u \in \Theta\}| = 2^m$。

另一方面，根据附录 B 中的引理 3，集合 $\{u \in \Theta :\ d(f_u(\mathbf{z}_i)\ \mathcal{E}) \le u\}$ 在 Θ中至多由 2 个边界点确定。因此， Θ 中至多存在 $2m$ 个点，记为 $u_1, u_2, \ldots, u_{2m}$，使得 $u_L = u_0 \le u_1 \le \cdots \le u_{2m} \le u_{2m+1} = u_U$，且对于任意满足$i = 0$的 $u \in (u_i, u_{i+1})$，向量 $(\mathbb{1}\{d(f_u(\mathbf{z}_1)\ \mathcal{E})^2 \le u\}, \ldots, \mathbb{1}\{d(f_u(\mathbf{z}_m)\ \mathcal{E})^2 \le u\})$ 保持不变，其中包含$i = 0, \ldots, 2m$。因此， $|\{(\mathbb{1}\{d(f_u(\mathbf{z}_1)\ \mathcal{E}) \le u\}, \ldots, \mathbb{1}\{d(f_u(\mathbf{z}_m)\ \mathcal{E}) \le u\}) :\ u \in \Theta\}| \le 2m + 1$。合-

并上述打散条件可得 $2^m \leq 2m+1$。因此， m 必为有限集，从而证明 $\mathcal{H}_2$ 是一个 VC类。此外，$\mathcal{H}_2$ 一致有界于 1。因此，Van Der Vaart 和 Wellner 的定理 2.6.7 和 2.5.2 [2023] 蕴含 $\mathcal{H}_2$ 是 Donsker。

对所有 $x, y \in \mathbb{R}$，令 $\phi(x, y) = xy$。由于 $\mathcal{H}_1$ 和 $\mathcal{H}_2$ 一致有界且为 Donsker 类，并且 $\mathcal{H} \subset \phi \circ (\mathcal{H}_1, \mathcal{H}_2) := \{\mathbf{z} \mapsto \phi(g_1(\mathbf{z}), g_2(\mathbf{z})) : g_1 \in \mathcal{H}_1, g_2 \in \mathcal{H}_2\}$，根据 Van Der Vaart 和 Wellner 的推论2.10.15和定理2.10.1， $\mathcal{H}$ 也是 Donsker 类 [2023]。

步骤 2。 $\widehat{u}_N$ 的收敛性。由于 $h(\cdot)$ 是满足 $h(u_*) = \delta^2$ 的严格递增函数，因此对于任意 $\varepsilon > 0$，我们有 $c(\varepsilon) := \inf_{|u-u_*|>\varepsilon} |h(u) - \delta^2|/2 > 0$。固定 $\varepsilon > 0$。若 $\sup_{u\in\Theta} |h(u) - \widehat{h}_N(u)| \leq c(\varepsilon)$，则 $|h(\widehat{u}_N) - \delta^2| \leq \max\{\lim_{u\uparrow\widehat{u}_N} |h(u) - \widehat{h}_N(u)|, \lim_{u\downarrow\widehat{u}_N} |h(u) - \widehat{h}_N(u)|\} \leq c(\varepsilon)$，这意味着 $|\widehat{u}_N - u_*| \leq \varepsilon$。因此，$\mathrm{P}(\sup_{u\in\Theta} |h(u) - h_N(u)| \leq c(\varepsilon)) \leq \mathrm{P}(|\widehat{u}_N - u_*| \leq \varepsilon)$。根据步骤 1 中 h_N 的一致收敛性，$\lim_{N\to\infty} \mathrm{P}(\sup_{u\in\Theta} |h(u) - h_N(u)| \leq c(\varepsilon)) = 1$。因此，当 $N \to \infty$ 时，$\widehat{u}_N \to u_*$ 依概率收敛。

步骤 3。 $\widehat{u}_N$的渐近正态性。我们定义 $H_1(\mathbf{z}, u) = (d(f_u(\mathbf{z}), \mathcal{E}) \wedge u_U)^2 L_u(\mathbf{z})$，以及$H_2(\mathbf{z}, u) = \mathbb{1}\{d(f_u(\mathbf{z}), \mathcal{E}) \leq u\}$，这意味着对于 $\mathbf{z} \in (0, \infty) \times \mathbb{R}^{n-1}$和 $u \in \Theta$， $H(\mathbf{z}, u) = H_1(\mathbf{z}, u)H_2(\mathbf{z}, u)$。我们注意到，当且仅当 $f_{u_*}(\mathbf{z})$ 位于$\{\mathbf{z} : d(\mathbf{z}, \mathcal{E}) \leq u_*\}$的边界上时， $d(f_{u_*}(\mathbf{z}), \mathcal{E}) = u_*$。此外，由于 f_{u_*} 是可逆仿射变换，可以验证 $\mathrm{P}(d(f_{u_*}(\mathbf{Z}), \mathcal{E}) = u_*) = 0$。

固定样本空间中的 ω，使得 $d(f_{u_*}(\mathbf{Z}(\omega)), \mathcal{E}) \neq u_*$。那么，由于 $u \mapsto d(f_u(\mathbf{Z}(\omega)), \mathcal{E}) - u$ 连续，存在 $\delta > 0$，使得对于任意 $|u - u_*| < \delta$， $H_2(\mathbf{Z}(\omega), u) = H_2(\mathbf{Z}(\omega), u_*)$。因此，当 $u \to u_*$ 时，$H_2(\mathbf{Z}, u) \to H_2(\mathbf{Z}, u_*)$ 几乎必然成立。于是，根据 $H_1(\mathbf{z}, \cdot)$ 的连续性和连续映射定理， $\|H(\cdot, u) - H(\cdot, u_*)\|_2^2 = \|H_1(\cdot, u)H_2(\cdot, u) - H_1(\cdot, u_*)H_2(\cdot, u_*)\|_2^2 \to 0$在 $u \to u_*$ 时成立。我们还注意到，对于某个 $\delta > 0$， $\{H(\cdot, u) - H(\cdot, u_*) : |u - u_*| < \delta, u \in \Theta\}$ 是 Donsker 类，因为 $\mathcal{H}$ 是 Donsker 类，且根据 Van Der Vaart 和 Wellner 的定理 2.10.8 [2023]。

令 $\Psi_N(u) := \widehat{h}_N(u) - \delta^2$ 且 $\Psi(u) := h(u) - \delta^2$。于是，根据中心极限定理，我们有$\sqrt{N}(\Psi_N - \Psi)(u_*) = N^{-1/2} \sum_{i=1}^N (H(\mathbf{Z}_i, u_*) - \mathrm{E}[H(\mathbf{Z}, u_*)]) \Rightarrow \mathcal{N}(0, \mathrm{Var}(H(\mathbf{Z}, u_*)))$。此外，根据附录 B 中的引理 4，$\Psi'(u_*) = h'(u_*) \neq 0$。而且，由于 $H(\mathbf{z}, u)$ 一致有界，可以根据 $\widehat{u}_N$ 的定义验证 $\Psi_N(\widehat{u}_N) = o_P(N^{-1/2})$，且 $\mathbf{Z}_i$ 是连续的

分布。结合 Van Der Vaart 和 Wellner 的引理 3.3.5 和定理 3.3.1 [2023],，我们得出 $\sqrt{N}h'(u_*)(\widehat{u}_N - u_*) = -\sqrt{N}(\widehat{h}_N - h)(u_*) + o_P(1)$。

步骤 4. 估计量的渐近正态性。使用与步骤 1 至 3 相同的论证，可以证明，对于某个 $\delta > 0$, $\{P(\cdot, u) - P(\cdot, u_*) : |u - u_*| < \delta, u \in \Theta\}$ 是 Donsker，且 $\|P(\cdot, u) - P(\cdot, u_*)\|_2^2 \to 0$ 当 $u \to u_*$ 时成立。因此，再次应用 Van Der Vaart 和 Wellner 的引理 3.3.5 [2023]，可得 $\sqrt{N}(\widehat{p}_N(\widehat{u}_N) - p(\widehat{u}_N)) = \sqrt{N}(\widehat{p}_N(u_*) - p(u_*)) + o_P(1)$。由于 $p(\cdot)$在 u_*处可微，泰勒展开表明 $\sqrt{N}(p(\widehat{u}_N) - p(u_*)) = \sqrt{N}p'(u_*)(\widehat{u}_N - u_*) + o_P(\sqrt{N}|\widehat{u}_N - u_*|)$。结合这些结果、步骤 3 的结论以及附录 B 中的引理 4, 我们得到

$$\begin{aligned}\sqrt{N}(\widehat{p}_N(\widehat{u}_N) - p(u_*)) &= \sqrt{N}(\widehat{p}_N - p)(u_*) + \sqrt{N}p'(u_*)(\widehat{u}_N - u_*) + o_P(\sqrt{N}|\widehat{u}_N - u_*|) + o_P(1) \\ &= \sqrt{N}(\widehat{p}_N - p)(u_*) - \frac{\sqrt{N}}{u_*^2}(\widehat{h}_N - h)(u_*) + o_P(1),\end{aligned}$$

其中，最后一个等式成立是因为根据步骤 3，$\sqrt{N}(\widehat{u}_N - u_*)$依概率有界。因此，由中心极限定理和 Slutsky 定理，(9) 中的所需结果得证。 □

定理 *3* 的证明。 由于$\mathbf{x}^* = \arg\min_{\mathbf{x}\in\mathcal{E}_r} \|\mathbf{x}\|^2 = r\mathbf{e}_1$ 和 $x_1^* = r$，我们观察到

$$\begin{aligned}\sigma_r^2 &= \mathsf{Var}\left(\mathbb{1}\{d(f_{u_r}(\mathbf{Z}), \mathcal{E}_r) \le u_r\} L_{u_r}(\mathbf{Z})\left(1 - \frac{d(f_{u_r}(\mathbf{Z}), \mathcal{E}_r)^2}{u_r^2}\right)\right) \\ &\le \mathsf{E}\left[L_{u_r}(\mathbf{Z})^2\left(1 - \frac{d(f_{u_r}(\mathbf{Z}), \mathcal{E}_r)^2}{u_r^2}\right)^2 ; d(f_{u_r}(\mathbf{Z}), \mathcal{E}_r) \le u_r\right] \\ &= \mathsf{E}_0\left[\ell_{u_r}(X_1)\left(1 - \frac{d(\mathbf{X}, \mathcal{E}_r)^2}{u_r^2}\right)^2 ; d(\mathbf{X}, \mathcal{E}_r) \le u_r\right] \\ &\le \left(\mathsf{E}_0\big[\ell_{u_r}(X_1)^2 : d(\mathbf{X}, \mathcal{E}_r) \le u_r\big]\mathsf{E}_0\left[\left(1 - \frac{d(\mathbf{X}, \mathcal{E}_r)^2}{u_r^2}\right)^4 : d(\mathbf{X}, \mathcal{E}_r) \le u_r\right]\right)^{1/2},\end{aligned} \tag{11}$$

其中 $\ell_u(x) := e^{-x^2/2 + (r-u)(x-(r-u))}/((r-u)\sqrt{2\pi})\mathbb{1}\{x \ge r-u\}$，最后一个不等式由柯西–施瓦茨不等式得出。简单计算可得

$$\mathsf{E}_0\big[\ell_{u_r}(X_1)^2; d(\mathbf{X}, \mathcal{E}_r) \le u_r\big] \le \mathsf{E}_0[\ell_{u_r}(X_1)^2; X_1 \ge r - u_r] \le \frac{e^{-3(r-u_r)^2/2}}{(2\pi)^{3/2}(r-u_r)^3}. \tag{12}$$

此外，利用分部积分可得

$$\begin{aligned}
\mathsf{E}_0\left[\left(1-\frac{d(\mathbf{X},\mathcal{E}_r)^2}{u_r^2}\right)^4 : d(\mathbf{X},\mathcal{E}_r)\le u_r\right] &\le \frac{8}{u_r^2}\int_0^{u_r} t\left(1-\frac{t^2}{u_r^2}\right)^3 \mathsf{P}_0(d(\mathbf{X},\mathcal{E}_r)\le t)\mathrm{d}t \\
&\le \frac{8}{u_r^2}\int_0^{u_r} t\left(1-\frac{t^2}{u_r^2}\right)^3 \frac{e^{-(r-t)^2/2}}{\sqrt{2\pi}(r-t)}\mathrm{d}t \\
&\le \frac{64e^{-(r-u_r)^2/2}}{\sqrt{2\pi}(r-u_r)u_r^4}\int_0^{u_r}(u_r-t)^3 e^{-(r-u_r)(u_r-t)}\mathrm{d}t \qquad (13)\\
&\le \frac{64e^{-(r-u_r)^2/2}}{\sqrt{2\pi}(r-u_r)u_r^4}\int_0^{\infty} y^3 e^{-(r-u_r)y}\mathrm{d}y \\
&\le \frac{384e^{-(r-u_r)^2/2}}{\sqrt{2\pi}(r-u_r)^5u_r^4},
\end{aligned}$$

其中，第二个不等式成立是因为对于任意 $t\in[0,r]$，$\mathsf{P}_0(d(\mathbf{X}\ \mathcal{E}_r)\le t)\le\bar{\Phi}(r-t)\le(2\pi)^{-1/2}e^{-(r-t)^2/2}/(r-t)$；第三个不等式成立是因为对于所有 $t\in[0,u_r]$，$t(1-t^2/u_r^2)^3\le 8(u_r-t)^3/u_r^2$ and$e^{-(r-t)^2/2}/(r-t)\le e^{-(r-u_r)^2/2-(r-u_r)(u_r-t)}/(r-u_r)$。根据（11）、（12）和（13），我们有

$$\sigma_r^2\le\left(\frac{e^{-3(r-u_r)^2/2}}{(2\pi)^{3/2}(r-u_r)^3}\frac{384e^{-(r-u_r)^2/2}}{\sqrt{2\pi}(r-u_r)^5u_r^4}\right)^{1/2}=\frac{4\sqrt{6}e^{-(r-u_r)^2}}{\pi(r-u_r)^4u_r^2}. \qquad (14)$$

设 $n\ge 2$。固定满足 $w<u^2$ 的 w、$u>0$。假设 $r-u<x_1<r-\sqrt{w}$ 且$\|\mathbf{x}-r\mathbf{e}_1\|\le u$，其中 $\mathbf{x}\in\mathbb{R}^n$。由于 $r\mathbf{e}_1\in\mathcal{E}_r$，我们有 $d(\mathbf{x}\ \mathcal{E}_r)\le u$。令 $\bar{}\mathbf{x}=\arg\min_{\mathbf{y}\in\mathcal{E}_r}\|\mathbf{x}-\mathbf{y}\|$。于是，$\bar{x}_1\ge r$，从而 $d(\mathbf{x}\ \mathcal{E}_r)\ge\bar{x}_1-x_1>\sqrt{w}$。此外，$\mathsf{P}_0(\|\mathbf{X}-r\mathbf{e}_1\|\le u\,|\,X_1=x)$ 等于自由度为 $n-1$ 的卡方随机变量不超过$u^2-(x-r)^2$ 的概率，因为对于任意 $\mathbf{x}\in\mathbb{R}^n$，都有 $\|\mathbf{x}-r\mathbf{e}_1\|^2=(x_1-r)^2+\sum_{i=2}^n x_i^2$。因此，存在$C>0$ 使得

$$\begin{aligned}
\mathsf{P}_0(w<d(\mathbf{X},\mathcal{E}_r)^2\le u^2) &\ge \mathsf{P}_0(r-u<X_1<r-\sqrt{w},\|\mathbf{X}-r\mathbf{e}_1\|\le u) \\
&=\int_{r-u}^{r-\sqrt{w}}\mathsf{P}_0(\|\mathbf{X}-r\mathbf{e}_1\|\le u\,|\,X_1=x)\frac{e^{-x^2/2}}{\sqrt{2\pi}}\mathrm{d}x \qquad (15)\\
&=C\int_{r-u}^{r-\sqrt{w}}\int_0^{u^2-(x-r)^2} t^{(n-3)/2}e^{-t/2}e^{-x^2/2}\mathrm{d}t\mathrm{d}x.
\end{aligned}$$

利用分部积分，可以证明 $h_r(u) = \int_0^{u^2} \mathrm{P}(w < d(\mathbf{X}\ \mathcal{E}_r)^2 \le u^2)\mathrm{d}w$。然后，通过

(15),

$$\begin{aligned} h_r(u) &\ge C\int_0^{u^2}\int_{r-u}^{r-\sqrt{w}}\int_0^{u^2-(x-r)^2} t^{(n-3)/2}e^{-t/2}e^{-x^2/2}\mathrm{d}t\mathrm{d}x\mathrm{d}w \\ &= 2C\int_0^{u}\int_0^{y^2}\int_0^{(u^2-y^2)^{1/2}} s^{n-2}e^{-s^2/2}e^{-(r-y)^2/2}\mathrm{d}s\mathrm{d}w\mathrm{d}y \\ &= 2C\int_0^{u} y^2\int_0^{(u^2-y^2)^{1/2}} s^{n-2}e^{-s^2/2}e^{-(r-y)^2/2}\mathrm{d}s\mathrm{d}y \\ &= 2C\int_0^{u} \rho^{n+1}e^{-(r-\rho)^2/2}\int_0^{\pi/2}\cos(\theta)^2\sin(\theta)^{n-2}e^{-r\rho(1-\cos(\theta))}\mathrm{d}\theta\mathrm{d}\rho, \end{aligned} \tag{16}$$

其中，第一个等号是通过交换前两个积分并令 $s = \sqrt{t}$ 和$y = r - x$ 得到的，最后一个等号则是令 $s = \rho\sin(\theta)$ 和 $y = \rho\cos(\theta)$ 得到的。

令 $\varepsilon_r = 1/u_r$。那么，对于所有充分大的 r，都有 $0 \le \varepsilon_r/r \le 1$。因此，对于所有 $\rho \in (0, u)$，式 (16) 中最后一个表达式的内层积分满足

$$\begin{aligned} \int_0^{\pi/2}\cos(\theta)^2\sin(\theta)^{n-2}e^{-r\rho(1-\cos(\theta))}\mathrm{d}\theta &\ge \int_0^{\arccos(1-\varepsilon_r/r)}\cos(\theta)^2\sin(\theta)^{n-2}e^{-r\rho(1-\cos(\theta))}\mathrm{d}\theta \\ &\ge \left(1-\frac{\varepsilon_r}{r}\right)^2 e^{-\varepsilon_r\rho}\int_0^{\arccos(1-\varepsilon_r/r)}\sin(\theta)^{n-2}\mathrm{d}\theta \\ &= \left(1-\frac{\varepsilon_r}{r}\right)^2 e^{-\varepsilon_r\rho}\int_0^{\varepsilon_r/r}(2\alpha)^{(n-3)/2}(1-\alpha/2)^{(n-3)/2}\mathrm{d}\alpha \\ &\ge \kappa_r e^{-\varepsilon_r\rho}, \end{aligned}$$

其中，$\kappa_r = (1-\varepsilon_r/r)^2(1-\varepsilon_r/(2r))^{(n-3)/2}(2\varepsilon_r/r)^{(n-1)/2}/(n-1)$，该等式是通过令$\theta = \arccos(1-\alpha)$ 得到的。因此，由式 (16) 并两次使用分部积分，我们得到

$$\begin{aligned} h_r(u_r) &\ge 2C\kappa_r\int_0^{u_r}\rho^{n+1}e^{-(r-\rho)^2/2-\varepsilon_r\rho}\mathrm{d}\rho \\ &= 2C\kappa_r\left(I_r(u_r) + \int_0^{u_r}\frac{e^{-(r-\rho)^2/2-\varepsilon_r\rho}}{(r-\varepsilon_r-\rho)^2}\rho^{n-1}\left(n(n+1)+\frac{3(n+1)\rho}{r-\varepsilon_r-\rho}+\frac{3\rho^2}{(r-\varepsilon_r-\rho)^2}\right)\mathrm{d}\rho\right) \\ &\ge 2C\kappa_r I_r(u_r), \end{aligned}$$

其中

$$I_r(u_r) \coloneqq \frac{e^{-(r-u_r)^2/2-1}}{r-u_r^{-1}-u_r}u_r^{n+1}\left(1-\frac{n+1}{u_r(r-u_r^{-1}-u_r)}-\frac{1}{(r-u_r^{-1}-u_r)^2}\right).$$

回顾一下，根据引理 2 和定理 1 的证明，$r - u_r \to \infty$ 和 $u_r/r \to 1$ 当 $r \to \infty$ 时。因此，我们有 $\kappa_r \sim r^{-n+1}2^{(n-1)/2}/(n-1)$ 和 $I_r(u_r) \sim r^{n+1}e^{-(r-u_r)^2/2-1}/(r-u_r^{-1}-u_r)$，其中

$\sim$ 表示当 $r \to \infty$时渐近等价。由于对所有 r， $\delta^2 = h_r(u_r)$ 均成立，上述不等式意味着

$$\limsup_{r\to\infty} r^2 \frac{e^{-(r-u_r)^2/2-1}}{r - u_r^{-1} - u_r} \le C_* \delta^2,$$

其中 $C_* = (n-1)/(2^{(n+1)/2}C)$。最后，结合此结果、(14) 和定理 1，可得

$$\begin{aligned}
\limsup_{r\to\infty} r^2 (r-u_r)^2 \frac{\sigma_r^2}{p_r^2} &\le \frac{4\sqrt{6}}{\pi} \frac{1}{\liminf_{r\to\infty} r^4 p_r^2} \limsup_{r\to\infty} r^4 \frac{e^{-(r-u_r)^2}}{(r-u_r)^2} \\
&\le \frac{4e^2\sqrt{6}}{\pi\delta^4} \left(\limsup_{r\to\infty} r^2 \frac{e^{-(r-u_r)^2/2-1}}{r - u_r^{-1} - u_r} \right)^2 \\
&\le \frac{4C_*^2 e^2 \sqrt{6}}{\pi} < \infty.
\end{aligned} \tag{17}$$

当 $n = 1$ 时，使用与 (16) 中相同的论证，可得以下关系：

$$\begin{aligned}
h_r(u_r) &= \int_0^{u_r^2} \mathsf{P}(w < d(\mathbf{X}, \mathcal{E}_r)^2 \le u_r^2) \mathrm{d}z \\
&= (2\pi)^{-1/2} \int_0^{u_r^2} \int_{r-u_r}^{r-\sqrt{w}} e^{-x^2/2} \mathrm{d}x \mathrm{d}w \\
&= (2\pi)^{-1/2} \int_0^{u_r} y^2 e^{-(r-y)^2/2} dy.
\end{aligned}$$

利用分部积分，右侧有如下下界：

$$\frac{1}{\sqrt{2\pi}} \frac{e^{-(r-u_r)^2/2}}{r-u_r} \left(\frac{u_r}{r}\right)^2 r^2 \left(1 - \frac{2}{u_r(r-u_r)} - \frac{1}{(r-u_r)^2}\right) \sim \frac{r^2 e^{-(r-u_r)^2/2}}{\sqrt{2\pi}(r-u_r)}.$$

类似于 (17)，我们应用 $\delta^2 = h_r(u_r)$，得到

$$\limsup_{r\to\infty} r^2 (r-u_r)^2 \frac{\sigma_r^2}{p_r^2} \le \frac{4\sqrt{6}}{\pi} \frac{1}{\liminf_{r\to\infty} r^4 p_r^2} \limsup_{r\to\infty} r^4 \frac{e^{-(r-u_r)^2}}{(r-u_r)^2} \le 8\sqrt{6} < \infty.$$

证明至此完成。 □

B. 技术引理

引理 3。固定 $z \in (0,\infty) \times \mathbb{R}^{n-1}$，并令 $g(u) = d(f_u(z), \mathcal{E}) - u$，其中任意 u 都属于 $(0, x_1^*)$ 的紧区间 Θ。若 v 位于 Θ 的内部，且存在 $\delta > 0$ 使得对所有 $t \in (0,\delta)$ 都有 $\mathbb{1}\{g(v-t) \le 0\} \ne \mathbb{1}\{g(v+t) \le 0\}$，则称 v 为零交叉。于是，零交叉至多有两个。证明。令 $z_1(u) = x_1^* - u + z_1/(x_1^* - u)$ 为 $f_u(z)$ 的第一坐标，其中 $u \in \Theta$。我们将其写作 $\Theta = [u_L, u_U]$，其中某个 $0 < u_L \le u_U < x_1^*$。

令 $u_* = \min\{\max\{u_L, x_1^* - \sqrt{z_1}\}, u_U\}$, $I_1 := [u_L, u_*]$, 以及 $I_2 := (u_*, u_U]$。由于 $d(\cdot, \mathcal{E})$ 是 1-Lipschitz 函数，且 $|z_1'(\cdot)| \le 1$ 在 I_1 上成立，因此有 $|g(u) + u - g(v) - v| \le |z_1(u) - z_1(v)| \le |u - v|$，这意味着对于任意 $u, v \in I_1$，均有 $g(u) \le g(v) + v - u + |u - v|$。因此，若对于某个 $v \in I_1$ 有 $g(v) \le 0$，则对于所有 $u \in [v, u_*]$，都有$g(u) \le 0$。

另一方面， $z_1(\cdot)$ 在 I_2 上严格递增且为凸函数。此外，容易验证， $d(\mathbf{y}, \mathcal{E})$ 关于 y_1 为凸函数。因此， $d(f_u(\mathbf{z}), \mathcal{E})$ 关于 u 在 $(u_*, w]$ 上递减，并在 $(w, u_U]$ 上递增，其中某个 $w \in (u_*, u_U]$。这表明， $g(\cdot)$ 在 $(u_*, w]$ 上也递减。此外， $g(\cdot)$ 在 $(w, u_U]$ 上为凸函数。因此， Θ 中至多有两个零交叉。

□

引理 4. $h(\cdot)$ 和 $p(\cdot)$ 在 u_* 处可微，且 $h'(u_*) = u_*^2 p'(u_*) \neq 0$。

证明。回顾一下， $p(u) = \mathsf{P}(d(\mathbf{X}, \mathcal{E}) \le u)$。不难验证， $d(\cdot, \mathcal{E})$ 是 1-利普希茨函数，且几乎处处可微，并且 $\|\nabla d(\cdot, \mathcal{E})\| = 1$。于是，根据余面积公式 [Evans and Gariepy，1992年，定理 3.4.2]，我们有

$$
\begin{aligned}
p(u_* + \delta) - p(u_*) &= \mathsf{P}(u_* < d(\mathbf{X}, \mathcal{E}) \le u_* + \delta) \\
&= \int_{\mathbb{R}^d} \phi(\mathbf{x}) \mathbb{1}\{u_* < d(\mathbf{x}, \mathcal{E}) \le u_* + \delta\} \|\nabla d(\mathbf{x}, \mathcal{E})\| \mathrm{d}\mathbf{x} \\
&= \int_{\mathbb{R}} \left(\int_{\partial(\mathcal{E} + B(u_* + t))} \phi(\mathbf{z}) \mathbb{1}\{u_* < d(\mathbf{z}, \mathcal{E}) \le u_* + \delta\} \mathrm{d}\mathcal{H}(\mathbf{z}) \right) \mathrm{d}t \\
&= \int_0^{\delta} \left(\int_{\partial(\mathcal{E} + B(u_* + t))} \phi(\mathbf{z}) \mathrm{d}\mathcal{H}(\mathbf{z}) \right) \mathrm{d}t,
\end{aligned}
$$

其中， $\phi(\mathbf{z}) = (2\pi)^{-n/2} e^{-\|\mathbf{z}\|^2/2}$ 是 n维标准高斯分布的密度， $\mathcal{H}$ 是$(n-1)$维豪斯多夫测度。

记 $\mathcal{E}_u := \{\mathbf{x} : d(\mathbf{x}, \mathcal{E}) \le u\}$。根据微积分基本定理，只需证明 $g(u) := \int_{\partial \mathcal{E}_u} \phi(\mathbf{z}) \mathrm{d}\mathcal{H}(\mathbf{z})$在$(0, \infty)$上连续。为此，我们任意固定 $u > 0$，并将$\mathbf{z} \in \partial \mathcal{E}_u$处的外单位法向量记为 $n(\mathbf{z})$。然后，根据变量替换，

$$
g(u + t) = \int_{\partial \mathcal{E}_u} \phi(\mathbf{z} + t n(\mathbf{z})) J_t(\mathbf{z}) \mathrm{d}\mathcal{H}(\mathbf{z}),
$$

其中，对每个 $t \ge 0$， $J_t(\mathbf{z})$ 表示映射 $\mathbf{z} \mapsto \mathbf{z} + t n(\mathbf{z})$ 的雅可比行列式。由 $\partial \mathcal{E}_u$ 的光滑性和 $\mathcal{E}_u$ 的凸性，不难验证雅可比行列式 $J_t(\mathbf{z})$ 非负，且关于 $\mathbf{z}$ 和 t 连续；例如参见 Schneider [2013] 和 Cecil and Ryan [2015]。

将 $\epsilon > 0$ 取为足够小的正数。令 $\eta(\mathbf{z}\ t) = \phi(\mathbf{z} + tn(\mathbf{z}))J_t(\mathbf{z})$，其中 $(\mathbf{z}\ t) \in \partial\mathcal{E}_u \times [0,\infty)$。于是，我们可以选取一个紧集 $K \subset \partial\mathcal{E}_u$ 和一个常数 $t_K > 0$，使得对所有 $t \in [0, t_K]$，$|\int_K \eta(\mathbf{z}\ t)\mathrm{d}\mathcal{H}(\mathbf{z}) - \int_K \eta(\mathbf{z}, 0)\mathrm{d}\mathcal{H}(\mathbf{z})| < \epsilon/3$，且 $\int_{\partial\mathcal{E}_u \setminus K} \eta(\mathbf{z}\ t)\mathrm{d}\mathcal{H}(\mathbf{z}) < \epsilon/3$。这是可行的，因为 η 在 $K \times [0, t_K]$ 上一致连续，η 在 $\partial\mathcal{E}_u \times [0,\infty)$ 上具有非负性，且根据 Ball [1993]，g 一致有界。因此，对所有 $t \in [0, t_k]$,

$$
\begin{aligned}
&|g(u+t) - g(u)| \\
&\leq \left| \int_K \eta(\mathbf{z}, t)\mathrm{d}\mathcal{H}(\mathbf{z}) - \int_K \eta(\mathbf{z}, 0)\mathrm{d}\mathcal{H}(\mathbf{z}) \right| + \int_{\partial\mathcal{E}_u \setminus K} \eta(\mathbf{z}, t)\mathrm{d}\mathcal{H}(\mathbf{z}) + \int_{\partial\mathcal{E}_u \setminus K} \eta(\mathbf{z}, 0)\mathrm{d}\mathcal{H}(\mathbf{z}) \\
&< \epsilon.
\end{aligned}
$$

因此，$p'(u_*) = g(u_*) > 0$。根据 h的定义，对于任意足够小的 $\varepsilon > 0$，我们有$u_*^2(p(u_* + \varepsilon) - p(u_*)) \leq h(u_* + \varepsilon) - h(u_*) \leq (u_* + \varepsilon)^2(p(u_* + \varepsilon) - p(u_*))$。将所有表达式除以 ε 并令 $\varepsilon \to 0$ 趋于极限，得到 $h'(u_*) = u_*^2 p'(u_*) > 0$。

□

参考文献

D. Ahn 和 L. Zheng。针对凸稀有事件集的条件重要性采样。载于*2023*年冬季仿真会议（*WSC*），第363–374页，2023年。D. Ahn 和 L. Zheng。多面体期望的高效模拟及其在金融中的应用。《运筹学数学》，页码（在线先行发表），2025年。S. Asmussen 和 P. W. Glynn。《随机仿真：算法与分析》。Springer，纽约，2007年。

Y. Bai、Z. Huang、H. Lam 和 D. Zhao。神经网络和随机森林预测器的稀有事件模拟。*ACM*建模与计算机仿真汇刊，32(3):1–33，2022。

Y. Bai、H. Lam 和 X. Zhang。用于极端事件估计的分布鲁棒优化框架。*arXiv:2301.01360*，第 1–92 页，2023。

K. Ball. 高斯测度的反等周问题。离散与计算几何，10(4):411–420，1993。

A. Bassamboo、S. Juneja 和 A. Zeevi。具有极值依赖的投资组合信用风险：渐近分析与高效模拟。《运筹学》，56(3):593–606，2008年6月。J. Blanchet 和 H. Lam。多服务台队列的稀有事件模拟。《运筹学数学》，39(4):1142–1178，2014。J. Blanchet 和 K. Murthy。通过最优传输量化分布模型风险。《运筹学数学，44(2):565–600，2019。J. Blanchet、F. He 和 H. Lam。通过模拟计算给定边缘分布时的最坏情况期望。载于*2017*年冬季仿真会议论文集。IEEE，2017。J. Blanchet、F. He 和 K. Murthy。关于分布鲁棒极值分析。《极值》，23:317–347，2020。Z. I. Botev。线性约束下的正态分布：通过极小极大倾斜进行模拟与估计。《皇家统计学会会志*B*辑：统计方法学》，79(1):125–148，2017年1月。J. A. Bucklew。稀有事件仿真导论。Springer，纽约，2004。T. E. Cecil 和 P. J. Ryan。超曲面几何。Springer，纽约，2015。A. Deo 和 K. Murthy。利用自组织重要性采样器，在黑箱模拟中高效估计分布尾部。《运筹学》，73(1):325–343，2025。L. C. Evans 和 R. F. Gariepy。测度论与函数的精细性质。CRC出版社，博卡拉顿，1992年。P. Glasserman。金融工程中的蒙特卡罗方法。Springer，纽约，2003。P. Glasserman 和 X. Xu。稳健风险度量与模型风险。《定量金融》，14(1): 29–58，2014。P. Glasserman、P. Heidelberger 和 P. Shahabuddin。估计风险价值的方差缩减技术。《管理科学》，46(10):1349–1364，2000。

P. Glasserman、W. Kang 和 P. Shahabuddin。多因子投资组合信用风险的快速仿真。《运筹学》，56(5):1200–1217，2008。

S. He、G. Jiang、H. Lam 和 M. C. Fu。用于高效随机求根与分位数估计的自适应重要性抽样。《运筹学》，72(6):2612–2630，2024。

Z. Huang、H. Lam 和 Z. Liu。稀有事件估计中的输入尾部不确定性传播：轻尾与重尾二分。*arXiv*，第 1–46 页，2023。

S. Juneja 和 P. Shahabuddin。使用延迟危险率扭曲模拟重尾过程。*ACM*建模与计算机仿真汇刊，12(2):94–118，2002。

H. Lam 和 C. Mottet。无需参数模型的尾部分析：最坏情况视角。《运筹学》，65(6):1696–1711，2017。

M. K. Nakayama 和 B. Tuffin。稀有事件情境下均值函数估计的效率。载于*2023*年冬季仿真会议（*WSC*），第 351–362 页，2023。

M. Niewiadomska-Bugaj 和 R. Bartoszy´nski。概率与统计推断。威立，霍博肯 River Street，第三版，2021。

G. Rubino 和 B. Tuffin。使用蒙特卡罗方法进行稀有事件仿真。约翰·威立父子有限公司，英国奇切斯特，2009。

R. Schneider。凸体：布伦–闵可夫斯基理论。剑桥大学出版社，纽约州纽约市，第二版，2013年。

A. W. Van Der Vaart 和 J. A. Wellner。 弱收敛与经验过程：及其在统计学中的应用。Springer，第二版，2023。

L. Zhang、J. Yang 和 R. Gao。Wasserstein分布鲁棒优化简短且通用的对偶性证明。《运筹学》，73(4):2146–2155，2025。