arXiv:2302.12823v1 [cs.CC] 24 Feb 2023

超大对象的生成模型

Lunjia Hu* Inbal Livni-Navon† Omer Reingold‡

摘要

本研究开启了对显式分布的系统性研究，这类分布与一个单一 指数规模组合对象不可区分。在此，我们拓展了 Goldreich、Goldwasser 和 Nussboim（SICOMP 2010）的工作；该工作聚焦于实现与均匀分布不可区分、且满足某些全局性质（他们称之为真实性）的超大对象。与单一对象不可区分这一概念，源于学习理论中对生成模型的研究，以及图论中的正则性引理。在伪随机性的背景下已得到充分理解的问题，在考虑超大对象的生成模型时却会带来重大挑战，有时甚至无法解决。

我们通过为多个自然场景中的超大不可区分对象提供学习算法，展示了本研究的多功能性。这些场景包括：满足真实性要求的稠密函数和图，即要求函数中 1 的数量或图中边的数量符合真实性要求；以及适用于满足某些全局性质的稀疏图的弱正则性引理版本。这些结果及其他结果推广了基本的伪随机对象以及算法公平性中引入的概念。这些结果依赖于来自多个领域的概念和技术，包括学习理论、复杂性理论、密码学和博弈论。

*斯坦福大学。研究获得 Simons 基金会算法公平性理论合作项目、Omer Reingold 的 NSF 奖项 IIS-1908774，以及 Moses Charikar 的 Simons 研究员奖资助。

†斯坦福大学。获 Simons 基金会算法公平性理论合作项目、Sloan 基金会资助项目 2020-13941 及 Zuckerman STEM 领导力计划支持。

‡斯坦福大学。获 Simons 基金会算法公平性理论合作项目及 Simons 基金会研究员奖 689988 支持。

目录

1 引言 **3**

1.1 总体目标：超大对象的不可区分生成模型 4 1.2 我们的结果.5 1.3 相关工作.9

2 预备知识 **10**

3 学习函数 的指数级大定义域 13

3.1 学习样本访问二元函数.13 3.2 保留支撑集大小的真实性学习.15 3.2.1 3.2.2 第 1 步：具有 $2^{O(N)}$ 时间实现的精确模型.16 3.2.2 第 2 步：具有 poly(N) 时间实现的不可区分模型.17 3.2.3 步骤 3：具有 polylog(N) 时间实现的不可区分模型.20 3.3 学习支撑集访问二元函数.23 3.3.1 支持访问对象审计器.28 3.4 学习位串函数.30

4 学习指数规模图 **32**

4.1 学习稠密图.32 4.2 学习无稠密子图的稀疏图.33 4.3 学习稀疏均匀出度图.38 4.4 学习均匀度图.40

5 不可能性 **42**

1 引言

对于一组计算上有界的区分器而言，伪随机分布与均匀分布不可区分。伪随机性是计算机科学和数学许多领域的基石。伪随机分布的多样性源于它们能够生成的不同对象（比特串、函数、置换等），以及可对区分器施加的不同计算限制。在密码学领域，通常考虑至少为多项式时间的强大区分器，由此产生了伪随机生成器 [BM84、Yao82], 伪随机函数 [GGM86] 和伪随机置换 [LR88] 等核心概念。能力更有限的区分器则产生了其他基本概念，例如 k-重独立哈希和 ε-偏分布（参见 [WC81 、 NN93]）。在显式组合构造领域，我们通常试图用单个对象而非一个分布来模拟均匀分布。一个典型例子是扩展图这一基本概念（综述参见 [HLW06、Vad12] ），及其多个变体（包括各种随机性提取器的概念）。这些图与均匀选取的图对于有限的一组区分器而言不可区分，例如那些检查随机边是否跨越给定割集的区分器。

在伪随机性的这些经典领域中，人们会构造一个分布，甚至一个单独的对象，来模拟某种分布（通常是均匀分布）。本文则要求一个分布去模拟一个单独的对象 （表 1）。从伪随机性的角度看，这种逆转似乎荒谬，但从生成模型的角度看却完全合情合理。理论计算机科学（TOC）社群早期接触生成模型，与万维网有关。这些模型生成图的分布，以模仿万维网的某些特性，例如节点度数的幂律（综述见 [Mit03] ）。在任一给定时刻，万维网都是一张单独的图，但它也是一张规模极大、无法简单描述的图。生成模型为分析万维网上协议的预期性能，或通过实验估计其性能，提供了实用的方法。

其他研究充分的生成模型包括随机块模型 [HLL83] 以及更复杂的混合成员随机块模型 [ABFX08]。考虑一张表示个体之间某些联系的图，例如社交网络中的连接关系。随机块模型将顶点划分为互不相交的社群，并为每一对社群指定一个连接概率。该模型表示图上的一个分布：对于任意两个顶点，都根据其端点所属社群对所指定的概率，独立地放置一条边。（在混合成员模型中，每个顶点都被赋予一个社群上的分布。）这些模型有助于识别社交结构中的实用子结构，例如子社群或不同的社会角色。但给定一个社交网络 B^*，什么样的模型适合对其进行刻画？毕竟，模型描述的是网络上的一个分布，而非我们试图解释的单个网络。一种常见做法是寻求最大似然模型。在所有模型中，最大似然模型使抽样得到 B^* 的概率最大化。估计最大似然模型的启发式方法，在生成模型文献及其实际应用中一直发挥着重要作用。需要指出的是，模型生成 B^* 的概率通常非常小。因此，最大似然模型是否有意义，仍值得商榷，而且可能取决于具体情境。

从不可区分性的角度来看，寻求一种生成模型，使其产生的分布与 B^* 对于一组有意义的区分器而言不可区分，或许更加自然。例如，对于随机块模型，自然的区分器由两个顶点集合定义

	我们模仿什么?	我们构造什么?
伪随机性	对象分布	对象分布
显式构造（例如扩展图）	对象分布	单个对象
我们的设定：生成模型	单个对象	对象分布

表 1：问题设定之间的比较

U 和 V，并询问图中一条随机边从 U 跨到 V的概率是多少。能够欺骗所有这类区分器的随机块模型，正是 Frieze-Kannan 正则性引理（也称为弱正则性引理） [FK99]所给出的模型。生成模型的不可区分性视角，以及学习与伪随机性之间已知的联系（我们稍后将讨论），既是本研究的动机，也是本研究的出发点。

1.1 总体目标：超大对象的不可区分生成模型

在生成模型的许多应用中，例如对万维网或社交网络建模，所建模的对象都非常庞大。本文旨在建立一套高效学习和实现超大生成模型的系统理论。我们的模型将生成满足某些全局性质的对象分布，且该分布与一个固定的组合对象不可区分。这样的理论带来了非平凡的挑战：无论是生成超大伪随机对象，还是生成多项式规模对象的生成模型，都不会出现这些挑战。

具体而言，我们假设可以访问一个规模为指数级的对象 B^*。例如，B^* 可以是定义域规模为指数级的函数，也可以是具有指数多个顶点和边的图。我们最感兴趣的是对象 B^* 过于庞大，无法整体读取或处理的情况，因此必须通过采样来访问它：例如，当 B^* 表示函数 f：$X \to \{0,1\}$，且 $|X|$ 的规模为指数级时，我们可以通过请求随机对$(x,\ f(x)) \in X \times \{0,1\}$（样本访问），或请求以$f(x)$为条件的随机输入 $x \in X = 1$（支撑集访问），来访问 B^*。给定对超大对象 B^*的访问权限，我们的目标是为 B^*创建一个生成模型 M。在这里，我们的模型 M 表示对象上的一个分布，并且我们希望确保该分布与 B^* 对于某个类别 $\mathcal{D}$中的所有区分器 D 而言都是不可区分的。具体而言，如果用 $D^B \in \{$“接受”，“拒绝”$\}$ 表示区分器 D 在对对象 B进行样本/支持集访问时的输出，那么我们的不可区分性要求是：对于每个 $D \in \mathcal{D}$,

$$|\Pr[D^{B^*} = \text{“accept”}] - \mathbb{E}_{B\sim M}[\Pr[D^B = \text{“accept”}]]| \le \varepsilon.$$

我们的目标是构建一个高效的学习器 L，它在获得真实对象 B^* 的样本/支持集访问权限后，能够输出满足上述不可区分性要求的模型 M。当 B^* 规模呈指数级时（这正是我们关注的情形），输出模型 M 也需要生成规模呈指数级的对象，因此我们不能指望高效学习器直接输出 M。相反，我们希望学习器输出 M的高效实现，粗略地说，也就是一种随机化算法，能够高效地提供对从 M 中抽取的对象进行样本/支持集访问的能力（定义 2.9 和 2.10）。

我们学习一个与真实对象 B^* 不可区分的生成模型 M 的目标，与伪随机性领域中 Goldreich、Goldwasser 和 Nussboim [GGN10] 所研究的问题类似：

他们研究的是如何高效实现满足某些全局性质的超大对象分布，使其与此类对象的均匀分布不可区分。后续研究工作（例如 [GGN10]之后的 [NNT05、NN07]）探讨了如何高效实现与某些超大随机图分布不可区分的分布。所有这些工作旨在实现与已知对象分布的不可区分性，而在我们学习生成模型的问题中，我们假设真实对象 B^* 起初是未知的；为了收集有关 B^*的信息，我们还需要一个学习器 L ，能够利用对 B^*的样本/支持集访问，高效地 构造一个不可区分模型的实现。

除了不可区分性之外，我们还旨在实现 [GGN10] 引入的真实性概念。为说明这一概念，请考虑伪随机置换 $f_s: \{0,1\}^n \mapsto \{0,1\}^n$ [LR88]。如果一个置换分布与置换的均匀分布不可区分，那么它就是伪随机的。需要注意的是，伪随机 $\{0,1\}^n \mapsto \{0,1\}^n$函数 [GGM86] 也与 $\{0,1\}^n$ 上的随机置换不可区分（只要查询数量远小于 $2^{n/2}$）。然而，在伪随机置换的应用中，坚持要求伪随机对象满足作为置换这一全局条件至关重要。这促使 [GGN10] 区分不可区分性（伪随机对象对于一类区分器而言，与均匀对象不可区分）和真实性；真实性是一种全局性质，需要在统计意义上精确或近似地成立。在我们的设定中，如果从 M 表示的分布中抽取的每个对象 B 都满足某种全局性质，那么生成模型 M 就是真实的。例如，当真实对象 B^*是一个函数 $f^*: X \to \{0,1\}$，其支持集大小为 $|\{x \in X: f^*(x) = 1\}|$为 k时，针对 B^* 的生成模型 M 的真实性要求可能会限制 M ，使其始终生成支持集大小为 k的函数。

实现超大伪随机对象的研究 [NR02, GGN10, NNT05, NN07] 将伪随机函数和置换作为重要的构件。除了这些构件，我们用于学习超大对象生成模型的技术也源于正则性引理，尤其是 Trevisan、Tulsiani 和 Vadhan 的工作 [TTV09]。在 [TTV09],中，他们构造了一个可高效实现的函数 $f: X \mapsto [0,1]$ ，该函数对于由函数 $g: X \mapsto [0,1]$表示的区分器族而言，与某个 $f^*: X \mapsto [0,1]$ 不可区分。这里的不可区分性是指， $|\mathbb{E}[f(x)g(x)] - \mathbb{E}[f^*(x)g(x)]|$ 小于某个误差参数 ε。在 [TTV09], 之后，关于创建不可区分函数的问题及其在密码学中的应用，又在 [VZ13, JP14, Sk´o16a, Sk´o16b, Sk´o17, CCL18]中得到了进一步研究。这些工作假设真实函数 f^* 已知，并未明确处理学习 f^*的问题；不过，相关的学习任务已在算法公平性文献中通过多准确性、多重校准和结果不可区分性等概念得到研究[HJKRR18, KGZ19, DKR+21, GKSZ22, DLLT23]。将这些工作的技术用于解决我们研究场景中的某些问题时，还需要应对额外的挑战，例如我们希望生成模型满足的真实性要求。

1.2 我们的成果

本文的主要概念性贡献在于为不可区分性研究提出了一个新的方向，这一方向动机充分，且在技术上颇具挑战性。正如我们在第 1.1 节中所介绍的，从单个巨大对象出发的不可区分性概念，其核心结合了学习理论与伪随机性这两个领域。近期研究揭示了二者之间的深刻联系，为描述和解决一系列丰富的自然问题提供了途径。下文

我们总结了在这一新框架下要解决的主要问题。

保留支撑集大小的真实性学习

假设我们可以访问函数 $f^*\colon \{0,1\}^n \to \{0,1\}$ 的样本，并希望为 f^*构建一个不可区分的生成模型。这里的样本访问允许我们观察形如 $(x,\ f^*(x))$ 的数对，其中 x 是从 $\{0,1\}^n$中均匀随机抽取的；因此，我们假设每个区分器也会根据函数 f给出的随机数对 $(x,\ f(x))$ 来决定接受或拒绝，而该函数可能是也可能不是真正的 f^*。为二值函数学习生成模型的任务与 [DKR+21]研究的无访问结果不可区分性 任务密切相关，并且已有研究发现，该任务可以归约为多准确性。事实上，假设区分器的复杂度有界且能够被高效学习，利用 [HJKRR18、KGZ19、DKR+21]中的先前算法，我们可以设计一种高效学习器，构造出与 f^* 不可区分的生成模型（定理 3.2）。以这种方式构造的模型由预测器 $p\colon \{0,1\}^n \to [0,1]$指定，并表示函数 $f\colon \{0,1\}^n \to \{0,1\}$ 的分布，其中函数值 $f(x)$ 对每个 $x \in \{0,1\}^n$ 相互独立地服从均值为 $p(x)$ 的伯努利分布 $\mathrm{Ber}(p(x))$。

当我们额外施加真实性要求时，学习二值函数的生成模型会变得更具挑战性。一种自然的真实性要求是保留函数的支撑集大小。假设我们知道函数的支撑集大小 $|\{x \in \{0,1\}^n : f^*(x) = 1\}|$ of $f^*$is k，我们希望模型只生成支撑集大小也为 k 的函数。我们将展示如何构建一个高效学习器，使其能够输出这样一个真实的模型，并且该模型与真实函数 f^* 不可区分：

定理 1.1 (定理 3.4 的非正式陈述)。 令 $\mathcal{B}$ 为由满足 $|\text{支撑集}(f)| = k$ 的二值函数 $f\colon \{0,1\}^n \to \{0,1\}$ 所诱导的样本访问对象类，其中 支撑集$(f) := \{x \in \{0,1\}^n : \ f(x) = 1\}$。令 $\mathcal{D}$ 为一个可高效学习的区分器类。存在一个高效的 (ε, δ)-学习器 L，用于学习 $\mathcal{B}$ （相对于 $\mathcal{D}$），且该学习器始终输出模型 M 的高效实现，该模型相对于 $\mathcal{B}$ 满足真实性要求。

请注意，不能仅凭计算能力受限的区分器来强制满足关于支持集大小的真实性要求，因为要精确计算函数 f 的支持集大小，就必须读取所有指数多的输入 $x \in \{0,1\}^n$ 对应的函数值 $f(x)$。此外，由预测器 p 指定的生成模型无法直接满足这一真实性要求，因为函数值 $f(x)$ 按照 $\mathrm{Ber}(p(x))$ 分布，且与其他个体 $x' \neq x$ 的函数值 $f(x')$ 相互独立。要强制实现固定的支持集大小，不同个体的函数值必须以全局方式协调，因此我们需要采用新技术。我们创建一棵二叉树，其叶节点对应函数定义域 $\{0,1\}^n$，并借鉴先前工作（如 [GGN10], ）中的思路，将支持集大小预算从根节点分配到叶节点。然而，真实函数 f^* 是一个单一的未知对象，与 [GGN10], 中考虑的均匀分布大不相同，因此不存在闭式分布（例如 [GGN10] 中使用的二项分布）可以指导我们如何将一个节点的支持集大小预算分配给它的两个子节点。相反，我们需要估计 预算应如何划分，这会导致误差在向叶节点传播的过程中不断累积，并迫使我们在到达叶节点之前停止。为了高效地将预算传递到叶节点，我们求解一个零和博弈：玩家 C 为叶节点选择预算，玩家 D 则试图区分这些预算与目标。我们证明，如果玩家

D 使用乘法权重算法来最小化遗憾，我们就能根据玩家 C 最优响应的经验分布，创建一个不可区分且真实的模型。

通过支撑集访问学习函数

在经典的学习函数 $f\colon \{0,1\} \to \{0,1\}$ 的设定中，学习器会收到形如 $(x,\ f(x))$ 的随机样本，其中 $x \in \{0,1\}$ 是均匀分布的。在这项工作中，我们还考虑一种函数对象，在其中我们会收到一个随机正项。也就是说，学习器会收到满足 $f(x) = 1$ 的随机 x。这种随机访问在某些情境下很自然，例如，我们掌握了从某个项目毕业的个体的信息，却没有未毕业者的信息。

定理 1.2（定理 3.15 的非正式表述）。设 $\alpha > 0$，且设 $f\colon \{0,1\} \to \{0,1\}$ 是一个函数，使得 $\Pr[f(x) = 1] = \alpha$。设 $\mathcal{D}$ 是一个区分器集合，其中每个区分器 $D \in \mathcal{D}$ 都与一个集合 $S_D \subseteq [N]$ 相关联，并且当且仅当 $x \in S_D$ 时接受 x。如果存在针对 $\mathcal{D}$ 的弱无关学习器，那么就存在一个运行时间为 $\mathrm{poly}(n)$ 的学习算法 L，它从集合 $\{x | f(x) = 1\}$ 中接收随机元素，并输出一个模型 M，该模型对于所有 $D \in \mathcal{D}$ 而言都与 f 不可区分。

当存在一个针对区分器集合 $\mathcal{D}$ 的弱无关学习器，且其分布为随机支持集元素的分布（即随机 x，使得 $f(x) = 1$）时，该定理成立。我们在第 3.3.1 节中证明，如果区分器集合 $\mathcal{D}$ 在标准样本访问分布上具有一个弱无关学习器，且该学习器是统计查询算法，那么也存在一个针对 $\mathcal{D}$ 的学习器，其分布同样为随机支持集元素的分布。

该定理的证明类似于经典的提升论证，但额外增加了一步：学习器需要使模型的支持集大小大致保持为与 f 的支持集大小相同。此步骤是必要的，因为在随机支持集元素的分布下，只有当 f 与模型的支持集大小大致相等时，才能保证提升算法有效。

在随机支持集元素的分布下学习一个对象，对于 稀疏对象 的情况可能非常有用。对于稀疏函数 f，如果我们选择一个均匀的 $x \in \{0,1\}$，那么以高概率有 $f(x) = 0$，学习器不可能仅凭形式为 $(x,\ f(x))$ 的随机样本学到任何非平凡的信息。遗憾的是，上述定理不适用于稀疏函数，但在下一部分中，我们将展示如何利用该定理学习不同的稀疏对象——稀疏图。

在没有稠密子图的情况下学习稀疏图

假设 $G = ([N],\ E)$ 是一个图，用其邻接矩阵的 N^2 长度字符串表示。在这种表示中，从 G 接收一条随机边，相当于从表示其邻接矩阵的函数中接收一个随机支持集元素。因此，定理 1.2 意味着，我们可以学习一个关于 G 的模型，使其对于一组 $\mathcal{D}$ 区分器而言不可区分，而这些区分器具有弱无关学习器。该定理仅适用于含有常数比例 1 项的函数 f，这对应于稠密图。那么稀疏图呢?

学习稀疏图，或一般意义上的稀疏对象，是一项极具挑战性的任务，因为其定义域极其庞大。弱正则性引理 [FK99] 的误差与 N^2 成正比，这对于稀疏图而言过大（在这种误差下，空图与稀疏图无法区分）。因此，在稀疏图的情形下，更自然的做法是要求区分器在接收到随机边的分布下具有 ε 误差。在这种分布下，区分器的误差会随边的数量而变化。我们针对一类特定的稀疏图展示了一种学习器

，即不含稠密子图的稀疏图。我们注意到，随机稀疏图不含稠密子图，因此许多图都具有这一性质。

定理 1.3（定理 4.3 的非正式表述）。设 $G = ([N], E)$ 是一个没有稠密子图的稀疏图。设 $\mathcal{D}$ 是一个区分器集合，其中每个区分器 $D \in \mathcal{D}$ 都与两个集合 $U_D, V_D \subseteq [N]$ 相关联，并且当 $u \in U_D, v \in V_D$ 时接受边 (u, v)。如果存在一个针对 $\mathcal{D}$ 的弱无关学习器，那么就存在一个运行时间为 $\mathrm{polylog}(N)$ 的学习算法 L，它从 G 接收随机边，并输出一个模型 M，该模型对于所有 $D \in \mathcal{D}$ 而言都与 f 无法区分。

学习器输出的模型是稠密的，也就是说，该模型输出的图有 $\Theta(N^2)$ 条边。这样做是出于技术上的原因——为了在训练模型时能够使用拒绝采样。这引出了一个问题：是否存在一个与我们的稀疏图 G 无法区分的稠密图？这个问题的答案取决于 G；在断言 4.5 中，我们证明，如果一个稀疏图 G 有一个非常稠密的子图，那么就不存在与 G 无法区分的稠密图。

该定理的证明分为两部分。第一部分中，我们证明，对于每个不含稠密子图的稀疏图G，都存在一个与其不可区分的稠密图 H G。在这部分证明中，我们对稀疏图应用强正则性引理 [KR03、Sco11] on G，并利用由此得到的划分构造出不可区分的稠密图 H。这部分证明是存在性的，我们确实知道如何高效地找到 H，因为强正则性引理没有用于高效寻找该划分的算法。无法使用弱正则性引理或其变体 [FK99],，因为其误差过大。在证明的第二部分中，我们将学习 G 归约为学习 H，并证明所得模型 M 与 G 对所有区分器 $D \in \mathcal{D}$而言不可区分。

关于学习生成模型的其他结果

在这项工作中，我们还展示了在其他若干设定下的不可区分模型

- 令 $f: \{0,1\}^n \to \{0,1\}^n$ 是一个函数。学习此类函数比学习二值函数更困难，因为较大的定义域使其成为 f 稀疏对象（例如，将其视为图时，它是一个出度为一的图）。对于此类函数，我们证明存在一个学习器：给定来自分布 $(x, f(x))$ 的样本，输出一个模型，该模型相对于以下区分器集合 $\mathcal{D} = \{(S_D, j_D) | S_D \subset \{0,1\}^n, j_D \in [n]\}$ 是不可区分的，使得 $D = 1 \iff x \in S_D, f(x)_j = 1$。相关内容见第 3.4 节。

- 在第 4.1 节中，我们将函数定理应用于稠密图的邻接矩阵$G = ([N], E)$。对于 $\mathcal{D}$ 具有弱无偏学习器的区分器集合，我们有一个高效学习器，当 G 为以下情况时，该学习器会输出一个不可区分模型：

  1. 学习器接收随机邻接矩阵条目时的稠密图。2. 边总数固定的稠密图 $m = \Theta(N^2)$。3. 固定出度的有向图 $m = \Theta(N)$。4. 学习器接收随机边时的稠密图。

- 对于一个出度为常数的有向图 $G = (\{0,1\}^n, E)$，我们可以将每个出度视为一个函数 $f_i: \{0,1\}^n \to \{0,1\}^n$。对于同一组区分器

在长度保持函数（此列表中的第一项）的情况下，我们可以处理这组区分器，并提供一个高效学习器。请参见第 4.3 节。

- 对于均匀度固定的无向图，我们在第 4.4 节给出了一个针对不可区分模型的学习算法，不过它只适用于一个范围相对有限的区分器集合。

不可能性结果

如前所述，我们学习生成模型的目标与 [GGN10] 实现大型随机对象的目标密切相关，但关键区别在于，我们假设真实对象是一个未知的单一对象 B^*，而 [GGN10] 考虑的是对象上的已知均匀分布。这意味着我们需要额外的学习过程来收集关于 B^*的信息；我们将在第 5 节证明，只有当区分器类可被高效学习时，我们才可能高效地学习生成模型。

除了学习这一要求外，我们的设定在许多方面都比 [GGN10]中的设定更具挑战性。我们通过另外两个不可能性结果来说明这一点：无法欺骗具有条目访问能力的区分器，也无法欺骗比模型更强的区分器。当我们考虑伪随机函数 f_s 时，对于能够访问函数条目（可任意询问字符串 x 并获得 $f_s(x)$）的区分器而言，该函数与均匀分布不可区分。此外，尽管 f_s 可在固定的多项式时间内计算，区分器却可以在任意多项式时间内运行（在合理的假设下，甚至可以在指数时间内运行）。[GGN10]及后续研究都沿用了这两个性质——模型对于计算能力更强且具有模型条目访问能力的区分器而言不可区分。在第 5 节中，我们将论证这两个性质在我们的设定下都无法实现。

关于无法在条目访问的情况下愚弄区分器，定理 5.1 中我们给出了一个类 $\mathcal{D}$ 的例子，其中对每个输入 $x \in \{0,1\}^n$ 都包含一个区分器 D_x：该区分器会查询函数 f 的函数值 $f(x)$，并且当且仅当 $f(x) = 1$ 时输出“接受”。我们论证，对于区分器集合 $\mathcal{D}$，任何与真实 f^* 无法区分的模型 M 都必须与 f^* 非常接近。由于 f^* 的大小呈指数级，且 f^* 未知，因此没有高效学习器能够输出接近 f^* 的模型。我们还利用 [TTV09], 中的一个思路，展示了一个区分器和真实函数 f^* 的例子：该区分器能够将 f^* 与任何复杂度相对于该区分器较低的模型 M 区分开来（定理 5.2 ）。这凸显出，在我们的设定中，生成模型及构建该模型的学习器，其计算能力必须与区分器相当或更强。

1.3 相关工作

如第 1 节所述，1, [GGN10] 提出了为随机对象创建不可区分实现的问题。[GGN10] 以及后续工作 [NNT05、NN07]还针对稠密图和稀疏图或函数，展示了一系列积极结果，其中区分器具有多种真实性条件和访问模型。

生成模型与不可区分性之间的联系，已通过生成对抗网络（GAN） [GPAM$^+$14、AGL$^+$17]的发明得到体现。直观地说，GAN 的训练目标是模仿对象的分布（例如图像）。生成器与判别器协同训练，而判别器可以被理解为区分器。经过一系列轮次的训练，生成器不断学习欺骗判别器，而判别器则不断学习识破生成器。

GAN 突出了生成模型与不可区分性之间的联系 [Imp17], ，但它们并不自然地适用于我们的框架，因为对 GAN 的描述更直接地基于两个分布的不可区分性。

不可区分性与学习理论之间的联系已在许多先前的研究中得到确立（例如， [TTV09] 运用了学习理论中的提升技术）。近来，在算法公平性的背景下，多重校准和结果不可区分性的概念大幅拓展了学习理论与不可区分性之间的关系 [HJKRR18, DKR+21]；全能预测器和普适适应性的概念则将这一关系应用于学习和统计推断 [GKR+22, KKG+22, HLNRY22, GHK+23, KP23]；此外，从学习理论的视角研究不可区分性的样本复杂度，也催生了揭示其中错综而令人振奋的联系的研究 [HPR22, HP23]。

我们可以将所研究的学习情境视为学习器与区分器之间的双人零和博弈：学习器的目标是输出一个与某个对象不可区分的模型，而区分器则试图分辨输入与模型。在这一情境中，极小极大定理与正则性引理定理之间存在联系。这类定理证明，复杂对象可以表示为若干较简单对象的函数 f， g_1， $\dots g_t$ 而在我们的情境中，这些较简单的对象代表区分器 [TTV09, VZ13]。已有研究改进了相关参数，并将这类定理用于密码学应用 [VZ13, JP14, Sk´o16a, Sk´o16b, Sk´o17, CCL18]。在本研究中，我们考虑的情境略有不同：我们假设对象 f 是复杂且未知的，学习器必须对其进行学习。定理 1.3 的证明包含一个存在性中间步骤，其结构类似于弱正则性引理定理；但由于所需误差过小，我们通过稀疏强正则性引理来推导该步骤。

2 预备知识

本文始终关注函数和图等对象的学习，尤其关注规模为指数级的对象（例如定义域为指数规模的函数，以及顶点和边数量呈指数级的图）。我们通常用 B 表示一个对象，用 $\mathcal{B}$ 表示一类对象。我们将对象 B 视为一个函数$B\colon Q \to \Delta_A$，它将查询 $q \in Q$ 映射为定义在 A上的答案分布 $B(q)$。

函数。 当对象是函数 $f\colon X \to Y$时，我们考虑三种访问类型。对于样本访问， B返回一个随机对 $(x,\ f(x))$。对于支撑集访问，它返回一个随机的 x，使得$f(x) = 1$。对于条目访问，在查询 x后， B返回 $f(x)$。

定义 2.1 (函数诱导的样本访问对象)。 设 $f : X \to Y$是一个函数， $B :Q \to \Delta_A$是一个对象。若 $Q = \{\perp\}$、$A = X \times Y$，且 $B(\perp)$是$(x, f(x)) \in A$的分布，其中 x从 X中均匀抽取，则称 B为由 f诱导的样本访问对象 。

定义 2.2 （函数诱导的支持集访问对象）。 设 $f\colon X \to \{0,1\}$ 是一个二值函数。我们将 f 的支持集定义为 $\mathsf{supp}(f) := \{x \in X\colon f(x) = 1\}$。设 $B\colon Q \to \Delta_A$ 是一个对象。假设 $\mathsf{supp}(f) \neq \emptyset$，我们称 B 为由 f诱导的支持集访问 对象，如果 $Q = \{\perp\}$、$A = X$，且 $B(\perp)$ 是 $\mathsf{supp}(f) \subseteq X$上的均匀分布。

定义 2.3 (函数诱导的条目访问对象)。设 $f\colon X \to Y$ 为一个函数，且设 $B\colon Q \to \Delta_A$ 为一个对象。若 $Q = X, A = Y$，且对于每个 $q \in Q$, $B(q)$ 都是一个单点分布，使得 $a \sim B(q)$ 确定性地等于 $f(q)$，则称 B 是由 f诱导的条目访问对象。

本文展示了关于学习具有样本访问和支撑集访问的函数生成模型的正面结果（第 3 节），而对于条目访问则展示了不可能性结果（第 5 节）。这种分离主要是因为条目访问使区分器更强，从而使不可区分性更难实现（见下文定义 2.6 和 2.7）。

图。 对于图 $G = (V, E)$，假设 V 的规模呈指数级，我们定义两种访问类型：样本访问对应于随机的邻接矩阵条目，支持集访问对应于图中的随机边。

定义 2.4 (图诱导的样本访问对象**)。** 设 $G = (V, E)$ 是一个有向图或无向图，且设 $B : Q \to \Delta_A$ 是一个对象。若 $Q = \{\perp\}$、$A = V \times V \times \{0, 1\}$，且 $B(\perp)$ 是 $(u, v, y) \in A$ 的分布，其中 (u, v) 从 $V \times V$ 中均匀抽取（若满足 $(u, v) \in E$），否则为 $y = 0$，则称 B 为由 G 诱导的样本访问对象。

定义 2.5 (图诱导的支持集访问对象**)。** 设 $G = (V, E)$ 是一个有向图或无向图，且设 $B : Q \to \Delta_A$ 是一个对象。假设 $E \neq \emptyset$，若 $Q = \{\perp\}$、$A = V \times V$，且 $B(\perp)$ 是 $E \subseteq V \times V$ 上的均匀分布，则称 B 为由 G 诱导的支持集访问对象。

不可区分性。 本文设计的每个学习器都可以访问一个真实对象B^*，其目标是输出一个与 B^*不可区分 的对象 B 。在许多情况下，学习器输出的不只是单个对象 B，而是对象上的一个分布，我们将这类分布称为模型。下面我们正式定义不可区分性的概念。

定义 2.6 （区分器）。 区分器 D 是一种算法，在获得对对象 B ： $Q \to \Delta_A$的访问权限后，会输出“接受” 或“拒绝”。也就是说，区分器可以向模型发起查询 $q \in Q$ ，并且对于每个查询 q ，区分器都会收到一个从 B（q） $\in \Delta_A$中独立抽取的答复 $a \in A$ 。我们允许区分器 D 本身具有随机性，并用随机变量 D^B 表示区分器 D 在获得对 B的访问权限时输出“接受”、“拒绝”} 的结果。

定义 2.7 （不可区分性）。 设 B^* ： $Q \to \Delta_A$ 是一个对象，设模型 M 是对象 B 上的分布： $Q \to \Delta_A$。若对于区分器 D 而言，模型 M 与对象 B^*是 ε-不可区分的，则称模型 M 关于对象 B^*对区分器 D 是 ε-不可区分的，如果

$$|\Pr[D^{B^*} = \text{“accept”}] - \mathbb{E}_{B \sim M}[\Pr[D^B = \text{“accept”}]]| \leq \varepsilon. \tag{1}$$

如果对每个 $D \in \mathcal{D}$均满足（1），则称模型 M 关于对象 B^* 对区分器类 $\mathcal{D}$ 是 ε-不可区分的。

真实性。 除了不可区分性之外，模型的另一个理想属性是由 [GGN10] 引入的真实性。真实性要求从模型生成的每个对象都满足某个特定的（通常是全局的）性质，我们使用对象类 $\mathcal{B}$将其形式化：

定义 2.8 （真实性）。 如果模型 M 关于对象类 是真实的，我们称

$$\Pr_{B \sim M}[B \in \mathcal{B}] = 1.$$

实现。 我们的目标是设计高效学习器，因此不能指望学习器显式地输出模型 M，尤其是在从 M 中抽取的对象非常庞大时。相反，我们的学习器会输出模型的一种高效实现，定义如下。

定义 2.9（普通实现）。 对于 $\ell \in \mathbb{Z}_{\geq 0}$，设 T 是一个随机化算法，它以（r, q）$\in \{0,1\}^{\ell} \times Q$ 为输入，并输出 T（r, q）$\in A$。如果对于每个种子 $r \in \{0,1\}^{\ell}$，都存在一个对象 B_r： $Q \to \Delta_A$，使得 T 是具有种子长度 ℓ 的模型 M 的普通实现，则称 T 为该模型的普通实现。

1. 对于每个 $q \in Q$， T（r, q）服从 B_r（q）所指定的分布，其中 T（r, q）中的随机性来自算法 T 的内部随机性；
2. 当 r 从 $\{0,1\}^{\ell}$ 中均匀抽取时， B_r 服从 M 所指定的分布。

虽然我们的目标是输出一个具有多项式长度种子的普通实现，但遵循 [GGN10], 中的方法，先构造基于随机预言机的实现，再利用引理 2.1 将其转换为普通实现，会更加方便。

定义 2.10（随机预言机实现）。 设 T 是一个随机化算法，它以函数 $r\colon \{0,1\}^* \to \{0,1\}$ 作为预言机。对于输入 $q \in Q$，算法 T输出 $T^r(q) \in A$。如果对于每个 $r\colon \{0,1\}^* \to \{0,1\}$，都存在一个对象 $B_r\colon Q \to \Delta_A$，使得

1. 对于每个 $q \in Q$， $T^r(q)$ 的分布与 $B_r(q)$相同，其中 $T^r(q)$ 中的随机性来自算法 T的内部随机性；
2. B_r 的分布 与 M 相同，其中 r 是从 $\{0,1\}^*$ 到 $\{0,1$的均匀随机函数 $\}$.

引理 2.1（GGN 中的定理 2.9 [GGN10]）。 假设单向函数存在。则存在一个算法 H，具有以下性质。设 $\mathcal{D}$ 是一个区分器类，其中每个 $D \in \mathcal{D}$都是大小至多为 W 的电路，某个 $W \geq 1$满足该条件。设 T 是模型 M的随机预言机实现，其电路复杂度至多为 W。给定 W、T以及任意一个 $(0$、$1)$ 作为输入，算法 H在 $\mathrm{poly}(W$、$1/\varepsilon)$ 时间内运行，并输出模型 M'的一个普通实现 T'，其中 T' 的种子长度和电路复杂度均为 $\mathrm{poly}(W$、$1/\varepsilon)$，并且

$$|\mathbb{E}_{B' \sim M'} \Pr[D^{B'} = \text{“accept”}] - \mathbb{E}_{B \sim M} \Pr[D^{B} = \text{“accept”}]| \leq \varepsilon \quad \textit{for every } D \in \mathcal{D}.$$

此外，如果 M 关于对象类 $\mathcal{B}$ 是真实的，那么 M' 关于 $\mathcal{B}$ 也是真实的。

引理 2.1 可以通过使用伪随机函数来模拟随机预言机得以证明。

学习。 我们在下面的定义中描述了旨在设计的学习器。

定义 2.11（学习器）。设 $\mathcal{B}$ 为一类对象 B： $Q \to \Delta_A$，且 $\mathcal{D}$ 为一类区分器。对于类 $\mathcal{B}$，关于 $\mathcal{D}$ 的（ε, δ）学习器 L 是一个满足以下性质的算法。对于任意 $B^* \in \mathcal{B}$，在能够访问 B^*的情况下，该学习器输出一个模型 M 的实现 T，使得 M 以至少 $1 - \delta$的概率关于 $\mathcal{D}$与 B^* 是 ε-不可区分的。

其他记号。对于 $v \in \mathbb{R}$，我们将 $\mathsf{cap}(v)$ 定义为把它的值截断到 $[0,1]$，，即：

$$\mathsf{cap}(v) = \begin{cases} v, & \text{if } 0 \le v \le 1; \\ 1, & \text{if } v > 1; \\ 0, & \text{if } v < 0. \end{cases}$$

给定一个值列表（v_1，……，v_t），我们通过对列表求和，并在每次迭代中将结果截断至 $[0,1]$，来定义 $\mathsf{Lcap}(v_1$，……，$v_t)$。我们将其递归地形式化定义如下：

$$\begin{aligned} \mathsf{Lcap}(v_1) &= \mathsf{cap}(v_1), \\ \mathsf{Lcap}(v_1, \ldots, v_t) &= \mathsf{cap}(\mathsf{Lcap}(v_1, \ldots, v_{t-1}) + v_t). \end{aligned} \tag{2}$$

3 学习定义域呈指数级的函数

本节的目标是高效地学习一个生成模型，使其对于一类区分器 $\mathcal{D}$ 而言，与目标函数 $f^*: X \to Y$ 无法区分。我们允许函数的定义域 X 具有指数级大小 $N := |X|$，并要求学习器的运行时间为 $\mathsf{polylog}(N)$。这意味着学习器无法读取整个函数 f^*，只能通过随机样本访问它。本文中，我们的学习器都会输出模型 M 的高效随机预言机实现 T；根据引理 2.1，该实现可以转换为高效的普通实现。

3.1 学习可通过样本访问的二元函数

我们首先研究目标对象 B^* 是由二值函数 f^* 诱导的样本访问对象的情形：$X \to \{0,1\}$（定义 2.1）。我们假设每个区分器 $D \in \mathcal{D}$ 都满足以下条件：当获得由函数 $f: X \to \{0,1\}$ 诱导的样本访问对象 B 的访问权限时，区分器会提出一次查询 $\perp$，收到回答 $a = (x, y) \sim B(\perp)$，并输出 $D(x, y) \in$ {“接受”，“拒绝”}。我们允许区分器本身是随机化的，并且每个区分器 D 定义了一个函数 $g_D: X \to [-1,1]$，使得

$$g_D(x) = \Pr[D(x,1) = \text{“accept”}] - \Pr[D(x,0) = \text{“accept”}] \quad \text{for every } x \in X.$$

我们使用以下断言，将区分器 $D \in \mathcal{D}$ 与函数 $g_D: X \to [-1,1]$ 联系起来：

断言 3.1。对于每个区分器 $D \in \mathcal{D}$，以下等式对每个 $x \in X$ 和 $y_1, y_2 \in \{0,1\}$ 都成立：

$$\Pr[D(x,y_1) = \text{“accept”}] - \Pr[D(x,y_2) = \text{“accept”}] = y_1 g_D(x) - y_2 g_D(x).$$

通过考虑 $(y_1, y_2) \in \{0,1\} \times \{0,1\}$ 的四种可能选择，可以轻松证明这一断言。

正如我们稍后在第 5 节中所示，有必要对区分器施加某些可学习性假设。为此，我们假设函数类 $\mathcal{G} := \{g_D : D \in \mathcal{D}\}$ 存在一个**审计器**，定义如下：

定义 3.1（审计器）。令 $\mathcal{D}$ 和 $\mathcal{G}$ 如上所定义。若算法 Λ 满足以下性质，我们称其为针对 $\mathcal{G}$ 的 $(\varepsilon, \gamma, \delta)$ 审计器。给定由函数 B^* 诱导的样本访问对象 B^*

，并将预测器 p： $X \to [0,1]$ 作为预言机，若存在$g \in \mathcal{G}$ 使得

$$|\mathbb{E}[f^*(x)g(x)] - \mathbb{E}[p(x)g(x)]| \geq \varepsilon,$$

then Λ 个输出 $\hat{g}$： $X \to [-1,1]$ 满足以下条件的概率至少为 $1-\delta$ ：

$$\mathbb{E}[f^*(x)\hat{g}(x)] - \mathbb{E}[p(x)\hat{g}(x)] \geq \gamma.$$

上述定义的审计器可以视为函数类 $\mathcal{G}$的弱无关学习器。当定义域 X 是大小为 $N = 2^n$的 $\{0,1\}^n$ 时，许多函数类都允许高效的审计器，其运行时间为 $\mathrm{poly}(n) = \mathrm{polylog}(N)$。利用针对 $\mathcal{G}$的审计器，我们证明以下定理:

定理 *3.2*。设区分器类 $\mathcal{D}$ 和函数类 $\mathcal{G}$ 如上所定义。设$\varepsilon, \gamma, \delta, \delta' \in (0, 1/2)$ 是满足 $\gamma \leq \varepsilon$和 $\delta' \leq c\delta\gamma^2$ 的参数，其中 $c > 0$是充分小的绝对常数。设 $\mathcal{B}$ 是由二值函数$f\colon X \to \{0,1\}$诱导的样本访问对象类。设 Λ 是针对 $\mathcal{G}$ 的 $(\varepsilon, \gamma, \delta')$-审计器（定义 *3.1*）。则存在一个针对 $\mathcal{B}$ 、关于 $\mathcal{D}$ 的 (ε, δ)-学习器 L 。此外，若审计器 Λ 的运行时间至多为 W_1 ，且其输出的函数的电路规模始终至多为 W_2，则学习器 L 的运行时间为 $\mathrm{poly}(\gamma^{-1}, \log(\delta^{-1})\ W_1)$，并且始终输出电路复杂度为 $\tilde{O}(\gamma^{-2}W_2)$ 的实现。

我们应用算法公平性领域中的以下结果来证明定理 3.2：

引理 3.3 （学习多重准确预测器 [TTV09, HJKRR18, KGZ19]）。令区分器类 $\mathcal{D}$ 和函数类 $\mathcal{G}$ 如上定义。令 ε、γ、δ、$\delta' \in$（0、1/2）为满足 $\gamma \leq \varepsilon$ 和 $\delta' \leq c\delta\gamma^2$ 的参数，其中绝对常数 $c > 0$ 足够小。令 Λ 为针对 $\mathcal{G}$ 的一个（ε、γ、δ'）审计器。给定对由函数 f^*： $X \to \{0,1\}$ 诱导的样本访问对象 B^* 的访问权限，存在一个学习器 L，它输出一个预测器 p： $X \to [0,1]$，使得其成功概率至少为 $1-\delta$，

$$|\mathbb{E}[f^*(x)g(x)] - \mathbb{E}[p(x)g(x)]| \leq \varepsilon \quad \textit{for every } g \in \mathcal{G}. \tag{3}$$

此外，如果审计器 Λ 的运行时间至多为 W_1，并且始终输出电路规模至多为 W_2 的函数，那么学习器 L 的运行时间为 $\mathrm{poly}(\gamma^{-1}, \log(\delta^{-1})\ W_1)$，且始终输出电路复杂度为 $\tilde{O}(\gamma^{-2}W_2)$ 的实现。

定理 *3.2* 的证明。令 B^* 为由二值函数 $f^*\colon X \to \{0,1\}$ 诱导的样本访问对象。只需将引理 3.3 中的预测器 p 转换为模型 M 的高效随机预言机实现 T，使其相对于 $\mathcal{D}$ 与 B^* ε-不可区分。考虑一个随机函数 $f\colon X \to \{0,1\}$，使得对每个 $x \in X$，函数值 $f(x)$ 都独立地从伯努利分布 $\mathrm{Ber}(p(x))$ 中抽取。根据断言 3.1，对于每个区分器 $D \in \mathcal{D}$，我们有

$$\begin{aligned} &|\mathbb{E}_x[\Pr[D(x, f^*(x)) = \text{“accept”}]] - \mathbb{E}_{f,x}[\Pr[D(x, f(x)) = \text{“accept”}]]| \\ &= |\mathbb{E}_x[f^*(x)g_D(x)] - \mathbb{E}[p(x)g_D(x)]|. \end{aligned} \tag{4}$$

因此，若我们选择模型 M 为由随机函数 f 诱导的样本访问对象 B_f 上的分布，那么 (3) 和 (4) 意味着，关于 $\mathcal{D}$， M 与 B^* 是 ε-不可区分的。要使用随机预言机 r 高效地实现 M：对于每个 $x \in X$，我们指定一个集合 $S_x \subseteq \{0,1\}^*$，使得对于 $x \neq x'$，有 $S_x \cap S_{x'} = \emptyset$。我们首先从 X 中均匀随机生成一个 $x \in X$，然后利用来自 S_x 的随机性抽取 $y \sim \mathrm{Ber}(p(x))$。我们返回答案 $(x,\ y)$。基于以下原因，这是 M 的一种实现。对于每个固定的 r，存在一个函数 f： $X \to \{0,1\}$，使得答案 $(x,\ y)$ 总是满足 $y = f(x)$，且 x 的分布在 X 上均匀。此外，当 r 是随机的时，给定 x 的条件下， y 的条件分布为 $\mathrm{Ber}(p(x))$。

□

3.2 保留支撑集大小的真实性学习

生成模型的一些理想性质是全局性质，单靠计算能力受限的区分器无法保证这些性质。这促使 [GGN10]引入了真实性 的概念，以确保不可区分性之外的此类全局性质。这里我们关注二值函数 f 的一个自然的全局性质： $X \to \{0,1\}$：其支持集 $\mathrm{supp}(f)$ 的大小 $= \{x \in X : f(x) = 1\}$。我们在第 3.1 节中使用引理 3.3 中的多重准确预测器 p 构造的模型 M，可能生成支持集大小不同于目标函数 f^* 的函数 f。事实上，即使 $\sum_{x\in X} p(x) = |\mathrm{supp}(f^*)|$，每个条目 $f(x)$ 都独立地从 $\mathrm{Ber}(p(x))$ 中抽取的随机函数 f，也不一定满足 $|\mathrm{supp}(f)| = |\mathrm{supp}(f^*)|$。现在我们展示一种高效学习器，它能输出真实模型，并保留所生成函数的支持集大小。

定理 *3.4.* 设区分器类 $\mathcal{D}$ 和函数类 $\mathcal{G}$ 如第 *3.1* 节中所定义。设 $\varepsilon, \gamma, \delta, \delta' \in (0, 1/2)$ 是满足 $\gamma \le \varepsilon$ 和 $\delta' \le c\delta\gamma^2$ 的参数，其中绝对常数 $c > 0$ 足够小。设 $\mathcal{B}$ 是由满足 $|supp(f)| = k$ 的二值函数 $f: X \to \{0,1\}$ 所诱导的样本访问对象类，其中 $X = \{0,1\}^n$。设 Λ 是针对 $\mathcal{G}$ （定义 *3.1*）的 $(\varepsilon, \gamma, \delta')$-审计器。则存在一个针对 $\mathcal{B}$ 、相对于 $\mathcal{D}$ 的 (ε, δ)-学习器 L ，且该学习器始终输出模型 M 的随机预言机实现，该模型相对于 $\mathcal{B}$ 具有真实性。此外，若审计器 Λ 的运行时间至多为 W_1 ，且始终输出电路规模至多为 W_2 的函数，则学习器 L 的运行时间为 $poly(n, \gamma^{-1}, log(\delta^{-1}) W_1)$，并始终输出电路复杂度为 $poly(n, \gamma^{-1}, W_2)$ 的实现。

我们使用引理 3.3 和以下引理证明定理 3.4：

引理 3.5 （基于预测器的定权重实现）。 对于正整数 n，定义 $X = \{0,1\}^n$ 。令 k 为满足 $0 \le k \le |X|$ 的整数，并令 $\varepsilon \in$ （0,1） 为一个参数。存在一个在 n、ε^{-1} 上运行时间为多项式的实现 T，使得 T以任意预测器 p： $X \to [0,1]$ 作为预言机，并实现一个模型 M ，其中每个对象 $B \sim M$ 都是由二值函数 f_B： $X \to \{0,1\}$诱导的样本访问对象。此外，对于每个预测器 p： $X \to [0,1]$，所得模型 M 满足

1. 概率$_{B\sim M}[\sum_{x\in X} f_B(x) = k] = 1$;
2. $\sum_{x\in X} |\mathbb{E}_{B\sim M} f_B(x) - p(x)| \le |k - \sum_{x\in X} p(x)| + \varepsilon|X|$.

我们首先使用引理 3.5 证明定理 3.4，然后证明引理 3.5。

定理 *3.4* 的证明。 我们首先注意到，利用 Λ，我们可以为 $\mathcal{G} \cup \{g_1\}$ 构造一个审计器 Λ' ，其中 g_1 是常数 1 函数：对于每个 $x \in X$， $g_1(x) = 1$ 。这是因为我们可以使用从 X 均匀抽取的 x 的随机样本（x， $f^*(x)$ $p(x)$）准确估计 $\mathbb{E}[f^*(x)g_1(x)] - \mathbb{E}[p(x)g_1(x)]$。

设 B^* 为由函数 f^* 诱导的样本访问对象：满足 $\mathrm{supp}(f^*) = k$ 的 $X \to \{0,1\}$。根据引理 3.3，我们可以学习一个预测器 p： $X \to \{0,1\}$ ，使得

$$|\mathbb{E}_x[f^*(x)g(x)] - \mathbb{E}_x[p(x)g(x)]| \le \varepsilon/4 \quad \text{for every } g \in \mathcal{G} \cup \{g_1\}. \tag{5}$$

特别地，取 $g = g_1$，我们有

$$\left| k - \sum_{x\in X} p(x) \right| = |X| \cdot |\mathbb{E}[f^*(x)g_1(x)] - \mathbb{E}[p(x)g_1(x)]| \le \varepsilon|X|/4.$$

我们将引理 3.5 应用于 p，从而得到模型 M 的一种实现 T，使得每个对象 $B \sim M$ 都是由二值函数 $f_B : X \to \{0,1\}$ 所诱导的样本访问对象。引理 3.5 允许我们确保

$$\Pr_{B\sim M}[|\mathsf{supp}(f_B)| = k] = 1, \tag{6}$$

$$\sum_{x\in X} |\mathbb{E}_{B\sim M} f_B(x) - p(x)| \le \varepsilon |X|/2. \tag{7}$$

剩下要证明的是，M 关于 $\mathcal{D}$ 与 B^* 是 ε-不可区分的，并且 M 关于$\mathcal{B}$具有真实性。真实性可由 (6) 立即得出。下面证明不可区分性。对于每个 $g \in \mathcal{G}$，我们有

$$\begin{aligned} |\mathbb{E}_x[p(x)g(x)] - \mathbb{E}_{B\sim M}\mathbb{E}_x[f_B(x)g(x)]| &\le \left|\mathbb{E}_x\left[g(x)\,\mathbb{E}_{B\sim M}[f_B(x) - p(x)]\right]\right| \\ &\le \mathbb{E}_x\,|\mathbb{E}_{B\sim M} f_B(x) - p(x)| \\ &\le \varepsilon/2, \end{aligned} \tag{8}$$

wh这里，最后一个不等式由 (7) 得出。结合 (5) 和 (8)，对于每个 $g \in \mathcal{G}$ ，

$$|\mathbb{E}_x[f^*(x)g(x)] - \mathbb{E}_{B\sim M}\mathbb{E}_x[f_B(x)g(x)]| \le \varepsilon.$$

根据断言 3.1，这意味着 M 关于 $\mathcal{D}$与 B^* 是 ε-不可区分的。 □

在接下来的子章节中，我们从一个效率低下的实现开始，逐步构建效率越来越高的实现，以此证明引理 3.5；其中，我们用 $N = 2^n$ 表示定义域大小 $|X|$。为方便起见，有时我们会用 $[N] := \{1,\dots,N\}$ 代替 $X = \{0,1\}^n$作为定义域，此时允许 N 不一定是 2 的幂。

3.2.1 步骤 1：时间复杂度为 $2^{O(N)}$ 的精确模型实现

我们证明以下引理。它可视为引理 3.5 的一个版本，其中模型具有 $\alpha = \varepsilon = 0$，且不满足效率要求：

引理 3.6。设 $p\colon [N] \to [0,1]$ 是一个满足 $\sum_{x\in[N]} p(x) = k$ 的预测器，其中 k 是某个整数。存在一个定义在 二元 函数 $f\colon [N] \to \{0,1\}$ 上的分布 μ，使得

1. $\Pr_{f\sim\mu}[\sum_{x\in[N]} f(x) = k] = 1$，且 2.
对每个 $x \in [N]$，$\mathbb{E}_{f\sim\mu}[f(x)] = p(x)$。

对于两个函数 $p, p'\colon [N] \to \mathbb{R}$，我们将它们的内积定义为 $\langle p, p'\rangle := \sum_{x\in[N]} p(x)p'(x)$。引理 3.6 的证明基于以下断言：

断言 3.7. 设 N、k为满足 $N > 0$和 $0 \le k \le N$的整数。设 $p\colon [N] \to [0,1]$ 为满足$\sum_{x\in[N]} p(x) = k$的函数。对于任意函数 $w\colon [N] \to \mathbb{R}$，其函数值按 $w(\sigma(1)) \ge \cdots \ge w(\sigma(N))$ 的顺序排列，其中 $\sigma\colon [N] \to [N]$为一个双射；令 $f\colon [N] \to \{0,1\}$ 为满足 $f(\sigma(1)) = \cdots = f(\sigma(k)) = 1$且 $f(\sigma(k+1)) = \cdots = f(\sigma(N)) = 0$的函数。那么

$$\langle w, f\rangle \ge \langle w, p\rangle.$$

证明。通过重新排列坐标，不失一般性，可假设对每个 $i \in [N]$，都有 $\sigma(i) = i$。定义 $w(N+1) = 0$，我们有

$$\begin{aligned}\langle w, p\rangle &= \sum_{i\in[N]} w(i)p(i) \\ &= \sum_{i\in[N]} \sum_{j=i}^{N} (w(j) - w(j+1))p(i) \\ &= \sum_{j\in[N]} \sum_{i\in[j]} (w(j) - w(j+1)p(i) \\ &= \sum_{j\in[N]} (w(j) - w(j+1)) \sum_{i\in[j]} p(i)\end{aligned}$$

并且类似地

$$\langle w, f\rangle = \sum_{j\in[N]} (w(j) - w(j+1)) \sum_{i\in[j]} f(i).$$

显然，对于每个 $j = 1, \ldots\ldots, N$，$\sum_{i\in[j]} p(i) \leq \min\{j, k\} = \sum_{i\in[j]} f(i)$，此外 $\sum_{i\in[N]} p(i) = k = \sum_{i\in[N]} f(i)$。因此，

$$\langle w, p\rangle = \sum_{j\in[N]} (w(j) - w(j+1)) \sum_{i\in[j]} p(i) \leq \sum_{j\in[N]} (w(j) - w(j+1)) \sum_{i\in[j]} f(i) = \langle w, f\rangle. \qquad \square$$

引理 *3.6* 的证明。令 F表示满足 $|\mathrm{supp}(f)| = k$ 的函数 $f\colon [N] \to \{0,1\}$ 的集合。我们首先证明 p 属于 F 的凸包。假设情况并非如此，根据超平面分离定理，存在 $w\colon [N] \to \mathbb{R}$ 和 $b \in \mathbb{R}$，使得 $\langle w, p\rangle - b > 0$，并且对于每个 $f \in F$，都有 $\langle w, f\rangle - b < 0$。根据断言 3.7，存在 $f \in F$，使得 $\langle w, f\rangle \geq \langle w, p\rangle$，因此 $\langle w, f\rangle - b \geq \langle w, p\rangle - b > 0$，矛盾。

我们已经证明，p 属于 F的凸包。这意味着存在一个关于 F的概率分布 μ，使得对于每个 $x \in [N]$，$\mathbb{E}_{f\sim\mu}[f(x)] = p(x)$，如所需。 $\square$

显然，可以通过线性规划在 $2^{O(N)}$时间内显式计算出分布 μ，其中变量是每个 f上的概率质量：$[N] \to \{0,1\}$。

3.2.2 第 2 步：具有 $\mathrm{poly}(N)$ 时间实现的不可区分模型

我们证明下面这个引理 3.6 的变体：允许较小的 ε 误差，但要求使用一个 poly(N) 时间算法从分布 μ 中采样。

引理 3.8。设 $N > 1$ 为正整数，并设 $p\colon [N] \to [0,1]$ 是一个满足$\sum_{x\in[N]} p(x) = k$ 的预测器，其中某个整数为 k。对于任意 $\varepsilon \in (0, 1/2)$，存在一个关于二元 函数 $f\colon [N] \to \{0,1\}$ 的分布，使得

1. 概率$_{f\sim\mu}[\sum_{x\in[N]} f(x) = k] = 1$，以及 *2.* 对于每个 $x \in [N]$，$|\mathbb{E}_{f\sim\mu}[f(x)] - p(x)| \leq \varepsilon$。

此外，μ可以取为在 $f^{(1)}, \ldots\ldots, f^{(T)}$ 上的均匀分布（当 $T = O(\varepsilon^{-2}\log N)$ 时），并且在给定 p作为输入的情况下，列表 $f^{(1)}, \ldots\ldots, f^{(T)}$ 可以在时间 $\tilde{O}(\varepsilon^{-2}N)$ 内计算出来。

我们使用乘法权重算法证明引理 3.8。我们需要以下记号。我们将 Δ_N 定义为满足 $\sum_{x\in[N]} w(x) = 1$的所有函数 $w\colon [N] \to \mathbb{R}_{\geq 0}$ 的集合。对于两个函数 $w,\ w' \in \Delta_N$，假设对每个 $x \in [N]$ 都有 $w'(x) > 0$，我们定义 $\mathsf{D}_{\mathsf{KL}}(w\|w') = \sum_{x\in[N]} w(x)\ln(w(x)/w'(x))$，并约定 $0\ln 0 = 0$。对于函数 $r\colon [N] \to \mathbb{R}$，我们定义 $\|r\|_\infty := \max_{x\in[N]} |r(x)|$。对于函数 $r,\ r'\colon [N] \to \mathbb{R}$，我们定义 $r \diamond r'\colon [2N] \to \mathbb{R}$ 为满足以下条件的函数：

$$(r \diamond r')(x) = \begin{cases} r(x), & \text{if } x \leq N; \\ r'(x-N), & \text{if } x > N. \end{cases}$$

The $r \diamond r'$ 的定义使我们能够陈述下面这个将会很有用的基本事实：

命题 3.9。对于一个任意函数 $r\colon [N] \to \mathbb{R}$，存在 $w \in \Delta_{2N}$ 使得 $\langle w, r \diamond (-r)\rangle = -\|r\|_\infty$.

上述命题使我们能够将 $-\|r\|_\infty$ 表示为一个内积。请注意，将 r 扩展到 $r \diamond (-r)$ 至关重要。如果我们改为考虑形如 $\langle w, r\rangle$ 的内积，其中 $w \in \Delta_N$，那么由于以下限制，该命题将不成立：对于每个 $w \in \Delta_N$ 和 $x \in [N]$，$w(x) \geq 0$。

下面是分析乘法权重算法时使用的一个标准引理：

引理 3.10（乘法权重）。设 $w^{(1)} \in \Delta_N$ 和 $r\colon [N] \to \mathbb{R}$ 是两个函数。假设$w^{(1)}(x) > 0$对每个 $x \in [N]$都成立。定义 $\hat{w}\colon [N] \to \mathbb{R}$ 和 $w^{(2)} \in \Delta_N$，使得对每个 $x \in [N]$，

$$\hat{w}(x) = w^{(1)}(x)e^{-r(x)},$$
$$w^{(2)}(x) = \hat{w}(x) / \sum_{x'\in[N]} \hat{w}(x').$$

那么，对每个 $w \in \Delta_N$

$$\langle w^{(1)} - w, r\rangle \leq \mathsf{D}_{\mathsf{KL}}(w\|w^{(1)}) - \mathsf{D}_{\mathsf{KL}}(w\|w^{(2)}) + \frac{1}{2}\|r\|_\infty^2.$$

引理 *3.8* 的证明。我们初始化 $w^{(1)} \in \Delta_{2N}$，使得对每个 $x \in [2N]$，$w^{(1)}(x) = 1/(2N)$。在第 i 次迭代中，我们执行以下计算。我们将 $w^{(i)}$ 分解为两个函数$w^{(i,+)},\ w^{(i,-)}\colon [N] \to [0,1]$，使得 $w^{(i)} = w^{(i,+)} \diamond w^{(i,-)}$。利用断言 3.7，我们可以高效地计算 $f^{(i)} \in \{0,1\}^N$，使得 $|\mathrm{supp}(f^{(i)})| = k$且

$$\langle w^{(i,+)} - w^{(i,-)}, f^{(i)}\rangle \geq \langle w^{(i,+)} - w^{(i,-)}, p\rangle. \tag{9}$$

定义 $r^{(i)}\colon [2N] \to \mathbb{R}$ 使得 $r^{(i)} = (f^{(i)} - p) \diamond (p - f^{(i)})$。我们使用步长为 α 的乘法权重算法计算 $w^{(i+1)}$：

$$\hat{w}^{(i+1)}(x) \leftarrow w^{(i)}(x)e^{-\alpha r^{(i)}(x)},$$
$$w^{(i+1)}(x) \leftarrow \hat{w}^{(i+1)}(x) / \sum_{x'\in[N]} \hat{w}^{(i+1)}(x')$$

根据引理 3.10，对于每个 $w \in \Delta_{2N}$，

$$\alpha\langle w^{(i)} - w, r^{(i)}\rangle \leq \mathsf{D}_{\mathsf{KL}}(w\|w^{(i)}) - \mathsf{D}_{\mathsf{KL}}(w\|w^{(i+1)}) + \alpha^2\|r^{(i)}\|_\infty^2.$$

对 $i = 1$、.....、T求和，并注意到 $\mathsf{D}_{\mathsf{KL}}(w\|w^{(1)}) \le \ln(2N)$、$\mathsf{D}_{\mathsf{KL}}(w\|w^{(T+1)}) \ge 0$ 以及$\|r^{(i)}\|_\infty \le 1$，我们得到

$$\sum_{i\in[T]} \alpha\langle w^{(i)} - w, r^{(i)}\rangle \le \ln(2N) + \alpha^2 T,$$

这意味着

$$\frac{1}{T}\sum_{i\in[T]} \langle w^{(i)}, r^{(i)}\rangle \le \frac{\ln(2N)}{\alpha T} + \alpha + \frac{1}{T}\left\langle w, \sum_{i\in[T]} r^{(i)}\right\rangle. \tag{10}$$

根据 $w^{(i,+)}, w^{(i,-)}$ 和 $r^{(i)}$ 的定义，

$$\langle w^{(i)}, r^{(i)}\rangle = \langle w^{(i,+)}, f^{(i)} - p\rangle + \langle w^{(i,-)}, f - q^{(i)}\rangle = \langle w^{(i,+)} - w^{(i,-)}, f^{(i)} - p\rangle \ge 0,$$

其中最后一个不等式由 (9) 得出。将其代入 (10)，对每个 $w \in \Delta_{2N}$，

$$0 \le \frac{\ln(2N)}{\alpha T} + \alpha + \frac{1}{T}\left\langle w, \sum_{i\in[T]} r^{(i)}\right\rangle. \tag{11}$$

根据我们的定义 $r^{(i)} = (f^{(i)} - p) \diamond (p - f^{(i)})$ 和命题 3.9，存在 $w \in \Delta_{2N}$ 使得 $\langle w, \sum_{i\in[T]} r^{(i)}\rangle = -\|\sum_{i\in[T]}(f^{(i)} - p)\|_\infty$，因此 (11) 意味着

$$\left\|\frac{1}{T}\sum_{i\in[T]} f^{(i)} - p\right\|_\infty \le \frac{\ln(2N)}{\alpha T} + \alpha.$$

令 $T = \lceil \frac{4\ln(2N)}{\varepsilon^2} \rceil$ 和 $\alpha = \sqrt{(\ln(2N))/T}$，得到 $\|\frac{1}{T}\sum_{i=1}^{T} f^{(i)} - p\|_\infty \le 2\sqrt{(\ln(2N))/T} \le \varepsilon$。注意到乘法权重算法可以在 $\tilde{O}(NT) = \tilde{O}(\varepsilon^{-2}N)$ 时间内计算$f^{(1)}, \ldots, f^{(T)}$，证明即告完成。

□

在 $N \gg 1/\varepsilon$ 的情况下，引理 3.8 中的运行时间还可以进一步改进：

引理 3.11。设 $N > 1$ 为正整数，并设 $p\colon [N] \to [0,1]$ 为满足$\sum_{x\in[N]} p(x) = k$ 的预测器，其中某个整数为 k。对于任意 $\varepsilon \in (0, 1/2)$，存在一个分布 μ，定义在二元函数 $f\colon [N] \to \{0,1\}$ 上，使得

1. $\Pr_{M\sim\mu}[\sum_{x\in[N]} f(x) = k] = 1$，以及 2. 对于每个 $x \in [N]$，$|\mathbb{E}_{f\sim\mu}[f(x)] - p(x)| \le \varepsilon$。

此外，给定作为输入的 p，我们可以在 $\tilde{O}(\varepsilon^{-2}r + N)$ 时间内从分布 μ 中采样，其中$r = \min\{1/\varepsilon, N\}$。

证明。根据引理 3.8，不失一般性，我们可以假设 $N \ge 1/\varepsilon$。我们将 $[0,1]$ 划分为$u = \lceil 2/\varepsilon \rceil$ 个子区间 $\Gamma_1, \ldots, \Gamma_u$，每个区间的宽度至多为 $\varepsilon/2$。这些子区间将 $X := [N]$ 划分为 $X_1, \ldots, X_u$，其中 $X_i := \{x \in X\colon p(x) \in \Gamma_i\}$。定义$s(i) := \sum_{x\in X_i} p(x)$，并对每个 $i \in [u]$ 定义 $t(i) := s(i) - \lfloor s(i)\rfloor$。我们有 $t(i) \in [0,1]$，且$\sum_i t(i)$ 是一个整数 $k' = k - \sum_i \lfloor s(i)\rfloor$。定义 $I_0 := \{i \in [u]\colon t(i) = 0\}$。根据引理 3.8，我们可以找到一个关于 $\hat{t}\colon [u] \setminus I_0 \to \{0,1\}$ 的分布，使得对每个 $i \in [u] \setminus I_0$，$|\mathbb{E}[\hat{t}(i)] - t(i)| \le \varepsilon/2$，且 $\sum_{i\in[u]\setminus I_0} \hat{t}(i) = k'$。此外，给定 t，可以在 $\tilde{O}(\varepsilon^{-2}r)$ 时间内采样一个随机的 $\hat{t}$。我们定义

对每个 $i \in I_0$, $\hat{t}(i) = 0$。对于每个 $i \in [u]$, 我们定义 $\hat{s}(i) := \lfloor s(i) \rfloor + \hat{t}(i)$。显然, $\hat{s}(i) \le \lceil s(i) \rceil \le |X_i|$, 并且

$$\sum\nolimits_{i\in[u]} \hat{s}(i) = k, \quad \text{and} \tag{12}$$

$$|\mathbb{E}[\hat{s}(i)] - s(i)| \le \varepsilon/2. \tag{13}$$

给定 $\hat{t}$, 我们计算 $\hat{s}$, 并对每个 $i \in [u]$, 随机选取 $\hat{s}(i)$ 个个体 x, 使其位于 X_i 中, 并赋予 $f(x) = 1$; 否则, 我们赋予 $f(x) = 0$。这样, 给定 $\hat{t}$, 我们就能在 $\tilde{O}(N)$ 时间内对 f 进行采样。根据 (12), 我们有 $\sum_x f(x) = k$。对于每个个体 $x \in X_i$,

$$\mathbb{E}_f[f(x)] = \mathbb{E}[\hat{s}(i)]/|X_i|.$$

根据 (13), 我们有

$$|\mathbb{E}_f[f(x)] - s(i)/|X_i|| \le \varepsilon/2.$$

由于 Γ_i 的宽度至多为 $\varepsilon/2$, 根据 $s(i)$ 和 X_i 的定义, 我们有 $|s(i)/|X_i| - p(x)| \le \varepsilon/2$。结合上面的不等式, 我们得到 $|\mathbb{E}_f[f(x)] - p(x)| \le \varepsilon$, 证毕。 □

3.2.3 步骤 3：具有 $\mathrm{polylog}(N)$ 时间实现的不可区分模型 n

现在我们证明引理 3.5, 其中 $N = 2^n$ 呈指数级大, 而我们希望实现的运行时间为 $\mathrm{poly}(n, \varepsilon^{-1})$。由于这一效率要求, 我们不能直接应用引理 3.11, 因为它只能给出运行时间为 $\mathrm{poly}(N, \varepsilon^{-1})$ 的实现。我们的思路是将定义域 $X = \{0,1\}^n$ 划分为若干小组, 每组大小约为 $\Theta(\varepsilon^{-1})$, 并在每个组上分别应用引理 3.11。这要求我们为每个组设定一个目标支持集大小, 使这些大小之和等于整个定义域的总支持集大小 k。为此, 我们构造一棵二叉树, 其根节点对应整个定义域, 叶节点对应各个组。我们为树中从根节点到叶节点的每个节点分配一个目标支持集大小。对于每个非叶节点, 我们通过估算与其两个子节点所对应的子定义域中的 $p(x)$ 之和, 来确定如何将该节点的支持集大小分配给这两个子节点。我们确保树的深度至多为 n, 从而能够高效地从根节点到达任意叶节点, 进而构造出高效的实现。

根据引理 3.11, 不失一般性, 我们可以在引理 3.5 中假设 $2^n \ge 8/\varepsilon$。首先, 我们描述一种低效的随机化算法: 给定预测器 p, 它会生成函数 $f\colon X \to \{0,1\}$。随后, 我们将模型 M 选为随机函数 f 所诱导的样本访问模型 B 的分布, 并证明 M 可以高效实现。

生成函数 f 的随机化算法如下。对于每个比特串 $z \in \bigcup_{j=0}^{n} \{0,1\}^j$, 我们可以将 z 视为一棵二叉树中的一个节点, 这棵树的根节点为空字符串。从根节点开始, 我们按如下方式为树中的某些节点 z 赋予整数 k_z。对于根节点 z, 我们设 $k_z := k$。假设我们已为树中的一个节点 z 赋予整数 k_z, 则按如下方式为 z 的两个子节点 $z\|0$ 和 $z\|1$ 赋予整数 $k_{z\|0}$ 和 $k_{z\|1}$。定义 $z' := z\|0$, 并令 $X_{z'}$ 为由满足以下条件的 $x \in X$ 组成的集合: z' 是 x 的前缀。对于 $m = \lceil 16n^2/\varepsilon^2 \rceil$, 我们从 $X_{z'}$ 中独立且均匀地抽取 x_1、.....、x_m, 并定义 $\ell_{z'} := \frac{|X_{z'}|}{m} \sum_{i=1}^{m} p(x_i)$。我们设 $\hat{k}_{z\|0} := \lfloor \ell_{z'} \rfloor \in [0, |X_{z'}|]$ 以及 $k_{z\|1} := k_z - k_{z\|0}$。然后我们定义

$$(k_{z\|0}, k_{z\|1}) = \begin{cases} (\hat{k}_{z\|0}, \hat{k}_{z\|1}), & \text{if } 0 \le \hat{k}_{z\|1} \le |X_{z\|1}|, \\ (k_z, 0), & \text{if } \hat{k}_{z\|1} < 0, \\ (k_z - |X_{z\|1}|, |X_{z\|1}|), & \text{if } \hat{k}_{z\|1} > |X_{z\|1}|. \end{cases}$$

上述过程使我们能够为每个节点 z 赋予一个整数 k_z，并利用这些整数构造函数 $f\colon X \to \{0,1\}$。具体而言，选择满足$2^{n-n'} \geq 8/\varepsilon$的最大整数 n'，对于每个 $z \in \{0,1\}^{n'}$，我们构造函数 $f_z\colon X_z \to \{0,1\}$，然后将所有 $z \in \{0,1\}^{n'}$ 对应的 f_z 组合起来，构造函数 f。

为构造每个 f_z，我们首先构造如下预测器 $p'\colon X_z \to [0,1]$。如果 $k_z \geq \sum_{x\in X_z} p(x)$，我们将每个 $p(x)$ 增加到 $x \in X_z$ 至 $p'(x)$，使得 $\sum_{x\in X_z} p'(x) = k_z$。类似地，如果 $k_z < \sum_{x\in X_z} p(x)$，我们将 $p(x)$ 减少到 $p'(x)$。利用引理 3.11，我们可以在时间 $\mathrm{poly}(|X_z|, \varepsilon^{-1}) = \mathrm{poly}(\varepsilon^{-1})$ 内随机构造一个函数 $f_z\colon X_z \to \{0,1\}$，使得对于每个 $x \in X_z$，

$$|\mathbb{E}\, f_z(x) - p'(x)| \leq \varepsilon/2. \tag{14}$$

其中，期望值取自 f_z 的随机性。

我们在以下引理中分析上述生成函数 f 的过程：

引理 3.12。在上述过程中，如果我们选择 $m \geq Cn^2/\varepsilon^2$，其中 $C > 0$ 为充分大的绝对常数，则

1. $\Pr[\sum_{x\in X} f(x) = k] = 1$;
2. $\sum_{x\in X} |\mathbb{E}[f(x)] - p(x)| \leq |k - \sum_{x\in X} p(x)| + \varepsilon|X|$.

注意，概率和期望均取决于 f 中的随机性。

在证明引理 3.12 之前，我们先证明以下两个辅助引理。

引理 3.13。假设 $|k_z - \sum_{x\in X_z} p(x)| \leq \alpha$ 且 $|\ell_{z\|0} - \sum_{x\in X_{z\|0}} p(x)| \leq \beta$。则

$$\left|k_{z\|0} - \sum_{x\in X_{z\|0}} p(x)\right| + \left|k_{z\|1} - \sum_{x\in X_{z\|1}} p(x)\right| \leq \alpha + 2\beta + 2.$$

^^^证明。根据我们对 $k_{z\|0} = \lfloor \ell_{z\|0} \rfloor$ 和 $k_{z\|1} = k_z - k_{z\|0}$ 的选择，我们有

$$\left|\hat{k}_{z\circ 0} - \sum_{x\in X_{z\|0}} p(x)\right| \leq \beta + 1,$$

and

$$\left|\hat{k}_{z\circ 1} - \sum_{x\in X_{z\|1}} p(x)\right| \leq \left|k_z - \sum_{x\in X_z} p(x)\right| + \left|\hat{k}_{z\circ 0} - \sum_{x\in X_{z\|0}} p(x)\right| \leq \alpha + \beta + 1.$$

综上，

$$\left|\hat{k}_{z\circ 0} - \sum_{x\in X_{z\|0}} p(x)\right| + \left|\hat{k}_{z\circ 1} - \sum_{x\in X_{z\|1}} p(x)\right| \leq \alpha + 2\beta + 2.$$

^^该引理的证明基于以下观察：将（$k_{z\|0}$，$k_{z\|1}$）替换为（$k_{z\|0}$，$k_{z\|1}$）不会增加上述不等式左侧。 □

引理 3.14。 令 $n' \in \mathbb{Z}_{\geq 0}$ 满足 $2^{n-n'} \geq \frac{8}{\varepsilon}$。则

$$\sum_{z\in\{0,1\}^{n'}} \mathbb{E}\left|k_z - \sum_{x\in X_z} p(x)\right| \leq \left|k - \sum_{x\in X} p(x)\right| + \varepsilon|X|/2.$$

证明。 考虑任意一个 $z \in \bigcup_{j=0}^{n-1}\{0,1\}^j$，并定义 $z' = z\|0$。根据我们对 $\ell_{z'}$的定义：$=\frac{|X_{z'}|}{m}\sum_{i=1}^{m} p$ (x_i)，对于从 $X_{z'}$上的均匀分布中独立同分布抽取的 $x_1, \ldots, x_m$，我们有 $\mathbb{E}[\ell_{z'}] = \sum_{x\in X_{z'}} p(x)$，且 $\mathsf{Var}(\ell_{z'}) \leq \frac{|X_{z'}|^2}{m}$。因此，

$$\mathbb{E}\left|\ell_{z'} - \sum_{x\in X_{z'}} p(x)\right| \leq \mathbb{E}\left[\left(\ell_{z'} - \sum_{x\in X_{z'}} p(x)\right)\right]^{1/2} = \mathsf{Var}[\ell_{z'}]^{1/2} \leq \frac{|X_{z'}|}{\sqrt{m}}.$$

根据我们对 $m \geq 16n^2/\varepsilon^2$的选择，我们有

$$\mathbb{E}\left|\ell_{z'} - \sum_{x\in X_{z'}} p(x)\right| \leq \frac{\varepsilon|X_{z'}|}{4n} = \frac{\varepsilon|X_z|}{8n}. \tag{15}$$

我们通过对 $j = 0, \ldots, n$使用归纳法证明以下不等式：

$$\sum_{z\in\{0,1\}^j} \mathbb{E}\left|k_z - \sum_{x\in X_z} p(x)\right| \leq \left|k - \sum_{x\in X} p(x)\right| + \left(\frac{\varepsilon j}{4n} + 2^{-(n-j-1)}\right)|X|. \tag{16}$$

当 z 为空字符串时，由于 $k_z = k$ 且 $X_z = X$，因此该不等式对 $j = 0$ 显然成立。现在假设该不等式对某个 $j = 0, \ldots, n-1$成立，并证明将 j替换为 $j+1$后不等式仍成立。对于每个 $z \in \{0,1\}^j$，定义 $z' := z\|0$，根据引理 3.13，我们有

$$\left|k_{z\|0} - \sum_{x\in X_{z\|0}} p(x)\right| + \left|k_{z\|1} - \sum_{x\in X_{z\|1}} p(x)\right| \leq \left|k_z - \sum_{x\in X_z} p(x)\right| + 2\left|\ell_{z'} - \sum_{x\in X_{z'}} p(x)\right| + 2$$

对 $z \in \{0,1\}^j$ 求和并取期望，我们有

$$\begin{aligned}
&\sum_{z\in\{0,1\}^{j+1}} \mathbb{E}\left|k_z - \sum_{x\in X_z} p(x)\right| \\
&\leq \sum_{z\in\{0,1\}^j}\left(\mathbb{E}\left|k_z - \sum_{x\in X_z} p(x)\right|\right) + 2\sum_{z\in\{0,1\}^j} \mathbb{E}\left|\ell_{z'} - \sum_{x\in X_{z'}} p(x)\right| + 2^{j+1} \\
&\leq \left|k - \sum_{x\in X} p(x)\right| + \left(\frac{\varepsilon j}{4n} + 2^{-(n-j-1)}\right)|X| + \frac{\varepsilon}{4n}|X| + 2^{-(n-j-1)}|X| \\
&\qquad\qquad \text{(by induction hypoth假设及式（15）)} \\
&\leq \left|k - \sum_{x\in X} p(x)\right| + \left(\frac{\varepsilon(j+1)}{4n} + 2^{-(n-j-2)}\right)|X|.
\end{aligned}$$

通过在 (16) 中令 $j = n'$，并注意到 $\frac{\varepsilon n'}{4n} \leq \frac{\varepsilon}{4}$ 和 $2^{-(n-n'-1)} \leq \frac{\varepsilon}{4}$，即可证明该引理。 □

引理 *3.12* 的证明。 对于每个 $z \in \{0,1\}^{n'}$，根据 $p'(x)$ 的定义，

$$\sum_{x \in X_z} \left| p'(x) - p(x) \right| = \left| k_z - \sum_{x \in X_z} p'(x) \right|. \tag{17}$$

结合（17）和（14），

$$\begin{aligned}\sum_{x \in X_z} |\,\mathbb{E}[f_z(x)|k_z] - p(x)| &\le \sum_{x \in X_z} |\,\mathbb{E}[f_z(x)|k_z] - p'(x)| + \sum_{x \in X_z} |p'(x) - p(x)| \\ &\le \varepsilon |X_z|/2 + \left| k_z - \sum_{x \in X_z} p'(x) \right|.\end{aligned}$$

对 k_z 取期望，

$$\begin{aligned}\sum_{x \in X_z} |\,\mathbb{E}[f(x)] - p(x)| &= \sum_{x \in X_z} |\,\mathbb{E}[f_z(x)] - p(x)| \\ &= \sum_{x \in X_z} |\,\mathbb{E}_{k_z}\,\mathbb{E}[f_z(x)|k_z] - p(x)| \\ &\le \sum_{x \in X_z} \mathbb{E}_{k_z} |\,\mathbb{E}[f_z(x)|k_z] - p(x)| \\ &\le \varepsilon |X_z|/2 + \mathbb{E} \left| k_z - \sum_{x \in X_z} p(x) \right|.\end{aligned}$$

对 $z \in \{0,1\}^{n'}$ 求和并应用引理 3.14，

$$\sum_{x \in X} |\,\mathbb{E}[f(x)] - p(x)| \le \varepsilon |X|/2 + \left| k - \sum_{x \in X} p(x) \right| + \varepsilon |X|/2 = \left| k - \sum_{x \in X} p(x) \right| + \varepsilon |X|. \qquad \square$$

引理 *3.5* 的证明。 只需证明，将引理 3.12 中的随机函数 f 诱导的样本访问对象 B 的分布定义为 M 的模型可以高效实现，并且该实现本身可以根据 p 高效计算。

为说明 M 可以高效实现，我们首先通过查询随机预言机，生成创建 f 的过程中所需的全部随机性。然后注意到，为了计算 $f(x)$，只需计算 x 的每个前缀 z 所对应的 k_z，再计算 z 对应的 f_z，其中 z 是 x 的长度为 n' 的前缀。这可以在 $\mathrm{poly}(n, \varepsilon^{-1})$ 时间内高效完成。

$\square$

3.3 学习支持访问二元函数

在前几节中，我们展示了如何学习并构造由函数诱导的样本访问对象 B^* 的不可区分模型的实现，也就是说，存在一个函数 $f^*\colon X \to \{0,1\}$，并且区分器和学习器都会接收形式为 $(x,\ f^*(x))$ 的随机样本。在本节中，我们学习由函数诱导的支持集访问对象，见定义 2.2。对于二值函数 $f^*\colon X \to \{0,1\}$，由 f^* 诱导的支持集访问对象 B^* 从集合 $\{x | f^*(x) = 1\}$ 中输出随机样本。我们将展示如何构造一个与 B^* 不可区分的模型的高效实现。

区分器： 设 $\mathcal{D}$ 为区分器的集合，其中每个 $D \in \mathcal{D}$ 都关联有一个子集$S_D \subset X$。给定从对象 B 获取的样本 x，当 $x \in S_D$ 时，区分器接受。

我们假设，对于每个 $D \in \mathcal{D}$，集合 S_D 的大小已知且有高效描述；给定输入 x 时，它会回答 $x \in S_D$是否成立。

审计器： 令 $\mathcal{D}$如上所定义。若算法 $\Lambda^{B^*,p}$满足以下性质，则称其为集合族 $\mathcal{S} = \{S_D | D \in \mathcal{D}\}$ 的 $(\varepsilon, \gamma, \delta)$ 审计器。给定对函数诱导的支持访问对象 B^* 的访问权限，以及对预测器 p 的查询访问权限： $X \to [0,1]$。如果存在 $S \in \mathcal{S}$ 和 $b \in \{-1, 1\}$，使得

$$b\left(\Pr_{x \sim B^*(\perp)}[x \in S] - \Pr_{x \sim p}[x \in S]\right) > \varepsilon.$$

那么审计器会返回一个集合 S'，使得以概率 $1-\delta$,

$$b\left(\Pr_{x \sim B^*(\perp)}[x \in S'] - \Pr_{x \sim p}[x \in S']\right) > \gamma. \tag{18}$$

其中， $x \sim p$ 是由预测器 p生成的分布，即 $\mathbf{Pr}_{x \sim p}[x = x'] = p(x') \sum_{x'' \in X} p(x'')$。

假设存在一个用于集合族 $\mathcal{S}$ 的审计器，我们将展示一种针对支持访问函数对象 B^* 的学习算法。

定理 3.15。 设 $\alpha \in [0,1]$，并设 $\mathcal{B}$ 是由二值函数诱导的一组支持访问对象，其中 $\forall B \in \mathcal{B}$, $\mathbb{E}_{x \in X}[f_B(x)] = \alpha$。设 $\mathcal{D}$ 是如上所述的区分器集合。

设 $\varepsilon, \gamma, \delta', \delta''$ 为参数，且对于足够小的常数 c，有 $\delta' \leq c\delta\gamma^2\alpha^{-2}$。设 Λ 是针对 $\mathcal{D}$ 的一个 $(\varepsilon, \gamma, \delta')$ 审计器 Λ。那么，存在一个相对于区分器类 $\mathcal{D}$、针对 $\mathcal{B}$ 的 $(2\varepsilon, \delta)$ 学习器 L。学习器 L 的运行时间为 $\mathit{poly}(\gamma^{-1}\log(\delta^{-1})\alpha^{-1}, W_1, W_2)$，其中 W_1 是审计器 Λ 的运行时间，而 W_2 是其输出的电路复杂度。学习器输出的实现 T 的运行时间为 $\mathit{poly}(\gamma^{-1}\log(\delta^{-1})\alpha^{-1}, W_2)$

上述定理中的学习器 L 可以访问一个审计器 Λ，该审计器能够审计支持集访问对象。在第 3.3.1 节中，我们将证明，在与样本访问对象的审计器类似的条件下，这样的审计器确实存在。

该学习算法与引理 3.3 中的经典算法类似，但多了一个步骤：使模型的期望值近似为 α。要证明支持集访问对象的收敛性，就需要这一步。我们注意到，如果函数是稀疏的，也就是说 α 非常小，那么该算法将不再高效。

学习算法： 我们描述了多项式时间学习器 L。该学习器维护一个列表F或元组 (S, w)，用于 $S \subseteq X$ 和 $w \in [0,1]$。利用该列表，学习器通过为每个 $x \in X$ 设置 $X \to [0,1]$，计算出一个预测器p：

$$p(x) = \mathsf{Lcap}\left(w \cdot \mathbf{1}(x \in S) : (S, w) \in F\right), \tag{19}$$

其中，L截断函数在添加每个列表元素后，将值 $p(x)$ 截断到范围 $[0,1]$ ，如公式（2）所定义。例如，如果列表为 $(S_1,\ w_1)$，$(S_2,\ w_2)$，那么我们有 $p(x)$ $=$截断函数(截断函数($w_1 \cdot \mathbf{1}(x\ \in S_1)) + w_2 \cdot \mathbf{1}(x \in S_2)$)。

对于每个 $x \in X$，计算 $p(x)$ 时，对函数 $\mathbf{1}(x \in S)$ 需要进行 $|F|$ 次调用；这些集合 S 由审计器返回。

学习器使用以下算法创建列表 F 。

1. 初始化：设定 $F = (X,\ w_0 = \alpha)$。

2. 查询 $\Lambda^{B^*,\ p}$，其中 p 是由 F 生成的预测器。

   (a) 如果没有返回集合，则输出当前的 F。(b) 如果审计器返回集合 $S \subseteq X$ 和符号 $b \in \{-1, 1\}$，则设定 $F \leftarrow F \cup (S,\ b\gamma\alpha\frac{|X|}{|S|})$。 —

3. 从 X, $I = \{x_1, \ldots, x_t\}$ 中选择 $t = \log\ |X|$ 个随机元素。
   如果 $|\mathbb{E}_{x \in I}[p(x)] - \alpha| > \beta = \frac{1}{10}\gamma\alpha^2$，则将 $([N], \beta\text{sign}\ (\alpha - \mathbb{E}_{i \in I}[p_i]))$ 添加到 F。重复当前项目，直到满足条件。返回项目 2。

实现： 给定列表 F，随机预言机实现 T 通过拒绝采样实现对 $x \sim B(\perp)$ 的采样。将随机预言机记为 R ，并假设对于每个输入，随机预言机都会输出不同的随机字符串。

1. 选择一个随机 $x \in X$。

2. 使用 $R(x)$ 采样一个比特 b ，使得 $\Pr_{R(x)}[b = 1] = p(x)$。

3. 如果 $b = 1$，则返回 x。如果 $b = 0$，则返回项目 1。

实现中的循环迭代次数的期望值为 α，因为这是 p 的期望值。随机预言机用于创建并实现一个一致函数对象 B。也就是说，每次实现对 $x \in X$ 进行采样时，随机预言机都会返回相同的随机字符串$R(x)$。这意味着，对于任意随机性，该实现都会针对特定函数 f 实现一个支持集访问对象：$X \to \{0, 1\}$，正如所要求的那样。

定理 *3.15* 的证明。 我们先假设学习算法已运行完毕，并证明在这种情况下，输出对所有 $D \in \mathcal{D}$ 而言都与 B^* 无法区分。

算　　法在条目 3 后结束，这意味着以概率 $(\beta\,|X|)^{-1}$，我们有　　e

$$\left|\sum_{x \in X} p(x) - \alpha\,|X|\right| < 2\beta. \tag{20}$$

假设式 (20) 成立，我们的模型的实现 T 是高效的，并且在停止前大约运行$1/\alpha$ 步。它从 p 中采样一个随机项 $x \sim p$，使得审计器不会返回任何集合 S，这意味着对于所有 $D \in \mathcal{D}$,

$$\left|\Pr_{x \sim B^*(\perp)}[x \in S_D] - \frac{\sum_{x \in S_D} p(x)}{\sum_{x \in X} p(x)}\right| \leq \varepsilon.$$

综上可知，算法结束后，模型 M 满足

$$\left|\Pr_{x\in B^*(\perp)}[D(x)=1]-\Pr_{B\sim M,x\in B(\perp)}[D(x)=1]\right|\le\varepsilon+2\beta\le 2\varepsilon.$$

因此，如果学习算法结束，它将输出一个模型的实现，对于所有 $D\in\mathcal{D}$，该实现都与 B^* 无法区分。

接下来需要证明学习算法是高效的，这同时也限制了实现 T 的复杂度（因为它取决于 F 的大小）。我们使用一个势函数来证明这一点：

$$\varphi(B^*,p)=\sum_{x\in X}\left(f^*(x)-p(x)\right)^2.$$

由于 p 和 f^* 都属于 $[0,1],$，因此 φ 的值有 $|X|$的界。

根据定义，审计器以 $1-\delta'$ 的概率返回一个满足式 (18) 的集合 S。在这种情况下，根据命题 3.16，φ 的值在每次迭代中减少 $\Omega(\gamma^2\alpha^2|X|)$。此外，根据命题 3.17，在条目 3 的每次迭代中，φ 的值减少 $\Omega(\gamma^2\alpha^4|X|)$。综合起来，我们得到两类权重更新的总次数至多为 $O(\gamma^{-2}\alpha^{-4})$。 □

断言 3.16。 假设在算法的项目 *2b* 中，学习算法列表从F更新为 $F\cup(S,w)$，使得(S,w)满足等式(18)。设 p为 F所描述的预测器，p'为 $F\cup(S,w)$所描述的预测器。假设 $|\mathbb{E}_{x\in X}[p(x)]-\alpha|\le 2\delta$，则 $\varphi(B^*,p')\le\varphi(B^*,p)-\frac{1}{2}\gamma^2\alpha^2|X|$。

证明。 根据我们的假设，有

$$\left|\Pr_{x\sim B^*(\perp)}[x\in S]-\frac{\sum_{x\in S}p(x)}{\sum_{x\in X}p(x)}\right|>\gamma.$$

前一个不等式也可以写成

$$\left|\frac{\sum_{x\in S}f^*(x)}{\sum_{x\in X}f^*(x)}-\frac{\sum_{x\in S}p(x)}{\sum_{x\in X}p(x)}\right|>\gamma. \tag{21}$$

我们分为两种情况： $b=1$ 和 $b=-1$。这两种情况几乎相同，为完整起见，我们将分别写出。

若 $b=1$： 则等式 (21) 可以写为：

$$\begin{aligned}\gamma&<\frac{\sum_{x\in S}f^*(x)}{\sum_{x\in X}f^*(x)}-\frac{\sum_{x\in S}p(x)}{\sum_{x\in X}p(x)}\\&\le\frac{1}{\alpha|X|}\sum_{x\in S}f^*(x)-\frac{1}{|X|(\alpha+2\delta)}\sum_{x\in S}p(x)\\&\le\frac{1}{\alpha|X|}\sum_{x\in S}\left(f^*(x)-\left(1-\frac{2\delta}{\alpha+2\delta}\right)p(x)\right)\end{aligned} \tag{22}$$

这里我们利用了这样一个事实： $\mathbb{E}_{x\in X}[p(x)]\le\alpha+\beta$。

$$\varphi(B^*,p)-\varphi(B^*,p')=\sum_{x\in S}\left(2f^*(x)w-2wp(x)-w^2\right)$$

$$
\begin{aligned}
&= 2w \sum_{x \in S} (f^*(x) - p(x)) - |S| \, w^2. \\
&\geq 2w\gamma\alpha N - 2w \frac{2\delta}{\alpha + 2\delta} |S| - w^2 |S|, \\
&= 2\frac{\gamma^2\alpha^2 |X|^2}{|S|} - 4 |X| \gamma\delta - \varepsilon^2\alpha^2 \frac{|X|^2}{|S|} \\
&\geq \frac{1}{2} \frac{\gamma^2\alpha^2 |X|^2}{|S|} \geq \frac{1}{2}\gamma^2\alpha^2 |X|.
\end{aligned}
$$

这里我们利用了这样一个事实：$\beta = \frac{1}{8}\gamma\alpha^2$。

将 $p(x)$ 截断到 $[0,1]$ 只能降低势函数的值，因为对于每个 $x \in X$，我们都有 $f^*(x) \in \{0,1\}$。

若 $b = -1$： 这种情况与 $b = 1$ 情况几乎相同。在这种情况下，公式 (21) 可以写成

$$
\begin{aligned}
\gamma &< \frac{\sum_{x \in S} p(x)}{\sum_{x \in X} p(x)} - \frac{\sum_{x \in S} f^*(x)}{\sum_{x \in X} f^*(x)} \\
&\leq \frac{1}{|X| (\alpha - 2\beta)} \sum_{x \in S} p(x) - \frac{1}{\alpha |X|} \sum_{x \in S} f^*(x) \\
&\leq \frac{1}{\alpha |X|} \left(\sum_{x \in S} (p(x) - f^*(x)) + \frac{2\beta}{\alpha - 2\beta} |S| \right).
\end{aligned}
$$

如果算法没有执行任何截断：

$$
\begin{aligned}
\varphi(B^*, p) - \varphi(B^*, p') &= \sum_{x \in S} (2f^*(x)w - 2wp(x) - w^2) \\
&\geq \frac{\gamma^2\alpha^2 |X|^2}{|S|} - 8\gamma\beta |X| \geq \frac{1}{5}\gamma^2\alpha^2 |X|.
\end{aligned}
$$

与之前一样，将 $p(x)$ 截断到 $[0,1]$ 只能降低 φ 的值。 □

命题 3.17. 假设在条目 *3* 所述的步骤中，算法对列表 F 进行了更新。设 p 为更新前的预测器（如 (19) 中所定义），p' 为更新后的预测器。那么，以 $1 - O(|X|^{-1})$ 的概率，$\varphi(B^*, p') \leq \varphi(B^*, p) - \frac{9}{10}\beta^2 |X|$。

证明。这个证明是经典学习证明。利用切尔诺夫界，以概率 $O(|X|^{-1})$，我们有

$$
\left| \sum_{x \in I} \frac{1}{t} p(x) - \sum_{x \in X} \frac{1}{N} p(x) \right| \leq \frac{1}{100}\beta.
$$

假设上述条件成立，则

$$
\left| \sum_{x \in X} f^*(x) - \sum_{x \in X} p(x) \right| > \frac{99}{100}\beta |X|.
$$

不妨设 $\sum_{x\in X} f^*(x) \geq \sum_{x\in X} p(x)$，则我们将 $([N], \beta)$ 添加到列表中。我们有

$$\begin{aligned}\varphi(B^*, p) - \varphi(B^*, p') &= \sum_{x\in X} \left((f^*(x) - p(x))^2 - (f^*(x) - p'(x))^2\right) \\ &= \sum_{x\in X} \left(-2f^*(x)(p(x) - p'(x)) + p(x)^2 - p'(x)^2\right).\end{aligned}$$

暂时假设不对 p' 进行截断函数处理。在这种情况下，$p'(x) = p(x) + \beta$ 因此

$$\begin{aligned}\varphi(B^*, p) - \varphi(B^*, p') &\geq \sum_{x\in X} \left(2f^*(x)\beta - 2\beta p(x) - \beta^2\right) \\ &\geq 2 \cdot \frac{99}{100}\beta^2 |X| - \beta^2 |X| = \frac{98}{100}\beta^2 |X|.\end{aligned}$$

对 p' 进行截断函数处理只会减小 $\varphi(B^*, p')$ 的值，因为 $f^*(x) \in [0, 1]$。 □

3.3.1 支撑集访问对象的审计器

样本访问对象的审计器可以看作概念类 $\mathcal{G}$ 的弱无关学习器，该概念类在均匀分布下表示区分器类 $\mathcal{D}$。对于支撑集访问对象，审计器不会获得带标签的样本（x，$f^*(x)$），而是获得一个随机 x，使得 $f^*(x) = 1$。

我们证明，对于一大类学习算法和统计查询算法而言，从支撑集访问对象中学习并不比从样本访问对象中学习困难多少。

定义 3.2（统计查询预言机）。关于分布 P：$X \times \{0, 1\} \to [0, 1]$ 的统计查询预言机是一个函数：它接收一个函数 ϕ：$X \times \{0, 1\} \to \{0, 1\}$ 和误差参数 ε 作为输入，并输出一个值 $v \in [0, 1]$，满足 $\left|v - \mathbb{E}_{(x,y)\sim P}[\phi(x, y)]\right| \leq \varepsilon$。

定义 3.3（统计查询算法）。算法 $\mathcal{A}$ 是针对分布 P：$X \times \{0, 1\} \to [0, 1]$ 的统计查询算法，当且仅当它只通过针对分布 P的统计查询预言机访问对象。

命题 3.18。若存在一个统计查询算法 $\mathcal{A}$，能够在由样本访问对象生成的分布 P上学习假设 h，即对均匀随机的 $x \in X$输出$(x, f^*(x))$，那么存在一个算法 $\mathcal{A}'$，能够从由支持集访问对象生成的分布 P' 中学习 h，即对于随机的 $x \in X$，输出均匀的$(x, 1)$，使得 $f^*(x) = 1$。算法 $\mathcal{A}'$还需要将 $m = |\mathsf{supp}(f^*)| = \sum_{x\in X} f^*(x)$作为输入。

此外，$\mathcal{A}'$ 对统计预言机的查询次数相同，运行时间与 $\mathcal{A}$ 的运行时间相近，至多相差多对数 $|X|$ 因子，且成功概率相同，至多相差加性误差 $1/$对数 $|X|$。

证明。该算法 $\mathcal{A}'$ 模拟算法 $\mathcal{A}$，并使用针对 P'的预言机，模拟关于分布 P 的统计查询预言机。假设算法 $\mathcal{A}$ 向统计查询预言机查询函数 ϕ：$X \times \{0, 1\} \to \{0, 1\}$。由于第二个变量是二元的，我们可以写成，

$$\phi(x, y) = \varphi_1(x) \cdot y + \varphi_2(x)(1 - y), \tag{23}$$

$$\mathbb{E}_{x\in X}[\phi(x, f^*(x))] = \mathbb{E}_{x\in X}[(\varphi_1(x) - \varphi_2(x))f^*(x)] + \mathbb{E}_{x\in X}[\varphi_2(x)]. \tag{24}$$

值 $\mathbb{E}_{x\in X}[\varphi_2(x)]$ 不依赖函数 f^*，因此无需向 f^* 或预言机进行任何查询即可近似得到。鉴于函数 $\varphi_2(x)$ 有界， $\mathcal{A}'$ 可以在多对数级 $|X|$ 运行时间内，将$E_{x\in X}[\varphi_2(x)]$ 近似到加性误差不超过 $\varepsilon/2$，成功概率为 $1-|X|^{-2}$。我们将此近似记为 v_1。

对于第二个元素， $\mathbb{E}_{x\in X}[(\varphi_1(x) - \varphi_2(x))f^*(x)]$,我们有

$$\begin{aligned}\mathbb{E}_{x\in X}[(\varphi_1(x)-\varphi_2(x))f^*(x)] &= \Pr_{x\in X}[f^*(x)=1]\cdot\mathbb{E}_{x\in X}[(\varphi_1(x)-\varphi_2(x))|f^*(x)=1] \\ &= \frac{m}{|X|}\,\mathbb{E}_{x\in X}[(\varphi_1(x)-\varphi_2(x))|f^*(x)=1].\end{aligned}$$

注意，最后一项是在支持访问对象分布 P' 下的 $\varphi_1(x) - \varphi_2(x)$ 的期望值。因此，算法 $\mathcal{A}'$ 随后针对分布 P' 以及函数 $\phi'(x,y) = \varphi_1(x) - \varphi_2(x)$ 和误差参数 $\varepsilon/2$，向统计查询预言机发起查询。将返回值记为 v_2。然后， $\mathcal{A}'$ 模拟 $\mathcal{A}$，并假设统计查询预言机的返回值为 $\mathcal{A}$ 是$v_1 + \frac{m}{|X|}v_2$。由于 v_1, v_2 中的误差被 $\varepsilon/2$所界定，因此该返回值以 $1-|X|^{-2}$ 的概率成为统计预言机查询的有效返回值。

对 $\mathcal{A}$ 向统计查询预言机发起的所有查询求和， $\mathcal{A}'$失败的概率与 $\mathcal{A}$ 相同，另有一个小于 $1/\log|X|$ 的因子（这是对 $\mathbb{E}_{x\in X}[\varphi_2(x)]$进行近似时产生的误差）。

□

因此，如果标准样本访问对象中存在一个审计器，也就是一个统计查询算法，那么支持集访问对象中也存在一个审计器。

我们进一步证明，对于特定的区分器集合，存在支持集访问对象的审计器。

命题 3.19。 设 $\mathcal{D}$ 是支持集访问对象的一组 t 区分器，定义见第 *3.3* 节。那么，存在一个审计器 Λ，定义见第 *3.3* 节，其运行时间为 $\Theta(t)$，查询复杂度为 $O(\log(t)\mathrm{polylog}(|X|)1/(\delta(\varepsilon-\gamma)))$。

证明。 该审计器抽取 $k = c\cdot\log(t)\mathrm{polylog}(|X|)$个均匀样本 $x_1,\ldots,x_k \sim B^*(\perp)$。对于每个区分器 $D\in\mathcal{D}$ 及其关联集合 S_D，审计器计算

$$v_D = \Pr_{x\in\{x_1,\ldots,x_k\}}[x\in S_D] = \frac{|\{x_1,\ldots,x_k\}\cap S_D|}{k}.$$

如果存在 $D\in\mathcal{D}$ 使得 $|v_D - \mathbf{Pr}_{x\sim p}[x\in S_D]| > \frac{\varepsilon+\gamma}{2}$，审计器就返回 S_D。对于每个集合 $S_D\subseteq X$，根据切尔诺夫界，关于 $x_1,\ldots,x_k$的概率满足

$$\left|\Pr_{x\in\{x_1,\ldots,x_k\}}[x\in S_D] - \Pr_{x\in B^*(\perp)}[x\in S_D]\right| \geq \frac{\varepsilon-\gamma}{2}$$

至多为 $e^{-k(\varepsilon-\gamma)^2/10}$。因此，我们选择常数 c，使得即使对 t 个区分器求和，误差也小于 δ。

□

我们可以将上述算法计算出的值 v_D 视为 $\mathbf{Pr}_{x\in B^*(\perp)}[x\in S_D]$ 的估计器。上述论证表明， v_D 一致收敛于 $\mathbf{Pr}_{x\in B^*(\perp)}[x\in S_D]$。

定义 3.4 (一致收敛**)**。 如果存在函数 $k:(0,\ 1)\times(0,\ 1)\to\mathbb{N}$，使得对于每个 ε、$\delta>0$，若 $k>k(\varepsilon,\ \delta)$，则估计器 v 在假设类 $\mathcal{D}$ 上关于分布 B 一致收敛于 P。

$$\Pr_{x_1,\ldots,x_k\sim B}[\exists D\in\mathcal{D}\ s.t.\ |v_D(x_1,\ldots,x_k)-P_{D,x\sim B}|\geq\varepsilon]<\delta.$$

一致收敛意味着存在一个算法，可以对类 $\mathcal{D}$进行弱不可知学习，其样本复杂度为 poly($1/\varepsilon$, log($1/\delta$),VCdim($\mathcal{D}$)) [VC71、BEHW89、LMR91]。

3.4 学习位串函数

本节关注学习函数 f: $\{0,1\}^n\to\{0,1\}^n$。这比学习二值函数更难，因为该函数的值域非常大。在本节中，我们只学习一个相对于非常有限的一组区分器的不可区分模型，并且具有乘积结构。在这种设定下，对于随机输入 x，采样分布是一个对 $(x, f(x))$。

区分器： 设 $\mathcal{D}$，使得每个区分器 $D\in\mathcal{D}$ 都有一个集合 $S_D\subset\{0,1\}^n$ 和一个坐标 $j\in[n]$。若满足 $x\in S$且 $f(x)_j=1$，则区分器 D接受样本 $(x, f(x))$。

审计器： 令 $\mathcal{D}$ 如上所定义。若算法 $\Lambda^{B^*,p}$ 满足以下性质，则称其为集合族 $\mathcal{S}$ 的（$\varepsilon,\gamma,\delta$）审计器。给定对由函数诱导的支持访问对象 B^*的访问权限，以及对一组包含 n 个预测器的集合 p_1, $\ldots p_n$的查询访问权限，其中 p_j: $\{0,1\}^n\to[0,1]$。若存在 $S\in\mathcal{S}$和 $j\in[n]$ 使得

$$b\left(\Pr_x[x\in S, f(x)_j=1]-\mathbb{E}_x[p_j(x)\cdot\mathbf{1}(x\in S)]\right)>\varepsilon.$$

则审计器返回集合 $S'\subseteq\{0,1\}^n$ 和 $j\in[n]$，使得其以概率 $1-\delta$满足:

$$b\left(\Pr_x[x\in S', f(x)_j=1]-\mathbb{E}_x[p_j(x)\cdot\mathbf{1}(x\in S')]\right)>\gamma.$$

.

定理 3.20。令 $\mathcal{B}$ 为由函数 f: $\{0,1\}^n\to\{0,1\}^n$诱导的支持访问对象集合。令 $\mathcal{D}$ 为如上所述的区分器集合。

令 ε、γ、δ'、δ''为满足以下条件的参数: 当常数 c足够小时， $\delta'\leq c\delta\gamma^2n^{-1}$。令 Λ 为 $\mathcal{D}$的（ε, γ, δ'）审计器。则存在一个针对区分器类 $\mathcal{D}$ 的（2ε, δ）学习器 L，且该学习器 L 的运行时间为 $poly(\gamma^{-1}\ log(\delta^{-1})\alpha^{-1},\ W_1,\ W_2)$，其中 W_1 为审计器 Λ 的运行时间， W_2 为其输出的电路复杂度。学习器输出的实现 T 的运行时间为 $poly(\gamma^{-1}\ log(\delta^{-1})\alpha^{-1},\ W_2)$

我们介绍学习器 L 和实现 T。

实现： 给定元组列表 F （S, j, w），可访问随机预言机 R 的实现算法 T 会生成模型 $\hat{}B$。给定 $x\in\{0,1\}^n$ 算法 T，其输出v 如下:

1. 对每个 $j\in[n]$，按以下方式计算 $p_j(x)\in[0,1]$：

$$p_j(x)=\mathsf{Lcap}\left(w\cdot\mathbf{1}(x\in S):(S,j,w)\in F\right).$$

2. 通过为每个 $j \in [j]$ 设置 $v \in \{0,1\}^n$ 来进行选择：对每个 $v_j = 1$，以概率 $p_j(x)$ 独立地使用随机预言机 $R(x)$。

学习算法： 学习算法 L。

1. 初始化：设置 $F = \{(\{0,1\}^n,\ j,\ 1/2) | j \in [n]\}$。 2. 使用当前模型和 B^* 查询审计器。如果审计器返回 $(S,\ j,\ b)$，则将 $(S,\ j,\ b \cdot \gamma)$ 添加到 F，并重复上述步骤。

注意，学习器与经典提升算法相同，我们将该算法分别应用于 $f(x)$ 的每个坐标上。这就是为什么我们只能相对于一组区分器 D 实现不可区分性；这些区分器只检查 $f(x)$ 中的单个输出坐标是否等于 1。

定理 *3.20* 的证明。与之前的证明类似，如果审计器没有输出任何集合，那么对于所有 $D \in \mathcal{D}$，模型都是 ε-不可区分的。

我们使用势函数证明正确性，该势函数对每个坐标的误差求和$j \in [n]$。令

$$\varphi(B^*, p_1, \ldots, p_n) = \sum_{j \in [n]} \sum_{x \in \{0,1\}^n} \left(\mathbf{1}(f(x)_j = 1) - p_j(x)\right)^2.$$

每个 x 在总和中出现 n 次，因此 φ 的初始值以 $n \cdot 2^n$ 为界。

设 $p_1, \ldots, p_n$ 为执行项目 2 之前的预测器，$p'_1, \ldots, p'_n$ 为执行之后的预测器。我们声称，以概率 $1 - \delta'$，有 $\varphi(B^*, p_1, \ldots, p_n) - \varphi(B^*, p'_1, \ldots, p'_n) \geq \gamma^2 2^n$。

审计器返回了 (S, i, b)，然后

$$\left| \frac{1}{2^n} \sum_{x \in S} (\mathbf{1}(f(x)_j = 1) - p_j(x)) \right| \geq \gamma. \tag{25}$$

不失一般性，假设 $b = 1$。我们注意到，$p_1, \ldots, p_n$ 与$p'_1, \ldots, p'_n$ 之间的差异仅在于 p_i, p'_i，且仅针对 $x \in S$。

$$\begin{aligned}
\varphi(B^*, p_1, \ldots, p_n) - \varphi(B^*, p'_1, \ldots, p'_n) &= \sum_{x \in S} \left(\left(\mathbf{1}(f(x)_i = 1) - p_i(x)\right)^2 - \left(\mathbf{1}(f(x)_i = 1) - p'_i(x)\right)^2 \right) \\
&= \sum_{x \in S} \left(-2 \cdot \mathbf{1}(f(x)_i = 1)(p_i(x) - p'_i(x)) + p_i^2(x) - p'_i(x)^2 \right).
\end{aligned}$$

假设该算法不进行任何截断，

$$\begin{aligned}
\varphi(B^*, p_1, \ldots, p_n) - \varphi(B^*, p'_1, \ldots, p'_n) &= \sum_{x \in S} \left(2\gamma \cdot \mathbf{1}(f(x)_i = 1) - 2\gamma p_i(x) - \gamma^2 \right) \\
&\geq 2\gamma^2 2^n - \gamma^2 2^n \\
&= \gamma^2 2^n,
\end{aligned}$$

其中，最后一个不等式中使用了式（25）。

$\mathbf{1}(f(x)_j = 1)$ 的值位于 $\{0,1\}$ 中，因此将 p_i 的值截断至 $[0,1]$ 只能降低势函数的值。

□

4 学习指数规模的图

4.1 学习稠密图

图最基本的设置是稠密模型。在该模型中，由图 $G = ([N], E)$ 诱导的图诱导的样本访问对象 B，可以看作是从 G 的邻接矩阵中为 $u, v \in [N], b \in \{0, 1\}$ 获取一个随机项 $((u, v), b)$（见定义 2.4）。在这种设置下，我们可以将图的邻接矩阵视为一个函数，而图会对区分器施加一些额外结构。因此，第 3 节中的一些结果可以直接得出。

区分器： 设 $\mathcal{D}$ 为一个区分器集合。每个区分器都有两个顶点集合 U_D，$V_D \subset [N]$。从图诱导的样本访问对象 B^*中获取一个随机项 $((u, v), b)$ 时，当 $u \in U_D$、$v \in V_D$ 和 $b = 1$时，它接受该项。

审计器： 如果算法 Λ 满足以下条件，则称其为包含成对元组 $(U, V) \subset [N] \times [N]$ 的元组集合 $\mathcal{S}$ 的 $(\varepsilon, \gamma, \delta)$-审计器。该审计器可查询预测器$p$：$[N] \times [N] \to [0, 1]$，以及由图 $G^* = ([N], E^*)$ 诱导的图诱导的样本访问对象 B^*。如果存在 $(U, V) \in \mathcal{S}$ 使得

$$b \cdot \left(\Pr_{(u,v)\in[N]\times[N]}[u \in U, v \in V, (u, v) \in E^*] - \mathbb{E}_{(u,v)\in[N]\times[N]}[\mathbf{1}(u \in U, v \in V)p(u, v)] \right) \geq \varepsilon.$$

那么以概率 $(1 - \delta)$，它输出 U'，V' 使得

$$b \cdot \left(\Pr_{(u,v)\in[N]\times[N]}[u \in U', v \in V', (u, v) \in E^*] - \mathbb{E}_{(u,v)\in[N]\times[N]}[\mathbf{1}(u \in U', v \in V')p(u, v)] \right) \geq \gamma.$$

根据定理 3.2（该定理使用了来自 [TTV09、HJKRR18、KGZ19],的多重校准预测器），我们可得：假设拥有一个针对区分器类的审计器，就能学习指数图模型。具体做法是，将图的邻接矩阵视为二值函数，然后把定理 3.2 直接应用于图的样本访问对象，即应用于 f^*：$[N] \times [N] \to \{0, 1\}$使得 $f^*(u, v) = 1 \iff (u, v) \in E^*$。

推论 *4.1*（定理 *3.2* 的推论）。令区分器类 $\mathcal{D}$ 为上文定义的类。令ε、γ、δ、$\delta' > 0$ 为满足 $\delta' \leq c\delta\gamma^2$ 的参数，其中绝对常数 $c > 0$足够小。令$\mathcal{B}$ 为由顶点集 $[N]$上的图所诱导的样本访问对象类。令 Λ 为针对 $\mathcal{S} = \{(U_D, V_D) | D \in \mathcal{D}\}$的 $(\varepsilon, \gamma, \delta')$-审计器。那么，存在一个相对于 $\mathcal{D}$、针对 $\mathcal{B}$ 的 (ε, δ)-学习器 L。此外，如果审计器 Λ 的运行时间至多为 W_1，且其输出的函数的电路规模始终至多为W_2，那么学习器 L 的运行时间为 $poly(\gamma^{-1}, log(\delta^{-1}) W_1)$，并且其输出的实现的电路复杂度始终为 $O(\gamma^{-2}W_2)$。

对于学习具有固定数量边的图，同样成立。当 m 已知时，我们可以将定理 3.4应用于表示具有 m 条边的图的邻接矩阵的样本访问对象。定理 3.4表明，在推论 4.1的设置下，我们可以学习一个不可区分模型，同时还要求该模型仅包含具有 m 条边的图。不过请注意，只有将图的邻接矩阵作为函数使用时才能做到这一点，因此该结果仅对稠密图有意义，即当 $m = \Theta(N^2)$时。

这两个定理同样适用于学习有向图的模型，之前的所有结果仍然成立。具体而言，只需使用有向图的邻接矩阵即可，即 $f(u, v)$可能不同于 $f(v, u)$。此外，对于有向图，还可以应用第 3.2 节中使用的构造，使真实性要求规定出度固定为 d。也就是说，当给定一个由具有均匀出度 d 的有向图 G 所诱导的样本访问对象时，我们输出一个模型，使得对于所有 $B \sim M$，图 G_B 对所有顶点的出度均为 d。具体做法是像之前一样学习一个预测器，并将第 3.2 节中的构造用于采样算法。在采样算法中，不再以概率$p(u, v)$独立地包含有向边 (u, v)，而是将第 3.2 节中的算法应用于邻接矩阵中 u 所在行所表示的函数，其中该函数的值为 d。我们通过执行第 3.2 节中的算法来采样一条有向边。

4.2 学习无稠密子图的稀疏图

对于稀疏图，样本访问图对象并无用处。这是因为稀疏图的邻接矩阵中，随机项几乎总是 0，因此对于样本访问对象而言，任何稀疏图都无法与空图区分开来。对于稀疏图，我们研究图诱导的支撑集访问对象（定义 2.5）。由稀疏图 $G = ([N], E)$,$B(\perp)$ B^* 诱导的支撑集访问对象会输出图 $(u, v) \in E$ 中的一条随机边。

在第 3.3 节中，我们研究了函数诱导的支撑访问对象，即函数f的一个对象，它会返回一个随机输入 x，使得 $f(x) = 1$。对于图 $G = ([N], E)$，若将其邻接矩阵视为函数 $f\colon [N] \times [N] \to \{0, 1\}$，那么 f的支撑访问函数对象 B 会返回 f中的一个随机正值项，也就是一条随机边 $(u, v) \in E$。因此，我们可以将定理 3.15 应用于支撑访问图对象。遗憾的是，定理 3.15 仅适用于期望值为常数的函数 f，也就是说，f 表示稠密图的邻接矩阵时才适用（当 f 表示稀疏图时，学习和实现的效率都很低）。

在本节中，我们将展示如何为稀疏图创建一个稠密模型，前提是该稀疏图不存在过于稠密的子图。为此，我们使用稀疏图的强正则性引理 [KR03, Sco11], 。该引理意味着，对于 G 不存在稠密子图的稀疏图，存在一个与之不可区分的稠密图。然后，我们将问题简化为为稠密图寻找稠密模型，并应用定理 3.15。

区分器： 设 $\mathcal{D}$ 为一个区分器集合。每个区分器都有两个顶点集合 $U_D, V_D \subset [N]$。对于输入 (u, v)，区分器 D 在满足 $u \in U_D$ 且 $v \in V_D$时接受。

审计器： 若算法 Λ 满足以下条件，则称其为一组集合对 $\mathcal{S}$的 $(\varepsilon, \varepsilon', \delta)$-审计器，其中 $(U, V) \in \mathcal{S}$, $U, V \subset [N]$。审计器通过查询访问获得预测器 $p\colon [N] \times [N] \to [0, 1]$，并访问表示图$G^* = ([N], E^*)$ 的图诱导的支持访问对象 B^*。若存在一对集合 $(U, V) \subset \mathcal{S}$和一个比特 b，使得

$$b \cdot \left(\Pr_{(u,v)\sim B^*(\perp)}[u \in U, v \in V] - \Pr_{(u,v)\sim p}[u \in U, v \in V] \right) \geq \varepsilon,$$

然后，Λ输出集合 $U', V' \subset [N]$使得

$$b \cdot \left(\Pr_{(u,v)\sim B^*(\perp)}[u \in U', v \in V'] - \Pr_{(u,v)\sim p}[u \in U', v \in V'] \right) \geq \varepsilon'.$$

分布（u， v） $\sim p$ 由预测器 p定义，即对于所有 u， $v \in [N]$ 我们有

$$\Pr_{(u',v')\sim p}[u'=u, v'=v] = \frac{p(u,v)}{\sum_{u'',v''\in[N]} p(u'',v'')}.$$

图的记号与定义 对于图 $G=([N],E)$ 和 $U,V \subset [N]$，我们定义$E_G(U,V) = \{(u,v) \in E | u \in U, v \in V\}$ 为 G 中 U,V之间的边集合。我们用 $\rho_G(U,V)$ 表示 G 中 U,V之间的边密度， $\rho_G(U,V) = \frac{|E_G(U,V)|}{|U||V|}$。我们用 $\rho_G = \rho_G([N],[N])$ 表示该图的边密度。

我们采用 [KR03, Sco11] 对上均匀图的定义，来定义不含稠密子图的图。我们需要一个比 [KR03, Sco11], 中使用的定义稍强的定义，此外还要求对小于 ηN 的 U,V 作出要求。

定义 4.1 (上均匀图)。 图 $G=([N],E)$ 是(η,γ)-上均匀的，如果对于任意两个不相交集合 $U,V \subset [N]$, 且最小值 $\{|U|,|V|\} \geq \eta N$，都有

$$\rho_G(U,V) \leq \gamma \rho_G,$$

且对于满足 $\min\{|U|,|V|\} < \eta N$ 的 U,V，

$$|E(U,V)| \leq \gamma\eta\rho_G N^2.$$

我们指出，以高概率，随机稀疏图对于常数η， γ是上均匀的。

In 本节中，我们使用 [KR03, Sco11中的稀疏图正则性引理].

定义 4.2 (正则集)。 对于图 $G=([N],E)$，若对所有满足 $|U'| \geq \delta|U|, |V'| \geq \delta|V|$ 的 $U' \subset U, V' \subset V$, 都有.....，则顶点集 $V,U \subset [N]$ 是(δ)-正则的

$$\left|\rho_G(U',V') - \rho_G(U,V)\right| \leq \delta\rho_G.$$

定理 4.2 (稀疏上均匀图正则性引理 [KR03])。 对于每个 $\delta,\gamma > 0$，存在$(\eta,m) = \varphi(\delta,\gamma)$ $\eta > 0$, $m \in \mathbb{N}$，使得每个(η,γ)-上均匀图 G 都有一个划分 $U_0,U_1,\ldots,U_k$，其中 $k \leq m$，且对于所有 i, $j \geq 1$，都有 $|V_0| \leq \delta N$, $|V_i| = |V_j|$；划分中的 V_i,V_j 这些对中，除δk^2 外，均为(δ)-正则。

我们不现在陈述并证明一个定理，用于为稀疏图构造稠密模型 h.

定理 4.3。 对于每个参数 $\gamma,\varepsilon,\varepsilon',\lambda'\lambda'' > 0$，只要 $\lambda' \leq c\lambda\varepsilon'^2$ 对某个充分大的常数 c 成立，就存在 $\eta \in [0,1]$ ，使得以下结论成立。设 $\mathcal{B}$ 是一个图诱导的支持集访问对象集合族，且对每个 $B^* \in \mathcal{B}$，它所表示的图 G_{B^*} 都是(η,γ)-上均匀的。

设 $\mathcal{D}$ 是一个区分器集合。如果存在一个$(\varepsilon,\varepsilon',\delta')$-审计器 Λ ，用于集合族 $\mathcal{C} = \{(U_D,V_D) | D \in \mathcal{D}\}$，那么就存在一个$(\varepsilon,\delta'')$-学习算法 L ，适用于所有关于 $\mathcal{D}$ 的 $B^* \in \mathcal{B}$ 。

证明。 从整体上看，该定理的证明分为两部分。第一部分，即命题 4.4，我们使用稀疏图正则性引理证明，对于每个上均匀图 G ，都存在一个

稠密图 H，使其与 G无法区分。第二部分中，我们将寻找模型的问题转化为：在与我们的图 G^* 无法区分的稠密图 H 中寻找模型。

令 $\delta = \frac{\varepsilon'}{100\gamma}$。令 $(\eta', m) = \varphi(\delta, \gamma)$，其中 φ 是正则性引理（定理 4.3）中的函数。令 $\eta = \min\{\eta', 1/(2m)\ \delta\}$。

令 $B^* \in \mathcal{B}$ 为一个图诱导的支持访问对象，表示一个 (η, γ)-上均匀的稀疏图 G^*。我们指出，由于上正则性对 η 参数具有单调性， G^* 也是 (η', γ)-上均匀的。

Fro根据命题 4.4，存在一个图 H，满足 $\rho_H \approx 1/\gamma$，使得对所有 $U, V \subset [N]$ ，

$$\left| \frac{|E_G(U,V)|}{\gamma} - |E_H(U,V)|\, \rho_G \right| \leq 12\delta\rho_G N^2.$$

假设我们可以访问一个对象 B^H，它表示稠密图 H，而非表示稀疏图 G^*的图诱导支持访问对象 B^*。那么，我们就可以在 B^H 上运行第 3.3 节中的学习算法（将 B^H 视为函数诱导的支持集访问对象），并得到一个模型 M 的实现，该模型与 B^H是 2ε-不可区分的。由于不可区分性误差是可加的， M 也与 B^*是 $(2\varepsilon + \delta)$-不可区分的。

在我们的情形中，我们无法访问对象 B^H，也无法找到 H。相反，我们通过能够访问 B^* 的审计器，模拟一个能够访问 B^H 的审计器。每当算法向审计器 $\Lambda^{B^H,p}$发出查询时，我们便改为调用审计器 $\Lambda^{B^*,p}$，并将其回答转发给算法。

我们声称，模拟审计器是针对 B^H 的 $(2\varepsilon, \frac{1}{2}\varepsilon', \delta')$-审计器。事实上，一个稍弱的结论就足够了。我们声称，如果 $\Lambda^{B^*,p}$返回一个元组 $((U,V)\ b)$，那么它也是 $\Lambda^{B^H,p}$ 的有效返回值。

假设审计器 $\Lambda^{B^*,p}$ 返回了一个元组 $((U,V)\ b)$。那么，以概率 $1-\delta'$，我们有

$$\left| \Pr_{(u,v)\sim B^*}[u \in U, v \in V] - \Pr_{(u,v)\sim p}[u \in U, v \in V] \right| > \varepsilon'.$$

这与下式相同

$$\left| \frac{E_{G^*}(U,V)}{\rho_{G^*} \cdot N^2} - \Pr_{(u,v)\sim p}[u \in U, v \in V] \right| > \varepsilon'.$$

根据命题 4.4，我们有

$$\left| \frac{|E_{G^*}(U,V)|}{\rho_{G^*} \cdot N^2} - \frac{\gamma\,|E_H(U,V)|}{N^2} \right| \leq 12\delta\gamma.$$

结合 $\rho_H \in [1/\gamma - 12\delta, 1/\gamma + 12\delta]$ 这一事实，我们得到

$$\left| \Pr_{(u,v)\sim B^H}[u \in U, v \in V] - \Pr_{(u,v)\sim p}[u \in U, v \in V] \right| \geq \epsilon' - 12\delta\gamma \geq \frac{1}{2}\varepsilon'.$$

对于我们对 δ 的选择。也就是说，如果审计器 $\Lambda^{B^*,\ p}$ 返回了一个元组 $((U,\ V)\ b)$，那么以概率 $1-\delta'$，该答案对于 $\Lambda^{B^H,\ p}$ 也是一个有效的返回值。

如果审计器 $\Lambda^{B^*,\ p}$ 不返回答案，那么 p 是一个与 B^*不可区分的 2ε 模型（更多细节请参见定理 3.15 的证明），因此结论成立。 □

命题 4.4。设 γ、δ、$\eta > 0$满足 $\eta \leq \eta'$、$1/(2m)$ δ对于（η'、m）$= \varphi$（δ、γ）。那么对于顶点集为 $[N]$ 的任意（η、γ）-上均匀图 G，存在一个顶点集为 $[N]$的图 H，满足$\rho_H \in [1/\gamma - 12\delta, 1/\gamma + 12\delta]$，使得对于任意 U、$V \subset [N]$，均有

$$\left| \frac{|E_G(U,V)|}{\gamma} - |E_H(U,V)|\,\rho_G \right| \leq 12\delta\rho_G N^2.$$

证明。根据该断言的要求，η 是一个足够小的常数，因此稀疏图正则性引理（定理 4.2）在 γ, δ 的条件下成立。因此，任意一个（η, γ）-上均匀图 G都可以划分为 $V_0, V_1 \ldots, V_k$，其中 $k \leq m$满足：对于所有 $i, j \geq 1$，$|V_0| \leq \delta N$，除 δk^2 个 V_i, V_j对之外，所有这些对都是（δ）-正则的。由于 $\eta \leq \frac{1}{2k}$且 G是（η, γ）正则的，所有集合 V_i, V_j都满足 $\rho_G\left(V_i, V_j\right) \leq \gamma\rho_G$。

我们利用这个划分按如下方式构造一个稠密图 H。对于每一对 $u, v \in [N]$，以$p_{u,v}$ 所定义的概率，将边（u, v）加入 H：

$$p_{u,v} = \begin{cases} \frac{\rho_G(V_i,V_j)}{\gamma\rho_G} & \text{if } u \in V_i, v \in V_j \text{ for a regular pair } V_i, V_j \\ \frac{1}{\gamma} & \text{otherwise.} \end{cases}$$

由 $\rho_G\left(V_i, V_j\right) \leq \gamma\rho_G$这一事实可知，对于每个 u, v，都有 $p_{u,v} \leq 1$。

我们证明 H满足该命题的条件。任取两个集合 $U, V \subset [N]$。我们先对 $|E_G(U,V)|$进行界定，

$$|E_G(U,V)| = \sum_{i,j} |E_G(U \cap V_i, V \cap V_j)|\,.$$

我们利用正则性引理对该和进行界定，并分情况讨论。

1. 异常集：根据正则性引理，异常集的大小至多为 δN。利用上均匀性条件，与异常集的交集满足：

$$|E_G(V_0 \cap V, U)| \leq \delta\gamma\rho_G N^2,$$

对于 $E_G(V_0 \cap U, V)$，同样成立。

2. 非正则集合：划分中至多有 δk^2 个集合 V_i, V_j 不是正则的。我们对这些集合之间的边数上界进行估计：

$$\sum_{i,j \text{ non-regular}} |E_G(V_i, V_j)| \leq \sum_{i,j \text{ non-regular}} \gamma\rho_G \cdot |V_i|\,|V_j| \leq \gamma\delta\rho_G N^2.$$

3. 交集较小的集合：对于集合 U，若 $|U \cap V_i| \leq \delta\,|V_i|$，则称 U 与 V_i 的交集较小。令 $I_U \subset [k]$ 为与 U交集较小的部分，并令 $S_U = \cup_{i \in I_U}(U \cap V_i)$为它们的并集。于是 $|S_U| \leq \delta N$。根据上均匀性条件，我们有

$$|E_G(S_U, V)| \leq \gamma\delta\rho_G N^2,$$

对于 $E_G(S_V,\ U)$，结论同样成立。

4. 具有大交集的正则集：设 V_i，V_j 是一个 (δ)-正则对，满足 $|V_i \cap U| \geq \delta\,|V_i|$ 和 $|V_j \cap V| \geq \delta\,|V_j|$。那么根据正则性引理，

$$|\rho_G(V_i \cap U, V_j \cap V) - \rho_G(V_i, V_j)| \leq \delta\rho_G.$$

综合所有情况，如果我们称 i，$j \in [k]$ 为良好的，当且仅当 V_i，V_j 是正则的，并且 U，V 与 V_i，V_j 具有大交集，则有

$$\begin{aligned}|E_G(U,V)| &\le 5\gamma\delta\rho_G N^2 + \sum_{\text{good } i,j} |E_G(U\cap V_i, V\cap V_j)| \\ &\le 5\gamma\delta\rho_G N^2 + \sum_{\text{good } i,j} |U\cap V_i|\,|V\cap V_j|\,\rho_G(V_i,V_j)\,(1+\delta\rho_G)\end{aligned}$$

同理，我们得到一个下界

$$|E_G(U,V)| \ge \sum_{\text{good } i,j} |U\cap V_i|\,|V\cap V_j|\,\rho_G(V_i,V_j)(1-\delta\rho_G).$$

现在我们界定 $|E_H(U,\ V)|$。图 H 的生成方式是：以概率 $p_{u,\ v}$ 均匀随机地采样每条边。对于所有良好的 i，j，我们有：若 $u \in V_i$，$v \in V_j$，则 $p_{u,\ v} = p_{i,\ j} = \rho_G\ (V_i,\ V_j)\ \rho_G\gamma$。由于 $|V_i|$，$|V_j| = \Theta(N)$，根据切尔诺夫界，以高概率，

$$|\rho_H(V_i\cap U, V_j\cap V) - p_{i,j}| \le \delta p_{i,j}.$$

至多有 $5\delta N^2$ 对顶点 $(u,\ v)$ 不属于一对良好的部分V_i，V_j。

综上，我们得到

$$\begin{aligned}\left||E_H(U,V)|\,\rho_G - \frac{|E_G(U,V)|}{\gamma}\right| &\le \sum_{\text{good } i,j}\left||E_H(U\cap V_i, V\cap V_j)|\,\rho_G - \frac{|E_G(U\cap V_i, V\cap V_j)|}{\gamma}\right| \\ &\quad + 5\delta\rho_G N^2 + 5\delta\rho_G N^2 \\ &\le \sum_{\text{good } i,j} |U\cap V_i|\,|V\cap V_j|\,(\delta\rho_G(V_i,V_j) + \delta\rho_G) + 10\delta\rho_G N^2 \\ &\le 12\delta\rho_G N^2.\end{aligned}$$

上述等式还意味着 $\rho_H \in [1/\gamma - 12\delta, 1/\gamma + 12\delta]$，将其应用于 $U = V = [N]$ 即可得出。

□

接下来，我们证明，含有稠密子图的稀疏图没有稠密的不可区分模型。这并不意味着无法学习含有稠密子图的稀疏图，而是表明我们的方法存在一定局限。这个结论并不紧确，因为有些图不满足定理 4.3 的结论，而我们也不知道它们是否有稠密模型。

断言 4.5。设 G 是一个图，若存在一个区分器 D，其顶点集为 U、V，满足$|U|$、$|V| \ge \varepsilon N$，并且

$$\frac{\rho_G(U,V)}{\rho_G} > \gamma + \frac{2\delta}{\varepsilon^2},$$

那么就不存在图 H，使其与 G 到 D之间的 δ无法区分，且满足 $\rho_H = 1/\gamma$。

证明。 如果 H 和 G 无法区分，那么必定有

$$\left| \Pr_{(u,v)\sim G}[u \in U, v \in V] - \Pr_{(u,v)\sim H}[u \in U, v \in V] \right| \leq \delta,$$

这与

$$\left| \frac{\rho_G(U,V)}{\rho_G} - \frac{\rho_H(U,V)}{\rho_H} \right| \leq \delta \frac{N^2}{|U|\,|V|}.$$

这意味着

$$\frac{\rho_H(U,V)}{\rho_H} \geq \frac{\rho_G(U,V)}{\rho_G} - \frac{\delta}{\varepsilon^2} > \gamma,$$

这与 $\rho_H(U, V) \leq 1$ and $\rho_H = 1/\gamma$ 相矛盾。 □

4.3 学习稀疏均匀出度图

本节我们关注如何学习一个稀疏有向图 $G^* = (\{0,1\}^n, E^*)$，其中每个 $x \in G^*$ 的出度都是一个均匀常数 d。我们假设 B^* 是表示图 G^*的对象，并且查询 $B^*(\perp)$ 时，会按照以下分布输出一条随机边：先随机选择一个 $x \in [N]$，再随机选择 x的一个邻居。

如果我们将每个 x 的 d 出边表示为 d 个函数，f_1、……、f_d：$\{0,1\}^n \to \{0,1\}^n$，那么学习 B^* 与学习这些 d 函数相当类似。使用这一记法，从$B^*(\perp)$采样一条随机边，对应于随机选择一个 $x \in \{0,1\}^n$，$i \in [d]$，并得到一对 $(x, f_i(x))$。如果我们还假设样本包含索引 i，即 $B^*(\perp)$ 返回一个随机三元组 $(x, i, f_i(x))$，那么这就完全等同于学习函数 f_1、……、f_d，第 3.4 节的结果也就可以直接转用。若没有这一额外假设，我们就需要多做一些工作，才能将第 3.4 节的结果转用于图对象。

我们正式描述该对象、我们能够对抗的区分器类以及审计器。

对象 B^* 是一个针对有向图 $G = (\{0,1\}^n, E)$ 的图诱导的支持集访问对象，其出度为常数 d。边分布是先均匀选择一个 $x \in \{0,1\}^n$，再随机选择 x 的一个邻居 y。我们假设每个顶点的 d 邻居有一个顺序，并用 f_1，……，f_d：$\{0,1\}^n \to \{0,1\}^n$ 表示 d 个函数，它们代表图中各顶点的邻居。

区分器： 设 $\mathcal{D}$，使得每个区分器 $D \in \mathcal{D}$ 都有一个集合 $S_D \subset \{0,1\}^n$ 和一个坐标 $j \in [n]$。如果 $u \in S_D$ 且 $v_j = 1$，区分器 D 就接受边 (u, v)。

审计器： 令 $\mathcal{D}$ 如上所定义。若算法 $\Lambda^{B^*,p}$ 具有以下性质，则称其为集合族 $\mathcal{S}$ 的（$\varepsilon, \gamma, \delta$）审计器。给定对图诱导的支持访问对象 B^* 的访问权限以及对预测器 p：$\{0,1\}^n \to [0,1]$的查询访问权限。如果存在$S \in \mathcal{S}$ 和 $j \in [j]$，使得

$$b\left(\Pr_{x,i}[x \in S, f_i(x)_j = 1] - \mathbb{E}_x[p_j(x) \cdot \mathbf{1}(x \in S)] \right) > \varepsilon.$$

那么，审计器会返回一个集合 S', j，且以概率 $1 - \delta$满足

$$b\left(\Pr_{x,i}[x \in S', f_i(x)_j = 1] - \mathbb{E}_x[p_j(x) \cdot \mathbf{1}(x \in S')] \right) > \gamma.$$

在所有 $x \in [N]$，$i \in [d]$ 中，概率是均匀的。

定理 4.6。设 $\mathcal{B}$ 是由 d 个函数 f_1、.....、f_d： $\{0,1\}^n \to \{0,1\}^n$ 诱导的支持访问图对象集合。设 $\mathcal{D}$ 是如上所述的区分器集合。

设 $\varepsilon, \gamma, \delta', \delta''$ 为参数，满足 $\delta' \leq c\delta\gamma^2 n^{-1}$，其中 c 是一个充分小的常数。设 Λ 是针对 $\mathcal{D}$ 的一个 $(\varepsilon, \gamma, \delta')$ 审计器 Λ。那么，存在一个关于区分器类 $\mathcal{D}$ 的 $(2\varepsilon, \delta)$ 学习器 L，用于学习 $\mathcal{B}$；学习器 L 的运行时间为 $poly(\gamma^{-1}\, log(\delta^{-1})\alpha^{-1}, W_1, W_2)$，其中 W_1 是审计器 Λ 的运行时间，W_2 是其输出的电路复杂度。学习器输出的实现 T 的运行时间为 $poly(\gamma^{-1}\, log(\delta^{-1})\alpha^{-1}, W_2)$

我们以与定理 3.20 的证明类似的方式证明该定理。学习器 L 和实现 T 与函数中的相同；唯一不同的是，我们将同一个预测器 d 次，以获得图中每个顶点的 d 条出边，并验证这些边的端点互不相同。

实现： 给定一个元组列表 F （S, j, w），实现算法 T 通过访问随机预言机 R 生成一个模型 $\hat{}B$。给定 $x \in \{0,1\}^n$，算法 T 通过以下方式输出v：

1. 对每个 $j \in [n]$，按以下方式计算 $p_j \in [0,1]$：

$$p_j(x) = \mathsf{Lcap}\left(w \cdot \mathbf{1}(x \in S) : (S, j, w) \in F\right).$$

2. 对于 $t = 1$ 到 d，使用随机预言机 $R(v, t)$，以概率 $p_j(x)$ 通过设置 $v_j^{(t)} = 1$ 来选取一个顶点 $v^{(t)}$。3. 如果所有顶点 $v^{(t)}$ 互不相同，则均匀地选取一个 $i \in [d]$ 并返回 $v^{(i)}$。如果发生碰撞，则返回项目 2，并使用随机预言机产生的新随机性重复操作。

我们指出，由于 d 是常数，而定义域 $\{0,1\}^n$ 的规模呈指数级，因此发生碰撞的可能性极低。在分析模型在大型顶点集合上的分布时，我们可以忽略这一事件。

学习算法： 学习算法 L。

1. 初始化：设置 $F = \{(\{0,1\}^n,\ j,\ 1/2) | j \in [n]\}$。2. 使用当前模型和 B^* 向审计器发起查询。如果审计器返回 $(S,\ j,\ b)$，则将 $(S,\ j,\ b \cdot \gamma)$ 添加到 F 中并重复执行。

请注意，该学习器与经典提升算法完全相同，我们将其分别应用于 $\sum_{i \in [d]} f_i(x)$ 的每个坐标。这就是为什么我们只能相对于一组区分器 D 实现不可区分性；这些区分器只检查 $f(x)$ 中单个输出坐标是否等于 1。

定理 *4.6* 的证明。证明与定理 3.20 的证明相同，只是对 $i \in [d]$ 求平均。项目 2 上发生碰撞的概率可以忽略不计，因此我们可以放心地假设 d 出边之间相互独立。

如果听　众未输出任何集合，那么对于所有 $D \in \mathcal{D}$，模型都是 ε-不可区分的　　.
我们使用一个势函数证明其正确性，该势函数对每个坐标的误差求和$j \in [n]$。设

$$\varphi(B^*, p_1, \ldots, p_n) = \sum_{j \in [n]} \sum_{x \in \{0,1\}^n} \left(\sum_{i \in [d]} \frac{1}{d} \mathbf{1}(f_i(x)_j = 1) - p_j(x) \right)^2.$$

每个 $x$ap 在求和中出现 $n \cdot d$ 次，因此 φ 的初始值有界于 $n \cdot 2^n$.

设 $p_1, \ldots, p_n$ 为项目 2 执行前的预测器， $p'_1, \ldots, p'_n$ 为执行后的预测器，则我们声称，以 $1-\delta'$ 的概率，有 $\varphi(B^*, p_1, \ldots, p_n) - \varphi(B^*, p'_1, \ldots, p'_n) \geq \gamma^2 2^n$。

审计器返回了 (S, i, b)，然后

$$\left| \frac{1}{2^n} \sum_{x \in S} \mathbb{E}_{i \in [d]}[\mathbf{1}(f(x)_j = 1)] - p_j(x) \right| \geq \gamma. \tag{26}$$

不失一般性，假设 $b = 1$，则以 $1 - \delta$ 的概率，

$$\begin{aligned}
&\varphi(B^*, p_1, \ldots, p_n) - \varphi(B^*, p'_1, \ldots, p'_n) \\
&= \sum_{j \in [n]} \sum_{x \in S} \left(\left(\mathbb{E}_{i \in [d]}[\mathbf{1}(f(x)_j = 1)] - p_j(x) \right)^2 - \left(\mathbb{E}_{i \in [d]}[\mathbf{1}(f(x)_j = 1)] - p'_j(x) \right)^2 \right) \\
&= \sum_{x \in S, j \in [n]} \left(-2\, \mathbb{E}_{i \in [d]}[\mathbf{1}(f(x)_j = 1)](p_j(x) - p'_j(x)) + p_j^2(x) - p'_j(x)^2 \right).
\end{aligned}$$

如果没有应用截断函数，则有

$$\begin{aligned}
\varphi(B^*, p_1, \ldots, p_n) - \varphi(B^*, p'_1, \ldots, p'_n) &= \sum_{x \in S} \left(2\gamma \cdot \mathbb{E}_{i \in [d]}[\mathbf{1}(f(x)_j = 1)] - 2\gamma p_i(x) - \gamma^2 \right) \\
&\geq 2\gamma^2 2^n - \gamma^2 2^n \\
&= \gamma^2 2^n,
\end{aligned}$$

其中，在最后一个不等式中我们使用了公式 (25)。

$\mathbb{E}_{i \in [d]}[\mathbf{1}(f(x)_j = 1)]$ 的取值均为 $[0, 1],$，因此将 p_i 的值截断至 $[0, 1]$ 只会降低势函数的值。

□

4.4 学习均匀度图

假设我们想要为均匀度 d 图生成一个真实模型。也就是说，我们希望模型中的所有图都具有均匀度 d。在第 4.1 节中，我们讨论了如何将函数学习的结果应用于图学习，并在假设图是稠密的前提下，得到了一个针对具有均匀出度的有向图的真实模型。在上一节中，我们还介绍了如何构造具有均匀出度的稀疏有向图。

对于无向图，在 [GGN10] 中，作者提出了一种生成均匀度图的算法，其生成的图与随机均匀度图无法区分。该算法对一个大围长均匀度扩展图应用随机置换。

我们希望生成一个模型 M ，使其与特定的图诱导对象 B^* 不可区分。因此，不能采用相同的方法。相反，我们将区分器的集合限制为那些能够通过顶点集合上的划分来描述的区分器，并展示在这种情况下如何构造一个均匀度图。我们的构造很简单——我们对区分器所诱导的划分中每两部分之间的边密度进行采样和近似，而该模型则是一个具有正确边密度的确定性图上的随机置换。

区分器： 于是，区分器集合 $\mathcal{D}$ 包含带有集合 $(U_D,\ V_D)$ 的区分器 D，满足 $U,\ V \in \mathcal{U}$。每个区分器 D 会接受一条边 $(u,\ v)$，当且仅当 $u \in U,\ v \in V$。令 $\mathcal{U} = \{U \subset [N] | \exists D$ **使得** $U = U_D$ 或 $U = V_D\}$。我们假设 $\mathcal{U}$ 是一个具有 t 个部分的划分，并且 $|U_J|$ 关于 N 是线性的。

引理 4.7。设 $\mathcal{B}$ 是一组图诱导的支持访问对象，使得对于所有 $B^* \in \mathcal{B}$，图 G_{B^*} 都具有均匀度 d。设 $\mathcal{D}$ 是上文定义的区分器类。那么对于每个常数 ε，都存在一个针对类 $\mathcal{B}$、关于 $\mathcal{D}$ 的 $(\varepsilon, \delta)-$ 学习算法 L。该算法的运行时间为 $\mathrm{poly}(1/\varepsilon, \log(1/\delta))$。

我们介绍一种实现和一种学习算法：

表示： 模型 M 由划分 $\mathcal{U}$ 和函数 $k: \binom{t}{2} \to \mathbb{N}$ 表示。函数 k 表示任意两部分之间的边数。我们假设对于所有 $j \in [t]$，均有 $\sum_{i \in [t]} k(\{i, j\}) = d \cdot |U_j|$。

实现： 实现 T 使用一个随机预言机 R，其定义在划分 $\mathcal{U}$ 和函数 k 上。它随机采样一个 $u \in [N]$，并按以下方式输出一条边 (u, v)：

1. 令 $U_j \in \mathcal{U}$ 为划分中 u 所属的部分。我们假设顶点 U_j、u_1、……、$u_{|U_J|}$ 有一个确定的顺序。2. 随机选择一个 $\ell \in [d]$。3. 按以下顺序，令 E_{U_j} 为从 U_j 出发的所有出边组成的列表：

$$E_{U_j} = (u_1^1, \ldots, u_{|U_j|}^1, u_1^2, \ldots, u_{|U_j|}^d)$$

令 $F_{U_j} \in [t]^{|E_{U_j}|}$ 为向量：$F_{U_j} = (1^{k(\{j,1\})}、2^{k(\{j,2\})}、\ldots\ldots、t^{k(\{j,t\})})$。也就是说，1 在 F_{U_j} 中恰好出现 $k(j、1)$ 次。

4. 计算 $i = F_{U_j}(u^\ell)$。

5. 使用 R，在 $\left\{(u'、l') \in U_j | F_{U_j}((u')^{l'}) = i\right\}$ 与集合 $\left\{(u'、l') \in U_i | F_{U_i}((u')^{l'}) = j\right\}$ 之间进行随机匹配。注意，这两个集合的大小均为 $k(\{i, j\})$。对于 $i = j$，任选一个不会将任何元素映射到自身的匹配。如果某个顶点 u' 在集合中出现多次，则从 U_j 到 U_i 中的匹配集合选取一个匹配，然后对于 u' 的第二条边，选取该匹配的一个平移（反之亦然）。

6. 设 $v \in U_i$ 是上述随机匹配中与 $(u,\ \ell)$ 匹配的顶点，输出 v。

我们指出，随机置换被选为必然的，因此如果选中 v，那么对应的邻居将是 u。

学习算法： 我们描述一种用于学习函数 k 的算法。

1. 从 B^* 中采样边，令 $S = (u_1,\ v_1) \ldots, (u_m,\ v_m)$ 为该样本。对于每个 $i,\ j \in [t]$，按 $i \neq j$ 设置 $k(\{i, j\})$

$$k(\{i, j\}) = \frac{|\{(u, v) \in S | u \in U_i, v \in U_j\}|}{|S|} dN.$$

以及 $i = j$ 这个值的两倍。

2. 对每个 $i \in [t]$，计算 $e_i = \sum_{j\in[t]} k(\{i,j\}) - d\,|U_i|$。3. 如果存在 $i \in [t]$ 使得 $|e_i| \geq d\varepsilon/2\,|U_i|$，则退出算法并输出任意一个$k$ 使得对于所有 i，都有 $e_i = 0$ 。4. 当存在满足 $e_i \neq 0$的 i 时：选取满足 $e_i > 0$ 的 i 以及满足 $e_{i'} < 0$的 i' 。随机选取一个$j \in [t]$，使得可以将 $k(\{i,j\})$ 减 1，并将 $k(\{i',j\})$ 加 1，然后执行此操作。

注 4.1. 第一项上的初始值 $k(\{i,j\})$ 的构造方式使得 $\sum_{i,j\in[t]} k(i,j) = 2dN$，因此我们有 $\sum_{i\in[t]} e_i = 0$。这意味着，在项目中，如果存在 i 使得 $e_i > 0$，那么也必定存在 i' 使得 $e_{i'} < 0$。因此，校正算法总会终止。

实现 T 为每个顶点 $u \in U$恰好输出 d 个邻居。我们证明，对于每个 $D \in \mathcal{D}$，所得模型都是不可区分的。

引理 *4.7* 的证明。 我们首先证明，对于学习算法生成的样本，以高概率有 $|e_i| < \frac{\varepsilon}{2}\,|U_i|$ 对所有 $i \in [N]$都成立。

固定某个 $i \in [t]$。规范图 B^* 的均匀度为 d，这意味着

$$\sum_{j\in[t]} |\{(u,v) \in B | u \in U_i, v \in U_j\}| = d\,|U_i|\,.$$

对于每一条采样得到的边（u，v），令 $I_{(u,v)}$ 为指示随机变量，表示该边是否对某个 $j \in [t]$ 的 $k(\{i,j\})$ 有贡献。那么我们知道 $\sum_{(u,v)\in S} I_{(u,v)}$ 的期望值。因此，根据切尔诺夫界，以 $\exp(-\frac{|U_i|}{N} \cdot |S|\,\varepsilon^2)$ 的概率，所有近似值与期望值的差距都在 $\varepsilon/2$ 以内。

Item 4 中 k 的值所做的修正可以忽略不计，因为对于每个大小与 N 成线性的组 U_j，它们最多修正 $t^2 d$ 条边。

令 D 为一个区分器，其在划分中的集合为 U_i，U_j。那么根据模型，采样到一条边且满足 U_i，U_j 的概率为 $k(\{i,j\})/(dN)$。因此，如果 $k_{i,j}$ 的误差在 $\varepsilon/2Nd$ 因子以内，那么对于 D，该模型是不可区分的。□

5 不可能性

我们的工作与 [GGN10] 的一个主要区别，在于我们旨在与之不可区分的目标分布/对象。在 [GGN10]，中，目标分布是固定的，并且在许多对象上均匀分布；而在我们的设定中，目标是一个起初未知的单一对象，需要学习器访问目标对象，才能创建一个不可区分模型。这一区别使我们的设定更具挑战性。下文将展示一些由于这一差异而无法在我们的设定中完成的任务示例。

利用条目访问来欺骗区分器很难。 在 [GGN10]，中，区分器可以查询对象中的特定条目。在我们的设定中，这类区分器可能根本无法被欺骗。例如，假设目标对象 B^* 是由函数 f^*：$X \to \{0,1\}$ （定义 2.3）导出的条目访问对象，并且我们的学习器旨在输出一个模型 M，该模型属于由函数 f：$X \to \{0,1\}$ 导出的条目访问对象 B。对于每个 $x \in X$，假设都有一个区分器查询 $f(x)$ 的值，并且当且仅当 $f(x) = 1$ 时输出“接受”。要欺骗这些

区分器，我们必须精确学习目标函数 f^*；如果定义域 X 的大小呈指数级，而学习器只能进行多项式次查询，这显然是不可能的。

定理 5.1。设 X 是一个非空有限集。设 $\mathcal{B}$ 是由所有函数 $f: X \to \{0,1\}$所诱导的条目访问对象类。设 $\mathcal{D}$ 是区分器类，其中 D_x 适用于每个 $x \in X$：给定对象 B后，区分器 D_x 当且仅当答案 $a \sim B(x)$ 等于1时输出"接受"。设 L 是针对类 $\mathcal{B}$、关于 $\mathcal{D}$ 且适用于 ε、$\delta < 1/2$的一个(ε, δ)-学习器。那么在最坏情况下，L 需要查询每个输入 $x \in X$。

证明。反证法，假设每当 L 输出一个模型时，总存在一个未被 L查询的 $x \in X$。考虑目标对象 $B^* \in \mathcal{B}$ 是从均匀分布中抽取的情况。也就是说，B^* 由 f^* 诱导：随机选择 $X \to \{0,1\}$，使得对于每个 $x \in X$，$f^*(x)$ 都独立且均匀地服从 $\{0,1\}$。当 L 在获得随机目标模型 B^*的条目访问权限后输出模型 M 时，对于未被 L 查询的 $x \in X$，$f^*(x)$ 的条件分布仍是 $\{0,1\}$上的均匀分布，因此，关于 f^*中的随机性，其条件概率至少为 $1/2$，有 $\Pr_{B\sim M,\ a\sim B(x)}[a = f^*(x)] \le 1/2$，因此

$$|\Pr[D_x^{B^*} = \text{“accept”}] - \mathbb{E}_{B\sim M}[\Pr[D_x^B = \text{“accept”}]]| \ge 1/2.$$

根据全概率定律，以至少 $1/2$ 的（无条件）概率，存在一个区分器 $D \in \mathcal{D}$，使得

$$|\Pr[D^{B^*} = \text{“accept”}] - \mathbb{E}_{B\sim M}[\Pr[D^B = \text{“accept”}]]| \ge 1/2.$$

这与以下事实矛盾：L 是一个具有 ε，$\delta < 1/2$ 的（ε，δ）学习器。 □

学习到的模型需要比区分器更强。在 [GGN10] 中学习到的模型可以欺骗电路复杂度显著高于模型本身的区分器。下面我们将证明，在我们的设置中，如果目标是单个对象，这一点可能无法实现。

定理 5.2 ([TTV09] 中的注 **1.6)**。设 n、$W > 1$ 是满足 $W \log W \le 2^n/C$ 的正整数，其中 $C > 0$ 是一个充分大的绝对常数。存在一个由函数 $f^*: \{0,1\}^n \to \{0,1\}$ 诱导的样本访问对象 $B^{*\tilde{}}$，以及一个电路复杂度为 $O(nW)$ 的区分器 D，使得对于任何电路复杂度至多为 W 的模型 M，都有

$$|\Pr[D^{B^*} = \text{“accept”}] - \mathbb{E}_{B\sim M}[\Pr[D^B = \text{“accept”}]]| > 1/3. \tag{27}$$

证明。定义 $k := CW \log W \le 2^n$。根据已知的 k-路独立哈希函数构造[CW79, WC81],，存在一个关于 $f^*: \{0,1\}^n \to \{0,1\}$ 的分布，具有以下性质：

1. 对任意互不相同的 $x_1, \ldots, x_k \in \{0,1\}^n$，$k$ 元组 $(f^*(x_1) \ldots, f^*(x_k))$ 在 $\{0,1\}^k$ 上均匀分布；

2. $^{\tilde{}}$ $^{\tilde{}}$. 从该分布中抽取的每个 f^*，其电路复杂度为 $O(nk) = O(n \quad W)$.

考虑任意一个固定的模型 M。根据 k-wise 独立随机变量的集中不等式[BR94],，有

$$\Pr_{f^*}\left[\Pr_{B\sim M,(x,y)\sim B(\perp)}[y = f^*(x)] \ge 2/3\right] < 2^{-\Omega(k)} = 2^{-\Omega(CW\log W)}.$$

选取足够大的 C，并对所有电路规模至多为 S 的 $2^{O(W_{\text{对数}}\ W)}$ 个模型应用并集界，存在一个电路复杂度为 $O(nW)$ 的 f^*，使得对于每个电路规模至多为 W 的模型 M，

$$\Pr_{B\sim M,(x,y)\sim B(\perp)}[y = f^*(x)] < 2/3. \tag{28}$$

现在，我们选择 D 作为区分器：对于样本 $(x,\ y)$，当且仅当$y = f^*(x)$ 时输出“接受”。不等式 (28) 推出

$$\mathbb{E}_{B\sim M}[\Pr[D^B = \text{“accept”}]] < 2/3. \tag{29}$$

令 B^* 为由 f^*诱导的对象。根据我们对 D 和 B^*的定义，$\Pr[D^{B^*} = \text{“接受”}] = 1$。将其与 (29) 结合，便可得到所需的 (27)。 □

区分器类需要是可学习的。由于 [GGN10]中的目标分布是固定的，因此无需学习即可生成一个不可区分模型。在我们的设定中，学习任务通常由审计器执行；审计器可以被视为区分器类的弱无关学习器。一个自然的问题是：如果不存在这样的弱无关学习器，我们是否仍能实现不可区分性？先前的研究 [HJKRR18, GHK+23]针对某些不可区分性概念（例如校准多精度）给出了否定答案，方法是证明这些概念蕴含区分器类的（强）不可知学习。我们用于生成模型的不可区分性概念更接近多重准确性；下文将证明，要高效地实现这一概念，区分器类必须是可高效实现学习的。对于真实函数 f^*：$X \to \{0,1\}$，多重准确性要求预测器p：$X \to [0,1]$ 满足

$$|\,\mathbb{E}[(f^*(x) - p(x))g(x)]| \le \varepsilon \tag{30}$$

对于某个类 $\mathcal{G}$ 中的每个函数 g。现在考虑 $\mathcal{G}$ 由函数 g：$X \to \{-1,1\}$ 组成的情况。对于任意 $g^* \in \mathcal{G}$，假设真实函数 f^* 满足：若 $g^*(x) = 1$，则 $f^*(x) = 1$；若 $g^*(x) = -1$，则 $f^*(x) = 0$。那么 (30) 表明

$$\mathbb{E}\,|f^*(x) - p(x)| \le \varepsilon. \tag{31}$$

现在我们定义：若 $p(x) \ge 1/2$，则 $\hat{g}(x) = 1$；若 $p(x) < 1/2$，则 $\hat{g}(x) = -1$。容易验证，若对于某个 $x \in X$，$\hat{g}(x) \ne g^*(x)$，则 $|f^*(x) - p(x)| \ge 1/2$，因此 (31) 给出了类 $\mathcal{G}$ 的如下可实现学习保证：

$$\Pr[\hat{g}(x) \ne g^*(x)] \le 2\varepsilon.$$

参考文献

[ABFX08] Edoardo M Airoldi、David M Blei、Stephen E Fienberg 和 Eric P Xing。混合成员随机块模型。机器学习研究杂志：*JMLR*，9:1981–2014，2008。3

[AGL+17] Sanjeev Arora、Rong Ge、Yingyu Liang、Tengyu Ma 和 Yi Zhang。生成对抗网络（GANs）中的泛化与均衡。载于 Doina Precup 和 Yee Whye

Teh 编，《第34届国际机器学习会议论文集》，机器学习研究论文集第70卷，第224–232页。PMLR，2017年8月6日至11日。9[BEHW89] Anselm Blumer、Andrzej Ehrenfeucht、David Haussler 和 Manfred K. Warmuth。可学习性与 Vapnik-Chervonenkis 维数。《美国计算机学会杂志》36(4):929–965，1989年。30[BM84] Manuel Blum 和 Silvio Micali。如何生成密码学强序列的伪随机比特。《SIAM计算杂志》13(4):850–864，1984年。3[BR94] M. Bellare 和 J. Rompel。随机性高效的无关采样。载于《第35届计算机科学基础年度研讨会论文集》，第276–287页，1994年。43[CCL18] Yi-Hsiu Chen、Kai-Min Chung 和 Jyun-Jie Liao。关于辅助输入模拟的复杂性。载于《密码学进展——EUROCRYPT 2018》第III部分，计算机科学讲义第10822卷，第371–390页。施普林格，Cham，2018年。5, 10[CW79] J. Lawrence Carter 和 Mark N. Wegman。通用哈希函数类。《计算机与系统科学杂志》18(2):143–154，1979年。43[DKR+21] Cynthia Dwork、Michael P Kim、Omer Reingold、Guy N Rothblum 和 Gal Yona。结果不可区分性。载于《第53届年度ACM SIGACT理论计算研讨会论文集》，第1095–1108页，2021年。5, 6, 10[DLLT23] Cynthia Dwork、Daniel Lee、Huijia Lin 和 Pranay Tankala。关于多重校准的新见解。arXiv预印本 arXiv:2301.08837，2023年。5[FK99] Alan Frieze 和 Ravi Kannan。矩阵的快速近似及其应用。Combinatorica，19(2):175–220，1999年。4, 7, 8[GGM86] Oded Goldreich、Shafi Goldwasser 和 Silvio Micali。如何构造随机函数。《ACM杂志》33(4):792–807，1986年10月。3, 5[GGN10] Oded Goldreich、Shafi Goldwasser 和 Asaf Nussboim。大型随机对象的实现。《SIAM计算杂志》39(7):2761–2822，2010年。4, 5, 6, 9, 11, 12, 15, 40, 42, 43, 44[GHK+23] Parikshit Gopalan、Lunjia Hu、Michael P. Kim、Omer Reingold 和 Udi Wieder。通过结果不可区分性透视损失最小化。载于 Yael Tau-man Kalai 编，《第14届理论计算机科学创新会议（ITCS 2023）》，莱布尼茨国际信息学会议录（LIPIcs）第251卷，第60:1–60:20页，德国达格施图尔，2023年。Schloss Dagstuhl – Leibniz-Zentrum f¨ur Informatik。10, 44[GKR+22] Parikshit Gopalan、Adam Tauman Kalai、Omer Reingold、Vatsal Sharan 和 Udi Wieder。全能预测器。载于 Mark Braverman 编，《第13届理论计算机科学创新会议》，ITCS 2022，2022年1月31日至2月3日，美国加利福尼亚州伯克利，LIPIcs第215卷，第79:1–79:21页。Schloss Dagstuhl - Leibniz-Zentrum f¨ur Informatik，2022年。10

[GKSZ22] Parikshit Gopalan、Michael P Kim、Mihir A Singhal 和 Shengjia Zhao。低阶多重校准。载于 Po-Ling Loh 和 Maxim Raginsky 编，第三十五届学习理论会议论文集，计算机科学讲义第178卷之机器学习研究论文集，第3193–3234页。PMLR，2022年7月2日至5日。5[GPA M+14] Ian Goodfellow、Jean Pouget-Abadie、Mehdi Mirza、Bing Xu、David Warde-Farley、Sherjil Ozair、Aaron Courville 和 Yoshua Bengio。生成对抗网络。载于 Z. Ghahramani、M. Welling、C. Cortes、N. Lawrence 和 K.Q. Weinberger 编，神经信息处理系统进展，第27卷。Curran Associates公司，2014年。9[HJKRR18] Ursula H´ebert-Johnson、Michael Kim、Omer Reingold 和 Guy Rothblum。多重校准：针对（可计算识别的）群体进行校准。载于 国际机器学习会议，第1939–1948页。PMLR，2018年。5、6、10、14、32、44[HLL83] Paul W. Holland、Kathryn Blackmond Laskey 和 Samuel Leinhardt。随机块模型：初步研究。*Social Networks*，5(2):109–137，1983年。3[HLNRY22] Lunjia Hu、Inbal Livni-Navon、Omer Reingold 和 Chutong Yang。约束优化的全能预测器。*arXiv*预印本 *arXiv:2209.07463*，2022年。10[HLW06] Shlomo Hoory、Nathan Linial 和 Avi Wigderson。扩展图及其应用。美国数学学会公报，43(4):439–561，2006年。3[HP23] Lunjia Hu 和 Charlotte Peale。比较学习：两个假设类的样本复杂度理论。载于 Yael Tauman Kalai 编，第 *14*届理论计算机科学创新会议（*ITCS 2023*），莱布尼茨国际信息学会议录第251卷之莱布尼茨国际信息学会议录（*LIPIcs*），第72:1–72:30页，德国达格施图尔，2023年。Schloss Dagstuhl – Leibniz-Zentrum f¨ur Informatik。10[HPR22] Lunjia Hu、Charlotte Peale 和 Omer Reingold。度量熵对偶性与结果不可区分性的样本复杂度。载于 Sanjoy Dasgupta 和 Nika Haghtalab 编，第三十三届国际算法学习理论会议论文集，计算机科学讲义第167卷之 机器学习研究论文集，第515–552页。PMLR，2022年3月29日至4月1日。10[Imp17] Russell Impagliazzo。学习模型讲座：提升、困难分布、稠密模型、生成对抗网络和正则性之间的联系 I。https://www.ias.edu/ video/csdm/2017/1113-RussellImpagliazzo，2017年。10[JP14] Dimitar Jetchev 和 Krzysztof Pietrzak。如何伪造辅助输入。载于 密码学理论，计算机科学讲义第8349卷之计算机科学讲义，第566–590页。施普林格，海德堡，2014年。5、10[KGZ19] Michael P Kim、Amirata Ghorbani 和 James Zou。多重准确性：用于分类公平性的黑盒后处理。载于*2019*年*AAAI/ACM*人工智能、伦理与社会会议论文集，第247–254页，2019年。5、6、14、32

[KKG+22] Michael P. Kim、Christoph Kern、Shafi Goldwasser、Frauke Kreuter 和 O mer Reingold。普适适应性：与倾向评分竞争的目标无关推断。《美国国家科学院院刊》，119(4):e2108097119，2022年。10[KP23] Michael P. Kim 和 Juan C. Perdomo。在结果可执行性下进行决策。载于 Yael Tauman Kalai 编，第14届理论计算机科学创新会议（ITCS 2023），第251卷《莱布尼茨国际信息学会议录（LIPIcs）》，第79:1–79:15页，德国达格施图尔，2023年。Schloss Dagstuhl – Leibniz-Zentrum f¨ur Informatik。10[KR03] Yoshiharu Kohayakawa 和 Vojtech R¨odl。Szemer´edi 正则性引理与拟随机性。《算法与组合数学最新进展》，第 289–351页，2003年。8、33、34[LMR91] Nathan Linial、Yishay Mansour 和 Ronald L. Rivest。关于可学习性与 Vapnik-Chervonenkis 维数的结果。《Inform. and Comput.》，90(1):33–49，1991年。30[LR88] Michael Luby 和 Charles Rackoff。如何从伪随机函数构造伪随机置换。《SIAM计算杂志》，17(2):373–386，1988年。3、5[Mit03] Michael Mitzenmacher。幂律和对数正态分布生成模型简史。《互联网数学》，1(2):226–251，2003年。3[NN93] Joseph Naor 和 Moni Naor。小偏差概率空间：高效构造及应用。《SIAM计算杂志》，22(4):838–856，1993年8月。3[NN07] Moni Naor 和 Asaf Nussboim。实现大型稀疏随机图。载于 Moses Charikar、Klaus Jansen、Omer Reingold 和 Jos´e D. P. Rolim 编，《近似、随机化与组合优化：算法与技术》，第10届国际研讨会 APPROX 2007 和第11届国际研讨会 RANDOM 2007，美国新泽西州普林斯顿，2007年8月20日至22日，论文集，计算机科学讲义第4627卷，第596–608页。施普林格，2007年。5、9[NNT05] Moni Naor、Asaf Nussboim 和 Eran Tromer。高效可构造且保持随机图一阶性质的大型图。载于 Joe Kilian 编，《密码学理论》，第二届密码学理论会议，TCC 2005，美国马萨诸塞州剑桥，2005年2月10日至12日，论文集，计算机科学讲义第3378卷，第66–85页。施普林格，2005年。5、9[NR02] Moni Naor 和 Omer Reingold。构造具有预定结构的伪随机置换。《密码学杂志》，15(2):97–102，2002年1月。5[Sco11] Alexander Scott。矩阵和稀疏图的 Szemer´edi 正则性引理。《组合学、概率与计算》，20(3):455–466，2011年。8、33、34[Sk´o16a] Maciej Sk´orski。辅助输入模拟：再探。载于《密码学理论》，第一部分，计算机科学讲义第9985卷，第159–179页。施普林格，柏林，2016年。5、10

[Sk´o16b] Maciej Sk´orski. 计算距离的次梯度算法及其在密码学中的应用。密码学电子印本档案，2016。5，10[Sk´o17] Maciej Sk´orski. 正则性引理的密码学视角：更简单的统一证明与更精细的界。载于计算模型理论与应用，第10185卷计算机科学讲义，第586–599页。施普林格，Cham，2017。5，10[TTV09] Luca Trevisan、Madhur Tulsiani 和 Salil P. Vadhan。正则性、提升，以及对每一种高熵分布的高效模拟。载于第*24*届*IEEE*计算复杂性年度会议论文集，*CCC 2009*，巴黎，法国，*2009*年*7*月*15*日至*18*日，第126–136页。IEEE计算机学会，2009。5，9，10，14，32，43[Vad12] Salil P. Vadhan。伪随机性。理论计算机科学基础与趋势® 理论计算机科学，7(1–3)：1–336，2012。3[VC71] V. N. Vapnik 和 A. Ya. Chervonenkis。关于事件相对频率一致收敛到其概率。概率论及其应用，16(2)：264–280，1971。30[VZ13] Salil Vadhan 和 Colin Jia Zheng。具有密码学应用的一致极小极大定理。载于密码学进展–*CRYPTO 2013*：第*33*届密码学年度会议，美国加利福尼亚州圣巴巴拉，*2013*年*8*月*18*日至*22*日。论文集，第*I*部分，第93–110页。施普林格，2013。5，10[WC81] Mark N. Wegman 和 J. Lawrence Carter。新的哈希函数及其在身份验证和集合相等性中的应用。计算机与系统科学杂志，1981。3，43[Yao82] Andrew C. Yao。陷门函数理论与应用。载于第*23*届计算机科学基础年度研讨会（伊利诺伊州芝加哥，*1982*年），第80–91页。IEEE，纽约，1982。3