arXiv:2303.04845v1 [cs.LG] 8 Mar 2023

序贯概率赋值的平滑分析

Alankrita Bhatt
CMS，加州理工学院
abhatt@caltech.edu

Nika Haghtalab 电气工程与计算机科学，加州大学伯克利分校nika@berkeley.edu

Abhishek Shetty 电气工程与计算机科学，加州大学伯克利分校shetty@berkeley.edu

2023年3月10日

摘要

我们首次研究带上下文的序贯概率赋值问题的平滑分析。我们研究信息论最优极小极大速率，以及一种涉及最大似然估计预言机的算法化约框架。我们的方法建立了一种通用约化，将平滑对手下序贯概率赋值的极小极大速率转化为传导式学习的极小极大速率。这为参数类和 VC 维数有限的类带来了最优的（对数级）快速速率。在算法方面，我们针对一般函数类开发了一种能够高效利用最大似然估计预言机的算法。我们证明，在一般条件下，这种算法方法可实现次线性遗憾。

1 引言

序贯概率赋值——也称为对数损失下的在线学习——是一项基础问题，对信息论、统计学、金融、优化和序贯决策有着深远影响 [Rissanen, 1983, 1984, Cover, 1991, Feder 等人, 1992, Xie 和 Barron, 1997, Merhav 和 Feder, 1998, Xie 和 Barron, 2000, Yang 和 Barron, 1999, Jiao 等人, 2013年, Orabona 和 P´al, 2016年, Foster 等人，2018]。近年来，将上下文或侧信息纳入序贯概率赋值的方法受到广泛关注 [Rakhlin 和 Sridharan, 2015, Fogel 和 Feder, 2017, 2018, Foster 等人, 2018, Bhatt 和 Kim, 2021, Bilodeau 等人, 2021, Wu 等人, 2022a], ，部分原因在于这些方法与序贯决策应用、上下文老虎机问题以及马尔可夫决策过程（MDPs）中的学习新近建立了联系（例如，参见 [Foster 和 Krishnamurthy，2021] 以及 [Foster 等人, 2021a]）。在这一设定中，能够访问由上下文 x_τ（例如，某日 τ 的气象信息）和结果 $y_\tau \in \{0,1\}$（例如，τ 是否为雨天）组成的历史数据 $x_{1:t-1}, y_{1:t-1}$ 的预测者，希望在给定新上下文 x_t 的情况下预测 y_t。预测者使用概率赋值 p_t 来估计结果 $y_t = 1$ 的概率，并承受对数损失，即 $-\log p_t(y_t)$；该损失会奖励那些为实际出现的结果赋予较高概率的预测者。预测者的目标是相对于选定的预测器参考类，使自身遗憾尽可能低。

关于带上下文的序贯概率赋值，已有大量研究聚焦于上下文从未知分布中独立同分布抽取的情形（参见 [Fogel 和 Feder，2017年、Bhatt 和 Kim，2021年、Bilodeau 等人，2021年、Wu 等人，2022a] 及其中的参考文献）；这一问题也称为条件密度估计。在这些情形下，已知序贯概率

赋值在若干参考类（例如 Vapnik–Chervonenkis（VC）类）上能够实现较小的遗憾。另一方面，即使对于简单的预测器参考类，尝试考虑不可预测且具有对抗性地演变的上下文分布，也遭遇了有力的不可能性结果。例如，对于将 $p_t = \theta_0 \mathbb{1}\{x_t \leq a\} + \theta_1 \mathbb{1}\{x_t > a\}$ 赋予 $a \in [0,1]$的简单一维阈值参考类，在独立同分布情形下，遗憾有界于 $O(\log T)$ [Fogel 和 Feder， 2017] 但当上下文序列由对手选择时，其下界为 $\Omega(T)$（这是广为人知的结果，例如见 Littlestone [1988]）。面对现代应用中日益增长的适应演变中上下文的需求，这些不可能性结果表明，要获得能够指导实际应用中序贯概率赋值设计的严格保证，就必须考虑新的对抗行为模型。

近年来，平滑分析自适应对手 [Haghtalab 等人，2021年，Rakhlin 等人，2011] 已成为一种框架，使我们能够超越最坏情况对手，同时对生成序列的自适应过程仅作最少假设。在这一设定中，上下文从一系列不断演变的所谓 σ-平滑分布中选取，其密度相对于某个基准测度（例如均匀分布）有一个上界，且不超过该基准测度的 $1/\sigma$ 倍。值得注意的是，Haghtalab 等人针对 0-1 损失建立了这些方法，随后 Haghtalab 等人将其推广到回归问题；Block 等人则证明，序贯预测问题的性能保证可达到独立同分布情形下的最优性能。这引出了一个问题：对于平滑自适应序列，序贯概率赋值问题是否同样能够获得改进的极小极大遗憾界？

除了极小极大速率之外，概率赋值及其对应概念——密度估计——的一个重要特点，是存在最大似然估计（最大似然估计）等基本且自然的估计技术。对于独立同分布序列，在一般条件下，最大似然估计渐近最优，并且常被用作设计更复杂估计量的起点。超越独立同分布序列的情形，我们想知道最大似然估计能否在自适应序列上实现良好的统计表现。更广泛地说，算法视角对于序贯概率赋值问题及其在上下文老虎机和强化学习中的应用（其中算法设计与极小极大速率同样是根本性的考量）日益重要 [Agarwal 等人，2014年，Simchi-Levi 和 Xu，2022年，Foster 和 Rakhlin，2020年，Langford 等人，2007年，Foster 等人，2021b]。在这一领域，能够重新利用现有离线算法例程的预言机高效序贯决策算法备受关注 [Kalai 和 Vempala，2005年，Dud´ık 等人，2020年，Wang 等人，2022年，Kakade 等人，2007年，Simchi-Levi 和 Xu， 2022]。同样，近期关于采用 0-1 损失和回归损失的序贯预测平滑分析的进展 [Haghtalab 等人，2021年，Block 等人， 2022] 展现出潜力，有望弥合独立同分布情形与平滑自适应序列情形下可实现结果之间的计算差距和信息论差距。

本文首次研究序贯概率赋值的平滑分析，旨在理解极小极大遗憾的信息论基本极限，并为这一问题设计自然且预言机高效的算法。此外，我们还研究在平滑分析设定下，最大似然估计能否高效解决序贯概率赋值问题，同时实现较小的遗憾。据我们所知，我们的工作首次考虑了用于序贯概率赋值问题的预言机高效算法（尤其是最大似然估计）。

1.1 主要结果

化约为传导式学习。我们的第一个主要结果是：将针对平滑对手的遗憾界化约为针对（广义版本的）传导式

对手的遗憾界。也就是说，我们证明，在平滑分析设定下的极小极大遗憾，上界为如下设定下的极小极大遗憾：向学习者提供一组上下文，并限制对手只能从该集合中选取上下文。对于 $\mathcal{F}$，一个将上下文映射到 $[0,1]$，的假设类，令传导情形下、给定大小为 T 的上下文集时，学习者在 M 个时间步上的极小极大遗憾定义为 $\underline{\mathcal{R}}_T^M(\mathcal{F})$。我们在定理 3.1 中证明，对于所有 σ-平滑序列，极小极大遗憾 $\mathcal{R}_T(\mathcal{F},\sigma)$ 满足：对任意 $k>1$,

$$\mathcal{R}_T(\mathcal{F},\sigma) \le \underline{\mathcal{R}}_T^{kT}(\mathcal{F}) + T^2(1-\sigma)^k.$$

此外，在定理 3.4 中，我们通过将此设定下的最坏情形对抗性遗憾与 $\underline{\mathcal{R}}_T^{kT}(\mathcal{F})$ 的尺度敏感 VC 维数联系起来，给出了其上界；该维数是该类的一种典型离线复杂度。我们的结果对 $\mathcal{F}$ 呈对数依赖。具体而言，在 $1/\sigma$中，我们证明，对于 VC 类（以及参数类），遗憾有如下上界：推论 3.4.1，对于 VC 类（以及参数类），遗憾有如下上界：$\mathcal{R}_T(\mathcal{F},\ \sigma) \le O\left(d\log\left(\frac{T}{\sigma}\right)\right)$，其中 d是类 $\mathcal{F}$的 VC 维数。

从序贯概率赋值到最大似然估计的高效归约。我们的第二项贡献是开启了对序贯概率赋值的预言机高效算法设计的研究。具体而言，对于较小的字母表，我们设计了一种自然的算法（算法 1），该算法能够高效调用最大似然估计预言机，并在平滑设定下实现次线性遗憾。我们的定理 4.1 以类 $\mathcal{F}$ 的统计复杂度和平滑参数 σ为依据，给出了一个一般性的遗憾界。对于 VC 类，该界达到的遗憾速率为 $T^{4/5}\sqrt{\frac{d}{\sigma}}$。据我们所知，这是首个预言机高效的算法，也是针对对数损失的跟随扰动领导者类算法的首个分析。

VC 类的概率赋值。对于 VC 类 $\mathcal{F}$，我们在平滑设定下显式构造序贯概率赋值，并确立其遗憾保证。具体而言，我们基于对 $\mathcal{F}$ 的贝叶斯混合构造概率赋值，使其满足 $\mathcal{R}_T(\mathcal{F},\sigma) \le Cd\log\left(\frac{T}{\sigma}\right)$，其中 d 是类 $\mathcal{F}$ 的 VC 维数。虽然这种方法不能高效地使用预言机，但确实实现了对 T 和 σ 具有最优依赖关系的遗憾界。这启发了一个自然的未来研究方向：基于混合的方法能否以预言机高效的方式实现，或者序贯概率赋值的遗憾与计算复杂度之间是否存在权衡。

2 预备知识

设 $\mathcal{X}$ 是一个上下文集合，且 $\mathcal{Y}=\{0,1\}$。那么，所研究的问题涉及一个序贯博弈：在每个时间步 t，玩家根据上下文历史 $x_{1:t}:=(x_1,\dots,x_t)$（其中 $x_i\in\mathcal{X}$）以及相应的比特 $y_{1:t-1}\in\{0,1\}^{t-1}$，必须为下一位比特 y_t 将取何值赋予一个概率 $q(\cdot|x_{1:t},y_{1:t-1})$。一旦比特 y_t 被揭示（可能以对抗方式），玩家便承担损失 $-\log q(y_t|x_{1:t},y_{1:t-1})$，随后游戏进入下一步。对于假设类 $\mathcal{F}\subset\{\mathcal{X}\to[0,1]\}$，固定 $x_{1:T},y_{1:T}$ 时，概率赋值策略 $\mathcal{Q}=\{q(\cdot|x_{1:t},y_{1:t-1})\}_{t=1}^n$ 所对应的遗憾为

$$\mathcal{R}_T(\mathcal{F},x_{1:T},y_{1:T},\mathcal{Q}) = \sum_{t=1}^T \log\frac{1}{q(y_t|x_{1:t},y_{1:t-1})} - \inf_{f\in\mathcal{F}}\sum_{t=1}^T\log\frac{1}{p_f(y_t|x_t)} \tag{1}$$

其中 $p_f(1|x_t)=f(x_t)$；也就是说，函数 f在给定上下文 x_t的情况下，为下一位比特分配伯努利分布 $\mathrm{Bern}(f(x_t))$。

我们的统计结果适用于一般的损失函数 ℓ 以及学习者 a_t的一般动作。对于由对手指定的一组输入 $\{(x_i, y_i)\}_{i=1}^T$ 和学习者的一组动作 $\{a_i\}_{i=1}^T$，遗憾定义为

$$\mathcal{R}_T(\mathcal{F}, x_{1:T}, y_{1:T}, a_{1:T}) = \sum_{t=1}^{T} \ell(a_t, (x_t, y_t)) - \inf_{f\in\mathcal{F}} \sum_{t=1}^{T} \ell(f(x_t), (x_t, y_t)),$$

其中，对于对数损失，$a_t = q(\cdot|x_{1:t}, y_{1:t-1})$；因此，动作是$\{0,1\}$上的概率质量函数（pmf）。

式（1）中的遗憾通常在各种对手模型下进行研究；也就是说，对生成模型 x_t 和 y_t的概率假设（或缺乏此类假设）各不相同。在本文中，我们考虑最坏情况 y_t （与可实现情形形成对比，在该情形下，$Y_t \sim$ 伯努利分布$(f^*(x_t))$，其中未知量固定）以及 $X_t \sim \mathcal{D}_t$，其中 $\mathcal{D}_t$构成一个由平滑分布组成的自适应序列。

定义 2.1 (平滑分布与对手，[Haghtalab 等人，2020]）。考虑定义在 $\mathcal{X}$ 上的一个固定且已知的基分布 μ （例如，若 $\mathcal{X}$ 支持均匀分布，则可取均匀分布）。若定义在 $\mathcal{X}$ 上的分布$\mathcal{D}$ 满足：对于所有可测集合 $A \subseteq \mathcal{X}$，$\mathcal{D}(A) \leq \mu(A)/\sigma$，则称其为 σ-平滑分布。我们用 $\Delta_\sigma(\mu)$表示所有 σ-平滑分布的集合。若一个由联合分布 $\mathscr{D}$刻画的对手满足 $X_t \sim \mathcal{D}_t$ （其中 $\mathcal{D}_t$ 可能依赖于历史），且对所有 $t \in \{1, \ldots, T\}$均有 $\mathcal{D}_t \in \Delta_\sigma(\mu)$，则称其为 σ-平滑自适应对手。

σ-平滑自适应对手下的极小极大遗憾为

$$\mathcal{R}_T(\mathcal{F}, \sigma) = \inf_{\mathscr{Q}} \sup_{\sigma\text{-smoothed } \mathscr{D}} \mathbb{E}_{X_{1:T}\sim\mathscr{D}} \left[\max_{y_{1:T}} \mathcal{R}_T(\mathcal{F}, X_{1:T}, y_{1:T}, \mathscr{Q})\right],$$

其中，$\mathscr{Q}$ 是所有概率分配策略的集合。

我们尤其关注函数类 $\mathcal{F}$ 的几何性质如何影响 $\mathcal{R}_T(\mathcal{F}, \sigma)$。量化函数类“丰富性”和复杂性的覆盖数与组合维数有多种定义，但尺度敏感 VC 维数对我们尤为重要，并将在我们的结果中用到。

定义 2.2（尺度敏感 VC 维数）。设 $\mathcal{F}$ 为一个函数类。对于任意 $\alpha > 0$ 和点 $x_1, \ldots, x_m \in \mathcal{X}$，如果存在 $s_1 \ldots s_m \in \mathbb{R}$，使得对于每个 $\epsilon \in \{-1, 1\}^n$，都存在一个函数 $f \in \mathcal{F}$ 满足 $\epsilon_i \left(f(x_i) - s_i\right) \geq \frac{\alpha}{2}.$，则称 $\mathcal{F}$ 在尺度 α 下打散集合 $x_1, \ldots, x_m$。$\mathcal{F}$ 在尺度 α 下的尺度敏感 VC 维数记为 VC $(\mathcal{F}, \alpha)$，定义为满足以下条件的最大 m：存在由 m 个点 $x_1, \ldots, x_m \in \mathcal{X}$ 组成的集合，使得 $\mathcal{F}$ 在尺度 α 下打散该集合。二元函数类 $\mathcal{F}$ 的（传统）VC 维数定义为 VC $(\mathcal{F})$ $= \lim_{\alpha\to 0^+}$ VC $(\mathcal{F}, \alpha)$。

全文中，我们使用 Haghtalab 等人 [2021] 关于 σ-平滑分布的以下结果。该结果有助于我们将平滑学习约化为传导式学习。

定理 2.1 (Haghtalab 等人的耦合引理 [2021])。设 $\mathscr{D}_\sigma$ 是一个由 $t\sigma$-平滑分布构成的自适应序列，定义在 $\mathcal{X}$上。存在一个耦合 Π 使得

$$(X_1, Z_{1,1}, \ldots, Z_{1,K}, \ldots, X_t, Z_{t,1}, \ldots, Z_{t,K}) \sim \Pi$$

满足：

1. $X_1, \ldots, X_t$ 服从分布 $\mathscr{D}_\sigma$，
2. 对每个 $j \leq t$, $\{Z_{i,k}\}_{i\geq j,\ k\in[K]}$ 在给定 $X_1, \ldots, X_{j-1}$ 的条件下，在 $\mathcal{X}$ 上均匀且相互独立地分布。
3. 对任意 t，以至少 $1-(1-\sigma)^K$ 的概率，$X_t \in \{Z_{t,k}\}_{k=1:K}$。

3 到传导式学习的一般归约

本节将考察针对损失函数[1] ℓ 以及一般函数类 $\mathcal{F}$的平滑在线学习博弈中的极小极大遗憾。 第 3.1 节将说明，极小极大遗憾可以约化为针对相同损失函数和函数类的某个传导式学习版本的极小极大遗憾。 第 3.2 节将给出传导式设定下的一般上界。直接涉及遗憾的约化与近期研究（例如 Haghtalab 等人 [2021], Block 等人[2022], Haghtalab 等人 [2022]）中使用的约化之间存在一个微妙但重要的区别；后者是在遗憾的代理指标（如覆盖数和序贯复杂度）之间进行约化。这一点对于对数损失尤其重要，因为其复杂度无法用覆盖数刻画。我们将在 第 3.2.2 节进一步讨论这一点。

3.1 遗憾到遗憾的归约

我们考虑一般的损失函数 ℓ 以及学习者的一般动作 a_t。注意，我们可以将平滑设定下的扩展式极小极大值写为

$$\mathcal{R}_T(\mathcal{F},\sigma) = \sup_{\mathcal{D}_1\in\Delta_\sigma(\mu)} \mathbb{E}_{X_1\sim\mathcal{D}_1} \inf_{a_1}\sup_{y_1} \sup_{\mathcal{D}_2\in\Delta_\sigma(\mu)} \mathbb{E}_{X_2\sim\mathcal{D}_2} \inf_{a_2}\sup_{y_2}\dots \tag{2}$$
$$\dots \sup_{\mathcal{D}_T\in\Delta_\sigma(\mu)} \mathbb{E}_{X_T\sim\mathcal{D}_T} \inf_{a_T}\sup_{y_T} \mathcal{R}(\mathcal{F}, X_{1:T}, y_{1:T}, a_{1:T}).$$

为了对其进行界定，我们考虑在线学习这一概念的一种推广，称为传导式学习。在这种设定中，交互开始时，对手会为某个 $M \geq T$ 选择一组上下文 $X = \{X_i\}_{i=1}^M$，并将其提供给玩家。之后，博弈照常进行：对手在时间 t 选择 (x_t, y_t)，学习者选择一个动作a_t，并承受损失 $\ell(a_t, (x_t, y_t))$。不过，对手此时受到约束，在所有时刻 t都必须选择 $x_t \in X$。于是，我们可以将由 X 索引的极小极大遗憾定义为

$$\underline{\mathcal{R}}_T(\mathcal{F}, X) := \left[\max_{x_1\in X}\inf_{a_1}\sup_{y_1}\dots\max_{x_T\in X}\inf_{a_T}\sup_{y_T}\mathcal{R}(\mathcal{F}, x_{1:T}, y_{1:T}, a_{1:T})\right].$$

此外，将大小为 M 的集合对应的最坏情况传导式学习遗憾定义为 $\underline{\mathcal{R}}_T^M(\mathcal{F}) = \max_{X\subseteq\mathcal{X},\, |X|=M} \underline{\mathcal{R}}_T(\mathcal{F}, X)$。在下面的定理中，我们将证明，当上下文集合从基分布 μ 中抽取时，针对 σ-平滑对手的遗憾由传导式学习设定下的遗憾所界定。

定理 3.1. 设 $\mathcal{F}$ 是从 $\mathcal{X}$ 到 $\mathbb{R}$ 的任意函数类，并设 $\sigma \in (0, 1]$。那么，对于任意 T 和 k，我们有

$$\mathcal{R}_T(\mathcal{F},\sigma) \leq \underline{\mathcal{R}}_T^{kT}(\mathcal{F}) + T^2(1-\sigma)^k.$$

证明。为了用 $\mathcal{R}_T(\mathcal{F}, \sigma)$ 来给出 $\underline{\mathcal{R}}_T^{kT}(\mathcal{F}$的上界，对于某个 k，我们将考虑（2）并进行归纳推导。主要思路是注意到，由于 $\mathcal{D}_i$ 经过 σ平滑处理，在给定截至目前的历史条件下，我们可以应用定理 2.1中的耦合引理。

为便于说明，先考虑 $T = 1$这一简单情形。令 X_1, $Z_1 \dots Z_k$ 表示定理 2.1中提到的耦合。回顾一下， $X_1 \sim \mathcal{D}_1$ 且 $Z_{1:k} \sim \mu^k$。定义事件$E_1 := \{X_1 \in Z_{1:k}\}$，我们有

$$\mathcal{R}_1(\mathcal{F}, \mathscr{D}) = \mathbb{E}_{X_1\sim\mathcal{D}_1} \inf_{a_1}\sup_{y_1} \mathcal{R}_1(\mathcal{F}, X_1, y_1, a_1)$$

[1]该归约适用于满足如下性质的一般损失函数：时域为 T 时，最坏情况遗憾不超过 T。不过，为了便于说明，在本节中可以始终将 ℓ 视为对数损失。

$$
\begin{aligned}
&= \mathbb{E}_{X_1, Z_{1:k}}\left[\inf_{a_1}\sup_{y_1} \mathcal{R}_1(\mathcal{F}, X_1, y_1, a_1)\right] \\
&= \mathbb{E}_{X_1, Z_{1:k}}\left[\mathbb{1}\{E_1\}\inf_{a_1}\sup_{y_1} \mathcal{R}_1(\mathcal{F}, X_1, y_1, a_1)\right] \\
&\qquad + \mathbb{E}_{X_1, Z_{1:k}}\left[\mathbb{1}\{E_1^C\}\inf_{a_1}\sup_{y_1} \mathcal{R}_1(\mathcal{F}, X_1, y_1, a_1)\right] \\
&\le \mathbb{E}_{X_1, Z_{1:k}}\left[\mathbb{1}\{E_1\}\inf_{a_1}\sup_{y_1} \mathcal{R}_1(\mathcal{F}, X_1, y_1, a_1)\right] + \mathbb{P}(E_1^c) \qquad (3) \\
&\le \mathbb{E}_{Z_{1:k}}\left[\max_{X_1 \in Z_{1:k}} \inf_{a_1}\sup_{y_1} \mathcal{R}_1(\mathcal{F}, X_1, y_1, a_1)\right] + (1-\sigma)^k \qquad (4) \\
&= \underline{\mathcal{R}}_T^{kT}(\mathcal{F}) + (1-\sigma)^k, \qquad (5)
\end{aligned}
$$

其中（3）利用了 $\inf_{a_1}\sup_{y_1} \mathcal{R}_1(\mathcal{F},\ X_1,\ y_1,\ a_1)\ \le 1$ [2]，（4）由耦合引理推出，而（5）则根据转导学习遗憾的定义得出。

下一步是将其推广到任意 T。使这一推广成为可能的关键在于，对于所有 $t \le T$，即使在给定过去的条件下，我们也有 $D_t \in \Delta_\sigma(\mu)$，因此可以应用耦合引理。此外，我们还需要 $\mathcal{R}_T \le T$ 对任意序列都成立；正如上文所述，对于对数损失等合理的损失函数，这一点确实有保证。完整证明见附录 B。

□

定理 3.1 表明，我们可以将计算平滑对手的极小极大遗憾这一问题，归约为计算传导式学习的极小极大遗憾。注意，第二项 $(1-\sigma)^k \le e^{-k\sigma}$，因此，要得到次线性界，就需要考虑 $k = c\log T/\sigma$，其中 c 是适当的绝对常数。正如下一节将要看到的，这会使结果对 σ^{-1} 呈对数依赖。此外，根据 定理 3.1，且由于 $\mathbb{E}_{X\sim\mu^T}\underline{\mathcal{R}}_T(\mathcal{F},\ X) \le \mathcal{R}_T(\mathcal{F},\ \sigma)$，我们可以看出，除 polylog $\left(\frac{T}{\sigma}\right)$ 因子外，传导式学习遗憾恰好刻画了平滑遗憾。

3.2 转导学习遗憾的界

本节讨论如何为转导学习遗憾 $\underline{\mathcal{R}}_T^M(\mathcal{F})$ 给出上界，从而通过 $\mathcal{R}_T(\mathcal{F}, \sigma)$ 得到 定理 3.1 中的界。

3.2.1 使用覆盖数

在线学习中一种常见的方法，是根据函数类 $\mathcal{F}$ 的几何性质（例如覆盖数）来刻画遗憾。所需的覆盖概念因损失函数和数据的随机性质而异——通常，完全对抗性问题需要更强的序贯覆盖概念 [Rakhlin 等人，2015a，b] 而随机问题通常只需较弱的离线覆盖即可。在我们的平滑情形下，我们证明，定义定义 2.2 中所述的尺度敏感 VC 维数这一离线复杂度概念是

[2] 注意，通过在每一步使用均匀概率分配这一简单策略，这对数损失也成立。

适用的。Haghtalab 等人对回归和凸 Lipschitz 损失情形也研究了类似的思路 [2022], Block 等人 [2022]。

我们首先定义一种近似概念，并据此构造覆盖；这里我们将考虑逐点近似。这一概念类似于 Wu 等人 [2022b]中的全局序贯覆盖。

˜定义 3.1。设 $\mathcal{F}$ 为一个函数类。若函数集合 $\mathcal{F}$ 满足以下条件，则称其为 $\mathcal{F}$˜的 ϵ-覆盖：对于任意 $f \in \mathcal{F}$，存在 $g \in \mathcal{F}$，使得 $\sup_{x\in\mathcal{X}} |f(x) - g(x)| \leq \epsilon$。我们用 $\mathcal{N}(\mathcal{F}, \epsilon)$ 表示 $\mathcal{F}$的最小 ϵ-覆盖的大小。

请注意，尽管定义 3.1 中的度量相当严格，但在传导式学习情形下使用这种覆盖，只要求我们考虑具有有界定义域大小的函数类。我们通过以下定理来刻画这一点。

定理 3.2（传导式学习的上界）。设 $\mathcal{F}$ 为一个函数类，且 $\epsilon > 0$。那么，

$$\underline{\mathcal{R}}_T^{kT}(\mathcal{F}) \leq \inf_{\epsilon} \left\{ \sup_{Z\subset\mathcal{X}\downarrow|Z|=kT} \log \mathcal{N}\left(\mathcal{F}|_Z, \epsilon\right) + 2\epsilon T \right\},$$

其中， $\mathcal{F}|_Z$ 是假设类 $\mathcal{F}$ 在集合 Z 上的投影。

该定理的证明基于以下关系：利用 Wu 等人 [2022b] 中提出的形式化方法，将传导式学习与有限点集上的最坏序贯预测联系起来。该定理的证明推迟至 附录 C。

接下来，我们回顾一下，覆盖数 $\mathcal{N}(\mathcal{F}, \epsilon)$ 可根据该类 $\mathcal{F}$ 的尺度敏感 VC 维数以及定义域中的点数进行界定。

定理 3.3 （Rudelson 和 Vershynin [2006]）。存在通用常数 c、C，使得对于所有 $\alpha > 0$、任意定义在有限集合 $\mathcal{F}$ 上的函数类 $\mathcal{X}$以及 $\epsilon > 0$，都有

$$\log \mathcal{N}(\mathcal{F}, \epsilon) \leq C \cdot \mathrm{VC}(\mathcal{F}, c\alpha\epsilon) \log^{1+\alpha}\left(\frac{C|\mathcal{X}|}{\mathrm{VC}(\mathcal{F}, c\epsilon)\epsilon}\right).$$

最后，将 这些内容结合起来 定理 3.1、定理 3.2 和定理 3.3 可得以下结论。

T定理 3.4（极小极大平滑遗憾与尺度敏感 VC 维数）。

$$\mathcal{R}_T(\mathcal{F}, \sigma) \leq \inf_{k,\alpha,\epsilon>0} \left\{ C \cdot \mathrm{VC}(\mathcal{F}, c\alpha\epsilon) \log^{1+\alpha}\left(\frac{CkT}{\mathrm{VC}(\mathcal{F}, c\epsilon)\epsilon}\right) + 2\epsilon T + T^2 (1-\sigma)^k \right\}.$$

我们可以针对两个特别有趣的情形，用 T 和 σ 来具体化定理 3.4 中的界：当 $\mathrm{VC}(\mathcal{F}, \epsilon)$ 的增长阶为 $d\log\left(1/\epsilon\right)$ 时（通常称为参数类），以及当 $\mathrm{VC}(\mathcal{F}, \epsilon)$ 的增长阶为 ϵ^{-p} 时（通常称为非参数类）。前一种情形的典型例子是 VC 类；对于 VC 维数为 d 的类，$\mathrm{VC}(\mathcal{F}^{\mathrm{VC}}, \epsilon) = Cd\log\left(\frac{1}{\epsilon}\right)$（例如参见 [Vershynin，2018年，定理 8.3.18]）。后一种情形的典型例子是有界变差函数， $\mathcal{F}^{\mathrm{BV}}$ 其 $\mathrm{VC}(\mathcal{F}\mathrm{BV}\,\epsilon) = \frac{C}{\epsilon}$（例如参见 [Musayeva，2020年，Bartlett 等人，1997]）。已知该类的序贯覆盖数无界 [Rakhlin 等人， 2010]，因此在最坏情况对手下不可学习——这可以简单地由如下事实得出： $\mathcal{F}^{\mathrm{BV}}$ 包含所有一维阈值。

推论 3.4.1（参数类和非参数类的收敛速率）。若 $\mathrm{VC}(\mathcal{F},\epsilon) = d\log(1/\epsilon)$，则对于足够大的 T

$$\mathcal{R}_T(\mathcal{F},\sigma) \le O\left(d\cdot \mathrm{poly}\log\left(\frac{T}{\sigma}\right)\right).$$

如果 $\mathrm{VC}(\mathcal{F},\epsilon) = \epsilon^{-p}$，那么

$$\mathcal{R}_T(\mathcal{F},\sigma) \le O\left(T^{\frac{p}{p+1}}\cdot \mathrm{poly}\log\left(\frac{T}{\sigma}\right)\right).$$

特别要注意，推论 3.4.1 表明，对于 VC 类 $\mathcal{R}_T(\mathcal{F}^{\mathrm{VC}},\ \sigma) = \widetilde{O}\left(d\cdot\log\left(\frac{T}{\sigma}\right)\right)$(界是紧的，另见同期工作 [Wu 等人， 2023] 他们给出了类似的界），而对于有界变差函数 $\mathcal{R}_T(\mathcal{F}^{\mathrm{BV}},\ \sigma) = \widetilde{O}\left(\sqrt{T}\log\left(\frac{1}{\sigma}\right)\right)$；注意，在这两种情形下，最坏情况对手的极小极大速率均按 $\Omega(T)$ 的量级增长。我们注意到，对于一般的非参数类，上述界可能并不紧，但应可通过对定理 3.4采用多尺度（链式）版本加以改进；不过本文不着重讨论这一点。

尽管上述结果对于许多感兴趣的函数类给出了极小极大意义下令人满意的界 $\mathcal{F}$，但考虑概率分配规则的显式构造仍有其价值。对于 VC 维数有限的情形，我们通过对函数类进行离散化并使用混合概率分配规则，给出一个显式的概率分配规则。具体而言，这一策略（记作 $\mathscr{Q}^{\mathrm{VC}}$）可得到最优遗憾 $\mathcal{R}_T(\mathcal{F}^{\mathrm{VC}},\sigma,\mathscr{Q}^{\mathrm{VC}}) \le O\left(d\text{对数}\left(\frac{T}{\sigma}\right)\right)$。关于正式陈述、证明及详细讨论，请参见附录 D。

3.2.2 无覆盖数的示例

这种通过约化来刻画对数损失下极小极大遗憾的方法颇为有趣，因为此前所有方法都使用了某种覆盖数——无论是序贯覆盖数，还是更强的全局覆盖概念。然而，与 0/1 损失及其他若干损失函数 [Rakhlin 等人，2015b], 形成鲜明对比的是，至少在对抗情形下，覆盖数无法刻画对数损失的极小极大遗憾。考虑上下文集上的以下函数类$\mathcal{X} = \mathbb{B}_2$ （其中 $\mathbb{B}_2$ 表示单位 ℓ_2 欧几里得球）

$$\mathcal{F}^{\mathrm{Lin}} := \left\{ x \mapsto \frac{\langle x, w\rangle + 1}{2} \,\middle|\, w \in \mathbb{B}_2 \right\}.$$

对于这一类，Rakhlin 和 Sridharan [2015] 构造了一种基于跟随正则化领导者（FTRL）的算法，其遗憾达到 $O(\sqrt{T})$。然而，Bilodeau 等人 [2020] 证明了一个以序贯覆盖数表示的遗憾上界，而且一般而言该上界无法改进——这表明序贯覆盖数不足以刻画对数损失的极小极大遗憾率。Wu 等人 [2022b] 进一步通过考察以下类别来巩固这一结论，该类别与 $\mathcal{F}^{\mathrm{Lin}}$十分相似，

$$\mathcal{F}^{\mathrm{AbsLin}} := \left\{ x \mapsto \left|\langle x, w\rangle\right| \,\middle|\, w \in \mathbb{B}_2 \right\}.$$

Wu 等人 [2022b，示例2，定理 6] 证明了 $\mathcal{F}^{\mathrm{AbsLin}}$ 的极小极大遗憾为 $\widetilde{\Theta}(T^{2/3})$，揭示了一个令人惊讶的事实：通过对假设类进行简单的线性变换（这不会改变其覆盖数），便可得到相差一个多项式因子的极小极大速率！另一方面，我们基于归约的方法无需使用任何基于覆盖的论证，因此能够得出紧确的速率（至少在 $\mathrm{poly}\log(T/\sigma)$ 范围内）。

我们指出，在无上下文（对抗性）情形下，对数损失的极小极大遗憾（信息论文献中通常称为极小极大冗余）的精确刻画，通常是通过研究该类的所谓 随机复杂度 来计算的 $\mathcal{F}$ [Rissanen, 1996]。这一方法可以扩展到带上下文的（最坏情况）传导式学习 x_1，……，x_T；在这种情况下，固定时域下的极小极大最优遗憾由归一化最大似然（NML）概率赋值达到 [Shtar'kov, 1987], 并可表示为

$$\underline{\mathcal{R}}_T^T(\mathcal{F}) = \max_{x_{1:T}} \log \left(\sum_{y_{1:T} \in \{0,1\}^T} \max_{f \in \mathcal{F}} \prod_{t=1}^{T} p_f(y_t | x_t) \right).$$

这一表达式此前已针对在线逻辑回归进行过评估 [Jacquet 等人, 2021] 以及更一般的假设类 [Wu 等人, 2022b]。一个引人深思的问题是，在上述例子表明覆盖数无法刻画随机复杂度的情况下，究竟有哪些性质决定了 $\mathcal{F}$ 随机复杂度。我们的归约提供了一种技术，可以利用对遗憾的这种细粒度理解，直接将界推广到更一般的平滑对手设定。

4 高效预言机平滑序贯概率赋值

在上一节中，我们从纯统计的角度考察了平滑对手下序贯概率赋值问题的极小极大值。本节将聚焦于算法视角，并设计一种可通过调用最大似然估计预言机高效实现的算法。我们将关注基准测度在输入空间 $\mathcal{X}$ 和标签空间 $\mathcal{Y} = \{0, 1\}$上均为均匀测度的情形。在这种设定下，我们可以访问预言机 OPT；该预言机接收数据集 $S = \{x_i, y_i\}_{i=1}^m$，并输出一个使其在 S上的损失最小的假设。也就是说，

$$\mathrm{OPT}(S) = \operatorname*{argmin}_{h \in \mathcal{F}} \frac{1}{m} \sum_{i=1}^{m} \ell\left(h, (x_i, y_i)\right).$$

对于对数损失，这对应于最大似然估计。大部分分析适用于一般的损失函数 ℓ （遗憾的缩放与取值及导数的适当界相对应），但为清晰起见，可以将 ℓ 视为对数损失。特别是，算法 1是针对对数损失编写的。

我们采用的主要框架是跟随扰动领导者（FTPL）框架。在此框架中，我们的算法在一个数据集上使用预言机，该数据集由历史样本和一组幻觉样本组成。幻觉样本旨在"稳定"算法的预测。这为我们提供了一个概率赋值 $\mathscr{Q}^{\mathrm{FTPL}} = \{q^{\mathrm{FTPL}}(\cdot | x_{1:t}, y_{1:t-1})\}_{t=1}^T$。

为了陈述遗憾界，我们需要以下概念。对于任意类 $\mathcal{F}$，定义截断类 $\mathcal{F}_\alpha$ 为 $\mathcal{F}_\alpha = \left\{ f_\alpha : f_\alpha(x) = \frac{f(x) + \alpha}{1+2\alpha} \text{其中 } f \in \mathcal{F} \right\}$. 我们还需要 Rademacher 复杂度的概念 $\mathrm{Rad}(\mathcal{F}, T) = \sup_{X \subset \mathcal{X}, |X| = T} \mathbb{E}_\epsilon \left[\sup_{f \in \mathcal{F}} \frac{1}{T} \sum_{x \in X} \epsilon_x f(x) \right]$.

定理 4.1 (主要遗憾界)。 对于任意假设类 $\mathcal{F}$ 和参数 n， α，算法 算法 1在 σ-平滑对手下的遗憾有如下界

$$\mathcal{R}_T(\mathcal{F}, \sigma, \mathscr{Q}^{\mathrm{FTPL}}) \le n \log\left(\frac{1}{\alpha}\right) + \alpha T + T\sqrt{\log\left(\frac{1}{\alpha}\right) \cdot \frac{1}{\sigma n}}$$

算法 1：概率赋值 $\mathscr{Q}^{\mathrm{FTPL}}$

输入：时间范围 T，平滑参数 σ，VC 维数 d，每步样本数参数 n，截断参数 α

对于 $t \leftarrow 1$ 到 T，执行

  生成 $N^{(t)} \sim$ 个泊松分布(n)新鲜的幻觉样本 $(\widetilde{x}_1^{(t)}, \widetilde{y}_1^{(t)}) \cdots, (\widetilde{x}_N^{(t)}, \widetilde{y}_N^{(t)})$，这些样本是在给定 N 的条件下独立同分布，其参数为 $\widetilde{x}_i^{(t)} \sim \mathcal{U}(\mathcal{X})$ 和 $\widetilde{y}_i^{(t)} \sim \mathcal{U}(\{0,1\})$。调用预言机以计算 $\tilde{h}_t \leftarrow \mathrm{OPT}\left(\{(\widetilde{x}_i^{(t)}, \widetilde{y}_i^{(t)})\}_{i\in[N^{(t)}]} \cup \{x_\tau, y_\tau\}_{\tau\in[t-1]}\right)$

  观察 x_t

  分配概率 $q^{\mathrm{FTPL}}(1|x_{1:t}, y_{1:t-1}) = \frac{\tilde{h}_t(x_t)+\alpha}{1+2\alpha}$

  接收 y_t

结束

$$+ T \cdot \inf_{m\le n}\left\{\frac{1}{\alpha}\mathrm{Rad}\left(\mathcal{F}_\alpha, n/m\right) + \frac{n\left(1-\sigma\right)^m \log\left(1/\alpha\right)}{m} + e^{-n/8}\right\}.$$

我们将针对假设类 $\mathcal{F}$ 具有有界 VC 维数的情况应用此界。对于此类假设类，已知其 Rademacher 复杂度有界。我们在以下推论中陈述这一点。

推论 4.1.1. 设 $\mathcal{F}$ 是一个假设类，其 Rademacher 复杂度有界为 $\mathrm{Rad}(\mathcal{F}_\alpha, T) = cT^{-\omega}$，则有 $\mathcal{R}_T(\mathcal{F}, \sigma, \mathscr{Q}^{\mathrm{FTPL}}) \le T^{\frac{2}{2+\omega}}\sqrt{\frac{1}{\sigma}\cdot\mathrm{poly}\log\left(\frac{Td}{\sigma}\right)}$。

请注意，特别是对于 VC 类，我们有 $\mathcal{R}_T\left(\mathcal{F}^{\mathrm{VC}}, \sigma\right)$ 按 $T^{4/5}$ 的尺度增长。若要改进这一结果以达到 推论 3.4.1 中讨论的极小极大速率，仍是一个有趣的开放问题。

备注。通过假设可以访问一个预言机（该预言机能够优化一个混合目标函数，其中涉及对数损失以及类中各函数的带符号求和），可以将遗憾的尺度改进至 $T^{3/4}$。然而，这些预言机无法用最大似然估计来作自然解释。

4.1 分析

在跟随扰动领袖框架中设计算法的主要挑战，是设计幻觉样本的分布，从而权衡算法“稳定性”（即算法在相邻时间步之间改变其预测的程度）与“扰动”（即加入幻觉样本后，算法的预测相对于输出历史样本上最佳假设的变化程度）之间的取舍。以下引理刻画了这一点。

引理 4.2 (跟随扰动领袖界)。设 ℓ 为一个凸损失函数，且 $\mathcal{F}$ 为一个假设类。设 $\mathcal{D}_t$ 表示对手在时刻 t的分布，并设 $\mathcal{Q}_t$ 表示算法 h_t 输出的假设 1的分布。那么，算法的遗憾

1 （其中，我们用 $\widetilde{s}_t = (\widetilde{x}_t, \widetilde{y}_t)$ 表示一个幻觉数据点） [3] 有界 *by*

$$\sum_{i=1}^{T} \underbrace{\mathop{\mathbb{E}}_{s_t\sim\mathcal{D}_t}\left(\mathop{\mathbb{E}}_{h_t\sim\mathcal{Q}_t}[\ell(h_t,s_t)] - \mathop{\mathbb{E}}_{h_{t+1}\sim\mathcal{Q}_{t+1}}[\ell(h_{t+1},s_t)]\right)}_{Stability} + \underbrace{\mathbb{E}\left[\sup_{h\in\mathcal{F}_\alpha}\sum_{i=1}^{N}\ell(h,\tilde{s}_t) - \ell(h^*,\tilde{s}_t)\right]}_{Perturbation} + \alpha T.$$

其中 $h^* = \operatorname{argmin}_{h\in\mathcal{F}_\alpha}\sum_{i=1}^{T}\ell(h,s_t)$。

为完整起见，我们在 附录 E 中给出证明。根据遗憾的这一分解，我们需要仔细处理这两项。为了解权衡关系，请注意，随着我们增加虚构样本的数量，扰动项通常会增大，而稳定性项通常会减小。首先，我们来关注较难处理的稳定性项。我们将采用的主要方法，是对 Haghtalab 等人 [2022] 提出的稳定性项分解进行泛化，使其即使在损失函数无界时也适用，例如对数损失。其核心思想是，将稳定性项分解为下一时间步的平均预测分布与当前时间步的分布之间的距离（由 χ^2 距离刻画），以及一项用于刻画从同一分布重新采样样本时算法预测差异的项。证明见 附录 F。 [4]

引理 4.3（ χ^2 + 泛化 $\Rightarrow$稳定性）。令 $\mathcal{Q}_t$ 表示学习者在第 t 轮对 $\mathcal{F}$ 的分布， $\mathcal{D}_t$ 表示对手在时刻 t 的分布（给定历史 s_1、 $\cdots$、 s_{t-1}）， $s_t \sim \mathcal{D}_t$ 表示时刻 t 实际发生的对抗样本， s'_t 表示其独立副本 $s'_t \sim \mathcal{D}_t$。令 $R^{(t+1)}$ 表示算法在第 $t+1$ 轮使用的随机性。那么，

$$\begin{aligned}&\mathop{\mathbb{E}}_{s_t\sim\mathcal{D}_t}\left(\mathop{\mathbb{E}}_{h_t\sim\mathcal{Q}_t}[\ell(h_t,s_t)] - \mathop{\mathbb{E}}_{h_{t+1}\sim\mathcal{Q}_{t+1}}[\ell(h_{t+1},s_t)]\right)\\ &\quad\le \sqrt{\frac{1}{2}\chi^2(\mathop{\mathbb{E}}_{s_t\sim\mathcal{D}_t}[\mathcal{Q}_{t+1}],\mathcal{Q}_t)\cdot\log\left(\frac{1}{\alpha}\right)} + \mathop{\mathbb{E}}_{s_t,s'_t\sim\mathcal{D}_t;R^{(t+1)}}[\ell(h_{t+1},s'_t) - \ell(h_{t+1},s_t)].\end{aligned}$$

根据这个引理，我们接下来要界定 χ^2 下一时刻平均预测的分布与当前时刻分布之间的散度。这通过使用 Ingster 方法界定混合分布的散度来实现。为完整起见，我们在附录 G中给出证明。

引理 4.4（ χ^2的界）。$\chi^2\left(\mathbb{E}_{s_t\sim\mathcal{D}_t}[\mathcal{Q}_{t+1}],\ \mathcal{Q}_t\right) \le \frac{2}{\sigma n}$.

接下来，我们转向引理 4.3中的第二项。注意，这一项涉及在两个独立点 $t+1$ 上评估时刻 s_t 的假设输出的损失之差，这两个点 s'_t 和 $\mathcal{D}_t$ 均取自。界定该项的主要思路是使用更强版本的耦合引理定理 2.1，该定理允许我们从基准测度的独立同分布样本中，提取按照平滑分布抽取的点所构成的子序列。这使我们能够将所需的泛化界与假设类和损失函数复合后的 Rademacher 复杂度联系起来。通过使用截断和压缩原理，我们可以得到所需的界。证明见附录 H。

[3]稍微滥用一下记号，我们将 $\mathcal{D}_t$ 视为定义在 $\mathcal{X}\times\{0,1\}$上的量。[4]对于我们分析中的具体用途，类似于Haghtalab 等人 [2022] 的一个更简单版本的引理已经足够，但我们证明一个一般版本，因为我们认为这样的版本有助于为该问题给出更优的遗憾界。

引理 4.5 (泛化)。令 h_{t+1} 表示算法算法 1在时刻$t+1$输出的假设。那么，对于任意 $m \leq n$，都有

$$\mathop{\mathbb{E}}_{s_t, s_t' \sim \mathcal{D}_t; R^{(t+1)}} \left[\ell(h_{t+1}, s_t') - \ell(h_{t+1}, s_t)\right] \leq \frac{1}{\alpha} \mathrm{RAD}\left(\mathcal{F}_\alpha, n/m\right) + \frac{n\left(1-\sigma\right)^m \log\left(1/\alpha\right)}{m} + e^{-n/8}.$$

我们希望界定的最后一项是扰动项。为界定这一项，请注意，我们将截断参数 α 设定为使算法所作预测的损失有界，因此扰动项也有界，为 $n \log \frac{1}{\alpha}$，见附录 I。定理 4.1 由上述结果结合得出。

5 结论与开放问题

本文开启了对平滑对手下序贯概率赋值的研究。我们用传导式学习的极小极大遗憾刻画极小极大遗憾，并据此给出紧确的遗憾界，例如 VC 类的遗憾界。此外，我们开启了对这一设定下预言机效率的研究，并证明对于一般的类，可以实现次线性遗憾。

我们的工作提出了若干未来研究方向。一个有趣的方向是，能否使用预言机高效算法，在某些类（例如 VC 类）上实现最优的 $O(\log(T))$ 遗憾？更一般地说，在这一设定下，使用对数损失进行预测时，获得快速速率是否存在计算障碍？

6 致谢

本研究部分由美国国家科学基金会资助，项目编号为 CCF-2145898；同时获得 C3.AI 数字化转型研究所的资助以及伯克利人工智能研究共同体的资助。本研究的部分工作是在作者作为访问学者访问西蒙斯计算理论研究所期间完成的。

参考文献

Jorma Rissanen。《通用数据压缩系统》。IEEE 信息论汇刊，29(5)：656–664，1983 年。

Jorma Rissanen。《通用编码、信息、预测与估计》。IEEE 信息论汇刊，30(4)：629–636，1984 年。

Thomas M Cover。《通用投资组合》。《数学金融》，1(1)：1–29，1991 年。

Meir Feder、Neri Merhav 和 Michael Gutman。个体序列的通用预测。《IEEE 信息论汇刊》，38(4):1258–1270，1992。

Qun Xie 和 Andrew R Barron。无记忆信源类的极小极大冗余。《IEEE 信息论汇刊》，43(2):646–657，1997。

Neri Merhav 和 Meir Feder。通用预测。《IEEE 信息论汇刊》，44 (6):2124–2147，1998。

Qun Xie 和 Andrew R Barron。数据压缩、赌博与预测的渐近极小极大遗憾。《IEEE 信息论汇刊》，46(2):431–445，2000。

Yuhong Yang 和 Andrew Barron。信息论方法确定极小极大收敛速率。《统计年鉴》，第 1564–1599 页，1999年。Jiantao Jiao、Haim H Permuter、Lei Zhao、Young-Han Kim 和 Tsachy Weissman。定向信息的通用估计。《IEEE 信息论汇刊》，59(10):6220–6242，2013年。Francesco Orabona 和 D´avid P´al。抛硬币投注与无参数在线学习。《神经信息处理系统进展》，29，2016。Dylan J Foster、Satyen Kale、Haipeng Luo、Mehryar Mohri 和 Karthik Sridharan。逻辑回归：不拘泥于标准模型的重要性。载于《学习理论会议》，第 167–208 页。PMLR，2018。Alexander Rakhlin 和 Karthik Sridharan。二元字母表与大型专家集合下的序贯概率赋值。arXiv 预印本 arXiv:1501.07340，2015。Yaniv Fogel 和 Meir Feder。关于采用对数损失的在线学习问题。载于 2017 年 IEEE 国际信息论研讨会（ISIT），第 2995–2999 页。IEEE，2017年。Yaniv Fogel 和 Meir Feder。采用对数损失的通用批量学习。载于 2018 年 IEEE 国际信息论研讨会（ISIT），第 21–25 页。IEEE，2018年。Alankrita Bhatt 和 Young-Han Kim。带侧信息的对数损失序贯预测。载于《算法学习理论》，第 340–344 页。PMLR，2021年。Blair Bilodeau、Dylan J Foster 和 Daniel M Roy。通过经验熵进行条件密度估计的极小极大速率。arXiv 预印本 arXiv:2109.10461，2021年。Changlong Wu、Mohsen Heidari、Ananth Grama 和 Wojciech Szpankowski。通过随机序贯覆盖求期望最坏情况遗憾。arXiv 预印本 arXiv:2209.04417，2022a。Dylan J Foster 和 Akshay Krishnamurthy。高效的一阶上下文老虎机：预测、分配与三角判别。《神经信息处理系统进展》，34:18907–18919，2021年。Dylan J Foster、Sham M Kakade、Jian Qian 和 Alexander Rakhlin。交互式决策的统计复杂度。arXiv 预印本 arXiv:2112.13487，2021a。Nick Littlestone。当无关属性大量涌现时如何快速学习：一种新的线性阈值算法。《机器学习》，2:285–318，1988年。Nika Haghtalab、Tim Roughgarden 和 Abhishek Shetty。面向自适应对手的平滑分析。载于 2021 年 IEEE 第62届计算机科学基础年度研讨会（FOCS），第 942–953 页。IEEE，2021年。Alexander Rakhlin、Karthik Sridharan 和 Ambuj Tewari。在线学习：随机、受约束和平滑对手。载于《神经信息处理系统进展》，第 1764–1772 页，2011年。Nika Haghtalab、Yanjun Han、Abhishek Shetty 和 Kunhe Yang。面向超越最坏情况对手的预言机高效在线学习。《神经信息处理系统进展》，2022。

Adam Block、Yuval Dagan、Noah Golowich 和 Alexander Rakhlin。平滑在线学习与统计学习一样简单。载于学习理论会议，页码 1716–1786。PMLR，2022年。Alekh Agarwal、Daniel Hsu、Satyen Kale、John Langford、Lihong Li 和 Robert Schapire。驯服怪兽：一种快速而简单的上下文老虎机算法。载于国际机器学习会议，页码 1638–1646。PMLR，2014年。David Simchi-Levi 和 Yunzong Xu。绕过怪兽：一种在可实现性条件下用于上下文老虎机的更快、更简单的最优算法。《运筹学数学》，47(3):1904–1931，2022年。Dylan Foster 和 Alexander Rakhlin。超越 UCB：借助回归预言机实现最优且高效的上下文老虎机。载于国际机器学习会议，页码 3199–3210。PMLR，2020年。John Langford、Lihong Li 和 Alexander Strehl。Vowpal Wabbit 开源项目。2007年。Dylan J Foster、Alexander Rakhlin、David Simchi-Levi 和 Yunzong Xu。上下文老虎机与强化学习的实例依赖复杂度：一种基于分歧的视角。2021b。Adam Kalai 和 Santosh Vempala。在线决策问题的高效算法。《计算机与系统科学杂志》，71(3):291–307，2005年。Miroslav Dud´ık、Nika Haghtalab、Haipeng Luo、Robert E Schapire、Vasilis Syrgkanis 和 Jennifer Wortman Vaughan。预言机高效在线学习与拍卖设计。《ACM学报》（JACM），67(5):1–57，2020年。Guanghui Wang、Zihao Hu、Vidya Muthukumar 和 Jacob Abernethy。自适应预言机高效在线学习。《神经信息处理系统进展》，2022年。Sham M Kakade、Adam Tauman Kalai 和 Katrina Ligett。运用近似算法进行博弈。载于第三十九届美国计算机协会计算理论研讨会论文集，页码 546–555，2007年。Nika Haghtalab、Tim Roughgarden 和 Abhishek Shetty。在线学习与差分隐私学习的平滑分析。载于 H. Larochelle、M. Ranzato、R. Hadsell、M.F. Balcan 和 H. Lin 编，《神经信息处理系统进展》，第 33 卷，页码 9203–9215。Curran Associates, Inc.，2020年。Alexander Rakhlin、Karthik Sridharan 和 Ambuj Tewari。序贯复杂度与一致鞅大数定律。《概率论及相关领域》，161(1-2):111–153，2015a。Alexander Rakhlin、Karthik Sridharan 和 Ambuj Tewari。通过序贯复杂度进行在线学习。《机器学习研究期刊》，16(1):155–186，2015b。Changlong Wu、Mohsen Heidari、Ananth Grama 和 Wojciech Szpankowski。基于截断贝叶斯算法的对数损失精确遗憾界。《神经信息处理系统进展》，2022b。Mark Rudelson 和 Roman Vershynin。随机过程组合学与凸体截面。《数学年刊》，页码 603–648，2006年。

Roman Vershynin。高维概率：数据科学应用导论。剑桥大学出版社，2018年。Khadija Musayeva。有界变差多类别分类器的样本复杂度结果。arXiv 预印本 arXiv:2003.09176，2020年。Peter L Bartlett、Sanjeev R Kulkarni 和 S Eli Posner。实值函数类的覆盖数。《IEEE 信息论汇刊》，43(5):1721–1724，1997年。Alexander Rakhlin、Karthik Sridharan 和 Ambuj Tewari。在线学习：随机平均、组合参数与可学习性。《神经信息处理系统进展》，23，2010年。Changlong Wu、Ananth Grama 和 Wojciech Szpankowski。在动态变化环境中的在线学习。arXiv 预印本 arXiv:2302.00103，2023年。Blair Bilodeau、Dylan Foster 和 Daniel Roy。基于自协调性的对数损失极小极大遗憾紧界。载于国际机器学习会议，页码 919–929。PMLR，2020年。Jorma J Rissanen。费 Fisher 信息与随机复杂度。《IEEE 信息论汇刊》，42(1):40–47，1996年。Yurii Mikhailovich Shtar'kov。单条消息的通用序贯编码。《Problemy Peredachi Informatsii》，23(3):3–17，1987年。Philippe Jacquet、Gil Shamir 和 Wojciech Szpankowski。类别特征值逻辑回归的精确极小极大遗憾。载于算法学习理论，页码 755–771。PMLR，2021年。Raphail Krichevsky 和 Victor Trofimov。通用编码的性能。《IEEE 信息论汇刊》，27(2):199–207，1981年。Sudeep Kamath、Alon Orlitsky、Dheeraj Pichapati 和 Ananda Theertha Suresh。从样本中学习分布。载于学习理论会议，页码 1066–1100。PMLR，2015年。Daniel A Spielman 和 Shang-Hua Teng。算法的平滑分析：为什么单纯形算法通常能在多项式时间内运行。《ACM学报》（JACM），51(3):385–463，2004年。Adam Block 和 Max Simchowitz。广义线性函数的高效且近乎最优的平滑在线学习。《神经信息处理系统进展》，2022年。Adam Tauman Kalai 和 Santosh Vempala。通用投资组合的高效算法。《机器学习研究期刊》，页码 423–440，2002年。Haipeng Luo、Chen-Yu Wei 和 Kai Zheng。对数遗憾的高效在线投资组合。《神经信息处理系统进展》，31，2018年。Julian Zimmert、Naman Agarwal 和 Satyen Kale。推动在线学习投资组合与量子态的效率—遗憾帕累托前沿。载于学习理论会议，页码 182–226。PMLR，2022年。

R´emi J´ez´equel、Dmitrii M Ostrovskii 和 Pierre Gaillard。高效且近乎最优的在线投资组合选择。arXiv 预印本 arXiv:2209.13932，2022年。Tim Van Erven、Dirk Van der Hoeven、Wojciech Kotlowski 和 Wouter M Koolen。开放问题：快速且最优的在线投资组合选择。载于《学习理论会议》，第 3864–3869 页。PMLR，2020年。Thomas M Cover 和 Erik Ordentlich。带侧信息的通用投资组合。《IEEE 信息论汇刊》，42(2):348–363，1996年。Jason E Cross 和 Andrew R Barron。面向依赖过去的目标类的高效通用投资组合。《数学金融：数学、统计学与金融经济学国际期刊》，13(2):245–276，2003年。L´aszl´o Gy¨orfi、G´abor Lugosi 和 Frederic Udina。基于非参数核的序贯投资策略。《数学金融：数学、统计学与金融经济学国际期刊》，16(2):337–357，2006年。Alankrita Bhatt、J Jon Ryu 和 Young-Han Kim。关于带连续侧信息的通用投资组合。载于《人工智能与统计国际会议》。PMLR，2023年。

Nicolo Cesa-Bianchi 和 G´abor Lugosi。《预测、学习与博弈》。剑桥大学出版社，2006。

A 相关工作

序贯概率赋值是信息论中的经典主题，相关文献十分丰富，参见 Merhav 和 Feder [1998] 的综述及其中的参考文献。尤其是，针对分布参考类上的贝叶斯混合概率赋值这一思想 [Krichevsky 和 Trofimov, 1981] 具有核心重要性——这类混合概率赋值是若干实际信息论和统计问题的最优解 [Kamath 等人, 2015]。众所周知，贝叶斯混合方法的表现往往优于“插件”方法：先从参考类中估计一个预测器，再使用该预测器 [Merhav 和 Feder, 1998]。Bhatt 和 Kim [2021], 将类似的贝叶斯混合概率赋值用于上下文概率赋值问题，并以数据依赖方式对 VC 函数类进行覆盖。随后，Bilodeau 等人 [2021]采用了将经验覆盖上的混合与所谓的“add-β”概率赋值相结合的思路。结合根据海灵格散度诱导的度量对函数类进行离散化这一关键思想，他们为若干有趣的类在可实现情形下获得了匹配的速率（即，$y_t|x_t$ 由参考类中的某个固定未知分布生成）；关于海灵格覆盖对随机数据的用途，另请参见 Yang 和 Barron [1999]，其中提供了更多直观解释。Wu 等人 [2022a,b] 的近期研究也采用了经验覆盖和 add-β 概率赋值，适用于随机对手和对抗性对手。

在线学习文献中更常见的一种互补方法，是以非构造性的方式研究序贯决策问题的基本极限（即给出极小极大遗憾的界限，却不提供能够达到该遗憾的概率赋值）。Rakhlin 等人 [2015b, a] 基于序贯复杂度的方法已被用于研究对数损失Rakhlin 和 Sridharan [2015] 和Bilodeau 等人 [2020]；不过，后者指出，序贯复杂度可能无法充分刻画对数损失问题。

Spielman 和 Teng [2004] 为研究单纯形法等算法的效率而开创的平滑分析，最近已被证明能够有效绕过在线学习中分类和回归的统计与计算下界Haghtalab 等人[2021], Rakhlin 等人 [2011], Block 等人 [2022], Haghtalab 等人 [2020], Block 和 Simchowitz [2022]。这项研究表明，平滑分析是构造具有统计效率和计算效率的序贯决策问题算法的一种可行途径。

由于该问题具有基础性，文献中一直在研究序列概率赋值的计算效率概念，以及与之密切相关的投资组合选择问题。Kalai 和 Vempala [2002] 利用马尔可夫链蒙特卡罗技术，实现了 Cover 的通用投资组合算法的高效版本。近年来，人们对使用跟随正则化领导者（FTRL）类技术，同时实现低遗憾和低复杂度产生了浓厚兴趣 [Luo 等人，2018年，Zimmert 等人，2022年，J´ez´equel 等人， 2022], 另见 Van Erven 等人 [2020] 及其中的参考文献。然而，这些方法都没有考虑该问题的上下文版本，并且与预言机高效方法存在显著差异。另一方面，研究带有上下文的投资组合选择的工作 [Cover 和 Ordentlich，1996年，Cross 和 Barron，2003年，Gy¨orfi 等人，2006，Bhatt 等人， 2023]并未将预言机效率纳入考虑。

同期工作： Wu 等人 [2023] 也研究了序贯概率赋值问题（以及一般的可混合损失），并针对 VC 类实现了最优遗憾 $O(d\log(T/\sigma))$。除了平滑对手之外，他们还研究了能够刻画基准测度未知情形的一般模型。他们的工作主要处于信息论框架下，并未给出有关高效算法的任何结果。

B 第 3 节的延后证明

现在我们转向一般情形，并将通过归纳法证明。假设截至时间 $t-1$，我们已使用耦合引理，并将来自平滑分布的样本替换为来自均匀分布的样本。也就是说，假设时间 t 的归纳假设如下

$$\mathcal{R}_T \le \mathop{\mathbb{E}}_{\{Z_{1:k}\}\sim\mu} \max_{X_1\in Z_1^k} \inf_{a_1}\sup_{y_1}\dots\sup_{\mathcal{D}_t}\mathop{\mathbb{E}}_{X_t\sim\mathcal{D}_t}\inf_{a_t}\sup_{y_t}\dots \\ \dots\sup_{\mathcal{D}_T}\mathop{\mathbb{E}}_{X_T\sim\mathcal{D}_T}\inf_{a_T}\sup_{y_T}\mathcal{R}(\mathcal{F},X_{1:T},y_{1:T},a_{1:T}) + T(t-1)(1-\sigma)^k$$

根据耦合引理，存在一个耦合 Π_t，使得 $X_t, Z_{t,1}\dots Z_{t,k}\sim\Pi_t$，以及一个事件 $E_t=\left\{X_t\in\left\{Z_{t,1}\dots Z_{t,k}\right\}\right\}$，其发生概率为 $1-(1-\sigma)^k$。利用 $Z_t:=\left\{Z_{t,1}\dots Z_{t,k}\right\}$，我们有

$$\mathcal{R}(\mathcal{F},X_{1:T},y_{1:T},a_{1:T}) \le \mathop{\mathbb{E}}_{Z_1\sim\mu^k}\text{最大}_{X_1\in Z_1}\text{下确界}_{a_1}\text{上确界}_{y_1}\cdots\text{上确界}_{\mathcal{D}_t}\mathop{\mathbb{E}}_{X_t\sim\mathcal{D}_t}\text{下确界}_{a_t}\text{上确界}_{y_t}\cdots\text{上确界}_{\mathcal{D}_T}\mathop{\mathbb{E}}_{X_T\sim\mathcal{D}_T}\text{下确界}_{a_T}\text{上确界}_{y_T}\mathcal{R}(\mathcal{F},X_{1:T},y_{1:T},a_{1:T}) \le \mathop{\mathbb{E}}_{Z_1\sim\mu^k}\text{最大}_{X_1\in Z_1}\text{下确界}_{a_1}\text{上确界}_{y_1}\cdots\text{上确界}_{\mathcal{D}_t}\mathop{\mathbb{E}}_{X_t,Z_t\sim\Pi_t}\left[\mathbb{1}[E_t]\left(\text{下确界}_{a_t}\text{上确界}_{y_t}\cdots\text{上确界}_{\mathcal{D}_T}\mathop{\mathbb{E}}_{X_T\sim\mathcal{D}_T}\text{下确界}_{a_T}\text{上确界}_{y_T}\mathcal{R}(\mathcal{F},X_{1:T},y_{1:T},a_{1:T})\right)\right] + \mathop{\mathbb{E}}_{Z_1\sim\mu^k}\text{最大}_{X_1\in Z_1}\text{下确界}_{a_1}\text{上确界}_{y_1}\cdots\text{上确界}_{\mathcal{D}_t}\mathop{\mathbb{E}}_{X_t,Z_t\sim\Pi_t}\left[\mathbb{1}[E_t^c]\left(\text{下确界}_{a_t}\text{上确界}_{y_t}\cdots\text{上确界}_{\mathcal{D}_T}\mathop{\mathbb{E}}_{X_T\sim\mathcal{D}_T}\text{下确界}_{a_T}\text{上确界}_{y_T}\mathcal{R}(\mathcal{F},X_{1:T},y_{1:T},a_{1:T})\right)\right]$$

$$\leq \mathop{\mathbb{E}}_{Z_1\sim\mu^k} \text{最大}_{X_1\in Z_1} \text{下确界}_{a_1} \text{上确界}_{y_1} \cdots \text{上确界}_{\mathcal{D}_t} \mathop{\mathbb{E}}_{X_t, Z_t\sim\Pi_t} \left[\mathbb{1}[E_t] \left(\text{下确界}_{a_t} \text{上确界}_{y_t} \cdots \text{上确界}_{\mathcal{D}_T} \mathop{\mathbb{E}}_{X_T\sim\mathcal{D}_T} \text{下确界}_{a_T} \text{上确界}_{y_T} \mathcal{R}(\mathcal{F}, X_{1:T}, y_{1:T}, a_{1:T}) \right) \right] + T(1-\sigma)^k \leq \mathop{\mathbb{E}}_{Z_1\sim\mu^k} \text{最大}_{X_1\in Z_1} \text{下确界}_{a_1} \text{上确界}_{y_1} \cdots \text{上确界}_{\mathcal{D}_t} \mathop{\mathbb{E}}_{Z_t\sim\Pi_t} \text{最大}_{X_t\in Z_t} \left(\text{下确界}_{a_t} \text{上确界}_{y_t} \cdots \text{上确界}_{\mathcal{D}_T} \mathop{\mathbb{E}}_{X_T\sim\mathcal{D}_T} \text{下确界}_{a_T} \text{上确界}_{y_T} \mathcal{R}(\mathcal{F}, X_{1:T}, y_{1:T}, a_{1:T}) \right) + T(1-\sigma)^k = \mathop{\mathbb{E}}_{Z_1\sim\mu^k} \text{最大}_{X_1\in Z_1} \text{下确界}_{a_1} \text{上确界}_{y_1} \cdots \mathop{\mathbb{E}}_{Z_t\sim\mu^k} \text{最大}_{X_t\in Z_t} \left(\text{下确界}_{a_t} \text{上确界}_{y_t} \cdots \text{上确界}_{\mathcal{D}_T} \mathop{\mathbb{E}}_{X_T\sim\mathcal{D}_T} \text{下确界}_{a_T} \text{上确界}_{y_T} \mathcal{R}(\mathcal{F}, X_{1:T}, y_{1:T}, a_{1:T}) \right) + T(1-\sigma)^k$$

结合归纳假设，即可得到下一步所需的归纳假设。所需结果可通过用所有大小为 kT 的子集上的上确界对平均值进行上界估计得到。

C 定理 3.2 的证明

首先，回顾 Wu 等人对类 [2022b]. 所定义的全局序贯覆盖概念。

定义 C.1 (Wu 等人的全局序贯覆盖 [2022b])。对于任意类 $\mathcal{F}$，我们称$\mathcal{F}'_\alpha \subset \mathcal{X}^* \to [0,1]$是 α 的全局序贯覆盖， $\mathcal{F}$其尺度为 T，如果对于任意序列 $x_{1:T}$和$h \in \mathcal{F}$，存在一个 $h' \in \mathcal{F}'$，使得对于所有 i,

$$\left|h(x_i) - h'(x_{1:i})\right| \leq \alpha.$$

定理 C.1 (Wu 等人 [2022b])。如果 $\mathcal{F}'_\alpha$是 α 的全局序贯覆盖， $\mathcal{F}$其尺度为 T，那么

$$\mathcal{R}_T(\mathcal{F}) \leq \inf_{\alpha>0} \left\{ 2\alpha T + \log\left|\mathcal{F}'_\alpha\right| \right\}.$$

证明的最后，注意到 ϵ-覆盖（定义见 定义 3.1 ）给出了一个全局序贯覆盖（定义见 定义 C.1）。

D VC 类

本节中，我们为 $\mathcal{F} \subset \{\mathcal{X} \to [0,1]\}$ 是一个VC 类 的情形构造一种概率赋值。为说明这种概率赋值的动机，我们考虑无上下文情形。这是信息论中的一个经典问题，其中（渐近）最优概率赋值已知为 Krichevsky 和 Trofimov [1981] （KT）概率赋值，它是一种如下形式的贝叶斯混合

$$q_{\mathrm{KT}}(y_{1:T}) = \int_0^1 p_\theta(y_{1:T}) w(\theta) d\theta$$

对于特定的先验 $w(\theta)$。这可以按序列形式写为 $q_{\mathrm{KT}}(1|y_{1:t-1}) = \frac{\sum_{i=1}^{t-1} y_i + 1/2}{t-1+1}$，因此有时称为加 1/2 概率赋值；通过选择 $w(\theta)$ 为 Beta(β,β) 先验，可以得到相应的加-β 概率赋值。我们将混合的思想

扩展到上下文情形。具体而言，对于函数 f_1、……、$f_m \in \mathcal{F}$，可以选择如下混合概率分配：[5]

$$\prod_{i=1}^{t} q(y_i|x_{1:i}, y_{1:i-1}) =: q(y_{1:t}\|x_{1:t}) = \frac{1}{m}\sum_{j=1}^{m}\prod_{i=1}^{t}\left(\frac{p_{f_j}(y_i|x_i)+\beta}{1+2\beta}\right).$$

这是目前采用的方法，并配以精心选择的 $f_1,\dots,f_m$。我们指出，对于 VC 类，这种混合方法可以扩展到任意 可混合的 [Cesa-Bianchi 和 Lugosi，2006，第 3] 损失。

先更仔细地考察 VC 类：也就是说，每个 $f \in \mathcal{F}^{\mathrm{VC}}$ 都由三项内容表征：一个集合 $A \subseteq \mathcal{X}$，其中 $A \in \mathcal{A} \subset 2^{\mathcal{X}}$，该集合的 VC 维数 $\mathcal{A}$ 为 $d < \infty$；以及两个数 θ_0、$\theta_1 \in [0,1]$。于是，我们有

$$f_{A,\theta_0,\theta_1}(x) = \theta_0 \mathbb{1}\{x \in A\} + \theta_1 \mathbb{1}\{x \in A^C\}.$$

下面这个假设类的等价表示形式更便于使用。我们将每个 f 视为由一个元组 $f = (g, \theta_0, \theta_1)$ 表征，其中

1. θ_0， $\theta_1 \in [0,1]$2.
   $g \in \mathcal{G} \subset \{\mathcal{X} \to \{0,1\}\}$。

换言之， g 属于一类 $\mathcal{G}$ 二元函数——这正是原记号中的函数类$\{x \mapsto \mathbb{1}\{x \notin A\} \big| A \in \mathcal{A}\}$；因此显然有 VCdim($\mathcal{G}$) = d。于是，我们有$p_f(\cdot|x) = p_{g,\ \theta_0,\ \theta_1}(\cdot|x)$ =服从伯努利分布(θ_0)，如果 $g(x) = 0$；而$p_{g,\ \theta_0,\ \theta_1}(\cdot|x)$ =否则服从伯努利分布(θ_1)。

回顾针对特定 $f = (g, \theta_0, \theta_1)$的遗憾定义，对于序贯概率赋值策略 $\mathcal{Q} = \{q(\cdot|x_{1:t}, y_{1:t-1})\}_{t=1}^{T}$

$$\begin{aligned}\mathcal{R}_T(f, x_{1:T}, y_{1:T}, \mathcal{Q}) &= \sum_{t=1}^{T}\log\frac{1}{q(y_t|x_{1:t}, y_{1:t-1})} - \sum_{t=1}^{T}\log\frac{1}{p_f(y_t|x_t)} \\ &= \log\frac{p_f(y_{1:T}|x_{1:T})}{q(y_{1:T}\|x_{1:T})}\end{aligned} \tag{6}$$

其中 $q(y_{1:T}\|x_{1:T}) := \prod_{t=1}^{T} q(y_t|x_{1:t}, y_{1:t-1})$。

在平滑分析情形下，我们有 $X_t \sim \mathcal{D}_t$ 其中 $\mathcal{D}_t$ 对所有 t 都是 σ-平滑的。回顾在这种情况下，我们关注的是遗憾

$$\begin{aligned}\mathcal{R}_T(\mathcal{F}, \sigma, \mathcal{Q}) &= \max_{\mathcal{D}:\sigma\text{-smoothed}} \mathbb{E}_{X_{1:T}}\left[\max_{y_{1:T}}\sup_{f\in\mathcal{F}}\frac{p_f(y_{1:T}|X_{1:T})}{q(y_{1:T}\|X_{1:T})}\right] \\ &= \max_{\mathcal{D}:\sigma\text{-smoothed}} \mathbb{E}_{X_{1:T}}\left[\max_{y_{1:T}}\sup_{g\in\mathcal{G}}\max_{\theta_0,\theta_1}\frac{p_{g,\theta_0,\theta_1}(y_{1:T}|X_{1:T})}{q(y_{1:T}\|X_{1:T})}\right].\end{aligned}$$

[5]注意，一旦为任意的 $x_{1:t}, y_{1:t}$ 定义了混合分布 $q(y_{1:t}\|x_{1:t})$，时刻 t 的概率赋值（等价地，预测下一位比特为 1 的概率）便可定义为 $q(1|x_{1:t}, y_{1:t-1})$ $=\frac{q(y_{1:t-1}1\|x_{1:t})}{q(y_{1:t-1}\|x_{1:t-1})}$；特别是，这一预测只取决于观测到的历史 $x_{1:t}$, $y_{1:t-1}$，而不取决于未来 y_t。

D.1 提出的概率赋值

令 μ 为 $X_{1:T}$ 的 σ-平滑分布的支配测度。令 $g_1, \ldots, g_{m_\epsilon} \in \mathcal{G}$ 为函数类 $\mathcal{G}$ 在度量 $\delta_\mu(g_1, g_2) = \Pr^{X\sim\mu}(g_1(X) \neq g_2(X))$ 下的一个 ϵ-覆盖。下面的引理给出了 m_ϵ 的界。

Lemma D.1 （VC 类在度量 δ 下的覆盖数，Vershynin [2018]）.

$$m_\epsilon \leq \left(\frac{1}{\epsilon}\right)^{cd}$$

其中 c 是一个绝对常数。

按照采用混合概率分配的思路，我们对 g_1、……、g_{m_ϵ} 以及 θ_0、θ_1 取均匀混合，使得

$$q(y_{1:t}\|x_{1:t}) = \frac{1}{m_\epsilon}\sum_{i=1}^{m_\epsilon}\int_0^1\int_0^1 p_{g_i,\theta_0,\theta_1}(y_{1:t}|x_{1:t})d\theta_0 d\theta_1$$

因此，序贯概率赋值（或等价地，赋予 1 的概率）为

$$q(1|x_{1:t}, y_{1:t-1}) = \frac{q(y_{1:t-1}1\|x_{1:t})}{q(y_{1:t-1}\|x_{1:t-1})}.$$

可以观察到，$q(0|x_{1:t}, y_{1:t-1})\, q(1|x_{1:t}, y_{1:t-1}) > 0$，并且 $q(0|x_{1:t}, y_{1:t-1}) + q(1|x_{1:t}, y_{1:t-1}) = 1$，因此 q 是一个合法的概率赋值。将此均匀混合所诱导的策略称为 $\mathscr{Q}^{\mathrm{VC}}$。

D.2 平滑对手的 $\mathscr{Q}^{\mathrm{VC}}$ 分析

由（6）可知，对于上一节中定义的 $\mathscr{Q}^{\mathrm{VC}}$，我们有

$$\begin{aligned}\mathcal{R}_T((g^*,\theta_0^*,\theta_1^*), x_{1:T}, y_{1:T}, \mathscr{Q}^{\mathrm{VC}}) &= \log m_\epsilon + \log\frac{p_{g^*,\theta_0^*,\theta_1^*}(y_{1:T}|x_{1:T})}{\sum_{i=1}^{m_\epsilon}\int_0^1\int_0^1 p_{g_i,\theta_0,\theta_1}(y_{1:T}|x_{1:T})d\theta_0 d\theta_1}\\ &\leq \log m_\epsilon + \log\frac{p_{g^*,\theta_0^*,\theta_1^*}(y_{1:T}|x_{1:T})}{\int_0^1\int_0^1 p_{g_{i^*},\theta_0,\theta_1}(y_{1:T}|x_{1:T})d\theta_0 d\theta_1} \qquad (7)\end{aligned}$$

其中，$g_{i^*} \in \{g_1, \ldots, g_{m_\epsilon}\}$ 是使二进制串 $(g_{i^*}(x_1)\ldots, g_{i^*}(x_T))$ 与 $(g^*(x_1)\ldots, g^*(x_T))$ 之间的汉明距离最小的函数 $i^* \in [m]$。

现在我们仔细考察（7）中的第二项。首先，注意对于任意的 (g, θ_0, θ_1)，我们有 $p_{g,\theta_0,\theta_1}(y_{1:T}|x_{1:T})$ $=$

$$\begin{aligned}\prod_{t=1}^T p_{g,\theta_0,\theta_1}(y_t|x_t) &= \prod_{t=1}^T \theta_{g(x_t)}^{y_t}(1-\theta_{g(x_t)})^{1-y_t} = \prod_{t:g(x_t)=0}\theta_0^{y_t}(1-\theta_0)^{1-y_t}\prod_{t:g(x_t)=1}\theta_1^{y_t}(1-\theta_1)^{1-y_t}\\ &= \theta_0^{k_0(g;x_{1:T},y_{1:T})}(1-\theta_0)^{n_0(g;x_{1:T})-k_0(g;x_{1:T},y_{1:T})}\\ &\qquad \theta_1^{k_1(g;x_{1:T},y_{1:T})}(1-\theta_1)^{n_1(g;x_{1:T})-k_1(g;x_{1:T},y_{1:T})},\end{aligned}$$

其中，对于 $j \in \{0,1\}$

$$k_j(g; x_{1:T}, y_{1:T}) = |\{t : y_t = 1, g(x_t) = j\}|$$

$$n_0(g; x_{1:T}) = |\{t : g(x_t) = j\}|.$$

接下来，我们注意到，对于任意 $g \in \mathcal{G}$

$$\begin{aligned}
&\int_0^1 \int_0^1 p_{g,\theta_0,\theta_1}(y_{1:T}|x_{1:T}) d\theta_0 d\theta_1 \\
&= \left(\int_0^1 \theta_0^{k_0(g;x_{1:T},y_{1:T})} (1-\theta_0)^{n_0(g;x_{1:T}) - k_0(g;x_{1:T},y_{1:T})} d\theta_0 \right) \cdot \\
&\qquad \left(\int_0^1 \theta_1^{k_1(g;x_{1:T},y_{1:T})} (1-\theta_1)^{n_1(g;x_{1:T}) - k_0(g;x_{1:T},y_{1:T})} d\theta_1 \right) \\
&= \frac{1}{\binom{n_0(g;x_{1:T})}{k_0(g;x_{1:T},y_{1:T})}(n_0(g;x_{1:T})+1)} \frac{1}{\binom{n_1(g;x_{1:T})}{k_1(g;x_{1:T},y_{1:T})}(n_1(g;x_{1:T})+1)} \qquad (8) \\
&\geq \frac{1}{n^2 \binom{n_0(g;x_{1:T})}{k_0(g;x_{1:T},y_{1:T})} \binom{n_1(g;x_{1:T})}{k_1(g;x_{1:T},y_{1:T})}}
\end{aligned}$$

其中（8）可由拉普拉斯概率分配（或贝塔/伽马函数）的性质得出，相关性质由 引理 D.2给出。

引理 D.2。对于 $k \leq n \in \mathbb{N}$,

$$\int_0^1 t^k (1-t)^{n-k} dt = \frac{\Gamma(k+1)\Gamma(n-k+1)}{\Gamma(n+2)} = \frac{1}{(n+1)\binom{n}{k}}$$

其中 $\Gamma(\cdot)$ 表示伽马函数。

将其代回（7）中（并整理）后，得到

$$\begin{aligned}
&\mathcal{R}_T((g^*, \theta_0^*, \theta_1^*), x_{1:T}, y_{1:T}, \mathscr{Q}^{\mathrm{VC}}) - \log m_\epsilon - 2\log n \\
&\leq \sum_{j \in \{0,1\}} \log \left(\binom{n_j(g_{i^*}; x_{1:T})}{k_j(g_{i^*}; x_{1:T}, y_{1:T})} (\theta_j^*)^{k_j(g^*;x_{1:T},y_{1:T})} (1-\theta_j^*)^{n_j(g^*;x_{1:T}) - k_j(g^*;x_{1:T},y_{1:T})} \right) \\
&= \sum_{j \in \{0,1\}} \log \left(\binom{n_j(g_{i^*}; x_{1:T})}{k_j(g_{i^*}; x_{1:T}, y_{1:T})} \binom{n_j(g^*; x_{1:T})}{k_j(g^*; x_{1:T}, y_{1:T})}^{-1} \cdot \right. \\
&\qquad \left. \binom{n_j(g^*; x_{1:T})}{k_j(g^*; x_{1:T}, y_{1:T})} (\theta_j^*)^{k_j(g^*;x_{1:T},y_{1:T})} (1-\theta_j^*)^{n_j(g^*;x_{1:T}) - k_j(g^*;x_{1:T},y_{1:T})} \right) \\
&\leq \sum_{j \in \{0,1\}} \log \left(\binom{n_j(g_{i^*}; x_{1:T})}{k_j(g_{i^*}; x_{1:T}, y_{1:T})} \binom{n_j(g^*; x_{1:T})}{k_j(g^*; x_{1:T}, y_{1:T})}^{-1} \right) \qquad (9)
\end{aligned}$$

其中（9）成立，因为对于任意自然数 $k \leq n$ 和 $\theta \in [0,1]$，我们有 $\binom{n}{k}\theta^k (1-\theta)^{n-k} \leq 1$。现在注意到

$$\begin{aligned}
\log \frac{\binom{n}{k}}{\binom{n'}{k'}} &= \log \frac{n!}{n'!} + \log \frac{k'!}{k!} + \log \frac{(n'-k')!}{(n-k)!} \\
&\leq \log \frac{(n' + |n-n'|)!}{n'!} + \log \frac{(k + |k-k'|)!}{k!} + \log \frac{((n-k) + |n-n'| + |k-k'|)!}{(n-k)!}
\end{aligned}$$

若 $|k-k'|$, $|n-n'| \le \delta$, 且 $\max\{n, n'\} \le N$, 那么例如根据 [Bhatt 和 Kim, 2021年, 命题6] , 我们有

$$\begin{aligned}\log \frac{\binom{n}{k}}{\binom{n'}{k'}} &\le 2\delta \log(n' + 2\delta) + 2\delta \log(k + 2\delta) + 4\delta \log((n-k) + 4\delta) \\ &\le 16\delta \log N. \end{aligned} \tag{10}$$

现在我们希望在（9）中使用这个界。为此，我们将回顾特定函数 g 的 $n_0(g;\, x_{1:T})$ 和 $k_0(g;\, x_{1:T},\, y_{1:T})$ 的定义，并注意到对于两个函数 g、g'，有 $|n_0(g;\, x_{1:T}) - n_0(g';\, x_{1:T})|$、$|k_0(g;\, x_{1:T}, y_{1:T}) - k_0(g';\, x_{1:T}, y_{1:T})| \le d_H(g(x_{1:T})\ g'(x_{1:T}))$，其中 $d_H(\cdot,\cdot)$ 表示汉明距离，而 $g(x_{1:T}) := (g(x_1) \ldots, g(x_T)) \in \{0,1\}^T$。因此，在（9）中使用（10），并令 $\delta = d_H(g^*(x_{1:T})\ g_{i^*}(x_{1:T}))$ $N = T$，我们得到

$$\mathcal{R}_T((g^*, \theta_0^*, \theta_1^*), x_{1:T}, y_{1:T}, \mathscr{Q}^{\mathrm{VC}}) \le \log m_\epsilon + 2\log T + 32 d_H(g^*(x_{1:T}), g_{i^*}(x_{1:T})) \log T. \tag{11}$$

注意到 (11) 实际上消除了对 $y, \theta_0^*, \theta_1^*$ 的任何依赖。因此，对于某个绝对常数 C，我们有（回顾前面对 i^* 和 $\mathcal{F}$ 的定义）

$$\mathcal{R}_T(\mathcal{F}, \sigma, \mathscr{Q}^{\mathrm{VC}}) \le C \log m_\epsilon + C \log T \max_{\mathscr{D}:\sigma\text{-smoothed}} \mathbb{E}\left[\sup_{g^* \in \mathcal{G}} \min_{i \in [m_\epsilon]} d_H(g^*(X_{1:T}), g_i(X_{1:T}))\right]. \tag{12}$$

最后，我们可以通过以下结果控制 (12) 中的最后一项；该结果由耦合引理以及关于 VC 类上确界的方差敏感上界推出。

引理 D.3 (Haghtalab 等人的 引理 3.3 [2021])。

$$\mathbb{E}\left[\sup_{g^* \in \mathcal{G}} \min_{i \in [m]} d_H(g^*(X_{1:T}), g_i(X_{1:T}))\right] \le \sqrt{\frac{\epsilon}{\sigma} T \log T d \log\left(\frac{1}{\epsilon}\right)} + T \log T \frac{\epsilon}{\sigma}$$

将上述结果代入 (12)，并取 $\epsilon = \frac{\sigma}{T^2}$，得到

$$\mathcal{R}_T(\mathcal{F}, \sigma, \mathscr{Q}^{\mathrm{VC}}) \le O\left(d \log\left(\frac{T}{\sigma}\right)\right).$$

E 引理 4.2 的证明

证明。注意，这个证明适用于一般的损失函数。令 $\mathcal{R}_T$ 表示遗憾。

$$\begin{aligned}\mathcal{R}_T &\le \mathbb{E}\left[\sum_{i=1}^T \ell(h_t, s_t) - \inf_{h \in \mathcal{F}} \sum_{i=1}^T \ell(h, s_t)\right] \\ &= \mathbb{E}\left[\sum_{i=1}^T \ell(h_t, s_t) - \sum_{t=1}^T \ell(h_{t+1}, s_t) + \sum_{t=1}^T \ell(h_{t+1}, s_t) - \inf_{h \in \mathcal{F}} \sum_{i=1}^T \ell(h, s_t)\right] \\ &= \mathbb{E}\left[\sum_{i=1}^T \ell(h_t, s_t) - \sum_{t=1}^T \ell(h_{t+1}, s_t)\right] + \mathbb{E}\left[\sum_{t=1}^T \ell(h_{t+1}, s_t) - \inf_{h \in \mathcal{F}} \sum_{i=1}^T \ell(h, s_t)\right]\end{aligned}$$

我们来关注第二项。

$$
\begin{aligned}
&\mathbb{E}\left[\sum_{t=1}^{T}\ell(h_{t+1},s_t)-\inf_{h\in\mathcal{F}}\sum_{i=1}^{T}\ell(h,s_t)\right]\\
&\le\mathbb{E}\left[\sum_{t=1}^{T}\ell(h_{t+1},s_t)-\inf_{h\in\mathcal{F}_\alpha}\sum_{i=1}^{T}\ell(h,s_t)+\inf_{h\in\mathcal{F}_\alpha}\sum_{i=1}^{T}\ell(h,s_t)-\inf_{h\in\mathcal{F}}\sum_{i=1}^{T}\ell(h,s_t)\right]\\
&\le 2\alpha T+\mathbb{E}\left[\sum_{t=1}^{T}\ell(h_{t+1},s_t)-\inf_{h\in\mathcal{F}_\alpha}\sum_{i=1}^{T}\ell(h,s_t)\right] \qquad (13)\\
&\le 2\alpha T+\mathbb{E}\left[\sum_{t=1}^{N}\ell(h_t,\tilde{s}_t)-\ell\left(h^*,\tilde{s}_t\right)\right] \qquad (14)\\
&\le 2\alpha T+\mathbb{E}\left[\sup_{h\in\mathcal{F}_\alpha}\sum_{t=1}^{N}\ell(h,\tilde{s}_t)-\ell\left(h^*,\tilde{s}_t\right)\right]
\end{aligned}
$$

其中 $h^*=$下确界$_{h\in\mathcal{F}_\alpha}\sum_{i=1}^{T}\ell(h,\ s_t)$。(13) 可通过比较截断类与整个类的最优值推出，参见 [Cesa-Bianchi 和 Lugosi，2006，引理 9。5]。(14) 可由领头者获胜引理推出，参见 Cesa-Bianchi 和 Lugosi [2006]。 □

F 引理 4.3 的证明

记 $R^{(t)}=(N^{(t)},\ \{\widetilde{s}_i\}_{i\in N^{(t)}})$ 为时刻 t开始时生成的新鲜随机性，它与 $\{s_\tau\}_{\tau<t}$ 对手生成的随机性相互独立。令 $\mathcal{Q}_t$ 表示学习者在 $h_t\in\mathcal{H}$ 算法 1中的行动分布，形式化地说，

$$
r^t(x)=\sum_{i=1}^{N^{(t+1)}}\widetilde{y}_i^{(t+1)}\cdot\mathbf{1}(\widetilde{x}_i^{(t+1)}=x)+\sum_{\tau=1}^{t}y_\tau\cdot\mathbf{1}(x_\tau=x).
$$

令 $\mathcal{P}^t$ 表示 r^t的分布。引入这一概念的原因是，h_t 在算法 1 中只依赖于向量 r^{t-1}。

证明的关键步骤是从分布 $\mathcal{D}_t$ 中引入一个独立样本，以便解耦分布 $\mathcal{Q}_{t+1}$ 对测试点 s_t的依赖。

$$
\begin{aligned}
&\mathop{\mathbb{E}}_{s_t\sim\mathcal{D}_t}\mathop{\mathbb{E}}_{h_t\sim\mathcal{Q}_t}[\ell(h_t,s_t)]-\mathop{\mathbb{E}}_{s_t\sim\mathcal{D}_t}\mathop{\mathbb{E}}_{h_{t+1}\sim\mathcal{Q}_{t+1}}[\ell(h_{t+1},s_t)]\\
&=\mathop{\mathbb{E}}_{s_t\sim\mathcal{D}_t}\mathop{\mathbb{E}}_{h_t\sim\mathcal{Q}_t}[\ell(h_t,s_t)]-\mathop{\mathbb{E}}_{s_t,s_t'\sim\mathcal{D}_t}\mathop{\mathbb{E}}_{h_{t+1}\sim\mathcal{Q}_{t+1}}[\ell(h_{t+1},s_t')]\\
&\quad+\mathop{\mathbb{E}}_{s_t,s_t'\sim\mathcal{D}_t}\mathop{\mathbb{E}}_{h_{t+1}\sim\mathcal{Q}_{t+1}}[\ell(h_{t+1},s_t')]-\mathop{\mathbb{E}}_{s_t\sim\mathcal{D}_t}\mathop{\mathbb{E}}_{h_{t+1}\sim\mathcal{Q}_{t+1}}[\ell(h_{t+1},s_t)] \qquad (15)\\
&=\mathop{\mathbb{E}}_{s_t'\sim\mathcal{D}_t}\mathop{\mathbb{E}}_{h_t\sim\mathcal{Q}_t}[\ell(h_t,s_t')]-\mathop{\mathbb{E}}_{s_t'\sim\mathcal{D}_t}\mathop{\mathbb{E}}_{h_{t+1}\sim\mathbb{E}_{s_t\sim\mathcal{D}_t}[\mathcal{Q}_{t+1}]}[\ell(h_{t+1},s_t')] \qquad (16)\\
&\quad+\mathop{\mathbb{E}}_{s_t,s_t'\sim\mathcal{D}_t}\mathop{\mathbb{E}}_{h_{t+1}\sim\mathcal{Q}_{t+1}}[\ell(h_{t+1},s_t')]-\mathop{\mathbb{E}}_{s_t\sim\mathcal{D}_t}\mathop{\mathbb{E}}_{h_{t+1}\sim\mathcal{Q}_{t+1}}[\ell(h_{t+1},s_t)]
\end{aligned}
$$

其中，15）是通过加上再减去中间项得到的，该中间项对应于在独立样本上计算损失 s_t' 而（16）则是通过观察到 s_t 和 s_t' 服从相同的分布得到的。由于

第二项与所需等式中的相同，因此我们可以关注第一项。

$$\mathop{\mathbb{E}}_{h_t\sim\mathcal{Q}_t}\left[\mathop{\mathbb{E}}_{s'_t\sim\mathcal{D}_t}[\ell(h_t,s'_t)]\right]-\mathop{\mathbb{E}}_{h_{t+1}\sim\widetilde{\mathcal{Q}_{t+1}}}\left[\mathop{\mathbb{E}}_{s'_t\sim\mathcal{D}_t}[\ell(h_{t+1},s'_t)]\right]. \tag{17}$$

这里我们用记号 $\widetilde{\mathcal{Q}_{t+1}}=\mathbb{E}_{s_t\sim\mathcal{D}_t}[\mathcal{Q}_{t+1}]$ 表示混合分布。为了对其进行界定，我们考察两个分布 P 和 Q之间 χ^2 距离的一种变分解释。

引理F.1（Hammersley–Chapman–Robbins界）。对于任意一对测度 P 和 Q，以及任意可测函数 $h\colon\mathcal{X}\to\mathbb{R}$，有

$$\begin{aligned}\left|\mathbb{E}_{X\sim P}[h(X)]-\mathbb{E}_{X\sim Q}[h(X)]\right|&\le\sqrt{\chi^2\left(P,Q\right)\cdot\mathrm{Var}_{X\sim Q}\left(h\left(X\right)\right)}\\&\le\sqrt{\frac{1}{2}\chi^2\left(P,Q\right)\cdot\mathbb{E}_{X,X'\sim Q}\left(h\left(X\right)-h\left(X'\right)\right)^2}.\end{aligned}$$

将此应用于（17），得到

$$\begin{aligned}&\mathop{\mathbb{E}}_{h_t\sim\mathcal{Q}_t}\left[\mathop{\mathbb{E}}_{s'_t\sim\mathcal{D}_t}[\ell(h_t,s'_t)]\right]-\mathop{\mathbb{E}}_{h_{t+1}\sim\widetilde{\mathcal{Q}_{t+1}}}\left[\mathop{\mathbb{E}}_{s'_t\sim\mathcal{D}_t}[\ell(h_{t+1},s'_t)]\right]\\&\le\sqrt{\frac{1}{2}\chi^2(\mathop{\mathbb{E}}_{s_t\sim\mathcal{D}_t}[\mathcal{Q}_{t+1}],\mathcal{Q}_t)\cdot\mathbb{E}_{h_t,h'_t\sim\mathcal{Q}_t}\left(\mathbb{E}_{s_t\sim\mathcal{D}_t}\left[\ell(h_t,s_t)-\ell(h'_t,s_t)\right]\right)^2}.\end{aligned}$$

这就得到了所需结论。正如前文所述，对于我们分析中的具体应用，类似于 Haghtalab 等人 [2022] 提出的一个更简单的引理版本就足够了；不过我们仍给出一般版本，因为我们认为，这种版本有助于为该问题给出更优的遗憾界。

G 上界估计 χ^2 距离： 引理 4.4的证明

本节将重点界定 χ^2 各时间步动作分布之间的距离。本节的推理基本遵循 Haghtalab 等人 [2022] 的工作。为完整起见，我们在此重述其内容。

我们假设 $\mathcal{X}$ 是离散的。定义

$$n_0(x)=\sum_{i=1}^{N}\mathbf{1}(\widetilde{x}_i=x,\widetilde{y}_i=0)\quad\text{and}\quad n_1(x)=\sum_{i=1}^{N}\mathbf{1}(\widetilde{x}_i=x,\widetilde{y}_i=1).$$

由于每个 $\widetilde{x}_i$都服从 $\mathcal{X}$上的均匀分布，且 $\widetilde{y}_i\sim\mathcal{U}(\{0,1\})$，根据泊松分布的子抽样性质， $2|\mathcal{X}|$随机变量 $\{n_0(x)\ n_1(x)\}_{x\in\mathcal{X}}$独立同分布，且服从参数为 $n/2|\mathcal{X}|$的泊松分布。

由于历史数据只是一个平移，只需考虑时刻 $t=0$和 $t=1$的分布。令 $n_0^1(x)=n_0(x)+\mathbf{1}(x_1=x,y_1=0)$，并以类似方式定义 n_1^1。令 P和 Q 分别为 $\{n_0(x)\ n_1(x)\}_{x\in\mathcal{X}}$和 $\{n_0^1(x)\ n_1^1(x)\}_{x\in\mathcal{X}}$的概率分布。注意，预言机的输出仅取决于该向量，因此根据数据处理不等式，只需界定 $\chi^2(P,Q)$。

注意，分布 P是乘积泊松分布：

$$P(\{n_0(x),n_1(x)\})=\prod_{x\in\mathcal{X}}\prod_{y\in\{0,1\}}\mathbb{P}(\mathrm{Poi}(n/2|\mathcal{X}|)=n_y(x)).$$

至于分布 Q，可以通过以下方式从 P 得到：平滑对手抽取 $x^\star \sim \mathcal{D}$，且它与 $\{n_0(x)\ n_1(x)\}_{x\in\mathcal{X}} \sim P$相互独立，其中 σ-平滑分布 $\mathcal{D} \in \Delta_\sigma(\mathcal{X})$。然后，他将标签 $y^\star = y(x^\star) \in \{0,1\}$ 选为 $x^\star$的函数，并设置

$$n^1_{y(x^\star)}(x^\star) = n_{y(x^\star)}(x^\star) + 1, \qquad \text{and} \qquad n^1_y(x) = n_y(x), \quad \forall (x,y) \neq (x^\star, y(x^\star)).$$

因此，给定一个 σ-平滑分布 $\mathcal{D}$ 以及对手使用的标记函数 $y : \mathcal{X} \to \{0,1\}$，分布 Q 是一个混合分布 $Q = \mathbb{E}_{x^\star\sim\mathcal{D}^{\mathcal{X}}}[Q_{x^\star}]$，其中

$$Q_{x^\star}(\{n^1_0(x), n^1_1(x)\}) = \mathbb{P}(\mathrm{Poi}(n/2|\mathcal{X}|) = n_{y(x^\star)}(x^\star) - 1) \times \prod_{(x,y)\neq(x^\star,y(x^\star))} \mathbb{P}(\mathrm{Poi}(n/2|\mathcal{X}|) = n_y(x)).$$

我们将使用 Ingster 方法来控制混合分布 Q 与基分布 P之间的 χ^2 。

引理 G.1（Ingster 的 χ^2 方法）。对于混合分布 $\mathbb{E}_{\theta\sim\pi}[Q_\theta]$ 和任意分布P，以下恒等式成立：

$$\chi^2\left(\mathop{\mathbb{E}}_{\theta\sim\pi}[Q_\theta], P\right) = \mathop{\mathbb{E}}_{\theta,\theta'\sim\pi}\left[\mathop{\mathbb{E}}_{x\sim P}\left(\frac{Q_\theta(x)Q_{\theta'}(x)}{P(x)^2}\right)\right] - 1,$$

其中，θ' 是 θ的一个独立副本。

令 $x^\star_1, x^\star_2$ 为任意一对实例。利用分布P 和 $Q_{x^\star}$的闭式表达式，可得

$$\frac{Q_{x^\star_1}Q_{x^\star_2}}{P^2} = \frac{2|\mathcal{X}|n_{y(x^\star_1)}(x^\star_1)}{n} \cdot \frac{2|\mathcal{X}|n_{y(x^\star_2)}(x^\star_2)}{n}.$$

利用在 P下，$\{n_0(x)\ n_1(x)\}_{x\in\mathcal{X}}$独立同分布且服从泊松分布 $\mathrm{Poi}(n/2|\mathcal{X}|)$这一事实，我们有

$$\mathop{\mathbb{E}}_{\{n_0(x),n_1(x)\}\sim P}\left(\frac{Q_{x^\star_1}(\{n^1_0(x),n^1_1(x)\})Q_{x^\star_2}(\{n^1_0(x),n^1_1(x)\})}{P(\{n_0(x),n_1(x)\})^2}\right) = 1 + \frac{2|\mathcal{X}|}{n}\cdot\mathbf{1}(x^\star_1 = x^\star_2).$$

我们将利用这样一个事实：两个独立抽样结果发生 $x^\star_1$，$x^\star_2 \sim \mathcal{D}$ 碰撞的概率很小。也就是说，根据引理 G.1，我们有

$$\sqrt{\frac{\chi^2(Q,P)}{2}} = \sqrt{\frac{\chi^2(\mathbb{E}_{x^\star\sim\mathcal{D}}[Q_{x^\star}],P)}{2}} = \sqrt{\frac{|\mathcal{X}|}{n}\cdot\mathop{\mathbb{E}}_{x^\star_1,x^\star_2\sim\mathcal{D}}[\mathbf{1}(x^\star_1 = x^\star_2)]}$$
$$= \sqrt{\frac{|\mathcal{X}|}{n}\sum_{x\in\mathcal{X}}\mathcal{D}(x)^2} \leq \sqrt{\frac{|\mathcal{X}|}{n}\sum_{x\in\mathcal{X}}\mathcal{D}(x)\cdot\frac{1}{\sigma|\mathcal{X}|}} = \frac{1}{\sqrt{\sigma n}},$$

其中，最后一个不等式由 σ-平滑分布的定义得出 .

H 泛化误差的上界： 引理 4.5的证明

该定理的证明与 [Haghtalab 等人，2022年第 4.2 节中的证明类似。2]在我们的设定中，需要处理一般损失。我们将需要以下关于平滑分布的性质，它是定理 2.1 中所示Haghtalab 等人[2022]提出的耦合引理的一个略微加强的版本。

引理 H.1. 设 X_1、$\cdots$、$X_m \sim Q$ 和 P 是另一个似然比有界的分布：$dP/dQ \le 1/\sigma$。那么，利用外部随机性 R，存在一个指标 $I = I(X_1, \cdots, X_m, R) \in [m]$ 以及一个成功事件 $E = E(X_1, \cdots, X_m, R)$，使得 $\Pr[E^c] \le (1-\sigma)^m$，并且

$$(X_I \mid E, X_{\setminus I}) \sim P.$$

固定泊松化样本量 $N \sim \mathrm{Poi}(n)$的任意一个取值。根据 m 在引理 H.1中的条件。由于对任意 σ-光滑的 $\mathcal{D}_t$，均有

$$\frac{\mathcal{D}_t(s)}{\mathcal{U}(\mathcal{X}\times\{0,1\})(s)} = \frac{\mathcal{D}_t(x)}{\mathcal{U}(\mathcal{X})(x)} \cdot \frac{\mathcal{D}_t(y \mid x)}{\mathcal{U}(\{0,1\})(y)} \le \frac{2}{\sigma},$$

引理 H.1 的前提以参数 $\sigma/2$ 对 $P = \mathcal{D}_t, Q = \mathcal{U}(\mathcal{X}\times\{0,1\})$ 成立。因此，将自生成样本 $\widetilde{s}_1, \cdots, \widetilde{s}_N$ 分成 N/m 组，每组大小为 m，并执行引理 H.1 中的过程，可得到 N/m 个独立事件 $E_1, \cdots, E_{N/m}$，每个事件发生的概率至少为 $1-(1-\sigma/2)^m \ge 1 - T^{-2}$。此外，在给定每个 E_j 的条件下，我们可以选取一个元素$u_j \in \{\widetilde{s}_{(j-1)m+1}, \cdots, \widetilde{s}_{jm}\}$ 使得

$$(u_j \mid E_j, \{\widetilde{s}_{(j-1)m+1}, \cdots, \widetilde{s}_{jm}\}\setminus\{u_j\}) \sim \mathcal{D}_t.$$

为简化记号，我们将未选取样本的集合 $\{\widetilde{s}_{(j-1)m+1}, \cdots, \widetilde{s}_{jm}\}\setminus\{u_j\}$ 记为 v_j。因此，得益于不同组之间的相互独立性，以及在给定$s_{1:t-1}$ 的条件下的 $s_t \sim \mathcal{D}_t$ （注意，我们在每一轮都会抽取新鲜随机性），对于 E，$\cap_{j\in[N/m]} E_j$，有

$$(u_1, \cdots, u_{N/m}, s_t) \mid (E, s_{1:t-1}, v_1, \cdots, v_{N/m}) \overset{\text{iid}}{\sim} \mathcal{D}_t.$$

记 $h_{t+1} = O_{t+1}(\widetilde{s}_1, \cdots, \widetilde{s}_N, s_{1:t-1}, s_t)$ 为算法在时间 t 的输出，其中 $\widetilde{s}_1, \cdots, \widetilde{s}_N$ 表示幻觉数据点，而 $s_{1:t-1}, s_t$ 表示观测数据点。我们将使用 O_{t+1} 是置换不变函数这一事实。因此，对于每个 $j \in [N/m]$，我们有

$$\begin{aligned}
&\mathop{\mathbb{E}}_{s_t\sim\mathcal{D}_t, R^{(t+1)}}[\ell(h_{t+1}, s_t) \mid E] \\
&= \mathop{\mathbb{E}}_{s_t\sim\mathcal{D}_t, \widetilde{s}_1, \cdots, \widetilde{s}_N}\left[\ell(O_{t+1}(\widetilde{s}_1, \cdots, \widetilde{s}_N, s_{1:t-1}, s_t), s_t) \mid E\right] \\
&= \mathop{\mathbb{E}}_{v, s_{1:t-1}\mid E}\left(\mathop{\mathbb{E}}_{s_t, u_1, \cdots, u_{N/m}}\left[\ell(O_{t+1}(s_{1:t-1}, v, u_1, \cdots, u_{N/m}, s_t), s_t) \mid E, s_{1:t-1}, v\right]\right) \\
&= \mathop{\mathbb{E}}_{v, s_{1:t-1}\mid E}\left(\mathop{\mathbb{E}}_{s_t, u_1, \cdots, u_{N/m}}\left[\ell(O_{t+1}(s_{1:t-1}, v, u_1, \cdots, u_{j-1}, s_t, u_{j+1}, \cdots, u_{N/m}, u_j), u_j) \mid E, s_{1:t-1}, v\right]\right) \quad (18) \\
&= \mathop{\mathbb{E}}_{v, s_{1:t-1}\mid E}\left(\mathop{\mathbb{E}}_{s_t, u_1, \cdots, u_{N/m}}[\ell(O_{t+1}(s_{1:t-1}, v, u_1, \cdots, u_{N/m}, s_t), u_j) \mid E, s_{1:t-1}, v]\right) \quad (19) \\
&= \mathop{\mathbb{E}}_{s_t\sim\mathcal{D}_t, R^{(t+1)}}[\ell(h_{t+1}, u_j) \mid E],
\end{aligned}$$

其中，(18) 是由条件化后（u_1、$\cdots$、$u_{N/m}$、s_t) 满足条件独立同分布（因而可交换）性质得出的，而 (19) 则是由于 O_{t+1} 在输入的任意置换下保持不变。另一方面，若 s'_t、u'_1、$\cdots$、$u'_{N/m}$ 是 $s_t \sim \mathcal{D}_t$的独立副本，由独立性可知

$$\mathop{\mathbb{E}}_{s_t, s'_t\sim\mathcal{D}_t, R^{(t+1)}}[\ell(h_{t+1}, s'_t) \mid E] = \mathop{\mathbb{E}}_{s_t, s'_t\sim\mathcal{D}_t, R^{(t+1)}}[\ell(h_{t+1}, u'_j) \mid E], \quad \forall j \in [N/m].$$

因此，采用简写 $u_0 = s_t, u_0' = s_t'$，我们有

$$\begin{aligned}
&\mathop{\mathbb{E}}_{s_t, s_t' \sim \mathcal{D}_t, R^{(t+1)}} [\ell(h_{t+1}, s_t') - \ell(h_{t+1}, s_t) \mid E] \\
&= \frac{1}{N/m+1} \mathop{\mathbb{E}}_{s_t, s_t' \sim \mathcal{D}_t, R^{(t+1)}} \left[\sum_{j=0}^{N/m} (\ell(h_{t+1}, u_j') - \ell(h_{t+1}, u_j)) \;\middle|\; E \right] \\
&\le \frac{1}{N/m+1} \mathop{\mathbb{E}}_{u_0, \cdots, u_{N/m}, u_0', \cdots, u_{N/m}' \sim \mathcal{D}_t} \left[\sup_{h \in \mathcal{F}_\alpha} \sum_{j=0}^{N/m} (\ell(h, u_j') - \ell(h, u_j)) \right] \\
&\le \frac{2\alpha}{N/m+1} \mathop{\mathbb{E}}_{u_0, \cdots, u_{N/m} \sim \mathcal{D}_t} \mathop{\mathbb{E}}_{\epsilon_1 \ldots \epsilon_{N/m}} \left[\sup_{h \in \mathcal{F}_\alpha} \sum_{j=0}^{N/m} \epsilon_j h(u_j) \right] \\
&\le \frac{1}{\alpha} \mathrm{Rad}\left(\mathcal{F}_\alpha, N/m\right).
\end{aligned}$$

最后一个不等式利用了算法始终输出属于 $\mathcal{F}_\alpha$ 的函数这一事实。此外，我们还使用了 Ledoux-Talagrand 收缩不等式。

定理 H.2（Ledoux-Talagrand 收缩）。设 $g : \mathbb{R} \to \mathbb{R}$ 是一个 L-Lipschitz 函数。对于一个函数类 $\mathcal{F}$，记 $g \circ \mathcal{F}$ 为 $\mathcal{F}$ 中的函数与 g 的复合函数。则对于所有 n，

$$\mathrm{Rad}\left(g \circ \mathcal{F}, n\right) \le L \cdot \mathrm{Rad}\left(\mathcal{F}, n\right).$$

最后一个不等式来自以下事实：对数损失在截断至水平 α 时，其导数有界于 $1/\alpha$。注意，并集界给出

$$\Pr[E^c] \le \sum_{j=1}^{N/m} \Pr[E_j^c] \le \frac{N\left(1-\sigma\right)^m}{m}.$$

因此，全期望定律给出

$$\begin{aligned}
&\mathop{\mathbb{E}}_{s_t, s_t' \sim \mathcal{D}_t, R^{(t+1)}} [\ell(h_{t+1}, s_t') - \ell(h_{t+1}, s_t)] \\
&\le \mathop{\mathbb{E}}_{s_t, s_t' \sim \mathcal{D}_t, R^{(t+1)}} [\ell(h_{t+1}, s_t') - \ell(h_{t+1}, s_t) \mid E] + \Pr[E^c] \log(1/\alpha) \\
&\le \frac{1}{\alpha} \mathrm{Rad}\left(\mathcal{F}_\alpha, N/m\right) + \frac{N\left(1-\sigma\right)^m \log\left(1/\alpha\right)}{m}.
\end{aligned}$$

最后一个等式来自以下事实：算法的输出的损失始终不超过 $\log\left(1/\alpha\right)$。

对 $N \sim \mathrm{Poi}(n)$ 取期望即可得到所需结果；在上述不等式中使用 $\Pr[N > n/2] \ge 1 - e^{-n/8}$ 即完成证明。

扰动项的界

引理 I.1（扰动）。

$$\mathbb{E}\left[\sum_{i=1}^{N} L(\hat{h}, \tilde{s}_t) - L(h^*, \tilde{s}_t)\right] \le n \log \alpha$$

ˆ证明。 由 算法 1 中的截断步骤可知， $L(h, \tilde{s}_t) \leq \log(\alpha)$。取期望即可得到所需的界。 □