arXiv:2304.11735v1 [econ.EM] 23 Apr 2023

偏置样本选择下的策略学习

LihuaLei * Roshni Sahoo Stefan Wager
lihualei@stanford.edu rsahoo@stanford.edu swager@stanford.edu

斯坦福大学

摘要

实践者通常利用随机对照试验的数据，学习可应用于目标总体的治疗分配策略。一个反复出现的问题是，即使随机试验设计和执行得当（即内部有效性成立），研究参与者也可能并非目标总体的随机样本（即外部有效性不成立），这可能导致策略在目标总体中的表现不尽理想。我们考虑这样一种模型：可观测属性对样本选择概率的影响可以是任意的，而不可观测属性的影响则有一个常数上界；我们的目标是学习策略，使其在任何此类抽样偏差下都能获得尽可能好的性能保证。具体而言，我们推导了存在抽样偏差时最坏情况福利的部分识别结果，并表明，最优的最大-最小策略、最大-最小增益策略和极小化最大后悔策略，取决于条件平均处理效应（CATE）以及给定协变量时潜在结果的条件风险价值（CVaR）。为避免插件估计量造成的有限样本效率损失，我们进一步提出一种端到端方法，用于学习最优的最大-最小策略和最大-最小增益策略，而无需分别估计干扰参数。

1 引言

实践者常使用随机对照试验（RCT）的数据，学习一种可部署于目标总体的治疗分配策略。形式上，在研究人群中，每个个体 i 都具有协变量 $X_i \in \mathcal{X}$ 和潜在结果 $Y_i(0)$ $Y_i(1) \in \mathcal{Y}$，分别对应对照组和处理，其分布遵循研究潜在结果分布 P。试验实施后，实践者可以获得样本 $(X_i,\ Y_i,\ W_i) \sim P_{\text{obs}}$，其中 P_{obs} 是与 P 一致的观测数据分布，前提是随机对照试验在内部有效；这意味着处理 $W_i \in \{0,1\}$ 是随机分配的，且观测结果 Y_i 等于 $Y_i(W_i)$，即按指定处理对应的潜在结果。

对随机对照试验的一种常见批评是其缺乏外部有效性。外部有效性体现了从随机对照试验研究人群得出的结论能够推广到更广泛总体或其他目标人群的程度。随机对照试验可能因多种原因缺乏外部有效性。例如，Bell 等人 [2016] 表明，在评估 Reading First 教育项目时，非随机的地点选择会比具有代表性的地点选择得出更低的影响估计值。此外，Wang 等人 [2018] 指出，用于衡量抗抑郁药效果的随机试验依赖志愿者自愿报名参加，因此这些研究的发现可能不适用于非志愿者。此外，从数据收集到实际部署之间的时间间隔，可能会因时间偏移而导致目标总体与研究人群不同。

标准策略学习方法忽略了随机对照试验可能缺乏外部有效性这一问题。这些方法使用来自 P_{obs} 的数据，来学习能够使研究人群的平均结果最大化的策略 [Atheyand Wager，2021年，Bhattacharya and Dupas，2012年，Kitagawa and Tetenov，2018年，Manski，2004]：

$$\pi^* = \underset{\pi \in \Pi}{\operatorname{argmax}}\, \mathbb{E}_P\left[Y(\pi(X))\right], \tag{1}$$

其中，$\pi : \mathcal{X} \mapsto \{0,1\}$ 表示治疗分配策略，Π 表示策略类。这些方法的隐含假设是，目标总体与研究人群相同。所学得的

*作者按字母顺序排列。

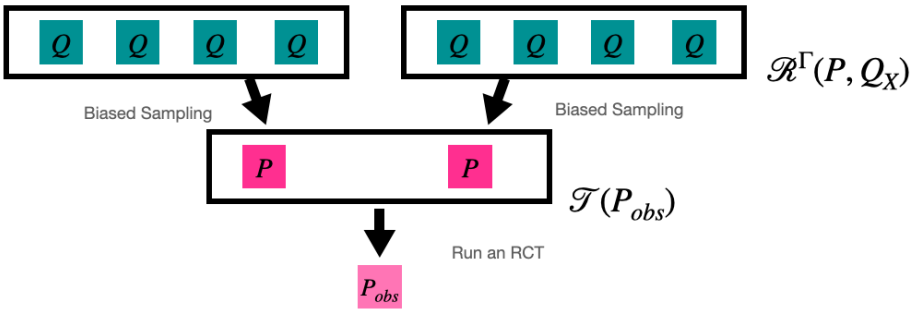

图1：我们将合理目标分布的集合可视化为 $\mathcal{S}^{\Gamma}(P_{\text{obs}}, Q_X)$。集合 $\mathcal{T}(P_{\text{obs}})$ 是与内部有效的随机对照试验下的 P_{obs} 相一致的研究潜在结果分布 P 的集合。集合 $\mathcal{R}^{\Gamma}(P, Q_X)$ 是在 Γ-有偏抽样下能够生成 P 且具有协变量分布 Q_X 的目标潜在结果分布集合。

策略（1）对不同于 Q 的 P目标潜在结果分布不提供性能保证。

在这项工作中，我们旨在识别并学习能够应对某些外部有效性失效情形的稳健策略，具体而言，是研究参与者选择中的抽样偏差。与此前一样，我们假设实践者能够获取来自 P_{obs} 的数据，但我们的目标是学习一种策略 π，使其在未知目标分布 Q 上具有性能保证，而研究潜在结果分布 P 可能因抽样偏差而相对于 Q 存在偏差。遵循 Manski [2003]，的做法，我们使用二元选择指标对样本选择建模。我们通过 Γ-有偏抽样模型［来量化样本选择偏差的程度（Aronow and Lee，2013年；Nie等人，2021年；Sahoo等人， 2022]）。该模型允许每个个体的样本选择概率任意依赖于协变量，但其对不可观测变量的依赖程度仅受有限限制。参数 $\Gamma \geq 1$ 刻画了由不可观测变量导致的抽样偏差程度， Γ 的值越大，允许的偏差程度就越大。请注意，当 $\Gamma = 1$ 时，该模型对应于“无混杂样本选择”，这一情形已在可推广性相关文献中得到研究［（例如，Stuart等人，2011年；Tipton， 2013, 2014]）。

定义1。设 $\Gamma \geq 1$。对于定义在 $(X, Y(0)\ Y(1))$ 上的任意一对分布 P 和 Q，如果存在定义在 $(X, Y(0)\ Y(1)\ S)$ 上的分布 Q，使得 Q 可以在 Γ-有偏抽样下生成 P，其中 $S \in \{0, 1\}$ 是满足以下性质的“选择指标”：Q 的 $(X, Y(0)\ Y(1))$-边际分布等于 Q，给定 $S = 1$ 时 Q 的 $(X, Y(0)\ Y(1))$-边际分布等于 P，并且

$$\frac{\mathbb{P}_{\tilde{Q}}\left[S = 1 \mid X = x, Y(0) = y_0, Y(1) = y_1\right]}{\mathbb{P}_{\tilde{Q}}\left[S = 1 \mid X = x\right]} \in [\Gamma^{-1}, \Gamma] \quad \forall x \in \mathcal{X}, y_0, y_1 \in \mathcal{Y}. \tag{2}$$

在学习能够抵御有偏样本选择的策略时，存在两个关键挑战。第一个挑战是标准策略学习设置中出现的“缺失数据问题”。在随机对照试验中，研究单位会被分配到对照组或处理组，因此对于任何研究单位而言，$Y_i(0)\ Y_i(1)$ 不会被同时观测到。尽管我们可以从 P_{obs} 中识别研究潜在结果分布的边际分布 $P_{X,Y(0)}\ P_{X,Y(1)}$，但无法从 P_{obs} 中识别联合分布 P。因此，真实的研究潜在结果分布是未知的，并且存在许多潜在结果分布 P，它们在内部有效的随机对照试验下都会产生 P_{obs}。令 $\mathcal{T}(P_{\text{obs}})$ 表示在内部有效的随机对照试验下产生 P_{obs} 的研究潜在结果分布集合，即

$$\mathcal{T}(P_{\text{obs}}) = \{P : P_{X,Y(0)} = P_{\text{obs},X,Y|W=0}, P_{X,Y(1)} = P_{\text{obs},X,Y|W=1}\}.$$

第二个挑战源于有偏抽样。我们注意到，真目标潜在结果分布也未知；给定一个特定的研究潜在结果分布 P，在针对 $\Gamma > 1$ 的 Γ-有偏抽样下，可以生成 P 的目标分布 Q 有许多种可能。令 $\mathcal{R}^{\Gamma}(P, Q_X)$ 表示这样一组潜在结果分布 Q：它们具有协变量分布 Q_X，并且可以通过 Γ-有偏

抽样生成。因此，我们可以将 $\mathcal{S}^{\Gamma}(P_{\text{obs}}, Q_X)$ 定义为稳健性集合，即与 P_{obs} 相一致、在 Γ-有偏抽样和内部有效的随机对照试验条件下所有合理的目标潜在结果分布集合：

$$\mathcal{S}^{\Gamma}(P_{\text{obs}}, Q_X) = \{Q \mid Q \in \mathcal{R}^{\Gamma}(P, Q_X) \text{ where } P \in \mathcal{T}(P_{\text{obs}})\}.$$

该集合在图1中进行了可视化。

为了获得关于未知目标总体的性能保证，我们采用分布鲁棒优化（DRO）框架 [Ben-Tal等人， 2013]以学习一种对稳健性集合中所有合理目标分布均稳健的策略。与Manski [2011], 一样，我们考虑多种不同方法来衡量策略的良好表现，并将部分识别下的模糊性纳入考量。最大-最小目标函数考察部分识别集合上的最坏情况福利，而最大最小策略则是

$$\pi^*_{\Gamma,\text{maxmin}} \in \operatorname*{argmax}_{\pi \in \Pi} \inf_{Q \in \mathcal{S}^{\Gamma}(P_{\text{obs}}, Q_X)} \mathbb{E}_Q \left[Y(\pi(X)) \right]. \tag{3}$$

优化最大-最小保证在概念上很简单；然而，在许多情况下，最大-最小目标函数被认为过于悲观 [Manski, 2011年， Savage， 1951]。为避免这种悲观，我们还考虑另外两个目标函数。当存在自然的基准策略 π_0 时（例如，现状明确），我们可以优化相对于基准策略的最大-最小改进；最大-最小增益策略则是

$$\pi^*_{\Gamma,\text{gain}} \in \operatorname*{argmax}_{\pi \in \Pi} \inf_{Q \in \mathcal{S}^{\Gamma}(P_{\text{obs}}, Q_X)} \mathbb{E}_Q \left[Y(\pi(X)) \right] - \mathbb{E}_Q \left[Y(\pi_0(X)) \right]. \tag{4}$$

正如Kallus和Zhou [2021], 所讨论的，当现状策略已被视为一个相当不错的决策规则时，最大-最小增益方法可能是理想的，因为它能在现状基础上带来“安全”的改进；也就是说，除非数据让我们能够毫无疑问地偏好另一种行动，否则我们将维持现状。最后，我们还考虑最小最大遗憾准则；极小极大遗憾策略则是

$$\pi^*_{\Gamma,\text{regret}} \in \operatorname*{argmin}_{\pi \in \Pi} \sup_{Q \in \mathcal{S}^{\Gamma}(P_{\text{obs}}, Q_X)} R_Q(\pi(X)), \tag{5}$$

其中，分布 Q 下的遗憾是指某一策略的均值结果与分布 Q下表现最佳策略的均值结果之差：

$$R_Q(\pi(X)) = \sup_{\pi' \in \Pi} \mathbb{E}_Q \left[Y(\pi'(X)) \right] - \mathbb{E}_Q \left[Y(\pi(X)) \right]. \tag{6}$$

多位作者认为，极小极大遗憾是指导模糊性情境下行动的一项相关准则 [Manski， 2011年、Savage， 1951]; 这一准则不会受到最大-最小准则悲观失效模式的影响，而且在缺乏有力基准策略时也可使用。

我们的工作有两项主要贡献。第一项贡献是，我们证明了在3、4和5中定义的最优策略是可识别的 P_{obs} ，并给出了这些策略的闭式表达式，条件是策略类 Π 由确定性、无约束的二值函数构成。第二项贡献是，我们提出了端到端流程，利用来自 P_{obs} 的数据学习最优最大-最小策略和最大-最小增益策略，而无需单独估计干扰参数。

在第 2 节中，我们证明，当 Π 由确定性、无约束、二元值函数构成时，仅利用来自 P_{obs}的数据，就可以识别最大-最小、最大-最小增益和极小化最大后悔目标下的最优策略，并给出这些策略的闭式表达式。在第 3 节中，我们给出一个等价性结果，这对于证明我们的学习程序能够得出最优策略至关重要。该结果表明，若观测数据价值函数在P_{obs} 下的均值等于潜在结果价值函数在任意潜在结果分布$P \in \mathcal{T}(P_{\text{obs}})$下的均值，那么在 $\mathcal{S}^{\Gamma}(P_{\text{obs}}, Q_X)$上对潜在结果价值函数 $v^*(\pi(X); X, Y(0)\ Y(1))$进行稳健优化，等价于在观测数据分布的稳健性集合上对观测数据价值函数 $v(\pi(X); X, Y, W)$进行稳健优化。在第 4 节中，我们证明，可以通过扩展 Rockafellar-Uryasev（RU）回归 [Sahoo等人, 2022]，求解在观测数据分布的稳健性集合上对观测数据价值函数进行的稳健优化。我们将这一结果与上一节的等价性结果结合，给出一种学习最优最大-最小策略和最大-最小增益策略的方法。在第 5 节中，我们通过模拟实验以及使用 Gerber等人 [2008]的投票数据集开展的半合成实验，对学习程序进行实证评估。

1.1 相关工作

最优治疗分配的研究在经济学领域 [Athey and Wager, 2021年, Bhattacharya and Dupas, 2012年, Kitagawa and Tetenov, 2018年, Manski, 2004], 统计学领域 [Qian and Murphy, 2011年, Zhao等人, 2012], 以及计算机科学领域 [Swaminathan and Joachims, 2015]均受到关注。该领域的标准设定并未考虑研究人群可能不同于所关注的目标总体这一问题。我们设定的独特之处在于，允许研究潜在结果分布P与目标潜在结果分布 Q 不同，即存在 Γ-有偏抽样（定义1）。Sahoo等人 [2022]针对监督学习设定提出了 Γ-有偏抽样模型，并关注学习稳健回归模型时的统计和算法问题。在本研究中，我们基于潜在结果框架考察最优治疗选择，并识别各种稳健目标函数下的最优策略，包括极小化最大后悔目标。

我们关于偏倚样本选择下策略学习的贡献，属于部分识别下策略学习的相关文献 [Adjaho 和 Christensen，2022；Christensen 等，2022；Ben-Michael 等，2021；Hansen 和 Sargent，2001；Hatt 等，2022；Higbee，2022；Kallus 等，2022；Kallus 和 Zhou，2021；Mu 等，2022；Manski，2000，2007；Si 等，2020；Watson 和 Holmes， 2016]。这一领域的主要挑战在于，最优策略无法识别，因为它取决于一些未知量，例如结果缺失数据 [Ben-Michael 等，2021；Higbee，2022；Manski，2007], 未知的倾向权重 [Kallus 和 Zhou， 2021], 或未知的目标分布 [Adjaho 和 Christensen，2022；Hatt 等，2022；Si 等，2020；Mu 等， 2022]。这些研究采用稳健优化方法来解决这一问题：他们为未知量定义一个合理取值集合，并找出一个在未知量取最坏情形值时仍表现良好的策略。我们的研究与这些先前工作一脉相承，因为我们定义了一组与观测数据及假定的抽样偏差模型相一致的合理目标潜在结果分布，并识别出能够抵御最坏情形目标分布的策略。

在这方面的文献中，我们的工作与Adjaho 和 Christensen [2022]以及Si 等人 [2020]的研究最为相关。与我们的工作类似，他们采用稳健优化方法，制定能够为可能不同于研究人群的目标人群提供福利保证的策略。不过，本文考虑的稳健性集合与其他研究中的形式不同（例如，围绕研究分布构建的Wasserstein球 [Adjaho 和 Christensen，2022] 以及 KL散度球 [Si 等人， 2020] ）。不同于Adjaho 和 Christensen[2022] 以及Si 等人 [2020] 仅讨论最大最小策略，我们能够推导出最大最小增益策略和极小极大遗憾策略的闭式表达式；后两者在实践中可能更具相关性 [Manski， 2000]。

我们的工作与Adjaho 和 Christensen [2022] 以及Si 等人 [2020] 的研究的另一处区别在于，他们都通过对 X、$Y(0)$ $Y(1)$上的联合分布施加约束来定义稳健性集合，而我们考虑的稳健性集合则对条件分布 $Y(0)$ $Y(1)$ $\mid$ X施加约束。对结果和协变量的联合分布施加约束，在许多方面都过于保守。首先，我们通常会在测试时观察到协变量分布的偏移（Q_X 在测试时可被观察到），因此往往没有必要防范任意的协变量偏移。其次，当策略类 Π 由不受约束的二值函数构成时，对于每一个 $x \in \mathcal{X}$，最优策略都能求解最优治疗分配问题，因此协变量偏移不会影响最优策略。第三，实证评估表明，与分别处理协变量偏移和条件偏移相比，对结果和协变量的联合分布施加约束可能会导致过于保守的结果 [Mu 等人，2022]。

2 识别

首先，我们定义策略学习问题的记号和设定，并定义我们希望对其具有稳健性的目标分布集合。其次，我们定义本文将考察的各种稳健性标准。最后，我们确定每种稳健性标准下的最优策略，并讨论它们之间的联系。

2.1 稳健性集合

我们考虑以下策略学习问题。在研究人群中，每个个体 i 都与协变量 $X_i \in \mathcal{X}$ 和潜在结果 $Y_i(0)$ $Y_i(1)$ $\in \mathcal{Y}$ 相关联，这些变量均独立地取自一个（未知的）研究分布 P。在全文中，我们假设

$$P_{(Y(0)\ Y(\ 1))|X=x} \text{ 对每个 } x \in \mathcal{X}, \text{ 相对于勒贝格测度绝对连续} \quad .$$

为简洁起见，之后不再提及这一假设。我们可以获得来自随机对照试验的数据（X_i， Y_i， W_i） $\sim P_{\text{obs}}$，其中 $X_i \in \mathcal{X}$ 为协变量， $W_i \in \{0,1\}$ 为处理分配， $Y_i \in \mathcal{Y}$ 为观测结果。我们假设，在处理概率为 e 的随机对照试验中， P_{obs} 与 P 一致。这意味着处理按照伯努利分布(e)随机分配，且观测结果满足稳定单位处理值分配假设（SUTVA），即 $Y_i = Y_i(W_i)$。 换言之，如果 P_{obs} 是通过内部有效的随机对照试验生成的，那么

$$P_{\text{obs},X,Y|W=w} = P_{X,Y(w)} \quad \forall w \in \{0,1\}, \tag{7}$$

$$P_{\text{obs},W|X=x} = \text{Bernoulli}(e) \quad \forall x \in \mathcal{X}. \tag{8}$$

令 Q 为目标潜在结果分布。我们旨在学习一个策略 $\pi:\ \mathcal{X} \to \{0,1\}$ ，使得当从 Q中抽取(X， $Y(0)$ $Y(1)$)时，平均结果 $\mathbb{E}_Q\left[Y(\pi(X))\right]$ 较大。在本文中，我们假设 $\pi \in \Pi$ ，其中

$$\Pi = \{\pi : \mathcal{X} \mapsto \{0,1\} : \pi \text{ is measurable and deterministic}\}.$$

这一设定的两项主要挑战包括：无法从观测数据中识别出真实的研究潜在结果分布 P，而且目标潜在结果分布 Q 与 P并不相同。具体而言，我们假设 Q 未知，且在 P 下生成 Γ-有偏抽样，其含义见定义 1。

对于任意边际协变量分布 Q_X，我们将 $\mathcal{S}^\Gamma(P_{\text{obs}}, Q_X)$ 定义为所有满足以下条件的目标潜在结果分布的集合：能够在 Γ-有偏抽样和处理概率为 e 的内部有效的随机对照试验下生成 P_{obs} ，具有协变量分布 Q_X，且相对于勒贝格测度绝对连续。令 $\mathcal{T}(P_{\text{obs}})$ 表示能够在处理概率为 e 的内部有效的随机对照试验下生成 P_{obs} ，且相对于勒贝格测度绝对连续的潜在结果分布集合。本质上，这是满足式（7）的 $P_{X、Y}(0)$ 和 $P_{X、Y}(1)$ 的耦合集合。此外，我们可以将 $\mathcal{R}^\Gamma(P, Q_X)$ 定义为能够在 Γ-有偏抽样下生成 P，且具有协变量分布 Q_X 的潜在结果分布集合。虽然可以对 P 和 Q 的可能取值施加额外限制，但我们关注最一般的设定，其中

$$\mathcal{S}^\Gamma(P_{\text{obs}}, Q_X) = \{Q \mid Q \in \mathcal{R}^\Gamma(P, Q_X) \text{ where } P \in \mathcal{T}(P_{\text{obs}})\}. \tag{9}$$

我们的目标是学习对于 $\mathcal{S}^\Gamma(P_{\text{obs}}$， $Q_X)$ 中所有目标潜在结果分布都具有鲁棒性的策略。在下一小节中，我们定义了在 $\mathcal{S}^\Gamma(P_{\text{obs}}$， $Q_X)$ 上进行优化时能够产生不同性能保证的目标函数。

2.2 目标函数

我们考察由三种不同目标函数给出的鲁棒策略：最大-最小、最大-最小增益和极小极大遗憾。我们参考 Manski [2011] 对最大-最小目标函数和极小化最大后悔目标进行了详细讨论。

学习鲁棒策略时，一个自然的第一步是考虑最大最小策略：

$$\pi^*_{\Gamma,\text{maxmin}} \in \operatorname*{argmax}_{\pi \in \Pi} \inf_{Q \in \mathcal{S}^\Gamma(P_{\text{obs}}, Q_X)} \mathbb{E}_Q\left[Y(\pi(X))\right]. \tag{10}$$

最大最小策略在所有自然状态下（在所有分布 $Q \in \mathcal{S}^\Gamma$（P_{obs}， Q_X）的平均结果上给出最大下界。最大-最小目标是部分识别下策略学习文献中研究最为广泛的目标函数 [Adjaho 和 Christensen， 2022， Mu 等， 2022， Si 等， 2020]。然而，已知它往往过于悲观 [Savage， 1951]因为可能存在某些分布 Q ，使得所有策略的表现都很差，而最大-最小目标函数将聚焦于这些分布

情形。在某些情况下，决策者可能希望防范这些最坏情形，因此，最大-最小目标函数可能是合理的选择 [Manski, 2011]。

作为最大-最小目标的替代方案，我们还考虑使相对于基线 π_0 的最坏情形增益最大化的策略，其定义如下：

$$\pi^*_{\Gamma,\text{gain}} \in \operatorname*{argmax}_{\pi\in\Pi} \inf_{Q\in\mathcal{S}^\Gamma(P_{\text{obs}},Q_X)} \mathbb{E}_Q\left[Y(\pi(X))\right] - \mathbb{E}_Q\left[Y(\pi_0(X))\right]. \tag{11}$$

如果我们的目标是相对于现状策略提高稳健性，那么最大-最小增益策略是一个自然的选择。Ben-Michael等人近期研究了这类目标 Ben-Michael 等人 [2021], Kallus和Zhou [2021]。

由 Savage [1951], 提出的极小化最大后悔目标与最大-最小增益目标密切相关。在极小化最大后悔问题中，我们的目标实质上是最大化相对于最优基准策略的最坏情形增益；对于每一种 Q选择，该基准策略都能取得最大的平均结果。

极小极大遗憾策略定义为

$$\pi^*_{\Gamma,\text{regret}} \in \operatorname*{argmin}_{\pi\in\Pi} \sup_{Q\in\mathcal{S}^\Gamma(P_{\text{obs}},Q_X)} R_Q(\pi(X)), \tag{12}$$

其中，策略 $\pi(X)$ 在分布 Q 下的遗憾定义为

$$R_Q(\pi(X)) = \sup_{\pi'\in\Pi} \mathbb{E}_Q\left[Y(\pi'(X))\right] - \mathbb{E}_Q\left[Y(\pi(X))\right].$$

当考虑包含非确定性策略的策略类时，我们通常会发现极小极大遗憾策略往往是非确定性的 [Manski, 2011]。我们强调，在本文中，我们将极小化最大后悔目标定义为相对于 Π而言；该策略类由不受约束的确定性二值函数构成，因此我们的目标是在确定性函数类中寻找极小极大遗憾策略。

极小化最大后悔规则给出了因未知 Q 而导致的平均结果损失的最小上界，并且被认为比最大-最小规则更不悲观 [Manski, 2011年, Savage, 1951]。然而，由于极小极大遗憾策略的结构复杂，识别和学习这一策略往往很困难。我们的工作的一项重要贡献是，针对确定性、无约束、二元取值策略这一类策略，推导出极小极大遗憾策略的闭式表达式。

2.3 结果

我们的识别结果依赖于潜在结果的条件平均处理效应（CATE）函数和条件风险价值 [Rockafellar 和 Uryasev, 2000] 。我们在下文定义 CATE

$$\tau(x) = \mathbb{E}_P\left[Y(1) - Y(0) \mid X = x\right]. \tag{13}$$

我们考虑特定分位数处的条件风险价值

$$\zeta(\Gamma) = \frac{1}{\Gamma+1}. \tag{14}$$

回顾一下，对于 具有分位数函数（累积分布函数的反函数） q_Z 和 $\zeta \in (0,\ 1)$ 的连续随机变量 Z

$$\text{CVaR}_\zeta(Z) = \mathbb{E}\left[Z \mid Z \geq q_\zeta(Z)\right].$$

此外，我们回顾一下，在不存在抽样偏差的情况下，应部署的最优策略是

$$\pi_{\text{non-robust}}(x) = \mathbb{I}(\tau(x) \geq 0), \tag{15}$$

对在 P [Kitagawa andTetenov, 2018] 下具有非负条件平均处理效应的研究单位进行处理。请注意，该策略在 P_{obs} 下是可识别的，因为条件平均处理效应在 P_{obs}. 下是可识别的

定理 1。定义

$$H_\Gamma(x) = (1-\Gamma^{-1}) \cdot \left(\text{CVaR}_{\zeta(\Gamma)}(Y(1) \mid x) - \text{CVaR}_{\zeta(\Gamma)}(Y(0) \mid x)\right). \tag{16}$$

该策略

$$\pi^*_{\Gamma,\text{maxmin}}(x) = \mathbb{I}(\tau(x) \geq H_\Gamma(x)) \tag{17}$$

满足 (10) 对于任意 Q_X 满足 $Q_X \ll P_{\text{obs},\ X}$ 且 上确界 $\sup_{x\in\mathcal{X}} dP_{\text{obs},\ X}(x)\, dQ_X(x) < \infty$. 证明见附录 D。 3.

我们回顾一下， $\Gamma = 1$ 对应于这样一种情况：不可观测变量不会导致样本选择概率发生变化。我们发现，（17）在 $\pi_{\text{non-robust}}$ 时可化简为 $\Gamma = 1$. 有趣的是，我们注意到，当 $\Gamma > 1$时， $H_\Gamma(x)$ 不一定为正，这意味着（17）不一定会产生比 $\pi_{\text{non-robust}}$. 更高的处理阈值。因此，根据 $Y(1)$ $Y(0)$ 的尾部行为，最大-最小规则所处理的研究单位数量可能多于 $\pi_{\text{non-robust}}$ ，后者假设不可观测变量不会造成抽样偏差。

我们指出，对最大最小策略的另一种解读是：它比较稳健性集合上$\mathbb{E}_Q[Y(1) \mid X = x]$ 和 $\mathbb{E}_Q[Y(0) \mid X = x]$ 的下界，并选择能带来较高下界的处理。

引理 2. 最大最小策略 $\pi^*_{\Gamma,\ \text{maxmin}}$ 可写为

$$\mathbb{I}\left(\inf_{Q\in\mathcal{S}^\Gamma(P_{\text{obs}},Q_X)} \mathbb{E}_Q[Y(1) \mid X = x] \geq \inf_{Q\in\mathcal{S}^\Gamma(P_{\text{obs}},Q_X)} \mathbb{E}_Q[Y(0) \mid X = x]\right). \tag{18}$$

附录*D.4*中的证明。

我们将最大最小策略与最大化相对于基准策略的最坏情形增益的策略，以及极小极大遗憾策略进行比较和对照。

定理 3. 定义

$$H^+_\Gamma(x) = (1-\Gamma^{-1})\cdot\left(\text{CVaR}_{\zeta(\Gamma)}(Y(1) \mid x) + \text{CVaR}_{\zeta(\Gamma)}(-Y(0) \mid x)\right), \tag{19}$$
$$H^-_\Gamma(x) = (1-\Gamma^{-1})\cdot\left(-\text{CVaR}_{\zeta(\Gamma)}(-Y(1) \mid x) - \text{CVaR}_{\zeta(\Gamma)}(Y(0) \mid x)\right). \tag{20}$$

该策略

$$\pi^*_{\Gamma,\text{gain}}(x) = \mathbb{I}(\pi_0(x) = 0)\cdot\mathbb{I}(\tau(x) \geq H^+_\Gamma(x)) + \mathbb{I}(\pi_0(x) = 1)\cdot\mathbb{I}(\tau(x) \geq H^-_\Gamma(x)) \tag{21}$$

求解 (11) 对于任意 Q_X 使得 $Q_X \ll P_{\text{obs},X}$ 且 上确界$_{x\in\mathcal{X}}\ dP_{\text{obs},X}(x)\,dQ_X(x) < \infty$. 证明见附录*D.5*。

请注意，当 $\Gamma = 1$ 时，最优最大-最小增益策略默认为 $\pi_{\text{non-robust}}$。当 $\Gamma > 1$ 时，式（21）中的最优最大-最小增益策略会根据基准策略建议处理还是对照组而采用不同的治疗阈值。直观而言，当基准策略建议对照组时（$\pi_0(x) = 0$），我们预期式（21）的治疗阈值会高于基准策略建议处理时（$\pi_0(x) = 1$）的阈值。下面的引理证实了这一直觉，证明了 $H^+_\Gamma(x) \geq H^-_\Gamma(x)$。我们还注意到，最大-最小策略（17）的治疗阈值介于 $H^-_\Gamma(x)$ 和 $H^+_\Gamma(x)$ 之间。

引理 4. 对于 $x \in \text{supp}(P_{\text{obs},X})$,

$$H^-_\Gamma(x) \leq H_\Gamma(x) \leq H^+_\Gamma(x). \tag{22}$$

证明见附录*D.6*。

上述引理的另一个推论是，如果基准策略由 $\pi_0(x) = \mathbb{I}(\tau(x) \geq b(x))$ 给出，其中对于所有 $x \in \mathcal{X}$， $H^-_\Gamma(x) \leq b(x) \leq H^+_\Gamma(x)$，那么基准策略 π_0 最大化最大-最小增益目标，即，

$$\pi^*_{\Gamma,\text{gain}} = \pi_0.$$

最后，我们考虑极小化最大后悔规则。

定理 5. 定义 $H^+_\Gamma(\cdot)$ $H^-_\Gamma(\cdot)$ 如 (19)、(20)。策略

$$\pi^*_{\Gamma,\text{regret}}(x) = \mathbb{I}\left(\tau(x) \geq \frac{H^+_\Gamma(x) + H^-_\Gamma(x)}{2}\right) \tag{23}$$

求解 (12) 对于任意 $Q_X \ll P_{\text{obs},X}$ 以及 上确界$_{x\in\mathcal{X}}\ \frac{dP_{\text{obs},X}(x)}{}\,dQ_X(x) < \infty$. 证明见附录 *D.7*。

我们给出了极小极大遗憾策略的另一种刻画，这种刻画与标准策略学习问题相关。在不存在抽样偏差的情况下，求解以下优化问题

$$\sup_{\pi\in\Pi} \mathbb{E}_P[(2\pi(X) - 1)\cdot(Y(1) - Y(0))] \tag{24}$$

可得到 $\pi_{\text{non-robust}}$，当不存在抽样偏差 [Zhao et al., 2012] 时，它就是最优策略。在下面的定理中，我们将看到，在稳健性集合上最大化该目标函数的最大化者给出了极小极大遗憾策略。我们发现，在稳健性集合 $\mathcal{S}^{\Gamma}(P_{\text{obs}}, Q_X)$ 上对目标函数 $\mathbb{E}_Q\left[v^*(\pi(X); X, Y(0)\ Y(1))\right] = \mathbb{E}_Q\left[(2\pi(X) - 1) \cdot (Y(1) - Y(0))\right]$ 进行稳健优化，可得到极小极大遗憾策略。

定理 6. 设

$$\tilde{v}_{\text{regret}}(z; x, y_0, y_1) = (2z - 1) \cdot (y_1 - y_0).$$

求解该问题的策略 π

$$\sup_{\pi \in \Pi} \inf_{Q \in \mathcal{S}^{\Gamma}(P_{\text{obs}}, Q_X)} \mathbb{E}_Q\left[\tilde{v}_{\text{regret}}(\pi(X); X, Y(0), Y(1))\right] \tag{25}$$

等 于极小化最大后悔目标下的最优策略 (23)。 证明见附录 *D*。 *8.*

注意，(17)、(21) 和 (23) 中的策略可以通过先估计干扰参数$\tau(\cdot)$ $H_{\Gamma}(\cdot)$ $H_{\Gamma}^{+}(\cdot)$ $H_{\Gamma}^{-}(\cdot)$，然后构造定义好的策略来估计。虽然这是一种学习策略的潜在方法，但实施这种方法可能十分繁琐，且实际表现不佳。在接下来的章节中，我们将探讨一种学习最大-最小和最大-最小增益策略的替代方法。

3 等价性

在本节中，我们将证明一个等价性结果，它将潜在结果分布上的稳健优化与观测数据分布上的稳健优化联系起来。当我们旨在仅使用来自 P_{obs} 的数据学习鲁棒策略，而不掌握真实研究总体的相关知识时，这种等价性很有用P。

潜在结果分布上的稳健优化通常可以表述为以下问题

$$\sup_{\pi \in \Pi} \inf_{Q \in \mathcal{S}^{\Gamma}(P_{\text{obs}}, Q_X)} \mathbb{E}_Q\left[v^*(\pi(X); X, Y(0), Y(1))\right]. \tag{26}$$

请注意， v^* 是潜在结果价值函数，它根据 X、$Y(0)$ $Y(1)$。来定义。现在，我们考虑观测数据分布上的稳健优化。令 $\mathcal{S}^{\Gamma}_{\text{obs}}(P_{\text{obs}}$、 $Q_X)$ 为由观测数据分布构成的稳健性集合。若满足 Q_{obs}、$X = Q_X$，则分布 $Q_{\text{obs}} \in \mathcal{S}^{\Gamma}_{\text{obs}}(P_{\text{obs}}$、 $Q_X)$ 且

$$\Gamma^{-1} \leq \frac{dQ_{\text{obs},Y|W=w,X=x}(y)}{dP_{\text{obs},Y|W=w,X=x}(y)} \leq \Gamma \quad \forall w \in \{0,1\}, x \in \mathcal{X}, y \in \mathcal{Y}. \tag{27}$$

我们考虑

$$\sup_{\pi \in \Pi} \inf_{Q_{\text{obs}} \in \mathcal{S}^{\Gamma}_{\text{obs}}(P_{\text{obs}}, Q_X)} \mathbb{E}_{Q_{\text{obs}}}\left[v(\pi(X); X, Y, W)\right]. \tag{28}$$

请注意， v 是观测数据价值函数，它根据 X、Y、W来定义 .

为建立潜在结果分布上的稳健优化与观测数据分布上的稳健优化之间的联系，我们需要以下假设。

假设1。 我们假设 v 的定义满足

$$\mathbb{E}_{P_{\text{obs}}}\left[v(\pi(X); X, Y, W)\right] = \mathbb{E}_P\left[v^*(\pi(X); X, Y(0), Y(1))\right],$$

对于任意 $P \in \mathcal{T}(P_{\text{obs}})$ 以及任意的 π。

定理 7。 设 v 是满足假设 *1* 的价值函数，其中 v^*。那么，满足(28) 的策略 π 也满足(26)，对于任意 $Q_X \ll P_{obs}$、X 以及 $\sup_{x \in \mathcal{X}} dP_{\text{obs}}$、$X(x)\, dQ_X(x) < \infty$. 证明见附录~~*D.9*~~。

定理7是一个非平凡且有些出人意料的结果。虽然 $Q \in \mathcal{S}^{\Gamma}(P_{\text{obs}}$， $Q_X)$ 蕴含 $Q_{\text{obs}} \in \mathcal{S}^{\Gamma}_{\text{obs}}(P_{\text{obs}}$， $Q_X)$，但反过来通常并不成立，因为 Q_{obs} 并未包含关于潜在结果之间连接函数的信息（在给定 X的条件下）。尽管如此，我们证明，这个优化问题的最坏情形连接函数会通过一种类似于 Neyman-Pearson引理的论证，得到一个属于 $\mathcal{S}^{\Gamma}(P_{\text{obs}}$， $Q_X)$ 的联合分布。

我们指出，当对连续值函数 h进行优化时，定理7 同样成立。令 $\mathcal{H}=L^2(P_{\text{obs}},\ X,\ \mathcal{X})$。我们可以扩展假设1：假设 $v,\ v^*$ 的定义满足

$$\mathbb{E}_{P_{\text{obs}}}\left[v(h(X);X,Y,W)\right]=\mathbb{E}_{P}\left[v^*(h(X);X,Y(0),Y(1))\right],$$

对于 $P\in\mathcal{T}(P_{\text{obs}})$ 和任意的 h。在此假设下，满足以下条件的函数 h

$$\sup_{h\in\mathcal{H}}\inf_{Q_{\text{obs}}\in\mathcal{S}^{\Gamma}_{\text{obs}}(P_{\text{obs}},Q_X)}\mathbb{E}_{Q_{\text{obs}}}\left[v(h(X);X,Y,W)\right]. \tag{29}$$

也满足以下条件

$$\sup_{h\in\mathcal{H}}\inf_{Q\in\mathcal{S}^{\Gamma}(P_{\text{obs}},Q_X)}\mathbb{E}_{Q}\left[v^*(h(X);X,Y(0),Y(1))\right]. \tag{30}$$

4 通过 RU 回归的端到端策略学习

在本节中，我们首先证明（29）可以通过 RU 回归 [Sahoo等人， 2022]求解。接下来，我们定义平滑的观测数据价值函数，并将其与 RU 回归过程结合，用于学习最优的最大-最小策略和最大最小增益策略。我们还定义了一个用于学习最优极小极大遗憾策略的观测数据价值函数，但该函数是非凹的，因此会导致非凸 RU 回归问题。

首先，我们证明 RU 回归可用于求解（29）。

定理 8。 令 v（z； x,y,w）为一个价值函数。对于任意 $\Gamma>1$，我们定义如下增广损失函数

$$L^{\Gamma}_{\text{RU}}(z,a;x,y,w)=-\Gamma^{-1}v(z;x,y,w)+(1-\Gamma^{-1})\cdot a+(\Gamma-\Gamma^{-1})(-v(z;x,y,w)-a)_{+}. \tag{31}$$

那么，任意解

$$\{h^*_{\Gamma},\alpha^*_{\Gamma}\}\in\underset{h,\alpha}{\operatorname{argmin}}\,\mathbb{E}_{P_{\text{obs}}}\left[L^{\Gamma}_{RU}(h(X),\alpha(X,W);X,Y,W)\right] \tag{32}$$

也是 (29) 的解，对于任意 $Q_{\text{obs}}\ll P_{\text{obs}}$。 证明见附录 *D.10*。

备注 9。 如果 $v(z;\ x,\ y,\ w)$ 对于任意 $(x,\ y,\ w)\ \in\ \mathcal{X}\times\mathcal{Y}\times\{0,1\}$ 关于 z 是凹的，那么增广损失 $L^{\Gamma}_{\text{RU}}(z,\ a;\ x,\ y,\ w)$ 对于任意 $(z,\ a)$ 关于 $(x,\ y,\ w)\ \in\mathcal{X}\times\mathcal{Y}\times\{0,1\}$. 是联合凸的

现在，还需要定义适当的观测数据价值函数 v，将其代入 RU回归 程序后，可以得到最优的最大-最小策略和最优的最大-最小增益策略。

4.1 策略学习的观测数据价值函数

首先，我们定义潜在结果价值函数，对其进行最大化可以得到最大-最小策略、最大-最小增益策略以及极小极大遗憾策略。

定理 10. 定义 $\mathcal{H}=\{h\in L^2(P_{\text{obs},X},\mathcal{X})\ |\ 0\leq h(x)\ \leq 1\}$. 令

$$v^*_{maxmin}(z;x,y_0,y_1)=\log(1+\exp(2z-1))\cdot y_1+\log(1+\exp(-2z+1))\cdot y_0.$$

那么策略 $\pi(x)\ =\mathbb{I}\left(h^*(x)\ \geq\ \frac{1}{2}\right)$，其中

$$h^*\in\arg\sup_{h\in\mathcal{H}}\inf_{Q\in\mathcal{S}_{\Gamma}(P_{\text{obs}},Q_X)}\mathbb{E}_{Q}\left[v^*(h(X);X,Y(0),Y(1))\right], \tag{33}$$

是最大-最小目标函数下的最优策略 (17)。 证明见附录 *D.11*。

定理 11. 定义 $\mathcal{H} = \{h \in L^2(P_{\text{obs}},\ X,\ \mathcal{X}) \mid 0 \le h(x) \le 1\}$. 令

$$v^*_{gain}(z; x,\ y_0,\ y_1) = (1 - \pi_0(x)) \log(1 + \exp(2z - 1)) \cdot (y_1 - y_0) + \pi_0(x) \log(1 + \exp(-2z + 1)) \cdot (y_0 - y_1$$

然后 $\pi\ (x)\ = \mathbb{I}\left(h^*\ (x) \ge \frac{1}{2}\right)$, 其中

$$h^* \in \arg\sup_{h \in \mathcal{H}} \inf_{Q \in \mathcal{S}_\Gamma(P_{\text{obs}}, Q_X)} \mathbb{E}_Q \left[v^*_{gain}(\pi(X); X, Y(0), Y(1))\right] \tag{34}$$

是最大-最小增益目标下的最优策略（21）。证明见附录 *D.1* *2.*

定理 12。定义 $\mathcal{H} = \{h \in L^2\ (P_{\text{obs}},\ X,\ \mathcal{X})\ \mid 0 \le h\ (x)\ \le 1\}$. 令

$$v^*_{regret}(z; x, y_0, y_1) = (2z - 1)(y_1 - y_0) + \log\left(|2z - 1|\right).$$

然后 $\pi\ (x)\ = \mathbb{I}\ (h^*\ (x)\ \ge \frac{1}{2}\ )$ </style>, 其中

$$h^* \in \arg\sup_{h \in \mathcal{H}} \inf_{Q \in \mathcal{S}_\Gamma(P_{\text{obs}}, Q_X)} \mathbb{E}_Q \left[v^*_{regret}(\pi(X); X, Y(0), Y(1))\right] \tag{35}$$

是极小化最大后悔目标下的最优策略（23）。证明见附录 *D.13*。

在以下引理中，我们定义满足假设 1 的观测数据价值函数 v^*_{maxmin}, v^*_{gain}, v^*_{regret}.

引理 13。定义

$$v_{maxmin}(z; x, y, w) = \log(1 + \exp((2z - 1) \cdot (2w - 1))) \cdot \left(\frac{y}{we + (1 - w)(1 - e)}\right). \tag{36}$$

$$\begin{aligned} v_{gain}(z; x, y, w) = (1 - \pi_0(x)) \log(1 + \exp(2z - 1)) \cdot \left(\frac{y \cdot w}{e} - \frac{y \cdot (1 - w)}{1 - e}\right) \\ + \pi_0(x) \log(1 + \exp(-2z + 1)) \cdot \left(\frac{y \cdot (1 - w)}{1 - e} - \frac{y \cdot w}{e}\right). \end{aligned} \tag{37}$$

$$v_{regret}(z; x, y, w) = (2z - 1) \cdot \left(\frac{y \cdot w}{e} - \frac{y \cdot (1 - w)}{1 - e}\right) + \log\left(|2z - 1|\right). \tag{38}$$

价值函数 v_{maxmin}, v_{gain}, v_{regret} 分别在 *1* 假设下满足 v^*_{maxmin}, v^*_{gain}, v^*_{regret}。证明见附录 *D.14*。

由于观测数据价值函数 v_{maxmin}, v_{gain} 关于 z光滑且凹，我们可以结合第 3节的结果和定理8，说明求解最优最大-最小策略和最优最大-最小增益策略，相当于求解凸 RU 回归问题。因此，我们可以直接学习这些策略，而无需单独估计多个干扰参数。我们提出使用神经网络对函数 h, α 进行参数化，并通过 RU 损失（31）对其进行联合训练。

相比之下，我们发现 v_{regret} 关于 z 是非凹的，并且在$z = \frac{1}{2}$ 处存在奇点。因此，以 v_{regret} 作为观测数据价值函数的 RU 回归是一个非凸问题，可能难以通过连续优化方法求解。我们将求解极小极大遗憾问题的算法留待未来工作开发。

5 实验

在本节中，我们评估第4节提出的学习方法。首先，我们介绍如何实现用于策略学习的RU回归。其次，我们将RU回归在一维合成数据集上学得的最大-最小策略和最大-最小增益策略可视化，并将这些策略与式（17）和（21）给出的真实策略进行比较。我们还考察式（23）给出的真实极小极大遗憾策略。第三，在半合成实验中，我们将最大-最小策略和最大-最小增益策略的表现与非鲁棒策略（15）进行比较。

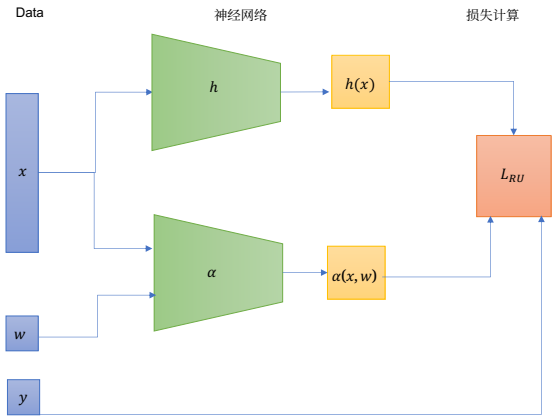

图 2：用于策略学习的 RU回归架构。我们注意到，函数 h 以 X 作为输入，而辅助函数 α 以 X、W 作为输入。

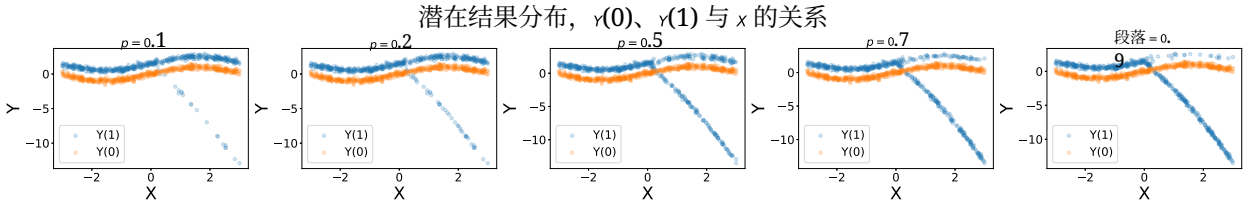

图 3：我们将潜在结果分布可视化为 p 变化时的情况。当 p 较低时，我们的潜在结果分布主要由满足 $Y_i(1) > Y_i(0)$ 的研究单位组成。随着 p 增加，满足$Y_i(0) > Y_i(1)$的研究单位比例也随之增加。

5.1 RU回归在策略学习中的实现

本研究中的RU回归实现与Sahoo等人 [2022]的方法有几处细微差异。根据定理8，在策略学习场景中，辅助函数 α 取决于协变量X 和处理分配 W；而在Sahoo等人 [2022]的回归场景中， α 仅取决于 X。这一点体现在我们用于学习 h的架构中， α （图2）。

我们注意到，价值函数 v_{maxmin} 在（36）中以及 v_{gain} 在（37）中在上方无界，因此在无界函数类上没有最大化解。为确保将 v_{maxmin} 或 v_{gain} 代入（31）时，RU 损失（31）不会在下方无界，我们在有界函数类上对 h 进行优化。具体而言，我们旨在学习回归函数 $h_\Gamma(x)$： $\mathcal{X} \to [0,1]$. 然后，我们通过计算来确定处理分配

$$\pi_\Gamma(x) = \mathbb{I}\Big(h_\Gamma(x) \geq \frac{1}{2}\Big).$$

由于策略仅取决于 h 是大于还是小于 $1\ 2$，因此将 h 限制在有界函数类中不会影响最优策略价值。为确保 h 的输出位于 $[0,1]$，我们将 Sigmoid 激活函数作为表示 h 的神经网络的最终激活函数，从而确保网络输出值位于 $[0,1]$.

5.2 玩具实验

首先，我们考虑一个简单的一维实验，以便将数据分布可视化，并将学习得到的策略与根据数据模型计算出的最优策略进行比较。

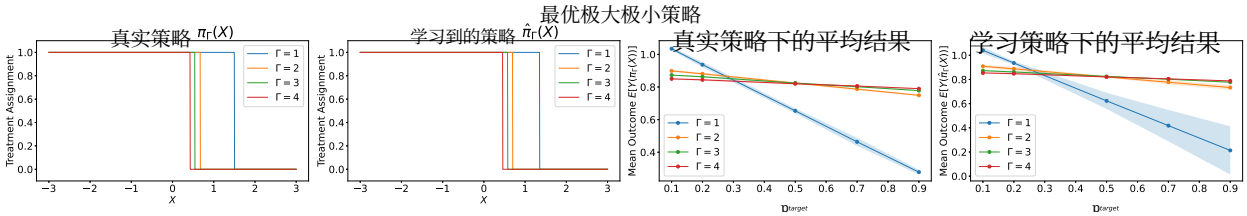

图4：一维玩具示例的最大最小策略结果。**左图**：我们展示了使用数据模型（39）计算出的真实最大最小策略。**中左图**：对于一个随机种子，我们展示了学习到的最大最小策略。**中右图**：我们展示了不同 p_{target}目标分布下真实策略的平均结果，汇总了6次随机试验，其中随机性来自数据集生成。**右图**：我们展示了不同 p_{target}目标分布下学习到的策略的平均结果，汇总了6次随机试验，其中随机性来自数据集生成。

5.2.1 数据

我们考虑一个玩具示例，其中研究和目标潜在结果分布按如下方式生成。设 U_i 为一个会影响 Y_i(1)。

$$\begin{aligned} X_i &\sim \text{Uniform}[-3,3] \\ U_i &\sim \text{Bernoulli}(p) \\ Y_i(0) \mid X_i &\sim N(\sin(X_i), \sigma^2) \\ Y_i(1) \mid X_i &\sim N(\sin(X_i) + 1.5 - U_i \cdot (5X_i)_+, \sigma^2), \end{aligned} \tag{39}$$

其中 $\sigma = 0.2$. 我们在图3中展示潜在结果分布。我们通过改变 p 的伯努利参数来生成研究潜在结果分布和目标潜在结果分布，其中 U_i的分布为 P 表示满足 $p_{\text{study}} = 0.2$的潜在结果分布。给定P，我们生成如下的观测数据分布 P_{obs}，其定义在（X_i，Y_i，W_i）上。

$$W_i \sim \text{Bernoulli}(0.5), \quad Y_i = Y_i(W_i).$$

目标潜在结果分布 Q 通过改变 $p_{\text{target}} \in [0,1]$.生成

5.2.2 策略

我们使用RU回归（其实现方式在第5.1节中有所介绍），仅利用来自观测数据分布的数据来学习最优的最大-最小策略和最大-最小增益策略。由于这是一个一维合成示例，我们还使用数据模型（39）和闭式公式（23）来计算极小极大遗憾策略。

我们评估以下4种策略。

1. 最大-最小策略（通过RU回归学习）。
2. 采用始终不治疗基线的最大-最小增益策略（通过RU回归学习 ）.
3. 带有始终处理基线的最大-最小增益策略（通过 RU回归学习）。
4. 极小极大遗憾（使用 (39) 和 (23) 计算）。

5.2.3 结果

我们使用 (39) 中的数据模型来计算干扰参数 $\tau(\cdot)$ $H_\Gamma(\cdot)$ $H_\Gamma^+(\cdot)$ $H_\Gamma^-(\cdot)$，并将它们代入 (17) 和 (21) 中，以计算真实最优策略。比较图 4、5、6中左侧和中左侧的图可以发现，我们学习得到的策略与真实策略一致。比较图4、5、6中中右侧和右侧的图可以看出，在 6 次随机试验中，学习得到的策略和真实策略取得了相近的平均结果。

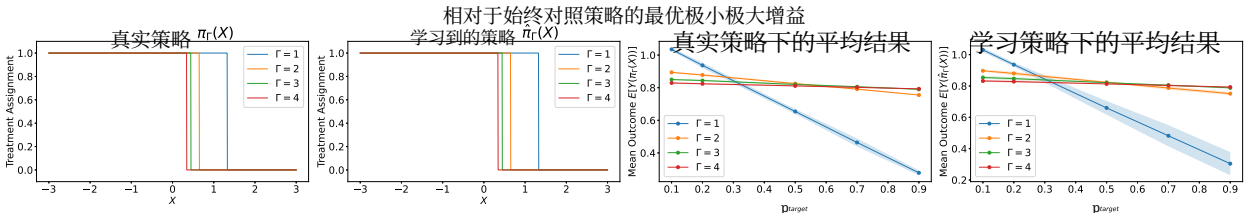

图5：一维玩具示例中最大-最小增益（相对于始终不治疗）的策略结果。**左图**：我们可视化了真实的最大-最小增益策略，其中基线为 $\pi_0(X) = 0$（始终不治疗），该策略使用数据模型（39）计算得到。**中左图**：对于一个随机种子，我们可视化了相对于始终不治疗策略的学习到的最大最小增益。**中右图**：我们可视化了真实策略在具有不同 p_{target}的目标分布下的平均结果，该结果汇总了6次随机试验，其中随机性来自数据集生成。**右图**：我们可视化了学习到的策略在具有不同 p_{target}的目标分布下的平均结果，该结果汇总了6次随机试验，其中随机性来自数据集生成。

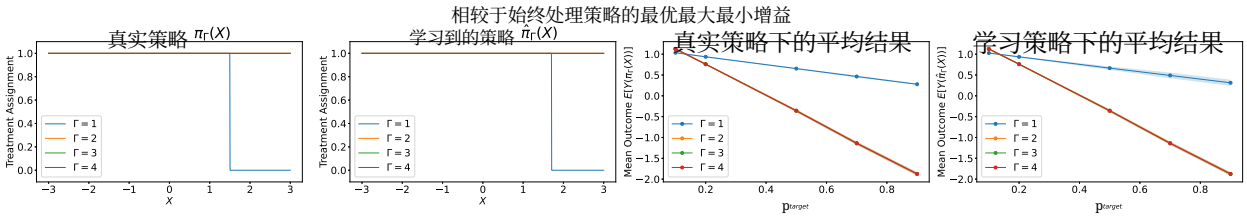

图6：一维玩具示例中最大-最小增益（相对于始终治疗）的策略结果。**左图**：我们可视化了真实的最大-最小增益策略，其中基线为 $\pi_0(X) = 1$（始终治疗），该策略使用数据模型（39）计算得到。**中左图**：对于一个随机种子，我们可视化了相对于始终不治疗策略的学习到的最大最小增益。**中右图**：我们可视化了真实策略在具有不同 p_{target}的目标分布下的平均结果，该结果汇总了6次随机试验，其中随机性来自数据集生成。**右图**：我们可视化了学习到的策略在具有不同 p_{target}的目标分布下的平均结果，该结果汇总了6次随机试验，其中随机性来自数据集生成。

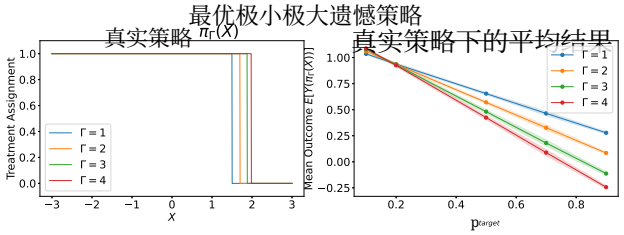

图7：一维玩具示例的极小极大遗憾策略结果。**左侧**：我们将使用数据模型（39）计算得到的真实极小极大遗憾策略可视化。**右侧**：我们将目标分布具有不同 p_{target}的学习到的策略的平均结果可视化，并汇总了6次随机试验，其中随机性来自数据集生成。

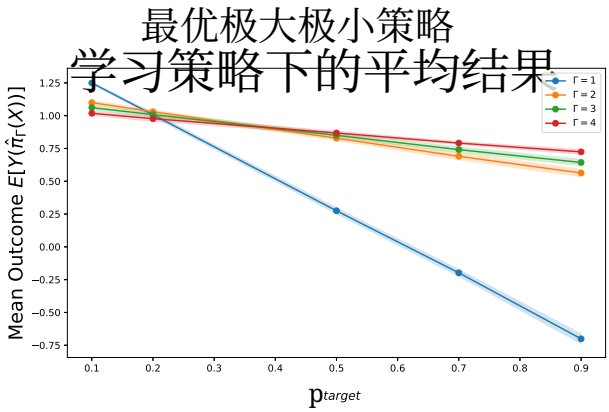

图8：高维模拟中的最大-最小策略结果。

注意 $\pi_{\text{non-robust}}$，式（15）在 $\Gamma = 1$ 时等于最优最大-最小策略、最大-最小增益策略或极小化最大后悔策略。非鲁棒策略建议对 CATE 非负的研究单位实施处理。当 $p_{\text{target}} < 0.3$ 时，该策略表现良好；但当 $p_{\text{target}} > 0.3.$ 时，其表现不如最大-最小策略以及相对于始终不治疗策略的最大-最小增益。我们注意到，相对于始终治疗策略的最大-最小增益在 $\Gamma > 1.$ 时不会偏离基线。有些出人意料的是，在这个特定示例中，极小化最大后悔策略比 $\pi_{\text{non-robust}}$（图 7）更不保守，因为随着 $\pi_{\text{non-robust}}$ 的 Γ 增加，它建议对总体中更高比例的研究单位实施处理。

5.3 高维实验

其次，我们在一个高维（$d = 10$）模拟中评估我们的方法。

5.3.1 数据

我们考虑一个高维（$d = 10$）示例，其中研究数据和目标潜在结果分布按如下方式生成。设 U_i 为一个会影响 Y_i (1)的未观测变量。

$$\begin{aligned} X_i &\sim \text{Uniform}[-3,3]^d \\ U_i &\sim \text{Bernoulli}(p) \\ Y_i(0) \mid X_i &\sim N(\sin(a^T X_i), \sigma^2) \\ Y_i(1) \mid X_i &\sim N(\sin(a^T X_i) + 1.5 - U_i \cdot (2(X_{i,1} + X_{i,2} + X_{i,3}))_+, \sigma^2), \end{aligned} \tag{40}$$

其中 $\sigma = 0.2$ 和 a 是第 B节中定义的常数向量。

5.3.2 策略

我们评估以下策略。

1. 最大-最小（通过RU回归学习）
2. 相对于基准策略的最大-最小增益（通过RU回归学习）

$$\pi_0(X) = \mathbb{I}(X_1 \leq 0).$$

5.3.3 结果

与玩具实验中一样，我们观察到当 $\Gamma = 1$ 最优鲁棒策略对应于 $\pi_{\text{non-robust}}$.与之前一样，我们观察到 $\pi_{\text{non-robust}}$ 在 p_{target} 较小时表现良好，但随着 $p_{\text{target}}.$ 取值增大，性能会下降。相比之下，当 $\Gamma > 1$时，最大-最小策略（图 8）和最大-最小增益策略（图 9）对 p_{target}的变化更加鲁棒。

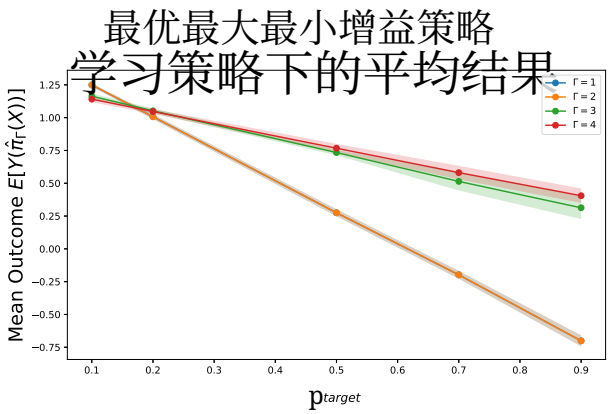

图9：高维模拟中的最大-最小增益策略结果。我们注意到，当 $\Gamma = 1{,}2$ 时，最大最小增益策略不会偏离 $\pi_{\text{non-robust}}$。当 Γ 的值较高时，对于较高的 p_{target}，最大最小增益策略优于 $\pi_{\text{non-robust}}$

5.4 使用 VotingDataset 的半合成实验

我们在一项半合成实验中比较最大-最小策略与最大-最小增益策略的表现，其中使用了 Gerber等人的 [2008]投票数据集。

5.4.1 数据

Gerber等人 [2008] 的投票数据集来自一项随机对照试验式研究，该研究旨在评估各种行动对选民投票率的影响。研究人员设计了一个对照组行动和四种处理行动，内容是在2006年密歇根州初选前向选定的研究单位寄送信件。在原始现场实验中，对照组行动的概率为$5\ 9$，每种处理行动的概率为$1\ 9$。

在我们的实验中，我们只关注“对照组”行动（$W_i = 0$）和“邻居”行动（$W_i = 1$）。请注意，这意味着处理的概率为 $e = \frac{1}{9}/(\frac{1}{9} + \frac{5}{9}) = \frac{1}{6}$。“邻居”行动包括向个人寄送一封信，信中包含该个人、其家庭其他成员以及邻居的投票参与记录。信中还提到，选举结束后会寄送一封后续信件，告知所有人的最新投票参与情况，因此该个人的投票参与情况会被邻居知晓。“对照组”行动则是不向个人寄送信件。

在我们的实验中，协变量 X_i 包括个人的家庭规模、年龄、性别，以及个人是否 i 在2000年至2002年的历次初选和2000年至2004年的历次大选中投过票。结果 Y_i 是个人 i 是否在2006年初选中投过票。尽管数据集还包含个人 i 是否在2004年初选中投过票的信息，但我们将其视为未观测变量 U_i，并利用它生成合成目标总体和研究人群。

通过附录B.3中所述的数据集生成流程，我们创建了一个合并的训练与验证数据集，其中67%的研究单位具有 $U_i = 1$。相比之下，具有 $U_i = 1$ 的研究单位仅占测试数据集的18%。训练集、验证数据集和测试集分别包含62044、41364和126036个样本。根据设计，与目标总体相比，在2004年大选中投过票的研究单位在研究人群中的比例更高。我们使用合成研究人群的数据拟合策略，并评估合成目标总体下的平均结果。

5.4.2 策略

我们评估以下策略。

1. 最大-最小（通过RU回归学习）

方法	目标策略价值	目标总体接受治疗的比例
非鲁棒	0.3106 ± 0.003	0.66
最大-最小（$\Gamma = 1.1$）	0.3272 ± 0.003	0.87
最大-最小 ($\Gamma = 1.2$)	0.3375 ± 0.004	1.0
最大-最小 ($\Gamma = 1.3$)	0.3375 ± 0.004	1.0
最大-最小 ($\Gamma = 1.5$)	0.3375 ± 0.004	1.0
相较于始终治疗的最大-最小增益（$\Gamma = 1.1$）	0.3375 ± 0.004	1.0
相较于始终治疗的最大-最小增益（$\Gamma = 1.2$）	0.3375 ± 0.004	1.0
相较于始终治疗的最大-最小增益（$\Gamma = 1.3$）	0.3375 ± 0.004	1.0
相较于始终治疗的最大-最小增益（$\Gamma = 1.5$）	0.3375 ± 0.004	1.0
相较于始终不治疗的最大-最小增益（$\Gamma = 1.1$）	0.2934 ± 0.003	0.44
相较于始终不治疗的最大-最小增益（$\Gamma = 1.2$）	0.2780 ± 0.002	0.25
相较于始终不治疗的最大-最小增益（$\Gamma = 1.3$）	0.2705 ± 0.002	0.15
相较于始终不治疗的最大-最小增益（$\Gamma = 1.5$）	0.2653 ± 0.001	0.0

表1：非鲁棒策略建议对66%的总体实施处理。最大-最小策略以及相对于始终治疗策略的最大-最小增益策略建议对整个总体实施处理。相比之下，相对于始终控制策略的最大-最小增益策略建议对总体中比例小得多的部分实施处理。

2. 使用始终不治疗基线（通过RU回归学习）的最大-最小增益

$$\pi_0(X) = 0.$$

3. 基于始终治疗基线的最大-最小增益（通过RU回归学习）

$$\pi_0(X) = 1.$$

5.4.3 结果

回顾一下，与之前一样，$\Gamma = 1$ 对应于 $\pi_{\text{non-robust}}$，因此我们报告 $\pi_{\text{non-robust}}$（通过RU回归学习，其中 v_{maxmin} 和 $\Gamma = 1$）。我们评估针对 $\Gamma = 1.1$、1.2、1.3、1.5、2、3、4学习得到的鲁棒策略。我们仅报告 $\Gamma = 1.1$、1.2、1.3、1.5的结果，因为当 $\Gamma \geq 1.5$ 时，我们观察到与 $\Gamma = 1.5$相同的结果。这可以由以下事实解释：结果 $Y_i(0)$ $Y_i(1)$是二值的，因此$q_{\zeta(\Gamma)}$（Y（w） $\mid X = x$）在许多 Γ取值下可能具有相同的值。请注意，我们的理论结果仅适用于条件研究潜在结果分布相对于勒贝格测度绝对连续的情况，但在本实验中，研究潜在结果分布是离散分布。尽管如此，我们观察到的结果与连续值结果下的结果一致。我们在表1中报告不同学习策略下的目标策略价值和目标总体接受治疗的比例。我们发现，非鲁棒策略建议为约66%的总体提供治疗。我们发现，最大-最小策略以及相对于始终治疗基准的最大-最小增益策略建议为整个总体提供治疗。相比之下，相对于始终不治疗基线的最大-最小增益策略建议，随着 Γ 增加，为越来越小比例的总体提供治疗。在这个特定示例中，我们观察到，不同的鲁棒目标函数可能产生截然不同的策略。

参考文献

Christopher Adjaho 和 TimothyChristensen。外部有效的治疗选择。*arXiv* 预印本 *arXiv:2205.05561*，2022。

Peter M Aronow 和 Donald KK Lee。在样本选择概率未知但有界的情况下估计总体均值。*Biometrika*，100(1):235–240，2013年。

Susan Athey 和 Ste fan Wager。基于观察性数据的策略学习。*Econometrica*，89(1):133–161，202 1.

Stephen H Bell、Robert B Olsen、Larry L Orr 和 Elizabeth A Stuart。当影响评估非随机地选择地点时外部有效性偏差的估计。教育评估与策略分析，38(2): 318–335，2016。

Eli Ben-Michael、D James Greiner、Kosuke Imai 和 Zhichao Jiang。通过外推进行安全策略学习：应用于审前风险评估。*arXiv* 预印本 *arXiv:2109.11679*，2021年。

Aharon Ben-Tal、Dick Den Hertog、Anja De Waegenaere、Bertrand Melenberg 和 Gijs Rennen。受不确定概率影响的优化问题的稳健解。《管理科学》，59(2)：341–357，2013年。Debopam Bhattacharya 和 Pascaline Dupas。在预算约束下推断福利最大化治疗分配。《计量经济学杂志》，167(1)：168–196，2012年。Timothy Christensen、Hyungsik Roger Moon 和 Frank Schorfheide。收益部分可识别时的最优离散决策。*arXiv* 预印本 *arXiv:2204.11748*，2022。Ivar Ekeland、Alfred Galichon 和 Marc Henry。多元风险的共单调测度。《数学金融：数学、统计学与金融经济学国际期刊》，22(1)：109–132，2012年。Alan S Gerber、Donald P Green 和 Christopher W Larimer。社会压力与选民投票率：来自一项大规模现场实验的证据。《美国政治学评论》，102(1)：33–48，2008年。Lars Peter Hansen 和 Thomas J Sargent。稳健控制与模型不确定性。《美国经济评论》，91(2)：60–66，2001年。Tobias Hatt、Daniel Tschernutter 和 Stefan Feuerriegel。样本选择偏差下的离策略学习推广。在《人工智能中的不确定性》中，第769–779页。PMLR，2022。Samuel Higbee。新治疗方案下的策略学习。*arXiv* 预印本 *arXiv:2210.04703*，2022。Nathan Kallus 和 Angela Zhou。未观测混杂下的极小极大最优策略学习。《管理科学》，67(5)：2870–2890，2021年。Nathan Kallus、Xiaojie Mao、Kaiwen Wang 和 Zhengyuan Zhou。双重稳健分布鲁棒离策略评估与学习。在《国际机器学习会议》中，第10598–10632页。PMLR，2022。Toru Kitagawa 和 Aleksey Tetenov。谁应当接受治疗？用于治疗选择的经验福利最大化方法。*Econometrica*，86(2)：591–616，2018年。Charles F Manski。识别问题与模糊性下的决策：治疗反应的经验分析与治疗选择的规范分析。《计量经济学杂志》，95(2)：415–442，2000年。Charles F Manski。概率分布的部分识别，第5卷。施普林格，2003年。Charles F Manski。异质性总体的统计治疗规则。*Econometrica*，72(4)：1221–1246，2004年。Charles F Manski。缺失结果数据下的极小极大后悔治疗选择。《计量经济学杂志》，139(1)：105–115，2007年。Charles F Manski。模糊性下的治疗政策选择。《经济学年度评论》，3(1)：25–49，2011年。Tong Mu、Yash Chandak、Tatsunori B Hashimoto 和 Emma Brunskill。因子化分布鲁棒优化：上下文老虎机的因子化分布鲁棒策略。《神经信息处理系统进展》，35：8318–8331，2022。Xinkun Nie、Guido Imbens 和 Stefan Wager。用于跨地点外推随机试验的协变量平衡敏感性分析。*arXiv* 预印本 *arXiv:2112.04723*，2021年。Georg Ch Pflug。关于风险价值和条件风险价值的几点评论。《概率约束优化：方法论与应用》，第272–281页，2000年。Min Qian 和 Susan A Murphy。个体化治疗规则的性能保证。《统计学年鉴》，39(2)：1180，2011年。

R Tyrrell Rockafellar、Stan Uryasev 和 Michael Zabarankin。关于未条件风险价值的优化。*Journal of Risk*，2：21–42，2000年。Roshni Sahoo、Lihua Lei 和 Stefan Wager。从有偏样本中学习。*arXiv* 预印本 *arXiv:2209.01754*，2022。Leonard J Savage。统计决策理论。美国统计学会期刊，46（253）：55–67，1951年。Nian Si、Fan Zhang、Zhengyuan Zhou 和 Jose Blanchet。分布稳健的批量上下文老虎机。*arXiv* 预印本 *arXiv:2006.05630*，2020。Elizabeth A Stuart、Stephen R Cole、Catherine P Bradshaw 和 Philip J Leaf。利用倾向得分评估随机试验结果的可推广性。皇家统计学会期刊：*A*辑（社会统计学），174(2):369–386，2011年。Adith Swaminathan 和 Thorsten Joachims。通过反事实风险最小化从日志赌博机反馈中进行批量学习。机器学习研究期刊，16(1):1731–1755，2015。Elizabeth Tipton。利用倾向得分亚分类改进实验的推广：假设、性质与情境。教育与行为统计学期刊，38(3):239–266，2013。Elizabeth Tipton。你的实验有多大可推广性？一个用于比较实验样本与总体的指数。教育与行为统计学期刊，39(6):478–501，2014。Sheng-Min Wang、Changsu Han、Soo-Jung Lee、Tae-Youn Jun、Ashwin A Patkar、Prakash S Masand 和 Chi-Un Pae。抗抑郁药的疗效：随机临床试验中的偏差及相关问题。临床药理学专家评论，11(1):15–25，2018。James Watson 和 Chris Holmes。近似模型与稳健决策。2016。Yingqi Zhao、Donglin Zeng、A John Rush 和 Michael R Kosorok。使用结果加权学习估计个体化治疗规则。美国统计学会期刊，107(499):1106–1118，2012年。

A 附加实验结果

方法	目标策略价值				
	$p = 0.1$	$p = 0.2$	$p = 0.5$	$p = 0.7$	$p = 0.9$
RU回归 ($\Gamma = 1$)	1.029 ± 0.016	0.930 ± 0.009	0.659 ± 0.015	0.491 ± 0.025	0.308 ± 0.062
True ($\Gamma = 1$)	1.035 ± 0.010	0.937 ± 0.016	0.654 ± 0.015	0.464 ± 0.025	0.278 ± 0.015
RU回归 ($\Gamma = 2$)	0.908 ± 0.015	0.882 ± 0.007	0.819 ± 0.003	0.780 ± 0.017	0.740 ± 0.019
True ($\Gamma = 2$)	0.899 ± 0.005	0.881 ± 0.007	0.824 ± 0.005	0.787 ± 0.006	0.749 ± 0.004
RU回归 ($\Gamma = 3$)	0.874 ± 0.008	0.861 ± 0.009	0.823 ± 0.003	0.801 ± 0.002	0.776 ± 0.009
True ($\Gamma = 3$)	0.873 ± 0.003	0.863 ± 0.005	0.824 ± 0.005	0.802 ± 0.004	0.778 ± 0.003
RU回归 ($\Gamma = 4$)	0.856 ± 0.010	0.847 ± 0.009	0.819 ± 0.001	0.804 ± 0.004	0.785 ± 0.008
True ($\Gamma = 4$)	0.850 ± 0.004	0.844 ± 0.006	0.820 ± 0.005	0.806 ± 0.004	0.790 ± 0.003

表 2：一维模拟实验中的最大-最小策略结果。我们报告了 6 次随机试验中目标策略价值的均值和标准差，其中随机性来自数据集生成。请注意，采用 $\Gamma = 1$ 的 RU回归会得到一个等价于 $\pi_{\text{non-robust}}$ 的策略，并且在 p 取值较低时产生较低的策略值。对于 $\Gamma > 1$，最大最小策略 $\hat{}\pi_\Gamma$ 在 p 取值较高时会产生更高的策略值。

方法	目标策略价值				
	$p = 0.1$	$p = 0.2$	$p = 0.5$	$p = 0.7$	$p = 0.9$
RU回归 ($\Gamma = 1$)	1.031 ± 0.016	0.930 ± 0.010	0.656 ± 0.019	0.484 ± 0.024	0.301 ± 0.057
True ($\Gamma = 1$)	1.035 ± 0.010	0.937 ± 0.016	0.654 ± 0.015	0.464 ± 0.025	0.278 ± 0.015
RU回归 ($\Gamma = 2$)	0.897 ± 0.011	0.876 ± 0.005	0.821 ± 0.003	0.788 ± 0.010	0.753 ± 0.011
True ($\Gamma = 2$)	0.894 ± 0.004	0.878 ± 0.006	0.825 ± 0.005	0.791 ± 0.005	0.755 ± 0.003
RU回归 ($\Gamma = 3$)	0.854 ± 0.009	0.846 ± 0.008	0.818 ± 0.002	0.804 ± 0.004	0.788 ± 0.008
True ($\Gamma = 3$)	0.850 ± 0.004	0.844 ± 0.006	0.820 ± 0.005	0.806 ± 0.004	0.790 ± 0.003
RU回归 ($\Gamma = 4$)	0.834 ± 0.006	0.827 ± 0.005	0.812 ± 0.001	0.801 ± 0.005	0.792 ± 0.004
True ($\Gamma = 4$)	0.829 ± 0.005	0.824 ± 0.005	0.811 ± 0.004	0.802 ± 0.004	0.793 ± 0.003

表3：一维模拟实验中相对于始终对照策略结果的最大-最小增益。我们报告了6次随机试验中目标策略价值的均值和标准差，其中随机性来自数据集生成。请注意，带有 $\Gamma = 1$ 的RU回归产生的策略等价于 $\pi_{\text{non-robust}}$，并且在 p取值较低时会产生较低的策略值。相对于始终控制策略的最大最小增益$\hat{}\pi_\Gamma$在 $\Gamma > 1$ 下会在 p取值较高时产生更高的策略值。

方法	目标策略价值				
	$p = 0.1$	$p = 0.2$	$p = 0.5$	$p = 0.7$	$p = 0.9$
RU回归 ($\Gamma = 1$)	1.028 ± 0.016	0.929 ± 0.009	0.660 ± 0.020	0.492 ± 0.029	0.311 ± 0.066
True ($\Gamma = 1$)	1.035 ± 0.010	0.937 ± 0.016	0.654 ± 0.015	0.464 ± 0.025	0.278 ± 0.015
RU回归 ($\Gamma = 2$)	1.124 ± 0.021	0.755 ± 0.019	-0.355 ± 0.035	-1.131 ± 0.043	-1.856 ± 0.032
True ($\Gamma = 2$)	1.126 ± 0.022	0.760 ± 0.019	-0.358 ± 0.031	-1.139 ± 0.037	-1.875 ± 0.039
RU回归 ($\Gamma = 3$)	1.124 ± 0.021	0.755 ± 0.019	-0.355 ± 0.035	-1.131 ± 0.043	-1.856 ± 0.032
True ($\Gamma = 3$)	1.126 ± 0.022	0.760 ± 0.019	-0.358 ± 0.031	-1.139 ± 0.037	-1.875 ± 0.039
RU回归 ($\Gamma = 4$)	1.121 ± 0.025	0.747 ± 0.013	-0.356 ± 0.043	-1.127 ± 0.052	-1.841 ± 0.012
True ($\Gamma = 4$)	1.126 ± 0.022	0.760 ± 0.019	-0.358 ± 0.031	-1.139 ± 0.037	-1.875 ± 0.039

表 4：一维模拟实验中最大最小增益相对于始终处理策略的结果。我们报告了 6 次随机试验中目标策略价值的均值和标准差，其中随机性来自数据集生成。请注意，使用 $\Gamma = 1$ 的 RU回归会得到与$\pi_{\text{non-robust}}$等效的策略。对于 $\Gamma > 1$，该最大最小增益策略默认采用基准策略 $\pi_0(x) = 1$ 。

方法	目标策略价值				
	$p = 0.1$	$p = 0.2$	$p = 0.5$	$p = 0.7$	$p = 0.9$
RU回归 ($\Gamma = 1$)	1.249 ± 0.012	1.007 ± 0.019	0.275 ± 0.027	-0.198 ± 0.030	-0.701 ± 0.042
RU回归 ($\Gamma = 2$)	1.249 ± 0.012	1.007 ± 0.019	0.275 ± 0.027	-0.198 ± 0.030	-0.701 ± 0.042
RU回归 ($\Gamma = 3$)	1.164 ± 0.020	1.052 ± 0.018	0.734 ± 0.021	0.514 ± 0.069	0.313 ± 0.086
RU回归 ($\Gamma = 4$)	1.141 ± 0.023	1.047 ± 0.018	0.767 ± 0.033	0.581 ± 0.052	0.405 ± 0.055

表5：高维模拟的最大最小增益策略结果。

方法	目标策略价值				
	$p = 0.1$	$p = 0.2$	$p = 0.5$	$p = 0.7$	$p = 0.9$
RU回归 ($\Gamma = 1$)	1.249 ± 0.012	1.007 ± 0.019	0.275 ± 0.027	-0.198 ± 0.030	-0.701 ± 0.042
RU回归 ($\Gamma = 2$)	1.099 ± 0.021	1.030 ± 0.016	0.828 ± 0.021	0.690 ± 0.030	0.563 ± 0.033
RU回归 ($\Gamma = 3$)	1.061 ± 0.020	1.008 ± 0.022	0.850 ± 0.020	0.741 ± 0.025	0.643 ± 0.028
RU回归 ($\Gamma = 4$)	1.018 ± 0.019	0.977 ± 0.024	0.868 ± 0.017	0.791 ± 0.014	0.724 ± 0.022

表6：高维模拟的最大最小增益策略结果。

B 实验细节

B.1 玩具示例

B.1.1 模型

对于 RU 回归模型，我们联合训练两个神经网络，分别学习回归函数 h 和分位数函数 α。图 2展示了 RU 回归模型架构的可视化结果。每个神经网络均包含 3 个隐藏层，每层有 64 个研究单位，并采用 ReLU 激活函数。RU 回归模型包含 17.1K 个可训练参数。

B.1.2 数据集划分

对于所有方法，训练集、验证集和测试集分别包含 20000、10000 和 10000 个样本。训练集和验证集通过（39）中指定的数据模型生成，其中 $p = 0.2$. 所有方法均在相同的测试集上进行评估，这些测试集通过数据模型（39）生成，其参数 p 取值范围为 $[0.1, 0.2, 0.5, 0.7, 0.9]$。对于 6 个随机种子中的每一个 $[0, 1, 2, 3, 4, 5],$，都会生成一个新的数据集（训练集、验证集和测试集）。

B.1.3 训练过程

模型最多训练 50 个训练轮次，批量大小设为 4000，并使用学习率为 $1e-2$ 的 Adam 优化器。在每个训练轮次中，我们都会检查模型在验证集上得到的损失，并选择在验证集上损失最小的模型。

B.2 高维实验

B.2.1 数据

在（40）中的数据模型里，我们设定

$$\begin{aligned} a =& [0.50067509, -0.67108696, -2.22006362, -0.37834265, 1.05841302, \\ & -0.4509034, 1.15857361, 0.62236239, -0.77458079, -0.74790281]. \end{aligned}$$

B.2.2 模型

我们使用与玩具实验中相同的模型（见第 B.1.1节）。

B.2.3 数据集划分

我们使用与玩具实验中相同的数据集划分（见第B.1.2节）

B.2.4 训练过程

我们采用与玩具实验中相同的训练过程（见第 B.1.3节）。

B.3 Voting半合成实验

B.3.1 模型

我们使用与玩具实验中相同的模型（见第 B.1.1节）。

B.3.2 数据集划分

我们随机选取带有 $U_i = 1$ 的75%研究单位和带有 $U_i = 0$ 的25%研究单位，创建合并的训练与验证数据集，该数据集代表研究人群。其余研究单位构成测试集，代表目标总体。

B.3.3 训练过程

我们采用与玩具实验中相同的训练过程（请参阅第B.1.3节）。

C 标准结果

引理 14. 令 $\zeta \in (0,1)$。条件风险价值具有以下性质：

1. 如果 Y 具有密度，那么

$$\mathbb{E}[Y] = (1-\zeta)\mathrm{CVaR}_{\zeta}(Y) - \zeta\mathrm{CVaR}_{1-\zeta}(-Y),$$

2. 条件风险价值 具有正齐次性，即。

$$\mathrm{CVaR}_{\zeta}(cY) = c\mathrm{CVaR}_{\zeta}(Y)$$

如果 $c > 0$,

[Pflug, 2000]。

引理 15. 对于具有有限二阶矩的概率测度，条件风险价值是一种强一致风险度量，这意味着它是凸的、连续的，并且对于随机变量 X, Y，若其具有绝对连续分布F_X, F_Y

$$\mathrm{CVaR}(X) + \mathrm{CVaR}(Y) = \sup\{\mathrm{CVaR}(\tilde{X}+\tilde{Y}) : (\tilde{X},\tilde{Y}) \sim \mu, \mu \in \mathcal{T}(F_X, F_Y)\}$$

其中 $\mathcal{T}$是所有边缘分布为 F_X的二元分布集合， F_Y *[Ekeland 等人, 2012]*。

命题 16. 设 f 为凸函数，且 C 为包含于 $\mathrm{dom} f$中的闭有界凸集。那么，相对于 f， C 的上确界有限，且在 C *[Rockafellar, 1997]*的某个极点处取得。

引理17. 设 $J(z) = A\log(1+\exp\{2z-1\}) + B\log(1+\exp\{-2z+1\})$。 设 $z^* = \mathrm{argmax}_{z\in[0,1]} J(z)$。$A \geq B$ 当且仅当$z^* \geq \frac{1}{2}$. 证明见附录 *E.1*。

D 主要结果的证明

D.1 记号

在证明中，我们使用以下记号。这里，我们假设 $Y(0)$ $Y(1)$ 服从潜在结果分布，其边缘分布与 P_{obs} 的边缘分布一致

$$\begin{aligned}
c_1^+(x) &= \mathrm{CVaR}_{\zeta(\Gamma)}(Y(1) \mid x)\\
c_1^-(x) &= \mathrm{CVaR}_{\zeta(\Gamma)}(-Y(1) \mid x)\\
c_0^+(x) &= \mathrm{CVaR}_{\zeta(\Gamma)}(Y(0) \mid x)\\
c_0^-(x) &= \mathrm{CVaR}_{\zeta(\Gamma)}(-Y(0) \mid x)\\
g_1(x) &= \Gamma\tau(x) + (1-\Gamma)(c_1^+(x) + c_0^-(x))\\
g_0(x) &= -\Gamma\tau(x) + (1-\Gamma)(c_0^+(x) + c_1^-(x)).
\end{aligned}$$

注意

$$\begin{aligned}
H_\Gamma^+(x) &= (1-\Gamma^{-1})\cdot(c_1^+(x) + c_0^-(x))\\
H_\Gamma^-(x) &= (1-\Gamma^{-1})\cdot(-c_1^-(x) - c_0^+(x))\\
\tau(x) \geq H_\Gamma^+(x) &\iff g_1(x) \geq 0\\
\tau(x) \geq H_\Gamma^-(x) &\iff g_0(x) \leq 0.
\end{aligned}$$

在本节中，我们假设对于任意 x，$P_{(Y(1)\ Y(0))|X=x}$ 几乎处处都关于勒贝格测度绝对连续。

D.2 技术引理

引理 18. 设 P, Q 为潜在结果分布。当且仅当以下条件成立时，Q 通过 Γ-有偏抽样生成 P

$$\Gamma^{-1} \leq \frac{dQ_{Y(0),Y(1)|X=x}(y_0, y_1)}{dP_{Y(0),Y(1)|X=x}(y_0, y_1)} \leq \Gamma \quad \forall x \in \mathcal{X}, y_0, y_1 \in \mathbb{R} \tag{41}$$

且 上确界$_{x\in\mathcal{X}} \frac{dP_X}{}(x)\ dQ_X(x) < \infty$. 证明见附录 *E.8*。

引理 19。 策略 π 对每个满足 $Q_X \ll P_{\text{obs},X}$ 且 上确界$_{x\in\mathcal{X}}\ dP_{\text{obs},X}(x)\ dQ_X(x) < \infty$ 的 Q_X，都能解决 (26)，当且仅当 π能够解决

$$\max_{\pi(x)\in\{0,1\}} \inf_{P\in\mathcal{T}(P_{\text{obs}})} \Gamma\mathbb{E}_P\left[v^*(\pi(x); X, Y(0), Y(1)) \mid X = x\right] + (1-\Gamma)\text{CVaR}_{\zeta(\Gamma)}(v^*(\pi(x); X, Y(0), Y(1)) \mid x) \tag{42}$$

对每个 $x \in \text{supp}(P_{\text{obs},X})$。类似地，函数 h 对每个满足 $Q_X \ll P_{\text{obs},X}$ 且上确界$_{x\in\mathcal{X}}\ dP_{\text{obs},X}(x)\ dQ_X(x) < \infty$ 的 Q_X，都能解决 (30)，当且仅当它能够解决

$$\max_{h(x)\in[0,1]} \inf_{P\in\mathcal{T}(P_{\text{obs}})} \Gamma\mathbb{E}_P\left[v^*(h(x); X, Y(0), Y(1)) \mid X = x\right] + (1-\Gamma)\text{CVaR}_{\zeta(\Gamma)}(v^*(h(x); X, Y(0), Y(1)) \mid x) \tag{43}$$

对每个 $x \in \text{supp}(P_{\text{obs},X})$。 证明见附录 *E.3*。

引理 20。 设 $v^*(z; x, y_0, y_1) = (z - \pi_0(x)) \cdot (y_1 - y_0)$。假设 Q_X 满足 $Q_X \ll P_{\text{obs},X}$ 且上确界$_{x\in\mathcal{X}}\ dP_{\text{obs},X}(x)\ dQ_X(x) < \infty$. 那么，

$$\begin{aligned}&\inf_{Q\in\mathcal{S}^\Gamma(P_{\text{obs}},Q_X)} \mathbb{E}_Q\left[v^*(\pi(X); X, Y(0), Y(1)) \mid X = x\right] \\ &= (\pi(x) - \pi_0(x))_+ g_1(x) + (\pi_0(x) - \pi(x))_+ g_0(x),\end{aligned} \tag{44}$$

对于任意 $x \in \text{supp}(P_{\text{obs},X})$。 证明见附录 *E.4*。

引理 21。 对于 $x \in \text{supp}(P_{\text{obs},X})$，$\min\{g_1(x)\ g_0(x)\} \leq 0$. 此外，如果 $g_1(x) = 0$，那么 $g_0(x) \leq 0$. 证明见附录 *E.5*。

引理 22。 设 P_{obs}为一个观测数据分布，其中 $P_{\text{obs},W|X}$ =伯努利分布(e)对应处理概率 e。设 $Q \in \mathcal{S}^\Gamma(P_{\text{obs}}, Q_X)$。假设在处理概率为 e 的随机对照试验下，潜在结果分布 Q 产生观测数据分布 Q_{obs}. 则，$Q_{\text{obs}} \in \mathcal{S}^\Gamma_{\text{obs}}(P_{\text{obs}}, Q_X)$。 证明见附录 *E.6*。

引理 23。 令 $\zeta(\Gamma) = \frac{1}{\Gamma+1}$. 令 $q_\zeta(v(h(x); X, Y, W) \mid x, w)$ 为 ζ-分位数，属于 $v(h(x); x, Y, w)$ 其中$Y \sim P_{\text{obs},Y|X=x,W=w}$. 求解下式的分布 dQ^*_{obs}

$$\underset{Q_{\text{obs}}\in\mathcal{S}^\Gamma_{\text{obs}}(P_{\text{obs}},Q_X)}{\text{argmin}} \mathbb{E}_{Q_{\text{obs}}}\left[v(h(x); X, Y, W)\right] \tag{45}$$

由 $dQ_X \cdot dP_{\text{obs},W|X} \cdot dQ^*_{\text{obs},Y|X,W}$ 给出，其中

$$dQ^*_{\text{obs},Y|X=x,W=w}(y) = \begin{cases}\Gamma^{-1} \cdot dP_{\text{obs},Y|X=x,W=w}(y) & v(h(x); x, y, w) \geq q_{\zeta(\Gamma)}(v(h(x); X, Y, W) \mid w, x) \\ \Gamma \cdot dP_{\text{obs},Y|X=x,W=w}(y) & \text{o.w.}.\end{cases} \tag{46}$$

附录 *E.7* 中的证明。

引理 24。 设 P、Q 为潜在结果分布。如果 Q 通过 Γ-有偏抽样生成 P，则

$$\Gamma^{-1} \leq \frac{dQ_{Y(w)|X=x}(y)}{dP_{Y(w)|X=x}(y)} \leq \Gamma \quad \forall x \in \mathcal{X}, y \in \mathbb{R}, w \in \{0, 1\} \tag{47}$$

以及 上确界$_{x\in\mathcal{X}}\ dP_{\text{obs},\ X}(x)\ dQ_X\ (x) < \infty$。 证明见附录 *E.8*。

D.3 定理1的证明

定义 $v^*(\pi(X);\ X,\ Y(0)\ Y(1)) = Y(\pi(X)) = \pi(X)\cdot Y(1) + (1-\pi(X))Y(0)$。我们可以应用引理19，由此可知，对于满足 $Q_X\, Q_X \ll P_{\text{obs},\ X}$ 且上确界 $\sup_{x\in\mathcal{X}} dP_{\text{obs},\ X}(x)\, dQ_X(x) < \infty$ 的任意 Q_X，求解（10）等价于求解以下条件最坏情况风险最小化问题

$$\max_{\pi(x)\in\{0,1\}} \inf_{P\in\mathcal{T}(P_{\text{obs}})} \Gamma\mathbb{E}_P\left[Y(\pi(x)) \mid X=x\right] + (1-\Gamma)\text{CVaR}_{\zeta(\Gamma)}(Y(\pi(x)) \mid x) \tag{48}$$

对于每个 $x \in \text{supp}(P_{\text{obs},X})$。

我们注意到

$$\mathbb{E}_P\left[Y(\pi(x)) \mid X=x\right] = \pi(x)\mathbb{E}_P\left[Y(1) \mid X=x\right] + (1-\pi(x))\mathbb{E}_P\left[Y(0) \mid X=x\right],$$

因此，该项在 P_{obs} 下是可识别的，且不依赖于 $P \in \mathcal{T}(P_{\text{obs}})$ 的选择。令 P 为任意一个潜在结果分布 $\mathcal{T}(P_{\text{obs}})$。因此，式（48）可写为

$$\max_{\pi(x)\in\{0,1\}} \Gamma(\pi(x)\mathbb{E}_P\left[Y(1) \mid X=x\right] + (1-\pi(x))\mathbb{E}_P\left[Y(1) \mid X=x\right]) + \inf_{P'\in\mathcal{T}(P_{\text{obs}})} (1-\Gamma)\text{CVaR}_{\zeta(\Gamma)}(Y(\pi(x)) \mid x). \tag{49}$$

化简右侧的项，可得

$$\begin{aligned}
&\inf_{P'\in\mathcal{T}(P_{\text{obs}})} (1-\Gamma)\text{CVaR}_{\zeta(\Gamma)}(Y(\pi(x)) \mid x) && (50)\\
&= (1-\Gamma) \sup_{P'\in\mathcal{T}(P_{\text{obs}})} \text{CVaR}_{\zeta(\Gamma)}(Y(\pi(x)) \mid x) && (51)\\
&= (1-\Gamma) \sup_{P'\in\mathcal{T}(P_{\text{obs}})} \text{CVaR}_{\zeta(\Gamma)}(\pi(x)Y(1) + (1-\pi(x))Y(0) \mid x) && (52)\\
&= (1-\Gamma)\Big(\text{CVaR}_{\zeta(\Gamma)}(\pi(x)Y(1)) + \text{CVaR}_{\zeta(\Gamma)}((1-\pi(x))Y(0))\Big) && (53)\\
&= (1-\Gamma)\Big(\pi(x)\text{CVaR}_{\zeta(\Gamma)}(Y(1) \mid x) + (1-\pi(x))\text{CVaR}_{\zeta(\Gamma)}(Y(0) \mid x)\Big) && (54)\\
&= (1-\Gamma)\Big(\pi(x)c_1^+(x) + (1-\pi(x))c_0^+(x)\Big). && (55)
\end{aligned}$$

（53）由引理15得出。（54）由引理14得出。

将上述表达式代入（49）可得

$$\max_{\pi(x)\in\{0,1\}} \pi(x)(\Gamma\mathbb{E}_P\left[Y(1) \mid X=x\right] + (1-\Gamma)c_1^+(x)) + (1-\pi(x))(\Gamma\mathbb{E}_P\left[Y(0) \mid X=x\right] + (1-\Gamma)c_0^+(x)). \tag{56}$$

根据上述最坏情况风险的形式，最优策略等于

$$\mathbb{I}(\Gamma\mathbb{E}_P\left[Y(1) \mid X=x\right] + (1-\Gamma)c_1^+(x) \geq \Gamma\mathbb{E}_P\left[Y(0) \mid X=x\right] + (1-\Gamma)c_0^+(x)).$$

重新整理各项，可得该策略等于

$$\mathbb{I}(\tau(x) \geq (1-\Gamma^{-1})(c_1^+(x) - c_0^+(x))).$$

注意，这等于（17）。

D.4 引理2的证明

我们有

$$\begin{aligned}
\inf_{Q\in\mathcal{S}^\Gamma(P_{\text{obs}},Q_X)} \mathbb{E}_Q\left[Y(w) \mid X=x\right] &= \inf_{P\in\mathcal{T}(P_{\text{obs}})} \inf_{Q\in\mathcal{R}^\Gamma(P,Q_X)} \mathbb{E}_Q\left[Y(w) \mid X=x\right] && (57)\\
&= \inf_{P\in\mathcal{T}(P_{\text{obs}})} \Gamma\mathbb{E}_P\left[Y(w) \mid X=x\right] + (1-\Gamma)\text{CVaR}_{\zeta(\Gamma)}(Y(1) \mid X=x) && (58)\\
&= \Gamma\mathbb{E}_P\left[Y(w) \mid X=x\right] + (1-\Gamma)c_w^+(x), && (59)
\end{aligned}$$

其中 P 是 $\mathcal{T}(P_{\text{obs}})$ 中任意一个研究潜在结果分布。在上述推导中，（58）由引理 19推出，而（59）由引理 15推出。我们考察（18）中的策略，并发现它可以写成如下形式

$$\begin{aligned}
&\mathbb{I}\left(\inf_{Q\in\mathcal{S}^\Gamma(P_{\text{obs}},Q_X)}\mathbb{E}_Q\left[Y(1)\mid X=x\right]\geq\inf_{Q\in\mathcal{S}^\Gamma(P_{\text{obs}},Q_X)}\mathbb{E}_Q\left[Y(0)\mid X=x\right]\right)\\
&=\mathbb{I}\left(\Gamma\mathbb{E}_P\left[Y(1)\mid X=x\right]+(1-\Gamma)c_1^+(x)\geq\Gamma\mathbb{E}_P\left[Y(0)\mid X=x\right]+(1-\Gamma)c_0^+(x)\right)\\
&=\mathbb{I}(\Gamma\tau(x)\geq(\Gamma-1)(c_1^+(x)-c_0^+(x)))\\
&=\mathbb{I}(\tau(x)\geq(1-\Gamma^{-1})(c_1^+(x)-c_0^+(x)))\\
&=\mathbb{I}(\tau(x)\geq H_\Gamma(x))\\
&=\pi^*_{\Gamma,\text{maxmin}}(x).
\end{aligned}$$

因此，我们得出结论：在稳健性集合 $\mathcal{S}^\Gamma(P_{\text{obs}},Q_X)$ 上比较 $\mathbb{E}_Q\left[Y(1)\mid X=x\right]$ 和 $\mathbb{E}_Q\left[Y(0)\mid X=x\right]$ 的下界的策略，就是最大最小策略 $\pi^*_{\Gamma,\text{maxmin}}$.

D.5 定理 3 的证明

注意到 $\mathbb{E}_Q\left[Y(\quad\pi(X))\right]-\mathbb{E}_Q\left[Y(\pi_0(X))\right]=\mathbb{E}_Q\left[(\pi(X)-\pi_0(X))\cdot(Y(1)-Y(0))\right]$,因此 (11) 等价于 o

$$\sup_{\pi\in\Pi}\inf_{Q\in\mathcal{S}^\Gamma(P_{\text{obs}},Q_X)}\mathbb{E}_Q\left[v^*(\pi(X);X,Y(0),Y(1))\right]$$

对于 $v^*(z;x,y_0,y_1)=(z-\pi_0(x))\cdot(y_1-y_0)$。对满足 $Q_X\ll P_{\text{obs},X}$ 且 $\sup_{x\in\mathcal{X}}dP_{\text{obs},X}(x)\,dQ_X(x)<\infty$ 的每个 Q_X 求解上~~述问题，~~等价于对 $\text{supp}(P_{\text{obs},X})$ 中的每个 $x\in$ 求解条件最坏情形收益问题（(44) 的左侧）。因此，根据引理 20 可知，求解 (11) 的策略 π 对 $\text{supp}(P_{\text{obs},X})$ 中的每个 x 都能求解 (44) 的右侧。

首先，我们考虑 $\pi_0(x)$ 的情况。$=0$. 在这种情况下，（44）可化简为

$$\max_{\pi(x)\in\{0,1\}}(\pi(x))_+(\Gamma\tau(x)+(1-\Gamma)(c_1^+(x)+c_0^-(x))).$$

因此，当 $\pi_0(x)=0$ 时，最优策略等价于 $\mathbb{I}(\Gamma\tau(x)+(1-\Gamma)(c_1^+(x)+c_0^-(x))\geq 0)$。重新整理各项后，这等价于 $\mathbb{I}(\tau(x)\geq H_\Gamma^+(x))$。

其次，我们考虑 $\pi_0(x)$ 的情况 $=1$. 在这种情况下，（44）可化简为

$$\max_{\pi(x)\in\{0,1\}}(1-\pi(x))_+(-\Gamma\tau(x)+(1-\Gamma)(c_0^+(x)+c_1^-(x))).$$

因此，当 $\pi_0(x)=1$ 时，最优策略等价于 $\mathbb{I}(-\Gamma\tau(x)+(1-\Gamma)(c_0^+(x)+c_1^-(x))\leq 0)$。重新整理各项后，这等价于 $\mathbb{I}(\tau(x)\geq H_\Gamma^-(x))$。

结合这两个结果，可得到（21）中的策略。

D.6 引理 4的证明

令 $q_\zeta(Z)$ 表示随机变量 Z 的 ζ 分位数。注意，对于随机变量 Z，$q_\zeta(Z)=-q_{1-\zeta}(-Z)$。

我们注意到

$$\begin{aligned}
-c_1^-(x)&=-\mathbb{E}_P\left[-Y(1)\mid -Y(1)\geq q_{\zeta(\Gamma)}(-Y(1)\mid x),X=x\right]\\
&=\mathbb{E}_P\left[Y(1)\mid Y(1)\leq -q_{\zeta(\Gamma)}(-Y(1)\mid x),X=x\right]\\
&=\mathbb{E}_P\left[Y(1)\mid Y(1)\leq q_{1-\zeta(\Gamma)}(Y(1)\mid x),X=x\right]\\
&\leq\mathbb{E}_P\left[Y(1)\mid Y(1)\geq q_{\zeta(\Gamma)}(Y(1)\mid x),X=x\right]\\
&=c_1^+(x).
\end{aligned}$$

这意味着

$$\begin{aligned} H_\Gamma^-(x) &= (1-\Gamma^{-1})\cdot(-c_1^-(x)-c_0^+(x)) \\ &\le (1-\Gamma^{-1})\cdot(c_1^+(x)-c_0^+(x)) \\ &= H_\Gamma(x). \end{aligned}$$

此外，我们有

$$\begin{aligned} -c_0^+(x) &= -\mathbb{E}_P\left[Y(0)\mid Y(0)\ge q_{\zeta(\Gamma)}(Y(0)\mid x), X=x\right] \\ &= \mathbb{E}_P\left[-Y(0)\mid -Y(0)\le -q_{\zeta(\Gamma)}(Y(0)\mid x), X=x\right] \\ &= \mathbb{E}_P\left[-Y(0)\mid -Y(0)\le q_{1-\zeta(\Gamma)}(-Y(0)\mid x), X=x\right] \\ &\le \mathbb{E}_P\left[-Y(0)\mid -Y(0)\ge q_{\zeta(\Gamma)}(-Y(0)\mid x), X=x\right] \\ &= c_0^-(x). \end{aligned}$$

这意味着

$$\begin{aligned} H_\Gamma(x) &= (1-\Gamma^{-1})(c_1^+(x)-c_0^+(x)) \\ &\le (1-\Gamma^{-1})(c_1^+(x)+c_0^-(x)) \\ &= H_\Gamma^+(x). \end{aligned}$$

结合这两个结果可得 $H_\Gamma^-(x)\le H_\Gamma(x)\le H_\Gamma^+(x)$。

定理5的证明

我们注意到，式（12）可表示如下。

$$\begin{aligned} \inf_{\pi\in\Pi}\sup_{Q\in\mathcal{S}^\Gamma(P_{\text{obs}},Q_X)} R_Q(\pi(X)) &= \inf_{\pi\in\Pi}\sup_{Q\in\mathcal{S}^\Gamma(P_{\text{obs}},Q_X)}\sup_{\pi'}\mathbb{E}_Q\left[Y(\pi'(X))\right]-\mathbb{E}_Q\left[Y(\pi(X))\right] \\ &= \sup_{\pi\in\Pi}\inf_{Q\in\mathcal{S}^\Gamma(P_{\text{obs}},Q_X)}\mathbb{E}_Q\left[Y(\pi(X))\right]-\inf_{\pi'\in\Pi}\mathbb{E}_Q\left[Y(\pi'(X))\right] \\ &= \sup_{\pi\in\Pi}\inf_{Q\in\mathcal{S}^\Gamma(P_{\text{obs}},Q_X)}\inf_{\pi'\in\Pi}\mathbb{E}_Q\left[Y(\pi(X))\right]-\mathbb{E}_Q\left[Y(\pi'(X))\right] \\ &= \sup_{\pi\in\Pi}\inf_{\pi'\in\Pi}\inf_{Q\in\mathcal{S}^\Gamma(P_{\text{obs}},Q_X)}\mathbb{E}_Q\left[Y(\pi(X))\right]-\mathbb{E}_Q\left[Y(\pi'(X))\right]. \end{aligned}$$

我们注意到，对于任意满足 Q_X 且 $Q_X\ll P_{\text{obs},\,X}$ 且 $\sup_{x\in\mathcal{X}} dP_{\text{obs},\,X}(x)\,dQ_X(x)<\infty$ 的 $x\in$，求解上述问题等价于对每个 $x\in\text{supp}(P_{\text{obs},\,X})$求解条件问题。

$$\sup_{\pi(x)\in\{0,1\}}\inf_{\pi'(x)\in\{0,1\}}\inf_{Q\in\mathcal{S}^\Gamma(P_{\text{obs}},Q_X)}\mathbb{E}_Q\left[Y(\pi(X))-Y(\pi'(X))\mid X=x\right]. \tag{60}$$

由于 $\mathbb{E}_Q\left[Y(\pi(X))\right]-\mathbb{E}_Q\left[Y(\pi'(X))\right]=\mathbb{E}_Q\left[(\pi(X)-\pi'(X))\cdot(Y(1)-Y(0))\right]$，我们可以将引理20 应用于（60），从而得到

$$\sup_{\pi(x)\in\{0,1\}}\inf_{\pi'(x)\in\{0,1\}}(\pi(x)-\pi'(x))_+g_1(x)+(\pi'(x)-\pi(x))_+g_0(x). \tag{61}$$

当 $\pi(x)=1$ 时，目标函数可化简为 $(1-\pi'(x))g_1(x)$，因此对 π' 的下确界为 $\min\{g_1(x),0\}=-(g_1(x))_-$。类似地，当 $\pi(x)=0$ 时，对 π' 的下确界为 $-(g_0(x))_-$。因此，最优策略 π 为 $\mathbb{I}(-(g_1(x))_-\ge-(g_0(x))_-)$。

为简化该策略的形式，我们考虑以下四种情况：(a) $g_1(x)\le 0$ 和 $g_0(x)\le 0$，(b)$g_1(x)\le 0$ 和 $g_0(x)>0$，(c) $g_1(x)>0$ 和 $g_0(x)\le 0$，以及 (d) $g_1(x)>0$ 和 $g_0(x)>0$.。根据引理 21，我们知道无需考虑情况 (d)，因为 $\min(g_1(x),g_0(x))\le 0$。此外，情况 (b) 中 $g_1(x)=0$ 和 $g_0(x)>0$ 的部分也无需考虑，因为这种情况不可能发生。

我们考虑情况 (a)，其中 $g_1(x) \le 0$ 和 $g_0(x) \le 0.$。那么

$$\begin{aligned}\pi(x) &= \mathbb{I}(-(g_1(x))_- \ge -(g_0(x))_-) \\ &= \mathbb{I}(g_1(x) \ge g_0(x)).\end{aligned}$$

在情况 (b) 中，我们只需考虑 $g_1(x) < 0$ 和 $g_0(x) > 0$ （因为根据引理 21 $g_1(x) = 0$, $g_0(x) > 0$不会发生）。因此，我们有

$$\begin{aligned}\pi(x) &= \mathbb{I}(-(g_1(x))_- \ge -(g_0(x))_-) \\ &= \mathbb{I}(g_1(x) \ge 0) \\ &= \mathbb{I}(g_1(x) \ge g_0(x)).\end{aligned}$$

在情况 (c) 中，其中 $g_1(x) > 0$ 和 $g_0(x) \le 0$。则

$$\begin{aligned}\pi(x) &= \mathbb{I}(-(g_1(x))_- \ge -(g_0(x))_-) \\ &= \mathbb{I}(0 \ge g_0(x)) \\ &= \mathbb{I}(g_1(x) \ge g_0(x)).\end{aligned}$$

根据引理 21，我们注意到情况 (d) 不会发生。

结合这些结果，我们有 $\pi(x) = \mathbb{I}(g_1(x) \ge g_0(x))$。我们可以将 $\mathbb{I}(g_1(x) \ge g_0(x))$ 表示为关于条件平均处理效应的阈值规则，如下所示。

$$\begin{aligned}\mathbb{I}(g_1(x) \ge g_0(x)) &= \mathbb{I}(\Gamma\tau(x) + (1-\Gamma)(c_1^+(x) + c_0^-(x)) \ge -\Gamma\tau(x) + (1-\Gamma)(c_0^+(x) + c_1^-(x))) \\ &= \mathbb{I}\left(2\Gamma\tau(x) \ge (1-\Gamma)\Big(c_0^+(x) + c_1^-(x) - c_1^+(x) - c_0^-(x)\Big)\right) \\ &= \mathbb{I}\Big(\tau(x) \ge \frac{1}{2}(1-\Gamma^{-1}) \cdot (c_1^+(x) + c_0^-(x) - c_0^+(x) - c_1^-(x))\Big) \\ &= \mathbb{I}\Big(\tau(x) \ge \frac{H_\Gamma^+(x) + H_\Gamma^-(x)}{2}\Big).\end{aligned}$$

D.8 定理 6 的证明

根据引理 19，如果对于任意满足 Q_X 使得 π 求解 (25)，且 $Q_X \ll P_{\text{obs},\ X}$ 以及 $\sup_{x\in\mathcal{X}} \frac{dP_{\text{obs},\ X}(x)}{dQ_X(x)} < \infty$，那么它求解

$$\max_{\pi(x)\in\{0,1\}} \inf_{P\in\mathcal{T}(P_{\text{obs}})} \Gamma\mathbb{E}_P\left[\tilde{v}_{\text{regret}}(\pi(x);\ X, Y(0)\ Y(1)) \mid X = x\right] + (1-\Gamma)\text{条件风险价值}_{\zeta(\Gamma)}(\tilde{v}_{\text{regret}}(\pi(x);\ X, Y(0)\ Y(1)) \mid x). \tag{62}$$

对于每个 $x \in \text{supp}(P_{\text{obs},\ X})$。

我们注意到，

$$\begin{aligned}\mathbb{E}_P\left[\tilde{v}_{\text{regret}}(\pi(x); X, Y(0), Y(1)) \mid X = x\right] &= (2\pi(x) - 1)\tau(x) \\ &= (2\pi(x) - 1)_+\tau(x) + (-2\pi(x) + 1)_+(-\tau(x)),\end{aligned}$$

且在 P_{obs} 下可识别，并且不依赖于 $P \in \mathcal{T}(P_{\text{obs}})$ 的选择。

此外，我们有

$$\inf_{P\in\mathcal{T}(P_{\text{obs}})}(1-\Gamma)\text{CVaR}_{\zeta(\Gamma)}(\tilde{v}_{\text{regret}}(\pi(x);X,Y(0),Y(1))\mid x) \tag{63}$$

$$=(1-\Gamma)\sup_{P\in\mathcal{T}(P_{\text{obs}})}\text{条件风险价值}_{\zeta(\Gamma)}(\tilde{v}_{\text{regret}}(\pi(x);X,\ Y(0)\ Y(1))\mid x) \tag{64}$$

$$=(1-\Gamma)\sup_{P\in\mathcal{T}(P_{\text{obs}})}\text{CVaR}_{\zeta(\Gamma)}((2\pi(x)-1)(Y(1)-Y(0))\mid x) \tag{65}$$

$$=(1-\Gamma)\sup_{P\in\mathcal{T}(\ P_{\text{obs}})}(2\pi(x)-1)_+\text{CVaR}_{\zeta(\Gamma)}(Y(1)-Y(0)\mid x)+(-2\pi(x)+1)_+\text{CVaR}_{\zeta(\Gamma)}(Y(0)-Y(1)\mid x) \tag{66}$$

$$=(1-\Gamma)\Big((2\pi(x)-1)_+(c_1^+(x)+c_0^-(x))+(-2\pi(x)+1)_+(c_0^+(x)+c_1^-(x))\Big). \tag{67}$$

（66）由引理 14推出，而（67）由引理 15推出。

因此，我们可以将（62）写成

$$\max_{\pi(x)\in\{0,1\}}(2\pi(x)-1)_+(\Gamma\tau(x)+(1-\Gamma)(c_1^+(x)+c_0^-(x)))+(-2\pi(x)+1)_+(-\Gamma\tau(x)+(1-\Gamma)(c_0^+(x)+c_1^-(x))). \tag{68}$$

由于最大化上述表达式的 $\pi(x)$ 必须为 0 或 1，我们可知，当 $\Gamma\tau(x)+(1-\Gamma)(c_1^+(x)+c_0^-(x)))\geq(-\Gamma\tau(x)+(1-\Gamma)(c_0^+(x)+c_1^-(x)))$ 时，$\pi(x)=1$；否则为 0。因此，最大化上述表达式的 π 等于 (23)。

D.9 定理7的证明

我们旨在证明，任何求解 (28) 的策略也能求解 (26)。引理 22表明，对于任意 $Q\in\mathcal{S}^\Gamma(P_{\text{obs}},Q_X)$，存在一个 $Q_{\text{obs}}\in\mathcal{S}^\Gamma(P_{\text{obs}},Q_X)$，其中 Q 在随机对照试验下与 Q_{obs} 一致。因此，根据假设 1，我们有

$$\inf_{Q\in\mathcal{S}^\Gamma(P_{\text{obs}},Q_X)}\mathbb{E}_Q\left[v^*(\pi(X);X,Y(0),Y(1))\right]\geq\inf_{Q\in\mathcal{S}^\Gamma_{\text{obs}}(P_{\text{obs}},Q_X)}\mathbb{E}_{Q_{\text{obs}}}\left[v(\pi(X);X,Y,W)\right]. \tag{69}$$

由此可得

$$\sup_{\pi\in\Pi}\inf_{Q\in\mathcal{S}^\Gamma(P_{\text{obs}},Q_X)}\mathbb{E}_Q\left[v^*(\pi(X);X,Y(0),Y(1))\right]\geq\sup_{\pi\in\Pi}\inf_{Q\in\mathcal{S}^\Gamma_{\text{obs}}(P_{\text{obs}},Q_X)}\mathbb{E}_{Q_{\text{obs}}}\left[v(\pi(X);X,Y,W)\right]. \tag{70}$$

还需证明

$$\sup_{\pi\in\Pi}\inf_{Q\in\mathcal{S}^\Gamma(P_{\text{obs}},Q_X)}\mathbb{E}_Q\left[v^*(\pi(X);X,Y(0),Y(1))\right]\leq\sup_{\pi\in\Pi}\inf_{Q\in\mathcal{S}^\Gamma_{\text{obs}}(P_{\text{obs}},Q_X)}\mathbb{E}_{Q_{\text{obs}}}\left[v(\pi(X);X,Y,W)\right]. \tag{71}$$

事实上，若将 (69) 的左端项和右端项分别记为 $f_L(\pi)$ 和 $f_R(\pi)$，则对每个 $\pi\in\Pi$ 都有 $f_L(\pi)\geq f_R(\pi)$。令 π_R^* 为 $f_R(\pi)$ 的任意最大化者。则由 (70) 和 (71) 可得 $\sup_{\pi\in\Pi}f_L(\pi)=f_R(\pi_R^*)$。若存在 $\pi\in\Pi$ 使得 $f_L(\pi)>f_L(\pi_R^*)$，则

$$f_L(\pi)>f_L(\pi_R^*)\geq f_R(\pi_R^*)=\sup_{\pi\in\Pi}f_L(\pi),$$

这就导出了矛盾。因此，π_R^* 必定是 $f_L(\pi)$的最大化点。

Let

$$\mathcal{A}_{x,w}=\{y\in\mathbb{R}\mid v(\pi(x);x,y,w)\geq q_{\zeta(\Gamma)}(v(\pi(x);X,Y,W)\mid x,w)\}.$$

回顾引理23，$q_{\zeta(\Gamma)}(v(\pi(X);\ X,Y,W)\ \mid x,w)$是 $v(\pi(x);\ x,Y,w)$的 $\zeta(\Gamma)$分位数，其中$Y\sim P_{\text{obs},Y|X=x,W=w}$。由于$P_{\text{obs},Y|X=x,W=w}$相对于勒贝格测度绝对连续，因此我们必有

$$\mathbb{P}_{P_{\text{obs},Y|X,W=w}}\left[Y\in\mathcal{A}_{X,w}\mid X=x\right]=\frac{\Gamma}{\Gamma+1}\quad\forall x\in\mathcal{X},w\in\{0,1\}. \tag{72}$$

根据引理 23，在（71）右侧达到下确界的分布 dQ^*_{obs}由 $dQ_X\cdot dP_{\text{obs}、\ W|X}\cdot dQ^*_{\text{obs}、\ Y|X、\ W}$ 给出，其中 $dQ^*_{\text{obs}、\ Y|X、\ W}=\theta^{(W)}\ (Y\mid X)\cdot dP_{\text{obs}、\ Y|X、\ W}$以及

$$\theta^{(w)}(y\mid x)=\Gamma^{-1}\cdot\mathbb{I}(y\in\mathcal{A}_{x,w})+\Gamma\cdot\mathbb{I}(y\notin\mathcal{A}_{x,w}). \tag{73}$$

为证明（71），我们证明存在一个潜在结果分布 $Q \in \mathcal{S}_\Gamma$（P_{obs}、Q_X），使其对所有 π 都产生与 Q^*_{obs} 相同的最坏情形值。在假设1下，只需证明存在一个与 Q^*_{obs} 在治疗倾向概率为 e 的随机对照试验下相一致的$Q \in \mathcal{S}_\Gamma$（P_{obs}、Q_X）。为此，我们首先构造一个与 P_{obs} 在随机对照试验下一致的潜在结果分布 P，并使其具有以下特殊性质：对于（X、Y（0） Y（1））$\sim P$，$\theta^{(0)}$（Y（0）$\mid X$）$= \theta^{(1)}$（Y（1）$\mid X$）。随后，我们根据 P 定义 Q，并检验 $Q \in \mathcal{S}^\Gamma$（P_{obs}、Q_X）。

我们按如下方式构造 P。首先抽取 $X \sim P_X$ 和 $Y(0) \sim P_{\text{obs},\, Y|X,\, W=0}$。然后按如下方式抽取 $Y(1)$。

- 如果 $Y(0) \in \mathcal{A}_{X,\,0}$，则抽取
$$Y(1) \sim P_{\text{obs},Y|X,W=1,Y\in\mathcal{A}_{X,1}}.$$
- 如果 $Y(0) \notin \mathcal{A}_{X,\,0}$，则抽取
$$Y(1) \sim P_{\text{obs},Y|X,W=1,Y\notin\mathcal{A}_{X,1}}.$$

现在我们证明$P_{(X,Y(0)\ Y(1))}$ 与 P_{obs}一致。根据构造，

$$P_X = P_{\text{obs},X}, \quad P_{Y(0)|X} = P_{\text{obs},Y|X,W=0},$$

and

$$\begin{aligned}
P_{Y(1)|X} &= \mathbb{P}\left[Y(0) \in \mathcal{A}_{X,0}\right] P_{\text{obs},Y|X,W=1,Y\in\mathcal{A}_{X,1}} + \mathbb{P}\left[Y(0) \notin \mathcal{A}_{X,0}\right] P_{\text{obs},Y|X,W=1,Y\notin\mathcal{A}_{X,1}} \\
&= \mathbb{P}_{P_{\text{obs},Y|X,W=0}}\left[Y \in \mathcal{A}_{X,0}\right] P_{\text{obs},Y|X,W=1,Y\in\mathcal{A}_{X,1}} + \mathbb{P}_{P_{\text{obs},Y|X,W=0}}\left[Y \notin \mathcal{A}_{X,0}\right] P_{\text{obs},Y|X,W=1,Y\notin\mathcal{A}_{X,1}} \\
&= \mathbb{P}_{P_{\text{obs},Y|X,W=1}}\left[Y \in \mathcal{A}_{X,1}\right] P_{\text{obs},Y|X,W=1,Y\in\mathcal{A}_{X,1}} + \mathbb{P}_{P_{\text{obs},Y|X,W=1}}\left[Y \notin \mathcal{A}_{X,1}\right] P_{\text{obs},Y|X,W=1,Y\notin\mathcal{A}_{X,1}} \\
&= P_{\text{obs},Y|X,W=1},
\end{aligned}$$

其中适用（72）。因此， P 在 P_{obs} 随机对照试验下是一致的。

给定这样的潜在结果分布 P，我们可以定义一个潜在结果分布 Q ，其协变量分布等于$Q^*_{\text{obs},\ X}$ 和

$$\frac{dQ_{Y(0),Y(1)|X=x}(y_0,y_1)}{dP_{Y(0),Y(1)|X=x}(y_0,y_1)} = \theta^{(0)}(y_0 \mid x). \tag{74}$$

我们可以证明， Q 由（74）定义的对于任意 x都是概率测度。

$$\begin{aligned}
\int dQ_{Y(0),Y(1)|X=x} &= \int \theta^{(0)}(y_0 \mid x) dP_{Y(0),Y(1)|X=x}(y_0,y_1) \\
&= \int \theta^{(0)}(y_0 \mid x) dP_{Y(0)|X=x}(y_0) \\
&= \int \theta^{(0)}(y_0 \mid x) dP_{\text{obs},Y|X=x,W=0}(y_0) \\
&= \int dQ^*_{\text{obs},Y|X=x,W=0}(y_0) \\
&= 1.
\end{aligned}$$

In add此外，由于在 P的支撑集上， $\theta^{(0)}(Y(0) \mid X) = \theta^{(1)}(Y(1) \mid X)$，我们还有

$$\frac{dQ_{Y(0),Y(1)|X=x}(y_0,y_1)}{dP_{Y(0),Y(1)|X=x}(y_0,y_1)} = \theta^{(0)}(y_0 \mid x) = \theta^{(1)}(y_1 \mid x). \tag{75}$$

分别对 $Y(0)$ 和 $Y(1)$ 进行边际化，我们得到

$$\frac{dQ_{Y(0)|X=x}(y_0)}{dP_{Y(0)|X=x}(y_0)} = \theta^{(0)}(y_0 \mid x), \quad \frac{dQ_{Y(1)|X=x}(y_1)}{dP_{Y(1)|X=x}(y_1)} = \theta^{(1)}(y_1 \mid x).$$

由于 $P_{Y(w)|X=x} = P_{\text{obs},\ Y|X=x,\ W=w}$，（73）意味着

$$Q_{Y(w)|X=x} = Q^*_{\text{obs},Y|X=x,W=w}.$$

这表明，在随机对照试验下，Q 与 Q_{obs} 一致。

此外，我们可以验证 $Q \in \mathcal{S}^\Gamma(P_{\text{obs}},\ Q_X)$。我们已经证明，在随机对照试验下，$P$与$P_{\text{obs}}$ 一致；并且可以验证，Q 通过 Γ有偏抽样生成 P。对于所有 $x \in \mathcal{X}$和 $y_0,\ y_1 \in \mathbb{R}$，

$$\frac{dQ_{Y(0),Y(1)|X=x}(y_0,y_1)}{dP_{Y(0),Y(1)|X=x}(y_0,y_1)} = \theta^{(0)}(y_0 \mid x) \tag{76}$$

$$= \frac{dQ^*_{\text{obs},Y|X=x,W=0}(y_0)}{dP_{\text{obs},Y|X=x,W=0}(y_0)} \tag{77}$$

$$\in [\Gamma^{-1},\Gamma]. \tag{78}$$

（76）由（75）推出。（77）根据（ $\theta^{(w)}(Y \mid X)$ 在（73）中的定义成立。最后，（78）成立，因为$Q^*_{\text{obs}} \in \mathcal{S}^\Gamma_{\text{obs}}$（$P_{\text{obs}},\ Q_X$）。根据假设 1，该结果意味着存在一个分布 $Q \in \mathcal{S}^\Gamma$（$P_{\text{obs}},\ Q_X$），它对所有 h 都能产生与 Q^*_{obs} 相同的策略值。因此，我们必须有

$$\inf_{Q\in\mathcal{S}^\Gamma(P_{\text{obs}},Q_X)} \mathbb{E}_Q\left[v^*(\pi(X);X,Y(0),Y(1))\right] \le \mathbb{E}_{Q^*_{\text{obs}}}\left[v(\pi(X);X,Y,W)\right]$$
$$= \inf_{Q_{\text{obs}}\in\mathcal{S}^\Gamma_{\text{obs}}(P_{\text{obs}},Q_X)} \mathbb{E}_{Q_{\text{obs}}}\left[v(\pi(X);X,Y,W)\right].$$

因此，我们必须有（71）。最后，（70）和（71）得出所需结论。

D.10 定理 8 的证明

我们注意到 th 在任何 h 满足（28）的解都能解决条件最坏情况价值最大化问题

$$\sup_{h(x)\in[0,1]} \inf_{Q_{\text{obs}}\in\mathcal{S}^\Gamma_{\text{obs}}(P_{\text{obs}},Q_X)} \mathbb{E}_{Q_{\text{obs}}}\left[v(h(x);X,Y,W) \mid X=x\right] \tag{79}$$

对于每个 $x \in \text{supp}(P_{\text{obs},X})$。

根据引理 23，我们可以将（79）写成等价形式

$$\sup_{h(x)\in[0,1]} \sum_{w\in\{0,1\}} \Gamma\mathbb{P}\left[W=w \mid X=x\right]\cdot\mathbb{E}_{P_{\text{obs}}}\left[v(h(x);X,Y,W) \mid X=x,W=w\right]$$
$$+ (\Gamma^{-1}-\Gamma)\mathbb{E}_{P_{\text{obs}}}\left[v(h(x);X,Y,W)\mathbb{I}(v(h(x);X,Y,W) \ge q_{\zeta(\Gamma)}(v(h(x);X,Y,W) \mid x,w)) \mid X=x,W=w\right]. \tag{80}$$

我们注意到

$$\mathbb{E}_{P_{\text{obs}}}\left[v(h(x);X,Y,W)\mathbb{I}(v(h(x);X,Y,W) \ge q_{\zeta(\Gamma)}(v(h(x);X,Y,W) \mid x,w)) \mid X=x,W=w\right]$$
$$= (1-\zeta(\Gamma))\cdot \text{CVaR}_{\zeta(\Gamma)}(v(h(x);X,Y,W) \mid X=x,W=w).$$

因此，我们有（80）可以写成

$$\sup_{h(x)\in[0,1]} \sum_{w\in\{0,1\}} \mathbb{P}\left[W=w \mid X=x\right]\Big(\Gamma\mathbb{E}_{P_{\text{obs}}}\left[v(h(x);X,Y,W) \mid X=x,W=w\right]$$
$$+ (1-\Gamma)\cdot \text{CVaR}_{\zeta(\Gamma)}(v(h(x);X,Y,W) \mid x,w)\Big). \tag{81}$$

我们可以应用引理14将此表达式改写为

$$\sup_{h(x)\in[0,1]} \sum_{w\in\{0,1\}} \mathbb{P}\left[W=w \mid X=x\right]\Big(\Gamma^{-1}\mathbb{E}_{P_{\text{obs}}}\left[v(h(x);X,Y,W) \mid X=x,W=w\right]$$
$$+ (\Gamma^{-1}-1)\cdot \text{CVaR}_{1-\zeta(\Gamma)}(-v(h(x);X,Y,W) \mid x,w)\Big). \tag{82}$$

我们可以将该问题转换为如下最小化问题。

$$\inf_{h(x)\in[0,1]} \sum_{w\in\{0,1\}} \mathbb{P}\left[W=w \mid X=x\right]\Big(\Gamma^{-1}\mathbb{E}_{P_{\text{obs}}}\left[-v(h(x);X,Y,W) \mid X=x, W=w\right] \\ + (1-\Gamma^{-1})\cdot \text{CVaR}_{1-\zeta(\Gamma)}(-v(h(x);X,Y,W) \mid X=x, W=w)\Big) \tag{83}$$

我们可以应用 Rockafellar 和 Uryasev [2000]，由此可知（83）可以写成

$$\begin{aligned}&\inf_{h(x)\in[0,1]} \Gamma^{-1}\mathbb{E}_{P_{\text{obs}}}\left[-v(h(x);X,Y,W) \mid X=x\right] \\ &\quad + e\cdot(1-\Gamma^{-1})\cdot \inf_{\alpha(x,1)\in\mathbb{R}} \Big(\alpha(x,1)+\zeta(\Gamma)^{-1}\mathbb{E}_{P_{\text{obs}}}\left[(-v(h(x);X,Y,W)-\alpha(x,1))_+ \mid X=x, W=1\right]\Big) \\ &\quad + (1-e)\cdot(1-\Gamma^{-1})\cdot \inf_{\alpha(x,0)\in\mathbb{R}} \Big(\alpha(x,0)+\zeta(\Gamma)^{-1}\mathbb{E}_{P_{\text{obs}}}\left[(-v(h(x);X,Y,W)-\alpha(x,0))_+ \mid X=x, W=0\right]\Big).\end{aligned} \tag{84}$$

因此，我们得到联合最小化问题

$$\begin{aligned}&\inf_{h(x)\in[0,1],\alpha(x,w)\in\mathbb{R}} \Gamma^{-1}\mathbb{E}_{P_{\text{obs}}}\left[-v(h(x);X,Y,W) \mid X=x\right] \\ &\quad + e\cdot(1-\Gamma^{-1})\cdot \Big(\alpha(x,1)+\zeta(\Gamma)^{-1}\mathbb{E}_{P_{\text{obs}}}\left[(-v(h(x);X,Y,W)-\alpha(x,1))_+ \mid X=x, W=1\right]\Big) \\ &\quad + (1-e)\cdot(1-\Gamma^{-1})\cdot \Big(\alpha(x,0)+\zeta(\Gamma)^{-1}\mathbb{E}_{P_{\text{obs}}}\left[(-v(h(x);X,Y,W)-\alpha(x,0))_+ \mid X=x, W=0\right]\Big).\end{aligned} \tag{85}$$

这类似于条件 RU 回归问题。

$$\begin{aligned}\inf_{h(x)\in[0,1],\alpha(x,w)\in\mathbb{R}} & e\mathbb{E}_{P_{\text{obs}}}\left[L_{\text{RU}}(h(X),\alpha(X,W);X,Y,W) \mid X=x, W=1\right] \\ & + (1-e)\mathbb{E}_{P_{\text{obs}}}\left[L_{\text{RU}}(h(X),\alpha(X,W);X,Y,W) \mid X=x, W=0\right].\end{aligned} \tag{86}$$

最后，任意函数 h，$\alpha(\cdot,1)$ $\alpha(\cdot,0)$，对于每个$x\in\text{supp}(P_{\text{obs},X})$，若能解决条件 RU 回归问题 (86)，则必须解决 (32)。因此，(32) 的任意解都是 (28) 的解。

D.11 定理 10 的证明

根据引理 19，如果 h 对于每个满足 Q_X 的 $Q_X \ll P_{\text{obs},X}$ 以及 $\sup_{x\in\mathcal{X}} \frac{dP_{\text{obs},X}(x)}{dQ_X(x)} < \infty$ 都能解决 (33)，那么它就能解决

$$\max_{h(x)\in[0,1]} \inf_{P\in\mathcal{T}(P_{\text{obs}})} \Gamma\mathbb{E}_P\left[v^*_{\text{maxmin}}(h(x);X,Y(0),Y(1)) \mid X=x\right] + (1-\Gamma)\text{CVaR}_{\zeta(\Gamma)}(v^*_{\text{maxmin}}(h(x);X,Y(0),Y(1)) \mid x) \tag{87}$$

对于每个 $x\in\text{supp}(P_{\text{obs},X})$. 我们注意到

$$\begin{aligned}&\mathbb{E}_P\left[v^*_{\text{maxmin}}(h(x);X,Y(0),Y(1)) \mid X=x\right] \\ &= \mathbb{E}_P\left[\log(1+\exp(2h(x)-1))Y(1)+\log(1+\exp(-2h(x)+1))Y(0) \mid X=x\right] \\ &= \log(1+\exp(2h(x)-1))\mathbb{E}_P\left[Y(1) \mid X=x\right]+\log(1+\exp(-2h(x)+1))\mathbb{E}_P\left[Y(0) \mid X=x\right].\end{aligned}$$

因此，(89) 的第一项在 P_{obs} 下是可识别的，并且不依赖于 $P\in\mathcal{T}(P_{\text{obs}})$ 的选择。接下来，应用引理 15 和引理 14，我们得到

$$\begin{aligned}&\inf_{P\in\mathcal{T}(P_{\text{obs}})}(1-\Gamma)\text{CVaR}_{\zeta(\Gamma)}(v^*_{\text{maxmin}}(h(x);X,Y(0),Y(1)) \mid x) \\ &= (1-\Gamma)\sup_{P\in\mathcal{T}(P_{\text{obs}})}\text{CVaR}_{\zeta(\Gamma)}(v^*_{\text{maxmin}}(h(x);X,Y(0),Y(1)) \mid x) \\ &= (1-\Gamma)\sup_{P\in\mathcal{T}(P_{\text{obs}})}\text{CVaR}_{\zeta(\Gamma)}\Big(\log(1+\exp(2h(x)-1))Y(1)+\log(1+\exp(-2h(x)+1))Y(0) \mid x\Big) \\ &= (1-\Gamma)\Big(\log(1+\exp(2h(x)-1))c_1^+(x)+\log(1+\exp(-2h(x)+1))c_0^+(x)\Big).\end{aligned}$$

令 P 为 $\mathcal{T}(P_{\text{obs}})$ 中的一个潜在结果分布。(87) 可以写成

$$\begin{aligned}\max_{h(x)\in[0,1]} &\log(1+\exp(2h(x)-1))\cdot(\Gamma\mathbb{E}_P[Y(1)\mid X=x]+(1-\Gamma)c_1^+(x))\\ &+\log(1+\exp(-2h(x)+1))\cdot(\Gamma\mathbb{E}_P[Y(0)\mid X=x]+(1-\Gamma)c_0^+(x)).\end{aligned}\tag{88}$$

根据引理17，$h^*(x)$ 满足(88)的解满足 $h^*(x)\geq\frac{1}{2}$ 当 $\Gamma\mathbb{E}_P[Y(1)\mid X=x]+(1-\Gamma)c_1^+(x)\geq\Gamma\mathbb{E}_P[Y(0)\mid X=x]+(1-\Gamma)c_0^+(x)$。换言之，$\pi(x)=\mathbb{I}(h^*(x)\geq\frac{1}{2})$ 等价于 $\mathbb{I}(\tau(x)\geq(1-\Gamma^{-1})\cdot(c_1^+(x)-c_0^+(x))$，这与(17)一致。

D.12 定理的证明11

根据引理 19，如果 h 对任意满足 $Q_X\ll P_{\text{obs},X}$ 且 $\sup_{x\in\mathcal{X}}\frac{dP_{\text{obs}}}{dQ_X}(x)<\infty$ 的 Q_X 求解(34)，那么它满足

$$\max_{h(x)\in[0,1]}\inf_{P\in\mathcal{T}(P_{\text{obs}})}\Gamma\mathbb{E}_P\left[v^*_{\text{gain}}(h(x);X,Y(0),Y(1))\mid X=x\right]+(1-\Gamma)\text{CVaR}_{\zeta(\Gamma)}(v^*_{\text{gain}}(h(x);X,Y(0),Y(1))\mid x).\tag{89}$$

对于每个 $x\in\text{supp}(Q_X)$。

我们注意到

$$\mathbb{E}_P\left[v^*_{\text{gain}}(h(x);X,Y(0)\ Y(1))\mid X=x\right]=\mathbb{E}_P[(1-\pi_0(X))\text{ 对数}(1+\text{指数}(2h(x)-1))(Y(1)-Y(0))+\pi_0(X)\text{ 对数}(1+\text{指数}(-2h(x)))(Y(0)-Y(1))\mid X=x]=(1-\pi_0(x))\text{ 对数}(1+\text{指数}(2h(x)-1))\tau(x)+\pi_0(x)\cdot\text{对数}(1+\text{指数}(-2h(x)+1))\cdot(-\tau(x))$$

因此，第一项在 P_{obs} 下是可识别的，并且不依赖于 $P\in\mathcal{T}(P_{\text{obs}})$。接下来，应用引理15 和引理 14 我们得到

$$\begin{aligned}&\inf_{P\in\mathcal{T}(P_{\text{obs}})}(1-\Gamma)\text{CVaR}_{\zeta(\Gamma)}(v^*_{\text{gain}}(h(x);X,Y(0),Y(1))\mid x)\\&=(1-\Gamma)\sup_{P\in\mathcal{T}(P_{\text{obs}})}\text{CVaR}_{\zeta(\Gamma)}(v^*_{\text{gain}}(h(x);X,Y(0),Y(1))\mid x)\\&=(1-\Gamma)\sup_{P\in\mathcal{T}(P_{\text{obs}})}\text{CVaR}_{\zeta(\Gamma)}((1-\pi_0(x))\log(1+\exp(2h(x)-1))(Y(1)-Y(0))\\&\qquad+\pi_0(x)\log(1+\exp(-2h(x)))(Y(0)-Y(1))\mid x)\\&=(1-\Gamma)\Big((1-\pi_0(x))\cdot\log(1+\exp(2h(x)-1))(c_1^+(x)+c_0^-(x))+\pi_0(x)\cdot\log(1+\exp(-2h(x)+1))(c_0^+(x)+c_1^-(x))\Big).\end{aligned}$$

因此，我们可以将（89）重写为

$$\begin{aligned}\max_{h(x)\in[0,1]}&(1-\pi_0(x))\log(1+\exp(2h(x)-1))\cdot\Big(\Gamma\tau(x)+(1-\Gamma)(c_1^+(x)+c_0^-(x))\Big)\\&+\pi_0(x)\log(1+\exp(-2h(x)+1))\Big(-\Gamma\tau(x)+(1-\Gamma)(c_1^-(x)+c_0^+(x))\Big).\end{aligned}\tag{90}$$

当 $\pi_0(x)=0$ 时，上述问题可化简为

$$\max_{h(x)\in[0,1]}\log(1+\exp(2h(x)-1))\cdot\Big(\Gamma\tau(x)+(1-\Gamma)(c_1^+(x)+c_0^-(x))\Big).$$

应用引理17，$h^*(x)$使上述方程取得最大值，当 $\Big(\Gamma\tau(x)+(1-\Gamma)(c_1^+(x)+c_0^-(x))\Big)\geq 0$.时满足 $h^*(x)\geq\frac{1}{2}$。因此，当 $\pi_0(x)=0$.时，$\mathbb{I}(h^*(x)\geq\frac{1}{2})$等于 $\mathbb{I}(\tau(x)\geq H^+(x))$。当 $\pi_0(x)=1$时，上述问题化简为

$$\max_{h(x)\in[0,1]}\log(1+\exp(-2h(x)+1))\cdot\Big(-\Gamma\tau(x)+(1-\Gamma)(c_1^-(x)+c_0^+(x))\Big).$$

应用引理17，$h^*(x)$ 使上述方程达到最大值，当 $0\geq-\Gamma\tau(x)+(1-\Gamma)(c_1^-(x)+c_0^+(x))$ 时，满足 $h^*(x)\geq\frac{1}{2}$。因此，当 $\pi_0(x)=1.$ 时，$\mathbb{I}(h^*(x)\geq\frac{1}{2})$ 等于 $\mathbb{I}(\tau(x)\geq H^-_\Gamma(x))$。由此，$\pi(x)=\mathbb{I}(h^*(x)\geq\frac{1}{2})$ 等价于（21）。

D.13 定理12的证明

根据引理19，如果 h s 为任意满足 $Q_X \ll P_{\text{obs},X}$ 且 $\sup_{x\in\mathcal{X}} dP_{\text{obs},X}(x)\, dQ_X(x) < \infty$ 的 Q_X 求解（35），则它求解 s

$$\max_{h(x)\in[0,1]} \inf_{P\in\mathcal{T}(P_{\text{obs}})} \Gamma\mathbb{E}_P\left[v^*_{\text{regret}}(h(x); X,\ Y(0)\ Y(1)) \mid X=x\right]+(1-\Gamma)条件风险价值^{\zeta(\Gamma)}(v^*_{\text{regret}}(h(x);\ X,\ Y(0)\ Y(1)) \mid x)。 \tag{91}$$

对每个 $x \in \text{supp}(Q_X)$。

我们注意到

$$\begin{aligned}
&\mathbb{E}_P\left[v^*_{\text{regret}}(h(x); X, Y(0), Y(1)) \mid X=x\right] \\
&= \mathbb{E}_P\left[(2h(x)-1)\cdot(Y(1)-Y(0)) + \log(|2h(x)-1|) \mid X=x\right] \\
&= (2h(x)-1)\tau(x) + \log(|2h(x)-1|)
\end{aligned}$$

因此，第一项在 P_{obs} 下可识别，且不依赖于 $P \in \mathcal{T}(P_{\text{obs}})$的选择。 接下来，应用引理15 和引理 14 可得

$$\begin{aligned}
&\inf_{P\in\mathcal{T}(P_{\text{obs}})} (1-\Gamma)\text{CVaR}_{\zeta(\Gamma)}(v^*_{\text{regret}}(h(x); X, Y(0), Y(1)) \mid x) \\
&= (1-\Gamma) \sup_{P\in\mathcal{T}(P_{\text{obs}})} \text{CVaR}_{\zeta(\Gamma)}(v^*_{\text{regret}}(h(x); X, Y(0), Y(1)) \mid x) \\
&= (1-\Gamma)\Big(\sup_{P\in\mathcal{T}(P_{\text{obs}})} \text{CVaR}_{\zeta(\Gamma)}((2h(x)-1)\cdot(Y(1)-Y(0)) \mid x) + \log(|2h(x)-1|)\Big) \\
&= (1-\Gamma)\Big((2h(x)-1)_+(c_1^+(x)+c_0^-(x)) + (-2h(x)+1)_+(c_0^+ + c_1^-(x)) + \log(|2h(x)-1|)\Big).
\end{aligned}$$

我们定义

$$J(z;x) = (2z-1)_+ \cdot g_1(x) + (-2z+1)_+ \cdot g_0(x) + \log(|2z-1|), \tag{92}$$

其中 $g_1(x)$ $g_0(x)$ 的定义见引理 21。当 $h(x) \geq \frac{1}{2}$, $J(h(x)\ x) = (2h(x)-1)g_1(x) + \log(|2h(x)$ $1|$. $h\ x\ \leq \frac{1}{2}\ J\ h\ x\ \ x\ =\ -2h\ x\ +1\ g_0\ x\ +\ \ |2h\ x\ -1|$. 91) 当 ( ) , (( )) (( )) ( ) log(( )) 我们可以将 ( ) 重写为

$$\max_{h(x)\in[0,1]} J(h(x);x). \tag{93}$$

令 $\tilde{h}(x)$ 使 (93) 最大。必有 $\tilde{h}(x) \neq \frac{1}{2}$，因为

$$\begin{aligned}
\lim_{h(x)\to\frac{1}{2}^+} J(h(x);x) &= -\infty \\
\lim_{h(x)\to\frac{1}{2}^-} J(h(x);x) &= -\infty.
\end{aligned}$$

我们可以证明，$\tilde{h}(x) > \frac{1}{2}$当且仅当 $g_1(x) > g_0(x)$。 注意，$\tilde{h}(x) > \frac{1}{2}$成立的一个充分条件是：对于任意 $z\in[0,\frac{1}{2})$，我们都能找到 $\tilde{z}\in(\frac{1}{2},1]$，使得 $J(\tilde{z},x) > J(z,\ x)$。 我们可以证明，$g_1(x) > g_0(x)$ 蕴含这一充分条件。令 $z \in [0,\frac{1}{2})$，则

$$\begin{aligned}
J(z,x) &= (-2z+1)\cdot g_0(x) + \log(|2z-1|) \\
&= (2(1-z)-1)\cdot g_0(x) + \log(|2(1-z)-1|) \\
&< (2(1-z)-1)\cdot g_1(x) + \log(|2(1-z)-1|) \\
&= J(1-z,x).
\end{aligned}$$

我们可以定义 $\tilde{z} = 1-z$.因此，如果 $g_1(x) > g_0(x)$，那么我们有 $\tilde{h}(x) > \frac{1}{2}$。我们还可以证明，$\tilde{h}(x) > \frac{1}{2}$意味着 $g_1(x) > g_0(x)$。 如果我们有 $\tilde{h}(x) > \frac{1}{2}$，那么我们有 $J(\tilde{h}(x);\ x) > J(1-\tilde{h}(x);\ x)$。这意味着

$$(2\tilde{h}(x)-1)_+ g_1(x) + \log(|2\tilde{h}(x)-1|) > (-2(1-\tilde{h}(x))-1)_+ g_0(x) + \log(|2(1-\tilde{h}(x))-1|).$$

约去各项并化简可得，$g_1(x) > g_0(x)$。

D.14 引理 13的证明

令 P 为在随机对照试验下与 P_{obs} 一致的任意潜在结果分布。我们检查每个价值函数，并认识到在 P_{obs} 下稳定单位处理值分配假设和随机处理分配成立。

首先，我们证明 v_{maxmin}、v^*_{maxmin} 满足假设 1。

$$
\begin{aligned}
&\mathbb{E}_{P_{\text{obs}}}\left[v_{\text{maxmin}}(z;X,Y,W)\right] \\
&= \mathbb{E}_{P_{\text{obs}}}\left[\log\left(1+\exp\{(2z-1)\cdot(2W-1)\}\right)\cdot\left(\frac{Y}{We+(1-W)(1-e)}\right)\right] \\
&= \mathbb{E}_{P_{\text{obs},X}}\left[\mathbb{E}_{P_{\text{obs},Y,W|X}}\left[\log(1+\exp\{(2z-1)\cdot(2W-1)\})\cdot\left(\frac{Y}{We+(1-W)(1-e)}\right)\mid X\right]\right] \\
&= \mathbb{E}_{P_{\text{obs},X}}\left[e\cdot\mathbb{E}_{P_{\text{obs},Y|X,W}}\left[\log\left(1+\exp\{2z-1\}\right)\cdot\frac{Y(1)}{e}\mid X=x,W=1\right]\right] \\
&\quad +\cdot\mathbb{E}_{P_{\text{obs}},X}\left[(1-e)\cdot\mathbb{E}_{P_{\text{obs},Y|X,W}}\left[\log\left(1+\exp\{-2z+1\}\right)\cdot\frac{Y(0)}{1-e}\mid X=x,W=0\right]\right] \\
&= \mathbb{E}_{P_{\text{obs},X}}\left[\log\left(1+\exp\{2z-1\}\right)\mathbb{E}_{P_{Y(1)|X}}\left[Y(1)\mid X=x\right]+\log\left(1+\exp\{-2z+1\}\right)\mathbb{E}_{P_{Y(0)|X}}\left[Y(0)\mid X=x\right]\right] \\
&= \mathbb{E}_{P}\left[v^*_{\text{maxmin}}(z;X,Y(0),Y(1))\right].
\end{aligned}
$$

其次，我们证明 v_{gain}、v^*_{gain} 满足假设 1。

$$
\begin{aligned}
&\mathbb{E}_{P_{\text{obs}}}\left[v_{\text{gain}}(z;X,Y,W)\right] \\
&= \mathbb{E}_{P_{\text{obs}}}\left[(1-\pi_0(X))\log(1+\exp\{2z-1\})\cdot\left(\frac{Y\cdot W}{e}-\frac{Y\cdot(1-W)}{1-e}\right)\right] \\
&\quad +\mathbb{E}_{P_{\text{obs}}}\left[\pi_0(X)\log(1+\exp\{-2z+1\})\cdot\left(\frac{Y\cdot(1-W)}{1-e}-\frac{Y\cdot W}{e}\right)\right] \\
&= \mathbb{E}_{P_{\text{obs},X}}\left[e\cdot\mathbb{E}_{P_{\text{obs},Y|X,W}}\left[(1-\pi_0(X))\log(1+\exp\{2z-1\})\cdot\frac{Y(1)}{e}\mid X=x,W=1\right]\right] \\
&\quad -\mathbb{E}_{P_{\text{obs},X}}\left[(1-e)\cdot\mathbb{E}_{P_{\text{obs},Y|X,W}}\left[(1-\pi_0(X))\log(1+\exp\{2z-1\})\cdot\frac{Y(0)}{1-e}\mid X=x,W=0\right]\right] \\
&\quad -\mathbb{E}_{P_{\text{obs},X}}\left[e\cdot\mathbb{E}_{P_{\text{obs},Y|X,W}}\left[\pi_0(X)\log(1+\exp\{-2z+1\})\cdot\frac{Y(1)}{e}\mid X=x,W=1\right]\right] \\
&\quad +\mathbb{E}_{P_{\text{obs},X}}\left[(1-e)\cdot\mathbb{E}_{P_{\text{obs},Y|X,W}}\left[\pi_0(X)\log(1+\exp\{-2z+1\})\cdot\frac{Y(0)}{1-e}\mid X=x,W=0\right]\right] \\
&= \mathbb{E}_{P_{\text{obs},X}}\left[(1-\pi_0(X))\log(1+\exp\{2z-1\})\cdot\tau(X)-\pi_0(X)\cdot\log(1+\exp\{-2z+1\})\cdot\tau(X)\right] \\
&= \mathbb{E}_{P}\left[v^*_{\text{gain}}(z;X,Y(0),Y(1))\right]
\end{aligned}
$$

第三，我们证明 v_{regret}、v^*_{regret} 满足假设 1。

$$
\begin{aligned}
&\mathbb{E}_{P_{\text{obs}}}\left[v_{\text{regret}}(z;X,Y,W)\right] \\
&= \mathbb{E}_{P_{\text{obs}}}\left[(2z-1)\cdot\left(\frac{Y\cdot W}{e}-\frac{Y\cdot(1-W)}{1-e}\right)+\log(|2z-1|)\right] \\
&= \mathbb{E}_{P_{\text{obs},X}}\left[e\cdot\mathbb{E}_{P_{\text{obs},Y|X,W}}\left[(2z-1)\cdot\frac{Y(1)}{e}\mid X=x,W=1\right]\right] \\
&\quad -\mathbb{E}_{P_{\text{obs},X}}\left[(1-e)\cdot\mathbb{E}_{P_{\text{obs},Y|X,W}}\left[(2z-1)\cdot\frac{Y(0)}{1-e}\mid X=x,W=0\right]\right]+\log(|2z-1|) \\
&= \mathbb{E}_{P_{\text{obs},X}}\left[(2z-1)\tau(X)\right]+\log(|2z-1|) \\
&= \mathbb{E}_{P}\left[v^*_{\text{regret}}(z;X,Y(0),Y(1))\right]
\end{aligned}
$$

E 技术结果的证明

E.1 引理 17 的证明

令 $f(z) = \log(1 + \exp\{2z - 1\})$。则

$$J(z) = Af(z) + Bf(1 - z).$$

显然， $f(z)$ 关于 z 严格递增。若 $A > B$，对于任意 $z > 1/2$， $J(z) > J(1 - z)$，因此

$$\max_{z>1/2} J(z) > \max_{z<1/2} J(z) \Longrightarrow z^* > 1/2.$$

同样，若 $A < B$,

$$\max_{z>1/2} J(z) < \max_{z<1/2} J(z) \Longrightarrow z^* < 1/2.$$

当 $A = B$ 时,

$$J(z) = A\left(f(z) + f(1 - z)\right).$$

容易验证， J 关于 z 是凹的，并且关于 $1/2$ 对称。因此， $z^* = 1/$ 2.

E.2 引理18的证明

该结果的证明与Sahoo等人 [2022中引理1的证明完全相同。].

E.3 引理19的证明

我们注意到，任何 π 对每个26都能求解（ Q_X 使得 $Q_X \ll P_X$ 的策略，都必须对每个 x求解条件问题，即

$$\sup_{\pi(x)\in\{0,1\}} \inf_{Q\in\mathcal{S}^\Gamma(P_{\text{obs}},Q_X)} \mathbb{E}_Q\left[v^*(\pi(x); x, Y(0), Y(1)) \mid X = x\right]. \tag{94}$$

根据 $\mathcal{S}^\Gamma(P_{\text{obs}}$， $Q_X)$的定义，我们有

$$\begin{aligned}&\inf_{Q\in\mathcal{S}^\Gamma(P_{\text{obs}},Q_X)} \mathbb{E}_Q\left[v^*(\pi(x); x, Y(0), Y(1)) \mid X = x\right]\\ &= \inf_{P\in\mathcal{T}(P_{\text{obs}})} \inf_{Q\in\mathcal{R}^\Gamma(P,Q_X)} \mathbb{E}_Q\left[v^*(\pi(x); x, Y(0), Y(1)) \mid X = x\right].\end{aligned}$$

回顾一下，如果 Q 具有协变量分布 Q_X，并通过 Γ-有偏抽样生成 P，则 $Q \in \mathcal{R}^\Gamma(P$， $Q_X)$。对于每个 Q_X 使得 $Q_X \ll P_X$ 且$\sup_{x\in\mathcal{X}} \frac{dP_{\text{obs},\ X}(x)}{dQ_X(x)} < \infty$,我们有（41）成立。利用这一结果，我们注意到，寻找满足以下条件的最坏情况分布$Q^*_{Y(0)\ Y(1)|X}$

$$\inf_{Q\in\mathcal{R}^\Gamma(P,Q_X)} \mathbb{E}_Q\left[v^*(\pi(x); x, Y(0), Y(1)) \mid X = x\right]$$

相当于在 Q 中，在 Q的凸（线性）约束下，最小化一个凹（线性）函数。根据命题 16，解必定出现在可行集的某个极点上。由于 $P_{Y(0),\ Y(1)|X=x}$相对于勒贝格测度绝对连续，我们可以证明$\frac{dQ^*_{Y(0)\ Y(1)|X=x}(y_0,\ y_1)}{dP_{Y(0)\ Y(1)|X=x}(y_0,\ y_1)} \in\{\Gamma^{-1}, \Gamma\}$.

为了确定最优解 $dQ^*_{Y(0)\ Y(1)|X}$ 的结构，我们考虑关于$v^*(\pi(x); x, Y(0)\ Y(1))$ 的分布，其中 $Y(0)\ Y(1) \sim P_{Y(0)\ Y(1)|X=x}\cdot$。我们必须为$Y(0)\ Y(1)$中对应于 $v^*(\pi(x);\ x, Y(0)\ Y(1))$高于某个 $\zeta \in\in(0, 1)$ 分位数的取值赋予权重 Γ^{-1}，并为 $Y(0)\ Y(1)$中对应于 $v^*(\pi(x); x, Y(0)\ Y(1))$低于第 ζ 分位数的取值赋予权重 Γ。否则，就会存在一个 Q ，其策略值低于 Q^*. 。分位数 ζ 的选择旨在确保 $dQ^*_{Y(0)\ Y(1)|X}$ 是一个有效的概率分布。我们选择满足以下条件的 ζ

$$\Gamma^{-1}\zeta + \Gamma(1 - \zeta) = 1.$$

因此，我们将 ζ 表示为 Γ 的函数，如下所示： $\zeta(\Gamma)=\frac{1}{\Gamma+1}$. 设 $v^*(\pi(x);\ x,\ Y(0)\ Y(1))$ 中 $Y(0)\ Y(1)\sim P_{Y(0)\ Y(1)|X=x}$ 的第 $\zeta(\Gamma)$ 分位数为 $q_{\zeta(\Gamma)}(v^*(\pi(x);\ X,\ Y(0)\ Y(1))\mid X=x)$。因此，我们可以得出结论，最坏情况分布为

$$dQ^*_{Y(0)\ Y(1)|X=x}(y_0,y_1)\ =\begin{cases}\Gamma^{-1}\cdot dP_{Y(0)\ Y(1)|X=x}(y_0,y_1)\ v^*(\pi(x);\ x,y_0,y_1)\ \geq q_{\zeta(\Gamma)}(v^*(\pi(x);X,Y(0)\ Y(1))\mid x)\\ \Gamma\cdot dP_{Y\ 0}\end{cases}()\ Y(1)|X=x(y_0,y_1)\ \text{否则}$$

我们可以写作娱乐主管

(95)

下确界 $_{Q\in\mathcal{R}^{\Gamma}(P,Q_X)}\mathbb{E}_P\left[v^*(\pi(x);x,Y(0)\ Y(1))\mid X=x\right]=\Gamma^{-1}\mathbb{E}_Q\left[v^*(\pi(x);\ x,Y(0)\ Y(1))\mathbb{I}(v^*(\pi(x);\ x,Y(0)\ Y(1))\geq q_{\zeta(\Gamma)}(v^*(\pi(x);\ X,Y(0)\ Y(1)))\mid x)\mid X=x\right]++\Gamma\mathbb{E}_P\left[v^*(\pi(x);x,Y(0)\ Y(1))\mathbb{I}(v^*(\pi(x);\ x,Y(0)\ Y(1))\leq q_{\zeta(\Gamma)}(v^*(\pi(x);\ X,Y(0)\ Y(1)))\mid x)\mid X=x\right]=\Gamma\mathbb{E}_P\left[v^*(\pi(x);\ x,Y(0)\ Y(1))\mid X=x\right]+(1-\Gamma)\cdot$条件风险价值$_{\zeta(\Gamma)}(v^*(\pi(x);x,Y(0)\ Y(1))\mid x)$。

因此， π 满足该条件的必须满足（94）

$$\sup_{\pi(x)\in\{0,1\}}\inf_{P\in\mathcal{T}(P_{\text{obs}})}\Gamma\mathbb{E}_P\left[v^*(\pi(x);x,Y(0),Y(1))\mid X=x\right]+(1-\Gamma)\cdot\text{CVaR}_{\zeta(\Gamma)}(v^*(\pi(x);x,Y(0),Y(1))\mid x).$$

如所需。

我们证明这一结果的逆命题同样成立。假设对于每个 $x\in$supp$(P_{\text{obs},X})$， $\tilde{\pi}$ 都满足（42）。根据前面的论证，这意味着对于每个 $x\in$supp$(P_{\text{obs},X})$， $\tilde{\pi}$都满足（94）。令 $\mathcal{T}$为关于 Q_X具有非零测度的任意集合。则 $\mathcal{T}$关于 P_{obs} 也必须具有非零测度，因为 $Q_X\ll P_X$. 所以，

$$\inf_{Q\in\mathcal{S}^{\Gamma}(P_{\text{obs}},Q_X)}\mathbb{E}_Q\left[v^*(\tilde{\pi}(X);X,Y(0),Y(1))\mid X\in\mathcal{T}\right]\geq\inf_{Q\in\mathcal{S}^{\Gamma}(P_{\text{obs}},Q_X)}\mathbb{E}_Q\left[v^*(\pi(X);X,Y(0),Y(1))\mid X\in\mathcal{T}\right]$$

对于任意 $\pi\in\Pi$ 以及关于 Q_X具有非零测度的任意集合 $\mathcal{T}$。这足以证明 $\tilde{\pi}$必须满足（26）。

同样的论证也适用于证明：对于满足 $Q_X\ll P_{\text{obs},X}$ 的每个 Q_X， h 满足（30），且 $\sup_{x\in\mathcal{X}}\frac{dP_X}{dQ_X}(x)<\infty$当且仅当它对于每个 $x\in$ supp$(P_{\text{obs},X})$满足（43）。

E.4 引理 20 的证明

由引理 19，我们可得

$$\inf_{P\in\mathcal{T}(P_{\text{obs}})}\Gamma\mathbb{E}_P\left[v^*(\pi(x),\pi_0(X);X,Y(0),Y(1))\mid X=x\right]+(1-\Gamma)\text{CVaR}(v^*(\pi(x),\pi_0(X);X,Y(0),Y(1))\mid x).\tag{96}$$

注意到

$$\begin{aligned}\mathbb{E}_P\left[v^*(\pi(x),\pi_0(X);X,Y(0),Y(1))\mid X=x\right]&=(\pi(x)-\pi_0(x))\cdot\tau(x)\\&=(\pi(x)-\pi_0(x))_+\tau(x)-(\pi_0(x)-\pi(x))_+\tau(x),\end{aligned}$$

其可识别为 u nder P_{obs} 且不取决于 $P\in\mathcal{T}(P_{\text{obs}})$。因此，我们可以将（96）重写为 a s

$$(\pi(x)-\pi_0)_+\tau(x)-(\pi_0(x)-\pi(x))_+\tau(x)+\inf_{P\in\mathcal{T}(P_{\text{obs}})}(1-\Gamma)\text{条件风险价值}_{\zeta(\Gamma)}(v^*(\pi(x)\ \pi_0(X);X,Y(0)\ Y(1))\mid x).\tag{97}$$

我们有

$$\inf_{P\in\mathcal{T}(P_{\text{obs}})}(1-\Gamma)\text{CVaR}_{\zeta(\Gamma)}(v^*(\pi(x),\pi_0(X);X,Y(0),Y(1))\mid x) \tag{98}$$

$$=(1-\Gamma)\sup_{P\in\mathcal{T}(P_{\text{obs}})}\text{条件风险价值}_{\zeta(\Gamma)}(v^*(\pi(x)\ \pi_0(X);\ X,\ \ Y(0)\ Y(1))\mid x) \tag{99}$$

$$=(1-\Gamma)\sup_{P\in\mathcal{T}(P_{\text{obs}})}\text{CVaR}_{\zeta(\Gamma)}((\pi(x)-\pi_0(x))\cdot(Y(1)-Y(0))\mid x) \tag{100}$$

$$=(1-\Gamma)\sup_{P\in\mathcal{T}(P_{\text{obs}})}(\pi(x)-\pi_0(x))_+\text{CVaR}_{\zeta(\Gamma)}(Y(1)-Y(0))\mid x)+(\pi_0(x)-\pi(x))_+\text{CVaR}_{\zeta(\Gamma)}(Y(0)-Y(1)\mid x) \tag{101}$$

$$=(1-\Gamma)\Big((\pi(x)-\pi_0(x))_+(c_1^+(x)+c_0^-(x))+(\pi_0(x)-\pi(x))_+(c_0^+(x)+c_1^-(x))\Big). \tag{102}$$

（101）由引理14可得，而（102）由引理15可得。将该表达式代入（97）得到（44）。

E.5 引理21的证明

我们有

$$\begin{aligned}
g_1(x)&=\Gamma\tau(x)+(1-\Gamma)(c_1^+(x)+c_0^-(x))\\
&=\Gamma\tau(x)+(1-\Gamma)\Big(\mathbb{E}_P\left[Y(1)\mid Y(1)\geq q_{\zeta(\Gamma)}(Y(1)\mid x),X=x\right]+\mathbb{E}_P\left[-Y(0)\mid -Y(0)\geq q_{\zeta(\Gamma)}(-Y(0)\mid x),X=x\right]\Big)\\
&\leq\Gamma\tau(x)+(1-\Gamma)\mathbb{E}_P\left[Y(1)-Y(0)\mid X=x\right]\\
&=\tau(x).
\end{aligned}$$

此外，我们还有

$$\begin{aligned}
g_0(x)&=-\Gamma\tau(x)+(1-\Gamma)(c_1^-(x)+c_0^+(x))\\
&=-\Gamma\tau(x)+(1-\Gamma)\Big(\mathbb{E}_P\left[Y(0)\mid Y(0)\geq q_{\zeta(\Gamma)}(Y(1)\mid x),X=x\right]+\mathbb{E}_P\left[-Y(1)\mid -Y(1)\geq q_{\zeta(\Gamma)}(-Y(1)\mid x),X=x\right]\Big)\\
&\leq-\Gamma\tau(x)+(1-\Gamma)\mathbb{E}_P\left[Y(0)-Y(1)\mid X=x\right]\\
&=-\Gamma\tau(x)+(1-\Gamma)(-\tau(x))\\
&=-\tau(x).
\end{aligned}$$

因此，我们有 $\min\{g_1(x)\ g_0(x)\}\leq\min\{\tau(x)\ -\tau(x)\}\leq 0$. 此外，如果 $g_1(x)=0$，那么 $\tau(x)\geq 0$，因此 $g_0(x)\leq 0$。

E.6 引理22的证明

为了证明 $Q_{\text{obs}}\in\mathcal{S}^{\Gamma}_{\text{obs}}(P_{\text{obs}},\ Q_X)$，我们首先验证（27）。由于 $Q\in\mathcal{S}^{\Gamma}(P_{\text{obs}},\ Q_X)$，$Q$通过 Γ-有偏抽样生成了潜在结果分布，且该分布与 P_{obs} 在治疗倾向函数为 e的随机对照试验下相一致。由于 Q 在 Γ-有偏抽样下生成了 P，因此根据引理24，等式（47）成立。由于P 与 P_{obs} 在随机对照试验下相一致，我们有

$$P_{\text{obs},Y|X=x,W=w}=P_{Y(w)|X=x}\quad\forall x\in\mathcal{X},w\in\{0,1\}.$$

此外，由于 Q 与 Q_{obs} 在随机对照试验下相一致，我们有

$$Q_{\text{obs},Y|X=x,W=w}=Q_{Y(w)|X=x}\quad\forall x\in\mathcal{X},w\in\{0,1\}.$$

将上述两个方程代入（47）得到（27），证毕。

接下来，我们验证 $Q_{\text{obs},W|X}=P_{\text{obs},W|X}$.根据假设，$P_{\text{obs},W|X}$ =伯努利分布(e)。此外，由于Q_{obs} 是通过处理倾向函数 e生成的，因此有 $Q_{\text{obs},W|X}$ = 伯努利分布(e)。因此，$Q_{\text{obs},W|X}=P_{\text{obs},W|X}$.

最后，我们注意到 $Q_{\text{obs},\ X}=Q_X$ 因为进行随机对照试验不会影响数据的协变量分布。

因此，我们得出结论： $Q_{\text{obs}}\in\mathcal{S}^{\Gamma}_{\text{obs}}(P_{\text{obs}},\ Q_X)$。

E.7 引理 23 的证明

我们可以针对每个 x条件求解（45）：

$$\inf_{Q_{\text{obs}} \in \mathcal{S}^{\Gamma}_{\text{obs}}(P_{\text{obs}}, Q_X)} \mathbb{E}_{Q_{\text{obs},Y,W|X=x}} \left[v(h(X); X, Y, W) \mid X = x \right]. \tag{103}$$

令优化变量为 $dQ_{\text{obs},Y,W|X}$,则可将上述问题视为在凸约束条件下最小化一个线性函数。这些约束包括:

1. $Q_{\text{obs},Y,W|X}$ 是一个有效的概率分布，因此 $dQ_{\text{obs},Y,W}(y,w) \geq 0$ 且 $\int dQ_{\text{obs},Y,W}(y,w) = 1$。

2. 处理分配是随机的。$dQ_{\text{obs},W|X} = dP_{\text{obs},W|X}$ =服从伯努利分布(e)。

3. $\Gamma^{-1} \leq \frac{dQ_{\text{obs},Y|X=x,W=w}(y)}{dP_{\text{obs},Y|X=x,W=w}(y)} \leq \Gamma$ for all $w \in \{0,1\}, x \in \mathcal{X}$.

凸函数在闭有界凸集上的上确界存在，且在可行集的某个极点处达到（命题 16）。由于 $P_{Y(0)\,Y(1)|X=x}$ 相对于勒贝格测度绝对连续，因此最优解必须满足 $\frac{dQ^*_{\text{obs}\ Y|X=x,W=w}(y)}{dP_{\text{obs},Y|X=x,W=w}(y)} \in \{\Gamma^{-1}, \Gamma\}$，并满足其余约束，尤其是 $dQ^*_{\text{obs},W|X}$ = 服从伯努利分布(e)。

为确定最优解 $dQ^*_{\text{obs}、Y|X、W}$ 的结构，我们可以考虑关于 $v(h(x); x、 Y、 w)$ 的分布，其中 $Y \sim P_{\text{obs}、Y|X=x、W=w}$。我们必须为 Y 的值分配权重 Γ^{-1}，这些值对应于 $v(h(x); x、 Y、 w)$ 的值，且其高于某个 ζ 分位数（其中 $\zeta \in (0,1)$）；同时，为 Y 的值分配权重 Γ，这些值对应于 $v(h(x); x、 Y、 w)$ 的值，且低于该阈值。否则，将存在一个 dQ_{obs}，使得其策略值低于 dQ^*_{obs}.。分位数 ζ 的选择应确保 dQ^*_{obs} 是一个有效的概率分布。我们选取满足以下条件的 ζ

$$\Gamma^{-1}\zeta + \Gamma(1-\zeta) = 1.$$

因此，我们将 ζ 表示为 Γ 的函数如下： $\zeta(\Gamma) = \frac{1}{\Gamma+1}$.此外，设 $v(h(x); x、 Y、 w)$ 的 $\zeta(\Gamma)$ 分位数为 $q_{\zeta(\Gamma)}(v(h(X); X、 Y、 W) \mid x、 w)$。因此，我们可以得出结论：最坏情况下的条件分布由（46）给出。因此，在（45）中取得下确界的分布 dQ^*_{obs} 由 $dQ_X \cdot dP_{\text{obs}W|X} \cdot dQ^*_{\text{obs}、Y|X、W}$ 给出，其中 $dQ^*_{\text{obs}、Y|X、W}$ 如（46）所定义。

引理 24的证明

首先，我们假设 Q 通过 P 进行 Γ有偏抽样生成。根据引理 18，我们可知协变量分布之间的密度比是有界的。此外，我们还有

$$\frac{dP_{Y(0)|X=x}(y)}{dQ_{Y(0)|X=x}(y)} = \frac{\mathbb{P}_{\tilde{Q}}\left[S=1 \mid X=x, Y(0)=y\right]}{\mathbb{P}_{\tilde{Q}}\left[S=1 \mid X=x\right]} \tag{104}$$

$$= \mathbb{E}_{\tilde{Q}_{Y(1)|X=x,Y(0)=y}}\left[\frac{\mathbb{P}_{\tilde{Q}}\left[S=1 \mid X=x, Y(0)=y, Y(1)\right]}{\mathbb{P}_{\tilde{Q}}\left[S=1 \mid X=x\right]}\right] \tag{105}$$

$$\in [\Gamma^{-1}, \Gamma]. \tag{106}$$

采用与Sahoo等人 [2022], 引理1的证明相同的论证，我们得到（104）。我们将（2）应用于（105）中期望内的量，从而得到（106）。同样的论证也适用于$dQ_{Y\ (1)\ |X=x}dP_{Y(1)|X=x}$。