

· 基础理论与方法 ·

针对连续型处理因素的广义倾向性评分估计方法简介

张雨 高倩 王彤

山西医科大学公共卫生学院卫生统计学教研室, 太原 030001

通信作者: 王彤, Email: tongwang@sxmu.edu.cn

【摘要】 在观察性研究中进行因果推断的众多方法中, 用于控制已测量混杂的倾向性评分方法应用越来越广泛。该类方法主要分为两步: 首先估计倾向性评分, 然后采取回归、加权、匹配和分层等手段进一步估计感兴趣的因果参数。不同于传统的二分类处理情况, 近年来针对连续型处理因素的广义倾向性评分方法被提出。目前已发展出了许多估计广义倾向性评分和直接估计均衡权重的方法, 本文将从基于模型和基于均衡性两个角度出发对现有方法进行介绍。

【关键词】 广义倾向性评分; 连续型处理因素; 稳定权重; 均衡性; 优化问题

基金项目: 国家自然科学基金 (82073674)

Overview on the generalized propensity scoring estimator for continuous treatment

Zhang Yu, Gao Qian, Wang Tong

Department of Health Statistics, School of Public Health, Shanxi Medical University, Taiyuan 030001, China

Corresponding author: Wang Tong, Email: tongwang@sxmu.edu.cn

【Abstract】 Among kinds of methods for causal inference in observational studies, the propensity score (PS) method to control measured confounding is more widely used. PS method usually consists of two critical steps: first, estimating the propensity score, followed by calculating the causal parameters of interest by regression, weighting, matching, and stratification. Unlike the traditional dichotomous treatment, the generalized propensity scoring estimator used for continuous treatment has been proposed in recent years. Many methods have been developed to estimate the generalized propensity score or even estimate the balancing weight directly. This paper introduces the existing estimators from both the model-based and balance-based perspectives.

【Key words】 Generalized propensity scoring estimator; Continuous treatment; Stabilized weight; Covariate balancing; Optimization problem

Fund program: National Natural Science Foundation of China (82073674)

在因果推断领域中, 使用基于反事实框架的倾向性评分(propensity score, PS)方法控制已测量混杂因素是由 Rosenbaum 和 Rubin^[1]于 1983 年首次提出。传统的 PS 方法主要针对的是二分类处理因素情况, 例如: 欲探究太极拳干预与不良心血管事件发生率的因果关系时^[2], 首先构建一个关于是否实

施太极拳干预与一系列已测量基线协变量包括性别、年龄、血脂水平等的 logistic 回归估计个体接受干预的概率, 即 PS。然后利用 PS 的均衡特性(在 PS 同质的子样本中, 已测量的混杂变量在干预组和对照组间均衡可比)通过匹配计算两组不良心血管事件发生率差的期望, 即为太极拳干预所致不良

DOI: 10.3760/cma.j.cn112338-20210827-00685

收稿日期 2021-08-27 本文编辑 万玉立

引用格式: 张雨, 高倩, 王彤. 针对连续型处理因素的广义倾向性评分估计方法简介[J]. 中华流行病学杂志, 2022, 43(4): 572-577. DOI: 10.3760/cma.j.cn112338-20210827-00685.

Zhang Y, Gao Q, Wang T. Overview on the generalized propensity scoring estimator for continuous treatment[J]. Chin J Epidemiol, 2022, 43(4): 572-577. DOI: 10.3760/cma.j.cn112338-20210827-00685.



心血管事件效应大小的无偏估计,实现了事后随机化也称伪随机化。为了应对多分类处理因素, Imbens^[3]于 2000 年开发了广义倾向性评分 (generalized propensity score, GPS) 方法。例如,欲探究文献发表期刊(9 种)与期刊影响因子的因果关系时^[4],与估计 PS 类似,可使用多分类 logistic 回归代替 logistic 回归来估计文献在给定处理前协变量包括出版年份、作者人数、页数等条件下发表于某期刊的概率,即 GPS。然后通过构建只包含期刊种类和 GPS 两个协变量的结局模型(回归调整)或先赋予文献 $\frac{1}{\text{GPS}}$ 大小的权重以产生一个协变量在处理组间均衡可比的虚拟文献集,再构建只包含处理因素的加权结局模型(逆概率加权)等手段对文献发表于该期刊的影响因子大小进行估计,其他期刊影响因子估计过程类似。在实际应用过程中,许多具有公共卫生意义的潜在可控因素比如吸烟数量(支/d)、药物剂量(mg/kg)、BMI(kg/m²)等多为连续型变量,若简单将其转换为分类变量进行分析,样本信息的损失可能导致效能降低。2004 年, Hirano 和 Imbens^[5]将 GPS 方法扩展到了连续型处理因素,与前述估计分类处理效应过程类似:首先估计个体 GPS(即给定基线协变量时,个体被分配到某一处理水平的条件概率密度),然后采用匹配、回归调整、逆概率加权等手段估计因果效应值,得到处理效应一致无偏估计的前提是准确估计 GPS。目前针对连续型处理因素的 GPS 估计方法大致可分为基于模型和基于均衡性两大类,均衡类方法由模型类方法发展而来,多数是在均衡条件的约束下直接估计均衡权重而非先估计 GPS 再估计逆概率权重,有效避免模型类方法可能存在模型误设问题的同时能更好地实现均衡性,保证因果效应参数估计的一致有效性。本文将简要介绍使用 GPS 方法估计连续处理效应的步骤以及需要满足的前提假设并举例说明,然后从模型和均衡性两方面对估计 GPS 或直接估计均衡权重的方法进行介绍。

1. 利用 GPS 估计处理效应:

(1)符号与假设:令 Z 为一个连续型处理因素, X 为一系列已测量基线协变量构成的向量。采用 Hirano 和 Imbens^[5]的术语,令 $f_{Z|X}(z|x)$ 为在给定系列已测量基线协变量 x 情况下连续型处理变量取特定值 z 时的条件概率密度,即为 GPS。潜在结果框架定义存在一系列潜在结局 $Y_i(z), z \in \psi$, 其中 $Y_i(z)$ 为第 i 个个体接受处理水平为 $Z = z$ 时的结

局, ψ 为处理变量全部取值的集合。在连续处理情况下, ψ 是一个实数集,即 $\psi \subset \mathbb{R}$ 。剂量反应函数 (dose-response function, DRF) 是指总体(或样本)中全部个体接受处理水平为 $Z = z$ 时的平均反应量,表示为 $\mu(z) = E[Y_i(z)]$ 。

利用 GPS 识别 DRF 需要满足 3 个前提假设:①弱无混杂假设:要求在给定基线协变量的条件下,处理因素与潜在结局相互独立。②一致性假设:要求观测到的结局等于相应处理水平的潜在结局。③阳性假设:要求连续处理变量的条件概率密度非负。

(2)DRF 的估计:回归调整: Hirano 和 Imbens^[5]提出将连续处理和 GPS 估计值作为协变量构建结局回归模型。要得到处理效应的无偏估计则需要正确指定结局模型,但是在复杂的现实数据情况下较难实现。基于此, Kreif 等^[6]于 2015 年利用机器学习算法提出了超级学习者,模拟研究显示该算法显著降低了结局模型错误指定的影响。

逆概率加权 (inverse probability weighting, IPW): Robins 等^[7]于 2000 年提出基于边际结构模型的 IPW 方法,其中逆概率权重 $w = \frac{1}{f_{Z|X}(z|x)}$,即 $\frac{1}{\text{GPS}}$ 。当处理因素与协变量强相关时,据此计算的 w 变异会较大而导致极端权重的出现。于是 Robins 等^[7]进一步提出将分子 1 替换成 Z 的边际概率,构成稳定的逆概率权重 (sw), $sw = \frac{f_Z(z)}{f_{Z|X}(z|x)}$ 。将此权重用于构建加权回归模型即可获得处理效应的估计值,但结局模型的正确指定仍是获得一致估计的前提。以上是当前基于 GPS 进行因果推断应用最为广泛的两类方法。此外, Bang 和 Robins^[8]于 2005 年提出将结局回归模型与 IPW 结合构建双稳健方法 (doubly robust, DR), DR 的优点在于只要 GPS 模型和结局模型之一正确指定即可获得处理效应的一致估计^[9]。

近年来,还有学者提出利用 GPS 得分进行匹配和分层估计 DRF。Wu 等^[10]于 2018 年提出联合 GPS 估计值和处理水平构建 GPS 卡钳匹配框架以充分控制混杂偏倚。Brown 等^[11]于 2019 年提出了两种针对连续处理因素的分层方法,分别是 GPS 累积分布函数和非参 GPS 累积分布函数,二者均为观测单位指定了具体的协变量分布用于分层。

(3)案例:欲探究彩票金额与中彩 6 年后劳动收入之间的因果关联^[5],考虑年龄、性别、文化程度

等基线协变量作为潜在的混杂因素。第一步:拟合一个彩票金额对基线协变量的正态线性回归模型,据此计算出彩票金额的条件概率密度,即 GPS;第二步:拟合一个结局回归模型估计个体在给定彩票金额和 GPS 条件下 6 年后劳动收入的条件期望(回归调整);第三步:计算研究群体中全部个体都获得某特定彩票金额时 6 年后劳动收入的期望,即为二者间的剂量反应关系。

2. 基于模型估计 GPS: Imai 和 van Dyk^[12]称条件概率密度函数 $f_{Z|X}$ 为倾向函数,倾向函数可通过构建 Z 对 X 的回归模型被估计,通常采用的是普通最小二乘法 (ordinary least squares, OLS) 回归。Austin^[13-14]、Brown 等^[11]假定 Z 是服从均数为 $\beta^T X$ 、方差为 σ^2 (其中 β 和 σ^2 是由 OLS 估计所得) 的正态分布, $f_{Z|X}(z|x)$ 即可估计为 $\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(z-\beta^T X)^2}{2\sigma^2}}$, 其中 z 是个个体处理水平的观测值,该估计值即为个体的 GPS。为满足正态性假定, Alejo 等^[15]根据 Wand 等^[16]对拟合模型前的原始数据转换原理对回归模型的两端进行 Box-Cox 转换,使得转换后模型残差项服从正态分布,采用极大似然法估计未知参数,进而得到 $\hat{f}_{Z|X}(z|x)$ 。Box-Cox 转换在应用时容易实现,在协变量个数较多时也能保证收敛速度,但是进行转换变量的取值要求非负。

正态性假定不是必要的, Bia 等^[17]采用一致函数、对数函数和幂函数作为连接函数并且各自考虑了高斯分布、逆高斯分布和伽马分布的情况,通过拟合广义线性模型 (generalized linear model, GLM) 对 GPS 进行估计。类似地, Guardabascio 和 Ventura^[18]假定 Z 服从指数家族分布,然后选择一个单调可微的连接函数对 $E(Z)$ 作非一致转换使其与 X 线性相关。

然而,上述方法使用时普遍存在一个问题:错误指定 GPS 模型会导致估计结果有偏。Kreif 等^[6]于 2015 年首次将超级学习者这一机器学习算法用于估计 GPS。该算法以数据驱动的方式通过交叉验证找到由数个预测算法加权组合的最优形式^[19-20],预测算法可包括参数和非参数形式,如简单线性回归、最小角回归、逻辑回归、删除/替换/增加算法、随机森林等。理论证明显示超级学习者表现渐近甚至优于组合算法中的任一算法^[21-22]。

2015 年, Zhu 等^[23]提出用 boosting 机器学习算法估计 GPS,其优点在于可以自动选择重要协变量,甚至考虑高阶和交互效应^[24]。boosting 算法拟

合的是一个可加模型,其中每一个加和项都是一棵回归树。树的数目 M 是该算法的一个重要调整参数, M 决定了进行因果推断时的偏差-方差平衡。Zhu 等^[23]提出最优的 M 应该使平均绝对相关系数 (average absolute correlation coefficient, AACC) 最小, AACC 是采用 bootstrap 计算得到的 Z 与 X 之间的加权平均相关系数,衡量均衡性。模拟研究显示:当 GPS 模型存在高阶或交互项时,boosting 算法估计因果效应的偏差、均方根误差以及 95% CI 覆盖率表现均显著优于线性拟合。

以上基于模型所得 GPS 用于估计 DRF 前需要检验均衡性,因为只有在协变量均衡情况下进行因果推断才能获得处理效应的无偏估计。然而上述方法并不能保证有效均衡协变量在不同处理水平间的分布,需要重复进行均衡性检验并对模型进行调整,直至达到预设的均衡标准^[25]。鉴于此,近年来发展出了许多直接基于均衡性估计 GPS 或者估计权重 w 的方法。

3. 基于均衡性估计 GPS 或权重: Fong 等^[26]提出了协变量均衡 GPS (covariate balancing generalized propensity score, CBGPS) 方法,该法首先将 Z 中心化得到 Z^* 、将 X 中心化并正交化后得到 X^* ,然后拟合 Z^* 关于 X^* 的回归模型,基于 Z^* 与 X^* 均服从正态分布构建稳定权重,同时满足 Z^* 与 X^* 间的加权相关系数等于 0。加权样本中转化后的协变量在不同处理水平间实现了均衡,同时原始协变量也满足此均衡性。此参数方法考虑到 GPS 的估计不确定性,据此可推导出估计因果效应的渐近方差,避免了 bootstrap 的繁复计算过程。然而 CBGPS 仍依赖于模型的正确设定和正态性假定,鉴于此, Fong 等^[26]还提出了一种非参 CBGPS (nonparametric CBGPS, npCBGPS), npCBGPS 通过构建一个经验似然函数并使其最大化得到权重 w_i 。同时, w_i 还需要满足以下约束条件:①为实现均衡,将 Z^* 与 X^* 间的加权相关系数设为 0;②加权后 Z^* 与 X^* 的边际概率不变;③加权后样本量不变;④权重为正。以上构成一个优化问题,可采用拉格朗日乘子法求解最优 w_i 。然而,由于经验似然函数并不总是凸的,所以并不一定能求得最优解。模拟研究表明:就均衡协变量均数方面而言, CBGPS 与 npCBGPS 表现优于 Zhu 等^[23]提出的 boosting 算法,且 npCBGPS 相对更佳。然而当样本量足够大时,boosting 估计 GPS 及处理效应相较 CBGPS 与 npCBGPS 更为准确^[27]。此外, CBGPS 与 npCBGPS 仍不具备变量选择的能力,

Gao 等^[28]等最近提出 GOAL 方法结合结局适应性 LASSO 技术在高维数据背景下解决了此问题。

与 npCBGPS 类似, Yiu 和 Su^[29]于 2018 年提出协变量关联消除权重 (covariate association eliminating weights, CAEW), 不同点: 为控制极端权重并保证效应估计的稳定性将目标函数设为最小化权重的方差。上述前两个约束条件利用极大似然法估计倾向函数参数原理保证加权后 Z 与 X 不相关的同时与加权前的边际分布保持一致。第四个约束条件权重非负, 即 $w_i \geq 0$ 可以构成一个凸二次规划问题便于求解。值得注意的是, 与 npCBGPS 不同, 此类凸优化问题目标函数并不是基于似然推断构建的, 但仍具有较好的统计特性^[30]。此外, npCBGPS 仅能消除协变量与连续处理均数之间的相关性, CAEW 方法还可以消除协变量与连续处理方差之间的相关性。但是此方法同等对待并且同时均衡全部协变量, 因此高维情况的适用性如何还有待研究。

同样通过最小化样本权重的方差构建凸约束二次优化问题求解权重的还有 Santacatterina^[31]于 2020 年提出的稳健正交权重 (robust orthogonality weights, ROW) 方法, 控制极端权重的同时最大化精度。与 npCBGPS 类似, ROW 采用相关性作为均衡性的度量指标, 并且未利用结局信息以模拟随机化过程。不同的地方在于 ROW 设定第一个约束条件 Z^* 与 X^* 间的加权相关系数为一个界值 δ 而非 0。但是 ROW 并未设置第二个约束条件, 即未考虑加权后 Z^* 与 X^* 的边际分布不变。合理选择参数 δ 是 ROW 方法的关键之一, 不严格为 0, 一是便于求解, 二是寻求偏差-方差平衡, 并且相较严格为 0 时会纳入更多协变量以尽可能满足不存在未测量混杂的假设^[32], 作者推荐 $\delta = 0.0001$ 或 $\delta = 0.001$ 。此外, ROW 还可看作是 CAEW 框架的一个特例。模拟研究表明: 不论处理因素分布如何以及处理与协变量间真实关系如何, ROW 都能够实现均衡。在实际阳性假设弱违背情况下, ROW 估计因果效应的偏差和均方根误差表现均优于 CBGPS 和 npCBGPS。这种优势在结局模型错误指定情况下表现更明显。

CAEW 和 ROW 关注与处理相关的协变量以期实现完全均衡, 但是当均衡协变量个数或形式被错误指定时会导致估计潜在结局的效率降低。因此 Greifer^[33]于 2020 年针对与结局相关的协变量构建优化权重以期实现充足均衡。基本原理: 首先构建处理与结局的一般线性模型, 然后将效应参数即模

型斜率的可能偏差分为五个可加部分: 已测量混杂、不对称性、协变量选择、处理样本选择以及未测量混杂并加以控制。优化权重仍以最小化权重方差作为目标函数, 除权重非负并且和为 n 外, 除未测量混杂外对可控的四部分偏差也逐一进行了约束。优化权重首次对 Z 和 X 分布的不对称性以及二者间的非线性关系可能导致的偏差进行控制。模拟研究表明: 当协变量个数与样本量接近并且适当放松对已测量混杂和不对称性的约束 (δ_1, δ_2) 时, 优化权重在均方根误差方面的表现优于结局回归、CBGPS 以及熵平衡加权方法。但是优化权重仅适用于结局变量为连续型并且与处理变量间线性相关也即 DRF 为直线型的情况, DRF 为非直线型时的权重估计还有待进一步研究。

Kallus 和 Santacatterina^[34]于 2019 年提出了核优化正交加权 (kernel optimal orthogonality weighting, KOOW) 方法, 其基本原理是通过求解一个线性约束凸优化问题找到使得最坏情况惩罚函数协方差最小的权重。使用最小化处理与混杂间协方差的总体思想源于最大化均值差异 (maximal mean discrepancies, MMD) 背景下进行的密度比估计^[35]。基于 MMD 理论, 可以证明 KOOW 方法估计剂量反应曲线的方差是有界的, 这个有界方差保证了估计的一致性^[36]。模拟结果显示: 与包括匹配^[10]、非参双稳健方法^[9]以及 CBGPS 和 npCBGPS 相比, 在不同处理设置情况下无论拟合参数或是非参结局模型, KOOW 估计偏差和均方根误差方面都表现良好。

Tübbicke^[37]于 2020 年基于 Lindley 和 Kullback^[38]的熵度量 $h(w_i) = w_i \ln(w_i/q_i)$ 构建损失函数, 其中 q_i 是由研究者指定的基权重, 一般为均匀权重 $q_i = 1/n$, 提出了针对连续处理的熵平衡 (entropy balancing for continuous treatments, EBCT) 方法。熵平衡被证明具有类似于双稳健估计器的性质^[39], EBCT 权重在保证与基权重尽可能接近以控制极端权重的同时实现加权后样本处理与协变量间的皮尔逊相关性为零, 并且熵度量权重自然非负。此外, 由于 EBCT 优化问题的凸性质使其算法收敛速度相较 npCBGPS 快得多。模拟研究表明: 与基于 GPS 再加权类方法相比, EBCT 消除有限样本处理与协变量间相关性的表现更优。在估计处理效应时比 boosting 算法、CBGPS 和 npCBGPS 有着接近甚至更小的偏差以及更小的均方根误差, 并且个体最大权重的占比也最小。同样基于熵原理计

算权重的还有 Vegetabile 等^[40]的研究,与 EBCT 不同之处在于:在熵平衡算法约束条件中考虑了更高阶处理因素和协变量的影响。然而,约束条件阶数升高对算法寻求最优解是一个挑战,并且会导致实际分析时有效样本量的减少进而造成阳性假设的违背,因此作者建议变量阶数一般不超过 3。模拟研究表明:Vegetabile 等^[40]的方法有着与 CBGPS 和 npCBGPS 可比的性能,甚至更优。但与 EBCT 相比优劣如何尚不清楚。

$$\text{基于稳定权重的定义 } sw = \frac{f(z)}{f(z|x)} = \frac{f(z)f(x)}{f(z,x)},$$

Arbour 等^[41]于 2020 年提出了通过求解一个二分类问题直接计算 sw 的置换加权(permutation weighting, PW)方法。PW 主要分为两个步骤:首先,将观测数据集中的 Z 变量值重新排列构成置换数据集,重新排列的目的在于保持 Z 的边际分布不变同时 Z 与 X 相互独立。然后定义观测集中 Z 与 X 的联合分布为 P ,置换集中 Z 与 X 分布的乘积为 Q ;其次,通过估计 P 与 Q 的密度比^[42-43]即可获得 $sw = \frac{dQ}{dP}$ 。第二步的密度比估计可基于分类器进行,该分类器将密度比估计问题转化为由观测集和置换集构成训练集的二分类问题。PW 的优点是可利用现成的分类器估计均衡权重,并且适用于任意类型的处理变量。此外, PW 可以通过交叉验证最小化分类器的误差直接实现最小化不均衡性和因果推断误差。

4. 总结:在因果推断领域, GPS 方法已被广泛用于控制已测量混杂进而一致估计连续型处理因素的效应。然而 GPS 方法的首要步骤就是正确估计 GPS 或均衡权重,基于此,已发展出了许多方法,大致可分为基于模型估计 GPS 和基于均衡性估计 GPS 或权重两大类。基于模型的方法包括假定原始或经过 Box-Cox 转换数据处理因素服从正态分布的一般线性模型以及假定处理因素服从指数分布族的广义线性模型,但是二者都面临模型错误指定的风险。于是有学者提出使用以数据驱动方式通过交叉验证找到由数个预测算法(模型)加权组合最优形式的超级学习者。此外,boosting 机器学习算法因可以实现自动选择重要协变量甚至高阶项和交互项纳入模型而受到广泛关注。上述基于模型的方法估计所得 GPS 用于随后的回归调整、逆概率加权、匹配和分层等手段进行处理效应的估计前需要进行协变量均衡性检验,处理与协变量间达到均衡是正确估计处理效应的前提。但是基于模型

的方法不一定能达到很好的均衡性,近年来直接在均衡约束下估计 GPS 或权重的方法被提出,包括先估计 GPS 再构建稳定权重的 CBGPS 以及系列直接通过求解一个优化问题获得均衡权重的方法: npCBGPS 的目标经验似然函数并不总是凸的,所以不一定能找到最优解; CAEW 方法利用最小化权重的方差作为目标函数有效控制了极端权重; ROW 方法在 CAEW 基础上考虑将加权相关系数设置为一个可接受的 δ 水平而非 0,这种近似均衡通过牺牲一定的准确性以获取更大的精确性,同时确保有最优解^[32]。但是 ROW 并未约束加权前后处理与协变量的边际分布不变;将偏差分解加以控制的优化权重方法;利用核空间进行优化的 KOOW 方法;利用熵平衡原理的 EBCT 方法和 Vegetabile 等^[40]提出约束高阶情况的方法;基于分类器估计密度比构建稳定权重的 PW 方法。以上方法估计内容及各自应用时的软件实现见表 1。上述两大类估计 GPS

表 1 估计 GPS 或权重的方法总结及其软件实现

方法	估计内容	软件实现
模型类		
基于正态性假定的 一般线性模型 (及 Box-Cox 转换)	GPS	R: <i>lm()</i> 函数
广义线性模型	GPS	R: <i>glm()</i> 函数
超级学习者	GPS	R: <i>superlearner</i> 包
boosting 算法	GPS	R: <i>gbm</i> 包
均衡类		
CBGPS	GPS	R: <i>CBPS</i> 包中 <i>CBPS.fit()</i> 函数
npCBGPS	均衡权重	R: <i>CBPS</i> 包中 <i>npCBPS.fit()</i> 函数
GOAL	均衡权重	R: <i>CBPS</i> 包和 <i>lqa</i> 包
CAEW	均衡权重	R: <i>limSolve</i> 包中 <i>lse()</i> 函数
ROW	均衡权重	R: <i>ROW</i> 包
优化权重	均衡权重	R: <i>osqp</i> 包
KOOW	均衡权重	R interface of <i>Gurobi</i>
EBCT	均衡权重	Stata: <i>EBCT</i> ado-package; R: <i>WeightIt/entbal</i> 包
PW	稳定权重	标准二进制概率分类器

注: GPS: 广义倾向性评分; CBGPS: 协变量均衡广义倾向性评分; npCBGPS: 非参协变量均衡广义倾向性评分; GOAL: 广义结局自适应 LASSO; CAEW: 协变量关联消除权重; ROW: 稳健正交权重; KOOW: 核优化正交加权; EBCT: 针对连续处理的熵平衡; PW: 置换加权

和权重的方法在一致有效估计 DRF 方面的表现孰优孰劣尚未可知,仍需进一步研究。尤其在大数据时代,高维数据的易获得性使得弱无混杂假设更为合理,与此同时带来的维数诅咒等问题^[44]对统计方

法提出了更高要求,因此探讨以上方法在高维情况的适用性也是未来的一个研究方向。

利益冲突 所有作者声明无利益冲突

作者贡献声明 张雨:文章撰写;高倩:论文修改;王彤:论文修改、经费支持

参 考 文 献

- [1] Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects[J]. *Biometrika*, 1983, 70(1): 41-55. DOI: 10.1093/biomet/70.1.41.
- [2] 潘玉璟, 史欣莹, 蒋文波, 等. 基于倾向性评分匹配评价太极拳干预稳定性冠心病疗效的回顾性分析[J]. *中国循证心血管医学杂志*, 2021, 13(2): 150-153. DOI: 10.3969/j.issn.1674-4055.2021.02.05.
- [3] Pan YJ, Shi XY, Jiang WB, et al. Reviewing curative effect of Tai Chi on stable coronary heart disease: a retrospective analysis based on propensity score matching[J]. *Chin J Evid Based Cardiovasc Med*, 2021, 13(2): 150-153. DOI: 10.3969/j.issn.1674-4055.2021.02.05.
- [4] Imbens GW. The role of the propensity score in estimating dose-response functions[J]. *Biometrika*, 2000, 87(3): 706-710. DOI: 10.1093/biomet/87.3.706.
- [5] Mutz R, Daniel HD. The generalized propensity score methodology for estimating unbiased journal impact factors[J]. *Scientometrics*, 2012, 92(2): 377-390. DOI: 10.1007/s11192-012-0670-4.
- [6] Hirano K, Imbens GW. The propensity score with continuous treatments[M]//Gelman A, Meng XL. *Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives: An Essential Journey with Donald Rubin's Statistical Family*. Chichester: John Wiley & Sons, Ltd, 2004: 73-84. DOI: 10.1002/0470090456.ch7.
- [7] Kreif N, Grieve R, Díaz I, et al. Evaluation of the effect of a continuous treatment: a machine learning approach with an application to treatment for traumatic brain injury[J]. *Health Econ*, 2015, 24(9): 1213-1228. DOI: 10.1002/hec.3189.
- [8] Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology[J]. *Epidemiology*, 2000, 11(5): 550-560. DOI: 10.1097/00001648-200009000-00011.
- [9] Bang H, Robins JM. Doubly robust estimation in missing data and causal inference models[J]. *Biometrics*, 2005, 61(4): 962-973. DOI: 10.1111/j.1541-0420.2005.00377.x.
- [10] Kennedy EH, Ma ZM, McHugh MD, et al. Nonparametric methods for doubly robust estimation of continuous treatment effects[J]. *J R Stat Soc B*, 2017, 79(4): 1229-1245. DOI: 10.1111/rssb.12212.
- [11] Wu X, Mealli F, Kioumourtzoglou MA, et al. Matching on generalized propensity scores with continuous exposures [Z]. *arXiv:1812.06575*, 2018.
- [12] Brown DW, Greene TJ, Swartz MD, et al. Propensity score stratification methods for continuous treatments[J]. *Stat Med*, 2021, 40(5): 1189-1203. DOI: 10.1002/sim.8835.
- [13] Imai K, van Dyk DA. Causal inference with general treatment regimes: generalizing the propensity score[J]. *J Am Stat Assoc*, 2004, 99(467): 854-866. DOI: 10.1198/016214504000001187.
- [14] Austin PC. Assessing the performance of the generalized propensity score for estimating the effect of quantitative or continuous exposures on survival or time-to-event outcomes[J]. *Stat Methods Med Res*, 2019, 28(8): 2348-2367. DOI: 10.1177/0962280218776690.
- [15] Austin PC. Assessing covariate balance when using the generalized propensity score with quantitative or continuous exposures[J]. *Stat Methods Med Res*, 2019, 28(5): 1365-1377. DOI: 10.1177/0962280218756159.
- [16] Alejo J, Galvao AF, Montes-Rojas G. A practical generalized propensity-score estimator for quantile continuous treatment effects[J]. *Stata J*, 2020, 20(2): 276-296. DOI: 10.1177/1536867X20930997.
- [17] Wand MP, Marron JS, Ruppert D. Transformations in density estimation[J]. *J Am Stat Assoc*, 1991, 86(414): 343-353. DOI: 10.1080/01621459.1991.10475041.
- [18] Bia M, Flores CA, Flores-Lagunes A, et al. A Stata package for the application of semiparametric estimators of dose-response functions[J]. *Stata J*, 2014, 14(3): 580-604. DOI: 10.1177/1536867X1401400307.
- [19] Guardabascio B, Ventura M. Estimating the dose-response function through a generalized linear model approach[J]. *Stata J*, 2014, 14(1): 141-158. DOI: 10.1177/1536867X1401400110.
- [20] van der Laan MJ, Dudoit S, Keles S. Asymptotic optimality of likelihood-based cross-validation[J]. *Stat Appl Genet Mol Biol*, 2004, 3(1): 1-23. DOI: 10.2202/1544-6115.1036.
- [21] van der Laan MJ, Dudoit S. Unified cross-validation methodology for selection among estimators and a general cross-validated adaptive epsilon-net estimator: finite sample oracle inequalities and examples[R]. Berkeley: University of California, 2003.
- [22] van der Laan MJ, Dudoit S, van der Vaart AW. The cross-validated adaptive epsilon-net estimator[J]. *Stat Decis*, 2006, 24(3): 373-395. DOI: 10.1524/STND.2006.24.3.373.
- [23] Zhu YY, Coffman DL, Ghosh D. A boosting algorithm for estimating generalized propensity scores with continuous treatments[J]. *J Causal Inference*, 2015, 3(1): 25-40. DOI: 10.1515/jci-2014-0022.
- [24] McCaffrey DF, Ridgeway G, Morral AR. Propensity score estimation with boosted regression for evaluating causal effects in observational studies[J]. *Psychol Methods*, 2004, 9(4): 403-425. DOI: 10.1037/1082-989X.9.4.403.
- [25] Chattopadhyay A, Hase CH, Zubizarreta JR. Balancing vs modeling approaches to weighting in practice[J]. *Stat Med*, 2020, 39(24): 3227-3254. DOI: 10.1002/sim.8659.
- [26] Fong C, Hazlett C, Imai K. Covariate balancing propensity score for a continuous treatment: Application to the efficacy of political advertisements[J]. *Ann Appl Stat*, 2018, 12(1): 156-177. DOI: 10.1214/17-AOAS1101.
- [27] 孙倩. 不同混杂结构下广义倾向性评分估计法模拟比较及应用研究[D]. 太原: 山西医科大学, 2020. DOI: 10.27288/d.cnki.gsxyu.2020.000061.
- [28] Sun Q. Simulation study of generalized propensity score estimation methods under different confounding structures and its application[D]. Taiyuan: Shanxi Medical University, 2020. DOI: 10.27288/d.cnki.gsxyu.2020.000061.
- [29] Gao Q, Zhang Y, Liang J, et al. High-dimensional generalized propensity score with application to omics data[J]. *Brief Bioinform*, 2021: bbab331. DOI: 10.1093/bib/bbab331.
- [30] Yiu S, Su L. Covariate association eliminating weights: a unified weighting framework for causal effect estimation[J]. *Biometrika*, 2018, 105(3): 709-722. DOI: 10.1093/biomet/asy015.
- [31] Zubizarreta JR. Stable weights that balance covariates for estimation with incomplete outcome data[J]. *J Am Stat Assoc*, 2015, 110(511): 910-922. DOI: 10.1080/01621459.2015.1023805.
- [32] Santacatterina M. Robust weights that optimally balance confounders for estimating marginal hazard ratios[Z]. *arXiv:2010.07695*, 2020.
- [33] Wang YX, Zubizarreta JR. Minimal dispersion approximately balancing weights: asymptotic properties and practical considerations[J]. *Biometrika*, 2020, 107(1): 93-105. DOI: 10.1093/biomet/asz050.
- [34] Greifer NH. Estimating balancing weights for continuous treatments using constrained optimization[D]. Church Hill: University of North Carolina at Chapel Hill, 2020. DOI: 10.17615/dyss-b342.
- [35] Kallus N, Santacatterina M. Kernel optimal orthogonality weighting: a balancing approach to estimating effects of continuous treatments[Z]. *arXiv:1910.11972v1*, 2019.
- [36] Sugiyama M, Suzuki T, Kanamori T. Density ratio estimation in machine learning[M]. New York: Cambridge University Press, 2012.
- [37] Sondhi A, Arbour D, Dimmery D. Balanced off-policy evaluation in general action spaces[C]//Proceedings of the twenty third international conference on artificial intelligence and statistics. Palermo: PMLR, 2020: 2413-2423.
- [38] Tibbick S. Entropy balancing for continuous treatments [Z]. *arXiv:2001.06281*, 2020.
- [39] Lindley DV, Kullback S. Information theory and statistics [J]. *J Am Stat Assoc*, 1959, 54(288): 825-827. DOI: 10.2307/2282528.
- [40] Zhao QY, Percival D. Entropy balancing is doubly robust[J]. *J Causal Inference*, 2017, 5(1): 1-19. DOI: 10.1515/JCI-2016-0010.
- [41] Vegetabile BG, Griffin BA, Coffman DL, et al. Nonparametric estimation of population average dose-response curves using entropy balancing weights for continuous exposures[J]. *Health Serv Outcome*, 2021, 21(1): 69-110. DOI: 10.1007/S10742-020-00236-2.
- [42] Arbour D, Dimmery D, Sondhi A. Permutation weighting [C]//Proceedings of the 38th international conference on machine learning. Virtual Event: PMLR, 2021.
- [43] Qin J. Inferences for case-control and semiparametric two-sample density ratio models[J]. *Biometrika*, 1998, 85(3): 619-630. DOI: 10.1093/biomet/85.3.619.
- [44] Cheng KF, Chu CK. Semiparametric density estimation under a two-sample density ratio model[J]. *Bernoulli*, 2004, 10(4): 583-604. DOI: 10.3150/BJ/1093265631.
- [45] Huang MY, Chan KCG. Joint sufficient dimension reduction for estimating continuous treatment effect functions[J]. *J Multivariate Anal*, 2018, 168: 48-62. DOI: 10.1016/j.jmva.2018.07.005.