flexoweight: flexible overlap weighting

a propensity score weighting family + a weighting family for combining groups

Trang Quynh Nguyen
trang.nguyen@jhu.edu

(joint work with Betsy Ogburn and Liz Stuart)

JSM2022 session 285
2022-08-09

Outline

Current overlap weighting

The *semi-flexible* overlap weighting family

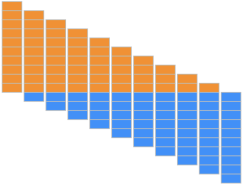
The *flexible* overlap weighting family

The rest

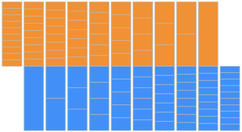
Problem: covariate nonoverlap

ATE estimation requires covariate overlap

but covariate nonoverlap is common



inverse probability weights (IPW) for covariate balancing are unbounded and may be extreme



What to do?

a common way to handle this is to drop observations in nonoverlap areas

usually based on propensity scores

- discrete set weight to zero
  - propensity scores δ-close to 0 or 1 (Crump et al. 2009)
  - a continuous-ized version (Yang and Ding 2018)
- smooth drop weight to zero
  - *original* overlap weights (Crump et al. 2006; Li, Morgan, and Zaslavsky 2018; Li and Li 2019)
  - other variants for the binary treatment case (Zhou, Matsouaka, and Thomas 2020; Matsouaka and Zhou 2020)
  - matching weights (Li and Greene 2013; Yoshida et al. 2017)

an alternative used in practice is trimming weights

Notation

original covariate density	$f_{\text{origin}}(x)$
target (ie weighted) covariate density	$f_{\text{target}}(x)$
tilting function	$h(x) = f_{\text{target}}(x)/f_{\text{origin}}(x)$
propensity score	$e_k(x)$ with k for group, or vector $\boldsymbol{e}(x)$

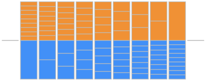
the weighting function $\propto h(x)/e_k(x)$

- balances covariates
- achieves $f_{\text{target}}(x)$

overlap weighting: $h(x) = 0$ if $\min[\boldsymbol{e}(x)] = 0$

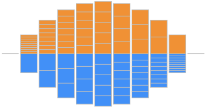
The original overlap weighting scheme (OOW)

for reference:
IPW if could restrict to overlap



original overlap weights (OOW)
designed to target an overlap subpop,
aka "moving the goalpost"

$$h(x) \propto \text{harmonic.mean}[\boldsymbol{e}(x)]$$



nice properties

- zeros out nonoverlap; shrink (amplify) where pscores are dissimilar (similar)
- if outcome Y homoscedastic conditional on treatment A and covariates X, leads to weighted ATE with the smallest variance bound among all WATEs

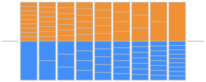
some reservations

- no pre-defined target population

  ... maybe should try to better characterize target population

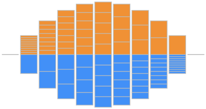
Consider 2 existing overlap weighting schemes

for reference:
IPW if could restrict to overlap



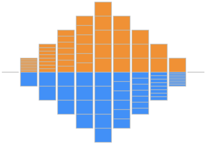
original overlap weights (OOW)
designed to target an overlap subpop,
aka "moving the goalpost"

$$h(x) \propto \text{harmonic.mean}[\boldsymbol{e}(x)]$$



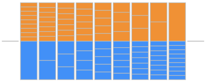
matching weights (MW)
designed to mimic 1-to-1 matching

$$h(x) \propto \min[\boldsymbol{e}(x)]$$

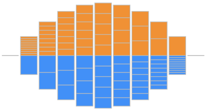


question 1

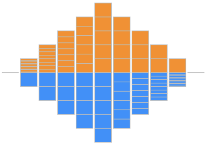
IPW if could restrict to overlap



original overlap weights (OOW)



matching weights (MW)



CAN WE HAVE SOME KIND OF CONTINUUM HERE?
– that gives more say on target population

Outline

Current overlap weighting

The *semi-flexible* overlap weighting family

The *flexible* overlap weighting family

The rest

Power means!

OOW: $h(x) \propto$ harmonic mean of $\boldsymbol{e}(x)$

- aka power mean with power param -1

MW: $h(x) \propto$ minimum of $\boldsymbol{e}(x)$

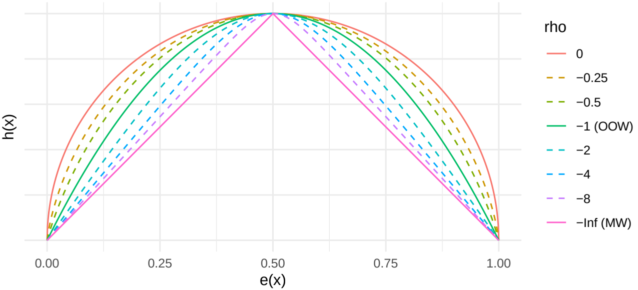
- aka power mean with power param $-\infty$

IPW w/in overlap area: $h(x) \propto$ arithmetic mean of $\boldsymbol{e}(x)$ $(= 1/K)$

- aka power mean with power param 1

Semi-flexible overlap weighting: a propensity score weighting family

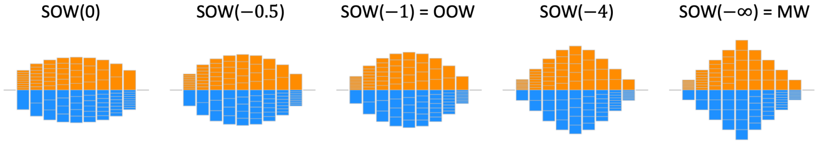
$$h(x) \propto \text{power.mean}[\boldsymbol{e}(x), \rho], \quad \rho \leq 0$$



power param ρ governs degree to which areas with dissimilar pscores are shrunk and areas with similar pscores areas are expanded

Semi-flexible overlap weighting: a propensity score weighting family

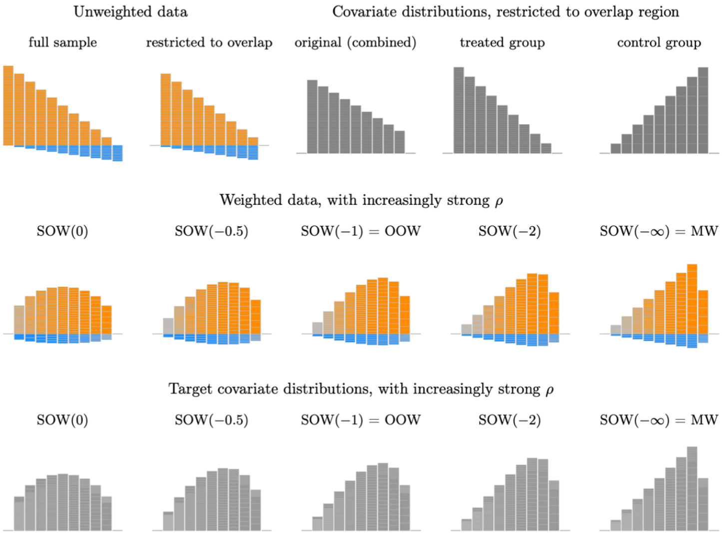
power param ρ governs degree to which areas with dissimilar pscores are shrunk and areas with similar pscores areas are expanded



the story looks simple so far, in the case where the groups are balanced

what would this look like if the groups are not balanced?

treat:control = 5:1



this moves the goalpost a lot!

An alternative view of the SOW family

$h(x)$ is a function of $\boldsymbol{e}(x)$,

which depends on both

(i) the relative density of x in the different groups – denoted $g_k(x)$
(ii) the relative sizes of the groups – denoted s_k

(in vector form, $\boldsymbol{g}(x)$ and $\boldsymbol{s}$)

could tease these apart

Dual representation

The weighting schemes above could be seen in two ways:

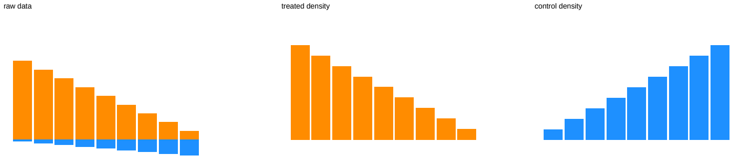
- as propensity score weighting (current view)
- as combination of groups, or "mixing" group covariate distributions (alternative view)

A motivation for the second view:

- sometimes we don't get a sample in which people are assigned one of several treatments
- rather we just get groups with different treatment conditions, eg a group on duloxetine and a group on vortioxetine from EHR data

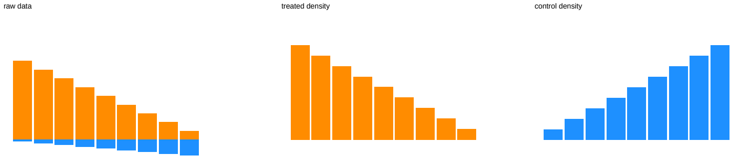
Teasing apart group density and group size

for simplicity, leave out nonoverlap



Teasing apart group density and group size

for simplicity, leave out nonoverlap



key connection between the two representations: $\frac{e_1(x)}{e_0(x)} = \frac{g_1(x)}{g_0(x)} \frac{s_1}{s_0}$

target density $\propto$ weighted power mean of group-specific densities where group-weight is power of group size

Powered-size-weighted power mean of group densities

OOW:

$$f_{\text{target}}(x) \propto \text{power.mean}(\boldsymbol{g}(x),\ \text{power} = -1,\ \text{group.weight} \propto \boldsymbol{s}^{-1})$$

ie inverse-size-weighted harmonic mean of group-specific densities

Powered-size-weighted power mean of group densities

OOW:

$$f_{\text{target}}(x) \propto \text{power.mean}(\boldsymbol{g}(x),\ \text{power} = -1,\ \text{group.weight} \propto \boldsymbol{s}^{-1})$$

ie inverse-size-weighted harmonic mean of group-specific densities

size ratio 5:1 $\longrightarrow$ combination ratio 1:5

Powered-size-weighted power mean of group densities

OOW:

$$f_{\text{target}}(x) \propto \text{power.mean}(\boldsymbol{g}(x),\ \text{power} = -1,\ \text{group.weight} \propto \boldsymbol{s}^{-1})$$

ie inverse-size-weighted harmonic mean of group-specific densities

simple combination:

$$f_{\text{target}}(x) \propto \text{power.mean}(\boldsymbol{g}(x),\ \text{power} = 1,\ \text{group.weight} \propto \boldsymbol{s})$$

ie size-weighted arithmetic mean of group-specific densities

Powered-size-weighted power mean of group densities

OOW:

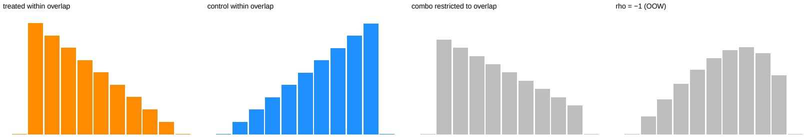
$$f_{\text{target}}(x) \propto \text{power.mean}(\boldsymbol{g}(x),\ \text{power} = -1,\ \text{group.weight} \propto \boldsymbol{s}^{-1})$$

ie inverse-size-weighted harmonic mean of group-specific densities

simple combination:

$$f_{\text{target}}(x) \propto \text{power.mean}(\boldsymbol{g}(x),\ \text{power} = 1,\ \text{group.weight} \propto \boldsymbol{s})$$

ie size-weighted arithmetic mean of group-specific densities



Second view of SOW family

$$f_{\text{target}}(x) \propto \text{power.mean}(\boldsymbol{g}(x),\ \text{power} = \rho,\ \text{group.weight} \propto \boldsymbol{s}^{\rho}), \quad \rho \leq 0$$

the group weights are tied to the power parameter!

Second view of SOW family

$$f_{\text{target}}(x) \propto \text{power.mean}(\boldsymbol{g}(x),\ \text{power} = \rho,\ \text{group.weight} \propto \boldsymbol{s}^{\rho}), \quad \rho \leq 0$$

the group weights are tied to the power parameter!

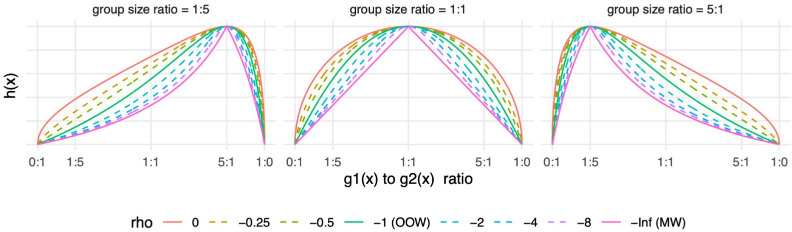
This family favors the smaller groups

Only $\rho = 0$ is group size "neutral":

$$f_{\text{target}}(x) \propto \text{power.mean}(\boldsymbol{g}(x),\ \text{power} = 0,\ \text{group.weight} \propto \boldsymbol{s}^{0} = 1)$$

ie unweighted geometric mean of group-specific densities

SOW: tilting function as a function of group density ratio, for varying group size ratio



question 2

Can we move the goalpost but not so much?

or CAN WE HAVE MORE SAY ABOUT WHERE THE GOALPOST GOES?

eg keep group weights proportional to size?

or something else?

Outline

Current overlap weighting

The *semi-flexible* overlap weighting family

The *flexible* overlap weighting family

The rest

Flexible overlap weighting: untie group weights from power parameter

$$f_{\text{target}}(x) \propto \text{weighted.power.mean}(\boldsymbol{g}(x),\ \text{power} = \rho,\ \text{group.weight} = \boldsymbol{\omega})$$

for $\rho \leq 0$

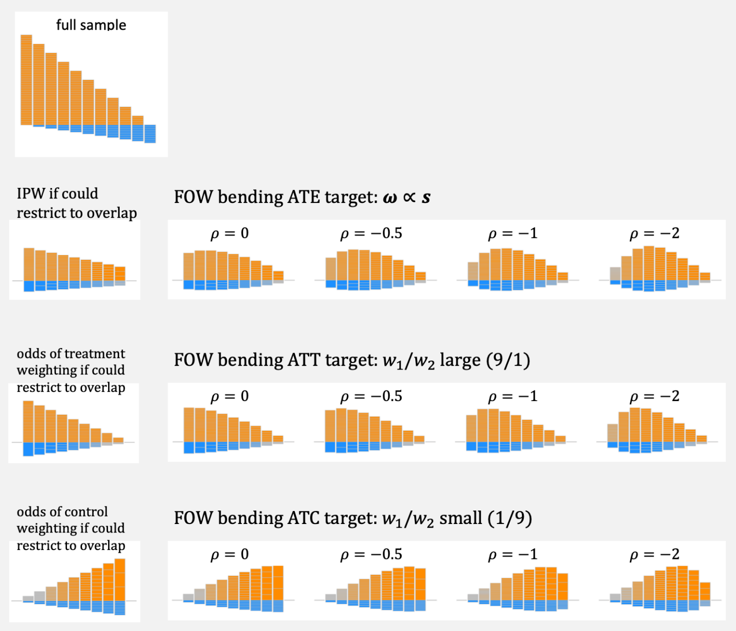
and $\boldsymbol{\omega} > 0$ specified based on intended/ideal target pop

How can this flexibility help?

roughly, we can have OW versions of any weighting scheme

(bending the goalpost but not quite moving it somewhere else)

example: ideal is ATE/ATT/ATC



example: ideal is ATE on an external target population

the usual way

- IPW within study sample to balance treatment groups
- odds weighting to mimic target covariate distribution

a flexoweight way

- stack samples, treat target sample as a group
- 3-way FOW, using $\boldsymbol{\omega}$ with large target weight, eg $(.8, .1, .1)$ or $(.9, .05, .05)$

how to choose ρ for FOW

$\rho = 0$ results in the least distortion of the target

dial it up (ie make it more negative) if some weights still too large

Outline

Current overlap weighting

The *semi-flexible* overlap weighting family

The *flexible* overlap weighting family

The rest

In progress

- weights estimation
- statistical inference
- illustrative applications
- package flexoweight

THANK YOU!

trang.nguyen@jhu.edu