

# Propensity score matching

10.05.21 07:20

- Content: calipers, propensity score overlap, trimming tails of a propensity score dist.

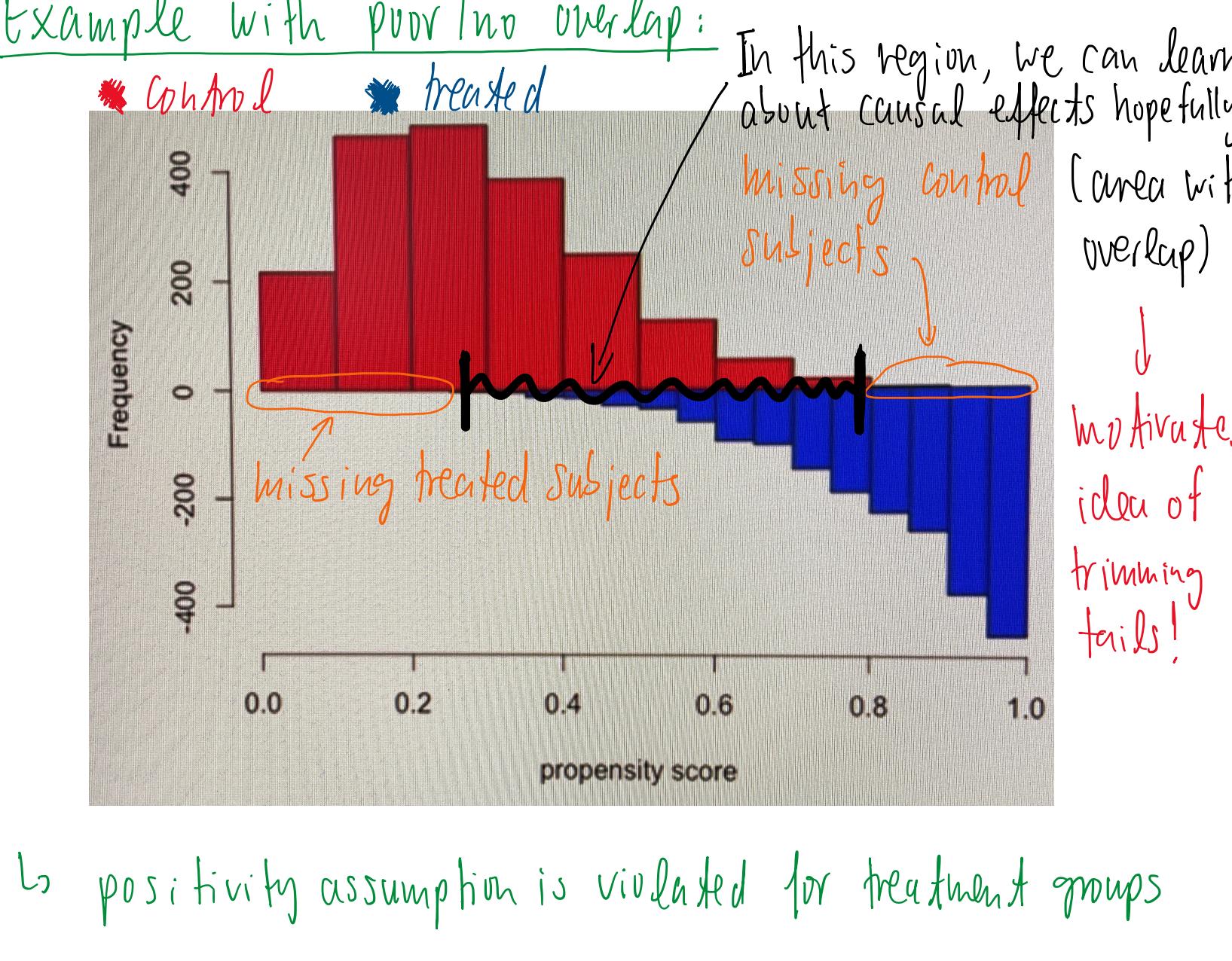
$\pi(x)$   
→ propensity score is a balancing score: so matching on the propensity score should achieve balance

- b/c the propensity score is a scalar, it lets us match subjects by solving the simplified problem based on a single variable  
↳ think of  $\pi(x)$  as a "summary" of the covariates  $X$ .

## Overlap:

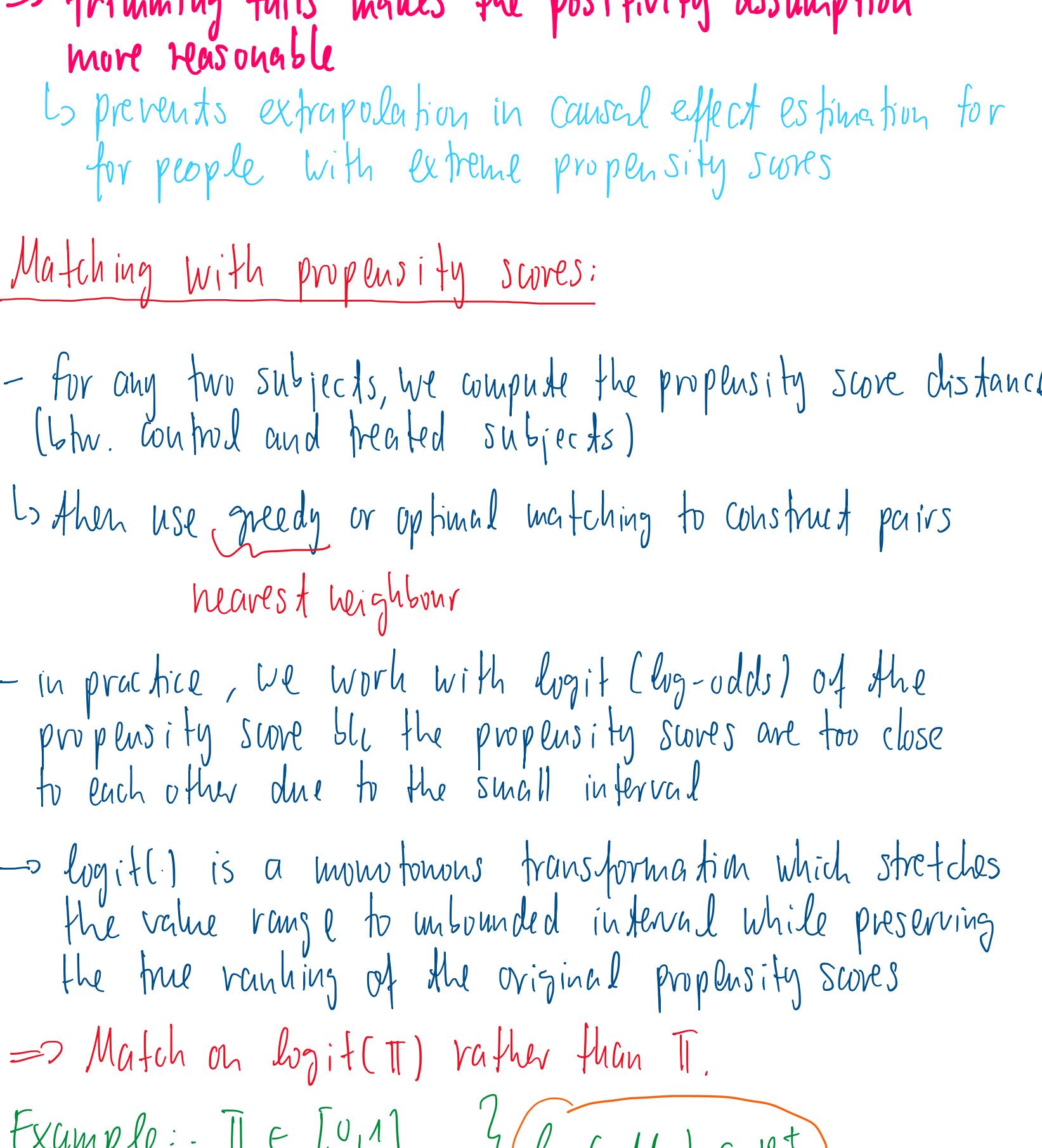
- after computing propensity score estimates
- before carrying out matching
  - ⇒ check for overlap in the propensity score dist.
  - use a plot to compare the dist. of prop. scores b/w. treated and control subjects

### Example:



Overlap: refers to the fact, that for any propensity score value, there are subjects available with this score in both treatment group.  
→ non-zero bars everywhere!

Overlap is a proxy for positivity assumption



↳ positivity assumption is violated for treatment groups

## Trimming tails:

- if we have a lack of overlap in the propensity score dist. then trimming tails is an option

removing subjects from dataset that have extreme propensity score values

→ e.g. control subjects whose prop. score is less than the minimum in the treatment group

→ e.g. treated subjects whose prop. score is greater than the maximum in the control group

⇒ trimming tails makes the positivity assumption more reasonable

↳ prevents extrapolation in causal effect estimation for people with extreme propensity scores

## Matching with propensity scores:

- for any two subjects, we compute the propensity score distance (b/w. control and treated subjects)

↳ then use greedy or optimal matching to construct pairs nearest neighbour

- in practice, we work with logit (log-odds) of the propensity score b/c the propensity scores are too close to each other due to the small interval

→ logit(.) is a monotonous transformation which stretches the value range to unbounded interval while preserving the true ranking of the original propensity scores

⇒ Match on  $\text{logit}(\pi)$  rather than  $\pi$ .

Example: -  $\pi \in [0, 1]$  {  
• odds  $\frac{\pi}{1-\pi} \in (0, +\infty)$  }  $\log(\text{odds}) \in \mathbb{R}^+$

unbounded from above.

$$\frac{0.1}{0.9} = \frac{1}{9} = 0.1111 \quad (0, +\infty)$$

$$\frac{0.9}{0.1} = \frac{9}{1} = 9 \quad (0, +\infty)$$

$$\frac{0.99}{0.01} = 99 \quad (0, +\infty)$$

$$\frac{0.999}{0.001} = 999 \quad (0, +\infty)$$

## Caliper:

- to ensure, that we do not accept "bad" matches, a caliper is the maximum distance that we are willing to tolerate b/w. two subjects.

- in practice a common caliper-value is  $0.2 \times \text{STD}$  of the logit of the propensity score.

### STEPS:

1. estimate prop. score (e.g. w/ logistic regression)

2. logit- transform each propensity score

3. take the STD of this transformed variable

4. set the caliper to  $0.2 \times \text{STD}$  from step 3.

⇒ **Smaller caliper**: ↗ less bias (i.e. better balance)

↳ more variance (b/c fewer matched pairs)

refers to the treatment effect estimate

## After propensity score matching:

- outcome analysis exactly the same as if we had matched directly on the covariates  $X$

↳ e.g. randomization tests (1)

conditional logistic regression (2)

stratified Cox model (3), etc.