## §4.5 Assessing balance

作法：检查 covariance balance 是否实现（在IPTW之后）

目标：Weighted sample ⟺ randomized trial

跟与之前 Propensity score 之后的检查方法相同，采用 standardized difference.

(Covariate balance can be checked on the weighted sample 这意味着检验对象. using standardized difference)

注意是对 weighting 之后的样本做检验，而不是原样本。

(1) Two ways: { Table 1 : summary statistics; plot

① Table 1: ① Weighted means mean (sd) of each covariate stratified on treated group and control group (using weights)

② standardized difference after weighting. (recall Smd is the difference in means between groups, divided by the pooled standard deviation.)

② plot

$$Smd = \left| \frac{\bar{X}_t - \bar{X}_c}{\sqrt{\frac{s_t^2 + s_c^2}{2}}} \right| \quad \text{绝对值}$$

(2) Standardized difference after weighting

- weighted means (variance) of each covariate stratified on treated group and control group.
- Take difference in weighted means and divide by an estimate of pooled weighted standardized deviation.

! Smd 越小越好， ~~>0.02~~ ~~0.4~~ | 0.2 < · < 0.4 | < 0.2 imbalance

| | Raw data | | | Weighted data | | |
|---|---|---|---|---|---|---|
| | control | treated | smd | control | treated | smd |
| n | | | | | | |
| v1 | | | | | | |
| v2 | | | | | | |
| ⋮ | | | | | | |

(3) If imbalance after weighting

① refine propensity score ~~model~~ model

- summary(weight)
- tail(sort(weight))
- head(sort(weight))

② Summary Statistics (quantail/min/max/mean/median)

A person who was likely to be treated (given covariates)

treated $\frac{9}{10}$ control $\frac{1}{10}$ but wasn't:

- ○ will have a large weight
- ○ will have a small weight

$P(A_i=1|X_i) = \frac{9}{10}$

$P(A_i=0|X_i) = \frac{1}{10} \Rightarrow \text{weight} = \frac{1}{\frac{1}{10}} = 10$

## §4.6 Distribution of weights

### 4.6.1 Intuition

大 weight ⟶ large standard error

- 极端的一个举例

1 person ⟶ weight 10000 ⟶ 1个人代表10000个人

↓

outcome: 1 (10000个人的outcome都依赖于1个人，有很大的噪声，很大的误差.)

相当于10000个人的 outcome 都是1.

如果1个人的 outcome data 很显著地影响了参数估计，那么 standard error 就会非常大.

Prefer person to have not too large weight.

### 4.6.2 Bootstrapping

Estimate standard error : bootstrapping

Step1: Randomly sample from original data (有放回)

Step2: Use sample data to estimate parameters

Step3: Repeat the procedure for k times.

Step4: Take the variance of k estimators as the estimate of standard error.

问题：当一些 weight 非常大的 sample 被采样时，该样本对估计量的作用非常显著.

当这些 weight 非常大的样本不被采样时，作用对估计量的作用非常明显，所以大 weight 的样本会使估计量 standard error 的波动很大，越会很大.

### 4.6.3 Relationship with positivity assumption.

$\text{weight} = \frac{1}{P(A_i=1|X_i)}$ 若 weight ↑, $P(A_i=1|X_i)$ ↓

假设 weight 非常非常大，那么 $P(A_i=1|X_i)$ 就会非常趋近于0. 这就可能违背 positivity assumption. (near violation)

### 4.6.4 Checking Weights

(1) Plot { density plot (所有样本的 weight 频数) 点(person) 频率; weight-index plot