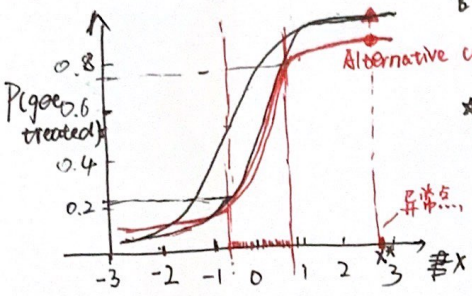
§4.7 Remedies for large weights.

§4.7.1 Very large weights: investigative step

First step: Why the weights are large?

☆识别有large weights的对象 { 查看他们的数据是否存在问题.
查看Propensity score是否有问题.



☆异常点...X^*在X上的取值非常大，此时P(get treated)趋向于1，若其最终没有get treated，那么就会导致large weight.

在两个vertical bar之间，alternative curve和Original curve几乎重合；但在该区间之外，X^*处的Propensity score明显小了，因此避免了large weight.

§4.7.2 Trimming the tails

在进行inverse probability之前基于propensity score的分布"修剪尾巴".

思路：trim the extreme data { People who are very likely to be treated
People who are very likely to be control.

trim the tails: get rid of some subjects from the data. → have extreme values of propensity score.

目的：Trim the tails can eliminate some of the extreme weights.

! Common trimming strategy:

(1) Remove treated subjects whose propensity scores >98th percentile from the distribution among controls.

(Ps >98th 的 treated subject 很难分配到control，因此去掉)

(2) control <2nd treated

(Ps <2nd 的 control subject 很难被treated，因此也去掉)

缺点：trimming the tails change the population

§4.7.3 Weight truncation.

思路：The weight for anyone who had a value higher than that truncation point, to set to the "maximum weight". (upper limit, if someone's weight is 1000, set it to ~~1000~~ instead.)
upper limit

Step1: Determine a maximum allowable weight.
① a specific value (e.g., 100)
② based on percentile (98th, 99th, ...)

Step2: 如果weight > 你定的maximum allowable weight, set it to the maximum allowable weight.

§4.7.4 bias-variance trade of

Truncation: bias, smaller variance
No truncation: unbiased, large variance

? Truncating extremely large weight ⇓ ~~results in~~ estimators with lower MSE.