

§ Data Example: Propensity score match in R.

o Right heart catheterization data

Step 1: load packages: tableone, MatchIt

read in data: load(url("website"))

view data: View(rhc)
 * 点击右侧全局变量中的Data=tableone即有

Step 2: Create new data set. only variables that will be used + change character to numeric.

e.g. ARF <- as.numeric(rhc\$arrest == 'ARF')

female <- as.numeric(rhc\$sex == 'Female') → 转换为0,1型

died <- as.numeric(rhc\$death == 'Yes') → 转换为0,1型

* treatment <- as.numeric(rhc\$swangl == 'RHC')

原本数值型的变量就不用变, 重命名即可.

age <- rhc\$age

meanbp1 <- rhc\$meanbp1

aps <- rhc\$aps1.

new dataset.

mydata <- cbind(ARF, ..., female, ~~meanbp1, aps, age~~, treatment, died)
 ← X → | ← A → ← outcome →

建立data frame形式.

mydata <- data.frame(mydata)

Step 3: Fit Pps model.

(1) 估计 Propensity score. logistic回归.

model <- glm(treatment ~ ARF +

... + aps,

family = binomial(),

data = mydata)

$T_i = P(A=1|X_i)$

Outcome: A=1

回归变量: X.

作用: Outcome is binary default = logit link → logistic回归

输出回归结果. (可查看系数)

summary(psmodel)

运行. 看内容包含什么 coefficients, p-values, std. error, fitted.value

创建 Pps.

pscore <- psmodel\$fitted.values
 predict values

Step 4: plot ⇒ check overlap.

Step 5.1: do Pps Match. (并不需要先进行 Pps Match 可替代 Step 3 + Matching, 也就是不必先分步计算 Ppscore.)

Use MatchIt for Pps, 使用最近邻 matching.

m.out <- matchit(treatment ~ ARF + ... + aps, data = mydata, method = "nearest")
 功能: 计算 Ppscore 做 Matching
 summary(m.out) ☆ 可以查看

plots

plot(m.out, type = "jitter")

plot(m.out, type = "hist")

Distribution of Ppscores
 Unmatched Treatment Units (未匹配的)
 通常为0, 因为每个 treated + treated subject 都会匹配上一个或多个 control

Matched Treatment Units (匹配的)
 Matched control Units (匹配的)
 主要看这俩是否相似, 代表相似

Unmatched control units
 倾向性小的 Pps. 因为没和 treated 匹配

Step 5: Matching Without Caliper

Match on logit(Ps) without caliper

load package: Match. *(不能自动计算Ppscore, 需要先计算好Ppscore再代入)*

do greedy matching on logit(Ps)

psmatch <- Match[Tr = mydata\$treatment,
M=1, *What the treatment is?*
Which type of Matching? M=1: One to One Matching

X=logit(pscore),

replace = FALSE)

people cannot be re-matched, once they matched, they are excluded in the data set.

matched <- mydata[unlist(psmatch[c("index.treated",

ID	treated	control
1	1	10
2	2	3
3	3	20

Xvars <- c("ARF", ..., "aps")

matchedtab1 <- CreateTableOne(vars = xvars, strata = "treatment",
data = ~~matched~~ *matched*, test = FALSE)

Print(matchedtab1, smd = TRUE)

Table 1. Stratified by treatment

	0	1	SMD
n	2184	2184	
ARF(mean(sdi)	0.4P(0.50)	0.42(0.4P)	
...			
aps	

Step 5.1: Match with caliper

psmatch <- Match(Tr = , M=1, X=logit),
replace = FALSE,
caliper = 0.2)

matched <- mydata[unlist(psmatch
[c("index.treated",
"index.control")])]

Xvars <- c("ARF", ..., "aps")

matchedtab1 <- CreateTableOne(vars = ,
strata = " ",
data = matched,
test = FALSE).

Step 6.

对比:
Without caliper
With caliper

2184 matched pairs
Causal risk difference:
0.04 (0.012, 0.068)
1000 matched pairs
Causal risk difference:
0.03 (0.004, 0.059)

$$SMD = \frac{\bar{X}_t - \bar{X}_c}{\sqrt{\frac{S_t^2 + S_c^2}{2}}}$$

Step 6: Outcomes analysis.

两样本 t-test.

y-trt <- matched\$died, y-con <- matched\$died [matched\$treatment == "RHF"]

~~t-test (outcome)~~

y* <- y-trt - y-con

t.test(y*)