Marginal Standardization for the ATE, ATT, and ATU

Ashley I Naimi

Spring 2024

Contents

1 Introduction

Here, we show how marginal standardization and IP weighting can be used to compute the ATE, ATT, and ATU on the risk difference, risk ratio, and odds ratio scales. Practical issues are emphasized, within some important theoretical contexts.

We'll use the NHEFS data:

```
file_loc <- url("https://bit.ly/47ECRcs")

#' This begins the process of cleaning and formatting the data
nhefs <- read_csv(file_loc) %>%
  select(qsmk, wt82_71, sex, age, race, alcoholfreq, income,
         marital, school, asthma, bronch,
         starts_with("price"),
         starts_with("tax"),
         starts_with("smoke"),
         smkintensity82_71) %>%
  mutate(wt_delta = as.numeric(wt82_71 >= median(wt82_71, na.rm = T)),
         income=as.numeric(income>15),
         marital=as.numeric(marital>2),
         alcoholfreq=as.numeric(alcoholfreq>1)) %>%
  select(-wt82_71) %>%
  na.omit(.)

dim(nhefs)
```

```
## [1] 1422   20
```

```
nrow(nhefs[nhefs$qsmk==1,])
```

```
## [1] 356
```

```
nrow(nhefs[nhefs$qsmk==0,])
```

```
## [1] 1066
```

```
names(nhefs)
```

```
##  [1] "qsmk"              "sex"               "age"
##  [4] "race"              "alcoholfreq"       "income"
##  [7] "marital"           "school"            "asthma"
## [10] "bronch"            "price71"           "price82"
## [13] "price71_82"        "tax71"             "tax82"
## [16] "tax71_82"          "smokeintensity"    "smokeyrs"
## [19] "smkintensity82_71" "wt_delta"
```

1.1 Marginal Standardization of the ATE

We'll start with marginal standardization for the ATE, which we've seen a few times before. We won't use the stratified modeling approach for the ATE, but we'll comment on why it's important to consider for the ATT and ATU:

```
formulaVars <- paste(names(nhefs)[c(1:11,14,17)],collapse = "+")
modelForm <- as.formula(paste0("wt_delta ~", formulaVars))
modelForm
```

```
## wt_delta ~ qsmk + sex + age + race + alcoholfreq + income + marital +
##     school + asthma + bronch + price71 + tax71 + smokeintensity
```

```
mod1 <- glm(modelForm, data = nhefs, family = binomial("logit"))
mu1 <- mean(predict(mod1, newdata = transform(nhefs, qsmk = 1), type = "response"))
mu0 <- mean(predict(mod1, newdata = transform(nhefs, qsmk = 0), type = "response"))
RD_ATE <- mu1 - mu0
RR_ATE <- mu1/mu0
OR_ATE <- (mu1/(1 - mu1))/(mu0/(1 - mu0))


## bootstrapping for CIs with a for loop instead of the boot package

RD_ATEb <- RR_ATEb <- OR_ATEb <- NULL
R = 2000
```

```
for(i in 1:R){
  index <- sample(1:nrow(nhefs), nrow(nhefs), replace = T)
  boot_dat <- nhefs[index,]
  mod1 <- glm(modelForm, data = boot_dat, family = binomial("logit"))
  mu1_ <- mean(predict(mod1, newdata = transform(boot_dat, qsmk = 1), type = "response"))
  mu0_ <- mean(predict(mod1, newdata = transform(boot_dat, qsmk = 0), type = "response"))
  RD_ATEb <- rbind(RD_ATEb, mu1_ - mu0_)
  RR_ATEb <- rbind(RR_ATEb, mu1_/mu0_)
  OR_ATEb <- rbind(OR_ATEb, (mu1_/(1 - mu1_))/(mu0_/(1 - mu0_)))
}



UCL_RD_ATE <- RD_ATE + 1.96*sd(RD_ATEb)
LCL_RD_ATE <- RD_ATE - 1.96*sd(RD_ATEb)

UCL_RR_ATE <- exp(log(RR_ATE + 1.96*sd(RR_ATEb)))
LCL_RR_ATE <- exp(log(RR_ATE - 1.96*sd(RR_ATEb)))

UCL_OR_ATE <- exp(log(OR_ATE + 1.96*sd(OR_ATEb)))
LCL_OR_ATE <- exp(log(OR_ATE + 1.96*sd(OR_ATEb)))

ate_tab <- rbind(c(RD_ATE,LCL_RD_ATE,UCL_RD_ATE),
                 c(RR_ATE,LCL_RR_ATE,UCL_RR_ATE),
                 c(OR_ATE,LCL_OR_ATE,UCL_OR_ATE))



rownames(ate_tab) = c("Risk Difference", "Risk Ratio", "Odds Ratio")
```

```
knitr::kable(round(ate_tab,2),
             caption = "Average Treatment Effect Estimates of Quitting Smoking
             on Greater than Median Weight Change among 1,422
             Individuals in the NHEFS Data, 1971-1982.",
             col.names = c("Estimate", "LCL", "UCL"),
             "simple")
```

Table 1: Average Treatment Effect Estimates of Quitting Smoking on Greater than Median Weight Change among 1,422 Individuals in the NHEFS Data, 1971-1982.

	Estimate	LCL	UCL
Risk Difference	0.14	0.08	0.20
Risk Ratio	1.31	1.17	1.45
Odds Ratio	1.78	2.23	2.23

1.2 Marginal Standardization of the ATT

To use marginal standardization to quantify the average treatment effect on the treated (ATT) , we need to observe a few conditions (**?**). First, on the difference scale, we note that ATT is defined as:

The effect of treatment on the treated is the same as the average treatment effect on the treated. The shorthand abbreviations are usually ETT or ATT, and these are the same quantity

$$E(Y^{x=1} - Y^{x=0} \mid X = 1)$$

which is equivalent to:

$$E(Y^{x=1} \mid X = 1) - E(Y^{x=0} \mid X = 1)$$

So we need two means to compute the ATT: the average outcome that would be observed if the exposure was set to 1 *among those who were exposed*, and the average outcome that would be observed if the exposure was set to 0 *among those who were exposed*. By counterfactual consistency and no interference, the average outcome that would be observed if the exposure was set to 1 among those who were actually exposed is the average of the observed outcomes. For this reason, we need not model the first term in the ATT equation, we can simply take the mean of the outcome among those who were exposed and use that as the first term in the ATT equation:

```
mean(nhefs[nhefs$qsmk == 1,]$wt_delta)
```

```
## [1] 0.5926966
```

However, for the second term in the ATT equation, we need the average outcome that would be observed if everyone's exposure was set to 0, *among*

those who were exposed. To get this mean, we need to model. But we have to consider how to specify this model carefully. To see why, consider this simple example:

```
## here's our dataset
head(dat)
```

```
##            y x c
## 1  5.6958622 1 0
## 2 -0.5373377 1 1
## 3  1.9574312 1 0
## 4  1.8488057 1 1
## 5 -1.0850572 1 1
## 6  0.8655124 0 0
```

```
dim(dat)
```

```
## [1] 10000000        3
```

```
# compute the first term of the ATT
y1_pred <- mean(dat[dat$x==1,]$y)

# fit a simple model
modmod <- lm(y ~ x + c, data = dat)

# compute the second term of the ATT
y0_pred <- mean(predict(modmod,
                        newdata = transform(subset(dat, x == 1), x = 0),
                        type = "response"))

# ATT from the first approach
y1_pred - y0_pred
```

```
## [1] 0.762865
```

```r
# fit a model stratified by the exposure, in this
# case among those with x = 0
modmod <- lm(y ~ c,
             data = subset(dat, x == 0))

# compute the second term of the ATT from the second model
y0_pred_ <- mean(predict(modmod,
                         newdata = transform(subset(dat, x == 1)),
                         type = "response"))

# ATT from the second approach
y1_pred - y0_pred_
```

```
## [1] -1.316083
```

```r
# models make a difference
```

When estimating the ATT, the example above shows that it's important to fit a flexible model. One way of doing this is fitting a (flexible) model among **unexposed** observations.

```r
formulaVars <- paste(names(nhefs)[c(2:11,14,17)],collapse = "+")
modelForm <- as.formula(paste0("wt_delta ~", formulaVars))
modelForm
```

```
## wt_delta ~ sex + age + race + alcoholfreq + income + marital +
##     school + asthma + bronch + price71 + tax71 + smokeintensity
```

```r
mod0 <- glm(modelForm, data = subset(nhefs, qsmk == 0),
            family = binomial("logit"))

mu1_att <- mean(nhefs[nhefs$qsmk == 1,]$wt_delta)
mu0_att <- mean(predict(mod0, newdata = subset(nhefs, qsmk == 1),
                        type = "response"))
```

```
RD_ATT <- mu1_att - mu0_att
RR_ATT <- mu1_att/mu0_att
OR_ATT <- (mu1_att/(1 - mu1_att))/(mu0_att/(1 - mu0_att))

## bootstrapping for CIs with a for loop instead of the boot package
RD_ATTb <- RR_ATTb <- OR_ATTb <- NULL
R = 2000
for(i in 1:R){
  index <- sample(1:nrow(nhefs), nrow(nhefs), replace = T)
  boot_dat <- nhefs[index,]

mod0 <- glm(modelForm, data = subset(boot_dat, qsmk == 0),
            family = binomial("logit"))

mu1_att_ <- mean(boot_dat[boot_dat$qsmk == 1,]$wt_delta)
mu0_att_ <- mean(predict(mod0, newdata = subset(boot_dat, qsmk == 1),
                         type = "response"))

  RD_ATTb <- rbind(RD_ATTb, mu1_att_ - mu0_att_)
  RR_ATTb <- rbind(RR_ATTb, mu1_att_/mu0_att_)
  OR_ATTb <- rbind(OR_ATTb, (mu1_att_/(1 - mu1_att_))/(mu0_att_/(1 - mu0_att_)))
}



UCL_RD_ATT <- RD_ATT + 1.96*sd(RD_ATTb)
LCL_RD_ATT <- RD_ATT - 1.96*sd(RD_ATTb)

UCL_RR_ATT <- exp(log(RR_ATT + 1.96*sd(RR_ATTb)))
LCL_RR_ATT <- exp(log(RR_ATT - 1.96*sd(RR_ATTb)))

UCL_OR_ATT <- exp(log(OR_ATT + 1.96*sd(OR_ATTb)))
LCL_OR_ATT <- exp(log(OR_ATT + 1.96*sd(OR_ATTb)))

att_tab <- rbind(c(RD_ATT,LCL_RD_ATT,UCL_RD_ATT),
```

```
                  c(RR_ATT,LCL_RR_ATT,UCL_RR_ATT),
                  c(OR_ATT,LCL_OR_ATT,UCL_OR_ATT))


rownames(att_tab) = c("Risk Difference", "Risk Ratio", "Odds Ratio")
```

```
knitr::kable(round(att_tab,2),
             caption = "Average Treatment Effect on the Treated Estimates
             of Quitting Smoking on Greater than Median Weight Change among 1,422
             Individuals in the NHEFS Data, 1971-1982.",
             col.names = c("Estimate", "LCL", "UCL"),
             "simple")
```

Table 2: Average Treatment Effect on the Treated Estimates of Quitting Smoking on Greater than Median Weight Change among 1,422 Individuals in the NHEFS Data, 1971-1982.

	Estimate	LCL	UCL
Risk Difference	0.14	0.08	0.20
Risk Ratio	1.31	1.16	1.45
Odds Ratio	1.75	2.19	2.19

1.3 Marginal Standardization of the ATU

The same strategy can be used to compute treatment effect on the untreated. The key here is to compute the mean outcome among the untreated, then fit a flexible model in the treated and use it to predict the outcome that would be observed if the untreated were treated.

References