G-Computation

Vishwajith Reddy Anagandula

12/5/2021

Read Data, pre processing, and omit NAs

```r
ObsData <- read.csv("slpexcov1517.csv")

ObsData <- ObsData %>% select(-SEQN, -exminwk, -slphrs, -household, -income, -snoring,
                              -apnea, -bmicat, -smoke, -alcohol, -phq9)

ObsData <- ObsData %>% mutate(A = targetex) %>% mutate(Y = targetslp) %>%
  select(-targetex, -targetslp)

ObsData <- na.omit(ObsData)
```

Regression Model

```r
reg.model <- glm(Y ~ age + factor(A) + factor(raceeth) + factor(educ) +
                   factor(marital) + bmi + waist + factor(depressed),
                 family = "binomial", data = ObsData)
summary(reg.model)
```

```
##
## Call:
## glm(formula = Y ~ age + factor(A) + factor(raceeth) + factor(educ) +
##     factor(marital) + bmi + waist + factor(depressed), family = "binomial",
##     data = ObsData)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.2416   0.4576   0.6234   0.7374   1.2798
##
## Coefficients:
##                     Estimate Std. Error z value Pr(>|z|)
## (Intercept)         1.367665   0.346292   3.949 7.83e-05 ***
## age                -0.008757   0.003680  -2.379 0.017341 *
## factor(A)1          0.328647   0.093760   3.505 0.000456 ***
## factor(raceeth)2   -0.155546   0.122161  -1.273 0.202915
## factor(raceeth)3   -0.704612   0.118080  -5.967 2.41e-09 ***
## factor(raceeth)4   -0.208653   0.136504  -1.529 0.126376
## factor(educ)2      -0.142141   0.124364  -1.143 0.253060
## factor(educ)3      -0.049859   0.124047  -0.402 0.687730
## factor(educ)4       0.447969   0.144199   3.107 0.001892 **
## factor(marital)2   -0.015536   0.092465  -0.168 0.866570
## bmi                -0.081716   0.021169  -3.860 0.000113 ***
## waist               0.027740   0.008479   3.272 0.001069 **
## factor(depressed)1 -0.475600   0.151744  -3.134 0.001723 **
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 3603.6  on 3405  degrees of freedom
## Residual deviance: 3472.4  on 3393  degrees of freedom
## AIC: 3498.4
##
## Number of Fisher Scoring iterations: 4
```

Predict the counterfactual outcomes when the $A = 1$ for everyone and $A = 0$

```r
txt <- control <- ObsData
txt$A <- 1
control$A <- 0

predictY.txt <- predict(reg.model, newdata = txt, type = "response")
predictY.control <- predict(reg.model, newdata = control, type = "response")
```

Evaluate the statistical parameter

```r
mean(predictY.txt - predictY.control)
```

```
## [1] 0.05367993
```