Sensitivity Analysis

Keyu Mao

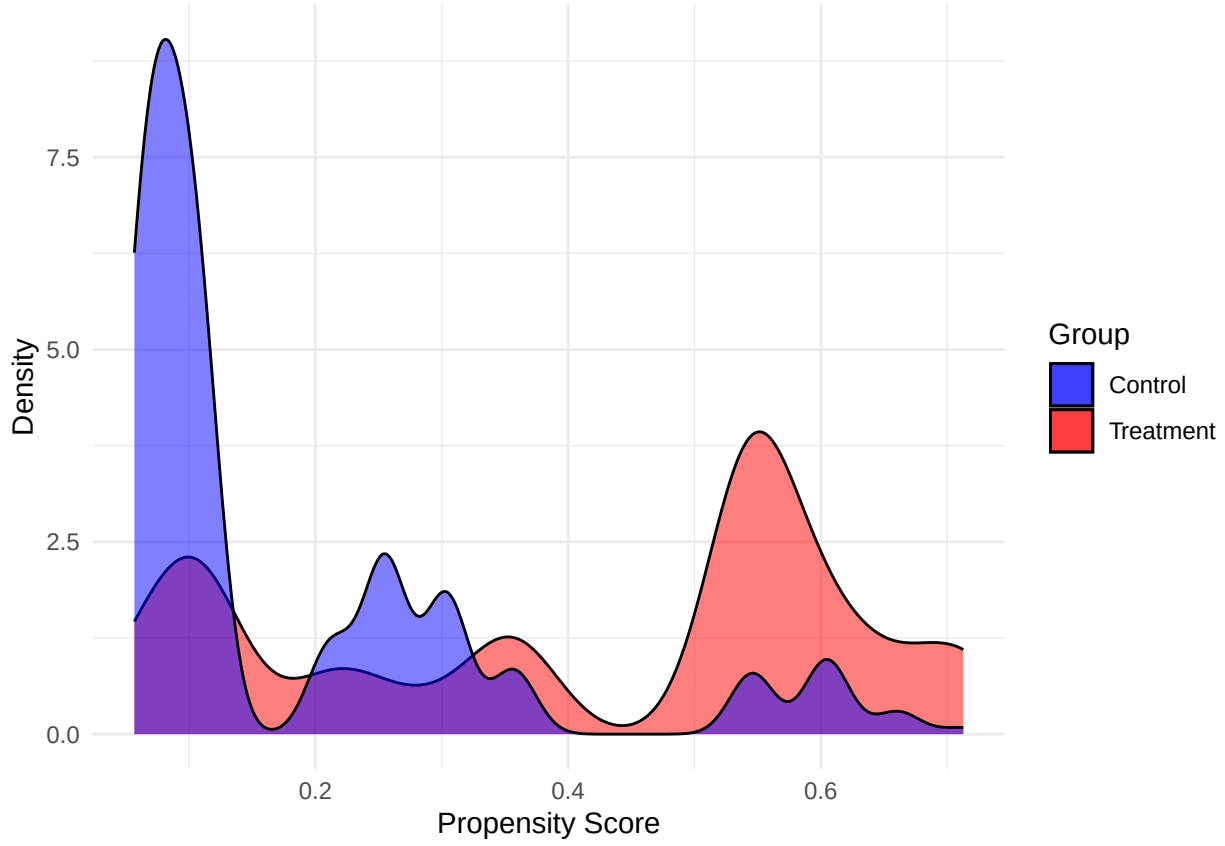
2023/6/8

```
data <- read.csv("final_data.csv")
head(data)
```

```
##   X Country AQI.Value AQI.Category CO.AQI.Value CO.AQI.Category Ozone.AQI.Value
## 1 0     123        51            1            1                0              36
## 2 1      71        41            0            1                0               5
## 3 2     108        66            1            1                0              39
## 4 3       0        34            0            1                0              34
## 5 4      37        22            0            0                0              22
## 6 5     164        54            1            1                0              14
##   Ozone.AQI.Category NO2.AQI.Value NO2.AQI.Category PM2.5.AQI.Value
## 1                  2             0                0              51
## 2                  0             1                0              41
## 3                  2             2                1              66
## 4                  2             0                0              20
## 5                  1             0                0               6
## 6                  0            11                2              54
##   PM2.5.AQI.Category
## 1                  1
## 2                  0
## 3                  1
## 4                  0
## 5                  0
## 6                  1
```

```
library(MatchIt)
library(ggplot2)
ps_formula <- CO.AQI.Category~Ozone.AQI.Category+NO2.AQI.Category
ps_model <- matchit(ps_formula, data = data, method = "nearest")
ps_scores <- ps_model$distance
data$propensity_score <- ps_scores
# Subset data for treatment and control groups
treatment_data <- data[data$CO.AQI.Category == 1, ]
control_data <- data[data$CO.AQI.Category == 0, ]

# Create density plots for propensity scores
ggplot() +
  geom_density(data = treatment_data, aes(x = propensity_score, fill = "Treatment"), alpha = 0.5) +
  geom_density(data = control_data, aes(x = propensity_score, fill = "Control"), alpha = 0.5) +
  labs(x = "Propensity Score", y = "Density", fill = "Group") +
  scale_fill_manual(values = c("blue", "red")) +
  theme_minimal()
```



Matching

```
match.method = matchit(CO.AQI.Category~propensity_score, data=data, method="nearest",ratio = 1)
match_data = match.data(match.method)
att <- function(match_data){
  ind_treat = which(match_data$CO.AQI.Category == 1)
  ind_control = which(match_data$CO.AQI.Category == 0)
  treat = match_data$PM2.5.AQI.Value[ind_treat]
  control = match_data$PM2.5.AQI.Value[ind_control]
  mean(treat)-mean(control)
}

att_v = att(match_data)
new_data = data
new_data$CO.AQI.Category = 1 - data$CO.AQI.Category
match.method = matchit(CO.AQI.Category~propensity_score, data=new_data, method="nearest",ratio = 1)
```

```
## Warning: Fewer control units than treated units; not all treated units will get
## a match.
```

```
match_data = match.data(match.method)
atc <- function(match_data){
  ind_treat = which(match_data$CO.AQI.Category == 0)
  ind_control = which(match_data$CO.AQI.Category == 1)
  treat = match_data$PM2.5.AQI.Value[ind_treat]
  control = match_data$PM2.5.AQI.Value[ind_control]
```

```
  mean(treat)-mean(control)
}
atc_v = atc(match_data)
prob = mean(data$CO.AQI.Category)
ATE = prob*att_v+(1-prob)*atc_v
cat("ATE is:", ATE)
```

```
## ATE is: 63.39343
```

Weighting

```
data$weights <- ifelse(data$CO.AQI.Category==1, 1/data$propensity_score, 1/(1-data$propensity_score))

E_Y_1 <- sum(data$PM2.5.AQI.Value*data$weights*data$CO.AQI.Category)/sum(data$weights*data$CO.AQI.Categ

E_Y_0 <- sum(data$PM2.5.AQI.Value*data$weights*(1-data$CO.AQI.Category))/sum(data$weights*(1-data$CO.AQ

cat("The ACE computed under weighted is:", E_Y_1-E_Y_0)
```

```
## The ACE computed under weighted is: 50.55991
```

Strata

```
model.stra <- matchit(ps_formula, data = data, method = "subclass")
```

```
## Warning: Due to discreteness in the distance measure, fewer subclasses were
## generated than were requested.
```

```
data$weights<-NULL
matched_data <- match.data(model.stra)
strata <- unique(matched_data$subclass)
ace <- NULL

for (i in 1:length(strata)) {
  stratum <- strata[i]

  treatment_group <- matched_data[matched_data$subclass == stratum & matched_data$CO.AQI.Category == 1,
  control_group <- matched_data[matched_data$subclass == stratum & matched_data$CO.AQI.Category == 0, ]

  mean_treatment <- mean(treatment_group$PM2.5.AQI.Value)
  mean_control <- mean(control_group$PM2.5.AQI.Value)

  ace[i] <- mean_treatment - mean_control
}
weighted_ace <- weighted.mean(ace, weights =table(matched_data$subclass))
weighted_ace
```

```
## [1] 54.26481
```

2

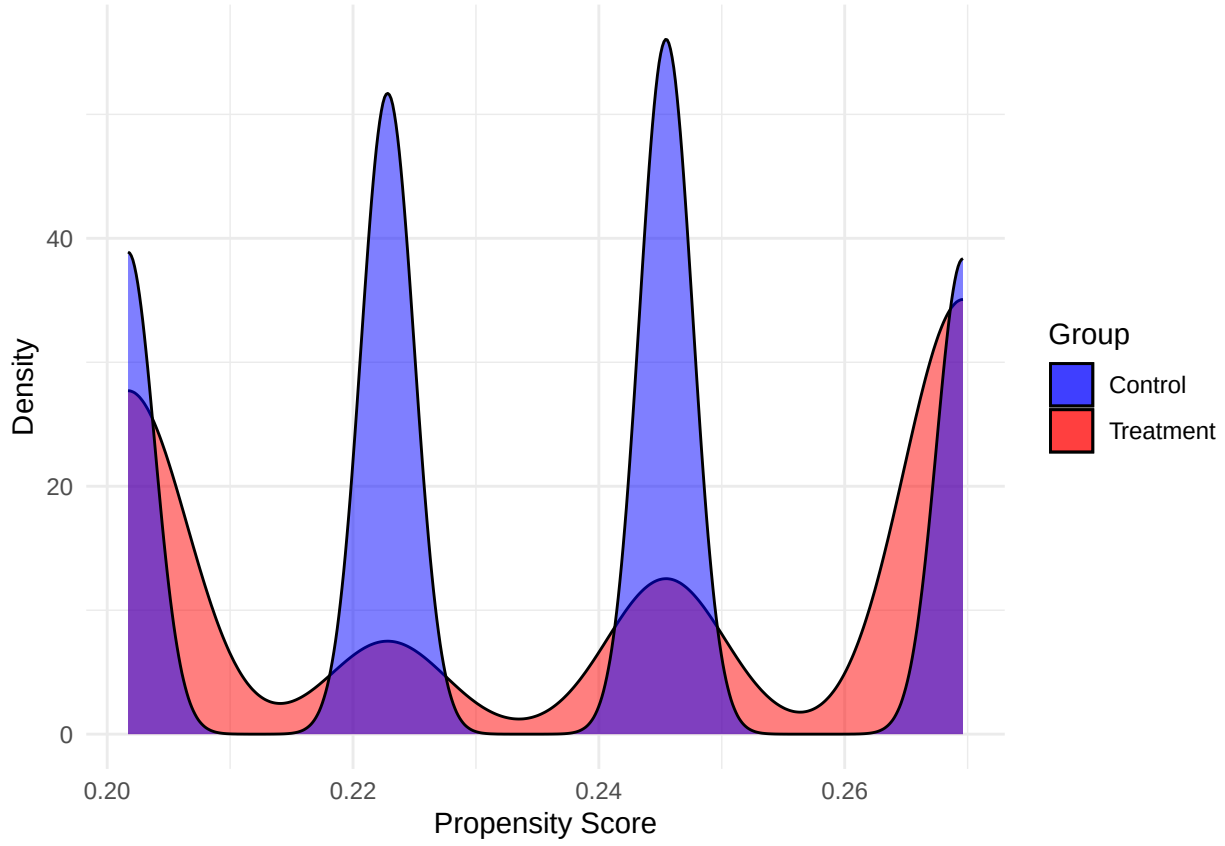
```
data <- read.csv("final_data.csv")
head(data)
```

```
##   X Country AQI.Value AQI.Category CO.AQI.Value CO.AQI.Category Ozone.AQI.Value
```

```
## 1 0     123        51            1            1               0              36
## 2 1      71        41            0            1               0               5
## 3 2     108        66            1            1               0              39
## 4 3       0        34            0            1               0              34
## 5 4      37        22            0            0               0              22
## 6 5     164        54            1            1               0              14
##   Ozone.AQI.Category NO2.AQI.Value NO2.AQI.Category PM2.5.AQI.Value
## 1                  2             0                 0              51
## 2                  0             1                 0              41
## 3                  2             2                 1              66
## 4                  2             0                 0              20
## 5                  1             0                 0               6
## 6                  0            11                 2              54
##   PM2.5.AQI.Category
## 1                  1
## 2                  0
## 3                  1
## 4                  0
## 5                  0
## 6                  1
```

```r
library(MatchIt)
library(ggplot2)
ps_formula <- CO.AQI.Category~Ozone.AQI.Category
ps_model <- matchit(ps_formula, data = data, method = "nearest")
ps_scores <- ps_model$distance
data$propensity_score <- ps_scores
# Subset data for treatment and control groups
treatment_data <- data[data$CO.AQI.Category == 1, ]
control_data <- data[data$CO.AQI.Category == 0, ]

# Create density plots for propensity scores
ggplot() +
  geom_density(data = treatment_data, aes(x = propensity_score, fill = "Treatment"), alpha = 0.5) +
  geom_density(data = control_data, aes(x = propensity_score, fill = "Control"), alpha = 0.5) +
  labs(x = "Propensity Score", y = "Density", fill = "Group") +
  scale_fill_manual(values = c("blue", "red")) +
  theme_minimal()
```



Matching

```
match.method = matchit(CO.AQI.Category~propensity_score, data=data, method="nearest",ratio = 1)
match_data = match.data(match.method)
att <- function(match_data){
  ind_treat = which(match_data$CO.AQI.Category == 1)
  ind_control = which(match_data$CO.AQI.Category == 0)
  treat = match_data$PM2.5.AQI.Value[ind_treat]
  control = match_data$PM2.5.AQI.Value[ind_control]
  mean(treat)-mean(control)
}

att_v = att(match_data)
new_data = data
new_data$CO.AQI.Category = 1 - data$CO.AQI.Category
match.method = matchit(CO.AQI.Category~propensity_score, data=new_data, method="nearest",ratio = 1)
```

```
## Warning: Fewer control units than treated units; not all treated units will get
## a match.
```

```
match_data = match.data(match.method)
atc <- function(match_data){
  ind_treat = which(match_data$CO.AQI.Category == 0)
  ind_control = which(match_data$CO.AQI.Category == 1)
  treat = match_data$PM2.5.AQI.Value[ind_treat]
  control = match_data$PM2.5.AQI.Value[ind_control]
```

```
  mean(treat)-mean(control)
}
atc_v = atc(match_data)
prob = mean(data$CO.AQI.Category)
ATE = prob*att_v+(1-prob)*atc_v
cat("ATE is:", ATE)
```

```
## ATE is: 46.39616
```

Weighting

```
data$weights <- ifelse(data$CO.AQI.Category==1, 1/data$propensity_score, 1/(1-data$propensity_score))

E_Y_1 <- sum(data$PM2.5.AQI.Value*data$weights*data$CO.AQI.Category)/sum(data$weights*data$CO.AQI.Categ

E_Y_0 <- sum(data$PM2.5.AQI.Value*data$weights*(1-data$CO.AQI.Category))/sum(data$weights*(1-data$CO.AQ

cat("The ACE computed under weighted is:", E_Y_1-E_Y_0)
```

```
## The ACE computed under weighted is: 59.93172
```

Strata

```
model.stra <- matchit(ps_formula, data = data, method = "subclass")
```

```
## Warning: Due to discreteness in the distance measure, fewer subclasses were
## generated than were requested.
```

```
data$weights<-NULL
matched_data <- match.data(model.stra)
strata <- unique(matched_data$subclass)
ace <- NULL

for (i in 1:length(strata)) {
  stratum <- strata[i]

  treatment_group <- matched_data[matched_data$subclass == stratum & matched_data$CO.AQI.Category == 1,
  control_group <- matched_data[matched_data$subclass == stratum & matched_data$CO.AQI.Category == 0, ]

  mean_treatment <- mean(treatment_group$PM2.5.AQI.Value)
  mean_control <- mean(control_group$PM2.5.AQI.Value)

  ace[i] <- mean_treatment - mean_control
}
weighted_ace <- weighted.mean(ace, weights =table(matched_data$subclass))
weighted_ace
```

```
## [1] 55.96216
```

3

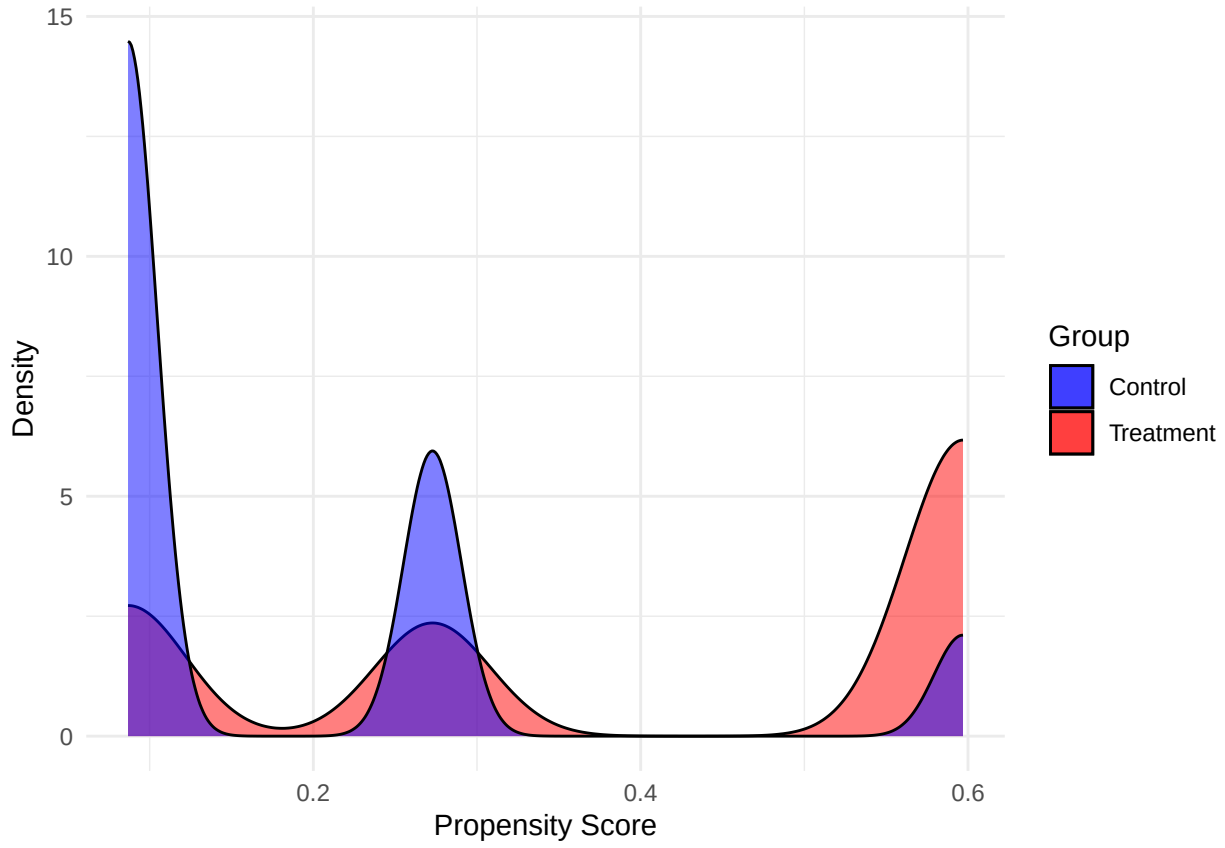
```
data <- read.csv("final_data.csv")
head(data)
```

```
##   X Country AQI.Value AQI.Category CO.AQI.Value CO.AQI.Category Ozone.AQI.Value
```

```
## 1 0     123        51            1            1               0              36
## 2 1      71        41            0            1               0               5
## 3 2     108        66            1            1               0              39
## 4 3       0        34            0            1               0              34
## 5 4      37        22            0            0               0              22
## 6 5     164        54            1            1               0              14
##   Ozone.AQI.Category NO2.AQI.Value NO2.AQI.Category PM2.5.AQI.Value
## 1                  2             0                0              51
## 2                  0             1                0              41
## 3                  2             2                1              66
## 4                  2             0                0              20
## 5                  1             0                0               6
## 6                  0            11                2              54
##   PM2.5.AQI.Category
## 1                  1
## 2                  0
## 3                  1
## 4                  0
## 5                  0
## 6                  1
```

```r
library(MatchIt)
library(ggplot2)
ps_formula <- CO.AQI.Category~NO2.AQI.Category
ps_model <- matchit(ps_formula, data = data, method = "nearest")
ps_scores <- ps_model$distance
data$propensity_score <- ps_scores
# Subset data for treatment and control groups
treatment_data <- data[data$CO.AQI.Category == 1, ]
control_data <- data[data$CO.AQI.Category == 0, ]

# Create density plots for propensity scores
ggplot() +
  geom_density(data = treatment_data, aes(x = propensity_score, fill = "Treatment"), alpha = 0.5) +
  geom_density(data = control_data, aes(x = propensity_score, fill = "Control"), alpha = 0.5) +
  labs(x = "Propensity Score", y = "Density", fill = "Group") +
  scale_fill_manual(values = c("blue", "red")) +
  theme_minimal()
```



Matching

```
match.method = matchit(CO.AQI.Category~propensity_score, data=data, method="nearest",ratio = 1)
match_data = match.data(match.method)
att <- function(match_data){
  ind_treat = which(match_data$CO.AQI.Category == 1)
  ind_control = which(match_data$CO.AQI.Category == 0)
  treat = match_data$PM2.5.AQI.Value[ind_treat]
  control = match_data$PM2.5.AQI.Value[ind_control]
  mean(treat)-mean(control)
}

att_v = att(match_data)
new_data = data
new_data$CO.AQI.Category = 1 - data$CO.AQI.Category
match.method = matchit(CO.AQI.Category~propensity_score, data=new_data, method="nearest",ratio = 1)
```

```
## Warning: Fewer control units than treated units; not all treated units will get
## a match.
```

```
match_data = match.data(match.method)
atc <- function(match_data){
  ind_treat = which(match_data$CO.AQI.Category == 0)
  ind_control = which(match_data$CO.AQI.Category == 1)
  treat = match_data$PM2.5.AQI.Value[ind_treat]
  control = match_data$PM2.5.AQI.Value[ind_control]
```

```
  mean(treat)-mean(control)
}
atc_v = atc(match_data)
prob = mean(data$CO.AQI.Category)
ATE = prob*att_v+(1-prob)*atc_v
cat("ATE is:", ATE)
```

```
## ATE is: 55.58115
```

Weighting

```
data$weights <- ifelse(data$CO.AQI.Category==1, 1/data$propensity_score, 1/(1-data$propensity_score))

E_Y_1 <- sum(data$PM2.5.AQI.Value*data$weights*data$CO.AQI.Category)/sum(data$weights*data$CO.AQI.Categ

E_Y_0 <- sum(data$PM2.5.AQI.Value*data$weights*(1-data$CO.AQI.Category))/sum(data$weights*(1-data$CO.AQ

cat("The ACE computed under weighted is:", E_Y_1-E_Y_0)
```

```
## The ACE computed under weighted is: 57.14648
```

Strata

```
model.stra <- matchit(ps_formula, data = data, method = "subclass")
```

```
## Warning: Due to discreteness in the distance measure, fewer subclasses were
## generated than were requested.
```

```
data$weights<-NULL
matched_data <- match.data(model.stra)
strata <- unique(matched_data$subclass)
ace <- NULL

for (i in 1:length(strata)) {
  stratum <- strata[i]

  treatment_group <- matched_data[matched_data$subclass == stratum & matched_data$CO.AQI.Category == 1,
  control_group <- matched_data[matched_data$subclass == stratum & matched_data$CO.AQI.Category == 0, ]

  mean_treatment <- mean(treatment_group$PM2.5.AQI.Value)
  mean_control <- mean(control_group$PM2.5.AQI.Value)

  ace[i] <- mean_treatment - mean_control
}
weighted_ace <- weighted.mean(ace, weights =table(matched_data$subclass))
weighted_ace
```

```
## [1] 59.0005
```