Treatment Effect Estimation

Susan Glenn, Jeanne Li, Jingcheng Xu, Stewart Kerr

12/2/2019

After matching, there were 2182 matched pairs. An outcome variable was assigned to each person; a person received a 1 if they outlived their matched counterpart and a 0 if they died first. If both people in a pair lived to the end of data collection or died in the same wave, this pair was removed from the data set. This removal may account for the balance issues seen later in the report. For future analysis, one possible option would be to look at exact date of death and assign the response to people who died in the same wave.

After this removal, 957 pairs were left. Out of these 957, 654 people in the treatment group died first and 571 people in the treatment group died second. The percentage of people who died last in the control group was 53% (47% for treatment); indicating that the control patients on average were living longer. The difference between these two means was 0.068 was significant with p-values of 0.01766 and 0.01772 for a paired t test and an exact Wilcoxin rank sum test respectively.

In an attempt to get an estimate for the average treatment effect, the IPW estimator was found to be -0.006. The propensity scores used, which were estimates of the true propensity scores, were obtained by regressing the covariates on the treatment assignment using a logistic model. The time varying covariates were changed for each pair to reflect the wave before a person was matched when calculating the propensity scores. As only the matched pairs were used, our propensity score might overestimate the odds of being treated. That is, in the control pool we have more controls than treated but in the propensity score model we have an equal number (because we used only matched pairs). The code used to generate these results is below.

Testing for Treatment Effect

```r
data <- read.csv("../data/results-final.csv")
library(exactRankTests)
```

```
## Warning: package 'exactRankTests' was built under R version 3.5.3
```

```
##  Package 'exactRankTests' is no longer under development.
##  Please consider using package 'coin' instead.
```

```r
dat1 <- subset(data, outcome==1)
dat0 <- subset(data, outcome==0)
t1 <- subset(dat1, treated==1)
c1 <- subset(dat1, treated==0)
t0 <- subset(dat0, treated==1)
c0 <- subset(dat0, treated==0)

ggg <- subset(data, outcome >= 0)
treated <- subset(ggg, treated==1)
control <- subset(ggg, treated==0)

t <- as.data.frame(rbind(t1, t0))
c <- as.data.frame(rbind(c0, c1))

# Set up outcome data for match
treated.y=t$outcome
control.y=c$outcome
mean(treated.y)
```

```
## [1] 0.4661224
```

```r
mean(control.y)
```

```
## [1] 0.5338776
```

```r
### Estimate ATE in pair matching
t.test(treated.y, control.y, paired=TRUE)
```

```
##
##  Paired t-test
##
## data:  treated.y and control.y
## t = -2.3759, df = 1224, p-value = 0.01766
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.12370353 -0.01180668
## sample estimates:
## mean of the differences
##              -0.0677551
```

```r
wilcox.exact(treated.y, control.y, paired=TRUE)
```

```
##
##  Asymptotic Wilcoxon signed rank test
##
## data:  treated.y and control.y
## V = 350020, p-value = 0.01772
## alternative hypothesis: true mu is not equal to 0
```

Overlap and IPW Estimated Treatment Effect

```r
library(sm)
```

```
## Warning: package 'sm' was built under R version 3.5.3
```

```
## Package 'sm', version 2.2-5.6: type help(sm) for summary information
```

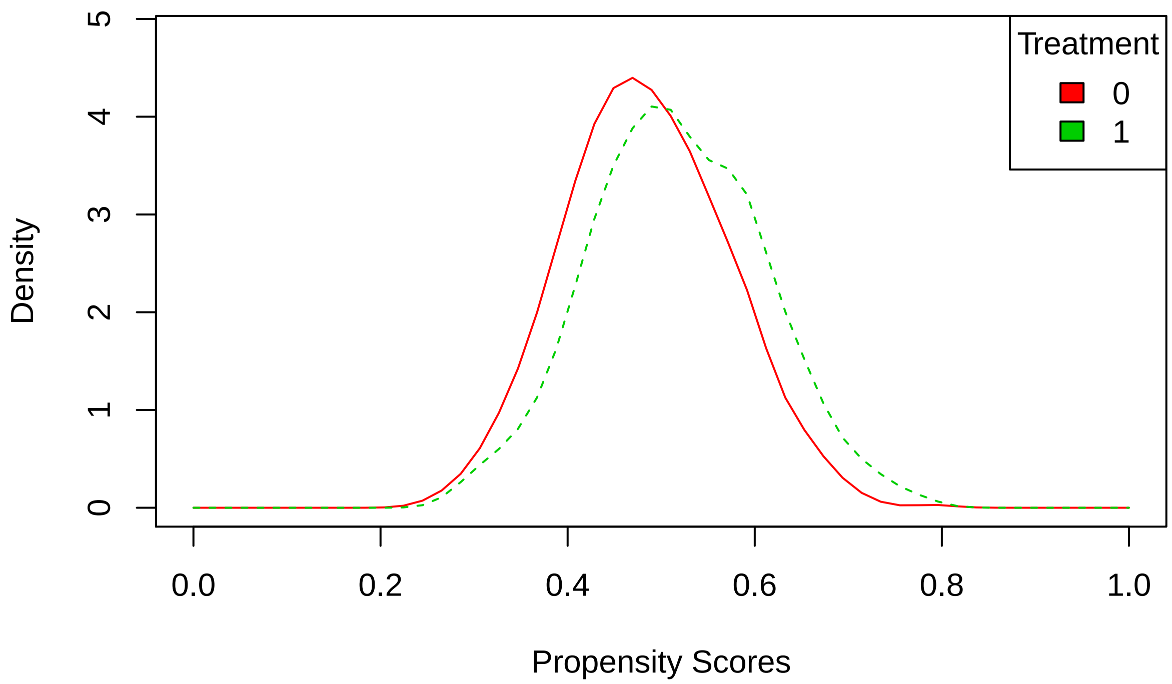
```r
X = rbind(treated[,c(8:18, 20:45)], c[,c(8:18, 20:45)])
M = rbind(treated[,c(8:9)], control[,(8:9)])
A = c(t$treated, c$treated)
Y = c(t$outcome, c$outcome)


new <- as.data.frame(cbind(X, A))

ehat = predict(glm(A~.,family="binomial", data=new),type="response")
ehat[ehat < 0.025] = 0.025
ehat[ehat > 0.975] = 0.975
paste0("The IPW estimated average treatment effect is: ",round(mean((Y*A)/ehat - (Y*(1-A))/(1-ehat)),5)
```

```
## [1] "The IPW estimated average treatment effect is: -0.06001"
```

```r
sm.density.compare(ehat, factor(A), xlab="Propensity Scores",main="Propensity Scores Between Treated an
legend("topright", levels(factor(A)), fill=2+(0:nlevels(factor(A))), title = "Treatment")
```


Density
0
1
2
3
4
5
0.0
0.2
0.4
0.6
0.8
1.0
Propensity Scores
Treatment
🟥 0
🟩 1