

Lab 12: Causal Forests

Research question

Does receiving financial aid reduce the probability of a convict being re-arrested after a release? This data set is originally from Rossi et al. (1980, Money, Work, and Crime: Some Experimental Results.). Observations are 432 convicts who were released from Maryland state prisons in the 1970s and who were followed up for one year after release. Half the released convicts were assigned at random to an experimental treatment in which they were given financial aid.

It is important to assess if financial aid helps with re-integration into the society. It is also important to assess the magnitude of the effect as financial aid programs can be costly to tax-payers. If the effect of the program can be heterogeneous, it is important to identify sub-populations where the program is more effective. This would allow policy makers to target released convicts for whom the effect is expected to be more substantial.

Data preparation

The required package is “carData” for the Rossi dataset.

```
library(carData)
```

Let's construct `Wks`: percentage of time unemployed before re-arrest or until the end of the period:

```
N=nrow(Rossi)
Rossi$Weeks_worked=0
for (i in 1:N){
  last_week_col=Rossi$week[i]+10 #last week before arrest column number
  for (j in 11:last_week_col){
    Rossi$Weeks_worked[i]=Rossi$Weeks_worked[i]+(Rossi[i,j]=="yes")
  }
}
Rossi$Wks=(Rossi$week-Rossi$Weeks_worked)/Rossi$week
```

Create a 0/1 treatment variable:

```
Rossi$D=(Rossi$fin=="yes")+0
```

Education is a factor:

```
Rossi$Grades=" "
Rossi$Grades[Rossi$educ==2]<-"_6"
Rossi$Grades[Rossi$educ==3]<-"7_9"
Rossi$Grades[Rossi$educ==4]<-"10_11"
Rossi$Grades[Rossi$educ==5]<-"12"
Rossi$Grades[Rossi$educ==6]<-"College"
```

The data set:

```
RData=subset(Rossi,select=c(arrest,D,age,race,wexp,mar,paro,prio,Wks))
RData$educ<-as.factor(Rossi$Grades)
```

Estimate the ATE

Compute the simple ATE estimator

The treatment was randomly assigned. We should be able to estimate ATE by just regressing `arrest` against `D`. The estimated ATE of financial aid is 8.3% reduction in probability of being re-arrested. We calculate the heteroskedasticity robust standard errors and find that the result is significant.

```
library(AER)

##      car
##      lmtest
##      zoo
##
##      'zoo'

## The following objects are masked from 'package:base':
##
##      as.Date, as.Date.numeric
##
##      sandwich
##
##      survival

ATEsimple=lm(arrest~D,data=RData)
coeftest(ATEsimple,vcov=vcovHC(ATEsimple, type = "HC"))

##
## t test of coefficients:
##
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.305556   0.031343   9.7488 < 2e-16 ***
## D            -0.083333   0.042220  -1.9738  0.04905 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Compute the IPW estimator

Since the treatment is randomized, there are not supposed to be controls explaining the treatment assignment. Let's check using the logit model:

```
logit=glm(D~(age+race+wexp+mar+paro+prio+educ+Wks)^2,data=RData,family = binomial(link = "logit"))
coeftest(logit,vcov=vcovHC(logit,type="HCO"))

##
## z test of coefficients:
##
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -7.0011e+00  4.1599e+00  -1.6830  0.0923761 .
## age          1.8340e-01  1.1578e-01   1.5839  0.1132069
## raceother    -3.1273e+00  3.1566e+00  -0.9907  0.3218273
## wexpyes      -4.7101e-01  2.7660e+00  -0.1703  0.8647846
## marnot married -4.0224e+00  3.5570e+00  -1.1308  0.2581205
## paroyes      1.7722e+00  2.1833e+00   0.8117  0.4169480
## prio         5.8832e-01  4.1311e-01   1.4241  0.1544039
## educ10_11    4.6079e+00  3.2665e+00   1.4107  0.1583474
## educ12       4.0261e+00  3.7154e+00   1.0836  0.2785317
```

```

## educ7_9          6.8662e+00  3.1623e+00  2.1712 0.0299125 *
## educCollege     1.5650e+02  7.6497e+00  20.4583 < 2.2e-16 ***
## Wks             2.6288e+00  3.6239e+00  0.7254 0.4681987
## age:raceother   1.7008e-02  7.4284e-02  0.2290 0.8188976
## age:wexpyes     6.0950e-03  5.5055e-02  0.1107 0.9118492
## age:marnot married 3.3636e-02  7.2572e-02  0.4635 0.6430218
## age:paroyes     -1.8435e-02  4.6440e-02  -0.3970 0.6913963
## age:prio        5.8193e-03  9.4815e-03  0.6138 0.5393786
## age:educ10_11   -1.6883e-01  9.4051e-02  -1.7951 0.0726349 .
## age:educ12      -1.7664e-01  1.0136e-01  -1.7427 0.0813812 .
## age:educ7_9     -2.0173e-01  8.3012e-02  -2.4302 0.0150914 *
## age:educCollege -3.6131e+00  2.3672e-01  -15.2633 < 2.2e-16 ***
## age:Wks         -1.5897e-02  5.8827e-02  -0.2702 0.7869858
## raceother:wexpyes 1.6040e-01  9.5713e-01  0.1676 0.8669104
## raceother:marnot married 1.5064e+00  1.2492e+00  1.2059 0.2278750
## raceother:paroyes 1.5773e+00  7.3784e-01  2.1377 0.0325409 *
## raceother:prio   -4.9654e-02  1.6167e-01  -0.3071 0.7587380
## raceother:educ10_11 -5.2417e-01  1.7236e+00  -0.3041 0.7610418
## raceother:educ12  6.3134e-02  1.8218e+00  0.0347 0.9723554
## raceother:educ7_9 -7.8363e-01  1.3777e+00  -0.5688 0.5694875
## raceother:educCollege 7.3793e+00  2.1404e+00  3.4476 0.0005656 ***
## raceother:Wks    1.4646e+00  1.1930e+00  1.2277 0.2195722
## wexpyes:marnot married -6.4887e-01  1.1027e+00  -0.5884 0.5562392
## wexpyes:paroyes  -3.8337e-01  5.3643e-01  -0.7147 0.4748105
## wexpyes:prio     3.7302e-02  9.1868e-02  0.4060 0.6847154
## wexpyes:educ10_11 1.4099e+00  1.6886e+00  0.8350 0.4037375
## wexpyes:educ12   2.5537e+00  1.8262e+00  1.3984 0.1620051
## wexpyes:educ7_9  1.6992e+00  1.6073e+00  1.0571 0.2904538
## wexpyes:educCollege -7.2536e+01  3.5762e+00  -20.2830 < 2.2e-16 ***
## wexpyes:Wks     -9.0347e-01  7.9669e-01  -1.1340 0.2567824
## marnot married:paroyes 8.3771e-01  8.4919e-01  0.9865 0.3238961
## marnot married:prio -1.5695e-01  1.7890e-01  -0.8773 0.3803398
## marnot married:educ10_11 4.4851e+00  2.8231e+00  1.5887 0.1121237
## marnot married:educ12 3.0887e+00  2.9747e+00  1.0383 0.2991141
## marnot married:educ7_9 4.1295e+00  2.7947e+00  1.4776 0.1395061
## marnot married:educCollege 5.5369e+01  3.6437e+00  15.1958 < 2.2e-16 ***
## marnot married:Wks -7.3695e-01  1.3320e+00  -0.5533 0.5800814
## paroyes:prio     -1.8080e-01  8.8226e-02  -2.0493 0.0404301 *
## paroyes:educ10_11 -1.5821e+00  1.3302e+00  -1.1893 0.2343143
## paroyes:educ12   -1.4427e+00  1.4764e+00  -0.9771 0.3284970
## paroyes:educ7_9  -2.1917e+00  1.2980e+00  -1.6885 0.0913174 .
## paroyes:educCollege 6.2094e+01  3.2015e+00  19.3952 < 2.2e-16 ***
## paroyes:Wks     6.8379e-01  7.3339e-01  0.9324 0.3511450
## prio:educ10_11   -4.4349e-01  2.9715e-01  -1.4925 0.1355735
## prio:educ12      -2.2440e-01  3.5738e-01  -0.6279 0.5300644
## prio:educ7_9     -4.2074e-01  2.9360e-01  -1.4330 0.1518447
## prio:educCollege -4.0299e+00  6.8431e-01  -5.8890 3.885e-09 ***
## prio:Wks         -1.2182e-01  1.3452e-01  -0.9056 0.3651312
## educ10_11:Wks    -4.5179e-01  2.9735e+00  -0.1519 0.8792332
## educ12:Wks       -7.1615e-02  3.2086e+00  -0.0223 0.9821928
## educ7_9:Wks      -2.0771e+00  2.9094e+00  -0.7139 0.4752845
## educCollege:Wks  -1.5351e+02  7.4152e+00  -20.7027 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

We find that some of the individual p-values are significant, possibly due to the multiple hypotheses testing problem, even if the treatment is truly randomized. However, it is possible that treatment assignment is not completely random. For example, convicts with some college seems to be more likely to be in the program. We also compute the IPW estimator of ATE using estimated propensity scores.

Let's construct a function that computes the IPW ATE estimator. We will estimate the propensity score by logit as above.

```
ipw.ate<-function(data,indices){
  #indices are needed later for bootstrap samples construction
  DATA=data[indices,]

  # Logit for propensity score
  logit=glm(D~(age+race+wexp+mar+paro+prio+educ+Wks)^2,data=DATA,family = binomial(link = "logit"))
  # predicted probabilities = estimated propensity score
  p=predict(logit,type="response")

  #Trim predictions close to zero or one
  ind=which(p>0.05 & p<0.95)
  p1=p[ind]
  DATA1=DATA[ind,]

  # The IPW ATE estimator
  Diff=(DATA1$arrest*DATA1$D)/p1 - (DATA1$arrest*(1-DATA1$D)/(1-p1))
  IPW_ATE=mean(Diff)
  return(IPW_ATE)
}
```

Let's apply the function to our data set:

```
ind=seq(1:nrow(RData))
IPW_ATE=ipw.ate(RData,indices=ind)
IPW_ATE
```

```
## [1] -0.0831134
```

We use the bootstrap for inference. We use 399 bootstrap samples:

```
library(boot)

##
## 'boot'
## The following object is masked from 'package:survival':
##
## aml
## The following object is masked from 'package:car':
##
## logit
set.seed(6064,sample.kind='Rejection')
IPW_ATE_boot=boot(data=RData,statistic=ipw.ate,R=399)
IPW_ATE_boot

##
## ORDINARY NONPARAMETRIC BOOTSTRAP
##
##
```

```
## Call:
## boot(data = RData, statistic = ipw.ate, R = 399)
##
##
## Bootstrap Statistics :
##      original      bias    std. error
## t1* -0.0831134 0.0109955 0.05140444
```

Similar results using the “causalweight” package:

```
library(causalweight)
```

```
##      ranger
X=model.matrix(D~(age + race + wexp + mar +
  paro + prio + educ + Wks)^2,data=RData)[-1]
Y=RData$arrest
D=RData$D
output=treatweight(d=D,y=Y,x=X,logit=TRUE,trim=0.05,boot=399)
```

```
## Warning: 9 bootstrap sample(s) dropped due to NA's
```

```
output
```

```
## $effect
## [1] -0.08020902
##
## $se
## [1] 0.04784987
##
## $pval
## [1] 0.09368645
##
## $y1
## [1] 0.2243621
##
## $y0
## [1] 0.3045712
##
## $ntrimmed
## [1] 17
```

Estimate CATE by causal forest

```
library(grf)
?causal_forest
```

Generate the output:

```
X=model.matrix(D~age+race+mar+paro+prio+educ+Wks,data=RData)[-1]
Y=RData$arrest
W=RData$D
CF=causal_forest(X,Y,W,seed = 6064)
```

Compute variables’ importance. We find that variables that have strongest explanatory power are **Wks**, **age** and **prio**. * **Wks**: percentage of time unemployed before re-arrest or until the end of the period; * **age**: age at time of release; * **prio**: number of convictions prior to current incarceration.

```
CF_var_imp<-variable_importance(CF)
selected.vars <- which(CF_var_imp > mean(CF_var_imp))
Importance=round(CF_var_imp[selected.vars],digits=3)
names(Importance)<-colnames(X)[selected.vars]
sort(Importance, decreasing=TRUE)
```

```
## Wks age prio
## 0.373 0.273 0.174
```

Compute CATE for age>=27 and age<27. We find that the program is effective for older convicts.

```
average_treatment_effect(CF,target.sample = "all",subset=(RData$age>=27))
```

```
## estimate std.err
## -0.18024396 0.06932013
```

```
average_treatment_effect(CF,target.sample = "all",subset=(RData$age<27))
```

```
## estimate std.err
## -0.02937981 0.05076452
```

Compute CATE for Wks>=0.85 and Wks<0.85. We find a stronger effect for those who are less employed.

```
average_treatment_effect(CF,target.sample = "all",subset=(RData$Wks>=0.85))
```

```
## estimate std.err
## -0.16766236 0.08949657
```

```
average_treatment_effect(CF,target.sample = "all",subset=(RData$Wks<0.85))
```

```
## estimate std.err
## -0.02765819 0.04591887
```

```
CATE_Wks=average_treatment_effect(CF,target.sample = "all",subset=(RData$Wks>=0.85))
```

```
Estimate=CATE_Wks[1]
```

```
Std.err=CATE_Wks[2]
```

```
T=abs(Estimate/Std.err)
```

```
pvalue=(2*(1-pnorm(T)))
```

```
cat("CATE for Wks>0.85: ",round(Estimate,digits=3), "
    Std.err: ",round(Std.err,digits=3), " p-value: ",round(pvalue,digits=3))
```

```
## CATE for Wks>0.85: -0.168
## Std.err: 0.089 p-value: 0.061
```

Compute CATE by prior convictions. There seems to be a stronger effect for those with multiple prior convictions

```
average_treatment_effect(CF,target.sample = "all",subset=(RData$prio>1))
```

```
## estimate std.err
## -0.10544583 0.05409921
```

```
average_treatment_effect(CF,target.sample = "all",subset=(RData$prio<=1))
```

```
## estimate std.err
## 0.0002742673 0.0641875634
```

```
CATE_prio=average_treatment_effect(CF,target.sample = "all",subset=(RData$prio>1))
```

```
Estimate=CATE_prio[1]
```

```
Std.err=CATE_prio[2]
```

```
T=abs(Estimate/Std.err)
```

```
pvalue=(2*(1-pnorm(T)))
cat("CATE for prio>1 : ",round(Estimate,digits=3), "
    Std.err: ",round(Std.err,digits=3), " p-value: ",round(pvalue,digits=3))
```

```
## CATE for prio>1 : -0.105
## Std.err: 0.054 p-value: 0.051
```

Compute CATE by marital status. The program might be more effective for single convicts.

```
average_treatment_effect(CF,target.sample = "all",subset=(RData$mar=="married"))
```

```
## estimate std.err
## -0.04822461 0.10534020
```

```
average_treatment_effect(CF,target.sample = "all",subset=(RData$mar=="not married"))
```

```
## estimate std.err
## -0.07132707 0.04530755
```

We can combine several conditions. We can find that the program is more effective for older single convicts with multiple prior convictions. They are less likely to have family support.

```
CATE_mult=average_treatment_effect(CF,target.sample = "all",subset=
    (RData$age>=27 & RData$mar=="not married"
    & RData$prio>1))
```

```
Estimate=CATE_mult[1]
Std.err=CATE_mult[2]
T=abs(Estimate/Std.err)
pvalue=(2*(1-pnorm(T)))
cat("CATE for age>=27, not married, with prio>1 : ",round(Estimate,digits=3),"
    Std.err: ",round(Std.err,digits=3), " p-value: ",round(pvalue,digits=3))
```

```
## CATE for age>=27, not married, with prio>1 : -0.313
## Std.err: 0.115 p-value: 0.006
```

Compute CATE by the three most important variables. We find that the program is more effective for employed at least 60% of the time older convicts with multiple prior convictions.

```
CATE_import=average_treatment_effect(CF,target.sample = "all",subset=
    (RData$Wks<0.40 & RData$age>=27 & RData$prio>1))
```

```
Estimate=CATE_import[1]
Std.err=CATE_import[2]
T=abs(Estimate/Std.err)
pvalue=(2*(1-pnorm(T)))
cat("CATE for Wks<0.4, age>=27, with prio>1 : ",round(Estimate,digits=3),"
    Std.err: ",round(Std.err,digits=3), " p-value: ",round(pvalue,digits=3))
```

```
## CATE for Wks<0.4, age>=27, with prio>1 : -0.359
## Std.err: 0.124 p-value: 0.004
```

Compute the confidence interval. We find that the program can be highly effective: up to 65% reduction in the probability of re-arrest for properly chosen subjects who keep employment, are older, and have multiple prior convictions.

```
c(Estimate-1.96*Std.err,Estimate+1.96*Std.err)
```

```
## estimate estimate
```

-0.6021438 -0.1164414