

pset 4 econ 21020

Matthew Zhao

5/22/2022

Problem 1

a)

False. Suppose W, U are iid on $\text{Uniform}[0, X]$ where X is a random variable that is equal to 1 with probability 0.5 and 2 with probability 0.5. In this case, W and U are independent by definition but W is not independent of $U|X$ since $U|X$ is a function of X and W is not independent of X .

b)

False.

c)

False. $E[\epsilon|X] = E[Y - \text{BLP}(Y|X)|X] = E[Y|X] - E[\text{BLP}(Y|X)|X] = E[Y|X] - \text{BLP}(Y|X)$. $E[\epsilon|X] = 0$ only when $E[Y|X] = \text{BLP}(Y|X)$ which is not always true.

Problem 2

$$\text{WTS } \frac{\hat{\text{ATE}}_n - \text{ATE}}{\text{se}(\hat{\text{ATE}}_n)} \xrightarrow{d} N(0, 1)$$

$$\text{where } \text{se}(\hat{\text{ATE}}_n) = \frac{1}{\sqrt{n}} \sqrt{\hat{\sigma}_{\text{ATE}}^2}$$

$$\frac{\hat{\text{ATE}}_n - \text{ATE}}{\text{se}(\hat{\text{ATE}}_n)} = \sqrt{n}(\hat{\text{ATE}}_n - \text{ATE}) \frac{1}{\sqrt{\hat{\sigma}_{\text{ATE}}^2}}$$

$$\text{Let } A_n = \sqrt{n}(\hat{\text{ATE}}_n - \text{ATE}), B_n = \hat{\sigma}_{\text{ATE}}^2$$

$$\text{By Lec 7 Theorem 4, } A_n \xrightarrow{d} N(0, \sigma_{\text{ATE}}^2)$$

$$B_n = \hat{\sigma}_{\text{CATE}}^2(1)\hat{p}_n(1)^2 + \hat{\sigma}_{\text{CATE}}^2(0)\hat{p}_n(0)^2 + (\hat{\text{CATE}}_n(1) - \hat{\text{CATE}}_n(0))\hat{p}_n(1)\hat{p}_n(0)$$

$$\text{where } \hat{p}_n(x) = \frac{1}{n} \sum_{i=1}^n 1(X_i = x)$$

$$\text{Let } C_n = \hat{\sigma}_{\text{CATE}}^2(1), D_n = \hat{\sigma}_{\text{CATE}}^2(0), E_n = \hat{p}_n(1), F_n = \hat{p}_n(0), G_n = \hat{\text{CATE}}_n(1), H_n = \hat{\text{CATE}}_n(0)$$

$$g(c, d, e, f, g, h) = ce^2 + df^2 + (g - h)ef$$

$$g(C_n, D_n, E_n, F_n, G_n, H_n) = B_n$$

$$\text{By WLLN, } C_n, D_n \xrightarrow{p} \sigma_{\text{CATE}}^2(x)$$

$$E_n, F_n \xrightarrow{p} p(x)$$

$$G_n, H_n \xrightarrow{p} \text{CATE}(x)$$

$$\text{By CMT, } B_n \xrightarrow{p} \sigma_{\text{ATE}}^2$$

By Slutsky's Theorem, since $A_n \xrightarrow{d} N(0, \sigma_{\text{ATE}}^2)$ and $B_n \xrightarrow{p} \sigma_{\text{ATE}}^2$,

$$\sqrt{n}(\hat{\text{ATE}}_n - \text{ATE}) \frac{1}{\sqrt{\hat{\sigma}_{\text{ATE}}^2}} \xrightarrow{d} \frac{1}{\sqrt{\sigma_{\text{ATE}}^2}} N(0, \sigma_{\text{ATE}}^2) \stackrel{d}{=} N(0, 1)$$

Problem 3

a)

$P(W = 1|X)$ is the probability of being treated given a set of observables X .

b)

$$\text{WTS } P(U \leq u, W = 1|p(X)) = P(U \leq u|p(X))P(W = 1|p(X))$$

$$P(U \leq u, W = 1|p(X)) = E[1\{U \leq u, W = 1\}|p(X)]$$

$$= E[1\{U \leq u\}1\{W = 1\}|p(X)] = E[1\{U \leq u\}W|p(X)]$$

$$= E[E[1\{U \leq u\}W|p(X), 1\{U \leq u\}]|p(X)]$$

$$= E[1\{U \leq u\}E[W|p(X), 1\{U \leq u\}]|p(X)]$$

$$= 1 \cdot E[W|p(X), 1\{U \leq u\}]P(U \leq u|p(X)) + 0 \cdot E[W|p(X), 1\{U \leq u\}]P(U > u|p(X))$$

$$= E[W|p(X), 1\{U \leq u\}]P(U \leq u|p(X))$$

$$\text{Now we just need to show that } E[W|p(X), 1\{U \leq u\}] = P(W = 1|p(X))$$

$$E[W|p(X), 1\{U \leq u\}] = E[E[W|p(X), 1\{U \leq u\}, X]|p(X), U \leq u]$$

$$= E[E[W|1\{U \leq u\}, X]|p(X), U \leq u]$$

$$= E[E[W|X]|p(X), U \leq u] \text{ by SO}$$

$$= E[p(X)|p(X), U \leq u] = p(X) \text{ where } p(X) = P(W = 1|p(X))$$

c)

This can be interpreted as the expected casual effect of treatment given that the probability of treatment is p .

d)

$$E[g(1, U) - g(0, U)|p(X) = p] = E[g(1, U) - g(0, U)|W, p(X) = p]$$

$$= E[g(1, U)|W, p(X) = p] - E[g(0, U)|W, p(X) = p] = E[g(W, U)|W = 1, p(X) = p] - E[g(W, U)|W = 0, p(X) = p]$$

$$= E[Y|W = 1, p(X) = p] - E[Y|W = 0, p(X) = p]$$

Problem 4

a)

$P(W = 1|X)$ is the likelihood of selecting into the policy given a set of observables/characteristics. It represents the probability of policy being selected as a function of the observables.

b)

An example would be winning big at the casino.

c)

A potential confounder would be family wealth. People who are wealthy are less likely to join the military and more likely to have higher lifetime earnings.

d)

The assumption states that the choice of serving the military is unrelated to unobserved determinants of lifetime income that the individual does not observe i.e. the choice of serving in the military is only based on factors the agent observes, here whether they are male or not.

This does not seem plausible here since it is likely that choosing to serve in the military is related to other factors besides your gender such as a family wealth or connections, which would allow you to dodge the draft, a frequent occurrence during the Vietnam War draft in particular.

e)

It does because it means that the event of serving in the military or not is completely random given gender, specifically that there is some random draft for males and none for females.

f)

$C\hat{ATE}(x) = E[g(1, U) - g(0, U)|X]$. The conditional average treatment effect here represents the expected lifetime earnings of an individual given their gender.

g)

Yes it is. Since we assume that W is independent of $U|X$, $C\hat{ATE}(1) = E[g(1, U) - g(0, U)|1] = E[g(1, U)|W = 1, x = 1] - E[g(0, U)|W = 0, x = 1] = E[g(W, U)|W = 1, x = 1] - E[g(W, U)|W = 0, x = 1] = E[Y|W = 1, x = 1] - E[Y|W = 0, x = 1]$

h)

Since we assume that W is independent of $U|X$, $C\hat{ATE}(0) = E[g(1, U) - g(0, U)|0] = E[g(1, U)|W = 1, x = 0] - E[g(0, U)|W = 0, x = 0] = E[g(W, U)|W = 1, x = 0] - E[g(W, U)|W = 0, x = 0] = E[Y|W = 1, x = 0] - E[Y|W = 0, x = 0]$. However, because we do not observe lifetime earnings for females who served in the military since we do not have $W = w, x = 0$, this is not point identified.

i)

$ATE = E[g(1, U) - g(0, U)]$. The average treatment effect is the expected difference in lifetime earnings for an individual if they had served in the military and if they had not.

j)

The ATE is point identified. Since we assume that W is independent of $U|X$, $ATE = E[g(1, U) - g(0, U)] = E[E[g(1, U) - g(0, U)|X]] = E[C\hat{ATE}(x)]$. Since the conditional average treatment effect is point identified, ATE is point identified.

Problem 5

$$\begin{aligned} & \arg \min_{\beta \in \mathbb{R}^{k+1}} E[(E[Y|X] - X^T \beta)^2] \\ & E[(E[Y|X]^2 - 2E[Y|X]X^T \beta - (X^T \beta)^2)] \\ & = E[E[Y|X]^2] - 2E[E[YX^T|X]]\beta - \beta^T E[X^T X]\beta \end{aligned}$$

$$= E[E[Y|X]^2] - 2E[YX^T]\beta - \beta^T E[XX^T]\beta$$

FOC: $[\beta]$

$$-2E[YX^T] - \beta^T(E[X^T X]^T + E[XX^T]) = 0_{k+1}^T$$

$$\Rightarrow -2E[YX^T] + \beta^T E[XX^T] = 0_{k+1}^T$$

$$\Rightarrow -2E[XY] + E[XX^T]\beta = 0$$

$$\Rightarrow E[XY] = E[XX^T]\beta$$

$$\Rightarrow \beta = E[XX^T]^{-1}E[XY]$$

$$\arg \min_{\beta \in \mathbb{R}^{k+1}} E[(Y - X^T \beta)^2]$$

$$E[(Y^2 - 2YX^T\beta - (X^T\beta)^2)]$$

$$= E[Y^2] - 2E[YX^T]\beta - \beta^T E[XX^T]\beta$$

FOC: $[\beta]$

$$-2E[YX^T] - \beta^T(E[X^T X]^T + E[XX^T]) = 0_{k+1}^T$$

$$\Rightarrow -2E[YX^T] + \beta^T E[XX^T] = 0_{k+1}^T$$

$$\Rightarrow -2E[XY] + E[XX^T]\beta = 0$$

$$\Rightarrow E[XY] = E[XX^T]\beta$$

$$\Rightarrow \beta = E[XX^T]^{-1}E[XY]$$

Problem 6

```
dat<-read.csv("data/bw06.csv")
dat<-as.matrix(dat)
```

```
y<-dat[, "birthweight"]
w<-dat[, "cigsdaily"]
x<-cbind(1, dat[, c("boy", "age", "highschool",
                  "somecollege", "college")])
x_tld<-dat[, "married"]
```

a)

```
ww_inv <- solve(t(w) %*% w)
wy <- t(w) %*% y
beta <- ww_inv %*% wy # 172
```

Beta here represents the change in birthweight associated with an additional cigarette smoked per day by the mother.

b)

```
x <- cbind(x, w)
xx_inv <- solve(t(x) %*% x)
xy <- t(x) %*% y
beta <- xx_inv %*% xy
```

```
beta[length(beta)]
```

```
## [1] -11.06251
```

Beta here represents the change in birthweight associated with an additional cigarette smoked per day by the mother when controlling for gender, age, and level of education.

c)

It does differ because we are controlling for other variables in addition to cigarettes smoked per day.

d)

Our result only shows that an increase in the number of cigarettes smoked per day is associated with lower birthweight, not that it causes lower birthweight.

e)

Common support states that $suppX = suppX|W$ i.e. that for every set of observables (boy, age, highschool, somecollege, college) there are individuals at each level of $W = w'$, here number of cigarettes smoked per day.

Selection on observables states that W is independent of $U|X$. Here, it means that the weight of the baby at birth is independent of the unobservables i.e. only related to the observables (boy, age, highschool, somecollege, college).

f)

```
w<-dat[, "cigsdaily"]
x<-cbind(1, dat[, c("boy", "age", "highschool",
                  "somecollege", "college")])
X<-cbind(x, w)

all_true <- function(list){
  for(x in 1:length(list)){
    if(!list[x]){
      return(list[x])
    }
  }
  return(TRUE)
}

dupes <- X[duplicated(X),]
matches <- dupes[1,]
count<-0
for(row in 1:nrow(dupes)){
  if(all_true(dupes[1, -ncol(dupes)]==dupes[row, -ncol(dupes)]) & (count==1)){
    matches <- rbind(matches, dupes[row,])
  }
  else{
    if(count!=1){
      count<-1
    }
  }
}

sprintf("Number of observations that have the same x as observation 1: %i", nrow(matches))
```

```
## [1] "Number of observations that have the same x as observation 1: 156"
```

```
sprintf("Unique w for the same x: %i",nrow(unique(matches)))
```

```
## [1] "Unique w for the same x: 6"
```

```
sprintf("Total number of unique w: %i",length(unique(w)))
```

```
## [1] "Total number of unique w: 25"
```

We cannot validate CS because for a given x, it does not cover all values of w in the sample.

Problem 7

a)

```
# Define a custom function to compute the ols estimates
my_coef<-function(y,X){
  xx_inv <- solve(t(X) %*% X )
  xy <- t(X) %*% y
  beta <- xx_inv %*% xy
  return(beta)
}#MY_COEF
# Test the function using your solution to Problem 6
coef<-my_coef(y,X)
coef
```

```
##                [,1]
##          3036.552960
## boy          113.250406
## age           3.865582
## highschool    90.509604
## somecollege  144.353669
## college      157.879709
## w            -11.062508
```

b)

```
# Define a custom function to compute the blp estimates
my_blp <- function(coef,x) {
  blp_yx <- x %*% coef
  return(blp_yx)
}#MY_BLP

# Test the function
mean(y-my_blp(coef,X)) # 0
```

```
## [1] 1.241753e-11
```

c)

```
# Define a custom function to compute the standard error
my_se <- function(coef,y,X) {
  n <- nrow(X)
  epsilon <- c(y - my_blp(coef,X))
  XX_inv <- solve(t(X) %*% X)
```

```

XX_eps2<-t(X * epsilon) %*% (X*epsilon)
Sigma <- (XX_inv %*% XX_eps2 %*% XX_inv)*n
se <- sqrt(diag(Sigma))/sqrt(n)
return(se)
}#MY_SE
# Test the function using your solution to Problem 6
se <- my_se(coef,y,X)
se

```

```

##           boy           age  highschool  somecollege      college
## 30.021512  12.643432   1.319258   18.191861   20.572197   24.341979
##           w
## 1.349021

```

d)

```

# Define a custom function to compute the test stat and p-value
my_teststat <- function(beta,se) {
  t <- abs(beta/se)
  p <- 2*(1-pnorm(t))
  return(list(t,p))
}#MY_TESTSTAT
# Test the function
my_teststat(coef,se)

```

```

## [[1]]
##           [,1]
##           101.145904
## boy           8.957252
## age           2.930118
## highschool    4.975280
## somecollege   7.016930
## college       6.485903
## w            8.200395
##
## [[2]]
##           [,1]
##           0.000000e+00
## boy           0.000000e+00
## age           3.388334e-03
## highschool    6.515341e-07
## somecollege   2.267964e-12
## college       8.820211e-11
## w            2.220446e-16

```

e)

```

# Define a custom function to compute and characterize ols estimates
my_ols <- function(y,X){
  coef <- my_coef(y,X)
  se <- my_se(coef,y,X)
  tp <- my_teststat(coef,se)
  return(list(coef,se,tp[1],tp[2]))
}#MY_OLS

```

```
# Test the function using your solution to Problem 6
my_ols(y,X)
```

```
## [[1]]
##                [,1]
##          3036.552960
## boy          113.250406
## age           3.865582
## highschool    90.509604
## somecollege  144.353669
## college      157.879709
## w            -11.062508
##
## [[2]]
##                boy          age highschool somecollege college
## 30.021512 12.643432 1.319258 18.191861 20.572197 24.341979
##                w
## 1.349021
##
## [[3]]
## [[3]][[1]]
##                [,1]
##          101.145904
## boy          8.957252
## age           2.930118
## highschool    4.975280
## somecollege   7.016930
## college       6.485903
## w             8.200395
##
##
## [[4]]
## [[4]][[1]]
##                [,1]
##          0.000000e+00
## boy          0.000000e+00
## age           3.388334e-03
## highschool    6.515341e-07
## somecollege   2.267964e-12
## college       8.820211e-11
## w             2.220446e-16
```