Class 4 — iTV Experiments
Power, Effects, Covariates

Jake Bowers

Here are a few of the questions that arose during my reading. You will have others, so this list is not meant to restrict our discussion.

1. What is statistical power?
2. To talk about power, what do we need?
3. Gerber and Green do not dwell on power per se. What do they focus their attention on instead? Why might they do this?
4. Why block?
5. What problems arise when blocking? How might one work around those problems?
6. Conceptually, how might you compare experimental designs if you had some background information on your experimental pool (or a sample like your experimental pool)? What steps would you take?
7. Review: Imagine we wanted to make a policy, and we had money to randomly sample people from the population potentially subject to the policy. We then ran a randomized experiment on the random sample. How might or might not results like an ATE or a hypothesis test from this procedure be a "reliable guide to policy in the parent population"?
8. When would a structural model be useful even given a randomized experiment?
9. Many of you expressed interest in individual attitudes and behaviors as outcomes from your research projects. Today, we will play around with designing an experiment focused on individual racial attitudes using the 2008 American National Election Survey.

```
> ##Load a small version of the 2008 American National Election Study data
> load(url("http://jakebowers.org/PS230Data/nes08sm.rda"))
> ## the name of the dataframe is nes08sm
```

Here are the variables in the dataset:

V085065c	D3c. Feeling thermometer: WHITES
V085064y	D2y. Feeling thermometer: BLACKS
V085064a	D2a. Feeling thermometer: HISPANICS
V085064v	D2v. Feeling thermometer: ASIAN-AMERICANS
V085063b	D1b. Feeling thermometer: Democratic Presidential candidate
V085063c	D1c. Feeling thermometer: Republican Presidential candidate
V085084	G1a. Liberal-Conservate: self placement
V085084a	G1b. If had to choose, liberal or conservative
V083098x	J1x. SUMMARY: R Party Identification
V081101	HHList.1. Respondent: gender
V081102	HHList.2. Respondent: race
V081103	HHList.3. Respondent: Latino status
V081103a	HHList.3a. Respondent: race and Latino status
V081104	HHList.4. Respondent: age
V083265a	Y31a. How many children in HH age 10 and younger
V083265b	Y31b. How many children in HH age 11-17
V083248x	Y21ax. SUMMARY: HOUSEHOLD INCOME
V083218x	Y3x. SUMMARY: R educational attainment

```
> ## Some variables that I created and added to the dataset with this command
> ## no need to run this
> nes08sm<-within(nes08sm,{
+   ftdemrep<-V085063b-V085063c
+   ideo<-V085084
+   ideo[ideo==4&(V085084a==1&!is.na(V085084a))]<-3
+   ideo[ideo==4&(V085084a==3&!is.na(V085084a))]<-5
+   ideo[ideo==4&(V085084a==5&!is.na(V085084a))]<-4
+   pid<-V083098x
+   female<-as.numeric(V081101==2)})
```

You can see the codebook for this little file here http://jakebowers.org/PS230Data/nes08cbk-small.txt.

10. The outcome for today is a the difference in feeling thermometer ratings between "Blacks" and "Asian-Americans":

```
> nes08sm$ftblkasn<-with(nes08sm,V085064y-V085064v)
> summary(nes08sm[,c("ftblkasn","V085064y","V085064v")])

    ftblkasn          V085064y         V085064v
 Min.   :-85.000   Min.   :  0.00   Min.   :  0.00
 1st Qu.:  0.000   1st Qu.: 50.00   1st Qu.: 50.00
```

```
 Median :  0.000   Median : 70.00   Median : 60.00
 Mean   :  6.083   Mean   : 72.04   Mean   : 65.78
 3rd Qu.: 15.000   3rd Qu.: 85.00   3rd Qu.: 85.00
 Max.   :100.000   Max.   :100.00   Max.   :100.00
 NA's   :318       NA's   :267      NA's   :311
```

```
> nes08sm$black1<-with(nes08sm,ifelse(V081102 %in% c(2,3,6,7),1,0))
> with(nes08sm,table(V081102,black1,useNA="ifany"))
```

```
       black1
V081102    0    1
   1    1442    0
   2       0  583
   4     262    0
   5      16    0
   6       0    6
   7       0    2
   <NA>   11    0
```

Also, now let us work with a smaller number of cases, like 100:

```
> set.seed(123456)
> nes08small<-nes08sm[sample(1:nrow(nes08sm),100,replace=FALSE),]
> nes08small<-nes08small[!is.na(nes08small$ftblkasn),] ## no missing outcomes
> summary(nes08small[,c("ftblkasn","V085064y","V085064v")])
```

```
    ftblkasn          V085064y         V085064v
 Min.   :-50.000   Min.   :  0.00   Min.   : 30.00
 1st Qu.:  0.000   1st Qu.: 50.00   1st Qu.: 50.00
 Median :  0.000   Median : 70.00   Median : 60.00
 Mean   :  5.581   Mean   : 70.99   Mean   : 65.41
 3rd Qu.: 15.000   3rd Qu.: 85.00   3rd Qu.: 83.75
 Max.   : 50.000   Max.   :100.00   Max.   :100.00
```

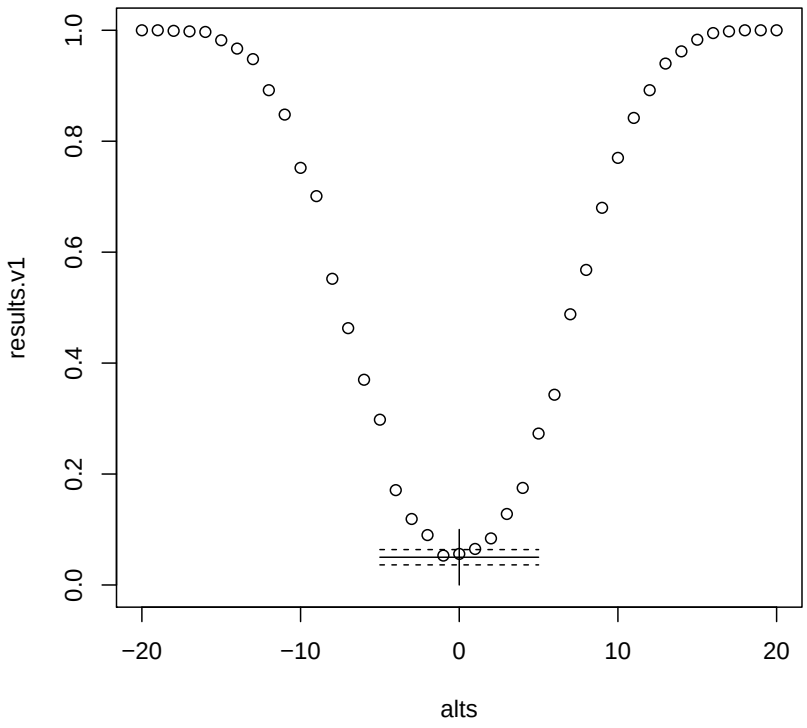
Imagine that I had a binary treatment in mind. And my first proposal for a design was to assign half to treatment and half to control:

```
> nes08small$z1<-sample(rep(c(0,1),nrow(nes08small)/2))
```

Explain what is happening here (understand the code line by line). How would this be useful for designing a study? What could be criticized about what is going on here?

```
> my.fn.v1<-function(z,y,alt){
+   newz<-sample(z)
+   newy<-y - newz*alt
+   thelm<-lm(newy~newz)
+   return(summary(thelm)$coef["newz","Pr(>|t|)"])
+ }
> alts<-seq(-20,20,1)
> nsims<-1000
> alpha<-.05
> tryCatch(load("results.v1.rda"),error=function(e){
+   results.v1<-sapply(alts,function(a){
+     message("Working on ",a)
+     thereps<-replicate(nsims,
+                        my.fn.v1(z=nes08small$z1,
+                                 y=nes08small$ftblkasn,
+                                 alt=a))
+     return(mean(thereps < alpha))
+   })
+   names(results.v1)<-alts
+   save(results.v1,file="results.v1.rda")
+ },
+          finally=load("results.v1.rda")
+          )
>
>
>
```

```
> plot(alts,results.v1,ylim=c(0,1))
> segments(rep(-5,3),
+          c(.05+c(-2,0,2)*sqrt( .05*(1-.05) / nsims)),
+          rep(5,3),
+          c(.05+c(-2,0,2)*sqrt( .05*(1-.05) / nsims)),
+          lty=c(2,1,2))
> segments(0,0,0,.1)
```



11. In case you missed it, we learn two important things about our design (and implicit model of causal effects) from the preceding figure. What are those two operating characteristics? Does this figure look like what we would like to see?

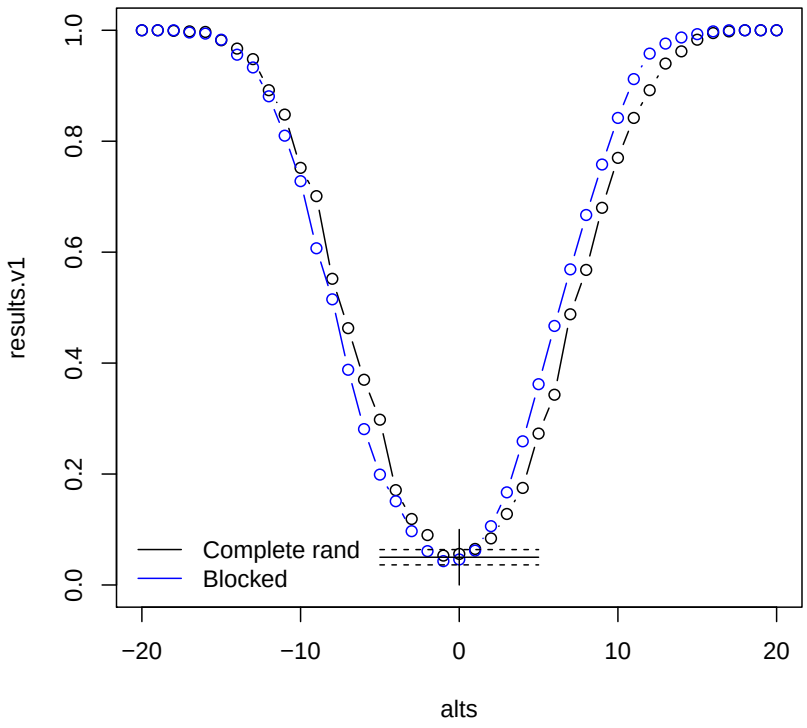
12. So, might have have said both "Yes (it does look like a well-operating test)." and "Compared to what?" for the last question. What other design might be more statistically powerful than the one we just evaluated? Why?

    Here is a function that uses simple blocking:

```
> my.fn.v2<-function(z,y,b,alt){
+   b<-factor(b)
+   newz<-unlist(lapply(split(z,b),sample),b)
+   newy<-y - newz*alt
+   thelm<-lm(newy~newz+b)
+   return(summary(thelm)$coef["newz","Pr(>|t|)"])
+ }
>
>
```

    Choose one variable (or more than one, and use `interaction` or some recoding to combine them into a single factor variable) and see if you can improve on the design I summarize below.

```
> plot(alts,results.v1,ylim=c(0,1),type="b")
> lines(alts,results.v2,col="blue",type="b")
> segments(rep(-5,3),
+          c(.05+c(-2,0,2)*sqrt( .05*(1-.05) / nsims)),
+          rep(5,3),
+          c(.05+c(-2,0,2)*sqrt( .05*(1-.05) / nsims)),
+          lty=c(2,1,2))
> segments(0,0,0,.1)
> legend(x="bottomleft",legend=c("Complete rand","Blocked"),
+        col=c("black","blue"),lty=c(1,1),bty="n")
```



13. How else might you summarize information the precision of your design (other than these power curves)? [Think about the Gerber and Green approach in their book]. How would you change the code above to do this?

14. Can we successfully use Moore's approach to improve precision? What is his suggestion?

```
> ## Just a start on this.
> library(blockTools)
> nes08small$id<-row.names(nes08small)
> tempdat<-na.omit(nes08small[,c("id","V081101","V083098x","V081104",
+                   "V083248x", "V083218x")])
> myblocks<-block(tempdat,id.vars="id",
+                 block.vars=c("V081101","V083098x","V081104",
+                   "V083248x", "V083218x"),
+                 )
>
```

15. Now (and previewing next week) if you haven't already mentioned it, notice that I've been playing a bit fast and loose with estimation of effects and/or testing hypotheses about effects. Why would we believe that (1) $\hat{\beta}$ in an ols model estimates the ATE well and/or that (2) $\hat{SE}(\hat{\beta})$ estimates the spread of the randomization distribution of the ATE well and/or that (3) a t distribution with such and such degrees of freedom is a good approximation for this distribution?