Class 1 — iTV Experiments
Basic Statistical Inference for Causal Quantities

Jake Bowers and Corrine McConnaughy

0. Get ready to work. Today we will be working to deepen our understanding of potential outcomes and the link between counterfactual causal quantities (which, in this course, we define formally as functions of potential outcomes) and statistical inference.

1. In 2005 Costas Panagopoulos randomly assigned treatment with a non-partisan Get-out-the-vote newspaper advertisement in low-salience mayoral elections in two of these four cities ($Z \in \{0, 1\}$ is assignment to treatment with advertisements and Y is observed proportion of the city turning out to vote).

	Z	Y
Oxford	0	23
Lowell	1	27
Yakima	0	58
Richland	1	61

   Assuming SUTVA, please write down the potential outcomes for a given unit i?

2. How would you write the sharp null hypothesis of no effects using your notation for potential outcomes?

3. Explain in words what the sharp null hypothesis of no effects means in terms of this particular research design.

4. Here is how the observed outcome relates to the potential outcomes for a unit i (using one notation scheme, yours may be different):

$$Y_i = Z_i y_{i,1} + (1 - Z_i) y_{i,0} \tag{1}$$

   Explain this equation in words.

5. How does Y_i (what we observe) relate to $y_{i,0}$ (the partially observed control potential outcomes) if we take $H_0 : y_{i,1} = y_{i,0}$ seriously? *Hint:* Try to get rid of $y_{i,1}$ from equation 1.

6. How many elements does the set that Rosenbuam calls Ω have in this design? Write them out.

7. What does Ω represent?

8. Say we summarized the relationship between treatment and outcomes using a difference of means (mean of those assigned to treatment minus mean of those assigned to control). And, say we repeated this experiment on these same cities (at the same moment in time, such that Y_i reveals the same y_i in each experiment) and calculated the difference of means each time:

```
Y<-c(23,27,58,61)
thedist<-apply(Om,2,function(z){ mean(Y[z==1])-mean(Y[z==0]) })
print(thedist)
```

```
[1] -34.5  -3.5  -0.5   0.5   3.5  34.5
```

```
table(thedist)/6
```

```
thedist
-34.5  -3.5  -0.5   0.5   3.5  34.5
0.167 0.167 0.167 0.167 0.167 0.167
```

   What does the variation in these means represent? What is this distribution called? What does this distribution represent?

9. Use this distribution to calculate a one-sided upper-tailed p-value for H_0. *Hint:* You'll need to know the difference of means for the observed treatment assignment.

```
obsZ<-c(0,1,0,1)
```

10. What did we need to assume to make this p-value valid and meaningful?

11. How, let us loosen one of those assumptions: Please write down the potential outcomes for a unit i when we do not make the SUTVA assumption.

12. Challenge question: Write down a sharp null hypothesis of no effects without SUTVA. Test this hypothesis as we tested the sharp null of no effects under SUTVA above. A bit of code below generates Ω:

```
Om<-combn(4,2,FUN=function(x){ tmp<-rep(0,4); tmp[x]<-1; return(tmp)})
```