1. 第一周

变量和被看作为潜在结果(potential outcome)或者反事实结(counterfactual outcome)。这两个值其实是一个假想的值。

当的时候就表明这个处理有因果效应。

1. 为什么随机的就是因果的 --- 数据中不存在混杂因素

因果推断解决的是数据缺失的问题，而随机化是保证数据是随机缺失的。随机化的条件是进行干预的结果不受被实施干预的群体差异所影响。

分组不影响反事实的结果,所以在随机试验中相关关系就是因果关系。

在不存在混杂因素是数据中，相关关系 = 因果关系

1. 有向无环图

有向无环图表示两个变量的作用关系是单向的不能是双向的或者是间接双向的（绕了一圈回去了），一旦有环因果就区分不开了。

1. Confusion over causality

·假相关性：有两个不相关的变量，但是二者在大多数时间呈现一种相关的表现。

·反向因果关系：两个变量之间的因果关系箭头可以为任意的指向

·因果关系的迷惑性：因果关系的箭头可以为不确定性方向、人们根据自己已获取的知识进行判断、…

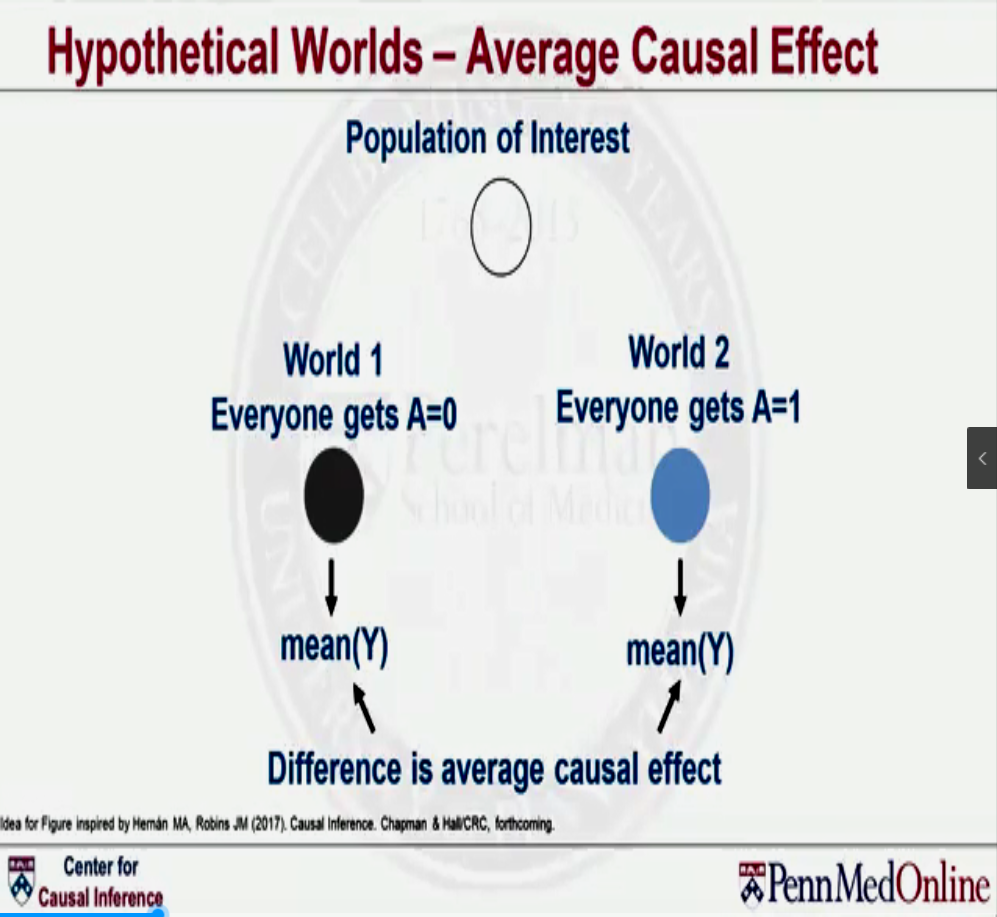
·敏感性分析：敏感性分析就是看你的结果对可能违反我们的因果假设的敏感程度。

sensitivity analysis has to do with is looking at how sensitive your results are to possible violations of some of our causal assumptions.

1. Potential outcomes and counterfactuals
2. Hypothetical interventions
3. Causal effects

两种典型的因果效应：平均因果效应、

1. 平均因果效应 average causal effect



，average value of Y if everyone was treated with A=1 minus the average value of Y if everyone was treated with A=0.

If Y is binary this is a risk difference.

举一个例子，假如 = -20mm Hg. 如果那些服用噻嗪类利尿剂的高血压患者，平均来说，他们的收缩压值会比不服用这种药物时低20个单位。

的意思是在这个集合中所有人都得到了治疗A=1的Y的期望;

而的意思是在整个集合中取出得到了治疗A=1的Y的期望。

1. Causal assumption
2. SUTVA

第一个假设是单位（Units）与单位之间互不干扰

施加在一个单位上的治疗不会影响到另一个单位的治疗结果。

第二个假设是只进行一种治疗。

1. Consistency

在方案A=a下的潜在结果与观测到的数据中接受A=a治疗的结果一致。

1. Ignorability --- 可交换性：接受的处理和最终的结果没有关联

在控制混淆变量的前提下，认为治疗是被随机分配的。

And the basic idea is that treatment

assignment is assumed to be independent from potential outcomes, conditional

on these pre-treatment variables.

