ORIE 6745 **September 6, 2017**

Lecture 6: Randomized Experiments

Lecturer: Nathan Kallus *Scribe: Amani Moin*

1 Completely Randomized Experiments

Classic Design

Choose $z_{1:n}$ uniformly from $CR_n = \{z_{1:n} : \sum_i z_i = n/2\}$
Note that by choosing $z_{1:n}$ randomly, there may be 2 sources of randomness in $\hat{\tau}$

1. Randomness from random sampling of units $Y_{1:n}(0,1)$ from population
2. Randomness due to random assignment $z_{1:n}$

Sometimes we consider units to be fixed, so we don't consider (1) to be a source of randomness. This is known as finite sample analysis. (2) is still a source of randomness here. Under complete randomization,

$$\begin{aligned}
\hat{\tau} &= \frac{2}{n}\sum_i (-1)^{1+z_i} Y_i \\
&= \frac{2}{n}\sum_i (z_i Y_i - (1-z_i)Y_i) \\
&= \frac{2}{n}\sum_i (z_i Y_i(1) - (1-z_i)Y_i(0)) \; (because\, of\, consistency)
\end{aligned}$$

$$\begin{aligned}
E[\hat{\tau}|Y_{1:n}(0,1)] &= \frac{2}{n}\sum_i (Y_i(1)E[z_i|Y_{1:n}(0,1)] - Y_i(0)E[1-z_i|Y_{1:n}(0,1)]) \\
&= \frac{2}{n}\sum_i (1/2Y_i(1) - 1/2Y_i(0)) \\
&= \frac{1}{n}\sum_i (Y_i(1) - Y_i(0))
\end{aligned}$$

$$\begin{aligned}
Var(\hat{\tau}|Y_{1:n}(0,1)) &= 1/n(2S_0^2 + 2S_1^2 - S_d^2) \\
where\, S_z^2 &= \frac{1}{n-1}\sum_i (Y_i(z) - \bar{Y}(z))^2 \\
and\, \bar{Y}(z) &= \frac{1}{n}\sum_i Y_i(z) \\
where\, S_d^2 &= \frac{1}{n-1}\sum_i (TE_i - SATE)^2 \\
&= \frac{1}{n-1}\sum_i (Y_i(1) - Y_i(0) - \bar{Y}(1) + \bar{Y}(0))^2
\end{aligned}$$

1.1 Asymptotics

If S_z^2 for z=0,1 and S_d^2 are bounded sequences (a form of the strong Lyapunov condition), then $Var(\hat{\tau}|Y_{1:n}(0,1)) = O(1/n)$
For fixed units, $\hat{\tau} - SATE = O_p(1/sqrt(n))$ so $\hat{\tau}$ converges to SATE as long as outcomes don't go to infinity and the sample size is large
$\hat{\tau}$ is finite sample unbiased and $\sqrt{n}$ consistent

When we say $z_n = O_p(a_n) : z_n/a_n$ is stochastically bounded
This means that $\forall\, \epsilon > 0\, \exists\, N, M\; st\; \mathbb{P}(z_n/a_n \geq M) \leq \epsilon$ When we say
$z_n = o_p(a_n) : z_n/a_n \to 0\; in\; \mathbb{P}$
This means $\forall\, \epsilon, \delta > 0\, \exists\, N\; st\; \mathbb{P}(z_n/a_n \geq \epsilon) \leq \delta\, \forall\, n \geq N$

This is why controlled experiments are the gold standard. There is no sampling, no counterfactuals. **If you have controls+noninterference+consistency you can estimate causal effects.**

When units are sampled iid from a population,

$$E[\hat{\tau} - PATE] = E[\hat{\tau} - SATE] = E[E[\hat{\tau} - SATE|Y_{1:n}(0,1)]] = 0$$
$$Var(\hat{\tau}) = \frac{4}{n}Var(\frac{Y(1) - Y(0)}{2}) + \frac{4}{n}Var(\frac{Y(1) + Y(0)}{2})$$

$\hat{\tau}$ is unbiased (marginally) and $\hat{\tau} - PATE = O_p(1/\sqrt{n})$ (marginally)

2 Inference on Effect

$\hat{\tau}$ is a point estimate for SATE or PATE.
Suppose $\hat{\tau} \neq 0$. Can we conclude the treatment has a real effect?
Inference=assessing the validity of a hypothesis

3 Permutation Tests, Fisher's Exact p-value

Given a fixed finite sample, we want to assess the validity of the Sharp(Fisherian) null hypothesis:
$H_0 : Y_i(0) = Y_i(1) \forall i = 1, 2, ..., n$ or $TE_i = 0$
"Is $\hat{\tau}$ big enough to eliminate the possibility of H_0?"
How do we test H_0? How big is big enough?

Suppose H_0 is true. How typical is $\hat{\tau}$? ie how extreme is $\hat{\tau}$ in magnitude in it's own distribution? If it is not extreme then $\hat{\tau}$ may be nonzero by chance. If it is extreme it is unlikely to be so big due to chance alone.

What is the distribution of $\hat{\tau}$ under H_0?
$\hat{\tau} = \frac{2}{n}\sum_i(z_iY_i - (1 - z_i)Y_i(0))$
For $z_{1:n}$ drawn uniformly from CR_n we can't compute this distribution because we don't know $Y_i(0)$ and $Y_i(1)$. BUT under H_0, $Y_i(0) = Y_i(1) = Y_i = Y_i(z_i')$ so we can fill in the counterfactuals. Under H_0, $\hat{tau} = \frac{2}{n}\sum_i(-1)^{1+z_i}Y_i(z_i')$ for any $z_{i:n}'$

3.1 Permutation Test Procedure

Generally:

1. Draw $z_{1:n} \in CR_n$
2. assign and apply treatments
3. measure $Y_{1:n}$
4. simulate distribution of $\hat{\tau}$ over new randomization under H_0
   (a) draw new $z'_{1:n}$
   (b) $\hat{\tau} = \frac{2}{n}\sum_i (-1)^{1+z_i} Y_i$

Specifically:

1. Fix significance level $\alpha \in (0,1)$
2. For each $z'_{1:n} \in CR_n$ compute $\hat{\tau}_{z'_{1:n}} = \frac{2}{n}\sum_i (-1)^{1+z'_i} Y_i$
3. set $p = \frac{\sum_i z'_{1:n} \in CR_n \mathbb{I}[|\hat{\tau}_{z'_{1:n}}| \geq |\hat{\tau}|]}{|CR_n|}$. This p is Fisher's exact p-value.
4. If $p \leq \alpha$ then reject H_0. In other words, $|\hat{\tau}|$ is so large that it is bigger than the $|CR_n|(1-\alpha)^{th}$ largest $|\hat{\tau}_{z'}|$.
   (a) draw new $z'_{1:n}$
   (b) $\hat{\tau} = \frac{2}{n}\sum_i (-1)^{1+z_i} Y_i$

3.2 Significance of the permutation test

If H_0 were true we'd only reject it incorrectly no more than α fraction of trials Exact significance: if $|CR_n|\alpha$ is integral then if H_0 were true we would reject it exactly α fraction of trials

3.3 Randomization test

If we have $\binom{n}{n/2}$ permutations, we try all of them. For n=100, $\binom{n}{n/2} = \binom{100}{50} > 10^29$. This makes the permutation test infeasible so we consider just a few of the permutations. Choose $z'_{1,1:n}, ...z'_{B,1:n}$ uniformly at random from CR_n and let $p = \frac{1+\mathbb{I}[|\hat{\tau}_{z'_{1:n}}| \geq |\hat{\tau}|]}{1+B}$