Randomization Inference

September 20, 2012

Randomization
Inference

Basic Setup

The Null
Hypothesis

Test Statistics

Basic Setup

- Using Rosenbaum's notation: there are N units divided into S strata or *blocks*, which are formed on the basis of pre-treatment characteristics.

Randomization
Inference

Basic Setup

The Null
Hypothesis

Test Statistics

Basic Setup

- Using Rosenbaum's notation: there are N units divided into S strata or *blocks*, which are formed on the basis of pre-treatment characteristics.

- A unit is an opportunity to apply or withhold the treatment.

Randomization
Inference

Basic Setup

The Null
Hypothesis

Test Statistics

Basic Setup

- Using Rosenbaum's notation: there are N units divided into S strata or *blocks*, which are formed on the basis of pre-treatment characteristics.

- A unit is an opportunity to apply or withhold the treatment.

- There are n_s units in stratum s for $s = 1, ..., S$, so $N = \sum n_s$.

Randomization
Inference

Basic Setup

The Null
Hypothesis

Test Statistics

Basic Setup

- Using Rosenbaum's notation: there are N units divided into S strata or *blocks*, which are formed on the basis of pre-treatment characteristics.

- A unit is an opportunity to apply or withhold the treatment.

- There are n_s units in stratum s for $s = 1, ..., S$, so $N = \sum n_s$.

- Write $Z_{si} = 1$ if the ith unit in stratum s receives the treatment and write $Z_{si} = 0$ if this unit receives control.

Randomization
Inference

Basic Setup

The Null
Hypothesis

Test Statistics

Basic Setup

- Using Rosenbaum's notation: there are N units divided into S strata or *blocks*, which are formed on the basis of pre-treatment characteristics.

- A unit is an opportunity to apply or withhold the treatment.

- There are n_s units in stratum s for $s = 1, ..., S$, so $N = \sum n_s$.

- Write $Z_{si} = 1$ if the ith unit in stratum s receives the treatment and write $Z_{si} = 0$ if this unit receives control.

- Write m_s for the number of treated units in stratum s, so $m_s = \sum_{i=0}^{n_s} Z_{si}$ and $0 \leq m_s \leq n_s$.

Randomization
Inference

Basic Setup

The Null
Hypothesis

Test Statistics

Treatment Assignment

- The most common assignment mechanism fixes the number of m_s in stratum s.

Randomization
Inference

Basic Setup

The Null
Hypothesis

Test Statistics

Treatment Assignment

- The most common assignment mechanism fixes the number of m_s in stratum s.
- Let Ω be the set containing $K = \Pi_{s=1}^{s} \binom{n_s}{m_s}$ possible treatment assigments **z**.

Randomization
Inference

Basic Setup

The Null
Hypothesis

Test Statistics

Treatment Assignment

- The most common assignment mechanism fixes the number of m_s in stratum s.

- Let Ω be the set containing $K = \Pi_{s=1}^{s} \binom{n_s}{m_s}$ possible treatment assigments $\mathbf{z}$.

- In the most common experiments, each of these K possible assignments is given the same probability, $\mathrm{prob}(\mathbf{Z} = \mathbf{z}) = 1/K$ for all $\mathbf{z}$ in Ω.

Randomization
Inference

Basic Setup

The Null
Hypothesis

Test Statistics

Treatment Assignment

- The most common assignment mechanism fixes the number of m_s in stratum s.

- Let Ω be the set containing $K = \Pi_{s=1}^{s} \binom{n_s}{m_s}$ possible treatment assignments $\mathbf{z}$.

- In the most common experiments, each of these K possible assignments is given the same probability, $\text{prob}(\mathbf{Z} = \mathbf{z}) = 1/K$ for all $\mathbf{z}$ in Ω.

- For example:

$$\Omega = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} \\ \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

Randomization
Inference

Basic Setup

The Null
Hypothesis

Test Statistics

Questions on Design

- When is m a random variable? $Pr(m = n) > 0$?
- What about fixed m_s?
- What are the units in a cluster randomized experiments?
- SUTVA
    - interference outside of the design?
- Underscores importance of known assignment and study design

Randomization
Inference

Basic Setup

The Null
Hypothesis

Test Statistics

The Sharp Null

- The most common hypothesis associated with randomization inference is the sharp null of no effect for all units.

Randomization
Inference

Basic Setup

The Null
Hypothesis

Test Statistics

The Sharp Null

- The most common hypothesis associated with randomization inference is the sharp null of no effect for all units.

- A unit labeled as "treated" will have the exact same outcome as a unit labeled as "control".

Randomization
Inference

Basic Setup

The Null
Hypothesis

Test Statistics

The Sharp Null

- The most common hypothesis associated with randomization inference is the sharp null of no effect for all units.

- A unit labeled as "treated" will have the exact same outcome as a unit labeled as "control".

- Under the null, the units' responses are *fixed* and the only random element is the meaningless rotation of labels.

Randomization
Inference

Basic Setup

The Null
Hypothesis

Test Statistics

The Sharp Null

- The most common hypothesis associated with randomization inference is the sharp null of no effect for all units.

- A unit labeled as "treated" will have the exact same outcome as a unit labeled as "control".

- Under the null, the units' responses are *fixed* and the only random element is the meaningless rotation of labels.

- When testing the null hypothesis of no effect, the response of the ith unit in stratum s can be written r_{si} and the vector of responses is **r**.

$$r_{1si} = r_{0si} + \tau_0$$

$$r_{si} = r_{0si} Z_{si} + r_{0si}(1 - Z_{si})$$

Randomization
Inference

Basic Setup

The Null
Hypothesis

Test Statistics

The Test Statistic

- A **test statistic** $t(\mathbf{Z}, r)$ is a quantity computed from the treatment assignment $\mathbf{Z}$ and the response r.

Randomization
Inference

Basic Setup

The Null
Hypothesis

Test Statistics

The Test Statistic

- A **test statistic** $t(\mathbf{Z}, r)$ is a quantity computed from the treatment assignment $\mathbf{Z}$ and the response r.

- The most commonly used test-statistic is the point estimate for the average treatment effect (weighted by strata proportion).

$$\sum_{s=1}^{S} \frac{n_s}{N} \sum_{i=1}^{n_s} \left\{ \frac{Z_{si} r_{si}}{m_s} - \frac{(1 - Z_{si}) r_{si}}{n_s - m_s} \right\}$$

Randomization
Inference

Basic Setup

The Null
Hypothesis

Test Statistics

Significance Test

- To compute the *p*-value for any given test statistic, we simply calculate the proportion of treatment assignments **z** in Ω giving values of $t(\mathbf{z}, \mathbf{r})$ greater than or equal to the observed T, namely:

$$\mathrm{prob}\{t(\mathbf{Z}, \mathbf{r} \geq T\} = \frac{|\{\mathbf{z} \in \Omega : t(\mathbf{z}, \mathbf{r}) \geq T\}|}{K}$$

Randomization
Inference

Basic Setup

The Null
Hypothesis

Test Statistics

Significance Test

- To compute the *p*-value for any given test statistic, we simply calculate the proportion of treatment assignments **z** in Ω giving values of $t(\mathbf{z}, \mathbf{r})$ greater than or equal to the observed T, namely:

$$\mathrm{prob}\{t(\mathbf{Z}, \mathbf{r} \geq T\} = \frac{|\{\mathbf{z} \in \Omega : t(\mathbf{z}, \mathbf{r}) \geq T\}|}{K}$$

- What happens when:

$$\mathrm{prob}(\mathbf{Z} = \mathbf{z}) \neq 1/K$$

Randomization
Inference

Basic Setup

The Null
Hypothesis

Test Statistics

Significance Test

- To compute the *p*-value for any given test statistic, we simply calculate the proportion of treatment assignments **z** in Ω giving values of $t(\mathbf{z}, \mathbf{r})$ greater than or equal to the observed T, namely:

$$\text{prob}\{t(\mathbf{Z}, \mathbf{r} \geq T\} = \frac{|\{\mathbf{z} \in \Omega : t(\mathbf{z}, \mathbf{r}) \geq T\}|}{K}$$

- What happens when:

$$\text{prob}(\mathbf{Z} = \mathbf{z}) \neq 1/K$$

- The above p-value is for a one-tailed test. What about a two-tailed test? There is some disagreement in the literature about this, but Rosenbaum recommends simply doubling the one-tailed p-value.

A Class of Test Statistics

- A widely-used statistic is the difference in means, but many others are possible.

Randomization
Inference

Basic Setup

The Null
Hypothesis

Test Statistics

A Class of Test Statistics

- A widely-used statistic is the difference in means, but many others are possible.

- Mean difference statistics will have low power given outliers, skewed distributions, or heavy tailed distributions.

Randomization
Inference

Basic Setup

The Null
Hypothesis

Test Statistics

A Class of Test Statistics

- A widely-used statistic is the difference in means, but many others are possible.

- Mean difference statistics will have low power given outliers, skewed distributions, or heavy tailed distributions.

- A class of non-parametric tests:

$$t(Z_{is}, r_{is}) = \sum_{s=1}^{S} d_s \sum_{i=1}^{n_s} c_{si} Z_{si}$$

Randomization
Inference

Basic Setup

The Null
Hypothesis

Test Statistics

A Class of Test Statistics

- A widely-used statistic is the difference in means, but many others are possible.

- Mean difference statistics will have low power given outliers, skewed distributions, or heavy tailed distributions.

- A class of non-parametric tests:

$$t(Z_{is}, r_{is}) = \sum_{s=1}^{S} d_s \sum_{i=1}^{n_s} c_{si} Z_{si}$$

- One common alternative is the **Wilcoxon rank sum test**. In a completely randomized experiment ($S=1$), the **r** responses are ranked from smallest to largest to produce **d** (with average ranks for ties). Sum ranks for all units assigned **Z** $= 1$

Randomization
Inference

Basic Setup

The Null
Hypothesis

Test Statistics

Rank Tests

- **Stratified rank sum test**: For block randomized experiments, one easy extension of the rank sum test is to calculate the rank sum test separately in each strata and take the sum of these S rank sums as the test statistic.

Randomization
Inference

Basic Setup

The Null
Hypothesis

Test Statistics

Rank Tests

- **Stratified rank sum test**: For block randomized experiments, one easy extension of the rank sum test is to calculate the rank sum test separately in each strata and take the sum of these S rank sums as the test statistic.

- **Aligned rank test**: Hodges and Lehmann (1962) find, when S is large relative to N, a more efficient rank test for block randomized experiment is the aligned rank statistic. For this statistic, subtract the mean of each stratum from the responses in that stratum, creating "aligned responses". Rank the aligned responses without regard to block. The aligned rank statistic is the sum of the aligned ranks in the treated group.

Randomization
Inference

Basic Setup

The Null
Hypothesis

Test Statistics

Covariate Adjustment

- Write $\tilde{\epsilon}(\cdot)$ for a function that creates residuals ($\tilde{\epsilon}(\mathbf{r}) = \mathbf{e}$) from $\mathbf{r}$, which are the oucomes under the null hypothesis, and $\mathbf{X}$, which is a matrix of covariates.

Randomization
Inference

Basic Setup

The Null
Hypothesis

Test Statistics

Covariate Adjustment

- Write $\tilde{\epsilon}(\cdot)$ for a function that creates residuals ($\tilde{\epsilon}(\mathbf{r}) = \mathbf{e}$) from $\mathbf{r}$, which are the oucomes under the null hypothesis, and $\mathbf{X}$, which is a matrix of covariates.

- $\tilde{\epsilon}(\cdot)$ can be a simple linear model, some non-parametric smoother such as lowess, a robust linear model, etc.

Randomization
Inference

Basic Setup

The Null
Hypothesis

Test Statistics

Covariate Adjustment

- Write $\tilde{\epsilon}(\cdot)$ for a function that creates residuals ($\tilde{\epsilon}(\mathbf{r}) = \mathbf{e}$) from $\mathbf{r}$, which are the oucomes under the null hypothesis, and $\mathbf{X}$, which is a matrix of covariates.

- $\tilde{\epsilon}(\cdot)$ can be a simple linear model, some non-parametric smoother such as lowess, a robust linear model, etc.

- The point of adjustment is to reduce dispersion in $\mathbf{r}$, so choose $\tilde{\epsilon}(\cdot)$ with that goal in mind.

Randomization
Inference

Basic Setup

The Null
Hypothesis

Test Statistics

Covariate Adjustment

- Write $\tilde{\epsilon}(\cdot)$ for a function that creates residuals ($\tilde{\epsilon}(\mathbf{r}) = \mathbf{e}$) from $\mathbf{r}$, which are the oucomes under the null hypothesis, and $\mathbf{X}$, which is a matrix of covariates.

- $\tilde{\epsilon}(\cdot)$ can be a simple linear model, some non-parametric smoother such as lowess, a robust linear model, etc.

- The point of adjustment is to reduce dispersion in $\mathbf{r}$, so choose $\tilde{\epsilon}(\cdot)$ with that goal in mind.

- Remember that under the null hypothesis, nothing is stochastic except for the shuffling of treatment assignment labels. As a result $\mathbf{e}$ is a fixed quantity, not a random variable or a by-product of estimation.

Randomization
Inference

Basic Setup

The Null
Hypothesis

Test Statistics

Covariate Adjustment

- Write $\tilde{\epsilon}(\cdot)$ for a function that creates residuals ($\tilde{\epsilon}(\mathbf{r}) = \mathbf{e}$) from $\mathbf{r}$, which are the oucomes under the null hypothesis, and $\mathbf{X}$, which is a matrix of covariates.

- $\tilde{\epsilon}(\cdot)$ can be a simple linear model, some non-parametric smoother such as lowess, a robust linear model, etc.

- The point of adjustment is to reduce dispersion in $\mathbf{r}$, so choose $\tilde{\epsilon}(\cdot)$ with that goal in mind.

- Remember that under the null hypothesis, nothing is stochastic except for the shuffling of treatment assignment labels. As a result $\mathbf{e}$ is a fixed quantity, not a random variable or a by-product of estimation.

- $\mathbf{e}$, however, may be less dispersed than $\mathbf{r}$ because some of the variation in $\mathbf{r}$ will have been captured by $\mathbf{X}$.

Randomization
Inference

Basic Setup

The Null
Hypothesis

Test Statistics

Covariate Adjustment

- Under the null hypothesis, since $r_T = r_c$, $e_T = e_C$.

Randomization
Inference

Basic Setup

The Null
Hypothesis

Test Statistics

Covariate Adjustment

- Under the null hypothesis, since $\mathbf{r_T} = \mathbf{r_c}$, $\mathbf{e_T} = \mathbf{e_C}$.

- So once can simply use the test statistic $t(\mathbf{z}, \mathbf{e})$ instead of $t(\mathbf{z}, \mathbf{r})$.

Randomization
Inference

Basic Setup

The Null
Hypothesis

Test Statistics

Covariate Adjustment

- Under the null hypothesis, since $r_T = r_c$, $e_T = e_C$.

- So once can simply use the test statistic $t(\mathbf{z}, \mathbf{e})$ instead of $t(\mathbf{z}, \mathbf{r})$.

- With $\mathbf{e}$ in hand, just proceed as you would with $\mathbf{r}$.

Randomization
Inference

Basic Setup

The Null
Hypothesis

Test Statistics

Example: Democratization Aid in
the Republic of Georigia

- The Republic of Georgia: recipient of US "democratization aid", foreign aid intended to bolster democratic processes.

Randomization
Inference

Basic Setup

The Null
Hypothesis

Test Statistics

Example: Democratization Aid in
the Republic of Georigia

- The Republic of Georgia: recipient of US "democratization aid", foreign aid intended to bolster democratic processes.

- Due to a previous history of fraudulent elections, US government and civil society groups wanted to encourage citizen monitoring of elections.
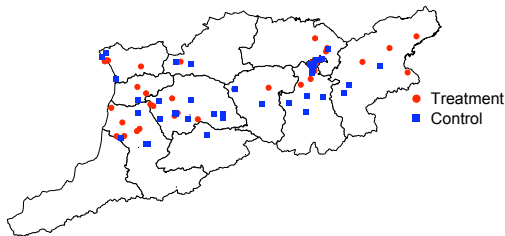
Randomization
Inference

Basic Setup

The Null
Hypothesis

Test Statistics

Example: Democratization Aid in the Republic of Georigia

- The Republic of Georgia: recipient of US "democratization aid", foreign aid intended to bolster democratic processes.

- Due to a previous history of fraudulent elections, US government and civil society groups wanted to encourage citizen monitoring of elections.

- Danny Hidalgo (and coauthors) conducted a program evaluation of one such effort: a simple information campaign to give voters the information necessary for filing a formal complaint with civil society groups or election officials if they witnessed problems on election day.

Randomization
Inference

Basic Setup

The Null
Hypothesis

Test Statistics

Example: Democratization Aid in
the Republic of Georigia

- The Republic of Georgia: recipient of US "democratization aid", foreign aid intended to bolster democratic processes.

- Due to a previous history of fraudulent elections, US government and civil society groups wanted to encourage citizen monitoring of elections.

- Danny Hidalgo (and coauthors) conducted a program evaluation of one such effort: a simple information campaign to give voters the information necessary for filing a formal complaint with civil society groups or election officials if they witnessed problems on election day.

- The intervention consisted of sending canvassers to knock on doors and hand out fliers in randomly selected precincts.

Randomization
Inference

Basic Setup

The Null
Hypothesis

Test Statistics

Example: Randomization Procedure

● Treatment
■ Control

Ths structure of randomization was as follows.

- 36 rural precincts were in blocks of 2, one treatment and one control. So for these precincts, $m_s = 1$ and $n_s = 2$.

- 48 urban precincts were in blocks of 4, two in treatment and and two in control ($m_s = 2$ and $n_s = 4$).

Randomization
Inference

Basic Setup

The Null
Hypothesis

Test Statistics

Some R code

How big is Ω?

```
choose(2,1)^18 * choose(4,2)^12
[1] 5.706304e+14
```

Let's create a function that will assign treatment repeatedly.

```
treat.assign <- function(treat,blocks=NA){
  if(length(unique(blocks))==1){
    treat.vector <- sample(treat)
  }
  else{
  treat.vector <- tapply(treat,blocks,sample)
  treat.vector <- unlist(treat.vector)
  }
  return(treat.vector)
}
```

Randomization
Inference

Basic Setup

The Null
Hypothesis

Test Statistics

R Code

Let's create our distribution of treatment vectors. We could compute all 5.7×10^{14} treatment vectors, but to save on computing time, we can sample a large number of possible treatment vectors to get "close-to-exact" p-values. If our experiment were smaller, then exhaustive enumeration would be better.

Let's use the `replicate` function to assign treatment 5,000 times and generate our Ω:

```
omega <- replicate(5000,
             treat.assign(treat,blocks))
omega <- unique(omega,MARGIN=2)
```