Randomization

Fill In Your Name

12 February 2021

Randomization

Some Common Types of Randomization

Best practices

Some Experimental Designs

Some Limits

Randomization

Randomization of treatment assignment

Start with a group units who you can directly observe: like the people who showed up to a lab, villages in a region, respondents willing to take a survey.

▶ Every observation has a known probability of treatment assignment between 0 and 1.

   ▶ No unit in the experimental sample is assigned to treatment or to control with certainty.

▶ Units can vary in their probability of treatment assignment.

   ▶ For example, the probability might vary by group: women might have a 75% probability of being assigned to treatment.

   ▶ Assignment probabilities can even vary across units **as long as you know the probability for each and every unit**, though that would complicate your analysis.

Random assignment ≠ random sampling

▶ Random sampling (from population): selecting subjects into your sample from a population with known probability. You cannot directly observe the whole population so you draw a sample.

▶ Randomization (of treatment): assigning subjects from an existing group of subjects with known probability to experimental conditions.

 ▶ You need to be able to see outcomes and assign treatment for every subject in the group
 ▶ This random assignment of treatment can be combined with any kind of sample (random sample, convenience sample, etc.) or even with the whole population (all the hospitals, all the judges).

Random assignment $\neq$ random sampling

- How you recruit your initial group (or experimental pool) does matter:
    - a big initial group means a big experiment and more statistical power to detect small effects
    - an initial group which is a random sample from a known population helps you make the argument that your effects might be the same or similar if you did this experiment with another sample from that population.
    - an experimental pool that is a random sample might also help with the argument that the effect should be similar if you scaled up the intervention to the whole population — depending on the factors driving a global equilibrium.

Randomization is closely linked to your research design

▶ See module on Research Design

▶ Remember that you need to define and justify your control condition,

   ▶ this may be no intervention
   ▶ or a placebo

Level of treatments and outcomes

▶ Treatment can be assigned at different levels: individuals, groups, institutions, communities, time periods, or many different levels.

▶ You may be constrained in what level you can *assign* treatment and *measure* outcomes.

  ▶ Example: Treatment at the classroom level, but outcomes at the student level.

  ▶ Example: Treatment at the district level, but outcomes at the community level.

▶ The level at which treatment is assigned and at which outcomes are measured affect what your study can demonstrate.

Some Common Types of Randomization

Simple or Coin-Flipping

▶ For each unit, flip a coin to see if it will be treated. Then you measure outcomes at the same level as the coin.

▶ The coins don't have to be fair (50-50), but you have to know the probability of treatment assignment.

▶ You can't guarantee a specific number of treated units and control units.

▶ Example: If you have 6 units and you flip a fair coin for each, you have about a 3% chance to assign **all** units to treatment or **all** units assigned to control.

Example code for simple randomization

```
# set the random number seed to make this replicable
set.seed(12345)

# set a sample size
N <- 200

# Generate the simple random assignment
# (Notice that in an experiment we have a single
# trial and thus size=1)
# Our object with N people total is called simple.ra
simple.ra <- rbinom(n = N, size = 1, prob = .5)

# 112 people ended up in the treatment group
sum(simple.ra)
```

```
[1] 112
```

Example code for simple randomization

```r
# you can also use the randomizr package
library(randomizr)

# for replicability
set.seed(23456)
# Simple random assignment uses the simple_ra function
# Our object with N people total is called treatment
treatment <- simple_ra(
  N = N, # total sample size
  prob = .5
) # probability of receiving treatment

sum(treatment)
```

```
[1] 96
```

Complete or Urn-Drawing

- A fixed number m out of N units are assigned to treatment. No more or less than $m \pm 1$ units are assigned treatment (depending on whether N is an even or odd number).

- The probability a unit is assigned to treatment is m/N.

- This is like having an urn or bucket with N balls, of which m are marked as treatment and $N - m$ are marked as control. Public lotteries use this method.

Example code for complete randomization

```r
# set sample size N
N <- 200
# set number of treated units m
m <- 100

# create a vector of m 1's and N-m 0's
complete.ra <- c(rep(1, m), rep(0, N - m))

# And then scramble it randomly using sample()
# The default is sampling without replacement

set.seed(12345) # for replicability
complete.ra <- sample(complete.ra)

sum(complete.ra)
```

```
[1] 100
```

Example code for complete randomization

```r
# you can also use the randomizr package
library(randomizr)

# for replicability
set.seed(23456)

# Complete random assignment:
treatment <- complete_ra(
  N = 200, # total sample size
  m = 100
) # number to assign to treatment

sum(treatment)
```

```
[1] 100
```

Block or Stratified Randomization

▶ We create blocks of units and randomize separately within each block. We are doing mini-experiments in each block.

   ▶ Example: block = district, units = communities. We randomize treatment at the community level **within district** and also measure outcomes at the community level.

▶ Blocks that represent a substantively meaningful subgroup can help you to learn about how effects might differ by subgroup.

   ▶ By controlling number of subjects per subgroup, you ensure that you have enough subjects in each group
   ▶ Especially useful when you have a rare group — by chance you might get very few of them in treatment or control even under random assignment (or you might have some imbalance)

▶ Blocks that are homogeneous on a given outcome increase precision of estimation for that outcome compared to the experiment without blocks. (We will address this more in the power analysis section)

egap EVIDENCE IN GOVERNANCE AND POLITICS

Cluster Randomization

▶ A cluster is a **group of units**. In a cluster-randomized trial all units in the cluster get the same treatment status.

▶ Use if the intervention has to work at the cluster level.

  ▶ For example, if the intervention is about school yards, then the school is the unit of assignment even if student health may be an outcome measured at level of students.

Cluster Randomization

- ▶ Having fewer clusters hurts your ability to detect treatment effects and make cause misleading *p*-values and confidence intervals (or even estimates). *How much* depends on the intra-cluster correlation (ICC or ρ).
  - ▶ Higher ρ is worse:
    - ▶ When $\rho = 1$ then every person in the village would give exactly the same answer (i.e. even if you have 100 people, they are identical from the perspective of the outcome).
    - ▶ When $\rho = 0$ then the village doesn't matter for the behavior of the people.
  - ▶ For the same number of units, having **more clusters** with fewer units per cluster can help.
  - ▶ Trade off spillover and power.
  - ▶ If you would not like an experiment with 10 units, then you should not be happy with an experiment with 10 clusters of 100 units. The effective sample size of this cluster randomized experiment is between 10 and $10 \times 100 = 1000$, but closer to 10 the higher the ρ.

You can combine blocks and clusters

► You can have clusters within blocks.

  ► Example: block = district, cluster = communities, units = individuals. You are measuring outcomes at the individual level.

  ► Example: block = province, cluster = district, units = communities. You are measuring outcomes at the community level.

► Note that you can't have blocks within clusters.

► For block and cluster randomization, you can use block_ra and cluster_ra in the randomizr package in R.

► For more complicated designs, you might find DeclareDesign helpful. (https://declaredesign.org)

Best practices

Best practices: replicability

▶ EGAP Methods Guide on Randomization
  (https://egap.org/resource/10-things-to-know-about-randomization/)

▶ Set a seed and save your code and random assignment column

▶ Verify

▶ Sometimes increased transparency $>$ replicability

Best practices: balance

▶ Check overall balance with a D-square test using `xBalance` in the `RItools` package (Hansen and Bowers (2008)) (large sample randomization inference):

```
xBalance(treatment ~ x1 + x2, data = dat, report = "chisquare")
```

▶ See also the `coin` package `independence_test` for permutation based version

▶ Use an F-test for a regression of treatment assignment on LHS and covariates on RHS (large sample approximate to randomization inference):

```
anova(lm(treatment ~ 1, data = dat),
  lm(treatment ~ x1 + x2 + x3, data = dat),
  test = "F"
)
```

Best practices: balance

▶ Random assignment gives us, in expectation, **overall balance** on the many covariates. It does not guarantee that all covariate to treatment relationships will be zero. In fact, in a small experiment, the magnitudes of imbalance may be high even if the randomization occurred perfectly.

▶ You will see t-tests of covariates one by one. Just by chance, you might get statistically significant differences on a variable. If you check balance on 100 variables, you will reject the null of no relationship in 5 of them even if there truly is no relationship.

Some Experimental Designs

Access

▶ Randomly select a treatment group through a lottery or equivalent mechanism, which randomizes **access** to the program

▶ Useful when you do not have enough resources to treat everyone

▶ Sometimes, some units (peoples, communities) must have access to a program.

   ▶ For example: a partner organization doesn't want to risk a vulnerable community NOT getting a program (want a guarantee that they will be always be treated).
   ▶ You can exclude those units from the experiment, and do random assignment among the remaining units that have a probability of assignment strictly between (and not including) 0 and 1.

Delayed Access (Phase-in or Wait list)

▶ Randomize *timing* of access to the program.

▶ When an intervention can be or must be rolled out in stages, you can randomize the order in which units are treated.

▶ Often you do not have the capacity to implement the treatment in a lot of places at once.

▶ Your control group are the as-yet untreated units.

▶ Be careful: the probability of assignment to treatment will vary over time because units that are assigned to treatment in earlier stages are not eligible to be assigned to treatment in later stages.

Factorial or Crossed-Assignment

- ▶ Factorial design enables testing of more than one treatment.

- ▶ You can analyze one treatment at a time.

- ▶ Or combinations thereof. Example:

	$X_1 = 0$	$X_1 = 1$
$X_2 = 1$	A	C
$X_2 = 0$	B	D

We might focus on an estimand like
$\mathbb{E}[Y(X_1 = 1, X_2 = 1)] - \mathbb{E}[Y(X_1 = 0, X_2 = 0)]$.

Encouragement

- Randomize **encouragement** to take the treatment, such as an invitation or subsidy to participate in a program.

- Useful when you cannot force a subject to participate.

- Estimands:

    - the ATE of the encouragement for your experimental sample.

    - the ATE of participation (not the encouragement) for the units who would participate when encouraged and wouldn't participate when not encouraged (compliers).

- Instrumental variables, exclusion restriction

Some Limits

Limits to randomized studies

- ▶ Ethics—is this sort of manipulation ethical? Sometimes not

- ▶ Must be done in real time, ahead of the intervention roll-out

- ▶ Reduced flexibility for a partner organization (problem for any prospective evaluation)

- ▶ Limits to the size of the experimental pool

- ▶ Cost

- ▶ The power constraint — you need a lot of units (problem for many statistical approaches)

- ▶ Violations of the key assumptions (spillovers; violation of the second key assumption for Causal Inference)

- ▶ External validity (problem for any evaluation and social science in general)

References

Hansen, Ben B., and Jake Bowers. 2008. "Covariate Balance in
    Simple, Stratified and Clustered Comparative Studies."
    Statistical Science 23 (2): 219–36.