Rerandomization in Randomized Experiments

STAT 320
Duke University
Kari Lock Morgan

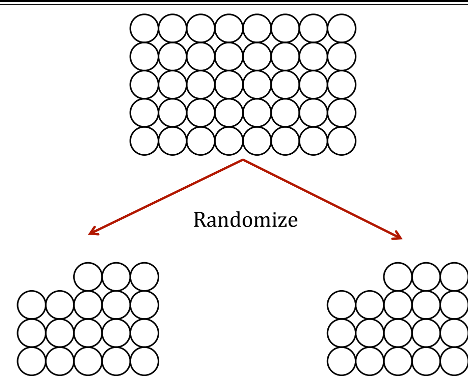
Comments on Homework 2

- When interpreting p-value:
  - chance of getting results as extreme IF NO DIFFERENCE
  - can't accept null even if insignificant
- Interpretation of a confidence interval
- Can't ever compute estimand (need theory to show unbiased)
- Standard error: standard deviation of statistics
- More possible allocations does not necessarily mean higher SE (think of increasing sample size: more allocations but lower SE)

The Gold Standard

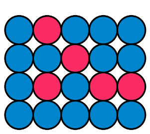
- Randomized experiments are the "gold standard" for estimating causal effects
- **WHY?**

1. They yield unbiased estimates
2. They balance covariates across treatment groups

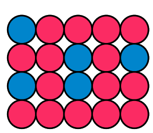


Covariate Balance - Gender

- What if you get a "bad" randomization?
- What would you do???
- Can you *rerandomize*? When? How?



5 Females, 15 Males



15 Females, 5 Males

The Origin of the Idea

Rubin: What if, in a randomized experiment, the chosen randomized allocation exhibited substantial imbalance on a prognostically important baseline covariate?

Cochran: Why didn't you block on that variable?

Rubin: Well, there were many baseline covariates, and the correct blocking wasn't obvious; and I was lazy at that time.

Cochran: This is a question that I once asked Fisher, and his reply was unequivocal:

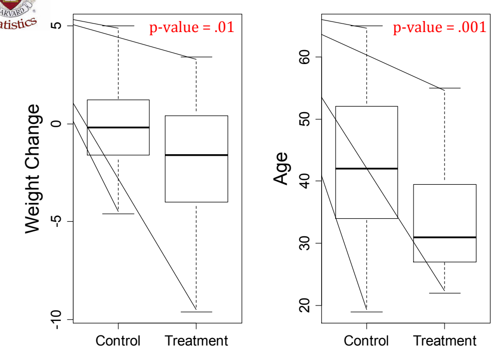
Fisher (recreated via Cochran): Of course, if the experiment had not been started, I would rerandomize.

Mind-Set Matters

- In 2007, Dr. Ellen Langer tested her hypothesis that "mind-set matters" with a randomized experiment
- She recruited 84 maids working at 7 different hotels, and randomly assigned half to a treatment group and half to control
- The "treatment" was simply informing the maids that their work satisfies the Surgeon General's recommendations for an active lifestyle

Crum, A.J. and Langer, E.J. (2007). "Mind-Set Matters: Exercise and the Placebo Effect," *Psychological Science*, **18**:165-171.

Mind-Set Matters



Covariate Imbalance

- The more covariates, the more likely at least one covariate will be imbalanced across treatment groups
- With just 10 independent covariates, the probability of a significant difference ($\alpha = .05$) for at least one covariate is $1 - (1-.05)^{10} = 40\%$!
- Covariate imbalance is not limited to rare "unlucky" randomizations

Covariate Imbalance

- Randomize 20 cards to two treatment groups
- Any differences?

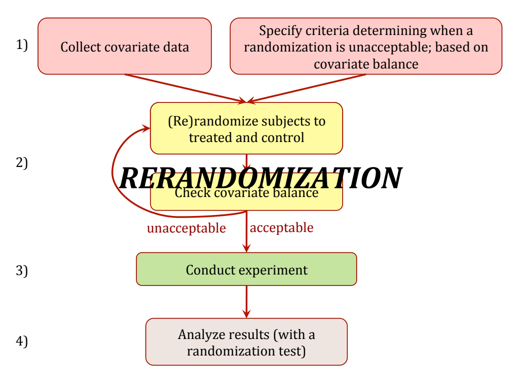
The Gold Standard

- Randomized experiments are the "gold standard" for estimating causal effects
- **WHY?**

1. They yield unbiased estimates
2. They eliminate confounding factors

… on average!

For any particular experiment, covariate imbalance is possible (and likely!), and conditional bias exists



Rerandomization Criterion

- Let **x** be the covariate matrix
- Let **W** be the vector of treatment assignments

$$W_i = \begin{cases} 1 & \text{if } i^{th} \text{ subject is treated} \\ 0 & \text{if } i^{th} \text{ subject is control} \end{cases}$$

- The criterion determining whether a randomization, **W**, is acceptable should be some function of **x** and **W**

Potential Outcomes

- Let $y_i(W_i)$ denote the i^{th} unit's potential outcome under treatment group W_i

$$\tau = \frac{\sum_{i=1}^{n} y_i(1)}{n} - \frac{\sum_{i=1}^{n} y_i(0)}{n}$$

$$\hat{\tau} = \frac{\sum_{i=1}^{n} W_i y_i(1)}{\sum_{i=1}^{n} W_i} - \frac{\sum_{i=1}^{n} (1-W_i) y_i(0)}{\sum_{i=1}^{n} (1-W_i)}$$

Unbiased

If the treated and control groups are the same size, and if the criteria for rerandomization treats the treated and control groups equally, then

$$E(\hat{\tau}) = \tau$$

Intuition: For every randomization, ***W***, thrown away, there is an exact opposite randomization, **1-*W***, that is also thrown away

Unbiased

If treated and control groups are not the same size, then rerandomization may not yield an unbiased estimate.

Example: Suppose you have one covariate *x*, and the criteria for rerandomization is $|\bar{X}_T - \bar{X}_C| < 1$

Units with more extreme *x* values will be more likely to be in the larger treatment group.

Rerandomization Test

- *Randomization Test*:
  - Simulate randomizations to see what the statistic would look like just by random chance, if the null hypothesis were true
- *Rerandomization Test*:
  - A randomization test, but for each simulated randomization, follow the same rerandomization criteria used in the experiment
- As long as the simulated randomizations are done using the same randomization scheme used in the experiment, this will give accurate p-values

Alternatives for Analysis

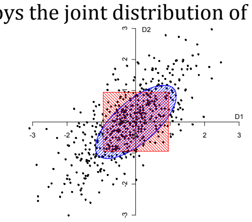
- **t-test:**
  - Too conservative
  - Significant results can be trusted
- **Regression**:
  - Regression including the covariates that were balanced on using rerandomization more accurately estimates the true precision of the estimated treatment effect
  - Assumptions are less dangerous after rerandomization because groups should be well balanced

Criteria for Acceptable Balance

If you have more than one covariate that you think may be associated with the outcome, what would you use as criteria for an acceptable randomization?

One Measure of Balance

- The obvious choice may be to set limits on the acceptable balance for each covariate individually
- This destroys the joint distribution of the covariates



Criteria for Acceptable Balance

We use *Mahalanobis Distance, M*, to represent multivariate distance between group means:

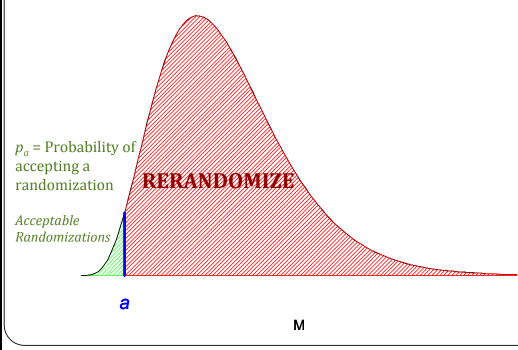
$$M = \left(\bar{\mathbf{X}}_T - \bar{\mathbf{X}}_C\right)' \mathrm{cov}\left(\bar{\mathbf{X}}_T - \bar{\mathbf{X}}_C\right)^{-1} \left(\bar{\mathbf{X}}_T - \bar{\mathbf{X}}_C\right)$$
$$= \left(\frac{1}{n_T} + \frac{1}{n_C}\right)^{-1} \left(\bar{\mathbf{X}}_T - \bar{\mathbf{X}}_C\right)' \mathrm{cov}\left(\mathbf{X}\right)^{-1} \left(\bar{\mathbf{X}}_T - \bar{\mathbf{X}}_C\right)$$

Under adequate sample sizes and pure randomization: $M \sim \chi^2_k$

k : Number of covariates to be balanced

Choose a and rerandomize when $M > a$

Distribution of M



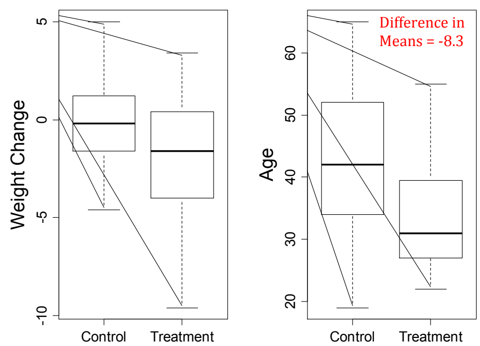
Choosing p_a

- Choosing the acceptance probability is a tradeoff between a desire for better balance and computational time
- The number of randomizations needed to get one successful one is Geometric(p_a), so the expected number needed is $1/p_a$
- Computational time must be considered in advance, since many simulated acceptable randomizations are needed to conduct the randomization test

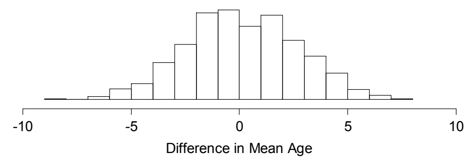
Rerandomization Based on *M*

- Since *M* follows a known distribution, easy to specify the proportion of accepted randomizations
- *M* is affinely invariant (unaffected by affine transformations of the covariates)
- Correlations between covariates are maintained
- The balance improvement for each covariate is the same (and known)…
- … and is the same for any linear combination of the covariates

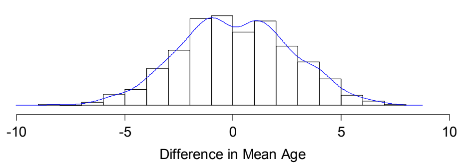
Mind-Set Matters



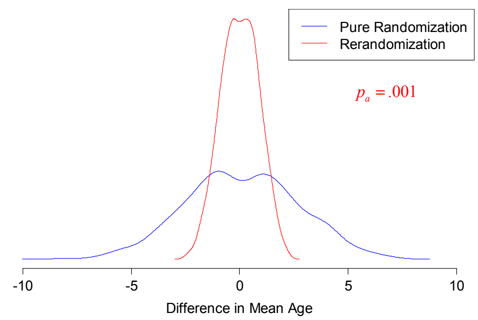
1000 Randomizations

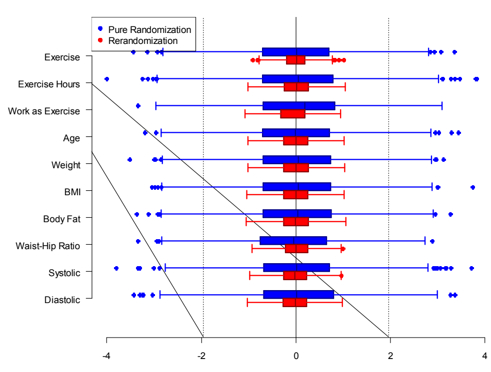


1000 Randomizations



1000 Randomizations





Covariates After Rerandomization

Theorem: If $n_T = n_C$, the covariate means are normally distributed, and rerandomization occurs when $M > a$, then

$$E\left(\overline{\mathbf{X}}_T - \overline{\mathbf{X}}_C \mid M \le a\right) = \mathbf{0}$$

and

$$\operatorname{cov}\left(\overline{\mathbf{X}}_T - \overline{\mathbf{X}}_C \mid M \le a\right) = v_a \operatorname{cov}\left(\overline{\mathbf{X}}_T - \overline{\mathbf{X}}_C\right).$$

$$v_a \equiv \frac{2}{k}\times\frac{\gamma\left(\frac{k}{2}+1,\frac{a}{2}\right)}{\gamma\left(\frac{k}{2},\frac{a}{2}\right)},$$

where γ is the incomplete gamma function: $\gamma(b,c) \equiv \int_0^c y^{b-1}e^{-y}\,dy$

Percent Reduction in Variance

- Define the *percent reduction in variance* for each covariate to be the percentage by which rerandomization reduces the randomization variance for the difference in means:

$$\frac{\operatorname{var}\left(\bar{X}_{j,T}-\bar{X}_{j,C} \mid \textit{rerandomization}\right)-\operatorname{var}\left(\bar{X}_{j,T}-\bar{X}_{j,C}\right)}{\operatorname{var}\left(\bar{X}_{j,T}-\bar{X}_{j,C}\right)}$$

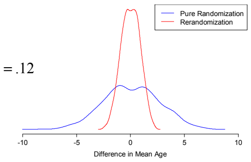
- For rerandomization when $M \leq a$, the percent reduction in variance for each covariate is $1-v_a$

Percent Reduction in Variance

$k=10,\ p_a=.001 \quad \Rightarrow a=1.48$

$$v_a=\frac{2}{k}\times\frac{\gamma\left(\frac{k}{2}+1,\frac{a}{2}\right)}{\gamma\left(\frac{k}{2},\frac{a}{2}\right)}=\frac{2}{10}\times\frac{\gamma\left(\frac{10}{2}+1,\frac{1.48}{2}\right)}{\gamma\left(\frac{10}{2},\frac{1.48}{2}\right)}=.12$$

Percent reduction in variance = 88%

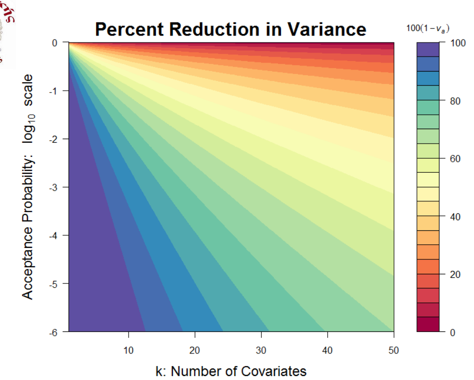


What if we increase p_a?

$k=10,\ p_a=.1 \quad \Rightarrow a=4.87$

$$v_a=\frac{2}{k}\times\frac{\gamma\left(\frac{k}{2}+1,\frac{a}{2}\right)}{\gamma\left(\frac{k}{2},\frac{a}{2}\right)}=\frac{2}{10}\times\frac{\gamma\left(\frac{10}{2}+1,\frac{4.87}{2}\right)}{\gamma\left(\frac{10}{2},\frac{4.87}{2}\right)}=.38$$

Percent reduction in variance = 62%



Estimated Treatment Effect

Theorem: If $n_T = n_C$, the covariate means are normally distributed, and rerandomization occurs when $M > a$, then

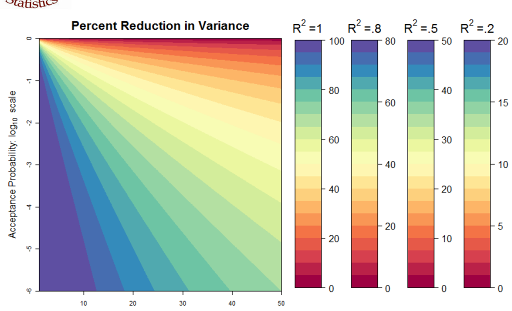
$$E\left(\bar{Y}_T-\bar{Y}_C \mid M \leq a\right)=0$$

and the percent reduction in variance for the outcome difference in means is

$$(1-v_a)R^2,$$

where R^2 is the coefficient of determination (squared canonical correlation).

Outcome Variance Reduction



Outcome Variance Reduction

$$k=10,\ p_a=.001 \quad \Rightarrow v_a=.12$$

Outcome: Weight Change
$R^2 = .1$
Percent Reduction in Variance = $(1 - v_a)R^2$ = (1 – .12)(.1) = 8.8%

Outcome: Change in Diastolic Blood Pressure
$R^2 = .64$
Percent Reduction in Variance = $(1 - v_a)R^2$ = (1 – .12)(.64) = 56%

Equivalent to increasing the sample size by 1/(1 - .56) = 2.27

Conclusion

• Rerandomization improves covariate balance between the treatment groups, giving the researcher more faith that an observed effect is really due to the treatment

• If the covariates are correlated with the outcome, rerandomization also increases precision in estimating the treatment effect, giving the researcher more power to detect a significant result