Integer Programming for Generalized Causal Bootstrap Designs

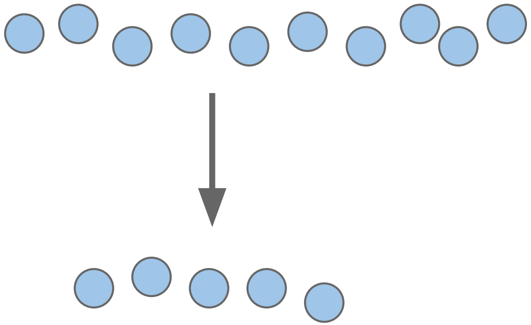
Jennifer Brennan[1], Sébastien Lahaie[1], Adel Javanmard[1,2], Nick Doudchenko[3], Jean Pouget-Abadie[1]

[1]Google Research, [2]University of Southern California, [3]Uber (work done while at Google Research)

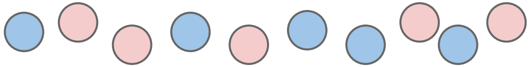
Google Research

ICML
International Conference On Machine Learning

Two sources of uncertainty in randomized experiments



Sampling Uncertainty: What if we had drawn a different set of units?



Design Uncertainty: What if we had assigned different units to treatment vs. control?

Claim: In randomized experiments, it is important to assess design uncertainty — and doing so may give tighter CI's than incorrectly applying the standard bootstrap.

Contributions

Previous work

A causal bootstrap procedure to estimate **design uncertainty** in the difference in means estimator under complete randomization.

This work

- **Integer programming** implementation of the causal bootstrap.
- Extensions to **linear-** and **quadratic-in-treatment estimators.**
- Extensions to **general randomized designs** where treatment probabilities and covariance are known.
- Application to **geographic experiments**.

Measuring design uncertainty

What we want
"If I were to repeat this experiment many times, what would the 5th and 95th percentiles on the test statistic be?"

i	$Y_i(0)$	$Y_i(1)$
1	9	10
2	10	11.5
3	8.5	10
4	12	13

i	$Y_i(0)$	$Y_i(1)$
1	?	10
2	?	11.5
3	8.5	?
4	12	?

$\hat{\tau} = -0.5$

i	$Y_i(0)$	$Y_i(1)$
1	?	10
2	10	?
3	?	10
4	12	?

$\hat{\tau} = -1$

i	$Y_i(0)$	$Y_i(1)$
1	9	?
2	?	11.5
3	?	10
4	12	?

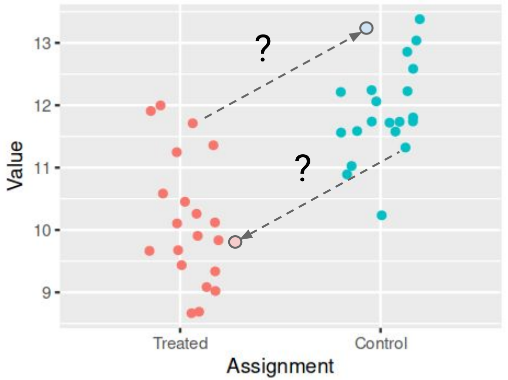
$\hat{\tau} = 0.25$

What we have
A single experimental observation

i	$Y_i(0)$	$Y_i(1)$
1	9	?
2	?	11.5
3	?	10
4	12	?

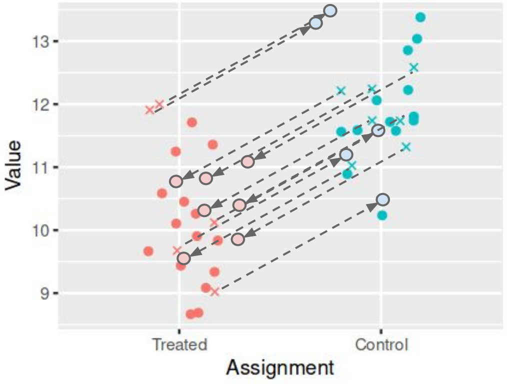
The Causal Bootstrap [Imbens-Menzel '21]

To generate one bootstrap replicate…



i	$Y_i(0)$	$Y_i(1)$
1	9	10
2	10	11.5
3	8.5	10
4	12	13

i	$Y_i(0)$	$Y_i(1)$
1	9	10
2	10	11.5
3	8.5	10
4	12	13



1) Impute the outcome of each unit under the unobserved condition

2) Draw a new randomization of the same units, using observed data if available and imputed data otherwise

Key Idea [Robins '88]: Impute via the joint distribution that maximizes the variance of your estimator, subject to matching the observed control and treatment marginal outcomes.

For diff-in-means + complete randomization, this is the assortative copula [Aronow et al '14].

Integer Programming Formulation

- Let $Z_i \in \{0, 1\}$ denote the binary treatment indicator for unit *i*.
- Let $X_{ik}^{(a)} \in \{0, 1\}$ be a binary optimization variable, indicating whether unit *i* has outcome *k* under treatment status *a*.
- $\mathbf{Q}$ is a symmetric positive definite matrix that depends on $\Pr(Z_i = 1)$ and $\mathrm{Cov}(Z_i, Z_j)$ for all pairs of units *i*, *j*.

$$
\begin{aligned}
&\max \quad \mathbf{X}^{\mathbf{T}}\mathbf{Q}\mathbf{X} && (4) \\
&\text{s.t.} \quad \mathbf{X} \in \{0,1\}^{N \cdot |\mathcal{Y}|} && \\
&\forall i, k, \quad X_{ik}^{(Z_i)} = 1 \text{ iff } Y_i^{obs} = y_k && \text{(a)} \\
&\forall a, i, k, \quad X_{ik}^{(a)} = 0 \text{ iff } y_k \notin \mathrm{supp}(F_a) && \text{(b)} \\
&\forall a, i, \quad \sum_{k=1}^{K} X_{ik}^{(a)} = 1 && \text{(c)} \\
&\forall a, k, b, \quad (-1)^b \sum_{i=1}^{N} X_{ik}^{(a)} \left(\frac{Z_i}{N_1} - \frac{1 - Z_i}{N_0} \right) \leq \epsilon && \text{(d)}
\end{aligned}
$$

- (4) → maximize estimator variance
- (a) → assign unit to observed outcome
- (b) → avoid outcomes not in support
- (c) → each unit is assigned to at most one outcome for *a* = 0,1
- (d) → marginals must stay the same

Simulation Result

We test our method on a **simulated geographical experiment** estimating the effect of an intervention on countries' GDP.

We consider two designs: **complete randomization** and **matched pairs**, and two estimators: **difference in means** and **model imputation**.

Our causal bootstrap achieves **almost nominal coverage**, with **up to 45% narrower confidence intervals** compared to the standard bootstrap.