Notation and Setup: The Finite Population Potential Outcomes Framework*

Thomas Leavitt

This guide collects, in one place, the notation and formal setup shared by the companion guides on (i) the unbiasedness of the Difference-in-Means estimator, (ii) the variance of the Difference-in-Means estimator and its conservative estimation, and (iii) the corresponding results for a superpopulation (the Population Average Treatment Effect). Each of those guides is written to be self-contained and recaps the pieces it needs, but this document develops the framework in full detail and serves as the canonical reference. A reader who is already comfortable with the potential outcomes framework for randomized experiments can safely skim it.

1 Units, assignments, and the assignment mechanism

Let the index $i \in \{1, \ldots, n\}$ run over n units in a finite sample, $\mathcal{S}_n$, where $n \geq 4$. Of these n units, $n_T \geq 2$ are assigned to the treatment condition and $n_C \geq 2$ are assigned to the control condition, where $n_T + n_C = n$. The bounds $n_T \geq 2$ and $n_C \geq 2$ (and hence $n \geq 4$) are not needed for the unbiasedness of the estimator or for the variance formula itself; they ensure that the conservative (Neyman) estimator of the variance (developed in a companion guide) is well defined. That estimator computes a separate sample variance within each treatment arm, dividing by $n_T - 1$ for the treated units and by $n_C - 1$ for the control units, and a sample variance requires at least two observations; we therefore need at least two units in each arm, i.e., $n_T \geq 2$ and $n_C \geq 2$, so that $n = n_T + n_C \geq 4$.

*This is a live document that is subject to updating at any time.

Let the binary indicator variable $Z_i \in \{0,1\}$ denote whether unit i is assigned to treatment ($Z_i = 1$) or control ($Z_i = 0$), and collect these indicators in the random assignment vector $\boldsymbol{Z} = \begin{bmatrix} Z_1, \ldots, Z_n \end{bmatrix}^\top$. The set of all logically possible assignment vectors is $\{0,1\}^n$. The assignment mechanism, however, places positive probability only on a subset of these vectors. Under *complete random assignment* with exactly n_T treated units, that subset is

$$\Omega = \left\{ \boldsymbol{z} \in \{0,1\}^n : \sum_{i=1}^{n} z_i = n_T \right\},$$

so that Ω is the support of $\boldsymbol{Z}$. Under complete random assignment, $\boldsymbol{Z}$ is distributed uniformly on Ω, and the number of elements in Ω, denoted $|\Omega|$,[1] is $\binom{n}{n_T}$. By contrast, under n independent Bernoulli assignments, there would be 2^n possible assignment vectors.

It helps to distinguish three objects that are easily conflated. The *assignment space* $\{0,1\}^n$ is the set of all conceivable assignments. The *support* $\Omega \subseteq \{0,1\}^n$ is the (typically much smaller) set of assignments the mechanism can actually produce. And the *randomization distribution* is the probability distribution the mechanism places on Ω (uniform, under complete random assignment). All of the randomness in what follows comes from this distribution.

The same framework accommodates *simple random assignment*, under which each of the n units is independently assigned to treatment, so that the number of treated units is itself random. In that case the support is $\Omega = \left\{ \boldsymbol{z} \in \{0,1\}^n : 0 < \sum_{i=1}^{n} z_i < n \right\}$, which has $2^n - 2$ elements (excluding the all-treated and all-control vectors, which would leave one group empty and the estimator undefined). Even when n_T is not fixed by design, one can fix it by conditioning on its observed value. The randomization distribution conditional on the realized n_T is exactly the complete random assignment distribution above. Hence results derived under complete random assignment carry over to simple random assignment after conditioning on the realized number of treated units.

2 The potential outcomes schedule and SUTVA

Adopting the terminology of Freedman (2009) and later Gerber and Green (2012), define a potential outcomes schedule as a vector-valued function $\boldsymbol{y} : \{0,1\}^n \to \mathbb{R}^n$ that maps each possible assignment vector $\boldsymbol{z} \in \{0,1\}^n$ to an n-dimensional vector of real-valued outcomes. More intuitively, a potential outcomes schedule lists how each of the n study participants would respond to every assignment $\boldsymbol{z}$ that the experiment could in principle produce. The ith entry of $\boldsymbol{y}(\boldsymbol{z})$ is the outcome that unit i would exhibit under assignment $\boldsymbol{z}$. The potential outcomes are *fixed* features of the units; they do

[1] For an arbitrary set W, let $|W|$ denote the cardinality of (i.e., the number of elements in) the set W.

not vary with the random assignment. Randomness enters only later, through which assignment $\boldsymbol{z}$ the mechanism happens to select.

Since the assignment space $\{0,1\}^n$ contains 2^n assignment vectors, the schedule specifies, in principle, 2^n outcome vectors. However, under the Stable Unit Treatment Value Assumption (SUTVA)[2] (Cox, 1958; Rubin, 1980, 1986), unit i's outcome depends only on its own assignment z_i and not on the assignments of the other units. Accordingly, let y_{Ti} denote the common outcome value of unit i across all assignments $\boldsymbol{z}$ with $z_i = 1$, and let y_{Ci} denote the common outcome value of unit i across all assignments $\boldsymbol{z}$ with $z_i = 0$. SUTVA thus collapses the entire schedule down to just two fixed numbers per unit, y_{Ti} and y_{Ci}. The individual causal effect for unit i on the additive scale is $\tau_i = y_{Ti} - y_{Ci}$. The vectors $\boldsymbol{y_T}$ and $\boldsymbol{y_C}$ collect the treatment and control potential outcomes, respectively, for all n units, and $\boldsymbol{\tau}$ collects the n individual, additive effects. These are all *fixed*, finite population quantities.

The observed outcome for unit $i \in \{1, \ldots, n\}$ is

$$Y_i = Z_i y_{Ti} + (1 - Z_i)\, y_{Ci},$$

which equals y_{Ti} when $Z_i = 1$ and y_{Ci} when $Z_i = 0$. Because y_{Ti} and y_{Ci} are fixed, the *only* source of randomness in the observed outcome Y_i is the assignment indicator Z_i. This distinction should be kept in mind throughout: Potential outcomes are fixed, the assignment is random, and observed outcomes are random only because a random assignment selects them from fixed potential outcomes.

3 Estimands and the Difference-in-Means estimator

It will be convenient to write the finite population means of the treatment and control potential outcomes as $\bar{y}_T \coloneqq n^{-1} \sum_{i=1}^n y_{Ti}$ and $\bar{y}_C \coloneqq n^{-1} \sum_{i=1}^n y_{Ci}$, both of which are fixed quantities. The target of interest is the Sample Average Treatment Effect (SATE),

$$\tau_{\text{SATE}} \coloneqq n^{-1} \sum_{i=1}^{n} \tau_i = n^{-1} \sum_{i=1}^{n} (y_{Ti} - y_{Ci}) = \bar{y}_T - \bar{y}_C,$$

[2]SUTVA implies that (1) units in the experiment respond to only the treatment condition to which each unit is individually assigned and (2) the treatment condition is actually the same treatment for all units assigned to treatment and the control condition is the same for all units assigned to control.

which is likewise a fixed (though unknown) quantity. The Difference-in-Means estimator of τ_{SATE} is

$$\hat{\tau} := \frac{\sum_{i=1}^{n} Z_i Y_i}{\sum_{i=1}^{n} Z_i} - \frac{\sum_{i=1}^{n} (1 - Z_i) Y_i}{\sum_{i=1}^{n} (1 - Z_i)}.$$

Unlike the estimand τ_{SATE}, the estimator $\hat{\tau}$ is a *random variable*, since it is a function of the random assignment vector $\boldsymbol{Z}$ and inherits all of its randomness from $\boldsymbol{Z}$. Under complete random assignment the denominators are fixed, $\sum_{i=1}^{n} Z_i = n_T$ and $\sum_{i=1}^{n} (1 - Z_i) = n_C$, so $\hat{\tau}$ can equivalently be written

$$\hat{\tau} = \frac{1}{n_T} \sum_{i=1}^{n} Z_i Y_i - \frac{1}{n_C} \sum_{i=1}^{n} (1 - Z_i) Y_i.$$

We use the ratio form when the number of treated units may be random (e.g., under simple random assignment) and the form with fixed denominators when n_T and n_C are fixed.

Because the only randomness is the assignment, we write $\mathrm{E}_{\Omega}[\cdot]$ and $\mathrm{Var}_{\Omega}[\cdot]$ for expectations and variances taken over the distribution of $\boldsymbol{Z}$ on Ω, to emphasize that they pertain to the randomization distribution alone.

4 Finite population variances of the potential outcomes

Two equivalent families of finite population variances appear in the variance results. The first uses the divisor $n - 1$:

$$S_n^2(\boldsymbol{y_T}) = \left(\frac{1}{n-1}\right) \sum_{i=1}^{n} (y_{Ti} - \bar{y}_T)^2$$
$$S_n^2(\boldsymbol{y_C}) = \left(\frac{1}{n-1}\right) \sum_{i=1}^{n} (y_{Ci} - \bar{y}_C)^2$$
$$S_n^2(\boldsymbol{\tau}) = \left(\frac{1}{n-1}\right) \sum_{i=1}^{n} (\tau_i - \tau_{\text{SATE}})^2.$$

The second uses the divisor n and is the form found in, e.g., Gerber and Green (2012, Equation 3.4):

$$\sigma_n^2(\boldsymbol{y_T}) = \left(\frac{1}{n}\right) \sum_{i=1}^{n} (y_{Ti} - \bar{y}_T)^2$$

$$\sigma_n^2(\boldsymbol{y_C}) = \left(\frac{1}{n}\right)\sum_{i=1}^{n}\left(y_{Ci} - \bar{y}_C\right)^2$$
$$\sigma_n(\boldsymbol{y_C}, \boldsymbol{y_T}) = \left(\frac{1}{n}\right)\sum_{i=1}^{n}\left(y_{Ci} - \bar{y}_C\right)\left(y_{Ti} - \bar{y}_T\right).$$

The two families are related by $\sigma_n^2(\cdot) = \frac{n-1}{n}S_n^2(\cdot)$, and likewise for the covariance term. All of these are fixed, finite population quantities. They will be the building blocks of the variance of $\hat{\tau}$.

5 Extension to a superpopulation: the Population Average Treatment Effect

For results about the Population Average Treatment Effect (PATE), we embed the finite sample in a superpopulation. Consider a superpopulation $\mathcal{P}_N$ of size $N \geq n \geq 4$, with index $i \in \{1, \ldots, N\}$, and let n units be drawn from $\mathcal{P}_N$ by simple random sampling into the experimental sample $\mathcal{S}_n$, while the remaining $N - n$ units are unsampled. Among the n sampled units, $n_T \geq 2$ are assigned to treatment and $n_C \geq 2$ to control by complete random assignment, exactly as above.

The binary indicator $R_i \in \{0, 1\}$ records whether unit i is included ($R_i = 1$) or excluded ($R_i = 0$) from the sample, and we collect these in $\boldsymbol{R} = \left[R_1, \ldots, R_N\right]^\top$, whose support is $\Pi = \left\{\boldsymbol{r} \in \{0,1\}^N : \sum_{i=1}^N r_i = n\right\}$ The estimand is the PATE,

$$\tau_{\text{PATE}} \coloneqq N^{-1}\sum_{i=1}^{N}\tau_i,$$

and the corresponding population variances (divisor N) are

$$\sigma_N^2(\boldsymbol{y_T}) = \left(\frac{1}{N}\right)\sum_{i=1}^{N}\left(y_{Ti} - \frac{1}{N}\sum_{i=1}^{N}y_{Ti}\right)^2$$
$$\sigma_N^2(\boldsymbol{y_C}) = \left(\frac{1}{N}\right)\sum_{i=1}^{N}\left(y_{Ci} - \frac{1}{N}\sum_{i=1}^{N}y_{Ci}\right)^2$$
$$\sigma_N^2(\boldsymbol{\tau}) = \left(\frac{1}{N}\right)\sum_{i=1}^{N}\left(\tau_i - \tau_{\text{PATE}}\right)^2.$$

In this setting the Difference-in-Means estimator has *two* sources of randomness, sampling $\{R_i\}_{i=1}^N$ and assignment $\{Z_i\}_{i=1}^n$. We write $\mathrm{E}_\Pi[\cdot]$ and $\mathrm{Var}_\Pi[\cdot]$ for expectations and variances over the sampling distribution, $\mathrm{E}_\Omega[\cdot]$ and $\mathrm{Var}_\Omega[\cdot]$ for those over the assignment distribution (conditional on the sample), and the unsubscripted $\mathrm{E}[\cdot]$ and $\mathrm{Var}[\cdot]$ for those over both. The two are connected by

the law of total variance,

$$\mathrm{Var}\left[\hat{\tau}\right] = \mathrm{E}_{\Pi}\left[\mathrm{Var}_{\Omega}\left[\hat{\tau}\right]\right] + \mathrm{Var}_{\Pi}\left[\mathrm{E}_{\Omega}\left[\hat{\tau}\right]\right],$$

which is the organizing identity for the PATE variance derivation.

References

Cox, D. R. (1958). *Planning of Experiments*. New York: John Wiley & Sons.

Freedman, D. A. (2009). *Statistical Models: Theory and Practice* (Revised ed.). Cambridge: Cambridge University Press.

Gerber, A. S. and D. P. Green (2012). *Field Experiments: Design, Analysis, and Interpretation*. New York: W. W. Norton.

Rubin, D. B. (1980). Comment on "randomization analysis of experimental data: The Fisher randomization test" by D. Basu. *Journal of the American Statistical Association 75*(371), 591–593.

Rubin, D. B. (1986). Comment: Which ifs have causal answers. *Journal of the American Statistical Association 81*(396), 961–962.