Segment 1: Fundamentals of Causal Inference

Section 08: “The Science Table”

Cut to Rubin 1978

The Annals of Statistics
1978, Vol. 6, No. 1, 34–58

BAYESIAN INFERENCE FOR CAUSAL EFFECTS: THE ROLE OF RANDOMIZATION

BY DONALD B. RUBIN

Educational Testing Service, Princeton, New Jersey

Causal effects are comparisons among values that would have been observed under all possible assignments of treatments to experimental units. In an experiment, one assignment of treatments is chosen and only the values under that assignment can be observed. Bayesian inference for causal effects follows from finding the predictive distribution of the values under the other assignments of treatments. This perspective makes clear the role of mechanisms that sample experimental units, assign treatments and record data. Unless these mechanisms are ignorable (known probabilistic functions of recorded values), the Bayesian must model them in the data analysis and, consequently, confront inferences for causal effects that are sensitive to the specification of the prior distribution of the data. Moreover, not all ignorable mechanisms can yield data from which inferences for causal effects are insensitive to prior specifications. Classical randomized designs stand out as especially appealing assignment mechanisms designed to make inference for causal effects straightforward by limiting the sensitivity of a valid Bayesian analysis.

The Potential Outcomes Framework

Three main components of the framework:

1. Formulate *potential outcomes* corresponding to various levels of a "treatment"
2. Formulate the *assignment mechanism* governing how treatments are assigned

The Potential Outcomes Framework

Three main components of the framework:

1. Formulate *potential outcomes* corresponding to various levels of a “treatment”
2. Formulate the *assignment mechanism* governing how treatments are assigned
3. Formulate a (Bayesian) *model* for “the science” represented by all of the potential outcomes and covariates
   - This last part is actually optional...

The Typical Setup

- Data: sample of $i = 1, 2, \ldots, n$ units from a target population
- A treatment with levels, $Z \in \{c, t\}$ or $Z \in \{0, 1\}$
  - More generally, with T levels
- For each i, two potential outcomes exist, (Y_i^c, Y_i^t)
- For each i, we observe
  - Z_i denoting treatment of i
  - X_i a vector of pre-treatment covariates
  - At most 1 of the potential outcomes is observed

$$Y_i^{obs} = Y_i(Z_i) = Z_i \cdot Y_i^t + (1 - Z_i) \cdot Y_i^c$$
$$Y_i^{mis} = Y_i(1 - Z_i) = (1 - Z_i) \cdot Y_i^t + Z_i \cdot Y_i^c$$

Key Quantities in the Science Table

- Pretreatment values, AKA, covariates, X
  - Variables that describe the characteristics of the experimental units
  - Measured *before* treatment
  - Not affected by treatment
- Treatment assignment vector, Z
  - Variable indicating which treatment is assigned to each unit
  - Potentially manipulable
  - Under experimenter control
- Potential Outcomes, AKA, counterfactuals
  - Dependent variables, outcomes, response, etc.
  - Values of posttreatment variables that would be potentially observed *after* assignment to treatment
  - Not possible to observe all of these!

Rubin 1978 and the "Science Table"

"The explicit representation of all potentially observable values leads to substantial notation, but once established, the notation permits important conclusions to be drawn almost immediately"

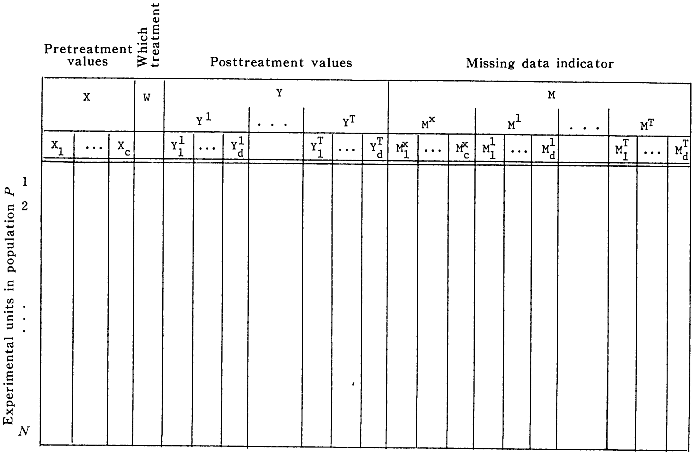

FIG. 1. All values in a study of T treatments.