Introduction to Designing Experiments

Gary W. Oehlert

School of Statistics
University of Minnesota

January 17, 2016

What do we mean by experiment?

An experiment applies treatments to experimental units and measures responses. We use the responses to learn about the treatments.

We want to learn about treatments. (dose of drug; amount of butter in a cookie; nano-tech coating for a fabric)

Responses tell us how the treatment worked. (patient better; tastiness of cookie; stain resistance)

Experimental units are the "guinea pigs" or test cases. (a patient; a batch of cookies; a bolt of fabric)

Experimenter assigns treatments to units.

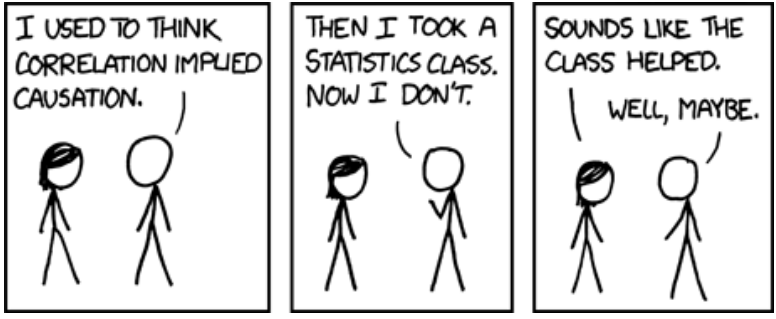
What makes an experiment special is control. The experimenter gets to control the assignment of treatments to units.

In contrast, an observational study has the same triple of treatment, unit, and response, but the experimenter observes the assignment of treatments to units.

For example, human health studies on smoking are observational.

Experiments can make causal inference. Observational studies find association (correlation).

Plenty of reasons to use observational studies too, including pre-existing data and ethical constraints on experimentation.


I USED TO THINK CORRELATION IMPLIED CAUSATION.
THEN I TOOK A STATISTICS CLASS. NOW I DON'T.
SOUNDS LIKE THE CLASS HELPED.
WELL, MAYBE.

Why designed experiments?

1. Design for direct comparison of treatments.
2. Design to reduce bias in comparisons (be accurate, avoid systematic errors).
3. Design to reduce variability (be precise).
4. Design to make variability estimable.
5. Design to widen validity.
6. Experiments support causal inference.

Some design goals may be in opposition, so compromise is often needed.

Some terminology

Treatments The procedures to learn about or compare.

Experimental units The things to which treatments are applied.

Responses Outcomes we measure to judge what the treatments do.

Randomization Using a known probabilistic mechanism in the assignment of treatments to units. (Also used in other ways.)

Experimental error Variation in response outcomes (modeled as random).

Control treatment A standard or baseline treatment.

Placebo A "null" treatment as a special type of control.

Blinding Evaluators of responses and/or human units do not know treatment assignments.

Factors Aspects of treatments that, when combined, form the complete treatment. Settings of factors are called *levels*.

Confounding When the effect of one factor (treatment, etc) cannot be distinguished from that of another. (Generally bad news.)

Measurement unit The part(s) of an e.u. actually used in measuring the response.

More on units

Experimental units vs measurement units. If a group of "units" **must** have the same treatment, they are likely m.u.s rather than e.u.s.

Physical size and shape of units. Edge often different from center.

More experimental units and fewer m.u.s per e.u. is generally better, but small e.u.s can also be a problem.

Need independence between e.u.s.

More on responses

Primary response.

Surrogate response.

Predictive response.

Audit response.