Principles of Experimental Design

Mark Conaway

October 1, 1999

Topics

- Basic principles applied to biomedical research.
- Examples of common experimental designs.

Focus on comparative experiments

- Def:Treatments can be allocated to the experimental units by the experimenter
  - A *treatment* is the diet, drug, device, delivery system, etc. that is:
    - under investigation and is under the control of the experimenter.
  - An *experimental unit* is the smallest division of the experimental material that can receive different treatments

Issues in entire process of design

- Which units?
- Which treatments?
- At what levels?
- Primary emphasis in statistical ED is on the question of how treatments should be allocated to units.

Requirements for a good experiment*

- Absence of systematic error
- Precision
- Range of validity
- Simplicity
- The calculation of uncertainty

*D.R. Cox, 1958, Planning of Experiments

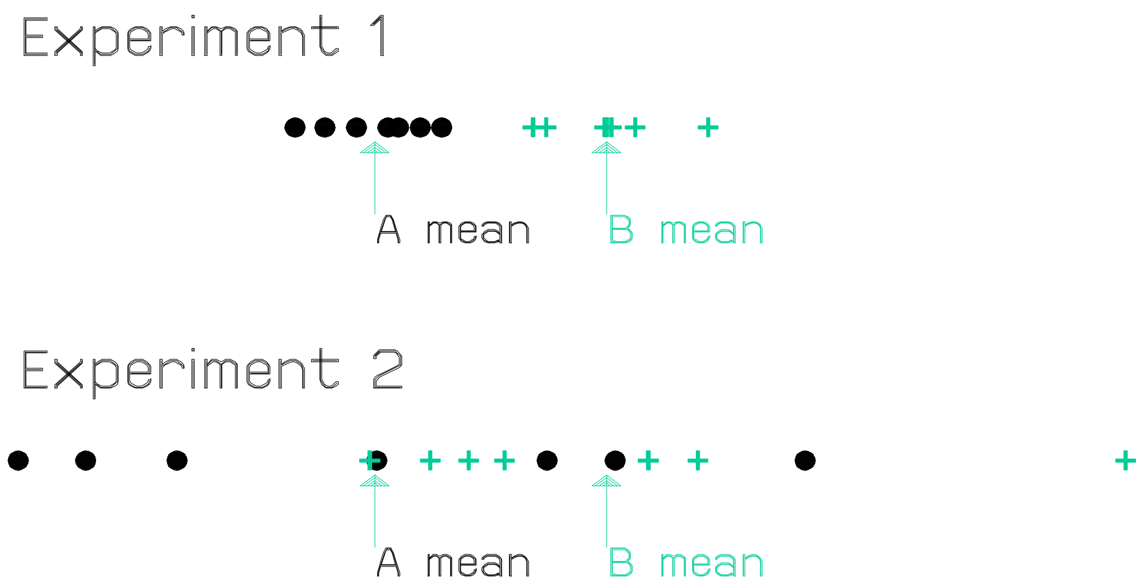
Requirements (cont'd)

- Absence of systematic error.
  - Why? Gives an unbiased estimate of effects of treatment
  - How? Compare equivalent groups under different treatments
    - Usually achieved through randomization

Requirements (con’td)

- Precision
  – If experiment has no systematic error, experimental results should differ from ``truth'' only by random variation
  – Would like to make amount of random variation as small as possible.

Precision:
Example. Which experiment is more convincing as to differences between A and B?



Requirements (cont'd)

- Range of validity:
  - Can the conclusions from the controlled conditions be applied in a larger context?
- Simplicity

Requirements (cont’d)

- Calculation of uncertainty
  - Do the results provide an assessment of the uncertainty associated with the estimated effects of treatments?

Ideal experiment

- Have units that
  - if treated with A, respond with r_A
  - if treated with B, respond with r_B
- Why is this ideal?
  - Gives an unbiased, precise (known) estimate of effects of treatment

Goal of statistical ED

- Get as close as possible to ideal experiment, given constraints:
  - Generalizability
  - Simplicity/Feasibility
  - Finite resources
  - Variable experimental material

General Rules

- Avoid systematic error:
  - Randomize
- Get precision:
  - Make treatment comparisons based on units that are as similar as possible.

General Rules

- Precision for comparing two treatments depends on:
  - Variation of units receiving same treatment
  - Number of units treated

Common Designs:
Completely Randomized Design

- Example:
- Assess effect of recombinant human growth hormone on recovery following bowel resection.
  - Treatments (diets): Chow, Standard, rHGH
  - Protocol: A total of 30 rats were randomized
  - Outcomes: Measure weight change,…,at day 8.

Liu, Liu and Jiang (1997) *Nutrition Research*, p. 303.

Check criteria for good experiment:

- Randomization
  - should provide protection against systematic biases in the treatment groups.
- Range of validity (?)
- Simple to implement.
- Precision and measuring uncertainty

ANOVA

- Results:

	Mean	SD
– Chow: 593, 587, 576, …	593	18.0
– Std : 525, 526, 540, …	533	16.5
– rHGH: 818, 785, 791, …	806	16.5

- Why do these outcomes vary?
  - Units received different treatments
  - Unknown/unexplained/natural variation

Common Designs: Completely Randomized Design with Factorial Structure

- Example: Effect of levels of dietary nitrogen, phosphorous on plasma Ca concentrations
- Treatments

Treatments	Nitro	Phos	n
1	Low	Low	6
2	High	Low	6
3	Low	High	6
4	High	High	6

Completely Randomized Design with Factorial Structure

- Factorial designs:
  - Treatments are combinations of factors.
  - Allow experimenter to answer questions about the *interaction* among factors.
    - Is the effect of increasing dietary nitrogen in a low-phosphorous diet equal to the effect of increasing dietary nitrogen in a high-phosphorous diet?
    - If answer is 'Yes', then there is no *statistical interaction*

Why is interaction important?

- With no interaction, estimate effect of nitrogen
  - Average of groups 2 and 4 (high nitro) minus average of groups 1 and 3 (low nitro)
  - Comparison based on 12 animals vs 12 animals
- Same precision as if you had devoted all 24 animals to a study of nitrogen effect

Why is interaction important?

- With no interaction, estimate effect of phosphorous
  - Average of groups 3 and 4 (high phos) minus average of groups 1 and 2 (low phos)
  - Comparison based on 12 animals vs 12 animals
- Same precision as if you had devoted all 24 animals to a study of phosphorous effect

Why is interaction important?

- With no interaction, factorial leads to a “2 for 1” efficiency:
- 24 animals gives same information as if had done two separate experiments of 24 animals each

What if there is interaction present?

- Example: Weight gains
- Trts Nitro Phos Avg Gain

Trts	Nitro	Phos	Avg Gain
1	Low	Low	35
2	High	Low	35
3	Low	High	36
4	High	High	42

- Adding nitrogen to diet low in phosphorous doesn’t make much difference

Advantage of Factorials

- Interaction present:
  - Best way to discover this is with a factorial
- Interaction absent:
  - 2 for 1 efficiency

Common design: randomized block designs

- Often called a “variance reduction design’’
- Group units into ‘’blocks’’such that units within blocks are relatively similar to each other

Common design: randomized block designs

- Example.
  - 10 animals assigned to each of three diets
  - Outcome: Weight change at day 8
- Line up 30 animals according to initial wgt

- A1,A2, A3, A4, A5, A6, …., A28, A29, A30
- Block1 Block 2 Block 10

Common design: randomized block designs

- Randomly assign diets to animals *within blocks*
  - Compares diets based on groups that are similar
  - Variability in units treated alike can be small

Common design: randomized block designs

- Can lead to highly efficient designs
  - E.g., a RB design that gives same precision of a CR design more than 2 times as large.