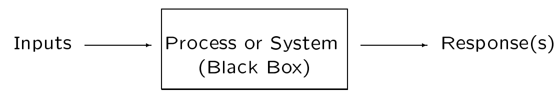
Introduction

Design of Experiments - Montgomery
Chapter 1

Overview of Experimental Design



- Have process quantified by specific response(s)
- Uncertainty in system response to fixed inputs
- Uncertainty due to nuisance factors/inherent noise
- Interested in studying process
  - Which inputs affect process response?
  - What level of inputs for specific response?
  - What input combination results in low uncertainty?

Design of experiments combines strategies of running an experiment with statistical tools for decision making

- Develop empirical **linear model** to explain process
- Plan experiment to obtain objective conclusions from model

Features of experiment to consider

- Statement of problem
- What response will be used?
- What change in response is important?
- What inputs to study?
- How many observations to be taken?
- What are resources? costs?
- Which inputs are most important?
- Are there uncontrollable nuisance factors?
- Block on controllable nuisance factors?
- What is experimental unit?

Design of Experiments

- Statement of the problem
  - What is the experiment intended to do?
  - Obvious question but often overlooked
  - Sound question goes long way towards solution
- Response(s) to be studied
  - Are variables measurable?
  - What sort of response is expected?
  - How accurately can response be measured?
- Inputs to be studied
  - What inputs may affect response?
  - What inputs are of interest?
  - Are factors to be held constant?
  - Varied at specific levels?

- Number of observations to be taken
  - How large a difference in response is important?
  - How much variation is present?
  - What costs and resources are available?
- Order of experiment
  - What is the timing of the experiment?
  - Is whole experiment to be randomized?
  - Are different factors randomized differently?

What is the experimental unit?

Experimental Unit - Material to which a treatment is applied in a single trial of the experiment

- Need to know EUs in order to do proper analysis
- May be different for different inputs

Example

3 fertilizers (A, B, C) and 3 seed varieties (**1**, 2̈, 3) thought to affect corn yield

C̈	Ä	A
B	B̈	**C**
A	C	**B**

C	Ä	A
B	B̈	C
A	C̈	B

- Experimental Variability
  - location in field / weather
  - differences in fertilizers or varieties
  - application of fertilizer or variety
  - measurement error
  - management of plot

Key Tools in Design/Analysis

- **Replication** - decrease uncertainty by averaging out experimental variability

  If Y_i has mean μ and variance σ^2 then E$(\overline{Y}) = \mu$ and Var$(\overline{Y}) = \sigma^2/n$

- **Blocking** - decrease uncertainty by adjusting for (controlling) specific nuisance factors
- **Randomization** - provides stronger basis for use of coincidence argument
  - Protection - averages out unknown factors
  - Independence of trials / Avoids biases
  - Randomization test $\longleftrightarrow$ Anova F-test

Randomization

- Random allocation of EUs to trts (or trts to EUs)
- All possible assignments are equally likely

Why not just try to be fair?

- Very difficult, can often bias results
- Subjective assignment inevitably avoids allocations
- One overcompensates and avoids specific patterns

Example: Consider assignment of three treatments to a field of nine regions (i.e., EUs). It is unknown that certain areas of the field are more fertile than others (hidden nuisance factor)

0	2	2
-2	0	2
-2	-2	0

Represent location effects in field

- Allocation
  - Fair: Randomly assign one rep per column/row
  - Random: All assignments equally likely

Possible Randomizations

Fair

C	B	A
B	A	C
A	C	B

A	C	B
B	A	C
C	B	A

Square 1: $\overline{A} = 0/3$ $\overline{B} = -2/3$ $\overline{C} = 2/3$
Square 2: $\overline{A} = 0/3$ $\overline{B} = 0/3$ $\overline{C} = 0/3$

All Possible

C	B	A
B	A	C
A	B	C

C	C	B
A	A	C
B	B	A

Square 1: $\overline{A} = 0/3$ $\overline{B} = -2/3$ $\overline{C} = 2/3$
Square 2: $\overline{A} = -2/3$ $\overline{B} = -2/3$ $\overline{C} = 4/3$

Assume there is **no difference** among trts (H_0 true). Still expect to reject Null hypothesis 100α% of the time.

- Fair
  - Can show F-statistic either 0.000 or 0.375 w/prob 50%
  - Depending on critical value, will reject 0 or 50% of time
  - Loss of sensitivity/specificity
- Random
  - Can show distribution of F-statistic is

F-stat	0.000	0.375	1.500	2.400	10.500	∞
Prob	.1286	.4179	.1929	.1929	.0643	.0036

  - Even with hidden trend, more like F-dist
  - Can set more reasonable α

Randomization/Permutation Test

Montgomery: page 73

H_0: Differences among trts only due to location

C	A	A
B	B	C
A	C	B

2	8	9
4	3	2
5	6	1

$\overline{A} = 22/3$ $\overline{B} = 8/3$ $\overline{C} = 10/3$

How unlikely is this based solely on chance?

$$\frac{9!}{3!3!3!} = 1680 \text{ orderings}$$

all equally likely

- Compute F statistic for each ordering (i.e., reference dist)
- Compare observed result with distribution
- Test does not rely on an underlying distribution (e.g., Normal)
- Uses observed data to construct reference distribution
- Reference distribution closely represented by usual F dist
- Gives justification for usual ANOVA approach

For example, randomization may result in

1

C	A	A
B	B	C
A	C	B

2

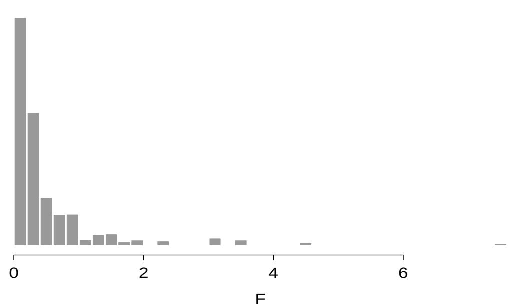
C	A	C
B	C	A
B	A	B

Use F test statistic (Analysis of Variance)

Box 1: F=1.59 Box 2: F=0.11

Randomly generate 5000 possible orderings

$\widehat{\Pr}(F \geq 1.59) = .075$ $\widehat{\Pr}(F \geq 0.11) = .735$



Histogram of 5000 random orderings

Summary

- **Don't forget non statistical knowledge**: Statistical techniques are most effective when combined with problem specific knowledge. Ask questions to discover as much about the problem as possible.
- **Keep the design and analysis simple**: Can often answer questions with sound straightforward approach. Complex designs more sensitive to problems.
- **Practical vs statistical significance**: Need to initially consider what is an "important" difference. Helps determine appropriate sample size. A statistical difference may not be anything of value.
- **Experiments often iterative**: Often little knowledge of problem and variability. Pilot studies can be done to obtain information and/or used to ensure experiment can be run as planned. Additional experiments may focus on new levels of important factors or include a new factor.
- **Randomization**: Provides justification for usual F-test analysis. Helps avoid unintentional subjective biases in assignments.

Examples of Experiments

For each of the experiments below, determine the various treatment factors and experimental units. Also describe differences in how the experiment was randomized.

Exp 1 To study the effects of pesticides on birds, an experimenter randomly (and equally) allocated sixty-five chicks to five diets (a control and four with a different pesticide included). After a month each chick's calcium content (mg) in one cm length of bone was measured.

Exp 2 A psychologist is interested in studying the IQs of 1st grade children from the low income areas of several major cities. Six grade schools were randomly chosen (from the low income areas) and from each of these schools, five 1st grade children were randomly chosen and had their IQs measured.

Exp 3 Brewer's malt is produced from germinating barley. The following is an experiment to determine the best conditions to germinate the barley. A total of thirty lots of barley seeds (100 seeds per lot) were equally and randomly assigned to ten germination conditions. The conditions are combinations of the week after harvest (1, 3, 6, 9, or 12 weeks) and the amount of water used in the process (4 ml or 8 ml).

Exp 4. Winter treatments to clear ice and snow can damage roads. An experiment was conducted comparing four treatments, each consisting of different combinations of salt and sand. Because traffic level also damages the roads, four roads were selected for the study and each treatment was randomly assigned to a portion of each road.

Exp 5. A researcher is interested in assessing a new fitness regimen. Thirty subjects were randomly selected to participate with fifteen each assigned to the control or treatment group. Prior to the regimen, a pre-test of the subject's fitness was performed. Blood measurements were taken 1, 5, 10, 30, and 60 minutes into this fitness test. After the six week treatment program, a similar post-test of fitness was performed with blood measurements again taken at the same five time points.