Introduction to Experimental Design

Raj Jain
Washington University in Saint Louis
Saint Louis, MO 63130
Jain@cse.wustl.edu

These slides are available on-line at:
http://www.cse.wustl.edu/~jain/cse567-06/
Overview

- What is experimental design?
- Terminology
- Common mistakes
- Sample designs
Experimental Design and Analysis

How to:

- Design a proper set of experiments for measurement or simulation.
- Develop a model that best describes the data obtained.
- Estimate the contribution of each alternative to the performance.
- Isolate the measurement errors.
- Estimate confidence intervals for model parameters.
- Check if the alternatives are significantly different.
- Check if the model is adequate.
Example

Personal workstation design

1. Processor: 68000, Z80, or 8086.
2. Memory size: 512K, 2M, or 8M bytes
3. Number of Disks: One, two, three, or four
4. Workload: Secretarial, managerial, or scientific.
5. User education: High school, college, or post-graduate level.

Five **Factors** at 3x3x4x3x3 **levels**
Cartoon

"If this doesn't help you don't worry, it's a placebo."
Terminology

- **Response Variable**: Outcome.
  E.g., throughput, response time

- **Factors**: Variables that affect the response variable.
  E.g., CPU type, memory size, number of disk drives, workload used, and user's educational level.
  Also called predictor variables or predictors.

- **Levels**: The values that a factor can assume, E.g., the CPU type has three levels: 68000, 8080, or Z80.
  # of disk drives has four levels.
  Also called *treatment*.

- **Primary Factors**: The factors whose effects need to be quantified.
  E.g., CPU type, memory size only, and number of disk drives.
Terminology (Cont)

- **Secondary Factors**: Factors whose impact need not be quantified.
  E.g., the workloads.

- **Replication**: Repetition of all or some experiments.

- **Design**: The number of experiments, the factor level and number of replications for each experiment.
  E.g., Full Factorial Design with 5 replications: $3 \times 3 \times 4 \times 3 \times 3$ or 324 experiments, each repeated five times.

- **Experimental Unit**: Any entity that is used for experiments.
  E.g., users. Generally, no interest in comparing the units.

- Goal - minimize the impact of variation among the units.
Terminology (Cont)

- **Interaction** ⇒ Effect of one factor depends upon the level of the other.

Table 1: Noninteracting Factors

<table>
<thead>
<tr>
<th></th>
<th>$A_1$</th>
<th>$A_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B_1$</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>$B_2$</td>
<td>6</td>
<td>8</td>
</tr>
</tbody>
</table>

Table 2: Interacting Factors

<table>
<thead>
<tr>
<th></th>
<th>$A_1$</th>
<th>$A_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B_1$</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>$B_2$</td>
<td>6</td>
<td>9</td>
</tr>
</tbody>
</table>
Common Mistakes in Experimentation

- The variation due to experimental error is ignored.
- Important parameters are not controlled.
- Effects of different factors are not isolated
- Simple one-factor-at-a-time designs are used
- Interactions are ignored
- Too many experiments are conducted.
  
  Better: two phases.
Types of Experimental Designs

- **Simple Designs**: Vary one factor at a time
  \[
  \# \text{ of Experiments} = 1 + \sum_{i=1}^{k} (n_i - 1)
  \]
  - Not statistically efficient.
  - Wrong conclusions if the factors have interaction.
  - Not recommended.

- **Full Factorial Design**: All combinations.
  \[
  \# \text{ of Experiments} = \prod_{i=1}^{k} n_i
  \]
  - Can find the effect of all factors.
  - Too much time and money.
  - May try $2^k$ design first.
Types of Experimental Designs (Cont)

- Fractional Factorial Designs: Less than Full Factorial
  - Save time and expense.
  - Less information.
  - May not get all interactions.
  - Not a problem if negligible interactions.
A Sample Fractional Factorial Design

- Workstation Design:
  (3 CPUs)(3 Memory levels)(3 workloads)(3 ed levels)
  = 81 experiments

<table>
<thead>
<tr>
<th>Experiment Number</th>
<th>CPU</th>
<th>Memory Level</th>
<th>Workload Type</th>
<th>Educational Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>68000</td>
<td>512K</td>
<td>Managerial</td>
<td>High School</td>
</tr>
<tr>
<td>2</td>
<td>68000</td>
<td>2M</td>
<td>Scientific</td>
<td>Post-graduate</td>
</tr>
<tr>
<td>3</td>
<td>68000</td>
<td>8M</td>
<td>Secretarial</td>
<td>College</td>
</tr>
<tr>
<td>4</td>
<td>Z80</td>
<td>512K</td>
<td>Scientific</td>
<td>College</td>
</tr>
<tr>
<td>5</td>
<td>Z80</td>
<td>2M</td>
<td>Secretarial</td>
<td>High School</td>
</tr>
<tr>
<td>6</td>
<td>Z80</td>
<td>8M</td>
<td>Managerial</td>
<td>Post-graduate</td>
</tr>
<tr>
<td>7</td>
<td>8086</td>
<td>512K</td>
<td>Secretarial</td>
<td>Post-graduate</td>
</tr>
<tr>
<td>8</td>
<td>8086</td>
<td>2M</td>
<td>Managerial</td>
<td>College</td>
</tr>
<tr>
<td>9</td>
<td>8086</td>
<td>8M</td>
<td>Scientific</td>
<td>High School</td>
</tr>
</tbody>
</table>
Summary

- Goal of proper experimental design is to get the maximum information with minimum number of experiments
- Factors, levels, full-factorial designs
Exercise 16.1

The performance of a system being designed depends upon the following three factors:

- CPU type: 68000, 8086, 80286
- Operating System type: CPM, MS-DOS, UNIX
- Disk drive type: A, B, C

a. There is significant interaction among factors.
b. There is no interaction among factors.
c. The interactions are small compared to main effects.