Introduction to Experimental Design

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.
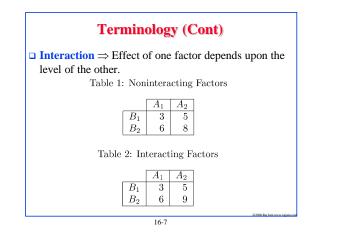
Five **Factors** at 3x3x4x3x3 **levels**
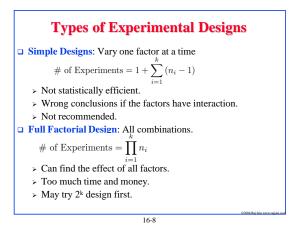

Terminology

- **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.
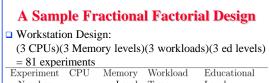- Goal - minimize the impact of variation among the units.

Terminology (Cont)

❑ **Interaction** $\Rightarrow$ Effect of one factor depends upon the level of the other.

Table 1: Noninteracting Factors

	A_1	A_2
B_1	3	5
B_2	6	8

Table 2: Interacting Factors

	A_1	A_2
B_1	3	5
B_2	6	9

16-7

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.

16-8

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

16-9

A Sample Fractional Factorial Design

❑ Workstation Design:

(3 CPUs)(3 Memory levels)(3 workloads)(3 ed levels) = 81 experiments

Experiment Number	CPU	Memory Level	Workload Type	Educational Level
1	68000	512K	Managerial	High School
2	68000	2M	Scientific	Post-graduate
3	68000	8M	Secretarial	College
4	Z80	512K	Scientific	College
5	Z80	2M	Secretarial	High School
6	Z80	8M	Managerial	Post-graduate
7	8086	512K	Secretarial	Post-graduate
8	8086	2M	Managerial	College
9	8086	8M	Scientific	High School

16-10

Summary

❑ Goal of proper experimental design is to get the maximum information with minimum number of experiments

❑ Factors, levels, full-factorial designs

16-11

2