Co-funded by the Erasmus+ Programme of the European Union

DeCAIR

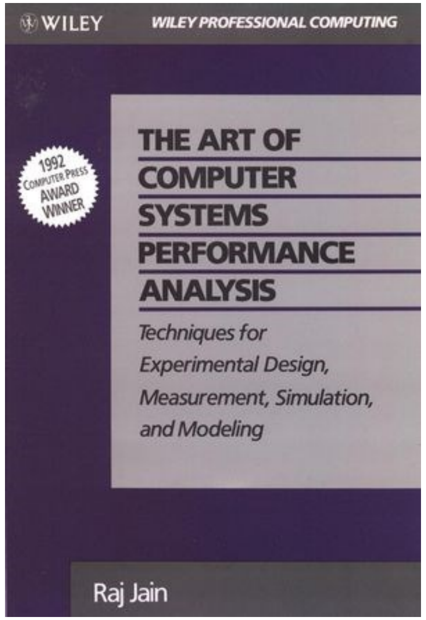
Experimental Design

Prof. Gheith Abandah

Developing Curricula for Artificial Intelligence and Robotics (DeCAIR)
618535-EPP-1-2020-1-JO-EPPKA2-CBHE-JP

References

- Raj Jain, **The Art of Computer Systems Performance Analysis**, Wiley, 1991.
  - Part I: An Overview of Performance Evaluation
  - Part II: Measurement Techniques and Tools
  - Part III: Probability Theory and Statistics
  - Part IV: Experimental Design and Analysis
  - Part V: Simulation



Outline

- Introduction and Terminology
- Types of Experimental Design
- 2^k Factorial Designs and 2^2 Example
- General 2^k Factorial Designs
- Other Designs

Introduction and Terminology

- **Objectives**
  - 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.
- **Response Variable**: **Outcome**, *e.g.*, throughput, response time.
- **Factors**: Variables that affect the response variable, *e.g.*, CPU type, memory size, workload used. Also called **predictor variables** or **predictors**.
- **Levels**: The values that a factor can assume, *e.g.*, the CPU type has three levels: i3, i5, or i7. Also called **treatments**.

Introduction and Terminology

- **Primary Factors**: The factors whose effects need to be quantified, *e.g.*, CPU type and memory size only.
- **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 levels, and the number of replications for each experiment.

Introduction and Terminology

- **Interaction**: When the 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

Outline

- Introduction and Terminology
- Types of Experimental Design
- 2^k Factorial Designs and 2^2 Example
- General 2^k Factorial Designs
- Other Designs

Types of Experimental Designs

1. **Simple Designs**: Vary one factor at a time.
   - Not statistically efficient.
   - Wrong conclusions if the factors have interactions.
   - Not recommended.

$$\#\text{ of Experiments} = 1 + \sum_{i=1}^{k} (n_i - 1)$$

2. **Full Factorial Design**: All combinations.
   - Can find the effect of all factors.
   - Too much time and money.
   - May try 2^k design first.

$$\#\text{ of Experiments} = \prod_{i=1}^{k} n_i$$

Types of Experimental Designs

- **Fractional Factorial Designs**: Less than Full Factorial.
  - Saves time and expense.
  - Less information.
  - May not get all interactions.
  - Not a problem if negligible interactions.
- **Example**: Workstation Design (3 CPUs)(3 Memory levels) (3 workloads)(3 levels) = 81 experiments!

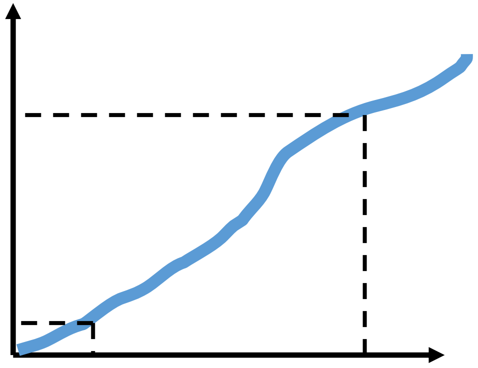
Experiment Number	CPU	Memory Level	Workload Type	Educational Level
1	68000	512K	Managerial	High School
2	68000	2M	Scientific	Post-graduate
3	68000	8M	Secretarial	College
4	Z80	512K	Scientific	College
5	Z80	2M	Secretarial	High School
6	Z80	8M	Managerial	Post-graduate
7	8086	512K	Secretarial	Post-graduate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

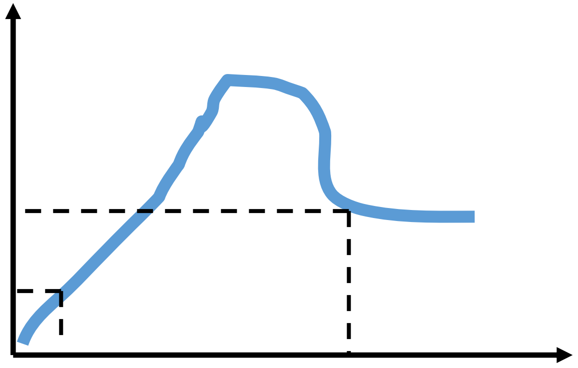
Outline

- Introduction and Terminology
- Types of Experimental Design
- 2^k Factorial Designs and 2^2 Example
- General 2^k Factorial Designs
- Other Designs

2^k Factorial Designs and 2^2 Example

- ***k* factors**, **each at two levels**.
- Easy to analyze.
- Helps in sorting out the impact of factors (**sensitivity analysis**).
- Good at the beginning of a study.
- Valid only if the effect is **unidirectional**, *e.g.*, memory size, the number of disk drives.





2^2 Factorial Designs

- **Two factors**, each at two levels.

Performance in MIPS

Cache Size	Memory Size	
	4M Bytes	16M Bytes
1K	15	45
2K	25	75

$$x_A = \begin{cases} -1 & \text{if 4M bytes memory} \\ 1 & \text{if 16M bytes memory} \end{cases}$$

$$x_B = \begin{cases} -1 & \text{if 1K bytes cache} \\ 1 & \text{if 2K bytes cache} \end{cases}$$

Model

$$y = q_0 + q_A x_A + q_B x_B + q_{AB} x_A x_B$$

Observations:

$$15 = q_0 - q_A - q_B + q_{AB}$$

$$45 = q_0 + q_A - q_B - q_{AB}$$

$$25 = q_0 - q_A + q_B - q_{AB}$$

$$75 = q_0 + q_A + q_B + q_{AB}$$

Solution:

$$y = 40 + 20x_A + 10x_B + 5x_A x_B$$

- **Interpretation**:
  - Mean performance = 40 MIPS
  - Effect of memory = 20 MIPS
  - Effect of cache = 10 MIPS
  - Interaction between memory and cache = 5 MIPS

Computation of Effects

Experiment	A	B	y
1	-1	-1	y_1
2	1	-1	y_2
3	-1	1	y_3
4	1	1	y_4

$$y = q_0 + q_A x_A + q_B x_B + q_{AB} x_A x_B$$

$$y_1 = q_0 - q_A - q_B + q_{AB}$$

$$y_2 = q_0 + q_A - q_B - q_{AB}$$

$$y_3 = q_0 - q_A + q_B - q_{AB}$$

$$y_4 = q_0 + q_A + q_B + q_{AB}$$

- **Solution**

$$q_0 = \frac{1}{4}(y_1 + y_2 + y_3 + y_4)$$

$$q_A = \frac{1}{4}(-y_1 + y_2 - y_3 + y_4)$$

$$q_B = \frac{1}{4}(-y_1 - y_2 + y_3 + y_4)$$

$$q_{AB} = \frac{1}{4}(y_1 - y_2 - y_3 + y_4)$$

Sign Table Method

I	A	B	AB	y	q_0	q_A	q_B	q_{AB}
1	-1	-1	1	15	15	-15	-15	15
1	1	-1	-1	45	45	45	-45	-45
1	-1	1	-1	25	25	-25	25	-25
1	1	1	1	75	75	75	75	75
				Total	160	80	40	20
				Total/4	**40**	**20**	**10**	**5**

Allocation of Variation

- **Importance of a factor** = proportion of the *variation* explained by this factor.

$$\text{Sample Variance of } y = s_y^2 = \frac{\sum_{i=1}^{2^2}(y_i - \bar{y})^2}{2^2 - 1}$$

$$\text{Total Variation of } y = \text{SST} = \sum_{i=1}^{2^2}(y_i - \bar{y})^2$$

- **For a 2^2 design**: $\text{SST} = 2^2 q_A^2 + 2^2 q_B^2 + 2^2 q_{AB}^2 = \text{SSA} + \text{SSB} + \text{SSAB}$
- Variation due to A = SSA = $2^2 q_A^2$
- Variation due to B = SSB = $2^2 q_B^2$
- Variation due to interaction = SSAB = $2^2 q_{AB}^2$
- **Fraction explained by A** = $\frac{SSA}{SST}$ Variation $\neq$ Variance

Example

- **Memory-cache study**:

$$\bar{y} = \frac{1}{4}(15 + 55 + 25 + 75) = 40$$

$$\begin{aligned} \text{Total Variation} &= \sum_{i=1}^{4} (y_i - \bar{y})^2 \\ &= (25^2 + 15^2 + 15^2 + 35^2) \\ &= 2100 \\ &= 4 \times 20^2 + 4 \times 10^2 + 4 \times 5^2 \end{aligned}$$

- Total variation (SST) = 2100

  Variation due to Memory (SSA) = 1600 (76%)

  Variation due to cache (SSB) = 400 (19%)

  Variation due to interaction (SSAB) = 100 (5%)

Sign Table Method

y	q_0	q_A	q_B	q_{AB}	
15	15	-15	-15	15	
45	45	45	-45	-45	
25	25	-25	25	-25	
75	75	75	75	75	
Total	160	80	40	20	
Total/4	**40**	**20**	**10**	**5**	
q^2		400	100	25	
$4q^2$		1600	400	100	2100
$4q^2$/SST		76%	19%	5%	100%
		SSA	SSB	SSAB	SST

Outline

- Introduction and Terminology
- Types of Experimental Design
- 2^k Factorial Designs and 2^2 Example
- General 2^k Factorial Designs
- Other Designs

General 2^k Factorial Designs

- ***k* factors** at **two levels each**.

  2^k experiments

  2^k effects:

  $$\text{k main effects}$$
  $$\binom{k}{2} \text{two factor interactions}$$
  $$\binom{k}{3} \text{three factor interactions...}$$

$$\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k(k-1)\cdots 1}$$

2^3 Design Example

- **Three factors** in designing a machine:
  - Cache size
  - Memory size
  - Number of processors

	Factor	Level -1	Level 1
A	Memory Size	4MB	16MB
B	Cache Size	1kB	2kB
C	Number of Processors	1	2

2^3 Design Example (cont.)

Cache Size	4M Bytes		16M Bytes	
	1 Proc	2 Proc	1 Proc	2 Proc
1K Byte	14	46	22	58
2K Byte	10	50	34	86

I	A	B	C	AB	AC	BC	ABC	y
1	-1	-1	-1	1	1	1	-1	14
1	1	-1	-1	-1	-1	1	1	22
1	-1	1	-1	-1	1	-1	1	10
1	1	1	-1	1	-1	-1	-1	34
1	-1	-1	1	1	-1	-1	1	46
1	1	-1	1	-1	1	-1	-1	58
1	-1	1	1	-1	-1	1	-1	50
1	1	1	1	1	1	1	1	86
320	80	40	160	40	16	24	9	Total
40	10	5	20	5	2	3	1	Total/8

2^3 Design Example (cont.)

$$\begin{aligned}
\text{SST} &= 2^3(q_A^2 + q_B^2 + q_C^2 + q_{AB}^2 + q_{AC}^2 + q_{BC}^2 + q_{ABC}^2) \\
&= 8(10^2 + 5^2 + 20^2 + 5^2 + 2^2 + 3^2 + 1^2) \\
&= 800 + 200 + 3200 + 200 + 32 + 72 + 8 = 4512 \\
&= 18\% + 4\% + 71\% + 4\% + 1\% + 2\% + 0\% \\
&= 100\%
\end{aligned}$$

- Number of Processors (C) is the most important factor.

2^4 Design Example

I	A	B	C	D											y

2^4 Design Example

I	A	B	C	D	AB	AC	AD	BC	BD	CD	ABC	ABD	BCD	ABCD	y
1	-1	-1	-1	-1	1	1	1	1	1	1	-1	-1	-1	1	y1
1	1	-1	-1	-1	-1	-1	-1	1	1	1	1	1	-1	-1	y2
1	-1	1	-1	-1	-1	1	1	-1	-1	1	1	1	1	-1	y3
1	1	1	-1	-1	1	-1	-1	-1	-1	1	-1	-1	1	1	y4
1	-1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1	y5
1	1	-1	1	-1	-1	1	-1	-1	1	-1	-1	1	1	1	y6
1	-1	1	1	-1	-1	-1	1	1	-1	-1	-1	1	-1	1	y7
1	1	1	1	-1	1	1	-1	1	-1	-1	1	-1	-1	-1	y8
1	-1	-1	-1	1	1	1	-1	1	-1	-1	-1	1	1	-1	y9
1	1	-1	-1	1	-1	-1	1	1	-1	-1	1	-1	1	1	y10
1	-1	1	-1	1	-1	1	-1	-1	1	-1	1	-1	-1	1	y11
1	1	1	-1	1	1	-1	1	-1	1	-1	-1	1	-1	-1	y12
1	-1	-1	1	1	1	-1	-1	-1	-1	1	1	1	-1	1	y13
1	1	-1	1	1	-1	1	1	-1	-1	1	-1	-1	-1	-1	y14
1	-1	1	1	1	-1	-1	-1	1	1	1	-1	-1	1	-1	y15
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	y16

Outline

- Introduction and Terminology
- Types of Experimental Design
- 2^k Factorial Designs and 2^2 Example
- General 2^k Factorial Designs
- Other Designs

Other Designs

- **$2^k r$ Factorial Designs with Replications**: Good for estimating the experimental errors.
- **2^{k-p} Fractional Factorial Designs**: Requires fewer experiments, but with **confounding**: Some of the effects and interactions cannot be separated.
- **One-Factor Experiment**: Used to compare several alternatives of a single categorical variable.
- **Two-factor Full Factorial Design Without Replications**
- **Two-factor Full Factorial Design with Replications**
- **General Full Factorial Designs with *k* Factors**

Summary

- Introduction and Terminology
- Types of Experimental Design
- 2^k Factorial Designs and 2^2 Example
- General 2^k Factorial Designs
- Other Designs