Introduction to Experimental Design

CS 700

1

Acknowledgement

Some of the following slides are based on presentations created and copyrighted by Prof. Daniel Menasce (GMU)

2

Basic Notions in Design of Experiments

- Response: what you want to measure.
- Factor: what affects the response.
- Level: value of a factor.

	Factors			Response
	CPU Clock Frequency (MHz)	Number of CPUs	Main Memory (MB)	Benchmark Execution Time (sec)
Levels	550	1	128	25
	750	1	128	32
	1000	1	128	48
	550	2	128	19
	750	2	128	14
	1000	2	128	10
	550	1	256	23
	750	1	256	29
	1000	1	256	45

3

Basic Notions in Design of Experiments

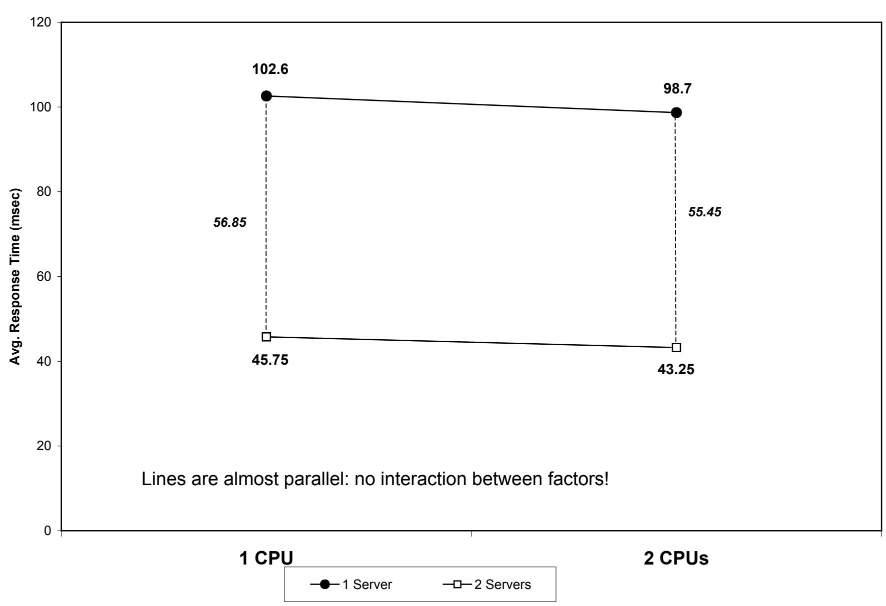
- Primary Factors: those whose effects need to be quantified.
- Secondary factors: not interested in quantifying effects.
- Replication: repetition of some or all of the experiments.
- Design: (no. of experiments, factor level combination, no. of replications per experiments).

4

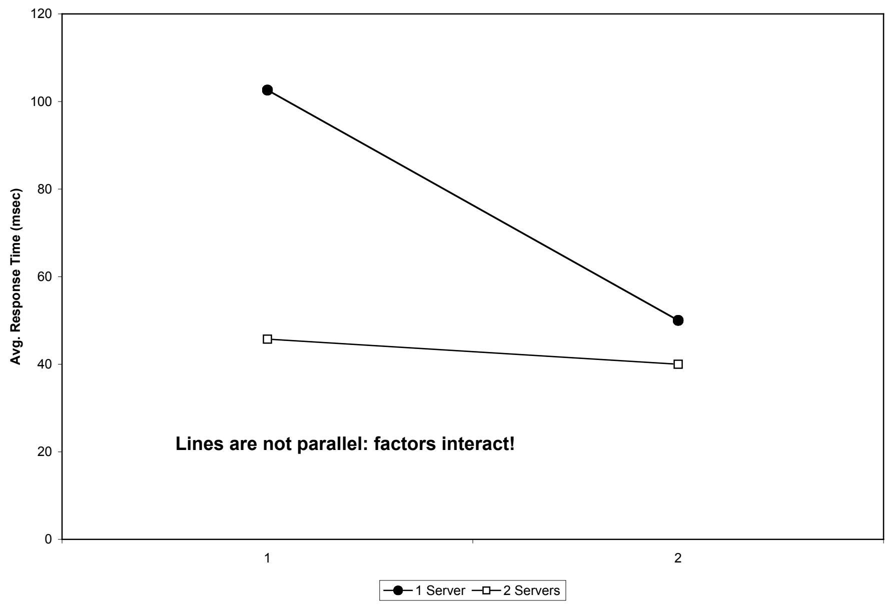
Basic Notions in Design of Experiments

- Experimental unit: entity used for an experiment.
- Interaction: Factors A and B interact if the effect of one depends upon the level of the other.

5



6



7

Common Errors in Experimentation

- Variation due to experimental error is ignored.
- Important parameters are not controlled.
- Effects of different factors not isolated.
- Simple one-factor-at-a-time designs.
- Interactions are ignored.
- Too many experiments are conducted.

8

Types of Experimental Designs

- Simple Designs:
  - Start with base combination of factor levels and vary one factor at a time.

	Factors			Response
	CPU Clock Frequency (MHz)	Number of CPUs	Main Memory (MB)	Benchmark Execution Time (sec)
Levels	550	1	128	25.0
	750	1	128	32.0
	1000	1	128	48.0
	550	2	128	19.0
	750	2	128	13.5
	1000	2	128	10.0
	550	1	256	23.0
	750	1	256	29.0
	1000	1	256	45.0
	550	2	256	16.5
	750	2	256	11.8
	1000	2	256	8.8

factors

$$n = 1 + \sum_{i=1}^{k} (n_i - 1)$$

exp.   # levels of factor i.

n=1+(3-1)+(2-1)+(2-1)= 5

Not good if factors interact.

9

Types of Experimental Designs

- Full Factorial Design:
  - Uses all possible combinations of all levels of all factors.

	Factors			Response
	CPU Clock Frequency (MHz)	Number of CPUs	Main Memory (MB)	Benchmark Execution Time (sec)
Levels	550	1	128	25.0
	750	1	128	32.0
	1000	1	128	48.0
	550	2	128	19.0
	750	2	128	13.5
	1000	2	128	10.0
	550	1	256	23.0
	750	1	256	29.0
	1000	1	256	45.0
	550	2	256	16.5
	750	2	256	11.8
	1000	2	256	8.8

$$n = \prod_{i=1}^{k} n_i$$

n=3*2*2=12

Too costly!

10

Types of Experimental Designs

- Reducing Cost of Full Factorial Design:
  - Reduce the no. of levels of each factor. If all factor have 2 levels, we have a 2^k factorial design.
  - Reduce the number of factors.
  - Use fractional factorial designs.

11

Types of Experimental Designs

- Fractional Factorial Design:
  - Use a fraction of the full factorial design.

	Factors		Response
	CPU Clock Frequency (MHz)	Number of CPUs	Benchmark Execution Time (sec)
Levels	550	1	25.0
	750	1	32.0
	1000	1	48.0
	550	2	19.0
	750	2	13.5
	1000	2	10.0

The factor memory size was eliminated.

$$n = \prod_{i=1}^{k-p} n_i$$

n=3*2=6

Some interactions among factors may be lost!

12

2^k Factorial Designs

13

2^k Factorial Designs

- 2^k designs are used to determine the effects of k factors, each of which have two alternatives or levels
  - Easier to analyze than full factorial designs
  - Help sort out factors in the order of their impact, especially when there are a large number of factors

14

2^2 Factorial Designs

- Special case of 2^k design with k = 2
- Example: impact of cache size and memory size on the performance of a computer

Cache size (KB)	Memory Size 4 MB	16 MB
1	15	45
2	25	75

Performance in MIPS

15

Regression Model

Let y denote the performance of the computer. We can model y using a non-linear regression model as follows:

$$y = q_0 + q_a x_a + q_b x_b + q_{ab} x_a x_b$$

where x_a and x_b are variables that represent the factors memory size and cache size respectively, and the q's are called effects

16

Regression Model (con'td)

Let

$$x_a = \begin{cases} -1 & \text{if memory size } = 4 \text{ MB} \\ 1 & \text{if memory size } = 16 \text{ MB} \end{cases}$$

$$x_b = \begin{cases} -1 & \text{if cache size } = 1 \text{ KB} \\ 1 & \text{if cache size } = 2 \text{ KB} \end{cases}$$

Substituting the four observations yields

$$15 = q_0 - q_a - q_b + q_{ab}$$
$$45 = q_0 + q_a - q_b - q_{ab}$$
$$25 = q_0 - q_a + q_b - q_{ab}$$
$$75 = q_0 + q_a + q_b + q_{ab}$$

17

Regression Model (cont'd)

There is a unique solution for the four effects: $q_0 = 40$, $q_a = 20$, $q_b = 10$, $q_{ab} = 5$. So, we have:

$$y = 40 + 20x_a + 10x_b + 5x_a x_b$$

Thus the mean performance is 40 MIPS, the effect of memory size is 20 MIPS, the effect of cache size 10 MIPS, and the interaction between memory and cache size accounts for 5 MIPS

18

Computing effects

In general, the model for a 2^2 design can be solved to obtain:

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

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

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

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

19

Sign table method for calculating effects

I	A	B	AB	Y
1	-1	-1	1	15
1	1	-1	-1	45
1	-1	1	-1	25
1	1	1	1	75
160	80	40	20	Total
40	20	10	5	Total/4

20

Allocation of variation

- The importance of a factor is explained by the proportion of the total variation in the response that is explained by the factor
  - If one factor contributes 90% of the variation whereas another factor contributes only 5%, then the second factor may be considered relatively unimportant
- The total variation of y or Sum of Squares Total (SST) is given by:

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

21

Allocation of variation (cont'd)

For a 2^2 design, the variation can be divided into three parts:

$$SST = 2^2 q_A^2 + 2^2 q_B^2 + 2^2 q_{AB}^2$$

where the three terms on the RHS represent the portion of the variation that is explained by the effects of A, B, and the interaction AB respectively (see derivation in textbook). Thus

$$SST = SSA + SSB + SSAB$$

22

Example

For the memory-cache example,

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

$$\text{Total variation } = \sum_{i=1}^{4}(y_i - \bar{y})^2 = (25^2 + 15^2 + 15^2 + 35^2)$$

$$= 2100 = 4 \times 20^2 + 4 \times 10^2 + 4 \times 5^2$$

Thus, 76% (1600) of the total variation can be attributed to memory size, 19% (400) can be attributed to cache, and only 5% (100) can be attributed to the interaction between memory and cache.

General 2^k Factorial Designs

- The analysis techniques described so far for 2^2 designs can be extended to that for a 2^k design
  - Given k factors with two levels each, 2^k experiments are required and there are 2^k effects
    - k main effects, $\binom{k}{2}$ two-factor interactions, $\binom{k}{3}$ three-factor interactions, and so on
    - The sign table technique for calculating effects and allocating the variation is applicable

$2^k r$ Factorial Designs with Replications

25

$2^k r$ Factorial Designs

- Shortcoming of 2^k design: cannot estimate experimental error because no experiment is repeated
- Solution: Replication
  - Repeat each experiment in a 2^k design r times yielding $2^k r$ observations (called a $2^k r$ design)
  - We will discuss k = 2
    - Can be generalized as in 2^k design (see section 18.9 of textbook)

26

2^2r Factorial Design

- Model for 2^2 design is extended to add an error term

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

where the q's are effects as before and e is the experimental error

- Holding the factor level constant and repeating the experiment yields samples of the response y_i
- Statistical analysis of the y_i's yields the fraction of variation due to experimental error, and confidence intervals for y

27

Computation of Effects

The effects can be calculated using a sign table as before except that in the y column we put the sample mean of r measurements at the given factor level

I	A	B	AB	Y	Mean $\bar{y}$
1	-1	-1	1	(15,18,12)	15
1	1	-1	-1	(45,48,51)	48
1	-1	1	-1	(25,28,19)	24
1	1	1	1	(75,75,81)	77
164	86	38	20		Total
41	21.5	9.5	5		Total/4

28

Estimation of experimental errors

Once the effects have been computed, the model can be used to estimate the response for any given factor values as:

$$\hat{y}_i = q_0 + q_A x_{Ai} + q_B x_{Bi} + q_{AB} x_{Ai} x_{Bi}$$

The difference between the estimate and the measured value y_{ij} in the *j*th replication of the *i*th experiment represents the experimental error

$$e_{ij} = y_{ij} - \hat{y}_i = y_{ij} - q_0 + q_A x_{Ai} + q_B x_{Bi} + q_{AB} x_{Ai} x_{Bi}$$

29

Sign table augmented with errors

I	A	B	AB	(y_{i1}, y_{i2}, y_{i3})	$\hat{y}_i$	e_{i1}	e_{i2}	e_{i3}
1	-1	-1	1	(15,18,12)	15	0	3	-3
1	1	-1	-1	(45,48,51)	48	-3	0	3
1	-1	1	-1	(25,28,19)	24	1	4	-5
1	1	1	1	(75,75,81)	77	-2	-2	4
41	21.5	9.5	5					

$$SSE = \sum_{i=1}^{2^2} \sum_{j=1}^{r} e_{ij}^2 = 102$$

30

Allocation of variation

Let $\bar{y}_{..}$ represent the mean of responses from all replications of all experiments. Then

$$SST = \sum_{i,j}(y_{ij} - \bar{y}_{..}) = 2^2 r q_A^2 + 2^2 r q_B^2 + 2^2 r q_{AB}^2 + \sum_{i,j} e_{ij}^2$$

$$SST = SSA + SSB + SSAB + SSE$$

We can also show that SST = SSY - SS0, giving us an easier way to compute SSE:

$$SSE = SSY - (SS0 + SSA + SSB + SSAB)$$

$$\text{where } SSY = \sum_{i,j} y_{ij}^2 \text{ and } SS0 = 2^2 r q_0^2$$

31

Example

For our memory-cache example

SSY = $15^2+18^2+12^2+45^2+\ldots+75^2+75^2+81^2$= 27,204

SS0 = 12 x 41^2= 20,172

SSA = 5547, SSB = 1083, SSAB = 300

SSE = 27,204 - 2^2x3($41^2+21.5^2+9.5^2+5^2$) = 102

SST = SSY - SS0 = 27,204 - 20,172 = 7032

Thus, factor A explains 5547/7032 or 78.88% of the variation, factor B explains 15.4%, interaction AB explains 4.27% of the variation.

The remaining 1.45% is attributed to experimental errors

32

Confidence intervals for effects

- The effects computed from a sample are random variables
- We can calculate the confidence intervals for the effects if we assume that
  - Errors are normally distributed with zero mean
  - Errors for different experiments are independent
  - Errors have a constant standard deviation
- Under these assumptions, we can use the t-distribution to compute confidence intervals as discussed previously (Chapter 13)

33

Confidence intervals for effects

- The sample variance of errors can be estimated from the SSE as follows:

$$s_e^2 = \frac{SSE}{2^2(r-1)}$$

$$s_{q_0}^2 = s_{q_A}^2 = s_{q_B}^2 = s_{q_{AB}}^2 = s_e^2/(2^2 r)$$

- The confidence intervals for effects can then be computed as $q_i \pm t_{[1-\alpha/2;2^2(r-1)]} s_{q_i}$

34

Example

- For the memory-cache example,

$$s_e = \sqrt{\frac{SSE}{2^2(r-1)}} = \sqrt{\frac{102}{8}} = \sqrt{12.75} = 3.57$$

$$s_{q_i} = s_e / \sqrt{(2^2 r)} = 3.57/\sqrt{12} = 1.03$$

- The t-value for 8 degrees of freedom and 90% confidence is 1.86. Thus, the confidence intervals for the effects are $q_i \pm (1.86)(1.03) = q_i \pm 1.92$ that is (39.08,42.91), (19.58,23.41), (7.58,11.41), (3.08,6.91) for q_0, q_A, q_B, and q_{AB} respectively

35

Additional Reading

- Confidence intervals for predicted responses: Section 18.6
- Visual tests for verifying the model: Section 18.7
- Multiplicative models: Section 18.8
- General $2^k r$ factorial design: Section 18.9

36