ECE560: Computer Systems Performance Evaluation

Lecture #3 – Experimental Design and Analysis

Instructor: Dr. Liudong Xing

Administration Issues (1/29)

- Last day to add/drop: **Today**
- Project team setup due **Today**
  - Two students per team
- Homework #1 assigned
  - Please download problems from course website
  - Due: **February 5, Monday**

Review of Lecture #2

- 3 key issues related to performance measurements are presented
  - 5 types of workloads that have been commonly used to compare CS
    - Addition instruction, Instruction mixes, Kernels, Synthetic programs, Application benchmarks
  - Performance monitoring techniques
  - Summarizing measured data
    - By a single number: sample mean, sample median, sample mode, geometric mean
    - Variability: sample range, sample variance, sample standard deviation, sample COV, percentiles, SIQR

Topics

- Terminology in experimental design and analysis
- Types of experimental designs
- 2^k factorial design and analysis

Reference:

R. Jain, "The Art of Computer Systems Performance Analysis: Techniques for Experimental Design, Measurement, Simulation, and Modeling", Wiley, John & Sons, 1990 (Part IV)

Introduction

- Involved in both Measurement and Simulation
- The goal of a proper experimental design is to obtain the maximum information with the minimum number of experiments
- A careful experiment design and analysis are essential for reducing cost and drawing meaningful conclusion

Terminology

- An *obsolete* example: choices to make for designing a personal workstation
  - Microprocessor: 68000, Z80, 8086
  - Memory size: 512KB, 2MB, 8MB
  - Disk drives: 1, 2, 3, 4
  - Workload types: secretarial, managerial, scientific
  - User education level: high school, college, postgraduate

Terminology (I)

- **Response variable**
  - The outcome of an experiment
  - The measured performance of the system
  - Example
    - Throughput in tasks completed per unit time
    - Response time of tasks

Terminology (II)

- **Factors**
  - Each variable that affects the response variable and has several alternatives is a factor
  - Also called *predictor variables* or *predictors*
  - Example: five factors in workstation design study
    - CPU types, memory size, number of disk drives, workload types, user's education levels

Terminology (III)

- **Primary factors**
  - The factors whose effects need to be quantified
  - Example: one may be primarily interested in quantifying the effects of *CPU type*, *memory size*, and number of *disk drives*
- **Secondary factors**
  - Factors that impact the performance but whose impact we are not interested in quantifying
  - Example: one may not be interested in determining whether performance with postgraduate is better than that with college students; quantifying the difference between the three *workload*s

Terminology (IV)

- **Levels**
  - The values that a factor can assume are called its levels
  - Each factor level constitutes one alternative for that factor
  - Also referred to as *treatments*
  - Example: CPU type has three levels: Z80, 68000, 8086 etc.
- **Replication**
  - Repetition of all or some experiments is called replication
  - Example: the study has three replications if all experiments in the study are repeated 3 times

Terminology (V)

- **Design**: consisting of specifying
  - The number of experiments
  - The factor level combinations for each experiment
  - The number of replications of each experiment
  - Example:

    perform 3*3*4*3*3=324 experiments and 5 replications → 1620 observations

Terminology (VI)

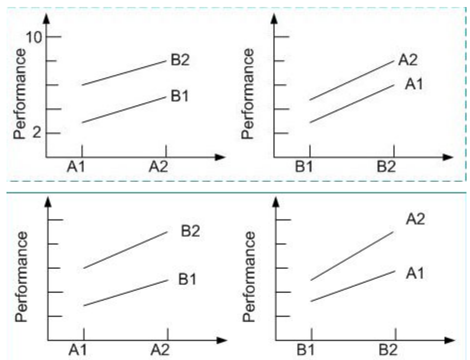
- **Experimental unit**
  - Any entity used for the experiment
  - Generally, only those units considered as one of the factors in the study are of interest
  - Example:
    - Users hired to use the workstation while measurements are being performed
    - Patients in medical experiments
    - Land used in agricultural experiments

Terminology (V)

- **Interaction**
  - Two factors A and B are said to interact if the effect of one depends on the level of the other
  - Example: the performance of a system with two factors A, B, each with two levels

Non-Interacting Factors		
	A1	A2
B1	3	5
B2	6	8

Interacting Factors		
	A1	A2
B1	3	5
B2	6	9



Agenda

- Terminology in experimental design and analysis
- **Types of experimental designs**
  - Simple designs,
  - Full factorial designs,
  - Fractional factorial designs
- 2^k factorial design and analysis

Simple Designs (I)

- Start with a typical configuration and vary *one* factor at a time to see how that factor affects the performance
- Example: revisit the workstation design
  - Initial configuration: *Z80* CPU + *2 disk drives* + *2MB memory* running a *managerial* task by a *college* student
  - Measure the performance of the initial configuration
  - Vary the first factor of CPU type, compare performance to decide which CPU is the best
  - Then change the number of drives to 1, 3, 4, compare performance to find the optimal number.

Simple Designs (II)

- Given k factors (the ith factor having n_i levels), a simple design requires only n experiments

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

- Disadvantages
  - Simple design may lead to wrong conclusions if the factors have interaction
  - Example: effects of the CPU depends on the size of the memory, need to try all possibilities to decide the optimal combination
  - Simple designs are not recommended

Full Factorial Designs (I)

- Utilizes every possible combination at all levels of all factors
- Example:
  - n = (3 CPUs)(3 memory sizes)(4 disk drives)(3 workloads)(3 educational levels) = 324 experiments
- Given k factors (the ith factor having n_i levels), a performance study using a full factorial design requires n experiments

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

Full Factorial Designs (II)

- Advantages
  - Every possible combination is examined
  - The effects of every factor including the secondary factors and their interactions are examined
- Disadvantages
  - Cost of the study is high, especially when each experiment has to be repeated several times

Full Factorial Designs (III)

- Three ways to reduce the number of experiments
  - Reduce the number of levels for each factor
    - Example: 2^k factorial design
  - Reduce the number of factors: by choosing the factors that have significant impact on performance
  - Use fractional factorial designs →

Fractional Factorial Designs (I)

- Use a fraction of the full factorial design
- *Pros* and *cons*
  - Save time and expense as compared to full factorial designs
  - Information obtained is less: may not be possible to incorporate interactions among all factors

Fractional Factorial Designs (II)

- Example: revisit the workstation design ignoring the number of disk drives for this example
  - A full factorial design:
    n = (3 CPUs)(3 memory sizes)(3 workloads)(3 educational levels) = 81 experiments → 3^4 design
  - A 3^{4-2} fractional factorial design: 9 experiments, each of the four factors is used three times at each of its three levels →

Fractional Factorial Designs (III)

- An example fractional factorial design (Table 1)

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

Hands-on Exercise

- The performance of a system being designed depends on 3 factors
  - CPU type: 68000, 8086, Z80
  - Operating system type: CP/M, MS-DOS, UNIX
  - Disk drive type: A, B, C
- Q: How many experiments are required to analyze the performance if
  - There is significant interaction among factors
  - There is no interaction among factors
  - The interactions are small compared to main effects

Agenda

- Terminology in experimental design and analysis
- Types of experimental designs
- **2^k factorial design and analysis**

2^k Factorial Designs

- A 2^k factorial design is used to determine the effect of k factors, each with two alternatives or levels
  - The effect of a factor is unidirectional (very often)
    - Performance either continuously decreases or increases as the factor is increased from min to max
    - Can begin by experimenting at the *min* and *max* level of the factor
    - Decide if the difference in performance is significant enough to justify detailed examination
  - Can help in sorting out factors in the order of impact on performance
  - Try more levels for factors that have significant impact

An Example: 2^2 Factorial Designs

- Study/quantify the impact of memory size and cache size on the performance of a workstation being designed. 2 levels of each factor are chosen for the initial simulation.

Performance in MIPS		
	A: Memory Size	
B: Cache Size	4MB	16MB
1KB	***15***	***45***
2KB	***25***	***75***

- Define two variables

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

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

2^2 Factorial Designs (Example Cont'd)

- Regression analysis allows one to estimate or predict a *r.v.* as a function of several other variables
  - Response variable – estimated variable
  - Factors/predictors/predictor variables – variables used to predict the response
  - Assume all predictor variables are quantitative so that arithmetic operations (+,*) are meaningful

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

- y: response variable, the system performance in MIPS
- q_0: the mean performance
- q_A: the effect of memory on the performance
- q_B: the effect of cache on the performance
- q_{AB}: the effect of the interaction between memory and cache

2^2 Factorial Designs (Example Cont'd)

- Substituting the four observations

$$\begin{aligned} 15 &= q_0 - q_A - q_B + q_{AB} \\ 45 &= q_0 + q_A - q_B - q_{AB} \\ 25 &= q_0 - q_A + q_B - q_{AB} \\ 75 &= q_0 + q_A + q_B + q_{AB} \end{aligned}$$

- Solving the equations

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

What if there are 3 factors? 5 factors? 10 factors?

Sign Table Method

Quantify Effects for General 2^2 Design

- Suppose y_1 (-1,-1), y_2 (1, -1), y_3 (-1,1), y_4 (1,1) represent the four observed response of a system with two factors A (-1,1) and B (-1,1).

Exp.	A	B	y
1	-1	-1	y1
2	1	-1	y2
3	-1	1	y3
4	1	1	y4

- Use regression model for 2^2 design

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

- We have

$$y_1 = q_0 - q_A - q_B + q_{AB}$$
$$y_2 = q_0 + q_A - q_B - q_{AB}$$
$$y_3 = q_0 - q_A + q_B - q_{AB}$$
$$y_4 = q_0 + q_A + q_B + q_{AB}$$

$$q_0 = \frac{1}{4}(y_1 + y_2 + y_3 + y_4)$$
$$q_A = \frac{1}{4}(-y_1 + y_2 - y_3 + y_4)$$
$$q_B = \frac{1}{4}(-y_1 - y_2 + y_3 + y_4)$$
$$q_{AB} = \frac{1}{4}(y_1 - y_2 - y_3 + y_4)$$

Sign Table Method for Calculating Effects

- A 4 X 4 sign matrix for the example:

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

1: not explicitly written out
+: 1
-: -1

- Apply column multiplication operation
  - E.g.: Multiply entries in column I by those in column y, put the sum under column I
- Each sum is divided by 4 to obtain the coefficients of the regression model

Agenda

- Terminology in experimental design and analysis
- Types of experimental designs
- **2^k factorial designs**
  - **2^2 factorial designs**
    - **Regression equations** $y = q_0 + q_A x_A + q_B x_B + q_{AB} x_A x_B$
    - **Sign table**
  - **General 2^k factorial designs**

General 2^k Factorial Designs

- 2^k factorial design
  - Require 2^k experiments for *k* factors, each having 2 levels
  - The analysis produces 2^k effects: C_k^1 main effects; C_k^2 two-factor interactions; C_k^3 3-factor-interactions……
- The techniques of quantifying effects can be easily extended from 2^2 to 2^k design

General 2^k Factorial Designs -- Example

- 3 factors need to study for designing a machine

Factor	Level -1	Level 1
Memory size, A	**4MB**	**16MB**
Cache size, B	**1KB**	**2KB**
# of processors, C	**1**	**2**

- The measured performance in MIPS

Cache size (KB)	Memory size 4MB		Memory size 16MB	
	1 processor	2 processor	1 processor	2 processor
1	***14***	***46***	***22***	***58***
2	***10***	***50***	***34***	***86***

Example (Cont'd)

- Hands-on problem: quantify the main effects and all interactions using
  - Regression equations
  - Sign table

Summary

- Frequently-used terms in experimental designs
  - Response variables, factors (predictors, predictor variables), levels (treatments), primary factors, secondary factors, replication, designs, experimental units, interactions
- Types of experimental designs
  - Simple designs, full factorial designs, fractional factorial designs
- A closer look at 2^k factorial designs
  - Regression equations and sign table methods to quantify the effects of the factors on the system performance

Next Topic: Simulations Overview