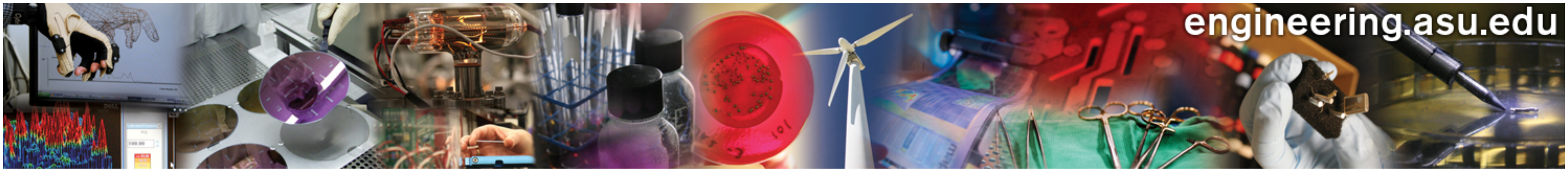


IRA A. FULTON SCHOOLS OF ENGINEERING

Leading engineering discovery and innovative education for global impact on quality of life.

Introduction to the Design and Analysis of Experiments

Violet R. Syrotiuk

School of Computing, Informatics, and Decision Systems Engineering

Arizona State University

Complex Engineered Systems

- What makes an engineered system complex?
  - Its large size.
  - Humans control their structure, operation, and evolution over time.
- Examples of complex engineering networks:
  - The power grid.
  - Transportation networks.
  - Computer networks (e.g., the Internet, etc.).

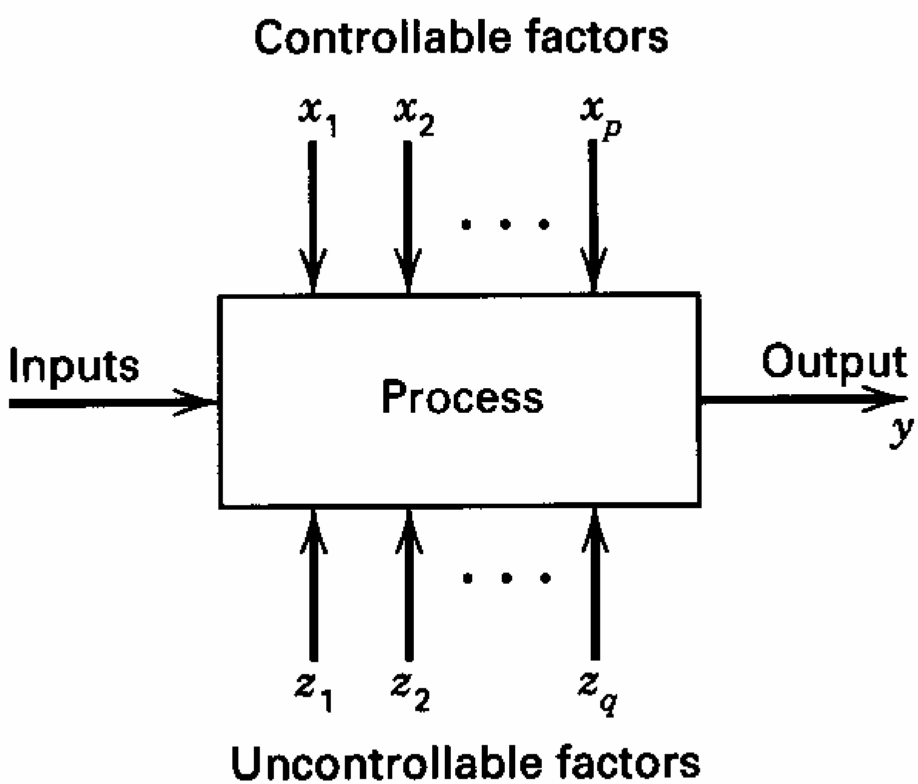
Experimentation

- Experimentation is a way to improve our understanding of CESs.
- Experiments are used widely for, e.g.:
  - Process characterization and optimization.
  - Improve the reliability and performance of products and processes.
  - Product/process design and development.
  - Achieve product and process robustness.

"All experiments are designed experiments;
some are poorly designed,
some are well-designed."

George E. P. Box

A General Model of a Process/System†



† "Design and Analysis of Experiments," by Douglas C. Montgomery, Wiley, 8th edition, 2013.

Four Eras in the History of DoE

- The agricultural origins, 1918-1940s:
  - R.A. Fisher and his co-workers.
  - Profound impact on agricultural science.
  - Factorial designs, ANOVA.
- The first industrial era, 1951- late 1970s:
  - Box and Wilson, response surfaces.
  - Applications in the chemical and process industries.

Four Eras in the History of DoE (cont'd)

- The second industrial era, late 1970s - 1990:
  - Quality improvement initiatives in many companies.
  - Taguchi and robust parameter design, process robustness.
- The modern era, beginning circa 1990.

Experimentation in CENs

- From a recent workshop report†:

  "The science of experiment design is widely used in science and engineering disciplines, but is often ignored in the study of complex engineered networks. This in turn has led to a shortage of simulations that we can believe in, of experiments driven by empirical data, and of results that are statistically illuminating and reproducible in this field."

† Networking and Information Technology Research and Development (NITRD), Large Scale Networking (LSN), Workshop Report on Complex Engineered Networks, September 2012.

Factorial Designs

- In a factorial experiment, all possible combinations of factor levels are tested.
- The golf experiment:
  - Type of driver.
  - Type of ball.
  - Walking versus riding.
  - Type of beverage.
  - Time of round, etc.

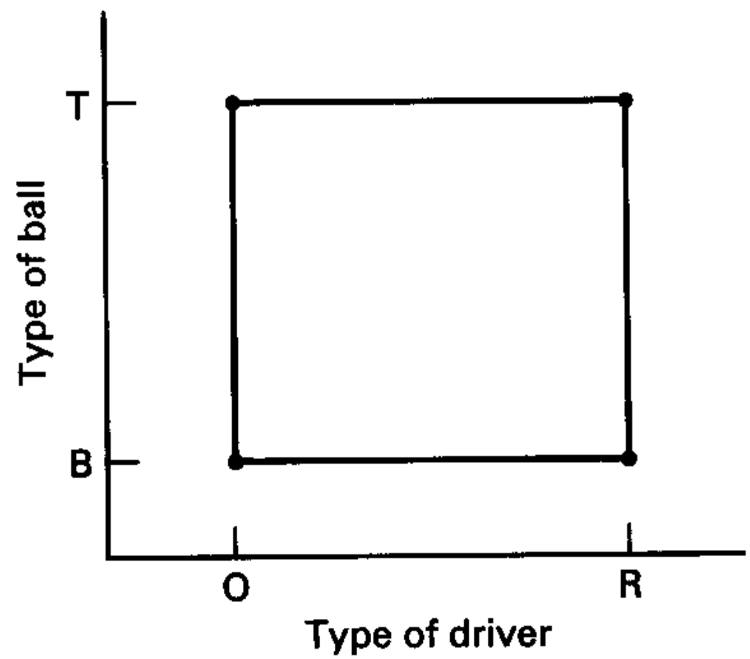


Figure 1-4 A two-factor factorial experiment involving type of driver and type of ball.

The Experimental Design

- An experiment is given by an N × k array.
  - The k columns correspond to the factors.
    - Each factor F_i, $1 \le i \le k$ has a set of levels L_i.
- Each of the N rows corresponds to a test in which each factor F_i is set to a level in L_i.
- For the two-factor factorial experiment:

	Ball	Driver
1	B	O
2	B	R
3	T	O
4	T	R

Factorial Designs with Several Factors

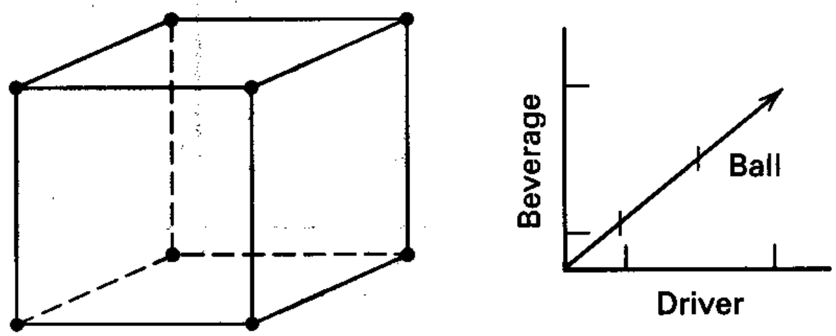


Figure 1-6 A three-factor factorial experiment involving type of driver, type of ball, and type of beverage.

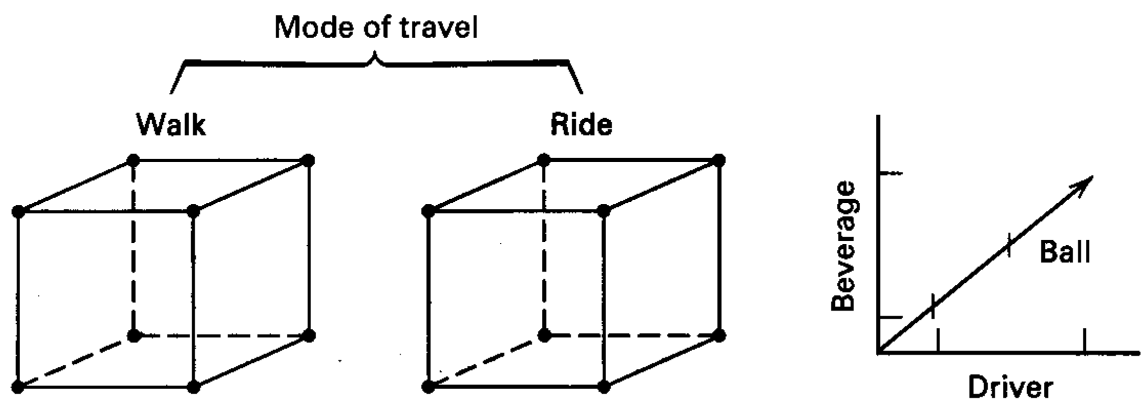


Figure 1-7 A four-factor factorial experiment involving type of driver, type of ball, type of beverage, and mode of travel.

A Fractional Factorial

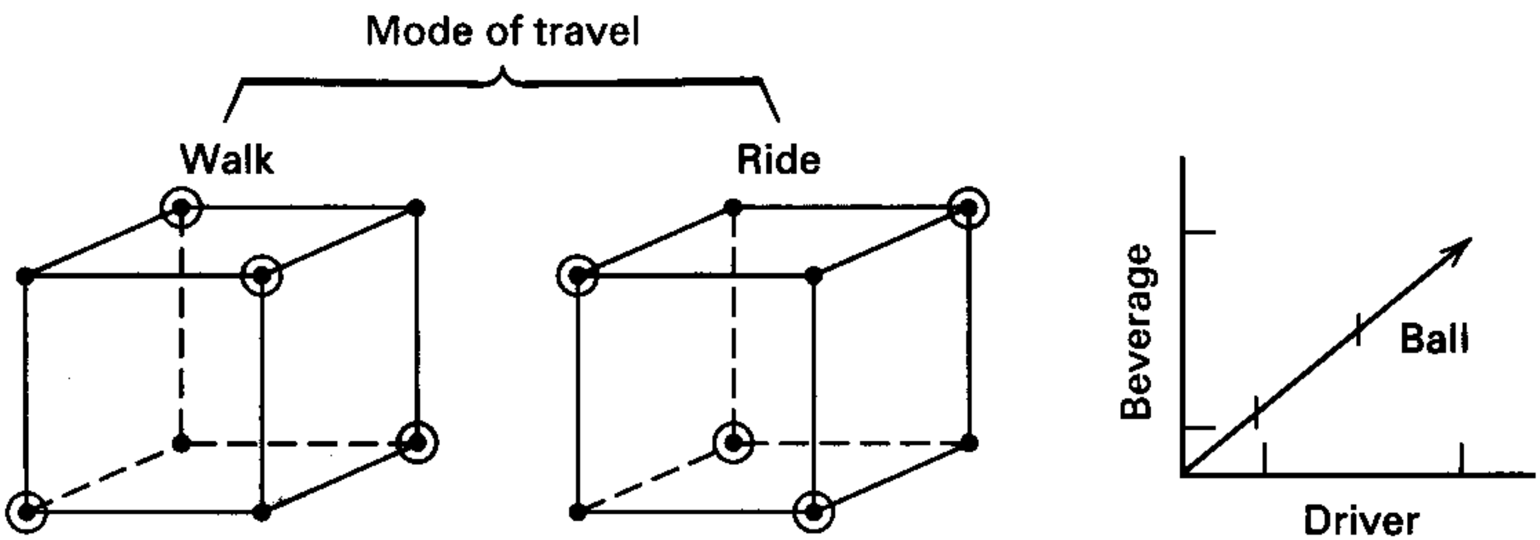


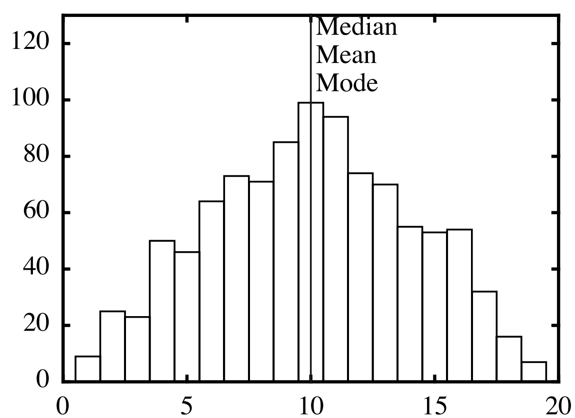
Figure 1-8 A four-factor fractional factorial experiment involving type of driver, type of ball, type of beverage, and mode of travel.

Statistical Rigour

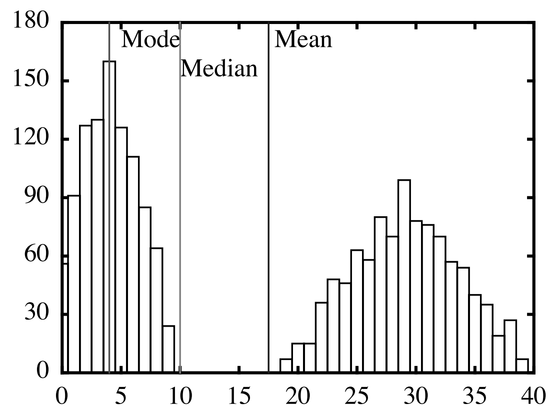
- Understanding common statistical methods is invaluable in being able to represent results coherently and accurately.
- When measuring a system that does not have fixed behaviour:
  - Perform multiple measurements (replicates).
  - Statistical error arises from variation that is uncontrolled; it is generally unavoidable.

Sample Distributions

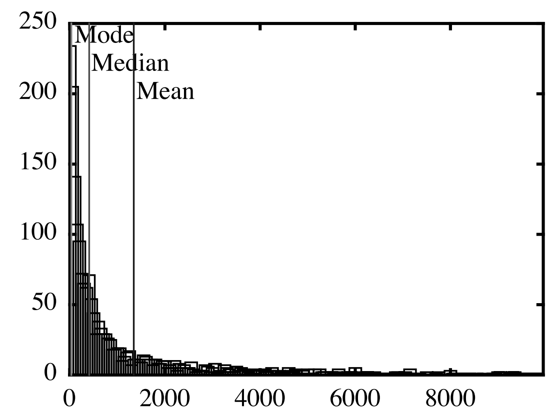
- The relationship between the measurements of centrality (mean, median, and mode) give hints about the distribution of the data collected.



A: Normal Distribution



B: Bimodal Distribution



C: Exponential Distribution

Expressing Variation

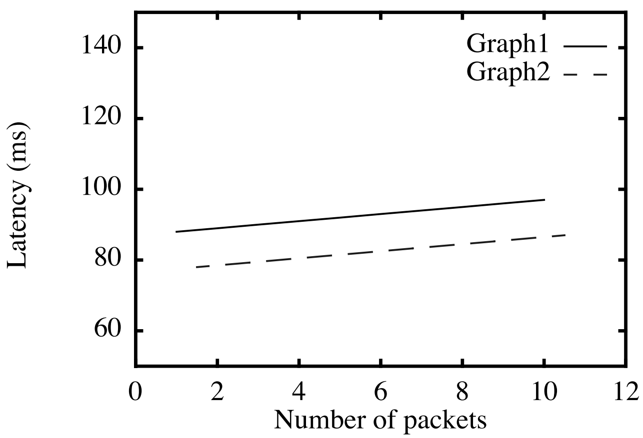
- Measures of centrality are not sufficient to completely describe a data set.
- It is often helpful to include a measure of the variance of the data.
  - A small variance implies that the mean is a good representative of the data, whereas a large variance implies that it is a poor one.
- The most commonly used measure of variance is the standard deviation.

Margin of Error

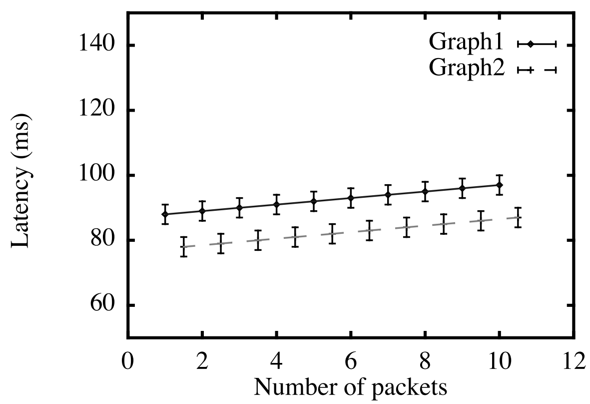
- Another metric for analyzing the usefulness of the mean is the margin of error.
  - The margin of error expresses a range of values about the mean in which there is a high level of confidence that the true value falls.

Graphing and Error Margins

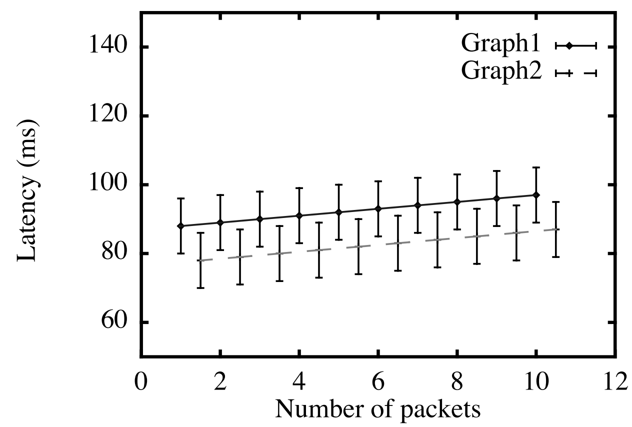
- The value of the error margins depict the results in completely different ways.†



A: Latency Improvement without error margins



B: Latency Improvement with small error margins



C: Latency Improvement with large error margins

† C. Small, N. Ghosh, H. Saleeb, M. Selzter, and K. Smith, “Does Systems Research Measure Up?” Harvard Computer Science Group, Technical Report TR-16-97.

Probability Distributions & Testing

- Plotting a histogram of the values in a sampled data set is easy way to get an idea of what type of distribution the data follows.
- The X^2 test can be used to determine if sampled data follows a specific distribution.
  - X^2 can be used to obtain a p-value from a family of X^2 distributions; the larger the p-value, the higher the probability that the measured distribution matches the candidate distribution.

Basic Statistical Concepts

- Hypothesis testing: A statement either about the parameters of a probability distribution or the parameters of a model.

  $H_0: \mu_1 = \mu_2$ (null hypothesis)

  $H_1: \mu_1 \neq \mu_2$ (alternative hypothesis)

- If the null hypothesis is rejected when it is true, a type I error has occurred.
- If the null hypothesis is not rejected when it is false, a type II error has been made.

Analysis of Variance (ANOVA)

- Analysis of the fixed effects model.
  - Estimation of model parameters.
- Model adequacy checking.
  - The normality assumption.
  - Residuals. Plots in time, versus fitted values, versus other variables.
- Practical interpretation of results.

Response Surface Methodology Framework

- Factor screening.
- Finding the region of the optimum.
- Modelling and optimization of the response.

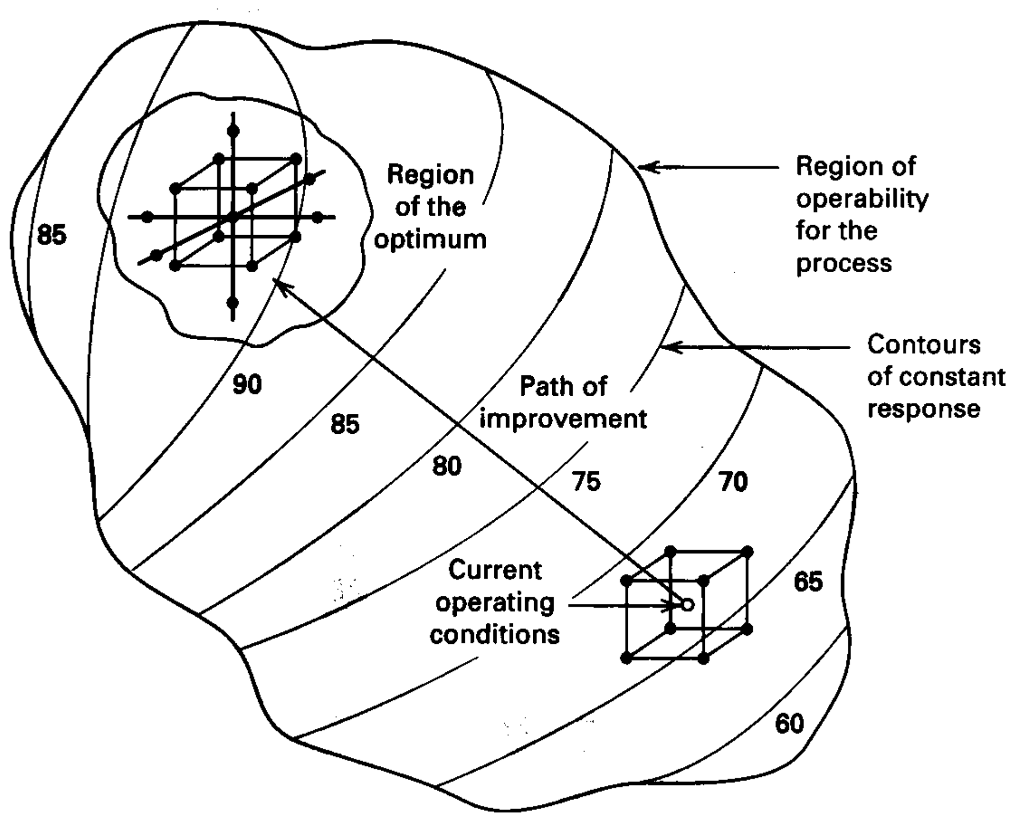


Figure 11-3 The sequential nature of RSM.

Other Aspects of RSM

- Robust parameter design and process robustness studies.
  - Find levels of controllable variables that optimize mean response and minimize variability in the response transmitted from "noise" variables.
  - Original approaches due to Taguchi (1980s).
  - Modern approach based on RSM.

Summary

- There is much known about designing and analyzing experiments!
  - Follow good practices, to improve repeatability and reproducibility of your experiments.