IE630: Simulation Modelling & Analysis
Comparing Systems

Saurabh Jain

Assistant Professor

Department of Industrial Engineering and Operations Research

IIT Bombay

Quick Recap

Comparison of Multiple Systems: Why?

- Key Utility of Simulation
  - Primarily used for comparing different alternatives to support decision-making.
- Nature of Alternative Scenarios
  - Minor variations: Changes in parameters (e.g., resource allocation, processing speed).
  - Major differences: Structural or logical modifications in decision rules or workflows.
- Ensuring Valid Comparisons
  - Apply appropriate statistical techniques to draw meaningful conclusions.
  - Avoid errors → misleading conclusions → poor decisions by ensuring:
    - Proper validation and verification of models.
    - Use of hypothesis testing and confidence intervals for comparison.

Comparison of Multiple Systems: How?

- Hypothesis Testing
  - Null Hypothesis (H_0): Both alternatives are statistically the same.
  - Alternative Hypothesis (H_1): Both alternatives are statistically different.

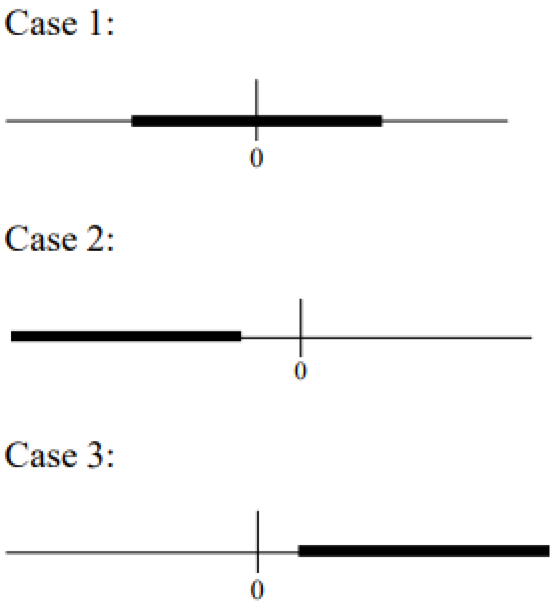
- Statistical Tests for Hypothesis Testing
  - When the number of replications is the same ($n_1 = n_2$):
    - Use Paired t-test to compare the means of two related datasets.

- When the number of replications is different ($n_1 \neq n_2$):
  - Use Welch's t-test (Welch's Confidence Interval Method) to account for unequal sample sizes and variances.

Hypothesis Testing

$$H_0: \mu_1 = \mu_2 \quad \rightarrow \quad \mu_1 - \mu_2 = (\qquad\qquad\qquad)$$

$$H_1: \mu_1 \neq \mu_2 \quad \rightarrow \quad \mu_1 - \mu_2 \neq (\qquad\qquad\qquad)$$



- Sample Confidence Intervals of (X_1 - X_2)

- Two statistical tests for this hypothesis testing
  - When the number of replications are same (n1 = n2)
  ( )
  - When the number of replications are different (n1 ≠ n2)
  ( )

Pairwise t-test

Replication (i)	Within run observations, y_{ij}	**Average,** $\bar{y}_{i1}$
1	$y_{11}, y_{12}, y_{13}, \quad \ldots, \quad y_{1m-1}, y_{1m}$	$\bar{y}_1$
2	$y_{21}, y_{22}, y_{23}, \quad \ldots, \quad y_{2m-1}, y_{2m}$	$\bar{y}_2$
3	$y_{31}, y_{32}, y_{33}, \quad \ldots, \quad y_{3m-1}, y_{3m}$	$\bar{y}_3$
4	$y_{41}, y_{42}, y_{43}, \quad \ldots, \quad y_{4m-1}, y_{4m}$	$\bar{y}_4$
…	……… $\mathbf{y_{ij}}$ ………	
n	$y_{n1}, y_{n2}, y_{n3}, \quad \ldots, \quad y_{nm-1}, y_{nm}$	$\bar{y}_n$

Replication (i)	Within run observations, y_{ij}	**Average,** $\bar{y}_{i1}$
1	$y_{11}, y_{12}, y_{13}, \quad \ldots, \quad y_{1m-1}, y_{1m}$	$\bar{y}_1$
2	$y_{21}, y_{22}, y_{23}, \quad \ldots, \quad y_{2m-1}, y_{2m}$	$\bar{y}_2$
3	$y_{31}, y_{32}, y_{33}, \quad \ldots, \quad y_{3m-1}, y_{3m}$	$\bar{y}_3$
4	$y_{41}, y_{42}, y_{43}, \quad \ldots, \quad y_{4m-1}, y_{4m}$	$\bar{y}_4$
…	……… $\mathbf{y_{ij}}$ ………	
n	$y_{n1}, y_{n2}, y_{n3}, \quad \ldots, \quad y_{nm-1}, y_{nm}$	$\bar{y}_n$

Pairwise t-test

Example

Replication (i)	First system, $X_{i1} = \overline{y}_{i1}$	Second system, $X_{i2} = \overline{y}_{i2}$	Difference, $X_{i(1-2)}$
1	54.48	56.01	-1.53
2	57.36	54.08	3.28
3	54.81	52.14	2.67
4	56.20	53.49	2.71
5	54.83	55.49	-0.66
6	57.69	55.00	2.69
7	58.33	54.88	3.45
8	57.19	54.47	2.72
9	56.84	54.93	1.91
10	55.29	55.84	-0.55
Sam. mean			1.67
Sam. S. D			1.85
Sam. Var			3.42

Welch's CI method

- Number of replications are different! (n1≠n2)
- You don't have to memorize these equations; but understand that we do need this statistical testing when replication numbers are different)
- Sample Mean and Sample Variance

$\boldsymbol{A1}$

$X_1, X_2, X_3, X_4 \; \dots\dots\dots\dots\dots \; X_n$

$\bar{\bar{X}}\ (n) \qquad S^2\ (n)$

$\boldsymbol{A2}$

$X_1, X_2, X_3, X_4 \; \dots\dots\dots\dots\dots \; X_m$

$\bar{\bar{X}}\ (m) \qquad S^2\ (m)$

Welch's CI method

- CI for 100(1-α) %

$$\left(\bar{\bar{y}}_{(A1)} - \bar{\bar{y}}_{(A2)}\right) \pm t_{\hat{f},\frac{\alpha}{2}} \sqrt{\frac{S_{A1}^2}{n_1} + \frac{S_{A2}^2}{n_2}}$$

- Estimated degree of freedom

$$\hat{f} = \frac{\left[\frac{S_{A1}^2}{n_1} + \frac{S_{A2}^2}{n_2}\right]^2}{\frac{\left(\frac{S_{A1}^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{S_{A2}^2}{n_2}\right)^2}{n_2 - 1}}$$

Dealing with multiple performance measures

- In many real-world simulations, several measures of performance are of interest simultaneously.
  - For e.g., Server Utilization, Average Time in system, & Average Number in Queue

Dealing with multiple performance measures