Comparing Systems and an Introduction to Design of Experiments

Victor S. Frost
Dan F. Servey Distinguished Professor
Electrical Engineering and Computer Science
University of Kansas
Phone: (785) 864-4833
e-mail: vsfrost@ku.edu
http://www.ittc.ku.edu/~frost

1

Comparing Systems Designs

- Goal: Determine if the performance of system 1 is different from the performance of system 2.

Let X_{1j} j=1...n be the average performance metric, e.g., delay, for system 1 replication j

Let X_{2j} j=1...n be the average performance metric, e.g., delay, for system 2 replication j

Assume:

X_{1j} are i.i.d and X_{2j} are i.i.d and n is large

Form

$Y_j = X_{1j} - X_{2j}$

2

Comparing Systems Designs

With

$Y_j = X_{1j} - X_{2j}$

The samples of the difference are y_j and the sample mean and variance of the difference is

$$\overline{Y} = \frac{1}{n}\sum_{j=1}^{n} y_j \text{ and } s_Y^2 = \frac{1}{n-1}\sum_{j=1}^{n}(y_j - Y)^2$$

$$P(\overline{Y} - z_\alpha\sqrt{\frac{s_Y^{\,2}}{n}} < E[Y] < \overline{Y} + z_\alpha\sqrt{\frac{s_Y^{\,2}}{n}})$$

$$Q(z_\alpha) = \frac{\alpha}{2}$$

- If the confidence interval does not contains 0 then we can say with $(1-\alpha)$% confidence that systems are different.
- Also if $\overline{X_1} > \overline{X_2}$ and the confidence interval does not contains 0 then we can say with $(1-\alpha)$% confidence that system 1 is better than system 2

3

Comparing Systems Designs

- Example: Compare FIFO with LIFO scheduling. Assume interarrival are exponentially distributed with mean 1 and fixed length messages, length .7 and link capacity is 1. (load=.7). Find 95% confidence, so z_α=1.96. Use n=30.
- Simulation results $\overline{Y} = 0.00843 \quad s_Y = 0.0756$

$$P(\overline{Y} - 0.023 < E[Y] < \overline{Y} + 0.023) = .95$$
$$P(-0.015 < E[Y] < 0.031) = .95$$

- Zero is contained in the confidence interval so the systems we can not say the systems are different.

4

Comparing Systems Designs

- Example: Compare FIFO scheduling with fixed and exponentially distributed messages at a load of .9. Assume interarrival are exponentially distributed with mean 1 and average messages length=0.9 and link capacity is 1. (load=.7). Find 95% confidence, so z_α=1.96. Use n=30.

5

Terminology of Design of Experiments

- Performance is affected by many factors k=number of factors, e.g.,
  - Scheduling
  - Routing
  - Traffic type
- Each factor can be at different levels, n_i=number of levels for i^{th} factor
  - Scheduling, 3- levels
    - FIFO
    - Non preemptive priority
    - WRR
  - Traffic, 6-levels
    - M/M, e.g., Exponential interarrivals and message lengths
    - M/G, e.g., Exponential interarrivals and fixed length messages
    - Load, e.g., low, medium, high
  - To explore the design space each factor is simulated for each level and replicated n times replications).
- Factors may interact

6

Question

- Determine which factors have greatest effect on the system response.
- For example: At a load of 0.9 does non preemptive priority have a statistically significant impact of the delay of the higher priority traffic.

Terminology of Design of Experiments

- Computational load can be significant for a full factorial experiment

$$\text{Number of experiments} = N_e = \prod_{i=}^{k} n_i$$

$$\text{Number of sumulation runs} = nN_e$$

- Example, an experiment with 5 factors at three levels per factor requires $3^5=243$ experiments with 30 replications; 7290 simulations are required.
- To reduce the computational load it is common to limit the number of levels for each factor to 2, creating a 2^k design
- 2^k design will be considered here.
- For other more details see:
  - “Simulation Modeling and Analysis” 5th Edition, Averill Law, 2013
  - "The Art of Computer Systems Performance Analysis: Techniques for Experimental Design Measurement, Simulation, and Modeling," R. Jain, Wiley- Interscience, New York, NY, April 1991.

2^k Factorial Designs

- For this discussion let k=2 and R_i=system response with factors and levels set to their i^{th} set of values, i.e., the i^{th} system configuration.
- Here there are 4 design points, i.e., system configurations.
- Each factor f_i can be at one of two levels, denote one level as "+" and the other as "-".
- To capture interactions define $f_i\ f_j$
- Using this information a Design Matrix is defined as

Factor Design Point	Factor 1	Factor 2	(Factor 1)(Factor 2)	Response
1	-	-	+	R_1
2	+	-	-	R_2
3	-	+	-	R_3
4	+	+	+	R_4

9

2^k Factorial Designs

- The main effect of factor j is defined as e_j
- e_j=average change of the response due to moving factor j from its level as "+" to its "-" holding all other factors constant
- Define e_{ij}=two-factor (two-way) interaction effect (by convention the difference is scaled by ½)
- Using the Design Matrix

$$e_1 = \frac{R_2 - R_1}{2} + \frac{R_4 - R_3}{2}$$

$$e_2 = \frac{R_3 - R_1}{2} + \frac{R_4 - R_2}{2}$$

$$e_{12} = \frac{1}{2}\left(\frac{R_1 + R_4}{2} - \frac{R_2 + R_3}{2}\right)$$

10

2^k Factorial Designs

- Example: At a load of 0.9 does non preemptive priority have a statistically significant impact of the delay of the higher priority traffic.
- Traffic:
  - Level 1: low priority traffic has exponential interarrival times (rate =1) and exponential length packets (length=1). High priority traffic has exponential interarrival times (rate =1) and exponential length packets (length=1).
  - Level 2: low priority traffic has exponential interarrival times (rate =1) and exponential length packets (length=1). High priority traffic has exponential interarrival times (rate =1) and gamma distributed message lengths with a

    shape parameter =k= 9 Expected value = $k\theta=1$

    scale parameter θ= 1/9 Variance =$k\theta^2=9$
- Total Offered load = 2 so set capacity = 2.222 for a load = 0.9
- Response = delay of the higher priority traffic
- Factors
  - Traffic with levels: M/M and M/G
  - Scheduling with levels: FIFO, non preemptive priority

11

2^k Factorial Designs

- One run with Tstop = 10000 (10,000 packets/traffic type)

  No replications.

Factor Design Point	Factor 1	Factor 2	(Factor 1)(Factor 2)	Response
1	- M/M	- FIFO	+	R_1=4.38
2	+ M/G	- FIFO	-	R_2=4.17
3	- M/M	+ NPP	-	R_3=1.18
4	+ M/G	+NPP	+	R_4=1.05

$$e_1 = \frac{R_2 - R_1}{2} + \frac{R_4 - R_3}{2} = \frac{4.17 - 4.38}{2} + \frac{1.05 - 1.18}{2} = 0.17$$

$$e_2 = \frac{R_3 - R_1}{2} + \frac{R_4 - R_2}{2} = \frac{1.18 - 4.38}{2} + \frac{1.05 - 4.17}{2} = -3.16$$

$$e_{12} = \frac{1}{2}\left(\frac{R_1 + R_4}{2} - \frac{R_2 + R_3}{2}\right) = \frac{1}{2}\left(\frac{4.38 + 1.05}{2} - \frac{4.17 + 1.18}{2}\right) = 0.02$$

Conclusion

- Effect of traffic is “small” 0.17
- Effect of Non preemptive priority is large reduces the delay by 3.16
- There “seems” to be no interactions.

12

2^k Factorial Designs

- Need replications to determine statistically significant impact.

Let e_{ji} =Main effect j observed from simulation replication i

Find

$$\overline{e_j} = \frac{1}{n}\sum_{i=1}^{n} e_{ji} \text{ and } s_j^2 = \frac{\sum_{i=1}^{n}(e_{ji} - \overline{e_j})^2}{n-1}$$

For n>~30

$$P(\overline{e_j} - z_\alpha\sqrt{\frac{s_j^2}{n}} < E[Y] < \overline{e_j} + z_\alpha\sqrt{\frac{s_j^2}{n}})$$

$$Q(z_\alpha) = \frac{\alpha}{2}$$

13

2^k Factorial Designs

- 30 runs with Tstop = 5000 (5,000 packets/traffic type)
- z_α=1.96: 95% confidence

$$\overline{e_1} = -0.08 \; \overline{e_2} = -3.021 \; \overline{e_{12}} = 0.209$$

$$s_1^2 = 0.349 \; s_2^2 = 0.350 \; s_{12}^2 = 0.0898$$

$$P(-0.29 < E[e_1] < 0.13)$$

$$P(-3.23 < E[e_2] < -2.80)$$

$$P(0.10 < E[e_{12}] < 0.31)$$

Conclusion

- Traffic factor CI includes 0, we can not say the traffic effect is statistically significant.
- The scheduling factor CI does not includes 0, we can say the priority service effect is statistically significant, real.
- The CI for the interactions does not include 0 so the interaction effect is statistically significant.

14