# General Full Factorial Designs With $k$ Factors

Raj Jain
Washington University in Saint Louis
Saint Louis, MO 63130
Jain@cse.wustl.edu

These slides are available on-line at:

http://www.cse.wustl.edu/~jain/cse567-06/

# Overview

- ❑ Model

- ❑ Analysis of a General Design

- ❑ Informal Methods

  - ➢ Observation Method

  - ➢ Ranking Method

  - ➢ Range Method

# General Full Factorial Designs With k Factors

❑ Model: k factors $\Rightarrow 2^k-1$ effects

$k$       main effects

$\binom{k}{2}$    two factor interactions,

$\binom{k}{3}$    three factor interactions,

and so on.

**Example**: 3 factors A, B, C:

$$y_{ijkl} = \mu + \alpha_i + \beta_j + \xi_k + \gamma_{ABij} + \gamma_{ACik} + \gamma_{BCjk} + \\ +\gamma_{ABCijk} + e_{ijkl}$$

$$i = 1, \ldots, a; \quad j = 1, \ldots, b; \quad k = 1, \ldots, c; \quad l = 1, \ldots, r;$$

# Model Parameters

$y_{ijkl}$ = Response in the lth replication with factors A, B, and C at levels i, j, and k, respectively.

$\mu$ = Mean response

$\alpha_i$ = Effect of factor A at level i

$\beta_j$ = Effect of factor B at level j

$\xi_k$ = Effect of factor C at level k

$\gamma_{ABij}$ = Interaction between A and B at levels i and j.

$\gamma_{ABCijk}$ = Interaction between A, B, C at levels i, j, and k.

and so on

❑ Analysis: Similar to that with two factors

$$\mu = \bar{y}....$$

$$\alpha_i = \bar{y}_{i...} - \bar{y}....$$

❑ The sums of squares, degrees of freedom, and F-test also extend as expected. }

# Case Study 23.1: Paging Process

Factors and Levels for Page Swap Study

| Symbol | Factor | Levels | | |
|--------|--------|--------|--------|--------|
| | | 1 | 2 | 3 |
| A | Page Replacement Algorithm | LRUV | FIFO | RAND |
| D | Deck Arrangement | GROUP | FREQY | ALPHA |
| P | Problem Program | Small | Medium | Large |
| M | Memory Pages | 24P | 20P | 16P |

❑ Total 81 experiments.

# Case Study 23.1 (Cont)

❑ Total Number of Page Swaps

| Algor- | Prog- | GROUP | | | FREQY | | | ALPHA | | |
|---|---|---|---|---|---|---|---|---|---|---|
| ithm | ram | 24P | 20P | 16P | 24P | 20P | 16P | 24P | 20P | 16P |
| LRUV | Small | 32 | 48 | 538 | 52 | 244 | 998 | 59 | 536 | 1348 |
| | Medium | 53 | 81 | 1901 | 112 | 776 | 3621 | 121 | 1879 | 4639 |
| | Large | 142 | 197 | 5689 | 262 | 2625 | 10012 | 980 | 5698 | 12880 |
| FIFO | Small | 49 | 67 | 789 | 79 | 390 | 1373 | 85 | 814 | 1693 |
| | Medium | 100 | 134 | 3152 | 164 | 1255 | 4912 | 206 | 3394 | 5838 |
| | Large | 233 | 350 | 9100 | 458 | 3688 | 13531 | 1633 | 10022 | 17117 |
| RAND | Small | 62 | 100 | 1103 | 111 | 480 | 1782 | 111 | 839 | 2190 |
| | Medium | 96 | 245 | 2807 | 237 | 1502 | 6007 | 286 | 3092 | 7654 |
| | Large | 265 | 2012 | 12429 | 517 | 4870 | 18602 | 1728 | 8834 | 23134 |

❑ $y_{max}/y_{min} = 23134/32 = 723 \Rightarrow$ log transformation

# Case Study 23.1 (Cont)

❑ Transformed Data For the Paging Study

| Algor-ithm | Prog-ram | GROUP | | | FREQY | | | ALPHA | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | 24P | 20P | 16P | 24P | 20P | 16P | 24P | 20P | 16P |
| LRUV | Small | 1.51 | 1.68 | 2.73 | 1.72 | 2.39 | 3.00 | 1.77 | 2.73 | 3.13 |
| | Medium | 1.72 | 1.91 | 3.28 | 2.05 | 2.89 | 3.56 | 2.08 | 3.27 | 3.67 |
| | Large | 2.15 | 2.29 | 3.76 | 2.42 | 3.42 | 4.00 | 2.99 | 3.76 | 4.11 |
| FIFO | Small | 1.69 | 1.83 | 2.90 | 1.90 | 2.59 | 3.14 | 1.93 | 2.91 | 3.23 |
| | Medium | 2.00 | 2.13 | 3.50 | 2.21 | 3.10 | 3.69 | 2.31 | 3.53 | 3.77 |
| | Large | 2.37 | 2.54 | 3.96 | 2.66 | 3.57 | 4.13 | 3.21 | 4.00 | 4.23 |
| RAND | Small | 1.79 | 2.00 | 3.04 | 2.05 | 2.68 | 3.25 | 2.05 | 2.92 | 3.34 |
| | Medium | 1.98 | 2.39 | 3.58 | 2.37 | 3.18 | 3.78 | 2.46 | 3.49 | 3.88 |
| | Large | 2.42 | 2.30 | 4.09 | 2.71 | 3.69 | 4.27 | 3.24 | 3.95 | 4.36 |

# Case Study 23.1 (Cont)

❑ Effects: $\qquad \alpha_1 = y_{1...} - y_{....} = 2.74 - 2.90 = -0.16$

| Factor | Main Effects | | |
|--------|------|------|------|
| | Level | | |
| | 1 | 2 | 3 |
| A | -0.16 | 0.02 | 0.14 |
| D | -0.36 | 0.07 | 0.29 |
| P | -0.47 | -0.02 | 0.49 |
| M | -0.69 | -0.01 | 0.70 |

❑ Also

➢ Six two-factor interactions,

➢ Four three-factor interactions, and

➢ One four-factor interaction.

# Case Study 23.1: ANOVA Table

| Component | Sum of Squares | %Variation | DF | Mean Square |
|---|---|---|---|---|
| y | 730.01 | | 81 | |
| $\bar{y}$ | 681.21 | | 1 | |
| $y-\bar{y}_{....}$ | 48.80 | 100% | 80 | |
| Main Effects | 45.80 | 93.85% | 8 | 5.7 |
| A | 1.30 | | 2 | |
| D | 6.10 | | 2 | |
| P | 12.30 | | 2 | |
| M | 26.20 | | 2 | |
| First-order Interactions | 2.40 | 4.91% | 24 | 0.1 |
| AD | 0.07 | | 4 | |
| AP | 0.02 | | 4 | |
| AM | 0.03 | | 4 | |
| DP | 0.15 | | 4 | |
| DM | 1.96 | | 4 | |
| PM | 0.14 | | 4 | |
| Second-order Interactions | 0.48 | 0.98% | 32 | 0.015 |
| ADP | 0.05 | | 8 | |
| ADM | 0.13 | | 8 | |
| APM | 0.04 | | 8 | |
| DPM | 0.26 | | 8 | |
| Third-order Interaction (ADPM) | 0.07 | 0.14% | 16 | 0.004 |

# Case Study 23.1: Simplified model

❑ Most interactions except DM are small.

$$y_{ijkl} = \mu + \alpha_i + \beta_j + \gamma_k + \delta_l + \xi_{jl}$$

Where,

$$
\begin{aligned}
\mu &= \text{grand mean} \\
\alpha_i &= \text{Effect of A} \\
\beta_j &= \text{Effect of D} \\
\gamma_k &= \text{Effect of P} \\
\delta_l &= \text{Effect of M} \\
\xi_{jl} &= \text{Interaction between D and M.}
\end{aligned}
$$

# Case Study 23.1: Simplified Model (Cont)

❑  Interactions Between Deck Arrangement and Memory Pages

|   |   | M | | |
|---|---|---|---|---|
|   |   | 1 | 2 | 3 |
| D | 1 | 0.11 | -0.30 | 0.19 |
|   | 2 | -0.05 | 0.09 | -0.04 |
|   | 3 | -0.06 | 0.21 | -0.15 |

# Case Study 23.1: Error Computation

| Algor-ithm | Prog-ram | GROUP | | | FREQY | | | ALPHA | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | 24P | 20P | 16P | 24P | 20P | 16P | 24P | 20P | 16P |
| LRUV | Small | 0.18 | 0.08 | -0.07 | 0.11 | -0.04 | -0.02 | -0.05 | -0.04 | 0.01 |
| | Medium | -0.05 | -0.13 | 0.04 | 0.01 | 0.02 | 0.10 | -0.18 | 0.07 | 0.11 |
| | Large | -0.13 | -0.26 | 0.01 | -0.14 | 0.04 | 0.03 | 0.22 | 0.04 | 0.04 |
| FIFO | Small | 0.17 | 0.04 | 0.09 | 0.11 | -0.02 | -0.07 | -0.08 | -0.04 | -0.08 |
| | Medium | 0.05 | -0.10 | 0.07 | -0.02 | 0.04 | 0.05 | -0.13 | 0.14 | 0.02 |
| | Large | -0.10 | -0.20 | 0.02 | -0.00 | 0.00 | -0.03 | 0.25 | 0.09 | -0.02 |
| RAND | Small | 0.16 | 0.09 | -0.06 | 0.14 | -0.05 | -0.07 | -0.08 | -0.08 | -0.08 |
| | Medium | -0.10 | 0.04 | 0.04 | -0.02 | 0.00 | 0.01 | -0.11 | -0.02 | -0.02 |
| | Large | -0.17 | 0.44 | 0.04 | -0.15 | 0.00 | -0.01 | 0.16 | -0.08 | -0.01 |

# Case Study 23.1: Visual Test
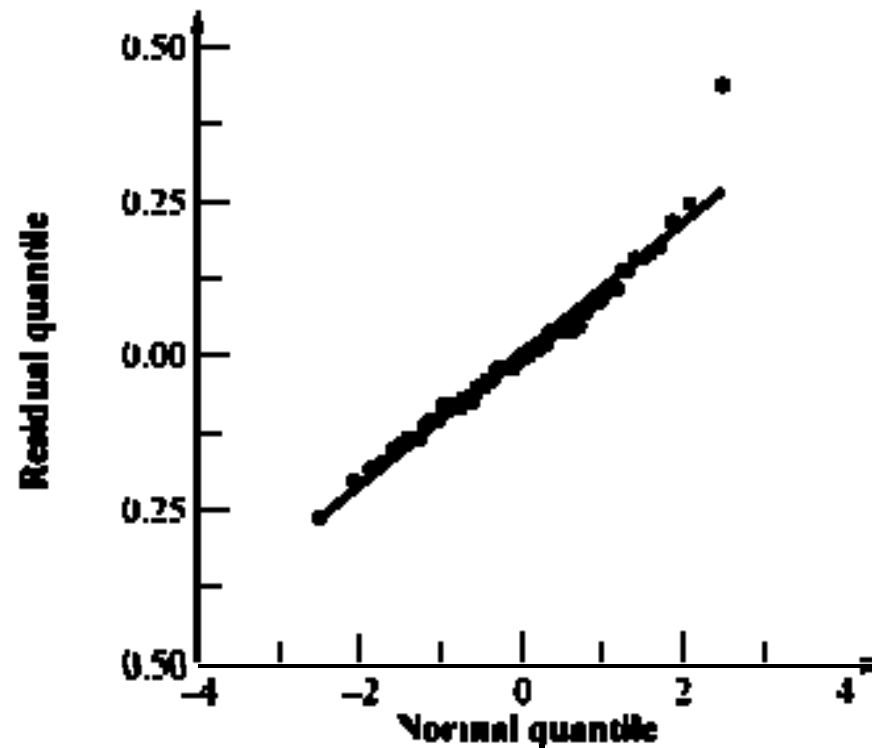


- ❏ Almost a straight line.
- ❏ Outlier was verified.

# Case Study 23.1: Final Model

$$\text{LPS} = 2.90 + A \begin{cases} -0.16 \\ +0.02 \\ +0.14 \end{cases} + D \begin{cases} -0.36 \\ +0.07 \\ +0.29 \end{cases} + P \begin{cases} -0.47 \\ -0.02 \\ +0.49 \end{cases} + M \begin{cases} -0.69 \\ -0.01 \\ +0.70 \end{cases}$$

Log (Pages Swapped)

Replacement Algorithm
1. LRUV
2. FIFO
3. RAND

1. GROUP
2. FREQY
3. ALPHA

1. Small
2. Medium
3. Large

1. 24P
2. 20P
3. 16P

$$+ D \overbrace{\begin{cases} -0.27 & 1.40 & 1.30 \\ -0.60 & 0.80 & -0.20 \\ 3.30 & -2.20 & -1.10 \end{cases}}^{M} \pm 0.12$$

Standard Error
= Stdv of sample mean
= Stdv of Error

# Observation Method

❑ To find the best combination.

❑ Example: Scheduler Design

❑ Three Classes of Jobs:

> ➢ Word processing

> ➢ Interactive data processing

> ➢ Background data processing

❑ Five Factors $2^{5-1}$ design

# Example 23.1: Measured Throughputs

| No. | A | B | C | D | E | $T_W$ | $T_I$ | $T_B$ |
|-----|-----|-----|-----|-----|-----|-------|-------|-------|
| 1 | -1 | -1 | -1 | -1 | 1 | 15.0 | 25.0 | 15.2 |
| 2 | 1 | -1 | -1 | -1 | -1 | 11.0 | 41.0 | 3.0 |
| 3 | -1 | 1 | -1 | -1 | -1 | 25.0 | 36.0 | 21.0 |
| 4 | 1 | 1 | -1 | -1 | 1 | 10.0 | 15.7 | 8.6 |
| 5 | -1 | -1 | 1 | -1 | -1 | 14.0 | 63.9 | 7.5 |
| 6 | 1 | -1 | 1 | -1 | 1 | 10.0 | 13.2 | 7.5 |
| 7 | -1 | 1 | 1 | -1 | 1 | 28.0 | 36.3 | 20.2 |
| 8 | 1 | 1 | 1 | -1 | -1 | 11.0 | 23.0 | 3.0 |
| 9 | -1 | -1 | -1 | 1 | -1 | 14.0 | 66.1 | 6.4 |
| 10 | 1 | -1 | -1 | 1 | 1 | 10.0 | 9.1 | 8.4 |
| 11 | -1 | 1 | -1 | 1 | 1 | 27.0 | 34.6 | 15.7 |
| 12 | 1 | 1 | -1 | 1 | -1 | 11.0 | 23.0 | 3.0 |
| 13 | -1 | -1 | 1 | 1 | 1 | 14.0 | 26.0 | 12.0 |
| 14 | 1 | -1 | 1 | 1 | -1 | 11.0 | 38.0 | 2.0 |
| 15 | -1 | 1 | 1 | 1 | -1 | 25.0 | 35.0 | 17.2 |
| 16 | 1 | 1 | 1 | 1 | 1 | 11.0 | 22.0 | 2.0 |

# Example 23.1: Conclusions

To get high throughput for word processing jobs,:

1. There should not be any preemption (A=-1)

2. The time slice should be large (B=1)

3. The fairness should be on (E=1)

4. The settings for queue assignment and re-queueing do not matter.

# Ranking Method

❑ Sort the experiments.

| No. | A | B | C | D | E | $T_W$ | $T_I$ | $T_B$ |
|---|---|---|---|---|---|---|---|---|
| 7 | -1 | 1 | 1 | -1 | 1 | 28.0 | 36.3 | 20.2 |
| 11 | -1 | 1 | -1 | 1 | 1 | 27.0 | 34.6 | 15.7 |
| 15 | -1 | 1 | 1 | 1 | -1 | 25.0 | 35.0 | 17.2 |
| 3 | -1 | 1 | -1 | -1 | -1 | 25.0 | 36.0 | 21.0 |
| 1 | -1 | -1 | -1 | -1 | 1 | 15.0 | 25.0 | 15.2 |
| 5 | -1 | -1 | 1 | -1 | -1 | 14.0 | 63.9 | 7.5 |
| 9 | -1 | -1 | -1 | 1 | -1 | 14.0 | 66.1 | 6.4 |
| 13 | -1 | -1 | 1 | 1 | 1 | 14.0 | 26.0 | 12.0 |
| 2 | 1 | -1 | -1 | -1 | -1 | 11.0 | 41.0 | 3.0 |
| 8 | 1 | 1 | 1 | -1 | -1 | 11.0 | 23.0 | 3.0 |
| 12 | 1 | 1 | -1 | 1 | -1 | 11.0 | 23.0 | 3.0 |
| 14 | 1 | -1 | 1 | 1 | -1 | 11.0 | 38.0 | 2.0 |
| 16 | 1 | 1 | 1 | 1 | 1 | 11.0 | 22.0 | 2.0 |
| 6 | 1 | -1 | 1 | -1 | 1 | 10.0 | 13.2 | 7.5 |
| 4 | 1 | 1 | -1 | -1 | 1 | 10.0 | 15.7 | 8.6 |
| 10 | 1 | -1 | -1 | 1 | 1 | 10.0 | 9.1 | 8.4 |

# Example 23.2: Conclusions

1. A=-1 (no preemption) is good for word processing jobs and also that A=1 is bad.

2. B=1 (large time slice) is good for such jobs. No strong negative comment can be made about B=-1.

3. Given a choice C should be chosen at 1, that is, there should be two queues.

4. The effect of E is not clear.

5. If top rows chosen, then E=1 is a good choice.

# Range Method

❑ Range = Maximum-Minimum

❑ Factors with large range are important.

| Factor | Level | | | Range of of Averages |
|---|---|---|---|---|
| | 1 | 2 | 3 | |
| Replacement Algorithm | 2056 | 2986 | 3781 | 1725 |
| Deck Arrangement | 1584 | 2913 | 4326 | 2742 |
| Problem Program | 592 | 2047 | 6185 | 5593 |
| Memory Size | 305 | 2006 | 6512 | 6207 |

❑ Memory size is the most influential factor.

❑ Problem program, deck arrangement, and replacement algorithm are next in order.

# Summary
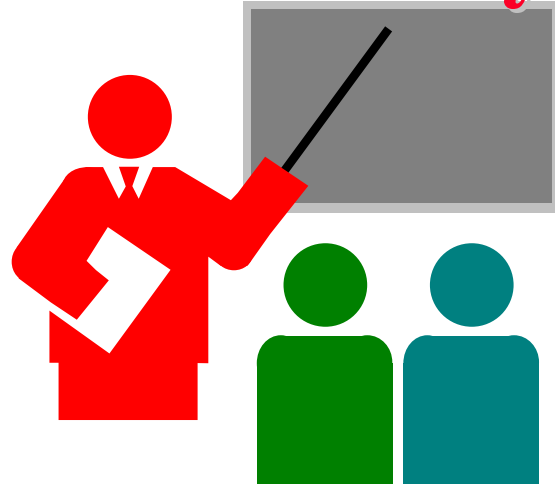


❑ A general k factor design can have k main effects, two factor interactions, three factor interactions, and so on.

❑ Information Methods:

  ➢ Observation: Find the highest or lowest response

  ➢ Ranking: Sort all responses

  ➢ Range: Largest - smallest average response

# Exercise 23.1

Using the observation method on data of Table 23.8, find the factor levels that maximize the throughput for interactive jobs ($T_I$). Repeat the problem for background jobs ($T_B$).

# Exercise 23.2

Repeat Exercise 23.1 using ranking method.

# Homework

❑ Analyze the following results using observation and ranking methods.

| No. | A | B | C | D | E | $T$ |
|-----|-----|-----|-----|-----|-----|------|
| 1 | -1 | -1 | -1 | -1 | 1 | 13.2 |
| 2 | 1 | -1 | -1 | -1 | -1 | 4.0 |
| 3 | -1 | 1 | -1 | -1 | -1 | 22.0 |
| 4 | 1 | 1 | -1 | -1 | 1 | 9.6 |
| 5 | -1 | -1 | 1 | -1 | -1 | 6.5 |
| 6 | 1 | -1 | 1 | -1 | 1 | 8.5 |
| 7 | -1 | 1 | 1 | -1 | 1 | 21.2 |
| 8 | 1 | 1 | 1 | -1 | -1 | 2.0 |
| 9 | -1 | -1 | -1 | 1 | -1 | 7.4 |
| 10 | 1 | -1 | -1 | 1 | 1 | 7.4 |
| 11 | -1 | 1 | -1 | 1 | 1 | 14.7 |
| 12 | 1 | 1 | -1 | 1 | -1 | 4.0 |
| 13 | -1 | -1 | 1 | 1 | 1 | 13.0 |
| 14 | 1 | -1 | 1 | 1 | -1 | 3.0 |
| 15 | -1 | 1 | 1 | 1 | -1 | 18.2 |
| 16 | 1 | 1 | 1 | 1 | 1 | 3.0 |