2^k Factorial Designs and $2^k r$ Designs with Replications

Dr. John Mellor-Crummey

Department of Computer Science
Rice University

johnmc@cs.rice.edu

Goals for Today

Understand

- 2^k factorial designs
  - —sign table method
  - —properties
  - —analysis
  - —allocating variation
- $2^k r$ designs with replications
  - —error estimation
  - —confidence intervals for effects
- Multiplicative models

2^k Factorial Designs

- **Determine effects of k factors, each at two levels**
- **Total of 2^k experiments required**
- **Total of 2^k effects**
  - **k main effects**
  - $\binom{k}{2}$ **two-factor interactions**
  - $\binom{k}{3}$ **three-factor interactions**
  - $\binom{k}{j}$ **j factor interactions, $j \leq k$**

Example: a 2^3 Design

Evaluate impact of memory, cache size, # processors

Factor	Level -1	Level 1
Memory Size, A	1GB	2GB
Cache Size, B	4MB	6MB
# processors, C	1	2

Model performance using a non-linear equation in x_A and x_B

$$y = q_0 + q_A x_A + q_B x_B + q_{AB} x_A x_B + q_{AC} x_A x_C + q_{BC} x_B x_C + q_{ABC} x_A x_B x_C$$

Solve using sign table method

Calculating Effects with a Sign Table

I	A	B	C	AB	AC	BC	ABC	y
1	-1	-1	-1	1	1	1	-1	14
1	1	-1	-1	-1	-1	1	1	22
1	-1	1	-1	-1	1	-1	1	10
1	1	1	-1	1	-1	-1	-1	34
1	-1	-1	1	1	-1	-1	1	46
1	1	-1	1	-1	1	-1	-1	58
1	-1	1	1	-1	-1	1	-1	50
1	1	1	1	1	1	1	1	86
320	80	40	160	40	16	24	8	total
40	10	5	20	5	2	3	1	total/8
q_0	q_A	q_B	q_C	q_{AB}	q_{AC}	q_{BC}	q_{ABC}	

Calculating Effects

40	10	5	20	5	2	3	1
q_0	q_A	q_B	q_C	q_{AB}	q_{AC}	q_{BC}	q_{ABC}

$$SST = 2^3(q_A^2 + q_B^2 + q_C^2 + q_{AB}^2 + q_{AC}^2 + q_{BC}^2 + q_{ABC}^2)$$

$$SST = 2^3(10^2 + 5^2 + 20^2 + 5^2 + 2^2 + 3^2 + 1^2) = 4512$$

Effects

- **SSA/SST = $2^3(10^2)/4512$ = .18 = 18%**
- **SSB/SST = $2^3(5^2)/4512$ = .04 = 4%**
- **SSC/SST = $2^3(20^2)/4512$ = .71= 71%**
- **SSAB/SST = $2^3(5^2)/4512$ = .04 = 4%**
- **SSAC/SST = $2^3(2^2)/4512$ = .01 = 1%**
- **SSBC/SST = $2^3(3^2)/4512$ = .02 = 2%**
- **SSABC/SST = $2^3(1^2)/4512$ = .00 = 0%**

$2^k r$ Factorial Designs with Replications

$2^k r$ Factorial Designs

- **Cannot estimate errors with 2^k factorial design**
  - **—no experiment is repeated**
- **To quantify experimental errors**
  - **—repeat measurements with same factor combinations**
  - **—analyze using sign table**
- **$2^k r$ design**
  - **—r replications of 2^k experiments (each of 2^k factor combinations)**
- **Model**

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

(e = experimental error)

Computing Effects for a 2^2r Design

I	A	B	AB	$\bar{y}_{i.}$
1	-1	-1	1	15
1	1	-1	-1	48
1	-1	1	-1	24
1	1	1	1	77
164	86	38	20	total
41	21.5	9.5	5	total/4
q_0	q_A	q_B	q_{AB}	

y_{ij}

15	18	12
45	48	51
25	28	19
75	75	81

- Factor combination = sign table row
- For each factor combination i, r replicants (here, r = 3)
- $\bar{y}_{i.}$ = mean observation for i^{th} factor combination

$$\bar{y}_{i.} = \frac{1}{r}\sum_j y_{ij}$$

- Solve for q_i coefficients using sign table and $\bar{y}_{i.}$ values

Estimating Experimental Errors for a 2^2r Design

- **Model predicts following value for factor combination i**

$$\hat{y}_i = q_0 + q_A x_{Ai} + q_B x_{Bi} + q_{AB} x_{Ai} x_{Bi}$$

- **Errors**

$$\begin{aligned} e_{ij} &= y_{ij} - \hat{y}_i \\ &= y_{ij} - q_0 + q_A x_{Ai} + q_B x_{Bi} + q_{AB} x_{Ai} x_{Bi} \\ &= y_{ij} - \bar{y}_{i.} \end{aligned}$$

- **Total error across row is 0 by design**
- **Corollary: total error (across all rows) is also 0**

Sum of Squared Errors for a $2^2 3$ Design

I	A	B	AB	$\bar{y}_{i.}$
1	-1	-1	1	15
1	1	-1	-1	48
1	-1	1	-1	24
1	1	1	1	77

y_{ij}

15	18	12
45	48	51
25	28	19
75	75	81

e_{ij}

0	3	-3
-3	0	3
1	4	-5
-2	-2	4

$$SSE = \sum_{i=1}^{2^2} \sum_{j=1}^{r} e_{ij}$$

$$SSE = 0^2 + 3^2 + (-3)^2 + (-3)^2 + 0^2 + 3^2 +$$

$$1^2 + 4^2 + (-5)^2 + (-2)^2 + (-2)^2 + 4^2 = 102$$

Mean Response for a 2^2r Design

- **Model for 2^2r design**

$$y_{ij} = q_0 + q_A x_{Ai} + q_B x_{Bi} + q_{AB} x_{Ai} x_{Bi} + e_{ij}$$

- **Sum across all 2^2r observations =**

$$\sum_{i,j} y_{ij} = \sum_{i,j} q_0 + \sum_{i,j} q_A x_{Ai} + \sum_{i,j} q_B x_{Bi} + \sum_{i,j} q_{AB} x_{Ai} x_{Bi} + \sum_{i,j} e_{ij}$$

- **Since sum of x's, their products and errors are all 0**

$$\sum_{i,j} y_{ij} = \sum_{i,j} q_0 = 2^2 r q_0$$

mean response $\bar{y}_{..} = \frac{1}{2^2 r} \sum_{i,j} y_{ij} = \frac{1}{2^2 r} 2^2 r q_0 = q_0$

SST and SSE for a 2^2r Design

$$SST = \sum_{i,j}\left(y_{ij} - \bar{y}_{..}\right)^2 = \sum_{i,j} y_{ij}^{\ 2} - \boxed{\sum_{i,j} \bar{y}_{..}^{\ 2}}$$

$$\boxed{SS0 = \sum_{i,j} \bar{y}_{..}^2 = \sum_{i,j} q_0^2}$$

$$SSY = \sum_{i,j} y_{ij}^2 = \boxed{\sum_{i,j} q_0^2} + \boxed{\sum_{i,j} q_A^2 x_{Ai}^2} + \boxed{\sum_{i,j} q_B^2 x_{Bi}^2} +$$

$$\boxed{\sum_{i,j} q_{AB}^2 x_{Ai}^2 x_{Bi}^2} + \boxed{\sum_{i,j} e_{ij}^2} +$$

product terms
(0 by design of sign table)

$$SSY = \boxed{SS0} + \boxed{SSA} + \boxed{SSB} + \boxed{SSAB} + \boxed{SSE}$$

$$SST = SSY - SS0 = SSA + SSB + SSAB + SSE$$

$$SSE = SSY - SS0 - SSA - SSB - SSAB$$

$$SSE = SSY - 2^2 r(q_0^2 + q_A^2 + q_B^2 + q_{AB}^2)$$

Allocating Variation for a 2^2r Design

- Fraction of variation attributed to A = SSA/SST
- Fraction of variation attributed to B = SSB/SST
- Fraction of variation attributed to interaction of A & B = SSAB/SST
- Fraction of variation attributed to error = SSE/SST

Variance for q_0

- **Effects computed from a sample are random variables**
- **Confidence intervals for effects**
  - **—compute from variance of sample estimates**
- **If errors ~** $N(0,\sigma_e^2)$**, model implies that y ~** $N(\bar{y}_{..},\sigma_e^2)$
- **Consider effect q_0**

$$q_0 = \frac{1}{2^2 r}\sum_{i,j} y_{ij}$$

  - **—q_0 = linear combination of normally distributed variables**
  - **—therefore, q_0 is normally distributed as well**
- **Variance of q_0**

$$\mathrm{var}(q_0) = \frac{\sigma_e^2}{2^2 r}$$

Confidence Intervals for Effects in 2^2r Designs

- **Estimate variance of errors from SSE**

$$s_e^2 = \frac{SSE}{\boxed{2^2(r-1)}} = \textbf{MSE}$$

- **Denominator = $2^2(r-1)$ = DOF = # independent terms in SSE**
  - **—error terms for r replications all add to 0**
  - **—only r-1 are independent**
- **Std dev of effects** $s_{q_0} = s_{q_A} = s_{q_B} = s_{q_{AB}} = \dfrac{s_e}{\sqrt{2^2 r}}$
- **Confidence intervals for effects** $q_i \pm t_{[1-\alpha/2;2^2(r-1)]} s_{q_i}$
- **Effect is significant if its confidence interval does not contain 0**

Returning to our Example …

Memory/Cache Study

- **Standard deviation of errors**

$$s_e = \sqrt{\frac{SSE}{2^2(r-1)}} = \sqrt{\frac{102}{8}} = \sqrt{12.75} = 3.57$$

41	21.5	9.5	5
q_0	q_A	q_B	q_{AB}

- **Standard deviation of effects**

$$s_{q_i} = s_e / \sqrt{2^2 r} = 3.57 / \sqrt{12} = 1.03$$

- **Confidence intervals for effects**
  - **90% confidence:** $t_{[1-\alpha/2;2^2(r-1)]} = t_{[.95;8]} = 1.86$

$$q_i \pm t_{[1-\alpha/2;2^2(r-1)]} s_{q_i} = q_i \pm (1.86)(1.03) = q_i \pm 1.92$$

q_0= (30.08,42.91), q_A= (19.58, 23.41), q_B= (7.58, 11.41), q_{AB}= (3.08, 6.91)

- **No confidence intervals include 0 ⇒ all significant**

Assumptions

- **Model errors are statistically independent**
- **Errors are additive**
- **Errors are normally distributed**
- **Errors have constant standard deviation** σ_e
- **Effects of factors are additive**
- **observations are independent and normally distributed with constant variance**

Visual Tests for Verifying Assumptions

- **Independent errors**
  - —scatter plot of residuals vs. predicted responses
    - should not have any trend
    - magnitude residuals < magnitude responses/10 ⇒ ignore trends
  - —plot residuals as a function of experiment number
    - trend up or down implies other factors or side effects
- **Normally distributed errors**
  - —normal quantile-quantile plot of errors should be linear
- **Constant standard deviation of errors**
  - —scatter plot of y for various levels of factors
  - —if spread at one level significantly different than at another
    - need transformation
- **What to do**
  - —if any test fails or y_{max}/y_{min} is large, investigate multiplicative model

Multiplicative Models for 2^2r Experiments

- **Additive model**

$$y_{ij} = q_0 + q_A x_{Ai} + q_B x_{Bi} + q_{AB} x_{Ai} x_{Bi} + e_{ij}$$

- **Not applicable if effects don't add**
  - —example: execution time of workloads
    - factor A: i^{th} processor speed = v_i instructions/sec
    - factor B: j^{th} workload size = w_j instructions
  - —the two effects multiply: execution time $y_{ij} = v_i \times w_j$
- **Use logarithm to produce additive model**
  - —$\log(y_{ij}) = \log(v_i) + \log(w_j)$
- **Better model** $y'_{ij} = q_0 + q_A x_{Ai} + q_B x_{Bi} + q_{AB} x_{Ai} x_{Bi} + e_{ij}$

  where $y'_{ij} = \log(y_{ij})$

Interpreting a Multiplicative Model

$$\log(y) = q_0 + q_A x_A + q_B x_B + q_{AB} x_A x_B + e$$

$$y = 10^{q_0} 10^{q_A x_A} 10^{q_B x_B} 10^{q_{AB} x_A x_B} 10^{e}$$

$$u_A = 10^{q_A}, u_B = 10^{q_B}, u_{AB} = 10^{q_{AB}},$$

- **u_A is ratio of MIPS rating of two processors**
- **u_B ratio of size of the two workloads**
- **Antilog of additive mean q_0 yields geometric mean**

$$\dot{y} = 10^{q_0} = (y_1 y_2 \ldots y_n)^{1/n}$$ **where $n = 2^2 r$**

Example: Execution Times

- **Analysis using an additive model**

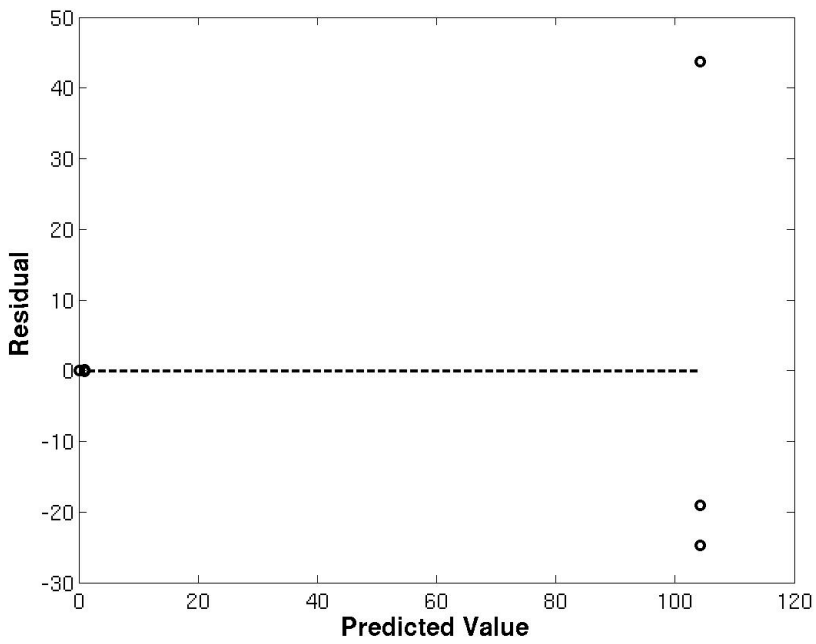
factor A = proc speed; factor B = instructions

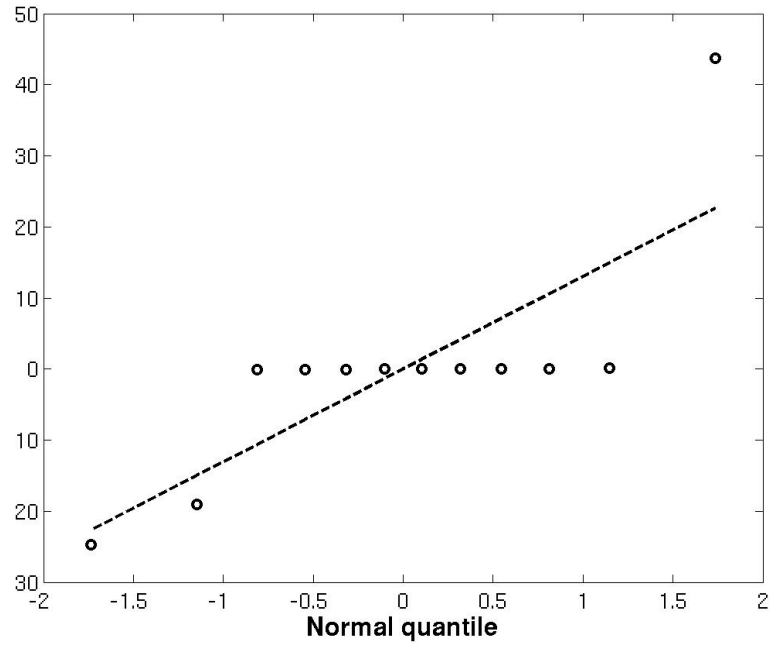
I	A	B	AB	$\bar{y}_{i.}$
1	-1	-1	1	104.170
1	1	-1	-1	1.003
1	-1	1	-1	1.003
1	1	1	1	0.013
106.19	-104.15	-104.15	102.17	total
26.55	-26.04	-26.04	25.54	total/4
q_0	q_A	q_B	q_{AB}	

y_{ij}

85.10	79.50	147.90
.891	1.047	1.072
.955	.933	1.122
.0148	.0126	.0118

Visual Tests for Additive Model





Why is This Model Invalid?

- **Physical consideration**
  - —workload & processor effects don't add
- **Large range for y**
  - — y_{max}/y_{min} = 147.90/.0118 = 12,534 ⇒ log transform
    - – taking arithmetic mean of 104.17 and .013 is inappropriate
- **Residuals are not small compared to response**
- **Spread of residuals is larger at larger value of response**
  - —suggests log transformation

Analysis using a Multiplicative Model

Data after log transformation

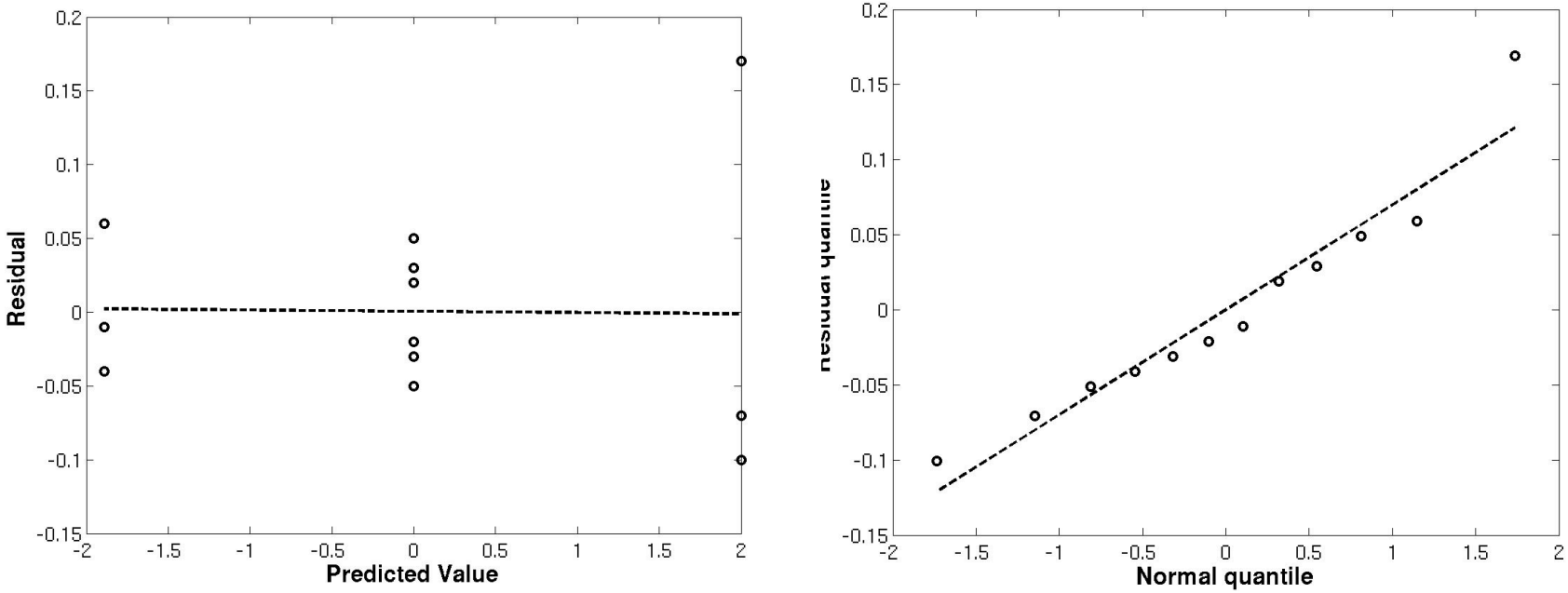
factor A = proc speed; factor B = instructions

I	A	B	AB	$\bar{y}_{i.}$
1	-1	-1	1	2.00
1	1	-1	-1	0.00
1	-1	1	-1	0.00
1	1	1	1	-1.89
.11	-3.89	-3.89	.11	total
.03	-0.97	-0.97	.03	total/4
q_0	q_A	q_B	q_{AB}	

y_{ij}

1.93	1.90	2.17
-.05	.02	.03
-.02	-.03	.05
-1.83	-1.90	-1.93

Visual Tests for Multiplicative Model



- **Conclusion: Multiplicative model is better than additive model**

Variation Explained by the Models

	Additive			Multiplicative		
Factor	Effect	%var	Conf Interval	Effect	%var	Conf Interval
I	26.55		(16.35,36.74)	.03		(-.02,.07)*
A	-26.04	30.1%	(-36.23,-15.84)	-.97	49.9%	(-1.02,-.93)
B	-26.04	30.1%	(-36.23,-15.84)	-.97	49.9%	(-1.02,-.93)
AB	25.54	29.0%	(15.35,35.74)	.03	0.0%	(-.02,.07)*
e		10.8%			.2%	

* = not significant

- With multiplicative model: interaction almost 0
- Little unexplained variation: .2%