$2^k r$ Factorial Designs

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/cse567-11/

- Computation of Effects
- Estimation of Experimental Errors
- Allocation of Variation
- Confidence Intervals for Effects
- Confidence Intervals for Predicted Responses
- Visual Tests for Verifying the assumptions
- Multiplicative Models

$2^k r$ Factorial Designs

- *r* replications of 2^k Experiments
  $\Rightarrow$ $2^k r$ observations.
  $\Rightarrow$ Allows estimation of experimental errors.
- Model:

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

- e = Experimental error

Computation of Effects

Simply use means of r measurements

I	A	B	A B	y	Mean $\bar{y}$
1	-1	-1	1	(15, 18, 12)	15
1	1	-1	-1	(45, 48, 51)	48
1	-1	1	-1	(25, 28, 19)	24
1	1	1	1	(75, 75, 81)	77
164	86	38	20		total
41	21.5	9.5	5		total/4

- Effects: q_0= 41, q_A= 21.5, q_B= 9.5, q_{AB}= 5.

Estimation of Experimental Errors

- Estimated Response:

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

Experimental Error = Estimated - Measured

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

$\sum_{i,j} e_{ij} = 0$

- Sum of Squared Errors: $\mathrm{SSE} = \sum_{i=1}^{2^2} \sum_{j=1}^{r} e_{ij}^2$

Experimental Errors: Example

- Estimated Response:

$$\hat{y}_1 = q_0 - q_A - q_B + q_{AB} = 41 - 21.5 - 9.5 + 5 = 15$$

- Experimental errors:

$$e_{11} = y_{11} - \hat{y}_1 = 15 - 15 = 0$$

i	Effect I	Effect A	Effect B	Effect A B	Estimated Response	Measured Responses			Errors		
	41	21.5	9.5	5	$\hat{y}_i$	y_{i1}	y_{i2}	y_{i3}	e_{i1}	e_{i2}	e_{i3}
1	1	-1	-1	1	15	15	18	12	0	3	-3
2	1	1	-1	-1	48	45	48	51	-3	0	3
3	1	-1	1	-1	24	25	28	19	1	4	-5
4	1	1	1	1	77	75	75	81	-2	-2	4

Allocation of Variation

- Total variation or total sum of squares:

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

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

$$\begin{array}{ccccccccc} \sum_{i,j}(y_{ij} - \bar{y}_{..})^2 & = & 2^2 r q_A^2 & + & 2^2 r q_B^2 & + & 2^2 r q_{AB}^2 & + & \sum_{i,j} e_{ij}^2 \\ \text{SST} & = & \text{SSA} & + & \text{SSB} & + & \text{SSAB} & + & \text{SSE} \end{array}$$

Example 18.3: Memory-Cache Study

$$
\begin{aligned}
\text{SSY} &= 15^2 + 18^2 + 12^2 + 45^2 + \cdots + 75^2 + 75^2 + 81^2 \\
&= 27204 \\
\text{SS0} &= 2^2 r q_0^2 = 12 \times 41^2 = 20172 \\
\text{SSA} &= 2^2 r q_A^2 = 12 \times (21.5)^2 = 5547 \\
\text{SSB} &= 2^2 r q_B^2 = 12 \times (9.5)^2 = 1083 \\
\text{SSAB} &= 2^2 r q_{AB}^2 = 12 \times 5^2 = 300 \\
\text{SSE} &= 27204 - 2^2 \times 3(41^2 + 21.5^2 + 9.5^2 + 5^2) \\
&= 102 \\
\text{SST} &= \text{SSY} - \text{SS0} \\
&= 27204 - 20172 = 7032
\end{aligned}
$$

Example 18.3 (Cont)

$$\begin{aligned} &\text{SSA} + \text{SSB} + \text{SSAB} + \text{SSE} \\ &= 5547 + 1083 + 300 + 102 \\ &= 7032 = \text{SST} \end{aligned}$$

Factor A explains 5547/7032 or 78.88%

Factor B explains 15.40%

Interaction AB explains 4.27%

1.45% is unexplained and is attributed to errors.

Confidence Intervals For Effects

- Effects are random variables.
- Errors $\sim N(0,\sigma_e) \Rightarrow y \sim N(\bar{y}_{\cdot\cdot}, \sigma_e)$

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

- q_0 = Linear combination of normal variates
  $\Rightarrow q_0$ is normal with variance $\sigma_e^2/(2^2 r)$

Variance of errors:

$$s_e^2 = \frac{1}{2^2(r-1)}\sum_{ij} e_{ij}^2 = \frac{\text{SSE}}{2^2(r-1)} \triangleq \text{MSE}$$

- Denominator = $2^2(r-1)$ = # of independent terms in SSE
  $\Rightarrow$ SSE has $2^2(r-1)$ degrees of freedom.
  Estimated variance of q_0: $s_{q_0}^2 = s_e^2/(2^2 r)$

Confidence Intervals For Effects (Cont)

- Similarly,

$$s_{q_A} = s_{q_B} = s_{q_{AB}} = \frac{s_e}{\sqrt{2^2 r}}$$

- Confidence intervals (CI) for the effects:

$$q_i \mp t_{[1-\alpha/2;2^2(r-1)]} s_{q_i}$$

- CI does not include a zero $\Rightarrow$ significant

Example 18.4

- For Memory-cache study: Standard deviation of errors:

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

- Standard deviation of effects:

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

- For 90% Confidence: $t_{[0.95,8]} = 1.86$

- Confidence intervals: $q_i \mp (1.86)(1.03) = q_i \mp 1.92$

q_0= (39.08, 42.91)

q_A=(19.58, 23.41)

q_B=(7.58, 11.41)

q_{AB}= (3.08, 6.91)

- No zero crossing ⇒ All effects are significant.

Confidence Intervals for Contrasts

- Contrast $\underline{\Delta}$ Linear combination with $\sum$ coefficients = 0
- Variance of $\sum h_i q_i$ $\quad s^2_{\sum h_i q_i} = \frac{s_e^2 \sum h_i^2}{2^2 r}$
- For 100(1-α)% confidence interval, use $t_{[1-\alpha/2;\, 2^2(r-1)]}$.

Example 18.5

Memory-cache study

$u = q_A + q_B - 2q_{AB}$

Coefficients= 0, 1, 1, and -2 ⇒ Contrast

$$\text{Mean } \bar{u} = 21.5 + 9.5 - 2 \times 5 = 11$$

$$\text{Variance } s_u^2 = \frac{s_e^2 \times 6}{2^2 \times 3} = 6.375$$

$$\text{Standard deviation } s_u = \sqrt{6.375} = 2.52$$

$t_{[0.95;8]}=1.86$

90% Confidence interval for u:

$$\bar{u} \mp t s_u = 11 \mp 1.86 \times 2.52 = (6.31,\ 15.69)$$

Conf. Interval For Predicted Responses

- Mean response $\hat{y}$:

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

- The standard deviation of the mean of m responses:

$$s_{\hat{y}_m} = s_e \left(\frac{1}{n_{\text{eff}}} + \frac{1}{m} \right)^{1/2}$$

$$\begin{aligned} n_{\text{eff}} &= \text{Effective deg of freedom} \\ &= \frac{\text{Total number of runs}}{1 + \text{Sum of DFs of params used in } \hat{y}} \\ &= \frac{2^2 r}{5} \end{aligned}$$

Conf. Interval for Predicted Responses (Cont)

100(1-α)% confidence interval:

$$\hat{y} \mp t_{[1-\alpha/2;2^2(r-1)]} s_{\hat{y}_m}$$

- A single run (m=1): $s_{\hat{y}_1} = s_e \left(\frac{5}{2^2 r} + 1 \right)^{1/2}$
- Population mean (m=∞): $s_{\hat{y}} = s_e \left(\frac{5}{2^2 r} \right)^{1/2}$

Example 18.6: Memory-cache Study

- For x_A= -1 and x_B = -1:
- A single confirmation experiment:

$$\begin{aligned}\hat{y}_1 &= q_0 - q_A - q_B + q_{AB}\\ &= 41 - 21.5 - 9.5 + 5 = 15\end{aligned}$$

- Standard deviation of the prediction:

$$s_{\hat{y}_1} = s_e\left(\frac{5}{2^2r}+1\right)^{1/2} = 3.57\sqrt{\frac{5}{12}+1} = 4.25$$

- Using $t_{[0.95;8]}$ = 1.86, the 90% confidence interval is:

$$15 \mp 1.86 \times 4.25 = (8.09,\ 22.91)$$

Example 18.6 (Cont)

- Mean response for 5 experiments in future:

$$s_{\hat{y}_1} = s_e \left(\frac{5}{2^2 r} + \frac{1}{m} \right)^{1/2}$$
$$= 3.57 \sqrt{\frac{5}{12} + \frac{1}{5}} = 2.20$$

- The 90% confidence interval is:

$$15 \mp 1.86 \times 2.20 = (10.91,\ 19.09)$$

Example 18.6 (Cont)

- Mean response for a large number of experiments in future:

$$s_{\hat{y}} = s_e \left(\frac{5}{2^2 r}\right)^{1/2} = 3.57\sqrt{\frac{5}{12}} = 2.30$$

- The 90% confidence interval is:

$$15 \mp 1.86 \times 2.30 = (10.72,\ 19.28)$$

- Current mean response: Not for future. Use contrasts formula.

$$s_{\hat{y}} = \sqrt{\frac{s_e \sum h_i^2}{2^2 r}} = \sqrt{\frac{12.75 \times 4}{12}} = 2.06$$

- 90% confidence interval:

$$15 \mp 1.86 \times 2.06 = (11.17,\ 18.83)$$

Homework 18A

Updated Exercise 18.1: The following table lists measured CPU times for two processors on two workloads. Each experiment was repeated three times. Determine the effects.

Table 18.12 $2^2 3$ Experimental Design Exercise

Workload	Processor	
	A	B
I	(41.16, 39.02, 42.56)	(65.17, 69.25, 64.23)
J	(53.50, 55.50, 50.50)	(50.08, 48.98, 47.10)

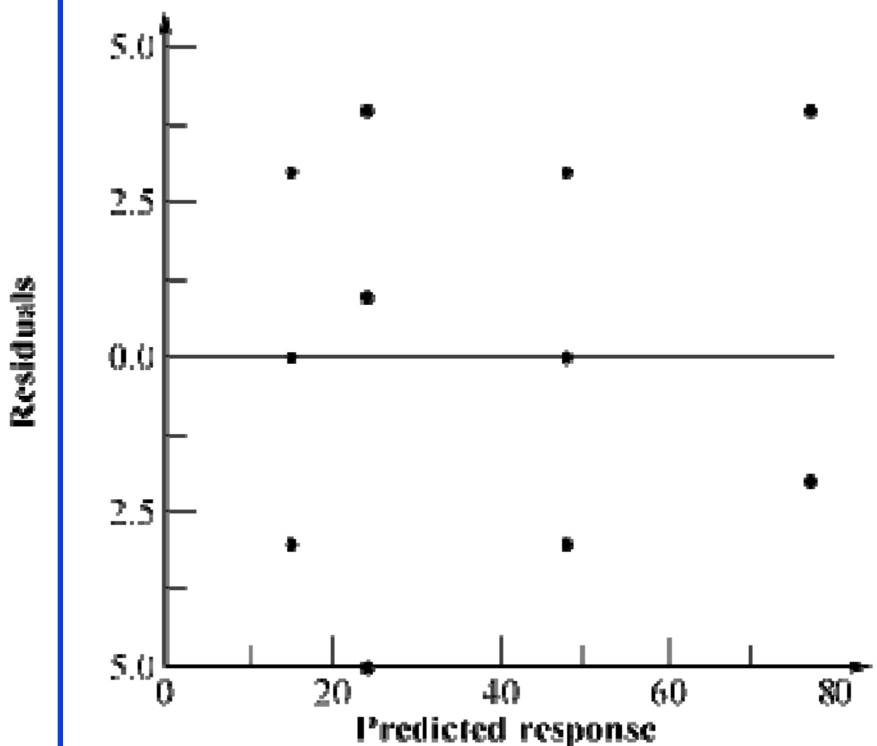
Assumptions

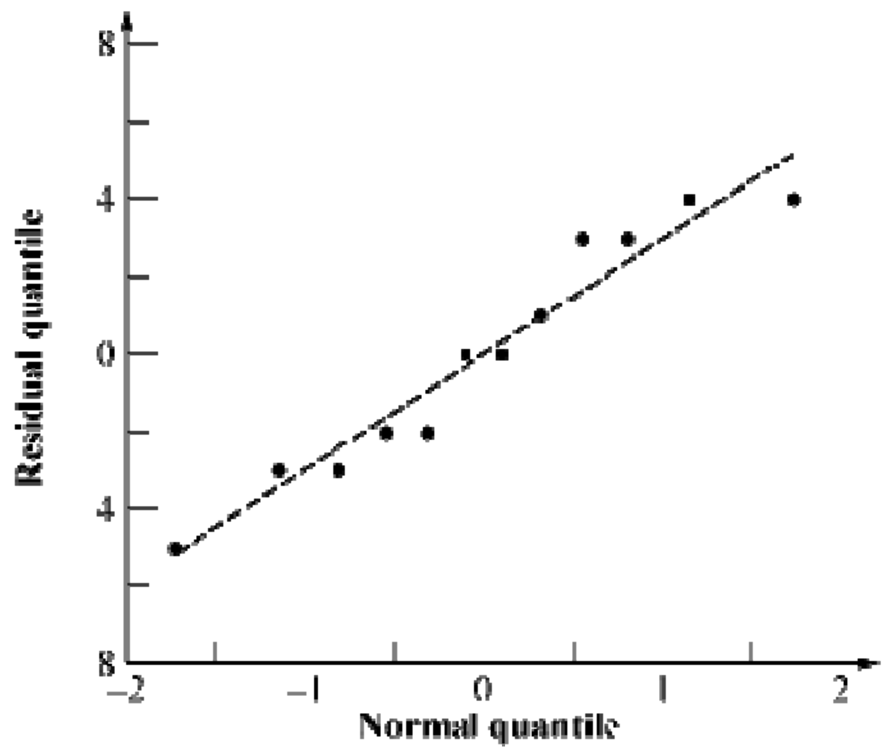
1. Errors are statistically independent.
2. Errors are additive.
3. Errors are normally distributed.
4. Errors have a constant standard deviation σ_e.
5. Effects of factors are additive
   $\Rightarrow$ observations are independent and normally distributed with constant variance.

Visual Tests

1. **Independent Errors**:
   - Scatter plot of residuals versus the predicted response $\hat{y}_i$
   - Magnitude of residuals < Magnitude of responses/10
     $\Rightarrow$ Ignore trends
   - Plot the residuals as a function of the experiment number
   - Trend up or down $\Rightarrow$ other factors or side effects
2. **Normally distributed errors**:
   Normal quantile-quantile plot of errors
3. **Constant Standard Deviation of Errors**:
   Scatter plot of y for various levels of the factor
   Spread at one level significantly different than that at other
   $\Rightarrow$ Need transformation

Example 18.7: Memory-cache





Multiplicative Models

- Additive model:

  $$y_{ij} = q_0 + q_A x_A + q_B x_B + q_{AB} x_A x_B + e_{ij}$$

- Not valid if effects do not add.
  E.g., execution time of workloads.
  *i*th processor speed= v_i instructions/second.
  *j*th workload Size= w_j instructions
- The two effects multiply. Logarithm ⇒ additive model:

  $$\text{Execution Time } y_{ij} = v_i \times w_j$$

  $$\log(y_{ij}) = \log(v_i) + \log(w_j)$$

- Correct Model:

  $$y'_{ij} = q_0 + q_A x_A + q_B x_B + q_{AB} x_A x_B + e_{ij}$$

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

Multiplicative Model (Cont)

- Taking an antilog of effects:

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

- u_A= ratio of MIPS rating of the two processors
- u_B= ratio of the size of the two workloads.
- Antilog of additive mean $q_0 \Rightarrow$ geometric mean

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

Example 18.8: Execution Times

Analysis Using an Additive Model

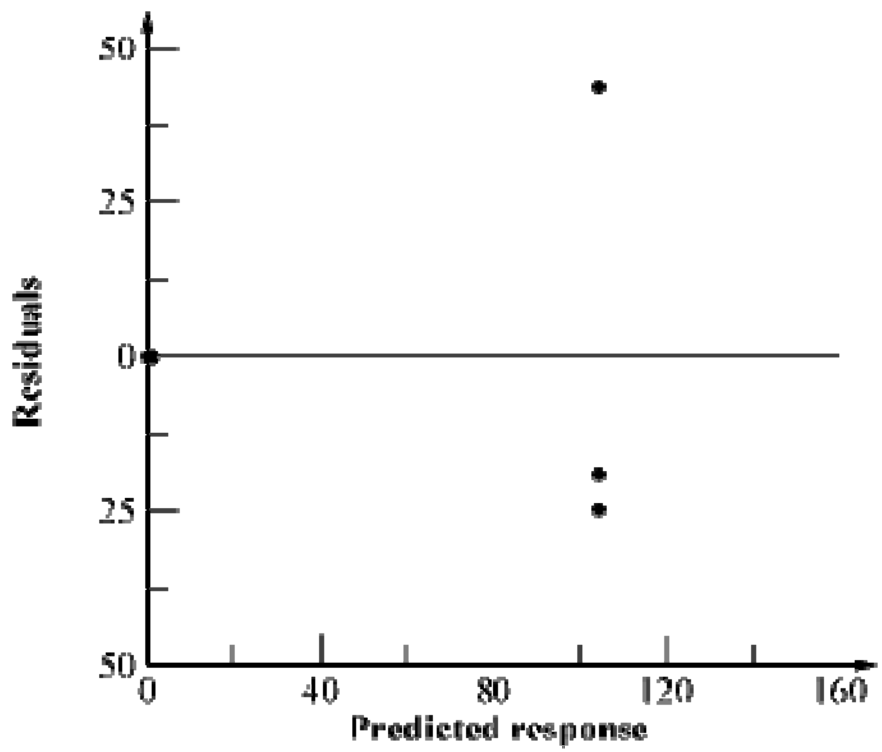
I	A	B	AB	y	Mean $\bar{y}$
1	-1	-1	1	(85.10, 79.50, 147.90)	104.170
1	1	-1	-1	(0.891, 1.047, 1.072)	1.003
1	-1	1	-1	(0.955, 0.933, 1.122)	1.003
1	1	1	1	(0.0148, 0.0126, 0.0118)	0.013
106.19	-104.15	-104.15	102.17	total	
26.55	-26.04	-26.04	25.54	total/4	

Additive model is not valid because:

- Physical consideration ⇒ effects of workload and processors do not add. They multiply.
- Large range for y. y_{max}/y_{min}= 147.90/0.0118 or 12,534
  ⇒ log transformation
- Taking an arithmetic mean of 114.17 and 0.013 is inappropriate.

Example 18.8 (Cont)

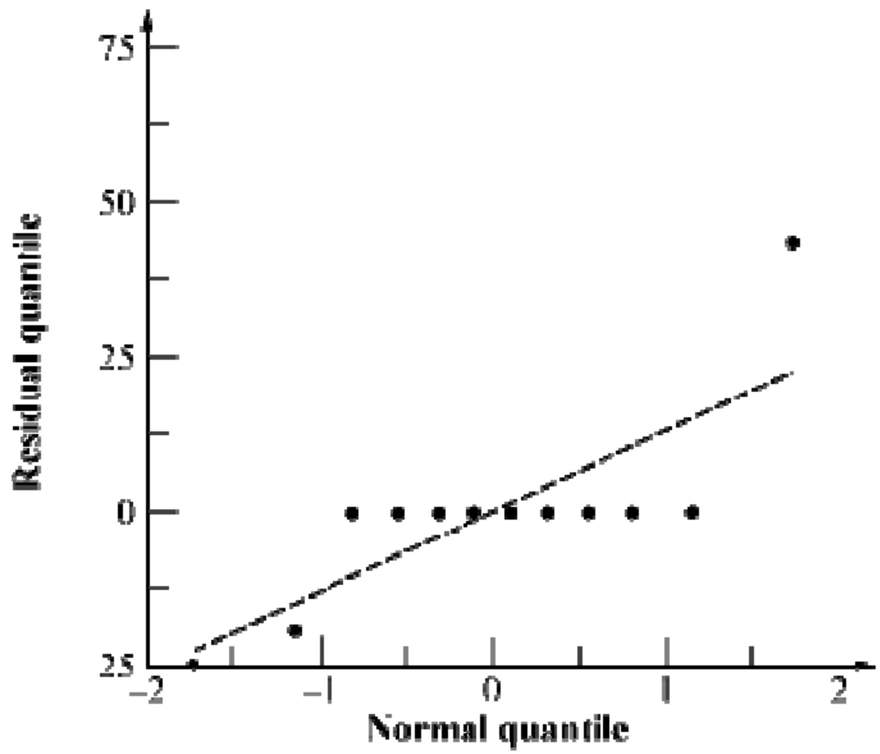
- The residuals are not small as compared to the response.



- The spread of residuals is large at larger value of the response. ⇒ log transformation

Example 18.8 (Cont)

- Residual distribution has a longer tail than normal



Analysis Using Multiplicative Model

Data After Log Transformation

I	A	B	AB	y	Mean $\bar{y}$
1	-1	-1	1	(1.93, 1.90, 2.17)	2.00
1	1	-1	-1	(-0.05, 0.02, 0.03)	0.00
1	-1	1	-1	(-0.02, -0.03, 0.05)	0.00
1	1	1	1	(-1.83, -1.90, -1.93)	-1.89
0.11	-3.89	-3.89	0.11	total	
0.03	-0.97	-0.97	0.03	total/4	

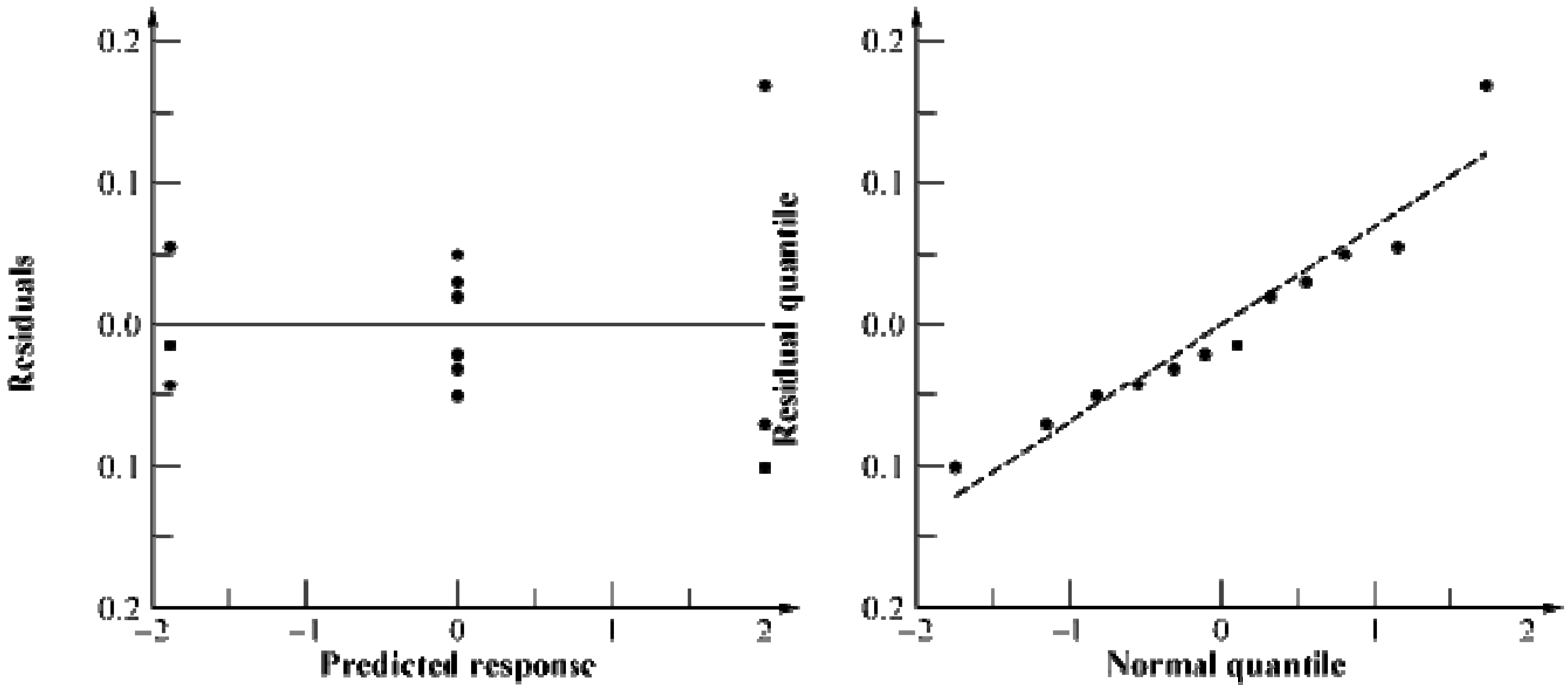
Variation Explained by the Two Models

	Additive Model			Multiplicative Model		
Factor	Effect	% Var.	Conf. Interval	Effect	% Var.	Conf. Interval
I	26.55		(16.35, 36.74)	0.03		(-0.02, 0.07)†
A	-26.04	30.1%	(-36.23, -15.84)	-0.97	49.9%	(-1.02, -0.93)
B	-26.04	30.1%	(-36.23, -15.84)	-0.97	49.9%	(-1.02, -0.93)
AB	25.54	29.0%	(15.35, 35.74)	0.03	0.0%	(-0.02, 0.07)†
e		10.8%			0.2%	

† $\Rightarrow$ Not Significant

- With multiplicative model:
  - Interaction is almost zero.
  - Unexplained variation is only 0.2%

Visual Tests



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

Interpretation of Results

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

$$\begin{aligned} \Rightarrow y &= 10^{q_0} 10^{q_A x_A} 10^{q_B x_B} 10^{q_{AB} x_A x_B} 10^e \\ &= 10^{0.03} 10^{-0.97 x_A} 10^{-0.97 x_B} 10^{0.03 x_A x_B} 10^e \\ &= 1.07 \times 0.107^{x_A} \times 0.107^{x_B} \times 1.07^{x_A x_B} 10^e \end{aligned}$$

- The time for an average processor on an average benchmark is 1.07.
- The time on processor A_1 is nine times (0.107^{-1}) that on an average processor. The time on A_2 is one ninth (0.107^{1}) of that on an average processor.
- MIPS rate for A_2 is 81 times that of A_1.
- Benchmark B_1 executes 81 times more instructions than B_2.
- The interaction is negligible.

  $\Rightarrow$ Results apply to all benchmarks and processors.

Transformation Considerations

- y_{max}/y_{min} small ⇒ Multiplicative model results similar to additive model.
- Many other transformations possible.
- Box-Cox family of transformations:

$$w = \begin{cases} \frac{y^a - 1}{a g^{a-1}}, & a \neq 0 \\ (\ln y) g, & a = 0 \end{cases}$$

- Where *g* is the geometric mean of the responses:

$$g = (y_1 y_2 \cdots y_n)^{1/n}$$

- w has the same units as y.
- *a* can have any real value, positive, negative, or zero.
- Plot SSE as a function of *a* ⇒ optimal *a*
- Knowledge about the system behavior should always take precedence over statistical considerations.

General $2^k r$ Factorial Design

- Model:

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

- Parameter estimation:

$$q_j = \frac{1}{2^k} \sum_{i=1}^{2^k} S_{ij} \bar{y}_i$$

S_{ij} = *(i,j)*th entry in the sign table.

- Sum of squares:

$$\mathrm{SSY} = \sum_{i=1}^{2^k} \sum_{j=1}^{r} y_{ij}^2$$
$$\mathrm{SS0} = 2^k r q_0^2$$
$$\mathrm{SST} = \mathrm{SSY} - \mathrm{SS0}$$
$$\mathrm{SS}j = 2^k r q_j^2 \mathrm{j} = 1, 2, \ldots, 2^k - 1$$
$$\mathrm{SSE} = \mathrm{SST} - \sum_{j=1}^{2^k-1} \mathrm{SS}j$$

General $2^k r$ Factorial Design (Cont)

- Percentage of y's variation explained by *j*th effect =

$$(\mathrm{SS}j/\mathrm{SST}) \times 100\%$$

- Standard deviation of errors:

$$s_e = \sqrt{\frac{\mathrm{SSE}}{2^k(r-1)}}$$

- Standard deviation of effects:

$$s_{q_0} = s_{q_A} = s_{q_B} = s_{q_{AB}} = s_e/\sqrt{2^k r}$$

- Variance of contrast $\sum h_i q_i$, where $\sum h_i = 0$ is:

$$s^2_{\sum h_i q_i} = (s_e^2 \sum h_i^2)/2^k r$$

General $2^k r$ Factorial Design (Cont)

- Standard deviation of the mean of m future responses:

$$s_{\hat{y}_p} = s_e \left(\frac{1+2^k}{2^k r} + \frac{1}{m} \right)^{1/2}$$

- Confidence intervals are calculated using $t_{[1-\alpha/2;2^k(r-1)]}$.
- Modeling assumptions:
  - Errors are IID normal variates with zero mean.
  - Errors have the same variance for all values of the predictors.
  - Effects and errors are additive.

Visual Tests for $2^k r$ Designs

- The scatter plot of errors versus predicted responses should not have any trend.
- The normal quantile-quantile plot of errors should be linear.
- Spread of y values in all experiments should be comparable.

Example 18.9: A $2^3 3$ Design

I	A	B	C	A B	A C	B C	A B C	y	Mean $\bar{\mathbf{y}}$
1	-1	-1	-1	1	1	1	-1	(14, 16, 12)	14
1	1	-1	-1	-1	-1	1	1	(22, 18, 20)	20
1	-1	1	-1	-1	1	-1	1	(11, 15, 19)	15
1	1	1	-1	1	-1	-1	-1	(34, 30, 35)	33
1	-1	-1	1	1	-1	-1	1	(46, 42, 44)	44
1	1	-1	1	-1	1	-1	-1	(58, 62, 60)	60
1	-1	1	1	-1	-1	1	-1	(50, 55, 54)	53
1	1	1	1	1	1	1	1	(86, 80, 74)	80
319	67	43	155	23	19	15	-1	total	
39.87	8.375	5.375	19.37	2.875	2.375	1.875	-0.125	total/8	

Example 18.9 (Cont)

- Sum of Squares:

Component	Sum of Squares	Percent Variation
y	4.9E4	
$\bar{y}$	3.8E4	
y-$\bar{y}$	1.1E4	100.00%
A	1683.0	14.06%
B	693.3	5.79%
C	9009.0	75.27%
AB	198.3	1.66%
AC	135.4	1.13%
BC	84.4	0.70%
ABC	0.4	0.00%
Errors	164.0	1.37%

Example 18.9 (Cont)

- The errors have $2^3(3-1)$ or 16 degrees of freedom. Standard deviation of errors:

$$s_e = \sqrt{\frac{\text{SSE}}{2^k(r-1)}} = \sqrt{\frac{164}{16}} = 3.20$$

- Standard deviation of effects:

$$s_{q_i} = s_e/\sqrt{(2^3 3)} = 3.20/\sqrt{24} = 0.654$$

Example 18.9 (Cont)

- % Variation:

Component	Sum of Squares	Percent Variation
y	4.9E4	
$\bar{y}$	3.8E4	
y-$\bar{y}$	1.1E4	100.00%
A	1683.0	14.06%
B	693.3	5.79%
C	9009.0	75.27%
AB	198.3	1.66%
AC	135.4	1.13%
BC	84.4	0.70%
ABC	0.4	0.00%
Errors	164.0	1.37%

Example 18.9 (Cont)

- $t_{[0.95,16]}$=1.337
- 90% confidence intervals for parameters: $q_i \mp (1.337)(0.654) = q_i \mp 0.874$

$$
\begin{aligned}
q_0 &= (39.00, 40.74)\\
q_A &= (7.50, 9.25)\\
q_B &= (4.50, 6.25)\\
q_C &= (18.50, 20.24)\\
q_{AB} &= (2.00, 3.75)\\
q_{AC} &= (1.50, 3.25)\\
q_{BC} &= (1.00, 2.75)\\
q_{ABC} &= (-1.00, 0.75)
\end{aligned}
$$

- All effects except q_{ABC} are significant.

Example 18.9 (Cont)

- For a single confirmation experiment (m = 1)
  With A = B = C = -1:

$$\hat{y} = 14$$

$$\begin{aligned} s_{\hat{y}} &= s_e\left(\frac{5}{2^k r} + \frac{1}{m}\right)^{1/2} \\ &= 3.2\left(\frac{5}{24} + 1\right)^{1/2} \\ &= 3.52 \end{aligned}$$

- 90% confidence interval:

$$14 \mp 1.337 \times 3.52 = 14 \mp 4.70 = (9.30,\ 18.70)$$

Case Study 18.1: Garbage collection

Factors and Levels

Variable	Factor	-1	1
A	Workload	Single Task	Several parallel tasks
B	Compiler	Simple	Deallocating
C	Limbo List	Enabled	Disabled
D	Chunk Size	4K bytes	16K bytes

Case Study 18.1 (Cont)

I	A	B	C	D	y	Mean $\bar{y}$
1	-1	-1	-1	-1	(97, 97, 97)	97.00
1	1	-1	-1	-1	(31, 31, 32)	31.33
1	-1	1	-1	-1	(97, 97, 97)	97.00
1	1	1	-1	-1	(31, 32, 31)	31.33
1	-1	-1	1	-1	(97, 97, 97)	97.00
1	1	-1	1	-1	(32, 32, 31)	31.67
1	-1	1	1	-1	(97, 97, 97)	97.00
1	1	1	1	-1	(32, 32, 32)	32.00
1	-1	-1	-1	1	(407, 407, 407)	407.00
1	1	-1	-1	1	(135, 136, 135)	135.33
1	-1	1	-1	1	(409, 409, 409)	409.00
1	1	1	-1	1	(135, 135, 136)	135.33
1	-1	-1	1	1	(407, 407, 407)	407.00
1	1	-1	1	1	(139, 140, 139)	139.33
1	-1	1	1	1	(409, 409, 409)	409.00
1	1	1	1	1	(139, 139, 140)	139.33
2695.67	-1344.33	4.33	9.00	1667.00		total
168.48	-84.02	0.27	0.56	104.19		total/8

Case Study 18.1 (Cont)

Factor	Effect	% Variation	Conf. Interval
I	168.48	138.1%	(168.386, 168.573)
A	-84.02	34.4%	(-84.114, -83.927)
B	0.27	0.0%	(0.177, 0.364)
C	0.56	0.0%	(0.469, 0.656)
D	104.19	52.8%	(104.094, 104.281)
AB	-0.23	0.0%	(-0.323, -0.136)
AC	0.56	0.0%	(0.469, 0.656)
AD	-51.31	12.8%	(-51.406, -51.219)
BC	0.02	0.0%	(-0.073, 0.114)†
BD	0.23	0.0%	(0.136, 0.323)
CD	0.44	0.0%	(0.344, 0.531)
ABC	0.02	0.0%	(-0.073, 0.114)†
ABD	-0.27	0.0%	(-0.364, -0.177)
ACD	0.44	0.0%	(0.344, 0.531)
BCD	-0.02	0.0%	(-0.114, 0.073)†
ABCD	-0.02	0.0%	(-0.114, 0.073)†

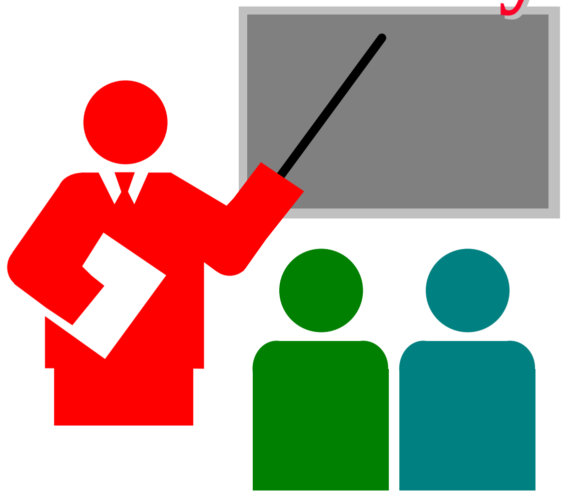
† $\Rightarrow$ Not Significant

Case Study 18.1: Conclusions

- Most of the variation is explained by factors A (Workload), D (Chunk size), and the interaction A D between the two.
- The variation due to experimental error is small

  ⇒ Several effects that explain less than 0.05% of variation (listed as 0.0%) are statistically significant.
- Only effects A, D, and AD are both practically significant and statistically significant.

Summary



- Replications allow estimation of measurement errors
  $\Rightarrow$ Confidence Intervals of parameters
  $\Rightarrow$ Confidence Intervals of predicted responses
- Allocation of variation is proportional to square of effects
- Multiplicative models are appropriate if the factors multiply
- Visual tests for independence normal errors

Exercise 18.1

Table 18.11 lists measured CPU times for two processors on two workloads. Each experiment was repeated three times. Analyze the design.

Table 18.11 $2^2 3$ Experimental Design Exercise

Workload	Processor	
	A	B
I	(41.16, 39.02, 42.56)	(63.17, 59.25, 64.23)
J	(51.50, 52.50, 50.50)	(48.08, 48.98, 47.10)

Homework 18B

Updated Exercise 18.1: For the data of Homework 18A, determine percentage of variation explained, find confidence intervals of the effects, and conduct visual tests.