Two Factors Full Factorial Design without Replications

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/cse567-13/

- Computation of Effects
- Estimating Experimental Errors
- Allocation of Variation
- ANOVA Table
- Visual Tests
- Confidence Intervals For Effects
- Multiplicative Models
- Missing Observations

Two Factors Full Factorial Design

- Used when there are two parameters that are carefully controlled
- Examples:
  - To compare several processors using several workloads.
  - To determining two configuration parameters, such as cache and memory sizes
- Assumes that the factors are categorical. For quantitative factors, use a regression model.
- A full factorial design with two factors *A* and *B* having *a* and *b* levels requires *ab* experiments.
- First consider the case where each experiment is conducted only once.

Model

$$y_{ij} = \mu + \alpha_j + \beta_i + e_{ij}$$

$$\begin{aligned}
y_{ij} &= \text{Observation with A at level j} \\
&\quad\ \text{and B at level i} \\
\mu &= \text{mean response} \\
\alpha_j &= \text{effect of factor A at level j} \\
\beta_i &= \text{effect of factor B at level i} \\
e_{ij} &= \text{error term}
\end{aligned}$$

Computation of Effects

- Averaging the jth column produces:

$$\bar{y}_{.j} = \mu + \alpha_j + \frac{1}{b}\sum_i \beta_i + \frac{1}{b}\sum_i e_{ij}$$

- Since the last two terms are zero, we have:

$$\bar{y}_{.j} = \mu + \alpha_j$$

- Similarly, averaging along rows produces:

$$\bar{y}_{i.} = \mu + \beta_i$$

- Averaging all observations produces

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

- Model parameters estimates are:

$$\begin{aligned} \mu &= \bar{y}_{..} \\ \alpha_j &= \bar{y}_{.j} - \bar{y}_{..} \\ \beta_i &= \bar{y}_{i.} - \bar{y}_{..} \end{aligned}$$

- Easily computed using a tabular arrangement.

Example 21.1: Cache Comparison

Workloads	Two Caches	One Cache	No Cache
ASM	54.0	55.0	106.0
TECO	60.0	60.0	123.0
SIEVE	43.0	43.0	120.0
DHRYSTONE	49.0	52.0	111.0
SORT	49.0	50.0	108.0

Example 21.1: Computation of Effects

Workloads	Two Caches	One Cache	No Cache	Row Sum	Row Mean	Row Effect
ASM	54.0	55.0	106.0	215.0	71.7	-0.5
TECO	60.0	60.0	123.0	243.0	81.0	8.8
SIEVE	43.0	43.0	120.0	206.0	68.7	-3.5
DHRYSTONE	49.0	52.0	111.0	212.0	70.7	-1.5
SORT	49.0	50.0	108.0	207.0	69.0	-3.2
Column Sum	255.0	260.0	568.0	1083.0		
Column Mean	51.0	52.0	113.6		72.2	
Column effect	-21.2	-20.2	41.4			

- An average workload on an average processor requires 72.2 ms of processor time.
- The time with two caches is 21.2 ms lower than that on an average processor
- The time with one cache is 20.2 ms lower than that on an average processor.
- The time without a cache is 41.4 ms higher than the average

Example 21.1 (Cont)

- Two-cache - One-cache = 1 ms.
- One-cache - No-cache = 41.4+20.2 or 61.6 ms.
- The workloads also affect the processor time required.
- The ASM workload takes 0.5 ms less than the average.
- TECO takes 8.8 ms higher than the average.

Estimating Experimental Errors

- Estimated response:
  $$\hat{y}_{ij} = \mu + \alpha_j + \beta_i$$
- Experimental error:
  $$e_{ij} = y_{ij} - \hat{y}_{ij} = y_{ij} - \mu - \alpha_j - \beta_i$$
- Sum of squared errors (SSE):
  $$\text{SSE} = \sum_{i=1}^{b} \sum_{j=1}^{a} e_{ij}^2$$
- Example: The estimated processor time is:
  $$\hat{y}_{11} = \mu + \alpha_1 + \beta_1 = 72.2 - 21.2 - 0.5 = 50.5$$
- Error = Measured-Estimated = 54-50.5 = 3.5

Example 21.2: Error Computation

Workloads	Two Caches	One Cache	No Cache
ASM	3.5	3.5	-7.1
TECO	0.2	-0.8	0.6
SIEVE	-4.5	-5.5	9.9
DHRYSTONE	-0.5	1.5	-1.1
SORT	1.2	1.2	-2.4

The sum of squared errors is:

$$\text{SSE} = (3.5)^2 + (0.2)^2 + \cdots + (-2.4)^2 = 2368.00$$

Example 21.2: Allocation of Variation

$$y_{ij} = \mu + \alpha_j + \beta_i + e_{ij}$$

- Squaring the model equation:

$$\begin{array}{ccccccccc} \sum_{ij} y_{ij}^2 & = & ab\mu^2 & + & b\sum_j \alpha_j^2 & + & a\sum_i \beta_i^2 & + & \sum_{ij} e_{ij}^2 \\ \text{SSY} & = & \text{SS0} & + & \text{SSA} & + & \text{SSB} & + & \text{SSE} \end{array}$$

$$\begin{array}{ccccccccccc} \text{SST} & = & \text{SSY} & - & \text{SS0} & = & \text{SSA} & + & \text{SSB} & + & \text{SSE} \\ 13402.41 & = & 91595 & - & 78192.59 & = & 12857.20 & + & 308.40 & + & 236.80 \\ 100\% & = & & & & = & 95.9\% & + & 2.3\% & + & 1.8\% \end{array}$$

- High percent variation explained
  $\Rightarrow$ Cache choice <u>important</u> in processor design.

Analysis of Variance

- Degrees of freedoms:

$$\begin{array}{ccccccccc} \mathrm{SSY} & = & \mathrm{SS0} & + & \mathrm{SSA} & + & \mathrm{SSB} & + & \mathrm{SSE} \\ ab & = & 1 & + & (a-1) & + & (b-1) & + & (a-1)(b-1) \end{array}$$

- Mean squares:

$$\mathrm{MSA} = \frac{\mathrm{SSA}}{a-1}$$

$$\mathrm{MSB} = \frac{\mathrm{SSB}}{b-1}$$

$$\mathrm{MSE} = \frac{\mathrm{SSE}}{(a-1)(b-1)}$$

$$\frac{MSA}{MSE} \sim F_{[1-\alpha;a-1,(a-1)(b-1)]}$$

- Computed ratio > $F_{[1-\alpha;a-1,(a-1)(b-1)]}$ $\Rightarrow$ A is <u>significant</u> at level α.

ANOVA Table

Component	Sum of Squares	%Variation	DF	Mean Square	F-Comp.	F-Table
y	$\text{SSY} = \sum y_{ij}^2$		ab			
$\bar{y}_{..}$	$\text{SS0} = ab\mu^2$		1			
$y - \bar{y}_{..}$	SST=SSY-SS0	100	$ab - 1$			
A	$\text{SSA} = b\Sigma\alpha_j^2$	$100\left(\frac{\text{SSA}}{\text{SST}}\right)$	$a - 1$	$\text{MSA} = \frac{\text{SSA}}{a-1}$	$\frac{\text{MSA}}{\text{MSE}}$	$F_{[1-\alpha, a-1, (a-1)(b-1)]}$
B	$\text{SSB} = a\Sigma\beta_i^2$	$100\left(\frac{\text{SSB}}{\text{SST}}\right)$	$b - 1$	$\text{MSB} = \frac{\text{SSB}}{b-1}$	$\frac{\text{MSB}}{\text{MSE}}$	$F_{[1-\alpha, b-1, (a-1)(b-1)]}$
e	$\text{SSE} = \text{SST} - (\text{SSA} + \text{SSB})$	$100\left(\frac{\text{SSE}}{\text{SST}}\right)$	$(a-1)(b-1)$	$\text{MSE} = \frac{\text{SSE}}{(a-1)(b-1)}$		

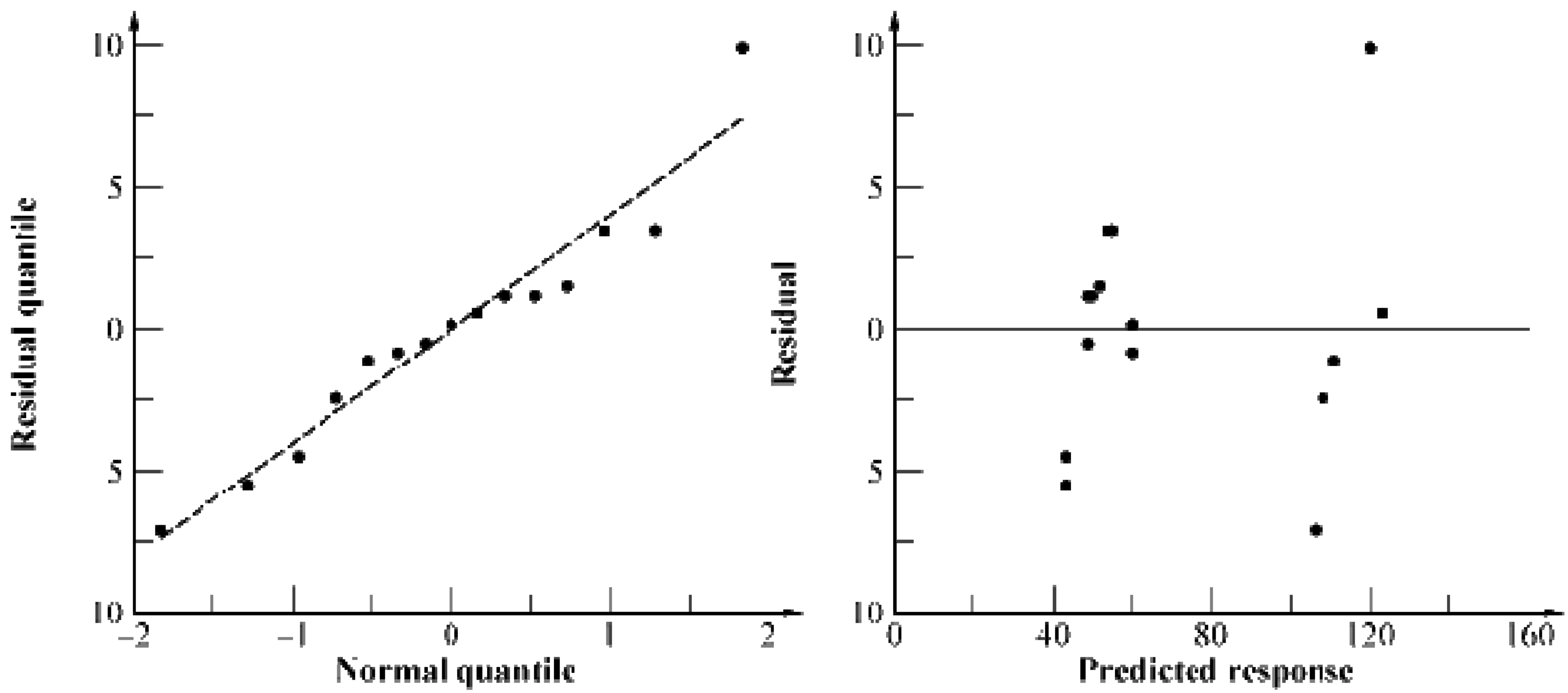
Example 21.3: Cache Comparison

Component	Sum of Squares	%Variation	DF	Mean Square	F-Comp.	F-Table
y	91595.00					
$\bar{y}_{..}$	78192.59					
y-$\bar{y}_{..}$	13402.41	100.0%	14			
Caches	12857.20	95.9%	2	6428.60	217.2	3.1
Workloads	308.40	2.3%	4	77.10	2.6	2.8
Errors	236.80	1.8%	8	29.60		

$$s_e = \sqrt{\text{MSE}} = \sqrt{29.60} = 5.44$$

- Cache choice significant.
- Workloads insignificant

Example 21.4: Visual Tests



Confidence Intervals For Effects

Parameter	Estimate	Variance
μ	$\bar{y}_{..}$	s_e^2/ab
α_j	$\bar{y}_{.j}$-$\bar{y}_{..}$	$s_e^2(a-1)/ab$
$\mu+\alpha_j$	$\bar{y}_{.j}$	s_e^2/b
β_i	$\bar{y}_{i.}$-$\bar{y}_{..}$	$s_e^2(b-1)/ab$
$\mu+\alpha_j+\beta_i$	$\bar{y}_{.j}+\bar{y}_{i.}$-$\bar{y}_{..}$	$s_e^2(a+b-1)/(ab)$
$\sum_{j=1}^{a} h_j\ \alpha_j,\ \sum_{j=1}^{a} h_j = 0$	$\sum_{j=1}^{a} h_j\ \bar{y}_{.j}$	$s_e^2 \sum_{j=1}^{a} h_j^2/b$
$\sum_{i=1}^{b} h_i\ \beta_i,\ \sum_{i=1}^{b} h_i = 0$	$\sum_{i=1}^{b} h_i\ \bar{y}_{i.}$	$s_e^2 \sum_{i=1}^{b} h_i^2/a$
s_e^2	$\{\sum_{j=1}^{a} \sum_{i=1}^{b} e_{ij}^2\}/\{(a-1)(b-1)\}$	

Degrees of freedom for errors = (a-1)(b-1)

- For confidence intervals use t values at (a-1)(b-1) degrees of freedom

Example 21.5: Cache Comparison

- Standard deviation of errors:

$$s_e = \sqrt{\text{MSE}} = \sqrt{29.60} = 5.4$$

- Standard deviation of the grand mean:

$$s_\mu = s_e/\sqrt{ab} = 5.4/\sqrt{15} = 1.4$$

- Standard deviation of α_j's:

$$s_{\alpha_j} = s_e\sqrt{(a-1)/ab} = 5.4\sqrt{\frac{2}{15}} = 2.0$$

- Standard deviation of β_i's:

$$s_{\beta_i} = s_e\sqrt{(b-1)/ab} = 5.4\sqrt{\frac{4}{15}} = 2.8$$

Example 21.5 (Cont)

- Degrees of freedom for the errors are (a-1)(b-1)=8. For 90% confidence interval, $t_{[0.95;8]} = 1.86$.
- Confidence interval for the grand mean:

$$72.2 \mp 1.86 \times 1.4 = 72.2 \mp 2.6 = (69.6,\ 74.8)$$

Parameter	Mean Effect	Std. Dev.	Confidence Interval
μ	72.2	1.4	(69.6, 74.8)
Caches			
Two Caches	-21.2	2.0	(-24.9, -17.5)
One Cache	-20.2	2.0	(-23.9, -16.5)
No Cache	41.4	2.0	(37.7, 45.1)

- All three cache alternatives are significantly different from the average.

Example 21.5 (Cont)

Parameter	Mean Effect	Std. Dev.	Confidence Interval
ASM	-0.5	2.8	(-5.8, 4.7)†
TECO	8.8	2.8	(3.6, 14.0)
SIEVE	-3.5	2.8	(-8.8, 1.7)†
DHRYSTONE	-1.5	2.8	(-6.8, 3.7)†
SORT	-3.2	2.8	(-8.4, 2.0)†

† $\Rightarrow$ Not significant

- All workloads, except TECO, are similar to the average and hence to each other.

Example 21.5: CI for Differences

	Two Caches	One Cache	No Cache
Two Caches		(-7.4, 5.4)†	(-69.0, -56.2)
One Cache			(-68.0, -55.2)

† $\Rightarrow$ Not significant

- Two-cache and one-cache alternatives are both significantly better than a no cache alternative.
- There is no significant difference between two-cache and one-cache alternatives.

Multiplicative Models

- Additive model:

  $$y_i = \mu + \alpha_j + \beta_i + e_{ij}$$

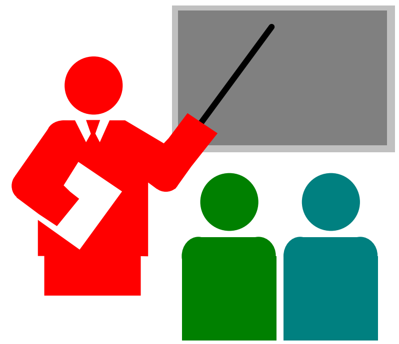
- If factors multiply $\Rightarrow$ Use multiplicative model
- Example: processors and workloads
  - Log of response follows an additive model
- If the spread in the residuals increases with the mean response

  $\Rightarrow$ Use transformation

Missing Observations

- Recommended Method:
  - Divide the sums by respective number of observations
  - Adjust the degrees of freedoms of sums of squares
  - Adjust formulas for standard deviations of effects
- Other Alternatives:
  - Replace the missing value by $\hat{y}$ such that the residual for the missing experiment is zero.
  - Use y such that SSE is minimum.

Summary



Two Factor Designs Without Replications

- Model:

$$y_{ij} = \mu + \alpha_j + \beta_i + e_{ij}$$

- Effects are computed so that:

$$\sum_{j=1}^{a} \alpha_j = 0$$
$$\sum_{i=1}^{b} \beta_i = 0$$

- Effects:

$$\mu = \bar{y}_{..};\ \alpha_j = \bar{y}_{.j} - \bar{y}_{..};\ \beta_i = \bar{y}_{i.} - \bar{y}_{..}$$

Summary (Cont)

- Allocation of variation: SSE can be calculated after computing

$$
\begin{array}{ccccccccc}
\sum_{ij} y_{ij}^2 & = & ab\mu^2 & + & b\sum_j \alpha_j^2 & + & a\sum_i \beta_i^2 & + & \sum_{ijk} e_{ijk}^2 \\
\text{SSY} & = & \text{SS0} & + & \text{SSA} & + & \text{SSB} & + & \text{SSE}
\end{array}
$$

Degrees of freedom:

$$
\begin{array}{ccccccccc}
\text{SSY} & = & \text{SS0} & + & \text{SSA} & + & \text{SSB} & + & \text{SSE} \\
ab & = & 1 & + & (a-1) & + & (b-1) & + & (a-1)(b-1)
\end{array}
$$

- Mean squares:

$$\text{MSA} = \frac{\text{SSA}}{a-1};\ \text{MSB} = \frac{\text{SSB}}{b-1};\ \text{MSE} = \frac{\text{SSE}}{(a-1)(b-1)}$$

- Analysis of variance:

MSA/MSE should be greater than $F_{[1-\alpha;a-1,(a-1)(b-1)]}$.
MSB/MSE should be greater than $F_{[1-\alpha;b-1,(a-1)(b-1)]}$.

Summary (Cont)

- Standard deviation of effects:

  $s_\mu^2 = s_e^2/ab;\ s_{\alpha_j}^2 = s_e^2(a-1)/ab;\ s_{\beta_i}^2 = s_e^2(b-1)/ab;$

- Contrasts:

  For $\sum_{j=1}^a h_j\ \alpha_j$, $\sum_{j=1}^a h_j = 0$: Mean $= \sum_{j=1}^a h_j\ \bar{y}_{.j}$; Variance $= s_e^2 \sum_{j=1}^a h_j^2/b$

  For $\sum_{i=1}^b h_i\ \beta_i$, $\sum_{i=1}^b h_i = 0$: Mean $= \sum_{i=1}^b h_i\ \bar{y}_{i.}$; Variance $= s_e^2 \sum_{i=1}^b h_i^2/a$

- All confidence intervals are calculated using $t_{[1-\alpha/2;(a-1)(b-1)]}$.
- Model assumptions:
  - Errors are IID normal variates with zero mean.
  - Errors have the same variance for all factor levels.
  - The effects of various factors and errors are additive.
- Visual tests:
  - No trend in scatter plot of errors versus predicted responses
  - The normal quantile-quantile plot of errors should be linear.

Homework 21: Exercise 21.1

Execution Times

Workloads	Processors		
	Scheme86	Spectrum125	Spectrum62.5
Garbage Collection	39.97	99.06	56.24
Pattern Match	0.958	1.672	1.252
Bignum Addition	0.01910	0.03175	0.01844
Bignum Multiplication	0.256	0.423	0.236
Fast Fourier Transform (1024)	10.21	20.28	10.14

Analyze the data of Case study 21.2 using a 2-factor additive model.

- Estimate effects and prepare ANOVA table
- Plot residuals as a function of predicted response.
- Also, plot a normal quantile-quantile plot for the residuals.
- Determine 90% confidence intervals for the paired differences. (Confidence intervals of α_1-α_2, $\alpha_1-\alpha_3$, α_2-α_3)
- Are the processors significantly different?
- Discuss what indicators in the data, analysis, or plot would suggest that this is not a good model.