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/cse567-08/

- 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 21.2 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:

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

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

- High percent variation explained
  ⇒ 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:

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

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

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

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

- 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	$SSY = \sum y_{ij}^2$		ab			
$\bar{y}_{\cdot\cdot}$	$SS0 = ab\mu^2$		1			
$y - \bar{y}_{\cdot\cdot}$	SST=SSY-SS0	100	$ab - 1$			
A	$SSA = b\Sigma\alpha_j^2$	$100\left(\frac{SSA}{SST}\right)$	$a - 1$	$MSA = \frac{SSA}{a-1}$	$\frac{MSA}{MSE}$	$F_{[1-\alpha, a-1, (a-1)(b-1)]}$
B	$SSB = a\Sigma\beta_i^2$	$100\left(\frac{SSB}{SST}\right)$	$b - 1$	$MSB = \frac{SSB}{b-1}$	$\frac{MSB}{MSE}$	$F_{[1-\alpha, b-1, (a-1)(b-1)]}$
e	$SSE = SST - (SSA + SSB)$	$100\left(\frac{SSE}{SST}\right)$	$(a - 1)(b - 1)$	$MSE = \frac{SSE}{(a-1)(b-1)}$		

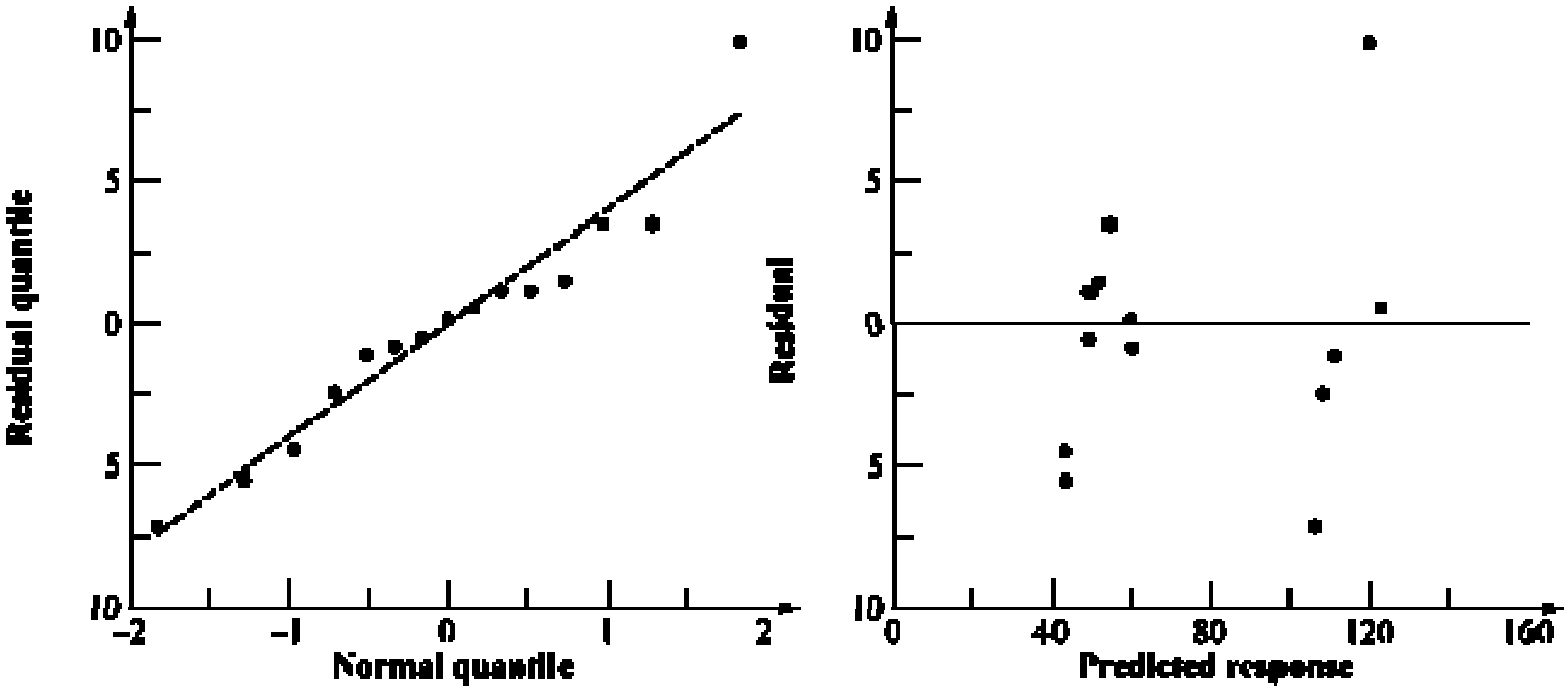
Example 21.3: Cache Comparison

Component	Sum of Squares	%Variation	DF	Mean Square	F-Comp.	F-Table
y	91595.00					
$y_{..}$	78192.59					
y-$y_{..}$	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	$\bar{y}_{.j}$-$\bar{y}_{..}$	$s_e^2(a-1)/ab$
μ+α_j	$\bar{y}_{.j}$	s_e^2/b
β_i	$\bar{y}_{i.}$-$\bar{y}_{..}$	$s_e^2(b-1)/ab$
μ+α_j+β_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:

$$s_{\alpha_j} = s_e\sqrt{(a-1)/ab} = 5.4\sqrt{\frac{2}{15}} = 2.8$$

- Standard deviation of β_i's:

$$s_{\beta_i} = s_e\sqrt{(b-1)/ab} = 5.4\sqrt{\frac{4}{15}} = 2.0$$

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	2.8	(-24.9, -17.5)
One Cache	-20.2	2.8	(-23.9, -16.5)
No Cache	41.4	2.8	(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	2.0	(-5.8, 4.7)†
TECO	8.8	2.0	(3.6, 14.0)
SIEVE	-3.5	2.0	(-8.8, 1.7)†
DHRYSTONE	-1.5	2.0	(-6.8, 3.7)†
SORT	-3.2	2.0	(-8.4, 2.0)†

† ⇒ 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.

Case Study 21.1: Cache Design Alternatives

- **Multiprocess environment**: Five jobs in parallel.

  ALL = ASM, TECO, SIEVE, DHRYSTONE, and SORT in parallel.
- **Processor Time:**

Workload	Two Caches	One Cache	No Cache
ASM5	231	262	489
TECO5	300	314	620
SIEVE5	213	214	604
DHRYSTONE5	245	263	564
ALL	229	242	551

Case Study 21.1 on Cache Design (Cont)

Confidence Intervals for Differences:

	One Cache	No Cache
Two Caches	(-51.6, 20.8)†	(-358.2, -285.8)
One Cache		(-342.8, -270.4)

† ⇒ Not significant

Conclusion: The two caches do not produce statistically better performance.

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

Case Study 21.2: RISC architectures

- Parallelism in time vs parallelism in space
- Pipelining vs several units in parallel
- Spectrum = HP9000/840 at 125 and 62.5 ns cycle
- Scheme86 = Designed at MIT

Cache Study 21.2: Simulation Results

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

- Additive model: $\Rightarrow$ No significant difference
- Easy to see that: Scheme86 = 2 or 3 × Spectrum125
- Spectrum62.5 = 2 × Spectrum125
- Execution Time = Processor Speed × Workload Size
  $\Rightarrow$ Multiplicative model.
- Observations skewed. $y_{max}/y_{min} > 1000$
  $\Rightarrow$ Adding not appropriate

Case Study 21.2: Multiplicative Model

- Log Transformation:

Workloads	Processors			Row	Row	Row
	Scheme86	Spectrum125	Spectrum62.5	Sum	Mean	Effect
Garbage Collect	1.6017	1.9959	1.7500	5.3477	1.7826	1.6212
Pattern Match	-0.0186	0.2232	0.0976	0.3022	0.1007	-0.0607
Bignum Add	-1.7212	-1.4949	-1.7447	-4.9608	-1.6536	-1.8150
Bignum Mult	-0.5918	-0.3737	-0.6271	-1.5925	-0.5308	-0.6922
FFT (1024)	1.0090	1.3092	1.0060	3.3243	1.1081	0.9467
Column Sum	0.2791	1.6598	0.4819	2.4208		
Column Mean	0.0558	0.3320	0.0964		0.1614	
Column effect	-0.1056	0.1706	-0.0650			

- Effect of the processors is significant.
- The model explains 99.9% of variation as compared to 88% in the additive model.

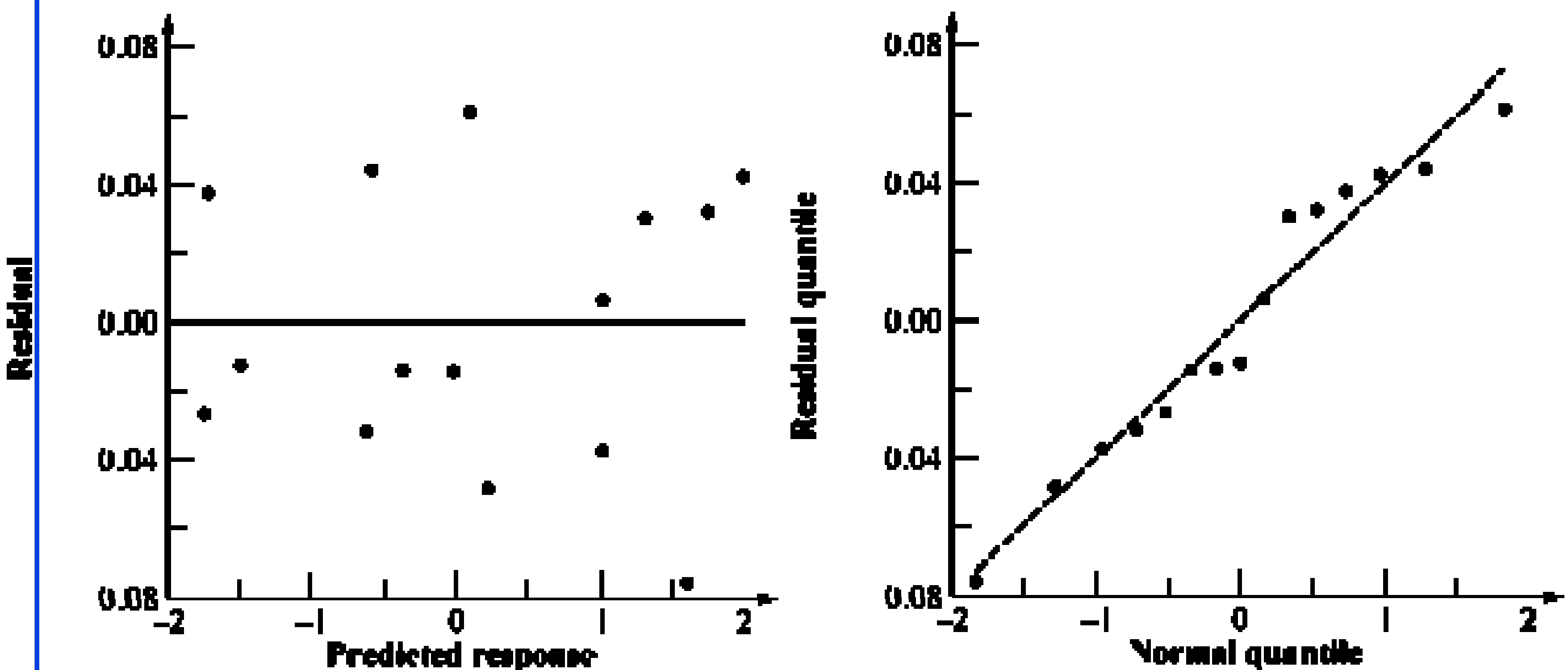
Case Study 21.2: Confidence Intervals

	Spectrum125	Spectrum62.5
Scheme86	(-0.3387, -0.2136)	(-0.1031, 0.0220)†
Spectrum125		(0.1730, 0.2982)

† ⇒ Not significant

- Scheme86 and Spectrum62.5 are of comparable speed.
- Spectrum125 is significantly slower than the other two processors.
- Scheme86's time is 0.4584 to 0.6115 times that of Spectrum125 and 0.7886 to 1.0520 times that of Spectrum62.5.
- The time on Spectrum125 is 1.4894 to 1.9868 times that on Spectrum62.5.

Cache Study 21.2: Visual Tests



Case Study 21.2: ANOVA

Component	Sum of Squares	%Variation	DF	Mean Square	F-Comp.	F-Table
y	22.54					
$y_{..}$	0.39					
y-$y_{..}$	22.15	100.00%	14			
Processors	0.22	1.00%	2	0.11	39.29	3.11
Workloads	21.90	98.89%	4	5.48	1935.48	2.81
Errors	0.02	0.10%	8	0.00		

$$s_e = \sqrt{\text{MSE}} = \sqrt{0.00} = 0.05$$

- Processors account for only 1% of the variation
- Differences in the workloads account for 99%.

  $\Rightarrow$ Workloads widely different

  $\Rightarrow$ Use more workloads or cover a smaller range.

Case Study 21.3: Processors

Measured Elasped Times

Workload	Processors			
	A	B	C	D
ASM	54	101	111	83
TECO	60	92	110	90
SIEVE	42	121	127	86
DHRYSTONE	49	97	122	81
SORT	52	100	107	82

Case Study 21.3: Additive Model

Component	Sum of Squares	%Variation	DF	Mean Square	F-Comp.	F-Table
y	168653.00					
$y_{..}$	156114.45					
y-$y_{..}$	12538.55	100.0%	19			
Processors	11522.15	91.9%	3	3840.72	54.9	2.6
Workloads	176.30	1.4%	4	44.07	0.6	2.5
Errors	840.10	6.7%	12	70.01		

- Workloads explain 1.4% of the variation.
- Only 6.7% of the variation is unexplained.

Case Study 21.3: Multiplicative Model

Component	Sum of Squares	%Variation	DF	Mean Square	F-Comp.	F-Table
y	74.56					
$y_{..}$	74.17					
y-$y_{..}$	0.39	100.00%	19			
Processors	0.36	93.15%	3	0.121	57.55	2.61
Workloads	0.00	0.37%	4	0.000	0.17	2.48
Errors	0.03	6.47%	12	0.002		

$$s_e=\sqrt{\text{MSE}}= \sqrt{0.002}= 0.05$$

- Both models pass the visual tests equally well.
- It is more appropriate to say that processor B takes twice as much time as processor A, than to say that processor B takes 50.7 ms more than processor A.

Case Study 21.3: Intel iAPX 432

System				Workload			
No.	Processor	Language	Word Size	484	Sieve	Puzzle	Acker
1	VAX-11/780	C	32	1.4	250.0	9400.0	4600.0
2		Pascal (UNIX)	32	1.6	220.0	11900.0	7800.0
3		Pascal (VMS)	32	1.4	259.0	11530.0	9850.0
4	68000 (8 MHz)	C	32	4.7	740.0	37100.0	7800.0
5		Pascal	16	5.3	810.0	32470.0	11480.0
6		Pascal	32	5.8	960.0	32520.0	12320.0
7	68000 (16 MHz)	Pascal	16	1.3	196.0	9180.0	2750.0
8		Pascal	32	1.5	246.0	9200.0	3080.0
9	8086 (5 MHz)	Pascal	16	7.3	764.0	44000.0	11100.0
10	432	Ada	16	35.0	3200.0	350000.0	260000.0
11		Ada	16	14.2	3200.0	165000.0	260000.0
12		Ada	32	16.1	3200.0	180000.0	260000.0

Case Study 21.3: ANOVA with Log

Component	Sum of Squares	%Variation	DF	Mean Square	F-Comp.	F-Table
y	576.64					
$y_{..}$	449.01					
y-$y_{..}$	127.63	100.0%	47			
Workload	113.01	88.5%	3	37.7	1158.5	2.3
System	13.55	10.6%	11	1.2	37.9	1.8
Errors	1.07	0.8%	33	0.03		

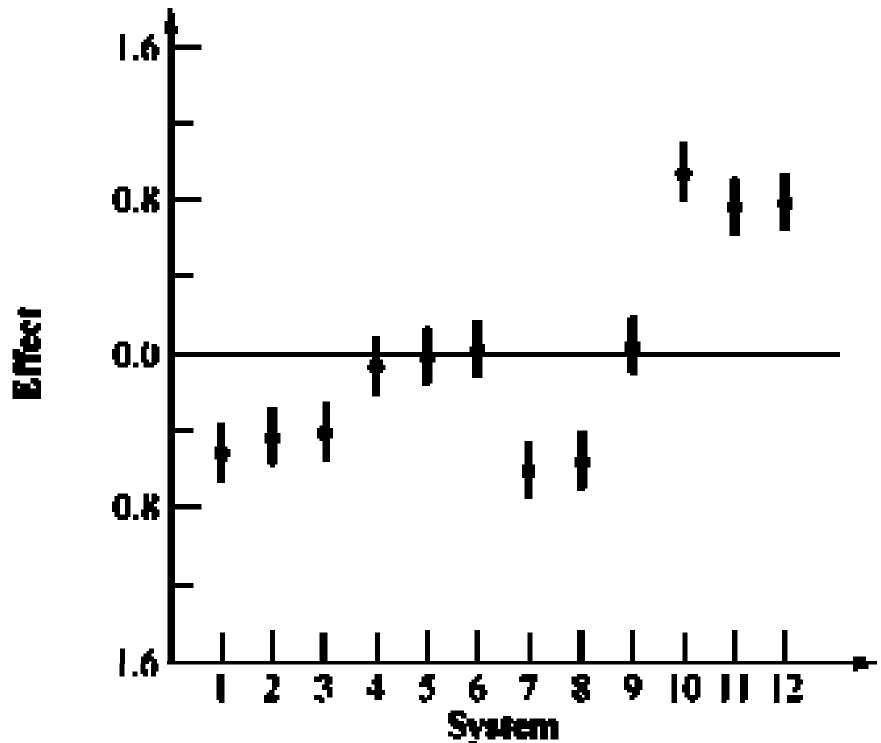
$$s_e = \sqrt{\text{MSE}} = \sqrt{0.03} = 0.18$$

- Only 0.8% of variation is unexplained.

  Workloads explain a much larger percentage of variation than the systems

  ⇒ the workload selection is poor.

Case Study 21.3: Confidence intervals



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.

Case Study 21.4: RISC-I Execution Times

	Processors					
Workloads	RISC I	68000	Z80002	VAX-11/780	PDP-11/70	C/70
E-String Search	0.46	1.29	0.74	0.60	0.41	1.01
F-Bit Test	0.06	0.29	0.43	0.29	0.37	0.55
H-Linked List	0.10	0.16	0.24	0.12	0.19	0.25
K-Bit Matrix	0.43	1.72	2.24	1.29	1.72	4.00
I-Quick Sort	50.40	206.64	262.08	151.20	181.44	292.32
Ackermann(3,6)	3200.00	-	8960.00	5120.00	5120.00	-
Recursive Qsort	800.00	-	4720.00	1840.00	2560.00	1040.00
Puzzle (Subscript)	4700.00	-	19740.00	9400.00	7520.00	15980.00
Puzzle (Pointer)	3200.00	13440.00	7360.00	4160.00	6400.00	6720.00
SED (Batch Editor)	5100.00	-	22440.00	5610.00	5610.00	13260.00
Towers Hanoi (18)	6800.00	-	28560.00	12240.00	15640.00	10880.00

Case Study 21.5: Using Multiplicative Model

Workloads	Processors: RISC I	68000	Z80002	VAX-11/780	PDP-11/70	C/70	Row Sum	Row Mean	Row Effect
E-String Search	-0.34	0.11	-0.13	-0.22	-0.38	0.01	-0.96	-0.16	-2.16
F-Bit Test	-1.22	-0.54	-0.36	-0.54	-0.43	-0.26	-3.36	-0.56	-2.55
H-Linked List	-1.00	-0.80	-0.62	-0.92	-0.72	-0.60	-4.66	-0.78	-2.77
K-Bit Matrix	-0.37	0.24	0.35	0.11	0.24	0.60	1.17	0.19	-1.80
I-Quick Sort	1.70	2.32	2.42	2.18	2.26	2.47	13.34	2.22	0.23
Ackermann(3,6)	3.51	-	3.95	3.71	3.71	-	14.88	3.72	1.72
Recursive Qsort	2.90	-	3.67	3.26	3.41	3.02	16.27	3.25	1.26
Puzzle (Subscript)	3.67	-	4.30	3.97	3.88	4.20	20.02	4.00	2.01
Puzzle (Pointer)	3.51	4.13	3.87	3.62	3.81	3.83	22.75	3.79	1.80
SED (Batch Editor)	3.71	-	4.35	3.75	3.75	4.12	19.68	3.94	1.94
Towers Hanoi (18)	3.83	-	4.46	4.09	4.19	4.04	20.61	4.12	2.13
Column Sum	19.90	5.45	26.25	23.01	23.70	21.42	119.73		
Column Mean	1.81	0.91	2.39	2.09	2.15	2.14		2.00	
Column effect	-0.19	-1.09	0.39	0.10	0.16	0.15			

Case Study 21.5: Experimental Errors

	Processors					
Workloads	RISC I	68000	Z80002	VAX-11/780	PDP-11/70	C/70
E-String Search	0.01	1.36	-0.36	-0.16	-0.38	0.02
F-Bit Test	-0.48	1.11	-0.20	-0.08	-0.03	0.15
H-Linked List	-0.04	1.07	-0.23	-0.24	-0.10	0.03
K-Bit Matrix	-0.37	1.13	-0.24	-0.18	-0.12	0.26
I-Quick Sort	-0.33	1.18	-0.20	-0.14	-0.12	0.10
Ackermann(3,6)	-0.03	-	-0.16	-0.11	-0.17	-
Recursive Qsort	-0.16	-	0.03	-0.08	-	-0.38
Puzzle (Subscript)	-0.15	-	-0.10	-0.13	-0.29	0.05
Puzzle (Pointer)	-0.10	1.42	-0.32	-0.27	-0.15	-0.11
SED (Batch Editor)	-0.04	-	0.02	-0.28	-0.35	0.04
Towers Hanoi (18)	-0.10	-	-0.06	-0.13	-0.09	-0.23

Case Study 21.5: Experimental Errors (Cont)

$$\text{SSE} = \sum_{i,j} e_{ij}^2 = 11.01$$

- 16 independent parameters (μ, α_j, and β_i) have been computed $\Rightarrow$ Errors have 60-1-5-10 or 44 degrees of freedom.

$$\text{MSE} = \frac{\text{SSE}}{\nu_e} = \frac{11.01}{44} = 0.25$$

- The standard deviation of errors is:

$$s_e = \sqrt{\text{MSE}} = \sqrt{0.25} = 0.05$$

- The standard deviation of α_j:

$$s_{\alpha_j} = s_e \sqrt{\frac{N - c_j}{N c_j}}$$

- c_j = number of observations in column c_j.

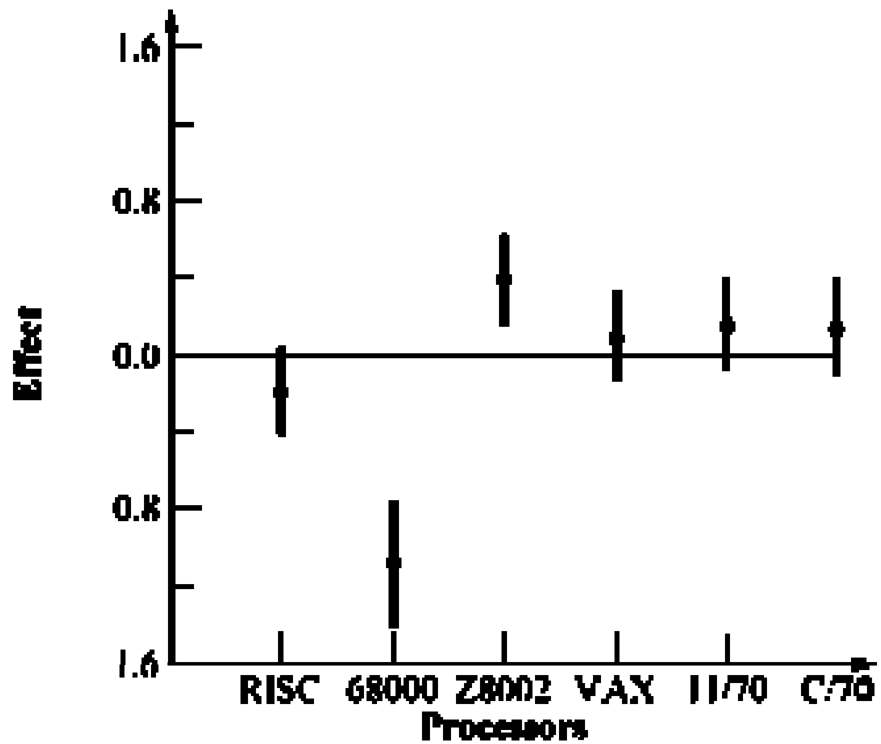
Case Study 21.5 (Cont)

- The standard deviation of the row effects:

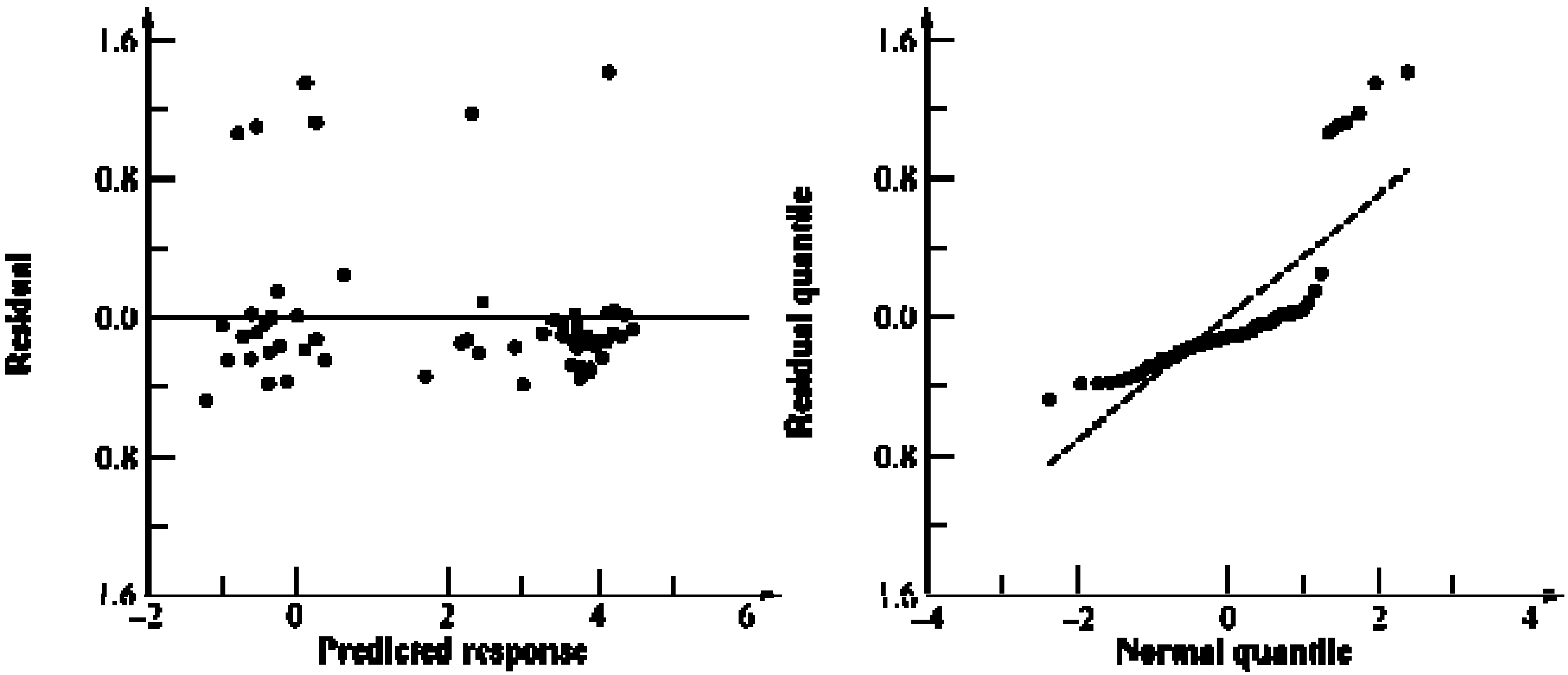
$$s_{\beta_i} = s_e \sqrt{\frac{N - r_i}{N r_i}}$$

r_i=number of observations in the *i*th row.

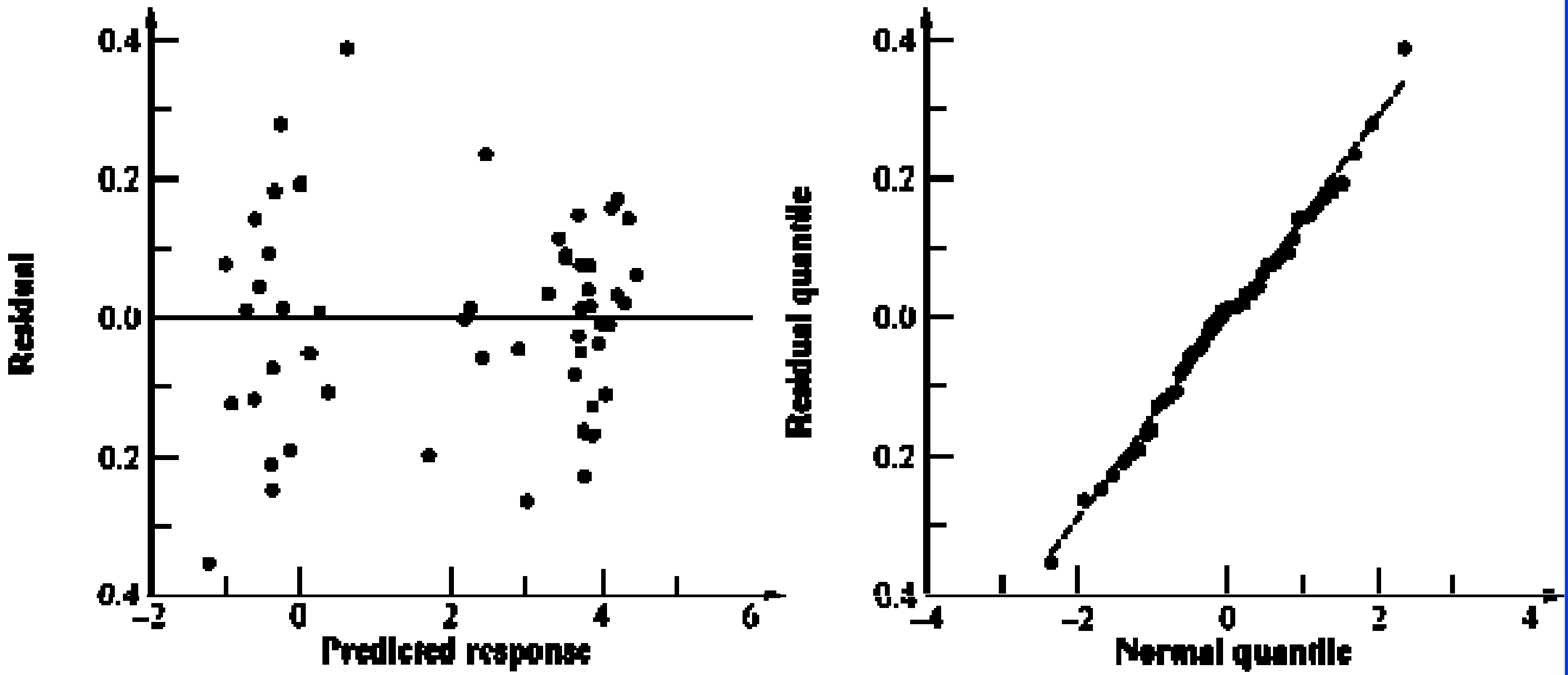
Case Study 21.5: CIs for Processor Effects



Case Study 21.5: Visual Tests



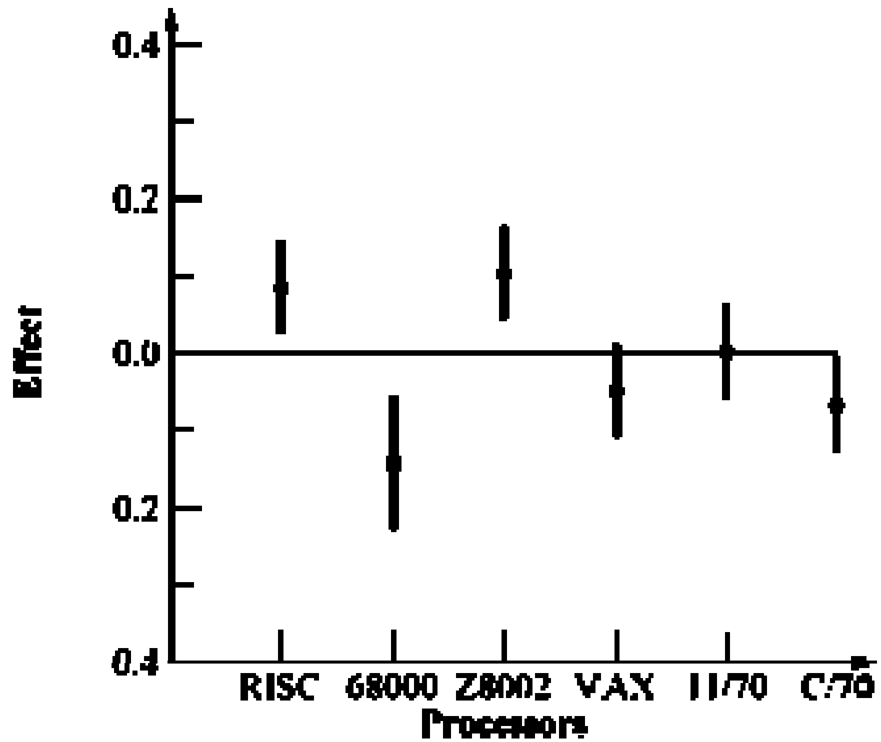
Case Study 21.5: Analysis without 68000



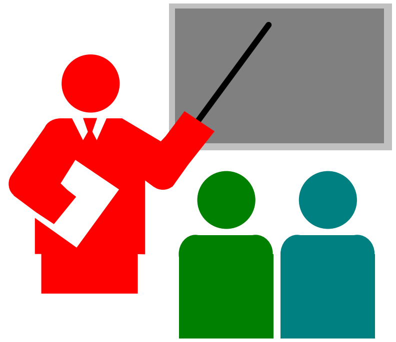
Case Study 21.5: RISC-I Code Size

	Processors					
Benchmarks	RISC I	68000	Z8002	VAX-11/780	PDP-11/70	C/70
E-String Search	140	112	126	98	112	98
F-Bit Test	120	144	180	144	168	120
H-Linked List	176	123	141	211	299	141
K-Bit Matrix	288	317	374	288	374	317
I-Quick Sort	992	694	1091	893	1091	893
Ackermann(3,6)	144	-	302	72	86	86
Recursive Qsort	2736	-	1368	1368	1642	1642
Puzzle (Subscript)	2796	2516	1398	1398	1398	1678
Puzzle (Pointer)	752	-	602	451	376	376
SED (Batch Editor)	17720	-	17720	10632	8860	8860
Towers Hanoi (18)	96	-	240	77	96	67

Case Study 21.5: Confidence Intervals



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.

Exercise 21.1

Analyze the data of Case study 21.2 using an additive model.

- 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.
- Are the processors significantly different?
- Discuss what indicators in the data, analysis, or plot would suggest that this is not a good model.

Exercise 21.2

Analyze the data of Table 21.18 using a multiplicative model and verify your analysis with the results presented in Table 21.19.

Exercise 21.3

Analyze the code size data of Table 21.23. Ignore the second column corresponding to 68000 for this exercise.

Answer the following:

a. What percentage of variation is explained by the processor?

b. What percentage of variation can be attributed to the workload?

c. Is there a significant (at 90% confidence) difference between any two processors?

Exercise 21.4

Repeat Exercise 21.3 with the 68000 column included.

Homework

- Submit answer to Exercise 21.1