Unit 6: Fractional Factorial Experiments at Three Levels

Source : Chapter 6 (Sections 6.1 - 6.6)

- Larger-the-better and smaller-the-better problems.
- Basic concepts for 3^k full factorial designs.
- Analysis of 3^k designs using orthogonal components system.
- Design of 3-level fractional factorials.
- Effect aliasing, resolution and minimum aberration in 3^{k-p} fractional factorial designs.
- An alternative analysis method using linear-quadratic system.

Seat Belt Experiment

- An experiment to study the effect of four factors on the pull strength of truck seat belts.
- Four factors, each at three levels (Table 1).
- Two responses : crimp tensile strength that must be at least 4000 lb and flash that cannot exceed 14 mm.
- 27 runs were conducted; each run was replicated three times as shown in Table 2.

Table 1: Factors and Levels, Seat-Belt Experiment

		Level		
	Factor	0	1	2
A.	pressure (psi)	1100	1400	1700
B.	die flat (mm)	10.0	10.2	10.4
C.	crimp length (mm)	18	23	27
D.	anchor lot (#)	P74	P75	P76

Design Matrix and Response Data, Seat-Belt Experiment

Table 2: Design Matrix and Response Data, Seat-Belt Experiment: first 14 runs

	Factor									
Run	*A*	*B*	*C*	*D*	Strength			Flash		
1	0	0	0	0	5164	6615	5959	12.89	12.70	12.74
2	0	0	1	1	5356	6117	5224	12.83	12.73	13.07
3	0	0	2	2	3070	3773	4257	12.37	12.47	12.44
4	0	1	0	1	5547	6566	6320	13.29	12.86	12.70
5	0	1	1	2	4754	4401	5436	12.64	12.50	12.61
6	0	1	2	0	5524	4050	4526	12.76	12.72	12.94
7	0	2	0	2	5684	6251	6214	13.17	13.33	13.98
8	0	2	1	0	5735	6271	5843	13.02	13.11	12.67
9	0	2	2	1	5744	4797	5416	12.37	12.67	12.54
10	1	0	0	1	6843	6895	6957	13.28	13.65	13.58
11	1	0	1	2	6538	6328	4784	12.62	14.07	13.38
12	1	0	2	0	6152	5819	5963	13.19	12.94	13.15
13	1	1	0	2	6854	6804	6907	14.65	14.98	14.40
14	1	1	1	0	6799	6703	6792	13.00	13.35	12.87

Design Matrix and Response Data, Seat-Belt Experiment (contd.)

Table 3: Design Matrix and Response Data, Seat-Belt Experiment: last 13 runs

	Factor									
Run	*A*	*B*	*C*	*D*		Strength			Flash	
15	1	1	2	1	6513	6503	6568	13.13	13.40	13.80
16	1	2	0	0	6473	6974	6712	13.55	14.10	14.41
17	1	2	1	1	6832	7034	5057	14.86	13.27	13.64
18	1	2	2	2	4968	5684	5761	13.00	13.58	13.45
19	2	0	0	2	7148	6920	6220	16.70	15.85	14.90
20	2	0	1	0	6905	7068	7156	14.70	13.97	13.66
21	2	0	2	1	6933	7194	6667	13.51	13.64	13.92
22	2	1	0	0	7227	7170	7015	15.54	16.16	16.14
23	2	1	1	1	7014	7040	7200	13.97	14.09	14.52
24	2	1	2	2	6215	6260	6488	14.35	13.56	13.00
25	2	2	0	1	7145	6868	6964	15.70	16.45	15.85
26	2	2	1	2	7161	7263	6937	15.21	13.77	14.34
27	2	2	2	0	7060	7050	6950	13.51	13.42	13.07

Larger-The-Better and Smaller-The-Better problems

- In the seat-belt experiment, the strength should be as high as possible and the flash as low as possible.
- There is no fixed nominal value for either strength or flash. Such type of problems are referred to as **larger-the-better** and **smaller-the-better** problems, respectively.
- For such problems increasing or decreasing the mean is more difficult than reducing the variation and should be done in the first step. (why?)
- Two-step procedure for larger-the-better problems:
  1. *Find factor settings that maximize E(y).*
  2. *Find other factor settings that minimize Var(y).*
- Two-step procedure for smaller-the-better problems:
  1. *Find factor settings that minimize E(y).*
  2. *Find other factor settings that minimize Var(y).*

Situations where three-level experiments are useful

- When there is a curvilinear relation between the response and a quantitative factor like temperature. It is not possible to detect such a curvature effect with two levels.
- A qualitative factor may have three levels (e.g., three types of machines or three suppliers).
- It is common to study the effect of a factor on the response at its current setting x_0 and two settings around x_0.

Analysis of 3^k designs using ANOVA

- We consider a simplified version of the seat-belt experiment as a 3^3 full factorial experiment with factors A, B, C.

- Since a 3^3 design is a special case of a multi-way layout, the analysis of variance method introduced in Section 3.5 can be applied to this experiment.

- We consider only the strength data for demonstration of the analysis.

- Using analysis of variance, we can compute the sum of squares for main effects A, B, C, interactions $A \times B$, $A \times C$, $B \times C$ and $A \times B \times C$ and the residual sum of squares. Details are given in Table 4.

- The break-up of the degrees of freedom will be as follows:
  - Each main effect has two degrees of freedom because each factor has three levels.
  - Each two-factor interaction has $(3-1) \times (3-1) = 4$ degrees of freedom.
  - The $A \times B \times C$ interaction has $(3-1) \times (3-1) \times (3-1) = 8$ degrees of freedom.
  - The residual degrees of freedom is $54 (= 27 \times (3-1))$, since there are three replicates.

Analysis of Simplified Seat-Belt Experiment

Table 4: ANOVA Table, Simplified Seat-Belt Experiment

Source	Degrees of Freedom	Sum of Squares	Mean Squares	F	p-value
A	2	34621746	17310873	85.58	0.000
B	2	938539	469270	2.32	0.108
C	2	9549481	4774741	23.61	0.000
$A \times B$	4	3298246	824561	4.08	0.006
$A \times C$	4	3872179	968045	4.79	0.002
$B \times C$	4	448348	112087	0.55	0.697
$A \times B \times C$	8	5206919	650865	3.22	0.005
residual	54	10922599	202270		
total	80	68858056			

Orthogonal Components System: Decomposition of $A \times B$ Interaction

- $A \times B$ has 4 degrees of freedom.
- $A \times B$ has two components denoted by AB and AB^2, each having 2 df.
- Let the levels of A and B be denoted by x_1 and x_2 respectively.
- AB represents the contrasts among the response values whose x_1 and x_2 satisfy

$$x_1 + x_2 = 0, 1, 2 (mod\ 3),$$

- AB^2 represents the contrasts among the response values whose x_1 and x_2 satisfy

$$x_1 + 2x_2 = 0, 1, 2 (mod\ 3).$$

Orthogonal Components System: Decomposition of $A \times B \times C$ Interaction

- $A \times B \times C$ has 8 degrees of freedom.
- It can be further split up into four components denoted by ABC, ABC^2, AB^2C and AB^2C^2, each having 2 df.
- Let the levels of A, B and C be denoted by x_1, x_2 and x_3 respectively.
- ABC, ABC^2, AB^2C and AB^2C^2 represent the contrasts among the three groups of (x_1, x_2, x_3) satisfying each of the four systems of equations,

$$\begin{aligned} x_1 + x_2 + x_3 &= 0, 1, 2 (mod 3), \\ x_1 + x_2 + 2x_3 &= 0, 1, 2 (mod 3), \\ x_1 + 2x_2 + x_3 &= 0, 1, 2 (mod 3), \\ x_1 + 2x_2 + 2x_3 &= 0, 1, 2 (mod 3). \end{aligned}$$

Uniqueness of Representation

- To avoid ambiguity, *the convention that the coefficient for the first nonzero factor is 1 will be used*.
- ABC^2 is used instead of A^2B^2C, even though the two are equivalent.
- For A^2B^2C, there are three groups satisfying

$$\begin{aligned} 2x_1+2x_2+x_3 &= 0,1,2(mod 3), \\ \text{equivalently,} \quad 2\times(2x_1+2x_2+x_3) &= 2\times(0,1,2)(mod 3), \\ \text{equivalently,} \quad x_1+x_2+2x_3 &= 0,2,1(mod 3), \end{aligned}$$

which corresponds to ABC^2 by relabeling of the groups. Hence ABC^2 and A^2B^2C are *equivalent*.

Representation of AB and AB^2

Table 5: Factor A and B Combinations (x_1 denotes the levels of factor A and x_2 denotes the levels of factor B)

x_1	x_2: 0	1	2
0	αi (y_{00})	βk (y_{01})	γj (y_{02})
1	βj (y_{10})	γi (y_{11})	αk (y_{12})
2	γk (y_{20})	αj (y_{21})	βi (y_{22})

- α, β, γ correspond to (x_1, x_2) with $x_1 + x_2 = 0, 1, 2 (mod 3)$ resp.
- i, j, k correspond to (x_1, x_2) with $x_1 + 2x_2 = 0, 1, 2 (mod 3)$ resp.

Connection with Graeco-Latin Square

- In Table 5, (α, β, γ) forms a Latin Square and (i, j, k) forms another Latin Square.
- (α, β, γ) and (i, j, k) jointly form a Graeco-Latin Square. This implies that SS for (α, β, γ) and SS for (i, j, k) are *orthogonal*.
- $SS_{AB} = 3n[(\bar{y}_\alpha - \bar{y}_\cdot)^2 + (\bar{y}_\beta - \bar{y}_\cdot)^2 + (\bar{y}_\gamma - \bar{y}_\cdot)^2]$,
  where $\bar{y}_\cdot = (\bar{y}_\alpha + \bar{y}_\beta + \bar{y}_\gamma)/3$ and n is the number of replicates,
  $\bar{y}_\alpha = \frac{1}{3}(y_{00} + y_{12} + y_{21})$, etc.
- Similarly, $SS_{AB^2} = 3n[(\bar{y}_i - \bar{y}_\cdot)^2 + (\bar{y}_j - \bar{y}_\cdot)^2 + (\bar{y}_k - \bar{y}_\cdot)^2]$.

Analysis using the Orthogonal components system

- For the simplified seat-belt experiment, $\bar{y}_\alpha = 6024.407$, $\bar{y}_\beta = 6177.815$ and $\bar{y}_\gamma = 6467.0$, so that $\bar{y}_{\cdot} = 6223.074$ and

$$\begin{aligned} SS_{AB} &= (3)(9)[(6024.407-6223.074)^2 + (6177.815-6223.074)^2 \\ &\quad +(6467.0-6223.074)^2] = 2727451. \end{aligned}$$

- Similarly, $SS_{AB^2} = 570795$.
- See ANOVA table on the next page.

ANOVA : Simplified Seat-Belt Experiment

Source	Degrees of Freedom	Sum of Squares	Mean Squares	F	p-value
A	2	34621746	17310873	85.58	0.000
B	2	938539	469270	2.32	0.108
C	2	9549481	4774741	23.61	0.000
$A \times B$	4	3298246	824561	4.08	0.006
AB	2	2727451	1363725	6.74	0.002
AB^2	2	570795	285397	1.41	0.253
$A \times C$	4	3872179	968045	4.79	0.002
AC	2	2985591	1492796	7.38	0.001
AC^2	2	886587	443294	2.19	0.122
$B \times C$	4	448348	112087	0.55	0.697
BC	2	427214	213607	1.06	0.355
BC^2	2	21134	10567	0.05	0.949
$A \times B \times C$	8	5206919	650865	3.22	0.005
ABC	2	4492927	2246464	11.11	0.000
ABC^2	2	263016	131508	0.65	0.526
AB^2C	2	205537	102768	0.51	0.605
AB^2C^2	2	245439	122720	0.61	0.549
residual	54	10922599	202270		
total	80	68858056			

Analysis of Simplified Seat-Belt Experiment (contd)

- The significant main effects are A and C.
- Among the interactions, $A \times B$, $A \times C$ and $A \times B \times C$ are significant.
- We have difficulty in interpretations when only one component of the interaction terms become significant. What is meant by "$A \times B$ is significant"?
  - Here AB is significant but AB^2 is not.
  - Is $A \times B$ significant because of the significance of AB alone ?
  - For the original Seat-Belt Experiment, we have $AB = CD^2$.
- Similarly, AC is significant, but not AC^2. How to interpret the significance of $A \times C$?
- This difficulty in interpreting the significant interaction effects can be avoided by using Linear-Quadratic Systems.

Why three-level fractional factorial ?

- Run size economy : it is not economical to use a 3^4 design with 81 runs unless the experiment is not costly.

- If a 3^4 design is used for the experiment, its 81 degrees of freedom would be allocated as follows:

	Main	Interactions		
	Effects	2-Factor	3-Factor	4-Factor
#	8	24	32	16

- Using effect hierarchy principle, one would argue that 3fi's and 4fi's are not likely to be important. Out of a total of 80 df, 48 correspond to such effects !

Defining a 3^{4-1} Experiment

- Returning to the original seat-belt experiment, it employs a one-third fraction of the 3^4 design. This is denoted as a 3^{4-1} design.
- The design is constructed by choosing the column for factor D (lot #) to be equal to Column A + Column B + Column $C (mod\ 3)$.
- This relationship can be represented by the notation

$$D = ABC.$$

- If $x_1, \ldots, x_4$ are used to represent these four columns, then

$$\begin{aligned} x_4 = x_1 + x_2 + x_3 (mod\ 3), &\text{or equivalently} \\ x_1 + x_2 + x_3 + 2x_4 = 0 (mod\ 3), & \end{aligned} \tag{1}$$

which can be represented by

$$\mathbf{I} = ABCD^2.$$

Aliasing Patterns of the Seat-Belt Experiment

- The aliasing patterns can be deduced from the defining relation. For example, by adding $2x_1$ to both sides of (1), we have

$$2x_1 = 3x_1 + x_2 + x_3 + 2x_4 = x_2 + x_3 + 2x_4 (mod\ 3),$$

- This means that A and BCD^2 are *aliased*. (Why?)
- By following the same derivation, it is easy to show that the following effects are aliased:

$$\begin{array}{lllll}
A & = & BCD^2 & = & AB^2C^2D, \\
B & = & ACD^2 & = & AB^2CD^2, \\
C & = & ABD^2 & = & ABC^2D^2, \\
D & = & ABC & = & ABCD, \\
AB & = & CD^2 & = & ABC^2D, \\
AB^2 & = & AC^2D & = & BC^2D, \\
AC & = & BD^2 & = & AB^2CD, \\
AC^2 & = & AB^2D & = & BC^2D^2, \\
AD & = & AB^2C^2 & = & BCD, \\
AD^2 & = & BC & = & AB^2C^2D^2, \\
BC^2 & = & AB^2D^2 & = & AC^2D^2, \\
BD & = & AB^2C & = & ACD, \\
CD & = & ABC^2 & = & ABD.
\end{array} \tag{2}$$

Clear and Strongly Clear Effects

- If three-factor interactions are assumed negligible, from the aliasing relations in (2), A, B, C, D, AB^2, AC^2, AD, BC^2, BD and CD can be estimated.
- These main effects or components of two-factor interactions are called **clear** because they are not aliased with any other main effects or two-factor interaction components.
- A two-factor interaction, say $A \times B$, is called *clear* if both of its components, AB and AB^2, are clear.
- Note that each of the six two-factor interactions has only one component that is clear; the other component is aliased with one component of another two-factor interaction. For example, for $A \times B$, AB^2 is clear but AB is aliased with CD^2.
- A main effect or two-factor interaction component is said to be **strongly clear** if it is not aliased with any other main effects, two-factor or three-factor interaction components. A two-factor interaction is said to be *strongly clear* if both of its components are strongly clear.

A 3^{5-2} Design

- 5 factors, 27 runs.
- The one-ninth fraction is defined by $\mathbf{I} = ABD^2 = AB^2CE^2$, from which two additional relations can be obtained:

$$\mathbf{I} = (ABD^2)(AB^2CE^2) = A^2CD^2E^2 = AC^2DE$$

and

$$\mathbf{I} = (ABD^2)(AB^2CE^2)^2 = B^2C^2D^2E = BCDE^2.$$

Therefore the defining contrast subgroup for this design consists of the following defining relation:

$$\mathbf{I} = ABD^2 = AB^2CE^2 = AC^2DE = BCDE^2. \tag{3}$$

Resolution and Minimum Aberration

- Let A_i be to denote the number of words of length i in the subgroup and $W = (A_3, A_4, \ldots)$ to denote the wordlength pattern.

- Based on W, the definitions of **resolution** and **minimum aberration** are the same as given before in Section 5.2.

- The subgroup defined in (3) has four words, whose lengths are 3, 4, 4, and 4. and hence $W = (1,3,0)$. Another 3^{5-2} design given by $D = AB, E = AB^2$ has the defining contrast subgroup,

$$\mathbf{I} = ABD^2 = AB^2E^2 = ADE = BDE^2,$$

  with the wordlength pattern $W = (4,0,0)$. According to the aberration criterion, the first design has less aberration than the second design.

- Moreover, it can be shown that the first design has minimum aberration.

General 3^{k-p} Design

- A 3^{k-p} design is a fractional factorial design with k factors in 3^{k-p} runs.
- It is a 3^{-p}th fraction of the 3^k design.
- The fractional plan is defined by p independent generators.
- How many factors can a 3^{k-p} design study?

  $$(3^n - 1)/2, \text{ where } n = k - p.$$

  This design has 3^n runs with the independent generators $x_1, x_2, \ldots, x_n$. We can obtain altogether $(3^n - 1)/2$ orthogonal columns as different combinations of $\sum_{i=1}^{n} \alpha_i x_i$ with $\alpha_i = 0$, 1 or 2, where at least one α_i should not be zero and the first nonzero α_i should be written as "1" to avoid duplication.
- For n=3, the $(3^n - 1)/2 = 13$ columns were given in Table 6.5 of WH book.
- A general algebraic treatment of 3^{k-p} designs can be found in Kempthorne (1952).

Simple Analysis Methods: Plots and ANOVA

- Start with making a main effects plot and interaction plots to see what effects might be important.
- This step can be followed by a formal analysis like analysis of variance and half-normal plots.

The strength data will be considered first. The location main effect and interaction plots are given in Figures 1 and 2. The main effects plot suggests that factor A is the most important followed by factors C and D. The interaction plots in Figure 2 suggest that there may be interactions because the lines are not parallel.

Main Effects Plot of Strength Location

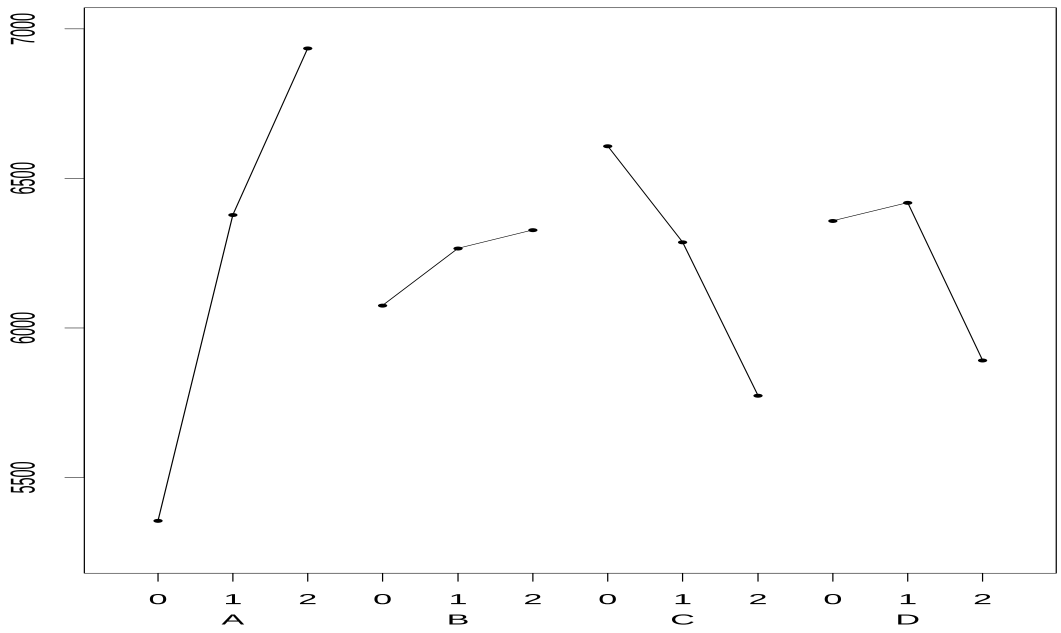

Figure 1: Main Effects Plot of Strength Location, Seat-Belt Experiment

Interaction Plots of Strength Location

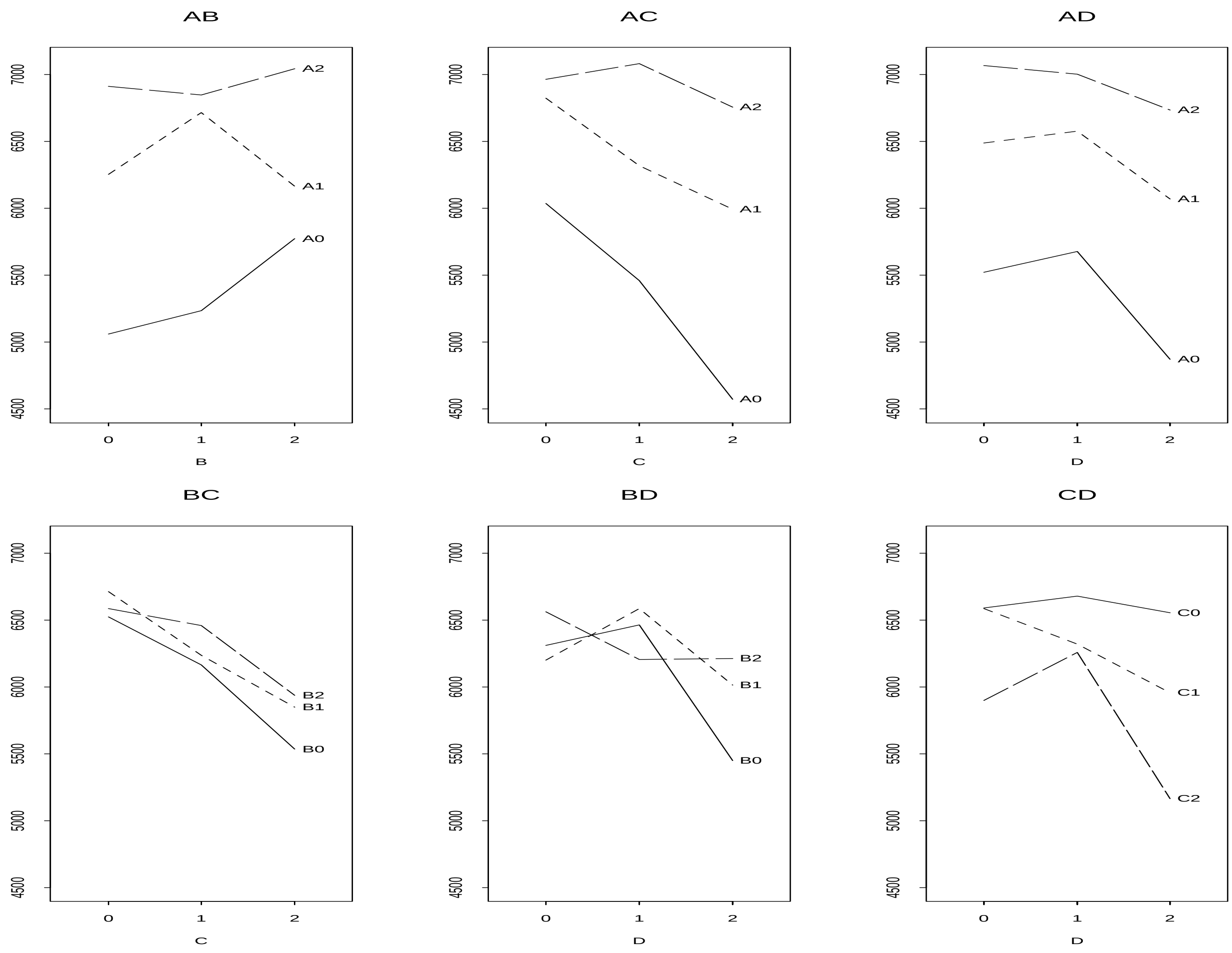

Figure 2: Interaction Plots of Strength Location, Seat-Belt Experiment

ANOVA Table for Strength Location

Source	Degrees of Freedom	Sum of Squares	Mean Squares	F	p-value
A	2	34621746	17310873	85.58	0.000
B	2	938539	469270	2.32	0.108
$AB = CD^2$	2	2727451	1363725	6.74	0.002
AB^2	2	570795	285397	1.41	0.253
C	2	9549481	4774741	23.61	0.000
$AC = BD^2$	2	2985591	1492796	7.38	0.001
AC^2	2	886587	443294	2.19	0.122
$BC = AD^2$	2	427214	213607	1.06	0.355
BC^2	2	21134	10567	0.05	0.949
D	2	4492927	2246464	11.11	0.000
AD	2	263016	131508	0.65	0.526
BD	2	205537	102768	0.51	0.605
CD	2	245439	122720	0.61	0.549
residual	54	10922599	202270		

Analysis of Strength Location, Seat-Belt Experiment

- In equation (2), the 26 degrees of freedom in the experiment were grouped into 13 sets of effects. The corresponding ANOVA table gives the SS values for these 13 effects.
- Based on the p values in the ANOVA Table, clearly the factor A, C and D main effects are significant.
- Also two aliased sets of effects are significant, $AB = CD^2$ and $AC = BD^2$.
- These findings are consistent with those based on the main effects plot and interaction plots. In particular, the significance of AB and CD^2 is supported by the $A \times B$ and $C \times D$ plots and the significance of AC and BD^2 by the $A \times C$ and $B \times D$ plots.

Analysis of Strength Dispersion (i.e., $\ln s^2$) Data

The corresponding strength main effects plot and interaction plots are displayed in Figures 3 and 4.

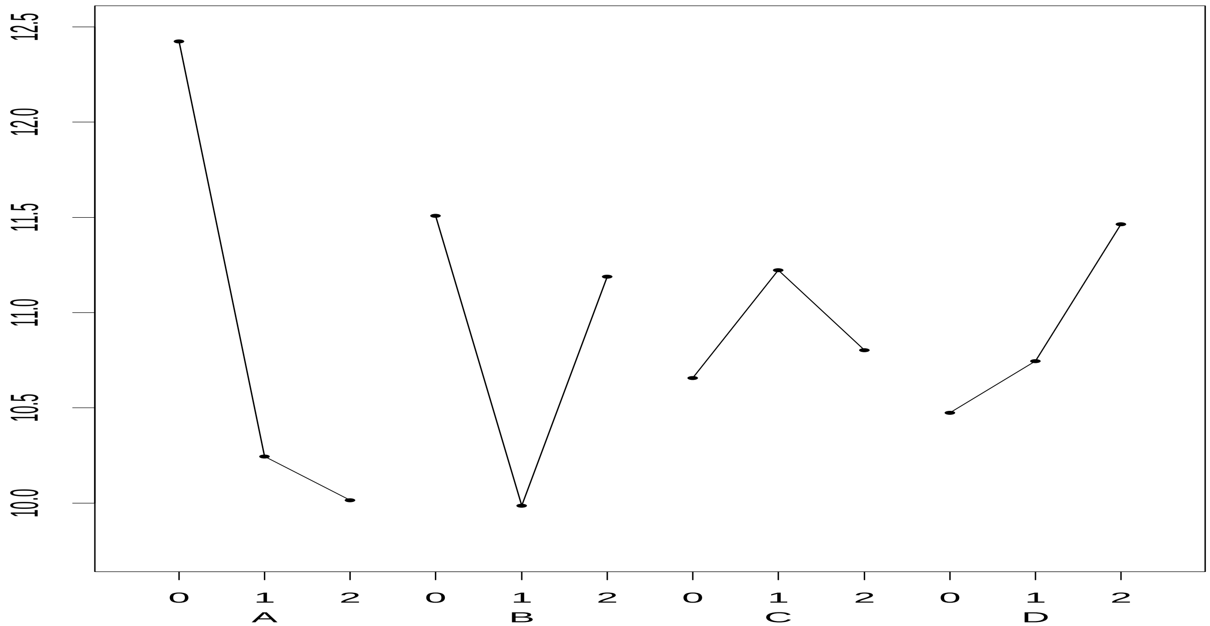

Figure 3: Main Effects Plot of Strength Dispersion, Seat-Belt Experiment

Interaction Plots of Strength Dispersion

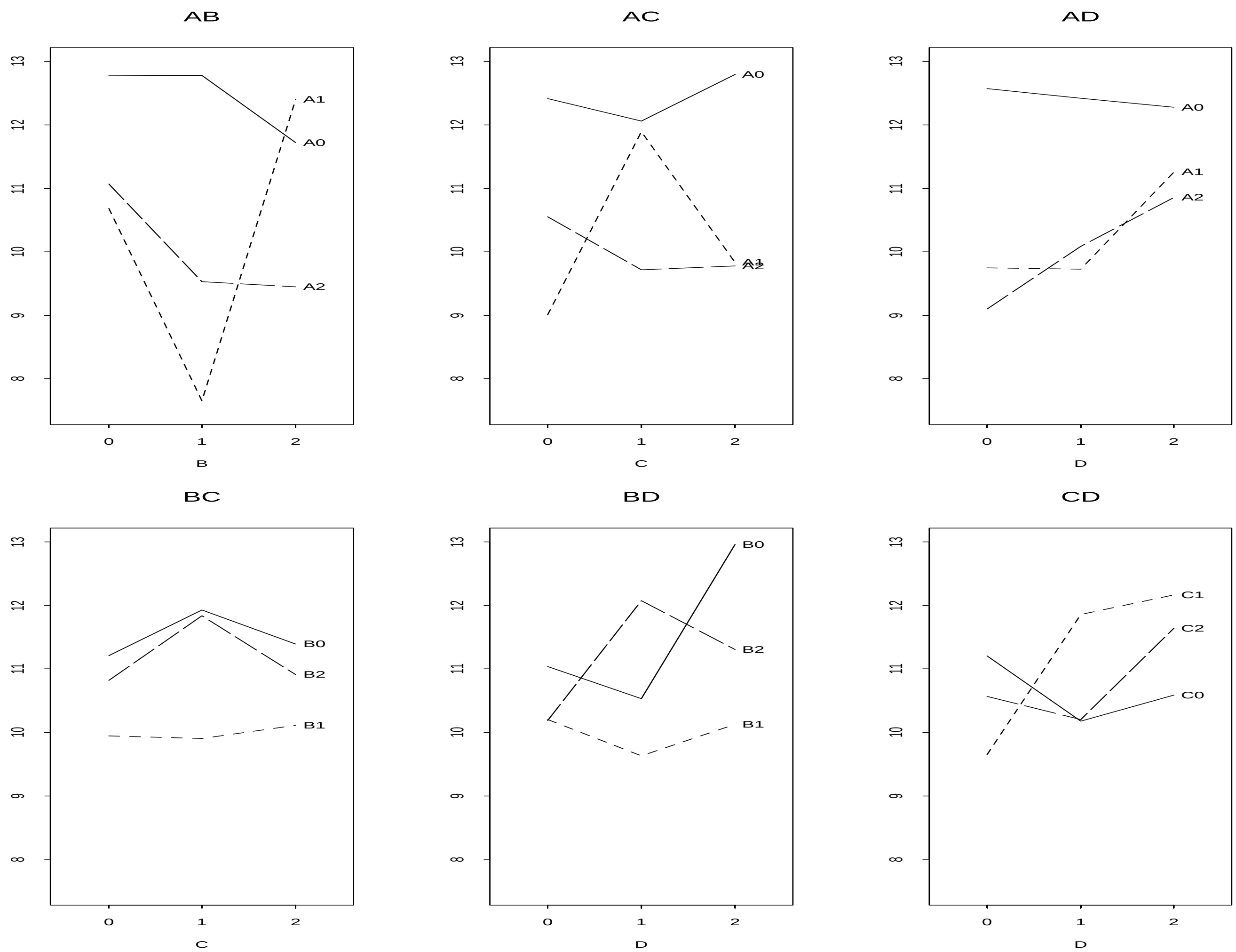

Figure 4: Interaction Plots of Strength Dispersion, Seat-Belt Experiment

Half-Normal Plots

- Since there is no replication for the dispersion analysis, ANOVA cannot be used to test effect significance.
- Instead, a half-normal plot can be drawn as follows. The 26 df's can be divided into 13 groups, each having two df's. These 13 groups correspond to the 13 rows in the ANOVA table of page 27.
- The two degrees of freedom in each group can be decomposed further into a linear effect and a quadratic effect with the contrast vectors $\frac{1}{\sqrt{2}}(-1, 0, 1)$ and $\frac{1}{\sqrt{6}}(1, -2, 1)$, respectively, where the values in the vectors are associated with the $\ln s^2$ values at the levels (0, 1, 2) for the group.
- Because the linear and quadratic effects are standardized and orthogonal to each other, these 26 effect estimates can be plotted on the half-normal probability scale as in Figure 5.

Half-Normal Plot

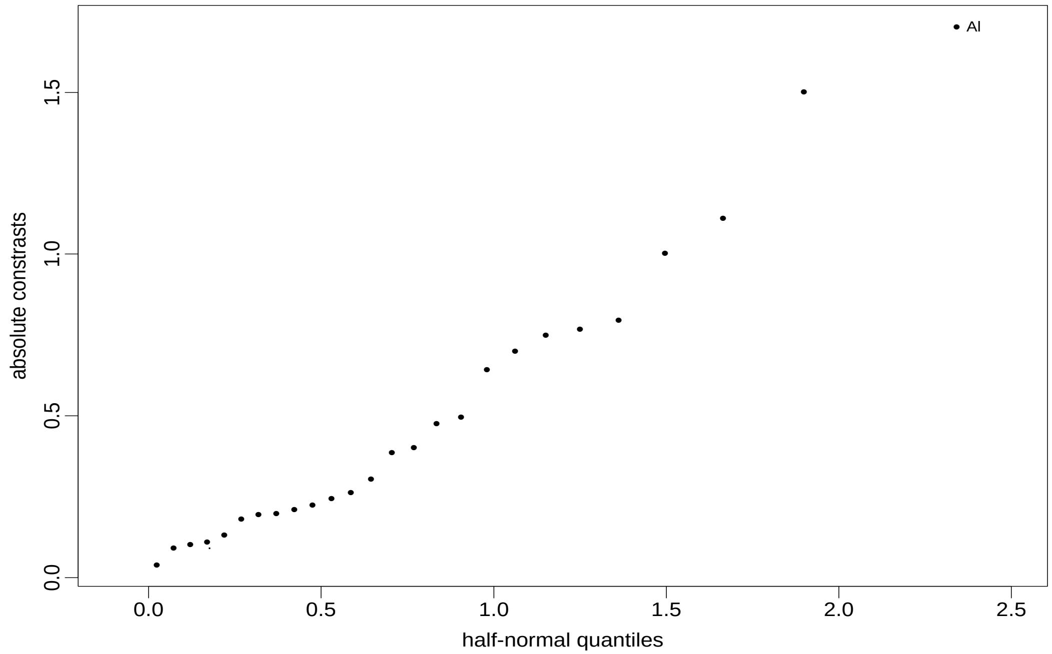

Figure 5: Half-Normal Plot of Strength Dispersion Effects, Seat-Belt Experiment

Informal analysis of the plot suggests that the factor *A* linear effect may be significant. This can be confirmed by using Lenth's method. The t_{PSE} value for the *A* linear effect is 3.99, which has a p value of 0.003(IER) and 0.050 (EER).

Analysis Summary

- A similar analysis can be performed to identify the flash location and dispersion effects. See Section 6.5 of WH book.
- We can determine the optimal factor settings that maximize the *strength location* by examining the main effects plot and interaction plots in Figures 1 and 2 that correspond to the significant effects identified in the ANOVA table.
- We can similarly determine the optimal factor settings that minimize the *strength dispersion*, *the flash location* and *flash dispersion*, respectively.
- The most obvious findings: level 2 of factor A be chosen to maximize strength while level 0 of factor A be chosen to minimize flash.
- There is an obvious conflict in meeting the two objectives. Trade-off strategies for handling multiple characteristics and conflicting objectives need to be considered (See Section 6.7 of WH).

An Alternative Analysis Method : Linear-Quadratic System

In the seat-belt experiment, the factors A, B and C are quantitative. The two degrees of freedom in a quantitative factor, say A, can be decomposed into the linear and quadratic components.

Letting y_0, y_1 and y_2 represent the observations at level 0, 1 and 2, then the *linear effect* is defined as

$$y_2 - y_0$$

and the *quadratic effect* as

$$(y_2 + y_0) - 2y_1,$$

which can be re-expressed as the difference between two consecutive linear effects $(y_2 - y_1) - (y_1 - y_0)$.

Linear and Quadratic Effects

Mathematically, the linear and quadratic effects are represented by two mutually orthogonal vectors:

$$
\begin{aligned}
A_l &= \frac{1}{\sqrt{2}}(-1, 0, 1), \\
A_q &= \frac{1}{\sqrt{6}}(1, -2, 1).
\end{aligned}
\tag{4}
$$

- For the sake of brevity, they are also referred to as the l and q effects.
- The scaling constants $\sqrt{2}$ and $\sqrt{6}$ yield vectors with unit length.
- The linear (or quadratic) effect is obtained by taking the inner product between A_l (or A_q) and the vector $\mathbf{y} = (y_0, y_1, y_2)$. For factor B, B_l and B_q are similarly defined.

Linear and Quadratic Effects (contd)

- Then the four degrees of freedom in the $A \times B$ interaction can be decomposed into four mutually orthogonal terms: $(AB)_{ll}, (AB)_{lq}, (AB)_{ql}, (AB)_{qq}$, which are defined as follows: for $i, j = 0, 1, 2$,

$$\begin{aligned}(AB)_{ll}(i,j) &= A_l(i)B_l(j),\\ (AB)_{lq}(i,j) &= A_l(i)B_q(j),\\ (AB)_{ql}(i,j) &= A_q(i)B_l(j),\\ (AB)_{qq}(i,j) &= A_q(i)B_q(j).\end{aligned} \tag{5}$$

  They are called the **linear-by-linear**, **linear-by-quadratic**, **quadratic-by-linear** and **quadratic-by-quadratic** interaction effects. They are also referred to as the $l \times l$, $l \times q$, $q \times l$ and $q \times q$ effects.

- It is easy to show that they are orthogonal to each other.

Linear and Quadratic Effects (contd)

Using the nine level combinations of factors A and B, $y_{00}, \ldots, y_{22}$ given in Table 5, the contrasts $(AB)_{ll}$, $(AB)_{lq}$, $(AB)_{ql}$, $(AB)_{qq}$ can be expressed as follows:

$(AB)_{ll}$: $\frac{1}{2}\{(y_{22} - y_{20}) - (y_{02} - y_{00})\}$,

$(AB)_{lq}$: $\frac{1}{2\sqrt{3}}\{(y_{22} + y_{20} - 2y_{21}) - (y_{02} + y_{00} - 2y_{01})\}$,

$(AB)_{ql}$: $\frac{1}{2\sqrt{3}}\{(y_{22} + y_{02} - 2y_{12}) - (y_{20} + y_{00} - 2y_{10})\}$,

$(AB)_{qq}$: $\frac{1}{6}\{(y_{22} + y_{20} - 2y_{21}) - 2(y_{12} + y_{10} - 2y_{11}) + (y_{02} + y_{00} - 2y_{01})\}$.

- An $(AB)_{ll}$ interaction effect measures the difference between the conditional linear B effects at levels 0 and 2 of factor A.
- A significant $(AB)_{ql}$ interaction effect means that there is curvature in the conditional linear B effect over the three levels of factor A.
- The other interaction effects $(AB)_{lq}$ and $(AB)_{qq}$ can be similarly interpreted.

Analysis of designs with resolution at least *V*

- For designs of at least resolution V, all the main effects and two-factor interactions are clear. Then, further decomposition of these effects according to the linear-quadratic system allows all the effects (each with one degree of freedom) to be compared in a half-normal plot.
- Note that for effects to be compared in a half-normal plot, they should be uncorrelated and have the same variance.

Analysis of designs with resolution smaller than V

- For designs with resolution III or IV, a more elaborate analysis method is required to extract the maximum amount of information from the data.
- Consider the 3^{3-1} design with $C = AB$ whose design matrix is given in Table 6.

Table 6: Design Matrix for the 3^{3-1} Design

Run	A	B	C
1	0	0	0
2	0	1	1
3	0	2	2
4	1	0	1
5	1	1	2
6	1	2	0
7	2	0	2
8	2	1	0
9	2	2	1

- Its main effects and two-factor interactions have the aliasing relations:

$$A = BC^2, B = AC^2, C = AB, AB^2 = BC = AC. \tag{6}$$

Analysis of designs with resolution III (contd)

- In addition to estimating the six degrees of freedom in the main effects A, B and C, there are two degrees of freedom left for estimating the three aliased effects AB^2, BC and AC, which, as discussed before, are difficult to interpret.
- Instead, consider using the remaining two degrees of freedom to estimate any pair of the $l \times l, l \times q, q \times l$ or $q \times q$ effects between A, B and C.
- Suppose that the two interaction effects taken are $(AB)_{ll}$ and $(AB)_{lq}$. Then the eight degrees of freedom can be represented by the model matrix given in Table 7.

Table 7: A System of Contrasts for the 3^{3-1} Design

Run	A_l	A_q	B_l	B_q	C_l	C_q	$(AB)_{ll}$	$(AB)_{lq}$
1	-1	1	-1	1	-1	1	1	-1
2	-1	1	0	-2	0	-2	0	2
3	-1	1	1	1	1	1	-1	-1
4	0	-2	-1	1	0	-2	0	0
5	0	-2	0	-2	1	1	0	0
6	0	-2	1	1	-1	1	0	0
7	1	1	-1	1	1	1	-1	1
8	1	1	0	-2	-1	1	0	-2
9	1	1	1	1	0	-2	1	1

Analysis of designs with resolution *III* (contd)

- Because any component of $A \times B$ is orthogonal to A and to B, there are only four non-orthogonal pairs of columns whose correlations are:

$$\begin{aligned} Corr((AB)_{ll}, C_l) &= -\sqrt{\tfrac{3}{8}}, \\ Corr((AB)_{ll}, C_q) &= -\tfrac{1}{\sqrt{8}}, \\ Corr((AB)_{lq}, C_l) &= \tfrac{1}{\sqrt{8}}, \\ Corr((AB)_{lq}, C_q) &= -\sqrt{\tfrac{3}{8}}. \end{aligned} \tag{7}$$

- Obviously, $(AB)_{ll}$ and $(AB)_{lq}$ can be estimated in addition to the three main effects.

- Because the last four columns are not mutually orthogonal, they cannot be estimated with full efficiency.

- The estimability of $(AB)_{ll}$ and $(AB)_{lq}$ demonstrates an advantage of the linear-quadratic system over the orthogonal components system. For the same design, the AB interaction component cannot be estimated because it is aliased with the main effect C.

Analysis Strategy for Qualitative Factors

- For a qualitative factor like factor D (lot number) in the seat-belt experiment, the linear contrast $(-1, 0, +1)$ may make sense because it represents the comparison between levels 0 and 2.

- On the other hand, the "quadratic" contrast $(+1, -2, +1)$, which compares level 1 with the average of levels 0 and 2, makes sense only if such a comparison is of practical interest. For example, if levels 0 and 2 represent two internal suppliers, then the "quadratic" contrast measures the difference between internal and external suppliers.

Analysis Strategy for Qualitative Factors (contd)

- When the quadratic contrast makes no sense, two out of the following three contrasts can be chosen to represent the two degrees of freedom for the main effect of a qualitative factor:

$$D_{01} = \begin{cases} -1 & & 0 & \\ 1 & \text{for level} & 1 & \text{of factor } D, \\ 0 & & 2 & \end{cases}$$

$$D_{02} = \begin{cases} -1 & & 0 & \\ 0 & \text{for level} & 1 & \text{of factor } D, \\ 1 & & 2 & \end{cases}$$

$$D_{12} = \begin{cases} 0 & & 0 & \\ -1 & \text{for level} & 1 & \text{of factor } D, \\ 1 & & 2 & \end{cases}$$

Analysis Strategy for Qualitative Factors (contd)

- Mathematically, they are represented by the standardized vectors:

$$D_{01} = \frac{1}{\sqrt{2}}(-1,1,0), D_{02} = \frac{1}{\sqrt{2}}(-1,0,1), D_{12} = \frac{1}{\sqrt{2}}(0,-1,1).$$

- These contrasts are not orthogonal to each other and have pairwise correlations of $1/2$ or $-1/2$.
- On the other hand, each of them is readily interpretable as a comparison between two of the three levels.
- The two contrasts should be chosen to be of interest to the investigator. For example, if level 0 is the main supplier and levels 1 and 2 are minor suppliers, then D_{01} and D_{02} should be used.

Qualitative and Quantitative Factors

- The interaction between a quantitative factor and a qualitative factor, say $A \times D$, can be decomposed into four effects.

- As in (5), we define the four interaction effects as follows:

$$
\begin{aligned}
(AD)_{l,01}(i,j) &= A_l(i)D_{01}(j), \\
(AD)_{l,02}(i,j) &= A_l(i)D_{02}(j), \\
(AD)_{q,01}(i,j) &= A_q(i)D_{01}(j), \\
(AD)_{q,02}(i,j) &= A_q(i)D_{02}(j).
\end{aligned}
\tag{8}
$$

Variable Selection Strategy

Since many of these contrasts are not mutually orthogonal, a general purpose analysis strategy cannot be based on the orthogonality assumption. Therefore, the following variable selection strategy is recommended.

(i) For a quantitative factor, say A, use A_l and A_q for the A main effect.

(ii) For a qualitative factor, say D, use D_l and D_q if D_q is interpretable; otherwise, select two contrasts from D_{01}, D_{02}, and D_{12} for the D main effect.

(iii) For a pair of factors, say X and Y, use the products of the two contrasts of X and the two contrasts of Y (chosen in (i) or (ii)) as defined in (5) or (8) to represent the four degrees of freedom in the interaction $X \times Y$.

(iv) Using the contrasts defined in (i)-(iii) for all the factors and their two-factor interactions as candidate variables, perform a stepwise regression or subset selection procedure to identify a suitable model. To avoid incompatible models, use the effect heredity principle to rule out interactions whose parent factors are both not significant.

(v) If all the factors are quantitative, use the original scale, say x_A, to represent the linear effect of A, x_A^2 the quadratic effect and $x_A^i x_B^j$ the interaction between x_A^i and x_B^j. This works particularly well if some factors have unevenly spaced levels.

Analysis of Seat-Belt Experiment

- Returning to the seat-belt experiment, although the original design has resolution IV, its capacity for estimating two-factor interactions is much better than what the definition of resolution IV would suggest.

- After estimating the four main effects, there are still 18 degrees of freedom available for estimating some components of the two-factor interactions.

- From (2), A, B, C and D are estimable and only one of the two components in each of the six interactions $A \times B$, $A \times C$, $A \times D$, $B \times C$, $B \times D$ and $C \times D$ is estimable.

- Because of the difficulty of providing a physical interpretation of an interaction component, a simple and efficient modeling strategy that does not throw away the information in the interactions is to consider the contrasts (A_l, A_q), (B_l, B_q), (C_l, C_q) and (D_{01}, D_{02}, D_{12}) for the main effects and the 30 products between these four groups of contrasts for the interactions.

Analysis of Seat-Belt Experiment (contd)

- Using these 39 contrasts as the candidate variables, the variable selection procedure was applied to the data.
- Performing a stepwise regression on the strength data (response y_1), the following model with an R^2 of 0.811 was identified:

$$\begin{aligned}\hat{y}_1 &= 6223.0741 + 1116.2859A_l - 190.2437A_q + 178.6885B_l \\ &\quad -589.5437C_l + 294.2883(AB)_{ql} + 627.9444(AC)_{ll} \\ &\quad -191.2855D_{01} - 468.4190D_{12} - 486.4444(CD)_{l,12}\end{aligned} \tag{9}$$

- Note that this model obeys effect heredity. The A, B, C and D main effects and $A \times B$, $A \times C$ and $C \times D$ interactions are significant. In contrast, the simple analysis from the previous section identified the A, C and D main effects and the $AC(=BD^2)$ and $AB(=CD^2)$ interaction components as significant.

Analysis of Seat-Belt Experiment (contd)

- Performing a stepwise regression on the flash data (response y_2), the following model with an R^2 of 0.857 was identified:

$$\begin{aligned} \hat{y}_2 &= 13.6657 + 1.2408A_l + 0.1857B_l \\ &\quad -0.8551C_l + 0.2043C_q - 0.9406(AC)_{ll} \\ &\quad -0.3775(AC)_{ql} - 0.3765(BC)_{qq} - 0.2978(CD)_{l,12} \end{aligned} \tag{10}$$

- Again, the identified model obeys effect heredity. The A, B, and C main effects and $A \times C$, $B \times C$ and $C \times D$ interactions are significant. In contrast, the simple analysis from the previous section identified the A and C main effects and the $AC(= BD^2)$, AC^2 and BC^2 interaction components as significant.

Comments on Board