model for × : $y \sim A+B+C+AB+AB^2+AC+AC^2+\cdots+\boxed{E}$ ← constant variance.

ANOVA : Simplified Seat-Belt Experiment

o: multi-way layout
x: orthogonal component.

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	448343	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			

can be further decomposed

span

span

span

span.

ω Ω

ω, Ω_1

Ω_1, Ω_2

Ω_2, Ω_3

Ω_3, Ω

⊕

⊕

Q: Why ⊕ hold?
① sequential ANOVA
② orthogonality

Q: why not so significant even though $SS_{A\times B} > SS_{AB}$?

$\{0, 1, 2\}$ $\boxed{1} \cdots \boxed{k}$

For a 3^k design, # of words $= 3^k$

# of OC's $= 3^k - 1/2$ ← two words form same OC.

2df

In 3^k, these OC's are mutually (design matrix) orthogonal.

Analysis of Simplified Seat-Belt Experiment (contd)

- The significant main effects are A and C.

C A

- Among the interactions, $A \times B$, $A \times C$ and $A \times B \times C$ are significant.

↳ multi-way layout o

for × ANOVA

- We have difficulty in interpretations when only one component of the interaction terms become significant. What is meant by "$A \times B$ is significant"?

why?

meaning? (check LNp.1-11) ← space spaned by AB, AB^2.

Q: should we give such conclusion?

  – 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 $A \times C$?

later lecture, it's coded in a more meanful way than orthogonal component.

- This difficulty in interpreting the significant interaction effects can be avoided by using Linear-Quadratic Systems.

v **Reading**: textbook, 6.3

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

$\left(\binom{4}{2}2^4/2\right)\times 2$ $\left(\binom{4}{3}2^3/2\right)\times 2$

- 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 !

higher-order interaction usually
① difficult to interpret
② insignificant.

Defining a 3^{4-1} Experiment

4 factors
3-level
run size
defining relation

- 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

A B C D=word of AB,C

full 3^3

$$D = ABC.$$

A B C D ABCD^2

full 3^4

0 0 0 1 1 1 2 2 2

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

$$x_4 = x_1 - x_2 + x_3 (mod\ 3), \text{or equivalently}$$

$0 = 3x_4 = +2x_4$ $+2x_4$

$$x_1 + x_2 + x_3 + 2x_4 = 0 (mod\ 3), \qquad (1)$$

which can be represented by

identity element "0" in "+"

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

defining relation: cut a design into 1/3,
the other 2/3 $\{ x_1+x_2+x_3+2x_4 = 1 (mod\ 3)$
$= 2 (\ \vdots \ )$

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

$$\overset{A^2}{2x_1} = 3x_1 + x_2 + x_3 + 2x_4 = \overset{BCD^2}{x_2 + x_3 + 2x_4} \ (mod\ 3),$$

① A & BCD^2 span same 2-dim space
② A & BCD^2 seperate all observations into same 3 groups

- 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:

※ orthogonal components are aliased, not interaction

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

※ independent defining relations cause 3^{k-p} OC's to be aliased together. Why?

① 1/3 of runs in original design matrix
② count d.f. $27-1/2=13$ alias sets. $3^4-1/2=40$, $40-1/13=3$

$I = ABCD^2$

$A = A^2BCD^2 \rightarrow \boxed{AB^2C^2D}$
$A^2 = \cancel{A^3}BCD^2 \rightarrow \boxed{BCD^2}$

$I = ABCD^2 \ (= A^2B^2C^2D)$

$A = A^2BCD^2 = \cancel{A^3}B^2C^2D$
$\boxed{AB^2C^2D}$ $\boxed{BCD^2}$

Clear and Strongly Clear Effects

Note: it can be defined on $AB, AB^2, \cdots$ or $A \times B, A \times C, \cdots$

- 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.
- for orthogonal components: 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.
- for interaction: A two-factor interaction, say $A \times B$, is called *clear* if both of its components, AB and AB^2, are clear.

none of those 6 2f.i. 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.

4 d.f. 6-dim space. $C \times D$ 4.d.f.

- 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

3^{k-p}

$D=AB$
$E=AB^2C$

$\begin{cases} x_1 + x_2 + \quad +2x_4 + \quad = 0 \text{ (eq1)} \\ x_1 + 2x_2 + x_3 + \quad +2x_5 = 0 \text{ (eq2)} \end{cases}$

eq1×2=0
eq2×2=0

- 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: indep. defining relations

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

$0 = (\text{eq1} + \text{eq2})$
$0 = (\text{eq1} + \text{eq2}) \times 2$

and

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

$0 = \text{eq1} + 2\,\text{eq2}$
$0 = (\text{eq1} + 2\,\text{eq2}) \times 2$

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

$A^2B^2D \quad A^2BC^2E \quad A^2CD^2E^2 \quad B^2C^2D^2E$

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

an alias set → $A = A^2BD^2 = B^2D = A^2B^2CE^2 = BC^2E = A^2C^2DE = CD^2E^2 = ABCDE^2 = AB^2C^2D^2E$

$ABD \quad BD^2 \quad ABC^2E \quad ACD^2E^2$

$3^p = 3^2 = 9$ OC's in the alias set

Resolution and Minimum Aberration

\# of factors involved in a word.

- 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. 2 levels, same argument
- 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,

$$\mathbf{I} = ABD^2 = AB^2E^2 = ADE = BDE^2, \text{ (exercise)}$$

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 $\boxed{k \text{ factors}}$ in 3^{k-p} runs.
- It is a 3^{-p}th fraction of the 3^k design. ← 3^{k-p}: run size.
- The fractional plan is defined by p independent generators. → 3^p OC's in an alias set → $(3^p-1)/2$ OC's in DCS
- How many factors can a 3^{k-p} design study?

  \# of OC's (including M.E): n factors can generate → $(3^n - 1)/2$, where $n = k - p$.

  $A_1 \cdots A_n$ OC OC ··· OC (A_{n+1}, A_{n+2}, ..., A); full 3^n

  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).

v Reading: textbook, 6.4

Simple Analysis Methods: Plots and ANOVA

for 3^{4-1} FFD.

Initial data analysis.

- 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. ← no replicates, all d.f. used by effects

  ↳ for data with replicates (constant variance)

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.

Be aware the danger of

true: $y=X_1\beta_1+X_2\beta_2+\varepsilon$ —(*)

fitted: $y=X_1\beta_1+\varepsilon$ —(Δ)

Note: $\hat{\beta}_1$ under (*) and (Δ) could be different if X_1 & X_2 are not orthogonal.

Main Effects Plot of Strength Location

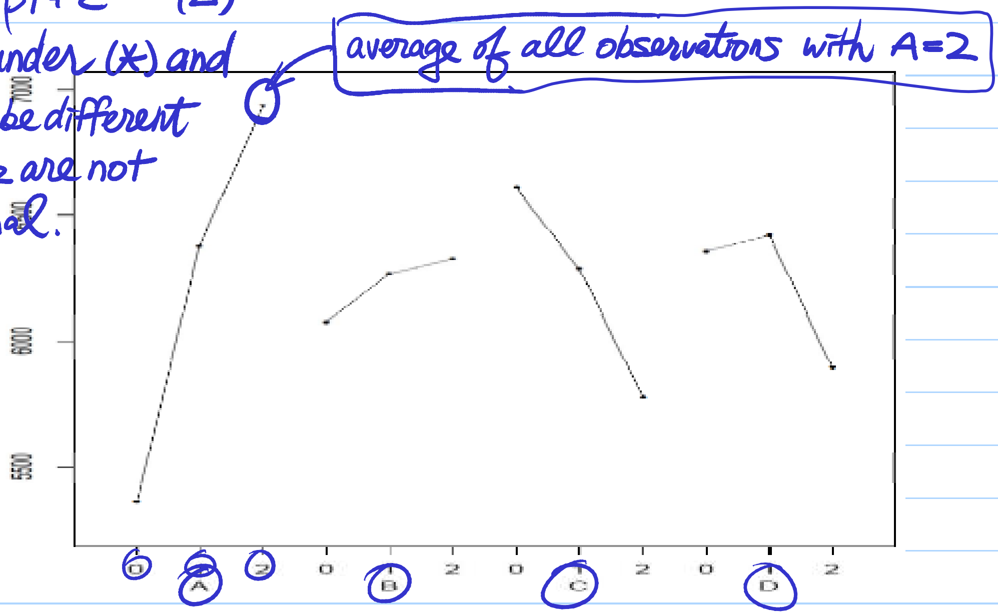

Figure 1: Main Effects Plot of Strength Location, Seat-Belt Experiment

lines are parallel or not? → if no parallel, interaction (4 d.f.) could be significant.

Interaction Plots of Strength Location

A×B SS: AB, AB^2

C×D: CD, CD^2 — If large

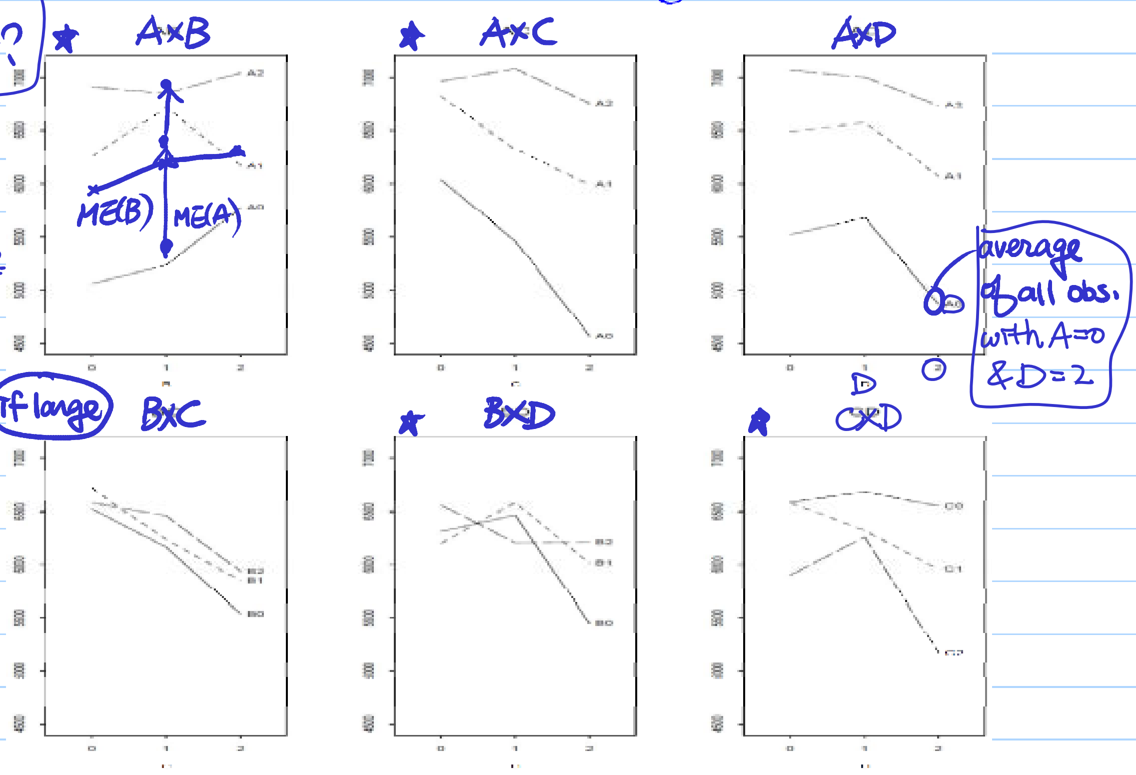

Figure 2: Interaction Plots of Strength Location, Seat-Belt Experiment