3 levels full & fractional factorial designs ⟷ (c.f.) 2 level case

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). ← all level combinations: $3^4 = 81$
- Two responses : crimp tensile strength that must be at least 4000 lb and flash that cannot exceed 14 mm.

  strength ≥ 4000, ↑ better

  flash ≤ 14, ↓ better
- 27 runs were conducted; each run was replicated three times as shown in Table 2.

Table 1: Factors and Levels, Seat-Belt Experiment

c.f. → 2 levels use (−1, 1)

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

A, B: ← equally spaced

C: ← not equally spaced

Design Matrix and Response Data, Seat-Belt Experiment

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

Run	Factor A	Factor B	Factor C	Factor 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	12.75	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

3 replicates for each run

→ carry more information about $\mathrm{Var}(y_{\underset{\sim}{x}})$

response

$\bar{y}_{\underset{\sim}{x}}$, $\ln(\widehat{\mathrm{Var}}(y_{\underset{\sim}{x}}))$

↳ no replicates

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

$3^4 - 1 = 27$ — not perform all level combinations

Q: what's a good choice of subset of all level combinations?

v **Reading**: textbook, 6.1

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

$\leftrightarrow$ nominal-the-best (c.f.)

Recall: ① min $Var(y_x)$ ② adjust $E(y_x)$

- 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?)

a or few factors that can easily increase/decrease the mean $E(y_x)$ is usually difficult to find in this case

- 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).*

sometimes, combined into one step.

$y_x \geq c$, $E(y_x - c)^2 = Var(y_x) + [E(y_x) - c]^2$

v **Reading**: textbook, 6.2

① quadratic effect
② increasing/decreasing rate not constant
③ increase then decrease
decrease " increase

Situations where three-level experiments are useful

Low M H

- 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). Note: VV cannot predict X.
- 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.

x_0- ← improve? — x_0 — improve? → x_0+

why study it? →
① x_0 could be a very good (best) setting
x_0- x_0 x_0+
② If x_0-/x_0+ better than x_0, how much better?

ANOVA-type → more suitable for qualitative factors.

Analysis of 3^k designs using ANOVA

← full factorial.

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

27×3=81 observations

26 effects

81−27=54 for residuals $=27\times(3-1)$

fit mean

Analysis of Simplified Seat-Belt Experiment

model: $y \sim A+B+C+A\times B+A\times C+B\times C+A\times B\times C+\varepsilon$

constant variance

Table 4: ANOVA Table, Simplified Seat-Belt Experiment

sequential ANOVA

$(RSS_\omega - RSS_\Omega)/(df_\omega - df_\Omega)$

$\dfrac{RSS_{full}}{df_{full}}$

with orthogonality

Source	Degrees of Freedom	Sum of Squares	Mean Squares	F	p-value
A (3−1)	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			

$\omega: y \sim A+B+C$

$\Omega: y \sim A+B+C+A\times B$

can be further decomposed into several smaller subspace

pay attention to what ω & Ω are → interpretation.

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

FYI, mod operation can work for s^k or s^{k-p} when s is a prime #. For s being prime power, $s=4,8,9$, use operation in Galois field.

- $A\times B$ 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

a linear equation in finite geometry → $x_1 + x_2 = 0, 1, 2 \ (mod\ 3).$

$\to 2x_1+2x_2 \ (mod\ 3)$

A	B	AB	A^2B^2	AB^2	A^2B
0	0	0	0	0	0
0	1	1 →	2	2	1
0	2	2	1	1 →	2
1	0	1 →	2	1 →	2
1	1	2	1	0	0
1	2	0	0	2	1
2	0	2	1	2	1
2	1	0	0	1 →	2
2	2	1 →	2	0	0

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

equivalent 2-level calculation. $0 \to 1$, $1 \to -1$ identity.

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

$2(x_1+2x_2) = 2x_1+4x_2 = 2x_1+x_2$

$AB = x_1+x_2 \ (mod\ 2)$

A		B		AB	
0	1	0	1	0	1
0	1	1	−1	1	−1
1	−1	0	1	1	−1
1	−1	1	−1	0	1

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

$A^2B^2C^2$ A^2B^2C

- $A \times B \times C$ 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.

A^2BC^2 A^2BC

- 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

they separate the data into 3 same groups

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

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

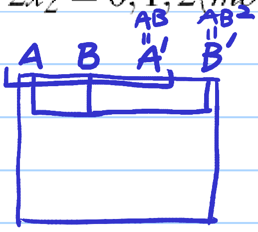
Representation of AB and AB^2

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})

A B

- 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, $\overset{A'}{(\alpha, \beta, \gamma)}$ forms a Latin Square and $\overset{B'}{(i, j, k)}$ 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 $\boxed{orthogonal}$. ← design matrix orthogonal.
- $SS_{AB} = \overset{3 \text{ levels}}{3}n[(\bar{y}_{\alpha} - \bar{y}_{\cdot})^2 + (\bar{y}_{\beta} - \bar{y}_{\cdot})^2 + (\bar{y}_{\gamma} - \bar{y}_{\cdot})^2]$,
where $\bar{y}_{\cdot} = (\bar{y}_{\alpha} + \bar{y}_{\beta} + \bar{y}_{\gamma})/3$ and n is the number of replicates,
how many observations in y_{00}.
$\bar{y}_{\alpha} = \frac{1}{3}(y_{00} + y_{12} + y_{21})$, etc.
- Similarly, $SS_{AB^2} = 3n[(\bar{y}_i - \bar{y}_{\cdot})^2 + (\bar{y}_j - \bar{y}_{\cdot})^2 + (\bar{y}_k - \bar{y}_{\cdot})^2]$.

Its model matrix columns span a 2-dim space

AB		
0	$y, \cdots, y$	α
1	$y, \cdots, y$	β
2	$y, \cdots, y$	γ

treat the column (LNp. 1-8) of AB as a 3-level factor, SS_{AB} is the variation of y that can be explained by the factor.

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

$$SS_{AB} = (3)(9)[(6024.407-6223.074)^2+(6177.815-6223.074)^2 +(6467.0-6223.074)^2] = \underline{2727451}.$$

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

Why? ∵ orthogonality.

Note: $SS_{\underline{AB}} + SS_{\underline{AB^2}} \overset{?}{=} 3298246 = SS_{\underline{A \times B}}$

Note: notation: × → interaction

without × → orthogonal component.

Note: all orthogonal components are defined on full factorial design.