Leaf Spring Experiment

★ response : free height
★ treatment factors : B, C, D, E, Q
all 2 levels — −1, +1
★ Exp'tal units : a spring
homogeneous — EU1 · · · EU48

- Five factors at two levels, use a 16-run design with three replicates for each run. It is a 2^{5-1} design, 1/2 fraction of the 2^5 design.

full factorial 2^5: # of all level combinations $= 2\times2\times2\times2\times2 = 2^5 = 32$

2^{n-k} design: ① n: # of factors ② 2: 2 levels ③ k: # of independent defining words ④ 2^{n-k}: run size ⑤ 2^{-k}: fraction; $2^{n-k} = 2^n/2^k$

Q: There are enough EUs (48) for 2^5 design (32). Why not do 2^5?
2^{5-1} design repeat 3 times (r) — 2^5 design (16)
Hint. 2 types of df.

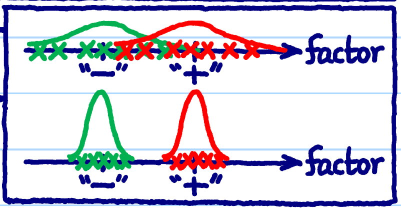
Table 1: Factors and Levels, Leaf Spring Experiment

	Factor (qualitative or quantitative ?)	Level −	Level +
B.	high heat temperature (°F)	1840	1880
C.	heating time (seconds)	23	25
D.	transfer time (seconds)	10	12
E.	hold down time (seconds)	2	3
Q.	quench oil temperature (°F)	130-150	150-170

heating stage: B, C
forming stage: D, E
quenching stage: Q

★ conceptual model :
2^5: $Z \sim \beta_0 +$ all factorial effects $+ \varepsilon$ (31)
2^{5-1}: $Z \sim \beta_0 +$ some factorial effects $+ \varepsilon$ (15)

measurement errors/errors in predictors (LM, LN p. 9-7)



- response y = free height of spring, target = 8.0 inches.

Goal : get y as close to 8.0 as possible (nominal-the-best problem). → build 2 models:
one for $\mu_x = E(y_x)$ ⟷ factors
one for $\sigma_x^2 = Var(y_x)$ ⟷ factors

Leaf Spring Experiment: Design Matrix and Data

Q: Which 16 level combinations should we choose out of the 32 ones?

Table 2: Design Matrix and Free Height Data, Leaf Spring Experiment

B	C	D	BCD	E	Q	Free Height (y) (replicated)			$\bar{y}_i$	s_i^2	$\ln s_i^2$
−	+	+	−	−	−	7.78	7.78	7.81	7.7900	0.0003	−8.1117
+	+	+	+	+	−	8.15	8.18	7.88	8.0700	0.0273	−3.6009
−	−	+	+	+	−	7.50	7.56	7.50	7.5200	0.0012	−6.7254
+	−	+	−	−	−	7.59	7.56	7.75	7.6333	0.0104	−4.5627
−	+	−	+	+	−	7.94	8.00	7.88	7.9400	0.0036	−5.6268
+	+	−	−	−	−	7.69	8.09	8.06	7.9467	0.0496	−3.0031
−	−	−	−	−	−	7.56	7.62	7.44	7.5400	0.0084	−4.7795
+	−	−	+	+	−	7.56	7.81	7.69	7.6867	0.0156	−4.1583
−	+	+	−	−	+	7.50	7.25	7.12	7.2900	0.0373	−3.2888
+	+	+	+	+	+	7.88	7.88	7.44	7.7333	0.0645	−2.7406
−	−	+	+	+	+	7.50	7.56	7.50	7.5200	0.0012	−6.7254
+	−	+	−	−	+	7.63	7.75	7.56	7.6467	0.0092	−4.6849
−	+	−	+	+	+	7.32	7.44	7.44	7.4000	0.0048	−5.3391
+	+	−	−	−	+	7.56	7.69	7.62	7.6233	0.0042	−5.4648
−	−	−	−	−	+	7.18	7.18	7.25	7.2033	0.0016	−6.4171
+	−	−	+	+	+	7.81	7.50	7.59	7.6333	0.0254	−3.6717

unreplicated: $\bar{y}_i$, s_i^2, $\ln s_i^2$

Only contains 16 out of all 32 level combanations
⇒ It is called a Fractional Factorial Design (FFD)
⇒ can estimate 15 factorial effects

Q: Which 15 effects?

Q: What is the model matrix of this FFD?

Under a conceptual model $\xrightarrow{\text{codings}}$ model matrix X
for replicated response : $y = X\beta^* + \varepsilon^*$
for unreplicate response : $\bar{y} = X\beta_1 + \varepsilon_1$
$\ln s^2 = X\beta_2 + \varepsilon_2$

Why Use Fractional Factorial Designs (FFDs)?

$2^5-1=$

- If a 2^5 design is used for the experiment, its 31 degrees of freedom would be allocated as follows:

	Main Effects	Interactions 2-Factor	3-Factor	4-Factor	5-Factor
#	5	$\binom{5}{2}=10$	$\binom{5}{3}=10$	$\binom{5}{4}=5$	$\binom{5}{5}=1$

more important ← 15 | 16 → less important

- Using effect hierarchy principle, one would argue that 4fi's, 5fi and even 3fi's are not likely to be important. There are 10+5+1 = 16 (16/32 = 1/2) such effects, half of the total runs! Using a 2^5 design can be wasteful (unless 32 runs cost about the same as 16 runs.) cf. $2^{5-1}=16$ runs → 15 effects
- Use of an FFD instead of full factorial design is usually done for economic reasons. Since there is **no free lunch** , what **price to pay**? See next slide.

★ one of the most important concepts in DOE

❖ **Reading**: textbook, 5.1

check LNp.6-2

Effect Aliasing and Defining Relation

- In the design matrix, col $B\times$ col $C\times$ col D = col E. That means,

It estimate $2(\beta_E+\beta_{BCD})$

$$2\hat{\beta}_E = ME(E) = \bar{y}(E+) - \bar{y}(E-) = \bar{y}(BCD+) - \bar{y}(BCD-) = INT(BCD) = 2\hat{\beta}_{BCD}$$

Therefore the design is not capable of distinguishing E from BCD. The main effect E is **aliased** with the interaction BCD. Notationally,

defined on model matrix → $E = BCD$ or $\mathbf{I} = BCDE$, ← Note. It is defined on design matrix.

$E^2=BCDE$; $B+C+D+E=0 \pmod 2$, where $B,C,D,E \in \{0,1\}$ (exercise) — group of additivity

Intercept in the model matrix

where $\mathbf{I}$ (= column of +'s) is the identity element in the group of multiplications. — group of multiplication ← $\{+1,-1\}$

(Notice the mathematical similarity between aliasing and confounding. What is the difference?) check LNp.5-36

btwn treatment & block effects | btwn treatment effects

# of all effects (including I): $2^n=2^5=32$ effects $= 16 \times 2$
In general, 2^{n-k} — # of alias sets ; In general, 2^k — # of effects in an alias set

- $\mathbf{I} = BCDE$ is the **defining relation** for the 2^{5-1} design. ($2^4=16$ runs → 15 df. for effect estimation → used on 15 alias sets)

It implies all the 15 effect aliasing relations :

an alias set ← $B = CDE$, $C = BDE$, $D = BCE$, $E = BCD$,
$BC = DE$, $BD = CE$, $BE = CD$,
$Q = BCDEQ$, $BQ = CDEQ$, $CQ = BDEQ$, $DQ = BCEQ$,
$EQ = BCDQ$, $BCQ = DEQ$, $BDQ = CEQ$, $BEQ = CDQ$.

effects formed by factors B, C, D, Q

In the model matrix of 2^{n-k}, each column corresponds to an alias set, & columns of different alias sets are orthogonal.

design matrix of a full factorial 2^3

Three 2-level factors

model matrix of 2^3 with all factorial effects

	A	B	C	a	b	c	ab	ac	bc	abc	I
1	−1	−1	−1	−1	−1	−1	1	1	1	−1	1
2	−1	−1	1	−1	−1	1	1	−1	−1	1	1
3	−1	1	−1	−1	1	−1	−1	1	−1	1	1
4	−1	1	1	−1	1	1	−1	−1	1	−1	1
5	1	−1	−1	1	−1	−1	−1	−1	1	1	1
6	1	−1	1	1	−1	1	−1	1	−1	−1	1
7	1	1	−1	1	1	−1	1	−1	−1	−1	1
8	1	1	1	1	1	1	1	1	1	1	1

Recall. The 3 Q in LNp.6-2

- effect^2=I

$a^2=b^2=c^2=(ab)^2=(ac)^2$

$=(bc)^2=(abc)^2=I$

- defining relation

$I = abc$ defining word

- effect aliasing

$a = a^2bc = bc$

$a=bc$

$b=ac$

$c=ab$

an alias set

a full factorial 2^2

	A	B	C	a	b	c	ab	ac	bc	abc	I
2	−1	−1	1	−1	−1	1	1	−1	−1	1	1
3	−1	1	−1	−1	1	−1	−1	1	−1	1	1
5	1	−1	−1	1	−1	−1	−1	−1	1	1	1
8	1	1	1	1	1	1	1	1	1	1	1

- un-aliased effects are orthogonal

$E(y)=\beta_0+\beta_a x_a+\beta_b x_b+\beta_c x_c$
$+\beta_{ab}x_{ab}+\beta_{bc}x_{bc}+\beta_{ac}x_{ac}+\beta_{abc}x_{abc}$
$\to(\beta_0+\beta_{abc})+(\beta_a+\beta_{bc})x_a$
$+(\beta_b+\beta_{ac})x_b+(\beta_c+\beta_{ab})x_{ab}$

design matrix of the 2^{3-1}, $I=abc$

model matrix — unidetifiable / identifiable

model matrix formed by all effects of A & B

the analyses for full factorial design can be applied on fractional factorial

design matrix of 2^4 full factorial

Four 2-level factors

model matrix of 2^4 with all factorial effects

$(abcd)\times(cd)=I$
$=abc^2d^2=ab$

	A	B	C	D	a	b	c	d	ab	ac	ad	bc	bd	cd	abc	abd	acd	bcd	abcd	I
1	−1	−1	−1	−1	−1	−1	−1	−1	1	1	1	1	1	1	−1	−1	−1	−1	1	1
2	−1	−1	−1	1	−1	−1	−1	1	1	1	−1	1	−1	−1	−1	1	1	1	−1	1
3	−1	−1	1	−1	−1	−1	1	−1	1	−1	1	−1	1	−1	1	−1	1	1	−1	1
4	−1	−1	1	1	−1	−1	1	1	1	−1	−1	−1	−1	1	1	1	−1	−1	1	1
5	−1	1	−1	−1	−1	1	−1	−1	−1	1	1	−1	−1	1	1	1	−1	1	−1	1
6	−1	1	−1	1	−1	1	−1	1	−1	1	−1	−1	1	−1	1	−1	1	−1	1	1
7	−1	1	1	−1	−1	1	1	−1	−1	−1	1	1	−1	−1	−1	1	1	−1	1	1
8	−1	1	1	1	−1	1	1	1	−1	−1	−1	1	1	1	−1	−1	−1	1	−1	1
9	1	−1	−1	−1	1	−1	−1	−1	−1	−1	−1	1	1	1	1	1	1	−1	−1	1
10	1	−1	−1	1	1	−1	−1	1	−1	−1	1	1	−1	−1	1	−1	−1	1	1	1
11	1	−1	1	−1	1	−1	1	−1	−1	1	−1	−1	1	−1	−1	1	−1	1	1	1
12	1	−1	1	1	1	−1	1	1	−1	1	1	−1	−1	1	−1	−1	1	−1	−1	1
13	1	1	−1	−1	1	1	−1	−1	1	−1	−1	−1	−1	1	−1	−1	1	1	1	1
14	1	1	−1	1	1	1	−1	1	1	−1	1	−1	1	−1	−1	1	−1	−1	−1	1
15	1	1	1	−1	1	1	1	−1	1	1	−1	1	−1	−1	1	−1	−1	−1	−1	1
16	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

How to obtain a $2^{4-2} = 2^4/2^2$ FFD?

$I=abcd$ & $I=cd$

design matrix of the 2^{4-2} {1,4,13,16}

4 runs ⇒ 4 df. ⇒ can etimate 3 (joint) effects

Which 3?

defining contrast subgroup

$I = abcd = cd = ab$ defining words

$a = a^2bcd = acd = a^2b$
$\quad\;\; = bcd \qquad\quad\; = b$

It estimates the parameter $\beta_a+\beta_{bcd}+\beta_{acd}+\beta_b$

$c = abd = d = abc$

$ac = bd = ad = bc$

In runs {1,4,13,16}, factors A & C form a 2^2 full factorial

Clear Effects

How to detect? The information is in the alias sets.

Intuition. For important effects, it would be better for them to be aliased with higher-order interactions (Why?) effect hierarchy principle

- A main effect or two-factor interaction (2fi) is called **clear** if it is not aliased with any other m.e.'s or 2fi's and **strongly clear** if it is not aliased with any other m.e.'s, 2fi's or 3fi's. Therefore a clear effect is *estimable* under the assumption of negligible 3-factor and higher-order interactions and a strongly clear effect is *estimable* under the weaker assumption of negligible 4-factor and higher-order interactions.

  $B_{ABC}=0$ ⋮ $B_{ABCD}=0$ ⋮

  $\to B_{ABCD}=0, \cdots, B_{ABCDE}=0, \cdots, \cdots$

- In the 2^{5-1} design with $\mathbf{I}=BCDE$, which effects are clear and strongly clear?

  $\mathbf{I}=BCD$ cf. cf.

  check LNp.6-4

  Ans: B, C, D, E are clear, Q, BQ, CQ, DQ, EQ are strongly clear.

- Consider the alternative plan 2^{5-1} design with $\mathbf{I}=BCDEQ$. (It is said to have resolution V because the length of the defining word is 5 while the previous plan has resolution IV.) It can be verified that all five main effects are strongly clear and all 10 2fi's are clear. (*Do the derivations*). This is a very good plan because *each* of the 15 degrees of freedom is either clear or strongly clear.

  # of letters (factors) involved in a word

  Intuition. It would be better to have longer words in DCSG.

DCSG Defining Contrast Subgroup for 2^{k-p} Designs

df. for effect (alias set) estimation

can be used to ① construct design matrix ② study effect aliasing ③ develop criteria for choosing "good" designs

check LNp.6-6. What can be the 3rd word?

- A 2^{k-p} design has k factors, 2^{k-p} runs, and it is a 2^{-p}th fraction of the 2^k design. The fraction is defined by p *independent* defining words. $\to W_1, \cdots, W_p$

  group defined in Algebra

  The group formed by these p words is called the **defining contrast subgroup**. It has 2^p-1 words plus the identity element $\mathbf{I}$.

  $\to \{\mathbf{I}, W_1, \cdots, W_p, W_1W_2, \cdots, W_{p-1}W_p,$
  generators $\to$ $W_1W_2W_3, \cdots, W_1W_2\cdots W_p\}$

- **Resolution** = shortest wordlength among the 2^p-1 words. ← Why? # of words in the group $=2^p$

  6 factors 1, 2, ⋯, 6 → 2 generators → generate $2^2=4$ words (including $\mathbf{I}$)

- Example: A $2^{6-\boxed{2}}$ design with $\mathbf{5}=\mathbf{12}$ and $\mathbf{6}=\mathbf{134}$. The two independent defining words are $\mathbf{I}=\mathbf{125}$ and $\mathbf{I}=\mathbf{1346}$. Then $\mathbf{I}=\mathbf{125}\times\mathbf{1346}=\mathbf{23456}$. The defining contrast subgroup $=\{\mathbf{I},\mathbf{125},\mathbf{1346},\mathbf{23456}\}$. The design has resolution III.

  $2^4=16$ runs

	WL
125	3
1346	4
23456	5

a column: an alias set

model matrix

a group (Algebra)

16 runs 2^4 full factorial

1	2	3	4	5 = 12	13	14	23	⋯34 123⋯	6 = 134	⋯234	1234	I
+	+	+	+	+	+	+	+		+		+	+
+	+	+	−	+	+	−	+		−		−	+
+	+	−	+	+	−	+	−		−		−	+
+	−	+	+	−	+	+	−		+		−	+
−	+	+	+	−	−	−	+		−		−	+
		⋮		⋮	⋮	⋮	⋮		⋮		⋮	⋮
−	−	−	−	+	+	+	+		−		+	+

2^{6-2} design matrix

$1+2+5=0 \pmod 2$
$1+3+4+6=0 \pmod 2$
$2+3+4+5+6=0 \pmod 2$
where $1,2,3,4,5,6 \in \{0,1\}$

Deriving Aliasing Relations for the 2^{6-2} design

$\to = 2^4$

- For the same 2^{6-2} design, the defining contrast subgroup is

word (defined on design matrix)

$$\mathbf{I} = 125 = 1346 = 23456.$$

$125 \times 1346 = 23456$

most serious aliasing?
1=25
2=15
5=12

most serious aliasing?
13=46
14=36
16=34

I	= 125	= 1346	= 23456,
×1	= ×25	= 346	= 123456,
×2	= ×15	= 12346	= 3456,
√3	= 1235	= 146	= 2456,
√4	= 1245	= 136	= 2356,
×5	= ×12	= 13456	= 2346,
√6	= 1256	= 134	= 2345,
×13	= 235	= ×46	= 12456,
×14	= 245	= ×36	= 12356,
×16	= 256	= ×34	= 12345,
√23	= 135	= 1246	= 456,
√24	= 145	= 1236	= 356,
√26	= 156	= 1234	= 345,
√35	= 123	= 1456	= 246,
√45	= 124	= 1356	= 236,
√56	= 126	= 1345	= 234

5=12 in LNp.6-8

cf.

an alias set

same as the defining contrast subgroup.

there are 16 alias sets

factors {1,2,3,4} forms a full factorial

{1,3,4,6} not form a full factorial

effect (appears in model matrix) (1)

All the 15 degrees of freedom (each is a coset in group theory) are identified.

- It has the clear effects: 3, 4, 6, 23, 24, 26, 35, 45, 56. It has resolution III.

WordLength Pattern and Resolution

$\to$ WLP $\leftarrow$ defined on defining contrast subgroup

- Define A_i = number of defining words of length i.
- $W = (A_3, A_4, A_5, \ldots)$ is called the **wordlength pattern**. $\xleftrightarrow{\text{cf.}}$ WLP of block scheme (LNp.5-40)
  - In this design, $W = (1, 1, 1, 0)$.
  - It is required that $A_1 = 0$ and $A_2 = 0$. (Why? No main effect is allowed to be aliased with the intercept or another main effect.)
- **Resolution** = smallest r such that $A_r \geq 1$. $\leftarrow$ resolution is a function of WLP
- **Maximum resolution criterion**: For fixed k and p, choose a 2^{k-p} design with maximum resolution.
- **Rules for Resolution IV and V Designs:**

It is better to have higher resolution

(i) In any resolution IV design, the main effects are clear.

resolution IV
e.g. I=ABCD
AB=CD
AC=BD
AD=BC

(ii) In any resolution V design, the main effects are strongly clear and the two-factor interactions are clear.

resolution V
e.g. I=ABCDE
AB=CDE, BD=ACE
AC=BDE, BE=ACD
AD=BCE, CD=ABE
AE=BCD, CE=ABD
BC=ADE, DE=ABC

(2)

(iii) Among the resolution IV designs with given k and p, those with the largest number of clear two-factor interactions are the best.

(*) A Projective Rationale for Resolution

projection of design, 2 viewpoints:
① matrix viewpoint ⟶ submatrix of design matrix
② geometric viewpoint

A constraint (e.g., I = a word) on a submatrix must be a constraint on the whole design matrix.

Why?

- For a resolution R design, its projection onto any $R-1$ factors is a full factorial in the $R-1$ factors. This would allow *effects of all orders among the $R-1$ factors to be estimable*. (**Caveat**: it makes the assumption that other factors are inert.) cf.

By (*) "some effects are inert" (effect hierarchy or effect sparsity)

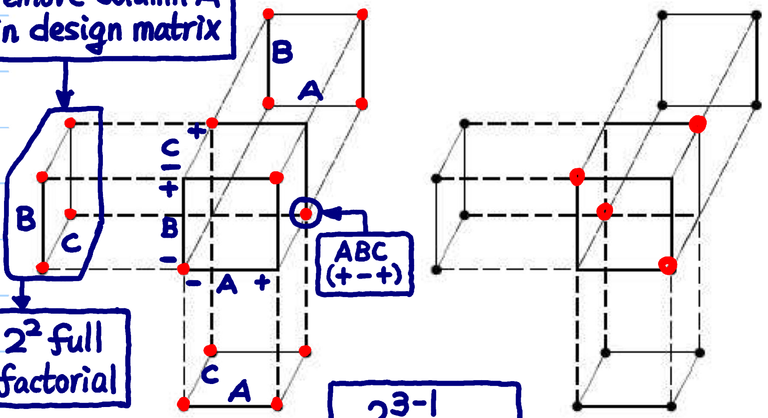
e.g. 2^{7-2} with $I = 1236 = 1457 = 23457$

projected onto $\{1,2,3,4\}$ ⇒ 2^4 full, 2 replicates

projected onto $\{1,2,3,4,5,6\}$ ⇒ 2^{6-1} with $I = 1236$

(exercise) $\{1,2,3,6\}$ $\{2,4,5,6,7\}$

remove column A in design matrix



2^2 full factorial

2^{3-1} resolution III

$\mathbf{I} = \mathbf{123}$ $\mathbf{I} = -\mathbf{123}$

$1+2+3=0 \pmod 2$ cf.

$1+2+3=1 \pmod 2$ cf.

Check LN p.6-5 → $1=-23$, $2=-13$, $3=-12$

Figure 1: 2^{3-1} Designs Using $\mathbf{I} = \pm\mathbf{123}$ and Their Projections to 2^2 Designs.

Q: Why is this property useful? Suppose that only $R-1$ or fewer "factors" are important. ⇒ all effects of these important factors are located in different alias sets.