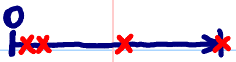
Factorial Effects, Adapted Epi-Layer Growth Experiment

Table 4: Factorial Effects, Adapted Epitaxial Layer Growth Experiment

single replicate (unreplicated), LNp.2

Effect		$\bar{y}$	$\ln s^2$
A	(½)ME(A)→	−0.078	0.016
B		0.173	−0.118
C		−0.078	−0.112
D		0.490	0.056
AB	(½)INT(A,B)→	0.008	0.045
AC		−0.093	−0.026
AD		−0.050	−0.029
BC		0.058	0.080
BD		−0.030	0.010
CD		−0.345	0.085
ABC		0.098	−0.032
ABD		0.025	0.042
ACD		−0.030	0.000
BCD		0.110	−0.003
ABCD		0.020	0.103

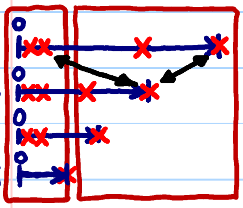
2^4 full factorial
= 16 distinct level combinations
→ can study 16−1=15 factorial effects



|$\hat{\beta}$effect|

Note. All $\hat{\beta}_j$'s have same s.e.($\hat{\beta}_j$)

ME's
2fi's
3fi's
4fi's



⇒ effect hierarchy principle

Q: Can we do t-tests for these $\hat{\beta}_j$'s?
# of distinct level combinations = 16
# of parameters in $\underline{\beta}$ = 16
⇒ residuals = 0
⇒ no d.f. left to estimate σ

Suppose we can peform t-tests (e.g., replicated exp't or unreplicated exp't using a simpler model, say only ME, 2fi)

sequential — drop-one
∵ orthogonality
ANOVA ←consistent→ t-test
F-statistic = (t-statistic)^2

❖ **Reading**: textbook, 4.4

Fundamental Principles in Factorial Design

define "more important" information in conceptual model

- **Effect Hierarchy Principle** (from experience)
  - Lower-order effects are more likely to be important than higher-order effects.
  - Effects of the same order are equally likely to be important. ← meaning? check LNp.5-11

identical order for 2-level factors

qualitative factor: # of factors involved in an effect, e.g., A = B = C ≫ AB=AC=BC ≫ ABC (more important)
quantative factor: (polynomial) power of an effect, e.g., $\ell \gg q = \ell\times\ell \gg \ell\times q = q\times\ell$

- **Effect Sparsity principle** (Box-Meyer): The number of relatively important effects in a factorial experiment is small. ← meaning? large/significant $\hat{\beta}_j$'s cf.

  e.g. fewer effects (say, 3~5) can explain 60% ~ 80% of variation in y, other effects (many) can explain only 10%.

  vital few ↕ trivial many

  ⊖ This is similar to the *Pareto Principle* in quality investigation.

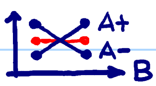
- Effect hierarchy and sparsity principles are more effective/relevant for *screening experiments* (Why?). (LNp.1-12) (many factors, small run size)

- ⊙ **Effect Heredity Principle** (Hamada-Wu): In order for an interaction to be significant, at least one of its parent factors should be significant. (useful in model selection) meaning? large $\hat{\beta}_j$ (≠0)
  - For modeling, McCullagh and Nelder called it the *Marginality Principle*.

$M_1: y \sim A+C+AB+\varepsilon$
$M_2: y \sim C+D+AB+\varepsilon$
same R^2
⇒ prefer M_1 (why?)

ME(A) ME(B)
↘ ↙
INT(A,B)

ME(A): not sig.
ME(B): not sig.
INT(A,B): sig.



❖ **Reading**: textbook, 4.6

One-Factor-At-A-Time (ofat) Approach

Table 5: Planning Matrix for 2^3 Design and Response Data For Comparison with One-Factor-At-A-Time Approach

P — 1200, 3 times; 1400, 1 time

S — fast, 3 times; slow, 1 time

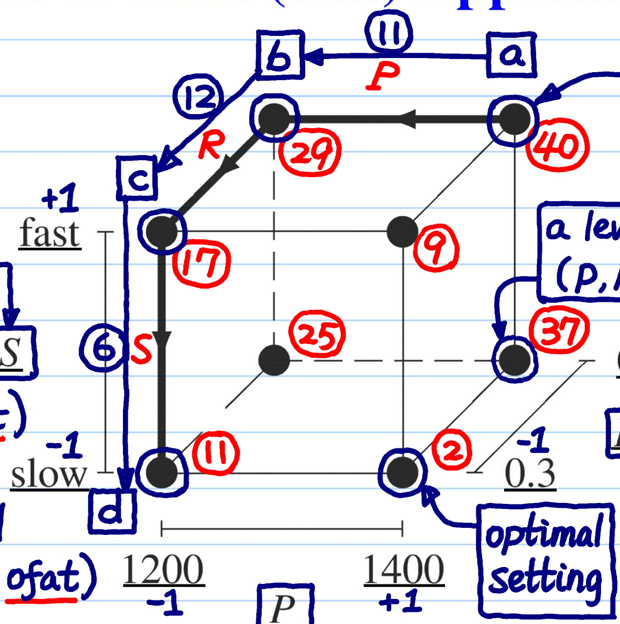
not a balance design (OA of strength 1) ⇒ cause collinearity

	Factor			Percent Burned
	P	R	S	
d	1200	0.3	slow	11
c	1200	0.3	fast	17
	1200	0.6	slow	25
b	1200	0.6	fast	29
	1400	0.3	slow	02
	1400	0.3	fast	09
	1400	0.6	slow	37
a	1400	0.6	fast	40

Percent Burned: response mean μ_x: smaller-the-better

02: optimal setting.

One-Factor-At-A-Time (ofat) Approach (Contd.)



design starting point

least important → S

a level combination (P, R, S) = (1400, 0.6, slow)

2nd most important → R

optimal setting

P ← most important

E.g., effect of P:
$ME(P \mid R=0.6, S=\text{fast})$
compare this effect (which may be claimed as the ME of P in ofat) with the definition of ME of P given in LNp.5-4.
(Q: Which has better reproducibility?)

Figure 4: The Path of an ofat Plan

Factorial effects θ

P	=	1.5
★ R	=	23
S	=	5
★ PR	=	10
PS	=	0
RS	=	−1.5
PRS	=	−0.5

- The three steps of ofat as illustrated in the arrows in Figure 4 are detailed in steps 1-3 on page 174 of textbook.

Disadvantages of ofat Approach Relative to Factorial Approach

For exp't with k factors, under unreplicated 2^k design, $Var(\hat{\theta})=\frac{1}{2^{k-2}}\sigma^2$ (LNp.5-8)

Let r be the # of replicates on each design point in an ofat. Then,

$$Var(\hat{\theta})=Var(\bar{z}_+-\bar{z}_-) \le \frac{\sigma^2}{kr}+\frac{\sigma^2}{r} = (k+1)\sigma^2/kr$$

k+1 design points

For r s.t. $\frac{k+1}{kr}\sigma^2 \approx \frac{1}{2^{k-2}}\sigma^2$,

$\Rightarrow r \approx \underbrace{(k+1/k)}_{\le 2}\cdot 2^{k-2}$ ($1 \le$)

can set $r=2\cdot 2^{k-2}=2^{k-1}$

$Var(\hat{\theta})$

1. *It requires more runs for the same precision in effect estimation.* In the example, the 2^3 design requires 8 runs. For ofat to have the same precision, each of the 4 corners on the ofat path needs to have 4 runs, totaling 16 runs. In general, to be comparable to a 2^k design, ofat would require 2^{k-1} runs at each of the $k+1$ corners on its path, totaling $(k+1)2^{k-1}$. The ratio is $(k+1)2^{k-1}/2^k=(k+1)/2$.

   $\hat{\theta}=\bar{z}_+-\bar{z}_-$ unreplicated replicates

2. *It cannot estimate some interactions.*

   e.g. INT(P,R) (1200, 0.3) ~~(1400, 0.3)~~ (1200, 0.6) (1400, 0.6) check LNp.13

3. *Conclusions for analysis not as general.*

   e.g., ofat → ME(P | R=0.6, S=fast)
   2^k → ME(P)

4. It can miss optimal settings. → ofat usually does not search the experimental region uniformly

- For points 2 – 4, see Figure 4.

Why Experimenters Continue to Use ofat?

- Most physical laws are taught by varying one factor at a time. Easier to think and focus on one factor each time.
- Experimenters often have good intuition about the problem when thinking in this mode.
  could be useful physical information
- No exposure to statistical design of experiments.
- **Challenges** for DOE researchers: To combine the factorial approach with the good intuition rendered by the the ofat approach. Needs a new outlook.

  physical knowledge / engineer knowledge / domain knowledge / … → conceptual model

  演繹法(deduction)
  歸納法(induction) ← DOE
  → problem solving

❖ **Reading**: textbook, 4.7