STA305/1004-Class 19

Nov. 28, 2019

Today's Class

- Lenth's Method for Assessing significance in unreplicated factorial designs
- Blocking factorial designs
  - Effect hierarchy principle
  - Generation of orthogonal blocks
  - Generators and deining relations

Factorial Notation - Design Matrix in Standard Order

A 2^4 design matrix in standard form is:

different from model matrix

1	2	3	4
-1	-1	-1	-1
1	-1	-1	-1
-1	1	-1	-1
1	1	-1	-1
-1	-1	1	-1
1	-1	1	-1
-1	1	1	-1
1	1	1	-1
-1	-1	-1	1
1	-1	-1	1
-1	1	-1	1
1	1	-1	1
-1	-1	1	1
1	-1	1	1
-1	1	1	1
1	1	1	1

2^2

Run		
1	-1	-1
2	+1	-1
3	-1	+1
4	+1	+1

2^3

-1	-1	-1
+1	-1	-1
-1	+1	-1
+1	+1	-1
-1	-1	+1
+1	-1	+1
-1	+1	+1
+1	+1	+1

Example - 2^3 design for studying a chemical reaction

A process development experiment studied four factors in a 2^4 factorial design.

- amount of catalyst charge `x1`,
- temperature `x2`,
- pressure `x3`,
- concentration of one of the reactants `x4`.
- The response y is the percent conversion at each of the 16 run conditions. The design is shown below.

Example - 2^4 design for studying a chemical reaction

run	x1	x2	x3	x4	conversion
1	-1	-1	-1	-1	70
2	1	-1	-1	-1	60
3	-1	1	-1	-1	89
4	1	1	-1	-1	81
5	-1	-1	1	-1	69
6	1	-1	1	-1	62
7	-1	1	1	-1	88
8	1	1	1	-1	81
9	-1	-1	-1	1	60
10	1	-1	-1	1	49
11	-1	1	-1	1	88
12	1	1	-1	1	82
13	-1	-1	1	1	60
14	1	-1	1	1	52
15	-1	1	1	1	86
16	1	1	1	1	79

repeated twice.

The design is not replicated so it's not possible to estimate the standard errors of the factorial effects.

Example - 2^4 design for studying a chemical reaction

```
fact1 <- lm(conversion~x1*x2*x3*x4,data=tab0510a)
round(2*fact1$coefficients,2)
```

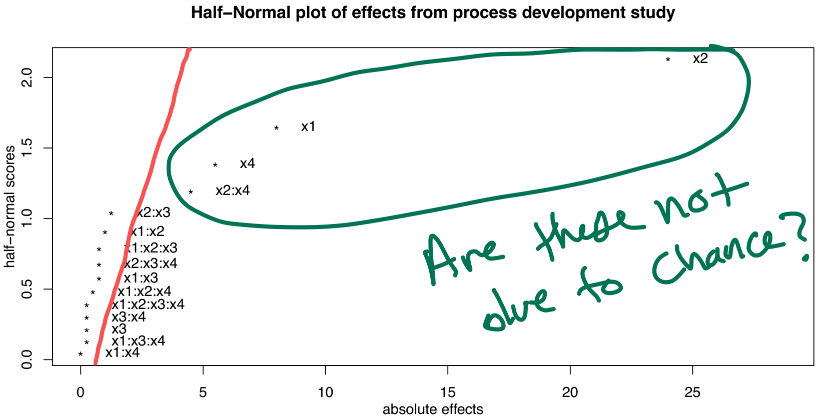
```
(Intercept)          x1          x2          x3          x4       x1:x2
     144.50       -8.00       24.00       -0.25       -5.50        1.00
      x1:x3       x2:x3       x1:x4       x2:x4       x3:x4    x1:x2:x3
       0.75       -1.25        0.00        4.50       -0.25       -0.75
   x1:x2:x4    x1:x3:x4    x2:x3:x4 x1:x2:x3:x4
       0.50       -0.25       -0.75       -0.25
```

s.e. of coefficients would be NA
– not available ∵ lm() won't
be able to compute ∵ not replicated.

Half-Normal Plots

- An advantage of this plot is that all the large estimated effects appear in the upper right hand corner and fall above the line.
- The half-normal plot for the effects in the process development example can be obtained with `DanielPlot()` with the option `half=TRUE`.

```
library(FrF2)
DanielPlot(fact1,half=TRUE,autolab=F,
          main="Half-Normal plot of effects from process development study")
```



Lenth's method: testing significance for experiments without variance estimates

- Half-normal and normal plots are informal graphical methods involving visual judgement.
- It's desirable to judge a deviation from a straight line quantitatively based on a formal test of significance.
- Lenth (1989) proposed a method that is simple to compute and performs well.

Lenth's method

$K=2 \quad 2^2$

then $\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3$ — interaction

— main effects

- Let $\hat{\theta}_{(1)}, \ldots, \hat{\theta}_{(N)}$ be $N = 2^k - 1$ factorial effects in a 2^k design.

Lenth's method

- Let $\hat{\theta}_{(1)}, \ldots, \hat{\theta}_{(N)}$ be $N = 2^k - 1$ factorial effects in a 2^k design.
- Let m be the median of $|\hat{\theta}|_{(1)}, \ldots, |\hat{\theta}|_{(N)}$.

Lenth's method

- Let $\hat{\theta}_{(1)}, ..., \hat{\theta}_{(N)}$ be $N = 2^k - 1$ factorial effects in a 2^k design.
- Let m be the median of $|\hat{\theta}|_{(1)}, ..., |\hat{\theta}|_{(N)}$.
- An estimate of the standard error of an effect, called the pseudo standard error, is then given by, $s_0 = 1.5m$.

Lenth's method

- Let $\hat{\theta}_{(1)}, ..., \hat{\theta}_{(N)}$ be $N = 2^k - 1$ factorial effects in a 2^k design.
- Let m be the median of $|\hat{\theta}|_{(1)}, ..., |\hat{\theta}|_{(N)}$.
- An estimate of the standard error of an effect, called the pseudo standard error, is then given by, $s_0 = 1.5m$.
- Any estimated effect exceeding $2.5s_0$ is excluded, and, if needed m and s_0 are recomputed.

Lenth's method

- Let $\hat{\theta}_{(1)}, ..., \hat{\theta}_{(N)}$ be $N = 2^k - 1$ factorial effects in a 2^k design.
- Let m be the median of $|\hat{\theta}|_{(1)}, ..., |\hat{\theta}|_{(N)}$.
- An estimate of the standard error of an effect, called the pseudo standard error, is then given by, $s_0 = 1.5m$.
- Any estimated effect exceeding $2.5s_0$ is excluded, and, if needed m and s_0 are recomputed.
- A margin of error is then given by $ME = t_{1-\alpha/2,d} \cdot s_0$, where $d = N/3$.

$\alpha = .95$ $1 - \frac{.95}{2}$

Lenth's method

- Let $\hat{\theta}_{(1)}, ..., \hat{\theta}_{(N)}$ be $N = 2^k - 1$ factorial effects in a 2^k design.
- Let m be the median of $|\hat{\theta}|_{(1)}, ..., |\hat{\theta}|_{(N)}$.
- An estimate of the standard error of an effect, called the pseudo standard error, is then given by, $s_0 = 1.5m$.
- Any estimated effect exceeding $2.5s_0$ is excluded, and, if needed m and s_0 are recomputed.
- A margin of error is then given by $ME = t_{1-\alpha/2,d} \cdot s_0$, where $d = N/3$.
- All estimates greater than ME may be viewed as "significant".

Lenth's method

- Let $\hat{\theta}_{(1)}, \ldots, \hat{\theta}_{(N)}$ be $N = 2^k - 1$ factorial effects in a 2^k design.
- Let m be the median of $|\hat{\theta}|_{(1)}, \ldots, |\hat{\theta}|_{(N)}$.
- An estimate of the standard error of an effect, called the pseudo standard error, is then given by, $s_0 = 1.5m$.
- Any estimated effect exceeding $2.5s_0$ is excluded, and, if needed m and s_0 are recomputed.
- A margin of error is then given by $ME = t_{1-\alpha/2,d} \cdot s_0$, where $d = N/3$.
- All estimates greater than ME may be viewed as "significant".
- But, with so many estimates some will be falsely identified.

2^3

$2^3 - 1$

$= 7$

Lenth's method

- Let $\hat{\theta}_{(1)}, \ldots, \hat{\theta}_{(N)}$ be $N = 2^k - 1$ factorial effects in a 2^k design.
- Let m be the median of $|\hat{\theta}|_{(1)}, \ldots, |\hat{\theta}|_{(N)}$.
- An estimate of the standard error of an effect, called the pseudo standard error, is then given by, $s_0 = 1.5m$.
- Any estimated effect exceeding $2.5s_0$ is excluded, and, if needed m and s_0 are recomputed.
- A margin of error is then given by $ME = t_{1-\alpha/2,d} \cdot s_0$, where $d = N/3$.
- All estimates greater than ME may be viewed as "significant".
- But, with so many estimates some will be falsely identified.
- A simultaneous margin of error is: $SME = t_{\gamma,s} \cdot s_0$, where $\gamma = (1 + (1-\alpha)^{1/N})/2$.

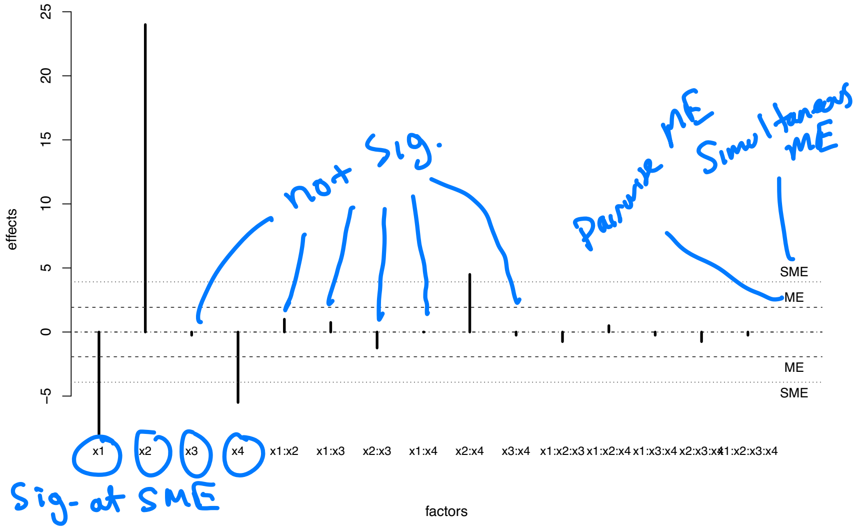
Bonferroni type correction

Lenth's method

- Let $\hat{\theta}_{(1)}, ..., \hat{\theta}_{(N)}$ be $N = 2^k - 1$ factorial effects in a 2^k design.
- Let m be the median of $|\hat{\theta}|_{(1)}, ..., |\hat{\theta}|_{(N)}$.
- An estimate of the standard error of an effect, called the pseudo standard error, is then given by, $s_0 = 1.5m$.
- Any estimated effect exceeding $2.5s_0$ is excluded, and, if needed m and s_0 are recomputed.
- A margin of error is then given by $ME = t_{1-\alpha/2,d} \cdot s_0$, where $d = N/3$.
- All estimates greater than ME may be viewed as "significant".
- But, with so many estimates some will be falsely identified.
- A simultaneous margin of error is: $SME = t_{\gamma,s} \cdot s_0$, where $\gamma = (1 + (1-\alpha)^{1/N})/2$.
- Estimated effects exceeding SME are declared significant, but SME is adjusted for multiple comparison.

Lenth's method - Lenth Plot for process development example

```
LenthPlot(fact1,cex.fac = 0.8)
```




```
##    alpha      PSE       ME      SME
## 0.050000 0.750000 1.927936 3.913988
```

1. Full Normal plot
2. Half Normal plot
3. Lenth's Method

These three methods can assess if factorial effects in an unreplicated design are due to chance (e.g., significant)

Blocking Factorial Designs

Blocking factorial designs

- In a trial conducted using a 2^3 design it might be desirable to use the same batch of raw material to make all 8 runs.
- Suppose that batches of raw material were only large enough to make 4 runs. Then the concept of blocking could be used.

Blocking factorial designs

Consider the 2^3 design.

Run	1	2	3	12	13	23	123
1	-1	-1	-1	1	1	1	-1
2	1	-1	-1	-1	-1	1	1
3	-1	1	-1	-1	1	-1	1
4	1	1	-1	1	-1	-1	-1
5	-1	-1	1	1	-1	-1	1
6	1	-1	1	-1	1	-1	-1
7	-1	1	1	-1	-1	1	-1
8	1	1	1	1	1	1	1

Runs	Block
1, 4, 6, 7	I
2, 3, 5, 8	II

How are the runs assigned to the blocks?

factor 1, 2, 3; 12 interaction; 13 interaction; 123 interaction

123: −, +

Confounded (mixed up) with 123 term or three-way interaction.

Blocking factorial designs

Run	1	2	3	12	13	23	123
1	-1	-1	-1	1	1	1	-1
2	1	-1	-1	-1	-1	1	1
3	-1	1	-1	-1	1	-1	1
4	1	1	-1	1	-1	-1	-1
5	-1	-1	1	1	-1	-1	1
6	1	-1	1	-1	1	-1	-1
7	-1	1	1	-1	-1	1	-1
8	1	1	1	1	1	1	1

Runs	Block	sign of 123
1, 4, 6, 7	I	$-$
2, 3, 5, 8	II	$+$

Blocking factorial designs

- Any systematic differences between the two blocks of four runs will be eliminated from all the main effects and two factor interactions.

Blocking factorial designs

- Any systematic differences between the two blocks of four runs will be eliminated from all the main effects and two factor interactions.
- What you gain is the elimination of systematic differences between blocks.

Blocking factorial designs

- Any systematic differences between the two blocks of four runs will be eliminated from all the main effects and two factor interactions.
- What you gain is the elimination of systematic differences between blocks.
- But now the three factor interaction is confounded with any batch (block) difference.

Blocking factorial designs

- Any systematic differences between the two blocks of four runs will be eliminated from all the main effects and two factor interactions.
- What you gain is the elimination of systematic differences between blocks.
- But now the three factor interaction is confounded with any batch (block) difference.
- The ability to estimate the three factor interaction separately from the block effect is lost.

Effect hierarchy principle

e.g., main effects, two-way interactions.

1. Lower-order effects are more likely to be important than higher-order effects.
2. Effects of the same order are equally likely to be important.

- One reason that many accept this principle is that higher order interactions are more difficult to interpret or justify physically.
- Investigators are less interested in estimating the magnitudes of these effects even when they are statistically significant.

Generating Factorial Blocks

In the 2^3 example suppose that the block variable is given the identifying number 4.

Run	1	2	3	4=123
1	-1	-1	-1	-1
2	1	-1	-1	1
3	-1	1	-1	1
4	1	1	-1	-1
5	-1	-1	1	1
6	1	-1	1	-1
7	-1	1	1	-1
8	1	1	1	1

- Think of your experiment as containing four factors.
- The fourth factor will have the special property that it does not interact with other factors.
- If this new factor is introduced by having its levels coincide exactly with the plus and minus signs attributed to 123 then the blocking is said to be **generated** by the relationship 4=123.
- This idea can be used to derive more sophisticated blocking arrangements.

An example of how not to block

Suppose we would like to arrange the 2^3 design into four blocks.

Run	1	2	3	4=123	5=23	45=1
1	-1	-1	-1	-1	1	-1
2	1	-1	-1	1	1	1
3	-1	1	-1	1	-1	-1
4	1	1	-1	-1	-1	1
5	-1	-1	1	1	-1	-1
6	1	-1	1	-1	-1	1
7	-1	1	1	-1	1	-1
8	1	1	1	1	1	1

- Runs are placed in different blocks depending on the signs of the block variables in columns 4 and 5.
- Consider two block factors called 4 and 5.
- 4 is associated with ? 123
- 5 is associated ? 23

An example of how not to block

$123 \cdot 23 = 1\,22\,33 = 1$ (22 = I, 33 = I)

Run	1	2	3	4=123	5=23	45=1
1	-1	-1	-1	-1	1	-1
2	1	-1	-1	1	1	1
3	-1	1	-1	1	-1	-1
4	1	1	-1	-1	-1	1
5	-1	-1	1	1	-1	-1
6	1	-1	1	-1	-1	1
7	-1	1	1	-1	1	-1
8	1	1	1	1	1	1

Block	Run
I	2,8
II	1,7
III	3,5
IV	4,6

4	5
+	+
−	+
+	−
−	−

An example of how not to block

- 45 is confounded with the main effect of 1.
- Therefore, if we use 4 and 5 as blocking variables we will not be able to separately estimate the main effect 1.
- Main effects should not be confounded with block effects.

Violates the effect hierarchy principle!

An example of how not to block

- Any blocking scheme that confounds main effects with blocks should not be used.
- This is based on the assumption:

  The block-by-treatment interactions are negligible.

- This assumption states that treatment effects do not vary from block to block.
- Without this assumption estimability of the factorial effects will be very complicated.

An example of how not to block

- For example, if $B_1 = 12$ then this implies two other relations:

$$1B_1 = 112 = 2 \text{ and } B_1 2 = 122 = 122 = 1.$$

- If there is a significant interaction between the block effect B_1 and the main effect 1 then the main effect 2 is confounded with $1B_1$.
- If there is a significant interaction between the block effect B_1 and the main effect 2 then the main effect 1 is confounded with $B_1 2$.

How to do it

Run	1	2	3	4=12	5=13
1	-1	-1	-1	1	1
2	1	-1	-1	-1	-1
3	-1	1	-1	-1	1
4	1	1	-1	1	-1
5	-1	-1	1	1	-1
6	1	-1	1	-1	1
7	-1	1	1	-1	-1
8	1	1	1	1	1

- Set 4=12, 5=13.
- Then $I = 124 = 135 = 2345$.
- Estimated block effects 4, 5, 45 are assoicated with the estimated two-factor interaction effects 12, 13, 23 and not any main effects.
- Which runs are assigned to which blocks?

Generators and Defining Relations

- A simple calculus is available to show the consequences of any proposed blocking arrangement.
- If any column in a 2^k design are multiplied by themselves a column of plus signs is obtained. This is denoted by the symbol I.

$$I = 11 = 22 = 33 = 44 = 55,$$

where, for example, 22 means the product of the elements of column 2 with itself.

- Any column multiplied by I leaves the elements unchanged. So, $I3 = 3$.

Generators and Defining Relations

- A general approach for arranging a 2^k design in 2^q blocks of size 2^{k-q} is as follows.
- Let $B_1, B_2, ..., B_q$ be the block variables and the factorial effect v_i is confounded with B_i,

$B_1 = 12$, $B_2 = 13$ $k = 3$, $q = 2$

$$B_1 = v_1, B_2 = v_2, ..., B_q = v_q.$$

- The block effects are obtained by multiplying the B_i's:

$$B_1 B_2 = v_1 v_2, B_1 B_3 = v_1 v_3, ..., B_1 B_2 \cdots B_q = v_1 v_2 \cdots v_q$$

- There are $2^q - 1$ possible products of the B_i's and the I (whose components are $+$).

$2^2 - 1 = 4 - 1 = 3$

Generators and Defining Relations

2-2way
4-3way
1-4 way
interactions

two-way → $B_1B_2 = 135 \cdot 235 = 12$

$B_1B_3 = 135 \cdot 1234 = 245$

$B_2B_3 = 235 \cdot 1234 = 145$

Example: A 2^5 design can be arranged in 8 blocks of size $2^{5-3} = 4$.

Consider two blocking schemes.

→ $B_1B_2B_3 = 135 \cdot 235 \cdot 1234 = 34$

1. Define the blocks as

The blocks are confounded with:

$$B_1 = 135,\ B_2 = 235,\ B_3 = 1234.$$

and 12, 245, 145, 34

The remaining blocks are confounded with the following interactions:

$B_1B_2 = 12 \cdot 13 = 23$

$B_1B_3 = 12 \cdot 45 = 1245$

$B_2B_3 = 13 \cdot 45 = 1345$

$B_1B_2B_3 = 12 \cdot 13 \cdot 45 = 2345$

2. Define the blocks as:

These blocks are confounded with:

$$B_1 = 12,\ B_2 = 13,\ B_3 = 45,$$

23, 1245, 1345, 2345

Which is a better blocking scheme?

4-2way
3-4 way interactions

First Scheme is Superior ∵ it only confounds 2-2way interactions.