STA305/1004-Class 25

April 4, 2016

Today's Class

- Fractional factorial design
- Split plot designs

Reminder
exam review
session
April 11 @11-12

TA office hours
Thurs 1-3
Extend deadline for
hw #4
April 8 @ 10:00pm

Effect Aliasing and Design Resolution

Consider a design that studies five factors in 16 run. A half fraction of a 2^5 or 2^{5-1}.

32

use 16 runs 2^{5-1} to estimate 5 main effects and

Run	B	C	D	E	Q
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
9	-1	1	1	-1	1
10	1	1	1	1	1
11	-1	-1	1	1	1
12	1	-1	1	-1	1
13	-1	1	-1	1	1
14	1	1	-1	-1	1
15	-1	-1	-1	-1	1
16	1	-1	-1	1	1

estimates same fact. effect

- The factor E is assigned to the column BCD.
- The column for E is used to estimate the main effect of E and also for BCD.
- The main factor E is said to be **aliased** with the BCD interaction.

Effect Aliasing and Design Resolution

- This aliasing relation is denoted by

$$E = BCD \text{ or } I = BCDE,$$

  where I denotes the column of all $+$'s.
- This uses same mathematical definition as the confounding of a block effect with a factorial effect.
- Aliasing of the effects is a price one must pay for choosing a smaller design.
- The 2^{5-1} design has only 15 degrees of freedom for estimating factorial effects, it cannot estimate all 31 factorial effects among the factors B, C, D, E, Q.

study 5 factors
in 16 runs vs. 32 runs

Example - Leaf spring experiment

= spring free ht.

B	C	D	E	Q	y
-1	1	1	-1	-1	7.7900
1	1	1	1	-1	8.0700
-1	-1	1	1	-1	7.5200
1	-1	1	-1	-1	7.6333
-1	1	-1	1	-1	7.9400
1	1	-1	-1	-1	7.9467
-1	-1	-1	-1	-1	7.5400
1	-1	-1	1	-1	7.6867
-1	1	1	-1	1	7.2900
1	1	1	1	1	7.7333
-1	-1	1	1	1	7.5200
1	-1	1	-1	1	7.6467
-1	1	-1	1	1	7.4000
1	1	-1	-1	1	7.6233
-1	-1	-1	-1	1	7.2033
1	-1	-1	1	1	7.6333

Example - Leaf spring experiment

The factorial effects are estimated as before.

```
library(FrF2)
fact.leaf <- lm(y~B*C*D*E*Q,data=leafspring)
fact.leaf2 <- aov(y~B*C*D*E*Q,data=leafspring)
round(2*fact.leaf$coefficients,2)
```

```
(Intercept)           B           C           D           E           Q
      15.27        0.22        0.18        0.03        0.10       -0.26
        B:C         B:D         C:D         B:E         C:E         D:E
       0.02        0.02       -0.04          NA          NA          NA
        B:Q         C:Q         D:Q         E:Q       B:C:D       B:C:E
       0.08       -0.17        0.05        0.03          NA          NA
      B:D:E       C:D:E       B:C:Q       B:D:Q       C:D:Q       B:E:Q
         NA          NA        0.01       -0.04       -0.05          NA
      C:E:Q       D:E:Q     B:C:D:E     B:C:D:Q     B:C:E:Q     B:D:E:Q
         NA          NA          NA          NA          NA          NA
    C:D:E:Q   B:C:D:E:Q
         NA          NA
```

The effects with NA are aliased

Effect Aliasing and Design Resolution

- The equation $I = BCDE$ is called the **defining relation** of the 2^{5-1} design.
- The design is said to have resolution IV because the defining relation consists of the "word" BCDE, which has "length" 4.
- Multiplying both sides of $I = BCDE$ by column B

$$B = B \times I = B \times BCDE = CDE,$$

  the relation $B = CDE$ is obtained.
- B is aliased with the CDE interaction. Following the same method all 15 aliasing relations can be obtained.
- Therefore, the main effects of B, C, D, E are estimable only if the aforementioned three-factor interactions are negligible.
- The other factorial effects have analogous aliasing properties.

$16 - 1 = df$

Effect Aliasing and Design Resolution

Run	A	B	C	AB	AC	BC	ABC
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

Suppose that the 2^{5-1} design defines $Q = BCDE$.

1. What is the defining relation?
2. What is the design resolution?
3. What are the aliasing relations?

$I = BCDEQ$

design res. is V

BCDEQ has 5 letters

$B = CDEQ$,

$E = BCDQ$

$QB = CDE$

still be 15 pairs of aliased effects

Which design is better
$Q = BCDE$ or $E = BCD$?

- which design has more higher order interactions ~~that~~ are aliased?
- depends on which interactions are considered more important !
- No clear answer ∵ it depends on experimental objectives.

Split plot designs

- These designs were originally developed for agriculture by R.A. Fisher and F. Yates.
- Due to their applicability outside agriculture they could be called split-unit designs.
- But we will use split-plot . . .

Split plot designs

- Some factors need to be applied to larger plots as compared to other factors.
- For example, type of irrigation method and the type of fertilizer used are two factors then irrigation requires a larger plot.
- One has to apply a specific irrigation to a large plot.

Split plot designs - corrosion study

- An experiment of corrosion resistance of steel bars treated with four different coatings C_1, C_2, C_3, C_4 was conducted.
- Three furnace temperatures were investigated and four differently coated bars randomly arranged in the furnace within each heat.
- Positions of the coated steel bars in the furnace randomized within each heat.
- Furnace temperature was difficult to change so heats were run in systematic order shown.

Temperature	Position 1	Position 2	Position 3	Position 4
360°				
370°				
380°				

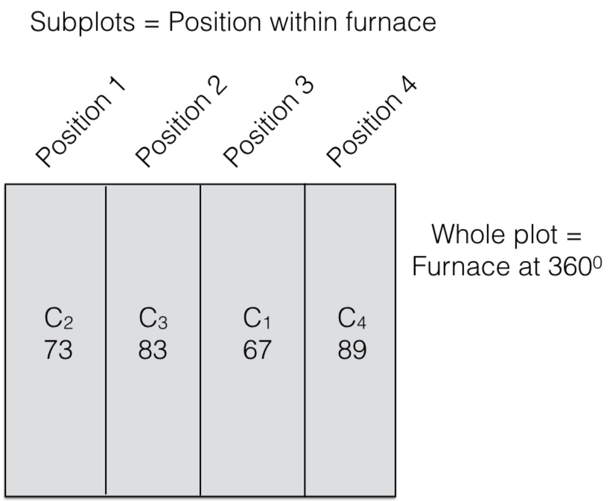
Split plot designs - corrosion study

The primary interest were the comparison of coatings and how they interacted with temperature.
Some of the data from the experiment is shown below.

run	heats	coating	position	replication	resistance
r1	T360	C2	1	1	73
r1	T360	C3	2	1	83
r1	T360	C1	3	1	67
r1	T360	C4	4	1	89
r2	T370	C1	1	1	65
r2	T370	C3	2	1	87

Split plot designs - corrosion study

The split-plot experiment of corrosion resistance is shown for the first replicate at 360.



Split-plot designs versus Factorial Designs

- How does the split-plot design compare with a 3x4 factorial design of coating and temperature?
- In the factorial design an oven temperature-coating combination would be randomly selected then we would obtain a corrosion resistance measure.
- Then randomly select another oven temperature-coating combination and obtain another corrosion resistance measure until we have a resistance measure for all 12 oven temperature-coating combinations.
- To run each combination in random order would require adjusting the furnace temperature up to 24 times (since there were two replicates) and would have resulted in a much larger variance.
- The split plot is like a randomized block design (with whole plots as blocks) in which the opportunity is taken to introduce additional factors between blocks.
- In this design there is only one source of error influencing the resistance.

Split-plot designs versus Factorial Designs

There are two different experimental units:

- The six different furnace heats, called whole plots.
- The four positions within each furnace heat, called subplots, where the differently coated bars could be placed in the furnace.
- Misleading to treat as if only one error source and one variance.
- Two different experimental units: six furnace heats (whole plots); and four positions (subplots) where differently coated bars placed in furnace.
- Two different variances: σ_W^2 for whole plots and σ_S^2 for subplots.
- It would be misleading to treat as if only one error source and one variance.

Split plot designs versus Factorial Designs

- Achieving and maintaining a given temperature in this furnace was very imprecise.
- The whole plot variance, measuring variation from one heat to another, was expected to be large.
- The subplot variance measuring variation from position to position, within a given heat, was expected to be small.
- The subplot effects and subplot-main plot interaction are estimated using with the same subplot error.

Why choose a split plot design?

- Two considerations important in choosing an experimental design are feasibility and efficiency.
- In industrial experimentation a split-plot design is often convenient and the only practical possibility.
- This is the case whenever there are certain factors that are *difficult to change* and others that are *easy to change*.
- In this example changing the *furnace temperature was difficult to change*; rearranging the *positions of the coated bars in the furnace was easy to change*.

ANOVA table for split plot experiment

- The numerical calculations for the ANOVA of a split-plot design are the same as for other balanced designs (designs where all treatment combinations have the same number of observations) and can be performed in R or with other statistical software.
- Experimenters sometimes have difficulty identifying appropriate error terms.

```
spfurcoat1 <- aov(resistance~ replication*heats*coating,data=tab0901)
summary(spfurcoat1)
```

```
                          Df Sum Sq Mean Sq
replication                1    782     782
heats                      2  26519   13260
coating                    3   4289    1430
replication:heats          2  13658    6829
replication:coating        3    254      85
heats:coating              6   3270     545
replication:heats:coating  6    867     144
```

ANOVA table for split plot experiment - whole plot

```
spfurcoat1 <- aov(resistance~ replication*heats*coating,data=tab0901)
summary(spfurcoat1)
```

```
                          Df Sum Sq Mean Sq
replication                1    782     782
heats                      2  26519   13260
coating                    3   4289    1430
replication:heats          2  13658    6829
replication:coating        3    254      85
heats:coating              6   3270     545
replication:heats:coating  6    867     144
```

The whole plot effects are `replication` and `replication:heats`. So the ANOVA table for the whole plots is:

Source	DF	SS	MS
replication	1	782	782
replication $\times$ heats	2	13658	6829

The whole plot mean square error is 6829. This measures the differences between the replicated heats at the three different temperatures.

ANOVA table for split plot experiment - sub plot

```
spfurcoat1 <- aov(resistance~ replication*heats*coating,data=tab0901)
summary(spfurcoat1)
```

```
                          Df Sum Sq Mean Sq
replication                1    782     782
heats                      2  26519   13260
coating                    3   4289    1430
replication:heats          2  13658    6829
replication:coating        3    254      85
heats:coating              6   3270     545
replication:heats:coating  6    867     144
```

The subplot effects are:

Source	DF	SS	MS
coating	3	4289	1430
coating $\times$ heats	6	3270	545

- The subplot mean square error is $(254 + 867)/(3 + 6) = 124.6$.
- The sum of squares for the subplot error is the sum of interaction between replicate and coating (`replication:coating`) and the three way interaction of replication, heats and coating (`replication:heats:coating`).
- The subplot error measures to what extent the coatings give dissimilar results within each of the replicated temperatures.

ANOVA table for split plot experiment - using aov() with Error()

In R the ANOVA table for whole plot and sub plot effects can obtained by specifying the subplot error structure explicit using Error().

```
spfurcoat <- aov(resistance~ replication + heats + replication:heats
                 + coating + heats:coating
                 +Error(heats/replication),data=tab0901)
summary(spfurcoat)
```

```
Error: heats
      Df Sum Sq Mean Sq
heats  2  26519   13260

Error: heats:replication
                  Df Sum Sq Mean Sq
replication        1    782     782
replication:heats  2  13658    6829

Error: Within
              Df Sum Sq Mean Sq F value  Pr(>F)
coating        3   4289  1429.7  11.480 0.00198 **
heats:coating  6   3270   545.0   4.376 0.02407 *
Residuals      9   1121   124.5
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

ANOVA for split plot experiment - using aov() with Error()- whole plot

Under the heading Error: heats is mean square error for a one-way ANOVA model comparing heats.

```
summary( aov(resistance~heats,tab0901))
```

```
            Df Sum Sq Mean Sq F value   Pr(>F)    
heats        2  26519   13260   12.04 0.000328 ***
Residuals   21  23119    1101                     
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

ANOVA for split plot experiment - using `aov()` with `Error()`- whole plot

- The ratio of mean square errors follows an $F_{2,2}$.
- The F statistic for whole plots is 13260/6829= 1.94.
- The p-value of the test

$$H_0 : \mu_{360} = \mu_{370} = \mu_{380}$$

```
1-pf(q = 13260/6829,df1 = 2,df2 = 2)
```

```
[1] 0.3399373
```

ANOVA for split plot experiment - using `aov()` with `Error()`- sub plot

- The subplot effects of coating and the interaction of temperature and coating can be tested by forming F statistics using the subplot mean square error.
- These tests are given in the ANOVA table under the heading `Error: Within`.
- There are statistically significant differences between coatings and the interaction between temperature and coating.

ANOVA for split plot experiment

The values for the split plot experiment can be put into one ANOVA table.

Source	DF	SS	MS	F	P
Whole plot:					
replication	1	782	782	782/6829=0.12	0.77
heats	2	26519	13260	13260/6829=1.9	0.34
replication $\times$ heats	2	13658	6829		
(whole plot error)					
Subplot:					
coating	3	4289	1430	11.48	0.002
coating $\times$ heats	6	3270	545	4.376	0.02
Subplot error	9	1121	124.5		