Design of Engineering Experiments Part 3 – Analysis of Variance (4)

- Comparing more that two factor levels…**the analysis of variance**
  - ANOVA decomposition of total variability
  - Statistical testing & analysis
  - Checking assumptions, model validity
  - Post-ANOVA testing of means
  - Box-Cox method
- **Sample size** determination

Contrasts

- Many multiple comparison methods use the idea of a **contrast.**
- In the plasma etching experiment, if we had suspected at the start of the experiment that the *average* of the lowest levels of power did not differ from the *average* of the highest levels of power, then the hypothesis would have been

$H_0: \mu_1 + \mu_2 = \mu_3 + \mu_4$ $\quad$ $H_0: \mu_1 + \mu_2 - \mu_3 - \mu_4 = 0$

$H_a: \mu_1 + \mu_2 \neq \mu_3 + \mu_4$ $\quad$ *or* $\quad$ $H_a: \mu_1 + \mu_2 - \mu_3 - \mu_4 \neq 0$

- In general, a **contrast** Γ is a linear combination of parameters of the form

$$\Gamma = \sum_{i=1}^{a} c_i \mu_i, \text{ where } \sum_{i=1}^{a} c_i = 0$$

Hypothesis Testing using Contrast

- Hypothesis: $H_0 : \sum_{i=1}^{a} c_i \mu_i = 0 \text{ v.s. } H_a : \sum_{i=1}^{a} c_i \mu_i \neq 0$
- t-test based testing:

$$t_0 = \frac{\sum_{i=1}^{a} c_i \bar{y}_{i.}}{\sqrt{\frac{MS_E}{n} \sum_{i=1}^{a} c_i^2}} \overset{\text{under } H_0}{\sim} t(N-a)$$

- F-test based testing:

$$F_0 = t_0^2 = \frac{\left(\sum_{i=1}^{a} c_i \bar{y}_{i.} \right)^2}{\frac{MS_E}{n} \sum_{i=1}^{a} c_i^2} = \frac{\left(\sum_{i=1}^{a} c_i \bar{y}_{i.} \right)^2 / \left(\frac{1}{n} \sum_{i=1}^{a} c_i^2 \right)}{MS_E} \overset{\text{under } H_0}{\sim} F(1, N-a)$$

Confidence Interval for a Contrast

$$\sum_{i=1}^{a} c_i \bar{y}_{i.} - t_{\alpha/2, N-a} \sqrt{\frac{MS_E}{n} \sum_{i=1}^{a} c_i^2} \le \sum_{i=1}^{a} c_i \mu_i$$

$$\le \sum_{i=1}^{a} c_i \bar{y}_{i.} + t_{\alpha/2, N-a} \sqrt{\frac{MS_E}{n} \sum_{i=1}^{a} c_i^2}$$

Orthogonal Contrasts

- A useful special case of contrast is **orthogonal contrasts**

$$C_1 = \sum_{i=1}^{a} c_i \mu_i, \text{ where } \sum_{i=1}^{a} c_i = 0$$

$$C_2 = \sum_{i=1}^{a} d_i \mu_i, \text{ where } \sum_{i=1}^{a} d_i = 0$$

$$\text{if } \sum_{i=1}^{a} c_i d_i = 0, C_1 \text{ and } C_1 \text{ are orthogonal contrasts.}$$

- For *a* treatments, there are *a-1* orthogonal contrasts that partition the sum of squares due to treatments into *a-1* independent single-degree-of-freedom components. Thus, tests performed on orthogonal contrasts are independent.

Example 3-6

- Prior to running the experiment, it is suspected that the *average* of the lowest levels of power did not differ from the *average* of the highest levels of power.

Hypothesis:

$$H_0 : \mu_1 = \mu_2$$

$$H_0 : \mu_1 + \mu_2 = \mu_3 + \mu_4$$

$$H_0 : \mu_3 = \mu_4$$

Orthogonal Contrast

$$\mathrm{C}_1 = \bar{y}_{1.} - \bar{y}_{2.}$$

$$\mathrm{C}_2 = \bar{y}_{1.} + \bar{y}_{2.} - \bar{y}_{3.} - \bar{y}_{4.}$$

$$\mathrm{C}_3 = \qquad \bar{y}_{3.} + \bar{y}_{4.}$$

Table 3-1 Etch Rate Data (in Å/min) from the Plasma Etching Experiment

Power (W)	Observations 1	2	3	4	5	Totals	Averages
160	575	542	530	539	570	2756	551.2
180	565	593	590	579	610	2937	587.4
200	600	651	610	637	629	3127	625.4
220	725	700	715	685	710	3535	707.0

$$C_1 = +1(551.2) - 1(587.4) = -36.2 \qquad SS_{C_1} = \frac{(-36.2)^2}{\frac{1}{5}(2)} = 3276.10$$

$$C_2 = +1(551.2) + 1(587.4) - 1(625.4) - 1(707.0) = -193.8 \qquad SS_{C_2} = \frac{(-193.8)^2}{\frac{1}{5}(4)} = 46948.05$$

$$C_3 = +1(625.4) + 1(707.0) \qquad SS_{C_3} = \frac{(-81.6)^2}{\frac{1}{5}(2)} = 16646.4$$

Table 3-11 Analysis of Variance for the Plasma Etching Experiment

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0	P-Value
Power Setting	66,870.55	3	22,290.18	66.80	<0.001
Orthogonal contrasts					
$C_1: \mu_1 = \mu_2$	(3276.10)	1	3276.10	9.82	<0.01
$C_2: \mu_1 + \mu_3 = \mu_3 + \mu_4$	(46,948.05)	1	46,948.05	140.69	<0.001
$C_3: \mu_3 = \mu_4$	(16,646.40)	1	16,646.40	49.88	<0.001
Error	5,339.20	16	333.70		
Total	72,209.75	19			

R code

```
#orthogonal contrasts
>C<-matrix(c(1,1,0,-1,1,0,0,-1,1,0,-1,-1),nc=4)
>simtest(EtchRate~Power,cmatrix=C)

#althernatively, we can use “glht” function for
#orthogonal contrasts
>C<-matrix(c(1,1,0,-1,1,0,0,-1,1,0,-1,-1),nc=4)
>summary(glht(plasma.aov, linfct = mcp(Power = C)))
```

R output

```
Simultaneous Tests for General Linear Hypotheses

Multiple Comparisons of Means: User-defined Contrasts


Fit: aov(formula = EtchRate ~ Power)

Linear Hypotheses:
        Estimate Std. Error t value p value
1 == 0   -36.20      11.55  -3.133  0.0190 *
2 == 0  -193.80      16.34 -11.861  <0.001 ***
3 == 0   -81.60      11.55  -7.063  <0.001 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Adjusted p values reported)
```