Design of Engineering Experiments Part 4 – The Blocking Principle (2)

- Text Reference, Chapter 4
- **Blocking** and **nuisance factors**
- The randomized complete block design or the **RCBD**
- Extension of the ANOVA to the RCBD
- Other blocking **scenarios…Latin square designs**

Introduction

- **RCBD** -reduce the experimental error by removing variability due to a known and controllable nuisance variable.
- Two or more blocking variables can be used simultaneously in **RCBD.**
- The full use of two blocking variables in a **RCBD** often requires too many experimental units.
  - For example, if the two blocking variables have six levels each, 36 blocks would be required. If six treatments were to be studied, 216 subjects would be needed for the experiment

- Cost considerations may not permit the use of this many experimental units, yet precision and range of validity considerations may require the simultaneous use of two blocking variables, each with six levels, in order to reduce the experimental error variance sufficiently.
- An incomplete block design may be helpful. In such a design, all 36 blocks in our example would still be used, but now each block would not contain all six treatments.
- Latin Square Designs are a special type of incomplete block design in which only one treatment is run in each block. There is another reason, besides economy, why a latin square design is used, that blocks sometimes cannot contain more than one treatment.

Example 4-3

- Interested in the effects of five different formulations (Factor I) of a rocket propellant used in aircrew escape systems on the observed burning rate.
- Each formulation is mixed from a batch of raw material that is only large enough for five formulations to be tested. There may be substantial differences among batches (Factor I).
- The formulations are prepared by five operators (Factor III), and there may be substantial differences in the skills and experience of the operators.

The Rocket Propellant Problem – A Latin Square Design

Table 4-8 Latin Square Design for the Rocket Propellant Problem

Batches of Raw Material	Operators				
	1	2	3	4	5
1	$A = 24$	$B = 20$	$C = 19$	$D = 24$	$E = 24$
2	$B = 17$	$C = 24$	$D = 30$	$E = 27$	$A = 36$
3	$C = 18$	$D = 38$	$E = 26$	$A = 27$	$B = 21$
4	$D = 26$	$E = 31$	$A = 26$	$B = 23$	$C = 22$
5	$E = 22$	$A = 30$	$B = 20$	$C = 29$	$D = 31$

- This is a 5×5 Latin square design

A Latin square for p fators, or a p×p Latin square, is a square containing p rows and p columns. Each of the resulting p^2 cells contains one of the p letters that corresponds to the treatments, and each letter occurs once and only once in each row and column.

Table 4-12 Standard Latin Squares and Number of Latin Squares of Various Sizes[a]

Size	3 × 3	4 × 4	5 × 5	6 × 6	7 × 7	$p \times p$
Examples of standard squares	*A B C*	*A B C D*	*A B C D E*	*A B C D E F*	*A B C D E F G*	*ABC . . . P*
	B C A	*B C D A*	*B A E C D*	*B C F A D E*	*B C D E F G A*	*BCD . . . A*
	C A B	*C D A B*	*C D A E B*	*C F B E A D*	*C D E F G A B*	*CDE . . . B*
		D A B C	*D E B A C*	*D E A B F C*	*D E F G A B C*	⋮
			E C D B A	*E A D F C B*	*E F G A B C D*	
				F D E C B A	*F G A B C D E*	*PAB . . .* $(P-1)$
					G A B C D E F	
Number of standard squares	1	4	56	9408	16,942,080	—
Total number of Latin squares	12	576	161,280	818,851,200	61,479,419,904,000	$p!(p-1)! \times$ (number of standard squares)

[a]Some of the information in this table is found in *Statistical Tables for Biological, Agricultural and Medical Research*, 4th edition, by R. A. Fisher and F. Yates, Oliver and Boyd, Edinburgh, 1953. Little is known about the properties of Latin squares larger than 7 × 7.

- A Latin square in which the first row and column consists of the letters written in alphabetical order is called a **standard Latin Square**

- The observations in the Latin square should be taken in random order. The usual procedure is to select a Latin square from a table like above, and then arrange the order of the rows , columns and letters at random. (Fisher and Yates 1953)

The Latin Square Design

- These designs are used to simultaneously control (or eliminate) **two sources of nuisance variability by systematically allows blocking in two directions**
- A significant assumption is that the three factors (treatments, nuisance factors) **do not interact**
- If this assumption is violated, the Latin square design will not produce valid results
- Latin squares are not used as much as the RCBD in industrial experimentation

Statistical Analysis of the Latin Square Design

- The statistical (effects) model is

$$y_{ijk} = \mu + \alpha_i + \tau_j + \beta_k + \varepsilon_{ijk} \begin{cases} i = 1, 2, \ldots, p \\ j = 1, 2, \ldots, p \\ k = 1, 2, \ldots, p \end{cases}$$

- The statistical analysis (ANOVA) is much like the analysis for the RCBD.

Table 4-9 Analysis of Variance for the Latin Square Design

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0
Treatments	$SS_{\text{Treatments}} = \frac{1}{p}\sum_{j=1}^{p} y_{.j.}^2 - \frac{y_{...}^2}{N}$	$p-1$	$\frac{SS_{\text{Treatments}}}{p-1}$	$F_0 = \frac{MS_{\text{Treatments}}}{MS_E}$
Rows	$SS_{\text{Rows}} = \frac{1}{p}\sum_{i=1}^{p} y_{i..}^2 - \frac{y_{...}^2}{N}$	$p-1$	$\frac{SS_{\text{Rows}}}{p-1}$	
Columns	$SS_{\text{Columns}} = \frac{1}{p}\sum_{k=1}^{p} y_{..k}^2 - \frac{y_{...}^2}{N}$	$p-1$	$\frac{SS_{\text{Columns}}}{p-1}$	
Error	SS_E (by subtraction)	$(p-2)(p-1)$	$\frac{SS_E}{(p-2)(p-1)}$	
Total	$SS_T = \sum_i \sum_j \sum_k y_{ijk}^2 - \frac{y_{...}^2}{N}$	p^2-1		

Lecture notes on Experiment Design & Data Analysis

Table 4-11 Analysis of Variance for the Rocket Propellant Experiment

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0	P-Value
Formulations	330.00	4	82.50	7.73	0.0025
Batches of raw material	68.00	4	17.00		
Operators	150.00	4	37.50		
Error	128.00	12	10.67		
Total	676.00	24			

$$SS_T = \sum_i \sum_j \sum_k y_{ijk}^2 - \frac{y_{...}^2}{N} = 16805 - \frac{403225}{25} = 676$$

$$SS_{Batches} = \frac{1}{p}\sum_{i=1}^{p} y_{i..}^2 - \frac{y_{...}^2}{N} = \frac{111^2 + 134^2 + 130^2 + 128^2 + 132^2}{5} - \frac{403225}{25} = 68$$

$$SS_{Operator} = \frac{1}{p}\sum_{i=1}^{p} y_{..k}^2 - \frac{y_{...}^2}{N} = \frac{107^2 + 143^2 + 121^2 + 130^2 + 134^2}{5} - \frac{403225}{25} = 150$$

$$SS_{Formula} = \frac{1}{p}\sum_{i=1}^{p} y_{.j.}^2 - \frac{y_{...}^2}{N} = \frac{143^2 + 101^2 + 112^2 + 149^2 + 130^2}{5} - \frac{403225}{25} = 330$$

$$SS_E = 676 - 68 - 150 - 330 = 128$$

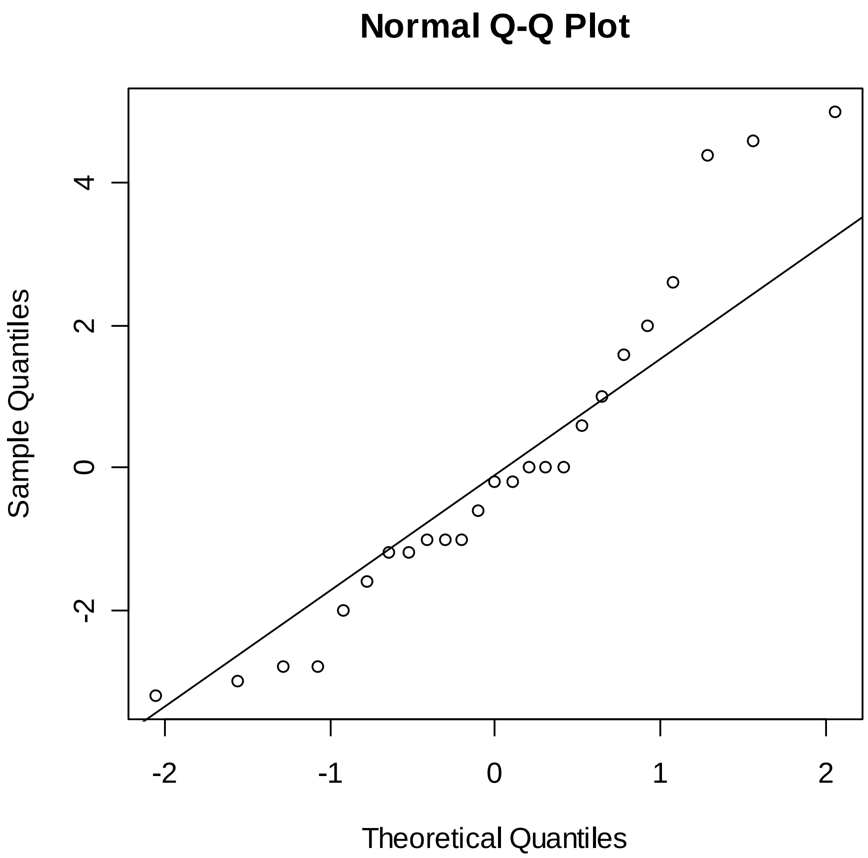
R code and output

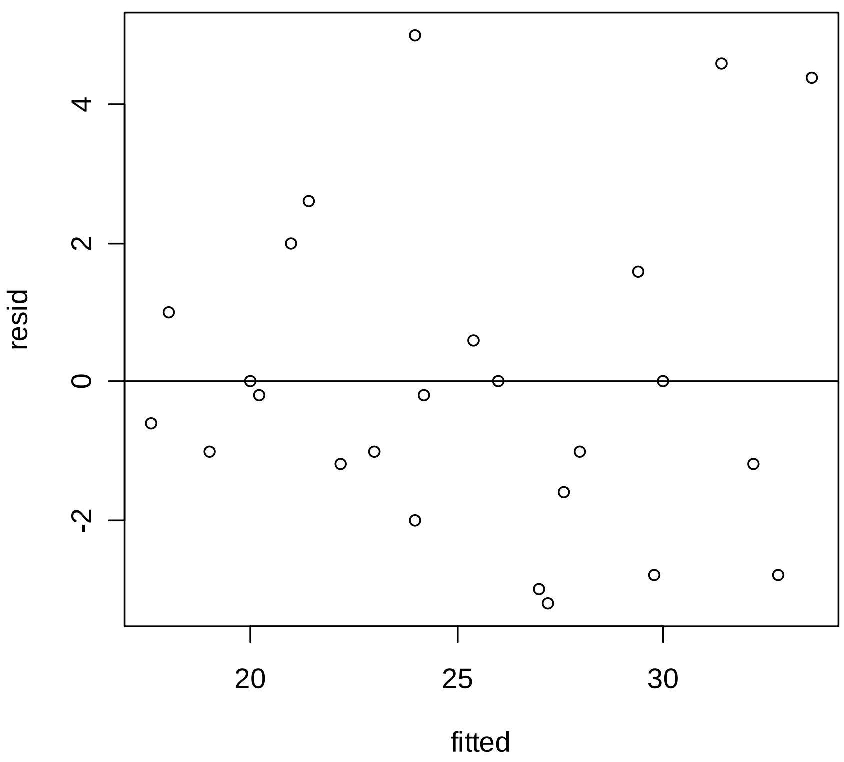
```
#####Latin square ANOVA # Read in Rocket Propellant Experiment data
>rocket=read.table("Rocket.txt", header=T)
>attach(rocket)
>Operator<-factor(Operator)
>Batch<-factor(Batch)
>Formula<-factor(Formula)
#ANOVA for latin square design
>rocket.fit<-lm(BurningRate~Formula+Batch+Operator)
>anova(rocket.fit)
```

```
Analysis of Variance Table

Response: BurningRate
          Df Sum Sq Mean Sq F value   Pr(>F)
Formula    4 330.00   82.50  7.7344 0.002537 **
Batch      4  68.00   17.00  1.5937 0.239059
Operator   4 150.00   37.50  3.5156 0.040373 *
Residuals 12 128.00   10.67
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Checking Assumption-R code

```
#Check Assumptions

>resid<-rocket.fit$resid
>fitted<-rocket.fit$fitted

#normality checking
>qqnorm(resid)
>qqline(resid)

#constant variances checking

>plot(fitted, resid)
>abline(h=0)
```


Normal Q-Q Plot
Sample Quantiles
Theoretical Quantiles
-2
0
2
4
-2
-1
0
1
2


resid
fitted
-2
0
2
4
20
25
30

Tukey's Multiple Comparison(1)

- Which formulas in general are significantly different from others? We can use Tukey pairwise interval to help determine this.

The width of the bands calculated as equation (3-37) in the textbook page 95

$$T_\alpha = \frac{q_\alpha(a, f)}{\sqrt{2}} \sqrt{MSE\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}$$

R code:

```
>qtukey(0.95,5,12)*sqrt(10.67*(1/5+1/5))/sqrt(2)
[1] 6.58496
```

Tukey's Multiple Comparison(2)

>rocket.fit$coef

```
(Intercept)   Formula2   Formula3   Formula4   Formula5
21.4               -8.4       -6.2       1.2        -2.6
Batch2   Batch3   Batch4   Batch5
4.6      3.8          3.4     4.2
Operator2   Operator3   Operator4 Operator5
7.2              2.8       4.6        5.4
```

>rocket.fit$coef[2:5]

```
Formula2 Formula3 Formula4 Formula5
    -8.4    -6.2       1.2      -2.6
```

Tukey’s Multiple Comparison(3)

```
#make a table for the formulation differencs
> scoefs<-c(0,rocket.fit$coef[2:5])
>outer(scoefs,scoefs, "-")
```

		Formula2	Formula3	Formula4	Formula5
	0.0	8.4	6.2	-1.2	2.6
Formula2	-8.4	0.0	-2.2	-9.6	-5.8
Formula3	-6.2	2.2	0.0	-7.4	-3.6
Formula4	1.2	9.6	7.4	0.0	3.8
Formula5	-2.6	5.8	3.6	-3.8	0.0