Unit 7: Orthogonal Arrays and Response Surface Methodology

Source : Sections 8.1 - 8.2, Appendix 8A, 8C; Sections 10.1 - 10.2

- In Tables 1 and 2, the design used does not belong to the 2^{k-p} series (Chapter 5) or the 3^{k-p} series (Chapter 6), because the latter would require run size as a power of 2 or 3. These designs belong to the class of orthogonal arrays.
- An **orthogonal array** array $OA(N, s_1^{m_1} \dots s_\gamma^{m_\gamma}, t)$ of strength t is an $N \times m$ matrix, $m = m_1 + \dots + m_\gamma$, in which m_i columns have $s_i (\geq 2)$ symbols or levels such that, for any t columns, all possible combinations of symbols appear equally often in the matrix.
- For OA of strength two, the index $t = 2$ is dropped for simplicity.
- An $OA(12, 2^{11})$ is used in Table 1 and an $OA(18, 2^1 3^7)$ is used in Table 2.

Example : $OA(12, 2^{11})$

Table 1: Design Matrix and Lifetime Data, Cast Fatigue Experiment

	Factor											Logged
Run	*A*	*B*	*C*	*D*	*E*	*F*	*G*	8	9	10	11	Lifetime
1	+	+	–	+	+	+	–	–	–	+	–	6.058
2	+	–	+	+	+	–	–	–	+	–	+	4.733
3	–	+	+	+	–	–	–	+	–	+	+	4.625
4	+	+	+	–	–	–	+	–	+	+	–	5.899
5	+	+	–	–	–	+	–	+	+	–	+	7.000
6	+	–	–	–	+	–	+	+	–	+	+	5.752
7	–	–	–	+	–	+	+	–	+	+	+	5.682
8	–	–	+	–	+	+	–	+	+	+	–	6.607
9	–	+	–	+	+	–	+	+	+	–	–	5.818
10	+	–	+	+	–	+	+	+	–	–	–	5.917
11	–	+	+	–	+	+	+	–	–	–	+	5.863
12	–	–	–	–	–	–	–	–	–	–	–	4.809

Example : $OA(18, 2^1 3^7)$

Table 2: Design Matrix and Response Data, Blood Glucose Experiment

	Factor								Mean
Run	*A*	*G*	*B*	*C*	*D*	*E*	*F*	*H*	Reading
1	0	0	0	0	0	0	0	0	97.94
2	0	0	1	1	1	1	1	1	83.40
3	0	0	2	2	2	2	2	2	95.88
4	0	1	0	0	1	1	2	2	88.86
5	0	1	1	1	2	2	0	0	106.58
6	0	1	2	2	0	0	1	1	89.57
7	0	2	0	1	0	2	1	2	91.98
8	0	2	1	2	1	0	2	0	98.41
9	0	2	2	0	2	1	0	1	87.56
10	1	0	0	2	2	1	1	0	88.11
11	1	0	1	0	0	2	2	1	83.81
12	1	0	2	1	1	0	0	2	98.27
13	1	1	0	1	2	0	2	1	115.52
14	1	1	1	2	0	1	0	2	94.89
15	1	1	2	0	1	2	1	0	94.70
16	1	2	0	2	1	2	0	1	121.62
17	1	2	1	0	2	0	1	2	93.86
18	1	2	2	1	0	1	2	0	96.10

Why Using Orthogonal Array

- **Run size economy.** Suppose 8-11 factors at two levels are to be studied. Using an $OA(12, 2^{11})$ will save 4 runs over a 16-run 2^{k-p} design. Similarly, suppose 5-7 factors at three levels are to be studied. Using an $OA(18, 3^{7})$ will save 9 runs over a 27-run 3^{k-p} design.
- **Flexibility.** Many *OA*'s exist for flexible combinations of factor levels. See the collection on next page.
- Analysis strategy for experiments based on *OA* can be found in Chapter 9 of WH.

Useful Orthogonal Arrays

- Collection in Appendix 8A and 8C of WH:

$$
\begin{array}{ccc}
{*}OA(12, 2^{11}) & OA(12, 3^1 2^4) & {*}OA(18, 2^1 3^7) \\
OA(18, 6^1 2^6) & OA(20, 2^{19}) & OA(24, 3^1 2^{16}) \\
OA(24, 6^1 2^{14}) & {*}OA(36, 2^{11} 3^{12}) & OA(36, 3^7 6^3) \\
OA(36, 2^8 6^3) & OA(48, 2^{11} 4^{12}) & OA(50, 2^1 5^{11}) \\
 & OA(54, 2^1 3^{25}) &
\end{array}
$$

  * especially useful

- Learn to choose and use the design tables in the collection.

$OA(18, 2^1 3^7)$ and $OA(18, 6^1 3^6)$

Table 3: $OA(18, 2^1 3^7)$ (columns 1–8) and $OA(18, 6^1 3^6)$ (columns $1'$ and 3–8)

Run	$1'$	1	2	3	4	5	6	7	8
1	0	0	0	0	0	0	0	0	0
2	0	0	0	1	1	1	1	1	1
3	0	0	0	2	2	2	2	2	2
4	1	0	1	0	0	1	1	2	2
5	1	0	1	1	1	2	2	0	0
6	1	0	1	2	2	0	0	1	1
7	2	0	2	0	1	0	2	1	2
8	2	0	2	1	2	1	0	2	0
9	2	0	2	2	0	2	1	0	1
10	3	1	0	0	2	2	1	1	0
11	3	1	0	1	0	0	2	2	1
12	3	1	0	2	1	1	0	0	2
13	4	1	1	0	1	2	0	2	1
14	4	1	1	1	2	0	1	0	2
15	4	1	1	2	0	1	2	1	0
16	5	1	2	0	2	1	2	0	1
17	5	1	2	1	0	2	0	1	2
18	5	1	2	2	1	0	1	2	0

$OA(36, 2^{11}3^{12})$

Table 4: $OA(36, 2^{11}3^{12})$

Run	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1
3	0	0	0	0	0	0	0	0	0	0	0	2	2	2	2	2	2	2	2	2	2	2	2
4	0	0	0	0	0	1	1	1	1	1	1	0	0	0	0	1	1	1	1	2	2	2	2
5	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	2	2	2	2	0	0	0	0
6	0	0	0	0	0	1	1	1	1	1	1	2	2	2	2	0	0	0	0	1	1	1	1
7	0	0	1	1	1	0	0	0	1	1	1	0	0	1	2	0	1	2	2	0	1	1	2
8	0	0	1	1	1	0	0	0	1	1	1	1	1	2	0	1	2	0	0	1	2	2	0
9	0	0	1	1	1	0	0	0	1	1	1	2	2	0	1	2	0	1	1	2	0	0	1
10	0	1	0	1	1	0	1	1	0	0	1	0	0	2	1	0	2	1	2	1	0	2	1
11	0	1	0	1	1	0	1	1	0	0	1	1	1	0	2	1	0	2	0	2	1	0	2
12	0	1	0	1	1	0	1	1	0	0	1	2	2	1	0	2	1	0	1	0	2	1	0
13	0	1	1	0	1	1	0	1	0	1	0	0	1	2	0	2	1	0	2	2	1	0	1
14	0	1	1	0	1	1	0	1	0	1	0	1	2	0	1	0	2	1	0	0	2	1	2
15	0	1	1	0	1	1	0	1	0	1	0	2	0	1	2	1	0	2	1	1	0	2	0
16	0	1	1	1	0	1	1	0	1	0	0	0	1	2	1	0	0	2	1	2	2	1	0
17	0	1	1	1	0	1	1	0	1	0	0	1	2	0	2	1	1	0	2	0	0	2	1
18	0	1	1	1	0	1	1	0	1	0	0	2	0	1	0	2	2	1	0	1	1	0	2

$OA(36, 2^{11}3^{12})$

Table 4: Continued

Run	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
19	1	0	1	1	0	0	1	1	0	1	0	0	1	0	2	2	2	0	1	1	0	1	2
20	1	0	1	1	0	0	1	1	0	1	0	1	2	1	0	0	0	1	2	2	1	2	0
21	1	0	1	1	0	0	1	1	0	1	0	2	0	2	1	1	1	2	0	0	2	0	1
22	1	0	1	0	1	1	1	0	0	0	1	0	1	1	2	2	0	1	0	0	2	2	1
23	1	0	1	0	1	1	1	0	0	0	1	1	2	2	0	0	1	2	1	1	0	0	2
24	1	0	1	0	1	1	1	0	0	0	1	2	0	0	1	1	2	0	2	2	1	1	0
25	1	0	0	1	1	1	0	1	1	0	0	0	2	1	0	1	2	2	0	2	0	1	1
26	1	0	0	1	1	1	0	1	1	0	0	1	0	2	1	2	0	0	1	0	1	2	2
27	1	0	0	1	1	1	0	1	1	0	0	2	1	0	2	0	1	1	2	1	2	0	0
28	1	1	1	0	0	0	0	1	1	0	1	0	2	1	1	1	0	0	2	1	2	0	2
29	1	1	1	0	0	0	0	1	1	0	1	1	0	2	2	2	1	1	0	2	0	1	0
30	1	1	1	0	0	0	0	1	1	0	1	2	1	0	0	0	2	2	1	0	1	2	1
31	1	1	0	1	0	1	0	0	0	1	1	0	2	2	2	1	2	1	1	0	1	0	0
32	1	1	0	1	0	1	0	0	0	1	1	1	0	0	0	2	0	2	2	1	2	1	1
33	1	1	0	1	0	1	0	0	0	1	1	2	1	1	1	0	1	0	0	2	0	2	2
34	1	1	0	0	1	0	1	0	1	1	0	0	2	0	1	2	1	2	0	1	1	2	0
35	1	1	0	0	1	0	1	0	1	1	0	1	0	1	2	0	2	0	1	2	2	0	1
36	1	1	0	0	1	0	1	0	1	1	0	2	1	2	0	1	0	1	2	0	0	1	2

Poorman's Response Surface Methodology

- Consider an experiment to study three quantitative factors with up to 5 levels.

Table 5: Factors and Levels, Ranitidine Experiment

Factor	Levels
A. pH	2, 3.42, 5.5, 7.58, 9
B. voltage (kV)	9.9, 14, 20, 26, 30.1
C. α-CD (mM)	0, 2, 5, 8, 10

- The design matrix and the data are given on the next page. The design differs from 2^{k-p} design in two respects :
  - 6 replicates at the center,
  - 6 runs along the three axes.

  It belongs to the class of *central composite designs*.

Ranitidine Experiment

Table 6: Design Matrix and Response Data

	Factor				
Run	*A*	*B*	*C*	CEF	ln CEF
1	−1	−1	−1	17.293	2.850
2	1	−1	−1	45.488	3.817
3	−1	1	−1	10.311	2.333
4	1	1	−1	11757.084	9.372
5	−1	−1	1	16.942	2.830
6	1	−1	1	25.400	3.235
7	−1	1	1	31697.199	10.364
8	1	1	1	12039.201	9.396
9	0	0	−1.67	7.474	2.011
10	0	0	1.67	6.312	1.842
11	0	−1.68	0	11.145	2.411
12	0	1.68	0	6.664	1.897
13	−1.68	0	0	16548.749	9.714
14	1.68	0	0	26351.811	10.179
15	0	0	0	9.854	2.288
16	0	0	0	9.606	2.262
17	0	0	0	8.863	2.182
18	0	0	0	8.783	2.173
19	0	0	0	8.013	2.081
20	0	0	0	8.059	2.087

Central Composite Designs

- General definition in Section 10.7 and designs in Table 10A.1. (Not required for the course).
- A simple CCD is shown graphically below. It has three parts
  (1) *cube* (or corner) points, (2) *axial* (or star) points, (3) *center* points.

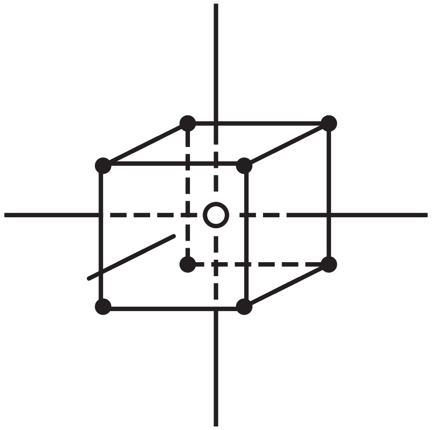

Figure 1: A Central Composite Design in Three Dimensions [cube point (dot), star point (cross), center point (circle)].

Sequential Nature of RSM

1. **Screening Experiment :** When many variables are considered, some are likely to be inert. Use *a* 2^{k-p} *design or an OA.* If the experimental region is far from the optimum, use the **first-order model**

   $$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \varepsilon, \qquad (1)$$

   to fit the data.

2. Based on the fitted model, find the steepest ascent direction and perform a search along this direction (called **steepest ascent search**).

   Steps 1 and 2 may be repeated until reaching the optimum region (e.g. peak of the surface).

Sequential Nature of RSM (Contd.)

3. To capture the curvature effects, use a **second-order design** (like the central composite design). Fit a **second-order model**

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i<j}^{k} \beta_{ij} x_i x_j + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \varepsilon \qquad (2)$$

   to data. Use the fitted model (with insignificant terms dropped) to do *contour plots* and find the *optimum* conditions.

   A graphical illustration of these steps is given on next page.

Sequential Exploration of Response Surface