均匀设计

方开泰

中国科学院应用数学研究所
概率统计咨询服务部 印

1983年元月

目 录

前　　言

在生产和科学研究中，经常需要做试验，如何做试验这里面大有学问。试验设计得好，试验次数不多，就能达到予期的目的；设计得不好，会事倍功半，甚到劳而无功。七十年代以来，在我国比较普及的优选法和正交试验法都是科学的试验的方法，相当多的试验可用它们来安排，并收到了良好的效果，取得了很大的成绩。但是世界上的事物是复杂的，没有能治百病的灵丹妙药，也没有任何情况都适用的试验设计方法，不同性质的对象要求不同的试验方法。在一些问题中，需要考察的因素较多，且每个因素变化的范围较大，因而每个因素需要分较多的等级（在试验设计中叫做水平），这一类问题用现有的试验设计方法（包括优选法和正交试验法）解决起来均不太方便，且需要的试验次数较多，而均匀设计法正是解决这一类试验的有效工具。

在均匀设计的研究中，曾得王元研究员的指导和大力帮助特此志谢。

本讲义于1979年曾由中国科学院数学研究所印出，三年多来收到许多读者来信，许多同志登门前来讨论，可惜因这段期间我大部分时间在国外，未能向这些同志询求意见，趁这次重印之机，希望得到广大读者的批评和指教。

方开泰

1983年元月

一、均匀设计的思想

先看一个例子

例1. 无粮上浆是纺织工业的一项重要改革，用化学原料来代替淀粉，可节省大批的粮食，CMC（羧甲基纤维素钠）就是一种代替淀粉的化学原料，今为了寻找CMC的最佳生产条件，考虑有关的三个因素：碱化时间（以A表示），烧碱浓度（B）和醚化时间（C），它们的变化范围选择为：

烧碱时间（A）120分～180分

烧碱浓度（B）25°～29°

醚化时间（C）90分～150分

由于时间变化范围较大，取五个等级（即水平）来考虑，为了醒目，一般列成因素水平表（表1.1）

表1.1　因素水平表

因素 \ 水平	1	2	3	4	5
碱化时间（分）	120分	135分	150分	165分	180分
烧碱浓度（°）	25°	26°	27°	28°	29°
醚化时间（分）	90分	105分	120分	135分	150分

为了叙述的方便，将因素A的五个水平分别记作 A_1、A_2、A_3、A_4、和 A_5；B和C也有类似的记号。例如 A_2 是A的第2个水平，$A_2 = 135$ 分；$B_4 = 28°$，$C_3 = 120$ 分

这样一项试验如何来设计呢？有如下几种常见的方法：

1．全面试验，就是让A，B，C的每一个水平彼此都有组合在一起做试验的机会。用组合 $A_i\ B_j\ C_K$ 表示在因素A取 i 水平，B取 B_j 和C取 C_K 的条件下所作的试验，对例1来说，全面试验共需要125次试验，它们是：$A_1B_1C_1$，$A_1B_1C_2$，

$A_1B_1C_3$, $A_1B_1C_4$, $A_1B_1C_5$, $A_1B_2C_1$, $A_1B_2C_2$, ……, $A_5B_5C_1$, $A_5B_5C_2$, $A_5B_5C_3$, $A_5B_5C_4$, $A_5B_5C_5$。

一般说来，如有 m 个因素，它们各有 l_1, l_2, ……, l_m 个水平，则全面试验共需做 $l_1 l_2 \cdots l_m$ 次试验，对例1，$m=3$，$l_1=l_2=l_3=5$，故共需试验次数 $5\cdot5\cdot5=5^3=125$ 次，全面试验的优点是分析的结果比较仔细，结论较精确；但由于它需要较多的试验，在多因多水平的场合常々是不可取的，有时甚至是不可能的。例如，有六个因素，每个都是五水平，则需 $5^6=15625$ 次试验，在实际中是不可能做到的。故需要一种试验次数较少，效果又与全面试验相近的试验设计方法。

2. 正交试验，一种直观而自然的想法是在全面试验中排出最有代表性的点子做一下，既可节省试验次数，又可达到既定目的。比如一个工厂让全厂职工讨论生产问题，如果让每个职工都发言，则会议势必很长，效果并不好。不如让职工选出各方面的代表（工人、干部、技术人员、妇女、青年等）来发言。会议时间可大大缩短，也可达到预期目的。但这些发言的代表必须是有充分代表性的，否则虽然减少了会议时间，但意见并没有充分反映上来。于是要解决如何选择最有代表的试验问题。

有一类正交表它可以帮助我们来挑选试验点，如表1.2是其中的一个正交表，每个正交表有一个代号，这个正交表的代号是 $L_{25}(5^6)$，其中"L"表示正交表，"25"表示总共要做25次试验，"5"表示每个因素都是5个水平，"6"表示这个表最多可以安排六个因素。常用的正交表除了 $L_{25}(5^6)$ 外还有 $L_4(2^3)$，$L_8(2^7)$，$L_{16}(2^{15})$，$L_{32}(2^{31})$，$L_9(3^4)$，$L_{27}(3^{13})$，$L_{16}(4^5)$ 等。这些正交表都具有从全面试验中挑择好的有代表性试验的能力。

对例1，如何用正交表来设计试验呢？大体有如下步骤：

(1) 选择合适的正交表。此例是三因素五水平试验，用 $L_{25}(5^6)$ 比较合适。

(2) 将 A, B、C 三因素放到 $L_{25}(5^6)$ 的任意三列的表头上，例如放在前三列。

(3) 把 A, B、C 对应三列的"1"，"2"，"3"翻译成具体的水平，如表 1.3 所示。

(4) 25 次试验方案是：第 1 号试验的工艺条件是 A_1 (120 分)，B_1 (25°)，C_1 (90 分)；第 2 号试验的工艺条件是 A_1 (120 分)，B_2 (26°) C_2 (105 分)；……，第 25 号试验的条件是：A_5 (180 分)，B_5 (29°)，C_4 (135 分)。这样试验方案就排好了。

正交表 $L_{25}(5^6)$ 从全面试验的 125 个试验点中挑选了 25 个点，这些点有什么特点呢？

表 1.2　正交表 $L_{25}(5^6)$

试验号 \ 列号	1	2	3	4	5	6
1	1	1	1	1	1	1
2	1	2	2	2	2	2
3	1	3	3	3	3	3
4	1	4	4	4	4	4
5	1	5	5	5	5	5
6	2	1	2	3	4	5
7	2	2	3	4	5	1
8	2	3	4	5	1	2
9	2	4	5	1	2	3
10	2	5	1	2	3	4
11	3	1	3	5	2	4
12	3	2	4	1	3	5
13	3	3	5	2	4	1
14	3	4	1	3	5	2
15	3	5	2	4	1	3
16	4	1	4	2	5	3
17	4	2	5	3	1	4
18	4	3	1	4	2	5
19	4	4	2	5	3	1
20	4	5	3	1	4	2
21	5	1	5	4	3	2
22	5	2	1	5	4	3
23	5	3	2	1	5	4
24	5	4	3	2	1	5
25	5	5	4	3	2	1

(A) 对任两个因素而言，这 25 次试验都是全面试验，这样可以保持全面试验的一些优点，并使得试验有可比性。

(B) 任一因素各水平试验的重复数都是相等的（此例每水平都重复五次）。

(C) 对绝大部分正交表而言，各列是完全等价的。例如此例用前三列或 1、2、4 三列或用 4、5、6 三列（即六列中任取三列），其试验效果是等价的。

表 1.3　正交试验方案

试验号 \ 列号	1 (A)	2 (B)	3 (C)
1	1 (120分)	1 (25°)	1 (90分)
2	1 〃	2 (26°)	2 (105分)
3	1 〃	3 (27°)	3 (120分)
4	1 〃	4 (28°)	4 (135分)
5	1 〃	5 (29°)	5 (150分)
6	2 (135分)	1 (25°)	2 (105分)
7	2 〃	2 (26°)	3 (120分)
8	2 〃	3 (27°)	4 (135分)
9	2 〃	4 (28°)	5 (150分)
10	2 〃	5 (29°)	1 (90分)
11	3 (150分)	1 (25°)	3 (120分)
12	3 〃	2 (26°)	4 (135分)
13	3 〃	3 (27°)	5 (150分)
14	3 〃	4 (28°)	1 (90分)
15	3 〃	5 (29°)	2 (105分)
16	4 (165分)	1 (25°)	4 (135分)
17	4 〃	2 (26°)	5 (150分)
18	4 〃	3 (27°)	1 (90分)
19	4 〃	4 (28°)	2 (105分)
20	4 〃	5 (29°)	3 (120分)
21	5 (180分)	1 (25°)	5 (150分)
22	5 〃	2 (26°)	1 (90分)
23	5 〃	3 (27°)	2 (105分)
24	5 〃	4 (28°)	3 (120分)
25	5 〃	5 (29°)	4 (135分)

这三条特点概括说来，对每个因素和每个水平都一视同仁，挑选的试验点在其试验范围内具有"均匀分散，整齐可比"的特性。由于有这些好的性质，正交试验可以用较少的试验来考察众多的因素，深受广大工人，农民和知识分子的欢迎，近若干年来，在我国已深入人心，并取得丰硕成果。

由正交表的这三个特点，容易看出，试验次数是水平数平

方的整数倍数，如例1，试验次数25是水平数与5平方的一倍；如用L_8（2^7），试验次数8是水平数2平方的二倍。因此正交试验只适合水平不太多（一般≤5）的多因素试验，当水平较多时，试验次数也还是很可观的，如11个水平的试验至少要$11^2=121$次，30个水平的试验至少要$30^2=900$次。而多水平的试验在农业，国防标准化工作和部分工业试验中是经常出现的，因此需要有另一种方法来解决这一类问题，这就产生了均匀设计。

3．均匀设计，上面我们说过，正交试验有"均匀分散，整齐可比"的特点。"均匀分散"性使试验点均衡地分布在试验范围内，让每个试验点有充分的代表性，因此，即使在正交表中各列都排满的情况，也能得到满意的结果；"整齐可比"性使试验结果的分析十分方便，可以估计各因素对指标的影响，找出影响事物变化的主要矛盾。但是我们注意到，为了有"整齐可比"性，对任两个因素它必须是全面试验，每个因素的各水平必须有重复，这样试验点在其试验范围内并不能做到充分"均匀分散"；为了有"整齐可比"性，试验的数目就必须比较多。这启示我们不考虑"整齐可比"，而让点子在试验范围内充分"均匀分散"，这样每个试验点可有更好的代表性。这种单纯从均匀性出发的设计称之为均匀设计。

如例1的试验，每个因素的每个水平都做了五次试验，如果每个水平只做一次，同样做25次试验，在试验范围内分成25个水平，则可将点子散布得更均匀。图1.1表示只有两因素情况下正交试验与均匀设计的区别，正交试验只取五个水平，每个水平重复五次，而均匀设计取25个水平，每个水平只做一次，由于这样我们看到均匀设计的试验点散布得更均匀，可以猜想它们更具有代表性，当因素更多时，均匀设计的这个优

点就为突出。如果这项试验费用很贵或由于其他原因希望减少次数，均匀设计法不减少水平数（至少有五个水平）的前提下，可以很方便的安排试验数（n）在 $5 \leq n < 25$ 的均匀试验。那么，均匀设计如何来安排试验呢？我们就来详细的介绍它的方法。

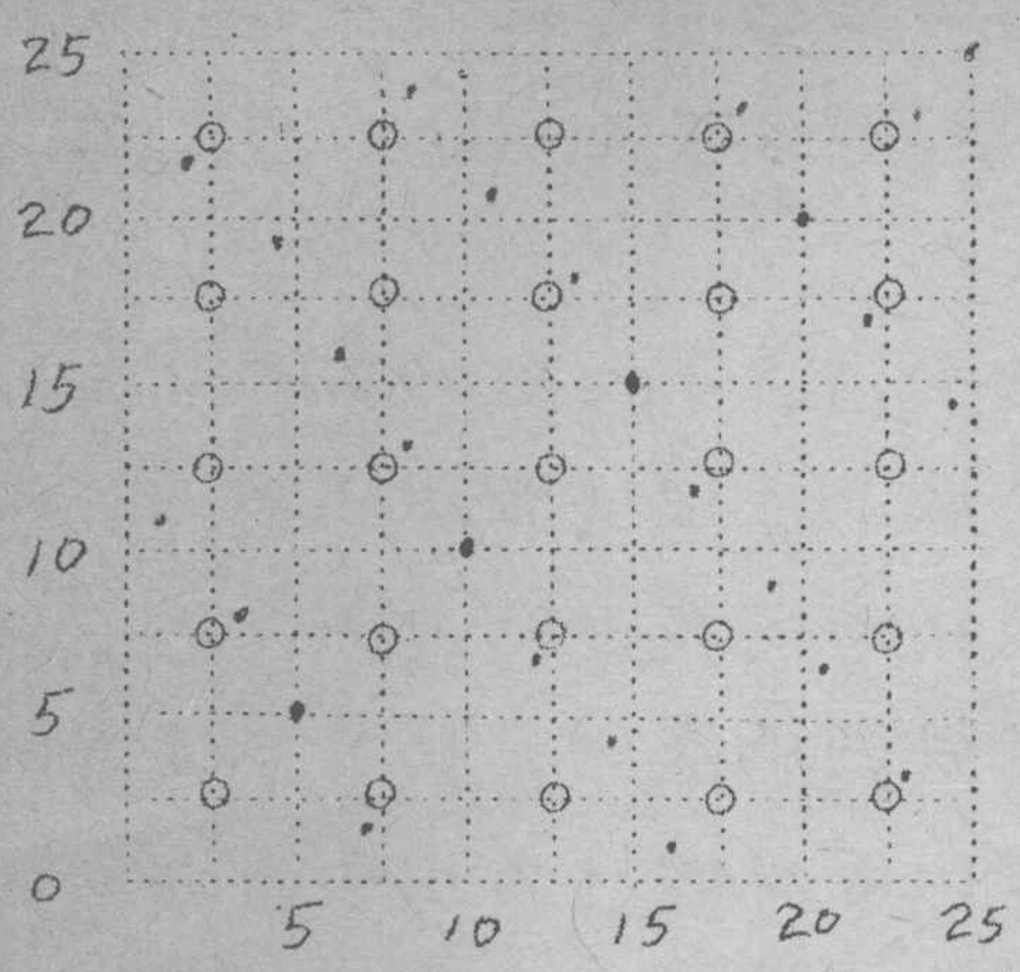


图 1.1 正交试验与均匀设计正交试验点用 ◦ 表示，均匀设计点用 • 表示。

二、试验的安排

根据均匀设计的思想也造了一套类似于正交表的均匀设计表，如表 2.1 和 2.2 就是其中的两个，每个表也有一个代号，表 2.1 的代号是 $U_5(5^4)$，其中"U"表示均匀设计表，U 的下标"5"表示要做五次试验，括号内的"5"表示每个因素都有五个水平，指数"4"表示最多只有允许安排四个因素。类似的表 2.2 的代号是 $U_6(6^6)$，它表示要做六次试验，每个因素有六个水平，最多可安排六个因素。我们知道，由于六阶正交拉丁方不存在，要安排四个六水平因素的试验，如用正交表，用 36 次试验都不行，至少需要 72 次试验，而用均匀设计六次试验就能安排了，这是均匀设计的很大优点。

那么均匀设计表有什么特点呢？

表 2.1 $U_5(5^4)$

试验号＼列号	1	2	3	4
1	1	2	3	4
2	2	4	1	3
3	3	1	4	2
4	4	3	2	1
5	5	5	5	5

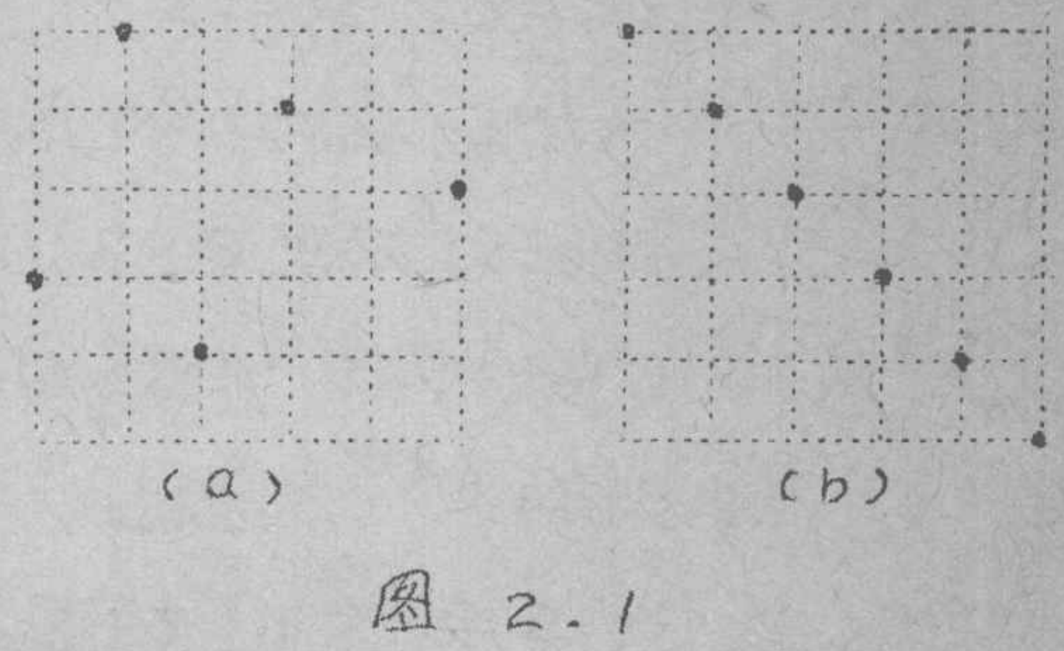
表 2.2 $U_6(6^6)$

试验号＼列号	1	2	3	4	5	6
1	1	2	3	4	5	6
2	2	4	6	1	3	5
3	3	6	2	5	1	4
4	4	1	5	2	6	3
5	5	3	1	6	4	2
6	6	5	4	3	2	1

1. 每个因素每个水平只做一次试验。从表 2.1 和表 2.2 看得很清楚，其它的均匀设计表（见附表）也有这个性质。

2. 任两个因素的试验画在平面的格子点上，每行每列恰好有一个试验点。如果表 2.2 的第一列和第三列点图得图 2.1(a) 我们看到每行每列恰有一个点。

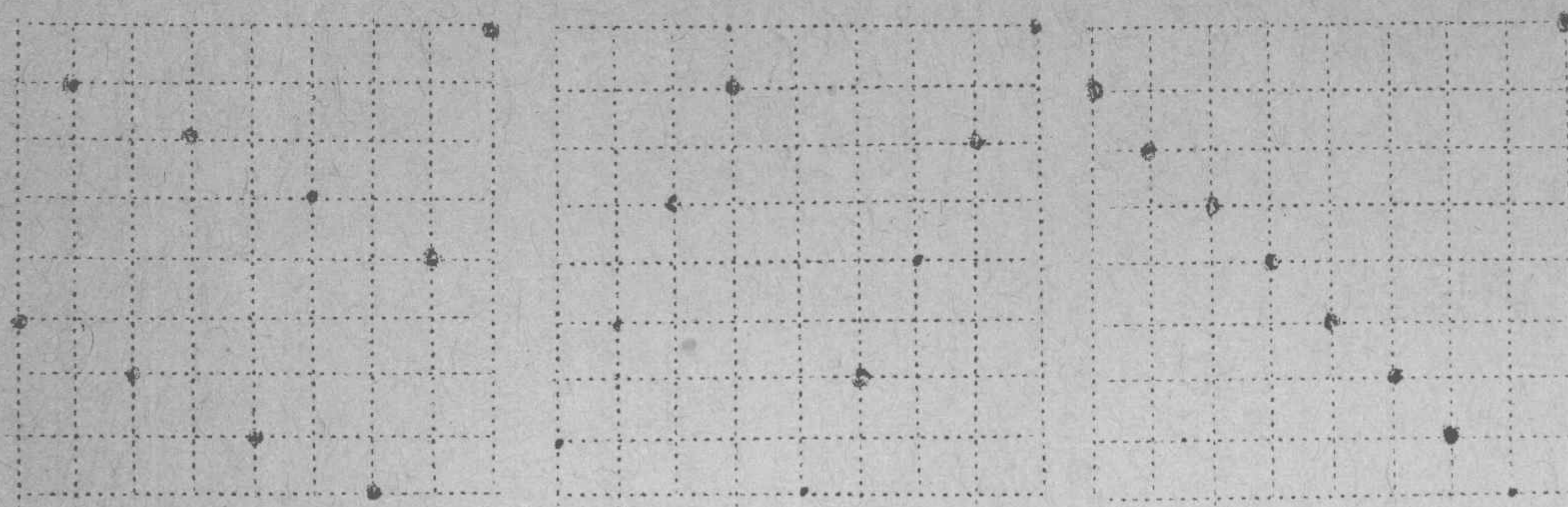
3. 均匀设计表任两列之间不一定是平等的，例如，用 $U_6(6^6)$ 的第一、三列和第一、六列分别画图，得图 2.1（a）和图 2.1（b），我们看到（a）的点很均匀，而（b）的点就不那么均匀了。又如将 $U_9(9^6)$（见附录）的第一、三列，第一、二列，第一、六列分别点图，得图 2.2 的（a），（b）和（c），我们看到（a）最均匀，（b）次之，（c）最不均匀。因此使用均匀设计表一般不宜



（a）　　（b）

图 2.1

随意排列，而应当选择均匀性搭配得比较好的列。于是，我们事先进行了精心挑选（按均匀性原则，然后用电子机째机计算，



（a）　　　　（b）　　　　（c）

图 2.2

详见第六节），挑选的结果立成了一个表，表 2.3是其中的一个；它告诉我们使用 $U_5(5^4)$时如何来挑选列：当只有两个因素时，用 1、2 列，当有三个因素时，用 1、2、4 列；当有四个因素时，四个列都用，自然没有选择的余地了。附录的均匀设计表，每个都附带一个使用表，在安排试验时，应当遵照使用表的规定，才能达到较好的效果。

表 2.3 $U_5(5^4)$的使用表

素数	列　号
2	1、2
3	1、2、4
4	1、2、3、4

那么是不是只能完全按照使用表的规定，不能有其它较好的安排呢？我们再回过来看图 2.1，那里，使用 1、6列效果不好，使用 1、3列的效果好，有没有与第 6 列配合得也很好的列呢？例如用第 2列和第 6 列画图，正好是图 2.1 的（a），也就是说，用 1、3 列与用 2、6 列是等价的；类似的我们容易发现与它们等价的还有 3、2列、4、5列、5、1列、6、4 列等。但对使用者来说，知道这么多等价的最好设计并无必要，故一般只

要按使用表执行就行了。

4．试验表之间的关系。附录只列了试验次数和水平数为奇数的表，$U_5(5^4)$，$U_7(7^6)$，$U_9(9^6)$……$U_{31}(31^{30})$，没有试验次数和水平数为偶数的表，那么如何如得到偶数次试验的表呢？这很简单，将奇数的表划去最后一行就得到比它次数少一的偶数表，而使用表不变。例如，将$U_7(7^6)$划去最后一行得到$U_6(6^6)$，使用表不变。

5．当水平数增加时试验数按水平数的增加量在增加。如当水平数从九水平增加到十水平时，试验数也从九增加到十。而正交试验当水平数增加时，试验数按平方的比例在增加，如当水平数从九增加到十时，试验数一般将从81增加到100。因此在正交试验中增加水平将使试验量有较大的增加，而均匀设计只有微少的增加，这是均匀设计的很大优点。

关于均匀设计的其它特点，以后再陆续介绍。现在转入试验的安排。例1的试验如用均匀设计表该如何安排呢？

显然它可以用$U_5(5^4)$来安排，由表2.3应选用1、2、4三列，将三个因素分别放在这三列的表头上，然后将表中抽象的水平换成因素具体的水平，得表2.4的试验方案。

三个因素只做了五次试验，在这么大试验范围内做这么几次试验是少了一点，试验点过少了会影响试验的结果，所得结论可靠性也差。于是可以考虑用$U_{10}(10^{10})$做十次试验，试验安排可用拟水平法，即原来每个因素都是五水平，要扩充成十水平，就将每个水平重复一次，比如A_1～A_5是原来的五个水平，A_6～A_{10}重复原来的五个水平，于是得到表2.5的试验方案。由于增加了五次试验，估计试验的结论会更加可靠。

细心的读者也许已经注意到，既然采用了表$U_{10}(10^{10})$，何不将每个因素在同一试验范围内分成十个水平，这样试验点

表 2.4 $U_5(5^3)$ 试验方案

列号 / 试验号	1 (A)	2 (B)	4 (C)
1	1 (120分)	2 (26°)	4 (135分)
2	2 (135分)	4 (28°)	3 (120分)
3	3 (150分)	1 (25°)	2 (105分)
4	4 (165分)	3 (27°)	1 (90分)
5	5 (180分)	5 (29°)	5 (150分)

表 2.5 $U_{10}(10^3)$ 拟水试验方案

列号 / 试验号	1 (A)	5 (B)	7 (C)
1	1 (120分)	5 (29°)	7 (105分)
2	2 (135分)	10 〃	3 (120分)
3	3 (150分)	4 (28°)	10 (150分)
4	4 (165分)	9 〃	6 (90分)
5	5 (180分)	3 (27°)	2 (105分)
6	6 (120分)	8 (27°)	9 (135分)
7	7 (135分)	2 (26°)	5 (150分)
8	8 (150分)	7 〃	1 (90分)
9	9 (165分)	1 (25°)	8 (120分)
10	10 (180分)	6 〃	4 (135分)

表 2.6 $U_{10}(10^3)$ 试验方案

试验号 \ 列号	1 (A)	5 (B)	7 (C)
1	1 (120分)	5 (27°)	7 (132分)
2	2 (127分)	10 (29.5°)	3 (104分)
3	3 (134分)	4 (26.5°)	10 (153分)
4	4 (141分)	9 (29°)	6 (125分)
5	5 (148分)	3 (26°)	2 (97分)
6	6 (155分)	8 (28.5°)	9 (146分)
7	7 (162分)	2 (25.5°)	5 (118分)
8	8 (169分)	7 (28°)	1 (90分)
9	9 (176分)	1 (25°)	8 (135分)
10	10 (183分)	6 (27°5)	4 (111分)

分布得不是更均匀吗？于是利用 $U_{10}(10^{10})$ 的同样三列得试验方案列于表 2.6。一般来说 2.6的方案比表 2.5均匀性要好一点。但是，在工业试验中，有时要增加水平并不那么方便，比如温度按半度来分水平，在工业控制中往往是做不到的。在刚才讨论的例子中，如果烧料温度不能按半度来分水平，这时可以因素A和C分成十个水平，而因素B分五个水平，用拟水平的方法来排，所得的试验方案正好是表 2.6的（A），（C）两列加表 2.5的（B)列。

还需要说明的，原来初定的是五水平，安排试验不一定非五的倍数不可，可改成任意的水平数，一切视试验的具体条件和要求而定。例如例 1的试验也可用 $U_{11}(11^{10})$ 来安排，这样因

素A和C取成6分一个间隔更为方便，也可用$U_{13}(13^{12})$来安排，A和C取成5分一个间隔，后者的试验方案见表2.7。

在用些试验，特别是化学试验中，所有因素的高档水平相迁，反应太剧烈，甚至会爆炸；所有低档水平相近，反应太慢，甚至不起反映而得不到试验结果。而在均匀设计表中，所有奇数试验的表最后一次试验都是所有的高档水平相迁，如表2.7最后一次试验是所有的第十三水平相迁。为了避免这个情况，可将水平次序作适当的调整，例如将因素A的水平作适当调整如下：

120分	125分	130分	135分	140分	145分	150分	155分	160分	165分	170分	175分	180分
7	8	9	10	11	12	13	1	2	3	4	5	6

表2.7 $U_{13}(13^3)$试验方案

列号 试验号	1 (A)	3 (B)	4 (C)
1	1 (120分)	3 (25°)	4 (105分)
2	2 (125分)	6 (26.5°)	8 (125分)
3	3 (130分)	9 (28°)	12 (145分)
4	4 (135分)	12 (29.5°)	3 (100分)
5	5 (140分)	2 (24.5°)	7 (120分)
6	6 (145分)	5 (26°)	11 (140分)
7	7 (150分)	8 (27.5°)	2 (95分)
8	8 (155分)	11 (29°)	6 (115分)
9	9 (160分)	1 (24°)	10 (135分)
10	10 (165分)	4 (25.5°)	1 (90分)
11	11 (170分)	7 (27°)	5 (110分)
12	12 (175分)	10 (28.5°)	9 (130分)
13	13 (180分)	13 (30°)	13 (150分)

根据均匀设计表制作的原理，水平不能象正交试验的水平那样任意改变次序，而只能按照原来的顺序进行平滑，也就是说，将原来的最后一水平与第一个水平接起来组成一个圈，然后从任一处开始定为第一个水平，按圈子的原方向（或相反方向）排第二个水平，第三个水平，……

图 2.3 表示的是刚才叙述的例子，“*”处表示从此处开始第一水平，箭头方向表示水平依次的顺序。

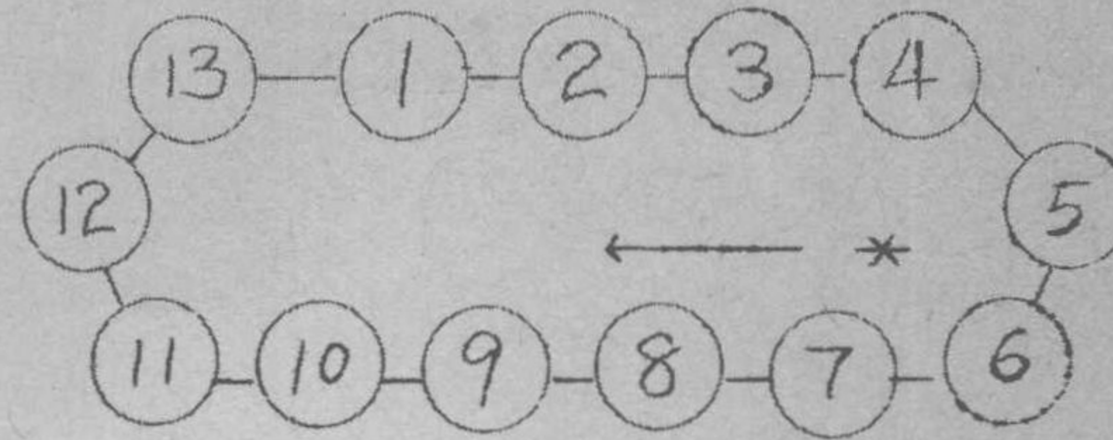


图 2.3

三、回归分析简介

用均匀设计表安排的试验如何分析呢？熟悉正交试验的同志都知道，正交试验的分析是很方便的。由于均匀设计舍弃了正交试验的“整齐可比”性，试验方案是不正交的，因而试验的分析不如正交试验那么便当，一般要用回归分析的工具，为此我们首先将回归分析作一个简单的介绍。

1. 一元线性回归

先看一个例子。

例 2：为了估计山上积雪融化后对下游灌溉的情况，在山上建立了一观测站，测量了最大积雪深度（x）与当年灌溉的面积（y），连续十年的数据如下：

最大积雪深度(x)尺	15.2	10.4	21.2	18.6	26.4	23.7	13.5	16.7	24.0	19.1
灌溉面积(y)千亩	28.6	19.3	40.5	35.6	48.9	45.0	29.2	34.1	46.7	37.4

首先，我们将数据点图，从图上看两者近似为线性关系，这是符合于经验的，如果我们能具体的求出这个线性关系，那么以后只要在冬天测出最大积雪深度，就能予报灌溉的面积，

以便及早采取措施。

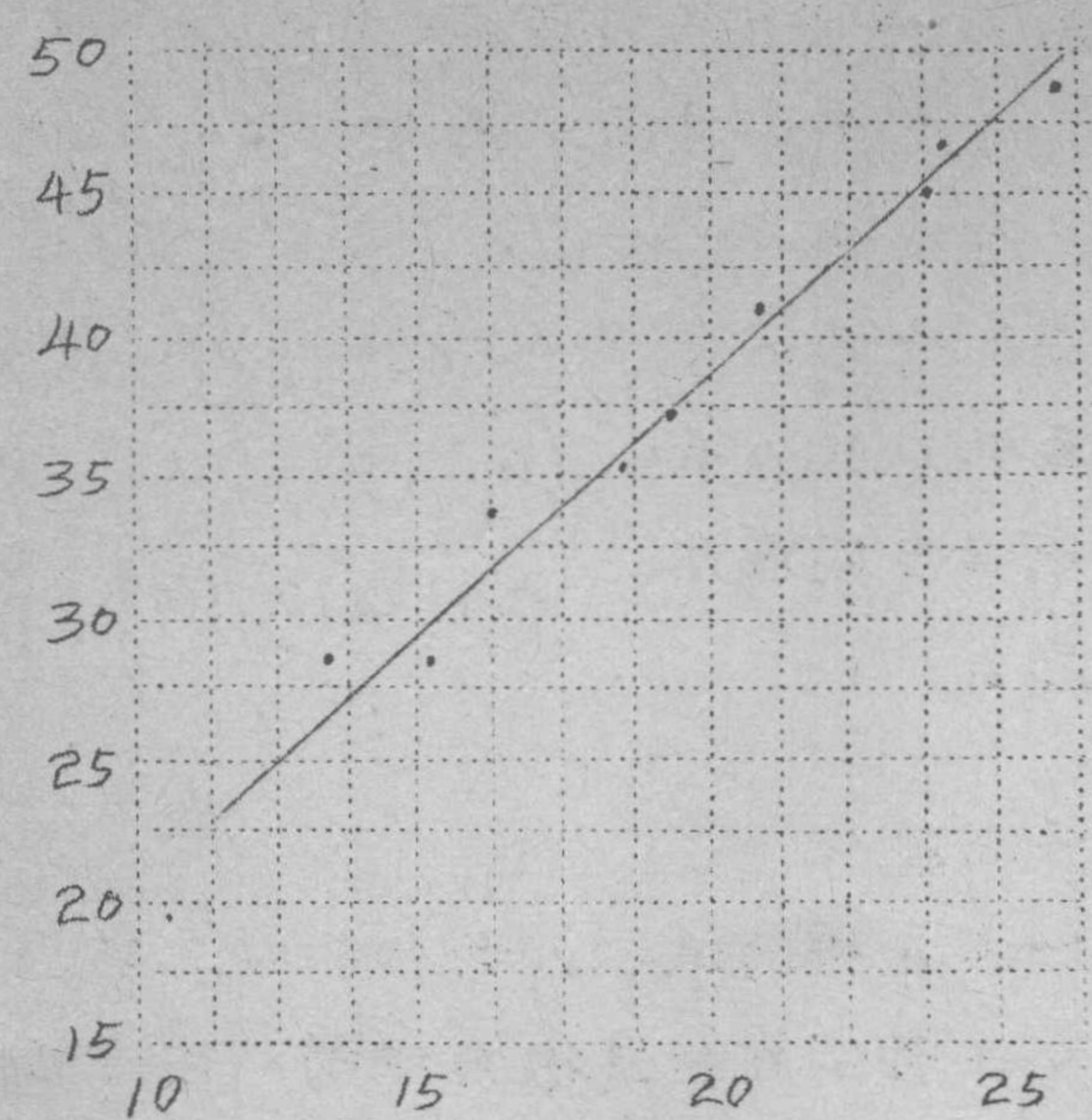


图 3.1

(A) 回归方程的求法

用n表示数据组数，(x_i, y_i) $i=1, 2, \cdots$，n表示测出的n组数据，如果它们有线性关系，应有

$$\hat{y} = a + bx \qquad (3.1)$$

它称为回归方程，此处因是直线可称为回归线，令

$$\hat{y}_i = a + bx_i \qquad i = 1, 2, \cdots n \qquad (3.2)$$

如果是精确的线性关系应有$y_i = \hat{y}_i$ $(i = 1, 2, \cdots n)$，但是由于测量误差以及其它因素的干扰，一般y_i不一定等于$\hat{y}_i$，例如图3.1上的点就不是在一条直线上，而是近似在一条直线上。因此我们要决定一条直线，与所有的点都比较接近。最简单的办法是凭目力用尺子画一条线，但这样会因人而异，