

正 交 设 计

(即正交试验法)

南京航空学院

1975.10.

目 录

第一章 正交设计的基本思想和方法

§ 1. 什么叫正交设计	1
§ 2. 正交表的极差分析	2
§ 3. 效应与工程平均	8
§ 4. 有交互作用的试验	13
§ 5. 多指标试验	16

第二章 正交表的方差分析

§ 6. 变动平方和	20
§ 7. 显著性检验	25
§ 8. 重复试验与重复取样	28

第三章 水平数不同的试验

§ 9. 直接使用混合型正交表	39
§ 10. 并列法	43
§ 11. 拟水平法	45

第四章 回归分析

§ 12. 什么叫回归分析	49
§ 13. 回归直线的求法	50
§ 14. 相关系数与显著性检验	57
§ 15. 利用回归方程进行预测和控制	58
§ 16. 化曲线为直线的回归问题	60

附 录

正交表	65
F 表	78
相关系数检验表	81

四季用 (田)

第一章 正交设计的基本思想和方法

§ 1. 什么叫正交设计?

在生产斗争和科学实验中,为了实现优质、高产、低消耗,必须以革旧工艺,研制新产品,这就需要选择最好的生产条件或者是最好的配方。要想实现这种选择,就离不开试验。但是,首先,如何安排试验?是一个很值得研究的问题,试验安排得好,通过为数不多的几次试验,就可以得到满意的结果。相反,试验安排得不好,试验次数既多,必然浪费大量的人力物力,有时由于时间拖长,试验条件改变,也会使试验失败。其次,如何对试验结果进行由表及里,去伪存真的科学分析?以求找到解决问题的有效办法,而不致于在一大堆试验数据面前束手无策,可见这也是一个很重要的问题。

通常在生产实践和科学实验中,影响产品的因素较多,要考察的指标往往也不止一个,而且有些生产或试验的周期较长,还要考虑到试验误差的影响,在这种错综复杂的情况下,迫切需要采用一种行之有效的试验方法,多因素优选法——正交设计(也叫正交试验法)正是为了适应生产建设的这种客观需要而产生的。

正交设计就是使用正交表来处理多因素试验的一种科学方法,即借助于正交表,在试验前科学地安排试验方案,以求尽量减少试验次数,在试验后经过简单的表格运算,正确地分析试验结果,就能分清哪些是影响试验结果的主要因素,然后找出最好的生产条或配方。

为了说明正交设计解决生产问题的实际效果,请看一些例子:

(1) 在表面处理方面:某厂 19 车间的镀锌任务占全车间任务的 70%。但自建厂以来,一直存在着一个老大难的问题,即镀锌零件的表面起泡(叫做镀锌爆皮)而报废,有时整批零件返修或报废,严重影响产品质量和生产任务的完成,去年应用正交设计仅安排两批试验,就找到了较好的工艺条件,经过两个月的生产考验,消除了成批爆皮的现象,爆皮零件也很少见。

(2) 在机械加工方面:某厂 15 车间的柱塞组合件,过去采用收口模多次收口,生产率低,模具耗费量大。该车间试制成功半自动液压收口机,能对多种类型的油泵柱塞组合件进行一次收口,提高工效 5~6 倍,大大减轻了劳动强度。但是,对某种燃油泵的柱塞组合件,设计要求拉脱力不小于 800 公斤,转角不小于 20°。但经长期试验,拉脱力一直在 700~1000 公斤之间波动,有时虽然拉脱力提高了,但转角又达不到 20°,因此使这项重大的技术革新无法验收投产。去年用正交设计分析了原有数据的规律,重新安排了探索性试验、验证性试验和稳定性试验,找到了较好的压力、滚轮方式和零件尺寸,使这种柱塞组合件的收口完全满足设计要求,而且比较稳定,该厂同志高兴地说:“这项成功的试验打响了 我厂三年技术改造的第一炮”。

(3) 在无损检验方面:某厂三车间用 X 光检验镁铸件的内部质量,有国产的和从苏修进口的两种 X 光探伤机。过去在使用国产的 X 光机时,由于不加区别地沿用了苏修 X 光

机的工艺参数，因此灵敏度不高，X光底片不清晰，铸件缺陷反映模糊，甚至发现不了，成了该厂的关键之一。去年使用正交设计和回归分析，经过47次试验，找出了相应于铸件各种厚度的合适电压与曝光量，使灵敏度由原来的1.5~2%提高到0.68~1.5%，经过两个月的试生产考验，效果良好，证明国产X光机的灵敏度超过了苏修的X光机。该厂主持试验的同志说：“正交设计试验次数少，效果好，是一个多快好省的方法”。由于国产X光机使用普遍，所以这项成果具有广泛推广的意义，去年十一月三机部在长沙召开的无损检验专业会议上获得了好评。

(4) 在钣金工艺方面：某厂的机身和机翼的框、肋骨架零件在拉弯成形加工中的主要问题是如何控制回弹，尽量减少校形的工作量。过去，为了掌握零件回弹的规律，需要甚至几年的时间不断地积累数据和总结经验，去年该厂运用正交设计对YL12一类硬铅型材拉弯进行试验，仅用一天的时间就找出了合理的加工过程和参数，基本上可以达到一次成型（即不需校正回弹），后来该厂又进行了三种不同型号的型材批试生产，效果良好。

(5) 在新产品的试制方面：球状热敏电阻是一种对温度特别敏感的半导体元件，由于它具有体积小、重量轻、结构简单、工作可靠以及电阻的变化量大等优点，因此在现代工业中得到了广泛的应用，但由于这种半导体元件的制造工艺比较复杂，影响性能的因素较多，某厂在试制过程中，发现其电阻值指标很难控制，同一炉生产出来的热敏电阻，其电阻值的差别也很大，合格率只有千分之几，远远不能满足生产的需要，去年经用正交设计安排试验，分析了十个因素的作用后，找出了合适的原料配比和工艺条件，使合格率提高到58%。

以上仅仅是某些工厂在生产和科学实验中应用正交设计所取得的一部分成果，其他的很多单位也在这方面作出了不少成绩，例如东北制药总厂对人工合成黄连素的主要原料儿茶酚，为了降低成本，用正交设计进行了试验，取得了可喜的成果，仅此一项每年就可为国家节约资金165万元。

实践证明，正交设计也像优选法一样，是在短时间内广大工农兵群众能够学得会、用得上，花钱少、收效快的一种好方法，它不需要增加设备、器材、人力、资金，就能取得优质、高产、低消耗的效果，达到多快好省的目的。是在现有条件下挖掘生产潜力的一种手段，并且在推广应用的过程中也促进了技术革新和技术革命的开展，有利于加速我国社会主义建设的步伐。

考虑到在生产斗争和科学实验中，也常出现这样的问题：即在进行了一组试验之后，我们总是希望把试验结果与影响这些结果的因素之间的关系用经验公式找出来，这就是回归分析的问题，因此把它作为一章的内容列在最后。

§ 2. 正交表的极差分析

为了说明正交设计的基本思想和方法，我们通过下面一个实例作为“解剖麻雀”，来了解用正交设计处理多因素试验的全过程，由此就可以看出用它来安排试验和分析结果的科学性。

【例1】某化工厂生产一种产品，收率较低且不稳定，一般在60%—80%之间波动，现问如何通过试验，摸清生产中存在的客观规律，然后找出好的生产条件，以实现稳产高产。

为此，首先分析影响该产品收率的各种因素，认为反应温度的高低，加碱量的多少和催

化剂种类的不同可能是造成收率不稳的较主要的原因。根据以往的生产经验，选了温度的三个状态：80℃，85℃，90℃；加碱量的三个状态：35公斤，48公斤，55公斤；催化剂的三个状态：甲种，乙种，丙种。

科学试验的目的，是为了通过试验来了解各种因素对产品的性能、成本、产量（收率）等的影响。为了今后叙述的方便起见，我们把性能、成本、产量等叫做试验的指标，影响指标的因素（如反应温度等）叫做因子，因子对指标的影响表现为，如果因子所处的状态发生了变化，就会引起指标的数据的变化，这里，因子所处的状态（如反应温度80℃，加碱量48公斤等）叫做水平，所以，这个例子是一个三因子、三水平的问题。通常我们用大写字母A、B、C、D等表示因子，比如用字母A代表反应温度，用A₁、A₂、A₃分别代表温度的三个水平，即

$$A_1=80^\circ, A_2=85^\circ, A_3=90^\circ;$$

类似地，用B代表加碱量，用C代表催化剂，它们各自的三个水平为

$$B_1=35 \text{ 公斤}, \quad B_2=48 \text{ 公斤}, \quad B_3=55 \text{ 公斤};$$

$$C_1: \text{甲种} \quad C_2: \text{乙种} \quad C_3: \text{丙种}$$

通常将所挑选的因子和水平列成因素水平表如下：

因子 水平	A 反应温度(℃)	B 加碱量(公斤)	C 催化剂种类
1	80	35	甲
2	85	48	乙
3	90	55	丙

任何一次试验都是由上述三个因子的一定水平搭配起来的，比如按照温度80°，加碱量55公斤，乙种催化剂的条件作试验，实际上也就是A₁、B₃和C₂的搭配，可以简记为A₁B₃C₂。现在为了找到最好的生产条件以使收率最高，当然可以考虑到上述各个因子的各个水平的全部不同的搭配，一共有 $3 \times 3 \times 3 = 3^3 = 27$ 个，即

- | | | |
|--|---|---|
| (1) A ₁ B ₁ C ₁ | (10) A ₂ B ₁ C ₁ | (19) A ₃ B ₁ C ₁ |
| (2) A ₁ B ₁ C ₂ | (11) A ₂ B ₁ C ₂ | (20) A ₃ B ₁ C ₂ |
| (3) A ₁ B ₁ C ₃ | (12) A ₂ B ₁ C ₃ | (21) A ₃ B ₁ C ₃ |
| (4) A ₁ B ₂ C ₁ | (13) A ₂ B ₂ C ₁ | (22) A ₃ B ₂ C ₁ |
| (5) A ₁ B ₂ C ₂ | (14) A ₂ B ₂ C ₂ | (23) A ₃ B ₂ C ₂ |
| (6) A ₁ B ₂ C ₃ | (15) A ₂ B ₂ C ₃ | (24) A ₃ B ₂ C ₃ |
| (7) A ₁ B ₃ C ₁ | (16) A ₂ B ₃ C ₁ | (25) A ₃ B ₃ C ₁ |
| (8) A ₁ B ₃ C ₂ | (17) A ₂ B ₃ C ₂ | (26) A ₃ B ₃ C ₂ |
| (9) A ₁ B ₃ C ₃ | (18) A ₂ B ₃ C ₃ | (27) A ₃ B ₃ C ₃ |

做完这 27 个试验并比较其结果，就可以找出最好的生产条件。像这样进行全面的试验，对试验项目的内部规律虽然揭示得比较清楚，但是试验次数多，工作量大，时间拖得长，给试验带来很多困难。特别是当一个试验的因子多，水平数也多的时候更是如此，比如 5 个因子，7 个水平的试验，全部作要作 $7 \times 7 \times 7 \times 7 \times 7 = 7^5 = 16807$ 次试验以后，才能找出最好的生产条件，这显然是很难办到的。因此我们设想，能不能从上述 27 个试验中挑选一部分有代表性的搭配作试验，比如说，选作 9 个，还希望它同样能够达到 27 个全面试验的效果，换句话说，从这 9 个试验结果中通过分析计算，希望同样能找出全部 27 个试验中那个最好的生产条件，这就是用正交设计解决生产实际问题的指导思想。

现问：在上述全部 27 个试验中，应该挑选什么样的 9 个试验，才能反映出很强的代表性呢？对此我们可以这样选，比如选下列 9 个试验，并注意各个试验的因子的水平号码构成一张表：

	A	B	C
(1) $A_1B_1C_1$	1	1	1
(5) $A_1B_2C_2$	1	2	2
(9) $A_1B_3C_3$	1	3	3
(11) $A_2B_1C_2$	2	1	2
(15) $A_2B_2C_3$ \longrightarrow	2	2	3
(16) $A_2B_3C_1$	2	3	1
(21) $A_3B_1C_3$	3	1	3
(22) $A_3B_2C_1$	3	2	1
(26) $A_3B_3C_2$	3	3	2

这个表在安排上有如下特点：

1. 从纵的来看，在任何一列中，每个因子的各个不同的水平“1”、“2”、“3”出现的个数相同，都是三个，既不多，也不少，这反映了被挑选出来的因子的水平具有均匀的代表性。

2. 从横的来看，在任何两列中，两个因子之间的各种水平的不同搭配的个数相同，比如在第一列和第二列的两个因子 A 和 B 之间的各种水平的不同搭配(1, 1)、(1, 2)、(1, 3)、(2, 1)、(2, 2)、(2, 3)、(3, 1)、(3, 2)、(3, 3)都有，而且都是一个搭配，既不遗漏，也不重复，这反映了在两个因子之间各种水平的搭配的均衡性。

具有上述这两个特点的试验方案叫做因子之间具有正交性，它表明我们从 27 个试验中挑选出来的这 9 个试验具有很强的代表性，即通过这 9 个试验对结果进行分析计算，就能找出全部 27 个试验中最好的结果，即使这个最好的结果并不在上述 9 个试验之中也行，关于这点我们很快就会看到。

这里也还表明，只要能满足上述任意两个因子之间的正交性，从全部的 27 个试验中任意挑选 9 个都行，但是在实际应用中，为了使挑选的工作简便起见，人们在长期的实践中总结出一套表格印好备用（见后），叫做正交表。

正交设计就是以正交表为工具来安排试验方案和分析试验结果的。为此我们先介绍一下正交表。

正交表有多种不同的规格，如：

$$L_4(2^3), L_8(2^7), L_{12}(2^{11}), L_{16}(2^{15}), \dots,$$

$$L_9(3^4), L_{27}(3^{13}), L_{16}(4^5), L_{25}(5^9), \dots.$$

今以 $L_9(3^4)$ 为例，其中 L 表示一张正交表，如右，用它来安排 3 水平的试验（表中右下方框内即表示水平数），因子的个数最多可以安排 4 个（即表上方列号的位置 1~4），全部作本来需要作 $3 \times 3 \times 3 \times 3 = 81$ 次试验，但用这张正交表来安排只要作 9 次就可以了（即表左边的试验号 1~9）。注意，水平数的排列，明显地反映了上述因子之间的正交性。实践证明，用正交表来安排多因子试验的问题是一个行之有效的办法，特别当因子越多，水平数越多，它的优点也就越明显。例如北京第二棉纺织厂在改进化学浆料 CMC 的产品质量的试验中，涉及到 5 个因子 5 个水平，全部需要作 $5 \times 5 \times 5 \times 5 \times 5 = 5^5 = 3125$ 次试验，但用正交表 $L_{25}(5^9)$ 来安排，只作了 25 次试验，就找出了全部 3125 个试验中最好的结果。北京维尼纶厂在

列号 试验号	1	2	3	4
1	1	1	1	1
2	1	2	2	2
3	1	3	3	3
4	2	1	2	3
5	2	2	3	1
6	2	3	1	2
7	3	1	3	2
8	3	2	1	3
9	3	3	2	1

解决维尼纶缩醛化工艺的试验中，涉及到 5 个因子 7 个水平 全部需要作 $7 \times 7 \times 7 \times 7 \times 7 = 7^5 = 16807$ 次试验，即使要求仅仅把这么多次的试验条件全部写出来，都很难办到，更不要说是把这么多次试验全部做完，但是用正交表 $L_{49}(7^8)$ 来安排，只作了 49 次试验就取得了原来 16807 次试验同样的效果，由此可见正交表的作用和用正交设计解决实际问题的具体效果。

要想用正交设计来解决例 1 的问题，必须选用合适的正交表，由于现在是 3 因子 3 水平的问题，应该到 3 水平的正交表 $L_9(3^4)$ 、 $L_{27}(3^{13})$ 中去找，由于 $L_9(3^4)$ 可以安排下最多 4 个因子，只须作 9 次试验，因此选用 $L_9(3^4)$ 来安排就行了。我们把三个因子 A 、 B 、 C 可以安排在表的上方四个列的任意三个位置上，这一步叫做表头设计。比如说，可以把 A 、 B 、 C 三个因子依次排在第 1、2、3 列上，再把各列的水平号码“翻译”成各因子的具体水平，比如反应温度 A 在第 1 列，就把第 1 列中的“1”水平换成 80° ，“2”水平换成 85° ，等等。第 4 列没有用上，可以不管它，或者将它抹去，这一步叫做列出试验方案，如下表。

这样一来，正交表就变成了试验方案表，即从横的看，表中的每一行都是一个试验操作条件，例如第 1 号试验是温度 80° ，加碱量 35 公斤，用甲种催化剂（即 $A_1B_1C_1$ ），第 4 号试验是温度 85° ，加碱量 35 公斤，用乙种催化剂（即 $A_2B_1C_2$ ），等等。稍加注意就会发现，当我们像这样把 A 、 B 、 C 三个因子依次安排在正交表 $L_9(3^4)$ 的第 1、2、3 列以后，由此得到的 9 次试验正好就是在前面提到的在全部 27 次试验中我们挑选的那 9 个试验，因为各个因子的水平数的排列完全一样。

列号 水平	A		B		C		4
	反应温度(°C)	1	加碱量(公斤)	2	催化 剂 种 类	3	
1	1	(80)	1	(35)	1	(甲)	1
2	1	(80)	2	(48)	2	(乙)	2
3	1	(80)	3	(55)	3	(丙)	3
4	2	(85)	1	(35)	2	(乙)	3
5	2	(85)	2	(48)	3	(丙)	1
6	2	(85)	3	(55)	1	(甲)	2
7	3	(90)	1	(35)	3	(丙)	2
8	3	(90)	2	(48)	1	(甲)	3
9	3	(90)	3	(55)	2	(乙)	1

以上我们完成了安排试验方案的工作，这是用正交设计解决生产实际问题的第一个阶段。

第二个阶段的工作要求按照已经安排好的方案作试验，并将9次试验的结果填在试验数据栏的相应位置上，以供下面分析试验结果之用。

第三个阶段是分析试验结果。为此首先观察，在这9个试验中，第8号试验的收率最高，达到85%，它的试验条件是 $A_3B_2C_1$ ，现问： A_3 、 B_2 、 C_1 是否就是A、B、C的最好水平的搭配？还有没有比这更好的生产条件能使收率更高呢？因为在全部27个试验中毕竟我们只选作了9个，在另外还没有作的18个试验中是否还有比这更好的结果？其次，第8号试验的收率高是偶然现象还是客观规律的反应？到现在我们都还不知道，这就需要对现有的9个试验结果进行科学的分析，然后才能得出正确的结论。

为此，仔细观察这9个试验结果即可看出，温度取90°一定比取80°，85°好，因为温度为90°的三个试验的收率平均来说，比80°的三个试验以及85°的三个试验的收率都要高。这启发我们，只要比较因子在各个水平下的平均收率，就能够找出该因子的使收率最高的水平来。

先看温度A，在 A_1 条件下的试验之和，反映在正交表上的就是在第1列中“1”水平所对应的试验数据之和，记为 K_1 ，即：

$$K_1(\text{第1列}) = 51 + 71 + 58 = 180(\%),$$

把180填在第1列下面的“ K_1 ”行上。又由于在第1列中的“1”水平重复了三次，即 K_1 (第1列)是三个试验结果之和，所以它的平均值(记为 k_1)是

$$k_1(\text{第1列}) = \frac{K_1(\text{第1列})}{\text{在第1列中“1”水平的重复次数}} = \frac{180}{3} = 60(\%).$$

这就是说，除了因子A都是“1”水平以外，其余的因子各个水平出现的次数都相同，于是这些因子水平变化的影响在取平均值后就自行抵销了，因此在温度为80°的条件下，三个试

列号 试验号	A 反应温度(°C)		B 加碱量(公斤)		C 催化剂种类		4	试验数据 收率(%)
	1		2		3			
1	1	(80)	1	(35)	1	(甲)	1	51
2	1	(80)	2	(48)	2	(乙)	2	71
3	1	(80)	3	(55)	3	(丙)	3	58
4	2	(85)	1	(35)	2	(乙)	3	82
5	2	(85)	2	(48)	3	(丙)	1	69
6	2	(85)	3	(55)	1	(甲)	2	59
7	3	(90)	1	(35)	3	(丙)	2	77
8	3	(90)	2	(48)	1	(甲)	3	85
9	3	(90)	3	(55)	2	(乙)	1	84
K_1		180		210		195		
K_2		210		225		237		
K_3		246		201		204		
k_1		60		70		65		
k_2		70		75		79		
k_3		82		67		68		
R		22		8		14		

验的平均收率是 60%，把 60 填在第 1 列下面的“ k_1 ”行上。同样对 A_2 、 A_3 进行计算，得：

$$\begin{aligned}
 K_2(\text{第 1 列}) &= \text{在 } A_2 \text{ 条件下的试验收率之和} \\
 &= \text{在第 1 列中“2”水平对应的试验数据之和} \\
 &= 82 + 69 + 59 = 210(\%),
 \end{aligned}$$

$$\begin{aligned}
 k_2(\text{第 1 列}) &= \frac{K_2(\text{第 1 列})}{\text{在第 1 列中“2”水平的重复次数}} = \frac{210}{3} = 70(\%) \\
 &= \text{在 } 85^\circ \text{ 条件下的平均收率;}
 \end{aligned}$$

$$\begin{aligned}
 K_3(\text{第 1 列}) &= \text{在 } A_3 \text{ 条件下的试验收率之和} \\
 &= \text{在第 1 列中“3”水平对应的试验数据之和} \\
 &= 77 + 85 + 84 = 246(\%),
 \end{aligned}$$

$$\begin{aligned}
 k_3(\text{第 1 列}) &= \frac{K_3(\text{第 1 列})}{\text{在第 1 列中“3”水平的重复次数}} = \frac{246}{3} = 82(\%) \\
 &= \text{在 } 90^\circ \text{ 条件下的平均收率。}
 \end{aligned}$$

然后再把 210、246 和 70、82 分别填在第 1 列下面的“ K_2 ”、“ K_3 ”和“ k_2 ”、“ k_3 ”行上。

对于第 2 列、第 3 列的计算法则完全相同，不再重复，建议学员自己进行验算。

根据以上的计算和记录在表上的结果可以看出，温度为 90° 时收率最高，加碱量以 48 公斤为宜，催化剂最好用乙种，把这三个因素的最好水平搭配起来，就得到一个好的生产条件 $A_3B_2C_2$ 。

为了进一步搞清问题，我们再问：在这三个因素中哪个是主要的，哪个是次要的呢？从上面的计算结果可以看到，温度在 $80^\circ \sim 90^\circ$ 之间变化时，收率在 $60\% \sim 82\%$ 的范围内变化，其变化的幅度是 $82 - 60 = 22(\%)$ ，我们把它叫做因素 A 的极差，或叫第 1 列的极差，记作 R ，即

$$R(\text{第 1 列}) = \text{在第 1 列的 } k_1, k_2, k_3 \text{ 中最大的减去最小的} \\ = 82 - 60 = 22(\%)$$

仿此也可以计算出第 2 列、第 3 列的极差，然后把它们填在表的最下面的“ R ”行上。极差的大小反应了某个因素的水平变化时影响到试验结果的变化幅度的大小，所以因素的极差越大，就说明该因素对收率的影响越大，它就越重要。因此，根据表中极差的大小可以看出，温度对收率的影响最大，是主要因素，催化剂次之，加碱量更次。即它们的主次顺序是：

主 \longrightarrow 次
A C B

由于主要因素的水平变化对指标的影响较大，所以必须控制它在最好的水平上，例如温度的变化为 $80^\circ \nearrow 85^\circ \nearrow 90^\circ$ 时，收率的变化是 $60\% \nearrow 70\% \nearrow 82\%$ ，相差较大，所以必须取温度为 90° 。催化剂种类是第二重要的因素，而且用乙种明显地比用甲种和丙种的收率要高，所以也应取乙种，至于加碱量则是次要因素，用 35 公斤时的平均收率为 70% ，用 48 公斤时的平均收率为 75% ，相比之下，收率提高不多，但加碱量却增加了 13 公斤。因此，考虑到既要保证收率又要兼顾节约用碱量起见，也可以取加碱量为 35 公斤，这说明 $A_3B_1C_2$ 也可能是较好的的水平搭配。

由此可见，正交设计在数据分析上的基本思想就是，利用在试验安排上均衡搭配的特点，把各个因素的影响从错综复杂的总的的影响中分别单独地分离出来，据此就可以看出各因素对试验结果的影响的规律性，然后抓住主要矛盾并选出较好的生产条件。

§ 3. 效应与工程平均

1. 效应

根据以上的讨论，虽然对收率试验的结果分析已经基本完成，但是试验的目的还没有完全达到，因为在 9 个试验中最好的生产条件是 $A_3B_2C_1$ ，而通过分析计算找出的好的生产条件却是 $A_3B_2C_2$ 或 $A_3B_1C_2$ 。正由于后面这两个生产条件并不在那已经作过的 9 个试验之中，因此它们的效果到底是好是坏？我们还不清楚。现在如问：在 $A_3B_2C_1$ 、 $A_3B_2C_2$ 和 $A_3B_1C_2$ 这三个搭配中究竟哪一个更好些？当然也就无法判断了。针对这种情况，客观上要求我们去作

进一步的分析计算，然后就可以对上述问题得到明确的结论。

为此我们先求出上面 9 个试验数据的总和，记如 T ，即：

$$T = 51 + 71 + 58 + 82 + 69 + 59 + 77 + 85 + 84 \\ = 636.$$

再求出这 9 个数据的平均值，叫做总平均，记如 \bar{y} ：

$$\bar{y} = \frac{T}{9} = \frac{636}{9} = 70.7(\%),$$

这个总平均 $\bar{y} = 70.7\%$ 意味着，设想如果我们把 A 、 B 、 C 中的每一个因子都既不取“1”水平，又不取“2”水平，也不取“3”水平，而是都取“中等”水平时将会得到的一个平均结果，这是因为在这 9 个试验中，全部因子都是“1”水平、“2”水平、“3”水平各取三次的原故。前面已经看到，当温度取 80° 的条件下作试验，平均收率为 k_1 (第 1 列) = 60%，今由 k_1 (第 1 列) 减去总平均 \bar{y} ，记为 ω_1 (第 1 列)，即

$$\omega_1(\text{第 1 列}) = k_1(\text{第 1 列}) - \bar{y} = 60 - 70.7 = -10.7(\%),$$

结果为负值，它表示由于因子 A 取“1”水平比取“中等”水平而引起收率降低了 10.7%，叫做因子 A 的“1”水平的效应。其次计算

$$\omega_2(\text{第 1 列}) = k_2(\text{第 1 列}) - \bar{y} = 70 - 70.7 = -0.7(\%),$$

$$\omega_3(\text{第 1 列}) = k_3(\text{第 1 列}) - \bar{y} = 82 - 70.7 = 11.3(\%),$$

其中 ω_3 (第 1 列) 为正值，它表示由于因子 A 取“3”水平比取“中等”水平而引起收率提高了 11.3%，这就是因子 A 的“3”水平的效应，仿此可以计算出其他因子的各个水平的效应值。我们也可以把这一部分计算结果列入表内(见下表)，这样，从表格上就能够更清楚地看出，当取某个因子的某个水平时，它对收率的影响如何？即和总平均来比较，看效应的正负号就知道它使收率提高了或者是降低了，看效应的绝对值大小就知道它使收率提高或者降低到底有多少。

2. 工程平均

当我们通过以上的分析确定了某个较好的生产条件以后，自然希望在事先能从数量上来估计一下，如果按照这个条件来进行长期而稳定的生产的话，收率有希望达到什么样的程度？这叫做在这个生产条件之下的工程平均，由于每一个效应都表示一个因子在取定某个水平后使数据比总平均增加了多少或者是减少了多少，所以我们可以把影响较大的因子的相应水平的效应和总平均统统加起来，用它来预测在某个试验条件下的指标的数值是很自然的。即

$$[\text{某个试验条件下的工程平均}] = [\text{总平均}] + [\text{主要因子在该条件下出现水平的效应}].$$

这里需要说明的是，任何试验条件不论控制得多么严格，观察或者测量试验结果多么认真，都免不了有误差，并且它最后必然要影响到试验数据，因此在计算工程平均时，由于那些次要的因子对试验结果的影响本来就很小，再考虑到误差的存在，可见它们的那一点点效应值是“靠不住的”，所以不应该加到工程平均里去。其次，我们算出的工程平均只是在某个试验条件下，根据现有的试验结果(它也可能有误差)算出的指标的理论值，也正是由于有误差的存在，所以如果我们按照这样同一个条件去作几个试验，就会发现得到的几个试验数据往往是在工程平均值的附近散布，而且和工程平均值差得越小的可能性越大，差得越大

列号 试验号	A		B		C		试验数据 收率(%)	
	反应温度(°C)	1	加碱量(公斤)	2	催化 剂 种 类	3		4
1	1	(80)	1	(35)	1	(甲)	1	51
2	1	(80)	2	(48)	2	(乙)	2	71
3	1	(80)	3	(55)	3	(丙)	3	58
4	2	(85)	1	(35)	2	(乙)	3	82
5	2	(85)	2	(48)	3	(丙)	1	69
6	2	(85)	3	(55)	1	(甲)	2	59
7	3	(90)	1	(35)	3	(丙)	2	77
8	3	(90)	2	(48)	1	(甲)	3	85
9	3	(90)	3	(55)	2	(乙)	1	84
K_1	180		210		195			
K_2	210		225		237			
K_3	246		201		204			
k_1	60		70		65			
k_2	70		75		79			
k_3	82		67		68			
ω_1	-10.7		-0.7		-5.7			
ω_2	-0.7		4.3		8.3			
ω_3	11.3		-3.7		-2.7			
R	22		8		14			

的可能性越小。因此工程平均的作用就在于帮助我们事先就能够估计到，如果要按照某个条件去作试验，那么试验的结果“大约”是多少；看看大约能够达到多高的指标。

在搞清了效应和工程平均的概念以后，回头来计算上面三个好的生产条件的工程平均，就能看出应该采用哪个条件更好些。先求 $A_3B_2C_1$ 的工程平均，记如 $\mu_{A_3B_2C_1}$ ，由于 B 不是主要因子，所以只应该取 A_3 、 C_1 的效应值和总平均相加，即

$$\mu_{A_3B_2C_1} = 70.7 + 11.3 - 5.7 = 76.3(\%);$$

同理可得

$$\mu_{A_3B_2C_2} = 70.7 + 11.3 + 8.3 = 90.3(\%);$$

$$\mu_{A_3B_1C_2} = 70.7 + 11.3 + 8.3 = 90.3(\%).$$

可见如果按照条件 $A_3B_2C_1$ 进行生产，收率只能达到大约 76.3%，这就说明第 8 号试验达到 85% 只是偶然的现象，而不是客观的规律。但是如果要按照条件 $A_3B_2C_2$ 或 $A_3B_1C_2$ 进行生产，则可使收率提高到 90.3% 左右，又由于条件 $A_3B_1C_2$ 既考虑到提高产量又兼顾了降低成本，所以最好是选取生产条件 $A_3B_1C_2$ 。

以上通过计算得到的结论是否合乎客观实际？还需要在实践中进行检验，这就是第四阶段的工作，叫做验证试验，即按照条件 $A_3B_2C_1$ 、 $A_3B_2C_2$ 和 $A_3B_1C_2$ 作试验，连同第 8 号在内，结果如下：

试验号	试验条件	收率 (%)
⑧, ⑩	$A_3B_2C_1$	85, 82
⑪, ⑫	$A_3B_2C_2$	95, 90
⑬, ⑭	$A_3B_1C_2$	93, 89

可见实际验证的结果和上面通过理论分析的结论是一致的，上述充分的科学根据促使我们最后拿定主意，确定以 $A_3B_1C_2$ 为正式的生产条件。

到此整个试验的过程才算全部完成了，当然旧的矛盾解决了，还有新的矛盾产生。比如注意观察温度的变化 $80^\circ \nearrow 85^\circ \nearrow 90^\circ$ ，相应的平均收率为 $60\% \nearrow 70\% \nearrow 82\%$ ，即温度越高，那么收率也越高，而温度又是影响收率最主要的因素，这就启发我们可以把其他因素全部固定下来，而就温度进行单因素优选，以便找出能够使收率进一步提高的合适的温度，但这毕竟是下次试验的任务了。

3. 小 结

归纳以上讨论，可知用正交设计处理多因素试验的一般步骤如下：

第一阶段 安排试验方案

第一步 明确试验目的，确定试验指标。

有时生产中存在的问题较多，在试验之前，应根据实际情况，确定这次试验主要解决什么问题，例如可能是提高质量的问题，提高产量的问题，降低成本问题或者是改善劳动条件问题，等等。

第二步 挑因素选水平。

试验指标确定以后，就可以着手分析影响指标的各种因素，如果对事物的变化规律了解不多，因素可以多取一些，如果对其规律已有相当了解，因素可以少取一些。对于那些对指标影响不大，或者已经掌握得较好的因素，可以把它们固定在适当的水平上，挑选那些对指标可能影响较大，但又没有把握的因素来进行考察。特别注意不能把重要因素固定，否则，由于把重要因素固定在不适当的水平上，将使得试验得不到应有的效果。对于选出的因素，根据生产经验或专业知识，定出它们的变化范围，在此范围内选出每个因素的水平，水平的间隔要适当。因素和水平选定后，列成因素水平表。

以上两步不是数学方法所能解决的，需要由试验人员按照实际情况，根据经验和专业知识确定。特别是因素和水平的挑选，是安排试验方案的重要环节，直接关系到试验的效果，

应该由直接从事这项工作的人员认真商定。

第三步 选用正交表。

如果是二水平的试验，就在 $L_4(2^3)$ 、 $L_8(2^7)$ 、 $L_{16}(2^{15})$ 等等中去选。比如因子的个数大于 3 个但不超过 7 个，就应该选用 $L_8(2^7)$ 。如果是三水平的试验，就在 $L_9(3^4)$ 、 $L_{27}(3^{13})$ 等等中去选。总之，根据试验要求，在保证因子和水平都能排得下的前提下，尽量选用较小的正交表，以便减少试验的工作量，尽快得到试验的结果。特别对于那些试验经费比较昂贵，或者试验周期比较长的情况，更应该注意选用试验次数尽可能少的正交表。在实际安排试验时，在挑选因子和水平的同时，就必然要考虑到选用什么正交表，因此这项工作往往是结合进行的。

第四步 表头设计。

就是把各个因素分别放在选用的正交表表头的适当的列上，在没有交互作用时，这一步很简单，怎么放都行。而在有交互作用的情况下，则是需要认真对待的，这将在 §4 中讨论有交互作用的试验时说明。

第五步 列出试验方案。

即对于表头上有因子的每一个列，把列中的数字，依次地都换成该因子的具体水平（在比较熟悉以后，这一步可以省去，但在思想上必须明确某个因子的某个水平具体表示什么），这样就把原来的正交表改换成了试验方案表。

第二阶段 按照安排好的方案作试验

这一阶段工作的目的是为了取得试验数据，以便为下一个阶段的分析计算提供依据。要想使试验的最后结果明确可靠，就必须尽量排除其它因素的变化对试验数据的干扰，并严格地控制各个因子在各个试验号所要求的水平条件下进行试验。试验的顺序可以按试验号逐个做，也可以根据操作方便跳着做。如果试验的目的只是为了寻找好的生产条件，也可以挑选那予计效果好的试验先做，找到了满意的生产条就不再往下做。

第三阶段 正确分析试验结果

第一步 计算各列的 K 、 k 和极差 R 。

计算公式是

$$K_i(\text{第 } j \text{ 列}) = \text{第 } j \text{ 列中 "i" 水平对应的试验数据之和} \quad (1.1)$$

$$k_i(\text{第 } j \text{ 列}) = \frac{K_i(\text{第 } j \text{ 列})}{\text{第 } j \text{ 列中 "i" 水平的重复次数}} \quad (1.2)$$

$$R(\text{第 } j \text{ 列}) = \text{第 } j \text{ 列中的 } k_1, k_2, \dots \text{中最大的减去最小的} \quad (1.3)$$

第二步 比较各因素的极差 R 的大小，排出因素的主次顺序（ R 越大的因素越重要）。

第三步 选取最好的水平搭配。

对主要因素的水平要仔细选取，应根据指标的要求（是越大越好还是越小越好），比较同一列中 k_1, k_2, \dots 的值来选取好的水平，对次要因素的水平可以随便选取，比如既可以按照上述同样的方法来选取，也可以从便于操作或节约原料的观点来选取水平。

如果需要我们z从理论上估计一下，按照以上选出的最好的水平搭配进行生产的话，能够使指标达到大约多么高的程度？这就需要计算各个因子水平的效应和最好的水平搭配的工程平均（简记为 $\mu_{\text{优}}$ ）：计算公式为

$$\omega_i(\text{第 } j \text{ 列}) = k_i(\text{第 } j \text{ 列}) - [\text{总平均}] \quad (1.4)$$

$$\mu_{\text{优}} = [\text{总平均}] + [\text{主要因子在该条件下出现水平的效应}] \quad (1.5)$$

第四阶段 验证试验

一个正确的认识往往需要从实践到认识、再实践，再认识，经过多次的反复才能得到。上述通过分析计算得到的最好条件可能不止一个，有的还可能是在试验中没有作过的，至于这些条件是不是真正好？还有待于在实践中来进一步考验，最后就可以确定应该选用哪个条件为最好。

在上述的整个过程中，正交设计主要是用在第一阶段和第三阶段，即用正交表来安排试验方案和分析试验结果，这样既可以尽量减少试验次数，只作一部分试验，又能找出全部试验的最好结果。

§ 4. 有交互作用的试验

在多因素试验中，某些因子对指标的影响往往有互相制约，互相联系的现象，这就需要分别研究各个因子的本身对指标的影响，而且需要考察这些因子的各个水平之间的搭配对指标也有关系，这就是说，因素之间会“联合”起来对指标产生作用。这种“联合”作用叫做交互作用。

例如生产队为了提高大豆的产量，在四块条件相同的试验田上，用不同的方式进行施肥试验，结果如下：

试验田号	施 肥 方 式	平均亩产(斤)	增产(斤)
1	不 施	400	
2	6 斤 氮 肥	430	30
3	4 斤 磷 肥	450	50
4	6 斤氮肥、4 斤磷肥	560	160

在这里值得注意的是第 4 块试验田上的情况，增产 160 斤是施两种肥料的结果，其中 30 斤是施氮肥后增产的，50 斤是施磷肥后增产的，而 $160 - 30 - 50 = 80$ (斤)，这里的 80 斤是什么原因促使增产的呢？这就是氮肥和磷肥搭配在一块以后所起的作用的结果，也就是交互作用的影响，这里的 80 斤表示了交互作用的大小。

在相当多的试验中，因素之间是有交互作用的，在安排正交试验时，应尽量根据以往的实践经验，结合上面介绍的方法，对初步选定的各因子之间有没有交互作用和交互作用的大小有一个估计，对交互作用小的可以忽略不管，而仅对较大的交互作用在安排试验时加以考虑。有交互作用的试验在步骤和计算等方面和例 1 基本相同，但毕竟由于出现了新的情况——交互作用，因此在选用正交表、安排试验和分析最好的方案时都带来一些新的问题，应很好注意。

【例 2】 电镀前金属零件要去油去锈，原来的工艺是去油去锈分别进行，现探索去油去

锈一步法，以求达到精简工序，节约时间的目的。

指标：去油去锈时间（越少越好）。

因子与水平的选取如下：

因子 水平	A	B	C	D
	硫酸 (毫升/升)	OP乳化剂(毫升/升)	硫 脲 (克)	温 度 (°C)
1	250	9	6	60
2	300	12	4	65

根据经验判定， D 和其他三个因素没有交互作用，而在 A 、 B 、 C 之间可能有交互作用。通常我们把 A 、 B 之间的交互作用记作 $A \times B$ ，在正交设计中，把交互作用看作是一个因子来安排试验和分析结果。在这里选用正交表时就要把 $A \times B$ 、 $A \times C$ 、 $B \times C$ 看成是三个因子，所以现在的问题应该看成是7个因子2个水平的试验，可选用正交表 $L_8(2^7)$ 。

前已指出，表头设计的过程，就是在正交表的表头上填写因子，在没有交互作用的情况下，因子是可以随便填在哪些列上的，只要不在同一个列上填上几个因子就行了。但是在有交互作用的情况下，就不能再这样随便填写了，因为当我们把两个因子填入正交表的某两个列以后，这两个列的交互作用列也就固定下来了。为了在实际应用上的方便起见，有一些正交表附有一个两列之间的交互作用表。例如 $L_8(2^7)$ 的交互作用表如下：

列号 列号	1	2	3	4	5	6	7
(1)		3	2	5	4	7	6
(2)			1	6	7	4	5
(3)				7	6	5	4
(4)					1	2	3
(5)						3	2
(6)							1
(7)							

利用交互作用表就可以找出正交表中任意两列的交互作用列。比如我们就例2进行表头设计时，可以先把因子 A 、 B 分别填在第1、2列上，要确定第1列和第2列的交互作用 $A \times B$ 在第几列，可以先在对角线上查出列号(1)和(2)，然后从(1)向右横看和从(2)向上竖看的交叉处是数字3，这就是说，第(1)列和第(2)列的交互作用列是第3列，即 $A \times B$ 对指标的影响大小，可由第3列计算出来。然后把因子 C 填在第4列，同样可以确定 $A \times C$ 在第5列， $B \times C$ 在第6列，剩下第7列填上因子 D 。即在需要考虑交互作用的试验中，必

须保证交互作用所占的列，不再排别的因子，以免混杂，这是表头设计的主要任务。

这里我们看到，在二水平的情况下，两个因子的交互作用只占有一个列，至于在三水平的情况下，两个因子的交互作用则占有两个列。由于交互作用不是具体的因素，而是因素之间的联合搭配作用，当然也就无所谓水平，因此我们在列出试验方案时，只需把各个因子所在的列的数字换作具体水平，而交互作用所在的列在试验方案中是不起作用的，它只是在分析试验结果时用到。

根据以上安排好的方案作试验，把试验结果（去油去锈时间）填入试验数据一栏，以下对于交互作用列的分析计算和普通列完全一样。

列号 试验号	A 硫酸(毫升/升) 1	B OP乳化剂(毫升/升) 2	A×B 3	C 硫脲(克) 4	A×C 5	B×C 6	D 温度(°C) 7	试验数据 去油去锈 时间(分)
1	1 (250)	1 (9)	1	1 (6)	1	1	1(60)	7.7
2	1 (250)	1 (9)	1	2 (4)	2	2	2(65)	6.1
3	1 (250)	2 (12)	2	1 (6)	1	2	2(65)	6.0
4	1 (250)	2 (12)	2	2 (4)	2	1	1(60)	17.7
5	2 (300)	1 (9)	2	1 (6)	2	1	2(65)	17.3
6	2 (300)	1 (9)	2	2 (4)	1	2	1(60)	10.5
7	2 (300)	2 (12)	1	1 (6)	2	2	1(60)	13.3
8	2 (300)	2 (12)	1	2 (4)	1	1	2(65)	16.2
K_1	37.5	41.6	43.3	44.3	40.4	58.9	49.2	
K_2	57.3	53.2	51.5	50.5	54.4	35.9	45.6	
k_1	9.375	10.4	10.825	11.075	10.1	14.725	12.3	
k_2	14.325	13.3	12.875	12.625	13.6	8.975	11.4	
R	4.95	2.9	2.05	1.55	3.5	5.75	0.9	

根据极差 R 的大小，可知各个因素和交互作用对指标影响的主次关系如下：

主 ————— → 次

$B \times C, A, A \times C, B, A \times B, C, D,$

由于 $B \times C$ 是影响指标的主要矛盾，其次是 A ，所以它们是选取好的生产条件的主要依据，再其次， $A \times C$ 也有影响，可作参考。

首先，考察交互作用 $B \times C$ ，要问在因子 B 和 C 的水平的全部搭配 $B_1 \times C_1$ 、 $B_1 \times C_2$ 、 $B_2 \times C_1$ 、 $B_2 \times C_2$ 中究竟哪一个的效果最好？结合上面的表格可以这样来计算：相应于 $B_1 \times C_1$ 的试验有第 1 号和第 5 号两次试验，数据的和为 $7.7 + 17.3 = 25$ ，因此 $B_1 \times C_1$ 的效果平