正交实验設计

郑州大学数学系数学专业

1976.2.

正交试验设计

试验设计是安排试验和分析所得数据的一种方法。

生产斗争和科学实验中经常要做许多试验。试验设计得好，就试验次数不多，能得到满意的结果；设计得不好，实验次数既多，结果还不一定满意。特别是在试验中，被考察的因素较多，又要考虑多个指标时，如何合理地设计试验。用较少的试验次数，对试验结果的数据进行分析，从达到预期的目的，是一个值得研究的问题。试验设计的理论和方法就是在解决这个问题的过程中，形成和发展起来的。目前已形成为数理统计学的一个重要分支。它在工农业生产和科学研究等方面已得到广泛的应用。

试验设计考虑两个问题：一、如何安排试验；（试验的设计），二、对试验的结果——数据——进行分析（试验的分析）。

对试验的结果进行正确的分析，要排除误差的干扰。有经验的同志都知道，试验的结果往往不能完全反映客观实际，它含有误差。即使在同一条件下，重复做若干次试验所得结果也不完全相同。因此，分析数据时，要估计出误差的影响，因素的影响，判断出哪些因素的影响是主要的，哪些因素的影响是次要的或不重要的。它们就是后面的方差分析和显著性检验的内容。

这里，我们介绍利用正交表（事先已制好）来设计多因素的试验。对试验结果——数据，仍利用正交表来计算，分清因素的主次。

下面我们先说明怎样用正交表安排试验，怎样直观分析（它只须经过简单的表格化的计算）。然后，解释它的理论根据，阐明统计分析的思想。

正交试验设计

试验设计是安排试验和分析所得数据的一种方法。

生产斗争和科学实验中经常要做许多试验。试验设计得好，就试验次数不多，能得到满意的结果；设计得不好，实验次数既多，结果还不一定满意。特别是在试验中，被考察的因素较多，又要考虑多个指标时，如何合理地设计试验。用较少的试验次数，对试验结果的数据进行分析，以达到预期的目的，是一个值得研究的问题。试验设计的理论和方法就是在解决这个问题的过程中，形成和发展起来的。目前已形成为数理统计学的一个重要分支。它在工农业生产和科学研究等方面已得到广泛的应用。

试验设计考虑两个问题：一、如何安排试验；（试验的设计），二、对试验的结果——数据——进行分析（试验的分析）。

对试验的结果进行正确的分析，要排除误差的干扰。有经验的同志都知道，试验的结果往往不能完全反映客观实际，它含有误差。即使在同一条件下，重复做若干次试验所得结果也不完全相同。因此，分析数据时，要估计出误差的影响，因素的影响，判断出哪些因素的影响是主要的，哪些因素的影响是次要的或不重要的。它们就是后面的方差分析和显著性检验的内容。

这里，我们介绍利用正交表（事先已制好）来设计多因素的试验。对试验结果——数据，仍利用正交表来计算，分清因素的主次。

下面我们先说明怎样用正交表安排试验，怎样直观分析（它只须经过简单的表格化的计算）。然后，解释它的理论根据，阐明统计分析的思想。

一.试验的设计与分析方法介绍

§1 正交表介绍

附录中列举了常用的正交表，例如 $L_4(2^3)$，$L_8(2^7)$，$L_9(3^4)$，$L_{16}(4^5)$，$L_8(4^1\times2^3)$ 等等。今从 $L_8(2^7)$ 说明符号的意义：

右表是 $L_8(2^7)$，"L"表示正交表，"L"右下角数字"8"表示这个正交表有8行，用该表来设计安排试验时，要做8个试验。括号内的数字"2"表示被考查的因素都取两个不同的位级。"

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

"2"的右上角数字"7"表示这个表有7列，用这个表最多可以同时考查7个不同因素。总之，

$$L_8(2^7) \longrightarrow L_{\text{试验次数}}(\text{位级数}^{\text{列数}})$$

正交表的特点是：

(1)．每一列中不同数字出现的个数相同。$L_8(2^7)$ 中，每列中出现数字"1"、"2"各四次。

(2)．任意两列中，不同数字组合出现的次数相同，$L_8(2^7)$ 中，任意两列中出现的不同数字组合(1,1)、(1,2)，(2,1)、(2,2)各二次，

所有的正交表都具有这样性质．

设有三因素，三位级的试验，要全面试验时，必须把因素A，B，C的各个位级 A_1，A_2，A_3；B_1，B_2，B_3；C_1，C_2，C_3 的所有可能27种组合，$A_1B_1C_1$，$A_1B_1C_2$，$A_1B_1C_3$，$A_1B_2C_1$，$A_1B_2C_2$，…，

$A_3B_3C_3$，全部试验过．从中挑出好的试验方案来。当然试验工作量是大的，如果有六个因素，每个因素是五个位级时，全面试验则需 $5^6=15625$ 次，是不可能全部去做的。利用正交试验设计来安排时，只需作25次试验，为全面试验次数的 $\frac{1}{625}$。

传统的简单比较法，是将其它因素固定，变化一个因素，例如固定B，C于B_1，C_1，让A变化

$$B_1C_1 \begin{cases} A_1 \\ A_2 \\ \boxed{A_3} \end{cases}$$

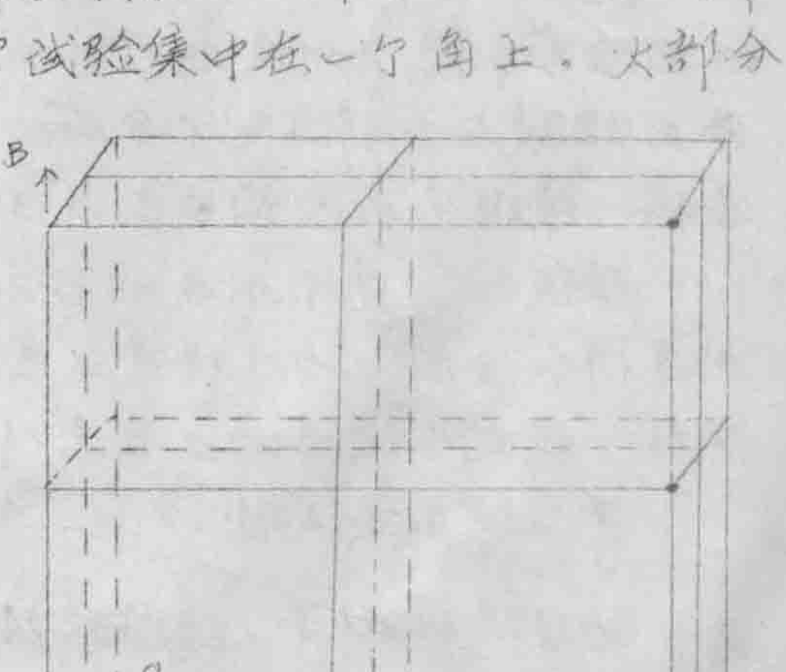
试验结果A_3最好，然后固定A于A_3，仍固定C于C_1，让B变化

$$A_3C_1 \begin{cases} \boxed{B_1} \\ B_2 \\ B_3 \end{cases}$$

以B_1最好，最后固定A，B于A_3，B_1，让C变化

$$A_3B_1 \begin{cases} C_1 \\ \boxed{C_2} \\ C_3 \end{cases}$$

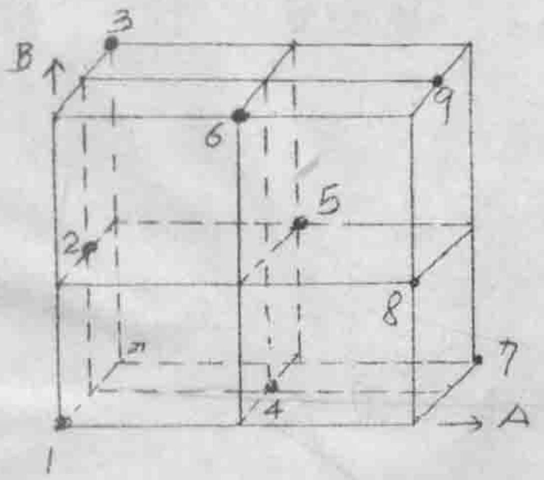
从上C_2为好，一般地认为$A_3B_1C_2$是较好的一种组合。上面共作了七次试验，在立体图上，这七个试验集中在一个角上，大部分范围内没有做试验，所以代表性不强，不能反映27个试验的客观情况。所以上述较好的组合在全体27个试验中就不一定是好的组合了。



当然，选取代表性强的试验来做，情况就会好些，用正交表来设计就解决了这个问题。我们用正交表 $L_9(3^4)$ 来设计这

试验时，把A、B，C排在$L_9(3^4)$的前三列上，正交表指示，要做九个试验$A_1B_1C_1$，$A_1B_2C_2$，$A_1B_3C_3$，$A_2B_1C_2$，$A_2B_2C_3$，$A_2B_3C_1$，$A_3B_1C_3$，$A_3B_2C_1$，$A_3B_3C_2$，在立体图上，这九个试验点均匀散布在立方体内，所以代表性强多了，这九个试验中的好结果，在27个试验中，也应当是较好的。这就是正交试验设计的均匀分散性，下面还要介绍它的整齐可比性，所用正交表来设计试验，

试验号 \ 列号	A 1	B 2	C 3
1	1	1	1
2	1	2	2
3	1	3	3
4	2	1	2
5	2	2	3
6	2	3	1
7	3	1	3
8	3	2	1
9	3	3	2



试验次数虽少，但能得到较多的信息。

§2. 用正交表设计，分析实例（直观分析法）。

例1. 北京化工厂生产的2.4二硝基甲苯肼是一种试剂，用于检验醛酮色分析及肝功能试验中，多年来采用2.4二硝基氯代苯与醋酸在乙醇溶剂中合成的工艺，工序繁杂，产品还是经常不合格，很难完成内销及出口任务。

经试验，采用水合肼与2.4二硝基氯代苯直接在乙醇水溶剂的条件下合成，为了稳定工艺，提高产率，用正交表做了两批小试验，找到了较佳的工艺条件。

第一批正交试验，

为了降低成本，提高收效率，经集体讨论，决定首批做以下大因素三位级的试验：

因素	(A)时间	(B)水合肼用量	(C)搅拌速度	(D)乙醇用量	(E)温度	(G)水合肼浓度
位级1	A1:2小时	B1:2倍(理论量)	C1:中快	D1:200ml	E1:回流	G1:20%
位级2	A2:4小时	B2:1.2倍	C2:快	D2:0ml	E2:60°C	G2:50%

五个师付在反应中

在小试验时，工人师付在反应中途看加200ml的乙醇，讨论时，大家提出这200ml的乙醇能否省下，所以对乙醇用量排了200ml和0ml两个位级。又如水和肼用量这个因素，知道相对于主料2.4二硝基氯代苯来说，应超过理论量，超过多少呢？由试验来鉴定，初步排了1.2倍和2倍两个位级。水合肼浓度20%和25%表示水合肼的粗品和精品，…

大因素二位级的全面试验要做64次试验，我们用$L_8(2^7)$来安排试验，只做38次，占全部的$\frac{1}{8}$，称为1/8部分实施。删去$L_8(2^7)$的第6列，把因素A，B，C，D，E，G依次排在表上，这时第1横行给出了第1号

试验计划							试验结果	
因素 / 试验号	(A)时间 1	(B)水合肼用量 2	(C)搅拌速度 3	(D)乙醇用量 4	(E)温度 5	(G)水合肼浓度 7	产率	外格，合溶，形，臭，量，醇度，
1	1(2小时)	1(2倍)	1(中快)	1(200ml)	1(回流)	1 20%	80%	合格
2	1	1	1	2(0ml)	2(60°C)	2 50%	93%	
3	1	2(1.2倍)	2(中快)	1	1	2	77%	
4	1	2	2	2	2	1	60%	
5	2(4小时)	1	2	1	2	2	90%	
6	2	1	2	2	1	1	86%	
7	2	2	1	1	2	1	60%	
8	2	2	1	2	1	2	60%	
Ⅰ：四个位级"1"产率之和	312	349	293	307	288	303	总和Σ=608	
Ⅱ：四个位级"2"产率之和	296	259	315	301	320	305		
Ⅲ=Ⅰ−Ⅱ	16	90	−22	6	−32	−2		
效应Ⅳ=$\frac{Ⅲ}{4}$	4	22.5	−4.5	1.5	−8	−0.5		

（表1）

试验计划A.B.C.D.E.G.即反应2小时，水合肼用量为理论用量的2倍，搅拌速度中快，乙醇用量为200ml，回流温度、20%的水合肼（粗品）下做试验。其余类推。

表1中列出了试验计划，试验结果，以及对结果——数据进行的分析，一般讲，用正交表安排试验，由于做的试验散布均匀，试验结果是相当分散的，有好有坏。上述八个试验中，第7.8号产品率只有60%，而第2.5号高达90%以上。比做试验之前大有提高。假如没有漏掉另外的重要因素，用量范围也没有多大出入。则八次试验中的最高产率在所有可能结果中也是相当高的。好的条件要给以足够重视，应重复试验进行核实。

另一方面，尽管这八个试验是按正交表排的，代表性强，但毕竟只占全部试验的1/8。比例是很小的，产率还有提高的可能。再从效果看，好的二号试验，第2.5号都是用50%水合肼（精品），相当于20%水合肼（粗品）来说，不仅成本高，而且不易找到。因此希望找到更好的条件。下面我们介绍正交表的整齐可比性，用它分析数据，分清因素的主次，为进一步安排试验提供信息。

效应值的计算，以第1列时间为例，位级"1"是反应时间2小时，共做了四次试验（1—4号），产率之和Ⅰ=80%+93%+77%+62%=312%，位级"2"是反应时间4小时，也做了四次（5—8号）试验，产率之和Ⅱ=90%+86%+60%+60%，两者产率之差Ⅲ=Ⅰ−Ⅱ=16%，效应Ⅳ=$\frac{Ⅲ}{4}$=4%反映了反应时间的不同对产品率的平均影响。因为其它五个因素的不同位级在前4次（1—4号）试验和后四次（5—8号）试验中各出现两次，它们不同位级对产品的影响在Ⅵ.Ⅳ中抵销了（这就是正交表的整齐可比性）。现在Ⅳ=4%表明位级"1"（2小时）比位级"2"（4小时）好，但好得不多，（也可能是误差的影响），不过操作时位级"1"（2时）比位级"2"（4小时）方便得多。我们愿意采用，为了慎重起见，决定在下一批试验中再比较一下2小时和4小时这两个不同位级。类似讨论其它五个因素的效应。从效应值的绝对值看，第2列的因素水合肼用量是首要因素，B_1（2倍）比B_2（1.2倍）要好得多，needs 在此处加强研究，重要因素要着重考察，第2批试

验时我们在水合肼用另为理论用另2倍的周围继续研究。其次，因素E（温度）的效应值为－8，看来60°C（E₂）比回流温度（E₁）要好，操作中60°C不易掌握，决定稳定在60°—70°C之间，因素C（搅拌）速度以快速搅拌为宜。第4列因素D（乙醇用另）和第7列因素G（水合肼浓度）效应值很小，是次要因素。从节约原则出发，我们选用D₂即不加乙醇，和G₁即用20%水合肼粗品，以降低成本。这说明，用水合肼粗品而不用精品，对产率影响不大，八次试验中，产率达90%以上的都是用的水合肼精品，上面的分析表明，用水合肼粗品也可能达到。为进一步安排试验提供了信息。

另外，在做上述八次试验时，第2、7号试验外形均为紫色，是不合格品，重新试验，又得到桔黄色的合格外形（表1中列出的是重做的合格品的指标）。因此，我们认为还有影响外形的重要因素未被注意和考察，于是界值转化为影响外形的重要因素。经过调查研究知道，出现紫色外形是此产品多年来一直未解决的老问题。工人师付把这两号试验前后两次具体情况进行分析，使想影响外形的主要因素可能是加料速度决定在第二批试验中来考察。

对主要因素和紫色因素研究

第二批正交试验

由第一批试验的试验情况及对试验结果分析提供的信息，决定第二批作以下三因素，二位级的试验。

因素	A 时间	B 水合肼用另	C 加料速度
位级"1"	2小时	1.7倍	快
位级"2"	4小时	2.3倍	慢

选用正交表$L_4(2^3)$，共四次试验，结果如下：

因素 / 试验号	A 时间 1	B 水合肼用量 2	C 加料速度 3	产率	外形	其它指标
1	1（2小时）	1（1.7倍）	1（快）	62%	不合格	合格
2	1	2（2.3倍）	2（慢）	86%	合格	
3	2（4小时）	1	2	70%	合格	
4	2	2	1	70%	不合格	
Ⅰ	148	132	132			
Ⅱ	140	156	156	总和 $\Sigma=288$		
Ⅲ	8	−24	−24			
$Ⅳ=\frac{Ⅲ}{2}$	4	−12	−12			

（表 2）

为简化工序，在第二批试验中，我们采用未经处理的工业氯代苯，而第一批是用的精制过的氯代苯（精制过程不但有可能发生危险，而且损失氯代苯30%）。经换算，第二批试验的产率普遍比第一批高。第2号试验的产品最高，换算后比第一批试验的最高产品率要好的多，验证了用水合肼粗品也有高的产率。

四次试验的结果表明，加料速度是影响外形的重要原因，慢速加料的都是合格的桔黄色，而快速加料的试验，外形是紫色的不合格品。由此得出结论应采用慢速加料。其它，水合肼用量是2.3倍好，反应时间还是2小时稍好。

用上述两批试验结果得到的工艺条件是：用工业氯化苯与水合肼在酒精溶液中合成，水合肼的用量为理论用量的2.3倍，反应二小时，温度在60—70°C之间。慢速加料，快速搅拌，水合肼粗品精品均可。用此工艺投产，效果很好，平均产率在80%以上，质量指标全部达二级标准，顺利完成出口任务。

例2. 污水去汞正交试验

含汞污水流入农田，危害很大。为了降低污水含汞率，我们

（北京化工厂）用离子交换纤维交换的办法来除去污水中的汞，作了以下四因素、三位级的正交试验，因素与位级如下：

因素	装填密度	接触时间	污水PH值及纤维型号	管型
位级 1	0.12	6分	PH7 H型	1:15
2	0.16	9分	PH7 Na型	1:5
3	0.18	12分	PH9 H型	1:10

污水PH值及纤维型号本来是两个因素，试验前估计它们不是重要因素，把它们合在一起分作三种状态，作为一个因素的三个位级。它和接触时间、装填密度（管内装填纤维的密度）和管型（直径与管长的比例）是四个因素。可以用正交表 $L_9(3^4)$ 来安排试验，否则五个因素三位级试验要用正交表 $L_{27}(3^{13})$，试验次数增加多了。（参看组合法）

试验的设计与结果如下：

试验计划					试验结果		
因素 / 试验号	A 密度 1	B 时间 2	C 纤维型号与PH值 3	D 管型 4	出水含汞 (mg/L)	最大流量 (ml/分)	综合评分
1	1 (0.12)	1 (6分)	1 (Na-7)	1 (1:15)	0.047	62	50
2	1	2 (9分)	2 (H-7)	2 (1:5)	0.048	262	90
3	1	3 (12分)	3 (H-9)	3 (1:10)	0.052	150	60
4	2 (0.16)	1	2	3	0.049	145	80
5	2	2	3	1	0.044	30	55
6	2	3	1	2	0.038	62	80
7	3 (0.18)	1	3	2	0.066	140	55
8	3	2	1	3	0.042	44	60
9	3	3	2	1	0.027	28	60
Ⅰ 三个位级1评分之和	200	185	190	165			
Ⅱ 三个位级2评分之和	215	205	230	225			
Ⅲ 三个位级3评分之和	175	200	170	200			
Ⅳ＝Ⅰ.Ⅱ.Ⅲ中最大－最小	40	20	60	60			
数值Ⅴ＝$\frac{Ⅳ}{3}$	13.3	6.7	20	20			

总和Σ＝590

综合评分是根据出水含汞、污水流量与管型（粗短的易于甚速，总分多些）估计的。

由效应分析看出纤维型号污水PH值和管径是重要因素，决定将废水调成中性，纤维采用H型，即H-7，管型比例取1:6，装填密度取0.16，接触时间取9—12分，进行了中型试验，效果较好。已处理过200吨污水，含汞量降到0.05mg/l以下。

一般来讲，影响试验结果的因素很多，但是起主要作用的因素并不多，经常是一两个主要因素起作用。用正交表来安排多因素试验，对所有因素而言，只进行了部分试验，然而对于两个任意因素来讲，它们的不同位级组合，都进行了试验，甚至是重复地进行试验。当然，在起主要作用的因素只有两个的情况下，对此主要因素是进行了全面的甚至是重复的试验。所以往往能得出较好的试验条件。

二 进一步的分析

上面已介绍了正交试验的设计与分析，我们一定会想到，进行试验都会产生误差，有些因素，它的效应值较小，即此因素的某一位级比其它位级稍好，能否说这种稍好一定是由条件改变引起的呢？当然不能。试验误差也会引起结果数据的波动。那末，效应值多大时？我们能够判断说，它是由于因素的条件变化引起的。

另外，由于因素的效应，误差不同，所以试验结果也不相同。这结果数据和效应，误差间的关系式是什么？即数据的结构也必须搞清楚。这是用数学方法进一步分析所必需的。

§1. 数据的结构

我们先从最简单的单因素试验来分析，设考察的因素只有一个，即因素A，取了m个不同位级$A_1, A_2, \cdots, A_m$。在条件A_i下做试验，理论上一定有一个数值μ_i，由于误差的影响，试验结果不一定就是μ_i，设在条件A_i下的第j次试验结果是y_{ij}，误差是ε_{ij}，

于是
$$y_{ij}=\mu_i+\varepsilon_{ij} \qquad (1.1)$$
就是数据的结构式。

数据结构式左端是可以通过试验测得的量，而右端则是尚不知道的客观存在的量，我们的问题是从数据y_{ij}来估计出理论值μ_i，由此找到较好的试验条件。

比较理论值μ_i，有两种方法：一种是直接比较μ_i的大小；另一种是比较μ_i-a（a是一个与i无关的常数）的大小，通常取$a=\bar{\mu}=\frac{1}{m}\sum_{i=1}^{m}\mu_i$是$m$个理论值的"平均值"。就是将每一个$\mu_i$与平均值$\bar{\mu}$比较，考虑$\mu_i-\bar{\mu}$的大小，判断$\mu_i$的大小。现在看起来不方便，但在多因子试验中，用后一种方法要方便得多。因为它有一个特点，在要比较的$\mu_i-\bar{\mu}$之间有恒等式

$$\sum_{i=1}^{m}(\mu_i-\bar{\mu})=0 \tag{1.2}$$

今后我们用 $a_i=\mu_i-\bar{\mu}$ 表示理论值偏离平均值的情况，当然有 $\sum_{i=1}^{m}a_i=0$.

二因素试验的数据结构．因素A.B对试验的指标数据有影响，A有 l 个不同位级 $A_1, A_2, \cdots, A_l$，B有 m 个不同位级 $B_1, B_2 \cdots B_m$，在条件 A_iB_j 下的试验理论上一定有一个指标 μ_{ij}，但是由于做试验时有试验误差 ε_{ij}，试验数据为 y_{ij}，于是：

$$y_{ij}=\mu_{ij}+\varepsilon_{ij} \qquad i=1,\cdots,l. \quad j=1,\cdots,m.$$

下面对 μ_{ij} 作进一步的分析．令

$$\mu=\frac{1}{lm}\sum_{i=1}^{l}\sum_{j=1}^{m}\mu_{ij}, \quad \mu_{i\cdot}=\frac{1}{m}\sum_{j=1}^{m}\mu_{ij}, \quad \mu_{\cdot j}=\frac{1}{l}\sum_{i=1}^{l}\mu_{ij}.$$

易见 μ 是总平均，μ_i 是在条件 A_i 下的平均指标，μ_j 是在条件 B_j 下的平均指标。它们反映因素A.B的不同位级的平均指标．令

$$a_i=\mu_{i\cdot}-\mu, \qquad b_j=\mu_{\cdot j}-\mu$$

当 $a_i>0$ 时，表明条件 A_i 使指标增大 a_i，$a_i<0$ 时．就表示条件 A_i 使指标减小，我们称 a_i 是因素A位级 A_i 的效应，同样 b_j 就是因素B位级 B_j 的效应。它们满足恒等式

$$\begin{aligned}\sum_{i=1}^{l}a_i&=\sum_{i=1}^{l}(\mu_i-\mu)=0\\ \sum_{j=1}^{m}b_j&=\sum_{j=1}^{m}(\mu_j-\mu)=0\end{aligned} \tag{1.3}$$

这时．分两种情况来讨论

① $\mu_{ij}=\mu+a_i+b_j \qquad i=1,\cdots,l. \quad j=1,\cdots,m,$

它表明条件 A_iB_j 的理论指标是由A的效应 a_i 与 B_j 的效应 b_j "叠加"到平均指标上而成的。这时数据结构式是

$$y_{ij}=\mu+a_i+b_j+\varepsilon_{ij} \qquad i=1,\cdots,l. \quad j=1,\cdots,m \tag{1.4}$$

这是一种比较简单的情况。

(2) $\mu_{ij} \neq \mu + a_i + b_j$

它表明条件 A_iB_j 的指标仅由 A_i、B_j 的效应 a_i、b_j 叠加于平均指标上还不行，还要考虑 A_iB_j 条件这一种搭配本身对指标的影响。若数 $\mu_{ij}-\mu-a_i-b_j$ 记为 $(ab)_{ij}$ 反映了这种影响的大小，称为因素A和B的交互作用效应。我们证明它们也满足恒等式：

$$\sum_{i=1}^{l}(ab)_{ij}=0 \qquad \sum_{j=1}^{m}(ab)_{ij}=0 \qquad (1.5)$$

证：$\sum_{i=1}^{l}(ab)_{ij}=\sum_{i=1}^{l}(\mu_{ij}-\mu-a_i-b_j)=l\mu_{\cdot j}-l\mu-lb_j=0$

同样 $\sum_{j=1}^{m}(ab)_{ij}=0$

这时，即有交互作用时，数据结构式为

$$y_{ij}=\mu+a_i+b_j+(ab)_{ij}+\varepsilon_{ij} \qquad i=1,\cdots,l \quad j=1,\cdots,m \quad (1.6)$$

搞清楚二因素的数据结构式之后，就不难写出多因子的数据结构式。对三因素，A、B、C数据结构式是*

$$y_{ijk}=\mu+a_i+b_j+c_k+(ab)_{ij}+(ac)_{ik}+(bc)_{jk}+\varepsilon_{ijk} \quad (1.7)$$

当某些因子之间没有交互作用时，就在结构式中将相应的效应删去。

例1的数据结构是（因子间没有交互作用）

$$\left\{\begin{aligned}
y_1&=\mu+a_1+b_1+c_1+d_1+e_1+g_1+\varepsilon_1\\
y_2&=\mu+a_1+b_1+c_1+d_2+e_2+g_2+\varepsilon_2\\
y_3&=\mu+a_1+b_2+c_2+d_1+e_1+g_2+\varepsilon_3\\
y_4&=\mu+a_1+b_2+c_2+d_2+e_2+g_1+\varepsilon_4\\
y_5&=\mu+a_2+b_1+c_2+d_1+e_2+g_2+\varepsilon_5\\
y_6&=\mu+a_2+b_1+c_2+d_2+e_1+g_1+\varepsilon_6\\
y_7&=\mu+a_2+b_2+c_1+d_1+e_2+g_1+\varepsilon_7\\
y_8&=\mu+a_2+b_2+c_1+d_2+e_1+g_2+\varepsilon_8
\end{aligned}\right. \qquad (1.8)$$

---

* 这里假设高级交互作用不存在。

其中 $a_1+a_2=0$，$b_1+b_2=0$，$\cdots$，$g_1+g_2=0$。

例2中当因子无交互作用时，数据结构式是

$$
\left\{
\begin{aligned}
y_1 &= \mu + a_1 + b_1 + c_1 + d_1 + \varepsilon_1 \\
y_2 &= \mu + a_1 + b_2 + c_2 + d_2 + \varepsilon_2 \\
y_3 &= \mu + a_1 + b_3 + c_3 + d_3 + \varepsilon_3 \\
y_4 &= \mu + a_2 + b_1 + c_2 + d_3 + \varepsilon_4 \\
y_5 &= \mu + a_2 + b_2 + c_3 + d_1 + \varepsilon_5 \\
y_6 &= \mu + a_2 + b_3 + c_1 + d_2 + \varepsilon_6 \\
y_7 &= \mu + a_3 + b_1 + c_3 + d_2 + \varepsilon_7 \\
y_8 &= \mu + a_3 + b_2 + c_1 + d_3 + \varepsilon_8 \\
y_9 &= \mu + a_3 + b_3 + c_2 + d_1 + \varepsilon_9
\end{aligned}
\right. \qquad (1.9)
$$

其中 $a_1+a_2+a_3=0$，$b_1+b_2+b_3=0$，$c_1+c_2+c_3=0$，$d_1+d_2+d_3=0$

§2 效应的估計（直观分析法）

仍以例1为例，已经知道了它的数据结构式（1.8），方程左端是已知的试验数值，右端是一些未知的，但是是客观存在的因素的效应值和试误误差。

下面用最小二乘估計来求出諸因素的估计值。μ的估計值是$\hat{\mu}$。设μ的估计值是$\hat{\mu}$，a_1的估计值是$\hat{a}_1$，a_2的估计值是$\hat{a}_2$，……，当然我们要求估计值也满足恒等式：

$$\hat{a}_1+\hat{a}_2=0,\quad \hat{b}_1+\hat{b}_2=0,\quad \cdots,\quad \hat{g}_1+\hat{g}_2=0$$

最小二乘法的原则是：估计值$\hat{\mu}$，$\hat{a}_i$，$\hat{b}_i$，…，$\hat{g}_i$使得误差的平方和最小。

误差的平方和是

$$\sum_{i=1}^{9}\varepsilon_i^2=(y_1-\hat{\mu}-a_1-b_1-c_1-d_1-e_1-g_1)^2+(y_2-\hat{\mu}-a_1-b_1-c_1-d_2-e_2-g_2)^2$$

$$+\cdots+(y_9-\mu-a_2-b_2-c_1-d_2-e_1-g_2)^2$$

它取最小值时，必须使諸偏導数为0。

由 $\frac{\partial\sum\varepsilon_i^2}{\partial\mu}=0$ 得 $\sum\limits_i y_i-9\mu=0$.

所以μ的估计值$\hat{\mu}=\frac{1}{8}\sum y_i=\bar{y}$，是数据的平均值.

由 $\frac{\partial\sum\varepsilon_i^2}{\partial a_1}=0$ 得

$$y_1+y_2+y_3+y_4-4\mu-4a_1=0$$

所以a_1的估计值 $\hat{a}_1=\frac{1}{4}(y_1+y_2+y_3+y_4)-\hat{\mu}$，用前面的记号，以$I_1$记因素A位级$A_1$的4个数据之和，即

$$I_1=y_1+y_2+y_3+y_4$$，所以

$$\hat{a}_1=\frac{I_1}{4}-\hat{\mu}$$

由 $\frac{\partial\sum\varepsilon_i^2}{\partial a_2}=0$ 得

$$y_5+y_6+y_7+y_8-4\mu-4a_2=0$$

所以a_2的估计值是

$$\hat{a}_2=\frac{1}{4}(y_5+y_6+y_7+y_8)-\mu=\frac{II_1}{4}-\hat{\mu}.$$

由其它偏导数等于0，类似求得各估计值是

$$\hat{b}_1=\frac{I_2}{4}-\hat{\mu},\qquad b_2=\frac{II_2}{4}-\hat{\mu},$$
$$\hat{c}_1=\frac{I_3}{4}-\hat{\mu},\qquad c_2=\frac{II_3}{4}-\hat{\mu},$$
$$\hat{d}_1=\frac{I_4}{4}-\hat{\mu},\qquad d_2=\frac{II_4}{4}-\hat{\mu},$$
$$\hat{e}_1=\frac{I_5}{4}-\hat{\mu},\qquad e_2=\frac{II_5}{4}-\hat{\mu},$$
$$\hat{g}_1=\frac{I_7}{4}-\hat{\mu},\qquad g_2=\frac{II_7}{4}-\hat{\mu}.$$

对照例1的分析，不难看出因素的两个效应估计之差就是Ⅳ栏内的因素效应。

$$\hat{a}_1-\hat{a}_2=(I_1-II_1)/4$$
$$\hat{b}_1-\hat{b}_2=(I_2-II_2)/4$$
$$\cdots$$
$$\hat{g}_1-\hat{g}_2=(I_7-II_7)/4.$$

对例2（假设它无交互作用）的分析也一样，设μ, a_i, b_i, c_i, d_i，对(1.9)用最小二乘估计，因