

正交试验在农业上的应用

西北农学院编

一九七六年十一月

前 言

无产阶级文化大革命和批林批孔运动，深刻地批判了刘少奇、林彪的反革命修正主义路线，有力地推动了各项工作。战斗在农业生产第一线的广大贫下中农和科技人员，坚持以阶级斗争为纲，坚持党的基本路线，坚持无产阶级专政下的继续革命，深入开展“农业学大寨”运动，抓革命，促生产，夺得了农业生产的连年丰收。农业科学实验，也取得了显著成绩。

“群众是真正的英雄。”在开展农业科学实验的过程中，广大群众成功地、创造性地应用了正交试验方法，把数学从课堂和书本上解放出来，变成群众科学实验的有力武器。实践证明，正确地应用正交试验，可减少试验处理，缩短试验周期，提高试验效果，方法简单，应用方便，深受广大群众欢迎。

为了适应日益发展的群众性科学实验的需要，我们挑选了一些农业试验的例子，说明正交试验的基本原理和方法，供同志们参考。由于我们水平所限，错漏难免，欢迎批评指正。

当前，举国欢腾。全国人民热烈庆祝华国锋同志任中共中央主席、中央军委主席，热烈欢呼粉碎“四人帮”篡党夺权阴谋的伟大胜利。我们坚信，在以华国锋同志为首的党中

在英明领导下，通过深入批判“四人帮”反革命的修正主义路线和他们篡党夺权、复辟资本主义的滔天罪行，人民群众的社会主义积极性，一定会象火山一样地迸发出来，推动各条战线社会主义革命和社会主义建设的飞速发展。群众性的农业科学实验也一定会出现一个崭新的局面。让我们紧密地团结在华国锋主席为首的党中央周围，搞好各方面的工作，把无产阶级革命事业进行到底。

编者

1976年11月于西北农学院

目 录

前 言

第一章 正交试验的直观分析法

§ 1 多因素试验	(1)
§ 2 合理安排试验的原则	(2)
§ 3 正交表	(6)
§ 4 正交试验	(7)
4—1 试验方案的设计	(7)
4—2 试验结果的分析	(11)
§ 5 多指标试验	(22)
5—1 混杂、区组和裂区	(23)
5—2 多指标试验的综合平分法	(24)
5—3 综合平衡法	(29)
§ 6 水平数不等的试验	(32)
§ 7 有交互作用的试验	(36)
7—1 交互作用概念	(36)
7—2 能考察交互作用的正交表	(37)
7—3 一个例子	(38)
7—4 有交互作用的试验的表头设计原则	(45)
7—5 一个五水平试验的例子	(49)
§ 8 有重复的试验	(54)

第二章 方差分析

§ 1	偏差平方和	(62)
§ 2	正交试验的方差分析	(65)
2—1	$S_{总}$ 的计算	(67)
2—2	各列偏差平方和的计算	(69)
2—3	F 检验	(71)
§ 3	有重复试验的方差分析	(81)
§ 4	没有空白列的试验的方差分析	(90)
§ 5	方差分析的简化方法	(97)
5—1	没有交互作用的试验的简化方差分析法	(97)
5—2	含有交互作用及重复试验的例子	(99)

第三章 $L_{27}(3^3)$ 型表的构造

§ 1	有限域简介	(104)
1—1	质数与合数	(104)
1—2	同余	(105)
1—3	数域	(105)
1—4	元素个数为质数的有限域	(106)
1—5	元素个数为质数幂的有限域	(111)
§ 2	二水平正交表的构造	(117)
2—1	水平运算规则	(117)
2—2	正交表与交互列表的构造	(117)
2—3	由二水平表变来的混合型表	(122)
§ 3	三水平正交表的构造	(125)
3—1	水平运算规则	(126)

3—2	正交表与交互列表的构造·····	(126)
3—3	由三水平表变来的混合型表·····	(131)
§ 4	构造 $L_{t^m}(t^q)$ 的一般方法·····	(132)
4—1	t 水平的运算·····	(132)
4—2	正交表与交互列表的构造·····	(132)
4—3	混合型表·····	(133)

附录:

F 表、 $d(n, k)$ 表、 $\Phi(n, k)$ 表及常用正交表。

第一章 正交试验的直观分析法

§ 1. 多因素试验

世界上一切事物都是互相联系互相制约的，一种事物的运动和发展，总是由多种原因造成的。由于这个缘故，多因素试验比起单因素试验，更有重大意义。

什么叫多因素试验呢？例如，根据生产实践，播期，密度和施肥方式这三个因素，对谷子产量影响很大。通过一个试验，同时确定合适的播期；合理的播量和较好的施肥方式，保证谷子高产，这就是一个多因素试验。这里，谷子亩产量是试验指标，用它衡量试验效果的好坏。由于产量可直接用数量来表示，称为定量指标。有些试验指标无法用数量表示，只能凭目测手模来评定，则称为定性指标。定性指标可以用评分、评等级的办法来表示，这样就把定性指标转化为定量指标了。谷子试验，指标只有亩产一个，称为单指标试验。试验指标如果是多项的，就称为多指标试验。播期，密度和施肥方式则是要通过试验来考察的影响试验指标的三个因素。因素一般用A，B，C，D……来表示。有一类因素在生产中可以人为地控制和调节，称为可控因素；另一类因素，暂时还不能人为地调节和控制，称为不可控因素，如降雨问题，目前人还不能控制。试验中要考察的因素只适用于可控因素。因素的变化可能引起指标的变化。因素在试验中的各种状态叫做因素的水平。变化了几种状态，就叫它是几水平

的因素。如播期可以取5月25日，5月28日和6月2日等三个水平（分别用 A_1 ， A_2 ， A_3 来表示），是一个三水平的因素。谷子试验中要考察三个因素，每个因素又可取几个水平。象这样考察多个因素，而每个因素又可以选取多个水平的试验称为多因素试验。三个三水平因素的试验称为 3^3 型试验。一般地， n 个 m 水平因素的试验称为 m^n 型试验。如果考察的因素有 $m+n$ 个，其中 m 个取2水平， n 个取3水平，则称为 $2^m \times 3^n$ 型试验。

通过多因素试验，可以分析因素与试验指标的关系，即当因素变化时，指标是如何变化的？找出这种变化规律，可以利用它能动地指导生产；又可分清因素的主次，如在谷子试验中，要分清播期、密度、施肥方式三种因素那个是主要的，那个是较次的，还可找到使试验指标最优的水平搭配，如播期、密度、施肥方式各取什么水平，使谷子产量最高。

§ 2 合理安排试验的原则

要做好一个多因素试验，首先要选择好有代表性的试验处理。实践证明，处理选的好，数目不多，就能分清因素主次，选出最优的水平搭配；选的不好，处理数目虽多，也解决不了问题。其次，要考虑到因素的各个水平在试验中要处于完全相似的状态，以便予分析比较。

上述两点是我们合理安排试验的基本原则。下面以三个三水平因素试验为例，说明这个原则。

设有三个因素 A ， B ， C 。 A 的三个水平记为 A_1 ， A_2 ， A_3 ； B 的三个水平记为 B_1 ， B_2 ， B_3 ； C 的三个水平记为 C_1 ， C_2 ， C_3 。如果按它们所有可能搭配的情况作试验，则需要做 $3^3 = 27$ 个处理（每一个水平搭配叫一个处理，如图2—1

中每一个“○”代表一个处理)，把27个处理结果通过分析比较，就能找到最优的水平搭配（这里最优是相对于27种搭配而言）。把各个因素的各个水平的所有可能搭配都拿来作试验，称为全面试验。

然而，随着因素个数的增加，处理个数会急剧增长。如8个7水平因素的全而试验，要作 $7^8 = 7564801$ 个处理。这样势必使试验周期拖长，费用增加，实际上这也是难以实现的。因此，全面试验虽然能解决问题，但并不是一种有效的办法，所以通常较少采用。

克服全面试验这种缺点的办法是部分实施试验。也就是从全面试验中选取一部分有代表性的处理来作，以达到我们的

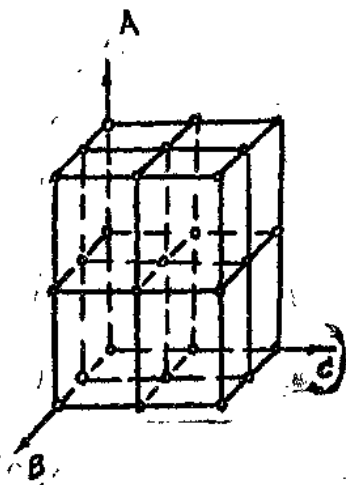


图2—1

目的。譬如，上述27个处理中，想用9个处理来进行试验，挑选那9个处理才有代表性呢？如选图2—2的9个黑点就好。好就好在立方体的每个面上（对应于 A_1 、 A_2 、 A_3 的是上、中、下三个面；对应于 B_1 、 B_2 、 B_3 的是前、中、后三个面；对应于 C_1 、 C_2 、 C_3 的是左、中、右三个面）都恰有三个点，而且立方体的每条线上也恰有一个点。9个点均衡地散布于整个立方体内。由于这9个处理是均衡分散的，它们的代表性就强，能够比较全面地反映整个立方体的大致

情况。所以这 9 个点中的好点即使不是全面试验的最好点，也往往是相当好的；类似地，其中的坏点也往往是相当坏的；而且由这 9 个点往往能分析出全面试验的好点。在多因素试验中，所谓处理点要选的具有代表性，就是指处理点是均衡分散的。如果选图 2—3 的九个黑点就不好。因为因素 C 的第一个水平有两个处理，第 2 个水平有五个处理，第三个水平有两个处理，分配的不均匀；同样，因素 B 的三个水平分配的也不均匀。因而用这 9 个处理就不能比较全面地代表整个立方体的情况，而且各因素各水平在试验中所处的状态不一样，也就不便于进行分析比较。所以，选用这 9 个处理，就不一定能得到正确的结论。这就是说图 2—3 所选的处理点不是均衡分散的，无代表性。下面我们吧图 2—2 所代表的试验方案列成表（表 2—1），进一步分析它的优点。

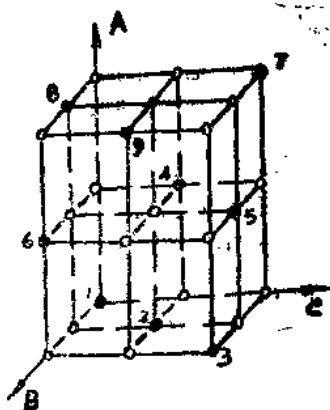


图 2—2

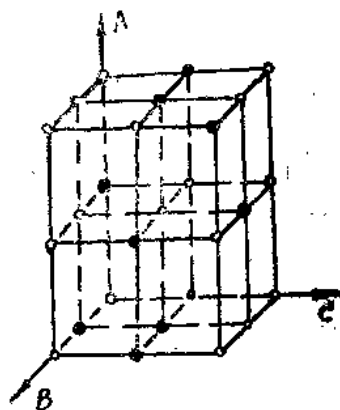


图 2—3

表 2—1

3³型试验的部分实施试验方案

因素 试验号	A	B	C	试验指标
1	A ₁	B ₁	C ₁	
2	A ₁	B ₂	C ₂	
3	A ₁	B ₃	C ₃	
4	A ₂	B ₁	C ₂	
5	A ₂	B ₂	C ₃	
6	A ₂	B ₃	C ₁	
7	A ₃	B ₁	C ₃	
8	A ₃	B ₂	C ₁	
9	A ₃	B ₃	C ₂	

由表2—1看出：每个因素的每个水平都与其它因素的每一水平相碰一次，且仅相碰一次，表现了它的均衡分散性；每个因素的每个水平参加处理的次数相等（都是三次），便于分析比较，具有整齐可比的特点。表2—1简记为表2—2。

2—2

3³型试验的部分实施试验方案

水平 因素 试验号	A	B	C	试验指标
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	

通过实践，人们摸索出了一系列如表2—2那样，具有均衡分散性和整齐可比性两个特点的表，用来安排多因素试验，使处理次数大大减少，这样的表叫正交表。

§ 3 正交表

正交表是由代表因素水平的数字组成的、具有某种数学性质的表格。

最简单的正交表是 $L_4(2^3)$ ，如表3—1所示。

表3—1

$L_4(2^3)$

列号 水平 试验号	1	2	3
1	1	1	1
2	1	2	2
3	2	1	2
4	2	2	1

记号 $L_4(2^3)$ 中， L 表示它是一张正交表；4表示它有4行，用它安排试验要有4个处理； 2^3 中的3表示它有3列，用它安排试验时，最多可以考察3个因素； 2^3 中的2表示每列恰有1，2两种数字，安排试验时，被考察的因素一般要求为二水平。

又如正交表 $L_8(2^7)$ ，表示这张表有8行7列，表中每列由1，2两种数字组成，用来安排试验，要作8个处理，最多能考察7个因素，每个因素一般要求为二水平。

表3—1具有如下两个性质：

(1) 每一列中，数字“1”和“2”出现的次数相同。

(2) 任意两列中, 把同一行的数字搭配成有序数对, 如 $(1, 1)$, $(1, 2)$, $(2, 1)$, $(2, 2)$, 每种数对出现的次数是相同的。

第一个性质是整齐可比性的体现; 第二个性质是均衡分散性的体现。这两个性质叫正交性, 这就是表 3—1 称为正交表的由来。由“1”和“2”两种数组成的具有正交性的数字表, 都叫做二水平的正交表。

用正交表安排多因素试验, 通常对全体因素来说是部分实施试验。如 2^3 型的全面试验要作 $2^3 = 8$ 个处理, 而用 $L_4(2^3)$ 来安排, 只须 4 个处理。又如 2^7 型的全面试验要有 $2^7 = 128$ 个处理。而用 $L_8(2^7)$ 安排只须 8 个处理。

象 $L_4(2^3)$, $L_8(2^7)$ 等叫做相同水平的正交表。还有一种是“混合水平”的正交表。如 $L_8(4 \times 2^4)$, 表示有 8 行 5 列, 其中第一列由 1, 2, 3, 4 四种数字组成, 后四列是 1, 2 两种数字组成; 用它安排试验时, 要作 8 个处理; 最多可以考察一个四水平和四个二水平的因素。

§ 4 正交试验

正交试验, 包括怎样安排试验(即如何设计试验方案)和怎样分析试验结果两部分内容。下面通过例子予以说明。

4—1 试验方案的设计

〔例 1〕陕西省吴旗县农技站, 为了寻找谷子丰产措施, 选用如表 4—1 所示的因素和水平进行了试验。

谷子试验是 3^3 型试验, 可选用三水平的正交表来安排这个试验。如可用 $L_9(3^4)$ 来安排这个试验。 $L_9(3^4)$ 有 4 列, 最多可安排 4 个 3 水平因素, 这个试验是可以用它来安排的。利用这个表, 只要作 9 个处理就可以了。而 3^3 型的全

面试验,要作 $3^3 = 27$ 个处理,因此9个处理是27个处理的 $1/3$ 部分实施试验。

表 4—1 谷子试验的因素水平表

因素 次序	播期(日/月) A	密度(万株/亩) B	施肥方式 C
1	23/5	2	每小区施“5406”菌肥6斤作底肥
2	28/5	2.5	每小区施“5406”6斤作底肥拔节追尿素0.4斤
3	2/6	3	拔节追尿素0.4斤抽穗前追硝酸铵0.8斤

(注:每小区面积为0.02亩)

把A, B, C三个因素任意填在 $L_9(3^4)$ 的表头的三列上,如依次填在1, 2, 3列上,这一步通常叫做表头设计(表4—2)。

表 4—2 谷子试验的表头设计

因素	播期 A	密度 B	施肥方式 C	
列号	1	2	3	4

然后,根据表头设计和 $L_9(3^4)$,列出试验方案(表4—3)。如播期(A)在 $L_9(3^4)$ 的第1列,就把第1列中的所有“1”换成 $A_1(23/5)$,“2”换成 $A_2(28/5)$,“3”换成 $A_3(2/6)$ 。同样可安排B(第2列)和C(第3

列)。至于 $L_3(3^4)$ 的第4列未用，先不管它。

表 4—3

谷子试验方案

试验号	播期(A) (1)	密度(B) (2)	施肥方式 (C) (3)	区组(D) 4	试验指标 $\bar{y}$ (斤/小区)
1	23/5 (1)	2 (1)	C_1 (1)	I (1)	
2	23/5 (1)	2.5 (2)	C_2 (2)	II (2)	
3	23/5 (1)	3 (3)	C_3 (3)	III (3)	
4	28/5 (2)	2 (1)	C_2 (2)	III (3)	
5	28/5 (2)	2.5 (2)	C_3 (3)	I (1)	
6	28/5 (2)	3 (3)	C_1 (1)	II (2)	
7	2/6 (3)	2 (1)	C_3 (3)	II (2)	
8	2/6 (3)	2.5 (2)	C_1 (1)	III (3)	
9	2/6 (3)	3 (3)	C_2 (2)	I (1)	

表 4—3 中每一行都是一个处理，如第 1 行（第 1 个试验号）为 $A_1B_1C_1$ （23/5 播种，密度 2 万株/亩，采用第 1 种施肥方式），第 2 行为 $A_1B_2C_2$ （23/5 播种，密度 2.5 万，采用第二种方式施肥）等等。如不设重复，只种 9 个小区。

在安排试验时，试验地绝不会完全均匀一致。这种不均匀性会影响试验的精确性。为了尽可能地避免这种干扰，可采用下面两种办法，来安排试验小区。

1. 随机化法：把 9 个处理，随机地安排在 9 个试验小区内（图 4—1）。

1	5	2
7	3	9
6	8	4

图 4—1

2. 区组因素法：把试验田作为一个因素 D 放到正交表中去，用正交表来帮助我们安排小区。如把试验田(D)分成三大块，代表 D 的三个水平(I、II、III)，把 D 放在 $L_9(3^4)$ 的第4列。根据表4—3，每一大块分成3个试验小区，安排三个处理(图4—2)。

I	1	5	9	I
II	6	2	7	II
III	8	4	3	III

图4—2

到此为止，试验方案的设计就算完成了。试验方案设计的步骤可概括如下：

1. 明确试验目的，确定试验指标。

在实际生产中，存在的问题较多，所有存在的问题不可能通过一次试验给予解决。因此，在试验之前，应根据实际情况，确定这次试验主要解决的一个问题或几个问题。针对这几个问题，确定相应的试验指标。

2. 挑选因素水平，列出因素水平表。

根据实践经验，挑选因素水平。因素水平直接关系到实验效果的好坏，必须认真对待。

3. 选用正交表。

根据因素和水平多少以及试验工作量大小，选用合适的正交表。如每因素都是三水平，当有3—4个因素时，一般用 $L_9(3^4)$ ；当有5—7个因素时，一般用 $L_{27}(3^7)$ 。

等。

4. 作表头设计。

把各个因素放在选用的正交表表头的适当列上，例如表 4—2 的表头设计。在简单的情况下，这一步是容易的；在复杂的情况下，要达到安排适当，还需要认真考虑。

5. 列出试验方案。

把表头上有因素的每个列的数字依次换成该因素的水平，最后列成如表 4—3 的试验方案。在列试验方案时，应同时考虑田间设计。

4—2 试验结果的分析

表 4—4 谷子试验结果分析表

因素 试验号	播期(A) (1)	密度(B) (2)	施肥方式 (C) (3)	(空) (4)	试验指标 斤/小区
1	1	1	1	1	14.9
2	1	2	2	2	13.7
3	1	3	3	3	12.4
4	2	1	2	3	12.4
5	2	2	3	1	12.4
6	2	3	1	2	10.85
7	3	1	3	2	12.2
8	3	2	1	3	11.0
9	3	3	2	1	10.5
K ₁	41	39.5	36.75	37.8	总和Σ = 110.35
K ₂	35.65	37.1	36.6	36.75	
K ₃	33.7	33.75	37.0	35.8	
k ₁	13.7	13.2	12.3	12.6	
k ₂	11.9	12.4	12.2	12.3	
k ₃	11.2	11.3	12.33	11.9	
R	2.5	1.9	0.12	0.7	