二重格子方区组法

15	10	1	8
12	13	6	3
5	4	11	14
2	7	16	9

12	15	2	5
1	6	11	16
14	9	8	3
7	4	13	10

12	14	1	7
13	11	8	2
6	4	15	9
3	5	10	16

11	9	10	12
7	5	6	8
3	1	2	4
15	13	14	16

10	14	6	2
12	16	8	4
9	13	5	1
11	15	7	3

山东省农业科学院情报资料研究所

1 9 8 0

目　　录

二重格子方区组法(简单格子法)

一、绪　言

我们如考虑把许多品种或处理加以比较鉴定，则在作物育种中对生产力的鉴定试验就是一例。在此种情况下，供试品系是很多的，且要求各品系之间的比较应尽量以同样的精确度来进行。为此广泛应用了F.Yates（1936）创始的所谓格子法(lattice design)，其中最简单的就是这里谈的二重格子法double (simple) lattice.

在试验田里，把100个以上的品系按排在一个区组里，最困难的是土壤条件不均。现在虽把100个品种重复一次作比较，即以10个品种为一组，共设立10个区组。这样每个区组因为不包含品种的一个重复，所以叫不完全区组。只有把10个不完全区组相加,才等于100个供试品种（系）的一次重复。这样做，在同一区组内10个品种虽能相互比较，但排列于不同区组里的品种之间的差异，却与区组之间的差异掺杂在一起了。为了避免上述缺点，须将上述排列重复2次以上，即采用部分混杂的原理才能找出不同区组里的品种间差异。为了控制试验地的不均匀，把品种分组排列，积极地利用混杂原理的排列法，可称为格子方排列法。为了导入混杂原理，在设计上要考虑拟定因子，兹简述其拟定程序如下：

此处说明只限定在二重格子法里，设供试品种数v为某个整数k的平方数$v=k^2$，即设v个品种能排列成k行(横向)k列（竖向）的正方形。在这里把假设具有k水平的2个因子x・y组合起来，其形式如图1。

图1 y 水平

		1	2	3	……	k
x	1	11	12	13	……	1k
水	2	21	22	23	……	2k
平	3	31	32	33	……	3k
	⋮	⋮	⋮	⋮		⋮
	k	k_1	k_2	k_3	……	kk

k^2个因子的组合可排列成$k\times k$的格子方形。现在如果把$v=k^2$个的品种排列在这样的格子点上，在因子组合$k\times k$和k^2个品种之间便出现1比1的对应。通过这样的格子排列，便具有了因子排列的一般性质，而实际上，只不过是v个单一品种的集合体而已，即实际上是由于巧妙的把非因子组合的东西，象因子组合那样进行了处理，所以也可称拟因子（quasifactor）形式的设计，因此，格子法，也叫拟因子设计法（quasifactorial design）。

二、二重格子法的排列与分析法

1．排列法。在上述的格子排列$k\times k$当中，把分配

在横行里的品种群（具有x同一水平的品种）排列在k个区组里，叫x群。把分配在竖行里的品种群（具有y同一水平的品种）排列在k个区组里，叫y群。即在各群中有k个区组，各区组中，包含k个品种。在这些排列中，在x群的重复中，拟因子x的主效被掺杂进去，同样，在y群的重复中，拟因子y的主效也被掺杂进去，这就构成了部分混杂排列。但是交互作用xy的自由度，与x、y两群的任何一个区组都没有混杂。这就是所说的二重格子或简单格子。见图2。

图2　二重格子的排列（随机排列之前）

区组	x群			区组	y群		
1	11	12	13……1k	1	11	21	31………k1
2	21	22	23……2k	2	12	22	32………k2
3	31	32	33……3k	3	13	23	33………k3
⋮				⋮			
K	k1	k2	k3……kk	K	1k	2k	3k……kk

在设计格子法时，双因子交互作用的自由度也能混杂。它的设计法勿庸详述。总之，要准备三个以上的区组群，除主效x、y的混杂外，将交互作用xy的自由度依次混杂起来。这样，就把主效的二个混杂和交互作用的一个(重)混杂相加而形成三个混杂的排列型；同样，主效的二个混杂与交互作用的二个混杂相加便构成四个混杂的排列型。这样依次分别设计成三重格子（triple lattice），四重格子（buadruple lattice）等。这样，当交互作用全部被混杂时，针对具有某种均衡的部分混杂排列法而设计成均衡格子

(balanced lattice)。

在均衡格子法上，最初二个品种的比较精确度，无论那一个都是同样的。因而，所说的二重格子法，是二重混杂排列型，特别是做为主效x、y的设计，就二个品种的任一品种来讲，均不能说比较的精确度是一样的。关于这个问题，下面还要说到。

这里，试举一个二重格子法排列的例子，我们在实际排列的时候，可按如下程序：

（1）把供试品种k^2个随机排列，从1到k^2编上号码。

（2）把已编上号的品种排成k行、k列的正方形图，把按行分组的k个品种群随机排列到x群的区组里，同样，把按列分组的k个品种群随机排列到y群区组里。

（3）把包含在各个品种群中的k个品种，随机排列到各个区组内的k个小区里。

（4）x群、y群各重复r次。

现设k＝5，举有25个品种的二重格子方法排列的例子。为了提供参考，试做三重、四重、五重、均衡（六重）格子的排列。如图3：

图3　　5×5均衡格子的排列

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

区组	x	群				区组	y	群			
(1)	10	7	9	6	8	(1)	13	23	18	8	3
(2)	13	15	14	12	11	(2)	9	14	19	24	4
(3)	2	3	5	4	1	(3)	7	12	2	17	22
(4)	22	23	21	24	25	(4)	20	10	15	5	25
(5)	19	17	18	16	20	(5)	16	21	11	1	6

区组	第	3	群			区组	第	4	群		
(1)	11	4	10	23	17	(1)	10	1	14	22	18
(2)	19	1	7	25	13	(2)	13	21	17	5	9
(3)	15	22	9	16	3	(3)	8	4	25	16	12
(4)	24	18	12	6	5	(4)	2	15	23	6	19
(5)	20	2	14	21	8	(5)	11	24	20	3	7

区组	第	5	群			区组	第	6	群		
(1)	9	23	12	1	20	(1)	19	10	12	3	21
(2)	8	11	19	22	5	(2)	17	1	8	24	15
(3)	21	15	4	18	7	(3)	6	20	4	13	22
(4)	13	2	16	10	24	(4)	11	9	2	25	18
(5)	25	3	6	14	17	(5)	5	14	23	7	16

在图3的排列中，用x、y 2个排列的是二重格子，用第3群的3个排列是三重格子，用到第6群全部排列的是均衡格子。

在二重格子法中，因为$v=k^2$的品种，被排列在x群和y群里，所以全部供试品种最少需要重复2次；且各重复被

排列成由k个小区组成的k个不完全区组，这种排列法的特征是：被排列在这些不完全区组里的k个品种，在其余群的k个区组里，正好一个一个地出现，这就是所谓x群和y群的区组排列的相互正交。小区与区组的形状和大小，因作物种类和试验田的具体情况而不同。但在设计时，区组内各小区之间，要尽量做到均匀，在重复内各区组之间也应尽量地保持地力的一致。

2．分析法。

分析格子法的试验数据，可以拟因子设计法中部分混杂排列为指导。即为了评价品种间比较的自由度(k^2-1)，作为主效和交互效应来评价时，改写为如下拟因子：

		自由度	
主　　效	x	$k-1$	（从y群评价）
主　　效	y	$k-1$	（从x群评价）
交互作用	xy	$\dfrac{(k-1)^2}{k^2-1}$	（从x、y群评价）

此分析法中的误差项是由区组内各个小区的地力差异及试验实施中出现的偶然误差所形成，叫做区组内误差。但在只注意区组内误差的上述分析中，基于不完全区组间比较的信息，作为包括品种间效应在内的信息没有得到利用，所以不能说是精确度很高的检测。这叫区组内分析（intra-block analysis），是格子法创始初期的分析法。现正致力于不完全区组间的品种效应的校正，推算其区组内小区间的地力差异，如果可以假定是根据某种误差因子得出的话，那么，除了上述区组内误差外，作为一个误差来评价不完全区组间的变量，还是可以利用的。现决定把它叫做区组间误差。推算区组内和区组间两种误差的变量，以提高格子法的

精确度的做法，称为区组间分析(inter-block analysis)。其分析法可就下面的例题加以阐述。

三、例题

二重格子法主要用于k = 5～13的场合。如果供试品种过多或不足时，可适当加减一些品种，使成平方数。这里举一个由k^2 = 25个品种（系）组成的玉米产量鉴定试验的计算实例。此例由4次重复组成，即重复1和3是x群的随机排列，重复2和4是y群的随机排列。每群各重复2次，各由k = 5区组组成；各区组由k = 5个小区组成。每小区分2畦，每畦各栽10株。

二重格子法的小区产量

表1　　**重　复　1　(x_1)**

区组						区组和
1	2	3	4	5	1	
	42.9	41.4	48.6	38.0	39.1	210.0
2	24	22	23	25	21	
	41.1	42.0	33.1	35.2	38.9	190.3
3	19	16	20	18	17	
	30.3	37.5	36.7	39.8	42.5	186.8
4	6	10	9	7	8	
	35.7	35.3	34.0	35.5	32.0	172.5
5	13	12	14	15	11	
	33.4	32.6	38.9	31.7	34.0	170.6
						930.2

重 复 2 (y_1)

区组						区组和
1	11	16	6	1	21	
	40.8	37.2	27.0	37.1	37.2	179.3
2	8	23	3	18	13	
	44.1	31.4	34.7	39.4	33.2	182.8
3	14	24	9	19	4	
	36.8	33.9	29.8	28.5	30.2	159.2
4	22	17	7	2	12	
	35.9	37.5	35.6	37.8	34.2	181.0
5	20	25	10	15	5	
	38.8	31.8	35.3	33.5	26.9	166.3
						868.6

重 复 3 (x_2)

区组						区组和
1	10	7	9	6	8	
	43.6	40.0	41.5	42.6	41.4	209.1
2	13	15	14	12	11	
	36.4	38.9	42.0	37.0	42.2	196.5
3	2	3	5	4	1	
	37.4	37.4	38.2	37.0	40.9	190.9
4	22	23	21	24	25	
	35.4	30.6	38.0	39.2	32.3	175.5
5	19	17	18	16	20	
	32.3	34.3	36.0	37.7	42.3	182.6
						954.6

重　复　4 (y_2)

区组						区组和
1	13	23	18	8	3	
	38.4	36.7	43.8	42.1	36.8	197.8
2	9	14	19	24	4	
	37.8	24.6	30.1	35.4	37.3	165.2
3	7	12	2	17	22	
	38.8	33.5	29.5	27.4	34.5	163.7
4	20	10	15	5	25	
	34.5	34.2	27.5	41.7	30.7	168.6
5	16	21	11	1	6	
	32.0	33.9	28.4	32.4	32.5	159.2
						854.5

各　群　品　种　之　和

表2　　x 群　(重复1＋3)

	1	2	3	4	5	合计
	80.0	80.3	78.8	85.6	76.2	400.9
	6	7	8	9	10	
	78.3	75.5	73.4	75.5	78.9	381.6
	11	12	13	14	15	
	76.2	69.6	69.8	80.9	70.6	367.1
	16	17	18	19	20	
	75.2	76.8	75.8	62.6	79.0	369.4
	21	22	23	24	25	
	76.9	77.4	63.7	80.3	67.5	365.8
合计	386.6	379.6	361.5	384.9	372.2	1884.8

y 群 （重复2＋4）

	1	**6**	**11**	**16**	**21**	合计
	69.5	59.5	69.2	69.2	71.1	338.5
	2	**7**	**12**	**17**	**22**	
	67.3	74.4	67.7	64.9	70.4	344.7
	3	**8**	**13**	**18**	**23**	
	71.5	86.2	71.6	83.2	68.1	380.6
	4	**9**	**14**	**19**	**24**	
	67.5	67.6	61.4	58.6	69.3	324.4
	5	**10**	**15**	**20**	**25**	
	68.6	69.5	61.0	73.3	62.5	334.9
合计	344.4	357.2	330.9	349.2	341.4	1723.1

表3 演 算 表

y / x	1	2	3	4	5	合计 (A)	区组和 x_1	区组和 x_2	区组和 合计(B)	区组和 差	C	μC
1	1 149.5	2 147.6	3 150.3	4 153.1	5 144.8	745.3	210.0	190.9	400.9	19.1	- 56.5	- 6.3
2	6 137.8	7 149.9	8 159.6	9 143.1	10 148.4	738.8	172.5	209.1	381.6	-36.6	- 24.4	- 2.7
3	11 145.4	12 137.3	13 141.4	14 142.3	15 131.6	698.0	170.6	196.5	367.1	-25.9	- 36.2	- 4.0
4	16 144.4	17 141.7	18 159.0	19 121.2	20 152.3	718.6	186.8	182.6	369.4	4.2	- 20.2	- 2.3
5	21 148.0	22 147.8	23 131.8	24 149.6	25 130.0	707.2	190.3	175.5	365.8	14.8	- 24.4	- 2.7
合计 (A)	725.1	724.3	742.1	709.3	707.1	3607.9	930.2	954.6	1884.8	-24.4	-161.7	-18.0
区组和 y_1	179.3	181.0	182.8	159.2	166.3	868.6						
区组和 y_2	159.2	163.7	197.8	165.2	168.6	854.5						
区组和 合计(B)	338.5	344.7	380.6	324.4	334.9	1723.1						
区组和 差	20.1	17.3	- 15	- 6.0	- 2.3	14.1						
C	48.1	34.9	-19.1	60.5	37.3	161.7						
μC	5.3	3.9	-2.1	6.7	4.2	18.0						

表 1 里把进行试验的区组和小区按原随机排列，用某种单位表示小区产量。在这个表中，列有区组和、重复和；表 2 里，列有x群的 2 次重复之和和y群的 2 次重复之和。即：

x群中的品种 1 是：$39.1+40.9=80.0$

y群中的品种 1 是：$37.1+32.4=69.5$

这是品种在同一群中的和，该表同时也列有它们的行和与列和。该表虽然从变量分析的计算程序看是不必要的，但这里是为了便于以后的说明和理解而举的例子。

表 3 的数值是将表 2 中x群和y群同一品种产量相加的总和。区组和是从x_1，x_2，y_1，y_2（表 1）里抄来的。此时，请注意x的区组和在重复 1 和重复 3 里，y的区组和在重复 2 和重复 4 里被混杂起来。在区组和的差项里所指示的同一品种群的重复间差，是用于区组平方和（a）计算的。$C=A-2B$是用于计算区组平方和（b）的（见表中有*的数字——校者注）。例如：

在x水平 1 中：$C=745.3-2(400.9)=-56.5$

在y水平 1 中：$C=725.1-2(338.5)=48.1$

如果把C在x上的合计与C在y上的合计相加，其和为 0，核对时可用之（见表右下斜线——校者注）。

μC项是用于产量校正的，以后另有说明。

（1）变量分析的演算：

矫正项：

$$C.T.=\frac{(3607.9)^2}{2rk^2}=\frac{(3607.9)^2}{2\times2\times25}=130169.42$$

r是区组重复数。

总平方和S是表 1 中各小区产量的平方和减去矫正项：

$$S=(42.9^2+\cdots+34.0^2+40.8^2+\cdots+26.9^2+43.6^2$$

$$+\cdots+42.3^2+38.4^2+\cdots+32.5^2)-C.T.$$

$$=132193.05-130169.42=2023.63$$

重复平方和：

$$S_R=\frac{1}{k^2}(930.2^2+868.6^2+954.6^2+854.5^2)-C.T.$$

$$=\frac{1}{25}(3261169.41)-C.T.$$

$$=130446.78-130169.42$$

$$=277.36$$

品种平方和（不看区组效果）是从表3（编号1—25下面数字的平方相加——校者注）计算而来。

$$S_V=\frac{1}{2r}(149.5^2+147.6^2+\cdots+130.0^2)-C.T.$$

$$=\frac{1}{4}(522614.41)-C.T.$$

$$=130653.60-130169.42$$

$$=484.18$$

区组平方和（除去品种效应）由(a)、(b)组成，是从表3中区组和之差及C项（见表中X_1，X_2，Y_1，Y_2差行及C行数字——校者注），分别依下式计算的。

$$S_{B(a)}=\frac{1}{2k}[19.1^2+(-36.6)^2+\cdots+14.8^2+20.1^2+17.3^2+\cdots+(-2.3)^2]-\frac{1}{2k^2}[(-24.4)^2+14.1^2]$$

$$=\frac{1}{10}(3581.45)-\frac{1}{50}(794.17)$$

$$=358.15-15.88=342.27$$

$$S_{B(b)}=\frac{1}{2rk}[(-56.5)^2+(-24.4)^2+\cdots+(-24.4)^2$$

$$+48.1^2+34.9^2+\cdots+37.3^2]-\frac{1}{2rk^2}$$

$$[(-161.7)^2+161.7^2]$$

$$=\frac{1}{20}(15049.42)-\frac{1}{100}(52293.78)$$

$$=752.47-522.94=\underline{229.53}$$

区组内误差平方和是由总平方和减去重复、品种（不管区组效应）和区组（不管品种效应）的平方和求得（见表4变量分析表——校者注）：

$$S_E=S-(S_R+S_V+S_{B(a)}+S_{B(b)})$$

$$=2023.63-(277.36+484.18+342.27+229.53)$$

$$=690.29$$

将以上结果整理如下变量分析表，这时，区组平方和的(a)和(b)相加，用Eb表示其平均方和，叫做区组间误差变量。Ee表示误差平均方和，叫做区组内误差变量。

表4　　变量分析表

变因	自由度		平方和	平均方和
重复	$2r-1$	3	277.36	92.453
品种(未校正)	k^2-1	24	484.18	20.174
区组(校正)	$2r(k-1)$	16	571.80	35.738＝E_b
(a)	$2(r-1)(k-1)$	8	342.27	
(d)	$2(k-1)$	8	229.53	
误差(区组内)	$(k-1)(2rk-k-1)$	56	690.29	12.327＝Ee
总计	$2rk^2-1$	99	2023.63	

注：r表示群内重复数，k^2表示品种数，r＝2，k＝5。

（2）变量分析的说明：在上述的变量分析中，着眼于演算的简化，未能表明分析的结构。为了加深对格子法的理解，在此对后者（指分析结构或步骤——校者注）做一补充说明。

用其总数除观测值总和的平方，求矫正项（见前C.T.式—校者注），计算重复平方和和普通的随机区组法相同。因重复相当于完整区组，品种平方和是没有去掉区组效应的未校正数，品种平均值是在评价区组间地力差异以后校正的。此处所说的分析法，即所以把一度混杂了的区组间差异作为误差加以重新利用的方法，可以说是把重点放在较品种间比较鉴定更为重要的推算上去的方法。

前面的分析，可认为是作为不受重复内区组框框限制的随机区组分析法。因此，只剩下误差的评价了。即如果当作随机区组法分析的话，变量分析的自由度应如下分配：

变因	自由度	
重复	$2r-1$	3
品种	k^2-1	24
机误	$(2r-1)(k^2-1)$	72
总和	$2rk^2-1$	99

但是在二重格子法的分析当中，虽然属于该误差的自由度是$(2r-1)(k^2-1)=72$，但却是被分成区组间误差的自由度16和区组内误差的自由度56来评价的。在前面的计算中是作为除去了品种效应的区组间变量（变异）对待的，但在这里只能叫区组间误差。而且，这样校正后的区组间变