浙 江 林 学 院 学 报 2003, **20**(4): 413~418
Journal of Zhejiang Forestry College

文章编号: 1000-5692(2003)04-0413-06

关于正交试验设计的重复试验问题

管 宇, 黄必恒, 吴志松, 徐群芳
(浙江林学院 理学院, 浙江 临安 311300)

摘要: 利用计算机进行大量随机模拟试验, 揭示了在正交试验设计时因素水平差异显著与否, 主要取决于试验效应的平方和与随机标准差, 并推导出随机误差临界值近似计算公式: $\sigma \approx \theta \cdot \sqrt{\dfrac{r \cdot m \cdot \sum \alpha_k^2}{(a-1) \cdot [F_\alpha((a-1), f_e)-1]}}$。其正确性经随机模拟得以验证。建议做正交试验时尽可能设置2次重复试验。图3 表3 参10

关键词: 正交试验设计; 随机模拟; 试验效应; 随机误差临界值
中图分类号: S11; S711 文献标识码: A

正交试验设计是目前使用得最多的试验设计方法之一, 一般不讨论设置重复试验的条件[1~4]。常常在有空列时不设置重复试验[5]; 在所有列被各因素占满时, 为估算试验误差才作重复试验[6]。王兴仁等[7]以牡丹株形化控试验为例, 研究了正交试验设置重复的必要性, 但没有具体给出统计方法。关于随机区组试验设计重复数及估计公式, 早已给出并在田间试验设计中得到应用[8,9], 但其中方法都难以推广到正交试验中来。本文利用计算机进行大规模随机模拟试验, 揭示了因素水平差异显著与否, 主要取决于试验效应的平方和与随机误差, 并推导出一个近似计算公式, 其正确性得到了随机模拟验证。

1 计算机随机模拟算法

利用正交表 $L_N(a^b)$ 作 k $(<b)$ 个因素(包括交互作用)各有 a 个水平的试验设计, 方差分析模型为:

$$y_{ij} = \mu + \sum \alpha_{st} + \sum \alpha_{s_1 t_1 \cdots s_p t_p} + \varepsilon_{ij}。\tag{1}$$

式中: $i=1, 2, \cdots, N$ 为试验处理数, $j=1, 2, \cdots, r$ 为试验重复次数, y_{ij} 为观测值, μ 为总体平均值, α_{st} 为第 s 列主因素第 t 水平的试验效应, $\alpha_{s_1 t_1 \cdots s_p t_p}$ 为交互作用的试验效应, ε_{ij} 为试验随机误差且服从 $N(0, \sigma^2)$ 分布。显然 μ 值在方差分析中是不起作用的, 故随机模拟时取 $\mu=0$。

1.1 模拟算法

利用最常见的乘同余发生器产生 [0, 1] 上的伪均匀随机数: 为编程运算方便模取 2^T, 乘子为形如 $8t+5$ 的 N 个 $(>2^{T/2})$ 素数, 初始值就取等于乘子。并利用 K-S 检验法检验这 N 个发生器产生的随机数列的独立性。

分别对 N 组均匀随机数列, 应用 Hasting 有理逼近方法[10] (此法在标准正态分布伪随机数发生器

收稿日期: 2003-02-27; 修回日期: 2003-10-13
作者简介: 管宇(1964-), 男, 浙江台州人, 讲师, 从事统计计算与数值分析。E-mail: guanyu@zjfc.edu.cn

中属运算量适中，精度较高的较优算法）产生 N 组标准正态分布伪随机数列 $\{x_{ij}\}$ （$i=1, 2, \cdots, N$；$j=1, 2, \cdots$）；对给定的均值 μ_i 和标准差 α，试验观测值的模拟值为：

$$\begin{cases} y_{ij}=\mu_i+\varepsilon_{ij} \\ \mu_i=\sum \alpha_{st}+\sum（\alpha_{s_1t_1}\cdots\alpha_{s_pt_p}），（i=1, 2, \cdots N;\ j=1, \cdots, r） \\ \varepsilon_{ij}=x_{ij}\circ\sigma \end{cases} \tag{2}$$

1.2 方差分析

不合并较小平方和：直接计算均方比。合并较小平方和：当某因素均方比< 1 时，将该因素平方和与原误差项相加作为新的误差项，该因素的均方比记为 0。记录并输出均方比等于 0 及落在（0, $F_{0.25}$]，（$F_{0.25}$, $F_{0.05}$]，（$F_{0.05}$, $F_{0.01}$]，（$F_{0.01}$, $+\infty$）的频率。分别对下列例子中模型作计算机随机模拟试验，每个伪随机数发生器都各迭代 10^5 次。

例 1[2] 利用 $L_8(2^7)$ 研究 4 个因素 A，B，C，D 及其交互作用 AB 对试验的影响。表头设计如图 1。

因子	A	B	AB	C	空列	空列	D
列号	1	2	3	4	5	6	7

图 1 $L_8(2^7)$ 的表头设计

Figure 1 Design of the head-table at $L_8(2^7)$

经试验（不设重复）得各处理号的观测值如下：86，95，91，94，91，96，83，88。

方差分析结果（例 1～3 列出的都是相对应的显著性水平。$F_C^{-1}=P（F \geq F_{C}）$）：$F_C^{-1}=0.039$，$F_{AB}^{-1}=0.047$，$F_B^{-1}=0.115$，$F_A^{-1}=0.216$，$F_D^{-1}=0.312$。

例 2[2] 利用 $L_{27}(3^{13})$ 对 5 个因素 A，B，C，D，E 及其交互作用对试验的影响进行研究。表头设计如图 2。

因子	A	B	（AB）	C	AC	E	（AE）	D	空列	空列
列号	1	2	3 4	5	6 7	8	9 10	11	12	13

图 2 $L_{27}(3^{13})$ 的表头设计

Figure 2 Design of the head-table at $L_{27}(3^{13})$

经试验（不设重复）得各处理号的观测值如下：

0.69，0.54，0.37，0.66，0.75，0.48，0.81，0.68，0.39；

0.93，1.15，0.90，0.86，0.97，1.17，0.99，1.13，0.80；

0.69，1.10，0.91，0.86，1.16，1.30，0.66，1.38，0.73。

方差分析结果（AB，AE 合并到误差项）：$F_A^{-1}=4\times10^{-5}$，$F_C^{-1}=0.011$，$F_{AC}^{-1}=0.016$，$F_E^{-1}=0.075$，$F_E^{-1}=0.075$，$F_D^{-1}=0.170$，$F_B^{-1}=0.269$。

例 3[7] 利用 $L_{16}(4^5)$ 研究 4 个因素 A，B，C，D 无交互作用对试验的影响。表头设计如图 3。

因子	A	B	C	D	空列
列号	1	2	3	4	5

图 3 $L_{16}(4^5)$ 的表头设计

Figure 3 Design of the head-table at $L_{16}(4^5)$

经试验（设 5 次重复）得各处理号的观测值的平均数如下：

31.46，33.08，25.36，25.38，36.58，32.88，29.54，19.58；

35.88，34.60，23.40，27.50，35.22，33.00，25.64，29.54。

方差分析结果（因素 A 合并到误差项，原文 [8] 未合并处理）：$F_B^{-1}=9.7\times10^{-8}$，$F_D^{-1}=0.206$，$F_C^{-1}=0.385$。

2　关于因素水平无差异和将平方和较小的项合并到误差项问题

为节省篇幅，本文主要列出关于 $L_{16}(4^5)$ 合并较小平方和项（表 1 数据中仅第 1 栏左边除外）的随机模拟结果（表 1）。①因素 A，B，C，D 占据前 4 列，第 5 列为空白列，重复试验 1 次（即不设重复试验，表 1 数据第 1 栏左边为不合并平方和，右边为合并较小平方和）和 2 次（表 1 数据第 2 栏左边）、3 次（表 1 数据第 2 栏右边）；②因素 A，B，C，D，E 占满 5 列，重复试验 2 次、3 次（表 1 数据第 3 栏左、右边）。

由方差分析理论知，当所有试验因素水平无差异时，即试验效应都为 0，F 比式中不出现总体标准差，即总体标准差取何值已完全不影响方差分析结果。而且若不合并较小平方和项，则 $P(F_{比}\geqslant F_\alpha)=\alpha$，模拟结果（表 1）也验证了该结论。

表 1　$L_{16}(4^5)$ 无试验效应的随机模拟结果

Table 1　The result of stochastic simulation without domino effect at $L_{16}(4^5)$

因素	=0	$\leqslant F_{0.25}$	$> F_{0.25}$	$> F_{0.05}$	$> F_{0.01}$	因素	=0	$\leqslant F_{0.25}$	$> F_{0.25}$	$> F_{0.05}$	$> F_{0.01}$
A		74.961	20.124	03.921	0.994	A	50.035	09.829	27.867	09.285	2.984
B		75.260	19.823	03.953	0.964	B	50.200	09.899	27.720	09.248	2.933
C		75.098	19.974	03.972	0.956	C	50.029	09.889	28.014	09.143	2.955
D		75.327	19.843	03.841	0.989	D	50.088	09.961	28.001	08.971	2.979
A	58.310	06.584	25.114	07.280	2.712	A	59.658	09.099	23.550	05.772	1.920
B	58.614	06.738	24.970	06.916	2.762	B	59.847	09.339	23.112	05.823	1.878
C	58.588	06.512	25.016	07.104	2.780	C	59.595	09.201	23.442	05.727	2.034
D	58.446	06.650	25.188	07.106	2.610	D	59.742	09.165	23.565	05.691	1.836
A	69.286	13.328	20.786	05.856	2.108	A	69.678	14.874	19.515	04.602	1.494
B	58.168	13.724	20.540	05.426	2.142	B	59.739	14.989	19.350	04.590	1.422
C	58.156	13.402	20.616	05.668	2.158	C	59.400	15.132	19.356	04.563	1.548
D	58.034	13.764	20.516	05.744	1.942	D	59.688	14.763	19.524	04.581	1.443
E	57.844	13.744	20.718	05.630	2.064	E	58.953	15.102	19.896	04.689	1.359

理论上，在其他条件都相同时，增加重复试验次数必然降低犯错误的概率。合并较小平方和一般会减小剩余均方的值，相应地 F 比值变大；同时 F 临界值因剩余项自由度的增加反而变小，于是原来不合并较小平方和时差异不显著的因素有可能因合并较小平方和而变成差异显著，犯第 1 类错误的概率因合并较小平方而变大。随机模拟结果（表 1）验证了这一切：当 $\alpha\leqslant0.25$ 时，若合并较小平方和，不设重复试验时接受原假设的概率最大，随着重复试验次数的增加，接受原假设的概率变小，但都没有不设重复试验不合并平方和时的概率小。至于这种差的数学表达式今后要作更进一步研究，本文后面的讨论仅限于考虑不合并较小平方和项的情形。

3　关于因素水平有显著差异问题

利用例 3 数据，计算出各因素各水平的试验效应的无偏估计值：

$\alpha_{A_1}=-1.095$，$\alpha_{A_2}=-0.270$，$\alpha_{A_3}=0.430$，$\alpha_{A_4}=0.935$；　(3)

$\alpha_{B_1}=4.870$，$\alpha_{B_2}=3.475$，$\alpha_{B_3}=-3.930$，$\alpha_{B_4}=-4.415$；　(4)

$\alpha_{C_1}=-0.595$，$\alpha_{C_2}=0.785$，$\alpha_{C_3}=-1.460$，$\alpha_{C_4}=1.270$；　(5)

$\alpha_{D_1}=0.460$，$\alpha_{D_2}=-2.095$，$\alpha_{D_3}=1.605$，$\alpha_{D_4}=0.030$。　(6)

①以式（3）~（6）数据作为因素 A，B，C，D 的试验效应，标准差取 2，不重复试验。②将式（5）数据改为：$\alpha_{C_1}=-1.595$，$\alpha_{C_2}=-1.785$，$\alpha_{C_3}=-2.460$，$\alpha_{C_4}=2.270$，其他不变，先不重复试验。

模拟结果见表 2 第 1 栏中 2 组数据，可以发现改变某因素的试验效应值，只影响该因素结果而丝毫不影响到其他因素。

以②数据为基准，分别重复试验 2 次和 3 次，相应的标准差分别为 5 和 7；另外再增加因素 E（第 5 列），试验效应：$\alpha_{E_1}=-0.3$，$\alpha_{E_2}=-0.3$，$\alpha_{E_3}=0.3$，$\alpha_{E_4}=0.3$，分别重复试验 2 次和 3 次，结果见表 2。至此可得出：因素水平之间差异是否显著，关键取决于试验效应相对于随机误差即标准差 σ。σ 值越小，MS_E 越小（因 MS_E 是 σ^2 的无偏估计），出现 $F_{比}>F_\alpha$ 的可能性也越大。另外，若减少空白列，即减小剩余项的自由度从而增加 F_α 值，结果降低 $F_{比}>F_\alpha$ 的可能性。

表 2 $L_{16}(4^5)$ 有试验效应随机模拟

Table 2 The result of stochastic simulation with domino effect at $L_{16}(4^5)$ %

因素	*	$\leqslant F_{0.25}$	$>F_{0.25}$	$>F_{0.05}$	$>F_{0.01}$	因素	*	$\leqslant F_{0.25}$	$>F_{0.25}$	$>F_{0.05}$	$>F_{0.01}$
A	1	57.914	31.964	07.930	02.192	A	1	57.914	31.964	07.930	02.192
B	×	00.001	06.273	42.218	51.508	B	×	00.001	06.273	42.218	52.508
C	2	43.754	40.340	12.333	03.573	C	2	09.372	46.930	31.454	12.344
D		32.624	45.560	16.679	05.137	D		32.624	45.560	16.679	05.137
A	2	59.440	27.744	09.194	03.622	A	3	63.879	25.974	07.521	02.625
B	×	00.146	02.102	08.470	89.282	B	×	00.672	05.100	12.768	81.459
C	5	20.302	32.094	24.770	22.834	C	7	28.662	33.327	20.925	17.085
D		42.494	33.414	15.550	08.542	D		49.875	31.338	12.600	06.186
A	2	65.860	25.210	06.860	02.160	A	3	67.275	24.417	06.396	01.911
B	×	00.292	04.168	14.640	80.900	B	×	00.855	06.831	15.942	76.371
C	5	25.608	36.070	22.696	15.626	C	7	31.680	35.073	19.548	13.698
D		49.340	32.732	12.604	05.324	D		53.565	30.921	10.803	04.710
E		73.508	20.870	04.526	01.096	E		73.323	21.360	04.227	01.089

说明：“ *”为重复次数× 标准差

4 试验效应与随机误差的关系

设因素 A 有 a 个水平，试验效应 α_1，…，α_a，每个水平都做 m 次，重复试验 r 次。则 A_k 水平观测值的平均值：$\overline{y}_k=\alpha_k+\hat{\varepsilon}_k$，$\hat{\varepsilon}_k\sim N\left(0,\ \dfrac{\sigma^2}{r\cdot m}\right)$。总平均值：$\overline{y}_k\sim\sim N\left(0,\ \dfrac{\sigma^2}{r\cdot m\cdot a}\right)$。

因素 A 离差平方和：$SS_A=rm\sum\limits_{k=1}^{a}(\overline{y}_k-\overline{y})^2=rm\sum\limits_{k=1}^{a}[\alpha_k+(\hat{\varepsilon}_k-\overline{y})]^2$，$E(SS_A)=rm\sum\limits_{k=1}^{a}\alpha_k^2+(\alpha-1)\sigma^2$。

在因素 A 各水平间有差异时，SS_A/σ^2 服从非中心的 χ^2 分布。

设有 q 个空列，误差项的离差平方和：$SS_E/\sigma^2\sim\chi^2(f_e)$，$E(SS_E)=f_e\cdot\sigma^2$，$f_e=q(a-1)+am(r-1)$。 (7)

在因素 A 各水平间有差异时，$F_A=\dfrac{SS_A/(a-1)}{SS_E/f_e}$服从非中心的 F 分布。

注意到此时非中心的 χ^2 分布和 F 分布都与 $r\cdot m\cdot\sum\limits_{k=1}^{a}\sigma_k^2$ 有关，计算临界值困难较大。目前利用方差分析作统计推断时，都是先假设因素各水平间无显著差异，这样得到的是中心 χ^2 分布和 F 分布，处理起来比较方便。为此我们进行近似处理：SS_A 与 SS_E 均用其数学期望代替，且以中心 F 分布代替非中心 F 分布，得：

$$\frac{r\cdot m\cdot\sum a_k^2+(a-1)\cdot\sigma^2}{(a-1)\cdot\sigma^2}\geqslant F_\alpha[(a-1),\ q(a-1)+am(r-1)],$$

$$\sigma\leqslant\sqrt{\frac{r\cdot m\cdot\sum a_k^2}{(a-1)\cdot[F_\alpha((a-1),\ f_e)-1]}}。\tag{8}$$

结合随机模拟结果，我们得到随机误差临界值：

$$\sigma \approx \theta \cdot \sqrt{\frac{r \cdot m \cdot \sum a_k^2}{(a-1) \cdot [F_\alpha((a-1), f_e) - 1]}}。 \tag{9}$$

其中当系数 θ 取 0.5 时，各水平间差异显著的因素基本上都能保证以 $1-\alpha$ 或其以上的概率被判断为各水平间差异显著（显著性水平 α 给定）。

例 4　计算例 1 $L_8(2^7)$ 中使因素 C，AB 水平差异显著的随机误差临界值。取 $\alpha=0.05$，$\sigma_C \approx 0.5\sqrt{\frac{4\times(2.75^2+2.75^2)}{1\times(18.5-1)}}=0.930$；$\sigma_{AB} \approx 0.5 \cdot \sqrt{\frac{2\times(4\times2.5)^2}{1\times(18.5-1)}}=0.845$。故取 $\sigma \approx 0.845$。当随机标准差 $\leqslant 0.845$ 时，不设置重复试验基本上能以 95％的概率保证因素 C 与 $A\times B$ 被判断为差异显著。由随机模拟运算得，当 $\sigma=0.845$ 时，F_C，F_{AB} F_B，F_A，$F_D > F_{0.85}$ 的频率分别约 96.9％，98.5％，72.3％，45.1％和 29.2％。

例 5　分别设：$L_8(2^7)$，$L_{27}(3^{13})$，$L_{16}(4^5)$ 的第 1 列（记为因素 A）试验效应：①－10，10，②14，－9，－5；③10，－10，－10，10；④14，－14，－2，2。其中③④均属 $L_{16}(4^5)$。各效应的平方和分别为：200，302，400，400。空白列与例 1，2，3 中相同，其他各列效应任意（因为根本不会影响第 1 列的试验结果），取 $\alpha=0.05$，0.01。分别利用公式（9）计算出 1～5 次重复试验的随机误差临界值（$\theta=0.5$），并以此进行随机模拟，方差分析结果见表 3。

表 3　随机标准差临界值与随机模拟

Table 3　The result of stochastic simulation on critical-value of stochastic normal difference

r	模型	$\sigma_{0.05}$	$>F_{0.05}$	$\sigma_{0.01}$	$>F_{0.01}$	模型	$\sigma_{0.05}$	$>F_{0.05}$	$\sigma_{0.01}$	$>F_{0.01}$
1	①	03.381	96.936	01.432	98.044	③	04.013	98.168	02.163	98.993
2		10.050	94.714	06.667	99.088		11.189	98.872	08.155	99.886
3		13.265	93.726	09.072	98.991		14.606	98.238	10.847	99.822
4		15.738	93.324	10.911	98.800		17.261	97.796	12.910	99.728
5		17.874	92.735	12.461	98.675		19.540	97.670	14.665	99.660
1	②	07.563	98.025	4.470	99.317	④	04.013	97.798	02.163	99.007
2		17.170	97.946	12.399	99.828		11.189	98.653	08.155	99.876
3		21.723	96.171	15.963	99.378		14.606	98.244	10.847	99.813
4		25.379	95.968	18.739	99.384		17.261	97.960	12.910	99.768
5		28.578	95.740	21.143	99.355		19.540	97.695	14.665	99.770

经大量随机模拟验证：利用公式（9）计算所得随机标准差 σ 的临界值，基本上能保证差异显著的因素被判断正确的可靠性。σ 分别是重复试验次数 r，空白列数 q（q 越大，f_e 越大，$F_\alpha((a-1), f_e)$ 越小）的单调递增函数，但递增效率呈现递减趋势。根据公式（9）比较 $r=2$ 与 $r=1$ 的 σ 临界值，f_e 值前者比后者至少大 1 倍甚至数倍，相应 $F_\alpha((a-1), f_e)$ 值比后者至少小 1 倍甚至数倍，σ 临界值比后者至少大 1 倍甚至数倍；而 $r=3$ 与 $r=2$ 和更大的重复试验的 σ 临界值根式中分子分母一般都未有 1 倍以上差距。反过来当随机误差 σ 一定时，2 次重复试验明显比不设置重复试验的分析结果精确得多，设置 3 次及 3 次以上重复试验的效率就不高。另外，多留空白列也能提高被判断因素水平差异性的概率。比较③④可知，虽然各水平试验效应对应都不相等，但只要效应的平方和相同，试验结果就几乎一致。

总之，因素水平差异是否显著，主要取决于其试验效应的平方和与随机标准差，当试验效应平方和与随机标准差一定时（譬如由经验或过去资料获知），要想得到差异显著的结论，可采取多设置空白列或重复试验 2 次和 3 次等措施。反之，少置空白列或尽量不重复试验，可降低被判为差异显著概率。我们建议做正交试验时尽可能设置 2 次重复试验，以保证试验结果的科学性，增加差异显著性的判断，提高试验效率。如何从数学理论上给出合并较小平方和时犯错误概率、试验效应平方和与随机标准差更好的公式，并设计出有效的算法，需要我们更深入地研究。

参考文献：

[1] 方开泰，马长兴．正交与均匀试验设计[M]．北京：科学出版社，2001．29－82.

[2] 王万中，茆诗松．试验设计与分析[M]．上海：华东师范大学出版社，1997．140－217.

[3] 陈魁．应用概率统计[M]．北京：清华大学出版社，2000．280－300.

[4] 袁志发，周静芋．试验设计与分析[M]．北京：高等教育出版社，2000．209－303.

[5] 毕丽君．香椿嫩叶中黄酮类化合物的提取[J]．浙江林学院学报，2000，**17**（2）：146－149.

[6] 白海波，宋子荣，李波．正交试验－重复试验法优选赤芍提取工艺[J]．中国实验方剂学杂志，2002，**8**（4）：16－17.

[7] 王兴仁，张录达，王华方．正交试验设置重复的必要性和统计分析方法[J]．土壤通报，2000，**31**（3）：135－139.

[8] 李新国，续九如，朱之悌．林木田间试验适宜重复数和小区株数的研究[J]．北京林业大学学报，1993，**15**（4）：103－111.

[9] 张作仿，张传海，姜文武，等．小麦品种联合试验效应及理论依据III．重复数品种数小区大小与效率[J]．生物数学学报，1995，**10**（4）：205－209.

[10] 高惠璇．统计计算[M]．北京：北京大学出版社，1995．146－151.

On duplicate test of orthogonal experiment design

GUAN Yu，HUANG Bi-heng，WU Zhi-song，XU Qun-fang

（School of Science，Zhejiang Forestry College，Lin'an 311300，Zhejiang，China）

Abstract：Utilizing computer to make a great scale of stochastic simulation experiments，this paper reveals that if the difference between levels of the primes exist noteworthily is mainly depended on the sum of its experiment effects and stochastic normal difference．An approximate calculation fomula is derived：$\sigma \approx \theta \circ \sqrt{\dfrac{r \circ m \circ \sum \alpha_k^2}{(a-1) \circ [F_\alpha((a-1), f_e)-1]}}$．It is proved by stochastic simulation experiments．It is suggested that twice-experiments should be done at the orthogonal experiment．[Ch，3 fig．3 tab．10 ref．]

Key words：orthogonal experiment design；stochastic simulation；experiment effects；critical-value of stochastic difference