# 冶金工程试验统计

中国金属学会　冶金继续工程教育丛书

冶金工业出版社

朱伟勇　傅连魁　编著

# 序

中国金属学会组织编写了《冶金继续工程教育丛书》，为大家办了一件好事。积极开展继续教育，对于提高冶金科技人员水平，促进冶金工业的发展具有重要意义。希望冶金战线各级领导重视这项工作，努力创造条件，为科技人员在职学习提供方便；同时也殷切希望广大冶金科技工作者坚持学习，不断吸收新知识，学习新技术，为实现四化，振兴中华做出更大贡献。

中国继续工程教育协会理事<br>冶金工业部副部长　徐大铨

一九八八年十二月

# 前　言

本书为冶金部中国金属学会组织编写的《冶金继续工程教育丛书》之一。出版该书的目的在于普及新的应用统计方法，提高解决实际问题的能力。书中介绍的“具有对数项的混料模型及其$D$-最优测度设计、$D_n$-最优确切设计”等内容，曾分别在1984年和1985年的美国第144届和第145届统计数学年会及1986年及 1989 年第2届和第 3 届日中统计讨论会上发表，并在《美国统计计算进展》等刊物上刊登。此外，还介绍了Kiefer(美国科学院院士）的$D$-最优设计、Box的旋转设计、Scheffe′的混料设计以及Draper的倒数项设计等，这些方法都是目前国内外应用效果较好的新的统计方法。它们在新材料、新产品的开发研究，在钢铁冶金专业领域建立数学模型、进行工艺研究和配方配比试验以及质量控制等方面的应用，都取得了明显的社会效益和经济效益。

为了便于读者使用，本书列举了冶金专业等领域内的应用实例30多个，通过实例由浅入深地分析和计算回归系数，介绍方差分析、极值判别、区间估计、获取最优工艺条件的线性与非线性规划的优化方法等等。其中，有些应用成果曾获冶金部科技成果二等奖及第35届世界博览会尤里卡银奖等，还有一些实例曾以论文形式分别在《应用数学学报》、《应用概率统计》、《数学的实践与认识》和《东北工学院学报》等刊物上发表。

本书可作为冶金系统的试验研究人员、工程技术人员和管理人员继续工程教育用书，也可作为冶金高等院校师生讲授和学习应用统计课程的教学参考书。

本书在编写过程中，曾先后得到国内外一些著名统计学家和教授的热情支持和帮助。在此谨向华东师范大学数理统计系魏宗舒教授、茆诗松教授，中国现场统计研究会理事长张里千教授和副理事长林少宫教授，中国科技大学数学系陈希孺教授，武汉大学统计系主任张尧庭教授，中国科学院系统研究所所长成平研究员和项可风研究员，鞍山钢铁研究所袁金海高级工程师，东北工学院金属陶瓷研究所所长李荣久副教授、东北工学院金属材料系姚祯等，以及美国北伊利诺斯大学美籍华人陈惠森教授，佛罗里达大学J.Cornell教授表示衷心的感谢！

本书的编写指导思想是理论与实际相结合，通过应用实例，由浅入深地介绍理论和方法，略去不必要的公式推导，力求使读者学习后会分析、会计算、会应用。但由于编者水平所限，有不妥之处，恳请读者批评指正。

编　者

1990年4月

# 目　　录

## 3 回归正交设计

# 1

# 正交试验设计

在工农业生产及科研工作中，经常会遇到多因素试验问题。在做试验时，若试验方案设计得好，则只要做少量试验就可得到满意结果；反之，试验次数就多，结果还不一定好。因此，要采用一种好的设计方法。正交设计法就是安排多因素试验的一种科学方法。该方法具有取点均衡、整齐可比、试验点代表性强、试验次数少、便于分析计算等优点，有重要的实用价值。

使用正交设计方法时，要按正交表来安排试验。下面首先介绍一些构造正交表的简单方法，然后介绍正交表试验设计和分析。

## 1.1 正交表的构造

### 1.1.1 以素数（3，5，7，11等）为水平数的正交表

凡是素数（3，5，7，11等）都可用拉丁方构造正交拉丁方，而由正交拉丁方形成正交表。如果把上述的素数 $n$ 作为水平数，则以 $n$ 为水平数的拉丁方就有$n-1$个。把$n-1$个拉丁方的相对应的元素组合在一起便构成了正交拉丁方。

例如，以素数 3 作为水平数，3 水平的拉丁方有$3-1=$

2个（表1-1）。把这两个拉丁方对应元素组合在一起，便构成了正交拉丁方（表1-2）。

表 1-1 拉 丁 方 表

| | | |
|---|---|---|
| 1 | 2 | 3 |
| 2 | 3 | 1 |
| 3 | 1 | 2 |

| | | |
|---|---|---|
| 1 | 2 | 3 |
| 3 | 1 | 2 |
| 2 | 3 | 1 |

表 1-2 正 交 拉 丁 方

| | | |
|---|---|---|
| 11 | 22 | 33 |
| 23 | 31 | 12 |
| 32 | 13 | 21 |

表 1-3 因素 $C$、$D$ 的水平表

| $A$ 的 水 平 | $B$ 的 水 平 | | |
|---|---|---|---|
| | 1 | 2 | 3 |
| 1 | 11 | 22 | 33 |
| 2 | 23 | 31 | 12 |
| 3 | 32 | 13 | 21 |

表 1-3a $L_9(3^4)$

| № | 因 | | 素 | |
|---|---|---|---|---|
| | $A$ | $B$ | $C$ | $D$ |
| 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 |

把正交拉丁方中的前列作为因素$C$，把后列作为因素$D$，然后用因素$A$和因素$B$与因素$C$和因素$D$均匀搭配，便构成了三水平四因素九次试验（即四列九行三水平）的正交表（表1-3、表1-3$a$）。

表1-3$a$记作$L_9(3^4)$，其含义为：

$L_9(3^4)$

→正交表有四列，每一列可排一个因素。最多可排四个因素。

→每个因素有三个水平。

→正交表有九行，每一行即为一组试验，共九组试验。

→正交表的开头字母。

由此可知素数5有（5-1）个拉丁方，素数7有（7-1）个拉丁方，……。由上述方法就可以得到以素数为水平数的正交表$L_{25}(5^6)$、$L_{49}(7^8)$，……。

## 1.1.2 二水平正交表

将Hadamard矩阵（$H_2$）用直积方法可得到二水平的正交表[24]。方法如下：

$$H_2=\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}$$

将$H_2$与$H_2$进行直积运算：

$$H_2\otimes H_2=\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}\otimes\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}$$

$$=\begin{bmatrix}1\otimes\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix} & 1\otimes\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}\\ 1\otimes\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix} & -1\otimes\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}\end{bmatrix}$$

$$=\begin{bmatrix}1 & 1 & 1 & 1\\ 1 & -1 & 1 & -1\\ 1 & 1 & -1 & -1\\ 1 & -1 & -1 & 1\end{bmatrix}$$

$H_2$与$H_2$直积运算后得出矩阵，记为$H_4$，将$H_4$的第一列去掉便得到正交表$L_4(2^3)$。

$H_2$与$H_4$再进行直积运算得出矩阵，记为$H_8$。将$H_8$的第一列去掉便得到正交表$L_8(2^7)$。

由此，对$H_2$与$H_8$、$H_2$与$H_{16}$……求直积，将结果矩阵的第一列去掉，便可得到正交表$L_{16}(2^{15})$、$L_{32}(2^{31})$……。

在二水平正交表中，也可将1作为水平1，−1作为水平2。

### 1.1.3 拟水平正交表

在参加试验的因素中，有的因素的水平数少于正交表中给出的水平数。这样的因素称为拟水平因素。

例如：在正交表$L_9(3^4)$中放入一个拟水平因素，其水平数是2。如果把拟水平因素放在正交表$L_9(3^4)$的第三列

表 1-4 $L_9(2\times3^3)$

| № | 因素 | | | |
|---|---|---|---|---|
| | A | B | C | D |
| 1 | 1 | 1 | 1 | 1 |
| 2 | 1 | 2 | 1 | 2 |
| 3 | 1 | 3 | 2 | 3 |
| 4 | 2 | 1 | 1 | 3 |
| 5 | 2 | 2 | 2 | 1 |
| 6 | 2 | 3 | 1 | 2 |
| 7 | 3 | 1 | 2 | 2 |
| 8 | 3 | 2 | 1 | 3 |
| 9 | 3 | 3 | 1 | 1 |

上，可根据下面方法重新安排：

（a）拟水平因素的水平数1放在正交表中水平1与水平2的位置上。

（b）拟水平因素的水平数2放在正交表中水平3的位置上。

这样便得到新的正交表$L_9(2\times3^3)$，见表1-4。

由上述方法可以得到正交表$L_9(2^2\times3^2)$、$L_9(2^3\times3)$、……。

### 1.1.4 拟因子正交表

参加试验的因素中，有的因素的水平数大于正交表中所给出的水平数。这样的因素叫拟因子。

1.1.4.1 拟因子为三水平的正交表

例如：在二水平正交表$L_{16}(2^{15})$中安排拟因子，拟因子的水平数为3。其安排方法是，选择正交表$L_{16}(2^{15})$中的两列，如果两列满足：

（a）各自的水平数为（1，1），则可存放拟因子的水平数2。

（b）各自的水平数为（2，2），则可存放拟因子的水平数2。

（c）各自的水平数为（1，2），则可存放拟因子的水平数2。

（d）各自的水平数为（2，1），则可存放拟因子的水平数3。

在正交表中安排三水平的拟因子，第一列不能与其它列合并作为拟因子列。它是赋闲列。

在正交表$L_{16}(2^{15})$中安排一个三水平的拟因子，选择2列与3列合并，这样便得到新的正交表$L_{16}(3\times2^{13})$。

如果继续将4列与5列，8列与9列，14列与15列合并，便可分别得到正交表$L_{16}(3^2\times2^{11})$，$L_{16}(3^3\times2^9)$，$L_{16}(3^4\times2^7)$。

1.1.4.2 拟因子为四水平的正交表

例如：在二水平正交表$L_{16}(2^{15})$中安排拟因子，拟因子的水平数为4。其安排方法是，选择正交表$L_{16}(2^{15})$中的三列进行合并，如果三列满足：

（a）各自的水平数为（1，1，1），则可存放拟因子的水平数1。

（b）各自的水平数为（1，2，2），则可存放拟因子的水平数2。

（c）各自的水平数为（2，1，2），则可存放拟因子的水平数3。

（d）各自的水平数为（2，2，1），则可存放拟因子的水平数4。

在二水平正交表$L_{16}(2^{15})$中安排一个四水平的拟因子，选择1列、2列、3列合并，便得到新的正交表$L_{16}(4\times2^{12})$。

如果继续将4列、8列、12列合并，5列、10列、15列合并，7列、9列、14列合并，6列、11列、13列合并，便可分别得到正交表$L_{16}(4^2\times2^9)$、$L_{16}(4^3\times2^6)$、$L_{16}(4^4\times2^3)$、$L_{16}(4^5)$。

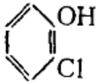
由此也可把其他二水平正交表合并成新的能够安排拟因子的正交表。

## 1.2 正交试验设计的分析

### 1.2.1 正交试验方案的设计

在试验项目和考察指标确定后，可根据经验及有关专业

知识找出对指标有影响的各种因素，然后选择每个因素的取值范围，制定因素水平表。

**例1.1** 某制药厂生产黄连素试验[20]，考察指标是回收率（%），参加试验的因素有（$A$）：NaOH；（$B$）温度（℃）；（$C$）时间（min）；（$D$）催化剂用量（%）。每个因素所取的水平如下：

因素$A$：$A_1=1:2.6$，$A_2=1:2.8$，$A_3=1:2.4$。

因素$B$：$B_1=240$℃，$B_2=225$℃，$B_3=210$℃。

因素$C$：$C_1=60$min，$C_2=40$min，$C_3=30$min。

因素$D$：$D_1=5\%$，$D_2=3\%$，$D_3=1\%$。

首先确定因素水平表（见表1-5）。

**表 1-5 因 素 水 平 表**

| № | 因素 | | | |
|---|---|---|---|---|
| | $A$ | $B$ | $C$ | $D$ |
| 1 | 1:2.6 | 240 | 60 | 5 |
| 2 | 1:2.8 | 225 | 40 | 3 |
| 3 | 1:2.4 | 210 | 30 | 1 |

然后选择合适的正交表来安排试验。选择正交表要满足两个条件：(a)正交表的行数$n \geqslant 1+\sum f_{因}$。$f_{因}$是因素的自由度，$\sum f_{因}$是各个因素自由度之和。(b)正交表的列数要大于或等于因素的个数，同时还要满足每个因素的水平数的要求。

选择正交表$L_9(3^4)$。因为每个因素是三水平，它们的自由度是：$f_A=f_B=f_C=f_D=2$。正交表的行数$n=1+f_A+f_B+f_C+f_D=9$。正交表的列数$m(=4)$与因素个数相同且每一列满足因素的水平数要求。

将正交表中的每一列对应于一个因素。将因素的每个水平值对应地排在表中，便构成四因素三水平的9组试验方案（见表1-6）。

**表1-6 试 验 方 案**

| № | 因素 | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | A | | B | | C | | D | |
| 1 | 1 | (1:2.6) | 1 | (240) | 1 | (60) | 1 | (5) |
| 2 | 1 | (1:2.6) | 2 | (225) | 2 | (40) | 2 | (3) |
| 3 | 1 | (1:2.6) | 3 | (210) | 3 | (30) | 3 | (1) |
| 4 | 2 | (1:2.8) | 1 | (240) | 2 | (40) | 3 | (1) |
| 5 | 2 | (1:2.8) | 2 | (225) | 3 | (30) | 1 | (5) |
| 6 | 2 | (1:2.8) | 3 | (210) | 1 | (60) | 2 | (3) |
| 7 | 3 | (1:2.4) | 1 | (240) | 3 | (30) | 2 | (3) |
| 8 | 3 | (1:2.4) | 2 | (225) | 1 | (60) | 3 | (1) |
| 9 | 3 | (1:2.4) | 3 | (210) | 2 | (40) | 1 | (5) |

根据表1-6进行试验；每组试验可以重复多次。

### 1.2.2 直观分析

把试验结果进行整理，列成计算表（见表1-7）。

#### 1.2.2.1 计算方法

在表1-7中，$\mathrm{I}_j$ 为第 $j$ 个因素第一水平所对应的数据之和。

$\mathrm{II}_j$ 为第 $j$ 个因素第二水平所对应的数据之和。

$\mathrm{III}_j$ 为第 $j$ 个因素第三水平所对应的数据之和。

$\overline{\mathrm{I}}_j$ 为第 $j$ 个因素第一水平所对应的数据之和的平均。

$\overline{\mathrm{II}}_j$ 为第 $j$ 个因素第二水平所对应的数据之和的平均。

$\overline{\mathrm{III}}_j$ 为第 $j$ 个因素第三水平所对应的数据之和的平均。

$\mathrm{I}_j^2$ 为第 $j$ 个因素第一水平所对应的数据之和的平方。

$\mathrm{II}_j^2$为第 $j$ 个因素第二水平所对应的数据之和的平方。

$\mathrm{III}_j^2$为第 $j$ 个因素第三水平所对应的数据之和的平方。

$R_j$ 为第 $j$ 个因素各水平的综合平均值的极差。即：

$$R_j=\max\{\overline{\mathrm{I}}_j,\ \overline{\mathrm{II}}_j,\ \overline{\mathrm{III}}_j\}-\min\{\overline{\mathrm{I}}_j,\ \overline{\mathrm{II}}_j,\ \overline{\mathrm{III}}_j\},\ j=1,\ \cdots,\ 4. \tag{1-1}$$

$\widehat{w}_{lj}$为第 $j$ 个因素第 $l$ 个水平的效应。它们分别是：

$$\widehat{w}_{1j}=\overline{\mathrm{I}}_j-\overline{y} \tag{1-2}$$

$$\widehat{w}_{2j}=\overline{\mathrm{II}}_j-\overline{y}\quad l=1,\ 2,\ 3 \tag{1-3}$$

$$\widehat{w}_{2j}=\overline{\mathrm{III}}_j-\overline{y} \tag{1-4}$$

$$j=1,\ 2,\ 3,\ 4$$

式中 $\overline{y}$——总平均值。

$$\overline{y}=\frac{1}{9}\sum_{i=1}^{9}y_i \tag{1-5}$$

1.2.2.2 直观分析

计算上面有关参数之后，可以作直观分析：

（a）通过每个因素各水平的平均值作出直观分析图。

（b）选取最优生产条件（即选取最优水平组合）。若考察指标越大越好，则选取每个因素各个水平的均值最大值所对应的水平组合作为最优生产条件；若考察指标越小越好，则选取每个因素各个水平的均值最小值所对应的水平组合作为最优生产条件。

（c）分析因素的主次关系。如果某个因素是主要因素，则在数量关系上表现出极差很大。如果极差很小，则说明该因素不是主要因素。主要因素是潜力很大的因素，要想进一步提高（或降低）指标，首先要在主要因素上考虑，或安排下一轮试验方案。