試驗設計教科书

产品设计开发质量管理基础教材之一

田口玄一
横山巽子 合著

上海机械工业质量管理协会

再版说明

为了吸取国外先进技术，提高科学管理水平，促进生产发展和改善产品质量，采用举办函授教育的方式，是一种普及与提高相结合的短期奏效的好方法。为此，我们翻译出版了日本规格（标准化）协会一九八一年五月一日出版的，日本青山大学教授田口玄一博士和青山大学讲师横山巽子女士合著的《试验设计法教科书》一书。

本书主要介绍试验设计法及其应用技巧。特别是结合工艺应用的实例，作了深入浅出的说明，适合我国作为质量管理基础教育的教材和供广大管理工作者参考。

一九八二年十月第一版发行后，经过试办函授教育和多次培训班作为教材，深受学员和广大质量管理工作者的喜爱，并在生产实践和设计试验中获得了一批满意的课题成果。

由于初次翻译编排出版，也发现了一些差错除已勘误外，在本书再版时由本协会质量研究小组（QRC）成员作了必要的修改。名词术语均按上海市第一机电工业局的企业标准沪Q/JBZ 1—85《试验设计名词术语》逐一订正，并此说明。

上海机械工业质量管理协会

一九八四年十二月

原　　序

试验设计法的内容按其方法来分：包括（１）特性指标的确定方法；（２）表头设计方法；（３）试验数据计算分析方法三部分。

当已经确定了试验的特性指标、设计好了表头并进行了试验后，怎样通过统计分析来获得正确的结论，这就是试验数据的计算分析。而对试验的各种因素对特性指标影响的大小进行定量分析的方法就是方差分析。试验数据的计算分析就是由这种方差分析、因素效应的估计、最佳条件和最佳公式的确定以及预测等内容所组成的。

在本书的第一至第十二章中，用极为简单的数据对试验数据计算分析的基础作了说明。

本书的内容重点放在试验设计法应用的通用性和技术上的实践意义方面，这是本书和其他试验设计法著作的不同之处，对于这部分，希望读者务必做一下练习。

书中习题的数据都比较简单，但读者为了解决实际问题而进行试验和调查研究时，取得的数据规模至少要比书中的问题大一倍。

当已经确定影响特性指标（可以有几个）的各个因素时，采用怎样的因素水平组合来进行试验才能有效地得到足以信赖的结论，这称为试验的表头设计，这是试验设计法的主要方法。表头设计方面，除因素少的场合，作为对上述问题的最有效的方法，现在广泛地开始采用正交表的表头设计。

在第十二章以后，尽可能抓住要点介绍采用正交表的表头设计，其中最重要的裂区法、拟因素法。裂区法和拟因素法是表头设计中最重要的方法，如能自由地应用这些方法将大大有助于解决实际问题。本书后半部分介绍的主要就是表头设计方法和试验数据的计算分析。

实际设计中的另一个问题是如何改进不同指标特性数据使具有效应的可加性。这个问题是一个与专业技术有关的边缘问题，也是试验设计中最难说清楚的问题。本书限于篇幅仅在第九章的特性数据的分类中涉及了这个问题，未能详细展开。并且，对于表示测量、试验、分析、动态特性等项目好坏的尺度的SN比和对上述内容进行改进的试验设计法也限于篇幅而不得不从略。另外，也不得不删去在计算试验（模拟试验的扩大应用）等方面中的应用部分。对此，请读者参阅本书的参考文献有关内容。

本书所采用的试验数据，可以采用日本电电公司的计算机服务DANEX程序进行分析，关于DANEX程序的具体内容可以向电电公司和日本规格协会咨询。

最后，对本书的校对、改正计算错误等方面给予大力帮助的日本规格协会的川村正信、若园叔邦两位先生表示深切的谢意。

作　　者

目　　录

第一章　总和、平均值、偏差平方和、方差……………………………………（1）

1．1　总和与平均值……………………………………………………（1）
1．2　偏　　差……………………………………………………………（2）
1．3　偏差平方和与方差………………………………………………（3）
练习（1）……………………………………………………………（7）

第二章　方差分析入门……………………………………………………（8）

2．1　偏差平方和的分解………………………………………………（8）
2．2　方差分析表………………………………………………………（11）
2．3　要因效应的估计…………………………………………………（14）
练习（2）……………………………………………………………（15）

第三章　试验的随机化……………………………………………………（16）

3．1　试验的随机化及其实施方法……………………………………（16）
3．2　计算公式…………………………………………………………（21）
练习（3）……………………………………………………………（22）

第四章　单因素设计………………………………………………………（23）

4．1　重复取样数相等时………………………………………………（23）
4．2　重复取样数不等时………………………………………………（26）
练习（4）……………………………………………………………（27）

第五章　自由度为1的分量分解…………………………………………（28）

5．1　对比及其偏差平方和……………………………………………（28）
5．2　采用一次回归式时………………………………………………（32）
5．3　采用正交多项式时………………………………………………（38）
练习（5）……………………………………………………………（41）

第六章　双因素设计………………………………………………………（42）

6．1　因素和水平………………………………………………………（42）
6．2　化学反应的例子…………………………………………………（44）

练习（6）……（54）

第七章　效应分解的双因素设计……（55）

7.1　只有一个因素是连续变量时……（55）

7.2　二个因素是连续变量时……（58）

练习（7）……（62）

第八章　重复取样的双因素设计……（63）

8.1　重复取样数相等的场合……（63）

8.2　重复取样数不等的场合……（66）

练习（8）……（70）

第九章　特性值的分类……（71）

9.1　特性值和单调性……（71）

9.2　特性值的分类……（74）

补充注释：奥米茄变换……（78）

第十章　0和1数据的计算分析法……（84）

10.1　重复取样数相等的场合……（84）

10.2　重复取样数不等的场合……（86）

练习（10）……（88）

十一章　累积法、频数法……（89）

11.1　累积法……（89）

11.2　频数法……（104）

练习（11）……（105）

第十二章　正交表入门……（107）

12.1 L_8 正交表……（107）

12.2　正交试验的目的……（110）

12.3　正交试验的计算……（111）

12.4　正交表的种类……（118）

练习(12)……（119）

第十三章　点线图及其应用……（121）

13.1 L_8 正交表的的点线图……（121）

13.2 L_{16} 正交表的的点线图及其使用方法……（124）

13.3 其它正交表的点线图……（139）

第十四章 并 列 法……（140）

14.1 二水平正交表的并列……（140）
14.2 三水平正交表的并列……（141）
14.3 汽车轮胎寿命试验的表头设计……（141）
14.4 喷淋洗涤试验数据的分析……（144）
练 习（14）……（149）

第十五章 拟水平法 组合法……（150）

15.1 拟水平法……（150）
15.2 组 合 法……（151）
15.3 碳粉的试验……（151）
15.4 铸铝零件的试验……（155）
练 习（15）……（160）

第十六章 确定试验规模的方法……（162）

16.1 理想的试验规模—某个简单的例子……（162）
16.2 正交试验的最佳重复取样数……（162）
16.3 正交试验的最佳重复试验次数……（164）

第十七章 裂 区 法……（167）

17.1 层压板的试验……（167）
17.2 正交表各列的分群……（169）
17.3 简单的例子说明……（170）
17.4 试验数据计算分析……（171）
17.5 层压板试验的表头设计……（172）
17.6 方差分析……（174）
17.7 显著性要因的估计……（176）
17.8 三水平正交表的裂区法—碳粉的试验……（179）
17.9 因素搭配和试验数据……（180）
17.10 试验数据计算分析……（183）
练 习（17）……（187）

第十八章 拟因素法……（188）

18.1 所谓拟因素法……（188）
18.2 拟因素法［1］变因素法……（188）
18.3 碳膜电阻的试验及数据分析……（191）

18.4　拟因素法[2]赋闲列法……………………………………………………（194）
18.5　电话用碳粉的试验……………………………………………………（195）
18.6　碳粉试验的表头设计……………………………………………………（196）
18.7　方差分析……………………………………………………（198）
18.8　显著性因素的估计和最佳条件……………………………………………………（201）
18.9　当混有各种不同水平的因素时……………………………………………………（204）
18.10　试验数据计算分析……………………………………………………（206）
18.11　赋闲列的设计方法……………………………………………………（208）
18.12　存在区组因素时……………………………………………………（209）
练　习（18）……………………………………………………（213）

附　录……………………………………………………（217）

附表一　平方表、平方根表……………………………………………………（217）
附表二　F表（5%，1%）……………………………………………………（224）
附表三　F表（2.5%）……………………………………………………（228）
附表四　分贝值表 $1 \leqslant \eta \leqslant 10.00$（$10\log 10\eta$）……………………………………………………（229）
附表五　奥米茄变换表……………………………………………………（232）
附表六　正交多项式系数表……………………………………………………（236）
附表七　正交表和点线图……………………………………………………（238）

第 一 章
总和、平均值、偏差平方和、方差

本章将对总和、平均值、偏差平方和、方差这一类数据解析的基础进行说明。

1.1 总和与平均值

下面是十个人的身长数据（单位：厘米）。

163.2	171.6	156.3	159.2	160.0
158.7	167.5	175.4	172.8	153.5

设这十个人的身长总和为T，则

$$T = 163.2 + 171.6 + \cdots\cdots + 153.5 = 1638.2 \tag{1.1}$$

并设其平均值为$\overline{Y}$，则

$$\overline{Y} = \frac{1}{10} \times T = \frac{1638.2}{10} = 163.82 \tag{1.2}$$

在实际计算时，为了简化计算，减去一个接近于平均值的数据（称为虚拟值）来求总和与平均值。这种方法在加算时位数较少，所以计算比较简单。

有时在减去虚拟值以后，再乘上0.1、10、100等，使之去掉小数部份。

这样的方法称为数据变换（一次变换），或称数据处理。

若从上述身长的数据减去虚拟值160.0厘米，则其数值如下：

3.2	11.6	−3.7	−0.8	0.0
−1.3	7.5	15.4	12.8	−6.5

这时，总和T与平均值$\overline{Y}$可用下式求得，

$$T = 160.0 \times 10 + [3.2 + 11.6 + \cdots\cdots + (-6.5)] = 1600 + 38.2 = 1638.2 \tag{1.3}$$

$$\overline{Y} = 160.0 + \frac{38.2}{10} = 163.82 \tag{1.4}$$

一般来说，当有n个测量值Y_1、Y_2、……，Y_n时，Y_1、Y_2、……，Y_n的总和T用下式定义，

$$T = Y_1 + Y_2 + \cdots\cdots + Y_n \tag{1.5}$$

其平均值$\overline{Y}$用下式定义，

$$\overline{Y} = \frac{1}{n}(Y_1 + Y_2 + \cdots\cdots Y_n) \tag{1.6}$$

设虚拟值为Y_0，则 T和$\overline{Y}$也可写成

$$T = Y_0 \times n + [(Y_1 - Y_0) + (Y_2 - Y_0) + \cdots\cdots + (Y_n - Y_0)] \tag{1.7}$$

$$\overline{Y} = Y_0 + \frac{1}{n}[(Y_1 - Y_0) + (Y_2 - Y_0) + \cdots\cdots + (Y_n - Y_0)] \qquad (1.8)$$

1.2 偏　　差

这里讲的偏差有两种，一种是偏离目标值Y_0的偏差，另一种是偏离平均值$\overline{Y}$的偏差。先讲偏离目标值Y_0的偏差。

设n个测量值为Y_1，Y_2，……，Y_n。如果n个测量值已有某个目标值Y_0，那末把Y_1，Y_2，……，Y_n与目标值Y_0的差值叫做偏离目标值Y_0的偏差。

例如：试制了12个10千欧姆的碳膜电阻，测得阻值如下：

10.3	9.9	10.5	11.0	10.0	10.3
10.2	10.3	9.8	9.5	10.1	10.6

偏离目标值10千欧的差值为

0.3	－0.1	0.5	1.0	0.0	0.3
0.2	0.3	0.2	0.5	0.1	0.6

目标值不一定只限于是目标本身的数值，有时也把规格值、标称值（例如标称容量为100克时的100克等）、理论值（由理论公式计算的值，用公司企业常用的标准方法计算的值），预测值及比较值（例如其它国家或其它公司同类产品的数值）等作为目标值Y_0。

例如，在某个电路中，采用五种不同的电阻值R（欧姆）和电压值E（伏特）时，测得的电流值Y（安倍）如表1.1所示。

表1.1　观测值Y和理论值Y_0

R	E	Y	Y_0	偏差
10	50	5.10	5.00	0.10
20	20	0.98	1.00	－0.02
20	40	2.04	2.00	0.04
40	30	0.75	0.75	0.00
40	60	1.52	1.50	0.02

这时，根据欧姆定律的理论值为5.00、1.00、2.00、0.75、1.50，所以，观测值Y与理论值Y_0的偏差可见表1.1的最后一项。

又如，以某地小学六年级学生的身长来说，有时又要计算与全国小学六年级学生身长平均值的差值。

这时，全国小学六年级学生身长的平均值就相当于目标值Y_0，因此有必要计算偏离目标值的差值。

下面来讨论偏离平均值的偏差。

例如，当不存在目标值和理论值Y_0而只有n个观测值Y_1、Y_2、……Y_n时，需要计算与其平均值$\overline{Y}$的差值。

我们把它叫做偏离平均值的偏差。

在1.1节的身长数据中，十个人的平均身长为$\overline{Y}=163.82$，各个数值与平均值的偏差为

$$\left.\begin{aligned}163.2-163.82&=-0.62\\171.6-163.82&=7.78\\166.3-163.82&=-7.52\\&\vdots\\153.5-163.82&=-10.32\end{aligned}\right\}\tag{1.9}$$

偏离平均值的偏差总和将为零。但是，很多场合并不讲明这是偏离算术平均值$\overline{Y}$的偏差，而只是称为偏差。因此重要的是要明确区别这是偏离平均值的偏差，还是偏离理论值的偏差。

对于上述身长的数据，若这十个人是偏僻地区的青年，有时除考虑上述的偏差外，还要考虑这些青年的身长比全国青年的平均身长高还是矮。这时需要计算与全国青年身长的平均值Y_0的偏差。这与计算十个人身长偏离平均值$\overline{Y}$的偏差是不同的。

1.3 偏差平方和与方差

以偏差来说，不管是偏离平均值的偏差，还是偏离目标值、理论值的偏差，其值将是正值、负值或零。

当偏差为一组数据时，为了用一个数值表示此偏差的大小，一般采用偏差的平方和或其平均值来表示。

这个偏差的平方和叫做偏差平方和，其值的平均值叫方差。它们都是定量地表示这些数据变化大小的值。

如有n个测量值$Y_1, Y_2, \cdots\cdots$、Y_n，设Y的目标值为Y_0，则偏差(Y_1-Y_0)，$(Y_2-Y_0)\cdots\cdots(Y_n-Y_0)$的平方和叫总偏差平方和（或叫总变差），用$S_T$表示。

$$S_T=(Y_1-Y_0)^2+(Y_2-Y_0)^2+\cdots\cdots+(Y_n-Y_0)^2\tag{1.10}$$

总偏差平方和S_T中独立的平方数称为它的自由度。（1.10）式的自由度为n。

第1.2节中碳膜电阻的偏差就是一组数据，最小为-0.5，最大为1.0。当存在一组不同数值的数据据时，为了用一个数值表示这些数据波动的大小，就要计算它们的偏差平方和。

$$\begin{aligned}S_T&=(0.3)^2+(-0.1)^2+(0.5)^2+\cdots\cdots+(0.6)^2\\&=0.09+0.01+0.25+\cdots\cdots+0.36=2.23\end{aligned}\tag{1.11}$$

（1.11）式中的差值偏离目标值10千欧的偏差，所以其自由度为12。

表1.1的总偏差平方和S_T为

$$\begin{aligned}S_T&=(0.10)^2+(-0.02)^2+(0.04)^2+(0.00)^2+(0.02)^2\\&=0.0124\end{aligned}\tag{1.12}$$

其自由度为5。

设n个测量值的平均值为$\overline{Y}$时，严格地讲应把偏离平均值$\overline{Y}$的差值的平方和称之为偏离平均值的总偏差平方和，但一般往往简称为总平方和（总变差）。

$$S_T=(Y_1-\overline{Y})^2+(Y_2-\overline{Y})^2+\cdots\cdots+(Y_n-\overline{Y})^2\tag{1.13}$$

在（1.13）式中，尽管有n个平方，但其总平方和的自由度为（n—1）。

这是因为在n个偏差$(Y_1-\overline{Y})$，$(Y_2-\overline{Y})$，……，$(Y_n-\overline{Y})$间存在着其和为零的下述线性关系式。

$$(Y_1-\overline{Y})+(Y_2-\overline{Y})+\cdots\cdots+(Y_n-\overline{Y})=0 \qquad (1.14)$$

一般说，n个偏差平方和的自由度，等于从n中减去平方前的偏差间关系式（恒等的）的数目。

在（1.13）式的总平方和S_T中，虽然平方的数目为n，但其偏差之间存在一个（1.14）式的关系式，所以其自由度为（n－1）。

一般说，当有n个观测值时，偏离目标值或理论值的偏差平方和的自由度为n，偏离平均值$\overline{Y}$的偏差平方和的自由度为（n－1）。

（1.9）式的总平方和S_T为

$$\begin{aligned}S_T&=(-0.62)^2+7.78^2+(-7.52)^2+\cdots\cdots+(-10.32)^2\\&=0.3844+60.5284+56.5504+\cdots\cdots+106.5204\\&=514.3860\end{aligned} \qquad (1.15)$$

这是自由度为9的偏离平均值$\overline{Y}=163.82$的偏差平方和。

平方值的计算可以利用平方表（附表1），或采用台式计算器计算。

在求偏离算术平均值$\overline{Y}$的偏差平方和时，可以从Y_1、Y_2、……、Y_n中减去一个虚拟值再取其平方值，然后再减去一个修正项求得，这种方法较简便。

$$\begin{aligned}S_T&=(Y_1-\overline{Y})^2+(Y_2-\overline{Y})^2+\cdots\cdots+(Y_n-\overline{Y})^2\\&=(Y_1')^2+(Y_2')^2+\cdots\cdots+(Y_n')^2-\frac{(Y_1'+Y_2'+\cdots\cdots Y_n')^2}{n}\end{aligned} \qquad (1.16)$$

式中，Y_1'、Y_2'、……Y_n'是从Y_1、Y_2、……、Y_n中减去虚拟值而得到的值。

在（1.16）式中，设

$$S_T=\frac{(Y_1'+Y_2'+\cdots\cdots Y_n')^2}{n} \qquad (1.17)$$

这称为修正项，用CF，CT表示。

现在来求上述身长的总平方和，虚拟值取160.0厘米，代入（1.16）式中进行计算。

$$\begin{aligned}S_T&=\text{（原数据减去虚拟值后的平方和}-\frac{\text{（原数据减去虚拟值后的平方和）}^2}{10}\\&=(163.2-160.0)^2+(171.6-160.0)^2+\cdots\cdots+(153.5-160.0)^2\\&\quad-[(163.2-160.0)+(171.6-160.0)+\cdots\cdots+(153.5-160.0)]^2/10\\&=3.2^2+11.6^2+\cdots\cdots+(-6.5)^2-\frac{[3.2+11.6+\cdots\cdots+(-6.5)]^2}{10}\\&=660.32-\frac{38.2^2}{10}=660.32-145.924=514.396\end{aligned} \qquad (1.18)$$

计算修正项与偏差平方和时，小数点一般算到原来数据的平方值的最后一位就可以了，在本例中为小数二位。因此小数点三位以下四舍五入。即

$$S_T=660.32-145.92=514.40\text{（自由度为9）} \qquad (1.19)$$

这时修正项CT为，

$$CT=\frac{38.2^2}{10}=145.92 \qquad (1.20)$$

下面对方差进行说明。

偏差平方和除以自由度就叫方差。

$$方差 = \frac{偏差平方和}{自由度} \tag{1.21}$$

方差是表示每个自由度的偏差大小的平均值，是意味着每个单位的误差、差值、波动、变化等的大小值，是相当于物理学的单位时间的功、能等的测度。

现假设有n个测量值Y_1、Y_2、……、Y_n，目标值为Y_0，则方差V为

$$V = \frac{(Y_1 - Y_0)^2 + (Y_2 - Y_0)^2 + \cdots\cdots + (Y_n - Y_0)^2}{n} \tag{1.22}$$

（1.11）式中碳膜电阻值的总平方和S_T的自由度是12，所以其方差为

$$V = \frac{S_T}{自由度} = \frac{2.23}{12} = 0.186 \tag{1.23}$$

有时此值也叫做自由度为12的误差方差。误差方差值的小数点位数最好比原来多一位。

在身长的例子中，当偏差平方和等于偏离平均值$\overline{Y}$的偏差平方和时，偏差平方和除以其自由度，就是方差。

即，设Y_1、Y_2、……、Y_n的算术平均值为$\overline{Y}$，则方差V为

$$V = \frac{(Y_1 - \overline{Y})^2 + (Y_2 - \overline{Y})^2 + \cdots\cdots + (Y_n - \overline{Y})^2}{(n-1)} = \frac{S_T}{(n-1)} \tag{1.24}$$

然而，有时当目标值和理论值已经确定，则一开始就把与目标值或理论值的差值作为试验数据。

例如，生产Y_0千欧的碳膜电阻时，把$(Y_1 - Y_0)$，$(Y_2 - Y_0)$……，$(Y_n - Y_0)$作为数据，有时把钟表指示的时间与标准时间的差（快、慢）作为数据。

若设定偏离目标值或理论值的差值作为Y_1，Y_2，……Y_n时，则平均值$\overline{Y}$为

$$\overline{Y} = \frac{Y_1 + Y_2 + \cdots\cdots + Y_n}{n} \tag{1.25}$$

我们把这个$\overline{Y}$值称为偏倚，正确地说，应称为偏倚的估计值。当减去目标值或理论值前的最初观测值比目标值或理论值大时，则（1.25）式的偏倚$\overline{Y}$是正值且较大，当观测值比目标值或理论值小时，则将是负值且较大。

在1.2节的碳膜电阻例子中，偏离目标值的偏差为0.3　−0.1　0.5　1.0　0.0　0.3　0.2　0.3　−0.2　−0.5　0.1　0.6，所以偏倚$\overline{Y}$为

$$\overline{Y} = \frac{1}{12}(0.3 - 0.1 + 0.5 + \cdots\cdots + 0.6) = \frac{2.5}{12} = 0.208 \tag{1.26}$$

$\overline{Y}$值的小数点位数最好也要比原来数据多求1～2位。

在本例中，偏倚为正，一般说，碳膜电阻的阻值往往比目标值大。

不论正负，为了表示偏倚的绝对值大小，可以把$\overline{Y}$值平方后乘以n倍。

$$偏倚大小 = n(\overline{Y})^2 = n\left(\frac{Y_1 + Y_2 + \cdots\cdots + Y_n}{n}\right)^2 = \frac{(Y_1 + Y_2 + \cdots\cdots + Y_n)^2}{n} \tag{1.27}$$

此值相当于把偏离目标值或理论值作为Y_1，Y_2，……，$+Y_n$时的平均值m的平方和。因此，下列的分解式成立。

$$Y_1^2 + Y_2^2 + \cdots\cdots + Y_n^2 = \frac{(Y_1 + Y_2 + \cdots\cdots + Y_n)^2}{n} + [(Y_1 - \overline{Y})^2 + (Y_1 - \overline{Y})^2$$

$+\cdots\cdots+(Y_n-\overline{Y})^2$ (1.28)

如用文字表示，则为

总平方和 =（平均值平方和）+（偏离平均值的偏差平方和） (1.29)

（注）当试验数据为偏离目标值或理论值的差值（Y_1，Y_2，……，Y_n）的场合，计算平均值平方和不能从Y_1，Y_2，……，Y_n中减去虚拟值，必须用偏离理论值或目标值的差值求其平均值平方和。

在碳膜电阻例子中，由（1.11）式求出的S_T为2.23，平均值m的平方和S_m为

$$S_m=\frac{(0.3+0.1+\cdots\cdots+0.6)^2}{12}=\frac{2.5^2}{12}=0.52 \tag{1.30}$$

因此，下式成立。

$$2.23=0.52+\text{偏离平均值}\overline{Y}\text{的偏差平方和} \tag{1.31}$$

偏离平均值的偏差平方和Sc可由下式求得。

$$Sc=S_T-S_m=2.23-0.52=1.71 \tag{1.32}$$

这种场合，把Se称为误差平方和。误差平方和的自由度为（12－1）=11。

误差平方和除以自由度的商称为误差方差，用Ve表示。

$$Ve=\frac{Sc}{\text{自由度}}=\frac{1.71}{11}=0.155 \tag{1.33}$$

Ve表示各个碳膜电阻值的波动大小，当此值较大时，意味着各个产品之间的差值较大。

另一方面，当平均值平方和 S_m 较大时，则表示产品的平均值偏离目标值较大。要将其减小，多数的场合是比较简单的。

例如在碳膜电阻的例子中，只要延长碳化时间或改变切深方法就可以进行修正。

象尺寸等特性值，其情况也与此相同，对偏倚进行修正比较容易，但是要减小波动，则须化很多工夫。

在身长的例子中，之所以把总平方和S_T作为偏离平均值的偏差平方和来求，是因为身长无目标值或理论值的缘故。若设理论值为零，把偏离零的偏差作为Y_1，Y_2，……，Y_n，则还可以求出S_m表示偏倚的大小。

当采用交直流混合的电话时，其电流功率的计算式如下，（1.29）的关系式就相当于这个概念。

总电流的功率 = 直流电压的平方 + 交流电的功率

这时，$\overline{Y}$就是直流电压（电话中为48伏），交流电的功率与交流电压的平方是同一回事。

在电的范畴中，不存在加法性，但是在电压平方的功率范畴中存在加法性，这是非常重要的。

在工程学范畴内，我们把这样的分解称为能谱分析。

（1.28）式的分解在数学上是有限元的分解，它是属于线性数学中称为二项式理论的概念。但是，我们不必拘泥于这种物理、数学的概念。

因为只要熟悉了实际的计算便能自己理解其意义的缘故。

在电话的例子中，直流电压对传递声音毫无作用。也就是说，只是交流部分才对声音的传递起作用。

与此同时，当采用偏离目标值或理论值的偏差数据时，平均值平方和Sm与剩余的偏差平方和的技术意义及采取的措施往往是完全不同的，所以区别这两者的不同是非常重要的。

练　习　（1）

1、制作了八个目标值为100Ω的电阻，实测的电阻值如下：

98	103	100	101
102	98	99	105

试求偏离目标值的偏差，平均值平方和，误差平方和，误差方差。

第二章 方差分析入门

本章讲解观测值的常数系数线性式与其偏差平方和的求法，并对方差分析入门进行介绍。

2.1 偏差平方和的分解

设n个观测值为Y_1，Y_2，……Y_n，则Y_1，Y_2，……，Y_n的常数系数的一次式称为线性式。

设n个的常数系数为C_1，C_2，……，C_n，线性式为L。

$$L = C_1Y_1 + C_2Y_2 + \cdots\cdots + C_nY_n \quad (2.1)$$

假设在系数C_1，C_2，……，C_n中，可以有若干个为零，但不是全部为零。这些系数C_1，C_2，……，C_n的平方和叫做线性式L的单位数D，即

$$D = C_1^2 + C_2^2 + \cdots\cdots + C_n^2 \quad (2.2)$$

例如，因为n个测量值的平均值（偏倚）$\overline{Y}$为

$$\overline{Y} = \frac{Y_1 + Y_2 + \cdots\cdots + Y_n}{n} = \frac{1}{n}Y_1 + \frac{1}{n}Y_2 + \cdots\cdots + \frac{1}{n}Y_n \quad (2.3)$$

$$C_1 = C_2 = \cdots\cdots = C_n = \frac{1}{n} \quad (2.4)$$

其单位D为

$$D = \left(\frac{1}{n}\right)^2 \times n = \frac{1}{n} \quad (2.5)$$

线性式的平方除以单位数D就得到自由度为1的偏差平方和。把它记作S_L，则

$$S_L = \frac{L^2}{C_1^2 + C_2^2 + \cdots\cdots + C_n^2} \quad (2.6)$$

Y_1，Y_2，……，Y_n的和也可称$C_1 = C_2 = \cdots\cdots C_n = 1$的线性式，它的单位数为n。

这时的偏差平方和S_L用（2.6）式计算，则

$$S_L = \frac{(Y_1 + Y_2 + \cdots\cdots + Y_n)^2}{n} \quad (2.7)$$

这相当于平均值平方和Sm或修正项。

为了加深对线性式与线性式的偏差平方和的理解，下面举一个例子。

计算例1

有六个日本人，四个美国人，其身高分别为（单位：厘米）

A_1（日本人）	158	162	155	172	160	168
A_2（美国人）	186	172	176	180		

这里列出十个人的身高，但这十个人身高的总平方和可由偏离平均值的偏差平方和求得。此时不存在目标值或理论值，所以可减去虚拟值170厘米来计算这十个人的总平方和。

A_1与虚拟值的偏差：$-12 \quad -8 \quad -15 \quad 2 \quad -10 \quad -2$（合计$-45$）

A_2与虚拟值的偏差：16　2　6　10（合计34）

因此

$$CT=\frac{(-45+34)^2}{10}=\frac{(-11)^2}{10}=12 \qquad (2.8)$$

$$S_T=(-12)^2+(-8)^2+\cdots\cdots+10^2-CT=937-12=925 \qquad (2.9)$$

但是，研究一下这些数据谁都会发现，这十个人身高的总平方和无疑主要是由日本、美国两民族不同所产生的。那么，在这十个人身高的总平方和925中，民族差别占多少呢？

大概无论谁都会用日本人的平均身高与美国人的平均身高的差值L来表示这两个不同民族的差值。

$$L=（日本人平均身高）-（美国人平均身高）=\left(170+\frac{-45}{6}\right)-\left(170+\frac{34}{4}\right)$$

$$=\frac{-45}{6}-\frac{34}{4} \qquad (2.10)$$

我们常把A_1（日本人）的合计值仍用A_1来表示，把A_2（美国人）的合计值仍用A_2来表示。

因此上述线性式还可写成

$$L=\frac{A_1}{6}-\frac{A_2}{4} \qquad (2.11)$$

现设这十个人的身高为Y_1、Y_2、Y_3、Y_4、Y_5、Y_6，（以上为日本人），Y_7、Y_8、Y_9、Y_{10}（以上为美国人），则总平方和S_T是这十个人身高测量值Y_1、Y_2……Y_{10}的总平方和，（2.10）式或（2.11）式是这十个观测值Y_1、Y_2、……Y_{10}的常数系数的线性式。

因而（2.11）式的L偏差平方和S_L为

$$S_A=S_L=\frac{\left(\frac{A_1}{6}-\frac{A_2}{4}\right)^2}{\left(\frac{1}{6}\right)^2\times6+\left(-\frac{1}{4}\right)^2\times4}=\frac{\left(\frac{-45}{6}-\frac{34}{4}\right)^2}{\frac{1}{6}+\frac{1}{4}}=\frac{[4\times(-45)-6\times34]^2}{4^2\times6+(-6)^2\times4}$$

$$=\frac{(-384)^2}{240}=614 \qquad (2.12)$$

结果也就是说，在十个人身高的总平方和925中，有614是由不同民族产生的。从925减去614的值则表示同一民族中个人差的大小。

这种差值称为误差平方和S_e。

$$S_e=S_T-S_A=925-614=311 \qquad (2.13)$$

S_T的自由度是9，S_A的自由是1，S_e的自由度是（$9-1$）$=8$，故误差方差V_e为

$$V_e=S_e/8=311/8=38.9 \qquad (2.14)$$

误差方差是表示同一民族内每个人身高的差值大小的尺度。

误差方差V_e的平方根称为每个人身高差值的标准偏差，大多数场合记作S。

$$S=\sqrt{V_e}=\sqrt{38.9}=6.24\text{厘米} \qquad (2.15)$$

这时就可以把十个人身高的总平方和S_T分解为日、美两民族间的偏差平方和与个人之间的偏差平方和两部分。

$$S_T=S_A+S_e$$

$$925=614+311 \quad （偏差平方和） \qquad (2.16)$$

$9=1+8$　（自由度）　（2.17）

此个人间的差值是因为存在着这十个人，如果没有这十个数据，当然是不合理的。

实际在S_e中，只包含八个人差，余下两个，一个在修正项CT中，一个在S_A中。

换句话说，在S_A的614中，多了一个平均的个人差，所以真正的民族间差值应从S_A中扣除一个个人差，即减去一个误差方差。

把扣除后的值叫做两民族间的纯偏差平方和，以S_A'来表示。

$$S_A'=S_A-V_e=614-38.9=575.1 \tag{2.18}$$

虽然还有一个个人间的差值在修正项中，但在计算身高这类的偏差平方和时，没有必要计算平均值平方和，所以让它在修正项算了。

也就是说，数据的总平方和始终是按自由度为9来考虑。

因此，误差的纯平方和S_e'，应在S_e上加上由S_A减去的一个误差V_e求得。

$$S_e'=S_e+V_e=311+38.9=349.9 \tag{2.19}$$

也就是说，下式成立。

$$S_T=S_A'+S_e' \tag{2.20}$$

$$925=575.1+349.9 \tag{2.21}$$

将两边除以S_T，再扩大100倍后的数值称为贡献率的分解。

此时，下式成立。

$$100=P_A+P_e=\frac{575.1}{925}\times 100+\frac{349.9}{925}\times 100=62.2+37.8 \tag{2.22}$$

也就是说，十个人身高的总平方和是925，而其中日美民族间的差值占62.2%，个人差占37.8%〔计算例1的计算完毕〕。

这样，当存在一组数值不同的数据时，则用总平方和来表示这组数据的变化。为了求出总平方和中各种因素影响的贡献率，就必须对总平方和进行分解。

为了能自由地对总平方和进行分解，必须掌握方差分析的计算方法。

方差分析的计算虽然是数学上的二项式表示，但是最重要的是把什么列为因素项。列什么样的因素项，这由研究人员的专业技术决定，所以不论对数学上的二项式理论怎么熟悉，还不能解决方差分析问题。

这样还不如说，重要的是研究人员和工程技术人员要学会方差分析的计算，对任何数据都能进行偏差平方和的分解，能立即计算出自己要研究的因素项的影响大小。

下面表示有目标值或理论值时的偏差平方和分解的计算例。

计算例2

为了提高某产品的耐磨性，准备了两种产品的原料，一种加有某种添加剂，一种未加。

A_1　未加添加剂的原料

A_2　加了添加剂的原料

现采用A_1、A_2分别制造了六个试样，并在某种条件下进行了磨损试验。磨损量的测量结果如下（单位：毫克）。

A_1	26	18	19	21	15	29
A_2	15	8	14	13	16	9

磨损量、变化量、变化率一类的数据目标值为零，所以可以直接求修正项的平方和。

$$A_1=26+18+\cdots\cdots+29=128 \tag{2.23}$$