

上海财经大学统计系列丛书

试验设计

• 苏均和



SYSJ

立信会计出版社

前 言

随着我国社会主义市场经济的确立,对统计提出了更高的要求,同时也为统计在市场经济与管理中的应用提供了一个广阔的前景。在充满竞争的市场经济环境中,统计定量分析已成为一种不可缺少的工具。

试验设计是统计学的一个重要分支,在经济与管理中具有广泛的应用价值。例如,试验设计在工业生产中可对产品的质量作出测定和分析;在农业中可分析影响某作物产量的诸多因素所产生的作用。这对提供质量和产量是极为重要的。

同样,在工农业生产和科学研究中,常常要进行各种试验,如果不了解统计推断的方法,可能会由于思考方法的错误而付出不必要的劳动。

例如,药厂在制造某种药品时,假定影响收率(即实际所得量与理论应得量的百分比)的因素有反应温度、反应压力及溶剂用量等。当希望知道反应温度上升 10°C 收率变化多少时,有人会想:在其它条件不变(反应压力、溶剂用量等因素必定要固定下来)的情况下,应该只使温度变动 10°C 。具有这样想法的人,实际上就是在付出不必要的劳动。读完本书后就能认识到,利用统计推断法,通过少数几次试验就能高效率地获得大量的情报知识。

同样,在社会科学中也可以有所应用。例如,从经济变动中分离出长期倾向与季节变动,以预测将来;分析经营数据,

查知与企业利益有关的主要因素关系,以决定将来的经营目标等。

试验设计的历史不长,但发展迅速,正渗透在工农业和科学研究等各个领域,体现了它的强大生命力。本书着重介绍试验设计的方法及这些方法的实际应用,目的在于使更多的人掌握这个工具,以适应目前改革开放形势的需要。

本书共分八章,这些章节是试验设计和方差分析的基本内容。读者只需有些高等数学和数理统计的基本知识就能顺利学完本书。相信读完此书,必有所获益。

在本书编写过程中,始终得到徐国祥教授的关心和帮助,吴永发在编纂习题及其它方面也做了许多工作,在此一并表示感谢。

由于编者水平有限,一定会存在不少缺点,恳请读者批评指正。

编 者

1994. 12

目 录

第一章 概述	1
§ 1.1 方差分析	1
§ 1.2 试验设计	2
§ 1.3 随机化原则	5
§ 1.4 试验设计的必要性和基本原理	6
习题一	11
第二章 线性模型简介	13
§ 2.1 线性模型.....	13
§ 2.2 多元正态及二次型分布.....	18
§ 2.3 最小二乘法.....	21
§ 2.4 最大似然估计.....	23
习题二	25
第三章 单因子试验设计	27
§ 3.1 单因子试验的数学模型.....	27
§ 3.2 固定效应模型的统计分析.....	30
§ 3.3 关于个别处理均值的检验.....	46
§ 3.4 随机效应模型的统计分析.....	48
习题三	53

第四章 双因子试验设计	56
§ 4.1 双因子试验的数学模型	56
§ 4.2 固定效应模型的统计分析	60
§ 4.3 随机效应模型的统计分析	77
§ 4.4 混合模型的统计分析	81
习题四	83
第五章 多因子试验设计	86
§ 5.1 多因子试验的数学模型和方差分析	86
§ 5.2 平方和与均方和的期望值规则	98
习题五	105
第六章 正交试验设计	108
§ 6.1 正交试验设计及其直观分析法	108
§ 6.2 正交试验的方差分析法	125
§ 6.3 正交表的灵活运用	133
习题六	137
第七章 随机化区组、拉丁方等设计	140
§ 7.1 随机化完全区组设计	140
§ 7.2 拉丁方设计	153
§ 7.3 希腊——拉丁方设计	168
习题七	176
第八章 不完全区组设计	177
§ 8.1 平衡不完全区组设计	177

§ 8.2 平衡不完全区组设计的统计分析 (区组内分析)	189
§ 8.3 平衡不完全区组设计的统计分析 (区组间分析)	198
§ 8.4 部分平衡不完全区组设计	202
§ 8.5 尤敦(Youden)方设计	212
习题八	217
附录	219
1. χ^2 分布表	219
2. t 分布表(双边)	221
3. F 分布表	222
4. 随机数字表	226
5. Duncan 多重显著性极差表	229
6. 正交表	231

第一章 概 述

§ 1.1 方差分析

方差(Variance)和方差分析(Analysis of Variance)两词,由 R. 费喧于 1918 年的《孟德尔遗传实验设计间的相对关系》一文中所首创。方差分析,也称变异数分析,其系统研究开始于 1923 年 R. 费喧在与麦凯基(W. A. Mackenzie)合写的《对收获量变化的研究》(Studies on crop variation)一文中。而于 1925 年, R. 费喧在《供研究人员用的统计方法》中对方差分析以及协方差分析(Analysis of covariance)进一步作了完整的论述。

R. 费喧最初在马铃薯实验中引入方差分析方法,是假定观察值为各试验处理(Treatment, 这里指品种)的效应与试验误差效应的总和。有了这个基本假定以后,就可使各组内的方差加以合并计算,从而使试验误差所含有的自由度增加,以有利提高分析的精确性和灵敏度。方差分析的基本思想是从不同角度,算出各种方差值,然后通过组间方差与组内方差的对比,来分析条件误差和随机误差,从而分析出实验数据中必然性因素与偶然性因素的影响大小。所以, R. 费喧指出,方差分析是一种在若干个能互相比较的资料组中,把产生变异的原因加以区分开来的方法与技术。正如英国耶特斯 1946 年 1

月在英国皇家统计学会会上所推荐的那样：“方差分析由于它使误差估计值的联合和使误差的不同值的各个构成部分的分开成为可能，就使从每一批资料中取出的独立抽样单位数能够减少到一个较小的数目，从而容许运用往往包含两个或多个阶段较为复杂的抽样设计。”

由于方差分析还能够同时比较任意多个平均数和适用于分析多个项目的试验要求，因此，它打破“除试验项目外，其余项目都控制不变”的传统方法，从而大大提高了试验分析的效率。而且这种方法并不限制样本的大小，即对大样本或小样本都可以使用。更有意思的是，方差分析简单实用，没有什么深奥的理论，正如 R. 费喧自己在 1934 年所说，方差分析“不是一种数学定理，宁可说是一种计算安排上的便利方法。……它的引人注目之处在于它的便利性质：一是因它使得全部资料纲举目张地一览无余，其逻辑结构令人感到层次井然；二是因它除了有助于逻辑进程外，它使得一切需要进行的显著性检验简化和归结为一种共同的格式。”

有关方差分析原理、方法及它的应用，我们将会在以后各章节中，具体展开和介绍。

§ 1.2 试验设计

R. 费喧自 1919 年起在农业实验站工作了 14 年。在此期间，由于试验的环境很不容易得到控制，试验的次数也必然大大地增加。因此，引起试验结果差异的主要有两个因素：第一，在田间试验中，土质、光照等客观条件的不同；第二，在田间试验中，所使用方法的的不同。由于两者经常同时在起作用，因此，

如何从总的差异中分析出这两种因素各自的影响以及如何测定它们,这是 R. 费喧所面临的迫切需要解决的问题,否则就无法进行试验分析。

R. 费喧于1922年8月与哥塞特在卢桑姆斯坦德第一次相见,并在试验设计思想上受其影响。后来,R. 费喧把自己扎实的数学知识和一系列统计手段运用到生产实践中去,并在实践中汲取前辈农业试验专家贝威(贝威在20世纪的最初10年曾设计过正方形试验区的棋盘法和半条播法等,对费喧影响很大)的成果,自1923年起,陆续发表了关于在农业试验中控制误差的论文。同年,他与麦凯基合作发表了有关试验设计的最初例子。自1925年来,他的完整的方差分析方法形成,给试验设计提供了有效地搜集原始数据以及更好地进行推断的方法。到了1926年,R. 费喧发表了试验设计方法的梗概,这些方法后来在1935年进一步得到完善。这样,R. 费喧的主要试验设计方法在20至40年代完成,并首先在卢桑姆斯坦德农业实验站得到检验与应用,后来又被他的学生推广到许多其它科学领域。

下面例举其中1925年提出的两种试验设计方法。

一、随机块区法(Randomized blocks)

此方法就是把试验区划分为若干块区,每一块区都试验全部项目。而试验项目的安排完全是随机的,这样就可以确保地力的均匀性,使影响试验结果的其它因素的作用相互抵消,达到突出需要反映的那个因素效果的目的。例如,A、B、C、D、E、F、G、H八个品种的产量试验,在每块区中,再分八小块作为基本单位,每个品种反复试验四次,如图1-1所示。

G	A	H	D	B	H	D	F
F	B	C	E	C	E	A	G
E	D	H	A	F	G	C	A
G	C	F	B	E	D	B	H

图 1-1 随机块区法

二、拉丁方格法(Latin square)

拉丁方格法是 R. 费喧为克服因随机块区法完全随机不易把握而设计的。他把各种试验项目在每一行和每一列中都控制在只出现一次,以防止某品种有可能较集中播种在一个位置上,图 1-2 所示为 4×4 拉丁方格法。

A	B	C	D	A	B	C	D	A	B	C	D	A	B	C	D
B	A	D	C	B	C	D	A	B	D	A	C	B	A	D	C
C	D	B	A	C	D	A	B	C	A	D	B	C	D	A	B
D	C	A	B	D	A	B	C	D	C	B	A	D	C	B	A

图 1-2 拉丁格法(4×4)

继 R. 费喧以后,试验设计理论的杰出贡献者是卢桑姆斯坦德农业实验站的统计部主任、R. 费喧的学生耶特斯。1936 年,他发表了“众多变量中安排各种实验的新方法”,创造了格子设计(Lattice design),并最先称之为拟因子实验。该设计因哈希巴格(B. Harshbarger)于 1947 年首创了方型格子设计而获得发展。

出身于澳大利亚的统计学家尤顿于 1937 年首次提出不完全拉丁方格设计,这是他受到 R. 费喧上一年在美国讲学后启发的结果。1938 年至 1939 年,印度统计学家鲍斯提出论文,对不完全设计与正交拉丁方格法作了进一步研究。其后,他与其他印度同行研究了“部分配置法”与“交错法”。1946 年,菲内(D. J. Finney)发表了《田间实验设计的新发展》一书,正式提出了“部分配置法”。1951 年,印度希里哈

德将尤顿的不完全拉丁方格设计加以一般化处理，这对于试验的进行，具有很重要的价值。

第二次世界大战期间，英美等国将试验设计应用于工业试验，取得了良好效果。并在工业传统的多因素试验方法（因素轮换法）的基础上发表了多种“正交试验法”。虽然英美统计学家在试验设计的理论和应用中做了许多工作，但其中有不少形式主义和繁琐哲学的东西，应用起来很不方便。

第二次世界大战结束以后，日本、印度等国学者结合实际应用，将其中某些方法加以改造，进一步发展提倡正交表方法，虽然还带有一些繁琐成份，但已实用多了。

目前，正交试验法在日本的工程技术界使用比较普遍。日本统计学家在试验设计中，对于实际的应用，成绩很突出，已经把这方面的理论与方法运用到经济的各个领域。

我国在这方面研究起步较晚，但在近10年，试验设计愈来愈被人们所重视，试验设计方法和理论正在不断地被研究和应用。

§ 1. 3 随机化原则

R. 费喧在创建试验设计理论的过程中，提出了十分重要的“随机化”（Randomization）原则。他认为这是保证取得无偏估计的有效措施，也是进行可靠的显著性检验的必要基础。所以，他把随机化原则放在极其重要的地位。所谓随机化原则，在R. 费喧看来就是保证有意义的检验不因为没有除掉一些扰乱原因而损坏其检验正确的措施。按他在《试验设计》一书中所指出：“随机化多少有点像保险，它是一种对付也许会、

也许不会发生的那种扰乱，防患于未然之中的措施”。

他认为，严格地说来，那种也许会、也许不会发生的扰乱，即影响试验结果而又无法控制的因素，无论何时何地总是相当多的。比如，从袋中随机取球，实际上球的大小、形状、光滑度等不可能绝对相同。其它如掷骰子、抛铜币、抽签等同样如此。同样在农业实验站，所谓随机地从小麦田抽出若干条，来测量小麦的穗高，不同的人的操作就存在不同程度的偏倚。当然增加劳力、经费和时间或许能改善试验的结果，但 R. 费喧认为既为试验设计，其实质就在于事先探索有效和妥当的措施。

而所谓随机化就是提供机会均等的措施，它是通过把各种“处理”完全随机地配置在各试验单元上来完成的。具体办法是通过随机数字表 (Random unnumber) 来保证总体中每一元素有同等被抽取的机会，舍弃不合手续的观察值等，以防止抽取时的人为偏见。1938 年，R. 费喧和他的学生耶特斯合作，根据 19 世纪英国汤姆生所编制的 20 行对数表的资料，随机取出填入预先制作好的 50 页（每页的空格为 40×8 ）表格中，总共 16,000 字，这就是有名的 Fisher—Yates 随机数字表。这样，R. 费喧就把随机化原则以最明确、最具体化的形式引入统计工作与统计研究中。

§ 1.4 试验设计的必要性和基本原理

在实际工作中常会遇到这样的问题：产品质量是否受到配料、设备、人工等的差别的影响；农作物产量是否受肥料、土壤、品种等的差异的影响；商品需求量是否受到不同地区、

不同季节的影响等等。为了比较不同的生产方法，鉴定各个因素的作用，试验设计与方差分析就是一个有效的统计方法。

下面就试验设计的必要性及其基本原理作简略的一般性说明。

一、试验设计的必要性

在科学实验和生产实践的各个领域经常要作试验，其目的或者是要寻找某个特定过程中的规律，或者是要比较某个现象的若干不同条件的效果。

例如，某研究所希望对生产工业耐火砖的两种不同配料：配料 A 和配料 B 的抗强度进行比较。试验的目的是要确定哪种配料的抗强度更大。试验可按如下方法进行：分别用两种配料生产工业耐火砖若干块，然后对其进行测试，记录抗强度测试的数据，再比较两种配料的平均抗强度。

在这样一个试验问题中，自然会提出很多有关的重要问题。例如：

- (1) 生产该耐火砖的配料是否只有这两种；
- (2) 有没有其它因素会影响抗强度？如有的话，我们要不要研究它对抗强度的影响？如我们的试验中不准备研究这种影响，那么在试验过程中就必须对它加以控制；
- (3) 每种配料要多少耐火砖来作试验；
- (4) 不同类型的耐火砖用何种配料？按什么次序收集数据；
- (5) 分析数据时采用什么统计方法；
- (6) 两种配料的耐火砖平均抗强度有多大差别时才能被认为是重要的。

所有这些问题,也许还有其它一些问题在试验之前就必须作出明确回答。

这里要强调的是,分析数据所得到的结论是依赖于收集数据的方法。例如,比较两种配料制成耐火砖的抗强度试验问题中,耐火砖的个体差异(各耐火砖在炉中烧烤的温度、压强、时间等差异)是应该引起注意的。如果不同的温度、压强、烧烤时间适合于不同的配料,当你用平均抗强度作比较时,就不容易分辨出平均抗强度的差异中有多少是由配料不同引起的,有多少是由个体差异引起的。这正是上述第(2)个问题所提到的,在比较两种配料制成的耐火砖抗强度时,个体差异是一个“外来因子”,试验中应该适当控制个体差异,使两种配料对个体差异大致相同的耐火砖进行试验。

上述问题概括起来包括两个方面:一是如何合理地制订试验方案;二是如何科学地分析试验结果。这就是试验设计和方差分析所要研究的对象。

二、试验设计的基本原理

试验设计是数理统计中一个重要的分支,是关于试验与分析的统计理论。我们在制订了试验计划并实施以后,必然会收集到适当的数据,然后使用数理统计方法(即方差分析)分析出客观而正确的结论。试验问题的两个方面,即试验的设计与数据的统计分析是紧密相联的。试验设计应保证:试验次数尽量少,而试验结果包含的信息尽量多,同时在进行数据的统计分析时便于提取这些信息。数据的统计分析方法是由试验的设计方法相应的数学模型所决定的。

试验设计中的两个基本原理是重复和随机化。

所谓重复是指基本试验的重复。例如,比较两种配料制成

的耐火砖抗强度试验中，一块耐火砖就意味着一个重复，若每种配料都对 5 块耐火砖进行测试，则称这个试验有 5 个重复。重复有两个重要性质，一是可由重复试验的结果估计误差，只有将观察所得试验结果的差别与误差作比较，才能得出客观而正确的结论。例如，在测试两种配料耐火砖抗强度试验中，只有当平均抗强度的差别超过误差所容许的差别时，才能断言两种配料的抗强度不同，二是在用样本均值估计因子效应（例如配料引起的抗强度）时，重复可使估计精确。例如，设样本为 y_1, y_2, \dots, y_n , $V(y_j) = \sigma^2, j=1, 2, \dots, n$, 则 $V(\bar{y}) = \frac{\sigma^2}{n}$, 其中 $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$ 。由此可见一个观察的方差较大，若干个观察的均值方差大大缩小。因此，利用试验的重复，由平均抗强度比较两种配料，所得结论更为可靠。

随机化是试验设计中使用数理统计方法的基础。所谓随机化是指试验器材的布置和试验的进行次序是随机确定的。数理统计方法往往要求观察值（包含随机误差）是独立的随机变量，随机化保证了这种独立性。而且，采用随机化还可以使“外来因子”得到平衡。例如，在测试两种配料制成耐火砖抗强度时，假如某些耐火砖受到温度和压强较大，而另外一些耐火砖烧烤的时间较长，则在分析数据时会带来困难，这时个体差异和配料对抗强度的影响就混杂了。由于我们的试验目的是要比较两种配料，不准备研究个体差异的作用，采用随机化的方法能使两种配料都能适合各种不同差异的耐火砖，使两种配料在个体差异方面取得某种“平衡”，就可以抵消个体差异对平均抗强度的影响。

为了更好地进行试验的设计和分析，必须明确如下问题：

试验要研究什么问题,数据如何收集,所得数据又如何进行分析等等。完整地解决一个试验问题,包括三个方面:

(一) 试验

1. 问题的叙述。

把问题叙述清楚,粗看起来似乎很容易,实际上往往很简单。只有对所要研究的现象有相当了解,才能把试验的目的及所要解决的问题叙述清楚,才能有利于问题的解决。

2. 响应变量的选取。

变量试验的观察结果的量称为响应变量,相当于数学中的因变量,而试验中的因子相当于自变量。观察所得的响应变量的值提供了所要研究问题的有关信息,我们需要明确怎样去测量响应变量的值,测量的精度如何。

3. 因子的选取。

试验者感兴趣的对试验结果有影响的因素都称为试验问题中的因子。那些虽对试验结果有影响,但试验者不准备加以考虑或者试验过程中固定其状态的因素不成为该问题中的因子。例如,测试配料影响抗强度试验中,配料的品种是一个因子,但个体差异不是因子。

4. 诸因子的水平的选取。

因子在试验中所处的状态称为它的水平。选取因子的水平时要区分以下两点:

(1) 因子是定性的还是定量的。

如果因子的水平是用数量表示的,则称该因子为定量因子,如温度、压力、时间等都是定量因子。凡不是用数量表示水平的因子称为定性因子,例如,在测试配料影响抗强度的试验中,配料就是一个定性因子。

(2) 固定的或随机的。

试验过程中因子水平的选取可有两种方法,一种是事先选定,试验中按预定的要求控制。另一种是随机而定,试验中测定其水平。

5. 各因子的诸水平如何组合。

(二) 设计

1. 确定试验的总观察次数(即试验总次数)。

对试验方案的要求是总试验次数尽量少,包含的信息尽量多且达到较高的统计精度。这两者之间往往有矛盾,试验方案应使两者得到平衡。

2. 试验按怎样的次序进行。

3. 采用什么随机方法。

4. 写出描述试验的数学模型。

(三) 分析

1. 收集与整理数据。

2. 计算检验的统计量和参数的估计值。

3. 解释结果,提出建议。

对数据分析的结果作出合理的定性解释,并提出符合实际情况的建议。

以上就是解决一个实际的试验问题的全过程。本书主要讨论其中的统计方法。

习 题 一

1. 何为方差分析?它的作用和意义是什么?

2. 试验设计源于何处?这个理论的创始者是谁?目前在哪些方面