

·数理统计丛书·



王万中 茆诗松 编

试验的设计与分析

华东师范大学出版社

• 数理统计丛书 •

试验的设计与分析

王万中 茆诗松 编



华东师范大学出版社

封面设计 高山

试验的设计与分析

王万中 茆诗松 编

华东师范大学出版社出版发行

(上海中山北路 3663 号 邮政编码 200062)

新华书店上海发行所经销

上海市印刷七厂印刷

开本 850×1168 1/32 印张 15.75 字数 420 千字

1997 年 5 月第 1 版 1997 年 5 月第 1 次印刷

印数 001—2 000 本

ISBN 7-5617-1658-3/N·092

定价 22.00 元

总 序

数理统计是一门应用性很强的学科.它是研究如何有效地收集、整理和分析受随机影响的数据,并对所考虑的问题作出推断或预测,直至为采取决策和行动提供依据和建议的一门学科.凡是有大量数据出现的地方,都要用到数理统计.人口调查、税收预算、测量误差、出生与死亡统计、保险业中赔款额和保险金的确定等,这些数理统计早期主要研究的问题,直到现在仍值得认真研究.建立在现代数学和概率论基础上的数理统计,在近半个世纪以来,在理论、方法、应用上都有较大的发展.抽样调查、试验设计、回归分析与回归诊断、多元分析、时间序列分析、非参数统计、统计决策函数、统计计算、随机模拟、探索性数据分析等统计方法相继产生并在实践中普遍使用,把以描述为主的统计发展到以推断为主的统计.今天,数理统计的内容已异常丰富,应用面广量大,成为当前最活跃的学科之一.

我国科技和经济的发展,需要大量的经过系统训练的数理统计专业人才.最近数年中,国家教育委员会已在一些高校先后设立各种统计专业,这将为我国数理统计发展开创新的局面.为了促进数理统计人才的培养,国家教育委员会于1984年召开了“数理统计教学座谈会”,会上交流了各校培养数理统计人才的经验,同时还指出,组织国内专家编写和出版一套数理统计专业的教材是当务之急.

我们认为,一本好的数理统计教材,首先应讲清统计思想与统计方法,所讲的理论应是在实践中有用的统计方法所必需的理论,在阐明各种统计方法时,应给出足够的问题背景和有关的数据,能使学生对数理统计有一个系统、全面的认识,并培养学生对统计实

践的兴趣.另外,文字流畅,带有趣味性,适合大学生阅读,当然也是一本好教材的必要条件.

鉴于国内对数理统计教材的急需,我们约请国内一些颇有造诣的专家,按照上述对教材的特定要求,编写了这套“数理统计丛书”,作为高等院校数理统计专业的基本教材,并将由华东师范大学出版社陆续出版.《试验的设计与分析》即为该丛书之一.我们希望,这套丛书能对我们数理统计事业的发展起到一定的促进作用,能成为我国年轻一代统计学家的引路之石.

茆诗松

1986年11月于华东师范大学

序 言

《试验的设计与分析》是数理统计的一个重要分支,它研究正确地设计试验计划与分析试验数据的理论和方法.自从1935年英国著名统计学家费希尔(R. A. Fisher)的书“The Design of Experiments”出版以来,试验设计这门学科由于其应用的广泛性,倍受统计学家和实际工作者的重视,并获得迅速发展.近三十多年来,我国数理统计学者与广大工程技术人员在试验设计的理论研究与实际应用两个方面都取得了很大的进展,尤其是正交试验设计法得到了相当广泛的推广与普及.这就提出了在高等学校加强试验设计教学的要求.

本书是在华东师范大学数理统计系八次讲授试验设计的讲义基础上经过三次较大修改而成的.本书的取材强调实用,因此正交设计、参数设计、回归设计和平衡与部分平衡不完全区组设计是本书的主要内容.本书在方法叙述上力求便于操作,理论叙述上力求简明严谨,但以说明方法所需理论为限.所以本书不仅适宜作大学生的教材,也适宜实际工作者阅读与使用.

在本书编写过程中吸收了听课同学很多有益建议.得到了华东师大数理统计系教师很多帮助.曹良同志选编了部分习题,陈信漪和周纪芎同志仔细审阅了全书,并提出很多宝贵意见.在此对他们表示衷心的感谢.茆诗松同志编写了第四章,其余各章由王万中同志编写.

限于水平,错误与不当之处在所难免,恳请专家、读者给予批评指正,我们不胜感激.

编者 1995.4

目 录

引言.....	1
第一章 单因子试验.....	7
§ 1 单因子试验的统计模型	7
1.1 完全随机化设计	7
1.2 试验的统计模型	7
§ 2 固定效应模型的统计分析.....	10
2.1 方差分析.....	10
2.2 参数估计.....	21
2.3 等重复情形.....	31
2.4 从回归角度看方差分析中的因子平方和.....	36
§ 3 多重比较方法.....	39
3.1 对比.....	39
3.2 邓肯多重比较法.....	44
3.3 谢菲多重比较法.....	48
§ 4 随机效应模型.....	52
4.1 试验设计与统计模型.....	52
4.2 统计分析.....	53
§ 5 模型恰当吗?	57
5.1 方差齐性检验.....	57
5.2 正态性检验.....	59
5.3 非齐性方差数据的变换.....	60
习题一	61
第二章 多因子试验引论	66
§ 1 两因子试验的统计模型.....	66
§ 2 固定效应模型的统计分析.....	70
2.1 可加效应模型的统计分析.....	70

2.2 交互效应模型的统计分析	82
§ 3 随机效应模型与混合模型的统计分析	94
3.1 随机效应模型的统计分析	94
3.2 混合模型的统计分析	99
§ 4 多因子试验的设计与分析	101
4.1 统计模型	101
4.2 固定效应模型的统计分析	103
4.3 随机效应模型与混合模型的方差分析	115
§ 5 拉丁方设计与正交拉丁方设计	117
5.1 拉丁方设计及其统计模型	117
5.2 统计分析	120
5.3 希腊-拉丁方设计	127
习题二	134
第三章 析因试验的部分实施与正交表	140
§ 1 2^k 设计的部分实施	140
1.1 2^2 设计与正交表 $L_4(2^3)$	140
1.2 2^3 设计与正交表 $L_8(2^7)$	144
1.3 2^k 设计与正交表 $L_{2^k}(2^{2^k-1})$	150
1.4 例	156
§ 2 3^k 设计的部分实施	161
2.1 3^2 设计与正交表 $L_9(3^4)$	161
2.2 3^3 设计与正交表 $L_{27}(3^{13})$	165
2.3 3^k 设计与正交表 $L_{3^k}(3^{\frac{3^k-1}{2}})$	172
2.4 例	174
§ 3 p^k 设计的部分实施	181
3.1 p^k 设计与正交表 $L_{p^k}(p^{\frac{p^k-1}{p-1}})$	181
3.2 p^k 设计的部分实施	182
3.3 关于正交设计	183

§ 4 正交表的并列	183
4.1 并列的方法	183
4.2 例	184
4.3 进一步的例	194
§ 5 拟水平法	197
5.1 试验设计	197
5.2 统计分析	198
5.3 含交互作用的例	204
§ 6 赋闲列法	205
6.1 赋闲列	205
6.2 统计分析	208
6.3 一个实例	210
6.4 含交互作用的例	216
习题三	217
第四章 参数设计	223
§ 1 田口的基本思想	223
§ 2 稳健性设计与分析	225
§ 3 灵敏度分析	240
§ 4 综合噪声因子	245
§ 5 动态特性的参数设计	254
5.1 动态特性	254
5.2 信号因子	255
5.3 动态特性参数设计的要求	256
5.4 动态特性参数设计的试验安排	257
5.5 SN 比的估计	258
5.6 动态特性的参数设计	263
习题四	272
第五章 不完全区组设计	278
§ 1 平衡不完全区组设计	278

1.1	平衡不完全区组设计的概念	278
1.2	平衡不完全区组设计的参数间的关系	281
1.3	互补设计、导出设计、剩余设计	285
§ 2	平衡不完全区组设计的统计分析(区组内分析)	290
2.1	参数估计	290
2.2	方差分析	294
§ 3	平衡不完全区组设计的统计分析(区组间分析)	301
§ 4	部分平衡不完全区组设计	306
4.1	问题的提出	306
4.2	结合类和部分平衡不完全区组设计的概念	307
4.3	统计分析	311
§ 5	尤登方设计	317
	习题五	321
第六章 回归设计与响应曲面分析		324
§ 1	正交回归设计的概念	326
1.1	编码变换	326
1.2	正交回归设计的定义	327
1.3	线性回归正交设计的统计分析	330
§ 2	用正交表构造线性回归的正交设计	334
2.1	使用正交表构造试验设计、作统计分析	334
2.2	添加中心点的重复试验	340
§ 3	用单纯形法构造线性回归的正交设计	344
3.1	单纯形的概念	344
3.2	第一种方法	345
3.3	由正交矩阵构造单纯形设计	347
§ 4	旋转回归设计的概念	349
§ 5	多项式回归的试验设计的旋转性条件	352
5.1	多项式回归的设计的信息矩阵元素的一般形式 ..	353
5.2	旋转性条件	355

§ 6 二次回归的旋转设计	359
6.1 二次回归旋转设计的试验点必须处于不同球面 ...	359
6.2 两个自变量的二次回归旋转设计	361
6.3 二次回归的旋转中心组合设计	362
6.4 二次回归的均匀精度旋转中心组合设计	372
§ 7 最速上升法	375
7.1 最速上升法的步骤	375
7.2 最速上升路线的确定	376
§ 8 二次响应曲面分析	379
8.1 响应曲面的等高线表示法	379
8.2 稳定点	380
8.3 二次回归方程的典范形式	383
习题六	390
第七章 最优设计	393
§ 1 设计的概念,信息矩阵的性质	393
1.1 模型	393
1.2 设计的概念	394*
1.3 信息矩阵的性质	398
§ 2 优良性准则	401
2.1 D 最优性	401
2.2 A 最优性	403
2.3 线性最优性	404
2.4 G 最优性	405
2.5 E 最优性	407
§ 3 等价性定理	408
3.1 引理	408
3.2 等价性定理	411
3.3 用等价性定理验证设计的 D 最优性	415
§ 4 费多洛夫迭代算法	418

习题七	423
附录 方差分析中的有关分布	424
§1 多维正态分布	424
§2 χ^2 分布	425
2.1 χ^2 分布的概念	425
2.2 χ^2 分布的基本性质	429
§3 正态变量的二次型	429
3.1 正态变量的二次型服从 χ^2 分布的条件	429
3.2 正态变量的二次型的独立性	433
3.3 正态变量的二次型与线性型独立的条件	436
§4 t 分布	438
4.1 t 分布的概念	438
4.2 t 分布的基本性质	439
§5 F 分布	442
5.1 F 分布的概念	442
5.2 F 分布的基本性质	442
参考书目	444
附表	446
1 正态分布表	446
2. χ^2 分布的上侧分位数(χ^2_α)表	450
3. t 分布表	452
4. t 分布的双侧分位数(t_α)表	454
5. F 检验的临界值(F_α)表	456
6. 邓肯多重比较的显著性极差的系数	466
7. 多重比较中的 S 表	468
8. 正交表	470
9. 随机数表	486

引 言

我们要说明一下《试验的设计与分析》这门学科的研究对象,简单地解释试验设计的两个重要手段的基本作用,介绍解决试验问题的主要步骤.

一、研究 对 象

在科学的各个领域和生产的各行各业经常要作试验. 试验的目的可能是要比较某个现象中各个因素的重要性以及它们的不同状态的效果;也可能是要寻找某个特定过程中各个变量之间的数量规律.

例一 某个医师希望对治疗扁桃腺炎的两种不同的药物——洁霉素和青霉素的疗效进行比较. 试验的目的是要确定哪种药物的疗效好. 试验可按如下方法进行,选取扁桃腺炎患者若干人,分别用两种药物治疗,并记录疗效. 然后比较两种药物的平均疗效.

例二 某个化工工程师希望研究某种产品的产量和化学反应的温度、化学反应的时间、催化剂的用量之间的关系. 试验的目的是要寻找最佳生产条件. 试验可按如下方法进行,选用三种反应温度,例如 60°C 、 70°C 、 80°C ; 选用三种反应时间,例如 1 小时、1.5 小时、2 小时; 选用三种催化剂的用量,例如 2 公斤、2.5 公斤、3 公斤. 对温度、时间、用量的 27 种组合的每一种作一次试验,记录产品的产量,根据试验结果寻找产量与温度、时间、用量之间的数量规律,然后根据所找到的规律确定最佳生产条件.

我们将衡量试验结果好坏的指标称为**响应变量**,例一中的疗效与例二中的产量就分别是这两个试验问题的响应变量. 很多情

况下的响应变量可用数值表示,称其为**定量的响应变量**,例二中的产量就是一个定量的响应变量.有时响应变量不是用数值表示的,称其为**定性的响应变量**,例一中的疗效可用显效、有效、无效表示,它是一个定性的响应变量.有些试验问题还可能包含不止一个响应变量.但是,本书只讨论单响应变量的情况,并且主要讨论定量的响应变量情形.

我们将影响响应变量的因素称为试验问题中的**因子**.例一那个试验问题中有一个因子,即药物的品种 A ;例二那个试验问题中有三个因子,即反应温度 A 、反应时间 B 和催化剂用量 C .在试验中因子所处的各个状态称为**因子的水平**.例一中的洁霉素 A_1 与青霉素 A_2 就是药物品种 A 这个因子的两个水平;例二中 $A_1: 60^\circ\text{C}$ 、 $A_2: 70^\circ\text{C}$ 、 $A_3: 80^\circ\text{C}$ 就是反应温度 A 这个因子的三个水平.因子也有**定量因子**与**定性因子**之分,例一中的那个因子是定性因子,例二中的三个因子都是定量因子.

解决任何一个试验问题都有三个阶段:制订试验计划;实施试验计划,记录试验结果;分析试验数据.《试验的设计与分析》这门学科的研究对象就是第一、三两个阶段中的数学和统计问题.具体地说,就是研究如何合理地制订试验计划(设计问题)和如何科学地分析试验结果(分析问题).

设计问题与分析问题之间是有制约关系的.对设计的要求是省(人力、财物、时间、……等),即试验次数要尽量少;而所包含的有用信息要尽量多,并且能有方便的分析试验结果的方法.分析方法又是依赖于设计方法的,不同的设计方法的试验结果需要采用不同的分析方法.

二、试验设计中的两个基本手段

为了保证试验结果中包含尽量多的有用信息,并且能够方便地将它们分析提取出来,在对试验作设计时,总要采用如下两种基

本手段。

1. 重复

重复是指一个基本试验(或试验条件)重复进行若干次,即对应着某因子的诸水平或者某些因子的诸水平组合重复进行若干次试验。由于使用了重复这一手段,在分析试验结果时,就可以对误差作出估计,当因子诸水平或若干因子的诸水平组合的效应之间的差异超过误差时,我们才能对因子的诸水平或若干因子的诸水平组合的优劣作出比较和选择。而且,在用样本均值去估计效应时,由于使用了重复这一手段,可使估计更加准确,重复次数越多,估计量的方差越小。当然在作试验设计时,不能只追求分析试验结果中估计量的精度,还要考虑到重复带来试验时间延长、试验经费提高等问题。因此试验设计应该使两者取得平衡。

2. 随机化

随机化是指试验仪器、材料、人员……等的布置要随机地确定,诸试验单元的执行顺序要随机地确定。由于使用了随机化这一手段,就给分析试验结果时提供了使用统计方法的基础。试验结果是一个个随机变量,随机化保证了它们的独立性,在对效应作检验或估计时,就可以应用数理统计学中有关独立样本的基本理论。另一方面,在分析试验结果时,随机化还可以有效地消除“外来因子”对数据的干扰。例如,在药物疗效试验中,不使用随机化的话,可能导致一种药只对中青年扁桃腺炎患者使用,另一种药只对老人、幼儿扁桃腺炎患者使用,使得个体差异这一“外来因子”和药物对疗效的影响混杂了。给数据分析带来了困难。使用随机化这一手段,可以大体上保证两种药物都用到各种年龄、不同性别、各种健康状况的扁桃腺炎患者身上,使得两种药物在个体差异方面取得某种“平衡”,在对两种药物的平均疗效作比较时,消除个体差异这一“外来因子”对数据的干扰。

三、解决试验问题的主要步骤

概括地说,完整地解决一个试验问题的主要步骤如下:

1. 问题的叙述

粗看起来,将问题叙述清楚并不困难,实际上很不简单,只有对所要研究的现象有相当的了解,才能把试验目的,所要解决的具体问题叙述清楚. 主要包括:

(1) 响应变量的选取. 选定的响应变量不仅要与试验目的相一致,而且其分布还应该与试验的统计模型的假定一致或非常近似. 因此,确定响应变量是一件需要仔细推敲的事. 我们需要区分响应变量是定性的还是定量的. 更要决定测量响应变量的方法,要了解测量的精度.

(2) 因子的选取. 首先要将对响应变量有影响的因素尽量罗列出来. 然后根据试验目的将那些对响应变量影响很小的因素作为误差因子,在试验中不必控制而任其随机变化;将那些对响应变量影响较大而又不准备考察其影响的因素,在试验中或对它们加以控制、或制订一种试验计划使能保证在分析试验结果时可以消除它们的影响,试验中也不对它们加以控制,即不把它们作为试验问题中的因子. 只有那些对响应变量影响大,又希望通过试验对它们的影响加以研究、比较的因素才当作试验问题中的因子. 一个试验问题中因子的个数对试验次数是有影响的,选取因子时,也要顾及到试验次数能否承受得了.

(3) 各因子诸水平的选取. 像选定因子一样,确定各因子的诸水平对解决试验问题同样重要,既要根据试验目的和实践经验,又要注意因子水平的多少对总试验次数的影响更大. 因此,要根据试验问题的需要和经费、人力、时间的许可来确定各因子的诸水平. 这里还要注意两点:

(i) 因子是定性的还是定量的. 如果是定量的, 还要考虑在试验中如何控制它们的水平, 控制的精度如何.

(ii) 因子是固定的还是随机的. 如果试验中某因子的诸水平是按试验人的主观意图选定的, 则称该因子是**固定的**; 如果试验中某因子的诸水平是随机确定的, 则称该因子是**随机的**. 一个因子是随机的还是固定的, 这要由试验目的来确定.

一般都将各因子的诸水平列成一张因子水平表, 使人一目了然.

2. 试验计划的设计与实施

(1) 确定总试验次数. 在能包含试验目的所需要的尽可能多的信息并保证分析试验结果时的统计精度的前提下使试验次数尽量地少.

(2) 各因子诸水平怎样组合. 要明确试验计划中必须包含那些水平组合, 有时还要注意避免那些试验中不能实施的水平组合.

(3) 试验按怎样的顺序进行, 采用什么随机化的方法.

(4) 采用什么样的统计模型描述试验.

进行试验时, 要严格监控试验计划的要求得到实现, 并准确记录试验结果.

3. 试验结果的统计分析

(1) 搜集试验数据, 并作适当整理.

(2) 计算统计假设检验中的统计量和模型中诸参数的估计

(3) 对统计分析的结果作出科学而符合实际的解释, 并提出建议.

这就是解决一个实际试验问题的完整步骤. 其中一部份工作是由试验目的和实践经验决定的, 例如响应变量的确定、诸因子的确定、诸因子的诸水平的确定或诸因子水平范围的确定等等不属