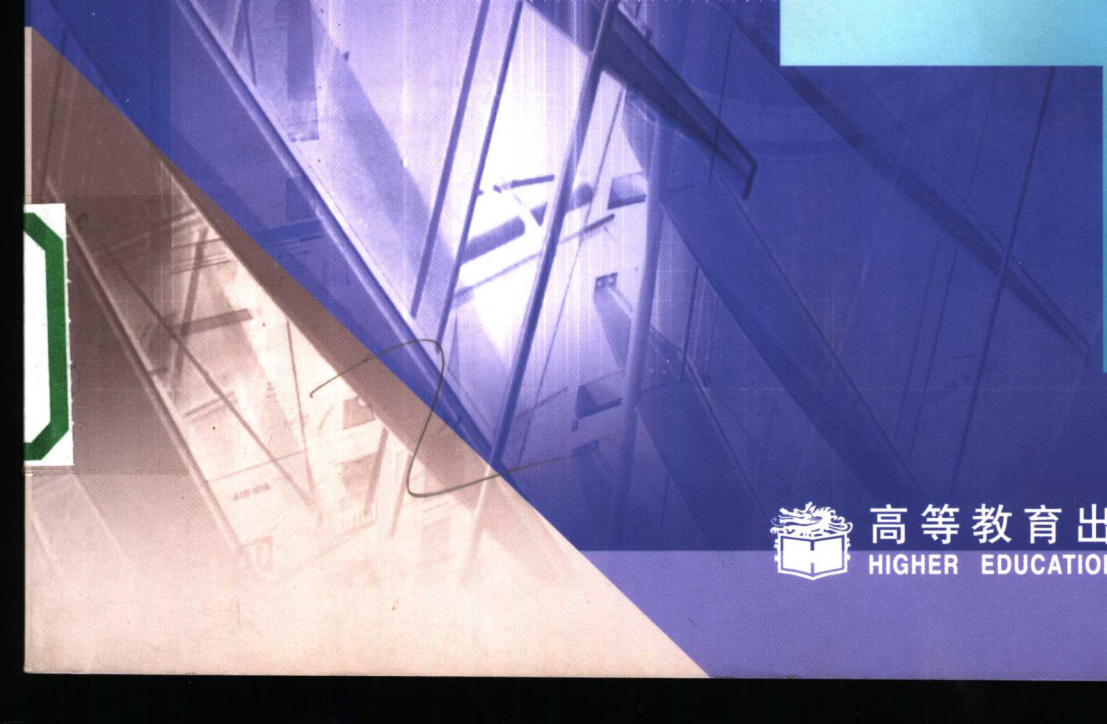
普通高等教育“十五”国家级规划教材

试验的设计与分析

王万中　主编



高等教育出版社
HIGHER EDUCATION PRESS

内 容 提 要

本书是自1987年起在华东师范大学统计系连续开设试验设计课程的教材基础上，经不断修改后，于1997年在华东师范大学出版社出版，同时，还在几个兄弟院校使用。近年来，在不断总结教学的基础上又做了较大修改，作为普通高等教育"十五"国家级规划教材出版。

本书是大学统计系本科专业基础课的教材，选材重点集中，包括单因子试验理论、多因子试验引论、析因试验的部分实施与正交表、参数设计、不完全区组设计、回归设计与响应曲面分析、最优设计，全书理论叙述较为严谨而且突出使用的可操作性，本书不仅可作为大学生的教材，又可供实际工作者参考使用。

图书在版编目(CIP)数据

试验的设计与分析/王万中主编. —北京：高等教育出版社，2004.6
ISBN 7-04-014366-6

Ⅰ.试… Ⅱ.王… Ⅲ.①试验设计(数学)-高等学校-教材②试验分析(数学)-高等学校-教材
Ⅳ.O212.6

中国版本图书馆CIP数据核字(2004)第016235号

策划编辑 李 蕊 **责任编辑** 高尚华 **封面设计** 王凌波 **责任绘图** 黄建英
版式设计 胡志萍 **责任校对** 尤 静 **责任印制** 孔 源

出版发行	高等教育出版社	**购书热线**	010-64054588
社 址	北京市西城区德外大街4号	**免费咨询**	800-810-0598
邮政编码	100011	**网 址**	http://www.hep.edu.cn
总 机	010-82028899		http://www.hep.com.cn
经 销	新华书店北京发行所		
印 刷	北京铭成印刷有限公司		
开 本	787×960 1/16	**版 次**	2004年6月第1版
印 张	27.75	**印 次**	2004年6月第1次印刷
字 数	520 000	**定 价**	34.40元

本书如有缺页、倒页、脱页等质量问题，请到所购图书销售部门联系调换。

序　言

“试验的设计与分析”是数理统计学的一个重要分支，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，备受统计学家和实际工作者的重视，并获得迅速发展。四十多年来，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。

本书是编者1986年起在华东师范大学统计系讲授试验设计的讲义基础上经过三次较大修改而成。本书取材强调实用，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。

这次作为教育部十五规划教材由高等教育出版社出版，编者又对全书做了仔细的审阅，改正了书中的错误，并增补了实用的两项新内容——调优操作和均匀设计。本书第六章中“均匀设计”一节(§9)由曾林蕊同志编写。第四章由茆诗松同志编写，其他部分都由王万中同志编写。

在本书编写过程中吸收了听课同学很多有益的建议。得到了华东师范大学统计系教师很多的帮助。陈信漪和周纪芗同志仔细审阅了全书，并提出很多宝贵意见。高等教育出版社和华东师范大学教务处都给予了极大支持和帮助。在此我们一并表示衷心感谢。

编　者

郑重声明

目　录

引　言

我们要首先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，并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 1　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 2　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℃、70℃、80℃；选用三种反应时间，例如 1 h、1.5 h、2 h；选用催化剂的三种用量，例如 2 kg、2.5 kg、3 kg.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 1 中的疗效与例 2 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 2 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 1 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 1 那个试验问题中有一个因子、即药物的品种 A；例 2 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 1 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 2 中 A_1:60 ℃、A_2:70 ℃、A_3:80 ℃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1中的那个因子是定性因子，例2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……)，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相应不同的分析方法.

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……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，实际上往往很不简单，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哪些水平组合,有时还要注意避免哪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量.

(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.其中一部分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本书讨论范围.本书主要讨论两个问题.一个是试验的设计问题.在比较试验的情形(目的是通过对诸水平或水平组合作出统计比较并得出最优水平或最优水平组合),确定要实施的水平的重复次数或水平组合及其重复次数,诸试验单元的实施顺序;在回归试验的情形(目的是在建立响应变量与诸因子间的数量联系以后,确定最优条件),由所给诸因子的试验范围确定诸因子的水平组合.另一个是讨论试验数据的统计分析方法.本书各章针对不同试验问题讨论解决上述两个问题的原理和方法.

第一章　单因子试验

本章将讨论单因子试验的设计与分析方法.单因子试验在实际生活中是经常碰到的.它不仅简单,而且典型,还涉及试验设计中许多基本概念与基本方法,也是后面学习多因子试验的基础.

§1　单因子试验的统计模型

1.1　完全随机化设计

在单因子试验问题中,应如何设计试验计划呢?设因子 A 共有 a 个不同的水平,分别记为 $A_1,A_2,\cdots,A_a$.有时,水平又称为处理.试验的目的在于比较这 a 个水平的优劣.如果试验中不存在误差或误差极其微小,则由各个水平下的一个试验的观察值即可对诸水平作出比较.但是在实际上,每个试验的观察值中不可避免地都存在着随机误差.因此,应通过试验的重复获得误差的估计,然后再对诸水平作比较.设在水平 A_i 下作 n_i 次重复试验,$i=1,\cdots,a$,则总共要作 $n=\sum_{i=1}^{a} n_i$ 个试验.为了使得原料、设备……方面对所采用的诸水平而言尽可能保持平衡,这 n 个试验应按随机顺序进行.这种试验设计的方法称为**完全随机化设计**.这里的随机化通常可借助于随机数表或抽签的方法实现.

1.2　试验的统计模型

设 n 个试验的数据如表 1.1 所示:

表 1.1　单因子试验的数据

水　平	观　察　值			
A_1	y_{11}	y_{12}	$\cdots$	y_{1n_1}
A_2	y_{21}	y_{22}	$\cdots$	y_{2n_2}
$\vdots$	$\vdots$	$\vdots$		$\vdots$
A_a	y_{a1}	y_{a2}	$\cdots$	y_{an_a}

其中 y_{ij} 表示在第 i 个水平 A_i 之下第 j 次重复试验的观察值.在统计学中,通常

假设一个数据 y 是由两部分组成的，其一是因子 A 的影响部分 μ，它是一个通常的变量，随因子 A 的水平变动而变动；另一是试验的随机误差 ε. 即

$$y=\mu+\varepsilon.$$

现因子 A 有 a 个水平，每个水平对数据 y 的影响是不同的，分别记为 $\mu_1,\mu_2,\cdots,\mu_a$. 试验的随机误差在每次试验里的观察值也是不同的，但常假设它们来自同一个正态总体 $N(0,\sigma^2)$，其中 σ^2 是未知的. 因此，数据 y_{ij} 有如下结构：

$$\begin{cases} y_{ij}=\mu_i+\varepsilon_{ij}, \\ \text{诸 } \varepsilon_{ij} \quad \text{i.i.d.} N(0,\sigma^2), \end{cases} \quad i=1,\cdots,a,\ j=1,\cdots,n_i, \tag{1.1}$$

其中 i.i.d. 表示独立同分布. 根据模型(1.1)可知，$y_{ij}\sim N(\mu_i,\sigma^2)$，$i=1,\cdots,a$，$j=1,\cdots,n_i$，并且诸 y_{ij} 相互独立. 这也说明，同一水平 A_i 下的 n_i 个重复试验的结果是来自同一正态总体的一个样本. 因此，在模型(1.1)所描述的试验问题中，我们涉及到 a 个正态总体，它们的均值是不同的，而它们的方差是相同的.

下面，我们将统计模型(1.1)中数据的结构写成意义更加清晰的形式. 记

$$\mu=\frac{1}{n}\sum_{i=1}^{a}n_i\mu_i, \tag{1.2}$$

其中 $n=\sum_{i=1}^{a}n_i$. 称 μ 为**一般平均**，又称**综合平均**，它表示全部数据(诸 y_{ij})的均值的算术平均. 记

$$\tau_i=\mu_i-\mu,\ i=1,\cdots,a. \tag{1.3}$$

称 τ_i 为因子 A 的第 i 个水平 A_i 的**效应**，它表示第 i 个水平 A_i 相应的第 i 个总体的均值比一般平均大(或小)多少. 由表示式(1.2)与(1.3)，不难看出，诸 τ_i 受到约束：

$$\sum_{i=1}^{a}n_i\tau_i=0. \tag{1.4}$$

于是，统计模型(1.1)可改写成

$$\begin{cases} y_{ij}=\mu+\tau_i+\varepsilon_{ij}, \\ \text{诸 } \varepsilon_{ij} \quad \text{i.i.d.} N(0,\sigma^2), \\ \text{约束条件：} \sum_{i=1}^{a}n_i\tau_i=0. \end{cases} \quad i=1,\cdots,a,\ j=1,\cdots,n_i. \tag{1.5}$$

观察值 y_{ij} 是一般平均、第 i 个水平 A_i 的效应和随机误差三者之和. 统计模型(1.5)将单因子试验的数据表 1.1 中的数据结构表示得更加清晰.

大家知道，一个试验中往往包含很多因子，它们的水平变动都会影响试验的结果. 有的影响大，有的影响小. 在预计影响大的因子中，如果我们感兴趣的只是其中的一个因子 A，那么在试验中就将那些预计影响大的其他因子控制在某固

定水平上，而让因子 A 在设定的诸水平 A_i，$i=1,\cdots,a$ 上变动，至于那些预计影响小的因子则任其在试验中随机变化. 这就成为单因子试验. 统计模型(1.5)中的 μ 就是那个预计影响大的因子参与试验的诸水平下试验数据的均值的算术平均；诸 τ_i 的大小则反映出因子 A 的诸水平 A_i 对数据影响的差异；而随机误差是预计影响小的那些因子的随机影响和预计影响大的那些因子(包括因子 A)在控制水平时的随机漂移的影响之总和. 由于(1.5)式中的参数 μ 和诸 τ_i 都是一次，所以(1.5)式表示的是一个**线性统计模型**. 单因子试验的统计模型是一般线性统计模型理论的特殊情形(参数 μ 与诸 τ_i 的系数都是 1).

在讨论单因子试验的统计模型(1.1)或(1.5)时要区分两种不同的情况，如果因子 A 的 a 个水平是根据试验的具体情况在试验前由试验人按主观意图指定好的，并在试验中得到很好的控制. 这种情况下的统计模型称为**固定效应模型**. 我们希望估计 μ、诸 τ_i 和 σ^2 并检验有关诸 τ_i 的各种假设，结论只适用于试验所使用过的因子 A 的各个水平，而不能推广至试验中未使用过的水平. 如果因子 A 的 a 个水平是从因子 A 的全部可能的水平这一总体中选取的一个随机样本，这种情况下的统计模型称为**随机效应模型**. 对它的进一步解释和讨论留到本章 §4 进行.

§2 固定效应模型的统计分析

2.1 方差分析

2.1.1 统计假设

我们先讨论方差分析问题. 方差分析的目的是要比较 a 个均值 $\mu_1,\mu_2,\cdots,\mu_a$ 是否相等，即比较 a 个效应 $\tau_1,\tau_2,\cdots,\tau_a$ 是否相等. 这是一个假设检验的问题. 一个可供选择的办法是对任意的 $i\neq j$，检验假设 $H_0:\tau_i=\tau_j$. 一共需要检验 C_a^2 个不同的假设. 如果检验其中的一个假设，犯第一类错误的概率是 $\alpha=0.05$，则正确地接受这个假设的概率是 0.95. 当 $a=5$ 时，共需检验 $C_5^2=10$ 个不同的假设. 如果这 10 个检验是相互独立的，则同时正确地接受这 10 个假设的概率等于 $(0.95)^{10}=0.60$，错误地拒绝这 10 个假设中的至少一个假设的概率等于 $1-0.60=0.40$. 这大大增加了犯第一类错误的概率. 因此，我们不能用这种分别检验上述 C_a^2 个假设的办法来解决比较 a 个均值的问题. 为了控制犯第一类错误的概率，我们应直接检验如下假设：

$$\begin{cases} H_0:\tau_1=\tau_2=\cdots=\tau_a=0, \\ H_1:\tau_i\neq 0 \quad \text{至少对一个 } i \text{ 成立}. \end{cases} \tag{2.1}$$

这就是单因子试验的方差分析所要检验的假设.检验这个假设的关键是作出一个适当的检验统计量,为此,我们将偏差平方和加以分解.

2.1.2 偏差平方和的分解

我们先引入如下记号:

$$y_{..}=\sum_{i=1}^{a}\sum_{j=1}^{n_i}y_{ij},\quad \bar{y}_{..}=\frac{y_{..}}{n}. \tag{2.2}$$

全部数据之间的差异可用下述**总偏差平方和**表示:

$$SS_T=\sum_{i=1}^{a}\sum_{j=1}^{n_i}(y_{ij}-\bar{y}_{..})^2. \tag{2.3}$$

引起诸数据 y_{ij} 之间差异的原因有二:一是因子 A 的 a 个水平不同,另一是试验误差.为了区分并比较这两个原因对数据的影响,需要将总偏差平方和 SS_T 进行分解,为此引入记号

$$y_{i.}=\sum_{j=1}^{n_i}y_{ij},\quad \bar{y}_{i.}=\frac{y_{i.}}{n_i},i=1,\cdots,a. \tag{2.4}$$

于是有

$$\begin{aligned}SS_T&=\sum_i\sum_j[(y_{ij}-\bar{y}_{i.})+(\bar{y}_{i.}-\bar{y}_{..})]^2\\&=\sum_i\sum_j(y_{ij}-\bar{y}_{i.})^2+\sum_i n_i(\bar{y}_{i.}-\bar{y}_{..})^2+\\&\quad 2\sum_i\sum_j(y_{ij}-\bar{y}_{i.})(\bar{y}_{i.}-\bar{y}_{..}).\end{aligned}$$

由表示式(2.4)可知交叉乘积项之和

$$\sum_i\sum_j(y_{ij}-\bar{y}_{i.})(\bar{y}_{i.}-\bar{y}_{..})=0,$$

故有

$$SS_T=\sum_i\sum_j(y_{ij}-\bar{y}_{i.})^2+\sum_i n_i(\bar{y}_{i.}-\bar{y}_{..})^2. \tag{2.5}$$

这就是著名的**偏差平方和分解公式**.下面解释右端两个平方和的含义.

在第一个平方和中,$\bar{y}_{i.}$ 表示第 i 个水平 A_i 下的数据的平均值,第 i 个水平 A_i 下诸数据 $y_{ij},j=1,\cdots,n_i$ 之间的差异可用偏差平方和 $\sum_{j=1}^{n_i}(y_{ij}-\bar{y}_{i.})^2$ 表示,它反映出第 i 个水平 A_i 下随机误差对数据的影响,在全部水平下,误差的影响之和是

$$\sum_{i=1}^{a}\sum_{j=1}^{n_i}(y_{ij}-\bar{y}_{i.})^2.$$

根据数据结构式(1.5)计算一下,上式的意义就更加清楚了.由于 $y_{ij}=\mu+\tau_i+$