

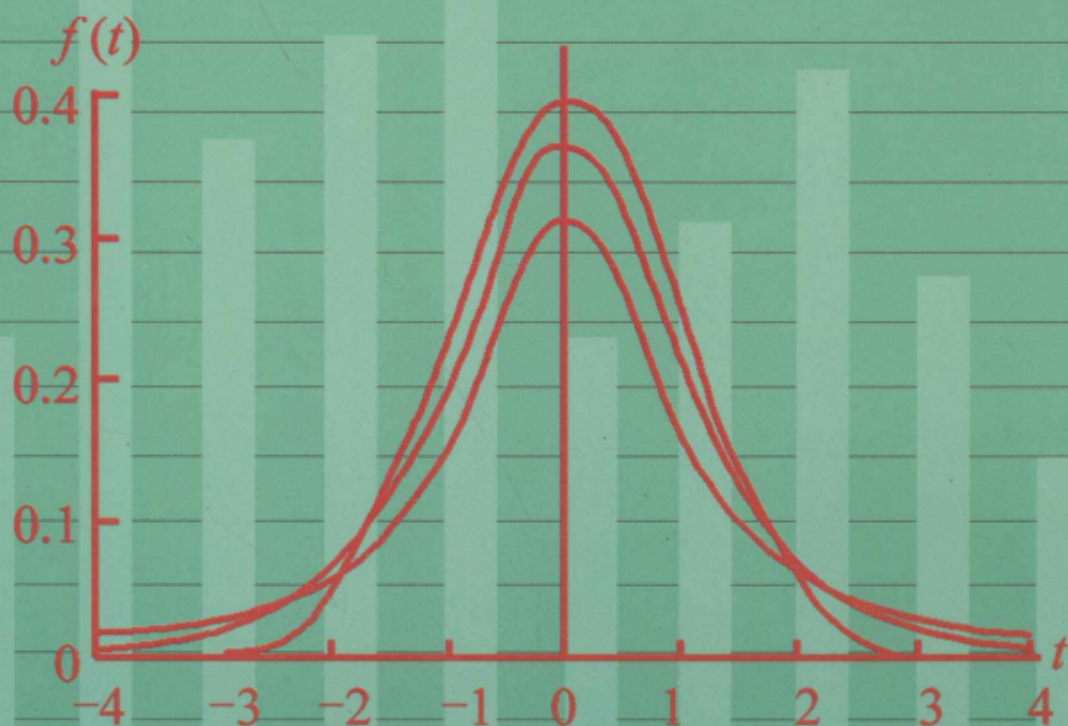
面向 21 世纪高等学校教科书

试验统计学

(第二版)

区靖祥 编著

 广东高等教育出版社



责任编辑：黄冬萍
责任技编：小 邱
封面设计：智 慧

ISBN 978-7-5361-3954-1



9 787536 139541 >

定价：28.00元

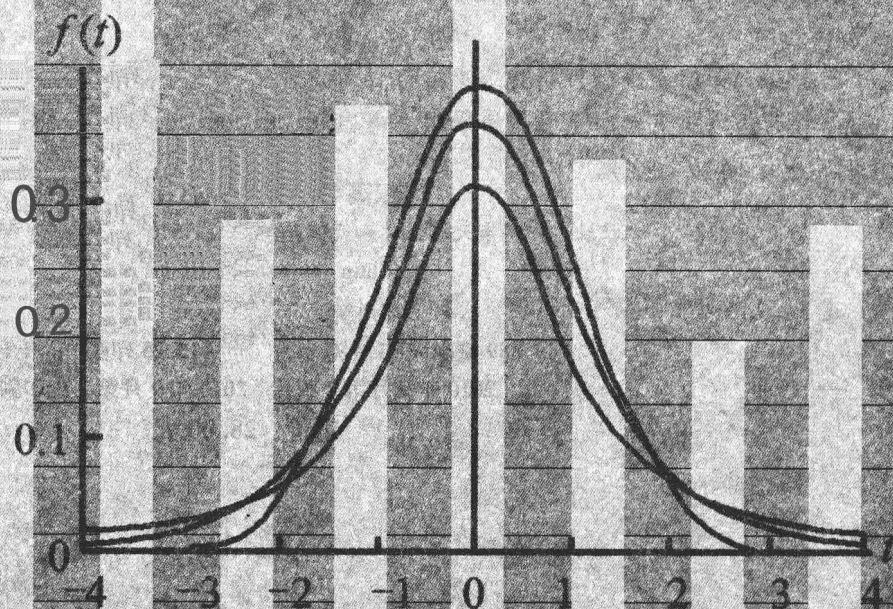
面向 21 世纪高等学校教科书

试验统计学

(第二版)

区靖祥 编著

广东高等教育出版社
广州



图书在版编目 (CIP) 数据

试验统计学/区靖祥编著. —2 版. —广州: 广东高等教育出版社, 2010. 9
ISBN 978 - 7 - 5361 - 3954 - 1

I. ①试… II. ①区… III. ①试验统计 - 统计学 - 高等学校 - 教材
IV. ①O212. 6

中国版本图书馆 CIP 数据核字 (2010) 第 145313 号

广东高等教育出版社出版发行

社址: 广州市天河区林和西横路 邮编: 510500 营销电话: 87557232

<http://www.gdgjs.com.cn>

广州新明光印刷有限公司印刷

889 毫米×1 194 毫米 1/16 15.25 印张 387 千字

2010 年 9 月第 2 版 2010 年 9 月第 2 次印刷

定价: 28.00 元

前 言

鉴于目前综合性大学和师范类高校中采用的生物统计学课本较少涉及试验设计的内容,大部分农业院校的专业也早已突破“农”的范围,开拓出更为宽广的专业领域,过去使用的田间试验设计与统计分析教材已不敷应用。作者编写这本试验统计学教材,举出了大量农学、林学、茶学、蚕学、园艺学、植物保护、环境资源、生物技术、动物科学、食品科学,甚至医学和工业方面的例子和习题,为学生将统计方法应用到这些领域提供了示范。本书可作为有关专业学生的专业基础课教材,也可用作有关研究领域的试验工作者的参考工具。

本书共分十章,每两章为一个单元。第一单元解释了科学试验的基本概念和方法;第二单元陈述了数据的整理方法、次数分布和概率分布;第三单元介绍了总体参数的置信区间和各种统计假设的测验;第四单元讨论了方差分析及各种设计的试验资料的分析方法;第五单元阐明了直线回归、相关和协方差分析的方法。

本书模仿了美国统计学家李景钧先生(C. C. Li)在编写 *Introduction to Experimental Statistics* (《试验统计学概论》) 中的写作风格。李先生在他的书中贯彻了这样的原则:他说,“我们在学解一元二次方程的时候,不会因为实践意义而先去解这样的方程: $683\,125x^2 + 1\,268.407x - 213.698\,25 = 0$,先学习解像 $x^2 + 3x - 2 = 0$ 这样的方程将会有更大的启发性”。在本书中,作者也遵照这样的原则,对于每一种统计方法,都举一个很简单的例子,并从头到尾详细解述各个计算步骤。学生完全可以跟着例子一步一步地学会每种方法。当然,在学习了某种方法并要将它应用到实际中去时,应该有足够的样本数目以及可靠的观察精度,才能产生理想的分析结果。没有一种统计方法能好到可以弥补试验观察原始资料的错误和缺陷。这一点,务必请读者注意。

作者还为本书的教学编制了一套多媒体教学课件,欢迎使用本教材进行教学的老师来函来电索取。作者的 E-mail 地址为 Oujingxiang@163.com, 电话号码为 020-85280761。

衷心感谢华南热带农业大学的林德光教授、美国堪萨斯州立大学统计系的彭先恐教授和 Michael Deaton 教授在作者学习统计学时所提供的悉心指导;衷心感谢华南农业大学卢永根院士、张桂权教授、刘向东教授,黎祖强、陈伟栋、刘桂富、黄成达等老师为本书的编写提供的热情支持和帮助。

殷切期望同行专家能给本书提出宝贵的意见,以便日后在重印或再版时改正错误、提高质量。也热切盼望学习本书的学生能学以致用,推陈出新,为祖国经济建设和中华民族的腾飞作出应有的贡献。

作 者

2003 年 1 月于广州

目 录

第一章 绪 论	(1)
第一节 科学试验的意义	(1)
一、农业和生物学领域的科学研究	(1)
二、科学研究的基本过程和方法	(1)
第二节 科学试验的基本概念和基本要求	(3)
一、科学试验中的因素及其水平的概念	(3)
二、农业科学试验的基本要求	(4)
第三节 试验误差及其控制	(4)
一、系统误差与随机误差	(4)
二、试验误差的来源	(5)
三、试验误差的控制	(6)
第四节 科学试验的一些基本类型和特点	(7)
一、田间试验和非田间试验	(7)
二、单因素试验和多因素试验	(8)
三、初级试验、高级试验和生产性试验	(12)
第五节 制订试验方案的要点	(12)
练习题	(13)
第二章 常用的试验设计	(15)
第一节 试验设计的基本原则	(15)
一、重复	(15)
二、局部控制	(16)
三、随机排列	(16)
四、三项基本原则之间的关系	(17)
第二节 常用的试验设计	(17)
一、成对法设计	(17)
二、完全随机设计	(18)
三、随机区组设计	(19)
四、拉丁方设计	(19)
五、裂区设计	(22)
六、条区设计	(24)
练习题	(24)

第三章 试验数据的整理	(26)
第一节 总体与样本的概念	(26)
一、总体	(26)
二、样本	(27)
第二节 次数分布	(27)
一、试验数据的性质及其分类	(27)
二、次数分布表	(28)
三、次数分布图	(31)
第三节 衡量资料中心位置的特征数——平均数	(32)
一、平均数的意义和种类	(32)
二、算术平均数	(34)
第四节 衡量资料分散程度的特征数——变异数	(35)
一、极差	(35)
二、方差	(36)
三、标准差	(37)
四、变异系数	(38)
第五节 用计算机和科学型计算器计算各种特征数	(38)
一、用 Excel 处理本章介绍的内容	(38)
二、科学型计算器的使用	(39)
练习题	(42)
第四章 概率论与数理统计学基础知识	(44)
第一节 概率论的一些基本概念	(44)
一、事件与事件之间的关系	(44)
二、随机事件发生的可能性——概率	(45)
三、事件概率的一些计算法则	(46)
第二节 随机变量的概率及概率分布	(48)
一、用随机变量来表示试验的结果	(48)
二、离散型随机变量数据资料的概率计算方法及概率分布	(48)
三、连续型随机变量数据资料的概率计算方法及概率分布	(49)
第三节 二项总体的理论分布	(50)
第四节 正态总体的理论分布	(51)
一、正态分布的密度函数和概率分布函数	(51)
二、正态分布曲线及其特性	(51)
三、利用正态分布计算概率	(52)
第五节 其他几种常用的理论分布	(55)
一、学生氏 t 分布	(55)
二、卡平方 (χ^2) 分布	(56)

三、费雪氏 F 分布	(57)
第六节 抽样总体的分布	(57)
一、从二项总体中抽样	(58)
二、从正态总体中抽样	(63)
练习题	(68)
第五章 参数区间估计	(69)
第一节 总体平均数的置信区间	(69)
一、单个总体平均数的置信区间	(69)
二、成对法试验资料中, 两处理观察值差数总体平均数的置信区间	(71)
三、两个独立样本所来自的总体平均数之差的置信区间	(73)
第二节 二项总体百分数的置信区间	(76)
一、单个二项总体百分数的置信区间	(76)
二、两个二项总体百分数之差的置信区间	(77)
第三节 正态总体方差的置信区间	(79)
一、单个总体方差的置信区间	(79)
二、两总体方差之比的置信区间	(79)
练习题	(80)
第六章 统计假设测验	(83)
第一节 统计假设测验的基本原理	(83)
一、统计假设的意义和步骤	(83)
二、两尾测验和一尾测验	(84)
三、假设测验的两种错误	(85)
第二节 关于总体方差的假设测验	(86)
一、关于一个总体方差的假设测验	(86)
二、关于两个总体方差的假设测验	(87)
三、关于多个总体方差的假设测验	(87)
第三节 关于总体平均数的假设测验	(88)
一、单个样本平均数与已知总体平均数之间的差异显著性测验	(89)
二、成对法试验资料中, 两处理观察值差数平均数的假设测验	(90)
三、两独立样本所来自的两总体平均数之间的差异显著性测验	(91)
第四节 关于二项资料百分数的假设测验	(93)
一、单个二项总体百分数的假设测验	(93)
二、两个二项总体百分数之差的假设测验	(94)
第五节 计数资料理论比率的适合性测验	(96)
一、资料分为两组的情况	(96)
二、资料分为多组的情况	(99)

三、总体概率分布的适合性测验	(100)
第六节 计数资料两属性间的独立性测验	(103)
一、 2×2 联列表的独立性测验	(103)
二、 $r \times 2$ 或 $2 \times c$ 联列表的独立性测验	(105)
三、 $r \times c$ 联列表的独立性测验	(105)
第七节 参数区间估计与统计假设测验的比较	(106)
练习题	(106)
 第七章 方差分析	(109)
第一节 方差分析的基本原理	(109)
一、方差分析的数学模型和基本的计算过程	(109)
二、方差分析的基本假定	(113)
三、处理效应的固定模型和随机模型	(114)
四、期望均方	(115)
第二节 处理平均数间的多重比较	(116)
一、最小显著差数法 (Least Significant Difference Method——LSD 法或 t 测验法)	(116)
二、最小显著极差法之一 (Least Significant Range——复极差法或 q 测验法)	(117)
三、最小显著极差法之二 (Least Significant Range——新复极差法或 Duncan's 法)	(118)
四、选择多重比较方法的原则	(119)
五、其他多重比较结果的表示方法	(119)
第三节 方差分量的估计	(120)
第四节 单向分类资料的方差分析	(121)
一、组内观察值数目相等的单向分类资料	(121)
二、组内观察值数目不等的单向分类资料	(123)
第五节 两向分类资料的方差分析	(125)
一、每处理组合内只有一个观察值的两向分类资料	(125)
二、每处理组合内有多于一个观察值的两向分类资料	(128)
第六节 系统分组资料的方差分析	(134)
第七节 方差分析的基本假定和数据变换	(139)
练习题	(143)
 第八章 常用试验设计的资料分析	(145)
第一节 完全随机设计试验的资料分析	(145)
一、单因素完全随机试验资料的统计分析	(145)
二、多因素完全随机试验资料的统计分析	(145)

第二节 随机区组设计试验的资料分析	(150)
一、单因素随机区组试验资料的统计分析	(151)
二、多因素随机区组试验资料的统计分析	(151)
第三节 拉丁方设计试验的资料分析	(156)
一、单因素拉丁方试验资料的统计分析	(156)
二、多因素拉丁方试验资料的统计分析	(158)
第四节 裂区设计试验的资料分析	(161)
一、主副处理都按随机区组安排的裂区设计试验资料的分析	(161)
二、主处理按拉丁方, 副处理按随机区组安排的裂区设计试验资料的分析	(166)
第五节 条区试验设计的资料分析	(167)
练习题	(171)
第九章 直线相关与回归	(172)
第一节 两随机变量之间的线性关系	(172)
一、两变量间的离均差乘积和和协方差	(172)
二、相关系数和决定系数	(174)
三、相关系数的假设测验	(175)
第二节 直线回归方程	(178)
一、直线回归方程的拟合	(178)
二、直线回归的显著性测验	(180)
三、直线回归的区间估计	(182)
四、两条回归直线的比较	(183)
练习题	(186)
第十章 协方差分析	(188)
第一节 协方差分析的意义和作用	(188)
第二节 单向分类资料的协方差分析	(188)
一、相关模型的协方差分析	(188)
二、回归模型的协方差分析	(191)
第三节 两向分类资料的协方差分析	(193)
一、相关模型的协方差分析	(193)
二、回归模型的协方差分析	(194)
练习题	(196)
附 表	(198)
附表 1 标准正态分布累计概率表	(199)
附表 2 正态离差 u_{α} 值表 (两尾)	(201)
附表 3 九种常用显著水准 α 下的 t_{α} 值表 (两尾)	(202)

附表 4	二项分布参数 p 的 95% 置信区间的上限和下限	(203)
附表 5	二项分布参数 p 的 99% 置信区间的上限和下限	(205)
附表 6	十三种常用显著水准 α 下的 χ^2 值表 (单尾)	(207)
附表 7	四种常用显著水准 α 下的 F 值表 (单尾)	(208)
附表 8	两种常用显著水准 α 下的 <i>Duncan</i> 氏 <i>SSR</i> 值表 (两尾)	(220)
附表 9	两种常用显著水准 α 下的 <i>SNK</i> 全距 q 值表 (两尾)	(221)
附表 10	二项百分数资料反正弦 ($\arcsin \sqrt{p}$) 变换表	(222)
附表 11	相关系数 r 与正态变量 z 的转换	(224)
附 录		(225)
附录 1	希腊字母表	(225)
附录 2	常用的统计符号	(225)
索 引		(228)
参考文献		(234)

第一章 绪 论

本章主要阐明田间试验的基本概念。着重讲述田间试验的任务、类型和有关试验误差的来源和控制等方面的内容。

第一节 科学试验的意义

一、农业和生物学领域的科学研究

邓小平同志指出：“科学技术是第一生产力。”按照传统的分类方法，科学体系分为自然科学和社会科学两大类*。而自然科学中又有理论科学和实验科学两大类。理论科学研究主要运用推理，包括演绎法和归纳法两种。演绎法是指从一般原理推断特殊事实，而归纳法是指由许多的事例概括出一般的原理。实验科学研究则主要通过周密设计的实验来探新。农业和生物学领域的科学研究包括农学、园艺、动物科学、生物技术、植物保护、资源与环境保护与利用等有关的专业所涉及的学科大都是实验科学。这些学科中使用的中科学研究的方法主要有两种：抽样调查和科学试验。前者使我们能够具有代表性地、准确地描述变化万端、千差万别的生物界中的各种自然现象，后者则通过严格控制试验条件，使所比较的对象间尽可能少地受非研究因素的干扰，并把所要研究的差异突出地显示出来。

二、科学研究的基本过程和方法

科学研究的目的在于探求新的知识、理论、方法、技术和产品。科学研究大致分为：基础性研究、应用基础性研究和应用性研究三类。前两类用于发现新的知识、理论和方法，而后者则应用于获得某种新的技术和产品。

（一）科学研究的基本过程

上述三种科学研究的基本过程都是一样的，即具有以下的三个基本环节：①根据前人的研究成果或研究者自己的观察，对所研究的命题形成一种认识或假说；②根据假说所涉及的内容安排相斥性的试验或抽样调查；③根据试验或调查所获得的资料进行推理，肯定或否定或修改假说，并形成结论，或开始新一轮的试验以验证修改完善后的假说。如此循环发展，使所获得的知识或理论逐步发展和深化。

（二）科学研究的基本方法

不管什么类型的科学试验都有些通用的基本方法。以下介绍一些科学试验工作者必须掌握的最基本的方法（参考文献[2]）。

（1）选题的方法：科学试验的基本目的是探新、创新。研究课题的选择决定了该项研究创新的潜在可能性。优秀的科学研究人员不仅要具有解决问题的能力，更重要的是应善于选取那些具有取得优秀成果可能性的课题。甚至有人认为，选到一个好的题目，相当于成功了一半。

* 到了21世纪，钱学森等人提出将科学技术体系分为自然科学、社会科学、数学科学、系统科学、文艺理论、思维科学、军事科学、行为科学、人体科学和地理科学等十个类别。（参考文献[1]）

最有效的研究是去开拓前人还未涉及过的领域。不论是理论性的研究还是应用性的研究,选题时都应明确它的意义或重要性。理论性研究应着重看所选课题在未来学科发展上的重要性,而应用性研究则应着重看它在未来生产发展的作用和潜力。

区别于其他类型的工作,科学研究需要进行独创性的思维。因此要求所选的课题能引起研究者的浓厚兴趣,促使研究者保持十分敏感的心理状态。如果所选的课题不能激发研究者的兴趣,这样的研究将不太可能获得新颖的见解和预期的成果。对于那些由资助者设定的课题,研究者必须认真体会它的确实意义并激发出对该项研究的热情和信心。

(2) 有关文献的收集和整理:科学发展是一个长期累积的过程。每项研究都是在前人建筑的大厦顶层上添砖加瓦。这就要求首先要登上顶层,然后才能增建新的层次。文献便是把研究工作者推到顶层,掌握大厦总体结构的通道。

选题要有文献作为依据,设计研究内容、路线和方法更需文献的启示。查阅文献可以减少走弯路的可能性。不要害怕查阅文献会花费时间和功夫。它能避免许多无效的重复工作,避免我们走弯路,因而所节省的时间会远远补偿查阅文献所花费的时间。

科学文献随着时代的发展越来越丰富。百科全书是最普通的资料来源。对于要进入一个新领域的人来说,它是取得最初了解的极为有用的工具。文献索引是帮助科学研究人员进入某一特定领域作广泛了解的重要工具。专业书籍可为科学研究人员进入这一领域提供一个基础性的了解。评论性杂志可使科学研究人员了解有关领域里已取得的主要成绩。文摘可帮助研究人员查找特定领域研究的结论性内容,使之跟上现代科学前进的步伐。科学期刊和杂志登载最新研究的论文,它介绍一项研究的目的、材料、方法以及由试验资料推论到结果的全过程,优秀的科学论文,可给人们以研究思路和方法上的启迪。各个有实力的研究机构都十分重视图书、期刊、文献的搜集,图书馆是研究人员工作的一个关键场所。现代通信和网络技术的发展,使一些书籍、杂志通过网络为研究人员提供服务,因此计算机及网络系统也是文献检索的主要工具。

(3) 假设的提出:在提出一项课题时,研究者事先一定对所研究的对象有些初步了解。有些可能来自以往观察的累积,有些可能来自文献分析。因而围绕研究对象和预期结果之间的关系,研究者常已有某种见解或想法,即已构成了某种假说。必须通过进一步的研究来证实或修改已有的假说。一项研究的目的和预期结果总是和假说相关联的,没有形成假说的研究,常常是含糊的、目的性不甚明确的。

即使最简单的研究,实际上也总会有某种假说。例如,引进若干外地品种,将它们与当地常用的品种进行比较试验,研究者心目中可能事先就会有这样的想法:“某些(个)引进品种可能优于当地常用的品种。”这就是一个假说,只不过它是比较简单的假说而已。

简单的假说只是某些现象的概括。复杂的假说则要进一步假定出各现象之间的联系,这种联系可能是平行的,也可能是因果的;复杂的假说中甚至还可能包含类推关系。假说只是一种尝试性的设想,即对于所研究对象的试探性概括,在它没有被证实之前,绝不能与真理、定律混为一谈。

科学的基本方法之一是归纳,从大量现象中归纳出真谛;演绎是科学的另一基本方法,当构思出一个符合客观事实的假说时,可据此推演出更广泛的结论。其中,形式逻辑是必要的演绎工具。自然科学研究人员应自觉地训练并用好归纳、演绎以及形式逻辑的方法。

(4) 假说的检验:假说有时也表示为假设。在许多研究中假设是简单的,它们的推论也很明确。对假说可以直接进行检验,也可以间接进行检验。所谓直接检验,是指重新对研

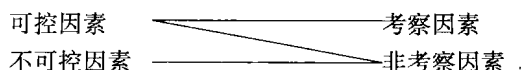
究对象进行观察、实验或试验。所谓间接检验是指安排一些试验来对假说的推理进行验证。验证了所有可能的推理的正确性，也就相当于验证假说本身。当然用这种间接方法进行检验时，要十分小心，防止漏洞。

(5) 试验的规划与设计：围绕检验假说而开展的试验，需要全面、仔细的规划与设计。试验所涉及的范围应能覆盖假说涉及的各个方面，以便能对待检验的假说作出无遗漏的判断。下一节中将介绍一些与科学试验有关的基本概念、对科学试验的基本要求以及制订试验方案的要点。

第二节 科学试验的基本概念和基本要求

一、科学试验中的因素及其水平的概念

科学试验 (scientific experiment) 是指人类为检验科学理论或假设而从事的专门实践活动 (参考文献 [3])。用于衡量试验效果 (effect) 的量称为试验指标 (experiment index)。对一个试验对象进行观察，得到的试验指标值称为观察值 (observed value)。在试验中对试验指标有影响的条件称为因素 (factor)，因素有时被翻译为因子。按能否通过人为手段进行控制，可以将因素分为可控因素 (controllable factor) 和不可控因素 (controllable or uncontrollable factors)。按是否试验考察的对象，可以将因素分为考察因素 (inspective factors) 和非考察因素 (inspective or non-inspective factors)。考察因素必须是可控因素，但可控因素不一定是考察因素；不可控因素只能作为非考察因素，因为不可控就无法进行考察。这四种因素的关系如下：



例如：在水稻品种比较试验中，若试验指标为产量。则对产量有影响的条件有无数方面，即有无数个因素。

品种、密度、肥料、管理等可能是可控因素；

天气、虫害、兽害等可能是不可控因素。

在这个例子中，只有品种是考察因素，其余都是非考察因素。

为了对因素进行考察，常把考察因素按数量分为不同的级别或者按质量分为不同的状态。这些级别或状态称为因素的不同水平 (levels of factor)。不同的考察因素可以有不同的水平。例如上述例子中，考察因素是水稻品种，如果试验中包括 5 个品种，那么品种因素就有 5 个质量水平。而非考察因素通常应该是处在同一水平上的。即除了考察因子要分为不同水平，以供考察之外，其余的因素都应保持一致、不分水平，这在试验设计上称为“唯一差异原则” (principle of unique variability)。

比较是科学研究中常用的方法，有比较才有鉴别。农业和生物学领域的研究中常常采用比较试验的方法，从比较中确定出最确凿的理论、方法和技术。比较研究中十分重要的是选定恰当的比较标准。

比较试验中比较的对象可以不止两个，而是一组对象间的比较。这组比较的对象是按假说的内涵选定的，称为一组处理 (treatment)。这一组处理可能是某一因素 (因子) 的水平，也可能是不同因素 (因子) 的不同水平的组合。作为比较标准的处理称为对照处理 (check)。

二、农业科学试验的基本要求

为了使农业科学试验能达到预期的目标,使试验结果能在农业科学理论或农业生产实践的发展中发挥作用,应对农业科学试验提出一些基本的要求。

(一) 试验的目的应明确

在大量阅读文献与社会调查的基础上,明确选题,制订合理的试验方案。对试验的预期结果及其在农业科学理论和农业生产实践中的作用要做到心中有数。试验项目首先应抓住当前生产实践和科学试验中亟需解决的问题;同时又能照顾到长远的和在不久的将来可能突出的问题。

(二) 试验的条件要有代表性

农业和生物学的试验中十分重视试验结果的代表性,从而可以明确研究结果的使用范围。试验的条件应能代表将来准备推广试验结果的地区的自然条件(如土壤类型、地势、土壤肥力、气象条件等)和农业条件(如农业结构、施肥水平、耕作制度等)。这样,新品种或新技术在试验中的表现才能真正反映今后推广地区实际生产中的表现。在进行试验时,既要考虑代表目前的条件,还应注意将来可能被广泛使用的条件。使试验结果既能符合当前的需要,又不落后于生产发展的要求。

(三) 试验的结果要可靠

试验结果的可靠性可以用准确性和精确性来衡量。试验结果的准确性(accuracy)是指观察值与其理论真值之间的符合程度,它衡量了试验观察结果与真实情况偏离的程度,这种偏离常常是由一些非考察的可控因素引起的,属于系统误差;试验结果的精确性(precision)是指重复观察值之间的符合程度,对同一对象进行多次观察,重复观察值之间的差异常常是由不可控因素引起的,属于偶然误差。第三节将详细讨论试验误差及其控制的方法。

(四) 试验的结果应能够重演

农业和生物学的试验同样十分重视试验结果的重演性。这是指在同样的条件下,再次进行试验或实验,应能获得与原试验相同的结果。这对于在生产实际中推广农业科学研究的成果极为重要。田间试验中不仅农作物本身有变异性,环境条件更是复杂多变。要保证试验结果能够重演,首先,要仔细明确地设定试验条件,包括田间管理措施等。试验实施过程中对试验条件(包括气象、土壤以及田间管理措施)和作物生育过程保持系统的纪录,以便创造相同的试验条件,重复验证,并在将试验结果应用于相应条件的农业生产时有相同的效应。其次,为了保证试验结果能重演,可将试验在多种试验条件下进行,以得到相应于各种可能条件的结果,例如品种区域试验,为了更全面地评价品种,常进行2~3年多个地点的试验,以明确品种的适用范围,将品种在适宜的地区和条件下推广应用,取得预期效果。

第三节 试验误差及其控制

为了保证试验结果的可靠性,必须尽量减少试验误差,为此有必要专门对试验误差及其控制进行更为深入的讨论。

一、系统误差与随机误差

试验误差是指试验指标的观察值与其真值之差。例如有一个物体,如果它的真正重量是

10 g, 但是试验中我们测得的观察值为 10.3 g。于是有了 $10.3\text{ g} - 10\text{ g} = 0.3\text{ g}$ 的差异。这 0.3 g 就是试验误差。为了提高试验的可靠性, 我们常常不只测量一次, 而是测量多次。每次测量都产生一个误差值, 这些误差值不尽相同, 而且大多数情况下是不相同的。误差有时为正, 有时为负。

事实上, 试验指标的真值是未知(或不可知)的。我们只能用试验的多次重复观察值的平均数来对此真值进行估计。这种估计就带有一定的误差。试验误差越小, 估计就越可靠。

这种误差是怎样产生的呢? 究其原因, 有两个方面: ①由一些非考察的可控因素引起的。例如我们用了一台不准确的天平来称量那个物体, 所以称起来总是偏高。只要换上一台准确的天平, 这类误差就可以减少甚至避免。②由不可控因素引起的。这类误差只可以减少而不可能完全避免。前一种是由一定的已知原因引起的误差, 称为系统误差(systematic error), 或称为偏差(bias); 后一种是完全偶然性的, 找不出确切原因的误差, 称为偶然误差(spontaneous error)或随机误差(random error)。

系统误差影响试验的准确性, 随机误差影响试验的精确性。拿打靶作比喻, 靶心是试验指标的理论真值。如果一枪命中了靶心, 这就是准确性高。如果这支枪有毛病, 用它打总是偏高, 这就是系统误差造成的偏差, 这时的准确性就低。如果同时打几枪, 结果都非常集中在某一点的附近, 这就说精确性高。如果连续打几枪, 命中点非常分散, 这就说精确性低。图 1.1 形象地展示了准确性与精确性的关系。其中图 1.1a 表示准确性和精确性都高; 图 1.1b 表示精确性高但准确性不高; 图 1.1c 表示准确性高但精确性不高; 图 1.1d 则表示准确性和精确性都不高。

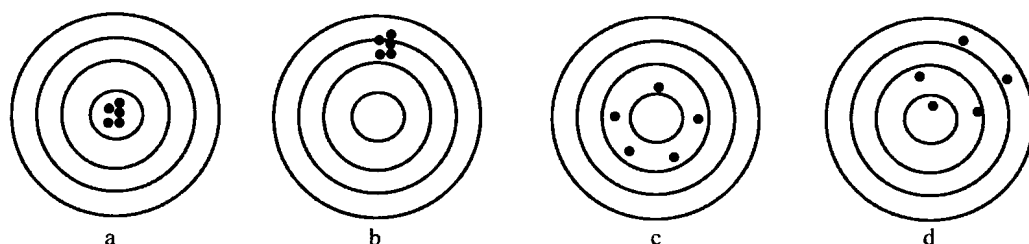


图 1.1 用打靶作比喻展示准确性与精确性的关系

在农业和生物学试验中, 常常采用比较试验来衡量试验的效应。如果两个处理均受到同一方向、同一大小的系统误差的干扰, 在比较这两个处理的效应时, 这种系统误差的影响不是很大。相反, 如果两个处理受到不同方向、不同大小的系统误差的干扰, 在比较这两个处理的效应时, 这种系统误差的影响就会非常严重。例如我们要比较两个物体的重量, 并且称量两个物体时, 都用同一台老是偏重的天平, 那么, 两物体重量之差, 可能不会非常离谱。但是, 如果称量 A 物体时用一台偏重的天平, 而称量 B 物体时用一台偏轻的天平, 那么, 比较这两个物体的重量时, 试验结果就会非常不可靠。但是在实际试验中, 只要尽量按前面说的“唯一差异原则”安排试验, 很少会出现后面这种情况。因此, 研究工作者在正确设计并实施试验计划的基础上, 十分重视精确性或对偶然误差的控制, 因为这直接影响到日后对试验资料的统计分析的正确性。

二、试验误差的来源

从原理上说, 试验误差有两个方面的来源: 由非考察的可控因素引起的系统误差和由不

可控因素引起的随机误差。但在实践中,由于有些非考察的可控因素难以得到很好的控制,它们的影响实际无法从总的误差中扣除出来。一般在试验分析中,常把误差中的一些主要偏差(系统误差)排除之后,把剩余的误差都归为随机误差,因而估计出来的随机误差有可能比想象的要大,甚至大得多。

在操作上,在农业和生物学领域的科学试验中,尤其是在田间试验中,试验误差的来源主要有以下三个方面:

(1) 试验材料固有的差异:在田间试验中供试的材料常常是植物或其他生物。它们在遗传性质和生长发育上往往会存在着差异。如果供试材料的基因型不一致,种子的生活力有差别,试验用的幼苗素质有差别等,均会造成试验结果的偏差。

(2) 试验操作和管理技术不一致所引起的差异:供试材料在田间的生长周期较长,在试验过程中的各个管理环节稍有不慎,都很容易增加试验误差。例如在播种前整地、施肥的不一致,播种的深浅不一致,中耕、除草、灌溉、追肥、防虫、治病、使用除草剂、使用生长剂等完成的时间和操作标准不一致,收获脱粒的操作质量不一致,观察测定的时间、人员、仪器不一致等,都有可能成为引起误差的原因。

(3) 试验中外部环境条件的差异:试验的外部条件错综复杂,千变万化。对于田间试验,最主要、最常见的外部条件是试验地的土壤差异。土壤的类型、肥力的差异是引起田间试验误差的重要原因。试验工作者要花大力气来了解和控制试验地的土壤差异,以便有效地减少试验误差。此外,病虫害侵袭、人畜践踏、风雨破坏等因素都带有随机性,使各种处理受到不同程度的影响,从而增加试验的误差。

这三方面的来源都包括了诸多因素,有些因素带来系统误差,有些因素带来随机误差。至于某一种因素带来的是系统误差还是随机误差,要视具体情况具体分析。

除了试验误差之外,试验中的失误操作,可能导致错误。试验中的错误是绝不允许的。一旦发生错误,必须马上纠正,无法纠正的,只能重做试验。绝不允许用错误的数据或虚假的试验结果欺骗和误导公众,这是科学工作者最基本的职业道德。

三、试验误差的控制

根据上述对试验误差来源的分析,为了控制试验误差,必须严格坚持“唯一差异原则”,逐一克服试验材料、试验操作和管理技术、试验条件等各种因素的不一致性。如果某一个非考察因素在整个试验环境不能完全达到一致,而是分为若干水平,那么,可以将整个试验范围按照此因素的水平分为几个局部区域,然后让每一个处理在各个局部区域中占同样分量的试验单元。这样的局部区域称为区组(block)。在田间试验中,常将一个区组按照试验处理的数目分为若干小区(plot),每个小区随机安排一个处理,每个小区通常收集一个观察值。在非田间试验中,收集一个试验指标观察值的试验单元称为“cell”。因为“cell”不太好翻译,因此也常用“小区”来称呼它。

针对上述误差来源,以田间试验为例,可采取以下措施:

(1) 选用质量相同的试验材料:必须严格要求试验材料的基因型同质一致;至于生长发育上的一致性,如苗木大小、壮弱不一致时,则可按大小、壮弱分档,然后将同一规格的苗木安排在同一区组的各处理小区中,或将各档次苗木按比例混合分配给不同的处理,从而减少试验的误差。

(2) 改进操作和管理技术,使之标准化:总的原则是,除了操作要仔细,一丝不苟,