



面向 21 世纪 课程 教材

Textbook Series for 21st Century

食品试验设计与统计分析

王钦德 杨 坚 主编



中国农业大学出版社

面向21世纪课程教材
Textbook Series for 21st Century

食品试验设计 与统计分析

王钦德 杨 坚 主编

中国农业大学出版社

图书在版编目(CIP)数据

食品试验设计与统计分析/王钦德,杨坚主编. —北京:中国农业大学出版社,
2003.2
ISBN 7-81066-438-7/TS·3

I. 食… II. ①王… ②杨… III. ①食品检验 ②食品分析:统计分析
IV. TS207.3

中国版本图书馆CIP数据核字(2002)第110489号

出版 中国农业大学出版社
发行 新华书店
经销 新华书店
印刷 莱芜市圣龙印务书刊有限责任公司
版次 2003年2月第1版
印次 2003年2月第1次印刷
开本 16 印张 33.5 千字 613
规格 787×980
印数 1~3 500
定价 37.00元

图书如有质量问题本社负责调换

社址 北京市海淀区圆明园西路2号 邮政编码 100094
电话 010-62892633 网址 www.cau.edu.cn

全国高等农业院校食品
专业“面向 21 世纪课程”系列教材
编审指导委员会委员

- 罗云波 中国农业大学教授博士生导师 (生物技术)
孙远明 华南农业大学教授博士生导师 (食品营养)
陈宗道 西南农业大学教授博士生导师 (食品化学)
李里特 中国农业大学教授博士生导师 (食品工程)
李新华 沈阳农业大学教授博士生导师 (粮油加工)
李士靖 中国食品科学技术学会副秘书长教授
李云飞 上海交通大学教授博士生导师 (食品工程)
何国庆 浙江大学教授博士生导师 (食品微生物)
杨公明 西北农林科技大学教授博士生导师 (食品工程)
周光宏 南京农业大学教授博士生导师 (畜产品加工)
林家栋 中国农业大学教授全国高等学校教学研究中心特聘专家
南庆贤 中国农业大学教授博士生导师 (畜产品加工)
谢笔钧 华中农业大学教授博士生导师 (食品化学)

出版说明并代序

我国农业结构的调整，解决农村、农业、农民的发展出路，已将农产品的贮藏加工及食品科学推到了举足轻重的位置，成为拉动农业产业化、提高农产品附加值以及实现国家现代化的牵引力。而大专院校食品科学各专业的教学工作为这种牵引力提供了人才保障。

全国高等农业院校的食品学科大多建立于 20 世纪 80 年代改革开放的初期，经过近 20 年的发展，现已成为我国食品科学人才培养的最为重要的人才基地。农业院校的食品学科之所以能快速发展，后来居上，成为我国食品科学的主要力量，其主要原因是：食品科学与生物学科广泛地联系在一起。农业院校的食品学科得益于它植根于生物科学学科群之中，借助于生物科学飞速发展的翅膀而不断地深化自己的研究内容，提高自己的学科水平。

在学科发展的起步阶段，教学工作一直沿用过去轻工院校所编写的食品工程专业教材。然而，经过 20 年的发展，这些教材已经远远不能适应今天的教学需要。虽然各院校针对这种情况也曾先后编写过一些教材，但终因不成体系，很难系统地将食品学科内容广泛的课程体系和教学内容很好地衔接起来。要培养面向 21 世纪的高素质食品科学人才，迫切地需要将现代生物学理论与食品科学紧密地结合在一起，编写一套理论性和实践性俱强的完整教材。

这套教材正是在这样的背景和需要的前提下，在教育部、农业部有关领导部门的指导下，通过全国 40 多所院校在第一线的教师的共同努力下，由中国农业大学出版社组织编写而成的。教材力求反映最新的食品科学的理论与实践，同时针对食品科学是多学科集成的优点，特别注重了教材的系统性，避免课程教学内容的重复；针对食品科学实践性强的特点，教材中使用了较多的案例分析。在写作方式上，力求教材能启发学生的主动思考能力，培养学生的创新思维能力。

这套教材还得到了食品学界一批有声望的老专家、老教授的关怀和指导。由于时间紧、任务重，加之该教材体系初次建立，使用效果怎样，还要在实践中去检验。随着学科的不断发展，其内容也需要不断地修改补充，编者真诚地期待着使用这套教材的教师和同学们能够提出宝贵意见，以使这套教材充实和得以完善。

罗云波

2002 年 7 月

于马连洼

前 言

《食品试验设计与统计分析》教材是根据“加强基础、强化专业、拓宽知识面和重视应用”的教改精神组织编写的,是高等教育面向21世纪教学内容和课程体系改革项目(04-18)成果。本教材编委会由山西农业大学王钦德、西南农业大学杨坚、福建农林大学庞杰、湖南农业大学谭敬军、四川农业大学单虹丽、山东农业大学乔旭光、西南农业大学童华荣和山西农业大学王如福、张吴平共9人组成,于2001年4月在中国农业大学召开了编写会议。在编写会上全体编委认真讨论、审定了编写大纲,确定了章节安排、内容取舍、深度、广度和详略,并进行了编写分工。初稿完成后,由主编王钦德、杨坚负责统稿,对书稿进行了必要的修改和增删。为了确保教材质量,特聘请四川农业大学明道绪教授为主审,对全部书稿进行了详尽审阅。这本教材得以问世,凝结了全体编、审人员的心血。

本教材包括绪论(杨坚、王钦德编写),数据资料的整理与特征数(王如福编写),理论分布与抽样分布(张吴平编写),统计假设检验(单虹丽编写),方差分析(王钦德编写),直线回归与相关(乔旭光编写),多元线性回归与相关(王钦德编写),非参数统计(童华荣编写),试验设计基础与抽样方法(庞杰、王钦德编写),随机区组设计(单虹丽编写),正交试验设计(谭敬军、杨坚编写),均匀设计(谭敬军编写),回归正交设计(杨坚编写),混料回归试验设计(童华荣编写)共14章(其中包括自选内容,已在有关章节的标题上用“*”注明),并附有统计处理软件(SAS)简介(张吴平编写)、英汉术语对照(庞杰整理)及常用统计数学用表。

本教材的编写,力求做到内容的科学性与实用性、先进性与针对性相统一;做到循序渐进,由浅入深,深入浅出,简明易懂;在正确阐述

重要的统计学原理的同时,着重于基本概念、基本方法的介绍,特别注意学生动手能力的培养;每一种设计或分析方法都安排有步骤完整、过程详细的实例子以说明;各章都有明确的教学目标且配备有习题(附简要答案)供读者练习;考虑到教学时数的限制和实际需要,教材中部分内容可作为选用内容(用“*”号标出)。

本教材在保持本学科的系统性和科学性的前提下,注意引入本学科发展的新知识、新成果;注重拓宽学生的知识面和提高实践能力,紧密联系食品科学生产、科研实际,以及统计分析与计算机科学的结合;避免与交叉学科有关知识的重复。力求体现“厚基础、强能力、高素质、广适应”和素质教育与创新教育的教学目标。

本教材除可作为高等农业院校食品科学类专业教学用书外,也可作为轻工、商学、水产、粮食等院校的食品科学、食品工程、发酵工程、生物工程等专业开设《生物统计》课程的教学用书,还可作为食品科学类专业成人教育教材。此外,对食品科技工作者亦有重要参考价值。

本教材在编写过程中参考了有关中外文献和专著,编者对这些文献和专著的作者,对大力支持编写和出版工作的中国农业大学出版社一并表示衷心感谢!

限于编者水平,错误、缺点在所难免,敬请统计学专家和广大读者批评指正,以便修订改正。

编者

2002.12.2

目 录

第1章 绪论	(1)
1 试验设计与统计分析概述	(2)
2 食品科学试验的特点与要求	(9)
第2章 数据资料的整理与特征数	(12)
1 常用术语	(13)
2 数字资料的性质	(15)
3 资料的整理	(17)
4 资料的特征数	(25)
5 异常数据的处理	(33)
习题	(38)
第3章 理论分布与抽样分布	(40)
1 理论分布	(41)
2 抽样分布	(55)
习题	(63)
第4章 统计假设检验	(65)
1 统计假设检验概述	(66)
2 样本平均数的假设检验	(75)
3 二项百分率的假设检验	(85)
4 统计假设检验中应注意的问题	(92)
5 参数的区间估计	(94)
习题	(100)
第5章 方差分析	(102)
1 方差分析的基本原理	(104)
2 多重比较	(115)
3 单向分组资料的方差分析	(122)
4 两向分组资料的方差分析	(127)
5 方差分析的基本假定和数据转换	(143)
习题	(148)

第6章 直线回归与相关	(152)
1 回归与相关的概念	(153)
2 直线回归	(154)
3 直线相关	(165)
4 应用直线回归与相关的注意事项	(169)
5 能直线化的曲线回归*	(172)
习题	(178)
第7章 多元线性回归和相关	(180)
1 多元线性回归	(181)
2 多元相关与偏相关*	(197)
3 通径分析*	(204)
习题	(219)
第8章 非参数统计	(222)
1 非参数统计的概念和特点	(223)
2 χ^2 检验	(224)
3 符号检验*	(232)
4 符号秩和检验*	(234)
5 秩和检验*	(236)
6 秩相关*	(245)
习题	(247)
第9章 试验设计基础与抽样方法	(249)
1 试验设计概述	(250)
2 基本概念	(251)
3 食品试验研究的主要内容	(254)
4 食品试验的基本要求和注意事项	(258)
5 试验设计的基本原则	(261)
6 试验计划与方案	(263)
7 试验误差及控制	(270)
8 完全随机设计	(274)
9 常用抽样方法及其抽样误差*	(277)
10 样本含量的确定	(292)
习题	(300)
第10章 随机区组试验设计及统计分析	(303)

1 随机区组试验设计方法	(304)
2 随机区组试验结果的统计分析	(308)
3 随机区组缺值估计及其统计分析	(322)
习题	(327)
第 11 章 正交试验设计	(330)
1 正交试验设计的意义	(331)
2 正交表	(336)
3 正交试验设计的基本步骤	(344)
4 正交试验的结果分析	(349)
5 正交设计的灵活运用*	(361)
习题	(366)
第 12 章 均匀试验设计	(368)
1 均匀设计的概念及特点	(369)
2 均匀设计表	(370)
3 均匀试验设计的基本方法	(375)
4 均匀试验设计实例	(377)
习题	(383)
第 13 章 回归的正交设计	(384)
1 回归正交试验设计简介	(385)
2 一次回归正交设计及统计分析	(386)
3 二次回归正交组合设计及其统计分析	(398)
4 回归的旋转设计*	(416)
习题	(434)
第 14 章 混料回归试验设计	(436)
1 混料设计的概念	(437)
2 单纯形格子设计	(438)
3 单纯形重心设计	(443)
习题	(445)
附录 统计处理软件(SAS)简介	(446)
附表	(459)
英汉术语对照	(516)
参考文献	(520)

绪 论

本章学习目的与要求

1. 了解试验设计与统计分析在食品科学研究中的应用及发展概况
2. 熟悉食品科学试验的特点与要求

1 试验设计与统计分析概述

1.1 试验设计与统计分析在食品科学研究中的应用

为了推动食品科学的发展,常常要进行科学研究。例如,食品原料资源的研究,新产品开发和新的加工工艺的研究,食品质量保持、贮藏方法、货架寿命、营养价值,安全性和经济特性等的研究以及卫生标准的制定等。这些研究都离不开调查和试验。进行调查或试验首先必须解决的问题是:如何合理地进行调查或试验设计。在实际研究工作中常常遇到这样的情况,即由于调查或试验设计不合理,以致无法从所获得的数据中提取有用的信息,造成人力、物力和时间的浪费。若调查或试验设计方法好,则用较少的人力、物力和时间即可收集到必要而有代表性的资料,从中获得可靠的结论,达到调查或试验的预期目的,收到事半功倍之效。

通过调查或试验获得一定数量的数据,这些数据常常表现出不同程度的变异。例如,测定某食品原料蛋白质含量获得100个数据。彼此不完全相同,表现出一定的变异;又如,同一生产工艺条件下生产某种食品其外观及内在质量不一定相同,也表现出一定的变异。产生这种变异的原因有的已被人们所了解。食品原料的来源种类不同蛋白质含量表现出差异;又如,食品配方,加工工艺贮藏方式不同使得食品质量有所不同。另外还有许多内在和外在的因素还未被人们所认识。由于这些人们已了解的因素和人们尚未认识而无法控制的因素的作用,使得通过调查或试验得来的数据普遍具有变异性。因此,进行调查或试验还必须解决的第二个问题是:如何科学地整理、分析所收集的具有变异的资料,揭示出隐藏在其内部的规律性。食品试验设计与统计分析总体上讲属于生物统计学的范畴。合理地进行调查或试验设计、科学地整理、分析所收集得来的资料是生物统计的根本任务,也是本门课程的根本任务。

食品试验设计与统计分析是整理统计原理和方法在食品科学研究中的应用,它在食品科学研究中的作用主要体现在以下两个方面。

(1)提供试验或调查设计的方法。广义的试验设计是指试验研究的课题设计,亦即整个试验计划的拟定。狭义的试验设计主要是指试验单位(如食品贮藏试验中心的包装件)的选取、重复数目的确定、试验单位的分组和试验处理的安排。通常讲的试验设计主要指的是狭义的试验设计。合理的试验设计能控制和降低试验误差,提高试验的精确性,为统计分析获得试验处理效应和试验误差的无偏估计提供必要的的数据。食品试验研究中常用的试验设计方法有安全随机设计、随机区组设计、正交设计、均匀设计、回归正交设计和混料设计等。

广义的调查设计是指整个调查计划的制定。狭义调查设计主要是指抽样方法的选取,抽样单位、抽样数目的确定等内容。通常讲的调查设计主要是指狭义的调查设计。合理的调查设计能控制与降低抽样误差、提高调查的精确性,为获得总体参数的可靠估计提供必要的的数据。

简言之,试验或调查设计主要解决合理地收集必要且有代表性资料的问题。

(2)提供整理,分析资料的方法。整理资料的基本方法是根据资料的特性将其整理成统计表,绘制成统计图。通过统计表与统计图可以大致看到资料集中,分散的情况,并利用所收集的数据资料计算出几个统计量,以表示该资料的数量特征,估计相应的总体参数。

统计分析的内容很多,最重要的是差异显著性检验,亦即假设检验。通过抽样调查或控制试验获得的是具有变异的资料。产生变异的主要原因是什么?是由于比较的处理间(如不同原料,不同工艺、不同配比间)有实质性差异,还是由于无法控制的偶然因素所引起?显著性检验的目的就在于承认并尽量排除无法控制的偶然因素的干扰将处理间是否存在本质差异揭示出来。显著性检验的方法很多,有 t 检验,方差分析, χ^2 检验等。

统计分析的另一重要内容是对变量间的关系进行研究,即研究它们之间联系的性质和程度或寻求它们之间联系的形式,亦即进行相关分析与回归分析。通过对资料进行相关、回归分析,可以揭示出变量间的内在联系,为食品新产品的研制开发、产品质量的预测和控制提供依据。

还有一类统计分析方法叫非参数检验法。这种分析方法不考虑资料的分布类型,也不事先对有关总体参数进行估计。非参数检验法计算简便。当通常的检验方法对某些试验或调查资料无能为力时,这类方法则正好发挥作用。

食品科学是涵盖了农副产品贮藏加工、生物科学、农业工程和轻工业等学科的综合性和交叉型学科。随着20世纪80年代以来世界食品工业的飞速发展,食品科学研究朝着自动化生产、计算机应用、系统工程、生物酶技术、基因工程等高新技术发展,逐步脱离了传统的加工方法,体现了科学化、集约化生产的特色,也对

食品科学研究的试验设计和统计方法提出了更高的要求。食品的试验研究已经由简单的假设测验、方差分析发展到多元分析、优化设计等高级试验设计分析方法,愈加显出了试验设计和统计分析在食品科学研究中的重要性。可以预言,随着生物科学、计算机和高级试验统计在食品工业中的广泛而深入的应用,食品科学将进入一个更快更新的发展阶段。

1.2 试验设计与统计学发展概况

试验统计的方法最早起源于对农业及生物遗传研究的应用统计方法,故一般称为生物统计学,它是应用数理统计学原理来研究生物界数量现象的科学方法,是一门数理统计学与生物科学相结合的应用边缘学科。20世纪以来,由于生物统计学的进展,使生物科学和农业科学逐渐成为可以用数学方法来处理和研究的科学。试验统计学作为一门系统的学科开始于1925年英国统计学家R. A. Fisher的著作*Statistical Methods for Research Workers*,该书形成了试验统计学较为完整的体系。以后随着农业和生物学研究的发展,生物统计、试验设计和抽样理论得到了快速的同步发展,并随着工业研究和数理科学研究的发展而进一步推动了应用数理统计学的发展,反过来又推动了试验统计学的不断发展。

统计(Statistics)是一个古老的政治术语,其原意是用于国家管理需要的统计数字,后来发展成为将统计学看做为科学实验数据收集、统计分析和推论的理论、方法和科学。试验统计是统计学的一个部分,应用数学的一个分支。它是统计学与试验设计相结合而发展起来的。试验统计学的发展也是与随机误差和误差控制的研究紧密相关的,试验统计学的发展,大致可划分为3个阶段。

1.2.1 古典记录统计学阶段

古典记录统计学形成期间大致在17世纪中叶至19世纪中叶。统计学在这个兴起阶段,还是一门意义和范围不太明确的学问,在它用文字或数字如实记录与分析国家社会经济状况的过程中,初步建立了统计研究的方法和规则。17世纪Pascal和Fermat的概率论;18世纪De Moivre、P. S. Laplace和Gauss的正态分布理论;最卓有成效地把古典概率论引进统计学的是法国天文学家、数学家、统计学家拉普拉斯(P. S. Laplace, 1749—1827)。因此,后来比利时大统计学家凯特勒指出,统计学应从Laplace开始。

1.2.1.1 拉普拉斯的主要贡献

拉普拉斯的主要贡献:①发展了概率论的研究,Laplace 第一篇关于概率论的表述发表于1774年。1812年发表的《概率分析理论》(先后出过4版)是他的代表作。Laplace 最早系统地把数学分析方法运用到概率论研究中,建立了严密的概率数学理论。②推广了概率论在统计中的应用,主要表现在人口统计、观察误差理论和概率论对于天文问题的应用。1809—1813年,Laplace 结合概率分布模型和中心极限思想来研究最小二乘法,首次为统计学中这项后来最常用的手段奠定了理论基础。③明确了统计学的大数法则,Laplace 发现在观察天体运动现象时,当次数足够多时,能使个体的特征趋于消失,而呈现出某种同一现象。Laplace 认为这其中一定存在着某些原因,而决非出于偶然。④进行了大样本推断的尝试,在统计发展史上,人口的推算问题,多少年来成为统计学家耿耿于怀的难题。直到19世纪初,Laplace 才用概率论的原理迈出了关键的一步。1781—1786年提出“拉普拉斯定理”(中心极限定理的一部分),初步建立了大样本推断的理论基础。在统计发展史上,他利用样本来推断总体的思想方法,开创了一条抽样调查的新路子。

1.2.1.2 高斯的主要贡献

另一位在概率论与统计学结合的研究上做出贡献的是德国大数学家高斯(C. F. Gauss, 1777—1855)。他的主要贡献有:①建立最小二乘法。1795年,Gauss 设想以残差平方和 $\sum (Y_i - a - bx_i)^2$ 为最小的情况下,求得的 a 与 b 来估计 α 与 β 。1798年完成最小二乘法的整个构思与结构,正式发表于1809年。②发现高斯分布。Gauss 以他丰富的数学实践经验,发现观察值 x 与真正值 μ 的误差变异大量服从现代人们最熟悉的正态分布。他运用极大似然法及其他数学方法,推导出测量误差的概率分布公式。“误差分布曲线”,这个术语就是Gauss 提出来的,后人为了纪念他,称该分布曲线为高斯曲线,也就是今天的正态分布曲线。Gauss 所发现的一般误差概率分布曲线以及据此来测定误差的方法,不仅在理论上,而且在应用上都有极为重要的意义。

1.2.2 近代描述统计学阶段

近代描述统计学形成期间大致在19世纪中叶至20世纪上半叶。由于生物学家们为了解决达尔文进化论中的复杂问题,经常需要借助统计学手段,而在这个过程中,原有的统计学方法的不足与局限性逐步地暴露出来。因此,许多学者在改善手段方面做了许多工作。19世纪达尔文应用统计方法研究生物界的连续性变异;孟德尔应用统计方法发现显性、分离、独立分配等遗传定律。由于这种“描

述”特色由一批研究生物进化的学者们提炼而成,因此历史上称他们为生物统计学派。生物统计学派的创始人是英国的高登(F. Galton, 1822—1911),主要发展是由Galton的得意门生K·泊松(K. Poisson, 1857—1936)完成的。

1.2.2.1 高登的主要贡献

高登的主要贡献:①初创生物统计学。Galton自1882年起开设“人体测量实验室”,在连续6年中,共测量了9 337人的“身高、体质量、阔度、呼吸力、效力和压力、手击的速率、听力、视力、色觉及个人的其他资料”,他深入钻研那些资料中隐藏着的内在联系,最终得出“祖先遗传法则”。在极其广泛地收集资料的同时,为了能使他的遗传理论建立在比较精确的基础上,出色地引入了中位数(Median)、百分位数(Persentile)、四分位数(Quartile)、四分位差(Quviation)以及分布、相关、回归等重要的统计学概念和方法。他在著作*Natural Inheritance*中首先提出了生物统计学(Biometry)一词,指出“所谓生物统计学,是应用于生物学中的现代统计方法”。②对统计学的贡献。变异是进化论中的重要概念,高登首先以统计方法加以处理,最终导致英国生物统计学派的创立。1889年, Galton把总体的定量测定法引入遗传研究中。通过总体测量发现,对动物或植物的每一个种别都可以决定一个平均类型。在一个种别中,所有个体都围绕着这个平均类型,并把它当做轴心向多方面变异。这就是他提出的“平均离差法则”。关于“相关”,统计相关法是由高登创造的。关于相关研究的起因,最早是他因度量甜豌豆的大小,觉察到子代在遗传后有“返于中亲”的现象。1877年,他搜集大量人体身高数据后,计算分析高个子父母、矮个子父母以及一高一矮父母的后代各有多少个高个子和矮个子女,从而把“父母高的后代高个子比较多,父母矮的其后代高个子比较少”这一定性认识具体化为父母与子女之间在身高方面的定量关系,并提出了相关函数(即现在常用的相关系数)的计算公式;关于“回归”,1870年,高登在研究人类身长的遗传时发现,高个子父母的子女,其身高有低于他们父母身高的趋势;相反,矮个子父母的子女,其身高却往往有高于他们父母身高的趋势。这就是统计学上“回归”的最初涵义。1886年,高登在论文“在遗传的身高中向中等身高的回归”中,正式提出了“回归”概念。

1.2.2.2 K·泊松的主要贡献

对生物统计学倾注心血,并把它上升到通用方法论高度的是K. Poisson,他对统计学的主要贡献有:①变异数据的处理。生物统计中所取得的数据常常是零乱的,很难看出其规律。K. Poisson首创的频数分布表与频数分布图成为统计方法中最基本的手段之一。②分布曲线的选配。19世纪以前,人们认为以频数分布描述变异值,最终都表现为正态分布曲线。但是,K. Poisson从生物统计资料的

经验分布中,注意到许多生物上的度量不具有正态分布,而常常呈偏态分布,甚至倾斜度很大,而且也不一定是单峰,也有非单峰的。1894年,他在“关于不对称频率曲线的分解”一文中首先把非对称的观察曲线分解为几个正态曲线。利用所谓“相对斜率”的方法得到12种分布函数型,其中包括正态分布、矩形分布、J型分布、U型分布或铃型分布等。③卡方检验的提出。1900年K. Poisson独立地又重新发现了 χ^2 分布,并提出了有名的“卡方检验法”(test of χ^2)。在自然现象的范围内, χ^2 检验法运用得很广泛。以后经R. A. Fisher补充,成了小样本推断统计的早期方法之一。④回归与相关的发展。回归与相关,经K. Poisson进一步发展后,这两个出自于生物统计学领域的概念,便被推广为一般统计方法论的重要概念。此外,K. Poisson还提出复相关、总相关、相关比等概念,不仅发展了Galton的相关理论,还为之建立了数学基础。

1.2.3 现代推断统计学阶段

现代推断统计学形成期间大致是20世纪初叶至20世纪中叶。人类历史进入20世纪后,无论社会领域还是自然领域都向统计学提出了更多的要求。各种事物与现象之间繁杂的数量关系以及一系列未知的数量变化,单靠记录或描述的统计方法已难以奏效。因此,相继产生了“推断”的方法来掌握事物总体的真正联系以及预测未来的发展。从描述统计学到推断统计学,这是统计发展过程中的一个大飞跃。统计学发展中的这场深刻变革是在农业田间试验领域中完成的因此,历史上称之为农业试验学派。对现代推断统计的建立贡献最大的是英国统计学家哥赛特(William Seely Gosset, 1876—1937)和R·费雪(Ronald Aylmer Fisher, 1890—1962)。

1.2.3.1 哥赛特的 t 检验与小样本思想

1908年,Gosset首次以Student为笔名,在*Biometrika*杂志上发表了“平均数的概率误差”一文。由于这篇文章提供了“学生 t 检验”的基础,为此,许多统计学家把1908年看做是统计推断理论发展史上的里程碑。此后,Gosset又连续发表了“相关系数的概率误差”(1909);“非随机抽样的样本平均数分布”(1909);“从无限总体随机抽样平均数的概率估算表”(1917)等。他在这些论文中,比较了平均误差与标准误差的两种计算方法,研究了泊松分布应用中的样本误差问题,建立了相关系数的抽样分布,导入了“学生氏分布”,即 t 分布。这些论文的完成,为“小样本理论”奠定了基础。由于Gosset开创的理论使统计学开始由大样本向小样本、由描述向推断发展,因此有人把哥赛特推崇为推断统计学的先驱者。人们认为Gosset研究成果的战略意义远比其战术意义大,它打开了人们的思路,