

试验统计分析

崔秀珍 黄中文 薛 香 编著



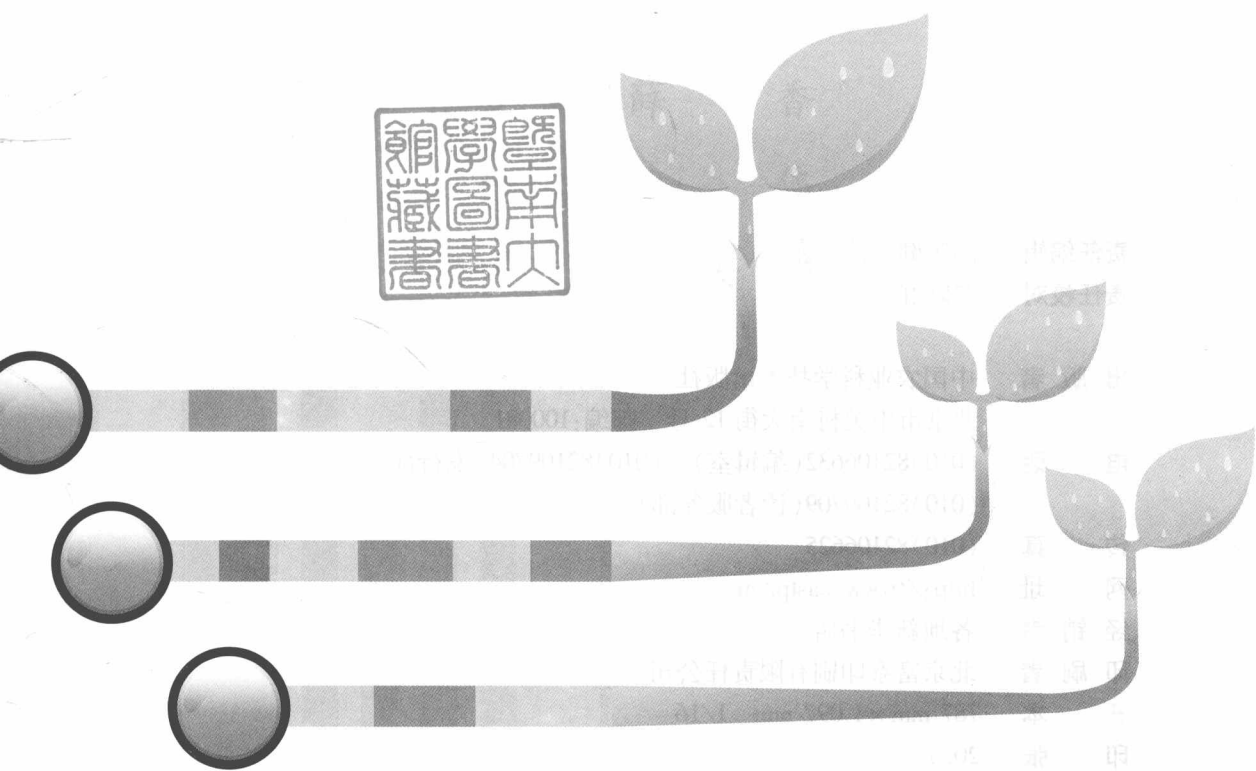
中国农业科学技术出版社

S3-33
2014.2

P₁

试验统计分析

崔秀珍 黄中文 薛 香 编著



中国农业科学技术出版社

图书在版编目(CIP)数据

试验统计分析 / 崔秀珍, 黄中文, 薛香编著. —北京: 中国农业科学技术出版社, 2013. 8

ISBN 978 - 7 - 5116 - 1333 - 2

I. ①试… II. ①崔…②黄…③薛… III. 生物统计 - 高等学校 - 教材 IV. ①Q - 332

中国版本图书馆 CIP 数据核字(2013)第 155873 号

责任编辑 闫庆健 范 潇
责任校对 贾晓红



出 版 者 中国农业科学技术出版社
北京市中关村南大街 12 号 邮编: 100081
电 话 (010)82106632(编辑室) (010)82109704(发行部)
(010)82109709(读者服务部)
传 真 (010)82106625
网 址 <http://www.castp.cn>
经 销 者 各地新华书店
印 刷 者 北京富泰印刷有限责任公司
开 本 787 mm × 1 092 mm 1/16
印 张 20.5
字 数 467 千字
版 次 2013 年 8 月第 2 版 2013 年 8 月第 3 次印刷
定 价 39.80 元

版权所有 · 翻印必究

《试验统计分析》

编著者：崔 秀 珍

黄 中 文

薛 香

王 伟

张 安 邦

内 容 简 介

该书系统地介绍了试验统计分析的原理和方法。全书十三章,可分为三个部分,第一部分(一至五章)试验统计分析基础,重点介绍试验资料整理与特征数、理论分布和抽样分布、平均数假设测验与区间估计、 χ^2 测验、方差分析等基本理论与统计方法;第二部分(六至八章)田间试验设计与统计分析,在介绍田间试验基本知识的基础上,介绍常用的随机区组、拉丁方、裂区、条区、正交设计等试验设计及统计分析方法;第三部分(九至十三章)回归及多元分析,主要介绍直线回归和相关分析、协方差分析、可线性化的曲线回归分析、多项式回归分析、多元回归与相关分析、通径分析等内容。全书内容力求简洁明快、言简意赅、深入浅出、循序渐进。强调实用性、通俗性、系统性,并具有一定的深广度。适合农科、生物学科工作者应用和参考,更适合作为农科及生物学科专业本、专科教材使用。

第一版前言

试验统计分析主要是把数理统计的原理和方法运用到农科及相关学科的研究中,解决试验设计、资料整理、统计分析等方面的问题。它是农科及相关学科本、专科学生的必修课、工具课、专业基础课,也是科技工作人员的必备工具书。

随着农业科学研究领域的不断扩展,试验统计分析方法除针对田间试验外也随之外延拓展,它包括对农科及相关学科各类模式调查和测定资料的处理和分析,也包括各种试验结果的统计分析。有鉴于此,我们把这本书定名为《试验统计分析》,以适合不同学科的需要。

该书的编著是在作者多年从事试验统计分析教学和应用研究的基础上完成的。书中所介绍的内容主要侧重于各种统计方法和应用,在统计原理方面,只作概念上的介绍,公式引出,尽量避开公式的繁琐推导,其目的是让读者在对统计原理有所了解的基础上,结合实例掌握常用的试验统计分析方法。全书共十三章,概括起来有三个部分,第一章至第五章在介绍试验资料整理与特征数、理论分布和抽样分布的基础上,重点介绍平均数假设测验、区间估计、 χ^2 测验、方差分析等基本的统计分析方法;第六章至第八章主要介绍田间试验基本知识和常用的试验设计及分析方法;第九章至第十三章介绍直线回归和相关分析、曲线回归分析、协方差分析、多元回归和相关及通径分析等内容。章后附有相应的习题和思考题,供读者参考。本书的例题一是来自于作者有关农学、植保、园林等学科的实际研究资料;二是其有关著作中的一些经典范例。

该书在内容安排上力求简洁明快、言简意赅、由浅入深、循序渐进,强调实用性、通俗性、系统性,并具有一定的深广度。适合农科、生物学科工作者应用和参考,更适合作为农科及生物学科专业本、专科教材使用。

全书由河南科技学院崔秀珍教授主编,黄中文、薛香等同志共同编写完成。崔秀珍教授负责全书的修改、润色、统稿、定稿等工作。在该书的出版过程中得到了中国农业科学技术出版社的大力支持,特别是闫庆健同志在书稿的编审和出版工作中作了大量工作。在此一并表示感谢。

由于作者水平有限,加之时间仓促。书中难免有误,敬请读者批评指正,以便再版修订完善。

崔秀珍

于河南科技学院

2002年6月

第二版前言

“试验统计分析”是我国高等农业院校农学、植保、生物类等各本科专业和综合、师范类本科专业以及成人教育、网络教育相应本科专业开设的一门重要的专业基础课和工具课,它提供常用的试验设计与资料统计分析的方法,培养学生研究分析的能力,也为“遗传学”、“育种学”、“数量遗传学”等后学科课程的学习打下统计学基础。

《试验统计分析》教材是河南科技学院生物统计教授崔秀珍,联合多位生物统计教师,结合当前农学、生物类院校相关学科本科生源特点、教学大纲,共同编写的本科专业基础课教材。该书自2002年6月出版至今,已在河南科技学院农学、种科、生技、生工、植保等专业及其他院校使用多年,由于该教材简洁明快、深入浅出,强调实用性、通俗性、系统性,并具有一定的深广度,深得学生和教师喜爱。在使用过程中,一些学者和教师也提出了一些修改意见,指出了本书第一版中出现的错误。本次再版根据读者的意见和教学的需要主要做了如下修订:第七、八章单因素、多因素随机区组试验分析增加了生物类专业的实例;第八章增加第五节“正交设计与统计分析”;对第五章方差分析、十二章多元回归和相关进行了系统简化;更换了各章的部分实例;同时订正了第一版中不妥和错误之处。

在本书第二版修订和出版过程中,得到了河南科技学院教务处和中国农业科学技术出版社的大力支持,在此一并感谢。

由于作者水平有限,不当之处敬请读者批评指正!

崔秀珍

于河南科技学院

2013年7月

目 录

绪 论	1
一、农业科学研究的任务	(1)
二、农业科学研究的基本方法	(1)
三、试验统计分析方法的功用	(3)
第一章 试验资料的整理与特征数	5
第一节 统计学中的几个基本术语	(5)
一、总体与样本	(5)
二、参数与统计数	(5)
三、机误与错误	(5)
四、准确性和精确性	(6)
第二节 试验资料的整理	(6)
一、试验资料的分类	(6)
二、次数分布表	(6)
三、次数分布图	(11)
第三节 试验资料的特征数	(12)
一、平均数	(12)
二、变异数	(15)
习 题	(18)
第二章 理论分布与抽样分布	20
第一节 概率基本知识	(20)
一、事件和概率的概念	(20)
二、计算概率的基本法则	(22)
第二节 二项分布	(23)
一、二项分布的概念	(23)
二、二项分布的概率计算	(24)
三、二项分布的形状和参数	(26)
四、二项分布的极限——泊松分布	(27)
第三节 正态分布	(29)
一、正态分布的概念	(29)
二、正态分布曲线的特征	(30)
三、正态分布的标准化	(31)
四、正态分布概率的计算	(31)

第四节 抽样分布	(35)
一、随机抽样和无偏估计	(35)
二、样本平均数的抽样分布	(37)
三、样本平均数差数的抽样分布	(39)
习 题	(42)
第三章 假设测验与区间估计	43
第一节 假设测验的基本原理	(43)
一、假设测验的意义	(43)
二、假设测验的推理步骤	(43)
三、两尾测验和一尾测验	(46)
第二节 单个平均数的假设测验和区间估计	(47)
一、学生氏 t 分布	(47)
二、单个平均数的假设测验	(48)
三、总体平均数的区间估计	(51)
四、假设测验与区间估计的关系	(53)
第三节 两个平均数的假设测验与区间估计	(54)
一、成组数据的比较	(54)
二、成对数据的比较	(61)
第四节 百分数的假设测验与区间估计	(63)
一、单个样本百分数的假设测验和区间估计	(63)
二、两个样本百分数的假设测验和区间估计	(67)
第五节 假设测验的两类错误	(70)
一、两类错误的概念	(70)
二、降低两类错误的方法	(71)
习 题	(72)
第四章 卡平方 (χ^2) 测验	74
第一节 χ^2 分布及 χ^2 测验	(74)
一、 χ^2 分布的概念	(74)
二、 χ^2 测验	(76)
第二节 方差同质性测验	(76)
一、样本方差与总体方差的同质性测验	(76)
二、多个样本方差的同质性测验	(77)
第三节 适合性测验	(78)
一、适合性 χ^2 测验的方法	(78)
二、各种遗传分离比适合性测验	(79)
三、配合理论分布的适合性测验	(80)
第四节 独立性测验	(82)
一、 2×2 表的独立性测验	(82)

二、 $2 \times c$ 表的独立性测验	(83)
三、 $r \times c$ 表的独立性测验	(85)
习 题	(86)
第五章 方差分析	87
第一节 方差分析的意义	(87)
一、方差分析的意义	(87)
二、方差分析的概念	(87)
第二节 方差分析的基本原理	(88)
一、自由度、平方和的分解	(88)
二、 F 分布及 F 测验	(89)
三、多重比较	(91)
四、多重比较结果的表示方法	(94)
第三节 单向分组资料的方差分析	(95)
一、组内观察值数目相等的单向分组资料的方差分析	(95)
二、组内观察值数目不相等的单向分组资料的方差分析	(97)
第四节 两向分组资料的方差分析	(99)
一、组合内只有单个观察值的两向分组资料的方差分析	(99)
二、组合内有重复观察值的两向分组资料的方差分析	(102)
第五节 系统分组资料的方差分析	(107)
第六节 方差分析的基本假定及数据转换	(111)
一、方差分析的数学模型和基本假定	(111)
二、不符合基本假定数据资料的转换	(114)
习 题	(116)
第六章 田间试验的基本知识	119
第一节 田间试验的种类和基本要求	(119)
一、田间试验的意义和种类	(119)
二、田间试验的基本要求	(120)
第二节 田间试验的误差及其控制	(121)
一、试验误差的概念和来源	(121)
二、试验误差的控制	(122)
三、试验地的选择和培养	(123)
四、控制土壤差异的小区技术	(124)
第三节 田间试验设计的基本原则	(127)
一、重复 (Replication)	(127)
二、随机 (Random)	(128)
三、局部控制 (Local Control)	(128)
第四节 试验方案的拟定	(129)
一、试验指标与效应	(129)

二、制定试验方案的要点	(130)
第五节 田间试验的实施	(132)
一、制定田间试验计划	(132)
二、试验地的准备和田间区划	(133)
三、种子准备	(133)
四、播种与移栽	(134)
五、栽培管理	(134)
六、田间试验的观察记载和测定	(134)
七、收获及脱粒	(136)
第六节 温室与实验室试验	(136)
习 题	(137)
第七章 单因素试验设计及统计分析.....	138
第一节 对比法、间比法试验设计及统计分析	(138)
一、对比法试验设计及统计分析	(138)
二、间比法试验设计及统计分析	(140)
第二节 完全随机和随机区组试验设计及统计分析	(141)
一、完全随机试验设计方法	(141)
二、随机区组试验设计方法	(142)
三、随机区组试验结果的统计分析	(143)
四、随机区组的线性模型与期望均方	(149)
五、随机区组试验的缺值估计与统计分析	(150)
第三节 拉丁方试验设计及统计分析	(154)
一、拉丁方试验设计方法	(154)
二、拉丁方试验结果的统计分析	(155)
三、拉丁方线性模型与期望均方	(159)
四、拉丁方试验的缺区估计和结果分析	(160)
习 题	(162)
第八章 多因素试验设计及统计分析.....	165
第一节 两因素随机区组试验设计及统计分析	(165)
一、两因素随机区组试验设计	(165)
二、两因素随机区组试验结果的统计分析	(165)
第二节 三因素随机区组试验设计及统计分析	(174)
一、三因素随机区组试验设计	(174)
二、三因素随机区组试验结果的统计分析	(174)
第三节 裂区试验设计及统计分析	(179)
一、裂区试验设计方法	(179)
二、裂区试验结果的统计分析	(180)
第四节 条区试验设计及统计分析	(184)

一、条区试验设计方法	(184)
二、条区试验结果的统计分析	(185)
第五节 正交设计及统计分析	(188)
一、正交设计的原理和方法	(188)
二、正交设计试验结果的统计分析	(192)
习 题	(199)
第九章 直线回归和相关分析.....	201
第一节 回归和相关的概念	(201)
第二节 直线回归方程	(203)
一、直线回归方程式	(203)
二、直线回归方程的计算和图示	(204)
三、直线回归的估计标准误	(206)
四、直线回归的数学模型和基本假定	(206)
第三节 直线回归的假设测验和区间估计	(207)
一、直线回归关系的假设测验	(207)
二、两个回归系数比较时的假设测验	(209)
三、回归系数和回归截距的区间估计	(210)
四、直线回归方程预测值的区间估计	(211)
第四节 有重复观察值的直线回归分析	(214)
第五节 直线相关	(218)
一、相关系数和决定系数	(218)
二、相关系数和决定系数的计算	(220)
三、相关系数的假设测验	(220)
习 题	(222)
第十章 协方差分析.....	223
第一节 协方差分析的意义和功用	(223)
一、协方差分析的意义	(223)
二、协方差分析的功用	(223)
第二节 单向分组资料的协方差分析	(223)
第三节 两向分组资料的协方差分析	(229)
习 题	(232)
第十一章 曲线回归分析.....	234
第一节 可线性化的曲线回归	(234)
一、对数函数曲线	(234)
二、指数函数曲线	(237)
三、幂函数曲线	(240)
四、Logistic 生长曲线	(242)
第二节 多项式曲线	(247)

一、二次多项式回归方程的建立	(248)
二、二次多项式回归方程的显著性测验和图示	(250)
习 题	(251)

第十二章 多元回归和相关..... 254

第一节 多元回归方程	(254)
一、多元回归方程和正规方程组的解	(254)
二、多元回归方程实例	(258)
三、多元线性回归方程的假设测验	(260)
四、自变量的相对重要性及取舍	(264)
第二节 多元相关分析	(266)
一、多元相关与偏相关	(266)
二、多元相关的假设测验	(269)
习 题	(271)

第十三章 通径分析..... 272

第一节 通径分析的基本原理	(272)
一、通径分析的意义和作用	(272)
二、通径系数的基本概念	(272)
三、通径系数的计算	(274)
四、通径系数的性质	(276)
第二节 通径分析的基本步骤	(279)
第三节 通径分析的显著性测验	(282)
一、线性关系的显著性测验	(282)
二、通径系数的显著性测验及不显著通径系数的剔除	(282)
三、通径系数差异显著性测验	(284)
习 题	(285)

附表..... 286

附表 1 10000 个随机数字	(286)
附表 2 累积正态分布 $F_N(u)$ 值表	(290)
附表 3 正态离差 u 值表(两尾)	(292)
附表 4 学生氏 t 值表(两尾)	(293)
附表 5 χ^2 值表(一尾)	(294)
附表 6 5%(上)和 1%(下)点 F 值(一尾)表	(295)
附表 7 Duncan's 新复极差测验 SSR 值	(301)
附表 8 5% 和 1% q 值表	(303)
附表 9 百分数反正弦($\sin^{-1}\sqrt{x}$)转换表	(305)
附表 10 r 与 R 的 5% 和 1% 显著值	(307)
附表 11 常用正交表	(308)

主要参考文献..... 312

绪 论

一、农业科学研究的任务

科学研究是人类认识自然,改造自然,服务社会的原动力,其目的在于探求新的知识、理论、方法、技术和产品。农业科学研究的根本任务是寻求提高农作物或其他植物的产量、品质,增加经济效益的理论、方法和技术。从而不断提高农业生产水平,改进人类生存环境,提高生存质量。农业科学研究和其他研究领域一样,有基础性和应用性研究,基础性或应用基础研究在于揭示新的知识、理论和方法;应用性研究在于获得某种新的技术或产品。它们都是通过一定的抽样调查、科学试验来实现的。

农业科学试验是农业科研中的一种自觉的、有计划的科研实践,其任务在于解决农业科研中提出的问题或农业生产中需要解决的问题。例如:优质小麦栽培研究中,如何应用正确的栽培方法使优质小麦既优质又高产的问题,就要设计一系列的栽培试验进行比较、研究,得到优质小麦优质高产的技术措施;又如:某地区棉花黄枯萎病流行严重,为了解决这个问题,就需要进行多方面农业试验,对各种防治病害措施进行鉴定,征集抗病品种进行比较试验,供生产择优选种,通过育种试验选育抗病品种等。农业科学试验通常是在易于控制的较小空间中进行的,因而,它有可能最大限度地排除各种非研究因素的干扰,将需要研究的问题充分地突出起来;同时,它又可以向各个方面试探解决问题的最佳方案,而不至于造成大的损失,所以农业试验能够有效地解决农业科学研究和农业生产中存在的问题,最终推动农业生产水平的提高。

二、农业科学研究的基本方法

农业科学研究的整个过程包括三个环节,首先根据研究者的观察或通过文献对所研究的内容形成一种设想或推测,然后根据对研究内容的设想安排农业试验或进行抽样调查,最后根据试验和调查所获得的资料进行推理总结,形成结论。或者进行新一轮试验验证或完善结论。如此循环发展,使获得的认识或理论逐步深化、发展。一般完成上述过程要做以下几个方面的工作。

(一)研究课题的选择与申报

1. 选题 在农业科学研究中,首先遇到的是选择什么样的课题也就是确定研究方向,这是搞好研究的第一步工作,是具有战略意义的大事。它直接关系到研究成果的质量水平,科研人员的远见卓识,首先就表在这里。优秀的科学人员主要在于选题时的明智,而不仅仅在于解决问题的能力。课题选择一般应遵循以下几个方面的原则:(1)具有借鉴创造性。农业科研选题应在比较鉴定前人研究的基础上,选本学科中前人没有做过的课题,或前人虽然做过但不完善的课题。(2)符合国民经济发展的需要。无论理论研究还是应用性研究,选题时必须明确其意义或重要性。理论性研究着重看所选课题在

未来国民经济发展上的重要性,而应用性研究着重看其对于未来发展的作用和潜力。应根据我国国情,国民经济的发展,选择有针对性的、关系到国计民生亟待解决的课题,自觉地把研究项目与发展国民经济、提高农业生产力联系起来。(3)课题要切实可行。确定课题必须量力而行,遵循可行性原则。即指资料、试验设备、经费、时间等能够满足研究的需要,选择的课题应在自己的知识领域范围内,便于发挥自己的业务专长。(4)课题大小适中。选择课题应根据研究者的知识和经验,注意课题大小适中,初学者以题目小一些为好,有经验的研究者可选择中等以上或跨学科的综合性研究课题。

2. 申报 课题确定以后,要向科研主管部门申报,申报人一般要填写科研项目申报书,申报书正文要写明以下几个方面的内容:(1)课题的目的和意义。目的是本研究课题要解决什么问题,达到什么目的。意义是课题成果对国民经济建设具有的贡献和应用价值,对科学和生产力的发展所起的促进作用。另外要明确本研究在国内外的的发展状况和达到的程度,这也体现该课题研究的价值、水平、难度。(2)研究的内容和预期目标。包括研究由哪几部分组成,各研究部分解决哪些问题及预期目标。(3)现有条件。一般从研究条件、技术、物质、经济等几方面阐述。(4)主要措施。包括人员组成、技术措施。(5)经费概算。(6)年度计划。按年度或阶段写出试验研究的进展安排。(7)效益分析。写明本研究完成后所产生的经济效益、社会效益和生态效益。

(二)资料搜集

搜集资料大体上有3种途径,即查阅文献资料、生产或田间调查、科学试验。

1. 查阅文献资料 科学的发展是累积性的,每一项研究都是在前人建筑的大厦顶层上添砖加瓦,这就要首先登上顶层,然后才能增加新的层次。研究工作者查阅文献就是便于把自己推上顶层,掌握大厦总体结构的通道。选题要有文献的依据,设计内容和方法要有文献的启示,研究过程需要文献资料的指引,查阅文献可以少走弯路,增加效率。文献资料通常包括图书、期刊、音像资料。要充分利用图书馆、资料室以及网上资料,掌握查找的方法,顺利获取自己所需要的资料。

2. 生产或田间调查 深入生产或田间进行调查,是获取资料的又一种途径,通过调查可以获得很多第一手资料。调查有两种,一种是普查,另一种是抽样调查。普查原则上是指对研究对象的每个个体都进行观测的一种全面调查。所以普查一般要规定观测的时间或时间范围,在农业研究方面有农业资源普查、土壤普查等。它主要在摸清家底方面中应用。但是农业科学研究的大量调查工作还是抽样调查,抽样调查往往是对有限总体而进行的。抽样调查的方法一般有系统抽样、典型抽样和随机抽样。(1)系统抽样也叫顺序抽样、机械抽样。它的特点是在总体中每隔若干单位抽取一个样本单位。例如:在一块麦田中,每隔若干距离抽取一个 1m^2 的面积作为样块,以便根据样块产量来计算麦田产量。又如调查果树单株产量,每隔10株抽取一株进行调查。一般田间试验用的三点、五点取样法,棋盘式取样法都属于这个类型。这种方法简单易行,但如果被调查的总体单位有某种周期性变异趋向时,样本会出现偏差。(2)典型抽样也叫经验抽样,主要根据调查者的经验和判断能力,有目的地在总体单位中抽选有代表性的单位或单位群。由于抽样带有主观认识,所以,此种方法在农业科学研究中不太常用。(3)随机抽样,随机抽样在农业科学研究中比较常用,它又分为简单随机抽样、分层随机抽样、整群随机抽样。简单随机抽样是总体中的每个单位都有

均等机会被取为样本单位。抽样的随机性由抽签法或查随机数字表法来满足。分层随机抽样是先将总体按变异性分为若干比较均匀的部分(或分层、分段),再在各部分随机抽取样本单位,这种抽样方法可以保证在各个变异层内都能抽取适当数目的样本单位。例如:要调查豫北地区土壤养分含量与产量的关系,可以先按该地区产量水平划分成低产区、中产区、高产区3个区段,然后再在每一个区段随机抽取样点进行调查。整群随机抽样是在总体中随机抽取包括若干样本单位在内的单位群,单位群内每个样本单位全部进行观测。如果把随机抽取的单位群视为初级样本单位,并且在初级样本单位内不作全面调查,而是再从中抽取一部分个体进行观测,这种抽样方法叫二级抽样。在具体抽样调查时,根据所调查总体的特点采用合适的抽样方法。

3. 科学试验 通过设计试验,在控制的环境条件下获得一些数据,这是农业科学研究的一个重要手段,大多数的农业科学研究必须通过严格的试验。

(三) 资料整理分析与总结

对搜集到的资料必须及时进行处理。调查资料可根据数据多少分为两种:第一种是简单平均数法,把同一性状调查的数据计算平均数,然后填入观察记载表,以待分析;第二种是次数分布法,此法适用于同一性状得到的数据比较多,把资料进行由小到大分组,统计数据在各组出现的次数,最后制成图表,直接分析。对产量资料还应该把小区产量或其他测产结果折算成公用的单位面积产量,在折算之前必须对原始资料进行充分核对,避免错误。对整理好的资料应进行相应的统计分析,通过分析,才能透过环境条件及其他偶然因素掩盖的表面现象揭示农作物现象自身所具有的规律。统计分析的方法很多,有差异显著性测验法、方差分析法、回归分析法等,这些是我们所涉及到的主要内容。

对一项研究项目告一段落或完成后,应将整理分析的资料有针对性的按试验的目的进行认真的总结分析,通过总结使资料系统化,从中找出有规律的东西,作出正确的结论,形成文字资料,作为进一步试验的依据或研究成果。如果试验结论重复率高,有着普遍指导意义或有突破性进展,可以写成论文在公开刊物上发表。

三、试验统计分析方法的功用

(一) 试验统计学的发展

统计(statistics)是一个古老的政治术语,原用于国家管理需要的统计数字,后来则将统计学看作实验数据搜集、分析及推理的理论方法的科学。试验统计学作为一门系统的学科是把统计学和试验设计相结合而发展起来的,它奠基于1925年R. A. Fisher出版的《Statistical Methods for Research Workers》一书,该书形成了试验统计学较为完整的体系,在这以前2~3个世纪内已经积累了有关概率、分布和统计方法的一些要素,例如:17世纪Pascal和Fermat的概率论;18世纪De Moivre、Laplace和Gauss的正态分布理论;19世纪达尔文应用统计方法研究生物界的连续性变异;孟德尔运用统计方法发现显性、分离、独立分配等遗传规律;20世纪以来Gosset用试验方法发现了 t 分布;Fisher提出了方差分析,建立了试验设计的三大原理,并提出了随机区组,拉丁方等试验设计,这些统计理论和方法的建立及在农业生物学中的应用,再加上统计学其他学科的发展,推动了试验统计学的发展。

(二) 试验统计分析方法的功用

在农业和生物学领域试验中,为了深刻认识研究对象或过程的表现和规律,往往需要认真地观察、测定、试验,搜集大量的资料。这些资料大多表现为一定的数量,杂乱无章,必须借助统计寻求其规律。试验统计分析方法作为农业试验的工具,可以有以下几个功用。

1. 提供整理和描述数据的科学方法 试验数据通常是有变异的,例如:同一品种 50 株梨的单株产量,相同条件下 100 个玉米果穗的穗重等往往各不相同,在统计上将某性状或特性的有变化的数据叫做变数(Variable),而变数的每一个具体值叫变量(Variate)或观察值(Observed value)。在农业试验中,可以研究一个变数,也可以联合研究两个或两个以上的有关变数。由于每一个变数包括很多变量,所以资料就显得庞杂零乱,而又不能说明任何问题。统计方法提供了整理资料、描述资料的科学程序,以及由几个特征数描述资料性质的统计方法,使研究者从少数几个统计数就能确定某性状变数的变异特征、典型水平,了解数据中蕴藏的信息。

2. 提供由样本推断总体的科学方法 在农业科学研究中经常要了解一些数量方面的问题,如某新品种亩产多少、生育期多长,某棉花品种绒长多少等。要得到这种数据,一般没有必要对研究对象所有个体都进行测定,而是从中抽取一部分进行研究。那么这一部分能否代表总体?我们知道生物界各种数量现象普遍存在着变异性,同一品种各植株的一些性状也不会完全相同,变异的原因是环境条件和其他偶然因素的影响,正因为这些变异的普遍存在,使抽取部分个体所研究的结论就存在着某种不确定性。试验统计方法就在于创造和发展一些理论和方法来度量这种不确定性,提供由样本推断总体的科学方法。

3. 提供通过误差分析测定处理效应的科学方法 农业试验通常是一种比较试验,有比较才有鉴别,从比较中确定最准确的理论、方法和技术。这种试验的基本特点是:除被比较因素可以取不同水平外,其余因素都作为试验条件要求保持一致,这样就能比较处理的效应,这是比较试验的“唯一差原则”。但实际上,无论试验控制得如何严格,处理所受到的试验条件决不可能完全相同,因而得到的处理间效应总是或多或少地混杂着一些试验条件不一致而产生的试验误差,表现出的处理间不同是处理本身不同还是误差所为尚待测验。试验统计方法提供了这种鉴定的科学方法。它可以透过环境条件及偶然因素掩盖的现象,揭示处理间的真实差异。

4. 提供科学试验设计的一些理论方法 为了以最少的人、物、财消耗而取得最多的试验信息和准确的处理比较,试验设计必须科学,而科学的设计又和相应的统计方法密切相关,因此,设计本身的科学性需要应用统计方法加以研究,每一种设计都必须有统计学依据,每一种设计的试验数据都必须配合相应的统计分析。根据试验结果的统计分析,又可为进一步试验或改进设计、减少误差和鉴别出较小的处理效应提供理论依据和实践指导。常有这样的情况,由于试验设计上的某些不当而丧失了大量的试验信息,究其原因,多半由于缺乏一定的统计学知识。

以上说明试验统计分析方法对农业及生物学领域科学研究全过程都是必不可少的工具,正确地使用这个工具,可以使科学试验结果更可靠,研究工作更加有效,所以它是农业科技工作者必须掌握的一门课程。