

JIANMING SHENGWU TONGJIXUE YUANLI
JI RUANJIAN YINGYONG JIAOCHENG

李一婧 马伟超 • 编著

简明生物统计学原理 及软件应用教程



兰州大学出版社

JIANMING SHENGWU TONGJIXUE YUANLI

JI RUANJIAN YINGYONG JIAOCHENG

兰州大学出版社

ISBN 978-7-311-04437-4

1. ①简... ②李... ③马... ④马... ⑤李... ⑥马... ⑦李... ⑧马... ⑨李... ⑩马...

①简... ②李... ③马... ④马... ⑤李... ⑥马... ⑦李... ⑧马... ⑨李... ⑩马...

①简... ②李... ③马... ④马... ⑤李... ⑥马... ⑦李... ⑧马... ⑨李... ⑩马...

①简... ②李... ③马... ④马... ⑤李... ⑥马... ⑦李... ⑧马... ⑨李... ⑩马...

李一婧 马伟超 · 编著

简明生物统计学原理 及软件应用教程



兰州大学出版社

图书在版编目(CIP)数据

简明生物统计学原理及软件应用教程 / 李一婧, 马伟超编著. —兰州: 兰州大学出版社, 2014. 4

ISBN 978-7-311-04437-4

I. ①简… II. ①李… ②马… III. ①生物统计—教材 ②生物统计—统计分析—软件包—教材 IV. ①Q-332

中国版本图书馆 CIP 数据核字(2014)第 073912 号

李一婧 · 马伟超 编著

策划编辑 宋 婷
责任编辑 郝可伟
封面设计 李鹏远

书 名 简明生物统计学原理及软件应用教程
作 者 李一婧 马伟超 编著
出版发行 兰州大学出版社 (地址: 兰州市天水南路 222 号 730000)
电 话 0931-8912613(总编办公室) 0931-8617156(营销中心)
0931-8914298(读者服务部)
网 址 <http://www.onbook.com.cn>
电子信箱 press@lzu.edu.cn
印 刷 兰州奥林印刷有限责任公司
开 本 787 mm × 1092 mm 1/16
印 张 19.5
字 数 445 千
版 次 2014 年 4 月第 1 版
印 次 2014 年 4 月第 1 次印刷
书 号 ISBN 978-7-311-04437-4
定 价 39.00 元

(图书若有破损、缺页、掉页可随时与本社联系)

前 言

生物统计学是各高校生物类、医学类及农学类专业的一门重要必修课程,也是一门实用性很强的工具性课程。学习生物统计学需要举一反三,既要对生物统计学的基本概念、基本内容有一定的理解和掌握,也要通过具体案例了解不同问题的解题思路和方法。另外,新形势对专业技术人才的计算机软件应用水平也提出了新的要求。

我们在长期的教学实践和深入调查研究中发现,各专业本科学生初次学习生物统计学时,迫切需要一本能将统计学基本理论与方法、简明统计基础知识、统计实际工作与 Excel、SPSS 等软件强大数据处理与分析能力紧密结合的教材,以便于更好地学习与理解,这是我们编写本教材的原因。我们在结合多年来本课程的教学经验,并参阅国内外相关资料的基础上,将教学设计理念与调动学生学习积极性相结合,编写成这本新颖且实用的教材,以供不同院校的同仁们选择,亦可作为生物技术、生物工程、农业等领域从事生产、科研等相关工作人员的参考书。

本书重点突出可操作性,各章节都具体介绍了基本操作步骤与方法,学生可以按教材指引的策略和具体方法来解决学习、工作中的数据处理与分析问题;同时本书力求通俗易懂,对统计学基本原理、基础知识、统计方法与计算公式只作简要介绍,以具体案例说明,针对生物统计学中经常碰到的数据处理与分析问题,配合软件的具体使用方法,使学生通过学习,不断提高统计应用能力和统计理论水平。

本教材第一章、第三章、第四章、第六章、第九章和附表由李一婧编写,第二章、第五章、第七章、第八章、第十章、第十一章由马伟超编写。

在编写中,编者虽然已尽心竭力,但限于学识水平与经验,书中难免有疏漏欠妥之处,恳请同行专家和读者批评指正。

编者

2014年3月

目 录

第一章 概述	001
1.1 生物统计学的概念	001
1.2 生物统计学的功能	001
1.3 生物统计学的学习方法与要求	002
第二章 试验设计	004
2.1 试验设计概述	004
2.2 试验方案的拟定	010
2.3 正交试验设计	012
第三章 利用 Excel 进行试验资料的整理	018
3.1 Excel 在统计中的应用基础	018
3.2 利用 Excel 进行试验资料的整理	029
第四章 利用 Excel 计算试验资料的特征数	046
4.1 利用 Excel 计算数值的平均指标	046
4.2 利用 Excel 计算位置平均指标	051
4.3 利用 Excel 计算变异指标	054
4.4 利用 Excel 进行描述性统计	057
第五章 概率计算与随机抽样	061
5.1 概率分布	061
5.2 统计数分布	070
5.3 利用 Excel 计算概率	076
5.4 利用 Excel 进行抽样设计	080
第六章 假设检验与参数估计	089
6.1 假设检验的原理与方法	089
6.2 常用参数检验实例	093
6.3 利用 Excel 进行参数检验	108
6.4 参数估计的原理与方法	116
6.5 利用 Excel 进行区间估计	120
第七章 抽样	128
7.1 抽样的基本方法	128
7.2 抽样误差的估计	129

7.3 样本容量的确定	130
第八章 χ^2 检验	135
8.1 χ^2 检验的原理与方法	135
8.2 适合性检验	136
8.3 独立性检验	138
8.4 利用 Excel 进行 χ^2 检验	143
第九章 方差分析	150
9.1 方差分析的基本原理	150
9.2 单因素方差分析	157
9.3 二因素方差分析	161
9.4 方差分析的基本假定和数据转换	170
9.5 利用 Excel 进行方差分析的实例	172
第十章 回归分析与相关分析	180
10.1 回归和相关的概念	180
10.2 直线回归分析	181
10.3 直线相关分析	184
10.4 利用 Excel 进行回归分析和相关分析	185
第十一章 SPSS 常用生物统计功能简介	194
11.1 SPSS 基本知识	194
11.2 t 检验	197
11.3 χ^2 检验	204
11.4 方差分析	211
11.5 协方差分析	222
11.6 线性相关分析	232
11.7 回归分析	234
参考文献	241
附表	242
英汉术语对照	281
索引	299

第一章 概述

1.1 生物统计学的概念

统计学是在相对有限的样本数据上,对特定的随机现象作推断的学科。统计学有两个主要的分支:数理统计学与应用统计学。数理统计学更关注统计推断中新方法的发展,要求有较多的抽象数学知识作为工具;应用统计学则关心如何把数理统计方法应用到特定的领域,比如经济学、心理学及公共卫生学。生物统计学是应用统计学的分支,它将统计方法应用到医学及生物学领域,在此,数理统计学和应用统计学有些重叠,例如在某些实例中,某个已有的标准统计方法不大适用就必须加以修正,在这种情形下,生物统计学就涉及如何去发展新的方法。

1.2 生物统计学的功能

1.2.1 生物统计学的内容

生物统计学的内容包括试验设计和统计分析。试验设计是指应用数理统计的原理与方法,制订试验方案,选择试验材料,合理分组,降低试验误差,使我们可以利用较少的人力、物力和时间,获得丰富而可靠的数据资料。统计分析是指应用数理统计的原理与方法对数据资料进行分析与推断,认识客观事物的本质和规律性,使我们对所研究的资料得出合理的结论。由于事物都是相互联系的,统计不能孤立地研究各种现象,而必须通过一定数量的观察,从这些观察结果中研究事物间的相互关系,揭示出事物客观存在的规律性。统计分析与试验设计是不可分割的两部分。试验设计须以统计分析的原理和方法为基础,而正确设计的试验又为统计分析提供了丰富、可靠的信息,两者紧密结合推断出合理的结论,不断地推动应用生物科学研究的发展。

1.2.2 生物统计学的功能

现代生物统计学已在科学研究和生产实践中得到极为广泛的应用。其基本功能有:

1.为科学地整理、分析数据提供方法

我们做任何工作,都必须掌握基本情况,做到心中有数,才能有的放矢,从而提高工作质量。在生物科学研究中也不例外,必须有计划地收集资料并进行合理的统计分析,通过调查得到数据,经过加工整理,从中归纳出事物的内在规律性,用于指导生产。例如,调查某水稻品

种的穗粒数,可以得到不同个体穗粒数的大量原始数据,从这些杂乱的数据中很难看出什么规律,若运用生物统计方法对这些数据进行加工整理,使之条理化,就可以大体了解该水稻品种穗粒数的一般情况及其变异特征,进而得到有用的信息。

2. 判断试验结果的可靠性

由于存在试验误差,从试验得到数据资料必须借助统计分析方法才能获得可靠的结论。例如,某农场要研究两种饲料对肉用仔鸡增重及饲料利用率的影响。选择同品种及体重接近的500只肉用雏鸡,半数饲以甲种饲料,半数饲以乙种饲料,8周后称其体重并结算饲料消耗,分析、比较这些资料,从中得出结论。这就要运用统计分析方法,以决定两群鸡体重及饲料消耗的差异,究竟是由本质原因造成的,抑或是由随机因素造成的,即判断是由不同饲料造成的,还是由于其他未经控制的偶然因素引起的。分析之后才能得出比较正确的结论。

3. 确定事物之间的相互关系

科学试验的目的,不仅是研究事物的特征、特性,同时还要研究事物间的相互关系。例如,测定某群奶牛第一胎的产乳量和以后几胎的产乳量之间的相互关系,就可以根据第一胎产乳量的高低来推断一生的产乳量,这样,就为早期选择和淘汰低产乳牛提供科学依据。这种研究事物之间的联系形式以及相关程度的方法是生物统计的一个重要部分。

4. 提供试验设计的原则和方法

做任何调查或试验工作,事先必须有周密的计划和合理的试验设计,它是决定科研工作成败的一个重要环节。一个好的试验设计,可以用较少的人力、物力和时间,最大限度获得丰富而可靠的资料,尽量降低试验误差,从试验所得的数据中能够无偏地估计处理效应和试验误差的估计值,以便从中得出正确的结论。相反,设计不周,不仅不能得到正确的试验结果,而且还会带来经济上的损失。

5. 为学习其他课程奠定基础

我们要学好遗传学、育种学等学科,就必须学好生物统计学。比如,数量遗传学就是应用生物统计方法研究数量性状遗传与变异规律的一门学科,如果不懂生物统计学,也就无法掌握遗传学。此外,阅读中外科技文献也常常会碰到统计分析的问题,也必须有生物统计的基础知识。因此,生物科学工作者必须学习和掌握统计方法,才能正确认识客观事物存在的规律性,提高工作质量。

生物统计学在生物科学研究中虽然有着重要的作用,但也不是万能的。因为生物科学中很多现象受物理学、化学、生物学规律的支配。所以,生物科学研究工作,应该在生物学相关专业理论的指导下进行。

总之,生物统计学是一种很有用的工具,正确使用这一工具可以使科学研究更加有效,使生产效益更高,所以,它是每位生物科学工作者必须掌握的基础知识。

1.3 生物统计学的学习方法与要求

生物统计学是数学与生物学相结合的一门交叉学科,所包含的公式很多,我们在学习过程中,首先要弄懂统计的基本原理和基本公式。要理解每一公式的含义和应用条件,可以不必

深究其数学推导。其次,要认真地做好习题作业,加深对公式及统计步骤的理解,达到能熟练地应用统计方法。第三,应注意培养科学的统计思维方法。生物统计意味着一种新的思考方法——从不确定性或概率的角度来思考问题和分析科学试验的结果。第四,必须联系实际,结合专业,了解统计方法的实际应用;平日要留意书籍和杂志中的表格、数据及其分析和解释,以熟悉表达方法及应用。另外,学习生物统计学的一个非常好的方法是在研究过程中,从开始计划阶段直到完成期间都能参与在内,这包括在研究完成时书写研究报告或为出版物写初稿。

第二章 试验设计

2.1 试验设计概述

试验设计 (design of experiment, DOE) 是以概率论与数理统计为理论基础, 合理安排试验的一种方法论, 它研究如何高效而经济地获取数据信息、科学地分析处理、得出正确的结论。试验设计的创始人是英国农学家、遗传学家、统计学家费舍尔 (R. Fisher, 1890—1962), 他于 20 世纪 20 年代至 40 年代, 建立了试验设计和方差分析方法。费舍尔的统计学研究始终和农业科学、生物学、遗传学研究紧密联系。20 世纪 40 年代, 英、美、日等工业化国家把试验设计及其分析方法推广到工业领域并予以发展。日本学者田口玄一等人相继提出了正交试验设计、参数设计和容差设计, 使多因素试验能够部分实施, 并解决了产品制造过程中的稳健性问题。美国学者 G. E. P. Box 提出了调优运算设计, 中国学者方开泰等提出了均匀设计, 这些都发展和丰富了试验设计的内容。本章将学习试验设计的基本要素、试验计划和试验方案的拟订、试验误差的控制、试验的评价和试验设计的基本原则等内容, 以讲解单因素试验和多因素试验设计的思路为主, 进而详细介绍正交设计方法。

2.1.1 试验设计的基本概念

我们先简要介绍几个试验设计的基本概念, 以便读者理解本章后面的内容。

1. 因素 (factor) 及其水平 (level)

生物学研究中, 不论农作物还是微生物, 其生长、发育以及最终所表现的产量等均受多种因素的影响, 其中有些属于自然的因素, 如光、温、湿、气、土、病、虫等, 有些属于栽培条件, 如肥料、水分、生长素、农药、除草剂等。进行科学试验时, 必须在固定大多数因素的条件下才能研究一个或几个因素的作用, 从变动这一个或几个因子的不同处理中比较、鉴别出最佳的一个或几个处理。这里被固定的因子在全试验中保持一致, 组成了相对一致的试验条件, 被变动并有待比较的一组处理的因子称为试验因素, 简称因素或因子, 试验因素量的不同级别或质的不同状态称为水平。试验因素的水平可以是定性的, 如不同品种, 具有质的区别, 称为质量水平; 也可以是定量的, 如喷施生长素的不同浓度, 具有量的差异, 称为数量水平。数量水平不同级别间的差异可以等间距, 也可以不等间距。

2. 主效应 (main effect) 与交互效应 (interaction effect)

由于试验因素的作用, 对试验指标所引起的增加或减少称为试验因素的效应, 估算和分析试验因素效应是科学试验的目标之一。

试验中由一个自变量的不同水平引起的变异叫作该自变量的主效应。在单因素试验

中,主效应就是自变量不同水平得到的数据的方差。例如在“父母受教育程度对家庭教养方式的影响”研究中,自变量“父母受教育程度”分成三个水平:“小学及以下”、“中学”和“本科及以上”,那么该自变量的主效应就是这三个水平下所得到的因变量数据的方差。在多因素试验中,一个自变量的主效应就是忽略其他自变量的差异,只看该自变量本身不同水平的方差。

当一个因素对因变量的影响大小受其他因素水平或安排的影响而不同时,就产生了交互作用,即一个因素的水平在另一个因素的不同水平上变化趋势不一致,称这两个因素之间存在交互作用。这种交互作用所产生的效应称为交互效应。

3. 多重比较(multiple comparisons)

方差分析的结果若是差异显著,则表明几个试验处理中至少有一对平均数间的差异显著,具体是哪对处理的差异显著,需要对每个自变量的各个水平之间的数据进行两两比较,这个过程叫作多重比较。试验目的、性质或条件不同,研究者所关心的多重比较内容也不同。例如在“家庭类型和归因方式对孤独感的影响”研究中,“家庭类型”有权威型、放任型和民主型3个水平,“归因方式”有外归因和内归因2个水平。数据统计结果表明两个自变量的交互作用显著。为了进一步探讨家庭教养方式的3个水平和归因方式的2个水平到底在哪种处理下的差异显著,需要对 3×2 共6种处理进行差异比较。

4. 总体(population)与样本(sample)

我们把研究对象的某种数量性质的一个数值集合,或简述为研究对象的全体,称为总体(或母体),而组成总体的每个单元(元素)称为个体。例如,研究泰山一号小麦的株高,所有泰山一号小麦品种的株高构成一个总体,同样,所有泰山一号小麦的穗重又可构成另一个总体。总体的性质取决于个体的性质,是一个与一定的数量指标相联系的概念,因此构成总体的单元应是清晰的,它构成了抽样的基础,总体中的个体数称为总体容量,用 N 表示。

若总体中的个体为有限个,如一块玉米田的所有玉米穗、一个畜群的头数等,则称总体为有限总体。若总体中的个体为无限多个,如某一品种粒重、一块农田中小麦的株数等,则称总体为无限总体。若从总体中随机地抽取 n 个个体,得到 n 个观察值 $y_1, y_2, \dots, y_n$,称 $y_1, y_2, \dots, y_n$ 为该总体容量为 n 的样本(或子样)。在统计上,当以平均数为考察对象时,一般规定:

$n < 30$ 为小样本;

$n \geq 30$ 为大样本。

样本是总体的缩影,统计分析的任务就是由样本推断总体,因此,任何试验都存在抽样问题。为使样本正确地反映总体,抽样时必须采用随机的方法。所谓随机抽样是指总体中的每一个体都有同等的机会被抽取,而且每次抽取时,总体中的个体成分不变。采用这种方法得到的样本称简单随机样本,这种抽样方法称简单随机抽样。显然,简单随机抽样就是一个独立地、重复地进行的同一随机试验,为了做到这一点,对有限总体要采用复置抽样(即放回抽样)法,对无限总体则无此必要了。样本毕竟只是总体的一部分个体,因此和总体的真实情况又有所出入。统计分析的核心在于由样本的信息推断总体的信息。因此,获得样本仅是一种手段,而推断总体才是真正目的。

2.1.2 试验设计的统计思想

在科学研究中,绝大多数情况下,研究者是不能、也没有必要观测所研究事件的总体的,

这就需要利用统计检验的知识,观测总体中的一部分成员(即样本),通过统计推论间接地对总体进行估计。这就是为什么必须用统计的思想进行试验设计的原因。

进行试验设计的第一步就是根据已有的理论和经验提出研究假设。如“联想教学法好于传统教学法”,这在统计检验中称为备择假设,记作 H_1 。研究者无法直接检验备择假设,而常常是检验它的无效形式,即“联想教学法与传统教学法没有差异”,这在统计检验中称为无效假设,记作 H_0 。这是两个对立的假设。统计检验的逻辑是,先假设 H_0 成立,并在此条件下依据抽样分布理论做出推论。如果两种教学方法教学效果的差别在统计学意义上显著,即这种差别的出现大于一定的概率,则说明两种教学方法等效的假设是以小概率出现的。通过小概率的实际推断原理,即“小概率事件在一次试验中不可能发生”,可以推断两种教学方法等效的假设是错误的。这样便推翻无效假设,接受备择假设,得到“联想教学法好于传统教学法”的结论。

试验中所获得的数据,通常是利用方差分析来得到各种处理效应。方差分析是将总变异的各个来源分离出来,将各种变异与误差变异相比较。如果某一变异显著地大于误差变异,说明它的存在不是由随机误差造成的,而是表现了这一变异来源的处理效应;如果某一变异与误差变异差异不显著,则说明这种变异可能是由随机误差造成的,并不是其处理效应。方差分析不仅能够同时检验多组平均数间的差异,还可以解释几个因素之间的交互作用,因此能够对试验设计中的假设进行显著性检验。

在基础研究中,试验用来验证理论是否成立,并提供用于解释客观现象的数据。试验设计的中心思想是操纵自变量、控制额外变量。理想的试验是除了自变量外其他因素都保持恒定,那么从逻辑上可以认为,结果的任何不同都是由自变量引起的。然后用方差分析对得到的数据进行统计检验,尽量分离出除试验处理外的其他无关变异,将各种变异与误差变异相比较,得到各种变异是否显著大于误差变异的结果,最后根据试验目的进行解释。

2.1.3 试验设计的基本原则

1. 对照原则

(1)空白对照指在不加任何处理的“空白”条件下进行观察的对照。例如研究某种可疑致癌物质对动物进行诱癌试验,设立与试验动物同种属、同年龄、同性别、同体重的动物做空白对照,使试验组和对照组在相同的饲养条件下饲养,观察两组的肿瘤发生情况。

(2)试验对照指在某种有关的试验条件下进行观察的对照。例如为了研究中草药烟熏剂的空气灭菌作用,如果只采用空白对照(不给任何处理),并不能最后说明是中草药的作用,还是烟本身的作用,所以还需设立不加中草药的单纯烟熏对照。

(3)标准对照指以标准值或正常值作为对照,以及在所谓标准的条件下进行的观察对照。例如研究某种手术方法的肝癌治疗效果,如果它声称“50%”治愈,可采用以往的治疗经验作为标准值(如:95%存活期在3年以内)来进行对比。

(4)相互对照指各试验组间互为对照。例如几种药物治疗同一种疾病的疗效观察,比较其疗效的差别。

(5)自身对照指对照与试验在同一受试对象进行。例如对某降压药的疗效观察,可用用药前后的血压值作对比。

2. 重复原则

重复指各处理组及对照组的例数(或试验次数)要有一定的数量。如果例数太少,有可能把个别情况误以为普遍现象,把偶然性或巧合的现象当作必然的规律性现象。

3. 随机化原则

所谓随机化,是保证研究对象有同样的机会进入试验组或对照组。在试验研究中,试验组和对照组除了处理因素外,其他可能产生混杂效应的非处理因素尽可能保持一致,即均衡性要好。贯彻随机化原则是提高组间均衡性的一个重要手段,同时也是资料统计分析时进行统计推断的前提。

4. 局部控制原则

局部控制是指在试验时采取一定的技术措施或方法来控制或降低非试验因素对试验结果的影响。在试验环境或试验单位差异大的情况下,根据局部控制的原则,可将整个试验环境或试验单位分成若干个小环境或小组,在小环境或小组内使非处理因素尽量一致。因为单位组之间的差异可在方差分析时从试验误差中分离出来,所以局部控制原则能较好地降低试验误差。

2.1.4 试验设计的步骤

1. 识别问题和提出问题

将问题摆明并变为大家可以接受的题目并不那么简单,需要弄清有关试验目的的全部想法。清晰的提问对更好地理解现象和最终求得问题的解答有重大帮助。

2. 选择因素和水平

试验者必须选择在试验中准备用来处理的因素,以及在做试验时规定这些因素的水平。还必须考虑如何将这些因素控制在所希望的数值上以及如何测量这些数值。

3. 选择响应变量

在选择响应变量时,试验者应确信这一变量真正会对所研究的过程提供有用的信息。最常用的是取测量特性的平均值或标准差(或两者)为响应变量。

4. 设计试验

选择试验方案涉及考虑样本量(重复次数)、对试验选择合适的试验次序、确定是否划分区组或是否涉及其他随机化约束。重要的是思想上总要关注试验目的。

5. 进行试验

当进行试验时,谨慎地监视试验的进程以确保每件事情都按计划做完,同时应当及时地将试验中的现象和数据记录下来。

6. 数据分析

分析数据应采用统计方法,使得结果和结论都是客观的,为试验结果的可靠性和有效性提供准则。

7. 得出结论和建议

根据试验观察的现象和记录的数据,通过分析、计算、图表、推理等处理,归纳出一般概括性判断,并用文字、表格、绘图等方法做一个简明的总结。有时还需进行跟踪试验和确认试验,以证实所得结论的正确性。

2.1.5 常用的几种试验设计方法

1.自身对照设计

又称单组比较设计,试验与对照是在同一受试个体上进行的。有三种形式:

- (1)第一次观察不给处理,第二次观察给予处理。
- (2)第一次给予A处理,第二次给予B处理。
- (3)交叉设计。例如测定两种减肥药的效果,拟采用自身对照的方法。如果在同一批受试对象上进行两种减肥药试验,先全部用A药一个月,再用B药一个月,可能会出现因先用后用产生的系统误差,为尽量避免产生此类误差,可采用交叉设计的方法,使得A、B两药在受试对象身上先用后用的比例达到均衡。具体做法如下:①将10名受试对象按体重排序;②采用随机数字表或其他随机化方法,确定第1名受试对象的测试顺序,例如先用A药后用B药,用“A-B”表示,那么第2名受试对象的用药顺序则为“B-A”;③依此类推。

2.完全随机设计

又称单因素设计,它是将受试对象按随机化的方法分配到每个处理组中,观察试验效应。每个处理组样本例数可以相等,也可以不等,但相等时效率较高。

例如将10头动物随机分配到甲、乙两组(见表2-1)。

表 2-1 完全随机法 10 头动物两组分配表

动物号	1	2	3	4	5	6	7	8	9	10
随机数字	4	26	21	46	87	77	33	94	1	70
分入组别	乙	乙	甲	乙	甲	甲	甲	乙	甲	乙

具体做法如下:①先将动物按体重排序,编号;②分配随机数字,如利用附表11随机数字表中第6行第5、6列开始(两位一组)向下抄10个随机数字;③按预先规定,随机数字为单数者分入甲组,双数者分入乙组。

结果:甲组包括第3、5、6、7、9号动物;乙组包括第1、2、4、8、10号动物。

3.配对设计

将受试对象按一定条件配成对子,再随机分配每对中的两个受试对象到不同处理组。

例如取20只小白鼠,按体重、窝别,配成10对,试用随机数字表将其分配到甲组和乙组中去(见表2-2)。

表 2-2 随机配对法 20 只小白鼠分组表

配对号	1	2	3	4	5	6	7	8	9	10
随机数字	60	83	95	53	23	50	2	50	40	91
第1号动物	乙	甲	甲	甲	甲	乙	乙	乙	乙	甲
第2号动物	甲	乙	乙	乙	乙	甲	甲	甲	甲	乙

具体做法如下:①先将10对小白鼠编上配对号,各对再按体重编成第1号动物和第2号

动物;②分配随机数字,如从附表11中第3行第9列开始,依次向后抄10个随机数字;③按预先规定,随机数字为单数者第1号动物分入甲组,第2号动物分入乙组;遇双数,第1号动物分入乙组,第2号动物分入甲组。

4. 配伍组设计

又称随机区组设计(randomized block design)。首先按一定条件(配伍条件)划分出若干配伍组,每个配伍组中随机抽取与处理组数相同的动物只数,然后将每个配伍组中的动物随机分配到各处理组。

例如将20只不同体重的动物随机分到4种处理组。首先按动物的体重等分为5个区组,每个区组各有4只体重基本相同的动物。依次编好号码,第一窝4只动物编为1、2、3、4号,第二窝编为5、6、7、8号,余类推。然后在附表11中任意指定一个点,假如指定第8行第4个数字为起点,依横向抄录数目,先抄录3个数目为89、87、77,为随机分配第一窝动物之用,再将这3个数目依次以4、3、2除之,第一数目89,除以4得余数1,将第一号动物分配给第1组(A组),第2个数目87除以3得余数0,将第2号动物分配给剩下的B、C、D三个组中的最后1组(即D组)去。第3个数目77除以2得余数1,将第3号动物分配到剩余的B、C两组中的第1组(即B组)去。第4号动物即分入剩余的C组。第一窝动物分配完了以后,再继续抄录随机数据,用同样方法把其余各窝动物分配到各组去,如表2-3。

表 2-3 随机区组法 20 只动物分组表

动物编号	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
随机数字	89	87	77	—	53	29	21	—	46	51	99	—	89	87	33	—	94	22	21	—
除数	4	3	2	—	4	3	2	—	4	3	2	—	4	3	2	—	4	3	2	—
余数	1	0	1	—	1	2	1	—	2	0	1	—	1	0	1	—	2	1	1	—
归组	A	D	B	C	A	C	B	D	B	D	A	C	A	D	B	C	B	A	C	D

结果:A组1、5、11、13、18;B组3、7、9、15、17;C组4、6、12、16、19;D组2、8、10、14、20。

假如这个试验的四种处理方法为甲、乙、丙、丁,哪一种方法用A组动物?哪一种方法用B组动物?……我们还可以用随机数字表进行分配(抄录三个随机数字,分别以4、3、2除之,按余数进行分配)。

随机区组设计中把条件一致的研究对象编入同一区组并分配于各研究组,使各研究组之间可比性更强,在最后的统计分析中由于扣除了各区组间不同条件产生的影响,因而随机误差比较小,研究的效率比较高。

5. 拉丁方设计

拉丁方设计(Latin square design)是从横行和直列两个方向进行双重局部控制,使得横行和直列两向皆成单位组,是比随机单位区组设计多一个单位组的设计。这种设计方法特别适用于有两个明显试验误差的试验(例如田间试验中时间、空间上的差异),实行双重局部控制,大大提高了分析的精确度。

6.析因设计
析因设计(factorial design)是一种多因素交叉分组设计,它不仅可检验每个因素各个水平之间的差异,而且可检验各因素间的交互作用。

7.正交设计
正交设计(orthogonal design)是一种研究与处理多因素的试验设计方法,它是按照一系列的规格化的正交表来安排试验,从中找到最优试验条件。其特点是样本量小、试验次数少、均衡性强。

2.2 试验方案的拟定

2.2.1 试验方案的基本概念

试验方案(experimental scheme)是指根据试验目的与要求而拟定的进行比较的一组试验处理的总称。试验方案是整个试验工作的核心部分,因此需周密考虑、慎重拟定。试验方案按照供试因素的多少可区分为单因素试验方案和多因素试验方案。

1.单因素试验方案

单因素试验(single-factor experiment)是指整个试验中只比较一个试验因素的不同水平的试验。单因素试验方案由该试验因素的所有水平构成。这是最基本、最简单的试验方案。例如配制4种不同pH值的模拟酸雨,进行樟树幼苗的胁迫试验。这是一个有4个水平的单因素试验,配置模拟酸雨的4种不同pH值,即该因素的4个水平就构成了试验方案。

2.多因素试验方案

多因素试验(multiple-factor or factorial experiment)是指在同一试验中同时研究两个或两个以上试验因素的试验。多因素试验方案由该试验的所有试验因素的水平组合(即处理)组成。多因素试验方案又分为完全方案和不完全方案两类。

(1)完全方案

在列出因素水平组合(即处理)时,要求每一个因素的每个水平都要相遇一次,这时,水平组合(即处理)数等于各个因素水平数的乘积。例如以3种不同强度的光对3种草本植物进行试验。两个因素分别为光强(A)、草本植物(B)。光强(A)分为A₁、A₂、A₃水平,草本植物(B)分为B₁、B₂、B₃水平。组合有A₁B₁、A₁B₂、A₁B₃、A₂B₁、A₂B₂、A₂B₃、A₃B₁、A₃B₂、A₃B₃,共3×3=9个处理。这9个处理(水平组合)就构成了这两个因素的试验方案。根据完全方案进行的试验称为全面试验。这既能考察试验因素对试验指标的影响,也能考察因素间的交互作用,并能选出最优水平组合,从而充分揭示事物的内部规律。多因素全面试验的效率高于多个单因素试验的效率。全面试验的主要不足是:当因素个数和水平数较多时,水平组合数太多,以至于在试验时人力、物力、财力、场地等都难以承受,试验误差也不易控制。因而全面试验宜在因素个数和水平数都较少时应用。

(2)不完全方案

这也是一种多因素试验方案,但与上述多因素试验完全方案不同,它是将试验因素的某些水平组合在一起形成少数几个水平组合。这种试验方案的目的在于探讨试验因素中某些

水平组合的综合作用,而不在于考察试验因素对试验指标的影响及因素间的交互作用。这种在全部水平组合中挑选部分水平组合获得的方案称为不完全方案。根据不完全方案进行的试验称为部分试验。动物试验的综合性试验(comprehensive experiment)、正交试验(orthogonal experiment)都属于部分试验。

综合性试验是针对起主导作用且相互关系已基本清楚的因素设置的试验,它的水平组合就是一系列经过实践初步证实的优良水平的配套。正交试验是在全部水平组合中选出有代表性的部分水平组合设置的试验。

一个周密、完善的试验方案不仅可以节省人力、物力,多快好省地完成试验任务,而且可以获得正确的试验结论。如果方案拟定不合理,如因素、水平选择不当,部分试验方案所包含的水平组合针对性或代表性差,试验将得不出应有的结果,甚至导致试验的失败。因此,试验方案的拟定在整个试验中占着极其重要的位置。

2.2.2 拟定试验方案的要点

为了拟定一个正确、切实可行的试验方案,应从以下几方面考虑:

1.明确研究工作的目的,根据试验的目的、任务和条件挑选试验因素

拟订方案时,在正确掌握生产中存在的问题后,对试验目的、任务进行仔细分析,抓住关键,突出重点。首先要挑选对试验指标影响较大的关键因素。且一个试验中研究的因素不宜过多,否则处理数太多,试验过于庞大,试验干扰因素难以控制。因此,凡是能用简单方案的,就不用复杂方案。

2.选择适当的因素和水平

安排一个试验之前,首先要考虑哪些因素会对试验结果产生影响。根据各试验因素的性质分清水平间的差异。

(1)水平的数目要适当。既不能贪大求全,又不能丢失主要影响因素。

(2)水平的确定和因素的类型密切相关。水平间的差异要合理。有些因素在数量等级上只需少量的差异就反映出不同处理的效应。如模拟重金属胁迫中金属离子的添加等。而有些则需较大的差异才能反映出不同处理效应来,如饲料用量等。

(3)试验方案中各因素水平的排列要灵活掌握。一般在文献数据或者预备试验的基础上,可采用等差法(等间距法)、等比法和随机法3种。

3.试验方案中必须设立作为比较标准的对照

试验的目的就是通过比较来鉴别处理效应的大小、影响等。为了达到这一目的,试验方案应当包括各试验处理,以及作为比较的对照。任何试验都不能缺少对照,否则就不能显示出试验的处理效果。根据研究的目的与内容,可选择不同的对照形式。

4.试验处理(包括对照)之间应遵循唯一差异原则

该原则是指在进行处理间比较时,除了试验处理不同外,其他所有条件应当尽量一致或相同,使其具有可比性,才能使处理间的比较结果可靠。

5.有的试验要设置预试期

所谓预试就是正式试验开始之前根据试验设计进行的过渡试验,为正式试验做好准备工作。通过预先种植,使供试的植物适应相同的环境,对不合适的供试植物进行调整和选