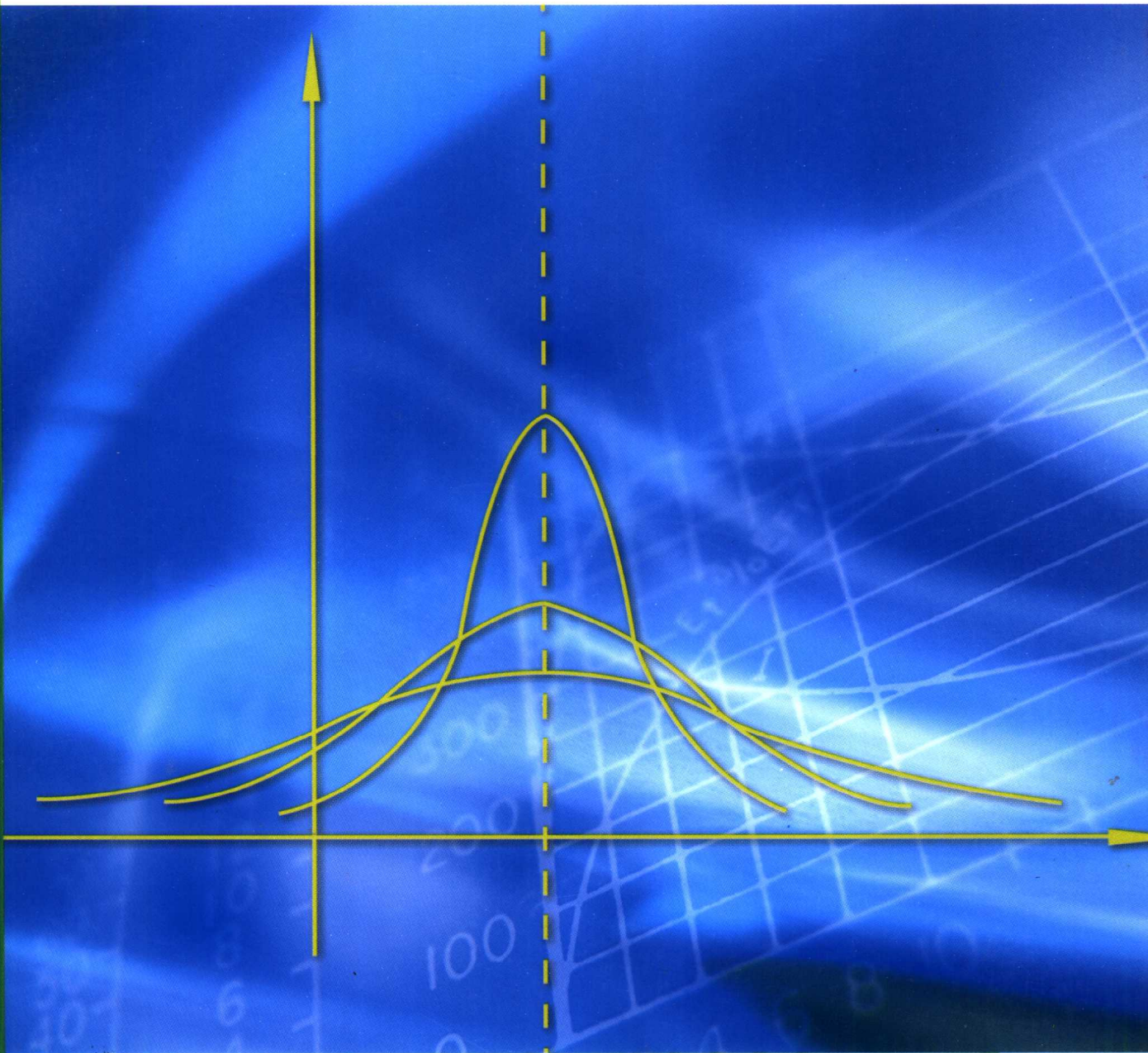


Biostatistics



生物统计学

陆建身 赖麟 主编



高等教育出版社
HIGHER EDUCATION PRESS

生物统计学

Biostatistics

陆建身 赖麟 主编



高等教育出版社

HIGHER EDUCATION PRESS

内容简介

本书包括绪论,资料整理和描述性统计,概率、随机变量及其分布,抽样分布及统计推断原理, t 检验, χ^2 检验,方差分析,直线回归与相关,非参数检验,试验设计,SPSS软件在生物统计学中的应用等内容。

本书可作为专升本教材。

图书在版编目(CIP)数据

生物统计学/陆建身,赖麟主编. —北京:高等教育出版社,2003.7

ISBN 7-04-012197-2

I. 生… II. ①陆…②赖… III. 生物统计-师范大学-教材 IV. Q-332

中国版本图书馆CIP数据核字(2003)第025469号

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	人民教育出版社印刷厂		
开 本	787×1092 1/16	版 次	2003年7月第1版
印 张	14.25	印 次	2003年7月第1次印刷
字 数	340 000	定 价	18.20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。

版权所有 侵权必究

本书编写作者、单位和编写内容分工如下：

陆建身 教授 华东师范大学生命科学学院 编写第1、第2、11章、附录和统稿

赖 麟 副教授 四川教育学院生物系 编写第8、10章和协助统稿

刘来福 教授 北京师范大学数学系 编写第3、4章

杨振才 副教授 河北师范大学生命科学学院 编写第5、7章

张尤凯 副教授 华南师范大学生物系 编写第6、9章

前 言

随着生命科学的发展,作为一门研究生物领域数量现象的科学——生物统计学,已越来越受到广大生物学工作者的重视。近年来,各地高等院校生物系多已将生物统计学列入本科教学计划,而以在职教师为对象的高师“专升本”生物学专业,也已把生物统计学作为专业基础必修课,以适应教师专业知识更新和提高科研能力的迫切需要。因此,编写一本既能满足现有生物学专业本科生教学需要,又适合在职教师学历教育使用的相关教材已十分必要。在教育部师范司的主持下,2000年在河北师范大学召开的“专升本”学科教材建设会议确定了由华东师范大学等五所院校组成《生物统计学》教材编写组。各方经过3年时间的努力,终于将《生物统计学》付诸出版。

根据生物学专业本科生培养目标和课程大纲的要求,本教材的编写在保证科学性、系统性的基础上,力求简明实用,由浅入深,强调针对性,注重实践性,加强自学性,突出应用性。遵循生物统计学的学科特点和教学规律,理论联系实际,同时最大程度地体现时代特征和先进性。

计算机普及和统计软件的问世给生物统计的应用提供了更大的发展空间和无比的优越性,为此,本书专门编写了有关计算机统计的章节,可供读者在掌握相关统计的原理和方法后对照使用。为了便于自学和复习,每章后有总结和习题,书末附有参考答案。

本教材共分为11章,由华东师范大学陆建身教授编写第1、第2、第11章和附录,北京师范大学刘来福教授编写第3、第4章,河北师范大学杨振才副教授编写第5、第7章,华南师范大学张允凯副教授编写第6、第9章,四川教育学院赖麟副教授编写第8、第10章。全书由陆建身教授统稿。

本教材的出版,不单纯是为了解决本课程教材的有无问题,更重要的是在新的历史条件下,为建设具有高校特色的教材探索路子,进行一些有益的尝试。但因条件和时间所限,缺点错误在所难免,恳望各校在使用中提出宝贵意见,以便再版时作进一步修改。

陆建身

2003年6月

责任编辑	张庆波
封面设计	张楠
责任绘图	郝林
版式设计	陆瑞红
责任校对	殷然
责任印制	杨明

郑重声明

高等教育出版社依法对本书享有专有出版权。任何未经许可的复制、销售行为均违反《中华人民共和国著作权法》，其行为人将承担相应的民事责任和行政责任，构成犯罪的，将被依法追究刑事责任。为了维护市场秩序，保护读者的合法权益，避免读者误用盗版书造成不良后果，我社将配合行政执法部门和司法机关对违法犯罪的单位和个人给予严厉打击。社会各界人士如发现上述侵权行为，希望及时举报，本社将奖励举报有功人员。

反盗版举报电话：(010) 82028899 转 6897 (010)82086060

传真：(010) 82086060

E-mail：dd@hep.com.cn

通信地址：北京市西城区德外大街4号

高等教育出版社法律事务部

邮编：100011

购书请拨打读者服务部电话：(010)64054588

目 录

第 1 章 绪论	1	附表 1b 正态分布的密度函数表	194
第 2 章 资料整理和描述性统计	9	附表 2 t 值表(两尾)	196
第 3 章 概率、随机变量及其分布	23	附表 3 χ^2 值表(一尾)	197
第 4 章 抽样分布及统计推断原理	53	附表 4 5%及 1% F 值表(一尾)	198
第 5 章 t 检验	68	附表 5a 5% q 值表(两尾)	204
第 6 章 χ^2 检验	77	附表 5b 1% q 值表(两尾)	205
第 7 章 方差分析	86	附表 6 SSR 表(邓肯氏新复极差值表)	
第 8 章 直线回归与相关	106	206
第 9 章 非参数检验	134	附表 7 积差相关系数 r 临界值表	
第 10 章 试验设计	142	208
第 11 章 SPSS 软件在生物统计学中		附表 8 r 值的 Z_r 转换表	211
的应用	153	附表 9 复相关系数临界值表	212
参考书目	188	附表 10 符号检验表	214
思考练习题参考答案	189	附表 11 游程总数检验表	215
附表 1a 正态分布在 $(-\infty, u)$ 区间		附表 12 秩和检验表	218
的概率表	192		

1

绪 论

生物统计学是应用数理统计的原理和方法来分析和解释生物界数量现象的科学,也可以说是数理统计在生物学研究中的应用,它是应用数学的一个分支,属于生物数学的范畴.生物统计学由于紧密结合生物学科研究和生物学教育,加上近年来计算机统计软件的支持,目前已被广泛地应用于生命科学的各个领域,并获得了迅速的发展.

第一节 生物统计学概述

生物统计学的研究内容包括统计原理、统计方法和试验设计.统计原理阐述统计理论和有关公式,以满足统计方法的需要.统计方法的应用,旨在对客观事物得出本质的和规律性的认识.试验设计是试验工作前应用统计原理,制订科学的试验方案和方法.

人们在从事科学研究时,总是通过事物的一部分(样本),来估计事物全体(总体)的性质特征,即由样本推断总体,从特殊推导一般,从而对所研究的总体得出正确的结论.在生物科学研究中,我们期望知道的是总体,而不是样本.可是在实际问题的调查和试验中,我们所得到的却只是样本资料.生物统计学从本质来看,实际上是研究如何从样本推断总体的一门科学.

生物统计学所研究的对象是有变异的总体,即使在同质的对象中往往也存在差异.例如,同一组小鼠,即使品种、性别、年龄相同的个体,它们的体重、尾长和各种生理指标的数值都会有所不同;同一栏猪,由于吃了腐烂的饲料而发生胃肠炎,病情的轻重也有所不同;对病情相同的病畜,用同一药物治疗,有的治愈,有的没有治愈,治愈的病例其病程也有长有短,不一而论.这些都是普遍存在的变异的实例.如何认识和了解这种有变异的总体,是生物统计学的基本任务之一.

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,这也是生物统计学的重要任务之一.

试验研究工作开展前进行试验设计,制订试验方案,选择试验动物,合理分组,可以利用较少的人力、物力和时间,获得更多更可靠的信息资料,从而得出科学的结论.生物统计与试验设计是

不可分割的两部分. 试验设计需要以统计的原理和方法为基础, 而正确设计的试验又为统计方法提供丰富可靠的信息, 两者紧密结合才能推断出较为客观的结论, 从而不断地推进生物科学研究的发展.

第二节 生物统计的功用

现代生物统计学已在生命科学研究、生产实践和生物教育领域中得到了极为广泛的应用. 其基本功用有:

一、科学地整理分析数据

做任何工作, 都必须掌握基本情况, 做到胸中有数, 才能有的放矢, 从而提高工作质量. 进行生物科学研究工作更不例外, 必须有计划地收集资料并进行合理的统计分析, 通过调查所得的数据, 经过加工整理, 从中归纳出事物的内在规律性, 用以指导实践. 例如, 调查某母鸡品种的产蛋数, 可以得到不同季节、不同年龄、不同个体、不同产蛋数的大量原始数据. 这些杂乱的数据开始时难以看出什么规律性, 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, 即可大体了解不同季节、不同年龄、不同母鸡产蛋数的一般情况及其变异特征, 得到母鸡在什么季节产蛋数最高, 什么年龄产蛋数开始下降, 以及母鸡适宜的利用年限等很多有用的信息.

二、判断试验结果的可靠性有效性

由于存在试验误差, 从试验所得的数据资料必须借助于统计分析方法才能获得可靠的结论. 例如, 某试验农场要研究两种饲料对肉用仔鸡增重及饲料报酬的影响. 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 周龄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, 以决定两群鸡体重与饲料消耗的差异, 究竟是属本质的, 还是属机遇的, 即判断是由于不同饲料造成的, 还是由于其他未经控制的偶然因素所引起的, 统计分析之后才能得出比较正确的结论.

三、确定事物之间的相互关系

科学实验的目的, 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的联系形式. 例如, 测定某群奶牛第一胎的产乳量和它以后几胎的产乳量之间的相关关系, 就可以根据第一胎产乳量的高低来推断它终生的产乳量, 这样, 就为早期选择和淘汰低产乳牛提供科学预见. 又如青少年近视的形成受很多因素的影响, 查明这些影响因素对降低青少年近视的发病率有很积极的意义. 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.

四、提供试验设计的原则

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,最大限度地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和其他方面的损失.

五、为学习相关学科提供基础

要学好遗传学、生态学和营养学等学科,也必须学好生物统计.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.如果不懂得生物统计,也就无法掌握遗传学.此外,阅读生物科技文献也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,生物科学工作者都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.

总之,生物统计是一种很有用的工具,正确使用这一工具可以使生物科学研究更加有效,使生产效益更高,使生物教育效果更好.所以,它是每位从事生物科学研究和生物学教育的工作者必须掌握的基本工具.

第三节 学习生物统计的方法和要求

生物统计学是数学与生物学相结合的一门边缘学科,与生物学的其他学科具有很大的不同,它所包含的公式很多,在性质上属于生物学领域内的应用数学.因此,在学习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,但一般可不必深究其数学推导、证明和数学原理.

其次,作为一门工具课,必须认真做好习题作业,加深对公式及统计步骤的理解.只有通过一定数量的实践和练习(没有其他途径可走),并能熟练使用函数型计算器操作,并在此基础上学会用计算机处理,才能达到熟练应用生物统计方法的目的.

其三,应注意培养科学的统计思维方法.生物统计意味着一种全新的思考方法,从不肯定性或概率的角度来思考问题和分析科学试验的结果,避免绝对肯定或绝对否定的武断结论,或单凭感觉不作检验的简单判断.

其四,必须联系实际,结合专业,了解生物统计方法的实际应用.平日要留意国内外书籍和杂志文献中的表格、数据及其分析和解释,熟悉规范的表达方法及其应用.

第四节 生物统计学的主要内容

生物统计学从统计方法的主要内容来讲,大致可分为以下三个方面.

一、描述性统计

描述性统计实际上就是对原始资料进行整理并作基本分析.

生物统计的基本特点,是以样本推断总体.如果在大样本情况下,就需要通过统计资料初步整理之后,再从资料中计算出三个主要的统计量,即平均数、标准差及标准误.以用来根据样本推断总体的特征,一是资料的集中性,以平均数来表示,二是资料的离中性,以标准差来表示,三是衡量平均数的可靠性——用标准误来表示.

二、显著性检验(又称假设检验)

1. 平均数间差异的比较

在进行生物科学研究工作中,经常会遇到两组或两组以上数据的平均数.它们之间进行比较,就有一个有无显著差异的问题或者说孰优孰劣的问题.平均数之间的比较,需要通过一定的统计方法,并且要注意抽样是否合理,否则它们将没有比较的基础.例如,有两批孵化期不同的两个月龄雏鸡的体重,甲组(3月8日出壳)共61只,平均体重为269.3g,乙组(4月15日出壳)共73只,平均体重为275.6g,试问这两批不同孵化时间,雏鸡平均体重之间的差异是由于不同孵化时期所造成,还是由于其他偶然因素所引起?这就需要应用均数差异显著性检验的统计分析方法,才能作出较可靠的判断,不致被某些偶然性因素所蒙蔽.

2. 属性资料的检验

生物学领域中有许多性状不能直接用测量的方法来加以衡量,一般称之为属性性状.例如,花的颜色、性别中的雌和雄,以及药物试验的治愈或无效等等,均可以应用属性统计的方法,通过对具有相同属性的计数来分析. χ^2 检验是最常用的属性资料的检验方法.

3. 方差分析

方差分析又名变异量分析.目的主要是为了进行多个平均数间的比较.它的原理是应用数学方法,把试验中总变异剖分为由不同变异原因所形成的各种变异并进行显著性检验与多重比较.

三、相关与回归

研究变量之间相互关系的密切程度,称为相关,以相关系数来表示.例如,人的身高与体重存在着一定程度的相关.一般身高越高,其体重可能越大.相关系数可用来表示两者间的相关程度.

回归是指两个或两个以上的变量存在着从属关系,即一个变量(x)变化时,引起另一变量(y)的相应变化的估计.它们的从属关系可以用回归分析的方法进行研究,根据实际数据建立的

关系式称为回归方程式,用以对某些指标进行预测和预报.例如,用胸围、体长来估计体重,或根据亲代生产力预测、预报后代的育种值等.

四、试验设计

本课程的另踴任务是讨论试验设计的原理和方法.所谓试验设计,主要是指如何选择试验动物,进行合理的分组和安排实验等,其目的是为了尽量减少和控制试验误差,并对试验误差作出无偏的估计.主要的设计方法有:① 完全随机设计;② 配对设计;③ 随机单位组设计;④ 拉丁方设计;⑤ 正交设计等.为了使试验结果成为有用而可靠的科学资料,在开始试验之前,认真地进行试验设计是非常必要的.

第五节 生物统计学的发展概况

生物统计学是一门比较年轻的学科.1870年,英国遗传学家高尔顿(F. Galton, 1822—1911)通过研究人类身高的遗传,认为子女的身高与父母的身高有着直接关系,发现子女的身高与他们父母的平均身高有回归的趋向.1889年,他发表了回归分析方法在遗传学上应用的论文.这就是在数理统计中,回归这个术语的由来,因而后人推崇他为生物统计学的创始人.

正态分布对研究生物统计的理论具有十分重要的意义,它早在1733年就被棣莫弗(De-Moivre)发现,而被后来的高斯(Gauss)所完成,因此,有人称它为高斯分布.

皮尔逊(K. Pearson, 1857—1936)是一位数学物理学家,他为将数学应用于生物学作了半个多世纪的努力,并创立了《生物统计学报》(Biometrics),对促进生物统计学发展作出了重要贡献.1899年他提出了一个测量实际数与预计数(或理论数)之间的偏离度的指数卡方(χ^2),在属性统计分析上有着广泛的应用.这种统计方法被同时代的孟德尔所掌握,成功地应用于豌豆杂交遗传学试验中,高豌豆品种与矮豌豆品种杂交之后,子一代均为高豌豆,高株性状属于显性,而矮株属于隐性,当子一代自交时,它的后代数预计有高3:矮1,但实际后代数,是否符合3:1规律,就需用 χ^2 进行检验.这种方法在遗传学科研方面至今还在广泛应用着.

戈塞特(W. S. Gosset, 1876—1937)是K·皮尔逊的学生,他对样本标准差等分布作了不少研究工作,并且于1908年用“学生氏”的笔名将《*t*检验》发表于《生物统计学报》.*t*检验已成为当代生物统计工作的基本工具之一.

1923年英国的费希尔(R. A. Fisher)第一个把变异来源不同的方差(S^2)比值称为*F*值,当*F*值大于理论上5%概率水准的*F*值时,该项变异来源的必然性效应就从偶然性变量中分析了出来.这个分析方法,被称为方差分析.在生物统计中,方差分析应用相当广泛,特别是在他发表了《试验研究工作中的统计方法》的专著后,对推动和促进农业科学、生物学和遗传学的研究和发 展,起了一定的奠基作用.

奈曼(J. Neyman)和E·S·皮尔逊分别在1936年和1938年提出统计假设检验学说,对促进理论研究以及对试验研究作出结论具有很大的实用价值.

试验设计是在20世纪初提出来的一种科研思路,30年代应用于农业科学实验,并使“试验

技术”成为一门专门学科,以后扩大应用于生物学、医学和工业领域.70年代提出的多因素试验的正交试验法,在工农业的试验研究中,也相继获得了可喜的成果.近年来,回归分析方法在工农业生产和科学实验中大量应用,如在进行数据处理时,寻求经验公式,探索新工艺、新配方,以及某些性状的预测预报,都积累了不少新经验,获得了许多新成果.特别是正交试验、回归试验设计是一个新的研究方向.计算机的广泛使用,使运算技术出现了一次革命,尤其是国际上出现了SAS、SPSS等大型统计软件以后,生物统计变得日益精确和迅速,从而进一步推动了生命科学研究向纵深发展.

第六节 常用统计术语

一、总体与样本

根据研究目的确定的、符合指定条件的全部观察对象称为总体.构成总体的每一个单元,称为该总体的个体.

例如,研究上海市16岁男中学生的身高,那么,凡是上海市16岁的男中学生,他们的身高构成一个总体,而每一个身高测量值则是一个个体.总体可以分为有限总体与无限总体两种.个体数有限的总体称为有限总体,个体数无限的总体称为无限总体.

从总体中抽取的一部分个体则构成样本.样本中所含个体的数目叫做样本容量.常以 n 表示.由于样本容量不同,一般又分大样本(例如样本容量 $n > 30$)和小样本(样本容量 $n < 30$).样本是总体的缩影,因此它应该能反映总体的特征和特性,但它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的目的.

例如,用某种新药治疗感冒100例,有效80例,即有效率80%.但我们所观察到的只是含量为100的一个样本,如果用这种新药治疗其他感冒患者,是否也会获得同样的效果呢?这就需要以样本的疗效来估测总体的疗效.如何正确地从样本来推断总体,这是统计所要解决的问题.

二、随机抽样与随机样本

所谓随机抽样是指抽样时,不掺杂人们的主观愿望,总体中每个个体被抽取的机会均等.由随机抽样而得的样本,称随机样本.一个总体所含的个体数目往往很多,甚至无穷,以致不可能一一加以考察.例如,我们研究羊毛细度这个总体,它的个体数目多得数不清,无法一一加以测定.有的时候,数据的测定是破坏性的,如研究猪的屠宰率、瘦肉率,测定一次就要杀掉一头.因此,即使个体的数目不多的有限总体,也不允许全都加以考察,我们只能通过样本来了解总体.同样的,在动物实验中,把动物分为几组,也必须用随机方法使每个动物有同等的机会被分配到各组去,这样就不至于人为地造成各组间动物的不相同.随机抽样的方法很多,其中最常见的是抽签和采用随机数字表.

三、变数与变异数列

统计学上把每一个观察值称为变数.把所有变数整理成为一个由小到大的序列,称为变异数列.常数表示能代表事物特征和性质的数值,通常由变数计算而来,在一定过程中是不变的,如某样本平均数、标准差、变异系数等.

四、参数与统计量

从总体中计算所得的特征数值,如总体平均数、总体标准差称为参数;而从样本中计算所得的特征数值称为统计量,它是总体参数的估计值.

人们在从事生物科学研究时,往往是通过某事物的一部分(样本),来估计事物全体(总体)的特征的,目的是为了由样本推断总体,从特殊推导一般,对所研究的总体,作出合乎逻辑的推论,得到对客观事物本质的和规律性的认识.实际研究过程由于资料庞大,常常不可能全部观察记载,而只能采用抽样的方法,计算样本的统计量.用样本的统计量,来估计总体的参数.

五、机误与错误

由样本推断总体,往往不可避免地会产生一部分误差,这一部分误差是由各种无法控制的随机因素所引起,我们把它叫做机误或试验误差.机误是试验中由于无法控制的随机因素所引起的差异,它是不可避免的,试验中只能设法减小,而不能完全消灭.增加抽样或试验次数,可以降低机误的数值.

错误是指在试验过程中,人为的作用所引起的差错.如试验人员粗心大意,使仪器校正不准、药品配制比例不当、秤量不准确、将数据抄错、计算出现错误等都是由于人为因素造成的,在试验中是完全可以避免的.

六、效应与互作

引起试验差异的作用称为效应,如不同饲料使动物的体重增加表现出差异,不同品种的玉米产量不同等.互作,也称连应,是指两个或两个以上处理因素间的相互作用产生的效应.如氮、磷肥并施会对作物产量产生互作效应.互作有正效应,也有负效应,如果氮、磷共施的产量效应大于氮、磷单施效应之和,说明氮磷互作为正效应,如果氮、磷共施的产量效应小于氮、磷单施效应之和,说明氮磷互作为负效应.

七、准确性与精确性

统计工作是用样本的统计数来推断总体的参数.我们用统计数接近参数真值的程度,来衡量统计数“准确性”的高低.用样本中各个变数间变异程度的大小,来衡量该样本“精确性”的高低.

因此,准确性就不等于精确性.准确性是说明测定值对真值的符合程度大小,而精确性却是多次测定值的变异程度.

生物统计学是建立在生物学和统计学(或数学)两个基础之上的学科.如果生物学本身的理论建立在不充分的基础上,即使再准确地计算也毫无意义;反之,正确的理论也会由于不精确的计算而导致错误的结论.可靠的判断方法是通过实践来检验的.因此,在科学研究中,在做出结论之后,也还必须再回到实践中加以验证.

本章小结

生物统计学是应用数理统计的原理和方法来分析和解释生物界数量现象的科学.生物统计学的研究内容包括统计原理、统计方法和试验设计.统计方法的主要内容可分为描述性统计、显著性检验(又称假设检验)、相关与回归三个方面.

生物统计学基本功用包括科学地整理分析数据、判断试验结果的可靠性、确定事物之间的相互关系、提供试验设计的原则,为学习相关学科提供基础.

生物统计学是一门比较年轻的学科,其产生发展不到 150 年的历史.近年来计算机统计软件的应用,推动了生物统计学的进一步发展.

思考练习题

1. 生物统计学的主要内容和作用是什么?
2. 举例说明什么叫总体.什么叫样本.
3. 对下列问题进行判断,哪些属于有限总体?哪些属于无限总体?
 - (1) 一批鱼苗的体长;
 - (2) 某地区的地下水位;
 - (3) 某年某地区棉铃虫的数量;
 - (4) 某林区的林木蓄积量;
 - (5) 全国所有自然保护区物种的数量.
4. 什么是参数?什么是统计量?二者有何联系和区别?
5. 什么是机误?什么是错误?如何避免试验错误的发生?
6. 练习计算机的使用方法.

2

资料整理和描述性统计

从生物科学研究或教学试验中取得的大量原始数据,是在某种具体条件下对某种具体现象的观察结果,但这些数据是一些零星的、孤立的现象记录,如要从其中揭露出它们的内部联系和规律性,必须先通过数据整理和基本分析.

第一节 资料的分类

资料的分类是统计归纳的基础,若不进行分类,大量的原始资料就不能系统化、规格化,只有根据科学原理来分类,才能使资料正确地反映出事物的本质和规律.

在调查或试验中,所观察、测量的数据必须加以记载.记载的方式,按观察或测量对象的性质,可分为数量性状资料与质量性状资料.

一、数量性状资料

数量性状资料简称数性资料,一般包括计量资料和计数资料两类.

1. 计量资料

计量资料是通过直接计量而得来的,即能用度量衡等计量工具直接测定的数性资料.其数据是用长度、容积、重量等来表示(如苗高、产奶量、体重、亩产量等).所测定的各个变数不一定是整数,在整数之间可以出现任何值.如测定奶牛的产奶量在 3 000 kg 至 6 000 kg 之间,可以有很多个变数,这些变数可以是整数,也可以是小数.在统计学中把这种资料称为计量资料,也称连续性变数资料.

2. 计数资料

计数资料是指用计数方式而得来的数性资料.在这类资料中,每一个变数必须以整数来表示,两整数间的数值是不连续的,因此不具有小数.计数方式表示的数据是以 1 为单位.如动物个体数、植株数、种子数、细菌数、产仔数、成活数、鸡的产蛋数等,这些变数都是以个数来表示,各变