



普通高等教育“十一五”国家级规划教材



普通高等教育农业部“十二五”规划教材



普通高等教育“十三五”规划教材

生物统计学

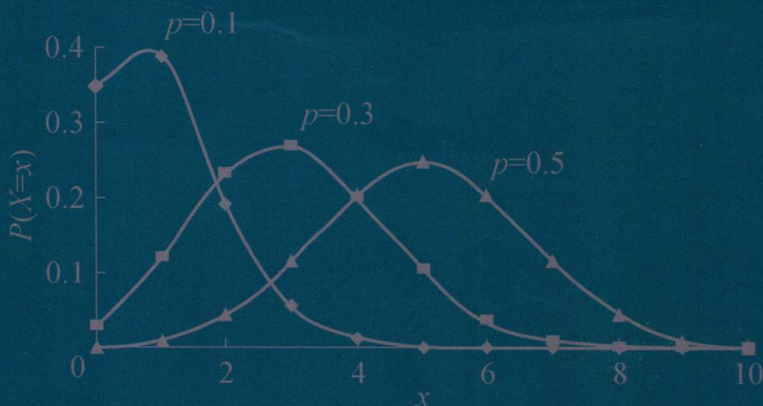
Biostatistics

第3版

动物科学、动物医学专业适用

张勤 主编

王雅春 徐宁迎 贾青 副主编



中国农业大学出版社

CHINA AGRICULTURAL UNIVERSITY PRESS



普通高等教育“十一五”国家级规划教材

内容简介



普通高等教育农业部“十二五”规划教材



普通高等教育“十三五”规划教材

生物统计学

(第3版)

(动物科学、动物医学专业适用)

张勤 主编

王雅春 徐宁迎 贾青 副主编

潘增祥(南京)

王志鹏(东北)

肖天放(福建)

余梅(华中)

王雅春(安徽)

戚慧玲(河北)

责任编辑 100132
发行部电话 010-84392728
印刷部电话 010-84392710
E-mail: cbszq@caup.cn

生物统计学
张 勤 主 编
王雅春 副主编
徐宁迎 副主编
贾 青 副主编
中国农业出版社
北京市东城区
安定门内大街
28号
100029
电话 010-84392728
http://www.cup.com.cn

中国农业大学出版社

· 北京 ·

内 容 简 介

本书重点介绍在动物科学和动物医学学科领域中常用的生物统计学的基本概念和基本方法。全书共分 15 章,主要包括生物统计学的基本概念和基本特点,数据异常值检验及描述性统计分析,随机变量与概率分布基础知识,假设检验的基本原理、基本步骤以及相关的概念, Z 检验、 t 检验、 F 检验、卡方检验等常用的假设检验方法,以及试验设计和抽样调查设计的基本概念及原理。在简单介绍分析原理的同时提供了大量与动物科学和动物医学相关的例题辅助理解。在附录中介绍了利用 Excel、SPSS 和 R 语言进行常见统计分析的步骤。每一章都给出了内容概要及必要的思考题,帮助学生更好地掌握该章的重点和难点。

本书不仅可以作为高等院校动物科学和动物医学等专业本科学生的教材,也可供动物科学和动物医学专业相关科研技术人员参考。

学 刊 农 业 大 学

(第 8 卷)

图书在版编目(CIP)数据

生物统计学/张勤主编.—3 版.—北京:中国农业大学出版社,2018.9

ISBN 978-7-5655-2042-6

I. ①生… II. ①张… III. ①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(2018)第 121451 号

书 名 生物统计学(第 3 版)

作 者 张 勤 主 编

策划编辑 宋俊果 荀红晓

责任编辑 郑万萍

封面设计 郑 川

出版发行 中国农业大学出版社

社 址 北京市海淀区圆明园西路 2 号

邮政编码 100193

电 话 发行部 010-62818525,8625

读者服务部 010-62732336

编辑部 010-62732617,2618

出 版 部 010-62733440

网 址 <http://www.caupress.cn>

E-mail cbsszs@cau.edu.cn

经 销 新华书店

印 刷 北京时代华都印刷有限公司

版 次 2018 年 9 月第 3 版 2018 年 9 月第 1 次印刷

规 格 787×1 092 16 开本 22 印张 540 千字

定 价 52.00 元

图书如有质量问题本社发行部负责调换

第3版编写人员

主 编：张 勤(中国农业大学) 兼：兼 主

副主编：王雅春(中国农业大学) 兼：兼 主 编

徐宁迎(浙江大学)

贾 青(河北农业大学)

参 编：(以姓氏拼音为序)

陈 静(沈阳农业大学)

戴国俊(扬州大学)

高会江(中国农业科学院北京畜牧兽医研究所)

姜 力(中国农业大学)

刘剑锋(中国农业大学)

刘小林(西北农林科技大学)

潘增祥(南京农业大学)

王志鹏(东北农业大学)

肖天放(福建农林大学)

余 梅(华中农业大学)

殷宗俊(安徽农业大学)

郑慧玲(西北农林科技大学)

第2版编写人员

主 编：张 勤(中国农业大学) 兼 主 审

副主编：徐宁迎(浙江大学) 曹 王 兼 主 审

贾 青(河北农业大学) 宁 霜

王雅春(中国农业大学) 贾

参 编：(以姓氏拼音为序) 陈 为 斌 等

戴国俊(扬州大学) 魏 翔

高会江(东北农业大学) 田 颖

刘小林(西北农林科技大学)

肖天放(福建农林大学) 姜

谢 庄(南京农业大学) 魏 欣

余 梅(华中农业大学) 小 政

殷宗俊(安徽农业大学) 谢 磊

(学大业农农农) 柳志王

(学大林农农农) 郑天肖

(学大业农中平) 薛 余

(学大业农康安) 魏宗魏

(学大林林林农北西) 刘慧联

书 名 生物统计学(第2版)

作 者 张 勤 主编

责任编辑 宋俊生 张有前

封面设计 邱 川

出版发行 中国农业大学出版社 (学大农林林农北西)

社 址 北京市海淀区圆明园西路2号

电 话 发行部(010)62217211, 62217213

编辑部(010)62217214, 62217215

网 址 <http://www.cupchina.com>

经 销 新华书店

印 刷 北京时代华都印刷有限公司

印 次 2018年5月第1次 2018年5月第1次印刷

开 本 787X1 092 1/16 印张 22 印定 40 元

定 价 58.00 元

第3版前言

《生物统计学》于2002年出版,2007年修订出版第2版。出版发行以来,教材被全国范围内数十所院校相关专业采用,赢得了使用院校师生的普遍肯定,反响良好。教材先后被教育部审批为“面向21世纪课程教材”“普通高等教育‘十一五’国家级规划教材”,被农业部审批为“普通高等教育农业部‘十二五’规划教材”。

为使教材适应新形势下的教学要求,做到教材内容与时俱进,我们于2016年9月组成新的编写组,对教材进行修订改版。第3版编写组在第2版人员的基础上增加了6位新成员,在广泛征求第2版教材使用者的意见和建议基础上,对整体结构、内容、细节等进行了研讨,形成了修改意见和编写分工安排。

第3版在保留第2版基本框架结构的基础上,在内容方面做的主要调整包括:

1)每章之前增加了“本章概要”,用简洁的文字梳理一章的梗概,帮助读者了解本章的重点内容。

2)由于近年来R语言统计软件在各个领域(尤其是生物领域)中的应用越来越广泛,显示出强大的生命力和广阔的应用前景,为适应这个发展趋势,我们增加了附录3——R语言统计软件简介,供生物统计实验课选择使用和读者参考。

3)对全书进行了全面细致的审核,改正了原书中存在的内容、文字、数字、计算等方面的错误,使之更加准确规范。

第3版的参编者是来自12所高等农业院校和科研单位的16位老师,他们都是多年从事生物统计学教学、活跃在教学科研第一线的教师,具有较扎实的生物统计学理论基础和丰富的生物统计学教学以及对实际资料统计分析的经验,这为保证本书的质量奠定了基础。尽管如此,由于水平有限,错误和不当之处仍在所难免,恳请使用本教材的教师和同学们以及其他读者提出批评指正。

我们衷心感谢在本书编写过程中给我们提出宝贵意见和建议的所有老师和学生,感谢中国农业大学出版社对本书的出版给予的大力协助。

编者

2018年4月

第2版前言

《生物统计学》第1版于2002年由中国农业大学出版社出版后,被全国十余所高等农业院校选用作为动物科学和动物医学专业本科生的生物统计学课程教材,并得到了普遍肯定。由于需求量较大,在短短的3年内,已经印刷了3次。2006年该教材被教育部审批为普通高等教育“十一五”国家级规划教材,并继续由中国农业大学出版社组织出版该书的第2版。为使第2版能在第1版的基础上进一步提高质量,使之更加适合我国动物科学和动物医学专业的师生、科研工作者和实际工作者对生物统计学的需要,在编写过程中,我们广泛征求了第1版使用者对本书的修改意见和建议,并于2007年4月在中国农业大学召开了编写组成员会议,对第1版的整体结构、内容的增减、细节的修改等进行了充分讨论,形成了详细的修改意见,并进行了编写分工。

第2版的基本结构与第1版相同,在内容上的主要调整如下:

1. 增加了“非参数检验”的内容,作为第2版的第14章。
2. 第1版的第2章“数据整理”更名为“资料的描述性统计分析”,并增加了“异常数据的判断和处理”的内容(见第2章2.1)。
3. 将第1版第7章中7.3“双向嵌套分组资料的方差分析”抽出作为第2版的第8章,并更名为“方差分析Ⅲ——两因子嵌套分组资料”,增加了对方差组分估计基本概念的介绍(见第8章8.5)。
4. 第1版的第8章“数据转换”并入第6章“单向分类资料的方差分析”,作为第6章的6.7节,内容上作了适当精简。
5. 在第10章“多元线性回归与相关”中增加了对最优回归方程的建立的基本思路的介绍(见第10章10.1.5)。
6. 在第11章“非线性回归”中增加了“Logistic生长曲线模型”的内容(见第11章11.2.6)。
7. 第1版中第14章“试验设计简介”改为第2版中的第15章,更名为“试验设计与抽样调查”,增加了反转设计(见第15章15.2.5)、重复测量设计(见第15章15.2.6)和析因设计(见第15章15.2.7)3种试验设计方法。

此外,对第1版全书中的文字表述进行了全面的审核和修饰,使之更加准确规范,并改正了原书中存在的文字、数字、计算等方面的错误。

第2版的参编者是来自10所高等农业院校的11位老师,他们都是多年从事生物统计学教学的活跃在教学科研第一线的教师,具有较扎实的生物统计学理论基础和丰富的生物统计学教学以及对实际资料统计分析的经验,这为保证本书的质量奠定了基础。尽管如此,由于我

们的水平有限,错误和不当之处仍在所难免,恳请使用本教材的教师和同学们以及其他读者提出批评指正。

原中国农业大学教授、我国著名动物遗传学家吴仲贤先生是我国畜牧生物统计的先驱者之一,他在生物统计学方面的造诣是我等晚辈迄今望尘莫及的。吴仲贤先生于2007年8月3日因病医治无效不幸去世,他生前对本书的编写给予了高度关注,值本书问世之际,我们对他表示崇高的敬意和深切的怀念。

衷心感谢在本书编写过程中给我们提出宝贵意见和建议的所有老师和学生,感谢中国农业大学出版社为本书出版给予了大力协助。

编者

2007年9月

目 录

第 1 章 绪论	1
1.1 什么是生物统计学	2
1.2 统计学的基本特点	2
1.3 为什么要学习生物统计学	2
1.4 常用术语	3
1.4.1 变量与观测值	3
1.4.2 总体和样本	3
1.4.3 参数与统计量	4
1.4.4 准确性和精确性	4
习题	6
第 2 章 资料的描述性统计分析	7
2.1 资料的分类	8
2.1.1 连续性资料	8
2.1.2 离散性资料	8
2.1.3 定量资料	9
2.1.4 定性资料	9
2.2 异常数据的判断和处理	9
2.2.1 四分位数检验	9
2.2.2 格拉布斯检验	9
2.3 数据的频数(率)分布	10
2.3.1 离散性资料的频数(率)分布	11
2.3.2 连续性资料的频数(率)分布	11
2.4 统计表	13
2.5 统计图	14
2.5.1 常用的统计图类型	14
2.5.2 统计图绘制基本要求	16
2.6 集中趋势的度量	16
2.6.1 算术平均数	16
2.6.2 几何平均数	17
2.6.3 调和平均数	18

2.6.4	中位数	19
2.6.5	众数	20
2.6.6	各个集中趋势度量指标之间的关系及评价	20
2.7	离散趋势的度量	21
2.7.1	方差和标准差	21
2.7.2	变异系数	22
2.7.3	范围	23
2.7.4	平均绝对离差	24
	习题	24
第3章	随机变量与概率分布	25
3.1	随机变量及其分类	26
3.2	概率分布	26
3.2.1	离散性随机变量的概率分布	26
3.2.2	连续性随机变量的概率分布	28
3.3	正态分布	29
3.3.1	正态分布的概率密度函数	29
3.3.2	标准正态分布	30
3.3.3	正态分布的概率计算	30
3.3.4	对偏离正态分布的度量——偏度和峰度	32
3.4	二项分布	33
3.5	泊松分布	34
	习题	35
第4章	统计推断概述	37
4.1	抽样分布	38
4.1.1	χ^2 分布、 t 分布和 F 分布	38
4.1.2	正态总体样本平均数的抽样分布	41
4.1.3	中心极限定理	42
4.1.4	正态总体样本方差的抽样分布	42
4.2	参数估计	42
4.2.1	点估计	42
4.2.2	区间估计	44
4.3	假设检验	46
4.3.1	假设检验的基本原理	47
4.3.2	假设检验的基本步骤	48
4.3.3	一些相关概念	49
4.3.4	几点说明	53
	习题	55

第 5 章 对单个和两个总体平均数的假设检验	56
5.1 对单个总体均数的检验	57
5.1.1 总体的方差 σ^2 已知时的检验——Z 检验	57
5.1.2 总体的方差 σ^2 未知时的检验——t 检验	57
5.2 对两个总体平均数的比较	58
5.2.1 随机分组资料的假设检验	58
5.2.2 配对资料的假设测验——t 测验	64
习题	66
第 6 章 方差分析 I——单向分类资料	68
6.1 单向分类资料的数据结构	69
6.2 数学模型	70
6.3 变异的分解	70
6.3.1 平方和的剖分	70
6.3.2 自由度的剖分	71
6.3.3 均方及均方的期望	72
6.4 假设检验	72
6.5 t 检验和 F 检验的关系	74
6.6 多重比较	75
6.6.1 LSD 法	75
6.6.2 Bonferroni t 检验	77
6.6.3 Duncan's 多重极差检验	78
6.7 方差分析的基本假定与数据转换	80
6.7.1 方差的同质性检验	80
6.7.2 方差稳定性转换	82
习题	84
第 7 章 方差分析 II——双向交叉分组资料	85
7.1 交叉分组无重复资料的方差分析	86
7.1.1 数据结构	86
7.1.2 数学模型	86
7.1.3 平方和与自由度的剖分	87
7.1.4 假设检验	88
7.1.5 多重比较	88
7.2 交叉分组等重复资料的方差分析	90
7.2.1 数据结构	90
7.2.2 主效应与交互效应	91
7.2.3 数学模型	92
7.2.4 平方和的剖分与自由度的剖分	92
7.2.5 假设检验	93

7.2.6 多重比较	93
习题	95
第8章 方差分析 III——两因子嵌套分组资料	97
8.1 数据结构	98
8.2 数学模型	98
8.3 平方和与自由度的剖分	99
8.4 假设检验	100
8.5 方差组分估计	104
习题	106
第9章 简单相关与回归	108
9.1 简单相关	109
9.1.1 样本相关系数的定义	109
9.1.2 样本相关系数的计算	111
9.1.3 相关关系的显著性检验	111
9.1.4 总体相关系数的置信区间	113
9.1.5 比较2个相关系数	113
9.2 简单回归	115
9.2.1 一元线性回归的数学模型	116
9.2.2 一元线性回归方程的建立	116
9.2.3 回归的显著性检验	117
9.2.4 对截距的检验	119
9.2.5 总体回归系数的置信区间	119
9.2.6 决定系数	120
9.2.7 两条回归直线的比较	121
9.2.8 利用回归方程进行估计和预测	123
9.2.9 制定校正系数	125
9.3 简单线性相关与回归的区别与联系	126
9.3.1 区别	126
9.3.2 联系	127
9.4 进行相关和回归分析应注意的问题	127
习题	128
第10章 多元线性回归与相关	130
10.1 多元线性回归	131
10.1.1 多元线性回归的数学模型	131
10.1.2 多元线性回归方程的建立	131
10.1.3 多元线性回归的显著性检验	135
10.1.4 偏回归系数的置信区间	137
10.1.5 决定系数	137

10.1.6	标准化的偏回归系数	138
10.1.7	最优回归方程的建立	138
10.2	复相关与偏相关	139
10.2.1	复相关	139
10.2.2	偏相关	141
	习题	143
第 11 章	非线性回归	145
11.1	非线性回归概述	146
11.2	曲线回归的线性化及其方法	146
11.3	未知曲线类型的回归分析——多项式回归	155
11.4	曲线配合的拟合度	157
	习题	159
第 12 章	协方差分析	160
12.1	协方差分析的模型和假定	161
12.2	单向分类资料的协方差分析	162
12.3	双向交叉分组资料的协方差分析	166
	习题	168
第 13 章	分类资料的假设检验	170
13.1	率的假设检验	171
13.1.1	对单个率的检验	171
13.1.2	率的区间估计	172
13.1.3	两个率的比较	173
13.2	卡方适合性检验	175
13.2.1	卡方检验的检验统计量	175
13.2.2	对不同分布类型的适合性检验	176
13.2.3	对分布类型的检验	177
13.3	卡方独立性检验	180
13.3.1	列联表	180
13.3.2	独立性检验的方法	181
13.3.3	2×2 表的精确概率检验法	183
13.3.4	配对 2×2 表的独立性检验	185
13.4	卡方检验的分解	186
13.5	Logistic 回归模型	187
13.5.1	二分类依变量的 Logistic 回归模型	188
13.5.2	Logistic 回归模型的参数估计	188
13.5.3	Logistic 回归模型参数的统计检验	189
	习题	190

第 14 章 非参数检验	193
14.1 非参数检验的意义	194
14.2 符号检验	194
14.3 符号秩和检验	196
14.4 二组非配对资料的秩和检验	197
14.5 多组资料的秩和检验	198
14.6 秩相关	200
14.7 Ridit 分析	202
习题	205
第 15 章 试验设计与抽样调查	207
15.1 试验设计概述	208
15.1.1 试验设计的有关概念	208
15.1.2 试验设计三原则	209
15.2 常用动物试验设计方法	210
15.2.1 完全随机设计	210
15.2.2 配对设计	211
15.2.3 随机区组设计	212
15.2.4 双向随机区组设计——拉丁方设计	214
15.2.5 交叉设计	218
15.2.6 重复测量设计	220
15.2.7 析因设计	222
15.2.8 正交试验设计	222
15.3 抽样调查设计	228
15.3.1 概率抽样	228
15.3.2 非概率抽样	229
15.4 样本含量的确定	230
15.4.1 抽样调查中样本含量的确定	230
15.4.2 试验设计中样本含量的估计	231
习题	234
参考文献	236
附录 1 Excel 电子表格统计功能简介	237
1 概述	238
2 常用概率分布的计算	242
3 数据整理	245
4 总体均数的假设检验	247
5 方差分析	249
6 回归与相关分析	251

附录 2 SPSS 统计软件简介	256
1 中文操作界面设置	257
2 数据管理及数据资料的基本统计分析	258
3 均数差异显著性检验	261
4 方差分析与协方差分析	262
5 回归与相关分析	265
6 卡方适合性检验与独立性检验	266
7 非参数检验	267
附录 3 R 语言统计软件简介	269
1 概述与安装使用	270
2 资料的描述性统计分析	277
3 常用概率分布的计算	279
4 总体均数的假设检验	280
5 方差分析与协方差分析	283
6 回归与相关分析	286
7 分类资料的检验	292
8 非参数检验	293
附录 4 常用统计用表	295
附表 1 标准正态分布的累积分布函数表	296
附表 2 标准正态分布的双侧分位数表	298
附表 3 χ^2 分布的上侧分位数表	299
附表 4 t 分布的双侧分位数表	301
附表 5 F 分布的上侧分位数表	303
附表 6 Duncan's 多重极差检验的 5% 和 1% SSR 值表	308
附表 7 Hertley 方差同质性检验临界值表	310
附表 8 Cochran 方差同质性检验临界值表	312
附表 9 相关系数检验的 5% 和 1% 临界值表	314
附表 10 Spearman 秩相关系数检验临界值表	316
附表 11 符号检验表	317
附表 12 符号秩和检验表(双尾)	317
附表 13 非配对资料秩和检验表	318
附表 14 1 万个随机数表	319
附表 15 常用正交表	327

生物统计学

第1章

绪论

本章概要

本章介绍了生物统计学的基本概念和基本特点,为什么要学习生物统计学,以及统计学中的一些常用术语。

生物统计学

总论和样本

生物统计学是研究生物群体中个体差异的规律性的一门科学。它主要研究生物群体中个体差异的规律性,以及如何利用这些规律性来推断生物群体的特征。

1.1 什么是生物统计学

我们先给统计学下一个定义。

统计学(statistics)是研究数据资料的收集、整理、分析和解释的科学。

收集数据是根据研究的目的收集相关的数据资料,通常是通过科学试验或抽样调查来获得。正确的结论只能来自高质量的数据资料,而高质量的数据资料需要通过对试验或抽样调查的合理科学设计才能获得。

整理资料是对数据资料进行初步归纳分析,判断并剔除异常数据,找出数据资料的基本特征,并以适当的形式(如表、图等)展示这些数据资料,以便对数据的基本特征有清晰、直观的了解。

分析资料是针对要研究的问题,通过对数据的深入分析,从数据资料中获取所需的相关信息。

解释资料是在分析结果的基础上对所研究的问题做出统计推断。

由以上定义可知统计学是与数据密切相关的科学,因而可将统计学看成是应用数学的一个分支。

将统计学应用于生物科学就称为生物统计学(biostatistics, biometry)。

本书将从以上4个方面介绍生物统计学的基本概念、理论和方法,由于科学试验和抽样调查的设计本身也要用到一些统计学知识,为叙述方便起见,本书将试验设计和抽样调查放在了全书的最后一章。但是我们必须记住,在实际应用中,我们是先通过试验或抽样调查获取数据,然后再进行统计分析。

1.2 统计学的基本特点

我国统计遗传学的奠基人吴仲贤教授在其《统计遗传学》一书中总结了统计学的三个基本特点:

概率性(probability) 在这里概率性是和决定性相对的,统计学以概率论为理论基础,这就决定了它不同于只进行描述、而不进行计算的学科(如解剖学);其次,它又不同于决定性的数量学科(如经典物理学),因为它的结论总有某种概率相伴随。

二元性(duality) 这里“二元”并无哲学的含义,只是指统计学并非一门唯理论的学科,也就是说,除了理论(概率论)之外,它还要有来源于实际的数据资料。如果要给统计学下一个定义,以区别概率论,我们可以说,“统计学是概率论应用于实际资料”。

归纳性(induction) 我们并不想排除演绎法在统计学中的作用,而只想说明统计学的主要精神是归纳性的,它的主要功能是从现实资料中归纳出一般的原理并把它应用于较大的范围。用逻辑学的说法,统计学是由特殊推导一般;用统计学自己的语言来说,统计学是由样本推断总体。

1.3 为什么要学习生物统计学

作为生物科学的研究者和实践者,我们之所以要学习统计学,是因为统计学是生物科学研