

全国中等农业学校教材

生物统计与畜牧兽医试验

上海市农业学校主编

畜牧兽医专业用

农业出版社

全国中等农业学校教材

生物统计与畜牧兽医试验

上海市农业学校 主编

全国中等农业学校教材
生物统计与畜牧兽医试验
上海市农业学校 主编

• • •
责任编辑 刘博浩

农业出版社出版(北京朝阳区枣营路)
新华书店北京发行所发行 农业出版社印刷厂印刷

787×1092mm32开本 7.5印张 1插图 151千字

1989年11月第1版 1989年11月北京第1次印刷

印数 1—11,300册 定价 1.45元

ISBN 7-109-01085-6/S·790

主 编 上海市农业学校 刘信福

编写人 辽宁省锦州畜牧兽医学校 金学志

广西柳州畜牧兽医学校 韦正宇

前 言

本教材是根据1986年5月全国中等农业学校畜牧兽医专业教材编写会的要求，结合当前全国各农业学校的教学实际编写的。

本课程着重讲述生物统计与畜牧兽医试验的基本理论、基本技能和基本方法及其在畜牧兽医实践中的应用。课程的内容在保持必要的系统性基础上，贯彻少而精、删繁就简的原则，尽可能地增加实践性内容，加强对学生实践技能的训练和培养。

本教材第二、三、四章由韦正宇编写，第五、七、八章由刘信福编写，第六、九章由金学志编写。最后由刘信福统稿。本教材承蒙东北农学院盛志廉、浙江农业大学徐继初参加审定，对初稿提出了宝贵的意见并作了认真的修改，谨此致谢。

在编写过程中，虽然我们力图提高质量，但限于水平，一定存在不少缺点和错误，恳请使用者批评指正。

编 者

1988年

目 录

前言

第一章 概论	1
第一节 概述	1
第二节 常用统计术语	3
习题	5
第二章 数据资料的来源和初步整理	6
第一节 数据资料的来源	6
第二节 数据资料的分类	6
第三节 数据资料的初步整理	7
习题	13
第三章 基本统计量的计算	14
第一节 平均数	14
第二节 标准差	21
第三节 变异系数	26
第四节 标准误	28
习题	30
第四章 畜牧兽医试验	32
第一节 畜牧兽医试验的任务与类型	32
第二节 试验计划和方案的拟定	37
第三节 试验设计的原则和方法	41
习题	52
第五章 均数差异显著性检验—— t 检验	54
第一节 均数差异显著性检验的意义和原理	54

第二节 t 检验	64
第三节 两个百分率资料差异显著性检验	70
第四节 平均数的置信区间和畜禽数量的确定	73
习题	79
第六章 方差分析	81
第一节 方差分析的意义和原理	81
第二节 方差分析的步骤	88
第三节 方差分析的基本类型	96
第四节 组内相关系数	120
*第五节 方差分析中的数据转换	124
习题	130
第七章 直线相关与回归	135
第一节 相关与回归的定义	135
第二节 直线相关	136
第三节 直线回归	142
第四节 直线相关与回归的应用	153
习题	156
第八章 χ^2 (卡方) 检验	158
第一节 χ^2 检验的意义和原理	158
第二节 适合性检验	164
第三节 独立性检验	167
习题	176
*第九章 协方差分析	178
习题	191
附录 1 电子计算器的使用方法	193
附录 2 《生物统计与畜牧兽医试验》基本技能考核项目	202
附表 1 t 值表 (两尾)	204
附表 2 配对设计资料样本含量 n 与 K 值的关系	206
附表 3 成组设计资料样本含量 n 与 K 值的关系	207

附表4	χ^2 值表 (一尾)	209
附表5	5%及1% F 值表 (一尾)	211
附表6	5% q 值表 (两尾)	213
附表7	1% q 值表 (两尾)	215
附表8	SSR 表 (邓肯氏新复极差值表)	217
附表9	百分数反正弦 ($\sin^{-1}/\sqrt{x}$) 转换表	221
附表10	r 显著数值表	227
附表11	随机数字表	229
参考文献		233

注,*符号表示为选学内容。

第一章 概 论

第一节 概 述

为了发展畜牧业生产，我们经常要进行调查研究和试验工作。调查和试验都须以生物统计作为研究工具。我们通过统计分析调查和试验所提供的大量数据，进行“量”的关系的研究，可以透过现象，认识事物的本质及其规律性。生物统计是应用数理统计的原理和方法来分析和解释生物界数量资料的一门科学。

在生物科学的研究中，我们所期望知道的是某事物全体（总体）的特征，而不是事物的一部分（样本）的特征，可是在具体实践中，我们所得到的却是样本而不是总体，因此，往往是通过样本来估计总体，达到由特殊推导一般，由样本推断总体的目的，获得对客观事物本质的规律性的认识。例如，在研究陆川猪品种母猪产仔数时，产仔的陆川母猪之多，资料之庞大，几乎不能一一观察记载，只能通过抽样，用样本统计量来估计总体参数。由此可见，生物统计的特点是从样本推断总体。

生物统计与畜牧兽医试验有着极其密切的关系。试验的设计及试验结果的分析与解释，需要运用生物统计的原理和方法。而试验方法的丰富，又促进了生物统计理论和方法的

发展。广义的试验设计是指试验前应用生物统计的原理，对整个试验研究的课题的设计。它包括拟定试验方案、选择试验单位与分组、搜集整理资料、确定统计分析方法，以及实现试验计划的方法和程序等内容。而狭义的试验设计专指对于试验方法的考虑。试验设计是决定试验工作成败的重要环节。一个复杂的试验，如果设计得好，就可以在较短的时间内利用较少的人力、物力，能获得多而可靠的资料，作出科学的结论。

本课程包括生物统计和畜牧兽医试验两大部分。

生物统计部分的主要内容有下列几个方面：数据资料的整理及其基本分析（计算平均数、标准差、变异系数和标准误）、显著性检验（ t 检验、 F 检验和 χ^2 检验）；相关与回归。

畜牧兽医试验部分的主要内容有下列几个方面：试验的任务；试验计划和方案的拟定；试验设计的基本原理以及常见的试验设计方法（完全随机设计、配对设计、随机类组设计和交叉设计等）。

学习生物统计对于促进畜牧生产和提高科研水平，都具有重要意义。畜牧兽医专业的学生学习这门课程的主要目的，在于掌握生物统计的基本原理和方法，应用于畜牧生产和科技推广工作。同时为进一步学习《家畜遗传育种学》、《家畜饲养学》和《兽医临床》等课程打好基础。

值得说明的是，生物统计作为一种工具是不能代替畜牧兽医科学本身的。所以在应用生物统计分析具体问题时，还必须遵循畜牧兽医科学的规律，才能得出正确的结论。

本课程是生物学与数学互相渗透和彼此结合的科学，理论性较强，计算公式较多，学习时要勤于思考，加深理解。

对于统计的基本公式，希望能在明了公式意义的基础上灵活应用。并且本课程的实践性强，运算较多，学习时要理论联系实际，加强练习，多做习题。总之，要以唯物辩证法为指导，以畜牧兽医科学为基础，这样才能充分发挥生物统计与畜牧兽医试验方法在畜牧业生产和科研中的重要作用。

第二节 常用统计术语

一、总体、样本与随机取样 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。组成总体的每一个成员，称为该总体的个体。总体可分为有限总体与无限总体两种，个体有限的总体称为有限总体，个体无限的总体称为无限总体。在畜牧兽医科研工作中所遇到的总体大多数是无限总体。即使是有限总体，也往往作为无限总体来加以考虑和处理。例如，在研究陆川猪品种母猪产仔数总体时，所有陆川母猪的产仔数就组成了一个无限总体，而其中每一陆川母猪产仔数，便是该总体中的一个个体。这样的产仔陆川母猪当然很多，我们无法全部调查得到其产仔数。但研究陆川母猪产仔数的目的，是要摸清这类未知总体的情况，所以，只能采用抽样的方法，使用数理统计方法从样本估计总体。从总体中抽取的若干个体称为样本。一般以含有 30 个以下变数的样本称为小样本，变数在 30 个及 30 个以上的样本称为大样本。样本中所含变数的个数称为样本含量。为了能通过抽取的样本来推断总体，在统计学中一般采用随机抽样的方法。所谓随机抽样，是指抽样时，不掺杂人们的主观愿望，总体中每一个体被抽作样本的机会均应相等，抽取的个体之

间是相互独立的。随机抽样一般可采用抽签、拈阄或者使用随机数字表等方法。

二、变数与变异数列 总体或样本中每一个观察值称为变数。这些观察值各不相同，表现为一定的变异，因而得名“变数”。变数通常用 x 、 y 等表示。按一定规则排列的一组变数称为变异数列。

三、参数与统计量 从总体数据资料中计算出的描述总体的特征数称为参数。例如总体平均数、总体标准差等均为参数。总体参数一般用希腊字母 μ 、 σ 等表示。而从样本数据资料中计算得出的描述样本的特征数称为统计量。统计量一般用英文字母 $\bar{x}$ 、 S 等表示。样本统计量是总体参数的估计值。

四、机误与错误 在试验过程中，由各种无法控制的随机因素所产生的误差，称为机误。例如，由于供试家畜个体特性不同、环境条件的微小差异、疾病的侵袭以及测量和称重时一些不可避免的损失等原因所造成的误差，都属于机误。机误又称为随机误差或偶然误差，它构成狭义的试验误差。在实际统计工作中把狭义的试验误差简称为试验误差，它服从正态分布，其大小是衡量试验精确性的依据。

在试验过程中，由于工作不认真，操作不正确等原因所产生的差错，称为错误。例如，在记载或统计时数字的错漏、结果弄错以及违反试验方案规定所造成的错误等。只要我们在工作中认真负责，错误是完全可以避免的。

五、准确性与精确性 生物统计就是要解决如何用样本来研究总体、了解总体和推断总体的问题。样本统计量接近总体参数的程度，可衡量样本统计量的准确性。准确性说明

了样本统计量对总体参数的符合程度。样本统计量与总体参数偏离程度越小，说明统计量的准确性越高；反之，则低。同一试验内容多次重复进行所得的测定值的变异程度，可衡量样本统计量的精确性。变异程度越小，说明样本统计量的精确性越高；反之，则低。

习 题

1. 什么叫做生物统计？生物统计与畜牧兽医试验或调查有何关系？

2. 什么叫做试验设计？试验设计在畜牧兽医试验中的地位如何？

3. 掌握下列统计术语的概念

总体与样本 随机取样 变数与变异数列 参数与统计量 机误与错误 准确性与精确性

第二章 数据资料的来源和初步整理

第一节 数据资料的来源

在进行生物统计的工作中，首先必须获得畜牧业生产和科学试验所提供的大量的原始数据资料。这些数据资料主要来源于科学试验记录、生产记录和调查资料。

为了获得完整的原始数据资料，试验记录必须根据试验的目的要求，列出在试验过程中所必须观察和记录的项目，并按试验计划执行。生产记录必须根据畜牧业生产发展的需要和存在的某些生产问题，列出测量和记录的项目。调查资料必须根据调查的任务和要求，采取科学的调查方法，有目的有计划地及时搜集。

第二节 数据资料的分类

数据资料按观察或测量对象的性质，可分为计量资料与计数资料。

计量资料是用计量方式得到的。例如体高、体长、体重、产奶量、产毛量等是直接用度量衡等计量工具测定的，其数据是用长度、重量等来表示的。各个数据不一定是整数，也可以是任何小数。在统计学中称这种资料为计量资料，

也称作连续性变数资料。

计数资料是用计数方式得到的。例如猪的头数、产仔数，牛的成活数，鸡的产蛋数等是直接计数得到的，其数据是整数，没有小数，两整数间的数值是不连续的。这种资料称为计数资料，也称作间断性变数资料。

第三节 数据资料的初步整理

数据资料得到后，根据生产或科研的需要，可进行初步整理。数据资料的初步整理，就是按统计分析任务的要求，对生产、试验或调查所获得的大量原始资料进行归类整理、综合汇总，使之条理化系统化，得出便于统计分析、能够反映总体特征的综合资料。

数据资料整理前，必须检查核对数据的完整性、真实性、同质性和准确性。对原始资料要认真、细致地检查和核对，检查数据本身有无差错；有无取样差错（如取样不全、非随机取样等）；有无数据不合理的合并和归类。对于极大或极小的数据，要反复核实，力求确实可靠，这样才能保证统计分析结果的正确性。

整理数据资料的作用就是按不同的标志（如产奶量的多少、产仔数的多少等）把数据的特征反映出来，以便进一步运用生物统计的方法来计算分析，从中找出规律性的东西，指导畜牧业生产和科研工作。

下面介绍数据资料的整理方法。

一、小样本的整理 整理小样本，可将原始数据按数值的大小顺序排列成依次表。例如，现有 21 头初产母猪产仔

数的原始资料 (表 2—1), 将它整理成由小到大排列的依次表 (表 2—2)。

表 2—1 21头初产母猪产仔数资料

(单位: 头)

8	6	7	13	11	9	7	9	8
9	6	8	9	10	12	7	10	8
9	11	9						

表 2—2 21头初产母猪产仔数依次表

(单位: 头)

6	6	7	7	7	8	8	8	8
9	9	9	9	9	9	10	10	11
11	12	13						

这个小样本, 在没有整理之前看不出资料的任何意义。整理成依次表后, 便可一目了然地看出资料的变异情况和集中趋势: 21 头初产母猪产仔数最高为 13 头, 最低为 6 头, 变异范围是 6 至 13 头; 集中趋势为 9 头。

二、大样本的整理 当资料的数据很多时, 为了能从庞大的资料中看出规律性, 可将资料中所有变数按其大小进行分组、归纳, 编制成次数分布表。例如, 现有 150 头母犊牛初生体重的原始资料 (表 2—3)。

将资料整理编制成次数分布表, 一般采用组距式分组法。具体的方法步骤如下:

(一) 求全距 全距是数据资料中的最大值与最小值之差。它表示了整个样本的变异范围。资料表 2—3 中上限为

表 2—3 150头母犍牛初生体重资料

(单位, kg)

25.2	34.9	34.3	38.1	41.3	27.8	33.8	37.7	28.4	33.5
38.0	43.8	40.9	37.5	36.6	33.4	47.4	36.4	41.4	36.5
39.6	33.2	32.1	29.9	43.7	33.8	35.1	37.8	32.4	38.5
27.8	42.6	23.1	37.1	44.0	35.6	44.5	46.5	35.0	31.8
37.1	29.2	38.2	41.1	36.2	43.5	32.8	36.3	31.8	30.6
42.9	20.0	35.4	26.5	42.0	39.6	38.7	35.4	51.2	31.4
33.8	24.5	29.2	45.9	32.5	23.5	36.8	27.2	34.0	34.7
30.0	31.4	40.8	27.3	48.6	35.8	29.7	45.6	41.8	33.0
37.6	25.5	37.2	37.5	41.5	38.4	44.2	43.2	31.5	40.2
33.4	32.4	32.8	36.8	45.7	41.2	40.9	36.5	47.9	35.7
47.3	34.8	30.5	36.2	51.0	34.1	25.3	29.6	38.2	43.7
42.5	33.7	29.3	39.6	37.5	44.4	41.2	35.3	42.6	34.1
28.2	36.5	23.4	35.8	34.1	28.3	33.3	35.1	40.6	38.2
36.4	36.2	47.9	38.7	20.5	34.5	37.4	37.8	33.4	32.2
38.5	39.6	28.7	33.7	35.1	39.3	42.2	35.3	30.1	27.2

51.2kg, 下限为 20.0kg, 全距 = $51.2 - 20.0 = 31.2$ kg.

(二) 确定组数 组数的多少是根据变数的多少来确定的, 以能够反映出变数分布的特点为原则。组数过少, 则由次数分布表计算出统计量的精确度较差; 组数过多, 又难以简化计算手续, 并且可能使某些组包含的次数很少, 甚至不含次数, 因而不利于图示。一般组数的确定, 可参考表 2—4。

表 2—4 确定组数参考标准

样 本 含 量	组 数
30—60	6—8
60—100	7—10
100—200	9—12
200—500	12—17
500以上	17—30