

农业大专院校兽医专业试用教材

# 兽医统计方法

明道绪 主编

成都科技大学出版社

.67  
52

农业大专院校兽医专业试用教材

# 兽医统计方法

明道绪 主编

成都科技大学出版社

S851.67  
6732

音

责任编辑 刘维庆

封面设计 秦昌明

兽 医 统 计 方 法

明道绪 主编

---

787×1092	1/16	18.5印张	465千字
成都科技大学出版社出版发行			雅安地区印刷厂印刷
1991年8月第1版			1991年8月第一次印刷
	印数1—2000册		

---

ISBN7-5616-0988-4/R·30

定价：7.55元

主 编 明道绪（四川农业大学）

编写人 钟运炎（广西农学院）

刘晓明（贵州农学院）

江志琴（西南民族学院）

胡清林（河南农业大学）

主 审 俞清江（贵州农学院）

荣廷昭（四川农业大学）

审稿人 张经夔（华中农业大学）

陈文广（华南农业大学）

张性敏（河北农业大学）

刘秉颖（石河子农学院）

马卫东（湖南农学院）

冯保华（西南民族学院）

杨秀琼（绵阳农业专科学校）

谢和芳（四川畜牧兽医学院）

刘 勇（西昌农业专科学校）

刘克惠（四川农业大学）

刘维庆（四川农业大学）

黄玉碧（四川农业大学）

# 序

建国四十年来，在党的领导下，我国建立了畜牧兽医教育和科研体系，培养了大批专业人材，形成了强大的科技队伍。尤其是兽医科学技术的提高与普及为促进我国畜牧业的迅速发展起了重要作用。

人们在长期实践中，逐渐认识到在兽医科学研究中试验设计与统计分析的重要性。为此，我国高等农业院校兽医专业于1978年决定将《兽医统计方法》作为一门专业基础课开设，以培养学生从事科学研究和分析问题、解决问题的能力。这一决定是十分正确的。

兽医统计方法不仅应用于疫病调查，药物检验，也应用于免疫学、诊断学的病因分析。随着计算技术的改进，其应用范围日益扩大。作为兽医科学研究的重要工具，兽医统计方法为提高兽医科学研究水平作出了贡献，间接地为提高畜禽疾病防治水平、促进畜牧业发展作出了贡献。

《兽医统计方法》由四川农业大学明道绪副教授主编，五所兄弟院校合作编写。在目前我国农业大专院校兽医专业缺少合适的教材，广大兽医科技工作者缺少统计方法工具书的情况下，编写和出版这本书是十分必要和及时的。通过集体审稿，集思广益，使本书编写质量得到可靠保证。本书的特点是：比较系统地叙述了兽医统计方法，而且保持了一定的理论性。例如，能够深入浅出地叙述概率分布，显著性检验的基本原理，数学模型；在一些内容上具有一定的新颖性。例如，在次数资料分析中引进了Ridit分析；在试验设计中介绍了序贯试验设计等，具有一定的实用价值。编者尽可能利用兽医应用实例，但限于我国兽医统计方法应用起步较晚，书中引用了少量人医试验资料或畜牧实例，这有待于再版时修改和充实。

编者本着由浅入深、循序渐进、通俗易懂、学以致用原则，尽可能使本书适合兽医专业的教学要求和广大兽医科技工作者实际需要。我确信，如果兽医科技工作者能在临床实践和试验研究中加强统计方法的学习和应用，必将取得“如虎添翼”之效。本人竭诚乐意推荐此书，特为之作序。

俞渭江

一九九〇年十月十一日于成都

# 前 言

为了满足农业大专院校兽医专业教学的需要、为了给广大兽医科技工作者提供一本介绍统计方法的工具书，我们五所院校：四川农业大学、广西农学院、贵州农学院、西南民族学院、河南农业大学联合编写了这本《兽医统计方法》。

编写组于1990年3月底在四川成都举行的《兽医统计方法》编写会议上正式成立。编写组由明道绪、钟运炎、刘晓明、江志琴、胡清林五人组成，明道绪副教授任主编。为了保证教材编写质量，在写出初稿后，于1990年10月上旬在四川成都召开了《兽医统计方法》审稿会。在会上，主审人：俞渭江教授、荣廷昭教授，审稿人：张经夔、陈文广、张性敏、刘秉颖、马卫东、冯保华、杨秀琼、谢和芳、刘勇、刘克惠、刘维庆、黄玉碧等同志和编写组成员一起对初稿逐章逐节进行了认真、仔细的审定，提出了十分中肯的修改建议。这本书的编写、出版，凝结了全体编、审人员的心血。

本书根据农业大专院校兽医专业的教学要求、针对学生已具有的高等教学基础，考虑到我国目前《兽医统计方法》师资队伍的实际状况、电子计算器已普遍使用并兼顾作为广大兽医科技工作者学习统计方法的中级读物的需要确定各章节的内容与深、广度，力求体现科学性、系统性、针对性与实用性相统一。全书包含概论，资料的来源与整理，平均数、标准差与变异系数，常用概率分布，均数差异显著性检验—t检验，参数的区间估计，方差分析，次数资料的分析、直线回归与直线相关，多元线性回归与多项式回归，非参数检验，半数致死量，试验设计十三章。个别章、节标记有“\*”，供教学时选用。本书在从理论上阐述统计学原理的同时，着重于基本概念、基本方法的介绍，每一种方法都列有步骤完整、过程详细的实例予以说明。各章有复习思考题与习题（附答案）。

本书在编写过程中曾参阅了有关中外文献与专著，并引用了其中的一些资料。编者对这些文献与专著的作者表示诚挚的谢意。

编者还对热情指导、大力支持编写工作的四川农业大学、广西农学院、贵州农学院、西南民族学院、河南农业大学的各级领导同志，对在本书出版发行上给予大力帮助的四川农业大学学报编辑部的刘维庆同志，为本书最后定稿作了大量具体工作的四川农业大学兽医系陈云同志和农学系生物统计教研室全体同志一并致以衷心感谢。

限于编者水平，书中难免有缺点和错误，敬请读者批评指正，以便再版时修改。

编 者

一九九一年三月

# 目 录

<b>第一章 概 论</b> .....	( 1 )
第一节 兽医统计方法在兽医科学研究中的作用.....	( 1 )
第二节 统计学的几个基本概念.....	( 2 )
习 题.....	( 4 )
<b>第二章 资料的来源及整理</b> .....	( 5 )
第一节 资料的来源.....	( 5 )
第二节 资料的分类.....	( 6 )
第三节 资料的整理.....	( 7 )
第四节 统计表及统计图.....	( 10 )
习 题.....	( 14 )
<b>第三章 平均数、标准差与变异系数</b> .....	( 16 )
第一节 平均数.....	( 16 )
第二节 标准差.....	( 22 )
第三节 变异系数.....	( 25 )
习 题.....	( 26 )
<b>第四章 常用概率分布</b> .....	( 28 )
第一节 二项分布.....	( 28 )
第二节 普阿松分布.....	( 30 )
第三节 正态分布.....	( 34 )
习 题.....	( 38 )
<b>第五章 均数差异显著性检验— t 检验</b> .....	( 39 )
第一节 假设检验概述.....	( 39 )
第二节 样本平均数的抽样分布与t 分布.....	( 41 )
第三节 假设检验的基本原理.....	( 44 )
第四节 样本均数与总体均数差异显著性检验.....	( 48 )
第五节 两样本均数差异显著性检验.....	( 49 )
第六节* 对数转换与t' 检验.....	( 52 )
习 题.....	( 55 )
<b>第六章 参数的区间估计</b> .....	( 58 )
第一节 正态总体平均数 $\mu$ 及二项总体百分率p的置信区间.....	( 58 )
第二节 正常值范围的确定.....	( 61 )
习 题.....	( 66 )
<b>第七章 方差分析</b> .....	( 68 )

第一节	方差分析的基本原理	(69)
第二节	单因素试验结果的方差分析	(80)
第三节	两因素试验结果的方差分析	(85)
第四节*	方差同质性检验与数据转换	(102)
	习 题	(107)
<b>第八章</b>	<b>次数资料的分析</b>	(111)
第一节	$\chi^2$ 检验	(111)
第二节	Ridit分析	(127)
第三节	相对数分析	(132)
	习 题	(137)
<b>第九章</b>	<b>直线回归与直线相关</b>	(140)
第一节	回归分析与相关分析概述	(140)
第二节	直线回归	(141)
第三节	直线相关	(152)
第四节	直线回归与相关的应用	(155)
第五节*	曲线回归	(156)
	习 题	(163)
<b>第十章*</b>	<b>多元线性回归与多项式回归</b>	(165)
第一节	多元线性回归	(165)
第二节	多项式回归	(177)
	习 题	(180)
<b>第十一章</b>	<b>非参数检验</b>	(182)
第一节	符号检验	(182)
第二节	秩和检验	(185)
第三节	等级相关分析	(193)
	习 题	(195)
<b>第十二章</b>	<b>半数致死量</b>	(198)
第一节	基本概念	(198)
第二节	半数致死量的测定方法	(199)
	习 题	(208)
<b>第十三章</b>	<b>试验设计</b>	(209)
第一节	试验设计概述	(209)
第二节	兽医试验计划	(210)
第三节	试验设计的基本原则	(212)
第四节	完全随机设计	(213)
第五节	配对设计	(215)
第六节	随机区组设计	(216)
第七节	拉丁方设计	(221)



第八节 序贯试验设计	(227)
第九节 正交试验设计简介	(231)
第十节 调查设计	(242)
习 题	(247)
主要参考文献	(249)
<b>附 表</b>	(250)
1. 正态分布表	( 1 )
2. 正态分布的双侧分位数 ( $u_\alpha$ ) 表	( 3 )
3. 质反应双向序贯试验数值表	( 3 )
4. t值表 (两尾)	( 5 )
5. F值表 (两尾, 方差齐性检验用)	( 6 )
6.1. 百分率的可信限 ( $1 \leq n \leq 50$ )	( 7 )
6.2. 百分率的可信限 ( $50 \leq n \leq 100$ )	(10)
6.3. 百分率的可信限 ( $100 \leq n \leq 1000$ )	(11)
7. 百分率与概率单位换算表	(12)
8. F值表 (方差分析用)	(14)
9. q值表	(18)
10. SSR值表	(20)
11. $\chi^2$ 值表	(22)
12. 百分率反正弦 ( $\sin^{-1}\sqrt{x}$ ) 转换表	(23)
13. r与R的显著数值表	(25)
14(1). 符号秩和检验用T临界值表	(26)
14(2). 秩和检验用T临界值表	(26)
14(3). 秩和检验用H临界值表	(27)
15. 符号检验用K临界值表	(27)
16. 等级相关系数 $r_s$ 临界值表	(28)
17. 加权系数(W)和计算矫正概率单位的( $y_c$ )的常数	(29)
18. 随机数字表(I)、(II)	(30)
19. 非配对试验样本含量表	(32)
20. 配对试验样本含量表	(32)
21. 单因素多组群样本含量表	(33)
22. 质反应单向序贯试验边界系数表	(34)
23. 对平均数作抽样调查, $S/\delta$ 取不同数值所需样本大小n, $\alpha = 0.05$	(35)
24. 对平均数作抽样调查, $S/\delta$ 取不同数值所需样本大小n, $\alpha = 0.01$	(35)
25. 对率作抽样调查时所需样本大小n, $\alpha = 0.05$	(36)
26. 对率作抽样调查时所需样本大小n, $\alpha = 0.01$	(36)
27. 常用正交表	(37)

# 第一章 概 论

## 第一节 兽医统计方法在兽医科学研究中的作用

在兽医科学研究中，常常要进行调查或试验。例如研究某种畜禽疾病在某地区的发病率、某种疾病病死率逐年的变化、新的疫苗预防某种疾病的效果、新的诊断手段是否比原有的更快速、准确，新的疗法、新的药物治疗某种疾病是否比原有的更有效；以及新的药物的筛选，半数致死量的测定，畜禽生理指标正常值范围的制定等都要通过调查或试验来解决。进行调查或试验首先必须解决的问题是：如何进行调查或试验设计？在实际研究工作中，常常碰见这样的情况：由于调查或试验设计不合理，以至于无法从所获得的数据提取有用的信息，造成人力、物力和时间的浪费。若调查或试验设计方法好，则可以收到事半功倍之效，即用较少的人力、物力和时间收集到必要而有代表性的资料，从中获得可靠的结论，达到调查或试验的预期目的。

兽医研究进行调查或试验常常获得的是具有一定变异的数据。例如测量100头内江猪的血红蛋白含量所得的100个数据，彼此不完全相同，表现出一定的变异；又如用同一种药物治疗若干头患有同种疾病的病畜，其反应不一定相同，也表现出一定的变异。产生这种变异的原因，有的已被人们所了解。例如品种、性别、年龄、体况、病程、药物种类、药物剂量等等不同。另外还有许多内在和外在的因素还未被人们所认识。由于这些人们尚未认识，因而无法控制的因素的作用，使得通过调查或试验得来的数据普遍具有变异性。所以在进行调查或试验时还必须解决的第二个问题是：如何整理、分析所收集得来的具有变异的资料，揭示出隐藏在其内部的规律性。这两个问题的解决，促使兽医统计方法的建立与发展。

兽医统计方法是统计学原理和方法在兽医科学研究中的应用，是一门应用数学。它在兽医科学研究中的作用是：

### 一、提供试验与调查设计方法

广义的试验设计是指试验研究的课题设计，也就是指整个试验计划的拟定。狭义的试验设计主要是指试验单位（如兽医临床试验的病畜）的选取、重复数目的确定及试验单位的分组。合理的试验设计能控制和降低试验误差，为统计分析获得试验处理效应和试验误差的无偏估计提供必要的的数据。

广义的调查设计是指整个调查计划的制定。狭义调查设计主要包含抽样方法的选取、抽样单位、抽样数目的确定等内容。合理的调查设计能控制与降低抽样误差，为获得总体参数的精确估计提供必要的的数据。

简而言之，试验与调查设计主要解决合理地收集资料的问题。

## 二、提供整理、分析资料的方法

整理资料的基本方法是根据资料的特性将其整理成统计表，绘制成统计图。通过统计表、图可以大致看到所得资料的变化规律。并利用所收集得来的数据计算几个统计量，以表达资料的数量特征，估计相应的总体参数。

统计分析最重要的内容是差异显著性检验。通过抽样调查或控制试验，获得的是具有变异的资料。产生变异的原因是什么？是由于进行比较处理间，例如不同疫苗、不同药物、不同疗法间有实质性的差异或是由于无法控制的偶然因素所引起？显著性检验的目的就在于承认并排除这些无法控制的偶然因素的干扰，将处理间是否存在本质差异揭示出来。显著性检验的方法很多，常用的有t检验、方差分析、 $\chi^2$ 检验等。

统计分析的另一个重要内容是对试验指标或畜禽性状间的关系进行研究：或者研究它们之间的联系性质和程度，或者寻求它们之间的联系形式，即相关分析与回归分析。通过对资料进行回归、相关分析，可以揭示出试验指标或性状之间的内在联系，为畜禽疾病的预测、控制、早期诊断、判别等提供强有力的依据。

还有一类统计分析方法不考虑资料的分布类型，也不事先对有关总体参数进行估算，这类统计分析方法叫非参数检验法。非参数检验法计算简便，当通常的方法对兽医科学研究中的某些资料束手无策时，它可以大显身手。此外，统计学还针对兽医科学研究的需要提供了正常值范围估计、Ridit分析、序贯试验、半数致死量测定等方法。

从以上对兽医统计方法在兽医科学研究中作用的介绍看出，兽医统计方法对于进行兽医科学研究是多么重要。它是每个兽医科技工作者必须掌握的其本工具。可喜的是，随着统计方法的普及、计算工具的改进，已有越来越多的兽医科技工作者掌握并在实际工作中应用了统计方法，取得了成效。

兽医统计方法除包含上述内容外，还有很多内容，例如平衡不完全区组设计、交叉设计、希腊拉丁方、挠敦方；回归分析的正交设计、旋转设计、D—最优设计；通径分析、主成分分析、聚类分析、判别分析、因子分析、典型相关分析等。本书限于篇幅未将这些内容编入。对这些内容感兴趣的读者，请查阅有关专著，例如郭祖超主编《医用数理统计方法》（1988年，第3版，人民卫生出版社）。

## 第二节 统计学的几个基本概念

兽医统计学是一门应用数学，其理论基础是概率论与数理统计。它涉及的数学概念多、计算公式多和数学用表多；还要求掌握一定的计算工具，例如掌握能进行统计运算的电子计算器的使用；从判断方式上要求摆脱传统的确定性推断方式而接受建立在概率论基础上的统计推断方式。这对初学者来说有一定的难度。为了便于初学者理解、掌握，除了结合实例、从应用的角度来介绍外，每章后还附有一定数量的习题供初学者练习。对于初学者来说，能正确理解统计学的基本概念、掌握并应用所介绍的基本的设计与分析方法解决兽医科学研究中收集、整理、分析资料的问题，也就达到预期目的了。

在这一节里首先介绍统计学中几个最常用的概念。

## 一、总体和样本

根据研究目的确定的研究对象的全体称为总体 (population), 其中的一个研究单位称为个体 (individual); 总体的一部分称为样本 (sample)。例如研究内江猪的血红蛋白含量, 每一头内江猪血红蛋白含量测定值的全体就构成内江猪血红蛋白含量的总体; 而测定100头内江猪血红蛋白含量所得的100个数据则是内江猪血红蛋白含量总体的一个样本, 这个样本包含有100个个体。包含有限个个体的总体称为有限总体。例如上述内江猪血红蛋白含量总体, 虽然包含的个体数很多, 但仍为有限总体。包含有无限多个个体的总体叫无限总体。例如在统计学理论研究上, 常常研究包含一切实数的总体, 这就是一个无限总体。在实际研究中还有一类假想总体。例如进行药物疗效试验, 实际上并不存在用这些药物进行治疗的总体, 只是假设有这样的总体存在, 把所进行的试验看成是假想总体的一个样本。样本中所包含的个体数目叫样本容量或大小 (sample size)。例如上述内江猪血红蛋白含量样本的容量为100。样本容量常记为 $n$ 。通常将 $n \leq 30$ 的样本叫小样本,  $n > 30$ 的样本叫大样本。

统计学总是通过样本来了解总体。这是因为或者总体是无限的、假想的; 即便是有限的, 但包含数目相当多的个体, 或者数据的获得带有破坏性, 因而不可能或者花费很多人力、物力和时间一一测量。研究的目的是要了解总体, 然而能观察到的却是样本。通过样本来推断总体这是统计分析的基本特点。为了能可靠地从样本来推断总体, 要求样本具有一定的含量和代表性。只有从总体随机抽取的样本才具有代表性。所谓随机抽取 (random sampling) 是指总体中的每一个个体都有同等的机会被抽取作为样本。然而样本毕竟只是总体的一部份, 尽管样本具有一定的数量也具有代表性, 通过样本来推断总体不可能是百分之百的正确。有很大的可靠性但有一定的错误率这是统计分析的又一特点。所以Lienert(1973)指出: 作为科学方法论的现代统计学究竟能提供什么? 它能回答在抽样调查中所发现的差异、联系和规律性以什么样的概率纯属偶然? 以及对于总体来说这些发现作为一般规律的可靠程度有多大?

## 二、参数与统计量

为了表示总体和样本的数量特征, 需要计算出几个特征数。由总体计算的特征数叫参数 (parameter); 由样本计算的特征数叫统计量 (statistic)。参数常用希腊字母表示, 例如总体平均数 $\mu$ , 总体标准差 $\sigma$ , 总体相关系数 $\rho$ 等。统计量常用拉丁字母表示, 例如样本平均数 $\bar{x}$ 、样本标准差 $s$ , 样本相关系数 $r$ 等。总体参数由相应的统计量来估计, 例如用 $\bar{x}$ 估计 $\mu$ 、用 $s$ 估计 $\sigma$ , 用 $r$ 估计 $\rho$ 等。

## 三、准确性与精确性

准确性 (accuracy) 是指在调查或试验中某一试验指标或性状的观察值与其真值接近的程度。设真值为 $\mu$ , 观察值为 $x$ 。若 $|x - \mu|$ 小, 则所得的观察值 $x$ 的准确性高; 反之则低。精确性 (precision) 指调查或试验中同一试验指标或性状的重复观察值彼此接近的程度。若观察值 $x$ 间彼此接近, 即 $|x_i - x_j|$ 小, 则所得的观察值精确性高; 反之则低。准确性、精确性的意义图示如下:

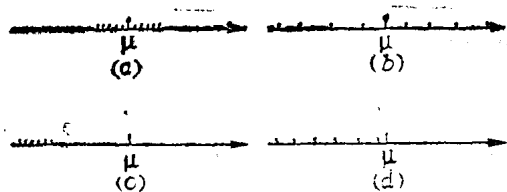


图1—1 准确性与精确性示意图

图1—1 (a) 表示观察值密集于真值  $\mu$  两侧，准确性高，精确性高；图1—1 (b) 表示观察值稀疏地分布于真值  $\mu$  两侧，准确性低，精确性低；图1—1 (c) 表示观察值密集于远离值  $\mu$  的一侧，准确性低，精确性高；图1—1 (d) 表示观察值稀疏地分布于远离真值  $\mu$  的一侧，准确性低、精确性低。

调查或试验的准确性与精确性合称为正确性，在调查或试验中应严格按照调查或试验计划进行，准确地进行观察记载，力求避免人为的差错，特别要注意试验条件的一致性，除所研究的各个处理外，供试畜禽的初始条件如品种、性别、年龄、健康状况、病情轻重等以及饲养条件，管理措施等应尽可能一致，并通过合理的调查或试验设计努力提高试验的准确性和精确性。由于真值  $\mu$  常常不知道，所以准确性不易度量，但利用统计方法可度量精确性。

#### 四、随机误差与系统误差

在兽医科学试验中，试验指标除受试验因素影响外，还受到许多其它非试验因素的干扰，从而产生误差。试验中出现的误差分为两类：随机误差 (random error) 与系统误差 (systematic error)。随机误差也叫抽样误差，这是由于许多无法控制的内在和外在外在的偶然因素如试验动物的初始条件、饲养条件、管理措施等尽管在试验中力求一致但不可能完全控制一致所造成。随机误差带有偶然性质，在试验中，即使十分小心也难以消除。随机误差影响试验的精确性。统计上的试验误差是指随机误差。这种误差越小，试验的精确性越高。系统误差也叫片面误差，这是由于试验动物初始条件不同，例如品种、年龄、性别、病程等不同，试验条件如饲料种类、品质、数量、管理措施相差较大，仪器不准、标准试剂未经校正，药品批次不同、药品用量与种类不符合试验计划的要求，以及观察、记载、抄录、计算中的错误所引起。系统误差影响试验的准确性。图1—1 (c)、(d) 所表示的情况，则是由于出现了系统误差的缘故。一般说来，只要试验工作做得精细，系统误差容易克服。图1—1 (a) 表示克服了系统误差的影响，且随机误差也较小，因而准确性高、精确性高。

#### 习 题

1. 什么是兽医统计方法？它在兽医科学研究中有何作用？
2. 什么是总体、个体、样本、样本含量、随机样本？统计分析的二个特点是什么？
3. 什么是参数、统计量？二者有何关系？
4. 什么是试验的准确性与精确性？如何提高试验的准确性与精确性？
5. 什么是随机误差与系统误差？如何控制、降低随机误差，避免系统误差？

## 第二章 资料的来源及整理

### 第一节 资料的来源

#### 一、资料的收集

科学研究的第一项工作是收集资料。为了保证收集资料的完整性和正确性，在收集资料前，必须根据研究的目的是要求，细致地考虑资料收集的方法、内容和项目。力求用少量的人力、物力和时间，收集到必要而可靠的数据。兽医研究资料可以由经常性工作、专题调查及控制试验获得。

##### (一) 经常性工作

1. 统计报表 指兽医工作情况的月、季度、年度报表，以及各种畜、禽疾病类别报表等，是定期取得系统而全面的统计资料的一种主要形式。统计报表能系统地反应各地区或各牧场在各个时期工作的基本情况，反应畜禽的健康和畜牧业的发展概况，是总结工作、拟定防治措施的重要依据。

2. 日常工作记录 指兽医院及各级兽防站的日常工作记录，如门诊、住院病历、临床化验记录以及一些专门内容的报告单，如产仔数、死亡数、传染病、流行病、中毒等报告单。依据这些资料，可以评价兽医工作的质量及效果，探索疾病的发生发展规律。

(二) 专题调查 指为了某一特定目的，如制定畜禽防治措施、了解防治效果、预测疫情发展趋势、探索发病原因等组织专人进行的调查。专题调查也是统计资料的一个重要来源，并且是研究和解决兽医工作中实际问题的一种重要方法。专题调查分为普查和抽样调查两种；普查是指对研究对象中的所有个体作调查，如对奶牛乳房炎的普查，对猪瘟病的普查等；抽样调查是在整个研究对象中，根据随机抽样的原则，从总体中抽取一定数量的个体组成样本进行观测。

(三) 控制试验 指在人为控制试验条件下，用少量的、初始条件大致相同的动物，在小范围内进行试验。如临床疗效试验、药物效果试验、动物生理生化指标测定试验等。

控制试验与专题调查有着十分密切的联系。专题调查从生产实际出发，为控制试验提出研究课题。通过控制试验，解决研究课题所提出的问题，得到规律性的认识，又再去指导生产，并在实践中接受检验。这又为未来的调查提出了新的任务。这二者密切配合，相互促进，推动着兽医科学研究和畜牧业的发展。

#### 二、资料的检查核对

收集来的资料在未整理之前称为原始资料。对原始资料的全部数据应作检查核对，分辨

真伪，力求完整、正确。对原始资料可以从以下两方面进行检查：

(一) **资料的完整性** 检查原始记录是否有遗漏或重复，各个项目是否填齐。

(二) **资料的正确性** 检查全部项目填写的数据是否正确，各项目间、数据间有无矛盾，是否合理。如研究疗效时，治愈和好转必须严格区分，好转不能作为治愈，否则治愈率就会偏高。

对于有重复或遗漏的资料，应予以删减或补充；对有错误、相互矛盾的资料应进行更正，必要时进行复查或试验。资料的检查核对是统计工作中极为重要的一步，一般是在收集资料的进程中同时进行，即边观察记载边检查核对，发现问题及时补救。

## 第二节 资料的分类

正确地进行资料的分类是对资料进行整理、分析的基础。不同类型的资料其整理、分析的方法不同。兽医统计资料按其性质可分为三类。

### 一、计量资料

计量资料 (quantitative data) 是指在调查或试验中，对所考察的性状或指标通过量测或直接计数得来的资料。如家畜的形态指标：体重、身高、体长、胸围等；生理生化指标：呼吸次数、脉搏次数、体温、白细胞数、红细胞数、血糖含量、血红蛋白含量等。由每一观察单位获得的计量资料值只有量的区别，没有质的区别。由量测方式得来的数量资料带有单位，如长度单位、重量单位、面积、体积单位等。在理论上，它们可以取数轴上某一区间内的一切数值。这一类数量资料在统计上称连续型变量。由直接计数得来的数量资料其数值为正整数，如呼吸次数、脉搏次数等。对于这一类数量资料在分析时按连续型变量对待。

### 二、次数资料

次数资料 (enumeration data) 是指先将观察单位按所考察的性状或指标的属性或类别分成若干组，然后清点各组观察单位的次数得到的资料。如通过检查家畜粪便中是否有寄生虫卵以研究感染情况，先将检查家畜按粪检阳性(有虫卵)和阴性(无虫卵)分组，再清点各组头数。又如研究家畜对几种传染病的致病情况，先将病畜按传染病类别归类，然后统计各类传染病的病畜数。在次数资料中，同组内各观察单位是同质的，但组与组之间则有质的区别。

次数资料的取值通常是正整数和零。如统计十个奶牛场二十年中的死胎数，用 0、1、2、……等整数表示，不会出现小数；统计家畜某种疾病的患病数，若包含过去已经发生的和将来可能发生的在内，则是无限的，但又是可数的。统计学上将次数资料称为离散型变量。

### 三、半定量(等级)资料

半定量或等级资料 (semi-quantitative or ranked data) 是指将观察单位按所考察的性状或指标的等级顺序分组，然后清点各组观察单位的次数而得的资料。这类资料既有次数资料的特点，又有程度或量的不同。如粪便潜血试验的阳性反应是在涂有粪便的棉签上加试剂后观察颜色出现的快慢及深浅程度分成六个等级；

加试剂后颜色变化	阳性反应	等级
立即呈现深兰或深绿色	最强性反应	卅
初现浅兰色半分钟后呈现深兰或深绿	强阳性反应	卅
半分钟到1分钟呈现绿兰色	阳性反应	++
1分钟到2分钟呈现绿色	弱阳性反应	+
2分钟到5分钟呈现浅绿色	痕迹反应	±
5分钟后不变色	阴性反应	-

又如用某种药物治疗家畜的某种疾病，疗效分为“无效”、“好转”、“显效”和“控制”四个级别；然后统计各级别的供试畜禽数或病畜数。半定量资料在兽医研究工作中是很常见的。

三种不同类型资料相互间是有区别的。但有时可根据研究的目的和统计方法的要求将一种类型的资料转化为另一种类型的资料。如白细胞总数属于计量资料。若以白细胞总数正常（4000~10000个/mm<sup>3</sup>）与白细胞总数不正常（大于10000个/mm<sup>3</sup>或小于4000个/mm<sup>3</sup>）分组，清点次数，于是计量资料就转化为次数资料。脉搏数是计量资料，若以小于60次/分为缓脉，以60~100次/分为正常脉，大于100次/分为速脉分组，清点次数，亦转化为等级资料。

### 第三节 资料的整理

当原始资料经检查核对以后，应根据资料的类型及研究目的对资料进行整理，为进一步统计分析打下基础。

#### 一、次数资料的整理

兽医统计资料大部分是次数资料。次数资料的整理方法是先将观察单位按所考察的性状或指标的属性或类别分组，清点各组的次数（频数）；根据需要进一步算出构成比或率；最后将整理结果列成表格形式。如整理仔猪死亡情况资料可按死亡原因将仔猪分组，并统计次数，计算出构成比，见表2—1。

表2—1 仔 猪 死 亡 情 况

死 因	死 亡 数	构 成 比(%)
冻 死	15	19.23
发 育 不 良	20	25.64
肺 炎	13	16.67
白 痢	10	12.82
寄 生 虫	20	25.64
合 计	78	100.00



又如调查四个品种牛的蹄病—趾间纤维瘤的发病情况，整理成表 2—2。

表 2—2 四个品种牛趾间纤维瘤的发病情况

品 种	调 查 头 数	发 病 头 数	发 病 率 (%)
海福特公牛	113	82	72.6
荷兰公牛	250	29	11.6
夏洛来种公牛	320	110	32.4
海福特繁殖母牛	156	6	3.8
合 计	836	337	27.1

## 二、计量资料的理整

对于计量资料，若资料包含的观察值个数少，即小样本，可不进行整理，直接进行统计分析。若资料包含的观察值个数多，即大样本，整理的方法是：将资料中的各个观察值按数值大小分成若干组，编制频数分布表。编制频数分布表分为：确定全距、组数、组距、组中值及组限，将全部观察值用划“正”字记数归组等几个步骤。下面举一实例说明编制频数分布表的具体过程。

〔例 2.1〕将 136 只绵羔羊红细胞含量资料（见表 2—3）整理成频数分布表。

表 2—3 136 只绵羔羊红细胞含量资料（单位： $\times 10^6/\text{mm}^3$ ）

13.5	12.5	12.1	9.2	9.8	12.7	9.9	12.8	12.0	15.4
10.7	11.0	10.9	13.5	10.6	12.9	12.5	14.5	12.2	15.4
10.0	14.9	10.3	10.8	8.5	9.5	7.4	9.7	15.2	12.9
14.9	12.6	8.4	10.5	13.4	12.4	8.5	14.0	10.3	13.8
9.8	14.1	10.4	11.5	15.0	12.8	15.5	13.5	9.5	15.6
11.3	12.2	14.0	9.5	12.5	15.9	7.3	11.3	14.8	15.4
9.3	15.6	10.9	11.0	9.2	8.3	12.8	12.9	11.9	14.5
7.6	11.3	12.5	12.8	11.8	14.3	8.8	10.9	8.0	10.8
9.7	12.3	10.5	13.6	10.8	13.7	14.2	14.7	12.4	12.6
12.7	12.8	12.9	14.2	14.5	13.0	13.7	11.1	12.5	14.1
14.9	9.7	11.6	10.5	16.5	11.3	15.8	8.1	13.8	14.2
12.8	13.4	7.6	9.3	13.8	8.5	12.6	13.8	11.9	11.8
12.0	12.1	12.2	12.4	12.5	11.1	11.2	11.3	11.4	11.5
11.6	8.7	8.8	8.9	11.2	11.4				

（一）求全距 全距是资料中最大值与最小值之差，又称为极差。表 2—3 中最大值为 16.5，最小值为 7.3，则全距 =  $16.5 - 7.3 = 9.2$  ( $\times 10^6/\text{mm}^3$ )。

（二）确定组数 组数要适当，不宜过多，也不可太少。一般原则是：资料变异幅度