21世纪高等学校规划教材

HUATENG EDUCATION

高等学校经济管理类核心课程教材

统 计 学

李静萍 主编

Statistics

上海交通大学出版社
SHANGHAI JIAO TONG UNIVERSITY PRESS

21世纪高等学校规划教材

高等学校经济管理类核心课程教材

统计学

主编　李静萍

上海交通大学出版社
SHANGHAI JIAO TONG UNIVERSITY PRESS

内容提要

本书是一部既吸收现有教材的优秀成果，又有所突破的新教材。主要介绍了统计学概述、统计数据的描述、概率与概率分布、参数估计、假设检验、随机变量间统计关联性分析、回归分析、时间序列分析、多元统计分析等内容。此外，本书特别安排了时间序列 ARMA 模型以及多元统计分析方法，特设了“软件操作指南”栏目，介绍了 Excel 软件及 R 软件的相关操作。

本书可供高等院校经济管理类专业（非统计学专业）本科生使用，也可作为社会相关人员的参考用书。

图书在版编目(CIP)数据

统计学/李静萍主编. —上海：上海交通大学出版社，2012

ISBN 978-7-313-08146-9

Ⅰ. ①统… Ⅱ. ①李… Ⅲ. ①统计学 Ⅳ. ①C8

中国版本图书馆 CIP 数据核字(2012)第 022221 号

统计学

李静萍 主编

上海交通大学出版社出版发行

（上海市番禺路 951 号 邮政编码：200030）

电话：64071208 出版人：韩建民

北京振兴源印务有限公司印刷 全国新华书店经销

开本：787mm×1 092mm 1/16 印张：17 字数：414 千字

2012 年 2 月第 1 版 2012 年 4 月第 1 次印刷

ISBN 978-7-313-08146-9/C 定价：32.00 元

前 言 Preface

自2000年任教以来，我几乎每年都承担非统计学专业统计学课程的教学工作。还记得初次讲授该课程时，总觉得如果不讲清楚公式的证明，就难以“证明”统计的科学性，就不能将统计学的真谛传授给学生，结果是黑板上密密麻麻的公式并没有换来理想的教学效果。我不得不开始思考一个问题：对于非统计学专业统计学课程的教学，如何才能让学生更好地理解统计学的基本思想、掌握统计学的基本方法并加以科学的应用？

随着教学和研究经验的积累，我逐渐意识到，统计学的生命在于应用，无论多么复杂高深的统计方法，都是来自于实际数据分析中提出的需求。因此，统计学的高明之处不在于数学公式有多复杂，而在于其解决实际问题的思路有多精妙。认识到这一点之后，我便在教学过程中适当弱化数学推导，同时更加强调统计学的基本思想。具体来说，我力图让学生明白，每一种统计方法背后对应着什么样的现实需求，每一种方法如何巧妙地将现实问题转化为统计问题，在这一转化的背后有什么样的假定，统计分析的结果又应如何转化为现实中的语言表达。

如今，我践行这一教学理念已近10年，可以说积累了一定的经验，遂萌生了自己编写一部统计学教材的想法，将自己的教学心得记录下来，与其他老师以及学生共享。但是，目前市场上已经有不少优秀的面向非统计学专业的统计学教材。编写一部既能吸收现有教材的优秀成果，又能有所突破的新教材，是我写作本书的初衷。

经过整整一年的写作，这部教材终于面世了。如果对比本教材与现有的同类其他教材，我想主要内容并无大不同，不过本教材还是有其独特之处，主要体现在以下几个方面。

第一，本教材强调统计学的应用背景。好的教材应当有好的例题和习题。很多教材为了表现统计方法的分析功能，常常构造一些符合需要的数据。这样做的结果是学生在演习教材上的例题和习题时，可以轻松得到漂亮的分析结果，但是一旦涉及现实中的实际数据，就常常束手无策。为了提高学生应用统计方法的能力，本教材绝大多数的例题和习题都使用真实数据，通过例题的分析和习题的训练，让学生切实掌握数据的处理和分析方法。此外，本教材的每一章都从一个实际案例出发，提出一个或几个实际问题，使学生带着问题进入每章的学习；同时在章末对案例进行分析，藉此与学生形成互动。

希望这种安排能够激发学生对统计学的兴趣。

第二，本教材在体系安排上力求突破。现有教材的体系安排大体上是一致的，即每章单独介绍一种方法。这样的安排在一定程度上割裂了统计方法与实际问题之间的关系，也割裂了不同的统计方法之间的关系。本教材试图以实际问题为主线，将不同的统计方法组合在一起。例如，第二章将图表显示方法和统计量计算方法放在一起，突出这些方法都是对原始数据的描述性分析。再如，第六章将列联分析、方差分析和相关分析放在一起，突出这些方法都是变量间关系的分析方法。希望通过这样的安排，能够让学生建立起问题与方法之间大体的对应关系，提高他们统计应用的能力。

第三，本教材在体例安排上比较灵活。如前所述，我认为面向非统计学专业学生的统计学教材不能是公式的堆砌，但同时我认为统计学的数理基础对于统计学学习的重要性无论如何强调都不为过。然而，在一部容量有限的教材中，既介绍方法原理与应用，又介绍其数理基础，实难兼顾。为此，本教材在正文之外设立了“专栏”，对一些与正文所介绍内容有关的数理统计的背景知识加以介绍。穿插“专栏”的做法有两方面的好处：一是坚持了统计教学的基本理念；二是适合不同读者群的需求，对那些学有余力的学生来说，适当自学专栏的内容会对他们的学习有所裨益。

此外，还有两点需要说明。

第一，从教学内容来看，本教材包含了在实际分析中非常有用的时间序列 ARMA 模型以及多元统计分析方法，现有同类教材多不涉及这两部分内容。相对于本教材的其他部分而言，ARMA 模型以及多元统计分析方法有一定的难度，老师可以根据教学的具体情况对此进行取舍。

第二，随着计算机技术的迅速发展，现代统计分析已经告别了手工计算的时代，而是多借助于各种专业的统计软件来完成，如 SAS、S-Plus、Stata、R、Eviews 等。毋庸置疑，对于非统计学专业的学生而言，如果能够熟练运用专业的统计软件，就可以大大提高其统计应用的能力和范围。但是，上述软件成本高昂，而且只有在高级的统计分析中才能显示出其优势。对于本教材所介绍的绝大多数统计方法而言，上述软件并不必要，学生只需要熟练掌握 Excel 就足够了。因此，本教材特设立了“软件操作指南”栏目，帮助学生利用 Excel 来分析数据。不过，Excel 没有建立 ARMA 模型以及进行多元统计分析的功能，因此本教材在最后两章介绍了 R 统计软件中的相关操作命令。R 软件是一款免费软件，学生可以在互联网上方便地下载它，然后使用。

由于作者水平所限，无论在体系设置、内容安排还是在行文的具体细节上，本书难免会有一些疏漏，欢迎广大读者提出宝贵意见。

李静萍

2011 年于世纪城时雨园

目 录 Contents

第一章
统计学概述

学习目标

了解统计学的研究对象；

掌握统计学的几个基本概念；

了解统计学的分支；

掌握统计数据的类型；

了解统计数据的收集方法；

理解统计学的正确应用。

引 例

“统计”这个词大家可能并不陌生，翻开报纸杂志，到处充斥着调查数据和用统计方法研究所得到的结论。在 Google 上检索“统计分析表明”一词，有很多结果出现，现摘引若干如下。

(1) 中国人民银行对我国 2002 年 2 月企业商品价格统计分析表明：我国企业商品价格总水平较上月上升 0.1%，较上年同期下降 2.7%，同比降幅较上月缩小 0.3 个百分点。

(2) 中华人民共和国国家知识产权局对我国 19 个省、50 个城市 2005—2007 年授权专利的个人、大专院校、科研单位和企业的专利实施状况进行了调查。统计分析表明：86%已经收回研发投入的成本，71.5%获得的收益大于研发投入。

(3) 全国 30 个省、自治区、直辖市消费者协会统计分析表明：2007 年共受理消费者投诉 656 863 件，为消费者挽回经济损失 83 964 万元。

(4) 2009 年就业蓝皮书《2009 年中国大学生就业报告》的统计分析表明：本科与高职高专的毕业生对雇主的满意程度分别为 70%和 68%，其中“工作要求与压力”不满意度最低，“薪资福利”和“个人发展空间”不满意度最高。

(5) 统计分析表明：在总体的社会化方面，流动儿童优于留守儿童，流动儿童社会化均值分高出留守儿童 1.819 个百分点，差异达到统计显著水平。

(6) 统计分析表明：在父母经常吵架的家庭中，孩子的心理问题检出率为 31.68%，离婚

家庭的为30.30%，和睦家庭的为18.88%。

(7) 气象卫星遥感监测统计分析表明：2010年1月23日前后，渤海全海域及辽东湾、渤海湾、莱州湾海冰面积均达2000年以来同期最大值，其中，莱州湾海冰面积较常年同期平均值偏大约5.6倍。

(8) 统计分析表明：胎次对母猪产仔总数具有影响，其中，1、2胎差异显著($P<0.05$)，1、3胎差异极显著($P<0.01$)，其他各组之间差异均不显著。

……

由上述资料至少可以激发以下四点思考。

第一，这些资料覆盖了经济、社会、科技、心理、气象等各个方面，由此统计应用范围之广可见一斑。

第二，对比不同的资料可以看到，有的统计分析是基于对基层数据的汇总，如资料(2)和资料(3)；有的来自专门组织的调查，如资料(1)、资料(4)、资料(5)和资料(6)；有的来自设备监测，如资料(7)；有的则来自科学实验，如资料(8)。由此可见，统计分析并不受数据来源的局限。

第三，前面七项资料对统计分析结果的陈述并无难理解之处，而第八项资料中则出现了一个符号P，而且资料(5)和资料(8)都用了一个看似普通，实则体现统计思想的术语——显著。

第四，以上资料大都指出“统计分析表明”，那么它们分别使用了哪种统计分析方法呢？

实际上，统计学的思想和方法远比上述资料所显示的深邃，其符号和术语也远比资料中用到的丰富。在统计学的学习起点，不妨先对统计学的基本内容和基本原理作一个剪影式的了解。

第一节 统计学及统计方法的基本思想

一、统计学及其若干基本概念

从字面来看，统计就是统而计之，即将个别数据综合起来得到结论。虽然统计活动是不分国界、人类早已有之的活动，但是对数据进行系统的描述，并在此基础上进行推断的科学原理和操作方法，即作为一个学科门类的统计学，却是一门舶来的学问。《不列颠百科全书》对统计学的定义为：“统计学是收集、分析、表述和解释数据的科学。”这一定义言简意赅，其中蕴涵了以下几层含义。

(1) 统计学的研究对象是数据。

(2) 统计学是一个围绕数据的全过程研究，该过程始于数据的获取，经过对数据的分析，最终从数据中提取信息并得出结论。

(3) 统计学是一门“硬”科学，因为统计方法具有坚实的数理基础。也有人认为统计学兼具艺术性，这是针对统计应用而言的。在分析实际数据时，需要灵活使用统计方法。实际

中，常常会遇到这样的情形：对于同样的数据，不同的分析者使用不同的统计分析方法，得到不同的结论。正是由于统计应用具有灵活性，才需要强调正确应用统计方法，避免对统计方法的误用与滥用。关于这一点，在第三节中有专门的论述。

统计涉及两对基本概念：一是总体和样本，二是参数和统计量。

所谓总体，是指研究所关注的全部单元组成的集合。例如，如果一个研究者希望了解某地区住户的收入，则该地区的全部住户就是总体。如果总体包含的单元很多，受时间和经费等条件的限制，往往不能对总体中的所有单元都进行调查，而只能抽取总体中的一部分单元进行调查。例如，若某地区的住户特别多，逐一进行调查的成本很高，此时可以考虑从中抽取一部分住户进行调查。还有些时候，为了获取所需数据，需要进行破坏性实验，为了减少损失，不必对总体中的每个单元都进行实验，而是抽取一部分单元进行实验。例如，为了确定某人的血铅含量是否超标，只需要抽取他的少量血液进行测量即可。这些从总体中抽取出来的单元所组成的集合就称为样本，而样本中所含单元的数目则称为样本容量。

对总体和样本的另外一种理解是将它们对应于研究所关注的具体特征。例如，在前面的例子中，总体为某地区住户的收入，样本则为抽出来接受调查的住户的收入。

在实际研究中，研究者常常需要了解总体的一些经过汇总后的数据特征，而不是每一个总体单元的具体特征。例如，在前面的例子中，研究者或许希望了解的是某地区全体住户的户均收入，即将全体住户的收入加总之后再除以全部住户数目所得到的数据。这种特征是总体的数学期望，在后面的章节中有专门介绍。同样，对样本也可以计算相应的数据特征。如果计算某地区样本住户的户均收入，则为样本均值。这种对总体数据加工出来的数据特征称为总体参数，对样本数据加工出来的数据特征称为样本统计量。

需要注意的是，研究者真正感兴趣的是总体或总体参数，而不是样本或样本统计量。如果数据只能在样本范围内获得，则对样本数据进行分析（描述统计）之后，研究并没有结束，还需要根据样本的分析结果，在一定的方法支持下对总体或总体参数进行推理和判断（统计推断）。例如，研究者需要依据样本住户的户均收入水平推断总体住户的户均收入水平。

由于样本只是总体的一部分，所以统计推断结果难免会有误差。为了控制推断误差，需要抽取对总体有较强代表性的样本。

二、统计方法的基本思想

统计方法是实证分析中收集和分析数据的重要工具，几乎所有科学都要运用统计方法。但是，在学习和应用统计方法的同时，一定要认识到“统计学不止是一种方法或技术，还含有世界观的成分——它是看待世界上万事万物的一种方法[①]”。

事实上，在统计模型成为科学研究的范式之前，科学界奉行的是一种固化的哲学观，即机械式宇宙观。这种哲学观认为，所有的物体都按照一定的规律运动，所有未来的事件都取

① 陈希孺. 数理统计学简史[M]. 长沙：湖南教育出版社，2002.

决于过去的事件。按照这种观点，用少数几个公式就可以描述现实世界的一切，而且只要有一套完整的公式和足够精确的数据，就可以对未来事件进行预测。

然而，随着科学的发展，人们发现无论是自然科学研究还是社会科学研究，都不可避免地存在误差与不确定性，因此任何对现实世界的描述以及对未来的预测都只能是一种逼近，其所能达到的最好的研究结果只能是给出现实状态与未来可能结果的概率分布，这种概率分布就是统计模型。参照统计模型，人们可以对不确定性有一个定量的把握，并据此作出各种决策。例如，两个企业生产同类产品，其次品率分别为5%和2%。这就是一个简单的概率分布，如表1-1所示。

表1-1 两个企业产品质量的概率分布

企 业	次品率	合格品率
企业1	0.05	0.95
企业2	0.02	0.98

显然，由于生产设备和生产环境的复杂性，没有哪个企业能够确定无疑地总是生产合格品。如果没有产品质量的概率分布，购买者将犹豫不决。然而，通过对比两个企业产品质量的概率分布，购买者很容易作出正确决策。当然，现实世界中的决策远不止这么简单，还要考虑各种因素。例如，在上面的例子中，购买者可能还需要考虑两个企业产品的价格和运输成本等，但毋庸置疑的是，产品质量的概率分布依然是购买决策中的关键信息。

一般来说，概率分布是科学研究中可以获得的最为全面的信息，研究者可以基于概率分布计算出其所感兴趣的任何参数。然而概率分布并不是总能获得的，此时，研究者会退而求其次，转向对关键参数的研究。不过即使是对关键参数的研究，依然离不开概率分布。仍以对某地区住户收入的研究为例，最全面的信息莫过于住户收入的概率分布，掌握了这个概率分布，研究者可以知晓该地区住户收入的各种特征，如中位数收入、贫困住户比重以及收入不均等程度等。如果该概率分布不可得，则研究者可能希望了解该地区住户收入的关键特征，如平均收入，并依据对样本平均收入的概率分布的假定推断出总体的平均收入。

三、统计学的大家族以及统计学与其他学科的关系

（一）统计学的大家族

自17世纪中叶一批数学家对概率的数学理论进行研究以来，经过三个多世纪的发展，统计学已经形成一个建立在数理统计学原理基础之上、集聚各种统计方法的庞大家族。限于篇幅和写作目的，本书不会也不可能涉及所有的数理统计原理和统计分析方法。下面对统计学的大家族进行简单介绍，以使读者更加清楚地认识统计学的功能。

(1) 根据统计分析的阶段不同，统计可以划分为描述统计和推断统计两个分支。相应地，统计学可以分为描述统计学和推断统计学。可以说，描述统计是推断统计的基础，推断统计是描述统计的高级阶段。两者的根本区别在于，描述统计是对确定的样本数据的分析，没有不确定性；而推断统计则是依据样本数据对总体特征进行推断，具有不确定性，需要借助概率这一工具。

描述统计和推断统计的关系如图 1-1 所示。

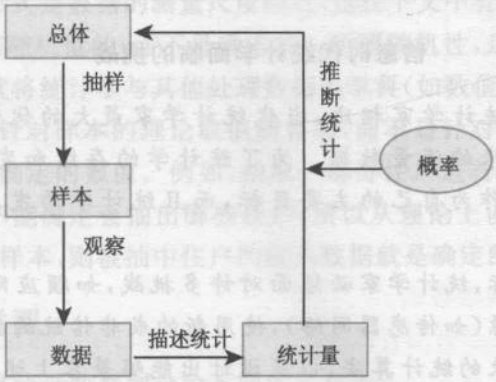

图 1-1　描述统计与推断统计

(2) 根据研究的内容不同，统计学可以划分为理论统计学和应用统计学两个分支。理论统计学在国外又称为数理统计学，其研究内容是统计学的概率，可视为一门纯数学。应用统计学则是在理论统计学的基础上对具体统计方法的研究。在应用统计学中，根据是否假定总体的概率分布只依赖于有限个实参数，又可以分为参数统计方法和非参数统计方法。前者有此假定，而后者则没有这个假定。

(3) 根据对概率和参数的理解不同，统计学可以划分为频率学派和贝叶斯学派两个分支。频率学派理解的概率是频率概念，而贝叶斯学派理解的概率是主观概念；频率学派认为参数是一个客观存在的常数，而贝叶斯学派则认为参数也是一个可以用概率分布刻画的随机变量。

总体来说，本书的内容包括基础的数理统计和常见的应用统计方法，主要介绍频率学派观点，以参数统计方法为主。如果读者对贝叶斯学派和现代非参数统计方法感兴趣，可以参见相关文献。

(二) 统计学与其他学科的关系

统计学在发展过程中，通过将统计方法应用于其他学科的研究，与其他学科保持了密切的联系。“在 20 世纪之前本无‘专职’的数理统计学家，统计学家都是某一专门学科领域的专家，因工作上的需要研究数据分析问题而介入统计学。”①例如，对统计学的发展作出重大贡献的罗纳德·爱尔默·费希尔(R. A. Fisher)，“在遗传学方面的名声不亚于统计方面，他的研究论文不少发表在《优生学杂志》”②。

今天，统计学继续保持着与其他学科交叉发展的关系，而且应用领域更加广泛。一方面，某些学科在解决本领域的数据分析问题时，借助于统计方法，产生出特定的交叉学科。例如，计量经济学、金融计量学、历史计量学、文献计量学等，都是此类学科。另一方面，自然科学和社会科学的各个方面的实证研究都需要应用统计方法。例如，近年来统计分析在生物医学、金融资产定价、质量管理过程等领域的应用均取得了丰硕的成果。

① 陈希孺. 数理统计学简史[M]. 长沙：湖南教育出版社，2002：270.

② 陈希孺. 数理统计学简史[M]. 长沙：湖南教育出版社，2002：270.

专栏 1-1

信息时代统计学面临的挑战①

与费希尔时代的统计学家相比，当代统计学家最大的优势在于掌握了强大的计算技术以及由此带来的海量数据。为了统计学的存续和发展，统计学家必须将解决实际的数据问题作为自己的主要目标，而且统计学的发展必须跟上信息技术的发展。

为了达到这一目标，统计学家必须面对许多挑战，如顺应网格计算趋势、有效利用数据库及其他数据源（如传感器网络），使用新的或非传统的数学结果，发展新的满足通信和计算能力约束的统计算法，以及设计出能够兼容上述努力的新的统计推断范式。

在组织或文化的层面，统计学界也面临诸多挑战。在许多大型的科学计划中，如大气科学（如模型模拟和遥感数据）、天文学（如数字巡天）以及生物学（如基因或脑切片扫描数据库）等，收集和管理着海量的数据。其中虽然可以发现单个统计学家的身影，但单个统计学家很难影响到这些计划的基础。例如，他们难以在收集数据和挖掘大型数据库时的算法选择等问题上拥有发言权。如果统计学想要在信息时代产生必要的影响，就需要进行集体思考，树立统计学界的领导力。

除了统计技术，要想与科学家们合作成功，并说服科学家们承认统计学在科学研究中的重要作用，还需要高超的社交技巧。这些专业外的技巧在跨学科研究中的重要性表明，统计学界需要转变文化。统计学家应当珍视这些非传统技巧，并承认它们在诸如终身教职审查、职称提升以及获奖等方面的作用。

最后，但不是最次要的，统计学家需要将相关的专业知识和社交技巧传授给本专业的研究生和本科生。

第二节 统计数据

一、统计数据的概念和特征

统计学的研究对象为数据。所谓数据，是指对研究对象的某种特征进行测量的结果。要想正确理解统计数据，应当注意统计数据的以下两个特征。

（1）统计数据的表现形式是多样的，既可以是数字，也可以是文字。而且，研究者根据需要，在对同一个对象进行测量时，可以灵活选择数据的形式。例如，对住户的收入进行测

① Bin Yu. Embracing Statistical Challenges in the Information Technology Age[J]. Technometrics. Vol. 49. No. 3. (August. ,2007): pp. 246-247.

量，其结果可以表现为具体的数字，如每月××元；也可以是收入水平的一个文字刻画，如高、中、低。数据的表现形式是数据的测量尺度问题，这在下文中有专门的介绍。

(2) 统计数据是带有随机性的，而不是确定的。所谓随机性，是指数据可以通过某种概率分布规律来描述。这就将统计学与其他处理数据的学科(如数值分析)区别开来。需要注意的是，数据的随机性是针对样本的理论取值而言的，而不是针对样本的具体取值而言的，后者一经测量，便可作为确定的数值。例如，抽取一部分住户进行观察，以研究某地区住户的平均收入。由于事先不能确定会抽出哪些住户，所以从理论上说，样本数据是随机的；然而，一旦抽出一个具体的样本，则被抽中住户的收入数据就是确定的。

二、统计数据的类型

从不同的角度可以将统计数据划分为不同的类型。

1. 根据测量尺度分类

根据数据的测量尺度(即精度)，可以将数据分为定性数据和定量数据。

定性数据是指以文字形式表现的数据。例如，性别的测量结果为男或女，满意度的测量结果为非常满意、比较满意、一般、比较不满意或者非常不满意。根据是否可以对数据进行排序，定性数据又分为两类：不能排序的是定类数据(categorical data，又称nominal data)，可以排序的是定序数据(rank data，又称 ordinal data)。显然，性别数据属于定类数据，而满意度数据则为定序数据。

为了方便数据分析，通常需要将定性数据转化为数字形式，这就是编码(coding)。例如，对性别进行测量时，用 1 表示男性，用 0 表示女性；对满意度进行测量时，用 5 表示非常满意，用 1 表示非常不满意，用 2、3 和 4 依次表示中间的满意程度。编码之后，定性数据就可以参加各种数学运算了。但一定要注意，这种编码数字的本质仍然是文字，不同于真正的数字。例如，性别的编码不能比较大小，相加没有意义；满意度的编码只能比较大小，相加同样没有意义。正因为如此，在进行统计分析时，对定性数据的编码要注意采用正确的方法，否则会出现不符合实际的分析结果。

定量数据是指以数字形式表现的数据。例如，考试成绩的测量结果为××分，收入的测量结果为××元，公路里程的测量结果为××千米，股票价格指数的测量结果为××点等。

显然，从测量精度或者从数据所包含的信息量来讲，定量数据要高(或多)于定性数据，而定序数据又高(或多)于定类数据。此外，测量尺度较高的数据可以转化为测量尺度较低的数据；反之，则行不通。例如，可以将考试成绩的定量数据转化为定序数据(如 98 分为优，65 分为及格等)，反过来却无法判断一个成绩为优的考生究竟考了多少分。因此，在收集数据时，要根据所测量对象的特点，尽量在较高级的尺度上测量，这样可以保留尽可能多的信息，也便于数据类型的转化。

2. 根据数据收集方法分类

根据数据的收集方法，可以将统计数据分为实验数据和观察数据。

实验数据是指在实验之前并不存在，需要通过事先的实验设计，在实验中控制实验对象而收集的数据。例如，在研究杀虫剂的剂量对虫子死亡的概率的影响时，会有严格的实验设

计，在此实验中收集的数据（包括杀虫剂剂量和虫子的生存状态等）就属于实验数据。一般来说，自然科学研究的数据多为实验数据。

观察数据是指客观上已经存在，但是需要经过观察或询问才能获得的数据。这类数据通常需要通过抽样调查来收集。例如，设计一个抽样调查方案，对某地区住户进行抽样调查，调查所收集到的住户信息（包括收入、家庭人口、户主职业等）就属于观察数据。一般来说，社会科学研究的数据多为观察数据。

在统计学中，因果关系推断是一个非常重要的研究目标。需要指出的是，在推断因果关系时，由于实验设计可以更好地控制实验环境和实验条件，从而更好地解决其他因素的干扰，所以实验数据要优于观察数据。

3. 根据数据结构分类

根据数据的结构，可以将数据分为截面数据和历时数据。

截面数据是指在同一时点或同一时期，就多个个体收集的数据。例如，某地区500名住户在2009年的年收入就属于截面数据。

历时数据是指在多个时点或多个时期，就一个或多个个体收集的数据。例如，某企业历年的净利润数据就属于历时数据。如果是对固定的多个个体收集的历时数据，则称为面板数据。例如，2000—2010年，某地区固定的500名住户的相关数据就属于面板数据。

由于因果关系推断的一个基本准则是时间顺序，即因在前、果在后，所以在因果关系推断中，历时数据要优于截面数据，其中面板数据是更为理想的数据结构。

三、统计数据的收集

（一）实验设计

实验设计是指对实验进行科学合理的安排，以达到最好的实验效果。一个科学的实验设计，能够合理地安排各种实验因素，严格地控制实验误差，从而获取有效的实验数据，为统计分析提供支持。

实验设计的基本步骤如下。

（1）随机选择实验对象。

（2）将实验对象随机分为两组：一组接受实验处理，即实验组；另一组不接受任何处理或接受一些对照处理[①]，即对照组。

（3）前测，即在实验前收集两组实验对象的有关数据。

（4）进行实验。

（5）后测，即在实验结束后再次收集两组实验对象的有关数据。

（6）分析前测和后测的数据，得出结论。

实验设计示意图如图1-2所示。

① 常用的对照处理包括安慰剂对照、实验条件对照、标准对照以及历史对照等。

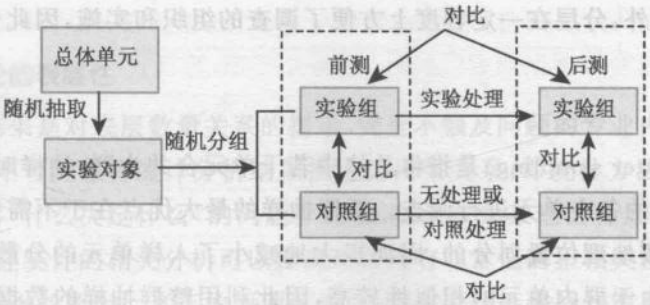

图 1-2　实验设计示意图

例如，要检验某种减肥药是否有效，有 n 名肥胖症患者同意接受实验。按照实验设计的要求，将 n 名患者随机分为两组，并测量每组患者的体重。尽量保证两组患者的前测结果是没有显著差异的。实验组患者服用该减肥药，对照组患者则服用外观相同的安慰剂①。一个疗程结束后，再次测量两组患者的体重。如果在后测中，两组患者出现了显著差异，则可以认为是减肥药产生的效果。

在实验设计中，设置实验组和对照组、将每个实验对象随机分入实验组或对照组、进行两次测量这三个要素非常重要。如果一个实验设计违背了随机化原则或对照原则或者没有进行前测，只能称为准实验设计。利用准实验数据进行分析，会面临更大的困难。

（二）抽样调查

抽样调查是指从总体中抽取一部分单元作为样本，根据对所抽取的样本进行调查，获得对总体的了解。按照抽样时是否遵循随机原则，即总体的每个单元是否都有非零的入样概率，抽样调查可分为概率抽样和非概率抽样两类。其中，概率抽样以随机原则抽取样本，非概率抽样则不然。一般而言，在推断统计中应使用概率抽样收集的数据。下面介绍几种基本的概率抽样方法。

1. 简单随机抽样

简单随机抽样（simple random sampling）是一种最基本的抽样方法，是指从抽样总体②的 N 个单元中随机地、一个一个地抽取 n 个单元作为样本。简单随机抽样具有简单直观、便于统计推断的优点，经典的统计推断理论——假定数据，就来自简单随机抽样。但是该方法要求有总体单元的名单，且入样单元比较分散，因而给调查带来一定的难度。通常大型调查很少直接采用简单随机抽样，而是将这种方法与其他抽样方法结合起来使用。

2. 分层抽样

分层抽样（stratified sampling）是指将抽样单元按照某种特征划分为不同的层，然后从各层中独立、随机地抽取样本。分层抽样保证了样本的结构与总体较为相似，有助于提高统

① 严格的实验设计还要做到令对照组患者不知道自己服用的是安慰剂。

② 抽样总体是指从中抽取样本的总体，与目标总体可能有出入。

计推断的精度。此外，分层在一定程度上方便了调查的组织和实施，因此分层抽样在实践中被广泛应用。

3. 整群抽样

整群抽样(cluster sampling)是指将总体中若干单元合并为群，抽样时直接抽取群，然后对抽中的群所包含的各个单元进行调查。整群抽样的最大优点在于不需要有总体单元的名单，且群通常都是按地理位置划分的，因而极大地减小了入样单元的分散性，有利于调查的组织实施。但是，由于群内单元的相似性较高，因此利用整群抽样的数据进行统计推断时，推断的精度较低。

4. 系统抽样

系统抽样(systematic sampling)是指将总体单元按照一定的顺序排列，先随机抽取一个样本单元，然后按照事先规定好的规则抽取其他样本单元。系统抽样虽然操作简便，但是对统计量方差的估计比较困难，不利于统计推断。

5. 多阶段抽样

两阶段抽样与整群抽样有些相似，第一阶段抽取群，但接下来并不是对群内所有单元进行调查，而是抽取若干个单元进行调查。如果第二阶段继续抽取群，接下来再抽取若干单元进行调查，依此类推，便形成了多阶段抽样(multi-stage sampling)。通常，大型调查大都采用多阶段抽样。

第三节 正确认识和使用统计方法

一、正确认识统计方法

(一) 统计方法的中立性

虽然统计方法广泛应用于各个领域的研究，但它只是一种中立的研究工具。在具体的应用中，统计方法不坚持任何学科中的任何观点。换句话说，统计方法只回答“是什么”的问题，而不回答“应该是什么”的问题。如果有人不同意使用统计方法，完全可以不用它，只作纯粹定性的讨论。但是，只要想进行实证分析，就必须按照统计学的规范来收集和分析数据，由数据来揭示结论。例如，20 世纪六七十年代，美国政界关于对付犯罪的根本途径分歧严重，自由派认为入狱刑罚具有破坏性，不利于改造罪犯、降低犯罪率，有些专家甚至认为某些犯罪行为不应该被看做犯罪。这些观点都只是定性的讨论，现实中入狱刑罚与犯罪之间究竟呈何种关系，需要进行客观的检验。一些犯罪学家的统计研究结果显示，入狱刑罚能够大大降低犯罪率。史蒂夫·莱维特(Steve Levitt)通过研究监狱诉讼的结果发现，从监狱里每释放一名犯人，每年的犯罪数量就会增加 15 起。这些实证研究的结论使得美国政界在审判政策和政策实施方面达成了共识，自由派不再像 20 世纪六七十年代那样抵制入狱刑