

21世纪 经济学教材

# 统计学

tatistics

金 秀 于春海 主 编



清华大学出版社

21世纪 经济学教材

# 统计学

tatistics

金 秀 于春海 主 编

清华大学出版社  
北 京

## 内 容 简 介

统计是认识客观世界数量规律的有力工具,统计学是高等学校经济管理类专业核心课程。考虑到理论的完整性,同时以经济管理实践需要为前提,本书精心选择了十二章内容,包括绪论、数据的收集、整理与显示、数据分布特征的描述、概率与概率分布、抽样分布、参数估计、假设检验、方差分析、相关与回归分析、列联表分析和对数线性模型、时间序列分析与预测、指数。为了让学习者更好地掌握知识点,每章后有本章小结和习题,并在书后附有参考答案。统计分析离不开统计软件的使用,本书各章所涉及的数据分析例题的计算都配备了相应的 Excel 软件或 SPSS 软件使用方法的解释说明。

本书适合统计学及其他财经类各专业本科生使用,也可供统计相关工作人员学习参考使用。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

统计学/金秀,于春海主编. —北京:清华大学出版社,2014  
21 世纪经济学教材

ISBN 978-7-302-34520-6

I. ①统… II. ①金… ②于… III. ①统计学-高等学校-教材 IV. ①C8

中国版本图书馆 CIP 数据核字(2013)第 276968 号

责任编辑:王文珠

封面设计:刘超

版式设计:文森时代

责任校对:马军令

责任印制:刘海龙

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦 A 座 邮 编:100084

社总机:010-62770175

邮 购:010-62786544

投稿与读者服务:010-62776969, [c-service@tup.tsinghua.edu.cn](mailto:c-service@tup.tsinghua.edu.cn)

质 量 反 馈:010-62772015, [zhiliang@tup.tsinghua.edu.cn](mailto:zhiliang@tup.tsinghua.edu.cn)

课 件 下 载: <http://www.tup.com.cn>, 010-62788951-223

印 刷 者:清华大学印刷厂

装 订 者:三河市兴旺装订有限公司

经 销:全国新华书店

开 本:185mm×230mm

印 张:25.5

字 数:555 千字

版 次:2014 年 3 月第 1 版

印 次:2014 年 3 月第 1 次印刷

印 数:1~4000

定 价:42.00 元

产品编号:052562-01

# 前 言

“统计学”课程是国家教育部批准、经济学教学指导委员会和工商管理类教学指导委员会讨论通过的经济学和工商管理类各类专业的核心课程之一。统计学是一门研究随机现象数量规律性的方法论科学，无论是国民经济管理和公司、企业的经营及决策，还是科学研究都越来越依赖于数量分析和统计分析方法，统计方法已经成为理、工、农、医、人文、社会、管理、军事等所有学科领域科学研究的基本方法。特别是经济管理类专业人员，统计学更是他们日常预测决策、分析问题所使用的主要手段之一。当今时代，统计思维应像读与写一样成为一个有效率公民的必备能力，而通俗易懂、精简而实用的集理论方法及实践应用需要并结合统计分析软件使用的教材是经济管理类专业的学生学好、用好统计学的必要条件之一。本教材正是基于此背景充分吸收国内外经典教材的优点而编撰，它适合作为统计学及其他财经类各专业本科生学习统计学课程的教材，也可供统计学领域的社会经济工作者参考。为了使学习者更好地掌握知识点，在各章都配有本章小结与习题并附有参考答案。此外，本书各章所涉及的数据分析例题的计算都配备了相应 Excel 软件或 SPSS 软件使用方法的解释说明。

统计应该是人人都能够理解和掌握的有用的数据分析方法，为此，本书主要特色体现在：

(1) 从使用者的角度出发，略去了大量的数学推导，每章的主要部分用通俗的语言引入和解释一些概念，并从经济管理的实践问题入手，将统计学的知识适时地渗透于实际案例中，便于学生对相关概念、理论与方法的理解和运用，使统计学知识生动易懂。

(2) 在内容安排上，将理论的完整性与实践应用的普遍性相结合，如在第九章“相关与回归分析”特别增加了以往很多同类教材忽略的包含分类变量情形下的回归分析方法介绍；其次，数据结构决定了统计方法的合理选择，在现实经济管理活动中存在大量的由分类变量整理得到的频次分布结果，而列联表分析和对数线性模型等正是针对该种数据结构有效地分析现象之间关系的方法，为此，本教材将此部分内容纳入其中。

(3) 现实中，面对大量的统计数据，统计分析实践必须依赖于计算机软件来帮助计算，而统计软件的正确选择与使用是统计分析方法能否在经济管理中发挥作用的关键之一。本教材选用了当今被广泛使用且功能越来越强大的 Excel 软件（包括在其环境下运行的插件 PHStat2）或 SPSS 软件作为各章统计方法实现数据分析的工具，并结合软件计算输出的结果给予必要的解释说明。

在本书的写作过程中，参考了有关学者的论著，得到了同行和朋友的支持与帮助。首先，感谢我们的研究生于晶贤、李想、王佳、刘家和、方钦榕、姜光华及邹吉娥等，他们参与了本书的一些章节的编写、排版和讨论。感谢清华大学出版社的编辑在这本书编写过程中提出的很多宝贵意见，为本书的顺利出版做了很多工作。

书中难免有纰漏之处，敬请读者批评指正。

# 目 录

|                        |    |
|------------------------|----|
| 第一章 绪论.....            | 1  |
| 第一节 统计学的研究方法和作用 .....  | 1  |
| 第二节 统计数据的计量尺度和类型 ..... | 4  |
| 第三节 统计学的基本概念 .....     | 8  |
| 本章小结 .....             | 11 |
| 本章习题 .....             | 12 |
| 第二章 数据的收集、整理与显示.....   | 14 |
| 第一节 统计数据的收集 .....      | 14 |
| 第二节 统计数据的整理 .....      | 20 |
| 第三节 统计数据的显示 .....      | 28 |
| 本章小结 .....             | 37 |
| 本章习题 .....             | 37 |
| 第三章 数据分布特征的描述.....     | 39 |
| 第一节 集中趋势的测度 .....      | 39 |
| 第二节 离散程度的测度 .....      | 52 |
| 第三节 偏度和峰度的测度 .....     | 59 |
| 本章小结 .....             | 62 |
| 本章习题 .....             | 62 |
| 第四章 概率与概率分布 .....      | 67 |
| 第一节 基本概念 .....         | 67 |
| 第二节 概率的定义及基本运算法则 ..... | 69 |
| 第三节 条件概率与事件的独立性 .....  | 72 |
| 第四节 全概率公式与贝叶斯定理 .....  | 74 |
| 第五节 随机变量及其概率分布 .....   | 77 |
| 第六节 常用的随机变量的概率分布 ..... | 80 |
| 本章小结 .....             | 92 |

|                          |            |
|--------------------------|------------|
| 本章习题 .....               | 92         |
| <b>第五章 抽样分布</b> .....    | <b>96</b>  |
| 第一节 抽样及抽样调查的相关概念 .....   | 96         |
| 第二节 几个重要参数的抽样分布 .....    | 100        |
| 第三节 一些重要的结论 .....        | 110        |
| 第四节 几种常用的抽样组织方式 .....    | 113        |
| 本章小结 .....               | 122        |
| 本章习题 .....               | 122        |
| <b>第六章 参数估计</b> .....    | <b>125</b> |
| 第一节 总体参数估计概述 .....       | 125        |
| 第二节 总体平均数和比例的区间估计 .....  | 127        |
| 第三节 两个总体平均数之差的区间估计 ..... | 133        |
| 第四节 两个总体比例之差的区间估计 .....  | 138        |
| 第五节 样本容量的确定 .....        | 139        |
| 第六节 总体方差和方差之比的区间估计 ..... | 141        |
| 本章小结 .....               | 146        |
| 本章习题 .....               | 147        |
| <b>第七章 假设检验</b> .....    | <b>150</b> |
| 第一节 假设检验概述 .....         | 150        |
| 第二节 总体均值和比例的假设检验 .....   | 156        |
| 第三节 两个总体均值之差的假设检验 .....  | 166        |
| 第四节 两个总体比例之差的假设检验 .....  | 171        |
| 第五节 总体方差及方差之比的假设检验 ..... | 172        |
| 本章小结 .....               | 178        |
| 本章习题 .....               | 178        |
| <b>第八章 方差分析</b> .....    | <b>182</b> |
| 第一节 方差分析的基本问题 .....      | 182        |
| 第二节 单因素方差分析 .....        | 185        |
| 第三节 方差分析中的多重比较 .....     | 191        |
| 第四节 双因素方差分析 .....        | 193        |
| 本章小结 .....               | 201        |
| 本章习题 .....               | 201        |

|                           |     |
|---------------------------|-----|
| 第九章 相关与回归分析 .....         | 204 |
| 第一节 相关与回归分析的基本概念 .....    | 204 |
| 第二节 简单线性相关分析 .....        | 206 |
| 第三节 一元线性回归分析 .....        | 210 |
| 第四节 一元线性回归模型预测 .....      | 221 |
| 第五节 多元线性回归分析 .....        | 226 |
| 第六节 多元线性回归模型中自变量的选择 ..... | 236 |
| 第七节 非线性回归分析 .....         | 246 |
| 第八节 定性自变量的回归分析 .....      | 251 |
| 本章小结 .....                | 254 |
| 本章习题 .....                | 254 |
| 第十章 列联表分析和对数线性模型 .....    | 258 |
| 第一节 列联表分析 .....           | 258 |
| 第二节 对数线性模型 .....          | 269 |
| 本章小结 .....                | 274 |
| 本章习题 .....                | 275 |
| 第十一章 时间序列分析与预测 .....      | 277 |
| 第一节 时间序列的基本问题 .....       | 277 |
| 第二节 时间序列的水平分析 .....       | 280 |
| 第三节 时间序列的速度分析 .....       | 283 |
| 第四节 时间序列的构成分析 .....       | 286 |
| 第五节 时间序列的预测方法 .....       | 296 |
| 本章小结 .....                | 299 |
| 本章习题 .....                | 300 |
| 第十二章 指数 .....             | 302 |
| 第一节 指数的基本问题 .....         | 302 |
| 第二节 总指数的编制 .....          | 305 |
| 第三节 指数体系与因素分析 .....       | 312 |
| 第四节 几种常用的价格指数 .....       | 316 |
| 本章小结 .....                | 321 |
| 本章习题 .....                | 322 |

|                        |     |
|------------------------|-----|
| 参考文献.....              | 325 |
| 参考答案.....              | 326 |
| 附录 1 标准正态分布表.....      | 359 |
| 附录 2 $\chi^2$ 分布表..... | 361 |
| 附录 3 $t$ 分布表.....      | 365 |
| 附录 4 $F$ 分布表.....      | 369 |

# 第一章 绪 论

统计学是一门处理数据中变异性的科学与艺术，内容包括收集、分析、解释和表达数据，目的是求得可靠的结果。在诺贝尔经济学奖获得者中，2/3 以上的研究成果与统计和定量分析有关。因此，著名经济学家萨缪尔森在其教科书《经济学》中提到：“在许多与经济学有关的科学中，统计学是特别重要的”。在生活中，统计学同样无处不在。例如，吸烟是对健康有害的，吸烟的男性减少寿命 2 250 天；不结婚的男性会减少寿命 3 500 天，不结婚的女性会减少寿命 1 600 天；身体超重 30%，会使寿命减少 1 300 天。从上面的例子可以看出，统计学对我们非常重要。本章主要介绍统计学的基本问题、统计数据的计量尺度和类型，以及其基本概念。

## 第一节 统计学的研究方法和作用

### 一、统计的含义

统计作为一种社会实践活动，已有悠久的历史，据历史记载，我国在西周时期就已建立了统计报告制度。在英文中，统计为 statistics，它与“国家”一词来自同一词根。可以说，自从有了国家，就有了统计实践活动。最初，统计只是作为一种计数活动，为满足统计者管理的需要而收集资料，通过统计技术弄清国家的人力、物力和财力，作为国家管理的依据。

今天，“统计”一词在日常生活、社会实践活动和科学研究领域中经常出现，但人们对“统计”一词常常有不同的理解。例如，企业每年要“统计”产品的产量和产值，这是将统计作为一项工作来看待；人们了解股票的交易状况要“统计”成交额和股票指数等相关信息，这是将统计作为数据资料来运用；教师向学生传授“统计”的知识和方法，则是将统计作为一门科学来看待。故“统计”通常有三种含义：统计工作、统计数据和统计学。

#### （一）统计工作

统计工作是指收集、整理、分析和提供统计数据的活动过程。它包括统计设计、统计调查、统计整理、统计分析、统计预测和统计决策六个阶段。

### （二）统计数据

统计数据是指统计工作中收集到的各种数字资料和相关其他资料。它既是统计工作的成果，也是进一步进行统计分析的材料。

### （三）统计学

统计学是随着人类社会的进步和社会管理的需要而逐步发展起来的。它是一门收集、整理和分析统计数据的方法论科学，其目的是探索数据的内在数量规律性，以达到对客观事物的科学认识。由统计学的含义可知，统计学主要包括以下三个方面的内容。

#### 1. 统计数据的收集

统计数据的收集是指取得统计数据的过程。它是进行统计分析的基础，离开了统计数据，统计方法就失去了根本。

#### 2. 统计数据的整理

统计数据的整理是指对统计数据加工整理的过程。它是介于数据收集和数据分析之间的一个必要环节，是统计学研究的重要内容之一。

#### 3. 统计数据的分析

统计数据的分析是统计学的核心内容，它是通过统计描述和统计推断的方法探索数据内在规律的过程。

通过对“统计”三种含义的解释，可以看出：统计工作是统计实践活动；统计数据是统计工作实践的成果；统计学是统计工作的理论概括，同时又反过来指导统计工作。统计学和统计工作的关系是理论与实践的关系。

## 二、统计学的研究方法

统计研究过程的起点是统计数据，终点是探索客观现象内在的数量规律性，而贯穿于这一过程的统计研究方法主要有大量观察法、统计分组法、综合分析法、统计推断法和统计模型法。

### （一）大量观察法

统计学研究社会经济现象的数量及其变化过程，要从总体上进行考察。对总体中的全部或足够多的单位进行观察和综合分析的方法，称为大量观察法。社会经济现象的总体是复杂的，是受多种因素影响的，而且总体各单位特征和数量表现也有很大差异，因此只有对足够多的单位进行观察，其观察值的平均结果才会趋于稳定，建立在大量观察法基础上的统计数据才会给出具有普遍意义的结论。统计学中的各种观测调查方法都属于大量观察法。

### （二）统计分组法

由于所研究现象本身的复杂性、差异性和多层次性，需要对研究现象进行分组或分类研究，用以区分研究现象的类型和形式，并在同质的基础上研究不同组或不同类之间的差异性。统计分组法在整个统计研究过程中占有重要的地位，在统计调查前必须对研究现象进行科学分组，才能确定调查对象和范围，才能搜集必要的统计资料。在统计整理环节也必须对原始资料进行分组加工整理。统计分组法有传统分组法、判别分析法和聚类分析法等。

### （三）综合分析法

综合是指对大量观察所获得的资料，运用各种综合指标来综合反映总体的一般数量特征。综合指标概括地描述了总体的综合数量特征和变动趋势。常用的综合指标有总量指标、相对指标、平均指标、离散程度指标和统计指数等。分析是指对综合指标进行分解和对比分析，以研究总体的差异和数量关系。在综合分析中，常用的统计分析方法有动态趋势分析法、因素影响分析法、相关与回归分析法、综合平衡分析法等。

### （四）统计推断法

以一定的置信标准，根据样本数据来推断总体数量特征的归纳推理方法，称为统计推断法。在统计研究过程中，我们所观察的往往只是研究现象总体中的一部分个体，掌握的只是具有随机性的样本观测数据，而认识总体数量特征才是统计研究的目的所在，这就需要我们采用统计推断法，由样本观察数据来推断总体数量特征。常用的统计推断法有重点调查、典型调查、抽样推断、统计预测和决策等。

### （五）统计模型法

除了使用统计指标来反映研究现象总体数量特征外，还需要对相关现象间的数量变动关系进行定量研究，以了解某一现象数量变动与另一现象数量变动间的关系及变动的影响程度。在研究这种数量变动关系时，需要根据具体的研究对象和一定的假定条件，用合适的数学方程来进行模拟，这种方法叫做统计模型法。统计模型的三个基本要素是变量、数学方程和模型参数。运用统计模型法，可以使统计分析更具广度和深度，提高统计的认识能力。

上述各种方法之间是相互联系、互相配合的，共同组成了统计学研究方法的体系。

## 三、统计学的作用

统计学作为一门应用性很强的学科，在认识事物、指导生产、经济管理和科学研究等各个方面都发挥着重要作用。

（1）统计学为认识自然和社会提供了必需的方法和途径。人们要认识自然和社会，就

离不开各种各样的数据资料，并需要对这些资料作出各种各样的分析研究，这就需要运用一整套的统计理论和方法作为指导。例如，要研究人口年龄分布的特征、经济运行的规律以及生活习惯与某些疾病间的关系等问题，统计数据及以此为基础的分析研究是必不可少的。如果说统计是认识自然和社会的手段，那么统计学则提供了理论和方法。

(2) 统计学在指导生产活动过程中发挥着重要作用。企业的生产任务是以尽量少的投入生产出数量多、质量好的产品，然而影响产品产量和质量的因素很多，这就需要对各种因素进行实验和观察，了解各种因素的影响方式和程度，从而确定企业的最佳生产条件和方式并使之始终处于科学的控制之中。这个过程就需要应用统计理论和方法，科学、合理地设计和安排实验并作出分析，同时对生产过程不断进行监测。最佳生产方案设计和最优质量控制是统计学的一大应用领域。

(3) 统计学在社会经济管理活动中的作用更为显著。无论从宏观角度还是微观角度上看，统计活动都具备收集信息、提供咨询、实施监督和支持决策的重要职能。在宏观上，政府的管理目标是保持国民经济持续稳定协调发展，实现劳动力的充分就业和人民生活水平的稳定提高。这就需要政府利用统计学提供的方法，科学、合理地收集数据资料，对国民经济和社会发展状况进行跟踪监测与预警，对各种社会经济问题进行定量模拟和分析，从而为制定和调整政策提供依据。在微观上，企业的管理目标是使生产要素达到最佳配置，取得最佳的经济效益，在激烈的市场竞争中保持优势。这就需要企业及时收集市场信息，科学地反映和分析企业的经营情况，准确地预测和判断市场变化的趋势，而这些都是必须运用各种统计理论和方法。

(4) 统计学为科学研究提供了有力手段。在科学研究中，其研究的任务是揭示客观事物的规律性，研究的方法一般是先根据若干观察或实验资料提出假设，然后通过各种途径进行观察或实验加以验证。显然，统计理论和方法在其中起着重要作用。一方面，它有助于集中并提取观察和实验数据中本质性的东西，从而有助于提出较正确的假设；另一方面，它又能知道研究人员如何去安排进一步的观察或实验，以判定所提出假设的正确性。

## 第二节 统计数据的计量尺度和类型

### 一、统计数据的计量尺度

统计数据是对客观现象进行计量的结果，因此，在收集数据之前要先对现象进行计量或测度，这就涉及计量尺度问题了。对于不同的事物，能够予以计量或测度的程度是不同的，有些事物只能对其属性进行分类，如人口的性别和文化程度、产品的型号及质量等级等；有些则可以用比较精确的数字加以计量，如物体的长度、产品的重量和价值等。显然，从对事

物计量的精确程度来看,采用数字计量要比采用分类计量更精确一些。根据计量学的一般分类方法,按照对事物计量的精确程度,可以将统计数据的计量尺度从低级到高级、由粗略到精确划分为四个层次,即定类尺度、定序尺度、定距尺度和定比尺度。

### (一) 定类尺度

定类尺度又称类别尺度或列名尺度,是按照某种属性对客观事物进行平行分类或分组的一种测度,它是最粗略、计量层次最低的计量尺度。例如,人口按照性别分为男、女两类;企业按照经济性质分为国有、集体、私营、合资、独资企业等。这些分类就是将所有观察的个体区分为不同的类型,属于定类尺度的测度。定类尺度只是测度了事物之间的类别差,对各类别之间的其他差别却没有反映。因此,使用该尺度对事物所作的分类,各类别之间是平等的并列关系,无法区分优劣或大小,各类别之间的顺序是可以改变的。由于定类尺度只能区分事物是同类或不同类,因此,它具有“=”与“ $\neq$ ”的数学特性。

定类尺度的值是以文字表述的,可以用数值标识,但仅起标签的作用。例如,在人口统计中按性别分组,并用数字作为代号,如用“1”表示男性人口、“2”表示女性人口等。定类尺度可用于计算各组数值占总体数值的比重和众数等,但不能对各类编号进行加减乘除计算。对定类尺度的计量结果,常常通过计算出每一类别中各元素或个体出现的频数或频率来进行分析。

定类尺度的主要特征表现为:只能区分事物的类别,但无法比较类别间的大小,各类别的顺序可以改变;对事物的区分必须符合穷尽和互斥的要求;对其进行分析的统计量主要是频数或频率。定类尺度是对事物最基本的测度,是其他计量尺度的基础,主要用于非参数的统计判断。

### (二) 定序尺度

定序尺度又称顺序尺度,是把各类事物按一定特征的大小、高低、强弱等顺序排列起来,构成定序数据,它是对事物之间的等级差别或顺序差别的一种测度。定序尺度的计量结果虽然也表现为类别,但这些类别之间是可以比较顺序的。例如,产品等级就是对产品质量好坏的一种次序测度,可以将产品分为一等品、二等品、三等品及次品等;考试成绩可以分为优、良、中、及格、不及格等;一个人的受教育水平可以分为小学、初中、高中、大学、大学以上等;一个人对某一事物的态度可以分为非常同意、同意、保持中立、不同意、非常不同意等。显然,定序尺度的计量精度要优于定类尺度。它不仅可以测度类别差,还可以测度次序差,并可以比较大小,但其序号仍不能进行加减乘除计算。与定类尺度不同,定序尺度中各类别的顺序不能改变。因此,定序尺度具有“ $>$ ”和“ $<$ ”的数学特性,当然,也包括了定类尺度的特性。

定序尺度除了可用来计量比重(频率)外,还可用来进行累计频数(率)、中位数等数值的计算。

### （三）定距尺度

定距尺度又称间隔尺度，是对事物类别或次序之间间距进行的一种测度。它是一种较定类尺度和定序尺度更为高级和精确的计量尺度。它不仅能将事物区分为不同的类型并进行排序，而且可以准确地指出类别之间的差距是多少。定距尺度的计量结果为数值，一般使用自然或度量衡单位作为计量尺度，如收入用人民币元度量、考试成绩用百分制度量、温度用摄氏度或华氏度度量、重量用克度量、长度用米度量等。定距尺度的每一间隔都是相等的，只要给出一个度量单位，就可以准确地指出两个计数之间的差值。例如，考试成绩每分之间的间隔是相等的，80分与90分的差距等同于80分与70分的差距。因此，定距尺度不仅具有定类尺度和定序尺度的特性，而且其计量结果可以进行加、减运算。

定距尺度的主要特征体现为：不仅能区分事物的类别、进行排序、比较大小，而且可以精确地计量大小的差异，即可以进行加减运算，但不能计算乘除；没有绝对零点，即可以设定任意一个点为零起点。这里的“零”表示一个数值，即“0”水平，而不表示“没有”或“不存在”，如温度为0度，成绩为0分等。

### （四）定比尺度

定比尺度又称比率尺度，是对事物之间比值的一种测度，可用于参数与非参数统计推断。它与定距尺度属于同一层次，一般可不作区分，其计量结果也表示为数值。它除了具有上述三种计量尺度的全部特性外，还具有一个特性，即可以计算两个测度值之间的比值。这就要求定比尺度中必须有一个绝对固定的“零点”，这也是它与定距尺度的差别。在定比尺度中，“0”表示“没有”或“不存在”。例如，一个人的收入为0，表示这个人没有收入；一种产品的产量为0，表示没有这种产品等。因此，采用定比尺度计量的结果通常不会出现“0”值。在现实生活中，大多数情况下使用的都是定比尺度。

对于定距尺度和定比尺度的区别，有些教材中形象地将定距尺度比喻成从桌面上开始测量高度，定比尺度则是从地面上开始测量高度。在定比尺度中，由于“0”不存在，因而其数值不仅可以比较大小、计算差值，还可以计算数值之间的比值。例如，甲的月工资收入为600元，乙的月工资收入为300元，可以得出甲的收入是乙的收入的2倍。而在定距尺度中，由于不存在绝对零点，因此只能比较数值差，而不能计算比值。例如，可以说30℃和15℃之差为15℃，而不能说30℃比15℃热1倍。可见，定比尺度除了能区分类别、排序、比较大小、求出大小差异、采用加减运算外，还可以进行乘除运算。

在计量尺度的应用中，需要注意的是，同类事物用不同的尺度计量会得到不同的尺度数据。例如，农民收入数据按实际填写就是定距尺度；按高、中、低收入水平分类就是定序尺度；按有无收入计量则是定类尺度；而说某人的收入是另一人的2倍，则是定比尺度。一般地，因为研究目的和内容的不同，计量尺度也会不同。如果不担心损失信息量，就可以降低计量层次。表1-1给出了统计数据计量尺度的特征、数学特性及运算功能。

表 1-1 统计数据的计量尺度

| 计 量 尺 度 | 特 征   | 数 学 特 性  | 运 算 功 能 |
|---------|-------|----------|---------|
| 定类尺度    | 分类    | =        | 计数      |
|         | 分组    | $\neq$   |         |
| 定序尺度    | 分类    | >        | 计数      |
|         | 排序    | <        | 排序      |
| 定距尺度    | 分类    | +        | 计数      |
|         | 排序    | -        | 排序      |
|         | 量的差距  |          | 加减      |
| 定比尺度    | 分类    | +        | 计数      |
|         | 排序    | -        | 排序      |
|         | 量的差距  | $\times$ | 加减      |
|         | 有绝对零点 | $\div$   | 乘除      |
|         |       |          |         |

在统计分析中，一般要求测量的层次越高越好，因为高层次的计量尺度包含更多的数学特性，所运用的统计分析方法越多，分析时也就越方便，因此应该尽可能使用高层次的计量尺度。

## 二、统计数据的类型

区分统计数据的类型是十分重要的，因为对于不同类型的统计数据，需要采用不同的统计方法进行收集、处理和分析。

### （一）直接数据和间接数据

根据获取统计数据的途径不同，可以将数据分为直接数据和间接数据。研究者直接对研究对象进行调查、观测和实验所获得的数据称为直接数据，也称为原始数据。研究者所采用的他人收集、整理和加工过的数据称为间接数据，也称次级数据。例如，人们对政府统计数据的应用就属于间接数据的利用。

### （二）品质数据和数量数据

根据数据的表现和用途不同，可以将数据分为品质数据和数量数据。品质数据，又称定性数据，是用于鉴别每一个个体品质的标记或名称，反映事物的品质特征。品质数据有数值型数据和非数值型数据两种。品质数据一般是非数值型数据，并且用文字表现，但为了便于计算机处理，有时也对品质数据赋值，用数值作为标签来表现。品质数据是由定类尺度和定序尺度计量所形成的数据，其中，定类尺度和定序尺度相当于通常所说的品质标志。

数量数据，又称定量数据，用于反映事物的数量特征，通常表现为数值，这种数值可以进行算术计算。数量数据是由定距尺度和定比尺度计量所形成的数据，其中，定距尺度和定比尺度相当于通常所说的数量标志。

### （三）横截面数据和时间序列数据

根据对客观现象观察的角度不同，可以将数据分为横截面数据和时间序列数据。横截面数据，又称静态数据，是指在同一时间对同一总体内不同单位的数量进行观测而获得的数据。时间序列数据，又称动态数据，是指在不同时间对同一总体的数量进行观测而获得的数据。例如，2006年全国各省、市、自治区的地区生产总值就是横截面数据，而“九五”期间我国历年的国内生产总值就是时间序列数据。

### （四）观察数据和实验数据

根据数据的收集方法不同，可以将数据分为观察数据和实验数据。观察数据是指在对现象进行人为控制的条件下，通过调查或观测而得到的数据，如有关社会经济的统计数据基本上是观察数据。而在实验中通过控制实验对象而收集到的数据便称为实验数据，如新药实验，医学、卫生以及自然科学的大多数数据都是实验数据。

## 第三节 统计学的基本概念

在开始学习统计学时，需要理解几个重要的概念，它们对统计学的学习和理解起到至关重要的作用。

### 一、总体、单位与样本

总体、单位与样本是统计学中最基本的一组概念，它们贯穿于统计学内容的始终。

#### （一）总体

总体，是统计总体的简称，是指根据一定目的确定的所要研究的对象的全体。它是由客观存在的、具有某种共同性质的若干个个体构成的集合。例如，在研究某市乡镇企业的发展状况时，该市的全部乡镇企业构成的集合是统计总体。

统计总体具有三个特性：同质性、变异性和大量性。同质性是指构成一个总体的所有个别事物至少在某一点上具有相同的性质。它是确定统计总体的基本标准，并根据统计的研究目的而定。例如，在研究城镇居民户的生活状况时，所有城镇居民户构成了统计总体，因为