

全国普通高等学校体育专业统编教材

体育统计方法与实例

◆ 雷福民 主编

**Sports
Statistics**

体育统计是运用统计学的原理和方法，
通过对体育教学、运动训练、
体育科研和体育管理中随机现象的描述、
推断和分析，揭示其数量规律的一门应用学科。

高等教育出版社

全国普通高等学校体育专业统编教材

体育统计方法与实例

主 编 雷福民

副主编 权德庆

郑 凯

唐东辉

高等教育出版社·北京

内容简介

本书以深入浅出的方式，紧密结合体育实例，经过全国体育统计学会多次研讨，由 15 所体育院校中具有丰富教学经验的专业教师共同编写完成。

本书结构合理，脉络清晰，逻辑性强，注重培养学生的应用能力，精选了大量体育科研、教学、训练、管理、选材中的应用实例。主要包括：统计数据的收集与整理、统计描述、概率分布及应用、参数估计与假设检验、方差分析、相对数与动态分析、相关与回归分析、多元线性回归分析、聚类分析、因子分析、各种方法的 SPSS 例解、34 个体育应用实例，以及体育统计中常见问题及解决策略等。

本书可作为体育学类各专业本科生与研究生教材，也可作为体育工作者的参考书。通过学习，可以帮助读者获得处理和分析体育数据的能力，建立数据驱动的统计思维模式。

图书在版编目（C I P）数据

体育统计方法与实例 / 雷福民主编. -- 北京：高等教育出版社，2017. 7
ISBN 978-7-04-047900-3
I. ①体… II. ①雷… III. ①体育统计-高等学校-教材 IV. ①G80-32

中国版本图书馆 CIP 数据核字（2017）第 130903 号

体育统计方法与实例
Tiyu Tongji Fangfa yu Shili

策划编辑 范 峰 责任编辑 范 峰 封面设计 姜 磊 版式设计 范晓红
责任校对 胡美萍 责任印制 赵义民

出版发行	高等教育出版社	网 址	http://www.hep.edu.cn
社 址	北京市西城区德外大街 4 号		http://www.hep.com.cn
邮政编码	100120	网上订购	http://www.hepmall.com.cn
印 刷	中国农业出版社印刷厂		http://www.hepmall.com
开 本	787 mm×960 mm 1/16		http://www.hepmall.cn
印 张	21.5	版 次	2017 年 7 月第 1 版
字 数	510 千字	印 次	2017 年 7 月第 1 次印刷
购书热线	010-58581118	定 价	36.00 元
咨询电话	400-810-0598		

本书如有缺页、倒页、脱页等质量问题，请到所购图书销售部门联系调换
版权所有 侵权必究
物 料 号 47900-00

编 委 会

主 编 雷福民
副主编 权德庆 郑 凯 唐东辉
编 委 (按章节编写顺序)

陈建华 (集美大学)
雷福民 (西安体育学院)
郑 凯 (沈阳体育学院)
谭延敏 (中南民族大学)
文 静 (杭州师范大学)
陈培友 (南京师范大学)
王秋茸 (西安体育学院)
夏成生 (成都体育学院)
唐东辉 (北京师范大学)
夏敏慧 (海南师范大学)
何国民 (武汉体育学院)
李世明 (中国海洋大学)
李旭芝 (西安体育学院)
黄德春 (安徽师范大学)
赵书祥 (北京体育大学)
隗金水 (广州体育学院)
魏德样 (福建师范大学)
史 进 (西安体育学院)
李伟平 (西安体育学院)

前言

随着信息技术的快速发展,电子商务、社交网络、云计算等新兴应用已经渗透到人们的日常生活中,人类已跨入大数据时代。面对大数据带来的挑战,统计学是一把获得数据分析能力的钥匙。

体育统计是体育学科与统计学科的相互融合。将实践性较强的体育学和理论性较强的统计学完美结合,找到理论与实践的结合点,可以使学生和体育工作者少花时间和精力,且学到统计学的精髓,从而在体育科学研究中发挥统计学的作用。

本书主要内容包括统计数据的收集与整理、统计描述、概率分布及应用、参数估计与假设检验、方差分析、相对数与动态分析、相关与回归分析、多元线性回归分析、聚类分析、因子分析以及体育统计常见问题及解决策略等。各章后附有大量练习题,并配有参考答案。

本书精选了大量体育科研、教学、训练、管理和选材中的应用实例,注重学生应用能力的培养,重点讲解应该选用何种统计方法解决体育科研中的实际问题,如何使用SPSS18.0统计软件,怎样解释统计结果。书中实例涵盖了调查设计、描述统计、频数分析、统计图表、体育评分、正态检验、均值比较、方差分析、动态分析、列联表分析、相关分析、回归分析、聚类分析、因子分析等。

本书由雷福民担任主编,各章编写分工如下:第一章(陈建华、雷福民)、第二章(郑凯)、第三章(谭延敏、雷福民、文静)、第四章(陈培友)、第五章(夏成生、唐东辉)、第六章(夏敏慧、何国民)、第七章(李世明、雷福民、李旭芝)、第八章(黄德春)、第九章(赵书祥)、第十章(隗金水)、第十一章(魏德样)、第十二章(雷福民)、各章实例(雷福民、郑凯、夏成生、唐东辉、王秋茸、李世明、史进、李伟平)。全书最终由雷福民统稿。

编写一本包含大量体育统计应用案例的教材,对于我们来说也是一种尝试。因此,书中不当之处在所难免,敬请读者批评指正,以便不断修改和完善。

编者

2017年5月

目 录

第一章 体育统计的基本概念	1	第二章 一元线性回归	216
第一节 统计与统计学	1	第九章 多元线性回归分析	226
第二节 体育统计	2	第一节 多元线性回归模型的求法	227
第三节 体育统计学的几个基本概念	5	第二节 多元线性回归模型的检验	228
第二章 统计数据的收集与整理	15	第三节 多元线性回归应用中应注意的 问题	230
第一节 统计数据的收集	15	第四节 多元线性回归分析的 SPSS 例解	233
第二节 样本设计与抽样方法	25	第十章 聚类分析	248
第三节 统计数据的整理	30	第一节 聚类分析概述	248
第三章 统计描述	45	第二节 系统聚类方法	249
第一节 描述统计	45	第三节 系统聚类应用	251
第二节 频数分析	61	第四节 系统聚类的 SPSS 操作步骤	257
第三节 统计图表	72	第十一章 因子分析	274
第四章 概率分布及应用	85	第一节 因子分析引例	274
第一节 概率基本知识	85	第二节 因子分析基本原理	275
第二节 离散型随机变量的概率分布	86	第三节 因子分析过程	277
第三节 连续型随机变量的概率分布	88	第四节 因子分析的适用条件	279
第四节 正态分布在体育中的应用	92	第十二章 体育统计常见问题及解决 策略	294
第五章 参数估计与假设检验	110	第一节 体育统计常见问题	294
第一节 抽样分布与参数估计	110	第二节 合理选择统计方法的策略	297
第二节 假设检验原理	122	练习题参考答案	313
第三节 单样本 t 检验	123	附录	319
第四节 两独立样本 t 检验	126	附表 1 (1) 标准正态分布表	319
第五节 配对样本 t 检验	131	附表 1 (2) 标准正态分布表	320
第六章 方差分析	142	附表 2 t 值表	321
第一节 单因素方差分析	142	附表 3 F 值表 (方差齐性检验用)	323
第二节 双因素方差分析	150	附表 4 (1) F 值表 (方差分析用)	325
第三节 协方差分析简介	156	附表 4 (2) F 值表 (方差分析用)	327
第七章 相对数与动态分析	175	附表 5 χ^2 值表	329
第一节 相对数	175	附表 6 相关系数界值表	330
第二节 动态分析	180	附表 7 等级相关系数界值表	332
第三节 χ^2 检验	191	参考文献	334
第八章 相关与回归分析	206		
第一节 相关分析	206		

第一章 体育统计的基本概念

在日常生活中，人们经常会接触到“统计”这一术语，各类媒体报道中也经常会使用统计数据、统计图表等。事实上，统计无处不在，统计学是应用最广泛、最具有普遍适用性的一门科学。只有学会一些统计分析方法，才会形成一套完全不同的看待问题的思维模式，从而帮助人们更加有效地制订各种统计决策。体育事业的快速发展，为体育统计学提供了难得的机遇，使之获得了更为广阔的发展空间，也使得其学科地位得到了进一步的巩固和强化。本章主要讨论体育统计学的基本概念。

第一节 统计与统计学

学习目标：

1. 理解统计的含义。
2. 了解统计学的发展阶段。
3. 掌握体育统计学的几个基本概念。

一、统计的含义

统计包括三个含义：统计工作、统计资料和统计学。

统计工作、统计资料、统计学三者之间的关系是相互联系、不可分割的。一方面，统计工作和统计资料是过程与结果的关系，统计资料是统计工作的成果或结晶，它来源于统计工作，又服务于统计工作；另一方面，统计学和统计工作是理论与实践的关系，统计学是从统计工作中总结出来的，反过来也是统计工作的理论基础，不断促进统计工作的发展。统计工作的发展又不断丰富和完善着统计学理论体系。

二、统计学的基本概念

统计学是指收集、整理、分析和解释数据的科学。

统计学的内容十分丰富，研究和应用的领域非常广泛。从统计教育的角度来看，统计学大致有两种分类：

（一）描述统计和推断统计

描述统计是用图形、表格和概括性的数字对数据进行描述的统计方法。常用直观的图形、汇总的表格和概括性的数字（如平均数、标准差等）表示数据的分布、形状等特征，并为进一步的推断提供依据。

推断统计是利用样本数据对总体进行估计、假设检验、预测或其他推断的统计方法。

（二）理论统计与应用统计

理论统计（数理统计）是研究统计学的一般理论，研究统计方法的数学原理。

应用统计是研究统计学在各领域的具体应用。例如，统计方法在物理研究中的应用就形成了物理统计，统计方法在生物学中的应用就形成了生物统计，统计方法在医学研究中的应用就形成了医疗卫生统计，还有教育统计、心理统计、社会经济统计、体育统计等。

概率论是数理统计的理论基础，体育统计是数理统计方法在体育中的应用。

三、统计学发展的三个阶段

（一）统计学的初级阶段

17 世纪中叶至 19 世纪末，描述统计学与概率论的内容基本形成。

（二）统计推断方法体系基本形成的阶段

20 世纪初至 20 世纪 40 年代末，推断统计迅速发展，概率论体系日益完善。

（三）统计方法与应用研究全面发展的阶段

20 世纪 50 年代至今，统计学受计算机、信息论等现代科学技术的影响，新的研究分支不断出现，统计应用领域不断扩展。

历史上著名的统计学家：

Jacob Bernoulli（伯努利）（1654—1705）

Edmond Halley（哈雷）（1656—1742）

De Moivre（棣莫弗）（1667—1754）

Thomas Bayes（贝叶斯）（1702—1761）

Leonhard Euler（欧拉）（1707—1783）

Pierre Simon Laplace（拉普拉斯）（1749—1827）

Adrien Marie Legendre（勒让德）（1752—1833）

Thomas Robert Malthus（马尔萨斯）（1766—1834）

Friedrich Gauss（高斯）（1777—1855）

Johann Gregor Mendel（孟德尔）（1822—1884）

Karl Pearson（皮尔逊）（1857—1936）

Ronald Aylmer Fisher（费希尔）（1890—1962）

Jerzy Neyman（奈曼）（1894—1981）

Egon Sharpe Pearson（皮尔逊）（1895—1980）

William Feller（费勒）（1906—1970）

第二节 体育统计

学习目标：

1. 了解体育统计学的发展及作用。
2. 熟悉体育统计分析过程。
3. 掌握体育统计的研究对象及特点。

体育统计是运用统计学的原理与方法,通过对体育教学、运动训练、体育科研和体育管理中随机现象的描述、推断和分析,揭示其数量规律的一门应用学科。体育统计是利用数理统计方法(理论)研究体育教学、运动训练和体育管理等体育专业基础课程,是现代体育的主要研究方法之一。

一、体育统计学科发展

早期的统计仅限于描述统计,19世纪中期开始进入推断统计阶段。20世纪30年代,以徐英超教授为代表的老一辈体育教育工作者将概率论引入体育研究中,提倡用统计方法研究体育问题;40年代至60年代,徐英超教授等在西北师范学院体育专业、北京体育学院本科班和助教班开设过“体育测验与统计”课程;80年代,随着全国体质测试、体育科学研究等对统计方法的需求,各院校陆续都开设了“体育统计”专业基础课。1982年,由华中师范大学刘厚生老师等学者发起,在武汉举办了第一届全国体育统计论文报告会,并成立了全国体育统计研究会。1984年,中国体育科学学会接纳体育统计研究会为体育统计专业委员会,从而为以后学科的发展奠定了坚实的基础;90年代初,体育统计学教材编写组编写完成了体育院(系)的《体育统计学》通用教材。

统计学在收集、整理和分析数据上所显示出的强大功能,已在我国改革开放30余年的历程中得到了充分的体现,并在我国得到了蓬勃发展。近年来,作为体育科学领域重要学科的体育统计已被广泛应用于体育运动实践,体育统计的原理和方法已被体育界广泛接受。

中国体育统计学科的建设,是跟随我国改革开放巨变而前行的。特别是1984年体育统计专业委员会(后改为体育统计分会)成立以来,刘厚生、王路德、徐迪生、陈及治、权德庆、程致屏、祁国鹰等体育统计工作者,依托体育统计学术影响,开展了大量卓有成效的工作,推动了体育统计学科的快速发展。在此期间,该学科发表了大量研究成果,涌现出各具特色的研究方向、创新团队和一批学术领军者,取得了一系列标志性成果。

目前,体育统计学科建设基本成型,内容体系不断完善,体育统计方法在体育实践中不断发展、提高和深化,新的统计方法不断引进,体育统计方法体系不断得到补充和更新,体育统计学科在体育科学体系中的地位得到确认,体育统计学科具有良好的发展前景和广阔的发展空间。

二、体育统计的研究对象及特点

体育统计主要是数理统计方法在体育领域中的应用,它以体育运动中随机现象的规律性为研究对象。数理统计方法之所以能够运用于体育运动中,是因为体育运动中的许多现象都是不确定的、随机性的,这正是以概率论为基础的数理统计学所研究的对象。体育统计学的研究对象是随机现象的数量方面,包括数量特征、数量关系和数量界限,通过这些数量方面的研究来探索现象变动的内在数量规律性。体育统计学的研究对象具有数量性、总体性、差异性的特点。

(一) 数量性

数量性是体育统计学研究对象的基本特点。这是因为数字是统计的语言,数据资料是统计的原料。任何客观事物都有质和量两个方面,事物的质与量总是密切联系的,共同规

定着事物的性质。事物的发展充满了不确定性，而体育统计学既研究如何从数据中把信息和规律提取出来，找出最优化的方案，也研究如何把数据中的不确定性量化出来。体育统计学总是与所研究体育随机现象的数量特征相联系，主要是从数量方面进行定量研究，揭示体育内在的数量规律性。

（二）总体性

总体性是指体育统计学所研究的是总体的数量规律性，而不是个体的数量规律性。体育统计学在认识和研究体育现象时，一般是从个体单位、个别现象、个别事物研究入手，但是体育统计学研究的目的并不仅仅是了解个别体育现象，而在于揭示体育的总体规律，从总体上研究体育现象的内在规律性。

（三）差异性

体育统计学研究同类体育现象总体的数量特征，其前提是总体各单位特征表现存在差异，而且这些差异并不是事先可以预知的。如各地区学生的身高、体重和体质指标每年不同，需要对其进行统计研究，编制体质健康指数等。对体育现象总体的数量研究，是从总体各单位的变异中归纳概括出它们的共同特征，显示出现象的普遍性和必然性的数量规律性。

三、体育统计研究

（一）体育统计分析过程

体育统计分析过程是对体育现象总体的数量方面进行调查或实验研究以及综合分析的过程，体育统计分析过程如图 1-2-1 所示。

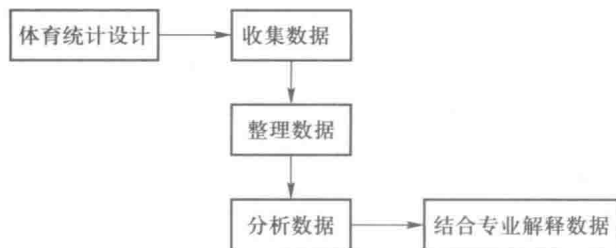


图 1-2-1 体育统计分析过程

（二）体育统计的作用

随着知识经济与体育事业的发展，当代体育领域大规模的信息处理所遇到的信息压缩、数量特征检测、可靠性分析、各种过程数理模型建立、事业未来发展定量预测等问题，都必须依靠统计理论与计算机技术加以解决。体育统计学可用于数据资料的描述统计分析，比较数据的整齐程度，检验可疑数据，制订体育项目考核标准，制订体质健康等级评价表，评价体育教学方法、训练手段的优劣，预测、控制、综合分析研究，识别训练中的疲劳程度，进行体质调研，确定选材指标，建立统计指标体系等。

随着计算机信息技术的发展和统计软件的广泛普及，体育统计的应用范围已涵盖体育学研究的众多领域。通过统计技术描述数据或进行推断分析，揭示事物的内在特征和联系，成为体育各学科实证研究的基本手段。在体育统计技术的运用中，古典统计方法，如描述统计、假设检验、非参数统计、相关分析、回归分析、方差分析、聚类分析、因子分析、判别分析等已成为量化分析的常用统计技术。此外，一些其他学科的数据处理技术和

新型统计技术被引入体育研究领域,如模糊数学、灰色系统理论、多层次分析、数据挖掘、结构方程模型等高级统计方法,以满足日益发展的体育研究需要。

第三节 体育统计学的几个基本概念

学习目标:

1. 掌握总体、样本的概念。
2. 理解参数、统计量的概念。
3. 熟悉变量与统计误差的概念。

一、总体与样本

根据研究目的而确定的同质对象的全体称为总体。如要研究北京市七年级男生速度素质状况,现将 50 m 跑的成绩作为测试项目进行分析,研究对象具有这样一些相同的属性:北京市学生,今年读七年级,性别是男性,且测试的是 50 m 跑的成绩。具有某种相同性质称为总体的“同质性”,这是形成统计总体的必要条件。组成总体的每个基本单位称为个体,从总体中抽取的部分个体组成的集合称为样本,样本中包含的个体数量称为样本量。总体与样本的概念是相对的,不是固定不变的,是由研究目的确定的。

例如,要调查某年某地区中学生的身高水平,则该地区全部中学生的身高就是一个总体,而该地区每一名中学生的身高就是个体。但要调查中学生的身高情况,并不一定要对该地区中学生身高的总体中所有的个体都测量到,而只要对其中的一部分个体进行测量。这种从总体中随机抽取部分个体的过程,称为“抽样”。当然所抽得的那一部分就叫样本。例如从该地区中学生中随机抽取 100 名进行身高测量,那么这 100 名中学生的身高就是一个样本,样本量为 100。

[例 1-3-1] 为了研究西安市 15 岁男生的身高发育情况,现从该市 20 所中学里随机抽取 300 名 15 岁男生,测其身高数据,问总体和样本分别是什么?样本量为多少?

答:总体——西安市 15 岁男生的身高全体,

样本——300 名 15 岁男生的身高,

样本量为 $n=300$ 。

[例 1-3-2] 某教师为了检验他所研究的中学女生俯卧式跳高教法的效果,用他所授课的八年级女生 80 人进行教法试验,问总体和样本各是什么?

答:总体——该教法适用范围内的中学女生的全体,

样本——八年级 80 名女生。

统计学的目的是通过对少量数据的观测、收集来印证尽可能多的发现。在人们获取和储存数据的能力都十分有限的年代,随机抽样是一种捷径,人们无须耗时耗力去观测总体中所有个体,可以通过样本对总体有一个大概的估计。用样本推断总体,就如窥一斑而见全豹,观滴水而知沧海。当然,为了使样本具有较好的代表性,必须注意抽样方法,同时要保证样本有适当的含量。

二、参数与统计量

参数是用来描述总体特征的统计指标，它是研究者想要了解总体的某种特征值。在统计中，总体参数通常用希腊字母表示。例如，总体平均数用 μ 表示，总体标准差用 σ 表示，总体比例用 π 表示，总体相关系数用 ρ 表示，等等。

统计量是用来描述样本特征的统计指标，它是样本数据计算出来的一个量。样本统计量通常用英文字母来表示，例如，样本平均数用 $\bar{x}$ 表示，样本标准差用 S 表示，样本比例用 P 表示，样本相关系数用 r 表示，等等。抽样的目的就是要根据样本统计量去估计总体参数。具体可参见图 1-3-1。

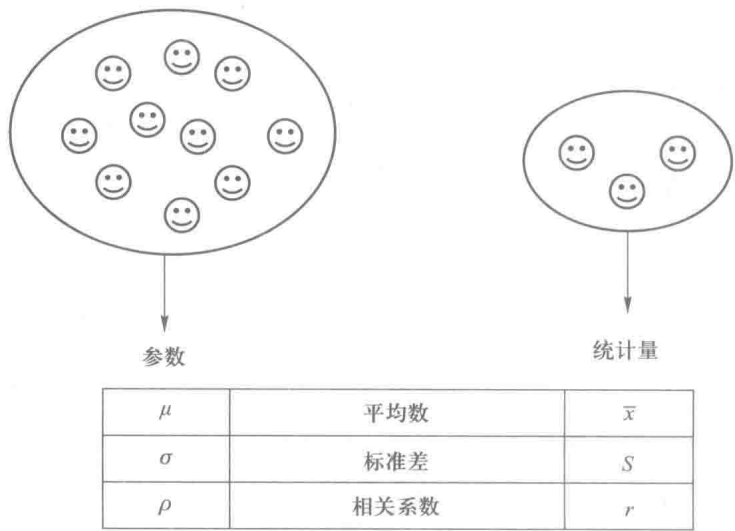


图 1-3-1 总体和样本、参数和统计量

三、变量

变量是说明现象某种特征的概念，其特点是从一次观察到下一次观察结果会呈现出差别或变化，如“运动员的身高”“50 m 跑的成绩”“受教育程度”等都是变量。变量的具体取值称为变量值，例如学生的考试成绩有 60 分、70 分、85 分、95 分、97 分等，这些数字就是变量值。统计数据就是统计变量的具体表现。变量可分为定量变量和定性变量。

(一) 定量变量

定量变量又称为数值变量，是用一定的测量工具进行测量所得的数据。如身高、体重、跳远成绩、学科成绩、反应时间等。

(二) 定性变量

定性变量又称为非数值变量或类别变量，是指通过计算个数所得的数据，如性别、学历、等级运动员人数、等级教练员人数等。

四、统计误差

测得值 = 真值 + 误差，所以误差是测量值与真值之差。

误差的种类主要有 4 种：

- (1) 随机误差：是一种不恒定的、随机变化的误差，由无法控制的偶然因素造成。
- (2) 系统误差：是由于仪器不准等原因导致的倾向性误差。系统误差恒定不变或遵循一定的变化规律。
- (3) 过失误差：是研究者偶然失误而造成的误差。
- (4) 抽样误差：是由于随机抽样造成的样本统计量与总体统计参数间的差别。

练习题 1

一、名词解释

1. 总体
2. 样本
3. 个体
4. 样本量
5. 统计量
6. 统计参数

二、思考题

1. 简述体育统计的主要研究过程。
2. 举例说明总体和样本的概念。

实例 1.1 建立 SPSS 数据文件

SPSS 是目前世界上最优秀的统计分析软件之一。SPSS 原意为 Statistical Package for the Science，即“社会科学统计软件包”，2000 年正式更名为 Statistical Product and Service Solutions，意为“统计产品与服务方案”，使用 Windows 操作系统的窗口方式展示各种管理和分析数据的方法，使用对话框展示各种功能选择项，其基本功能包括数据管理、统计分析、图表分析、输出管理等。迄今为止，SPSS 已广泛应用于自然科学和社会科学，其中涉及领域包括工程技术、应用数学、经济学、商业、金融、生物学、医疗卫生、体育、心理学、农林等。

SPSS 的基本功能包括数据管理、统计分析、输出管理等，其统计分析过程包括描述统计、均值比较、方差分析、相关分析、回归分析、聚类分析、因子分析、生存分析、时间序列分析、多重响应等。SPSS 还有专门的报表和绘图系统，可以产生各类统计表和绘制各种图形。

SPSS for Windows 界面友好，功能强大，主要版本有 SPSS7.0、……、SPSS17.0，SPSS18.0 等。本书以 SPSS18.0 for Windows 标准版为例，并在本书中简称为 SPSS。对于使用其他版本的用户，本书的内容也基本适用。

一、SPSS 的主界面

1. 启动与退出

SPSS 安装完毕后，系统会自动在 Windows 的开始菜单中创建快捷方式。选择“开始”菜单/“程序”/“ PASW Statistics 18”命令，即可启动 SPSS。

完成 SPSS 的统计分析后，退出该系统的方法是：选择“文件”菜单/“退出”命令，或单击标题栏上的关闭按钮“”，即可退出 SPSS。

2. SPSS 的数据编辑窗口

SPSS 的数据编辑窗口与微软公司的 Excel 窗口有些相似, 不过 SPSS 的统计功能要比 Excel 强得多。

数据编辑窗口由标题栏、菜单栏、工具栏、编辑栏、变量名栏、内容区和状态栏组成, 如实例图 1-1-1 所示。



实例图 1-1-1 数据编辑窗口

标题栏中显示编辑的数据文件名, 实例图 1-1-1 中所编辑的数据文件为“实例 1.1 某地区 13~15 岁学生资料”。

SPSS 自动命名变量名为“变量”。实例图 1-1-1 中有 5 个变量: 编号、性别、年龄、身高和体重。

数据编辑窗口下方有两个标签: “数据视图”和“变量视图”。“数据视图”对应的表格用于查看、录入和修改数据; “变量视图”对应的表格用于输入和修改变量的定义。

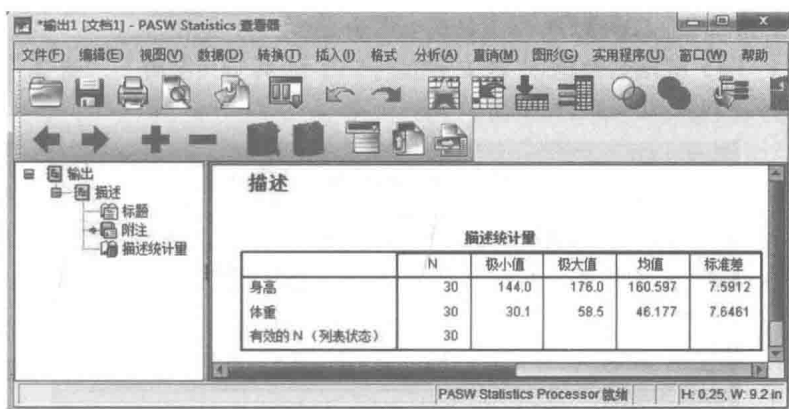
完成变量定义后, 在数据编辑窗口中输入数据, 选择统计分析菜单的某个子菜单项, SPSS 会自动完成统计分析, 并将弹出结果输出窗口, 显示数据分析结果。

3. SPSS 的结果输出窗口

SPSS 的结果输出窗口名为“文档 1”“文档 2”等, 它是显示和管理 SPSS 统计分析结果、报表及图形的窗口。用户可以将此窗口的内容以结果文件“.spv”的形式保存。

标准的结果输出窗口如实例图 1-1-2 所示。

可以对输出区中的表格进行编辑等操作。单击鼠标选中某个表格时, 相应表格的四周会出现黑色单线边框。如果要编辑某个表格, 可以双击该表格, 当表格四周出现黑色虚线边框时, 即可对表格内的数据进行修改。

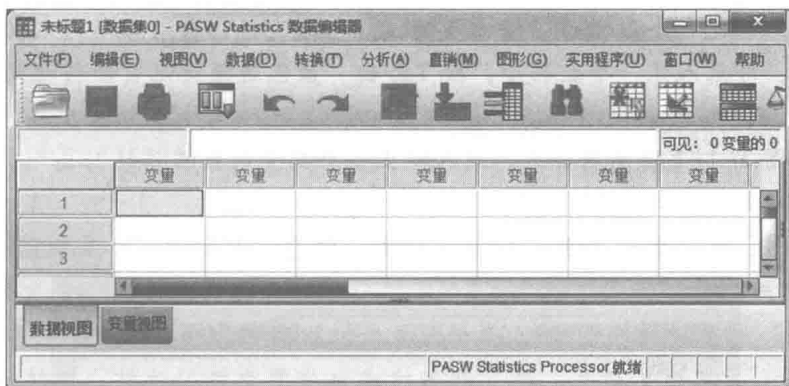


实例图 1-1-2 结果输出窗口

二、变量与数据文件

(一) 定义变量

启动 SPSS 后, 首先出现数据编辑窗口。在没有输入数据前, 显示的是一个空文件, 如实例图 1-1-3 所示。



实例图 1-1-3 数据编辑窗口 (空文件)

输入数据之前先要定义变量。定义变量包括定义变量名、变量类型、变量长度 (小数位数)、变量标签 (或值标签) 和变量的格式。

单击数据编辑窗口左下方的“变量视图”标签或双击列的题头 (变量), 进入如实例图 1-1-4 所示的变量定义视图窗口, 在此窗口中即可定义变量。

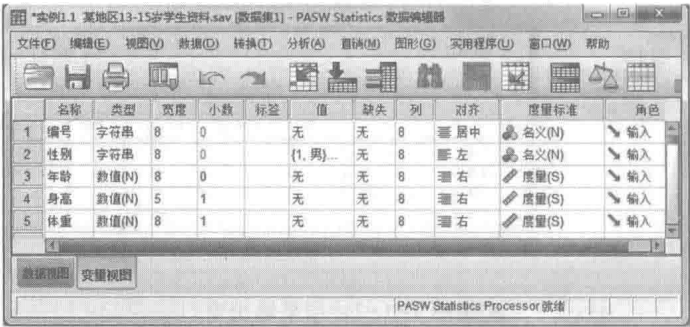


实例图 1-1-4 变量定义视图窗口

在实例图 1-1-4 窗口中, 每行表示一个变量信息, 包括变量名、变量类型、变量宽度、小数点位数、变量标签、变量值标签、缺失值、变量显示列宽、变量对齐方式、变量

度量标准、变量角色。

按实例图 1-1-5 所示的格式，定义“编号”“性别”“年龄”“身高”“体重”5 个变量。



实例图 1-1-5 定义变量后的视图窗口

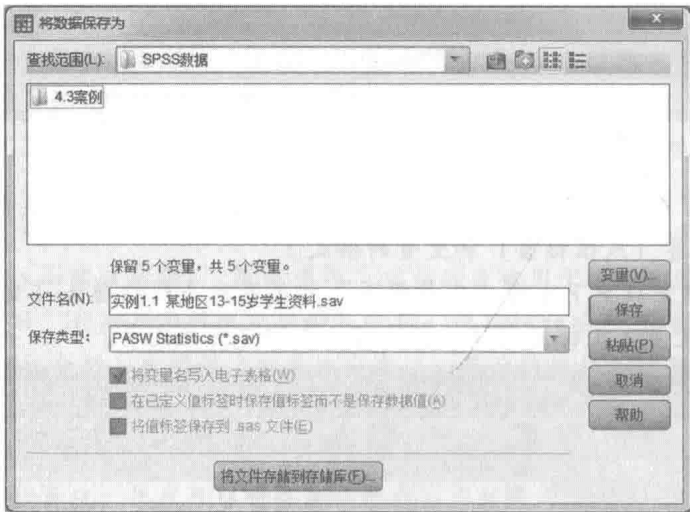
(二) 数据的输入与保存

定义了所有变量后，单击“数据视图”标签，即可在数据视图中输入数据。数据编辑窗口中蓝色边框标记的单元格为当前数据单元格。数据录入时可以逐列录入，即按照变量录入数据，录入完一个数据后，按回车键，蓝色底纹标记自动移动到本列的下一行。也可以逐行录入，即录入完一个数据后，按 Tab 键，蓝色边框标记自动移动到本行的右边一个变量列上。

按照实例图 1-1-1 所示的格式，录入学生 5 个变量的数据。

保存数据的具体操作如下：

选择“文件”菜单的“保存”命令，弹出“保存”对话框，如实例图 1-1-6 所示。



实例图 1-1-6 “保存”对话框

单击“保存类型”对话框，可以看到 SPSS 中数据可以保存的各种类型，这里将其保存为 SPSS 默认的数据格式（*.sav 文件）。在“文件名”框内键入“实例 1.1 某地区 13~15 岁学生资料”并单击“保存”按钮，系统返回到了数据编辑窗口，标题栏变成了“实例 1.1 某地区 13~15 岁学生资料.sav”，表明当前编辑的文件为“实例 1.1 某地区 13~15 岁学生资料.sav”。

实例 1.2 SPSS 数据文件的基本操作

一、数据文件的导入

数据文件的导入,即调用已建立的数据文件。SPSS 可以调用 SPSS (*.sav)、Excel (*.xls)、dBASE (*.dbf)、stata (*.dta)、Text (*.txt, *.dat) 等数据文件形式,详细过程可参阅其他参考书。

SPSS 的文件类型主要有:数据文件“.sav”,结果文件“.spv”,图形文件“.cht”。

二、数据文件的整理

对数据文件进行加工处理的功能基本都集中在“编辑”和“数据”菜单上,比如根据统计分析的要求对数据进行分组、合并、加权、筛选等操作。

(1) 插入变量。在数据编辑窗口中选定要插入变量位置的后一个变量,从工具栏上直接点击插入变量按钮;或选择“编辑”菜单/“插入变量”命令,数据编辑窗口便产生一个变量列。

(2) 记录插入。在数据编辑窗口中选定要插入记录位置的后一个记录,从工具栏上直接点击插入记录按钮;或选择“编辑”菜单/“插入个案”命令,数据编辑窗口便产生一个记录行。

(3) 到某一记录。从工具栏上直接点击按钮,或选择“编辑”菜单/“转至个案”命令,弹出“转到(G)”对话框,填入某一记录的编号,单击“转向”按钮,光标移到该记录。

(4) 变量值的排序。选择“数据”菜单/“排序个案”命令,弹出“排序个案”对话框,将排序变量选入“排序依据”列表框中,在“排列顺序”栏中选择升序(A)或降序(D),单击“确定”按钮,完成对记录的排序。

(5) 行列转置。将数值型变量的行和列进行互换。选择“数据”菜单/“转置”命令,弹出“转置”对话框。

(6) 数据文件的合并。数据文件的合并有两种方式:纵向合并和横向合并。

纵向合并就是将一个 SPSS 数据文件的内容追加到数据编辑窗口当前数据的后面。选择菜单“数据”/“合并文件”/“添加个案”命令。

横向合并也就是变量的合并,使用横向合并可以将多个数据文件连在一起。选择菜单“数据”/“合并文件”/“添加变量”命令。

(7) 数据文件的分组。数据文件的分组是按某个变量进行分组,随后的统计分析将对各个组进行。例如,想分别了解男生和女生的成绩情况,可按性别变量进行数据文件的分组。选择“数据”菜单/“拆分文件”命令。

(8) 数据文件的分类汇总。数据文件的分类汇总就是按指定变量的数值进行归类分组汇总。选择“数据”菜单/“分类汇总”命令。可将分类变量的均值、标准差、最大值、最小值等形成一个新的数据文件。

(9) 选取记录子集。在数据统计时,可从所有资料中选择符合条件的记录进行统计分析。选择“数据”菜单/“选择个案”命令。例如,可对满足条件为(性别=“男”)并且(身高>155)的数据进行统计分析。