

全国普通高等教育医学类系列教材

供研究生和7年制各专业用

中医药统计学

*Statistics for Traditional
Chinese Medicine*

史周华 张雪飞 主编

申 杰 主审



科学出版社

www.sciencep.com

中国医药统计学会
CHINA SOCIETY OF PHARMACOSTATISTICS

CHINA SOCIETY OF PHARMACOSTATISTICS

中国药统计学

Journal for Pharmaceutical
Statistics

1982年10月 第1期
第1卷第1期



中国医药统计学会
CHINA SOCIETY OF PHARMACOSTATISTICS

全国普通高等教育医学类系列教材
供研究生和7年制各专业用

中医药统计学

史周华 张雪飞 主 编
申 杰 主 审

科学出版社

北 京

内 容 简 介

本书是由科学出版社牵头组织全国 12 所高等中医药院校统计学教学一线、有丰富经验的教师历时一年精心完成的创新教材。

全书共 18 章,内容包括绪论、统计描述、概率分布、参数估计与假设检验、 t 检验、方差分析、双变量相关与回归分析、 χ^2 检验、基于秩次的非参数检验、实验设计、调查设计、协方差分析、多重线性回归分析、Logistic 回归分析、生存分析、聚类分析和判别分析、主成分分析与因子分析、Meta 分析,以及统计用表、希腊字母表和常用统计符号等附表。

本书重视“三基”,突出重点,强调统计理论与计算机实验有机结合,兼顾统计知识与技能的培养,深入浅出地介绍统计概念、方法与原理、中医药统计应用实例与 SPSS 电脑实验。本书既可作为研究生与 7 年制各专业学生的教材,又可作为从事中医药科研人员的参考用书。

图书在版编目(CIP)数据

中医药统计学/史周华,张雪飞主编.—北京:科学出版社,2009

全国普通高等教育医学类系列教材
ISBN 978-7-03-025013-1

I. 中… II. ①史…②张… III. 中国医药学—医学统计—医学院—教材 IV. R2-32

中国版本图书馆 CIP 数据核字(2009)第 119824 号

责任编辑:潘志坚 钱 鑫 / 责任校对:刘珊珊
责任印制:刘 学 / 封面设计:殷 靛

科 学 出 版 社 出 版

北京东黄城根北街 16 号
邮政编码:100717

<http://www.sciencep.com>

上海敬民实业有限公司长阳印刷厂印刷
科学出版社发行 各地新华书店经销

*

2009 年 8 月第 一 版 开本:889×1194 1/16
2009 年 8 月第一次印刷 印张:18
印数:1—3 300 字数:634 000

定价:34.00 元

全国普通高等教育医学类系列教材
供研究生和7年制各专业用

《中医药统计学》编委会名单

主 编 史周华(山东中医药大学)
张雪飞(湖北中医学院)

主 审 申 杰(河南中医学院)

副主编 刘仁权(北京中医药大学)
莫传伟(广州中医药大学)
李秀昌(长春中医药大学)
李国春(南京中医药大学)

编 委 (以姓氏笔画为序)
王成岗(山东中医药大学)
邢建民(北京中医药大学)
孙春阳(河南中医学院)
张跃林(陕西中医学院)
陈晓凡(江西中医学院)
郑国华(福建中医学院)
黄 杏(湖北中医学院)
黄品贤(上海中医药大学)
崔 宁(山东中医药大学)

前 言

中医药统计学是基于概率论和数理统计的基本原理和方法,研究中医药领域中科研设计、数据收集、数据整理和数据分析的一门应用性科学,是医疗卫生人员正确认识中医药领域及其相关领域的客观规律、总结工作经验、进行中医药科学研究和疾病诊断防治工作的重要工具。现代中医药事业的发展必须注重统计艺术和科学证据。人类社会已进入信息时代,要在大量的信息中获得有价值的结果,需要对信息进行科学的统计分析,这就要求医学生具有扎实的中医药统计学知识和技能。

中医药统计学涉及概率论和数理统计等数学知识,具有抽象性、综合性和灵活性等特点,对医学生来讲有一定的难度,不过通过计算机知识的普及、优秀统计软件的应用使得统计复杂的数值计算问题可以迎刃而解。医学生学习中医药统计学,不必死记数学公式,只需对概率论和数理统计知识有所了解,领会统计基本概念、基本原理与方法,掌握技能,利用统计软件对实际数据进行统计分析并正确报告结果。做到统计理论指导实践应用,实践应用提高对统计理论方法的理解与把握。基于上述编写指导思想,本着“通俗易懂、着力实用”和“知其然,知其所以然”的编写原则,本书具有如下特点:

1. 本书按基础统计、研究设计、多元统计分析三大部分,由浅入深,由简单到复杂比较系统地介绍了统计方法在中医药科学研究的应用,内容简洁严谨、重点突出、实用性强,适用于研究生和7年制专业统计学课程教学。

2. 力争深入浅出地对中医药统计学基本概念、基本原理和基本方法进行阐述,淡化数学公式的推导与证明,着力各种基本统计方法的基本思想或原理的通俗解释与理解,强调设计优先,明确资料属性,侧重统计方法的应用条件、具体应用和统计结果的正确解释与表达。

3. 每章设有例题与 SPSS 电脑实验,包括建立数据集、具体分析步骤、主要结果与解释及结论等。这样,既可训练读者 SPSS 统计软件的应用技能,又可使读者及时加深对统计概念、原理方法的理解与把握。

本书的编写得到了山东中医药大学、湖北中医学院、河南中医学院、上海中医药大学、北京中医药大学、广州中医药大学、长春中医药大学、南京中医药大学、陕西中医学院、福建中医学院、江西中医学院等院校领导的大力支持。科学出版社上海分社、湖北中医学院和上海中医药大学对教材编委会、定稿会给予了热情接待、周到安排。在本书编写过程中,借鉴和吸收了国内外有关文献和科研资料。在审稿期间,刘仁权教授、莫传伟教授、李秀昌教授付出了艰辛而富有智慧的劳动,山东中医药大学的李运伦教授、郝涛老师和研究生李青、王若维、史林等在审稿编辑校对等方面做了大量细致的工作。在此,我们谨代表编委会一并对他们的关心、支持、帮助和贡献表示最诚挚的敬意和谢意!

本书汇集了所有编委成员的统计学教学科研经验与智慧,但限于我们的水平、能力和经验,书中难免存在不妥或错漏,恳请同行、广大师生、中医药工作者不吝指正,以便我们今后进一步改善和提高本教材的质量。

为方便大家使用,与本书配套的 SPSS 数据文件集可以登陆“www.sciencep.sh.cn”下载,文件名为“中医药统计学——研究生和 7 年制 SPSS 数据文件”。

史周华 张雪飞

2009 年 5 月

目 录

前言

第一章 绪论

1

- 第一节 中医药统计学概述 /1
 - 一、统计学与中医药统计学的定义 /1
 - 二、统计学的发展简史 /1
 - 三、统计学的基本思想、研究方法和特点 /1
 - 四、中医药统计学的主要内容 /2
- 第二节 统计工作的基本步骤 /2
 - 一、研究设计 /2
 - 二、收集资料 /2
 - 三、整理资料 /3
 - 四、分析资料 /3
- 第三节 统计指标与资料的分类 /3
 - 一、变量与变量值 /3
 - 二、变量转换 /4
 - 三、统计资料的分类 /4
- 第四节 统计学的几个基本概念 /5
 - 一、同质与变异 /5
 - 二、总体与样本 /5
 - 三、参数与统计量 /5
 - 四、误差 /5
 - 五、概率 /6
- 第五节 如何学习中医药统计学 /6
 - 一、思想重视 /6

- 二、观念正确 /7
- 三、方法得当 /7

第二章 统计描述

8

- 第一节 频数分布 /8
 - 一、频数表 /8
 - 二、频数图 /10
 - 三、频数表和频数图的用途 /11
- 第二节 数值资料的统计描述 /12
 - 一、集中趋势的统计描述 /12
 - 二、离散趋势的统计描述 /15
- 第三节 分类资料的统计描述 /17
 - 一、分类资料的相对数描述指标 /18
 - 二、率的标准化的 /21
- 第四节 统计表和统计图 /23
 - 一、统计表 /23
 - 二、统计图 /25
- 第五节 例题和 SPSS 电脑实验 /27

第三章 概率分布

29

- 第一节 正态分布 /29
 - 一、正态分布的概念和特征 /29
 - 二、正态分布面积分布规律 /30
 - 三、正态分布的应用 /31
- 第二节 二项分布和 Poisson 分布 /31
 - 一、二项分布 /31
 - 二、Poisson 分布 /34
- 第三节 抽样分布 /35
 - 一、 t 分布 /35
 - 二、 χ^2 分布 /36
 - 三、 F 分布 /36
- 第四节 医学参考值范围的制定 /37
- 第五节 例题和 SPSS 电脑实验 /38

第四章 参数估计与假设检验

40

- 第一节 参数估计 /40
 - 一、抽样误差与标准误 /40
 - 二、总体均数的估计 /41
 - 三、总体率的估计 /42
- 第二节 假设检验 /43
 - 一、假设检验的基本思想 /43
 - 二、假设检验的步骤 /43

- 三、假设检验的两类错误 /44
- 四、假设检验的注意事项 /45
- 五、假设检验的分类 /47
- 六、常用假设检验方法的选择 /47
- 第三节 正态性检验与数据转换 /48
 - 一、正态性检验 /48
 - 二、数据转换 /49
- 第四节 例题和 SPSS 电脑实验 /50

第五章 t 检验

52

- 第一节 样本均数与总体均数比较的 t 检验 /52
- 第二节 配对设计资料比较的 t 检验 /53
- 第三节 两独立样本均数比较的 t 检验 /54
 - 一、两小样本均数的比较 /54
 - 二、两大样本均数的比较(u 检验) /55
 - 三、方差不齐时的 t 检验 /55
 - 四、两样本几何均数的比较 /56
- 第四节 例题和 SPSS 电脑实验 /57

第六章 方差分析

59

- 第一节 方差分析概述 /59
 - 一、方差分析的基本思想 /59
 - 二、方差分析的应用条件 /61
- 第二节 单因素和双因素方差分析 /62
 - 一、完全随机设计的单因素方差分析 /62
 - 二、随机区组设计的方差分析 /65
 - 三、多个样本均数间的多重比较 /67
- 第三节 多因素方差分析 /69
 - 一、拉丁方设计的三因素方差分析 /69
 - 二、交叉设计的方差分析 /70
 - 三、析因设计的方差分析 /72
 - 四、正交设计的方差分析 /75
- 第四节 重复测量资料的方差分析 /78
 - 一、一元方差分析 /78
 - 二、多元方差分析 /80
 - 三、注意事项 /81
- 第五节 例题和 SPSS 电脑实验 /81

第七章 双变量相关与回归分析

88

- 第一节 直线相关 /88
 - 一、直线相关的概念及应用条件 /88
 - 二、相关系数的意义及计算 /88

- 三、相关系数的假设检验 /89
- 四、直线相关分析的步骤 /89
- 五、注意事项 /90
- 第二节 秩相关 /90
- 第三节 直线回归 /91
 - 一、直线回归的概念 /91
 - 二、直线回归分析的应用条件 /91
 - 三、直线回归分析的一般步骤 /92
 - 四、直线回归方程及其求法 /92
 - 五、直线回归方程和回归系数的假设检验 /92
 - 六、决定系数 /94
 - 七、在 X 实测范围内绘制出回归直线 /94
 - 八、残差分析 /94
 - 九、直线回归方程的应用 /94
 - 十、回归分析效果的评价 /95
 - 十一、注意事项 /96
 - 十二、直线相关与回归的区别与联系 /97
- 第四节 曲线拟合 /97
 - 一、曲线直线化的意义 /97
 - 二、常用的非线性函数 /97
 - 三、利用线性回归拟合曲线的一般步骤 /98
- 第五节 例题和 SPSS 电脑实验 /99

第八章 χ^2 检验

102

- 第一节 χ^2 检验概述 /102
 - 一、 χ^2 检验的基本思想 /102
 - 二、 χ^2 检验的基本公式 /102
- 第二节 独立样本四格表资料的 χ^2 检验 /103
 - 一、不校正的 χ^2 检验 /103
 - 二、校正的 χ^2 检验 /104
- 第三节 四格表资料的 Fisher 确切概率法 /105
- 第四节 独立样本多个率和构成比资料的 χ^2 检验 /106
 - 一、多个样本率比较的 χ^2 检验 /107
 - 二、多个构成比比较的 χ^2 检验 /107
 - 三、行 \times 列表资料 χ^2 检验的注意事项 /108
 - 四、多重比较 /108
- 第五节 配对设计分类资料的 χ^2 检验 /109
 - 一、配对设计分类资料概述 /109
 - 二、配对设计分类资料分析方法 /110
 - 三、配对设计四格表资料的假设检验 /111
 - 四、配对设计 $k \times k$ 表资料的假设检验 /113
- 第六节 行 \times 列表资料假设检验方法选择 /114
 - 一、双向无序的行 \times 列表 /114
 - 二、单向有序的行 \times 列表 /114
 - 三、双向有序属性相同的行 \times 列表 /115

- 四、双向有序属性不同的行×列表 /115
- 第七节 拟合优度的 χ^2 检验 /115
 - 一、单个样本两个或多个构成比的分析 /116
 - 二、二项分布的拟合优度检验 /116
- 第八节 多维分类资料的 CMH 统计分析 /117
- 第九节 例题和 SPSS 电脑实验 /119

第九章 基于秩次的非参数检验

123

- 第一节 非参数检验概述 /123
 - 一、非参数检验 /123
 - 二、秩和检验 /123
- 第二节 配对设计资料的符号秩和检验 /124
- 第三节 完全随机设计两样本比较的秩和检验 /125
 - 一、两组数值变量资料的秩和检验 /125
 - 二、两组有序分类变量资料(等级资料)的秩和检验 /126
- 第四节 完全随机设计多个样本比较的秩和检验 /126
 - 一、多组数值变量资料的秩和检验 /127
 - 二、多组有序分类变量资料的秩和检验 /127
- 第五节 随机区组设计资料的秩和检验 /128
- 第六节 多个样本两两比较的秩和检验 /130
 - 一、完全随机设计多个样本间的多重比较 /130
 - 二、随机区组设计资料的两两比较 /131
- 第七节 例题和 SPSS 电脑实验 /132

第十章 实验设计

134

- 第一节 实验设计的基本要素与基本原则 /134
 - 一、实验设计的基本要素 /134
 - 二、实验设计的基本原则 /135
- 第二节 常用实验设计类型 /138
 - 一、完全随机设计 /138
 - 二、配对设计 /139
 - 三、随机区组设计 /140
 - 四、拉丁方设计 /140
 - 五、交叉设计 /141
 - 六、析因设计 /142
 - 七、正交设计 /144
 - 八、重复测量设计 /147
- 第三节 临床试验设计简述 /148
 - 一、临床试验设计的原则 /148
 - 二、双盲临床试验操作步骤 /149
 - 三、新药临床试验的分期 /150
 - 四、临床试验的统计分析数据集 /150
 - 五、临床试验统计分析 /151
- 第四节 样本含量的估计 /152
 - 一、样本均数与总体均数比较时样本含量的估计方法 /152

- 二、配对设计和交叉设计数值资料样本含量(对子数)的估计方法 /153
- 三、完全随机设计的两样本均数比较时样本含量估计方法 /153
- 四、完全随机设计多个样本均数比较时样本含量估计方法 /153
- 五、样本率与总体率比较时的样本含量估计方法 /154
- 六、配对设计分类资料样本含量估计方法 /154
- 七、完全随机设计两样本率比较时的样本含量估计方法 /155
- 八、完全随机设计多个样本率比较样本含量的估计方法 /155
- 九、直线相关分析的样本含量估计方法 /156
- 十、多因素分析样本含量的估计方法 /156
- 十一、生存分析样本含量的估计 /156

第五节 例题和 SPSS 电脑实验 /156

第十一章 调查设计

158

第一节 概述 /158

- 一、确定调查对象和调查指标 /158
- 二、确定调查方法 /158
- 三、常用的概率抽样方法 /158
- 四、抽样调查的主要偏倚及其控制措施 /159
- 五、确定样本含量 /159
- 六、确定资料收集方式 /159
- 七、编制调查表或调查问卷 /160
- 八、制定资料整理分析计划 /160
- 九、制定调查的组织计划和质量控制方案 /160

第二节 调查问卷设计 /161

- 一、问卷的设计要求 /161
- 二、问卷方式 /161
- 三、统计尺度 /162
- 四、问卷设计 /162
- 五、调查表举例 /163
- 六、量表的应用 /164

第三节 调查问卷的质量评价 /164

- 一、信度 /165
- 二、效度 /167
- 三、信度和效度的关系 /168
- 四、反应度 /168
- 五、实用性和可接受性 /168
- 六、SPSS 统计软件实现效度与信度的分析 /168

第四节 样本含量的估计 /169

- 一、单纯随机抽样和系统抽样时样本含量的估计 /169
- 二、分层随机抽样时样本含量的估计 /170
- 三、整群抽样时样本含量的估计 /171

第十二章 协方差分析

172

第一节 协方差分析概述 /172

- 一、协方差分析的基本思想 /172
- 二、协方差分析的应用条件 /172
- 三、协方差分析的基本步骤 /172
- 第二节 协方差分析的应用 /176
 - 一、完全随机设计协方差分析 /176
 - 二、配伍组设计协方差分析 /176
- 第三节 例题和 SPSS 电脑实验 /178

第十三章 多重线性回归分析

181

- 第一节 多重线性回归 /181
 - 一、多重线性回归方程 /181
 - 二、多重线性回归方程的建立 /181
 - 三、多重线性回归方程的检验 /182
- 第二节 自变量的选择 /184
 - 一、选择自变量的标准与原则 /184
 - 二、选择自变量的常用算法 /184
- 第三节 多重线性回归的应用及其注意事项 /185
 - 一、多重线性回归的应用 /185
 - 二、各自变量作用大小的评价 /186
 - 三、应用多重线性回归时的注意事项 /186
- 第四节 例题和 SPSS 电脑实验 /187

第十四章 Logistic 回归分析

188

- 第一节 Logistic 回归分析概述 /188
 - 一、Logistic 回归模型 /188
 - 二、Logistic 回归参数的估计及其意义 /189
 - 三、Logistic 回归模型与回归系数的假设检验 /190
 - 四、标准偏回归系数 /190
 - 五、Logistic 回归模型的拟合优度 /191
 - 六、Logistic 回归模型的预测准确度 /191
 - 七、Logistic 回归的样本含量 /191
 - 八、共线性问题 /191
 - 九、异常值与强影响案例 /191
 - 十、筛选自变量与建立好的回归模型 /192
 - 十一、Logistic 回归的应用 /192
- 第二节 非条件 Logistic 回归 /192
 - 一、二分类资料的 Logistic 回归 /192
 - 二、多分类资料的 Logistic 回归 /195
- 第三节 条件 Logistic 回归 /198
 - 一、1:1 配对资料的条件 Logistic 回归 /198
 - 二、1:m 匹配资料的条件 Logistic 回归 /199
 - 三、关于 Logistic 回归的几点说明 /199
- 第四节 例题与 SPSS 实验 /199

第十五章 生存分析

204

- 第一节 生存分析概述 /204
 - 一、生存资料的基本概念 /204
 - 二、生存分析的基本概念与用途 /204
 - 三、生存分析的主要内容和基本方法 /205
 - 四、生存分析的常用指标 /205
- 第二节 生存资料的统计描述 /206
 - 一、乘积极限法 /206
 - 二、寿命表法 /207
- 第三节 生存率的估计 /208
 - 一、乘积极限法 /208
 - 二、寿命表法 /208
 - 三、尾部数据生存率可信区间的估计 /209
- 第四节 生存曲线的比较 /209
 - 一、对数秩检验 /210
 - 二、Breslow 检验 /211
 - 三、Gehan 比分检验 /211
 - 四、生存时间比较的假设检验注意事项 /213
- 第五节 Cox 比例风险回归模型 /213
 - 一、Cox 模型的基本原理 /213
 - 二、参数估计 /214
 - 三、假设检验 /215
 - 四、因素筛选与最佳模型的建立 /215
 - 五、应用实例 /216
 - 六、Cox 回归 PH 假定的判定方法 /218
 - 七、Cox 回归与多元线性回归、Logistic 回归的比较 /219
- 第六节 例题和 SPSS 电脑实验 /219

第十六章 聚类分析和判别分析

222

- 第一节 聚类分析 /222
 - 一、聚类分析概述 /222
 - 二、系统聚类 /223
 - 三、K 类中心聚类 /224
 - 四、两步聚类 /224
 - 五、讨论 /226
- 第二节 判别分析 /227
 - 一、Fisher 判别分析法 /227
 - 二、Bayes 判别分析法 /229
 - 三、逐步判别分析法 /230
 - 四、讨论 /231
- 第三节 例题和 SPSS 电脑实验 /231

第十七章 主成分分析与因子分析

233

- 第一节 主成分分析 /233

- 一、主成分分析的数学模型 /233
- 二、主成分的求法 /233
- 三、主成分的性质 /234
- 四、主成分个数的确定 /234
- 五、主成分得分的计算 /235
- 六、主成分分析的主要应用 /236
- 第二节 因子分析 /236**
 - 一、因子分析的数学模型 /236
 - 二、因子模型的性质 /237
 - 三、因子载荷阵的求解及公因子数目的确定 /237
 - 四、因子旋转 /237
 - 五、因子得分 /238
 - 六、因子分析的几点注意 /239
- 第三节 例题和 SPSS 电脑实验 /240**

第十八章 Meta 分析

243

- 第一节 Meta 分析概述 /243**
 - 一、Meta 分析的用途 /243
 - 二、Meta 分析的方法和步骤 /243
- 第二节 Meta 分析的基本统计方法 /244**
 - 一、效应指标的选择和表达 /244
 - 二、异质性检验和模型的选择 /245
 - 三、合并效应量的估计与统计推断 /245
 - 四、敏感性分析与亚组分析 /250
 - 五、Meta 分析应用注意事项 /250
- 附录 A 统计用表 /252**
 - 附表 1 百分率的 95% 可信区间 /252
 - 附表 2 Poisson 分布总体均数的可信区间 /253
 - 附表 3 q 界值表 /254
 - 附表 4 q' 检验界值表 /255
 - 附表 5 r 界值表 /256
 - 附表 6 r_s 界值表 /257
 - 附表 7 T 界值表(配对比较的符号秩和检验用) /258
 - 附表 8 T 界值表(两样本比较的秩和检验用) /259
 - 附表 9 H 界值表(三样本比较的秩和检验用) /260
 - 附表 10 M 界值表(随机区组比较的秩和检验用) /261
 - 附表 11 φ 值表(多个样本均数比较时所需样本含量的估计用) $\alpha=0.05, \beta=0.10$ /262
 - 附表 12 λ 值表(多组样本率检验时所需样本含量估计用) /264
 - 附表 13 随机数字表 /265
- 附表 B 希腊字母表 /266**
- 附表 C 常用统计符号 /267**

主要参考文献 /270

中医药统计学是研究医药科学中不确定现象客观规律的一门方法性学科,是中医药科学研究的重要工具。中医药科研的总体设计、资料收集、资料整理、资料分析和结果解释都与之关系密切,掌握了中医药统计学基本概念、基本原理和基本技能,就可以花费比较少的人力、物力、财力和时间获得比较可靠的研究结果,不然,就会事倍功半,甚至得出不合理或错误的结论。注重中医药统计学的学习与运用,对中医药工作者做好科研工作具有重要作用。

第一节 中医药统计学概述

一、统计学与中医药统计学的定义

1. 统计学(statistics) 是关于数据的学问,是一门从数据中提取信息、知识的科学与艺术,包括研究设计、收集资料、整理资料、分析资料和结果报告等步骤。它分为理论统计学和应用统计学两大类。

(1) 理论统计学(theoretical statistics):即数理统计学(mathematical statistics),主要探讨统计学的数学原理和统计公式的来源。它把研究对象一般化、抽象化,以概率论为基础,从纯理论的角度,对统计方法加以推导论证,其中心内容是统计推断问题,其实质是以归纳方法研究随机变量的一般规律。

(2) 应用统计学(applied statistics):是数理统计学的原理方法在不同学科领域中的具体应用。例如,数理统计学在生物学中的应用形成生物统计学(biostatistics);在医学中的应用形成医学统计学(medical statistics)、卫生统计学(health statistics)和中医药统计学(statistics for traditional Chinese medicine)。

2. 中医药统计学 是运用理论统计学(概率论和数理统计等)的基本原理与方法,结合中医药研究实际,阐述中医药领域研究设计,资料或信息的收集、整理、分析和结果报告的一门应用性学科。

二、统计学的发展简史

人类最初利用手指、石子、贝壳、小木棍以及绳索等工具进行的计数活动就蕴藏着统计学的种子。人类由统计实践上升到统计学,却只有300多年的历史。按统计方法及特征的历史演变顺序,统计学的发展史可分为古典统计学的萌芽时期(17世纪70年代至19世纪初期)、近代统计学的形成时期(19世纪初至20世纪初)和现代统计学的发展时期(20世纪初至今)三个阶段。正是由于古典统计学时期的政治算术学派和国势派以及近代统计学时期的数理统计学派和社会统计学派之间的相互争论、相互渗透,使数理统计学与社会统计学最后融合成为统一的现代统计学。

统计学的发展有4个明显趋势:①随着数学的发展,统计学吸收和运用的数学方法越来越多。②统计方法与计算机技术相结合,已渗透到所有的学科部门,以统计学为基础的边缘学科不断形成。③统计与实质性学科、统计软件、现代信息相结合,所发挥的功效日益增强。④统计学的作用与功能已从描述事物现状、反映事物规律,向抽样推断、预测未来变化的方向发展,已成为具有方法论性质的综合性学科。

三、统计学的基本思想、研究方法和特点

1. 统计学的基本思想 可归纳为变异的思想、概率的思想和随机抽样思想。正是由于客观事物的变异性和复杂性才需要统计学;由于统计学主要采用抽样研究方法探求总体的规律性,故统计结论具有概率性;为了得出正确的结论,在随机抽样时必须有效地控制各种误差。

2. 统计学的研究方法 统计学的基本研究方法是由观察到的个体特征归纳推断总体某种特征的归纳推理法。人们在统计实践经验的基础上不断地概括和总结,逐步形成了统计方法体系。统计方法体系包括以下4个方面:①研究设计方法有对照、随机、重复、均衡等方法。②收集资料方法有大量观测法、统计实验法、统计调查法等。③整理资料方法有统计审核法、统计分组法、统计汇总法等。④统计分析方法有统计描述法、统计推断法等。