

GAODENGYUANXIAO

TONGJIXUE

JINGPINJIAOCAI

高等院校统计学精品教材

统计学讲义 (第二版)

主 编/程士富 杜金柱

副主编/冯利英 田永祺 斯琴



中国统计出版社
China Statistics Press



高等院校统计学精品教材



统计学讲义

(第二版)

主 编/程士富 杜金柱
副主编/冯利英 田永祺 斯琴

常州大学图书馆
藏书章

(京)新登字 041 号

图书在版编目(CIP)数据

统计学讲义 / 程士富, 杜金柱主编. —2 版

—北京: 中国统计出版社, 2012. 3

ISBN 978-7-5037-6484-4

I. ①统… II. ①程… ②杜… III. ①统计学 IV.
①C8

中国版本图书馆 CIP 数据核字(2012)第 026567 号

统计学讲义(第二版)

作 者/程士富 杜金柱

责任编辑/陈悟朝

装帧设计/杨 超

出版发行/中国统计出版社

通信地址/北京市西城区月坛南街 57 号 邮政编码/100826

办公地址/北京市丰台区西三环南路甲 6 号

网 址/www.stats.gov.cn/tjshujia

电 话/邮购(010)63376907 书店(010)68783172

印 刷/河北天普润印刷厂

经 销/新华书店

开 本/787×1092mm 1/16

字 数/530 千字

印 张/22.75

版 别/2012 年 3 月第 1 版

版 次/2012 年 3 月第 1 次印刷

书 号/ISBN 978-7-5037-6484-4/C. 2603

定 价/44.00 元

中国统计版图书, 版权所有, 侵权必究。

中国统计版图书, 如有印装错误, 本社发行部负责调换。

第二版前言

《统计学讲义》第一版出版后,编者即结合教学实践对第一版内容进行了校勘和研讨。

经向使用者征询意见,主编、副主编对以下问题有共同认识:

1. 本讲义主要适用于经济学、管理学门类本科各专业统计学课程教学的需要,也可用于有关专业统计学课程教学或教学参考。
2. 本讲义考虑到相应专业的专业建设,以及统计学对相应专业发展的支撑需要,特别是能满足经济学、管理学门类本科各专业需要的统计学课程内容应包括统计决策、多元统计方法、非参数统计方法。
3. 统计学学科分类应考虑到统计学科的最新发展。
4. 属于应用统计范畴、与现实结合紧密的内容应以具体使用的方法为准。例如,我国政府统计中编制的各种指数。
5. 例题内容要丰富,但风格要一致。
6. 应坚持本讲义第一版立足课堂教学的原则,满足课堂教学“既要加大信息量,又要精讲多练”的要求。

为此,在这次新版的讲义中,增加了统计决策(斯琴执笔)、多元统计方法(陈志芳执笔)、非参数统计方法(王志刚执笔)三章内容;由程士富执笔对统计指数、相关与回归分析两章进行了修订;毛志勇对附录的概率与概率分布进行了修订;其余修订内容由五位主编、副主编负责执笔。最后,由主编负责对初稿和修订稿的修改、总纂和定稿。

编者仍然建议使用本讲义最好在三年级开设本课程(学期课),且先修课程至少包括:高等数学、概率论、经济学或管理学基础;一般为4学分,教学时数应在60—80学时。其中,可选讲多元统计方法、非参数统计方法、统计决策和国民经济统计概述中二至三章的内容。

李国晖、包月霞、郭继龙、陈丽老师等参加了本讲义新版的有关工作。同时,对本讲义能适时顺利出版而付出辛勤劳动的中国统计出版社、中国统计出版社陈悟朝博士表示感谢。

虽经修订再版,本讲义仍会有一些不足之处,恳请读者(特别是同行)不吝赐教。我们将不胜感激。

编者

2011年5月于呼和浩特

第一版前言

统计学是一门研究统计活动规律和方法的方法论科学。

在国家教育部分别于2000年9月、2001年12月公布的《(面向21世纪课程教材)全国普通高等学校工商管理类课程教学基本要求》、《(面向21世纪课程教材)全国普通高等学校经济学类课程教学基本要求》(以下简称《基本要求》)中都包括有统计学。在国家教育部的《普通高等学校本科专业目录和专业介绍(1998年颁布)》中,统计学是经济学门类、管理学门类各专业的主要课程或主干学科。1987年,原国家教委组织中外专家论证,将包括统计学在内的10门现代经济管理人才必须掌握的基本经济理论课、工具技能课和主要应用经济课列为“财经类专业课程”。当然,就课程本身内容来说,其重要性是不言而喻的。

但在教学实践中,我们又感到“强调思想,强调方法,重在讲授统计学中的基本概念、基本方法、基本思想”的课堂教学原则很难实现。因为,满足这一原则要求的教材难觅。这也是我们编写本讲义的初衷之一。我们力求按照课堂面授教学的逻辑安排相应的课程内容。当然,我们更考虑到了21世纪统计学科发展和统计科技发展的规律,参照公认的统计学学科体系,并结合我国实际组合课程内容,以适应专业、学科和课程内容的发展;并注意协调好相关课程的内容,做到不重不漏。其课程内容要兼顾两个方面的要求:一是满足本科专业教育的一般要求;二是满足相应专业教育所必需的专业基础要求。按照“厚基础、宽口径、重应用、适于专业方向展开”的思路确定讲义内容。

本讲义主要适用于非统计学专业本科相关专业统计学课程教学的需要。建议在三年级开设本课程,且先修课程至少包括:高等数学、概率论、经济学或管理学基础;一般为4学分,教学时数应在60—72学时。当然,本讲义也可做为统计学专业相应课程的参考教材。

通过对本课程的学习,应使学生了解统计学的性质、特点和基本学科分类;了解统计学的基本概念、基本思想,系统地掌握统计学的基本方法;能够将统计学的思想和方法运用于现实问题的分析研究之中,为进一步

深入系统地学习统计学专业知识、本专业的其他知识奠定基础。

本讲义在编写过程中参考并汲取了众多统计理论专著、论文、教材等的有益内容,在此恕不一一提及,谨一并表示诚挚的谢意。同时,也汲取了部分非公开出版物的有益内容。例如:1996年5月,由程士富、冯利英、郝晓兰编写的《内蒙古财经学院〈统计学教学大纲〉实施细则》;2004年3月由内蒙古财经学院统计学系组织制定,冯利英执笔、程士富审核,用于非统计学专业教学用的《(内蒙古财经学院)统计学教学大纲》。

根据《基本要求》本讲义应列其中的“概率与概率分布”部分的内容,考虑到先修课程《概率论》已经讲授,本讲义为保持内容的一致性,将其作为附录列出。考虑到与《计量经济学》教学内容的衔接,“相关与回归分析”中的多元线性回归分析、非线性回归分析部分的内容没有更多展开。

本讲义由主编、副主编提出编写大纲后,由内蒙古财经学院统计与数学学院的部分教师执笔各章节的初稿。具体分工如下:第一章:程士富;第二章第一、二、四节:程士富,第三节:曹美丽,第五节:王春枝;第三、四、五、六章:程士富;第七章:包一玫、程士富;第八章:冯利英;第九章:田永棋、程士富;第十章王来拴、程士富;第十一章:杜金柱;附录(概率与概率分布):孙春花。最后,由主编负责对初稿的修改、总纂和定稿。另外,王志刚、海小辉、李国晖、包月霞、郭继龙、苏娟老师等参加了本讲义数据搜集、编制统计表、绘制统计图、整理参考文献等工作。

同时,对本讲义能适时顺利出版而付出辛勤劳动的中国统计出版社、中国统计出版社杨映霜编审表示感谢。

本讲义会有一些不足之处,恳请读者(特别是同行)不吝赐教。我们将不胜感激。

编 者

2009年6月于呼和浩特

目 录

第一章 绪论	(1)
§ 1.1 统计学的对象和性质	(1)
§ 1.2 统计学的基本方法和学科体系	(6)
§ 1.3 统计学的基本概念	(8)
思考与练习题	(11)
第二章 统计数据的搜集与整理	(12)
§ 2.1 数据的计量与类型	(12)
§ 2.2 统计数据的搜集	(16)
§ 2.3 统计数据的整理	(19)
§ 2.4 统计表和统计图	(25)
§ 2.5 频数分布的图示和应用	(31)
思考与练习题	(38)
第三章 数据分布特征的描述与应用	(39)
§ 3.1 集中趋势的测度	(39)
§ 3.2 分布离散程度的测度	(51)
§ 3.3 偏度与峰度的测度	(58)
思考与练习题	(60)
第四章 抽样与抽样分布	(62)
§ 4.1 抽样中的基本概念	(62)
§ 4.2 常用的抽样方法	(63)
§ 4.3 抽样分布	(65)
§ 4.4 中心极限定理的应用	(77)
思考与练习题	(79)

第五章 参数估计	(80)
§ 5.1 参数估计的基本问题	(80)
§ 5.2 一个总体参数的区间估计	(84)
§ 5.3 两个总体参数的区间估计	(92)
§ 5.4 样本容量的确定	(102)
思考与练习题	(106)
第六章 假设检验	(108)
§ 6.1 假设检验的基本问题	(108)
§ 6.2 一个总体参数的检验	(114)
§ 6.3 两个总体参数的检验	(124)
思考与练习题	(136)
第七章 方差分析	(138)
§ 7.1 方差分析的意义	(138)
§ 7.2 方差分析方法	(141)
思考与练习题	(148)
第八章 相关与回归分析	(150)
§ 8.1 相关与回归分析的基本问题	(150)
§ 8.2 简单线性相关与回归分析	(153)
§ 8.3 多元线性相关与回归分析	(168)
§ 8.4 非线性相关分析与非线性回归分析	(173)
思考与练习题	(176)
第九章 时间序列分析初步	(178)
§ 9.1 时间序列概述	(178)
§ 9.2 时间序列的对比分析	(180)
§ 9.3 趋势变动分析	(188)
§ 9.4 季节变动分析	(193)
§ 9.5 循环变动和不规则变动分析	(196)
思考与练习题	(197)
第十章 统计指数	(201)
§ 10.1 统计指数的概念与分类	(201)
§ 10.2 总指数的编制方式	(204)
§ 10.3 统计指数体系及因素分析	(217)
§ 10.4 几种常用的经济指数	(224)
思考与练习题	(231)

第十一章 应用多元统计分析	(234)
§ 11.1 多元统计分析概述	(234)
§ 11.2 聚类分析	(237)
§ 11.3 判别分析	(241)
§ 11.4 主成分分析	(243)
§ 11.5 因子分析	(246)
第十二章 非参数统计初步	(256)
§ 12.1 关于非参数检验的一些常识	(256)
§ 12.2 单样本中心位置的非参数检验	(257)
§ 12.3 两独立样本的位置检验	(262)
§ 12.4 两相关样本位置参数的非参数检验	(266)
§ 12.5 K 个相关样本的 Kruskal—Wallis 检验	(268)
§ 12.6 两独立样本的 Siegel—Tukey 尺度检验	(269)
第十三章 统计决策	(272)
§ 13.1 统计决策的基本概念	(272)
§ 13.2 不确定型决策	(276)
§ 13.3 一般风险型决策	(278)
§ 13.4 贝叶斯决策	(285)
第十四章 国民经济统计概述	(295)
§ 14.1 国民经济统计的基本内容	(295)
§ 14.2 国民经济统计的主要分类	(299)
§ 14.3 国民经济统计的基本指标	(302)
§ 14.4 国民经济统计的常用分析指标	(309)
思考与练习题	(313)
附录	
概率与概率分布	(314)
标准正态分布表	(336)
t 分布表	(339)
χ^2 分布表	(341)
F 分布表	(344)
参考文献	(352)

学习内容提要:本章作为全书的总论,将阐述统计与统计学、统计的产生与发展、统计学的分科以及基于统计内容的统计学的几个基本概念。希望在此基础上,使学生能建立起对本课程全部内容的基本认识。

第一章 绪 论

§ 1.1 统计学的对象和性质

一、统计的三种涵义及其关系

统计一词在一般意义上可以有三种涵义,即:

(1)统计工作或称统计(工作)活动。它是搜集、整理和分析各种现象的现实方面——数量方面有关资料的一种实践活动及其过程。

(2)统计数据或称统计资料。它是包括现象的数量资料和有关情况在内的统计(工作)活动的结果,是一组有明确意义的数字或符号。其中,数量资料是它的主要内容。它可以单纯的表现为现象数量表现的多少,也可以表现为有关现象之间的数量对比关系。

(3)统计学或称统计理论。它是在统计实践的基础上概括出来的关于统计知识的系统论证和结论。例如,包括本讲义所述内容在内的统计一般理论和应用理论。

统计的三种涵义可概括为两个基本方面:统计实践和统计理论。

在我们的日常工作和生活中,到处都有统计实践,它包括统计活动和统计数据两个基本内容,它们是统计实践过程与结果的关系。例如,上课时老师要统计一下出席的人数;球类比赛时解说员通过统计竞赛双方的进攻次数和成功率等解说比赛进程和客观状况;企业管理人员每天都要掌握的生产销售情况和利润额;人们可以通过报刊和电视等各种媒体获得关于一国的国内生产总值、物价指数和经济增长率的数据;等等。

对统计理论或统计学的定义不一而足,例如:

统计学是一门收集、整理、显示和分析统计数据的科学,其目的是探索数据内在的数量规律性;

统计学是关于统计数据的科学;

统计学是研究搜集、整理和分析大量事物数量关系和变化的科学;

统计学是根据从总体中随机取出的样本里所获得的信息来推断关于总体的性质的一门学科;

.....

本讲义对统计理论或统计学的定义:它是一门研究统计全过程的活动规律和统计方法的方法论科学。其中,统计活动和统计方法的对象是以客观现象的存在为前提的现象的数量表现,即统计数据。在这个意义上可以说,统计学是关于统计数据的科学。

在当代科学体系中,当作一门学科,统计学有着基础学科的意义。

统计实践和统计理论具有一般意义上的实践与理论的关系。统计实践是统计理论产生和形成的前提和基础,统计理论是对统计实践的理论概括和知识体系的系统论述。

二、统计的产生和发展

作为一种社会活动的统计实践活动萌芽于奴隶社会初期,即产生于有阶级的社会。据可见的文字记载,公元前 2100 年左右,夏朝有文字根据的人口总数为 13553923 人;公元前 3050 年,古埃及为建造金字塔征集建筑费,曾进行过全国的人口与财产的普查。

总之,统计是适应国家管理的需要而产生的。作为一种社会实践活动的基本职能,统计实践活动用于国家管理;作为国家管理的工具或手段,在当今社会仍然存在。同时,这也意味着统计是以一定现象的存在为基础的。当然,这并不是说现代统计也只是社会经济统计。因为,就统计产生的历史范畴而言,无论中外,都滥觞于人口统计。而作为人口统计对象的人口现象具有双重属性,它首先是一种自然现象,同时也是一种社会现象。因此,作为一种社会实践活动而产生的统计活动,既适用于社会现象,也适用于自然现象。

统计实践活动在封建社会有所发展。其相应的标志性的表现有两个方面。一方面,其实践活动范围有所扩大。作为一种社会活动,它不仅用于国家管理,而且用于社会经济生活的其他方面。例如,中国在公元 3 世纪至 6 世纪的魏晋南北朝时期,为征收田赋所进行的调查统计活动;16 世纪末法国财政大臣编制的法国财政统计——《商情报告》;等等。另一方面是提出了一些最基本的统计思想。例如,我国商鞅(约前 395~前 338,战国时期政治家、思想家,先秦法家代表人物)的“强国知十三数”的思想:竟内仓口之数,壮男壮女之数,老弱之数,官士之数,以言说取食者之数,利民之数,马、牛、刍藁之数。欲强国不知十三数,地虽利,民虽众,国愈弱至削。(《汉书·艺文志·商君书·去强篇》)

在资本主义社会,统计实践活动有了较大发展。它有两个显著标志:第一,从统计实践活动本身来看,实践范围进一步扩大。它包括两个方面的内容:从统计实践活动产生最初,作为社会实践活动用于国家管理的单一功能,扩大为应用于社会经济生活的各个方面和其他方面;从行政地域来看,各国方法实现了交流。第二,从统计实践与统计理论的关系来看,统计实践为统计理论的产生提供了必要条件。它也包括两个方面的内容:统计实践活动提出了一些需要给予理论解释的实践问题;某些成功的经验需要理论的支持,以便进一步发展和完善实践活动。

在现代社会里,统计的存在基础是一切客观存在,统计实践活动已随处可见,无所不在。没有统计,现代社会就不是一个完整的社会。

当作理论的统计学源于 13 世纪末叶的国势调查,到 17 世纪中叶,从几个不同的研究领域汇聚而成。

在经济领域和社会领域的应用是其理论产生的第二个源头。其中,英国的威廉·配第(William Petty, 1623~1687, 英国经济学家, 科学家, 哲学家, 古典政治经济学创始人)的《政

治算术》提出了比较系统的统计思想。这是威廉·配第对统计学发展的主要贡献。一般,可以以《政治算术》做为统计理论产生的标志,我们将其所表述思想称之为统计学。威廉·配第在《政治算术》中用大量的数字对英国、法国、荷兰三国的经济实力进行比较,用数字、重量、尺度等定量的方法进行分析 and 比较,表达他的思想和观点。正如该书原序所说:“因为和只使用比较级或最高级的词语以及单纯作思维的论证相反,我却采用了这样的方法(作为我很久以来就想建立的政治算术的一个范例),即用数字、重量和尺度的词汇来表达我自己想说的问题,只进行能诉诸人们的感官的论证和考察在性质上有可见的根据的原因”(【英】威廉·配第,陈冬野(译),商务印书馆 1978 年 5 月第一版,正文第 8 页)。卡尔·马克思(Karl Marx, 1818~1883,全世界无产阶级的伟大导师、科学共产主义的创始人)在《资本论》中评价配第“是政治经济学之父。在某种程度上也可以说是统计学的创始人”。查理·达芬南(Charles D'Avenant, 1656~1714,英国晚期重商主义的代表人物)对《政治算术》的定义:关于政治事物的数字推理艺术。

统计学的另一个创始人是英国的约翰·格朗特(John Graunt, 1620~1674,英国经济学家、人口统计学创始人之一)。他在 1662 年出版的《对死亡表(或死亡率公报)的自然观察和政治观察(关于宗教、管理、商业、增殖、空气、疾病及本市各种变化)》一书中通过大量观察的方法,研究并发现了人口与社会现象中很重要的数量规律。如新生婴儿的性别比例稳定在 14:13;男性在各年龄组中死亡率高于女性;新生婴儿的死亡率较高;一般疾病与事故的死亡率较稳定,而传染病的死亡率波动较大;等等。在研究中,格朗特不但探索了人口变化和发展的一些数量规律,而且还对伦敦市总人口数量做出了较科学的估计。同时还提出了一些基本的统计思想:管理的艺术和真正的政治,就是使臣民过着和平生活和人丁兴旺;为使各个社会成员能够各得其所,搞清数目字是必要的。

统计学的第三个源头是古典概率论,奠基人包括法国的布莱士·帕斯卡((Blaise Pascal, 1623~1662,法国著名的数学家、物理学家、哲学家和散文家)和费马(Pierre de Fermat, 1601~1665,法国著名数学家,被誉为“业余数学家之王”)。早在他们之前,已有一些数学家在研究赌博中的数量规律了。意大利诗人阿利盖利·但丁(Dante Alighieri, 1265~1321)早在 15 世纪就讨论过掷三颗骰子可能出现的各种点数问题。16 世纪中叶,意大利科学家伽利略·伽利雷(Galileo Galilei, 1564~1642,物理学家、天文学家和哲学家,近代实验科学的先驱者)讨论了掷三颗骰子出现 10 点次数多于 9 点次数的原因。在数学家们对机会游戏研究的基础上,帕斯卡和费马通过通信的方式,将赌博中出现的各种具体问题,归纳为一般的概率原理,为后来概率论和统计学的发展奠定了重要的基础。自 17 世纪中叶上述几位科学家从不同角度开始了统计学研究后,经过几代统计学家的努力,历经两个半世纪,到 19 世纪末建成了古典统计学(主要是描述统计学)的基本框架。在此期间的 18 世纪下半叶,由拉普拉斯(Laplace, 1749~1827,数学家、天文学家,天体力学的主要奠基人、天体演化学的创立者之一,分析概率论的创始人)将概率论全面引入统计学的研究领域,才使得统计理论有了较大发展。但对统计学的发展更有影响的当首推比利时人——阿道夫·凯特勒(Adolphe Quetelet, 1796~1874,数学家、天文学家、人类学家、统计学家)。凯特勒运用概率论原理进行的研究主要是关于人口、犯罪、人体测量等问题。例如,犯罪常数。凯特勒综合几个方面的研究成果,提出了“平均人”的理论;等等。凯特勒对统计学发展的贡献,更在于为众多的研究者提出了统计学的一些基本问题:(1)统计学的应用领域应扩大到自然领域,并与数理

原则相联系;(2)对统计学的性质应有一定的认识:它是实质性科学还是方法论科学,抑或实质性科学兼方法论科学?

由于现代社会的发展和其他各种实质性理论的发展,统计理论已发展成一门独立的学科,并应用于自然科学、工程技术、农学、医学、军事科学等领域。甚至连纯文科领域的法律、历史、语言、新闻等都越来越重视对统计的使用。

20 世纪初,大工业的发展对产品质量检验问题提出了新的要求,即只抽取少量产品作为样本对全部产品的质量好坏做出推断。因为大批量产品要做全面的检验既费时、费钱、又费人力,加之有些产品质量的检验要做破坏性检验,全部检验已不可能。1907 年,英国的戈塞特(1876~1937,英国数学家)提出了小样本 t 统计量。利用 t 统计量就可以从大量的产品中只抽取较小的样本完成对全部产品质量的检验和推断。这样就使统计学进入了现代统计学(主要是推断统计学)的新阶段。以后经过著名统计学家费希尔(R. A. 费希尔, Ronald Aylmer Fisher, 1890~1962, 英国统计学家和遗传学家)给出了 F 统计量、最大似然估计、方差分析等方法 and 思想,奈曼(Neyman, 1894~1981, 统计学家)和卡尔·皮尔逊(Karl Pearson, 1857~1936, 英国数学家、哲学家,现代统计学的创始人之一,被尊称为统计学之父)的置信区间估计和假设检验,亚伯拉罕·瓦尔德(Abraham Wald, 1902~1950, 罗马尼亚裔美国统计学家)的序贯抽样和统计决策函数等,到 20 世纪中叶构筑完成了现代统计学的基本框架。

从 20 世纪 50 年代以来,统计理论、方法和应用进入了一个全面发展的新阶段。一方面,统计学受计算机科学、信息论、混沌理论、人工智能等现代科学技术的影响,新的研究领域层出不穷,如多元统计分析、现代时间序列分析、贝叶斯统计、非参数统计、线性统计模型、探索性数据分析、数据挖掘等。另一方面,统计方法的应用领域不断扩展,几乎所有的科学研究都离不开统计方法。因为不论是自然科学、工程技术、农学、医学、军事科学,还是社会科学都离不开数据,要对数据进行研究和分析就必然要用到统计方法。

我们可以看看下面的一些例子。

【例 1.1】 一个家庭新生婴儿的性别可能是男的,也可能是女的。在过去没有实行计划生育时,有的家庭几个孩子都是男孩,有的都是女孩,但多数是有男有女。从表面上看,新生婴儿的性别比例似乎没什么规律可循。如果对新生婴儿的性别进行大量观察,即观察成千上万个,就会发现性别比例还是有规律的,即婴儿总数中男孩要多于女孩,大致为每生育 100 个女孩就有 107 个左右的男孩。这个 107:100 的比例就是新生婴儿男女性别的数量规律,它是由人类社会长期遗传与发展的结果。因为人类社会要延续、要发展,就要保持男女人数的大致相同。若在新生儿中,男多于女,不是性别比例失衡了吗?是的。但男孩的死亡率高于女孩,到了中青年时,男女人数就大致相等了。进入中老年后,男性的死亡率仍然高于女性,导致男性的平均预期寿命比女性短,长寿男性要少于女性。但人口总数中的男女人数大体相当。对人类性别比例的研究是统计学的起源之一,也是统计方法探索的最早的数量规律之一。

【例 1.2】 我们知道,在掷硬币或掷骰子的游戏中,若随机地掷一次硬币或骰子,我们不能事先确定出现正面还是反面或者某个点数。也就是说个别游戏或试验中充满了不确定性或偶然性。机会游戏或赌博正是利用了这种不确定性和偶然性才能够吸引人。但当我们进行大量观察,即不断做重复试验时,就会发现掷一枚均匀硬币出现正面和反面的次数会大

体相同,即比值接近 $1:1$ 。试验的次数越多,出现正面和反面的可能性就越接近 $1/2$ 这一稳定的数值。同样,在掷一枚均匀骰子时,出现 $1、2、3、4、5、6$,其中任一点数的可能性也都接近 $1/6$ 。这里的 $1/2$ 和 $1/6$ 就是掷硬币和掷骰子出现某一特定结果的概率,也就是我们探索的现象数量表现的规律性。

【例 1.3】 在进行农作物试验时,如果其他试验条件相对固定,我们会发现某种粮食作物的产量会随着施用某种化肥数量的增加而增加。当开始增加施肥量时,产量增加较快。以后增加同样的施肥量,粮食产量的增加量逐渐减少。当施肥量增加到一定数量,产量不再增加。这时如果再增加施肥量,产量反而会减少。这一施肥量与粮食产量的数量关系就是我们要探索的规律性。当我们从大量试验数据中用统计方法探索到施肥量与产量之间的数量关系,就可以考虑施用肥料的费用并选择最佳的施肥效果了。

【例 1.4】 与例 1.3 相似但又不同的是商品广告费用与销售额的关系。一般而言,随着广告费用的增加,商品销售额会相应增加。对于不同的商品,投入相同的广告费用带来的销售额显然是不尽相同的。即使是同一种商品,投入的第一个 10 万元广告费、第二个 10 万元广告费和第三个 10 万元广告费等所带来的销售增量也是不同的。统计方法就是要从广告费与销售额的数据中找出其内在的数量规律,从而确定最佳广告策略。

三、统计学的研究对象、特点和性质

就统计学理论体系内容的逻辑而言,统计学是一门研究统计全过程的活动规律及其统计方法的方法论科学。或者说,统计学的研究对象是:以现象及其数量表现为基础的、可以表明统计活动规律的统计方法。其中,统计活动及其统计方法的对象是以客观现象存在为前提的现象的数量表现,即统计数据。

一般,统计活动过程包括统计设计、统计调查、统计整理、统计分析四个阶段。统计设计就是对统计工作各个方面和各个环节进行通盘考虑和安排,其结果形成工作方案。统计调查就是根据统计方案的要求采用科学的方法,有计划地搜集统计数据的工作过程。统计整理就是对调查资料加以综合汇总,使之条理化、系统化的工作过程,它是统计活动的初步成果。统计分析就是对经过加工汇总的资料进行分析研究,计算各项分析指标,揭示现象的发展趋势和比例关系,阐明现象和过程的特征,并做出科学解释的过程。它是统计工作的决定性阶段。

当作统计学对象的统计方法,既可以是适用于某一统计活动阶段的专门方法,也可以是适用于整个统计活动过程的通用方法。一般,有以下一些特点:

- a. 统计方法要以各种客观现象及其数量表现的存在为基础。
- b. 统计方法的具体形式是多种多样的,但总可以在一定条件下揭示和说明现象的统计规律。
- c. 统计方法总体及其方法体系是统计活动规律的客观反映。统计方法总体是指当作方法论的认识方法、工作方法、组织方法及其相应的计算方法;考虑到各种方法之间的联系,我们又称之为方法体系。
- d. 作为一般方法应适用于所有的统计活动领域。

由统计学的研究对象内容所决定,统计学是一门方法论性质的应用科学。统计学有数

量性、总体性和具体性的特点。统计实践的认识对象是现象的数量方面,包括现象表现的数量多少;现象数量表现之间的数量关系;质与量互变的数量界限。从而,以统计实践为基础的统计学也具有数量性的特点,即统计学是关于(统计)数据的科学。由于统计实践是以现象总体的数量特征作为自己的工作对象。从而,统计学也就必须在总体的范围内进行研究,也即统计学具有总体性的特点。统计所面对的现象数量表现总是现象具体条件、过程和结果的表现。因此,以现象数据为基础的统计学是抽象的,也是具体的。

§ 1.2 统计学的基本方法和学科体系

一、统计学的理论和方法论基础

统计学要以一切科学方法论为理论基础,特别要以质与量的辩证统一关系规律作为统计学最直接的理论基础。

同时,数学为统计学和统计方法提供了数学基础,并在一定程度上出现了某种融合。因此,在需要把统计方法、数学方法进行融合,或者用数学支持统计理论的方面,都要充分利用数学基础和数学思想、数学方法,以完善统计方法体系。当然,从思维方式来说,统计学和数学在研究目标和思想方法上是有差异的。数学有以公理系统为基础,以演绎为基本思想方法的逻辑体系。它属于少数可以和世界具体事物无关的自成体系的学科,可以完全脱离现实问题而存在。数学的应用也是演绎的。统计学是服务于客观存在的现实领域的,其理论本身是演绎的。但其理论前提既有相应的公理系统,也有对现实问题的归纳。并且,统计学的应用还是归纳的。因此,统计学以归纳为它的基本思想;在逻辑上,统计的实现要归纳和演绎并用。

另外,统计学在研究某一具体领域问题时,还要以相应领域的实质性科学作为统计学的理论基础。

二、统计学的基本方法

统计学的基本方法之一是大量观察法。它是说统计研究现象及其过程,要从总体上加以考察,要对现象总体中的全部或足够多数单位或具体表现进行调查研究,要占有充分的实际数据作为统计认识现象的基础。

统计学的另一基本方法是综合分析法。它包括两个方面的内容:综合——是指对于大量观察所获得的资料,要运用各种综合指标来反映总体的一般数量特征。例如,整理汇总出总量指标、相对指标、平均指标。分析是指对综合指标进行分解和计算,以便据以研究现象总体的数量差异和数量关系。例如,对现象的原始数据进行统计分组、动态趋势分析、因素分析、综合平衡分析。

再有就是归纳推断法。它是指根据从个别到一般、从具体事实到抽象概括的推理方法,由观察到的各单位特征进而归纳出关于总体的某种信息的统计方法。

应该说,统计的各种具体方法无不派生于此。当然在基本方法的意义上,这些方法也是研究统计理论的方法。

三、统计学学科体系

统计学的内容十分丰富,研究与应用的领域非常广泛。从不同的角度,可以有不同的分类。

(一)按照统计实践的发展阶段分

1. 描述统计学

它是研究如何取得反映客观现象的数据,并通过图表形式对搜集的数据进行加工处理和显示,进而通过综合、概括与分析得出反映客观现象的规律性数量特征,并为进一步的统计推断提供根据的统计理论。其内容包括统计数据的搜集方法、数据的加工处理方法、数据的显示方法、数据分布特征的概括与分析方法等。

2. 推断统计学

它是研究如何根据样本数据去推断总体数量特征的方法,并在对样本数据进行描述的基础上对统计总体的未知数量特征做出以概率形式表述的推断统计理论。其相应的推断统计方法包括参数估计、假设检验、预测和其他方法。

统计学分为描述统计学和推断统计学,一方面反映了统计发展的前后两个阶段,另一方面也反映了统计方法研究和探索客观事物内在数量规律性的先后两个过程。当然,在当代统计的实际应用中二者并非截然分开。统计研究过程的起点是数据,终点是探索到客观事物总体内在的数量规律性。要达到统计研究的目的,如果我们搜集到的是总体数据(如普查),则经过描述统计之后就可以达到探索内在数量规律性的目的了;但如果我们所获得的数据只是研究总体的一部分数据,要探索到总体的数量规律性,就必须应用概率论的理论,并根据样本整理出的信息对总体做出科学的推断。推断统计是现代统计学的核心和统计研究工作的关键环节。但如果没有描述统计收集可靠的数据并提供有效的样本信息,即使很高明的统计学家和很科学的推断方法也难于得出准确的结论。因此,在一定意义上说,推断统计又是以描述统计为前提的。或者说,这两者之间是互为条件、互相依存的。

(二)按照统计学的理论内容分

1. 理论统计学

理论统计学或称一般统计学。它是指统计学的数学原理,它主要研究包括概率论在内的统计学的一般理论和统计方法的数学理论。

2. 应用统计学

应用统计学是研究如何应用统计方法去解释、解决实际问题的统计理论。

统计学是一门数据科学。由于在自然科学、社会科学的所有研究和实际工作中都要通过数据来分析和解决问题,统计方法的应用就自然而然地扩展到几乎所有的研究领域。例如,统计方法在物理研究中的应用就形成了统计物理,统计方法在生物学中的应用形成了生物统计,统计方法在医学中的应用形成了医疗卫生统计,统计方法在风险管理与保险中的应用形成了保险精算学,统计方法在微观企业管理中的应用形成了管理统计,等等。这些应用统计学的不同分支所应用的基本方法都是一样的,即都是描述统计和推断统计的主要方法。

但由于各应用领域都有其特殊性,统计方法在应用中就具有了不同的特点及其相应的理论。

学习统计理论就应该注意到这两种理论,并把它们当作一个整体去学习。当然,在不同的学习阶段、针对不同的统计问题,对二者可以有所侧重。

(三)按照一般的学科构成内容分

1. 基础学科——概率论、数理统计学等。
2. 综合性学科——统计史、统计学原理等。
3. 部门或专门统计学科——其中,部门统计学科包括农业、工业、商业统计学等;或宏观统计学、微观统计学等;专门统计学科包括统计指标学、统计调查学等。
4. 边缘交叉学科——统计组织管理学、系统学与统计学等。

(四)我国的国家分类标准

我国的国家分类标准(GB/T13745 — 2009《学科分类与代码》)中的一级学科的统计学包括:统计学史、数理统计学、应用统计数学、经济统计学、科学技术统计学、社会统计学、人口统计学、环境与生态统计学、生物与医学统计学、统计学其他学科。

另有,属于理论统计学的概率论(数学和统计学的交叉学科):几何概率、概率分布、极限理论、随机过程(包括正态过程与平稳过程、点过程等)、马尔可夫过程、随机分析、鞅论、应用概率论、概率论其他学科;以及,属于心理学(一级学科)的二级学科——心理统计(包括心理统计原理和心理统计方法);属于经济学(一级学科)的三级学科——世界经济统计;属于社会学(一级学科)的三级学科——劳动统计学。

§ 1.3 统计学的基本概念

一、统计总体和样本

(一)统计总体

统计总体简称总体,它是由客观存在的、具有某种相同性质的许多个别事物(或个别单位)组成的整体。组成总体的每一个个别事物或个别单位称作总体单位,或称作单位、个体,它是组成总体的元素。

例如,总体可以是:①所有的中国人;②所有的内蒙古选民;③某电信经营商的所有客户;④某汽车制造厂某条生产线去年生产的所有汽车;⑤中国汽车维修行业的所有库存汽车零部件;⑥去年北京市某品牌快餐餐馆所有窗口销售的某一品种快餐;等等。同时,相应的总体单位是:①每一个中国人;②每一个内蒙古选民;③某电信经营商的每一位客户;④某汽车制造厂某条生产线去年生产的每一辆汽车;⑤中国汽车维修行业的每一个库存汽车零部件;⑥去年北京市某品牌快餐餐馆所有窗口销售的某一品种的每一份快餐;等等。

总体是在一定研究目的条件下,用来规范统计活动范围的。也就是说,它的作用在于确定统计研究范围。统计总体所包含的个体数量称作总体单位数即总体容量,一般用 N 表示。按照总体单位数的有限性进行区分,可以有有限总体和无限总体。

总体有三个由总体单位情况所决定,且同时具备的特点(或性质):大量性、同质性、变异