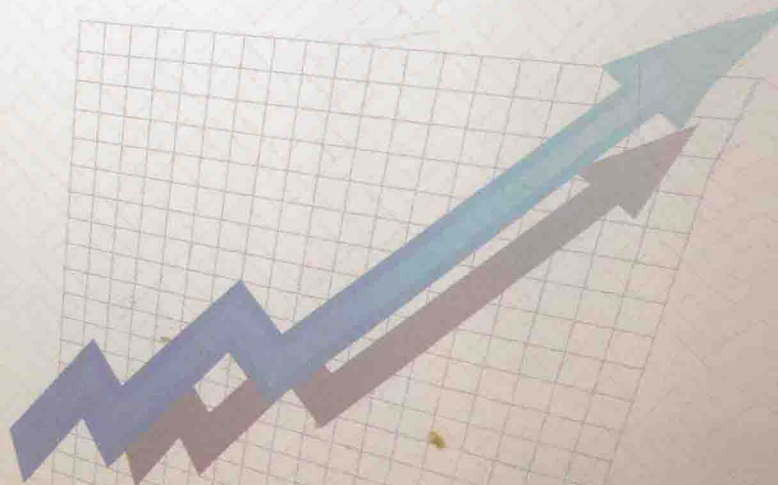


统计学

STATISTICS

甘平 主编



 北京理工大学出版社
BEIJING INSTITUTE OF TECHNOLOGY PRESS

统 计 学

主 编 甘 平
副主编 徐 瑞 冯金丽
张 涛 曹石云
参 编 詹浩勇

内 容 简 介

本书结合应用本科学生的特点,从“应用”角度为学生设计了引导案例和应用案例;在理论上增加了“文科性质”统计思想的辩证探讨;在数理上,为几大分布的关系进行了推演;在数据处理上,以 Excel 为工具。以此,增加同学们的学习兴趣和帮助学生理解。

全书共 12 章,主要内容包括绪论、统计数据的搜集、统计数据的整理与展示、数据的概括性度量、概率与概率分布、抽样推断及参数估计、假设检验、方差分析、相关分析与回归分析、时间序列分析、统计指数和统计决策。

本书可作为高等院校经济管理类、应用统计类各专业学习统计学课程的入门教材,也可供从事社会、经济和管理等研究和实际工作的人员阅读参考。

版权专有 侵权必究

图书在版编目(CIP)数据

统计学/甘平主编. —北京:北京理工大学出版社, 2014. 11

ISBN 978 - 7 - 5640 - 9926 - 8

I. ①统… II. ①甘… III. ①统计学 IV. ①C8

中国版本图书馆 CIP 数据核字(2014)第 257400 号

出版发行 / 北京理工大学出版社有限责任公司

社 址 / 北京市海淀区中关村南大街 5 号

邮 编 / 100081

电 话 / (010) 68914775 (总编室)

82562903 (教材售后服务热线)

68948351 (其他图书服务热线)

网 址 / <http://www.bitpress.com.cn>

经 销 / 全国各地新华书店

印 刷 / 三河市华骏印务包装有限公司

开 本 / 787 毫米 × 1092 毫米 1/16

印 张 / 19

字 数 / 469 千字

版 次 / 2014 年 11 月第 1 版 2014 年 11 月第 1 次印刷

定 价 / 49.00 元

责任编辑 / 陈莉华

文案编辑 / 陈莉华

责任校对 / 孟祥敬

责任印制 / 李志强

图书出现印装质量问题,请拨打售后服务热线,本社负责调换

前 言

统计学是一门提供数据信息搜集、处理、归纳和分析的理论与方法的科学。随着社会的发展,统计的运用领域越来越广泛,不管是在经济管理领域,还是自然、工程领域的研究,人们对数量分析与统计分析都提出更高的要求。现有的统计学教材种类繁多,从事统计学教学的教师背景也各异,但怎样设计适合不同学生层次的统计学教材?为此,我们汇集了本校长期从事统计学教学的几位老师,试图编写一本适合应用本科院校的统计学教材。

本书结合了应用本科学生的特点,从“应用”角度为学生设计了引导案例和应用案例;在理论上也增加了“文科性质”的统计思想、统计复杂性、统计科学性的辩证探讨;在数理上,为几大分布的关系进行了推演;在数据处理上,以 Excel 为工具。以此,增加同学们的学习兴趣和降低学习的难度。

全书共 12 章。第 1 章,绪论;第 2 章,统计数据的搜集;第 3 章,统计数据的整理与展示;第 4 章,数据的概括性度量;第 5 章,概率与概率分布;第 6 章,抽样推断及参数估计;第 7 章,假设检验;第 8 章,方差分析;第 9 章,相关分析与回归分析;第 10 章,时间序列分析;第 11 章,统计指数;第 12 章,统计决策。任课老师可结合专业和人才培养方案特点,酌情选择授课内容。

本书由甘平担任主编,由徐瑞、冯金丽、张涛和曹石云担任副主编,具体分工如下:第 1 章、第 2 章、第 4 章由甘平编写;第 3 章、第 9 章、第 12 章由徐瑞编写;第 5 章、第 6 章由张涛编写;第 7 章、第 8 章由曹石云编写;第 10 章、第 11 章由冯金丽编写。詹浩勇参与了本书的资料搜集与整理工作,并在编写过程中提出了许多宝贵的意见和建议。

在本书的写作过程中,参阅了许多国内外的参考文献资料,并引用了部分例题和习题,在此向这些文献的作者们表示感谢,同时感谢陈有禄教授给予本书的建议。

由于编者水平有限,书中难免有疏漏或错误之处,恳请同行专家和读者不吝赐教,以便我们修改与完善。

编 者

目 录

第1章 绪论	(1)
1.1 数学与统计学	(1)
1.2 统计之亚哲学	(2)
1.2.1 统计哲学思想的发展历程	(2)
1.2.2 现代五大统计学派	(4)
1.2.3 统计与统计学的含义	(4)
1.2.4 统计学的研究对象	(4)
1.2.5 统计学的性质	(6)
1.2.6 统计学的分科	(6)
1.2.7 统计学的研究方法	(7)
1.3 统计学的发展和应用	(8)
1.3.1 统计学的产生和发展	(8)
1.3.2 现代统计学的应用和发展趋势	(10)
1.4 统计学的基本概念	(12)
1.4.1 统计总体和总体单位	(12)
1.4.2 标志	(13)
1.4.3 统计指标和统计指标体系	(14)
1.4.4 变量	(16)
1.4.5 统计描述和统计推断	(17)
1.5 Excel 软件简介	(17)
1.5.1 中文 Excel 概述	(17)
1.5.2 Excel 基本操作	(20)
思考与练习	(24)
第2章 统计数据的搜集	(25)
2.1 统计数据的类型与来源	(25)
2.1.1 统计数据的类型	(25)

2 统计学

2.1.2 统计数据的来源	(28)
2.1.3 统计数据的搜集方法	(31)
2.2 调查方案设计	(34)
2.2.1 统计指标和指标体系的设计	(34)
2.2.2 统计数据采集方案设计	(37)
2.3 调查问卷设计	(38)
2.3.1 问卷的意义及内容	(38)
2.3.2 问卷回答的类型和方式	(39)
2.3.3 问卷调查方法的种类	(43)
2.3.4 问题的种类和设计原则	(43)
2.3.5 问题的表述	(45)
2.3.6 设计答案应该注意的原则	(45)
2.3.7 统计数据的误差	(45)
思考与练习	(57)
第3章 统计数据的整理与展示	(58)
3.1 统计数据的预处理	(59)
3.1.1 数据审核	(59)
3.1.2 数据筛选	(59)
3.1.3 数据排序	(59)
3.2 品质型数据的汇总	(59)
3.2.1 品质型数据的频数分布	(59)
3.2.2 品质型数据的图形展示	(61)
3.3 数值型数据的汇总	(63)
3.3.1 频数分布	(63)
3.3.2 数值型数据的图形展示	(65)
3.3.3 累积频数分布及其图形展示	(67)
3.3.4 探索性数据分析: 茎叶图	(68)
3.4 多变量之间关系的描述	(70)
3.4.1 列联表	(70)
3.4.2 散点图	(71)
3.4.3 时间序列图	(72)
3.5 Excel 在统计数据整理和展示中的应用	(73)
3.5.1 Excel 在数据预处理中的应用	(73)
3.5.2 Excel 在品质型数据汇总中的应用	(77)
3.5.3 Excel 在数值型数据汇总中的应用	(81)
思考与练习	(84)
第4章 数据的概括性度量	(88)
4.1 总规模度量	(89)
4.1.1 总量指标的意义	(89)

4.1.2	总量指标的种类	(89)
4.1.3	应用总量指标应注意的问题	(90)
4.2	相对度量	(90)
4.2.1	相对指标的意义	(90)
4.2.2	相对指标的种类及其计算方法	(90)
4.2.3	应用相对指标应注意的问题	(92)
4.3	集中趋势的度量	(92)
4.3.1	统计平均数的含义与作用	(92)
4.3.2	数值平均数	(93)
4.3.3	位置平均数	(100)
4.4	离散趋势的度量	(104)
4.4.1	变异指标的含义与作用	(104)
4.4.2	极差、四分位差与异众比率	(105)
4.4.3	方差与标准差	(106)
4.4.4	相对位置的测量: 标准分数	(108)
4.4.5	经验法则	(108)
4.4.6	变异系数	(108)
4.5	偏态和峰度的测度	(109)
4.5.1	统计动差	(109)
4.5.2	偏度	(110)
4.5.3	峰度	(110)
4.6	用 Excel 计算描述统计量	(111)
	思考与练习	(114)
第 5 章	概率与概率分布	(118)
5.1	事件及其概率	(118)
5.1.1	随机事件和样本空间	(118)
5.1.2	概率和频率	(119)
5.1.3	古典概型	(120)
5.2	离散型概率分布	(121)
5.2.1	随机变量	(121)
5.2.2	一维离散型随机变量的概念	(121)
5.2.3	常见的离散型随机变量及其分布	(122)
5.2.4	多维随机变量	(125)
5.2.5	边际分布列	(125)
5.2.6	随机变量的相互独立性	(126)
5.2.7	随机变量函数及其分布	(127)
5.2.8	数学期望的定义及性质	(128)
5.2.9	方差的定义及性质	(130)
5.3	连续型概率分布	(131)

4 统计学

5.3.1 问题的提出	(131)
5.3.2 分布函数的定义与性质	(132)
5.3.3 一维连续型随机变量及其分布	(134)
5.3.4 一维连续型随机变量函数的分布	(139)

思考与练习	(141)
-------------	-------

第6章 抽样推断及参数估计

(143)

6.1 统计推断的相关概念	(143)
6.2 抽样分布	(144)
6.2.1 统计量	(144)
6.2.2 样本统计量的分布 (常见的抽样分布)	(145)
6.3 抽样估计	(149)
6.3.1 点估计	(149)
6.3.2 点估计中的矩估计法及最大似然估计法	(150)
6.4 Excel 在抽样推断中的应用	(153)

思考与练习	(155)
-------------	-------

第7章 假设检验

(157)

7.1 假设检验的基本原理	(157)
7.1.1 问题的提出	(158)
7.1.2 假设检验的基本思想	(158)
7.1.3 假设检验的基本步骤	(159)
7.2 一个总体参数的假设检验	(159)
7.2.1 总体均值的检验	(159)
7.2.2 总体比例的检验	(161)
7.2.3 总体方差的检验	(162)
7.3 两个总体参数的假设检验	(163)
7.3.1 两个总体均值之差的检验	(163)
7.3.2 两个总体比例之差的检验	(166)
7.3.3 两个总体方差之比的检验	(166)
7.4 假设检验的错误分析与控制	(168)
7.5 Excel 在假设检验中的应用	(169)
7.5.1 两个总体均值差的检验	(169)
7.5.2 两个总体的方差比检验	(170)

思考与练习	(171)
-------------	-------

第8章 方差分析

(174)

8.1 方差分析的基本原理	(175)
8.1.1 基本概念	(175)
8.1.2 方差分析的基本思想	(176)
8.1.3 方差分析的适用条件	(177)

8.2 单因素的方差分析	(177)
8.2.1 提出假设	(177)
8.2.2 构造检验统计量	(177)
8.2.3 做出判断	(178)
8.3 双因素的方差分析	(178)
8.3.1 无交互作用的双因素方差分析	(179)
8.3.2 有交互作用的双因素方差分析	(181)
8.4 Excel 在方差分析中的应用	(183)
8.4.1 单因素方差分析的 Excel 操作	(183)
8.4.2 无交互作用的双因素方差分析的 Excel 操作	(184)
8.4.3 有交互作用的双因素方差分析的 Excel 操作	(185)
思考与练习	(187)
第 9 章 相关分析与回归分析	(191)
9.1 相关分析	(192)
9.1.1 基本概念	(192)
9.1.2 相关系数	(193)
9.1.3 相关系数的显著性检验	(194)
9.2 一元线性回归分析	(195)
9.2.1 回归分析	(195)
9.2.2 经典线性回归模型	(195)
9.2.3 一元线性回归模型的估计	(198)
9.2.4 一元线性回归模型的拟合优度	(200)
9.2.5 一元线性回归模型的假设检验	(201)
9.2.6 残差分析	(202)
9.2.7 回归结果的报告	(203)
9.2.8 一元线性回归模型的预测	(203)
9.3 多元线性回归分析	(204)
9.3.1 多元线性回归模型	(204)
9.3.2 多元线性回归模型的估计	(205)
9.3.3 多元线性回归模型的拟合优度	(206)
9.3.4 多元线性回归模型的假设检验	(207)
9.4 Excel 在相关与回归分析中的应用	(208)
9.4.1 Excel 在相关分析中的应用	(208)
9.4.2 Excel 在一元线性回归分析中的应用	(210)
9.4.3 Excel 在多元线性回归分析中的应用	(211)
思考与练习	(213)
第 10 章 时间序列分析	(218)
10.1 时间序列分析的基本问题	(219)
10.1.1 时间序列的概念及作用	(219)

10.1.2 时间序列的分类	(219)
10.1.3 时间序列的编制原则	(221)
10.2 时间序列的水平分析	(221)
10.2.1 发展水平和平均发展水平	(221)
10.2.2 增长量与平均增长量	(226)
10.3 时间序列的速度分析	(227)
10.3.1 发展速度和增长速度	(227)
10.3.2 平均发展速度和平均增长速度	(229)
10.4 时间序列的趋势分析	(230)
10.4.1 时间序列的4种成分	(230)
10.4.2 长期趋势测定与预测的意义	(231)
10.5 时间序列的季节变动分析	(235)
10.5.1 同期平均法	(235)
10.5.2 趋势剔除法	(237)
10.6 Excel在时间序列分解分析中的应用	(238)
思考与练习	(246)
第11章 统计指数	(250)
11.1 统计指数概述	(250)
11.1.1 指数的概念	(250)
11.1.2 指数的作用	(251)
11.1.3 指数的种类	(251)
11.2 综合指数法和平均指数法	(253)
11.2.1 综合指数法	(253)
11.2.2 平均指数法	(257)
11.3 指数体系与因素分析	(259)
11.3.1 指数体系的概念和作用	(259)
11.3.2 因素分析	(260)
11.3.3 综合指数因素分析	(260)
11.3.4 指数体系中的因素推算	(263)
11.4 常用的统计指数	(263)
11.4.1 工业生产指数	(264)
11.4.2 股票价格指数	(264)
11.4.3 居民消费价格指数(CPI)	(264)
11.5 Excel在统计指数中的应用	(266)
思考与练习	(270)
第12章 统计决策	(273)
12.1 统计决策概述	(273)
12.1.1 引言	(273)
12.1.2 统计决策的基本要素	(274)

12.1.3 收益矩阵	(274)
12.1.4 机会损失	(275)
12.2 决策标准	(275)
12.2.1 期望收益	(275)
12.2.2 期望机会损失	(276)
12.2.3 风险回报率	(276)
12.3 不确定性决策的策略	(277)
12.3.1 最大最大策略	(277)
12.3.2 最大最小策略	(278)
12.3.3 最小最大后悔值策略	(278)
12.3.4 等可能性策略	(278)
思考与练习	(280)
附录 常用统计表	(282)
附表1 正态分布概率表	(282)
附表2 t 分布临界值表	(284)
附表3 χ^2 分布临界值表	(285)
附表4 F 分布临界值表 ($\alpha=0.05$)	(286)
参考文献	(288)

第1章

绪论

导入案例

硬币实验提问：当我把一枚硬币抛向空中，请问同学们，硬币落下来在正面的概率是多少？

引出的思考：

(1) 硬币是一个高度抽象的概念。

(2) 硬币落下来有多种可能：除了正反还有可能立在地上；当然，还有可能粉碎……

(3) 即使硬币是绝对均匀对称的（实际上不是），也会受到风向、磁场、手法等诸多因素的影响。

(4) 我们可能还要回到古典概率的两个基本假设：实验结果是有限的；每种结果的可能性都一样。但假设成立吗？

(5) 可能我们要回到概率的频率定义了，比较“科学的回答”是“依据我们长期无数次观察、实验和‘统计’，发现正面和反面出现的次数相近”。可问题依然来了，你怎么能保证你观察到的“够全面”和“够深刻”？或者，你怎么能获得“完整的信息”？

1.1 数学与统计学

“在终极的分析中，一切知识都是历史；

在抽象的意义下，一切科学都是数学；

在理性的基础上，所有的判断都是统计学。”

——C·R·劳

统计学与数学都是研究数量的关系和数量的规律，都要与大量的数字打交道。现代统计学运用了大量的数学方法，如概率论、数理统计、模糊数学、线性代数和微积分等。有人认为统计学是数学的一个分支，这些人认为概率论是统计学的基础理论；也有人认为数学与统计学不同，统计学与数学有密切的联系，但两者存在本质的区别，这两个学科各有独立的研究领域和研究特点。

统计学和数学都是利用各种数学公式进行数字演算,但两者研究的数是存在差别的。统计学的数总是与所研究的客观对象联系在一起的,称之为“数据”,统计的过程就是从客观对象中抽取出其数量表现,得到有关的数据。统计数据是有具体的实际含义的,它反映着某一现象的质。数学所研究的数字,是高度抽象的数字,它并不反映现象的质。

统计学和数学都是研究数量规律的,统计学研究的是具体的实际现象的规律,它从客观实际中搜集数据,进行统计处理后又将这些处理结果返回到实际中,并解释这些结果的意义。而数学研究的是抽象的数量规律,它撇开具体的对象,以公理为基础,以最一般的研究探索数量的联系和空间的形式。

从研究方法看,统计学和数学的研究方法不尽相同,统计学根据实验或调查,观察大量的个别现象,对所观察的个别现象加以归纳,并判断总体的情况,实质上,统计学的研究方法是归纳与演绎相结合的方法,其中归纳占主要地位。数学的研究方法主要是逻辑推理和演绎论证的方法。

数学与统计学各成体系,两门学科各有各的研究对象和研究方法,但两者关系密切,数学为统计学提供了数量分析方法论的基础,尤其是数学中的概率论,研究的是随机现象的数量关系和变化规律,它从数量方面揭示了偶然与必然、个别与一般、局部与总体之间的辩证关系,为现代统计学奠定了基础。笔者的理解是:最原先的数学是为了求解寻找方向的,是为人类认知世界或科学决策搭建的“最高层次的逻辑”。当连续被打断,不确定性为决策和认知真理带来风险时;当变量越来越多,系统越来越复杂时,便出现了概率论和统计,指导人们在黑暗中摸索、判断和前行。更为具体的指导,便是各种应用统计学(如会计学)。

1.2 统计之亚哲学

“不确定性知识 + 所含不确定性量度的知识 = 可用的知识”。

——南丁格尔

知识存在于真实的或确定的谬误之中,推理是获得这样的知识的工具。而从科学家的观点来看,一切知识都不是绝对正确的,通过任何方式所得到的一个科学理论知识,如果能引导出可接受限度内的预示,就能获得认可。这个新的理论如果能提供更好的预示,就将取代已经存在的科学理论。而从统计学的角度来看,从经验或实验中获取的知识是不稳定的,但在实际生活中,不管这些已有的知识如何贫乏,也不得不以此做出决策。统计学关注的是如何探知由观察数据获取的知识中的不确定性的量度,以及如何明确在最小损失下的最优决策。

——C·R·劳

1.2.1 统计哲学思想的发展历程

统计学实际上是介于理性思维和艺术思维之间的一门学科。它告诉人们:如何在现实世界中搜集信息或搜集数据,组成有限的资料;如何使用这些有限的资料,兼顾理性思维与艺术思维的平衡,采取合适的行动。目的是更好地掌握环绕我们周围的不确定性。因此,统计学与人类的行为有关,是模拟人类行为的操作方式的科学。从这个观点来考查,统计思想在

人类文明之初就有了。但很长时间以来,统计的思想并未形成科学,统计学的形成是20世纪30年代的事。尽管人类文明的结晶,基本上是在人类经验的基础上得到的,已有许多可以由某些理性的思维所代替,但也有许多现象是人类统计行为的直接结果,目前还无法或者说很难用理性的思维所代替。统计学就是这样:根据以往的经验资料,断言采取某些行动是合适的,而其背后的内在规律以及此种行动的原因,并不是统计学重点关心的问题。

与统计学形成密切相关的学科之一是概率论。概率论是研究随机现象或不确定现象的科学,它起源于对赌博的研究。考古证据显示,赌博的历史源远流长。在人类之初,相信一切机会和命运皆由天定,当有争执或遇到难以决定的事,往往以抽签解决。我国古代的卜卦,西方古代的投掷兽骨,以对未来做一些预测,都是此类问题的古典例子。与概率同义的一个词——Stochastic (随机的),就是源自古希腊文,指一个人能预测未来。由于经验的积累也使人们得到一些简易的概率计算方法。不过直到17世纪中叶之前,概率论的进展并不是很大。由于James (Jakob) Bernoulli (1654—1705) 及 de Moivre (1667—1754) 等人的努力,在18世纪中叶前,一些关于求赌博中所涉及概率的计算方法得到了比较完善的发展。18—19世纪是概率论应用广泛的时期。不少人认为是因为人们对赌博的兴趣促使了概率论的发展,de Mere 也常被视为狂热的赌徒。事实上,以赌局来讨论概率,只是取其方便易懂而已。数学家探讨问题的原动力,主要还是出于对随机现象的研究在科学上的需要,这种需要直接来源于统计学的需要。

17世纪,统计学开始萌芽。为了征兵和税收,欧洲的各国政府开始搜集如出生、死亡及结婚的人口统计学的资料。1662年,英国的Graunt (1620—1674) 发表了首度关于出生与死亡的记录“统计式”的处理报告,显示了搜集这些数据的好处。同样在16—19世纪的时候,由于商业的兴起,为了减少船难等意外造成直接经济损失,保险及精算业产生了。精算师制作了寿命表及意外事件表,然后计算损失以定保险费。人口统计学家及精算师在搜集资料不久后便发现,他们资料中的变异形式,可与赌局中的一些结果相对应。例如,连续几个小孩都是男孩或都是女孩,与投掷铜板连续出现正面或反面类似。因此,自然现象的变异(如一个即将出生婴儿的性别),似乎可以比作上帝在玩一赌局,以此当作模式来分析问题是恰当的。若对赌局有充分的了解,有助于预测如男孩出生率等。反过来说,由于赌局中有些结果较易出现,由观测值也可推测上帝在玩什么赌局(此即统计推断,Statistical Inference)。因此发展概率理论,是为了适应统计上的需要,此角色概率论直到今日仍在扮演。

在18—19世纪,由于统计的需要使得概率论迅速发展,并形成公理化,使其能与数学中的一些传统领域,如代数、几何及分析等分庭抗礼、各领风骚了。但关于统计学的概念形成,还是没有多大发展。直到20世纪30—40年代,英国伦敦大学的Neyman (1894—1981)、Pearson (1895—1980)、Fisher (1890—1962),以及在此留学的中国统计学家许宝禄 (1910—1970)、美国多元分析专家 Hotelling (1895—1973) 等,才开始对统计学的基本哲学思想进行原始的思考。以人类行动为原始模式来分析,人们根据原始资料所进行的统计推断,无外乎对未知变量的估计、检验和预测。事实上,这几种行动,原始人类是在没有科学指导下做出的。统计学的目的,不单单是给出人们解决问题的可操作性方法,重要的是能够比较所用方法的好坏。统计推断的结论,不保证其结论是百分之百的自然规律,但它能证明他们的行为是现有资料下的可以采取的最佳行为。由此就奠基和形成了统计推断的基本哲学思想。

在统计推断的基本哲学思想形成以后,促使人们进行一种统计学的哲学思维,这种思维已远远不同于数学思维。数学讲究演绎,但统计学讲究的是归纳。这样也就使得统计学成了不是人们通常所理解的应用数学即数学的应用,而是有自己的独立哲学思想的学科。

目前人们把统计学的哲学思想建立的时期作为统计学形成的标志。为了有别于原始的只限于数据收集和数学应用的统计方法,此时的统计学被称为数理统计学。中国统计学家许宝禄在这一时期写出了一系列出色的论著,成为中国近代数学界中在概率统计方向达到世界水平的第一位杰出的学者。他也被作为数理统计学的创始人之一,在世界许多国家的大学里被挂像纪念。但在我国并没有得到应有的重视。这种现象产生的原因可能众多,而统计的哲学思想与数学哲学思维的不同,应是其中的主要一个。统计学在数学的哲学框架下,是很难想象能够正常发展的,统计学只有远离数学才可能生存。为了弄清统计的哲学思想与数学哲学思维的不同之处,必须先弄清统计学的各个学派的哲学思想基础。

1.2.2 现代五大统计学派

统计学是模拟人类思维的行为方式的。人类对有限资料的处理方式,或人类的思维行动准则,无外乎以下5种:

(1) 对未知系统进行信息的搜集,并研究数据的可识别性,从而认识系统的结构与性质。

(2) 弄清未知系统的变化波动范围,以求在没有充分了解未知系统的内部性质的情况下,达到对未知系统的控制。

(3) 总结自己的信息,使其与所能搜集的未知系统的数据结合,考虑它们之间的协调性质,以期使我们的行为与未知系统协调,或使未知系统的各个部分运行协调。

(4) 分析未知系统的功能输出,了解它们的共同特点及目标特性,以期最佳地利用未知系统的功能。

(5) 在对未知系统采取行动时,也可以不去研究未知系统本身,而是去研究对于未知系统,自己如果采取某些行动时,它可能对自己带来的效益和损失,以期使采取的行动对我们自己最有利。

1.2.3 统计与统计学的含义

在不同场合人们对统计主要有3种不同的理解,即统计工作、统计资料和统计学。统计工作就是搜集、整理和分析客观现象总体数量方面的活动过程。统计资料,就是统计实践活动所取得的各种信息,其中主要是反映统计对象总体数量特征的数字资料。统计学是关于统计的原理、原则和方法的科学。它主要研究和阐明搜集、整理与分析客观现象统计资料的原理和方法,同时也揭示统计对象的本质特征和发展规律。统计工作、统计资料和统计学三者之间既有区别,又有密切联系。三者以统计工作(统计活动)为基础或核心。统计资料是统计工作的成果;统计学是统计工作的经验总结和理论概括。反过来,统计学又是指导统计工作的原理、原则和方法,推动统计工作水平和质量不断提高,使统计资料更加准确、及时和全面。

1.2.4 统计学的研究对象

空间是三维的,时空是四维的,人们所面临的客观事物更是多维的,统计学正是要从复

杂多变的客观事物中,挖掘出其蕴含的客观规律,为人们的各种行为活动提供有力的参考,避免行为的盲目性。由于客观事物往往可以用其数量表现,因此,要找出客观事物的内在规律,首先要认识客观事物,那么就必须通过试验或调查来搜集有关数据,并且加以整理、归纳和分析,以便对客观事物规律性的数量表现做出合理的描述。

由此可见,统计学的研究对象是客观事物的数量特征和数量关系。统计学也就是关于数据搜集、整理、归纳、分析的方法论科学,其目的是探索数据的内在数量规律性。

统计学的研究对象具有以下特点:

1. 数量性

客观现象有着质和量两个方面的表现,根据质和量的辩证统一研究现象的数量特征,从数量上认识现象的性质和规律性,这是统计研究的基本特点。统计运用科学方法搜集、整理、分析反映现象特征的数据,并通过统计指标反映现象的规模、水平、比例、速率及其变动规律。认识现象的数量表现,是深入研究现象质的表现的前提和基础。现象的数量方面包括数量多少、数量关系、质和量互变的数量界限等。数量关系指各种平衡关系、比例关系和依存关系,如总供给与总需求的平衡关系、各产业间的比例关系、消费与收入之间的依存关系等。客观现象往往具有复杂性的特点,现象之间具有多方面的联系。在研究现象的数量方面时,必须把握现象的全貌,反映现象发展变化的过程,必须紧密联系现象的具体内容和质的特征,这是统计学与数学的一个重要区别。例如,一个国家的人口数量、结构和分布,国民经济的规模、发展速度,人们的生活水平等,都是反映基本国情和基本国力的指标,通过这样的一系列指标才能对整个国家有一个客观、清晰的认识。由此可见,数量性是统计研究对象的特点之一。对于用定性方式表述的客观现象,则应该将其转换为数量形式。例如,为了反映某产品的质量情况,用合格与不合格来表示,这时可分别记合格与不合格等。

2. 大量性

大量性也称总体性,统计研究的对象总是由大量同类事物构成的总体现象的数量特征。个别和单个事物的数量表现是可以直接获取的,一般不需运用统计研究方法。例如,要了解某名工人的工作情况,查一查生产记录就可以了,但如果要了解全体工人产量的分布、差异和一般水平等,就要用统计方法来进行计算和分析。统计对总体现象的数量特征进行研究时,是通过对组成总体的个别事物量的认识来实现的。例如,在人口普查中,通过对每一户家庭的人口状况进行调查,根据所取得的资料,编制人口总数、人口结构(性别、年龄、民族、职业等结构)、人口分布、人口出生率、人口死亡率等指标来反映一个国家或一个地区的人口总体状况。

个别事物有很大的偶然性,大量事物具有共性,统计学正是要从大量的客观事物中找出其共性,即规律性。从对个体数量特征的观测入手,运用科学的统计方法获得反映总体一般特征的综合数量,这是统计的又一基本特征。

3. 变异性

变异性是指组成研究对象的各个单位在特征表现上存在差异,并且这些差异是不可以按已知条件事先推断的。例如,要研究某地区大学生的消费行为,每个学生的家庭收入有差异,学生的消费偏好有差异,消费品的市场价格也不稳定。这时就需要研究大学生的平均消费、家庭平均收入、消费偏好和消费品的市场价格等因素,如果每个大学生不存在这些差

异,只要调查一个学生相关消费行为,就可以知道整个地区的大学生消费行为,这时也就不需要做统计了。正是因为研究对象的各单位存在差异性,统计方法才有了用武之地。

1.2.5 统计学的性质

根据统计学前面的定义,很容易知道统计学的性质:统计学是一门认识方法论科学,具体说它是研究如何搜集数据、整理数据并分析数据,以便从中做出正确推断的认识方法论科学。

之所以统计学具有这样的性质,是因为:首先,统计学是为了揭示客观事物的规律性;其次,为了达到这个目的,需要各种统计方法来认识事物的真面目。因此,统计学是认识客观事物的方法论科学。

统计学和数学都是研究数量关系的科学,它们之间既有联系又有区别。一方面,数学以抽象的概念和方法研究各种数量关系和空间形式,而统计学则是对客观现象在质和量的相互联系中研究其数量方面,揭示其数量变动的规律性,这是它们之间的本质区别;另一方面,数学又为统计学提供大量的计算分析方法,尤其是数理统计不仅用于研究社会经济现象,也可用于研究自然技术现象。工业产品、农副产品的抽样调查、生产过程的检验和控制等就是数理统计方法在社会经济领域中的应用。

统计学在研究客观现象的数量特征和数量关系时,必然要以相关的科学的基本理论和基本知识为指导,如经济学、社会学、物理学、生物学、心理学等。而且,统计学的基本理论在各个领域中的应用形成了各种专门统计学,如经济统计学、人口统计学、科技统计学、金融统计学、经营统计学、心理统计学等。统计学与相关科学的结合同时也促进了统计理论和方法的发展。

1.2.6 统计学的分科

从研究的层次和方法来看,统计学可以分为描述统计学和推断统计学;从研究的主要目的来看,统计学可以分为理论统计学和应用统计学。

1. 描述统计学和推断统计学

描述统计学研究如何取得反映客观现象的数据,并通过图表形式对所收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。推断统计学则是研究如何根据样本数据去推断总体数量特征的方法,它是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。描述统计和推断统计是统计学的两个组成部分。描述统计是整个统计学的基础,推断统计则是现代统计学的主要内容。由于在对现实问题的研究中,所获得的数据主要是样本数据,因此,推断统计在现代统计学中的地位和作用越来越重要,已成为统计学的核心内容。

2. 理论统计学和应用统计学

理论统计学是指主要研究统计学的一般理论、方法和数理基础的统计学。应用统计学是研究如何应用统计理论和方法去解决实际问题的统计学。由于在各种科学研究中都需要通过数据分析来解决实际问题,因而,统计学的应用几乎扩展到了所有的科学研究领域,从而形成了一个庞大的应用统计科学体系。例如,统计理论和方法在生物学中的应用形成了生物统计学,在医学中的应用形成了医疗卫生统计学,在农业试验、育种等方面的应用形成了农业统计学,在经济和社会科学研究领域的应用形成了经济统计学、管理统计学、社会统计学、