



高职高专教学改革系列规划教材

TONGJIXUE JICHU YU YINGYONG

统计学基础与应用

张文法 孔建华 主编



化学工业出版社

高职高专教学改革系列规划教材

统计学基础与应用

张文法 孔建华 主编

王雅平 刘旭 耿万里 闫克信 副主编



化学工业出版社

· 北京 ·

《统计学基础与应用》内容简单实用,充分考虑到统计工作的作业流程,并考虑到企业管理对于统计职能的要求。本书由两部分内容组成,上篇为统计学基础,主要介绍统计学认知、统计设计与调查、统计整理和统计分析方法等。下篇为统计应用,主要介绍企业经营环境统计分析、企业经营投入统计、企业经营产出统计和企业经营综合统计分析等。

本书体例设计充分考虑到学习者的需求。每一项目开篇部分设有知识目标、能力目标和项目导入,使学习者在短时间内对后续学习的内容有初步宏观的认识,帮助其合理安排学习计划。每一项目结尾部分设有思考题和课内实训,促使学习者对前面学习的知识进行巩固和提高。在理论介绍部分还穿插动脑筋和扩展阅读等栏目,促使做中学,拓展学生的视野。

本书不仅可以作为高职高专院校经管类学生使用的教材,也可以作为成人教育、电大、函授大学的教材及企业管理人员、统计工作者的自学参考书。

图书在版编目 (CIP) 数据

统计学基础与应用/张文法,孔建华主编. —北京:
化学工业出版社, 2015. 7

ISBN 978-7-122-23548-0

I. ①统… II. ①张…②孔… III. ①统计学 IV. ①C8

中国版本图书馆 CIP 数据核字 (2015) 第 067866 号

责任编辑:蔡洪伟 于 卉 王 可
责任校对:王素芹

装帧设计:王晓宇

出版发行:化学工业出版社(北京市东城区青年湖南街 13 号 邮政编码 100011)

印 装:三河市延风印装有限公司

787mm×1092mm 1/16 印张 12 字数 323 千字 2015 年 6 月北京第 1 版第 1 次印刷

购书咨询:010-64518888 (传真:010-64519680) 售后服务:010-64519661

网 址: <http://www.cip.com.cn>

凡购买本书,如有缺损质量问题,本社销售中心负责调换。

定 价: 26.00 元

版权所有 违者必究

前 言

企业管理工作就是要从繁杂的数据中探究规律,从错综交织的现象中发现问题。管理者在实现这个目标的过程中应用最多的工具就是统计,所以,学习统计学的基本理论,掌握统计应用的方法,具备收集、整理和分析统计数据的能力已成为人们职业素养的重要组成部分。现在大多数院校经管类专业都把统计能力的培养列入人才培养的计划中。本书编写就是基于这样的现实,努力为读者提供一本实用的参考书籍。

本书的编写充分考虑到统计工作的作业流程,并考虑到企业管理对于统计职能的要求,在新的高职高专人才培养的要求下,本着理论“必须、够用”、注重实践的原则,我们组织和编写了教材的内容。本教材由两部分组成。上篇为统计学基础,主要介绍统计学认知、统计设计与调查、统计整理和统计分析方法等。下篇为统计应用,主要介绍企业经营环境统计分析、企业经营投入统计、企业经营产出统计和企业经营综合统计分析等。

本书在体例设计上充分考虑到学习者的需求。每一项目开篇部分设有知识目标、能力目标和项目导入,使学习者在短时间内对后续学习的内容有初步宏观的认识,帮助其合理安排学习计划。每一项目结尾部分设有思考题和课内小实训,促使学习者对前面学习的知识进行巩固和提高。

本书由长期在高校从事统计工作的骨干教师和具有企业统计实践经验的专家共同编写。具体分工如下:山东华宇工学院的张文法、孔建华担任主编,中国人民财产保险股份有限公司的耿万里和山东华宇工学院的王雅平、刘旭和闫克信担任副主编。张文法负责全书编写大纲和体例的设计,并编写项目三、项目四、项目七和项目八;孔建华负责全书的统稿,并编写项目一;王雅平编写项目六和项目十;刘旭编写项目二和项目五;耿万里编写项目九;闫克信编写项目十一;此外,王钢、赵玉娥、孙强参加了本书部分内容的编写和资料整理工作。

由于编者水平有限,书中难免有不当之处,敬请广大读者和专家批评指正,以期不断改进和完善。

编者

2015年2月

目 录

项目 1 统计的认知 001

名人小传	001
知识目标	001
技能目标	001
项目导入	002
1.1 统计学的产生与发展	002
1.1.1 统计活动的产生和发展	002
1.1.2 统计学的产生和发展	002
1.1.3 我国的统计和统计学	003
1.1.4 统计学的分类	004
1.2 统计学的研究对象和方法	005
1.2.1 统计学的研究对象	005
1.2.2 统计学的研究方法	006
1.3 统计学的基本概念	008
1.3.1 统计总体与总体单位	008
1.3.2 样本	009
1.3.3 统计标志	009
1.3.4 统计指标	010
1.3.5 统计指标体系	011
1.3.6 变异、变量和变量值	012
思考与练习	012
课内实训	013

项目 2 统计设计与统计调查 015

名人小传	015
知识目标	015
技能目标	015
项目导入	016
2.1 统计设计概述	016
2.1.1 统计设计的概念、原则与意义	016
2.1.2 统计设计的类型	017
2.1.3 统计设计的内容	017
2.2 统计调查概述	019
2.2.1 统计调查的概念	019
2.2.2 统计调查的基本要求	020
2.2.3 统计调查的种类	020
2.2.4 统计调查的方法	021

2.3 统计调查方案	022
2.4 统计调查的组织方式	023
2.4.1 统计报表	023
2.4.2 专门调查	026
2.5 调查问卷	027
2.5.1 问卷的结构与问题形式	027
2.5.2 问卷设计	028
思考与练习	030
课内实训	031

项目 3 统计整理 035

名人小传	035
知识目标	035
技能目标	035
项目导入	036
3.1 统计整理的含义和步骤	036
3.1.1 统计整理的概念和意义	036
3.1.2 统计整理的步骤	037
3.2 统计分组	038
3.2.1 统计分组的含义	038
3.2.2 组的拟定	040
3.3 次数分布	043
3.3.1 分布数列的概念与种类	044
3.3.2 变量数列的编制	045
3.3.3 次数分布的主要类型	049
3.4 统计表和统计图	050
3.4.1 统计表	050
3.4.2 统计图	052
思考与练习	054
课内实训	056

项目 4 统计分析之指标对比分析法 057

名人小传	057
知识目标	057
技能目标	057

项目导入	058
4.1 指标对比分析法的含义	058
4.1.1 指标对比分析法的概念	058
4.1.2 指标对比分析法的形式	058
4.1.3 注意事项	058
4.2 总量指标	059
4.2.1 总量指标的含义	059
4.2.2 总量指标的分类	059
4.2.3 总量指标的作用	060
4.2.4 总量指标的计量单位	060
4.2.5 总量指标的计算	060
4.2.6 使用总量指标时需要注意的问题	061
4.3 相对指标	061
4.3.1 相对指标的意义	061
4.3.2 相对指标的表现形式	062
4.3.3 相对指标的作用	062
4.3.4 相对指标的种类与计算	062
4.3.5 相对指标的应用原则	065
4.4 平均指标	066
4.4.1 平均指标的意义	066
4.4.2 算术平均数	066
4.4.3 调和平均数	067
4.4.4 几何平均数	068
4.4.5 众数和中位数	069
4.5 标志变异指标	072
4.5.1 标志变异指标的意义	072
4.5.2 标志变异指标的种类	072
思考与练习	076
课内实训	077

项目5 统计分析之动态分析法 ... 079

名人小传	079
知识目标	079
技能目标	080
项目导入	080
5.1 动态数列概述	080
5.1.1 动态数列的概念	080
5.1.2 动态数列的分类	080
5.1.3 编制动态数列的原则	081
5.2 动态数列的水平指标	082
5.2.1 发展水平与平均发展水平	082
5.2.2 增长量和平均增长量	086
5.3 动态数列的速度指标	087
5.3.1 发展速度	087
5.3.2 增长速度	088
5.3.3 平均发展速度和平均增长速度	089

5.4 趋势分析及季节分析	091
5.4.1 长期趋势分析	091
5.4.2 季节变动分析	093
思考与练习	095
课内实训	097

项目6 统计分析之指数分析法 098

名人小传	098
知识目标	098
技能目标	098
项目导入	099
6.1 统计指数的含义	099
6.1.1 统计指数的概念	099
6.1.2 统计指数的性质	099
6.1.3 统计指数的种类	100
6.1.4 统计指数的作用	101
6.2 综合指数	102
6.2.1 综合指数的含义	102
6.2.2 综合指数的编制方法	103
6.3 平均指数	105
6.3.1 平均指数的含义	105
6.3.2 平均指数的编制	105
6.4 指数体系及因素分析	107
6.4.1 指数体系的含义	108
6.4.2 指数体系的作用	108
6.4.3 因素分析	108
思考与练习	114
课内实训	115

项目7 统计分析之抽样推断分析法 117

名人小传	117
知识目标	117
技能目标	117
项目导入	118
7.1 抽样推断的概述	118
7.1.1 抽样推断的概念及特点	118
7.1.2 抽样推断的作用	118
7.1.3 抽样推断相关基本概念	119
7.2 抽样误差	120
7.2.1 抽样误差的含义	120
7.2.2 抽样平均误差	121
7.2.3 抽样极限误差	123

7.2.4 抽样估计	125
7.3 抽样方法	126
7.3.1 简单随机抽样	126
7.3.2 分层抽样	127
7.3.3 系统抽样	128
7.3.4 整群抽样	128
7.4 必要抽样容量的确定	129
7.4.1 必要抽样容量的概念	129
7.4.2 必要抽样容量的计算	129
思考与练习	130
课内实训	132

项目 8 企业经营环境统计分析

134

名人小传	134
知识目标	134
技能目标	134
项目导入	134
8.1 企业经营环境统计概述	135
8.1.1 企业经营环境的含义	135
8.1.2 企业经营环境统计的特点	135
8.2 企业经营的自然环境统计	135
8.2.1 企业自然环境统计的概述	135
8.2.2 企业自然资源的统计方法	136
8.3 企业经营的社会环境统计	136
8.3.1 企业社会环境统计的概述	136
8.3.2 企业社会环境的统计方法	137
8.4 市场环境统计	139
8.4.1 市场环境统计概述	139
8.4.2 市场需求环境调查	141
8.4.3 市场供给环境调查	142
8.4.4 竞争对手调查和竞争能力分析	142
思考与练习	146
课内实训	147

项目 9 企业经营投入统计

148

名人小传	148
知识目标	148
技能目标	148
项目导入	149
9.1 企业经营要素统计概述	149
9.1.1 企业经营要素统计的基本内容	149
9.1.2 企业经营要素统计指标体系	149
9.1.3 企业经营要素统计分析常用方法	149

9.2 企业经营要素统计指标	150
9.2.1 劳动力统计指标及统计分析	150
9.2.2 设备利用统计指标及统计分析	153
9.2.3 原材料、能源统计指标及应用 分析	155
9.2.4 技术投入效益的统计分析	157
9.2.5 财务统计分析	157
思考与练习	158
课内实训	159

项目 10 企业经营成果统计

161

名人小传	161
知识目标	161
技能目标	161
项目导入	162
10.1 企业经营成果统计的概述	162
10.1.1 企业经营成果概述	162
10.1.2 企业经营成果统计的意义	163
10.2 企业经营成果统计指标	163
10.2.1 企业经营成果的实物量统计 指标	163
10.2.2 企业经营成果价值量统计指标	166
思考与练习	172
课内实训	173

项目 11 企业经营综合统计评价

174

名人小传	174
知识目标	174
技能目标	174
项目导入	175
11.1 企业经营综合统计评价概述	175
11.1.1 企业经营综合统计评价的概念 和意义	177
11.1.2 企业经营综合统计评价的步骤	177
11.2 企业综合统计评价的指标 体系	180
11.2.1 企业经济效益综合评价 指标体系	180
11.2.2 企业市场竞争力综合统计评价 指标体系	182
思考与练习	183
课内实训	184
参考文献	186

项目 1

统计的认知

名人小传

威廉·配第（William Petty, 1623~1687）是英国古典政治经济学创始人，统计学家。一生著作颇丰，主要有《赋税论》（写于1662年，全名《关于税收与捐献的论文》），《献给英明人士》（1664），《政治算术》（1672），《爱尔兰政治剖析》（1674），《货币略论》等。

威廉·配第似乎被统计学遗忘了。由他所著《政治算术》，首先由于其诞生于亚里士多德之后，在当时，通过严格的统计调查得出的数据少，而根据经验得出的数据多。其次在人们对于统计学的认识中，人们往往只重视统计调查所得的数据，而忽视了收集数据和分析数据对于统计学的影响。虽说亚里士多德开创了统计学，但也只是单纯的统计调查，而威廉·配第则将统计学领入了收集数据和分析数据的新时代。

《政治算术》的问世，标志着统计学的诞生。《政治算术》是一部用数量方法（即“算数”）研究社会问题（即“政治”）的著作。在书中，威廉·配第以劳动价值论为基础，对英、法、荷三国进行了国情、国力（主要是经济实力）的数量对比分析，以此为依据，为当时英国社会经济的发展出谋划策。



威廉·配第

知识目标

1. 了解统计学的产生、发展。
2. 理解统计学的分类。
3. 掌握统计学研究的对象和方法。
4. 理解和掌握统计学的基本概念。

技能目标

1. 培养应用统计基本理论思想在实际工作中的能力。
2. 认知并能有意识地培养自己具备统计分析的方法。

平时的学习过程中,班长经常说:我们统计一下今天出勤的人数;我们统计一下咱们班学生的生源地。我们看新闻的时候经常听到这样的一句话:据统计,今年3月的CPI增长4.2%……“统计”这个词经常出现在我们的生活当中,那么统计到底是什么意思呢?

1.1 统计学的产生与发展

1.1.1 统计活动的产生和发展

统计是适应社会政治经济的发展和国家管理的需要而产生和发展起来的。统计实践活动先于统计学的产生,早在四五千年前,为适应当时社会经济发展的需要,就开始了各种各样的统计实践活动。

早在奴隶社会时期,当时的统治阶级为了征兵和收税,需要了解土地、人口、粮食和牲畜的数量,就有了人口、土地等政府记录。我国在公元前22世纪已有人口、土地的记载。

封建社会统计有了进一步发展。在中国,历代封建王朝都十分重视统计,战国时期商鞅提出强国应了解13个方面的数字资料,其中包括粮食、各类人口、农业生产资料及自然资源等。

统计的广泛发展开始于资本主义社会。从18世纪起,许多国家先后设立了专门的统计机构,搜集各个方面的统计资料,出版统计刊物,倡议建立国际统计组织,积极推动召开国际统计会议。

1.1.2 统计学的产生和发展

一般来说,统计学的产生和发展可分为三个时期。不同时期各学派之间的相互争论,相互渗透,使统计理论最后发展成为统一的现代统计学。

1.1.2.1 古典统计学时期

从17世纪中叶到18世纪中叶是古典统计学时期,当时主要有政治算术学派和国势学派。

(1) 政治算术学派。该学派产生于17世纪中叶的英国,主要代表人物是威廉·配第。英国古典政治经济学的创始人威廉·配第的代表作《政治算术》一书,是经济学和统计学史上的重要著作。书中用“数字、重量、尺度”等定量的分析工具,对英国和当时的主要发达国家的经济实力进行了比较分析。该书的出版标志着统计学的产生,为此威廉·配第被推举为统计学的创始人,并将其所代表的学派命名为政治算术学派。

(2) 国势学派。该学派产生于德国国势派又称记述学派,创始人是德国的海尔曼·康令,国势学派搜集大量实际资料,分门别类系统地记述了有关国情国力的重要事项,如人口、领土、政治、军事、经济、宗教、地理、风俗、货币等。使用了“统计学”这个名称。但几乎不用数字而只用文字形式对国情国力进行系统地描述,所以人们也把它叫做记述学派,并认为国势学派有统计学之名,而无统计学之实。

1.1.2.2 近代统计学时期

统计学的近代时期是18世纪末到19世纪末。该时期的主要贡献是建立和完善了统计学的理论体系,并逐步形成了以随机现象的推断统计为主要内容的数理统计学和传统的政治经

济现象描述为主要内容的社会统计两大学派。

(1) 数理统计学派。产生于 19 世纪中叶, 创始人是比利时的阿道夫·凯特勒, 在统计学发展中的最大贡献是把概率论引入了统计学, 从而使统计学产生了质的飞跃。凯特勒的研究成果在自然科学、经济学、生物学等科学中得到不断的应用, 逐渐形成一门独立的学科。因此, 被称为数理统计学的奠基人, “近代统计学之父”。

(2) 社会统计学派。产生于 19 世纪后半叶, 创始人是德国的尼克斯, 尼克斯在《独立科学的统计学》中, 提出了把“国势论”作为“国势学”的科学命名, 把“统计学”作为“政治算术”的科学命名, 从而结束了对统计学研究对象长达 200 年之久的争论。

1.1.2.3 现代统计学的发展时期

从 19 世纪末开始, 统计学进入了现代统计学时期。在这个时期, 数理统计学与社会统计学逐步融合成为统一的现代统计学。



扩展阅读

一首统计学的赞歌

在 1985 年, 研究莎士比亚的学者泰勒 (G. Taylor), 从保存在保德联 (Bodelian) 图书馆里的收藏品中, 发现保存了两百多年的一张纸片, 其上记有九节新诗, 此诗只有 429 个字, 没有记载诗的作者。于是提出疑问, 这首诗是莎士比亚的吗? 统计学家瑟思梯德 (R. Thisted) 和埃福伦 (B. Efron), 对于莎士比亚, 以及同时代的几位著名诗人的作品进行了研究, 用纯粹的统计学方法, 对他们在诗中用词规律作了比较分析, 发现那首无名诗与莎士比亚的作品非常一致, 这表明它的作者很可能是莎士比亚。获美国总统奖 (2002 年) 的统计学家 C. R. 劳 (Rao), 称这一结果为“一曲统计学的赞歌”。

两位统计学家考查了莎士比亚在其著作中用词总数量为: 884 647 个, 其中包括重复使用的; 所有不相同的用词总数量为 31 534 个。可见, 每个词平均被使用了 28 次 ($\approx 884\,647/31\,534$)。但是, 每个词被使用的次数并不相等。两位统计学家把被使用了 1 次, 2 次, 3 次…的用词个数逐一记录下来, 这就是他们获得的数据。对其他的诗人也可获得同样数据, 对上述的 429 个字的诗也记录同样数据。他们使用了费歇 (Fisher, A. R.) 1943 年提出的判别法则, 将此诗与诸位诗人的作品进行比较后, 得出了前述结果。

不难理解, 他们的数据记录了作者的风格特征, 与诗作内容联系不大。因此, 可用来解答判断作者是谁的问题。不无遗憾地说, 此方法不能简单地用于中文著作的同类判别问题, 因为中文与西文有本质性的组词差异。

虽然也有不少学者用数据的统计分析方法, 讨论《红楼梦》的前 80 回与后 40 回是否属于同一作者的问题, 但是他们获取数据的指标多与内容密切相关。其中使用某些虚词数量作为指标, 看起来好像与内容无关, 仔细想来并非如此。如此说来, 他们所判断的结果, 只是关于《红楼梦》前 80 回与后 40 回在这些指标上有无差别, 而不是作者是否有别的问题。如何寻找更能表现中文作者风格的指标, 还是有待探讨的问题。

1.1.3 我国的统计和统计学

我国的统计学界, 在新中国成立前也存在着数理统计学派和社会统计学派, 两派的观点都是从外国传来的。新中国成立初期, 认为只有社会经济统计学才是唯一的统计学, 从而在根本上否定了数理统计学是统计学的组成部分, 严重妨碍了整个统计学的发展。改革开放以来, 人们被禁锢的思想终于获得解放, 经过长期、广泛的认识和探讨, 我国统计学学科建设取得了重大突破和质的飞跃。

1996年10月,中国统计学会、中国数理统计学会、中国现场统计学会联合举办全国统计科学研讨会,这次会议达成了中国各统计学科,各统计学派之间相互借鉴、相互融合、共同发展的思想并确立了统计学科体系的基本框架,肯定了统计学是包括社会经济统计学和数理统计学在内的一般方法论性质的科学,这为今后我国统计学的发展奠定了坚实的基础。

1.1.4 统计学的分类

统计方法已被应用到自然科学和社会科学的众多领域,统计学也发展成为由若干分支学科组成的学科体系。

1.1.4.1 从统计方法角度分

从统计方法的构成来看,统计学可以分为描述统计学和推断统计学。

从统计方法研究和统计方法的应用角度来看,统计学可以分为理论统计学和应用统计学。

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。内容包括统计数据的收集方法、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。

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。

描述统计学和推断统计学的划分,一方面反映了统计方法发展的前后两个阶段,同时也反映了应用统计方法探索客观事物数量规律性的不同过程。统计研究过程的起点是统计数据,终点是探索出客观现象内在的数量规律性。在这一过程中,如果搜集到的是总体数据(如普查数据),则经过描述统计之后就可以达到认识总体数量规律性的目的了;如果所获得的只是研究总体的一部分数据(样本数据),要找到总体的数量规律性,则必须应用概率论的理论并根据样本信息对总体进行科学的推断。

显然,描述统计和推断统计是统计方法的两个组成部分。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地位和作用越来越重要,已成为统计学的核心内容。当然,这并不等于说描述统计不重要,如果没有描述统计收集可靠的统计数据并提供有效的样本信息,即使再科学的统计推断方法也难以得出切合实际的结论。从描述统计学发展到推断统计学,既反映了统计学发展的巨大成就,也是统计学发展成熟的重要标志。

1.1.4.2 从统计方法研究和统计方法的应用角度分

从统计方法研究和统计方法的应用角度来看,统计学可以分为理论统计学和应用统计学。

理论统计学是指统计学的数学原理,它主要研究统计学的一般理论和统计方法的数学理论。由于现代统计学用到了几乎所有方面的数学知识,从事统计理论和方法研究的人员需要有坚实的数学基础。此外,由于概率论是统计推断的数学和理论基础,因而广义地讲统计学也是应该包括概率论在内的。理论统计学是统计方法的理论基础,没有理论统计学的发展,统计学也不可能发展成为像今天这样一个完善的科学知识体系。

在统计研究领域,从事理论统计学研究的人相对是很少的一部分,而大部分则是从事应用统计学研究的。应用统计学是研究如何应用统计方法去解决实际问题的。统计学是一门收集和分析数据的科学。由于在自然科学及社会科学研究领域中,都需要通过数据分析来解决实际问题,因而,统计方法的应用几乎扩展到了所有的科学研究领域。例如,统计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。统计方法在经济和社会科学研究领域的应用也形成了若干分支学科。例如,统计方法在经济领域的应用形成了经济统计学及其若干分支,在管理领域的应用形成了管理统计学,在社会学研究和社會管理中的应用形成了社会统计学,在人口学中的应用形成了人口统计学,等等。以上这些应用统计学的不同分支所应用的基本统计方法都是一样的,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,统计方法在应用中又形成了一些不同的特点。

1.2 统计学的研究对象和方法

1.2.1 统计学的研究对象

世间所存在的客观事物是多维的,表现为错综复杂和千头万绪,统计学正是要从复杂多变的客观事物中,挖掘出其蕴涵的客观规律,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,那么就必須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作出合理的描述。

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。

1.2.1.1 数量性

统计学的研究对象是自然、社会经济领域中现象的数量方面,这一特点是统计学(定量分析学科)与其他定性分析学科的分界线。数量性是统计学研究对象的基本特点,因为,数字是统计的语言,数据资料是统计的原料。一切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、共同规定着事物的性质。没有无量的质,也没有无质的量。一定的质规定着一定的量,一定的量也表现为一定的质。但在认识的角度上,质和量是可以区分的,可以在一定的质的情况下,单独地研究数量方面,通过认识事物的量进而认识事物的质。因此,事物的数量是我们认识客观现实的重要方面,通过分析研究统计数据资料,研究和掌握统计规律性,就可以达到我们统计分析研究的目的。例如,要分析和研究国民生产总值,就要对其数量、构成及数量变化趋势等进行认识,这样才能正确地分析和研究国民生产总值的规律性。

1.2.1.2 总体性

统计学的研究对象是自然、社会经济领域中现象总体的数量方面,即统计的数量研究是对总体普遍存在着的事实进行大量观察和综合分析,得出反映现象总体的数量特征和资料规律性。自然、社会经济现象的数据资料和数量对比关系等一般是在一系列复杂因素的影响下形成的。在这些因素当中,有起着决定和普遍作用的主要因素,也有起着偶然和局部作用的次要因素。由于种种原因,在不同的个体中,它们相结合的方式和实际发生的作用都不可能完全相同。所以,对于每个个体来说,就具有一定的随机性质,而对于有足够多数个体的总体来说又具有相对稳定的共同趋势,显示出一定的规律性。例如,对工资的统计分析,我们并不是要分析和研究个别人的工资,而是要反映、分析和研究一个地区、一个部门、一个企业事业单位的总体的工资情况和显示出来的规律性。统计研究对象的总体性,是从个体的实际表现的研究过渡到对总体的数量表现的研究的。例如,工资统计分析,要反映、分析和

研究一个地区的工资情况，先要从每个职工的工资开始统计，然后再综合汇总得到该地区的工资情况，只有从个体开始，才能对总体进行分析研究。研究总体的统计数据资料，不排除对个别事物的深入调查研究，但它是为了更好地分析研究现象总体的统计规律性。

动脑筋：研究对象的总体性特征是否就意味着在统计研究中放弃对典型个体的研究？

1.2.1.3 具体性

统计研究对象是自然、社会经济领域中具体现象的数量方面。即它不是纯数量的研究，是具有明确的现实意义的，这一特点是统计学与数学的分水岭。数学是研究事物的抽象空间和抽象数量的科学，而统计学研究的数量是客观存在的、具体实在的数量表现。统计研究对象的这一特点，也正是统计工作必须遵循的基本原则。正因为统计的数量是客观存在的、具体实在的数量表现，它才能独立于客观世界，不以人们的主观意志为转移。统计资料作为主观对客观的反映，必然是存在第一性，意识第二性，存在决定意识，只有如实地反映具体的已经发生的客观事实，才能为我们进行统计分析研究提供可靠的基础，才能分析、探索和掌握事物的统计规律性。否则，虚假的统计数据资料是不能成为统计数据资料的，因为它违背了统计研究对象的这一特点。

1.2.1.4 变异性

统计研究对象的变异性是指构成统计研究对象的总体各单位，除了在某一方面必须是同质的以外，在其他方面又要有差异，而且这些差异并不是由某种特定的原因事先给定的。就是说，总体各单位除了必须有某一共同标志表现作为它们形成统计总体的客观依据以外，还必须要有所要研究的标志上存在变异的表现。否则，就没有必要进行统计分析研究了。例如，高等院校这个统计对象，除了都是从事高等教育的教学活动这一共同性质之外，各高等院校在隶属主管部门、院校性质、招生规模、专业设置等各方面又有差异。人作为统计数据资料对象，每个人在性别、年龄、工龄、工作性质、工资等方面是会有不同表现的。这样，统计分析研究才能对其表现出来的差异探索统计规律性。

1.2.2 统计学的研究方法

统计学根据研究对象的性质和特点，形成了它自己专门的研究方法，这些基本方法是：实验设计法、大量观察法、统计描述法和统计推断法。

1.2.2.1 实验设计法

统计是要分析数据的，但首先需要考察的是，数据的来源是否合适，实验采集的数据是否符合分析的目的要求。由于安排不科学，使实验数据不能反映现象的真实情况，或不能用以估计总体的数量特征，那么接着的一系列分析工作也就白费工夫了。例如要比较某农作物甲品种和乙品种的收获率高低，分别在两地段播种甲品种和乙品种，结果获得甲品种单位面积产量高于乙品种的数据。如果根据这个数据判断甲品种优于乙品种，这个结论就太不可靠了。原因是影响收获率高低的因素不但有种籽品种的差异，还有土地区位、肥沃程度等差异，所以我们需要事先做出安排，使实验结果数据的差异中排除可控因素（土地）的差异，而显示不可控因素（品种）的差异。所谓实验的统计设计就是指设计实验的合理程序，使得收集得到的数据符合统计分析方法的要求，以便得出有效的客观的结论。它主要适用于自然科学研究和工程技术领域的统计数据搜集。

实验的统计设计要遵循以下三个基本原则。

（1）重复性原则。即允许在相同条件下重复多次实验。如果只用一次实验所得的数据作为总体的估计量精度就很差，这时实验的误差等于观察的误差，观察误差可能是实验误差的结果，很难用观察的数据来代表总体情况。多次重复实验的好处是显然的，其一，可以获得更加精确的效应估计量，其二，可以获得实验误差的估计量。这些都是提高估计精度或缩小

误差范围所需要的。

(2) 随机性原则。随机性是指在实验设计中,对实验对象的分配和实验次序都是按随机安排的。这种安排可以使可控的影响因素作用均匀化,突出不可控影响因素的作用。例如在种子品种的实验中如果不是将 A 品种固定在甲地段、B 品种固定在乙地段,而是两地段随机地选择不同品种多次重复实验,可以断定这种安排在不同品种收获率的差异中,由于土地因素的影响大大减少了,而品种因素的影响大大提高了。所以随机化原则是实验设计的重要原则。

(3) 区组化原则。即利用类型分组技术,对实验对象按有关标志顺序排队,然后依次将各单位随机地分配到各处理组,使各处理组组内标志值的差异相对扩大,而处理组组间的差异相对缩小,这种实验设计安排称为随机区组设计。这样就可以提高处理组的估计精度。

1.2.2.2 大量观察法

所谓大量观察法,是指对所研究的事物的全部或足够数量进行观察的方法。社会现象或自然现象都受各种社会规律或自然规律相互交错作用的影响。在现象总体中,个别单位往往受偶然因素的影响,如果任选其中之一进行观察,其结果不足以代表总体的一般特征;只有观察全部或足够的单位并加以综合,影响个别单位的偶然因素才会相互抵消,现象的一般特征才能显示出来。大量观察的意义在于可使个体与总体之间在数量上的偏误相互抵消。

大量观察法的数学依据是大数定律。大数定律是随机现象的基本规律。大数定律的一般概念是:在观察过程中,每次取得的结果不同,这是由偶然性所致的,但大量、重复观察结果的平均值却几乎接近确定的数值。狭义的大数定律就是指概率论中反映上述规律性的一些定理,表述平均数的规律性与随机现象的概率关系。

大数定律的本质意义在于经过大量观察,把个别的、偶然的差异性相互抵消,而必然的、集体的规律性便显示出来。例如,当我们观察个别家庭或少数家庭的婴儿出生时,生男生女的比例极为参差不齐,有的是生男不生女,有的是生女不生男,有的是女多男少,有的是男多女少,然而经过大量观察,男婴、女婴的出生数则趋向均衡。也就是说,观察的次数越多,离差的差距就越小,或者说频率出现了稳定性。这就表明,同质的大量现象是有规律的,尽管个别现象受偶然性因素的影响出现偏差,但观察数量达到一定程度就呈现出规律性,这就是大数定律的作用。

扩展阅读

大数定律的发展历史

1733 年,德莫佛·拉普拉斯在分布的极限定理方面走出了根本性的一步,证明了二项分布的极限分布是正态分布。拉普拉斯改进了他的证明并把二项分布推广为更一般的分布。1900 年,李雅普诺夫进一步推广了他们的结论,并创立了特征函数法。这类分布极限问题是当时概率论研究的中心问题,卜里耶为之命名“中心极限定理”。20 世纪初,主要探讨使中心极限定理成立的最广泛的条件,二三十年代的林德贝尔格条件和费勒条件是独立随机变量序列情形下的显著进展。伯努利是第一个研究这一问题的数学家,他于 1713 年首先提出后人称之为“大数定律”的极限定理。

1.2.2.3 统计描述法

统计描述法是指对由实验或调查而得到的数据进行登记、审核、整理、归类、计算出各种能反映总体数量特征的综合指标,并加以分析从中抽出有用的信息,用表格或图像把它表示出来。统计描述是统计研究的基础,它为统计推断、统计咨询、统计决策提供必要的事实依据。统计描述也是对客观事物认识的不断深化过程。它通过对分散无序的原始资料的整理

归纳,运用分组法、综合指标法和统计模型法得到现象总体的数量特征,揭露客观事物内在数量规律性,达到认识的目的。

分组法是研究总体内部差异的重要方法,通过分组可以研究总体中不同类型的性质以及它们的分布情况,如产业的经济类型及其行业分布情况。可以研究总体中的构成和比例关系,如三次产业的构成,生产要素的比例等。可以研究总体中现象之间的相关依存关系,如企业经营规模和利润率之间的关系等。

综合指标法是指运用各种统计指标来反映和研究客观总体现象的一般数量特征和数量关系的方法。通过综合指标的计算可以显示出现象在具体时间、地点条件下的总量规模、相对水平、集中趋势、变异程度,并进一步从动态上研究现象的发展趋势和变化规律。

统计模型法则是综合指标法的扩展。它是根据一定的理论和假定条件,用数学方程去模拟客观现象相互关系的一种研究方法,利用这种方法,可以对客观现象和过程中存在的数量关系进行比较完整和全面的描述,凸显所研究的综合指标之间的关系,从而简化了客观存在的复杂的其他关系,以便利用模型对所关心的现象变化进行评估和预测。

1.2.2.4 统计推断法

统计在研究现象的总体数量关系时,需要了解的总体对象的范围往往是很大的,有时甚至是无限的,而由于经费、时间和精力等各种原因,以致有时在客观上只能从中观察部分单位或有限单位进行计算和分析,根据局部观察结果来推断总体。例如,要说明一批灯泡的平均使用寿命,只能从该批灯泡中抽取一小部分进行检验,推断这一批灯泡的平均使用寿命,并给出这种推断的置信程度。这种在一定置信程度下,根据样本资料的特征,对总体的特征做出估计和预测的方法称为统计推断法。统计推断是现代统计学的基本方法,在统计研究中得到了极为广泛的应用,它既可以用于对总体参数的估计,也可以用作对总体某些分布特征的假设检验。从这种意义上来说,统计学是在不确定条件下做出决策或推断的一种方法。

动脑筋:大量观察法和统计推断相比较各自有哪些优势?

1.3 统计学的基本概念

统计科学和其他科学一样,在论述本门科学的理论与方法时,要运用一些专门的概念,有些是基本的、常用的,有些是属于局部的,在论述专门问题时使用属于局部的,局部的概念在以后章节讲解。本节只就几个基本的、常用的概念加以阐述。

1.3.1 统计总体与总体单位

统计总体简称总体,是指客观存在的、在同一性质基础上结合起来的许多个别单位的整体。构成总体的这些个别单位称为总体单位。例如,所有的工业企业就是一个总体,这是因为在性质上每个工业企业的经济职能是相同的,即都是从事工业生产活动的基本单位,这就是说,它们是同性质的。这些工业企业的集合就构成了统计总体。对于该总体来说,每一个工业企业就是一个总体单位。

总体可以分为有限总体和无限总体。总体所包含的单位数是有限的,称为有限总体,如人口数、企业数、商店数等。总体所包含的单位数是无限的,称为无限总体,如连续生产的某种产品的生产数量、大海里的鱼资源数等。对有限总体可以进行全面调查,也可以进行非全面调查。但对无限总体只能抽取一部分单位进行非全面调查,据以推断总体。

确定总体与总体单位,必须注意两个方面:

(1) 构成总体的单位必须是同质的,不能把不同质的单位混在总体之中。例如,研究工

人的工资水平,就只能将靠工资收入的职工列入统计总体的范围。同时,也只能对职工的工资收入进行考察,对职工由其他方面取得的收入就要加以排除,这样才能正确反映职工的工资水平。

(2) 总体与总体单位具有相对性,随着研究任务的改变而改变。同一单位可以是总体也可以是总体单位。例如,要了解全国工业企业职工的工资收入情况,那么全部工厂是总体,各个工厂是总体单位。如果旨在了解某个企业职工的工资收入情况,则该企业就成了总体,每位职工的工资就是总体单位了。

动脑筋:要研究一下某学院大一学生的日常消费情况,其统计总体是什么?

1.3.2 样本

从总体中抽取的部分单位组成的集合称为样本。抽取样本的目的,就在于要用样本的数量特征来估计或推断总体的数量特征。对于无限总体我们不可能对每一单位进行观察,即使是有限总体,由于其大量性的特点,要对所有单位进行观察,要花费大量的人力、物力、财力和时间,是十分不经济的事情。因此,一般情况下,我们都是通过样本来推断总体的特征的。

既然抽样的目的是推断总体的特征,因而从总体中抽取样本时必须遵循随机原则,这样才能保证样本的代表性。例如,我们研究某大学学生课余时间的利用情况,随机抽取 100 名学生进行调查,则这 100 名学生就是一个样本。总体是统计研究的对象,样本作为总体的代表,也是统计研究的对象,因此样本也符合总体的概念,为了加以区别,通常将所要研究的事物全体构成的总体称为全及总体,而将样本单位组成的总体称为抽样总体。

抽取样本时要注意以下问题:

(1) 从一个总体中可以抽取许多个不同的样本。根据研究目的确定的统计总体是唯一确定的,而样本却是随机的,从一个总体中抽取不同样本的数目的多少与样本单位数和抽样方式有关。例如,某校有 $N=10000$ 名学生,从该校学生中抽取 $n=100$ 名构成样本,按重置抽样方式其可能的样本数目为: $N^n=(10000)^{100}=10^{400}$,按不重置抽样时的情形请读者推算。由此可见,从一个总体中抽样一定样本容量,不同的样本是非常多的。

(2) 样本的代表性。抽样的目的是用样本的数量特征去推断总体的数量特征,因此,就要求样本的指标与总体的指标的误差要小,即抽样误差小。抽样误差越小则样本的代表性就越强。样本的代表性的高低与样本单位数、抽样方式和抽样的组织形式有关。提高样本的代表性,降低抽样误差,是抽取样本时要高度关注的问题。

(3) 样本的客观性。在抽取样本时,要遵循随机原则,排除主观因素的影响,保持取样的客观性,从而提高样本的代表性。

动脑筋:抽取样本时,样本容量设定为多少比较合适?

1.3.3 统计标志

1.3.3.1 标志和标志表现

统计标志简称标志,是指统计总体各单位所具有的共同特征的名称。从不同角度考察,每个总体单位可以有许多特征。如每个学生可以有性别、年龄、民族等特征。这些都是学生的标志。

标志表现是标志特征在各单位的具体体现。学生的性别是“女”,年龄为 19 岁,民族为汉族等,这里“女”、“19 岁”、“汉族”就是性别、年龄、民族的具体体现,即标志表现。

1.3.3.2 标志的分类

(1) 标志按变异情况可分为不变标志和变异标志。当一个标志在各个单位的具体表现都

相同时,这个标志称为不变标志;当一个标志在各个单位的具体表现有可能不同时,这个标志称为可变标志或变异标志。如中国第五次人口普查规定:“人口普查的对象是具有中华人民共和国国籍并在中华人民共和国国境内常住的人。”按照这一规定,在作为调查对象的人口总体中,国籍和在国境内居住是不变标志,而性别、年龄、民族、职业等则是变异标志。不变标志是构成统计总体的基础,因为至少必须有一个不变标志将各总体单位联结在一起,才能使它具有“同质性”,从而构成一个总体。变异标志是统计研究的主要内容,因为如果标志在各总体单位之间的表现都相同,那就没有进行统计分析研究的必要了。

(2) 标志按其性质可以分为品质标志和数量标志。品质标志表示事物的质的特性,是不能用数值表示的,如学生的性别、民族、专业等。数量标志表示事物的量的特性,是可以数值表示的,如学生的身高、体重、年龄等。品质标志主要用于分组,将性质不相同的总体单位划分开来,便于计算各组的总体单位数,计算结构和比例指标。数量标志既可用于分组,也可用于计算标志总量以及其他各种质量指标。

1.3.4 统计指标

1.3.4.1 统计指标及其构成要素

对统计指标的含义,一般有两种理解和两种使用方法。

(1) 统计指标是指反映总体现象数量特征的概念。如人口数、商品销售额、劳动生产率等。它包括三个构成要素:指标名称,计量单位,计算方法。这是统计理论与统计设计上所使用的统计指标含义。

(2) 统计指标是反映总体现象数量特征的概念和具体数值。例如,2013年我国国内生产总值为568 845亿元。这个概念含义中包括了指标数值。按照这种理解,统计指标除了包括上述三个构成要素外,还包括时间限制、空间限制、指标数值。这是统计实际工作中经常使用的统计指标的含义。因此,统计指标包括六个具体的构成因素。

一般认为,对统计指标的这两种理解都是成立的。在做一般性统计设计时,只能设计统计指标的名称、内容、口径、计量单位和方法,这是不包括数值的统计指标。然后经过搜集资料、汇总整理、加工计算可以得到统计指标的具体数值,用来说明总体现象的实际数量状况及其发展变化的情况。从不包括数值的统计指标到包括数值的统计指标,在一定意义上反映了统计工作的过程。

1.3.4.2 统计指标的特点

(1) 数量性。即所有的统计指标都是可以用数值来表现的。这是统计指标最基本的特点。统计指标所反映的就是客观现象的数量特征,这种数量特征,是统计指标存在的形式,没有数量特征的统计指标是不存在的。正因为统计指标具有数量性的特点,它才能对客观总体进行量的描述,才使统计研究运用数学方法和现代计算技术成为可能。

(2) 综合性。这是指统计指标既是同质总体大量个别单位的总计,又是大量个别单位标志差异的综合,是许多个体现象数量综合的结果。例如,某人的年龄,某人的存款额不能叫做统计指标,一些人的平均年龄,一些人的储蓄总额,人均储蓄才叫做统计指标。统计指标的形成都必须经过从个体到总体的过程,它是通过个别单位数量差异的抽象化来体现总体综合数量的特点的。

(3) 具体性。统计指标的具体性有两方面的含义:一是统计指标不是抽象的概念和数字,而是一定的具体的社会经济现象的量的反映,是在质的基础上的量的集合。这一点使社会经济统计和数理统计、数学相区别。二是统计指标说明的是客观存在的,已经发生的事实,它反映了社会经济现象在具体地点、时间和条件下的数量变化。这一点又和计划指标相区别。统计指标反映的是过去的事实和根据这些事实综合计算出来的实际数量,而计划指标