

工商业和经济学中 应用的统计方法

〔美〕罗伯特·D·梅森 著

中国人民大学出版社

**工商业和经济学中
应用的统计方法**

〔美〕罗伯特·D·梅森 著
中国人民大学统计研究室 译

•
中国人民大学出版社出版
(北京西郊海淀路39号)
中国人民大学出版社印刷厂印刷
(北京鼓楼西大街石桥胡同61号)
新华书店北京发行所发行

•
开本: 850×1168毫米32开 印张: 17
1984年5月第1版 1984年5月第1次印刷
字数: 416,000 册数: 14,800
统一书号: 3011·176 定价: 2.10元

STATISTICAL TECHNIQUES
IN BUSINESS AND ECONOMICS
ROBERT D. MASON

Forth Edition 1978
RICHARD D. IRWIN, INC.
Homewood, Illinois

译者说明

数理统计方法既是许多社会科学和自然科学中进行科学研究的必要工具，又是在实际工作中用来决策的重要手段。目前为适应我国社会主义现代化建设的需要，大专院校各财经专业开设了数理统计应用的课程，随之提出了对这方面的教材和习题的需要。为此，我们翻译了本书。

本书是美国出版的《工商业数量分析丛书》之一，由托莱多大学梅森教授编写，曾在1967年、1970年、1974年和1978年先后四次出版。这个译本是根据1978年第四版翻译的，它反映了美国在社会经济方面应用的数理统计方法，值得我们借鉴。

本书有以下几个特点：（1）着重介绍在工商业和经济学中应用数理统计方法，在介绍各种方法时，一般都是根据具体例子提出问题，使读者容易理解和掌握。关于数学原理则尽量简略，具有中学的数学水平，基本上就可以看懂。（2）在讲述过程中，每个段落都有复习题，并附有答案，有利于读者及时巩固学到的知识。（3）包括的内容比较广泛，注重实际，所附的习题也比较多。因此，本书既可以作为大专院校开设数理统计方法应用课程的教材，又可以作为广大统计工作者学习数理统计方法的入门书。

本书刊行于美国，其内容不可避免地反映了资本主义国家的一些情况和特点；因此，请读者在阅读和应用时加以注意，要洋为中用，吸收其中有用的方法。

考虑到原书篇幅较多，在翻译时我们作了适当的调整，主要

是删去了每章后的简要概述，并把正文和习题（各章自我复习检查）分为两册出版。

由于我们的水平有限，译文中难免有一些错误之处，因此恳切希望读者提出批评、指正。

参加本书翻译和校对工作的有：戴世光、林茂美、周复恭、倪加勋、黄运成同志。高庆丰同志也参加了部分工作。

一九八二年十一月

前 言

正如书名所示，本书主要是作为工商业和经济统计学课程的一个初级课本。本书包括有足够讲授一学年课程所需要的教材，并且，全书课题范围较为广泛，也可以缩短教学时间，选用部分课题作为教材。本书涉及的数学比较浅，数学较差的学生也能够掌握。

第四版增加的新内容有：双向方差分析和属于非参数假设检验的“曼-惠特尼检验”。对于复回归和相关一章和决策一章，已作了修改和补充。在每一章之后，都另增加了简要的概述。学生能够用来概括一章的全部要点，而教师则可以用来证实已讲授过的该章中的每一课题。

根据学生们的反映，我将前一版书中每章所列的习题（备有答案）和自我复习检查（备有答案），在本版中都已作了补充。学生通过解答这些习题，就能够立即判断出来自己对本章概念的理解如何。在一定意义上，它也是帮助学习统计科学基础知识的一个有力的工具。在每一章后备有未解答的习题。本版在数量上已作了增加。

书中包括若干为社会科学统计项目用的电子计算机软件包，以说明计算机解题的能力。

另外还备有习题手册和题解手册，作为这本书的补充教材。

我感谢同事威·吉·马奇尔、道·爱·林德和帕·弗·斯科特在组织本书的第四版中给予的许多有益的帮助。特别要感谢在纳什维尔的田纳西大学的布·阿·博尔斯，在圣泽西的加里佛尼

亚州立大学的巴·克拉佛特，在印第安纳波利斯的印第安纳波迪大学的顾绍恩，在麦拉美克的圣路易斯公立学院的约·朗机，迈阿密的达德公立南方学院的哈·鲁·罗耶，和苏孚尔克大学约·西·香农。他们在审阅原稿时，提出了许多有价值的见解和批评。我还要特别感谢西伊利诺斯大学的吉·哈·沃克，他提出了许多有帮助的建议。最后，我十分感激已故的英国皇家学会会员罗·爱·弗希尔爵士的遗嘱执行者——儒萨斯替德皇家学会会员弗·耶茨博士和奥利弗·博伊德公司的先生们，他们允许我翻印《生物、农业及医学研究用的统计表》一书中统计表Ⅲ和统计表Ⅳ。

罗伯特·D·梅森

1978年1月

目 录

第一章 绪 论	1
定 义	1
描述的统计学	2
归纳的统计学	3
第二章 资料的搜集和描述	6
来源于出版物的原始资料	6
调查, 作为资料的一个来源	11
资料的描述	13
第三章 次数分配	28
次数分配的编制	29
次数分配的图示	38
第四章 集中趋势的计量	47
算术平均数	48
中位数	53
众 数	60
几何平均数	63
算术平均数、中位数和众数的关系	66
章内问题答案	69
第五章 离差和偏态的计量	71
绝对离差的计量	71
标准差的解释和使用	95
相对离差的计量	98
偏态的计量	100
关于非熟练工人统计分析的概述	101

电子计算机的应用	102
章内问题答案	103
第六章 指 数	105
引 言	105
指数的编制	108
消费价格指数	119
消费价格指数的特殊用途	120
改换基期和拼接	124
章内问题答案	129
第七章 时间数列分析	132
时间数列的组成部分	134
直线趋势	137
非线性趋势	146
章内问题答案	153
第八章 季节性分析	156
引 言	156
测定季节指数的方法	161
章内问题答案	174
第九章 概率概念的概述	176
引 言	176
概率定理	178
排列和组合	185
章内问题答案	190
第十章 概率分布	193
二项式概率分布	194
普哇松概率分布	201
正态概率分布	208
章内问题答案	214
第十一章 统计推论导论	217
抽样的原因	218

选择样本的方法	219
点估计与区间估计	227
关于总体平均数的推论	228
样本的容量	236
关于总体比例的推论	240
比例的样本容量	242
章内问题答案	243
第十二章 参数假设检验	245
假设检验的概念	245
假设检验的程序	246
计量水准	253
假设检验：一个样本，总体平均数	
μ 和标准差 σ 为已知	256
假设检验：一个样本，总体平均数 μ	
为已知，但标准差 σ 为	
未知	259
假设检验：两个平均数，大样本	262
假设检验：两个平均数，小样本	267
第二类错误，运算特性曲线和功效曲线	270
章内问题答案	278
第十三章 方差分析：单向和双向分类	280
引言	280
方差分析：一般概念	281
单向方差分析：相同的样本容量	283
计算F比率的另一种方法	290
单向方差分析：样本容量不等	294
双向方差分析	299
一个生产问题	303
章内问题答案	305
第十四章 非参数假设检验：单一样本和两个样本的检验	306

克方分布：单一样本的显著性检验·····	306
已知总体比例的单一样本检验·····	315
包含两个比例的两个样本的检验·····	318
曼一惠特尼U检验·····	324
柯尔莫哥洛夫—斯米尔诺夫的两个样本的检验·····	331
威尔科克森的配对带正负号的等级检验·····	337
章内问题答案·····	343
第十五章 非参数假设检验：两个以上样本·····	349
科可安Q检验·····	349
弗利德曼双向方差分析·····	354
K个样本的克方检验·····	357
章内问题答案·····	361
第十六章 统计质量控制·····	364
控制图的观点·····	365
章内问题答案·····	386
第十七章 简单回归和相关分析·····	389
回归分析·····	389
相关分析·····	405
应用电子计算机的例子·····	412
章内问题答案·····	414
第十八章 复回归和相关分析·····	416
复回归分析·····	416
复相关分析·····	426
电子计算机的应用·····	428
一个薪金预计问题·····	430
章内问题答案·····	442
第十九章 联系的非参数计量·····	445
斯皮尔曼等级次序相关系数·····	446
肯达尔的一致性系数：W·····	452
章内问题答案·····	458

第二十章 偶然不确定性情况下的决策导论	460
引 言	460
在偶然不确定情况下作出决策的一件事例	463
最大的最小、最大的最大及最小的最大遗憾的策略	469
完备的情报资料的价值	469
灵敏度分析	471
决策树枝图	473
章内问题答案	476
第二十一章 贝叶斯定理和决策	478
贝叶斯规则	480
关于生产咯咯笑玩偶最佳水平的一个实例	483
章内问题答案	490

附 录

附录 1：常用对数
附录 2：对数表
附录 3：平方与平方根表
附录 4：随机数目表
附录 5：t 的临界值
附录 6：威尔科克森 T 值
附录 7：正态曲线下的面积和纵座标表
附录 8：克方的临界值
附录 9：二项分布
附录 10：关于选定 μ 值的普哇松分布
附录 11：控制图的系数
附录 12：曼·惠特尼检验：U 的临界值
附录 13：斯皮尔曼等级相关系数： ρ 的临界值
附录 14：肯达尔一致性系数，W；S 的临界值
附录 15：在 5 % 和 1 % 的显著性水准下，F 分布的临界值
附录 16：柯尔莫哥洛夫、斯米尔诺夫两个样本检验中 D 的临界值（大样本，双尾检验）

第一章

绪 论

定 义

爱看洛杉矶“公羊撞角”^{*}、辛辛那提“红布斗牛”^{*}、波士顿“凯尔特人”^{*}及其它专业和学院球队比赛的观众们,通常把统计人员理解为这样一种人:他坐在新闻记者席上并且记录着得分场次、棒球“本垒打”比赛的次数以及犯规、投篮未中的次数,然后,公布这些统计数字,譬如,吉博士在对运气不好的尼克斯的比赛中获得42点,尼克拉斯投篮67次等。

同样,参看《华尔街日报》、各种地方新闻报纸、《福布斯》、其他商业杂志以及各种工业年报,它们都揭示了这些各种各样的统计数字,例如罗伯特·吉·苏斯洛(萨克斯第五大街)的年收入为35万元,麦凯尔·西·伯杰瑞克(瑞尔温)的年收入为150万元^①;再如,阿拉斯加的教师们的年薪最初为1.8万元,年平均为2.1万元,以后增加到3万元^②。另外,一个地方报纸^③报导,到1986年买一所中等价值的新住宅将要花9万元(与目前的4.5万元相比较);消费者全部的应付未付的分期贷款的总额

• “公羊撞角”、“红布斗牛”、“凯尔特人”等都是美国球队的名字——译者注。

① 《福布斯》,1977年4月1日,第50页。

② 《华尔街日报》,1977年4月5日,第1页。

③ 《布累德》,多伦多,俄亥俄州,1977年4月1日。

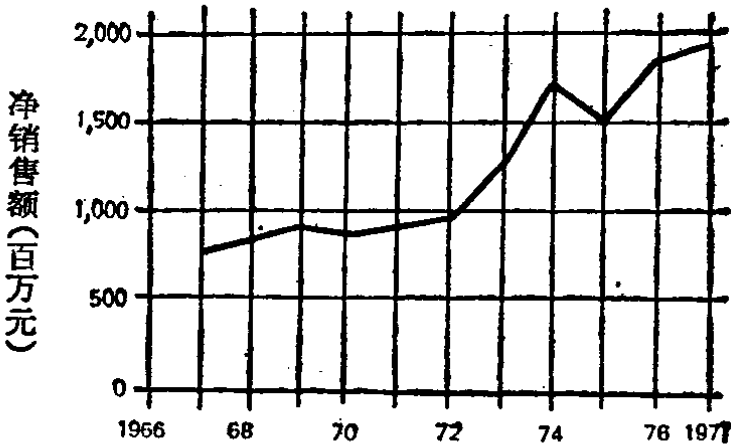
是1,783亿元；去年苏联生产了40万台彩色电视机和122.4万吨通心粉；在多伦多，去年每10万个人中就有 2.3 人犯谋杀罪，与其相应的是，在洛杉矶有14.3人，在美国的乡村地区有 8.0 人，而全美国平均为9.6人。国际卫生教育委员会在一份统计报告中说，“典型的美国人中有0.9人患感冒或流行性感冒，……在全部的被限制活动的10亿多天中，感冒或流行性感冒就剥夺了病人们大约 5 天的正常活动时间。”

然而，统计一词还有另外的含义，即统计学，它被定义为：在客观偶然不确定的情况下，为了做出较好的决策而对数字资料进行搜集、组织、描述、分析和说明的科学。本书将要探讨的正是这种含义的统计学。

描 述 的 统 计 学

从上述定义看来，统计学的一个方面是研究搜集、组织、描述的方法，也即用以描述数字资料的方法。例如，凯撒制铝化学公司要向股东表明本公司1967年以来销售额迅速增长的情况，就用一种叫作线图（line chart）的描述工具来描述（图 1—1）。

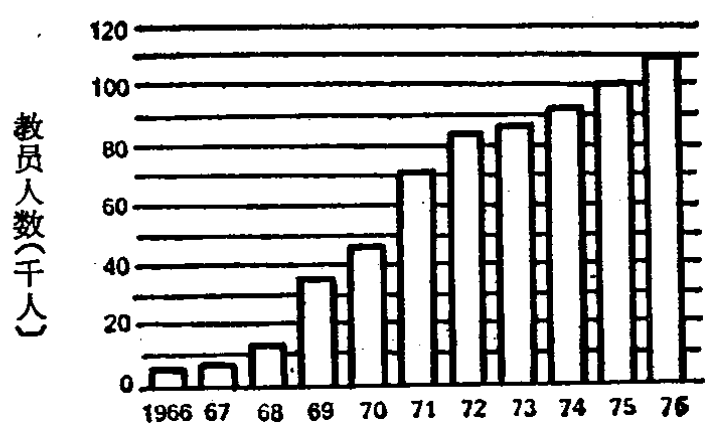
图 1—1 凯撒制铝化学公司的净销售额
(1967—1976年)



资料来源：凯撒制铝化学公司的1976年年报，第30页。

劳动统计局还用条形图的方法来描述学院和大学中参加工会的教员人数的显著增长（图 1—2）。

图 1—2 高等院校中参加工会的教员人数
(1966—1976年)



资料来源：《劳动评论月报》1976年10月号，第33页。

本书中有几章是论述描述统计学的，第二章考察统计资料的各种来源和图表的描述方法，第三章探讨资料的组织整理。还通过第四章计算一种或一种以上的集中趋势和第五章计算离散度来对统计资料作出分析。另外在第六章中把统计资料变换为指数并在其它各章都适当地运用了描述的统计方法。

归纳的统计学

归纳的统计学也叫做统计推论；它是统计学的另一方面。它超出单纯描述数量的情况。统计推论方法是通过样本来推论总体的方法。总体也叫做全域，例如：全部有投票资格的人，大学中新生总数，池塘中所有的鱼，卡车中全部麦子，以及一台机器一小时内制造的小齿轮数。换句话说，凡可以搜集到的一群人、一塘鱼、一堆加工零件等等统计数字的总数，都可以叫做总体。应该指出，“总体”这个字（Population 英文中也可以解释为“人

口”——译者) 在统计学中的含义, 并不是专指人口而言的。

从总体中选出来的若干个体单位, 我们把它叫做样本。专业民意测验的机构, 在选举之前, 从美国全国千百万选民中选出不到 3,000 人的选民作样本, 根据这种样本的回答, 便可以推论出全体选民将如何进行投票。一个汽车制造商可以从每小时生产的小齿轮中取出 5 个小齿轮作为样本, 并根据样本的质量情况, 确定所生产的全部小齿轮的质量。从停在谷仓前准备卸车的卡车中, 取出一小勺麦子, 根据对这勺麦子质量检验的结果, 便可以对整个卡车所载的麦子按质论价。上面列举的每个例子, 都不适于对整个总体进行调查。选民的样本是必要的, 因为在选举前夕, 要把全体选民聚集在一起, 其费用会大得惊人。又如, 检验全卡车麦子的水分质量, 会使全部麦子受到损害。因此, 对样本进行考察是绝对必要的。

再举一些广泛采用抽样法的例子。美国劳工部经常地调查就业、失业、工资等等情况。根据这些抽样调查, 他们发表了一系列统计资料, 如: 1976 年有 13.1% 的黑人和 7 % 的白人失业; 11 月份, 每 100 个工人中有 1.6 人离职; 矿工平均每周收入为 291.04 元, 而零售贸易工作人员则为 118.30 元^①。詹·爱·金和格·吉·史密斯采用“随机抽样”, 从精神病院中选出 2,737 个专业人员来进行研究^②。伦敦的人口普查调查局在国际机场及码头进行定期调查, 对离开本地的旅客询问下列有关事项, 如: 他们停留的期限; 花费了若干英镑; 游览了若干个旅游胜地; 以及其它等等。

统计推断、归纳、预测和预计, 全都超出了描述统计学的范围。对于未知的事情进行预计, 存在一定的不可靠性。例如, 从一小时内制造的全部小齿轮中, 随机抽选出来的 5 个小齿轮可能都是完好的。根据这个样本来推论总体, 就会推论出全部生产的

① 《劳动评论月报》1977 年 2 月号, 第 8、9、113 页。

② 《社会科学》1976 年夏季号, 第 161—169 页。

小齿轮都是完好合格的。但是，由于这种推论是以样本为根据的，因而还存在着并非全部生产的齿轮个个都完好合格的那种机会。再例如，以样本为根据的民意测验，对新的谷物、肥皂、牙膏进行的市场测验，通过样本来推断总体，也会存在一定的不可靠性；民意测验或市场调查可能指出，候选人甲将以绝对的优势获胜，或者在市场销售中，一种新的谷物将被绝大多数人所购买。然而，也有可能候选人乙获得胜利，或者这种新谷物被消费者所拒绝，从而使生产者蒙受了巨大的损失。

专栏作者菲·瓦格纳在论联邦政府按月公布的失业统计数字^①的这一重要课题中，曾经强调指出抽样调查所包含的不准确性。

“某些统计资料是可靠的。体育运动统计数字，如击球的平均数等类似的数字，通常是真实的。这里，我们要相信统计资料，但确实许多关系到更重要的事情的统计资料，并不是这样。例如，失业统计数字，首先，什么是失业，什么不是失业，是按武断的定义来确定的，多数定义很值得商榷。其次，有些样本是假设抽选的，并不知道它能否代表总体。另外，这种样本又是由具有不同工作能力的抽样工作者，从不同的可靠来源中抽选的。最后，把公式应用于这一群准确程度不同的原始资料。从而把这种混合物输给计算机，算出一系列的数字、比率、比例。如此，就会从最令人吃惊的不准确的材料中得出虚构的准确数字。

如果让我在这种每天输入和输出的统计数字与慎重的而又说清楚实情的估计数字二者之中进行选择，那么我宁愿每次都选择估计数字，因为至少你知道它仅仅是个估计，而一个统计数字则十分容易使人信以为是真的。”

本书首先要考察描述统计学。下一章将讨论某些重要工商业统计资料的来源和用图表的形式对资料进行的描述。

^① 《利剑》，托莱多，俄亥俄州，1977年1月27日，第31页。