

工商业和经济学用 统计方法技术

第 一 册

校 内 用 书
注 意 保 存

中国人民大学统计研究室译

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。该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：双向的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；在圣泽西的加利福尼亚州立大学的巴·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注解和批评。我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，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N。

罗·第·梅森

1978年1月

目 录

第一章	导 论	1
第二章	资料的搜集和表列	6
第三章	次数分配	34
第四章	集中趋势的计量	58
第五章	离差和偏态的计量	93
第六章	指 数	135
第七章	时间数列分析	171
第八章	季节性分析	205
第九章	概率概念的概述	231

第一章 导 论

定 义

爱看洛杉矶“公羊撞角”橄榄球比赛，辛辛那提“红布斗牛”棒球比赛，和波士顿“凯尔特人”篮球比赛，及其他专业或学院球队比赛的观众们，通常总认为统计人员就是坐在新闻记者席上记录得分，棒球本垒打的次数，和投篮未遂的次数，然后公布这些统计数字。譬如，吉博士在与运气不好的尼克斯比赛中获得42点，和尼古拉斯投篮得到67分。

同样地，参看《华尔街日报》、地方新闻报纸，《弗比斯》，及其它商业杂志，和工业部门的年度报告，也可以看到各式各样的统计数字，如罗·杰·苏斯洛（萨克斯第五大街）年收入为350,000元，米·西·伯杰瑞克（瑞尔温）年收入为150万元^①。再如，阿拉斯加的教师年薪最低者从18,000元开始，平均薪金21,000元，最高达到30,000元^②。并且，从地方报纸^③看到的，1986年买一所中等价格的新房子将要用90,000元，（与目前相比约差45,000元）；消费者应付未付的债务总额为1783亿元；去年苏联生产了40万台彩色电视机和122.4万吨通心粉；去年在多伦多，每10万人中有2.3件谋杀案，与之对比，计：洛杉矶为14.3，美国乡村地区为8.0，而美国全国则为9.6。国家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患感冒或流行性感冒……这种病能使病人大约5天不能参加正常活动，而全国人不能参加活动的总天数竟达

① 《弗比斯》，1977年，4月1日第30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。

十亿天。此外再如《哥若吉》上提供的一项统计数字是：本星期的牛肉售价，每磅为69分钱。

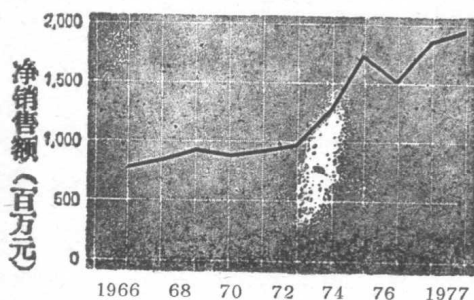
然而，上列例子指的是统计数字，另外还存在着一门科学，即统计学。统计学的定义是：为了对客观偶然性不确定情况作出更好决策的目的，而从事搜集，组织，表列，分析和解释变异数量资料的科学。本书所要探讨的正是这种含义的统计学。

描述的统计学

从上述的字义看来，统计学的一个方面为研究搜集、组织、表列的方法，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）的描述工具表列销售的情况。

（图1—1）

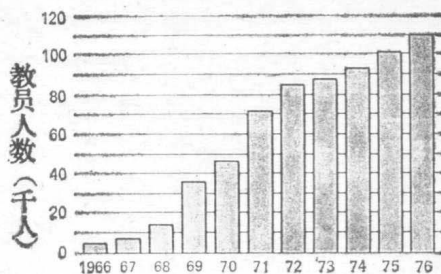
图1—1 凯撒制铝化学公司的净销售额
（1967—1976年）



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来表列学院和大学中参加工会的教员人数显著的增长。（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示的表列，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。在第六章进一步把适当的资料变算为指数；并在其它各章阐明对适当的资料所采用的方法技术。

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。统计推论方法为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，池塘中全部鱼的条数，卡车中全部麦子，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 英文中指的是‘人口’）在统计学中的含义，并不是专指人而言的。

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。一个汽车制造者可以从每小时中取出5个小齿轮作为样本，并根据样本的质量数字，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，取出一小构麦子，根据对这构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。正是这些情况，才有必要只对样本进行考察。

再举一些广泛采用抽样法的例子。美国劳动部经常地调查就业、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者则为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。

统计推论，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，涉及一定的不可靠性。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，通过样本来推论总体，也会存在一定的不可靠性，通过样本的

① 《劳动评论月报》 1977年2月份，第8、9、113页。

② 《社会科学》 1976年，夏季号第161—169页。

民意测验可能指明某甲取得了压倒的胜利；通过市场的样本测验，推论出市场上新的谷物有极大部分会销售出去。然而，也有可能，某乙会取得了胜利；在市场上消费者会拒绝购买这种新的谷物，从而导致生产者将受巨大的损失。

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，如球击的平均数，和类似的数字等等，通常是真实的。从这里，我们学习到：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，而一个统计量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表列，统计资料的图示法。

① 《利剑》托莱多，俄亥俄州，1977年1月27日，第31页。

第二章 资料的搜集和表列

正如在导论一章中简略提到的，解决工商业和经济中许多问题的第一步，就是要搜集必要的资料。资料的分析 and 解释（加上判断）乃是经营管理决策的基础。与问题有关的资料，可以是已经公布的形式，也可以由于问题的性质还需要举行调查来搜集资料。

出版资料的来源

标准的工商业和经济资料的来源

按照问题而言，著名的资料来源，例如，《美国历史统计资料》《美国统计摘要》，或下列各节所开列的许多出版物之一。这许多列举的出版物，仅仅是大量出版物中有代表性者，并非全部。这些出版物可以大致地分为：政府的，私人的，及国际的出版物。并引用实际材料来说明其中统计资料所属的类型。

联邦政府的出版物

十年一度的出版物。毫无疑问，美国政府比任何其它机构或任何外国政府提供更多的统计资料。许多部和局，包括普查局、劳工部和联邦贸易委员会，每十年、每年、每月、每周、每天都公布和分发统计资料。普查局每十年举办一次人口普查和住房普查，从而得到大量的关于人口特征的详细报告，比如，年龄、收入、性别。普查局还提供人口按州按年龄、性别等等的年度估计和预计数字。例如：

年度和数列

美国总人口 (千人) *

1950.....152,271

1960.....180,667

1970.....204,800

预计:

1985—B.....257,903

C.....252,093

D.....246,265

E.....240,153

• 7月1日的人口数, 包括美国在国外的军队人数。

资料来源: 美国普查局, 现行人口报告丛辑, 25页, 第310号、470号、476号。

年度出版物。从1878年以来, 美国商业部每年出版《美国统计摘要》, 大多数统计资料是最近几年的, 但也有许多统计数列包括过去四、五年的数字。这本约 1,000 页出版物的统计表目录包括人口, 生命统计, 迁移, 教育, 法律实施, 劳动, 价格, 选举, 能源, 运输, 制造业, 渔业, 及农业等章节。

少数种种个别的数字。

夏威夷、火奴鲁鲁的最低温度是53度。

底特律市去年每一百人中罪犯2.1人。

14岁及以上的文盲占人口1.0%。

美国有107,371,000辆汽车 (包括出租汽车在内)。

按月的出版物。联邦政府大多数的部还按月出版统计资料。劳动部出版《劳动评论月报》, 大约用40页的篇幅刊登有关劳工周转, 收入, 工时, 消费者和批发价格, 罢工, 及工伤等资料。最近一个月的材料中就有:

加利福尼亚州就业人数为8,045,500人 (不包括农业工人在内)。

美国矿业就业人数为815,000人。

最近一个月, 每一百个雇佣员工中有1.5人被解雇。

在石油和煤炭的生产活动中，平均小时毛收入为7.23元。

商业部按月还出版《现代商业概览》其中包括二个部分主要内容：一为现行工商业统计资料；另一为工商业重要情况的评述：价格，商业倒闭，及工业前景。这其余一半的工商业的统计资料，大约包括2,500个各月的统计数字。在大多数情况下，主要是关于本月份和前十一个月的统计数字。

美国未偿付的分期信贷总额为1783亿元。这仅仅是联邦储备体系管理局每月出版在《联邦储备公报》上公布的上千金融数字中的一个。前一百页提供的为银行家和金融界其他人士所关心的课题。后50页，包括商业信贷，利率、资产和负债，及活期存款的统计数字。其中，还包括许多有关的统计数列，如零售销售额，生产指数，房屋建筑数，及出口额等。

私人组织的出版物

年度出版物。许多产业协会编制并出版的各种类型的统计资料。例如，机动车制造商产业协会，按年出版《机动车业的事实和数字》。它提供的数字，计有1900年的汽车登记数（4,192辆），年生产公共汽车数，按州登记的卡车和汽车数，通用汽车公司、福特汽车公司及其它制造商生产的机动车数，由交通事故造成的旅客死亡人数，等等。

《世界年鉴和事实》，《情况材料、地图、图表、和年鉴》这两本书，包括大量关于销售、体育运动等等各式各样的统计资料。在《情况材料年鉴》一书中（由西蒙和舒斯特出版），就列出如下的一些统计数字：

美国石油总产量的34.9%产自得克萨斯州。

圣母玛利亚大学的学费计为2,980元。

国际组织的出版物

年度的出版物。联合国统计局，经济和社会事务部是发送国际资料的领导机构，他们每年出版两本书。

《统计年鉴》，其中包括160个以上的国家或地区，通过调查

表直接提供各该国家的统计资料。而其它单位则利用已公布的文件来取得统计资料。书中的统计表列举了大多数国家有关人口、农业、制造业、交通运输、金融和教育各方面的统计资料。

《人口统计年鉴》，其中包括来自全世界大约 250 个单位的官方人口统计资料，提供有关民族、死亡率、结婚、离婚，和人口的年龄性别等等的资料。

月份出版物。《统计月报》包括大多数国家各种的统计资料，其中，有年度的数字，也有月份的数字，比如，机动车的生产、人口、住宅建筑、出口等等，兹列举几项数字如下：

3,967,000,000	世界人口总数
39.7	萨尔瓦多每1,000居民月出生率
93,200	苏联平均每月机动车产量（与之相比，美国为559,000）。

参 考 辅 助 材 料

如果查阅包括经济和工商业资料的标准参考书并不解决问题，那么，可以查阅下列一种或几种供参考的辅助材料：埃·特·科曼所编的《工商业情况的资料来源》，包括有将近 300 页的参考资料，分成下列几类：经营管理、对外贸易、市场、不动产，等等。核·韦·约翰逊所编的《如何利用工商业图书馆》，和玛·曼利所编的《工商业的查寻和使用》，为两本类似的出版物。保·沃塞曼所编的《工商业资料来源百科全书》第一卷，是关于工业资料来源（如冰淇淋工业，砒素工业等）的有价值的参考手册。此外，关于其它课题的参考资料（如维他命、工业卫生等）。

《工商业期刊索引》是许多索引中比较典型的一种。每月搜集了 100 种以上工商业期刊，其中包括《商业周刊》、《兴隆》，《销售管理》和《纺织界》。正如其序言所指出的，它是按月把有关会计、广告和公众关系、自动化、银行、交通、经济、金融和投资，保险、劳动、管理、市场、课税、及特定的商业、工业和贸易等领

域英文期刊题目索引汇编起来的。现在从一期月刊中摘引几个题目索引如下：

汽 车 工 业

《汽车安全仍然证明是难以捉摸的》题目符码（译者注下面为期刊名称、卷数、页数、时期的符码数字）：191:17—18○18'76

《福特罢工在拖延》题目符码：218:9;○11'76

《每周工作四天已在望》题目符码：218:1○;018'76

《周工作日缩短的前缘》题目符码：26.7○4'76

《汽车工作力量在缓慢下降》题目符码：114—18○25'76

《华厂迈出缩短周工作日的一小步》题目符码：44 + ○18'76

证 券

《黄金黯然失色的形象》题目符码：145;39S15'76

《黄金分摊的转变》题目符码：118;122O1'76

除去《工商业期刊索引》外，《纽约时报》和《华尔街日报》也出版了有关经济和工商业的题目的资料索引。

联邦政府从1895年起按月出版了政府出版物手册，手册名为《美国政府书刊目录月报》。

许多统计出版物还提供其它资料。《美国统计摘要》附录中的《资源指南》就很有用。

调查，作为资料的一个来源

一般说来，调查的目的在于搜集那些很难从已经公布的或任何其它形式中得到的统计资料，我们可以利用许多不同的方法，搜集所需要的资料；其中一种就是“通讯调查”。最近，林奇公司发给公司股东每人一张调查表，其内容如下：

亲爱的股东：

我们希望尽一切力量更好地了解我们的股东，并以尽可能最好的方式与他们互通消息。我们相信您对附上的调查表的回答，会在上述两方面对我们有所帮助。我们感谢您花费几分钟时间填写附去的调查表，并用寄上的信封，把调查表寄还给我们。我们将会很高兴地把调查的结果向您报告。

如果您不愿意透露您的姓名，可以不必填写姓名和地址。感谢您的合作。

**林奇公司
投资联络部**

通讯调查法是一种比较省钱的搜集资料的办法，尤其是回答者分布的地区比较辽阔，更是如此。但通讯调查法也有缺点，主要是收回调查表的可能性比较小。某些举办调查的企业报告说他们收到的回答低到10—15%。然而，回答者的多寡，取决于调查表的编排，以及被调查者对调查主题关心程度。例如，有一个15页的调查表，送给100个以上主要的石油公司，其结果收到了92%的回答。

一些大学和专业调查机构，例如，罗·波尔斯，通过亲自访问的办法，来测定建立市场中心的需要、政治候选人的声望以及其它等等。调查者与被调查者亲自接触，比邮寄调查表要多花钱，因为需要保持一批能胜任的调查人员，而且每一项调查都要给他们编调查须知。访问调查方法的主要优点为：可能得到回答的百分比很高。

许多调查研究机构采用通讯调查法与访问调查法相结合的办法。对那些未回答调查表的人，就要进行访问调查。

编制调查表的标准原则

主要根据公司、大学及个人长期积累的经验，业已总结发展出能使调查结果良好的若干标准原则。兹将几项主要准则分述如下

(列举的实例都是取自实际的调查)：

1. 进行调查的人必须作自我介绍，并说明调查的目的。(参考林奇公司的信)

2. 必须有充分的调查说明。兹以林奇公司的调查表为例：

你之所以购买这项股票是按照谁的推荐？(对下表中的一个或几个空白划√来表示回答)

☐ 经纪人	☐ 金融出版物
☐ 银行家	☐ 朋友和亲戚
☐ 投资顾问	☐ 自己的判断
☐ 金融情报服务机构	☐ 其它_____

3. 对于可能产生分歧意义的字句，必须避免使用，或者予以明确解释。

例如：

一份全国杂志向公司经理提出下列问题：“如果你们装备了现代化的办公设备和系统，那么你们公司的经营管理能够增进多大的生产率？_____元。”

对于上例所引的问题，你相信公司经理能用数字估计吗？究竟什么是经营管理的生产率？什么是现代化的办公设备和系统？

4. 必须避免会导致偏见倾向性的问题。过分的形容词，如“美丽”，往往使回答者发生倾向性；需要推论的问题会使回答者产生偏见。

例如：

你既然知道这种镇痛药的优点，那么你不感到应该随身携带它以便下次需要镇痛时用吗？

——1. 是的，疼痛令人极不愉快，切勿冒冒疼难忍的风险。

——2. 是的，我对这种镇痛药很感兴趣，它进入血液的速度要比一般阿斯匹林快两倍。

——3. 是的，这商标使我相信它的质量很高。

——4. 是的，由于我经常服用镇痛药，我认为我必须关心一般阿斯匹林会引起胃的不舒适。

——5. 不。

资料的表列图示

统计资料，如年销售额，或产量资料，能够以图或表的形式来表列它。一般说来，一张画得很好的图，（也叫做统计图），能引起许多读者的注意。《商业周刊》、《弗比斯》，《兴隆》和其它商业杂志赞同中国的一句谚语“一张图胜似千言万语”，它们广泛地采用了统计图。统计图的定义，大致地说，就是：“用图的形式来表列资料”。不过，从大多数统计图上，要得到确切的统计数字是困难的。例如，如果要显示出纽约证券交易所某一股票的确切的收盘价格，或过去四年中每一年公司出现赤字的准确元数，那么选用统计表来表列比用图更为合理。这里，我们将只考察制图的方法。

算 术 尺 度

大多数经济和工商业资料，都可以画在有算术尺度或比例尺度的纸上。算术尺度是比较常用的。我们首先对它加以考察。

算术图纸上尺度的划分是等距离的。从下面举的例子来看，垂直坐标（Y）轴上尺度划分（每单位）的距离是10百万元。横坐标（X）轴上行之间的距离是一年。算术图纸上的相同距离代表相同的数量。

在附加说明中，注意垂直尺度上的数目（销售额）是从原点零起向上增加的，“年”是在X轴上从左向右增加的。（图见下页）

回顾一下，代数中讲到图的坐标方格上有四个象限。注意下面的实例：X与Y值都是正的，在第一象限内。因为大多数工商业统计数字是正的，所以最经常使用的是第一象限。