

大 众 统 计 学

杜 之 道

经 济 科 学 出 版 社

一九八八年·北京

前 言

统计是认识社会的重要手段，是社会主义建设的一项重要工作，是社会经济信息的主体，也是国家管理和企业决策的工具。

为了普及统计知识，便于国民经济各部门和企业事业的经济工作人员、计划统计人员、管理人员有效地掌握统计这一重要武器，提高管理和决策水平，本书以比较通俗的形式，紧密联系实际，较系统地阐明社会经济统计学和数理统计学的基本原理和统计分析技巧，力求做到大众化、趣味性、实用性。并介绍了国内外统计理论的新观点和新方法以及建国以来统计工作的重要经验。

本书可作为中等专业统计学校或业余大学文科的统计教材。也适合作各部门的正规统计培训班的教材，同时，也是成人自学统计学的良好读物。

由于作者水平有限，本书错误在所难免，敬希读者指正。

目 录

第一章 统计——对客观现象进行数量概括的

认识活动	1
一、统计工作(对客观现象进行数量认识活动的过程)	3
二、统计学(对客观现象进行数量认识活动的理论)	4
三、统计资料(对客观现象进行数量认识活动的结果)	9
四、统计与管理	14
五、数字胜于雄辩	17
六、定性认识与定量认识	21
七、大量观察、小量观察与大数法则	24
八、打打算算是统计的特点吗	26
九、统计的几个基本概念	30

第二章 怎样搜集统计数据

一、要有周密的统计设计	34
二、确定反映客观事物数量的特征	
——建立统计指标和统计指标体系	37
三、没有万能的统计指标	44
四、统计指标与经济指标不能混淆	50
五、两种核算体系与两种比值	52
六、怎样进行全面调查	55
七、怎样进行非全面调查	60
八、进行调查要运用“十八般武艺”	63
九、怎样下户调查访问	65

第三章 怎样整理统计数据	68
一、统计误差与数据审查	68
二、统计特有的一种整理和分析方法——分组法	72
三、从马克思使用统计分组法说起	80
四、统计汇总技术	83
五、怎样编制和阅读统计报表	85
六、统计台帐、数据资料库与历史资料	90
七、统计图的学问	94
第四章 综合指标——对具体统计指标的抽象	
概括	98
一、反映现象总量的指标——绝对指标	99
二、两个有关指标对比的派生指标——相对指标	101
三、一道统计试题的争论	115
四、除法与减法之争	118
五、 Σ 与 σ	121
六、代表不同标志值的典型水平的指标——平均指标	123
七、出现次数最多的数值的指标——众数	132
八、位置上的平均数——中位数	136
九、卡尔·皮尔逊的贡献	138
十、调和平均数是不是独立的平均数	140
十一、对平均数要透视	148
第五章 从数量的动态变化看趋势——时间数列	156
一、四种动态比较指标	156
二、怎样正确计算增长速度	161
三、四种序时平均数	165
四、为什么两种算法结果相等	175
五、翻番的计算	177
六、怎样分析发展趋势	179

第六章 抽样调查——一种被广泛应用的科学

调查方法.....	187
一、抽样调查的理论依据——概率论	188
二、抽样调查能代替全面调查吗	195
三、样本与总体的奇妙关系	197
四、怎样组织抽样调查	205
五、抽样误差的计算	211
六、点估计和区间估计	215
七、怎样测定必要抽样数目	221

第七章 要使统计数字会说话——怎样分析统计

资料.....	223
一、揭示现象的内在联系，剖析事物的结构和特点	224
二、对比研究、分析各种比例关系	226
三、指数因素分析法的妙用	228
四、怎样运用相关与回归分析法	235
五、投入产出分析	250
六、分析过去和现在，预测未来	256
七、怎样分析企业潜力和经济效益	265
八、统计分析写作技巧	269
附：正态概率表	275

第一章 统计——对客观现象进行 数量概括的认识活动

什么是统计？人们有各种不同的回答。有的人一听到“统计”一词，便会产生“算盘加数字”的概念。远在1869年著名的统计学家凯特莱(A. Quetelet)在第七次国际统计会议作报告时，就提到“统计学”的定义有180多种。现在经过100多年，各种提法就更多。诚然，定义不是万能的，要用简短的一两句话来描述“统计”的定义是不容易的，但是我们毕竟还要回答这个问题：统计就是指对客观现象进行数量概括的认识活动。

客观现象错综复杂，但都有质量和数量两个方面，任何数量都表现为一定的质量，没有数量就没有质量。人们要认识它，就既要从质量方面认识也要从数量方面认识，只有从质与量的辩证统一中去认识数量，才能对质量的认识深化。一个婴儿出生不久，他抓到苹果往嘴里送，抓到木球也往嘴里送，这时他对苹果的质量毫无认识，等到他知道苹果可以吃而木球不能吃时，他已经开始有质量的认识了。当他吃了一个苹果还不满足，希望吃两个的时候，他开始有数量的认识。认识国家、社会，认识一个城市，认识一个企业，认识人的身体构造，同样离不开数量的认识活动。比如生活在广州的人去了北京，觉得北京的街道和地面要比广州大，但北京的平房又比广州多，街上行人和车辆好象没广州拥挤，如

果我们问，北京比广州到底大多少倍？这就要进行数量认识了，也就是要统计了。1984年的《城市统计年鉴》表明，北京市的土地面积为16 808平方公里，比广州市16 657平方公里只大1%，即等于广州市的1.01倍，但其中市区面积北京市为2 738平方公里，广州市只有1 345平方公里，北京等于广州的2.04倍；再从人口来看，北京市1984年末市区人口575.46万，等于广州市322.16万的1.79倍；从市区住宅的建筑总面积看，北京市为6 578万平方米，等于广州市2 182万平方米的3.01倍。从上述的初步数量认识活动中，我们就可以概括出北京市区大致比广州市区大1倍左右的认识。当然我们还可以用各种经济指标作比较，然后作综合的评价。另外，北京市区面积比广州大1倍，但人口只多79%，这是否说明北京市的人口没有广州市拥挤？仅凭这两个数字还不能下这样的结论，还要调查一下两个市的流动人口和常住而没有户口未统计在内的人数，再作进一步的分析，这样认识就更深化了。

“统计”这个词还是一种“进口货”，拉丁语叫STATUS(状况)或STATO(国家)，18世纪传到英国叫STATISTICS，即用数字表现事实的意思。我国过去“统计”一词一般是指数字加总的意思，现代涵义的“统计”是清末经由日本传到我国。

统计一般是指：统计工作(对客观现象进行数量认识活动的过程)、统计学(对客观现象进行数量认识活动的理论)、统计资料(对客观现象进行数量认识活动的结果)。

统计是认识社会的有力武器之一；统计是国家管理和企业管理的重要手段。

一、统计工作（对客观现象进行数量认识活动的过程）

要对客观现象进行数量认识，首先就要用各种有效的方式方法去搜集反映客观现象的数据资料，组织强有力的统计队伍进行调查了解，数据搜集回来后，进行汇总和整理，最后对统计资料进行分析研究和发表，以便发挥统计的作用，这就是统计工作的全过程。这个过程可以分为四个阶段：统计设计、统计调查、统计整理和统计分析。

统计工作的内容很广泛，如统计方针政策、统计体制、统计立法、统计调查方案和实施、报表管理、资料整理和提供、统计服务咨询、统计干部培训、统计组织建设等等都属于统计工作。

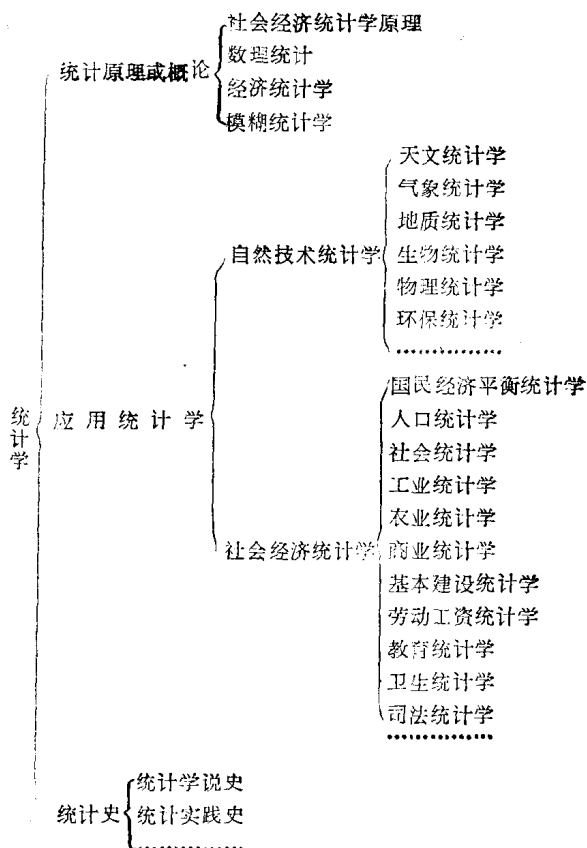
有人问：“统计不过是加加减减的工作，还要方针政策吗？”我们说“要”。统计是一门学问，而且是国家、企业、科研部门等都离不开的重要工作，当然有方针政策。比如根据党的十二届三中全会精神，国家统计局提出了由“封闭式”统计转变为“开放式”统计，这就是一项统计方针政策。

“封闭式”统计的主要特点是“自我封锁”，就是过分强调保密。本来某些统计资料是不必保密的，但过去层层保密，什么数字都保密，连全国人口总数都不敢公开，联合国统计全世界人口对我国的数字只能靠推算。有的单位接到统计资料，因为上面写有“绝密件”，按制度只好锁在保险柜里，长久见不到天日，发挥不了统计的作用。现在要转变为“开放式”统计并不是说统计资料一律要公开，而是要建立和健全各种保密和解密的制度，尽可能把各种信息和数据送到需要

者手里。因此，要贯彻这条方针，就要建立各种信息和数据的服务系统，加强对统计资料的整理和提供工作，这样也就把统计搞活了。

二、统计学(对客观现象进行数量认识活动的理论)

统计学就是指导统计工作的理论，是统计实践经验的升华。可是统计实践史和统计学说史的发展速度并不是同步的。统计的实践，可以追溯到伏羲画八卦和公元前2200年大禹治水时代，或公元前3050年埃及造金字塔时代，那时对主要的国情国力，已作了一些统计调查工作，但还没有形成统计理论。一直到资本主义社会，由于大生产和社会分工日益发达，随着文化和市场经济活动发展的需要，在社会经济和自然科学领域，统计都得到广泛的应用，特别在工业、商业、税收、航运和外贸等领域，统计发展更快。这样人们开始总结统计工作的经验和吸收其他学科的研究方法，形成了比较系统的统计理论知识，产生了统计学，到现在也只有300多年的历史。但经过几百年来的论争和发展，现在，在统计学的大家庭中，已有不少家庭成员，形成各种各样的统计学：有原理、概论性质的，有应用、实用性质的，有历史性质的，它们之间的关系，可以图示如下：



统计学在形成的过程中出现了三个学派：

第一个出现的学派，叫国势学派，又名记述学派。这一派的代表人物有海尔曼·康令(1606—1681年)、高特弗里德·阿享华尔(1719—1772年)等。他们都在德国的大学讲课，主要讲述欧洲各国的人口、领土、政治、财政、经济、宗教、地理、科研艺术等，用文字记述，并定名为统计学(德语STATISTIK)。可是只有统计学之名，没有统计学之实，因为他们

不讲数量表现和数量分析，实际上不是我们现在所讲的统计学的意思。

第二个出现的学派，叫政治算术学派。这一派的代表人物是英国的威廉·配第(1623—1687年)。他出生于英国一个水手家庭，当过水手、牛津大学教授、军队随从医生和土地分配总监等职，并于1671—1676年间写了一本叫《政治算术》的书，内容共有十章，首次采用数量对比分析的办法把英国和其敌对的荷兰及法国两强作了比较，用事实告诉英国人对其国家前途不必抱悲观情绪。这本书是用统计方法来研究政治经济的实质问题的。当然还不是统计理论的书，他也不叫统计或统计学，但实际上是我们现在所说的社会经济统计的创始，虽无统计之名，却有统计之实。所以马克思对威廉·配第评价很高，说他是“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”^①。政治算术学派可以说是社会经济统计学的来源。

第三个出现的学派叫数理统计学派。这一派的主要代表人物是比利时的阿道夫·凯特莱(1796—1874)。他是一个数学家、天文学家、生物学家和被称作统计学之父的博学多才的学者。他把数学中的概率论运用到统计学中来，从而使统计学迅速发展而形成数理统计学。主要的著作有《统计学的研究》、《社会物理学》等。

由于一开始，数理统计就是作为一种方法论提出来的，并且从生物学中吸收了相关与回归的理论，而且紧密结合自然科学的实践，不断充实新的内容，同时也向研究社会问题发展，因此随着资本主义的发展而很快地发展，使社会经济

① 马克思：《政治经济学批判》，人民出版社1955年版，第25—26页。

统计学的性质也逐渐发生变化，由原来的实质性科学转变为一种方法论的科学，即不是直接研究社会经济现象本身，而是提供研究的方法。数理统计在西方国家，一直占统治地位。

我国在解放前也存在社会经济统计学派和数理统计学派两派。我国最早编写统计学著作的人有彭祖植(1907年)和沈秉诚(1909年)主要是介绍德国社会统计学派的观点。

解放后，我国全面学习苏联的统计学，只承认社会经济统计学而否定数理统计学。谁如果不把社会经济统计学作为一门实质性的科学而看作一门方法论的科学时，谁就会被扣上“资产阶级观点”的帽子。一直到党的十一届三中全会以后，才逐步扭转这种偏向。但在80年代初期，又有人认为只有一门统计学即数理统计学；而社会经济统计学只是一种“政府统计”，不是一门统计学。经过激烈争论，现在比较一致的看法是两门统计学并存，要正确对待两门统计学，两门统计学发展的结果可能融合为一体。

为什么说两门统计学都独立存在呢？这是因为：

第一，两者的研究对象不同。社会经济统计学是一门社会科学，是研究如何做好社会经济统计工作的理论和方法，而数理统计学是研究客观现象中随机现象的具体规律，它属于自然科学。所谓随机现象就是指客观事物中存在着偶然性，有的现象可能出现也可能不出现。如今天可能下雨也可能不下雨；某种商品的价格可能上升也可能下降；某个企业可能盈利也可能亏本等。

第二，理论基础不同。由于前者是研究如何对社会现象进行数量概括的认识问题，因此有强烈的阶级性，以政治经济学作为理论基础；后者既研究社会现象，也研究自然现象，

它基本上和数学一样是研究空间形式及其数量关系的。因此，它没有阶级性。它的理论基础是概率论。

第三，研究方法不同。社会经济统计学的主要方法是大量观察法、综合指标法、分组法、指数因素分析法等。而数理统计学的主要方法则是大数法则、抽样推断、相关与回归、实验设计、时间序列、参数统计分析、方差分析、序贯分析等。

总之，这两门统计学都是客观存在的，在认识活动中是相辅相成的。目前，在社会经济统计学中已越来越多的使用数理统计方法。

除此以外，还有两门正在形成的新的统计学：经济统计学和模糊统计学。经济统计学是苏联统计学界近年的新作。它从统计的角度出发，对各种经济现象的质的规定性及某些规律性作理论的概述，既不同于社会经济统计学也不同于政治经济学。

模糊统计学是正在形成的新的学科。模糊这个词，也是从西方引进来的。主要是一个美国教授创建了模糊集合的数学理论。我国统计界有人联系对社会经济现象的观察，认为确实存在一些“亦此亦彼”的模糊现象，它们之间界限不太明确，如大小、长短、老少、先进与落后等。很多现象是不固定的、模糊的，对这些现象怎么研究呢？用传统的统计方法来确定它们的数量界限是不行的，所以要用模糊数学作为理论基础来研究，形成一门有其特定的研究对象和研究方法的新的统计学，但目前不少人对此还持怀疑态度。

三、统计资料 (对客观现象进行数量认识活动的结果)

统计资料的具体表现就是统计数字, 每个统计数字, 都表现某一客观现象的数量特征。尽管客观现象错综复杂, 千变万化, 但许多现象都可以通过统计数字来描述。当然, 有些现象是不能用数字来表示的, 比如齐白石是全国有名的大画家, 关山月是华南有名的大画家, 如果我们问, 齐白石比关山月水平高百分之几? 这就难以用统计数字来回答。

统计数字可以反映事物的数量表现、数量关系和数量界限。

数量表现, 指事物的量的多少。如反映事物的规模、水平、速度、程度等。1984年我国的国民经济和社会发展的统计公报说“全民所有制单位固定资产投资1160亿元”就是指主要的建设规模; “钢产量4337万吨”, 就是指生产水平; “全年社会总产值比上年增长13%”, 就是指生产发展速度。

数量关系, 指一事物与另一事物(或几种事物)的数量变化关系、比例、比率等。比如说多施一斤肥料可以增产3斤粮食; 1982年全国全民所有制工业产值利税率23.4%等等。为什么产值利税率23.4%是反映数量关系的数字呢? 这是因为这个数字是表明全民所有制工业企业的利润和税收总数与总产值的比率, 也就是说生产100元的工业产值, 可提交23元多的利润和税收, 这个数字越大, 说明经济效果越好。

数量界限, 指事物从量变到质变的临界数量。如水温在 0°C 以下时结冰, 100°C 时成蒸汽, 0°C 和 100°C 就是数量界限。从社会现象来说, 过去认为看一个国家是不是工业化国家,

就看它的工业总产值占工农业总产值的比重是不是超过70%，这这也是一个临界数量。1986年11月2日《羊城晚报》登载一条消息说：“据香港统计处发表的最新资料表明，今年头八个月，内地输往香港的商品总值占香港进口总额的比重由去年同期的23.7%上升到28.8%，居香港进口商品的首位……达492.63亿港元。……内地与香港之间的贸易由去年同期的逆差70.55亿港元，转为顺差143.32亿港元”。从逆差转为顺差，是由于输往香港的商品总值达到492.63亿港元的量变的结果，因此，这个数字就是数量界限。

有些数量界限，虽然要靠大量观察，但不必统计和计算，一般不作为统计资料。比如动物的寿命，据新加坡《联合早报》1986年10月20日报导可列表如下：

表1-1

大象	可以超过两百年
鳄鱼	可以超过一百年
鸚 鵡	可以活七、八十年
獅、虎	可以活四、五十年
狗	可以活十多年
兔 子	可以活八、九年
老 鼠	可以活几年
苍 蝇	不到一个月
蚊 子	不到一星期

统计资料反映的数量和数学上的数量是不同的，它不是研究抽象的数量关系，统计反映的数量是紧密联系现象的质的规定性来进行的。比如1983年全国平均每人消费粮食232.23公斤。这个数字的粮食消费量的性质是指贸易粮，它包括饮食业消费掉的粮食和其它用粮在内。

统计资料，如果按指标时间先后次序排列，可以形成一

种动态资料，而且可细分为时期资料、时点资料和次数资料三类。

时期资料，是指每个统计数字反映的是客观现象在一定时期(如一年、一个月)连续发生的总数量。如：

表 1-2 我国历年工农业总产值

年 份	工 农 业 总 产 值 (亿元)	年 份	工 农 业 总 产 值 (亿元)
1949	466	1971	3 482
1950	575	1972	3 640
1951	684	1973	3 967
1952	810	1974	4 007
1953	960	1975	4 467
1954	1 050	1976	4 530
1955	1 109	1977	4 978
1956	1 252	1978	5 634
1957	1 241	1979	6 379
⋮		1980	7 077
⋮		1981	7 580
⋮		1982	8 291
⋮		1983	9 209

资料来源：《中国统计年鉴——1984》

本表按当年价格计算

时期资料有两个特点：

(1) 数字可以加总或累积。因为每个时期的数字都是反映客观现象在该时期发生的数量总量。

(2) 数字受时间长短的影响。时间愈长，数字愈大。比如我国钢产量，1976年年产2 046万吨，1980年发展到3 712万吨，但是把“五五”时期5年产钢量加总，则达到14 758万吨。

时点资料 是指每个统计数字反映的是客观现象在某一时刻或瞬间发生的数量。如：

表 1-3

我国历年年末人口数

单位：万人

年	份	人	口	总	数
	1949			54	167
	1950			55	196
	1951			56	300
	1952			57	482
	⋮			⋮	
	1980			98	705
	1981			100	072
	1982			101	541
	1983			102	495

资料来源：同上

时点资料也有两个特点：

(1) 数字不能相加或累积。因为加总了不能说明什么问题。比如说1982年加1983年我国年末人口数为204 036万人，这毫无意义。

(2) 数字大小与时点的间隔长短无关，并不是间隔时间越长，数字就越大。比如1978年底我国猪的头数为30 129万头，相隔一年到1979年底，增加到31 971万头，可是相隔四年到1982年底，反而少了，只有30 078万头。

时点资料所讲的“瞬间”一般是指很短的时间里事物所达到的水平。例如我国第三次人口普查全国人口数1 031 882 511人，是指1982年7月1日零时这一瞬间的人口状态，但是瞬间的概念也不是太严格，比如我们讲人口或猪的年底数字，一般是指12月31日这一天的情况。

时点资料还有连续时点资料和间断时点资料之分。比如把全国建国以来每年末的人口数排列起来就叫连续时点资