

第 一 册

《统计学》辅导

曹毓侯 主编

中国展望出版社

《统计学》辅导

第一册

曹毓侯 主编

吕月瑞 李目珍 刘兰亭 任若恩

中国展望出版社

内 容 提 要

《〈统计学〉辅导》第一册的主要内容，除对统计学的发展历史作了简明的叙述之外，主要是阐述统计学的一般原理、原则和方法，并运用这些原理和方法对统计原始资料进行加工整理，说明次数分配的性质，反映统计数量的集中趋势和离散程度，以及用或然率（概率）的原理，判断统计推论的可靠程度和对统计总体估计的真确性等。所阐述的各种有关统计方法是先理论后应用，比较全面，系统和深入，适于从事统计理论研究、教学人员和统计工作者学习参考。

《统 计 学》辅 导

第 一 册

曹毓侯 主编

中国展望出版社出版

（北京市白塔寺太平桥大街4号）

北京市新华书店发行

太原新华印刷厂印刷

开本850×1168毫米1/32

8.125 印张

197,668 千字 1983年8月

太原第1版

第1次印刷

1—55,000册

统一书号：4271·097

定价 0.95 元

前 言

1976年再版的香港中文大学胡孝绳教授所著《统计学》一书，对统计学的基本原理、原则和方法作了全面、深入的论述，还对有关统计数学公式作了详细地推导，由浅入深，条理清楚，内容丰富，颇受广大读者欢迎。为了供学习《统计学》者参考之用，我们按照该书的章节顺序编写了《〈统计学辅导〉》。

原《统计学》一书共分为30章，第1—9章主要是阐述统计学的基本原理、原则和方法；第10—19章主要阐述概率分布的基本原理，并对计算公式进行推导；第20—30章主要是阐述统计学的一般方法及其在社会经济领域中的具体应用。

《〈统计学〉辅导》按照原书的三个基本段落分为三册编写。对原书内容作了全面、系统的叙述，并结合实际情况和当前所采用的有关统计方法作了详细的补充说明，力求通俗易懂。各章后面附有思考题和练习题，每册后面另附有答案。

参加《〈统计学〉辅导》编写的主要有：曹毓侯、吕月瑞、李目珍、刘兰亭、任若恩。第一册的统计学史略、第一章、第二章、第三章、第八章、第九章由任若恩提供初稿，第四章、第五章由李目珍提供初稿，第六章、第七章由吕月瑞提供初稿，经曹毓侯、吕月瑞、李目珍审编，最后由曹毓侯主笔修改定稿。庞风仙、任若恩编写了思考题及练习题答案。

在编写辅导材料的过程中，得到了山西省统计学会的鼓励和支持，特在此表示谢意。

由于我们的水平所限，加上编写时间仓促，书中难免存在缺点和错误，恳切希望广大读者批评指正。

编 者

1983年3月

目 录

统计学史略	1 ~ 6
第一章 绪论	7 ~ 22
一、统计学的性质	7
二、统计资料	10
三、变数与函数	17
四、数的概念	18
第二章 统计图示法	23 ~ 32
一、图示的意义与方法	23
二、直角坐标的图示	24
三、对数值的图示	24
四、对数与半对数的图示	26
五、各种图形	27
第三章 次数分配	33 ~ 60
一、原资料	33
二、次数表	34
三、组限的选定	37
四、组距的大小	41
五、不等组距的应用	42
六、组中点的位置	44
七、累积次数表	45
八、次数分配图示：直方图与次数折线	47
九、次数分配图示：次数曲线	51

十、累积次数分配图示：穹形	54
十一、洛伦兹 (Lorenz) 曲线	57
第四章 集中趋势	61 ~ 100
一、平均数	61
二、算术平均数	62
三、几何平均数	69
四、调和平均数	73
五、平方均数	76
六、中位数	78
七、众数	83
八、四分位数、十分位数及百分位数	87
九、各种平均数的关系与优劣	93
第五章 离势数量	101 ~ 119
一、分配的离势	101
二、全距	102
三、标准差与变易数	103
四、平均差	112
五、四分位差	115
六、相对离势的测定	116
第六章 分配的动势	120 ~ 130
一、动势的概念	120
二、动势与次数分配性质关系的分析	125
三、动势的校正	128
第七章 偏态与峰度	131 ~ 151
一、偏态的概述	131
二、偏态的测定方法	132
三、峰度的种类	144
四、峰度的测定	147

第八章 统计推论与或然率(概率)	152~175
一、演绎与归纳	152
二、统计的推论	153
三、或然率的意义	156
四、或然率的公理	157
五、加法定理与乘法定理	158
六、排列与组合	166
七、从或然率分析机遇变数与函数	169
第九章 选样方法	176~203
一、样本与选样	176
二、任意抽样	179
三、非任意抽样	190
四、总体数值的估计	192
五、真确性与样本大小	198
附: 思考题与练习题解答	205~252

统计学史略

《统计学》一书，在三十章之前，作者对统计学的发展历史作了简略的论述。

书中叙述统计学的历史，是从统计工作最原始的形态讲起。在中国，这种最原始形态的统计可追溯到伏羲画八卦、大禹治水的时代。随着封建制国家的建立、发展，统治者需要了解全国人口、土地、生产等方面的情况，所以，历代政府对统计工作都很重视，组织人力对这些资料进行搜集、整理，并用文字、图表描述下来，作为他们核定和征收赋役的依据。作者认为，这类统计工作，在我国历史上曾达到很大的规模，但因没有学者专门从事理论方面的研究，所以一直没有形成系统的统计学说。

在西方各国，最原始形态的统计，可追溯到公元前3050年，埃及建造金字塔时，对全国人口、财产所进行的普查，以及罗马二世对埃及土地所进行的调查等。

上述这一切，作者认为都属“统计事务”、“统计实务”或“统计工作”，都未涉及统计理论的研究。并认为统计学理论的研究开始于古希腊亚里斯多德时代，自那时起，统计学的发展可分为四个时期。以下分别说明这四个时期的具体情况及其对现代统计学的关系和影响。

第一个时期：约自公元前三世纪至公元十七世纪，经历了两千多年的时间。这一时期的统计具有浓厚的哲学色彩，主要是用文字记述政情，最早开始于古希腊的亚里斯多德，他用文字对希腊城邦政情进行比较叙述，编写了大量的城邦政情论文。书中称其为“城邦纪要”形态的统计工作。

“城邦纪要”形态的统计工作延续的时间很长。到十七世纪和十八世纪在欧洲特别是在德国受到空前的重视与广泛地应用，并形成了统计史上第一个学派——“记述统计学派”。这一学派的代表人物主要有德国的学者康令(Corning)和阿痕瓦尔(Achenwa-11)。他们继承和发扬了城邦纪要形态的统计思想与方法，广泛地用以研究国家的政治、社会和经济状况。康令将国家土地、生产及其它资源的数字资料作为“国情纪要”课程的教材，阿痕瓦尔所编统计讲义对欧洲各国的物产、人口、土地、兵力等状况也作了详细的叙述和比较，还将拉丁文的Status一词引用为专指统计学的名词。由此就产生了“国势学”(Staatenkunde)，这就是形成统计学(Statistics)一词的直接模型。在记述学派内部由于采用的记述形式不同，又形成了两派，即文字记述学派，主张专用文字来记述；图表统计学派主张用统计表和几何图形来记述。所以有文字派与图表派之争。正是由于这个学派的活动，使“城邦纪要”形态的统计工作发展达到顶峰，但也接近了尾声。作者认为，他们的努力“仍不脱”“城邦纪要”的巢臼”，仍表现为纯粹哲学形态的统计分析。

第二个时期：17世纪与18世纪欧洲各国，其中主要是英国的学者、教授对生命统计、保险统计和经济统计所做的部分数学研究，使得统计学的研究具有浓厚的社会学气质。这就标志着统计学的发展进入了一个新的历史阶段。

这一时期产生了统计史上的另一个重要学派，作者称其为数理统计学派。这一时期的代表人物是英国的统计学家戈兰特(Graunt)。17世纪中叶欧洲各国发生了周期性的瘟疫，人口大量死亡，严重地影响到农业生产，对外贸易及公共管理事业的发展。再加上工业革命的影响，就使得经济、社会和政治的管理工作日趋重要和复杂化。社会上下不仅对人口的出生死亡都非常重视，而且迫切要求统计工作能够提供可作为管理依据的统计资料

和分析方法。戈兰特的研究活动就是在这种情况下进行的。他在生命统计、保险统计、经济统计中都做了一些数学研究，即形成了“政治算术”学。特别要指出的是他在统计中发现人口的出生与死亡率经久不变，据此创立了“大量恒静”定律(Law of inertia of Large numbers)，并成为统计学的基本原理。“大量恒静”定律是指某一事物的某一性质或状态在反复观察或反复试验中保持不变。如某一位妇女所生婴儿是男的还是女的，这是十分偶然的，但一个国家、一个民族或一个地区某一时期出生婴儿的性别比例则保持在一个稳定的比例上。这种性质只有通过大量观察才能揭示出来。其他政治数学家也设计数学分析的方法来研究经济、社会和政治等方面的问题，使得统计学的数学性质日益增加，并使得这一时期的统计学研究充满了社会学气质，为现代统计方法奠定了基础。

第三个时期：统计学发展的这一时期是以或然率(概率)的数学理论的发展并成为统计学的一个基本理论为标志的。正如书中所说：“倘无或然率理论，统计学殊难有今日之成就，将不足获得普遍的信赖与发挥广泛的功用”。或然率早在16世纪初就由于应用数学方法解答机遇游戏中各项疑问而产生，所以作者认为第三个时期应“始自16世纪”，基于这种认识，作者在叙述这一阶段的情况时，即从十六世纪或然率的产生谈起，并对其发展过程作了比较详细的叙述。书中第三页的第三段到第五页的第一段所谈的内容就是或然率本身的发展史。而或然率真正成为统计学的一个基本理论，并形成适用于许多场合的一种统计方法则是19世纪的事情了。

这一时期创立的常态曲线、最小平方、相关分析与迴归分析法，以及中数、四分位数、百分位数等都是对统计学理论的新发展，都是构成现代统计学基本内容的重要组成部分。

在这一时期中对或然率的创立与发展贡献较大的学者，主要

有巴斯卡(Pascal), 费尔马(Fermate), 贝努里(Bernoulli), 德莫维尔(De Moivre)等人。贝努里系统地研究了或然率的理论, 并有专门著作出版, 专论或然率原理。他把或然率建立在数学的基础上, 并提出或然率原理可以应用于社会经济等各个方面。德莫维尔对或然率理论作过很多研究。他最早发现了“常态曲线”并计算出该曲线各不同间隔的或然率。这条曲线及其有关的理论是现代统计学的重要组织部分。正是由于他们的努力, 才使得或然率的理论系统化、一般化, 从而成为一种适合于许多场合的统计方法, 在自然科学与社会科学领域广为采用。如在物理学上被用以研究实验的观察误差; 在天文学上被用以研究天体运动或位置的误差; 在政府统计工作中大量资料被收集、组织、列表并加以解释。

19世纪中叶, 对统计学理论的新发展贡献较大的统计学家有: 比利时统计学家凯特勒(A. Quetelet), 他的主要功绩是把或然率的原理和方法引进统计学, 并将统计学发展成为一种可以应用于任何科学的一般方法; 英国人类学家高尔顿, 他在运用统计方法研究生物学的基础上创立了相关分析方法与回归分析方法及相关系数的计算, 并发明了中数、四分位数、百分位数和四分位差等; 皮尔逊(Pearson), 他对以常态分布为基础的大样本理论有重大贡献, 并已开始研究从样本中导得结论的抽样误差问题。

第四个时期: 自19世纪末到现在。这一时期统计学发展的特点是在综合了第二、三两个阶段研究内容的基础上, 进行了改进与发展, 着重研究如何从样本提供的信息推论总体。或然率理论得到极大地发展, 成为现代统计学的理论基础, 从而使统计学发展成为一门数学性质的学科, 成为一种分析性的科学。其应用范围越来越广, 成为一门几乎可以应用于一切科学领域的方法论科学。

这一时期主要代表人物是戈司特(Gosset), 他创立了以小

样本代替大样本的方法——“T分配”，为样本资料的统计分析与解释开辟了新纪元。他的思想为英国统计学家费歇（Fisher）所发展，费歇是对现代统计学有重大贡献的一个人，在很多现代统计学教科书中我们都会看到大量的以他的名字命名的定理和公式。他同奈曼（Neyman）、爱古·皮尔逊（Egon·Pearson）等人继续研究小样本的统计理论与方法并共同领导了统计学界关于“实验结果显著性的测定理论和估计理论”的研究活动。奠定了实验设计的基础。

以上是仅就书中叙述的内容所做的说明。此外，我们想补充几点意见供读者参考：

1、我国统计学界在叙述统计学发展史时通常都是按学派划分时期，以学派的活动为线索论述的。第一个出现的学派就是记述学派（又叫作国势学派），主要的代表人物是康令与阿痕瓦尔，（又译作握天华）。他们首先使用统计学这个术语。这个学派的活动从17世纪开始，一直延续到19世纪中叶。

第二个学派是政治算术学派，这个学派起源于英国。最初的研究活动也是在17世纪，主要代表人物是戈兰特与威廉·配第。这个学派的活动声势与影响都很大，到19世纪中叶逐渐占据了主导地位。

概率论本身的发展历史可以追溯到15世纪，初具形态也在17世纪。但是将概率论的思想引进统计学却是19世纪的事，是比利时统计学家凯特勒的贡献。他因此成为统计学史上第三个学派——数理统计学派的奠基人。

在18世纪末到20世纪初德国还出现过一个社会统计学派，这个学派认为统计的研究对象为社会集团，研究方法为大量观察法。他们主张统计学不是方法论的科学而是实质性的科学。主要代表人物是恩格尔（Engel）和梅尔（Mayr）。

从18世纪开始，统计学逐渐运用数学，但概率论对政治算术

学派的直接影响很小。一直到19世纪才使二者紧密联系起来产生了数理统计学派。这一学派的代表人物有高尔顿、皮尔逊、戈司特、费歇等人。这一学派在目前欧美统计学界仍占主导地位。

按照学派来叙述统计学的历史，从记述学派开始至今仅有三百多年的历史。

2、作者没有提到罗马尼亚著名统计学家瓦尔德(Wald)。他的两本著作《序贯分析》和《统计决策函数》开创了现代统计学两个崭新的部门，是统计学的最新发展。

读者在阅读这一节时，希望能抓住这几个问题理解：

一、统计工作与统计学两者从历史的发展看有什么联系与区别？

二、统计学发展中每个时期的主要特点是什么？有哪些代表人物？这些代表人物有什么重大贡献？

三、统计学史上的三个学派的主要特点是什么？对现代统计学有什么影响？

四、从历史发展的角度看，统计学如何从一门主要记录情况的科学发展成为一门统计分析的科学？如何从主要探求国家实力的科学发展为几乎可以应用到一切科学领域的方法论科学？

第一章 绪 论

第一章主要是论述统计学最基本最一般的概念。统计学属于一种方法论的科学，内容包括统计技术与统计原理。

1、统计学包含多种多样的统计方法，统计方法的三个主要用途是：取得未知数字，说明事物的真实状况；对原始数计算出来的概括结果进行比较；从资料导出推论和结论。

2、统计资料就是数字事实，统计群体是统计资料的实体，统计项目是统计资料的表现，统计数列是统计资料的组织。

3、变数就是数列，根据影响变数大小的因素不同，可分为时间数列，品类数列，次数数列。根据变数本身的特点，可分为自变数与倚变数，两者之间的联系就是函数关系。

4、统计学上的数，具有特殊的性质，一般都是近似数。需要特殊的处理方法。

一、统计学的性质

在这一节中，主要阐述了三个问题：1、统计学的研究对象和统计学的定义；2、统计学所包括的内容；3、统计方法的四种用途。

（一）作者认为统计学一词，最普遍的解释是：统计学也可称为统计方法，就是从事搜集、显示（即反映）、分析、解释被研究对象的数量资料，籍以表现事物真实情况的一种学问。并明确指出：不论是社会的、自然的、或实验的，按照特定性质，加以归类，而划分为各种不同群体（即总体）的客观事物，凡是能

够用数字事实表现的都可以作为统计学的研究对象。而根据统计的研究目的，所确定每个个体（即总体单位）都具有共同的条件或概念的群体（即同质总体）就叫作统计群体。但构成统计群体的每个个体又具有不同的性质。统计就是观察、研究每个个体不同性质的表现特征。由此可见，统计学就是以搜集、显示、分析、解释等统计技术为手段，而以揭示出统计群体的本质为目的。

关于统计学的定义，历来众说纷纭。1869年凯特勒（A. Quetelet）在第七次国际统计会议时的报告，就提到已有180多种。到了现在提法就更多了。例如瓦渥（Waugh）认为统计学是研究群体数字资料，从而得出简明概括事实的一类方法。金（King）则认为统计学是依据搜集事实或估计数字的分析结果，用来研究自然现象或社会现象的方法。而耶尔和肯达尔（Yule和Kendall）认为，统计学包括各种方法，适用于说明多种原因的影响的数量资料。波勒（Bowley）认为统计学是计算平均数的科学。

（二）就统计学的性质来说，是一种方法论，其内容包括统计技术和统计原理两个方面。而统计原理大部分是数学性质，主要是用数学的方法证明统计学中所用方法、公式的正确性，以满足统计技术（即搜集、显示、分析、解释）的需要。至于统计技术的应用，最终目的在于获得对于事物的认识。所以，统计学就成为揭示事物本质的一种手段或科学方法。也就是对从实验或其它来源所取得的观察数值加以描述或解释。

统计方法与实验方法是科学研究的主要方法，这两种方法的性质是不同的，应用的范围也不相同。在社会科学范围更多地使用统计方法而难以使用实验的方法。

实验方法就是控制其它事物的变化来观察某一种事物变化的情况。但是客观事物特别是社会经济领域内的许多现象受多种因

素的影响，各种因素之间交错发生作用，我们往往根本无法控制它，所以就任其变化，然后观察记录下所有作用力量的变化，进而判定每一个因素对事物变动结果所发生的影响。这就是统计方法。一般说来统计方法的结果不如实验方法的结果那样准确，但是在很多场合下，统计方法却是唯一能够获得结果的方法。在另外一些学科中，例如物理学与天文学这样一些所谓精密科学中，实验方法与统计方法同时使用，特别是运用统计方法来设计试验、分析试验结果与误差等方面，都发挥了重大作用。

作者认为，统计方法大致可以分为两类：一类是叙述统计方法，一类为分析统计方法。两者的意义和用途各不相同。叙述的统计方法往往是指的统计工作的初步阶段，主要是搜集与显示这两种技术，目的在于描述观测的结果；分析统计方法则包括分析与解释两种统计技术，也就是在叙述统计方法的基础上进行分析与解释以剖析状况或决定趋势与可能发展。这两种方法经常是结合起来使用的。

统计方法的具体用途有如下三个方面：

(一)统计方法的首要用途是说明事实与数字的意义，取得未知数字和对所研究事物的真实情况作出确切的说明。为达到这一目的，就需要采用不同的统计方法。至于选用哪种方法要根据所研究事物本身的性质和特点以及研究与调查的目的来决定。例如，根据25,000个婴儿的体重表来研究婴儿体重的具体状况，从中得出概括的结论，找出最高的体重与最低的体重，计算平均体重等，这些仅需采用叙述的统计方法。

再如，假定有100只兔子，50只喂甲种饲料，50只喂乙种饲料，三个月后秤一下体重并从秤量的结果来研究哪一种饲料使兔子增重的效果较好，这就要采用分析统计方法。

(二)统计方法的第二个用途是对原始数字计算出来的概括结果进行比较。而平均数就是根据原始数字所计算出来的概括结