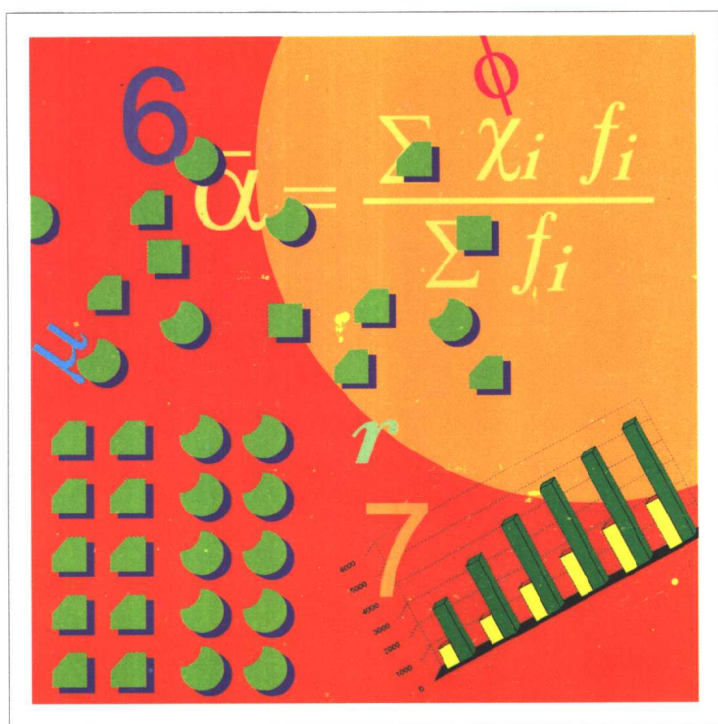


普通高等教育“九五”国家级重点教材

统计学原理



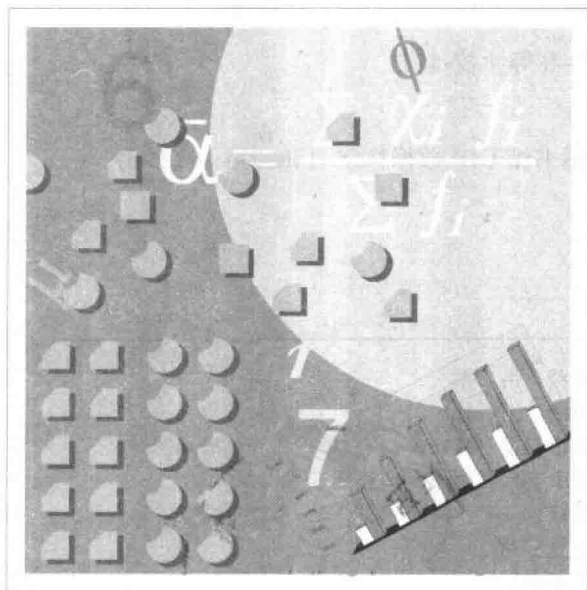
主编 黄良文 副主编 曾五一

中国统计出版社



普通高等教育“九五”国家级重点教材

统计学原理



主编 黄良文 副主编 曾五一

中国统计出版社

(京)新登字 041 号

图书在版编目(CIP)数据

统计学原理/黄良文主编.

—北京:中国统计出版社,2000.6

普通高等教育“九五”国家级重点教材

ISBN 7-5037-2922-8

I. 统…

I. 黄…

Ⅱ. 统计学-高等学校-教材

Ⅳ. F270

中国版本图书馆 CIP 数据核字(1999)第 10225 号

责任编辑/杨映霜 吕 军

出版发行/中国统计出版社

通信地址/北京市三里河月坛南街 75 号 邮政编码/100826

办公地址/北京市丰台区西三环南路甲 6 号

电 话/(010)63459084,63266600—22500(发行部)

印 刷/科伦克三莱印务(北京)有限公司

经 销/新华书店

开 本/787×1092mm 1/18

字 数/400 千字

印 张/22.5

印 数/35001~40000

版 别/2000 年 6 月第 1 版

版 次/2001 年 11 月第 6 次印刷

书 号/ISBN 7-5037-2922-8/F·1165

定 价/30.00 元

中国统计版图书,版权所有,侵权必究。

中国统计版图书,如有印装错误,本社发行部负责调换。

出版说明

根据原国家教育委员会《普通高等教育“九五”国家级重点教材立项、管理办法》的要求,国家统计局经过专家评审,五种教材被立项,其中:三种教材是在“八五”期间规划统计教材的基础上进行修订的,二种教材属“九五”期间新编规划统计教材,这五本教材的编审工作由全国统计教材编审委员会负责审定。

“抓重点,出精品”是“九五”期间普通高等教育教材建设与改革工作的核心。按照教育部的要求,国家级重点教材都应建设成“九五”普通高等教育的精品教材。根据这一精神,这批教材力求适应我国政治、经济、科技、教育等改革的形势,充分反映改革的成果,同时适应专业目录调整及专业面拓宽以后教学改革的实际需要,科学、合理地设置教材体系和安排教材内容,努力提高教材质量。

相信通过这批教材的出版、发行,对我国普通高等教育统计教材建设工作将起到较好的示范、导向作用;对提高统计教育水平和培养面向 21 世纪的统计人才也将发挥积极的促进作用。

限于水平和经验,这批教材的编审、出版工作还会有缺点和不足之处,诚恳欢迎教材的使用单位、广大教师和同学们提出批评和建议。

全国统计教材编审委员会

1999 年 3 月

本书初版《社会经济统计学原理》是高等学校统计专业主干课的教材。1997年经教育部审定,向全国高等学校推荐作为经济类和管理类专业核心课程的通用教材。同年,教育部将《统计学原理》列入“九五”普通高等教育国家级重点教材编写计划,重新立项组织编写,作为跨世纪核心课程教材。

近几年来,我国经济、教育以及统计学科有许多发展和重大变化。国家已经确定以普查为基础、以抽样调查为主体、多种调查方法结合运用的统计调查方法体系,相应进行了统计体制重大改革;在社会上,特别在市场经济活动中,统计方法在金融、证券、保险、投资、理财各方面的应用领域迅速扩大;统计学自身的发展形成了独立的统计学科。涵盖自然科学与社会科学的大统计思想,不仅提高了统计学的学科地位,而且也有利于统计学科间的融合、借鉴,促进统计科学的发展。基于以上认识,新编的《统计学原理》努力体现以下特点:

(一)科学性。作为一门认识方法论的科学,统计学可适用于自然现象和社会现象。由于统计对象都具有数量性、总体性和变异性,因此,在认识方法上也具有一定的统一性。即存在通用的数据搜集、整理和分析方法。《统计学原理》从统计学作为一门独立科学的认识高度,建立自己的科学体系。阐明统计思想、理论背景、方法思路和发展前景。加强统计理论和方法的研究,将大大提高对客观现实特别是对不确定现象规律性的认识能力。我们认为,一方面,数理统计和社会经济统计应该相互促进、相互补充,共同提高。社会经济统计应该尽可能运用数理统计方法来解决社会经济的实际问题。把两者割裂甚至对

立起来是不合适的。另一方面,由于社会经济现象的复杂性,社会经济统计除了数理方法外,还有自己独特的方法。根据这一认识,本书在体系内容上作如下安排:第一部分包括第一章导论、第二章数据收集、整理和显示、第三章统计分布的数值特征,研究统计学的一般问题。第二部分包括第四章抽样和抽样分布、第五章统计推断、第六章相关与回归分析,研究一般统计方法及其在社会经济领域的应用。第三部分包括第七章统计指数、第八章时间序列分析、第九章统计预测、第十章统计综合评价,主要是讨论社会经济统计方法的特有特点。体现全书既考虑共性也考虑个性的特点。

(二)实用性。根据全国统计教材编审委员会的要求,本教材是为授予经济学或管理学学位的高等学校各专业统计课程而编写的。因此,本书是以社会经济现象为研究对象的方法论基础教材。统计方法为认识社会经济规律服务,明显地受所研究对象的性质制约。作为社会经济统计学应该以相关的社会经济理论为指导,既要防止脱离社会经济现象实质的抽象数学化倾向,也要避免脱离数量分析的纯经济理论研究,而必须努力把两者有机地结合起来。

根据经济管理类教育培养目标的要求,通过本课程的教学要达到以下目的:(1)为经济管理统计提供统计调查、资料整理汇总和统计分析的一般原则和方法;(2)为进一步学习有关统计的专门知识,奠定理论和方法基础;(3)为学习其他经济、管理课程和从事经济研究工作提供数量分析的方法。

(三)基础性。《统计学原理》作为经济、管理类专业基础核心课程,注意基础理论、基础知识和基本技能的培养和锻炼,强调课程的基础地位。首先教材内容的取舍上应精选属于基础的统计知识,这次新编教材又在初版的基础上再作进一步精简压缩。基础知识是最重要的知识,有了牢固的基础,继续深造就不难了。其次,保持教材内容的相对稳定和连续性。应该说基础的、比较成熟的统计方法都是长期知识的积累,是相对稳定的,应该很好地继承。我们认为,教材的创新应该着重于能够应用各项基本的统计方法来研究新情况、新问题,提高同学对统计方法的应用能力。再次,注意与后继课程的衔接。《统计学原理》的后继课程有两类。一类属于经济管理类专业统计课,如企业经营统计、国民经济核算等,另一类属于统计方法的深化提高课,如多元统计分析、各种统计方法专题等。《统计学原理》为这些后继课程打好基础,同时从应用的角度介绍统计理论和方法的基本知识与发展前景,使同学带着问题学习后继各门课,以调动学习的积极性。

《统计学原理》加强统计思想和统计方法论的研究,以拓宽同学的思路,提高学科发展的适应能力。注意精简属于统计实务的内容,例如把统计对象方

法、指标分类等合并为一章,把统计调查、整理、显示等合并为一章,而且不再包括统计组织管理体制法制等方面的内容。当然这并不意味这些内容不重要,而是因为从学科性质来判断,这些内容由统计实务等类课程来承担更恰当。《统计学原理》还注意加强对同学动手操作能力,特别是应用计算机能力的培养,每章习题配备可供上机完成的案例,同时将配套编写《教学指导书》包括 *Excel* 在统计中的应用、习题解答与上机结果。

参加本书编写的有:黄良文(第一章、第四章、第五章),陈珍珍(第二章、第十章),杨灿(第三章、第七章),曾五一(第六章),王美今(第八章、第九章)。由黄良文教授担任主编、曾五一教授担任副主编,负责全书的计划、设计、审订、修改、总纂和定稿工作。

国家统计局有关领导、全国统计教材编审委员会对本书编写给予了组织和领导。中国统计出版社给予了大力支持。东北财经大学邱东教授和上海财经大学徐国祥教授,对本书进行了认真细致的审查,提出十分宝贵的意见,使本书生色不少。我们在此一并表示衷心的感谢。我们深感教材建设任重道远,还需要不懈的努力。由于水平有限,调查研究不够,教材缺点乃至错误在所难免,殷切希望同行专家和广大读者不吝指教。

编 著 者

2000 年 1 月

频数分布的基本概念(36)	变量数列的编制(37)	
累计频数与累计频率(38)	频数分布的类型(41)	
第四节 数据显示		(43)
统计表(43)	统计图(47)	
思考与练习		(50)
第三章 统计分布的数值特征		(52)
第一节 分布的集中趋势——数值平均数		(52)
统计平均数概述(52)	算术平均数(\bar{x})(53)	
调和平均数(H)(56)	几何平均数(G)(60)	
幂平均数和平均数方程(60)		
第二节 分布的集中趋势——位置平均数		(64)
众数(M_o)(65)	中位数(M_e)(66)	
其他分位数(69)	各种平均数的比较(72)	
第三节 分布的离散程度		(75)
变异指标概述(75)	极差与分位差(76)	
平均差与标准差(78)	变异系数(V)(84)	
第四节 分布的偏度与峰度		(86)
分布的矩(动差)(86)	分布的偏度(87)	
分布的峰度(89)	偏度和峰度的计算举例(90)	
思考与练习		(92)
第四章 抽样和抽样分布		(97)
第一节 样本空间、事件及其概率		(97)
样本空间(98)	事件及其概率(99)	
第二节 随机变量的概率分布		(103)
离散型随机变量的概率分布(104)	连续型随机变量的概率分布(106)	
随机变量的数值特征(108)		
第三节 抽样分布		(110)
重置抽样分布(111)	不重置抽样分布(115)	
第四节 正态分布和正态逼近		(120)
正态分布的密度函数(121)	正态分布函数及其标准化(122)	
关于正态分布的定理(126)	抽样分布的正态逼近(127)	

思考与练习.....	(130)
第五章 统计推断.....	(133)
第一节 总体参数估计.....	(133)
总体参数估计概述(133)	总体参数的点估计(135)
总体参数的区间估计(136)	总体平均数的估计(140)
总体成数的估计(142)	总体方差的估计(143)
第二节 总体参数检验.....	(146)
总体参数假设检验概述(146)	
双侧检验与单侧检验(148)	Z 检验与 t 检验(150)
总体平均数的检验(152)	总体成数的检验(155)
总体方差的检验(157)	假设检验的两类错误分析(159)
第三节 非参数检验.....	(161)
非参数检验概述(161)	符号检验(162)
	秩和检验(165)
第四节 抽样设计.....	(168)
抽样设计的基本原则(168)	抽样组织设计(170)
	抽样方案的检查(181)
思考与练习.....	(183)
第六章 相关与回归分析.....	(187)
第一节 相关与回归分析的基本概念.....	(187)
函数关系与相关关系(187)	相关关系的种类(188)
相关分析与回归分析(189)	相关表和相关图(190)
第二节 一元线性回归分析.....	(191)
标准的一元线性回归模型(191)	一元线性回归模型的估计(194)
一元线性回归模型的检验(199)	一元线性回归模型预测(203)
第三节 多元线性回归分析.....	(207)
标准的多元线性回归模型(207)	多元线性回归模型的估计(208)
多元线性回归模型的检验(212)	多元线性回归模型预测(215)
第四节 非线性回归分析.....	(216)
非线性回归分析的意义(216)	非线性函数形式的确定(216)
非线性回归模型的估计(219)	
第五节 相关分析.....	(222)
单相关系数及其检验(222)	等级相关系数及其检验(225)

复相关系数和偏相关系数(227)	相关指数(229)	
思考与练习		(230)
第七章 统计指数		(234)
第一节 统计指数及其种类		(234)
统计指数概述(234)	总指数编制的基本问题(235)	
统计指数的主要种类(238)		
第二节 综合指数及其应用		(240)
综合指数的编制原理(240)	拉氏指数(242)	
帕氏指数(243)	拉氏指数与帕氏指数的比较(244)	
综合指数的其他类型(246)	综合指数的主要应用(248)	
第三节 平均指数及其应用		(251)
平均指数的编制原理(251)	算术平均指数(253)	
调和平均指数(254)	几何平均指数(255)	平均指数的主要应用(256)
第四节 指数体系与因素分析		(259)
指数体系及其作用(259)	总量变动的因素分析(260)	
平均数变动的因素分析(265)		
第五节 指数数列		(270)
指数数列的构成和种类(270)	个体指数数列及其分析性质(270)	
总指数数列及其分析要求(271)	定基指数与环比指数的衔接方式(272)	
思考与练习		(275)
第八章 时间序列分析		(279)
第一节 时间序列编制及分析指标		(279)
时间序列的编制(279)	时间序列的水平指标(282)	
时间序列的速度指标(289)		
第二节 时间序列的分解分析		(296)
时间序列的构成因素和分析模型(296)	长期趋势的测定(300)	
季节变动的测定(306)	循环变动测定(311)	
思考与练习		(316)
第九章 统计预测		(318)
第一节 统计预测的基本问题		(318)

统计预测的概念和分类(318)	统计预测的原则和步骤(319)	
预测结果评价与误差分析(321)		
第二节 趋势外推预测.....		(323)
多项式模型(323)	成长曲线预测模型(336)	
第三节 季节变动预测.....		(344)
分解分析预测法(344)	合成分析预测法(347)	
季节模型的指数平滑法(349)		
思考与练习.....		(354)
第十章 统计综合评价.....		(357)
第一节 综合评价概述		(357)
统计综合评价的概念(357)	统计综合评价分析的一般程序(357)	
综合评价结果的局限性(359)		
第二节 综合评价的常规方法.....		(360)
总分评定法(360)	指数综合法和最优值距离法(362)	功效系数法(365)
第三节 模糊综合评价法.....		(369)
模糊综合评价的基本理论问题(368)	综合评判中权数的确定(372)	
综合评判的数学模型、步骤和应用(375)		
思考与练习.....		(381)
附表.....		(383)
附表 1 二项分布临界值表		(384)
附表 2 正态分布概率表		(385)
附表 3 t 分布临界值表		(387)
附表 4 χ^2 分布临界值表		(388)
附表 5 F 分布临界值表		(389)
附表 6 秩和检验表		(391)
主要参考书目.....		(392)

研究。例如政府要治理国家,作出决策,执行计划,检查监督,宏观调控等都需要以充分、灵通、可靠的国民经济和社会发展的统计资料为基础。所以各国都建立了完备强有力的统计信息网络和健全的统计制度。企业要开发产品,市场营销,生产管理,质量控制,投资评估,资金运用等等更需要统计资料和统计方法的支持。因此,社会上信息咨询公司,统计事务所,统计协会等机构纷纷成立。几乎所有的科学试验都应用统计方法作为有效的手段。它们的作用是:在试验之前依靠统计方法,作出试验的科学安排,这个安排不但影响到以后所搜集资料的代表性,而且直接决定试验的效率。在试验的过程中依靠统计分析显示事物的关系及其规律性。对试验结果作出解释,并检验它的正确性。所以统计方法广泛应用于自然科学、工程技术研究的各个领域。药剂师应用统计方法进行新药疗效的显著性检验,工程师应用统计方法测定新工艺、新材料的创新效果,天文学以统计方法为基础,预测星体的未来位置,生物学应用统计方法安排田间试验,和遗传工程研究等等。虽然所研究的问题属于不同领域,存在千差万别,但所根据的统计学理论和方法则是相同的。

随着社会经济发展,科学技术进步,人类由土地社会进入资本社会,进而迈向信息社会和知识经济社会,统计信息已成为社会、经济、科技信息的主体。统计专业机关,特别是国家统计局,作为智力的机构,其统计的功能大大扩张了,不仅发挥信息功能,而且发挥咨询功能、监督功能和辅助决策功能。统计学的研究的课题也随之扩大,并开辟了广阔的前景。

统计学的研究对象具有以下明显的特点:

(一) 数量性

数量性是统计学对象的基本特点,常言说:“数字是统计的语言”,“数据是统计的原料”,指的正是这个意思。但并不是任何一种数量都可以作为统计对象。统计不同于抽象的数学运算。统计数据总是客观事物量的反映,通过数来测度事物的类型、量的顺序、量的大小、量的关系,以认识客观规律的量的表现。而且统计定量认识必须建立在对客观事物定性认识的基础上,统计研究是密切联系现象的质来研究它的量,并通过量反映现象的质,这一点和数学研究抽象的数量关系是不同的。所以统计数据总是客观的、具体的、也是最有说服力的。早在300多年前英国的威廉·配弟在《政治算术》中首创社会经济现象的数量分析方法,他自己认为他的研究方法不是传统的,而立志要用数字、重量和尺度来说话。这些数据的特点是“从感观的经验中得出的论证”,是“自然界中具有可见的根据”,而坚决排斥那种“以个人的变化无常的意图、见解、爱

好和热情为依据的原因”。马克思给予很高评价，称配弟是统计学的创始人^①。因此统计对象的客观数量性是统计研究最基本的特点。

(二) 总体性

统计学是以客观现象总体的数量方面作为自己的研究对象，这就是说统计的数量研究是对现象总体中各单位普遍存在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。例如，进行城镇居民家计调查，目的不在于了解个别居民家庭的生活状况，而是要反映一个城市、一个市区、一个部门的居民收入水平、收入分配、消费水平、消费结构等等。客观事物的个别现象通常有其特殊性、偶然性，而总体现象则具有相对的普遍性、稳定性、是有规律可寻的，统计研究现象总体的数量特征，有助于我们对现象规律性的认识。

当然，统计研究是要从个别入手的，但对个别单位的具体事实的调查观察只是为了达到研究现象总体特征的目的。统计研究对象的总体性，也不排斥对个别典型单位的深入研究，但它也是为了更有效地掌握总体现象的规律性。

(三) 变异性

统计研究同类现象总体的数量特征，它的前提则是总体各单位的特征表现存在着差异，而且这些差异并不是由某种特定的原因事先给定的。例如一个企业职工的工龄长短有差异，文化水平高低有差异，工资报酬多少有差异等等。这才需要研究职工的平均工龄、文化结构、平均工资等等指标。如果各单位不存在这些差异，也就不需要做统计，如果各单位之间的差异是按已知条件事先可以推定的，也就不需要用统计方法。例如昼夜时间长短因季节变化而不同，这与统计无关，而江河水位高低随时间而不同则是统计研究的对象。统计上把总体各单位由于随机因素引起的某一标志表现的差异称为变异。

如果说，总体各单位的变异表现出个别现象的特殊性和偶然性，而对现象总体的数量研究，则是从各单位的变异中归纳概括出它们的共同特征，显示出现象的普遍性和必然性。这就是统计认识方法的特点。

二、量的尺度

统计离不开计算，这里存在着计算的对象、计算的手段、计算的结果之间的关系，即量、数、值的关系。事物的量是事物属性的重要方面，事物的量和事物的质是密切联系、相互制约、并且共同决定事物本体的性质和特征，所以它是客观存在的。数则是根据事物的性质和研究任务，以一定的计算规范确定某

^① 参见《马克思恩格斯全集》第13卷，第43页。

种数的形式,如自然数、相对数、平均数等等来反映客观事物量的特征、量的界限、量的关系等等。当给定某些数的具体数值,称为数的取值,如果这些数值是取自试验或调查的结果,我们就称这种数值为数据,当然,由数据再加工的数据值仍为数据。举例说,我们要研究城市居民的生活状况,反映城市居民生活状况有许多量的方面,我们可以选择城市常住人口数,常住人口的年龄构成,居民平均年收入,居民消费构成等数字来反映居民生活的基本情况。通过调查可以确定1999年底某市常住人口为579.8万人,其中劳动年龄组(男16~59岁,女16~54岁)占总人数的58.5%,全年居民人均可支配收入为6925元,在居民人均消费性支出中,食品衣着类占总支出的59.3%,这些数据反映该市居民生活的基本情况。

数作为量的尺度,可以分四个不同层次。

(一) 定类尺度

将数字作为现象总体中不同类别或不同组别的代码,这是最低层次的尺度。在这种情况下,不同的数字仅表示不同类(组)别的品质差别,而不表示它们之间量的顺序或量的大小。这种尺度的主要数学特征是“=”或“ \neq ”。例如将国民经济按其经济类型,可以分为国有经济、集体经济、私营经济、个体经济等类,并用(01)代码表示国有经济,(02)表示集体经济,(03)表示私营经济,(04)表示个体经济。并且用(011)代表国有经济中的国有企业,(012)代表国有联营企业;用(021)表示集体经济中集体企业,(022)表示集体联营企业;用(031)表示私营经济中的私营独资企业,(032)表示私人合伙企业,(033)表示私营有限责任公司;用(041)表示个体经济中的个体工商户,(042)表示个人合伙等等。其中两位代码表示经济大类,而三位代码则表示各类中的构成。不同代码反映同一水平的各类(组)别,并不反映其大小顺序。各类中虽然可以计算它的单位数,但不能反映第一类的一个单位可以相当于第二类的几个单位等等。

(二) 定序尺度

定序尺度不但可以用数表示量的不同类(组)别,而且也反映量的大小顺序关系,从而可以列出各单位、各类(组)的次序。这种尺度的主要数学特征是“ $>$ ”或“ $<$ ”。例如对合格的产品按其性能和好坏,分成一等品、二等品、三等品等等。这种尺度虽然也不能表明一个单位一等品等于几个单位二等品,但却明确表示一等品性能高于二等品,而二等品性能又高于三等品等等。定序尺度除了用于分类(组)外,在变量数列分析中还可以确定中位数、四分位数、众数等指标的位置。

(三) 定距尺度

定距尺度不但可以用数表示现象类(组) 别的不同和顺序大小的差异,而且可以确切的数值反映现象之间在量方面的差距。定距尺度的主要数学特征是“+”或“-”。定距尺度的精确性比定类尺度和定序尺度前进了一大步,它又是定比尺度的基础,所以在统计数据中定距尺度居于重要地位。凡是反映现象的总量规模指标都必须运用定距尺度。例如工厂生产的产品数量、职工人数、国民生产总值、城市居民常住人口数、城市流动人口数、科学试验数据、气候温度、湿度等等都用定距尺度计量。

(四) 定比尺度

定比尺度是在定距尺度的基础上,确定可以作为比较的基数,将两种相关的数加以对比,而形成新的相对数,用以反映现象的构成、比重、速度、密度等数量关系。由于它是在比较基数上形成的尺度,所以能够显示更加深刻的意义。定比尺度的主要数学特征是“÷”或“×”。例如将某地区人口数和土地面积对比计算人口密度指标,说明人口相对的密集程度。甲地区人口可能比乙地区多,但甲地区的土地更广阔,用人口密度指标就可以说明相对说来甲地区人口不是多了,而是少了。又如将一个国家(地区) 的国内生产总值与该国(地区) 居民对比。计算人均国内生产总值,可以反映国家(地区) 的综合经济能力。1998年我国国内生产总值约占世界生产总值的12%,排列世界第七位,堪称世界经济大国,但我国人口占世界总人口的21.2%,如果按人均国内生产总值计算,在世界各国中又居于比较落后的位次,说明我国仍属于发展中国家。

三、统计学的性质

统计学有自己的研究对象,也有自己的研究方法。统计学的任务就在于为统计活动提供数量研究和认识规律的科学方法,包括指导统计工作的原理原则,统计过程所应用的核算和分析方法,以及组织方法。其核心内容是数据的搜集、整理、归纳、分析的原理和方法。这些方法论构成了统计学的科学体系,所以统计学是一门认识方法论科学,它是研究如何搜集数据、分析数据、以便从中作出正确推断的认识方法论科学。

统计学和数学都是研究数量关系的,但两个学科有不同的性质特点。数学撇开具体的对象,以最一般的形式研究数量的联系和空间形式,数学的分析方法主要是逻辑推理和演绎论证的方法,从严格的定义、假设的命题和给定的条件去推证有关的结论。而统计学的数据则总是与客观的对象联系在一起,统计的过程就是从客观对象中抽出其数量表现,得到有关的数据,然后加以适当的