

1000

统计学理论与应用

肖昭媛 主编



A1012840

上海交通大学出版社

内 容 提 要

本书系统地阐述了统计学理论、方法和应用。全书共十三章,内容包括:统计指标分析、动态分析、指数分析、事件与概率、随机变量与概率分布、常用随机变量的概率分布、大数定律与抽样分布、参数估计、假设检验、方差分析、回归与相关分析、统计调查等。最后以若干统计分析实证,介绍了统计理论的实际应用。

本书可供高等学校经济、工商管理、财会和审计等专业大学本科及工商 MBA 的教材,同时也可作为企业管理、市场分析和商务信息咨询等领域有关人员学习参考。

图书在版编目(CIP)数据

统计学理论与应用 / 肖昭媛主编. - 上海:上海交通大学出版社, 2001

ISBN 7-313-02761-3

I. 统... II. 肖... III. 统计学-高等学校-教材
IV. C8

中国版本图书馆 CIP 数据核字(2001)第 052030 号

统计学理论与应用

肖昭媛 主编

上海交通大学出版社出版发行

(上海市番禺路 877 号 邮政编码:200030)

电话:64071208 出版人:张天蔚

太仓市印刷厂有限公司印刷 全国新华书店经销

开本:787mm×1092mm 1/16 印张:20.75 字数 509 千字

2001 年 8 月第 1 版 2001 年 8 月第 1 次印刷

印数:1-3050

ISBN 7-313-02761-3/C·055 定价:27.00 元

版权所有 侵权必究

前 言

随着社会主义市场经济的发展,统计方法在经济、管理等工作中的作用和重要性,愈来愈为广大从业人员所认识,人们迫切希望学习和掌握统计理论和统计方法,以应用于他们的实际工作。另外,我们从对经济、管理类等专业教学工作中发现,年轻的大学生们普遍认为,“描述统计”简单易学,但用于分析实际问题时感到困难;而对“推断统计”,则感到抽象难理解。针对上述问题,本书在编写过程中,始终坚持“理论联系实际”的原则。因此本书有以下特点:

1) 全书对有着广泛应用的概念和方法,均作严谨、细致的叙述(不作或很少作详细的数学论证),并配以大量的经济、管理中的实例,帮助读者理解基本理论和方法的应用。为方便教学的需要,每章后面配有练习。

2) 本书用大量的实例和确切的文字,阐述描述统计中每一种分析方法的用途和实际应用。在阐述推断统计的同时,都详细说明与描述统计中有关概念和方法的联系,使推断统计不再“抽象难理解”。

3) 本书用一章来专门介绍一些实际问题中统计分析的应用,可以帮助初学者学习和提高对实际问题进行统计分析的能力。

4) 本书在章节内容和编排顺序等方面对传统做法作了修正,使本书内容结构更合理、前后衔接更加自然和紧凑,避免了局部的重复和交叉。

全书共十三章。其中,第一、三章由余思勤教授编写,第二章主要由余思勤教授编写,第十一章由李序颖副教授编写,第十二章由余思勤、李序颖和肖昭媛合编,剩余其他各章皆由肖昭媛编写。全书由肖昭媛教授负责统稿。

目 录

绪 论	1
第一章 统计指标分析	7
§ 1.1 统计指标的意义	7
§ 1.2 统计分组和分组指标	9
§ 1.3 总量指标和相对指标	17
§ 1.4 平均指标	19
§ 1.5 变异指标	27
§ 1.6 统计表和统计图	31
§ 1.7 国民经济主要总量指标	35
练习题一	38
第二章 动态分析	44
§ 2.1 时间序列	44
§ 2.2 动态分析指标	46
§ 2.3 时间序列的四个分量	51
§ 2.4 趋势分析	53
§ 2.5 季节变动分析	58
练习题二	64
第三章 指数分析	69
§ 3.1 指数的意义	69
§ 3.2 综合指数	70
§ 3.3 平均数指数	73
§ 3.4 指数体系与因素分析	75
练习题三	79
第四章 事件与概率	85
§ 4.1 随机现象、随机试验和事件	85
§ 4.2 事件的关系	87
§ 4.3 概率	89
§ 4.4 条件概率、乘法公式和事件的独立性	94
§ 4.5 全概率公式和贝叶斯公式	98
练习题四	100
第五章 随机变量与概率分布	104
§ 5.1 随机变量	104
§ 5.2 随机变量的概率分布	105
§ 5.3 二维随机变量的概率分布	109

§ 5.4 随机变量的数字特征	114
§ 5.5 协方差和相关系数	120
练习题五	123
第六章 常用随机变量的概率分布	128
§ 6.1 离散型随机变量的概率分布	128
§ 6.2 连续型随机变量的概率分布	133
§ 6.3 中心极限定理	139
练习题六	142
第七章 大数定律与抽样分布	145
§ 7.1 基本概念	145
§ 7.2 大数定律	147
§ 7.3 均值 $\bar{X}$ 的抽样分布	149
§ 7.4 方差 S^2 的抽样分布	154
§ 7.5 比率 π 的抽样分布	155
练习题七	158
第八章 参数估计	162
§ 8.1 参数的点估计	162
§ 8.2 参数的区间估计	165
§ 8.3 总体均值 μ 的估计	169
§ 8.4 总体比率 p 的估计	173
§ 8.5 总体方差 σ^2 的估计	175
§ 8.6 两个总体差异的区间估计	180
练习题八	184
第九章 假设检验	187
§ 9.1 概念、原理和步骤	187
§ 9.2 两类错误与假设检验方法	190
§ 9.3 正态总体均值 μ 的检验	193
§ 9.4 总体比率 p 的检验	198
§ 9.5 总体方差 σ^2 的检验	200
§ 9.6 两个总体差异的检验	203
§ 9.7 II 型错误概率 β 的讨论	206
练习题九	209
第十章 方差分析	211
§ 10.1 概述	211
§ 10.2 单因素方差分析	212
§ 10.3 有交互作用的双因素方差分析	217
§ 10.4 无交互作用的双因素方差分析	222
练习题十	224
第十一章 回归与相关分析	227

§ 11.1	回归分析和相关分析的基本概念	227
§ 11.2	一元线性回归与相关分析	228
§ 11.3	多元线性回归与相关分析	235
§ 11.4	非线性回归分析简介	241
练习题十一		242
第十二章 统计调查		245
§ 12.1	统计调查的意义	245
§ 12.2	统计调查方法	247
§ 12.3	简单随机抽样	251
§ 12.4	等距抽样	258
§ 12.5	整群抽样	259
§ 12.6	分层抽样	262
§ 12.7	多阶段抽样	266
§ 12.8	不等概抽样	270
练习题十二		272
第十三章 应用与分析		276
§ 13.1	统计指标分析方法应用——中国与西方七国综合国力比较	276
§ 13.2	指数的编制——中国出口集装箱海运价格指数简介	282
§ 13.3	概率计算应用	286
§ 13.4	对我国水路运输量的回归分析	289
§ 13.5	全国港口普查方案	293
§ 13.6	水路运输量抽样调查方法的选择	297
附表 1	标准正态分布的累积概率表	302
附表 2	χ^2 分布累积概率表	304
附表 3	t 分布表	306
附表 4	F 分布表	307
附表 5	二项分布表	313
附表 6	泊松分布表	317
附表 7	随机数字表	319
参考文献		321

绪 论

一、统计学和它的研究对象

1. 什么是统计

从词意上解释,统计就是对大量事物进行数量上的综合、汇总和分析。而实际上,关于统计,人们有两种不同的理解:

(1) 统计工作,即搜集、整理、综合、汇总、分析和提供各种统计资料的工作过程。

(2) 统计学,即研究搜集、整理、分析和提供统计资料的理论和方法。

这两者之间存在着密切的关系:统计工作是统计学的实践过程,统计学是统计工作的经验总结和理论概括,它指导统计工作。统计学向一切自然科学和社会经济科学提供了一种普遍适用的研究方法和理论。

2. 统计学的研究对象

统计学研究对象的数量特征。应该进一步说明,统计学研究客观现象的数量,并不是任意的数字组合,而是在一定质的规定条件下的客观现象,在一定时间、空间中具体的数量表现。统计工作就是从研究现象的质与量的联系中,分析和揭示现象的本质。

统计学研究的对象,不是个别事物,而是由大量事物组成的现象总体(或全体)的数量特征。例如,某航运公司有 100 艘运输船,统计工作研究着眼于全公司 100 艘运输船舶完成的运输量,消耗的燃油量等等,而不对个别某一艘运输船舶的情况进行研究。因为个别运输船运输量不能说明全公司的运输量,即个体特征不能代表总体特征。统计学研究现象全体单位的数量特征,对认识现象的本质和指导工作具有重要意义。

一般情况下,客观现象中包含的各个单位,所处的环境有所不同,在一定的时间内,还可能会受到各种不同偶然因素的作用,因此,各个体的数值表现普遍存在差异,通常,这种差异呈现出随机性质。

二、统计学的研究内容

统计学为统计工作提供搜集、处理和分析大批数据的方法和理论。随着人类对自然现象的认识由浅到深,人类社会经济规模由小到大、由单一到复杂,统计学的研究内容,即统计工作方法和理论,也经历了由简单计数、文字记述事实、简单加总、对比到利用复杂数学模型预测未来的发展过程。迄至今日,概括起来,统计学的研究内容主要包括统计调查和统计分析两部分。

统计学研究现象总体的数量特征,首先从资料收集开始。统计资料主要通过统计调查取得。由于总体规模和特点、调查目的、对精度的要求以及其他诸多客观条件的差异,使统计调查方法不可能存在简单划一的模式,这就要求统计学研究适合各种研究目的、不同类型总体和条件的统计调查方法和对总体进行估计的理论与方法。

统计调查工作结束以后,统计工作进入了统计分析阶段。对统计资料进行整理、计算和分析,揭示现象的本质与变化规律的工作过程称为统计分析。根据分析方法的特点,统计分析研

究内容包括：

1) 描述统计。主要研究现象质的规定与其数量表现的联系方法。具体地说来，就是研究用什么方法，怎样来描述客观现象的数量特征。统计学对现象的描述方法或描述现象的主要工具是“统计指标”。统计指标是描述统计的中心内容。它包括许多理论与方法论的问题，如统计指标概念、计算方法、指标体系的设置等。

2) 推断统计。是研究如何根据部分数据推断出总体指标的一套方法，用统计的语言来说，就是研究如何根据随机样本提供的信息，推断出样本所来自的总体及其相应情况的一套统计分析方法。推断统计研究的内容主要包括：

(1) 根据随机样本资料对总体进行估计、检验、相关与回归分析。大多数此类实际问题是根据空间中某一时间截面上部分资料对同一时间截面总体情况进行推断。

(2) 根据所研究现象的历史资料推断它未来发展变化的趋势，即预测问题。统计预测的基本方法是相关与回归分析及其它时间序列模型等。

如果说，第(1)个是研究某一时间截面上的、静态总体的统计推断问题的话，那么，第(2)个预测问题，即是研究随着时间流动的、动态总体的统计推断。

随着社会的发展，推断统计越来越重要，内容上越来越占有较大的比重，成为统计学的核心部分。

统计学各部分的研究内容相互之间联系紧密，相互补充。例如，统计资料是统计分析的基础，没有科学的统计调查方法，就不可能得到优质的统计资料，在劣质的统计资料基础上进行的统计分析结果，不可能作出对现象的正确结论。反过来统计分析的目的和要求，限制和规定了统计调查的方法和内容。

统计描述为推断统计提供了研究的工具和手段，因此，是统计推断的基础。例如，统计描述样本的各种数字特征，是推断总体特征的依据。在统计描述力所不能及的领域，推断统计方法不仅是对描述统计方法的补充，而且突破了“统计是已发生了的事实的记录”的传统观念，从而大大扩展了统计学的研究领域。

三、统计学的研究方法和理论基础

由上述内容可知，统计学是一门方法论科学。然而，统计学之所以是一门科学，在于统计学研究各种方法的同时，还研究这些方法的理论依据和应用条件。

统计学研究现象总体的数量特征借助于大量观察法。现象总体一般包含着大量个体，现象总体特性是其全部个体特性综合的结果。每个个体处在不同的环境中，各受很多共同的和不同的因素影响。结果在数量上，各个体之间存在着大小不同的差异，只有对总体中大量个体进行观察、综合、分析，才能消除各个体间由于偶然的、不同的因素影响所造成的差异，从而得出总体的一般性质。

统计分析所依据的基础资料来自统计调查，因此，统计调查的方法和理论，是统计学的重要组成部分。统计报表制度和抽样调查是我国目前普遍采用的两个主要统计调查方法。统计报表制度是政府规定调查内容，由指定的基层单位按照一定的程序报送统计数据的方法。抽样调查是用随机取样的方法，从客观现象总体中抽取一部分单位进行调查，然后依据概率论和数理统计理论，对现象总体进行推断。显而易见，抽样调查既是统计调查方法又是统计分析方法。

描述统计方法是用确定的数值描述所研究的现象“是什么”，即采用对统计资料进行简单

加总、对比等算术运算方法,得到统计指标,用统计指标来描述接受调查的群体的、数量特征。这种确定性的研究方法是一种对现实已经发生了的社会经济现象的“映象”的简单加工。

显然,描述统计的理论基础,就是关于统计指标的概念(内涵及外延)的规定、计算、应用和分析的理论。统计描述简单现象采用一个或几个指标,应用统计指数(统计指标的一种类型)理论描述分析复杂现象,对于由相互依存又相互制约的若干现象组成的系统,统计描述分析依据指标体系或指数体系理论。

推断统计方法是把所研究现象在各个体上的具体表现看做随机变化的量,以概率论和数理统计学作为理论依据,利用对现象总体部分统计调查资料,推断和回答现象总体“可能是什么”。

由于现象总体中,各个体间在数量上存在着差异,因此,用对个别或部分单位的观察结果不能确定地描述现象总体。当客观上仅具备现象总体部分观察资料,而又必须对现象全体进行推断时,统计学研究不得不利用主要是描述和处理随机变量性质和变动规律的“概率论与数理统计学”。根据概率论与数理统计基本理论,统计学研究中假设,所研究现象总体,可以在相同条件下,重复无数次观察,观察结果是确定的;个体(或观察结果)之间存在数量上的差异的原因,主要是偶然因素作用的结果,其值可正可负,大量观察、综合加总后呈现一定的规律性。推断统计方法就是依据对现象总体进行的大量观察,给出有关现象总体性质的结论,因此,这个结论总是带有大量观察下的平均意义。

四、基本概念

下面是统计学中常用的一些基本概念。

1. 总体、个体

总体是由具有相同性质的个体组成的集合体。

总体是客观存在的,即构成总体的每个个体都是独立存在的实体。例如,某航运公司有两艘总吨位分别为 1000 吨的运输船。则该公司船舶总体由两艘船即两个个体组成,不能假设认为该公司船舶总体是两千艘总吨位是 1 吨的小船,因为这些总吨各为 1 吨的船舶都是事实上不存在的假想体。

总体中每个个体至少具有一个共同的性质,即总体的同质性。这是统计研究的基础。共性由所研究现象的性质和统计研究目的决定。例如,研究某港 1994 年实现的装卸产值。则统计研究的总体中各个体至少在以下方面是相同的:它们都是该港的一个下属单位;1994 年建制继续存在;它们都有共同的生产经营成果装卸产值。除了共同的性质以外,其他方面则不同:各个从事装卸业务活动的单位可能有不同的职工人数,不同的设备,不同的用工、分配方法等。因此,“共性”在各个体的数量表现,一般不会完全相同。例如,各单位实现的装卸产值大小,通常是不一样的。总体内部各个体间这种数量上的差异性,正是需要进行统计研究的原因。

由于统计研究目的不同,总体和个体具有相对性。例如,某港有 12 个装卸作业区,研究该港生产量时,该港 12 个装卸区组成了统计研究的总体。若研究全国港口生产量时,该港即是组成“全国港口”这个总体的一个个体。

在统计实践中“个体”又称为总体单位。统计学中所研究的总体总是包含着大量的个体。总体中包含个体的数量,称为总体容量,它是衡量总体大小的尺度,一般用符号 N 表示。统计研究的总体容量由统计范围和统计分析目的所决定,其值通常很大。包含无限多个个体的总体称

为无限总体,个体数量有限的总体为有限总体。例如,由流水生产线生产出来的产品所组成的总体是无限总体,某公司某一时期产品是一有限总体。

在统计工作的不同阶段,总体有两个不同的含义。统计调查阶段,人们最关心的是实在的总体,它是统计调查工作的对象,或者说,构成实在总体的每个个体都是统计调查的最小单元。在统计分析阶段,统计研究总体是上述实在总体的数量表现,称之为数值总体。统计调查的目的是为了取得数值总体,或其一部分,统计分析的目的是为了揭露实在总体的本质。

统计调查阶段,应该十分注意区分作业总体和目标总体。例如,欲调查运输船舶某时期完成的运输量。这些运输船一般属于不同的运输企业或个人。目标总体是统计研究的对象,由全部运输船舶组成。作业总体是承担填报统计数据的单位或个人组成的集合体,上例即运输船舶所属的运输企业或个人。显然,作业总体不总是等于目标总体。

统计实践中,总体中各单位,常常需要按某一种性质分成若干群体或组,组内各单位关于分组性质具有相近的特征。

2. 样本和样本单位

从总体中抽出来的一部分个体组成了一个样本,这些个体称为样本单位。样本的大小称为样本容量,一般用符号 n 表示。样本容量是样本中包含的个体的数量。样本是总体的一部分,具有总体的一部分特征,或者说,样本带有总体一部分信息;随着 n 的增大,样本带有愈来愈多的关于总体的信息,样本与总体愈来愈相似。当 n 一定时,从总体中可以抽出很多的样本,它们各自的特征可能相同,也可能不同。

样本的基本功能是为了被调查。样本单位是进行统计调查的最小、最基本单元。一般,对样本调查的目的,是为了对它所来自的总体的特征进行估计。例如,某市从全体居民家庭中抽出 100 户家庭,调查他们的月收入,用来估计全市居民家庭户平均月收入。这 100 户居民家庭是统计调查的一个样本,样本容量 $n=100$ 。

在统计理论和实践中,有两类不同的样本,一类称为随机样本,其中每个样本单位都是按随机原则从总体中抽出来的;另一类是非随机样本,它的每一个单位都是按确定的规律选出来的。实践中,从总体中随机抽取样本单位时,有两种操作方法:一是放回抽样,也叫重复抽样,指从同一总体抽取样本单位的过程中,任何一个单位被抽出后,仍放回总体内,参加下一次抽取。这样,在整个抽样过程中,每一次抽取,总体单位数始终保持不变,因此,每个单位被抽中的机会完全相等。二是不放回抽样,也叫不重复抽样,指从同一总体抽取样本单位过程中,任何一个单位被抽出后,就不再放回总体内,即不再参加下次抽取,这样总体单位数在逐渐减少,样本中各单位被抽中的机会不相等。

3. 指标、标志

指标是说明现象总体或统计数据群体的数量和性质的,标志则说明总体单位的性质。例如,某厂 800 名职工完成产量 1200 万吨,用等式表示这种说法:

职工总人数=800 人;

总产量=1200 万吨。

等号左边是反映职工总规模和生产总水平的科学概念和经济学定义,称为指标名称;等号右边是相应指标名称在该厂(总体)的具体的数量表现,称为指标值。指标名称和指标值一起说明了研究对象(该厂)在职工总规模、生产总水平等两个侧面的数量特性。所有的统计指标都有指标名称和以数值表示的指标值。如果统计研究一个复杂现象;例如职工状况。对于一定

空间的职工总体,可以用很多指标从不同的角度、侧面反映它的数量特征:例如:职工男女比例、技工人数、职工平均工资等。

说明总体单位特征的称为标志。对于统计研究的职工状况问题,年龄、性别、文化程度、工种、收入等等都能反映某位职工状况的某个方面的特性,因此都是标志,假如下面是一份该厂职工状况调查表:

姓 名	性别	年龄	工种	文化程度	工龄	收入	
张 三	男	26	司机	高中	5 年	250	
李 四	女	32	炊事员	初中	15 年	260	
⋮	⋮	⋮	⋮	⋮	⋮	⋮	

表中第一行各项反映职工总体中各个体的共同属性,称为标志名称,下面各行每一项都是标志名称对于不同职工的具体表现,称为标志值。标志值反映了个体的特性。与指标值都是数值的特性不同,标志值可以用数量表示,如年龄、收入,也可以是用文字说明的,如性别、工种等。

标志值可以用数值表示的称为数量标志,只能用文字表示其值的标志称为品质标志。各个体标志值不同,说明总体内部存在差异。指标和标志的关系很密切:所有的指标值都是由对标志值汇总得到的,例如,职工总人数是由对每一位职工计数得到的结果;对不同的研究目的和统计范围,指标和标志可以相互转换,例如,某港完成的吞吐量是该港生产量的一个统计指标,但相对全国港口完成的吞吐量指标而言,它又是一个标志值。

在推断统计理论中,反映总体特性的指标称为参数。

4. 变量、变量值

统计分析中,标志名称被抽象定义为变量,标志值即为变量值。总体中每一个个体的标志值,都是变量的所有可能取值,该变量的值域因此而被确定。

因为变量名即标志名称,反映了总体中所有个体的共性,即总体的属性,因此,变量名是总体的代表。变量所有可能取值,全面、详尽地记录了总体中各个单位的数量特性,表现了总体的内部变异性。

为了便于数学处理,一般用符号 $X, Y, \dots$ 表示变量或总体,用数学语言表达即变量 X , 变量 $Y, \dots$ 或总体 X , 总体 $Y, \dots$ 。显然,统计指标值,是对该变量所有可能取值某个侧面的概括。例如,研究某公司职工收入。用变量 X 代表该公司每位职工的收入,等同于说 X 代表该公司职工收入总体。第 i 位职工收入为变量 X 的第 i 个(或次)取值,用 x_i 表示;平均指标 $\bar{X}$ (职工平均收入),概括描述了该公司全体职工收入的平均水平。

按取值是否连续,变量分为连续变量和离散变量。在一定范围内,若变量可以取任意实数值,则这样的变量称为连续变量。若只能取有限个或可数多个值,则称为离散变量,大多数离散变量值为整数。一般品质标志总体可以处理为离散变量总体,例如,性别总体,可以用 1 代表男,0 代表女,值域为 0 和 1 两个数值,因此是离散变量总体。这是一个典型的离散变量。

按照对变量值能否确切地预测,变量分为确定性变量和随机变量。如果在对现象总体进行

观测以前,能够通过各种方法了解到它的确切值,此即确定性变量。通常,确定性变量的影响因素及其影响结果是确定的,因此可以通过对影响因素的观测,计算而得出变量的确定值。随机变量是这样一类变量,即它可能取的值无法确切的预言。造成变量值随机性的因素众多,且分不清主次。

第一章 统计指标分析

统计学的研究对象是现象总体的数量特征。总体的数量特征主要有：总体规模、总体水平、总体内部变异、总体中各部分间的比例、总体变化速度、总体之间的比较等等，在经济统计学中，把描述总体这些特征的概念和数值称为统计指标，或简称为指标。统计指标是统计分析的工具，是描述统计研究的中心内容，因此，要学会运用统计学方法认识现象的数量特征，首先必须掌握各类统计指标的设计和计算方法。

§ 1.1 统计指标的意义

一、概念

统计指标就是根据一定的统计方法对总体^① 各单位的标志表现进行登记、核算、汇总、综合而形成的能说明总体某一特征的概念和数值。统计指标主要包括指标概念和指标数值两方面内容。例如，我国 1993 年的人口数为 118 517 万人，国内生产总值为 31 380 亿元，货物周转量为 30 405 亿吨公里等，它们反映了一定时间和空间条件下总体现象的某一综合数量特征。通常要通过对原始资料的搜集、整理和分析计算，才能得出有关统计指标的具体数值，所以，统计指标也体现了统计工作过程。每个具体的统计指标都有其特定的作用，但总的说来，通过统计指标可以反映总体现象的规模、水平、比例和速度等，研究总体现象发展的规律的数量表现，因此，统计指标成为认识社会、管理经济、科学研究的基本依据，是体现统计咨询、服务、监督三大功能的工具和手段。

社会经济现象错综复杂，各种现象之间又存在着相互联系、相互制约的关系。每一个统计指标只能反映所研究现象的某一个侧面，若要反映现象的全部过程和现象的各个方面，就需要使用一系列的统计指标。统计学中把这一系列相互联系和相互制约的统计指标，叫做统计指标体系。统计工作只有通过统计指标体系，才有可能将所要研究的社会经济现象，全面而系统地反映出来，才有可能据此认识社会。所以统计指标体系可以比统计指标更广泛、更全面地分析现象总体存在的矛盾，更加深刻地认识现象总体的全貌和发展的全过程。

由于实际的总体现象常常很复杂，总体内部各部分之间存在不同程度的差异，为了更深刻地认识总体现象的本质，在对现象总体的认识过程中，仅有总体指标还不够，很多时候，还必须同时将总体加以分组，进一步制定分组指标体系。例如，在一个贸易公司，商品的购进、销售、库存等三个方面相互联系，相互制约。所以这三个指标就构成了商品流通的统计指标体系。缺少任何一个指标，都无法反映商品流通过程的全面情况。若要更深入地反映商品流转情况，还应分成不同的商品，分别研究其购、销、库存等分组指标。

^① 统计指标是反映统计数据所来自的群体特征的概念和数值。这个群体既可以是总体，也可能是样本，为了叙述的简单，本书中只讲是反映“总体”特征的，这种说法仅是为了使叙述简单化。

二、统计指标设计的一般原则

把统计调查资料综合起来,使它表现为反映现象总体各种特征的量,是统计工作必须解决的首要问题。要完成综合的任务,只有依靠科学的统计指标和严密的指标体系。当一项具体的统计分析任务确定之后,第一步工作就是根据研究对象的特点和研究任务的要求,制定和设计一系列用来反映各个有关现象总体数量方面的、相互联系、相互制约的、科学的统计指标(名称及计算方法),如果不进行这项工作,统计调查和统计分析工作,就无法进行下去。统计指标与指标体系设计时,应注意以下几个原则:

1. 需要与可能相结合

所谓需要,就是国家或企业为进行经济、管理、制定政策或决策等提出的统计分析任务。所谓可能,即条件,包括经济管理水平、统计部门或研究单位可以投入的人力、物力、财力和时间等。

设计统计指标与指标体系时,应该首先考虑需要,但也必须考虑到可能。凡属既有需要,又有可能的统计指标,则一定要设立;凡属经过努力,还不具备条件的,就不要勉强建立;如不顾客观条件,片面强调需要,勉强制定的统计指标,到头来,可能完全得不到该指标数值,或者是误差很大的指标值。

2. 要有科学依据,并符合实际

人们是通过统计指标体系来认识社会、经济现象的。因此,在统计指标体系中,各统计指标的名称和统计范围,一定要能准确地表达所要研究的社会经济现象的本质及彼此间的内在联系。只有依据科学的政治经济学有关原理,制定统计指标体系,才能准确地描述它们的本质及其内在联系。科学的政治经济学所说明的经济范畴和原理都是抽象的,而现实的社会经济现象非常复杂,要完全按照这些抽象的理论和原理,规定统计的口径和范围,往往无法取得可靠的统计数字,因此,必须结合研究对象的特点,规定切实可行的统计范围、计算口径和计算方法。

总之,设计统计指标体系的时候,一定要使理论和实际紧密联系,不得片面强调任何一方面。要尽可能地广泛征求用户和基层工作人员的意见,并深入实际调查研究和试点,以使指标解释、计算方法、统计范围等切合实际,方便易行。

3. 全局性和稳定性

企事业单位的统计资料既是政府统计的基础,又是对本单位工作情况进行统计分析的依据,统计工作要讲效益,必须争取国家和本单位统计资料共享。因此,为本单位进行统计分析而设计统计指标体系时,首先,应尽量和国家现有统计指标体系的指标名称、指标概念、统计范围、口径、计算方法和计算价格等保持一致,这样,所设计的指标体系,既能够同时满足本单位和国家的需要,又便于不同单位间进行对比和分析。其次,为了不同时间进行比较,指标体系不应经常变动。包括指标项目、指标概念、统计范围、口径、计算方法等,应尽量保持稳定不变。

总之,由于“比较”是统计分析最基本的方法,所以,在任何时候,设计统计指标与指标体系时,都应该严格遵守“可比性”原则。

三、统计指标的分类

统计指标按其反映的总体内容不同,可以分为数量指标和质量指标。

数量指标是说明总体现象规模、水平或总量的统计指标。例如:国民生产总值,港口吞吐

量、职工人数等。数量指标所反映的是总体的绝对数量,它的数值随总体范围大小的变化而增减。

质量指标是说明总体现象相对水平或工作质量的统计指标。例如:平均工资、人口密度、出勤率,单位产品原材料消耗等。质量指标不随总体范围大小的变化而增减,常用来反映总体现象之间的联系和对比关系。

统计指标按其作用和表现形式不同,可以分为总量指标、相对指标、平均指标和变异指标。

总量指标是反映总体现象规模的统计指标,常用来说明总体现象的广度,表明总体现象发展的结果。

相对指标是两个有联系的指标相比较而得出的统计指标。它常用来说明总体的构成,总体现象的发展变化等。

平均指标是反映总体单位某一数量标志一般水平的统计指标。它常用来说明总体现象的一般水平。

变异指标是反映总体内部和单位标志值的差异程度。它常用来说明现象总体的均衡性和发展变化的稳定性。

在以上各种指标中,总量指标是基础指标,其他指标都是由总量指标计算加工后而得出的派生指标。

§ 1.2 统计分组和分组指标

一、概述

1. 意义和作用

统计分组和分组指标是一种重要的统计分析方法。统计分组是根据统计研究的目的和任务,按照某一标志将总体划分成若干类型或组别,使组与组之间具有差别性,而同一组内的单位保持相对的同质性。统计分组的目的是要将同质总体中有差别的单位区分开来,同时又将性质相同的某些单位组合在一起,以便通过计算相应的分组指标,对总体中所有单位在质量上、数量上、空间上存在的差异进行研究和分析,进一步认识现象的本质特征及其发展的规律性。可见,统计分组和分组指标不仅只是将统计资料进行简单的组合、整理和分组汇总,也是使认识深化的重要手段。

统计分组和分组指标的作用主要有以下三个方面:

1) 划分事物类型,研究不同类型事物的特点。在统计总体中,各个单位总是客观地存在某种差异,这是统计研究的基本条件。但是,统计研究这种差异,不是专门针对个别单位的具体差异程度和变化进行研究,而是对这种差异加以抽象和概括。统计分组将总体按一定标志划分为不同类型的组,这就确定了总体各部分的范围和界限,通过计算各分组指标,就可以对总体有一个更深刻的认识。如,1993年我国各种运输工具完成的货物周转量分组资料如表1.1所列。

由上述分组资料可以看出,铁路和水路这两种运输工具完成了绝大部分的货物周转量,而水路中,又是远洋运输占了很大比重。民航运输尽管所占比重很小,但它比上年有了较大增长,管道运输比去年减少了1.8%。

2) 研究总体结构。通过统计分组,可以计算各组数值在总体中所占的比重或各组之间的

比例关系等分组指标,从而反映总体的结构情况,说明现象总体的性质和特点。例如,1993年我国国内生产总值的构成情况见表1.2,它可以反映不同产业的构成,说明第三产业还有待进一步发展。

表 1.1 1993 年我国各种运输工具完成的货物周转量分组表

运输工具	货物周转量(亿吨公里)	比上年增长(%)
铁路	11936	3.1
公路	4175	11.2
水路	13672	3.1
其中:远洋	9324	3.2
民航	16	22.4
管道	606	-1.8
合计	30405	4.1

资料来源:国家统计局 1993 年统计公报。

3) 研究现象之间的依存关系。在社会经济活动中,各种现象之间不是彼此孤立的,经常存在着不同程度的相互联系、相互依存关系。在统计研究中,把表现为现象发展变化原因的事项称作因素标志,把表现现象发展结果的标志叫结果标志。通过统计分组和分组指标可以研究两个标志的联系和方向,具体表明现象之间的相互依存关系的程度。例如,从表1.3中可以看出,工人工龄越长,工作越熟练,劳动生产率就越高。

表 1.2 1993 年我国不同产业构成

产 业	增加值(亿元)	结构(%)
第一产业	6650	21.2
第二产业	16245	51.8
第三产业	8485	27.0
合计	31380	100.0

资料来源:国家统计局 1993 年统计公报。

表 1.3 某企业工人工龄与劳动生产率关系

按工人工龄分组(年)	工人数(人)	平均劳动生产率(件/小时)
3~5	7	7.2
5~7	8	7.7
7~9	12	8.3
9~11	6	8.9

分组指标的内涵和数值取决于统计分组。

2. 统计分组的方法和步骤

1) 选择分组标志。分组标志是划分总体单位为各个性质不同的组的标准或依据。如果分组标志选择不当,分组指标就难以正确反映总体的特征,甚至会把不同性质的现象混淆在一起,歪曲现象的本来面目。因此,应根据统计研究的具体任务和目的,结合研究对象所处的具体历史条件,选择具有本质性的重要标志作为分组标志。例如,研究港口企业规模大小,可以按吞吐量来划分,而研究金融企业经营状况,则可从存贷业务量来划分。

2) 确定组界。在分组标志的差异范围内,划定各个相邻组之间的性质界限和数量界限。原则是,使总体各单位能够不重复不遗漏地被唯一确认属于某一组。

3) 将各单位归并为不同的组。