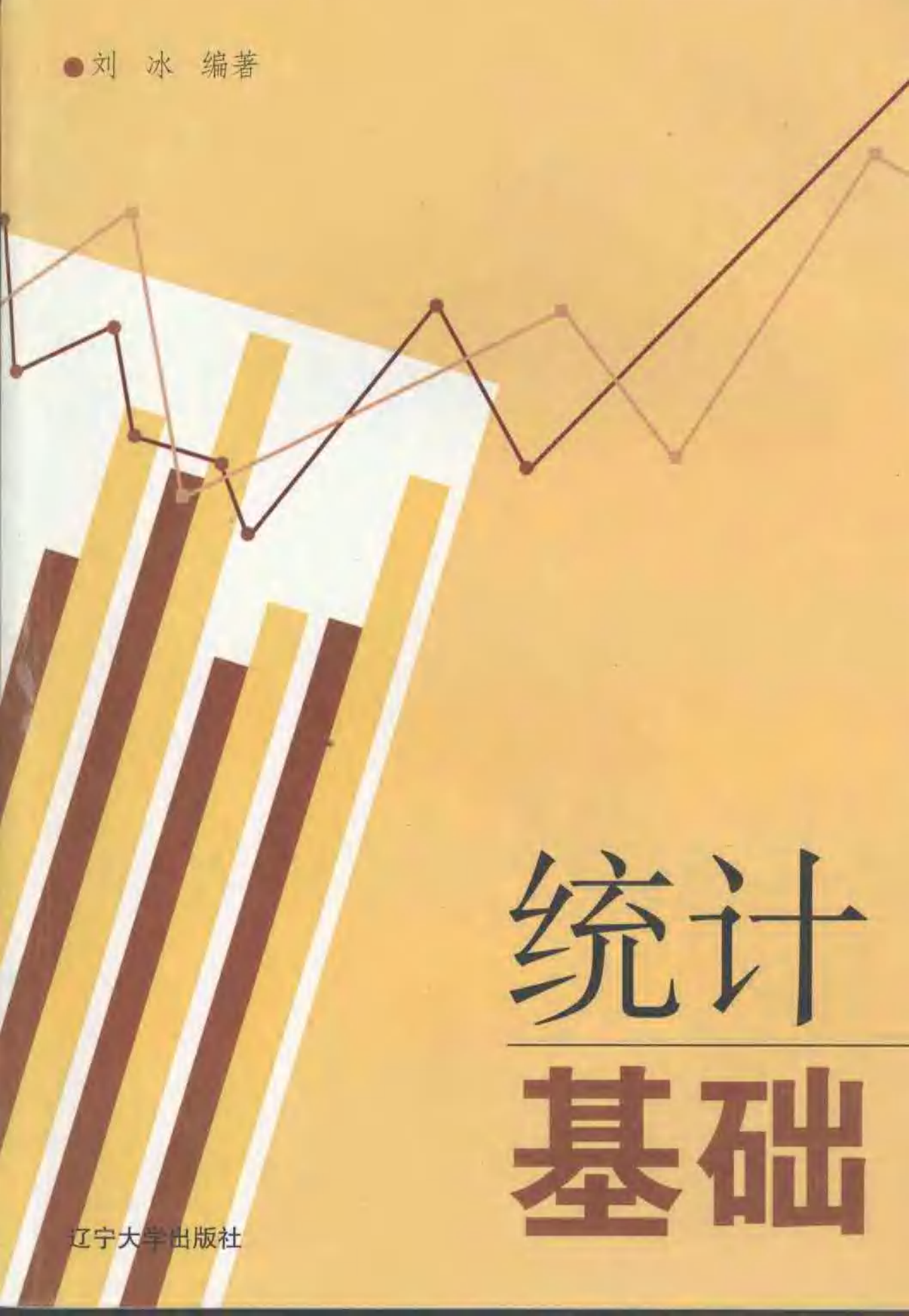


● 刘 冰 编著



The background of the cover features an abstract graphic design. On the left side, there is a bar chart with several vertical bars of varying heights, colored in shades of yellow and brown. Overlaid on this is a line graph with two lines, one dark brown and one light brown, both with circular markers at data points. The lines show fluctuating trends. The entire design is set against a solid light brown background.

统计 基础

辽宁大学出版社

统计基础

刘 冰 编著

辽宁大学出版社

©刘冰 2007

图书在版编目(CIP)数据

统计基础/刘冰编著. —沈阳: 辽宁大学出版社, 2007. 1

ISBN 978-7-5610-5257-0

I. 统... II. 刘... III. 统计学—高等学校—教材 IV. C8

中国版本图书馆 CIP 数据核字 (2007) 第 001072 号

出 版 者: 辽宁大学出版社

(地址: 沈阳市皇姑区崇山中路 66 号 邮政编码: 110036)

印 刷 者: 沈阳市政二公司印刷厂

发 行 者: 辽宁大学出版社

幅面尺寸: 185mm×260mm

印 张: 11.5

字 数: 270 千字

出版时间: 2007 年 1 月第 1 版

印刷时间: 2007 年 1 月第 1 次印刷

责任编辑: 刘东杰

封面设计: 邹本忠

责任校对: 何 成

定 价: 21.00 元

联系电话: 024-86864613

邮购热线: 024-86830665

网 址: <http://press.lnu.edu.cn>

电子邮件: Lnupress@vip.163.com

前 言

统计学是一门实用性很强的学科。它在收集大量信息的基础上,从看似杂乱无章的海量数据中,发掘隐藏在数据后面的规律,加以总结提炼后应用到各种工作领域。统计学的作用非常大,应用十分广泛,没有哪一门学科能如统计学这样具有如此强的实用性和包容性。统计学已经在诸如生活、工业、科技、军事、环境、气象等众多领域发挥极其重要的作用。例如,日常生活中我们关心商品的价格变动,究竟是上涨了还是下降了,变动的幅度有多少,我们可以从统计部门公布的物价指数中了解到。随着人们对统计方法认识的不断提高,统计调查得到越来越多的使用,如通过对众多顾客的调查,了解他们对商场所销售产品和服务的看法,经过统计分析取得客观意见和建议,从而改善现有状态。如果调查人员请求我们填写某些调查表,那就是在进行一项统计调查——问卷调查,填表人就是被调查者,我们的意见将受到足够的重视,影响最终调查结论。统计活动在实际工作中的应用事例不胜枚举,只要留心,我们会发现生活中处处有统计,统计已经在不知不觉中被我们使用,为我们服务。

编写本书的指导思想是:一要通俗易懂。全书在保证体系完整、理论不失严谨的前提下,突出统计的实用性,尽量避免数学公式的推导,主要介绍经济生活中常用统计方法的应用背景和使用条件;二要以方法使用为目的。在给出清晰的概念后,侧重介绍各类统计方法在实际中的应用状况。书中选取了一系列与实际紧密相关的前沿案例,结合案例明确统计方法,给出分析问题和解决问题的一系列统计思维步骤;三要注重计算机统计软件的利用。目前计算机在各领域的使用十分普遍,我们有必要了解如何使用最常用的统计软件进行必要的数据处理和统计分析,本书以附录的形式介绍了应用 Excel 实现绝大部分统计分析方法的功能,以期能够指导读者简化统计的应用过程。

本书共分九章。第一、二章介绍统计的基本理论和基本概念,是全书的基础;第三章介绍统计数据的描述,用最基本的指标来说明统计现象;第四、五、六章介绍概率抽样理论和进行参数估计及假设检验,偏于数理统计,但应用性比较强;第七、八、九章主要介绍经济统计方法,与经济生活实际结合更为紧密。各章之间相对独立,但又彼此联系,都是统计方法论的组成部分,学习中一定要注意知识的连贯性和互补性,这样才能较全面完整地理解统计基础知识并掌握一些重要的统计方法。

本书适用于大专院校经济类专业教学使用,也适合于统计工作者的岗位培训及个人丰富统计基础知识使用。

由于水平所限,书中难免有不妥或错误之处,请专家和读者不吝赐教。

编 者

2006 年 10 月

目 录

MULU =====

第一章 统计方法与统计应用	1
第一节 统计概念与统计应用	1
第二节 统计学中的几个基本概念	3
本章小结	7
思考与练习	7
部分练习题参考答案	8
第二章 统计数据的收集与整理	9
第一节 统计数据的收集	9
第二节 统计数据的整理	17
第三节 统计表与统计图	20
本章小结	25
思考与练习	25
部分练习题参考答案	26
第三章 统计数据的描述	28
第一节 对比分析	28
第二节 变量数列的集中趋势	32
第三节 变量数列的离中趋势	38
本章小结	44
思考与练习	45
部分练习题参考答案	45
第四章 概率与抽样分布	46
第一节 概率与概率分布	46
第二节 抽样方式	54
第三节 单一样本统计量的抽样分布	57
本章小结	60
思考与练习	61
部分练习题参考答案	62
第五章 参数估计	63
第一节 点估计	63
第二节 区间估计	64
第三节 样本容量的确定	69
本章小结	72

思考与练习	72
部分练习题参考答案	73
第六章 假设检验	74
第一节 假设检验的一般问题	74
第二节 总体平均数的假设检验	80
第三节 总体成数的假设检验	84
本章小结	85
思考与练习	85
部分练习题参考答案	86
第七章 相关与回归	87
第一节 相关分析	87
第二节 简单线性回归分析	92
第三节 多元线性回归分析	96
本章小结	99
思考与练习	99
部分练习题参考答案	100
第八章 时间序列分析和预测	101
第一节 时间序列分析概述	101
第二节 朴素的动态分析和预测方法	103
第三节 长期趋势分析和预测	109
第四节 季节性变动分析和预测	116
第五节 指数平滑预测法	121
本章小结	125
思考与练习	125
部分练习题参考答案	126
第九章 指数	127
第一节 指数的意义和分类	127
第二节 指数的编制	128
第三节 指数的应用	134
本章小结	142
思考与练习	142
部分练习题参考答案	144
附录一 用 Excel 对数据进行频数分布的统计整理	145
附录二 用 Excel 绘制统计图	155
附录三 用 Excel 产生随机数	164
附录四 用 Excel 描述总体和进行区间估计	166
附录五 用 Excel 绘制散点图并进行相关分析	169
附表一 标准正态分布	174
附表二 t 分布	176
参考文献	178

第一章 统计方法与统计应用

本章学习要点与要求

本章对统计的基本研究方法和统计方法的实际应用作了概括性介绍,侧重说明实际统计工作中常用的一些基本概念,为以后各章的学习奠定基础。通过本章学习,要求:

1. 正确理解统计的含义和统计研究对象的特点;
2. 了解统计研究的基本程序和基本方法;
3. 掌握统计学的基本概念,包括总体、总体单位、样本、标志、变量和指标等。

第一节 统计概念与统计应用

统计是人们既熟悉又陌生的概念,说熟悉是因为生活中我们会在不自觉中使用统计知识,如学生考试后计算平均分数、班级计算及格率等;说陌生是因为不具备统计知识的人把统计看成数字游戏,或者把统计看作是一门很玄妙的学科。实际上,我们生活中时时处处都会遇到统计,社会经济工作中更离不开统计。看似杂乱无章的统计数字,其实是有规律可循的,统计的价值就体现在透过现象看本质。

一、统计的含义

在日常生活中,统计的应用十分广泛,我们有必要对统计一词的含义加以说明。对统计含义的理解可以从三个方面进行:

(一)统计工作

统计工作是指统计的实践活动,包括收集、整理、研究和提供统计数据的活动。这层含义的统计应用范围很广,尤其是在经济领域,每一个管理者或研究者,都需要收集、整理和分析有关的统计信息,即从事相关的统计活动,以便加强对各领域的科学管理。实践证明,统计工作是管理工作的重要组成部分。

(二)统计数据

统计数据是指统计工作过程中收集到的各种数字资料和相关其他资料,是统计工作的成果。我们可以在统计年鉴、报刊、各种媒体上见到大量的统计数据。准确可靠的统计数据是进行有效的宏观管理和微观决策的重要依据。统计数据具有数量性、客观性、具体性和总体性的特点。

(三)统计学

统计学是指收集、整理和分析解释统计数据的方法与理论。统计学研究的对象主要是客观现象的数量方面,具体来说就是研究总体现象的数量特征与规律。由于各个领域都要研究事物的数量方法,所以统计理论和方法已经渗透到各个学科领域,成为科学研究的重要工具。

上述统计的含义,侧重点各有不同,如某人的职业是“做统计的”,这里的统计指的就是统计工作;报刊杂志上常说“据统计……”,这里的统计指的是统计数据;如学生所学专业是“统计

专业”,这里的统计则指的是统计学。在客观实际中,我们应该根据具体情况正确理解和使用“统计”一词。

二、统计学的应用

无论人们从事生产活动、科研活动还是社会活动,大多都离不开对数据的收集、整理、分析和解释。目前,统计方法已被应用到自然科学和社会科学的诸多领域,统计学在各个领域都有其分支,发展成为不同的应用统计学科,如生物统计学、人口统计学、管理统计学等。

需要指出的是,尽管统计学非常有用,应用领域也十分广泛,但是单纯靠统计学并不能解决实际工作中的所有问题。在使用统计学知识研究实际问题时,必须同时具备多方面的知识,既要掌握统计学的基本理论与方法,更要具有实际问题的知识背景,这样才能将定量分析和定性分析结合起来,在利用统计方法的基础上,应用各学科的专业知识对统计分析的结果做出合理的解释,得出令人满意的结论。

统计学是关于如何对统计数据进行收集、整理和分析,以获取信息用于支持决策的方法论科学。统计学是借助于观察和分析具体的总体现象来研究如何观察和分析总体现象数量特征的一般方法。它包括统计数据收集的方法;统计数据整理和描述的方法;对比分析方法;平均分析方法;相关和回归分析方法;时间序列分析方法;统计推断方法等。统计学提供一系列研究方法,使人们透过对总体现象数量特征的观察与分析,认识现象自身发展变化规律的存在,而解释为什么现象表现出这些规律,则由各门实质性科学来承担。

我国第五次全国人口普查资料表明:零岁人口性别比例为 112:100,由此往上,一直到 69 岁人口的性别比都是男性大于女性,自 70 岁开始,人口性别比开始逆转为 96:100,随着年龄增大,性别比急剧下降,一直到 85 岁以上人口性别比为 52:100,其他国家的人口统计数字也有类似分布。这些统计数据向人们揭示这样一条统计规律:初生男婴比例略大于女婴,在较小年龄段和中年年龄段,大体上保持这一比例。由于女性寿命平均高于男性,到了高年龄段则呈现女性比例越来越超过男性的趋势。统计学只揭示现象的内在特点,至于如何解释这条规律,则是生物学或人口学的事情,不属于统计学的范畴。

三、统计研究的基本程序和基本方法

应用统计方法研究问题,一般要经过三个步骤:

(一)统计数据的收集

统计数据依据其来源划分为原始数据和次级数据两大类。原始数据是第一手数据,是统计人员直接调查所获得的数据资料;其他渠道得来的数据都是次级数据。

原始数据和次级数据的划分是根据进行统计研究的单位或个人来划分的,自己亲自到实际中调查得来的数据就是原始数据,对于其他使用这些数据的人来说就属于次级数据。所有的统计数据都是通过统计调查得来的,进行统计调查时必须设计好统计调查方案,进行科学的调查,才能取得有用的统计数据。即使使用次级数据,也必须尽可能地了解其数据是怎样收集来的,可靠性如何,特别是要搞清楚数据的调查范围,使之与自己的需要相一致。

(二)统计数据的整理

对于原始数据,在统计调查之后,应该按照整理的方案,对调查数据进行审查、分组、汇总,直到编制成统计表,这个过程是整个统计工作过程的中间环节,起承上启下的作用。

对于次级数据,统计整理主要是审查数据的范围与所需要的范围是否一致。如果范围不

一致,首先应当加以换算调整使其一致;有些统计数据分组不一致,也必须设法调整使之具有可比性;还有的数据有缺口,应当进行必要的数据估算。因此,整理次级数据往往比整理原始数据需要更多的统计知识。相对来说,取得次级数据比取得原始数据要简洁方便,所以统计人员还是倾向于收集次级数据。

(三)统计分析

统计数据经过整理以后就可以进入统计分析阶段。统计调查的目的是为了取得分析所必需的统计数据,统计数据的整理则是对已经取得的统计数据加工,方便统计分析之用,统计分析才是统计研究的最终目的。

上述统计研究基本程序,对于各种统计活动都是适用的,只是在统计分析的深度上有所不同。旨在提供一般统计信息的统计活动,只进行统计数据的初步分析即可;旨在解决某一课题的统计活动,则需要进行较深层次的分析研究,直到找出解决问题的对策。

统计活动是对客观总体现象的数量特征进行研究,受统计方法的指导。大量观察法是统计研究的基本方法,其他统计研究方法都是建立在此基础上的。所谓大量观察法就是对同质客观现象所组成的集合中的大量个体逐一观察,由此取得反映所有个体现象某些特征的一组数据。作为加工、分析原始数据的过程,大量观察法并非要求一定要对总体现象中的全部个体无一遗漏地观察,被观察的个体数目只要足够多,能够代表总体的基本特征就可以。

大量观察法的意义在于,在同质的基础上,经过大量观察,把个别的、偶然的差异性相互抵消,将总体的、必然的规律性显示出来。如对新生儿性别比例进行观察,若只观察 10 个新生儿,其性别比例可能是 7:3,也可能是 5:5,或其他数值。但是,经过对新生儿的大量观察,发现其性别比稳定在 105:100 这个水平上,从而显现出新生儿性别比的自然规律。

第二节 统计学中的几个基本概念

一、总体、总体单位和样本

(一)总体和总体单位

我们把作为大量观察实施对象的同质个体集合称为总体,组成总体的集合成员称为总体单位。每一个个体都有一系列的属性,如体现在人身上的性别、年龄、身高、体重等,体现在一个企业上的职工人数、销售额、利润额等属性。总体是在某一属性相同(同质性)的基础上,结合起来的许多个别事物的整体。如要研究全市中学生的健康状况,全市所有的中学生就构成总体,他们具有该市中学生的相同属性,而其他的属性则可能相同,更可能不同,正是属性的不同才使总体内具有了差异性,使统计的存在具有必要性。

总体分为有限总体和无限总体,有限总体是指构成总体的个体数量是有限多个,是可数的,如某企业的所有职工是有限总体,因为职工人数的多少是可数的;无限总体是指构成总体的个体数量是无限多个,不可数的,如工厂流水线生产的产品,构成无限总体,因为截止到某一时间产品数量是有限的,但是生产还在继续,产品还在增加,工厂所生产的产品是不可数的。在社会经济现象中,统计总体大多数是有限的。总体是否有限,对统计调查方式的确定十分重要,显然对无限总体不可能采用全面调查方式收集统计数据,而对有限总体既可以采用全面调查方式,也可以采用非全面调查方式,这取决于统计研究的具体情况和目的。

总体具有同质性、变异性和大量性的特点。同质性指总体单位至少应具有一个相同的属性,并依据这一相同属性结合到一起,构成总体;变性指总体单位在除了同质属性之外的其他属性上存在着差异;大量性指总体单位数目众多,只有对大量个体进行观察,才能找到反映总体变化的数量特征和客观规律。

(二)样本

样本是指从总体中抽取的部分个体构成的集合。统计研究的目的是认识总体的数量特征,但有时构成总体的个体数量很大,实际工作中不可能也不必要对每个个体的数量特征逐一调查,通常是以某种方式,从总体中抽取一部分个体代表总体加以研究。如某外贸公司从供货方提供的一批电子元件中随机抽取 100 只,检验并推断该批电子元件的质量,这 100 只电子元件就构成一个样本,这一批电子元件是总体。由此可见,样本是总体的代表,从总体中抽取样本的目的就是用样本的特征去推断总体的特征。

二、标志和变量

表示总体单位属性的具体名称称为标志,在特定的时间、空间条件下对总体单位标志的观察结果称为标志表现。

按照标志表现的形式不同,可以把标志分为数量标志和品质标志。能够用数值描述的标志称为数量标志,其表现是数值型数据;只能用文字不能用数值描述的标志称为品质标志,其表现是品质型数据。如学生的身高、体重、年龄都是数量标志,性别、籍贯等是品质标志。

按照标志表现的种类不同,可以把标志分为不变标志和可变标志。一个总体中,当一个标志在各个总体单位上的表现都相同时,它就是不不变标志;而当一个标志在各总体单位上的表现可能不同时,它就是可变标志。在我国人口普查中,“国籍”是不变的标志,居住地、年龄、性别等都是可变标志。不变标志是构成总体的基础,在总体范围确定之后,它不再为统计研究感兴趣,可变标志才是统计研究感兴趣的对象。

总体单位、标志和标志表现三者密切联系,但又是不同的概念,总体单位是标志的承担者,而标志又需要由标志表现来具体体现。

数量标志作为一个取值可变的量,称为变量。数量标志的表现称为变量值,也称为标志值。

变量分为离散型变量和连续型变量。只能用整数表示其数值的变量是离散型变量,如一家公司各职能部门的职员人数、全国各省(市、自治区)的企业数等,都是离散型变量。可以用小数表示其值,其取值可能充满某一区间的变量是连续型变量,如一个车间的各个工人装备同样的机器所耗用的时间,不同地区人口自然增长率等,都是连续型变量。有的变量既可以作为连续型变量理解,又可以作为离散型变量处理。比如年龄,从出生开始严格计量,是一个连续型变量,但是人们通常都按实足整年头计算,又是一个离散型变量。

三、统计指标

(一)统计指标的含义

统计指标(简称指标)是用来反映社会经济现象总体数量特征的范畴。社会经济范畴是对现实生活中大量反复出现的许多经济现象的某种共同特征进行概括而形成的基本概念。而统计指标则是对社会经济范畴数量方面的规定性加以具体化,使得根据统计指标的含义对社会经济现象进行调查和综合汇总取得的统计指标的一定数值,能够具体反映这一现象在一定时

间、空间和一定条件下的数量表现。简单来说,统计指标就是说明总体现象数量特征的概念及其表现,统计指标的具体数值表现称为指标数值,简称指标值。

一个完整的指标由以下6个要素组成:①指标名称和含义;②指标数值;③时间;④空间;⑤计量单位;⑥计算方法。如2003年某企业实现利润200万元这一指标中,指标名称是利润,其含义源于所要说明现象的本质;指标数值是200万元;时间是2003年;空间是某企业;计量单位是万元;计算方法在统计上有统一规定。概括起来说,指标包括指标名称和指标数值两部分,具有可量性、总体性和具体性的特点。所有指标都能用数值表现,可以计量,指标是针对总体计算的,有了总体才有指标及指标数值,而不是缺乏所指对象的抽象的概念和数值。

统计指标值可由组成总体的各单位的相应标志值加总而得,或直接点数总体中所含总体单位数目而得,其结果为绝对数。这种统计指标称为总量指标。如2004年,我国出口总额为5934亿美元,它是把2004年内所有经海关实际出口货物金额加在一起而得。又如截止到2004年底,我国从事对外贸易的企业有403,426家,这也是一个总量指标,它是对全国范围内从事对外贸易的企业逐一点数得到的。

在总量指标值的基础上,经过计算所得到的相对数或平均数,也是统计指标,用来反映社会经济现象数量特征的相对状态和平均状态。如2003年,进出口税收占全国税收总额的13.6%,成为国家税收的第二大税源;大学毕业生在上海工作首月平均工资为1680元等。

(二)设置统计指标的原则

1.统计指标要符合客观规律

统计指标既然是反映客观存在的社会经济现象的数量表现,它的范畴就一定要如实地反映客观现实。如现在使用农村社会总产值指标来反映我国农村经济发展的客观现实,而不再使用原有的农业总产值。

2.统计指标要具有可量性

所设置的统计指标不仅在概念上要明确,而且在实践上还应当便于调查和计量。既要考虑统计研究的必要性,也要考虑实施的可能性。

3.统计指标的口径要尽可能地与业务核算口径相一致

这样要求的目的,不仅是为了进行比较,也是管理的需要。因为在基层单位,业务、财务和统计工作是以一套原始单据为依据的,核算口径一致后可以使工作更加流畅。

4.统计指标的设置要考虑到国际可比性

有些统计指标的设置,可以有不同的方法,但应尽量选用国际通用的标准,或选用可以较为容易改算为国际通用指标的标准。

社会经济的各种现象间存在着相互联系、相互制约的关系。一个统计指标只反映社会经济现象的某一个数量方面,要反映一系列的社会经济现象间数量方面的联系,必须把多个统计指标联系在一起研究,这就构成了统计指标体系。在人口统计中,人口指标、性比例指标、人口自然增长率和人口平均增长率等指标,构成了人口统计指标体系,用来反映人口数量特征的现状和发展趋势。

由于现象的复杂性和事物的多层次性,大指标体系之内还可以设计次一级指标体系,乃至更低阶次的指标体系。国民经济核算指标体系就是一个庞大的指标体系,它包含对外经济联系次级指标体系,而对外贸易统计指标体系又是对外经济联系指标体系的一个小体系。

(三)常用经济指标

1. 总产值

总产值是用货币表现的一定时期内全部产品的价值,反映生产单位、生产部门或国民经济在一定时期内生产活动的总成果。总产值是建立在马克思经济理论的基础上,认为劳动是创造价值的唯一源泉,因此,全社会中只有五大物质生产部门,即工业、农业、建筑业、运输邮电业、商业等部门能够创造价值,五大物质生产部门总产值之和即为社会总产值。

2. 增加值

增加值是指在报告期内常住单位所创造的最终价值。增加值有两种计算方法:一是生产法,用全部总产出减去为取得总产出所投入的中间消耗,所得结果就是增加值,即增加值 = 总产出 - 中间投入;二是收入法,也叫分配法,将各个生产要素提供者所获得的初次分配收入加总,得到增加值,增加值 = 固定资产折旧 + 劳动者报酬 + 生产税净额 + 营业盈余。生产法和收入法的角度不同,计算结果是相同的。

3. 国内生产总值(GDP)

国内生产总值(GDP)是指一个国家(或地区)所有常住单位在一定时期内生产活动的最终成果。GDP有三种表现形式:价值形态、收入形态和产品形态。从价值形态看,它是所有常住单位的增加值之和;从收入形态看,它是所有常住单位在一定时期内创造并分配给所有常住单位和非常住单位的初次收入分配之和;从产品形态看,它是所有常住单位在一定时期内最终使用的货物及服务价值与货物及服务的净出口之和。

国内生产总值可以从不同环节分别采用不同的方法来加以计算:

(1) 生产法

生产法的基本原理是:首先计算国民经济各部门的总产出,再从总产出中扣除相应部门的中间消耗,求得各部门的增加值,最后汇总所有部门的增加值,得到国内生产总值。

$$GDP = \sum (\text{各部门总产出} - \text{该部门的中间消耗}) = \sum \text{各部门的增加值}$$

(2) 收入法

收入法依然是将各部门的增加值加总求得国内生产总值,只是在计算各部门增加值时,使用的是收入法,而不是生产法。

$$GDP = \sum \text{各部门的增加值} = \sum (\text{固定资产折旧} + \text{劳动者报酬} + \text{生产税净额} + \text{营业盈余})$$

(3) 支出法

支出法是从全社会角度着眼,直接将各种货物及服务的最终使用额汇总起来,再减去商品进口,求得国内生产总值。

$$GDP = \text{总消费} + \text{总投资} + \text{净出口}$$

$$= (\text{居民消费} + \text{政府消费}) + (\text{固定资产投资} + \text{存货增加}) + (\text{出口} - \text{进口})$$

4. 国民总收入(GNI)

国民总收入是指一个国家(或地区)的国民一定时期内在国内外生产的最终产品及劳务的价值总和。

$$GNI = GDP + \text{来自国外的要素收入} - \text{付给国外的要素收入}$$

国民总收入是从国内生产总值中扣除了由国外生产要素所创造产值份额,加上由本国的生产要素在国外所创造的产值份额之后得到的结果。GNI实质上是一个收入分配指标,它表明本

国常住单位由于提供各种生产要素而从国内和国外的生产活动中所获得的全部收入的总和。

5. 可支配总收入(PDI)

可支配总收入是指一个国家(或地区)一定时期内可以由个人支配的全部收入。

可支配总收入 = 国民总收入 + 转移收入 - 转移支出

本章小结

1. 统计具有统计工作、统计数据和统计学三种含义,在不同的地方使用不同的含义。
2. 统计的基本程序是统计数据的收集、统计数据的整理和统计数据的分析。
3. 统计学中的基本概念包括总体、总体单位、样本、标志、变量和统计指标,经济中常用的统计指标有总产值、增加值、国内生产总值和国民总收入。

思考与练习

1. 统计一词有哪些含义?
2. 统计研究的基本程序有哪些?
3. 什么是总体、总体单位和样本?什么是标志、变量和指标?
4. 为什么统计研究的基本方法是大量观察法?
5. 一个完整的统计指标包含几个因素?统计指标有哪些特点?经济生活中常用经济指标有哪些?
6. 阅读以下数据,回答问题:

2003年中国企业(成长型)排名100强前10名排行榜

排名	营业收入排名	公司简称	00-02增长速度(%)	00年营业收入总额(万元)	02年营业收入总额(万元)	02年资产总额(万元)	02年员工总数(人)	省份	行业类别
1	19	内蒙古小肥羊	1519.5	251700.0	16560.0	129984.6	4000	内蒙	餐饮业
2	8	长城电器	1367.8	462817.0	33837.0	189504.0	5927	浙江	设备制造业
3	30	美特斯邦威	903.1	158034.0	17500.0	47060.3	3216	浙江	纺织业
4	96	北京宅急送	728.0	13125.0	1803.0	10318.9	4400	北京	快运物流
5	78	河北中旺	617.4	36842.0	5967.0	13055.0	5000	河北	食品
6	91	西安海天天线	614.8	16452.6	2676.0	24813.4	530	陕西	通信制造业
7	6	人民电器	509.6	621008.0	121873.0	157017.0	11062	浙江	电气机械
8	79	河北旭瑞	491.3	35089.0	7142.0	12971.0	680	河北	食品
9	88	山西恒康乳业	452.2	22625.0	5003.0	35615.0	1968	山西	食品
10	2	广东格兰仕	434.9	862379.0	198302.0	359127.0	16000	广东	家电行业

- (1)该资料中有多少个数据?哪些是品质数据?哪些是数值数据?
 (2)该资料中有多少个变量?哪些是离散的?哪些是连续的?
 (3)若以前10名企业作为样本,分别指出总体、总体单位和样本单位。

部分练习题参考答案

6.(1)100个数据,其中排名、营业收入排名、公司简称、省份、行业类别品质标志下的数据是品质型数据;增长速度、营业收入总额、资产总额、员工总数数量标志下的数据是数值型数据。

(2)10个变量,其中员工总数是离散的;增长速度、营业收入总额、资产总额是连续的。

(3)总体:2003年中国企业(成长型)排名100强的所有企业;

总体单位:2003年中国企业(成长型)排名100强的每一个企业;

样本单位:2003年中国企业(成长型)排名100强中前10名中的每一个企业。

第二章 统计数据的收集与整理

本章学习要点与要求

本章介绍统计数据收集和整理的理论和方法。通过本章学习,要求:

1. 了解统计数据的收集方法;
2. 了解统计报表制度、普查、抽样调查、重点调查和典型调查等调查方式的特点及适用环境;
3. 了解统计调查方案的基本内容;
4. 了解统计数列的种类;
5. 掌握统计分组的方法;
6. 掌握频数分布表的编制方法。

第一节 统计数据的收集

统计数据的收集,包括对原始数据的收集和对次级数据的收集。原始数据的收集也叫统计调查,是直接取得原始数据的活动。

一、原始数据的来源

收集原始数据是指对客观现象的个体表现进行有计划、有组织的调查。在某些情况下,统计调查可以借行政业务记录之便进行。政府机关和企事业单位为了行政管理 and 业务管理的需要,对关系到其工作进度和成果的社会经济现象,平时均做系统的记录或登记。当统计研究需要这方面的数据时,即可以通过查阅这些数据的途径取得,或者委托行政和业务单位按统计活动的要求定期报告有关数据。如载于交通安全管理部门的行政记录,可以用于交通统计研究工作等。

当没有行政业务记录可利用时,必须组织直接的调查活动获取原始数据。这又有两种取得原始数据的途径,一种是派员到待研究现象发生的现场去注视现象的即时状态和演化进程的特征,并根据事前拟订的调查表进行记录,我们称这种调查方式为观察法。如城市街道路口交通流量的研究、消费者逛商场和购买行为的研究、生产工人岗位动作及其时间的研究等,都可以采用观察的方法来取得数据。另一种调查途径是以交流的方式进行调查,通过调查人员与被调查者的口头或书面交流,完成对拟调查项目的调查。人口普查、社会舆论调查都属于交流调查。

二、次级数据的收集

对于研究者来说,收集次级数据可以节省更多的人力、物力和时间,次级数据一般包括各种公开出版物和某些部门的内部资料。

(一)公开出版物

世界各国或地区都有定期的统计出版物。在我国,国家统计局和各省、自治区、直辖市的

人民政府的统计机构,定期地公布国民经济及社会发展的各种统计信息,如反映我国及各省市、部门统计资料的各类统计年鉴;反映世界各国统计资料的公开出版物有《联合国统计年鉴》、《国际贸易统计年鉴》等。这些出版物均能提供系统、大量的数据,供国民经济各管理部门和社会使用,也是收集次级数据的主要渠道。

公开出版物除了各种统计年鉴外,我们还可以查阅某些专门调查的统计资料。如逢0年份进行的人口普查、逢3年份进行的第三产业普查、逢5年份进行的工业普查等。除此之外,各种报刊、杂志、图书、广播电视和互联网等传媒中的数据也是被广为引用的次级数据。特别是随着网络的发展,许多公司和部门建立了自己的网站,并且提供着越来越丰富的统计数据,如万方数据公司、中国在线等。

(二)来源于内部调查的数据

除了已公开发表的数据之外,还有许多他人已经收集,但没有公开发表的数据。如一些调查咨询公司的数据、气象台的气象数据、城市家庭的收支数据等,都收集掌握着大量的统计信息,但由于数据保密,或是需要有偿服务,或是受公开发表的篇幅限制等原因,这些统计数据不为社会通用。对于这些次级数据,应该充分挖掘和充分利用,因为获取这些资料比自己进行调查要节约很多资源。

三、统计调查的方式

按收集数据的组织方式不同,统计调查可分为全面统计报表和专门调查。全面统计报表是依据经济活动原始记录汇总成一定的报表,按期向上级报送,逐级汇总以取得全面数字的报告制度。专门调查是指为某一特定目的临时组织的一种调查活动,主要包括普查和抽样调查。

按实地观察的总体单位的全面与否,统计调查可以分为全面调查和非全面调查。尽管统计研究以总体为对象,但在调查实施过程中,却不一定对全部总体单位逐一进行观察。无一遗漏地观察全部总体单位为全面调查,全面统计报表制度和普查都是全面调查。只观察总体中部分单位的调查是非全面调查,主要有重点调查、典型调查和抽样调查,我们用得最普遍的是抽样调查。

(一)全面报表制度

全面报表制度是各基层单位及各主管机关用统计表格的形式,按照统一规定的项目、内容和上报时间,经常向上级报告经济活动情况的一种方式。为保证所取得的统计数据在全国范围内达到统一,国家对统计报表的实施范围、指标解释、计算方法、填报方法、报送期限、报送程序以及报表的制订和管理办法等方面做出统一规定,以政府的名义颁发实施。

原始记录是基层单位通过一定的表格形式,对生产和业务管理活动所进行的第一手记录,如工厂的产品产量记录、质量记录,工人出勤记录、原材料消耗记录;商品销售记录、现金支付凭证、库存物资收付记录等。

全面统计报表是我国特有的调查方式,是取得国民经济统计数据的重要手段。对需要定期地、系统地了解其情况的国民经济活动,使用这种调查方式比较合适。但随着经济环境的进一步多样化,某些数据已经不适合于用全面统计报表来收集,必须使用其他调查方式。

(二)普查

普查是为了某一特定目的而专门组织的一次性全面调查。有些社会经济现象不需要或不可能进行经常性调查,但又需要掌握它的全面情况,这就需要普查的方式来收集数据。普查

一般是对某一社会经济现象总体在一定时点上的状态所做的调查,如人口普查、物资库存普查、工业普查、基本建设项目普查等。普查也是一种重要的统计调查方式,世界各国对重大国情国力的调查,通常都采用普查的方式。近年来,我国已将普查作为最重要的统计调查方式,并逐渐规范化、制度化。经国务院批准,周期性普查项目包括:人口、工业、农业、第三产业和基本统计单位等。人口、第三产业、工业和农业普查每隔 10 年进行一次,分别在逢 0、3、5、7 年份进行;基本统计单位普查每隔 5 年进行一次,在逢 1、6 年份进行。

普查主要是用来收集较全面、细致的统计数据,为国家掌握有关国计民生的基本情况,制定发展规划和方针政策提供基本依据。

普查所耗用人、财、物力很大,占用时间较长,不宜经常采用,如人口普查一般国家都是每 10 年搞一次,其余年份以抽查数据加以修正和推算。

(三)抽样调查

抽样调查是一种非全面调查。它是按随机原则从调查对象中抽取一部分单位作为样本进行观察,然后根据样本数据去推算调查对象的总体特征。

抽样调查有如下几个特点:第一,样本单位按随机原则抽取,排除了主观因素对选样的影响。第二,根据部分调查的实际资料对调查对象总体的数量特征做出估计。根据数理统计的原理,抽样调查中样本指标和相对应的总体指标之间存在着内在联系,而且两者的误差分布也是有规律可循的,因而抽样调查提供了一种用实际调查所得的部分信息来推断总体数量特征的科学方法。第三,抽样误差可以事先计算并加以控制。以样本数据推算总体数量特征,不可避免地会产生误差。但这种误差与其他统计估算所产生的误差不同,它可以根据有关资料事先加以计算,并通过一定的途径来控制误差的范围,保证抽样推断结果达到预期的可靠程度。

抽样调查除随机抽样外,还有非随机抽样,任何不符合随机抽样要求的抽样都称为非随机抽样。非随机抽样方式因其简便易行,在实践中被广泛应用。但是,由于它的抽样结果在推断总体特征时不能说明其误差的大小,因此,使用非随机抽样对总体进行推断时要非常慎重。在选择非随机抽样后,应考虑以下几个问题:样本是否能够覆盖目标总体的各个部分,样本是否无目的地倾向于一些特殊方面,调查者之间存在的水平差异是否影响调查结果。回答这些问题将有助于减少非随机抽样时的误差。经常使用的非随机抽样包括:方便抽样、主观抽样、配额抽样和滚雪球抽样。

1. 方便抽样

方便抽样就是在选择样本时以方便为原则,是选择样本和收集数据的一种相对简单的方法。如某教授在大学里进行一项有关大学生的研究,他可以邀请在他的班级上课的学生作为研究项目的样本,因为他选择这个样本很方便。在某些特定情况下,我们也不得不使用方便抽样代替随机抽样,如对野生动物生存状况进行研究以及对消费者消费行为的研究,都只能采取方便抽样的形式。因为对消费者和野生动物使用随机抽样的方法是不可能的,无法将总体中的全部个体逐一编号。在类似情况下,进行方便抽样将是收集统计数据的有效而实用的方法。在许多试探性调研时,缺乏经验而又急需真实数据的近似值时,这种方法很实用。

2. 主观抽样

根据个人的主观意识来选择总体中有代表性个体的方法,称为主观抽样。如税务人员对于个体经营者赋税水平的抽样调查;市场调查人员在商场门前对消费者的抽样调查等,都属于