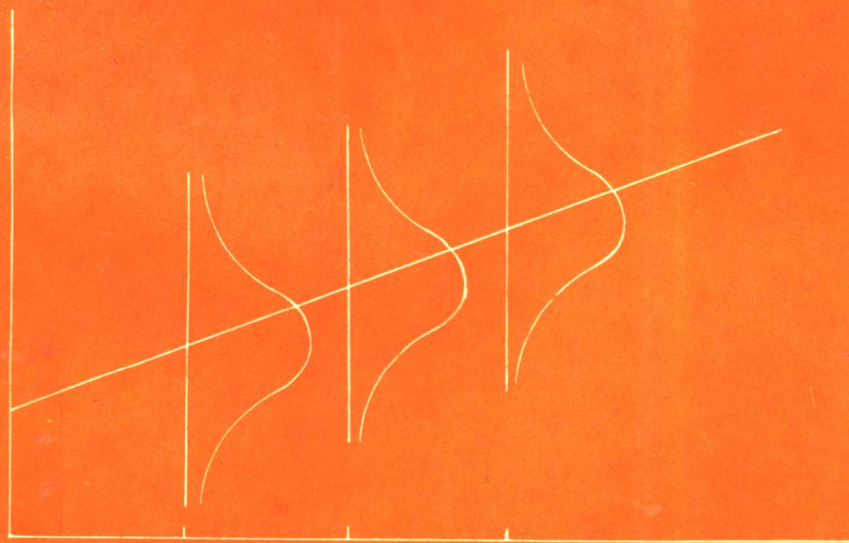


国际工商管理本科教材

应用统计

■ 贾怀勤 主编



对外经济贸易大学出版社

应 用 统 计

贾怀勤 主编

对外经济贸易大学出版社

(京) 新登字 182 号

图书在版编目 (CIP) 数据

应用统计/贾怀勤编著. —北京: 对外经济贸易大学出版社, 1994

ISBN 7-81000-666-5

I. 应… I. 贾… III. 统计—方法 IV. C81

中国版本图书馆 CIP 数据核字 (94) 第 04124 号

对外经济贸易大学出版社出版

社址: 北京惠新东街 12 号 邮政编码: 100029

北京市通州印刷厂印刷 新华书店北京发行所发行

850×1168 1/32 18.125 印张 468 千字 插表 1

1994 年 6 月北京第 1 版 1997 年 4 月北京第 2 次印刷

ISBN 7-81000-666-5/F·242 责任编辑 张孟秋

印数 3001—6000 册 定价: 23.50 元

目 录

绪论	(1)
第一章 统计设计	(1)
第一节 统计设计概述	(1)
第二节 统计指标的设计	(10)
第三节 实验设计	(16)
第二章 统计资料的搜集和整理	(21)
第一节 统计资料的类别和来源	(21)
第二节 统计调查的方式和方法	(23)
第三节 统计资料的整理和表述	(34)
第三章 数据的静态对比分析	(60)
第一节 相对数的意义和种类	(60)
第二节 结构相对数	(62)
第三节 其它静态相对数	(70)
第四章 单变量截面数据的描述性分析	(80)
第一节 频数分布数列的图形表述	(81)
第二节 集中趋势指标	(90)
第三节 离散趋势指标	(104)
第四节 偏度指标和峰态指标	(111)
第五节 位次指标	(114)
第五章 双变量截面数据的描述性分析	(129)
第一节 列联表	(129)
第二节 相关与回归分析	(137)
第六章 时间数列分析	(161)
第一节 时间数列的动态比较和平均分析	(161)

第二节	时间数列的解析分析	(176)
第七章	统计指数	(213)
第一节	指数的概念	(213)
第二节	指数的编制方法	(215)
第三节	指数的应用	(231)
第四节	因素分析	(241)
第八章	概率	(252)
第一节	概率的概念	(252)
第二节	概率的计算规则	(258)
第九章	概率分布	(283)
第一节	概率分布	(283)
第二节	常用的离散型概率分布	(295)
第三节	正态分布	(315)
第十章	抽样和抽样分布	(331)
第一节	抽样的基本概念	(331)
第二节	抽样分布的基本原理	(342)
第十一章	参数估计	(373)
第一节	点估计与区间估计	(373)
第二节	区间估计的应用	(384)
第三节	样本容量的确定	(393)
第十二章	假设检验	(402)
第一节	假设检验的有关概念	(402)
第二节	单一样本的假设检验	(419)
第三节	双样本的假设检验	(434)
第十三章	关于简单回归和相关的统计推断	(447)
第一节	关于因变量的估计和推算	(447)
第二节	关于回归和相关总体参数的推断	(458)
第三节	可决系数	(462)
第十四章	多元回归和多重相关分析	(474)

第一节	多元线性回归关系的描述和推断	(475)
第二节	多重相关分析	(485)
第三节	多元回归分析中的 F 检验	(493)
第四节	关于多元线性回归模型假定条件的讨论...	(497)
第五节	多元回归和多重相关分析的综合运用 ...	(503)
统计数表		(521)
附表 1	二项分布累积概率表	(521)
附表 2	普哇松分布累积概率表	(524)
附表 3	标准正态分布	(529)
附表 4	随机数字 4000 个	(530)
附表 5	t 分布	(533)
附表 6	F 分布	(534)
附表 7	杜宾-沃岑检验上下界	(544)
词汇中英文对照		(546)

第一章 统 计 设 计

第一节 统计设计概述

1.1.1 统计设计的意义和作用

统计活动,是依据科学的研究方法对客观总合现象的数量表现进行有目的、有组织的研究,是系统的活动。这样的系统活动,必须在一定的活动方案指导下,才能顺利开展。统计设计的任务就是制定一套为统计研究目的服务的、切实可行的活动方案。

统计设计搞得与不好,直接关系到统计研究的质量和效率,关系到统计活动费用的高低,最终关系到统计研究的价值——科学价值、经济价值和社会价值。统计设计搞得不好,就会造成所搜集的资料不真实、不全面、或不及时,或造成整理和分析方法欠缺、偏误,乃至误用,最后导致不可靠或全然错误的结论,从而使得整个统计研究失去了科学价值,也就没有经济价值和社会价值可言。

统计研究的科学价值、经济价值和社会价值应该是统一的。在市场经济中,除去最基本的研究活动之外,统计活动的成果应该既具有科学价值,同时又讲求经济价值,顾及社会价值。一项统计活动,如果其活动所费超过了其成果带来的经济利益,或者因为没能及时提出成果导致本来可以收到的经济利益丧失掉,就不能说其设计是成功的。当然,统计研究成果所能带来的经济效益是不易计算的,尤其对于大型的官方统计活动更是如此。总之,一个成功的统计设计,所采取的步骤和措施对统计研究目的而言,应该是充分而必要的,所搜集的资料应该是准确、及时的,所使用的方法应该是科学的、讲求效率的,所出现的误差应该控制在合理的程度以

下。

统计设计的作用表现在如下两个方面：

(1)它是统计活动由定性认识向定量认识的过渡。统计活动从研究对象总体的定性认识出发,经过在设计工作中对总体进行全面考虑,制定活动方案,为取得统计数据并作定量分析作出计划安排。

(2)它对统计活动起通盘安排的作用。在设计工作中,根据统计研究的需要和可能,对各种不同的局部方案和总体方案进行考虑和比较,以便最终拟定一套较为科学、合理的方案,并使得各个平行作业活动相互协调,各个递序作业活动相互衔接,使整个作业活动有条不紊地进行。

1.1.2 统计设计的内容

科学技术领域的统计活动,即在科学研究和技术开发项目中采集数据、加工数据和使用数据的活动,是整个项目密不可分的一部分。整个项目方案的设计之中,就融入了数据采集、加工和使用的方法、步骤等,而且这个设计过程的专业性非常强。本书不拟对此多加阐述。这里讲的“统计设计内容”和此后本章其余部分,以及下一章关于统计调查的内容,基本上是针对社会经济领域的统计活动讲的。

在阐述统计设计内容之前,有必要明确统计资料的加工者与需求者之间的关系,和前者对后者的服务方式。

统计资料,也就是统计信息。它是经济信息的主体。统计活动是生产统计信息的活动。从信息和系统的观点看,统计活动涉及到统计主体、统计客体和统计宿体三个部分。这三部分关系如图1—1所示。

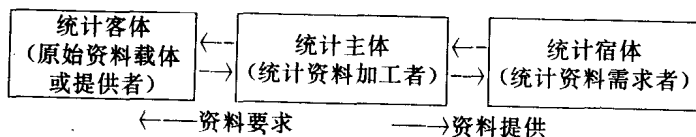


图 1—1 统计活动三部分相互关系

作为统计主体的有,官方统计机构,和企业、社团、科研院所、大专院校举办的,乃至民办的信息、咨询机构。其中,官方统计机构是统计信息的主要生产者。

作为统计客体的有:参与各种人生活、经济活动和社会活动的人,企、事业单位,国家机关等行为主体。

作为统计宿体的有:国家机关,企、事业单位,科研机构,居民,还有国外的机构、企业和人员。

三者的联系形式是:统计宿体提出统计研究的目的,或统计活动的成果要求(同时以某种渠道和方式提供统计活动所需人力、财力、物资,或支付统计咨询报酬);统计主体根据统计宿体的目的和要求,向统计客体搜集统计资料,再经过整理和分析,提供给统计宿体使用。

统计主体的统计(资料生产)活动,有两种类型:一种是通用型活动,另一种是专题型活动。前者提供的是一般统计信息,不以任何特定统计宿体为提供对象,这好象服装厂家按各种不同规格、式样进行批量生产一样,任何顾客中意都可以购买。后者提供的是专门加工的统计信息,而且还要在此基础上提出研究报告,对统计宿体的决策直接发挥咨询作用,这好象服装店为每个顾客量体裁衣,提供专门服务。从事通用型统计活动的,主要是官方统计机构,但是也有少数非官方统计机构发布某些经济的或社会的信息。从事专题型统计活动的,主要是官方统计机构以外的信息、咨询机构,但是官方统计机构在完成正常的通用型活动任务之余,也可能开展专题型统计活动。

一项统计设计的各阶段可以用图 1—2 来说明。

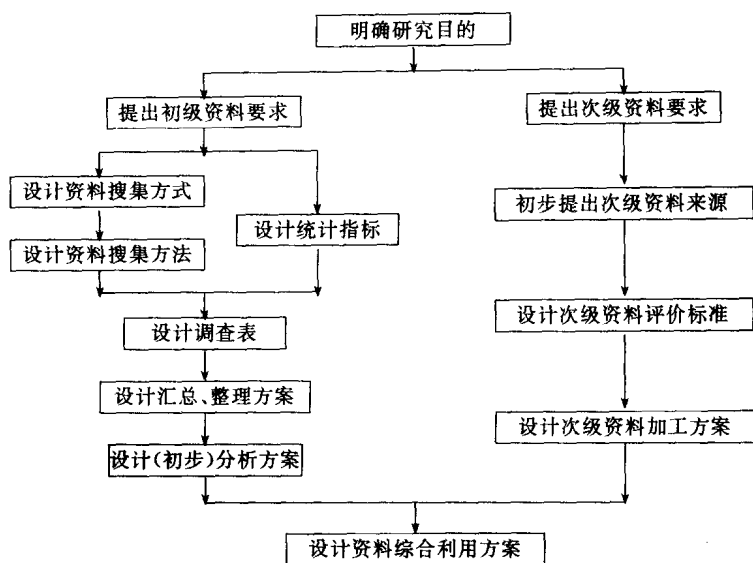


图 1—2 统计设计主要内容

(1)任何统计设计首先都要明确统计研究的目的。对于通用型统计活动,是把统计宿体提出的统计任务和要求,归纳为清晰的、具体的,并且是切合实际的统计研究目的。有些时候,或对于某些统计主体来说,则是根据设想的统计宿体的需求来规定统计目的。对于专题性统计活动,信息咨询的需求者一般只提出面临的业务问题,而要由信息咨询服务的提供者在深刻理解对方问题的基础上,自己提出统计研究的目的。总之,规定统计目的是统计设计的第一步,也是指导全局的关键一步,是设计具体方案的出发点。

对于通用型统计活动,除了最后要发布历史对比或横向对比数据外,并不需要已经形成的综合统计资料(即次级资料)。为了方便叙述,我们先抛开图 1—2 的右半边,专门讲通用型统计活动的

设计内容。

(2)根据研究目的,明确对要搜集的资料的要求。设想所要取得的资料反映哪些属性,这些属性由什么样的总体承载,总体的空间、时间如何界定。

(3)根据对统计资料的要求设计统计指标和指标体系。首先是确定用什么指标反映前述现象属性。然后是设计指标的涵义和口径等。在设计一系列指标所组成的指标体系时、要注意设计统一标准。

(4)设计搜集统计资料的途径和方式。

(5)设计搜集统计资料的具体方法。

对于可以选择的途径、方式和方法,要设计最佳的搭配,以利用以较小的花费、较短的时间搜集到较为准确、全面的,对统计研究充分而必要的资料。

(6)设计调查表。不同的资料搜集途径、方式和方法,需要设计不同类型、不同形式的调查表,以便于实施现场搜集活动。即使是同样的途径、方式和方法,又可以设计出风格不同的调查表,这项工作需要语言学、心理学等学科知识。

(7)设计资料汇总整理方案。根据对所研究现象的深入分析,设计出能反映问题本质的分组标志和指标体系,具体地落实到一套空白的汇总表、汇总表编制说明和各种分类目录。

(8)设计统计分析方案。根据要研究的问题,准备搜集的数据,确定使用统计分析方法。

对于专题型统计活动来说,这里的统计分析只是初步的,还必须综合运用其他方面的数据资料和文字资料进行全面分析,深入研究,才能实现研究目的。

对于通用型统计活动来说,到这里可告结束。将分析结果制成各种统计资料汇编表,上报或发布即可。

(9)提出对次级资料的要求。明确需要反映什么现象的资料,总体的涵义和口径是什么,资料所属时间如何。

(10)设想应该到什么地方去查阅或购买所需次级资料,使用什么样的联系方式。

(11)对拟取得的次级资料要设计一套大致的评价标准,以确定它们的使用价值。

(12)设计次级资料的加工方案,使其符合研究目的需要。

(9)一(12)步不可能由一次设计完成,只能是在发现一些主要的或易接触的次级资料来源,并对其作一定的浏览之后,才能进行第(11)和(12)步。

图 1—2 右半边关于搜集和加工次级资料的设计,尤其是第(9)、(10)两步,应该先于左半边的设计。对于专题型统计活动,利用次级资料能解决问题,就不需要初级资料;次级资料缺哪些,再由初级资料补哪些。图内所用内容方框的安排出于其他方面的考虑,没有突出这个设计顺序。此外,上边对于内容解释各段所加的序号,纯粹为了叙述层次分明,也不要理解为右半边内容接在左半边之后。

在图 1—2 所列统计设计内容之外,还有几项内容。它们分别是:

(13)复查方案的设计。在大型统计活动中,有时需要对资料搜集阶段的工作质量进行评估,并在资料分析时考虑到评估意见。这种评估建立在对资料复查的基础上。复查不是对原调查工作的逐步审核,而是通过一种较为简单迅捷的方式对研究对象进行局部调查,将对局部对象的两次调查结果进行对比,看其差距如何。

(14)设计试点调查和评估方案。大型统计活动,一般需要经过试点阶段。目的在于测试所设计方案的可行性,取得经验,指导全局。

(15)许多通用型统计活动和少数专题型统计活动是周期性的调研,而非一次性的调研。为此,要根据统计研究目的,规定调查(填报)频度和时间。

(16)制定统计活动所需人力、物力和时间的配置方案。其中尤

为注意的是工作人员的培训方案,和建立与调查单位的良好配合关系的方案。

(17)制定统计活动的组织机构和管理方案。

以上阐述的是单项统计研究的全阶段设计应该包括的内容。对于一些统计活动,有的时候只需要详细制定某一阶段(如统计资料的搜集)的活动方案,是为单阶段设计。

相对于单项统计设计的,是整体统计设计。这指的是通用型统计活动对其整个工作的全面设计。这样的设计所要考虑的方面较单项设计要广得多,但是从内容上看基本是一致的。

上边所述统计设计内容,其中许多基本设计方法都贯穿在以后各章展开的统计方法之中,唯有统计指标设计和实验设计是两个独特的内容,放在本章第二、三节阐述。

1.1.3 统计设计在整个统计活动程序中的位置

统计设计作为综览全局的一项工作,是整个统计活动的开始阶段,以下依次为统计资料的搜集、统计资料的整理、统计分析。在实际活动中,这个程序允许存在循环。这是因为在统计活动开展之前,人们对于所研究的现象的认识是不深入的、不全面的,甚至可能是错误的,基于这种预认识所设计的统计研究方案在执行中遇到意外情况是很难避免的,有些情况下原方案中的某一条、某一步不宜执行下去,或者根本不能执行下去,就需要对原方案进行局部的调整或修正。这种情况可能发生在搜集资料时,或发生在整理资料时,甚至发生在资料的分析和综合使用时。甚至有个别情况、发现局部的设计失误,需要重新设计。我们只能希望统计设计尽可能切合实际,但是不能排除实施中的循环。

另外,有些统计活动可以承袭和套用历史上也已形成的统计指标体系、分类方法,或搜集、整理和分析方法,在搜集资料开始前只作非常粗线条的设计,随着实际工作展开到某一步,根据接触到的实际情况,对下一步活动作局部设计也是可行的。

1.1.4 统计设计的原则

统计设计所应遵循的基本原则可以概括为可行性、准确性、时效性和经济性。这是由统计活动在国民经济管理、企业管理、科学研究和社会生活管理等诸领域中的作用,和统计活动的主体、客体与宿体三者关系决定的。

可行性原则

这里所说的可行性,指的是所设计的方案在实际上必须是具备实施的主、客观条件,即是要确认:

- (1)就逻辑而言、所设计的统计活动可以导出原定的目的;
- (2)就客观情况而言,所设计的统计活动可以取得必需的数据;
- (3)就主观条件而言,具备开展此项统计活动所需的资源——人员、资金、物资设备和充足的时间。

如果设计方案不能确保前两个条件,就是废案。这就要或者推倒重来,或者放弃研究。如果设计方案面临第(3)条的困难,可以仔细审查原设计,看看能否经细微调整即能降低资源消耗,或缩减时间;或者考虑在不影响总的研究目的的前提下,适当减少一点活动内容,作方案的局部修改,从而使不可行变为可行。

统计活动作为统计信息生产活动,一定要从用户(统计信息需求者)的需要出发,以满足他们的需求,解决他们提出的咨询问题为目的。因此,统计设计的整体思想和每一个步骤、每一条措施,都要紧紧扣住用户需求的主题。这对于信息咨询机构的统计设计尤为重要。

准确性原则

统计设计方案为求所获统计信息的准确性,一方面要减少资料搜集方案可能给实施现场调查带来的误差,另一方面要防止不恰当的整理方法和分析方法的误用,以免造成“高质量原料、低质量产品”的可悲局面。此外,指标和指标体系的设计,以及实验设计,对于统计设计的准确性也是非常重要的。

违背准确性原则有两种情况：一种是设计失误，另一种是设计误导。前者是设计者无意造成的，后者是设计者有意造成的。由此就引出了对统计研究的客观、真实的要求。统计设计乃意在客观地反映社会经济现象的数量特征，切忌先入为主，在设计指标、设计资料的搜集、整理和分析方案时下意识地设定某种导向。

时效性原则

统计资料是为决策服务的，因此必须讲求时效性。在市场经济体制下，统计资料的时效性尤为重要。在进行统计设计时，要力求以较短的时间提供所需的统计资料。因此要按照系统的观点设计总体方案和各阶段方案，合理的配置资源，在不影响工作质量的前提下，尽可能地实现时间的节约。

经济性原则

经济性原则，就是节约的原则。在作统计设计时，要研究如何针对统计宿体的需要，用尽可能少的人、财、物力，去取得尽可能多的高质量统计信息。

前面我们曾讲过提供充分而必要的统计资料。经济性原则除了体现在“节约”两字上，还体现在“必要”两字上。

在设计统计活动方案时，首先要确保满足统计宿体的管理决策的需求，提供充分的统计资料，非如此就不能作出合理的决策。然而在另一方面，也要力求所提供的统计资料对于统计宿体的管理决策是必要的。不必要的统计资料是对统计活动资源的浪费。这是因为，在一定的社会经济和科学技术条件限制下，在特定的时期和特定的地域内，过量的、或精确度过高的统计资料在决策中派不上用场，不能产生收益，因而是多余的。这就好比将高精度的计量器具提供给农民称量化肥，以便向大田里施用一样，超过普通计重称的精度部分对大田施肥不起任何作用。

第二节 统计指标的设计

1.2.1 统计指标的概念

统计研究客观总合现象的数量方面,而现象的数量特征要由相应的含有可计量内容的名词来提示,这样的名词在统计学中就称为指标。确切地说,统计指标,简称指标(**indicator**)是说明客观总合现象数量特征的范畴。例如,可以用“国民生产总值”这个范畴说明一个国家或地区国民经济活动总量规模,可以用“每万人中大学生数”这个范畴说明一个国家或地区高等教育的发达程度,等等。表明指标计量结果的数值称为指标值。在实际运用时,统计指标常常连同其指标值一起出现。在理论研究和统计设计中,只涉及到指标。

一个完整的指标由以下五个要素构成:(1)指标名称及其涵义;(2)时间;(3)空间;(4)计量单位;(5)计算方法。指标名称即是所使用的名词。许多指标及其涵义源出于关于他们要说明的现象的实质性科学,例如上边举的“国民生产总值”指标,即来源于经济学。另外一些指标出自经验和常识,无须有实质性科学的理论根据,例如上边举的“每万人中大学生数”指标,即属于后一种情况。客观总合现象总处于特定的时间、空间之中,因此统计指标又要和时间、空间限定词联系在一起。计量单位对于描述现象数量特征的重要性是显而易见的,计算方法更不容忽视。统计指标值之间的比较,只能是五要素中仅空间一项不同,或仅时间一项不同时的比较。通常还有“统计口径”的提法,它指的是指标涵义和计算方法两项。

统计指标有下述三个特点:

(1)可量性。统计指标作为说明客观总合现象数量特征的范畴,必须是可以测度的,即具有可操作的特点,这在设计指标时要十分注意。

(2)总体性。统计指标说明的不是个体,而是总体。它是对个体的标志表现和标志值进行数量综合的结果。

(3)具体性。统计指标和指标值总是联系于具体现象的,不是抽象的概念和抽象的数值。因此,在计算和使用统计指标和指标值,一定要把它们放到具体的客观环境和场景之中去对待,而不能用纯数量观点来对待。

1.2.2 统计指标的类别

从不同的角度,按不同的标准,可以对统计指标作不同的分类。

(一)总量指标、相对指标和平均指标

按指标的作用和表现形式不同,统计指标可以划分为总量指标、相对指标和平均指标。总量指标,又称绝对数(**absolute value**),是反映总合现象规模的指标,它是在同质前提下,对个体现象的品质特征或数量特征综合计量得到的。相对指标,又称相对数(**relative**),是反映现象对比关系的指标。平均指标,又称平均数(**average**),是反映同质总体内各单位某一数量标志(或标志间数量关系)的一般水平的指标。相对指标和平均指标都属于派生指标,一般由两个有联系的总量指标导出,也有的是由两个有联系的派生指标(或一个派生指标与一个总量指标)导出。相对指标和平均指标都是各自一类指标的统称,由于导出形式不同,各包括一系列的具体指标。本书第三章、第四章和第六章将安排专门内容阐述这两类派生指标的意义、计算方法和作用。

总量指标反映的现象带有不同的时间性质,因而它可以进一步分成时点指标和时期指标。描述瞬时现象数量特点的总量指标是时点指标,它表示在特定时间点上总体的静态规模,如第四次全国人口普查得到的某省人口总数描述人口数这一瞬时现象在普查标准时间——1990年7月1日零时在这一省份所达到的总量。描述持续现象动态规模的总量指标是时期指标,它表示在特定时间间隔之内总体数量的发展变化结果,如第四次全国人口普查取得