

社会调查中 样本容量的确定

耿修林 著



科学出版社
www.sciencep.com

C915/11

2008

社会调查中样本容量的确定

耿修林 著

科学出版社

北 京

内 容 简 介

本书利用势函数规则和统计决策的基本思想,对抽样推断时的样本规模进行了讨论,给出了相应的计算公式,并编制了样本容量的查对表。全书内容包括:样本容量确定的目标及影响因素,常用统计量的抽样分布,总体均值推断时的样本容量,总体成数推断时的样本容量,总体方差推断时的样本容量,方差分析时的样本容量,相关与回归分析时的样本容量等。本书的主要特点:不仅系统讨论了样本容量确定的一般原理,还编制了样本容量的查对表,实际应用时只需查表,就能大致确定应抽取的样本观察数目。

本书可作为统计学专业的教学教辅用书,也可供从事社会经济调查的有关人员参考阅读。

图书在版编目(CIP)数据

社会调查中样本容量的确定/耿修林著. —北京:科学出版社,2008

ISBN 978-7-03-021399-0

I. 社… II. 耿… III. 社会调查-样本-研究 IV. C915

中国版本图书馆 CIP 数据核字(2008)第 035689 号

责任编辑:孙 芳 王志欣 / 责任校对:陈丽珠

责任印制:刘士平 / 封面设计:耕者设计工作室

科 学 出 版 社 出 版

北京东黄城根北街16号

邮政编码:100717

<http://www.sciencep.com>

新 蕾 印 刷 厂 印 刷

科学出版社发行 各地新华书店经销

*

2008年5月第 一 版 开本:B5(720×1000)

2008年5月第一次印刷 印张:14

印数:1—3 000 字数:272 000

定价:38.00 元

(如有印装质量问题,我社负责调换〈明辉〉)

前 言

在教学工作中,经常会遇到应该发放多少份市场调查问卷比较合适这样的问题。抽样调查概念已为大多数人所熟悉,并且成为人们开展实证分析采集第一手资料的重要手段,但要具体弄清楚发放多少份问卷,即样本观察的规模,却不是一件简单的事情。直到现在,仍然有很多人重视样本调查单位数目的确定,有人认为是样本调查规模越大就越能印证研究的结论,越能让人信服研究结果真实性,所以在我们见到的研究论文和研究报告中,大多数都会指出样本调查规模非常之大。也有人在开展实证研究的时候,不事先评估现有的样本数据是否能达到研究需要的最低信息量,而是把数据资料不充分的责任归结到某种无法改变的客观原因上。既然是科学研究,就不能仅仅注重大体,同时也要重视细节。更何况,在开展调研和实证分析的时候,样本容量还不是一个细节问题,确切地讲,它应该属于统计调查研究过程中一个必不可少的重要环节,从某种程度上看,样本观察规模的大小最终会影响到我们研究的效果和决策分析的经济性。

样本容量是按照某种规则从总体中抽取出来的样本观察单位的数目。在进行统计推断分析的时候,究竟抽取多大规模的样本是个需要认真对待的问题。如果样本容量过大,会急剧增加调查费用,并且也有违抽样调查的初衷;当然,样本容量也不能过小,否则容易造成样本推断精度和可靠性降低,无助于揭示客观存在的统计规律性。正因为如此,在推断统计体系创立之初,人们就逐步明确地意识到需要解决样本观察的规模问题。只不过由于这个问题同数理统计中的重大命题相比不具有挑战性,因而在理论研究中没有得到应有的重视。

本书在综合考察多年来有关研究成果的基础上,对样本容量的确定进行了探讨,其主要目的在于:

- (1) 系统总结样本容量确定的一般原理。
- (2) 在综合考虑影响样本容量大小相关因素的前提下,给出符合统计分析要求、能有效达到统计认识目的的样本容量的计算公式。
- (3) 依据样本容量确定的原理,编制出样本容量查对表,以便能给实际工作者提供必要的方便。

全书共分 12 章,主要包括:样本容量确定的目标和样本容量确定的影响因素,样本统计量的抽样分布,总体均值估计时样本容量的确定,总体均值检验时样本容量的确定,总体成数估计时样本容量的确定,总体成数假设检验时的样本容量确定,决策规则下总体均值推断时的样本容量确定,总体方差估计时的样本容量

确定,总体方差假设检验时的样本容量确定,方差分析时的样本容量,相关与回归分析时的样本容量等。

由于作者水平有限,不妥之处在所难免,敬请读者批评指教。

作 者

2008 年新春于南京大学

目 录

前言

第一章 绪论	1
第一节 研究背景	1
第二节 发展回顾	4
第三节 本书体系	8
第二章 影响样本容量大小的因素	10
第一节 几个概念	10
第二节 样本容量确定的目标	14
第三节 影响样本容量大小的因素分析	18
第三章 抽样分布	24
第一节 正态分布	24
第二节 由正态分布导出的分布	26
第三节 常用的抽样分布	29
第四章 均值估计时样本容量的确定	37
第一节 引言	37
第二节 单总体均值估计时的样本容量	39
第三节 两总体均值差估计时的样本容量	43
第四节 均值估计时样本容量查对表	46
第五章 均值检验时样本容量的确定	49
第一节 假设检验的两类错误	49
第二节 单总体均值假设检验时的样本容量	51
第三节 两总体均值差假设检验时的样本容量	55
第四节 均值检验时样本容量查对表	57
第六章 成数估计时样本容量的确定	59
第一节 引言	59
第二节 单总体成数估计时的样本容量	62
第三节 两总体成数差估计时的样本容量	65
第七章 成数检验时样本容量的确定	70
第一节 单总体成数检验时的样本容量	70
第二节 两总体成数差检验时的样本容量	76

第三节 多个总体成数检验时的样本容量	78
第八章 统计决策规则下均值推断时的样本容量	81
第一节 综述	81
第二节 决策规则均值估计时的样本容量	83
第三节 决策规则均值检验时的样本容量	89
第九章 方差估计时样本容量的确定	98
第一节 方差的意义及其估计	98
第二节 单总体方差估计时的样本容量	103
第三节 两总体方差比估计时的样本容量	106
第十章 方差检验时样本容量的确定	110
第一节 引言	110
第二节 单总体方差检验时的样本容量	111
第三节 两总体方差比检验时的样本容量	115
第十一章 方差分析时样本容量的确定	120
第一节 方差分析的一般原理	120
第二节 单因素方差分析时的样本容量	122
第三节 双因素方差分析时的样本容量	127
第十二章 相关与回归分析时样本容量的确定	135
第一节 引言	135
第二节 相关分析时的样本容量	136
第三节 一元线性回归分析时的样本容量	139
参考文献	149
附录 A 均值估计时样本容量查对表	151
A-1 置信水平 99%	151
A-2 置信水平 95%	159
A-3 置信水平 90%	167
附录 B 均值检验时样本容量查对表	175
B-1 $\alpha=0.01, \beta=0.10$ 时假设检验样本容量	175
B-2 $\alpha=0.05, \beta=0.10$ 时假设检验样本容量	183
B-3 $\alpha=0.10, \beta=0.10$ 时假设检验样本容量	191
附录 C 成数估计时样本容量查对表	199
C-1 置信水平 99%	199
C-2 置信水平 95%	203
C-3 置信水平 90%	207
附录 D 成数检验时样本容量查对表	212

D-1	$\alpha=0.10, \beta=0.10$ 单总体成数单边检验	212
D-2	$\alpha=0.05, \beta=0.10$ 单总体成数单边检验	213
D-3	$\alpha=0.01, \beta=0.10$ 单总体成数单边检验	215

第一章 绪 论

第一节 研究背景

辩证唯物主义世界观告诉我们,任何事物都是由质量和数量两方面组成的,是二者的矛盾统一体。因此,对社会经济现象开展研究,既需要从质量方面出发,也需要从数量方面进行探讨,以对事物的质的研究为起点,辅之以数量分析,并将它们有机地结合起来。只有这样,才能更进一步地提高我们对社会经济问题的认识能力和认识水平。

社会科学中,传统的研究手段主要是诉诸于定性分析,从学科的性质特征来看,这本身没有什么不对的地方,但是单纯依靠定性分析,往往只能形成非常原则的结论,随着社会关系越来越复杂,一味依赖于定性分析很难能把研究的课题深化下去。例如,增加全社会投资会促进国民经济增长,可是在一定条件下,能够用于追加投资的资源是有限的,那么,将这些有限的投资资金用于哪些方面才能产生最大限度的乘数效应? 税收是宏观经济调控的重要杠杆,税率提高会压制社会消费水平,避免经济出现过热现象,税率降低会促进消费,从而通过消费拉动经济增长。现在的问题是,税率怎样变化才能保证整个社会经济稳定有序的发展? 产业结构的提升是我国近年来一直关注的课题,产业结构的演变有着自身的规律,国民经济各行业之间存在着关联效应,那么在现有的社会经济状态下,如何选择合适的具有所谓带动效应的产业,才能更好更快地促进国民经济整体素质的提高? 在市场营销活动中,我们经常面临着目标客户的确定问题,对此,如果能够根据过去的销售记录,通过分析找到重点客户,就非常有利于制定出更具针对性的营销策略。诸如此类的问题,如果仅一般性地给出一个方向性的建议,是很难让人感到满意的。

“对于一种科学,只有在成功地应用数学的时候,才算达到真正完善的地步。任何一门科学,只有当它真正与数学联系起来,它才算真正发展起来。”在社会科学的研究中,如果能积极借助于定量分析,可以达到以下几个方面的目的。

(1) 通过引进数量研究方法,能够在一定程度上改变社会科学的研究面貌。在当今这个时代,对于社会科学的研究活动,不应该仍然一味以文献资料加工为基础,而应该主动面向现实,以丰富、具体的事实资料为依据,通过对数据的探索性发掘,找到社会经济现象发展变化的规律和本质特征。

(2) 通过增加数量研究的成分,能够改变社会科学工作者的身份,提高社会科

学研究人员价值。社会科学工作者不能仅仅成为一般的研究型人员和专家,也应该像工厂的专业管理人员和工程技术人员一样,努力成为社会进步事业的“工程师”。

(3) 通过数量分析方法,能够拓宽社会科学研究渠道,丰富社会科学研究内容。因为对所研究对象的有关质的特征和关系,如果诉诸于定量的、精确化的描述,有可能会促进新的研究课题的发现,至于这一点,很容易从科学发展史中找到依据。

(4) 通过数量分析方法,能够为经济政策和管理措施的制定提供“实验室式”的论证与检验,这对于增强政策措施实际执行效果的预见性、排除主观认识误区大有裨益。

(5) 通过数量分析方法,能够提高研究结论的精确化水平。社会科学虽然不同于自然科学,但是适当地运用数量分析还是十分必要的。例如,商品价格与商品销售量存在一定的联系,价格上涨,销售量将趋于下降;反之,价格降低,需求量会增加。如果有该商品的价格与对应的销售量资料,然后在价格与销售量之间建立分析模型,便能进一步了解到价格变动对销售量边际影响或弹性效应的具体数值,这对于合理确定价格促销策略很有好处。

总之,在社会科学研究中适当进行数量分析无疑是有意义的。要实现从数量角度认识和研究问题的目的,掌握必要的资料是其基本条件。因为数量分析的基础,尤其是实证分析的场合,大量的数据资料是开展研究的主要“原料”。只有拥有准确、可靠、系统、完备的数据,才有可能获得客观的研究结论。

从形式上看,获取数据资料的途径不外乎有直接来源和间接来源。间接来源的数据资料主要是“第二手资料”,那么,凡是获取“第二手资料”的途径或来源就是所谓的间接来源。例如,从统计年鉴、统计摘要、统计资料汇编、互联网,以及报纸、杂志、新闻报道等渠道获得的资料,都属于间接来源的数据资料。直接来源的数据资料是指为了某一研究任务的需要,专门组织搜集的“第一手资料”。获得第一手资料的途径或方式有很多,其中,最主要的是统计调查和统计实验。间接来源的资料归根结底也是来自于统计调查或统计实验,也可以说,统计调查和统计实验是一切数据资料的根本源泉,它们共同构成统计资料来源的两大基本手段。

统计调查根据涉及的范围大小,可分为全面调查和非全面调查。全面调查有普查、全面统计报表等形式,非全面调查是指对总体部分单位进行的调查,包括随机抽样调查、判断样本调查、任意样本调查、定额样本调查等。全面调查虽然能取得全面的调查资料,但由于范围广、工作量大、需要人手多、时间长、费用高等原因,因而在实际应用中经常会受到一些限制。判断样本调查、任意样本调查、定额样本调查,由于样本观察单位的确定受到人为偏好的影响,如果处理不好,其代表性往往较差。

抽样调查是社会经济统计中用于搜集资料的重要手段,由于这种方法具有普查不可替代的优势,比如需要人手少,资料采集工作量不大,能够节约调查费用,节省资料搜集和处理时间,使用形式灵活多样等,从而成为社会经济统计领域获取数据资料的极为常用的组织方式。例如,人口调查、居民家庭收支状况调查、产品加工过程中的抽样检测、资源环境调查、疾病发生情况调查、一批货物的抽样接收检查、企业管理中的市场行情调查、社会舆论调查等,一般都采用抽样调查的方式。抽样调查不仅仅是搜集资料的方法,在现代统计方法体系中,它也是一种对统计总体进行认识的途径。从研究对象中按照某种规则抽取一定数量的总体单位组成样本,通过对样本实施观察得到样本数据,然后通过一定的方法对样本数据进行处理,最后上升到对统计总体的认识,可以说这是现代统计学的基本研究逻辑。这一过程,我们可以借助图 1.1 来说明。

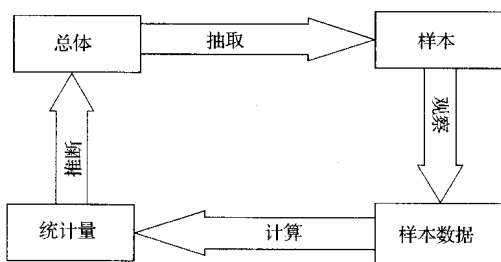


图 1.1 样本推断的研究逻辑

在上述图示中,每个环节都有值得重视的统计研究课题。如何从总体中抽取样本,按照什么样的规则抽取样本观察单位,这些都是一般抽样调查理论着力讨论和解决的问题。不同的资料采集方式,得到的回收率以及数据的可靠程度往往不一样,因此对产生出来的样本,如何选择合适的方法搜集资料,是采用当面访谈,还是邮寄报表,这应该成为样本数据采集方式重点关注的问题。对搜集来的样本数据,如何进行加工整理,从而提炼出有用的认识信息,这是统计分析环节需要做的工作。在这一过程中,怎样根据研究问题的需要、资料表现出来的特征,以及统计方法的适应条件,确定合适的统计分析、对比和计算的工具体或模型,是需要认真考虑的。样本本身不是目的,直观地理解,样本只不过是帮助我们实现对统计总体认识的途径,在统计活动的完整过程中,充当着过渡或桥梁的作用,所以在获得样本信息和认识之后,还需要继续上升到对统计总体的推断,以便能形成更为一般性的统计研究结论。

在从总体抽取样本的过程中,抽取多大规模的样本比较合适,这也是在抽样设计阶段必须给出科学回答的重要议题。我们知道,如果样本观察数目过小,由于样本信息不充分,由此会导致统计分析结果的稳定性差,而且也无助于揭示和掌握客观现象之间的统计关系。当然,样本观察数目也不能过大,如果盲目要求样本观察越多越好,一方面会增加不必要的数据搜集成本,另一方面也可能会造成数据信息上的冲突和干扰,反而不利于模拟出统计规律性。

所谓样本容量,是指按照某种规则从总体中抽取出来的样本观察单位的数目。正如前面所述,在进行抽样推断统计分析的时候,究竟抽取多大规模的样本是个需

要认真考虑的问题。样本容量过小,虽然能够节约调查费用和后续的数据分析成本,但容易造成推断精度降低,统计推断结果的概率把握程度下降。反之,样本容量过大,虽然会提高抽样估计的精度和分析结果的概率保证程度,但与此同时会急剧增加调查费用。

通过对样本容量确定问题的研究,主要目的在于:

- (1) 系统研究样本调查和统计推断中各种情况下的样本容量的确定原理。
- (2) 在综合考虑影响样本容量大小的相关因素的前提下,给出符合统计分析要求、能够有效达到统计认识目的的样本容量。
- (3) 编制样本容量查对表,给开展实际应用研究提供必需的参考。

第二节 发展回顾

在统计理论发展的过程中,人们很早就注意到了如何通过有限的样本观察来实现统计分析和认识功能的问题,如统计实验设计方法的兴起,主要就是要达到以有限的实验样本数据实现统计推断的目的。原则上讲,运用实验设计确定样本,这些方法在自然科学和工程技术领域应用比较有效,联系到社会经济领域,由于不具备统计实验所必需的条件,如果完全依据实验设计的办法确定最优样本容量,操作起来几乎没有多大的可能。由于没有一个客观的数量上的参照依据,在实际运用抽样的时候,有关样本容量的确定多半是比较随意的,或者是经验之谈。例如,样本容量不低于 30,最好能达到 50 甚至 100 以上,把能够搜集到的有用数据都找到等。

样本容量的确定问题是伴随着统计思想和哲学观的转变而逐渐引起人们重视的,从作为一门科学的统计学诞生的 17 世纪中叶起,经过后来的 Süssmilch 明确倡导大数法则,到 19 世纪中期,在将近 200 多年的时间里,由于大数法则在统计世界观中居于支配性质的地位,人们在从事统计活动时必做大量观察,一切从大样本出发,因此在这一期间,样本容量的确定可以说没有得到应有的重视。

样本容量确定问题的提出与科学的社会调查和作为统计学的一个分支学科——抽样调查的兴起有着非常密切的关系。1861 年,英国人 Farr 为估算当时英国人口的总数,经过慎重分析选取了 14 个地区作为样本,得到了后来被证实较为准确的英国人口数字。此外,Farr 还通过随机取样的办法,以 10 个随机抽取的样本地区,对全国的雇工人数做了估算。1878 年,俄国统计委员会主席塞廖诺夫在里亚赞州的穆拉耶文地区开展实地调查,按富裕程度对该地区居民家庭先进行详细的分类,然后再从各个类别中抽取部分居民家庭进行调查,当时该地区的实际居民家庭总户数为 800 多户,他确定了 11 户作为重点调查对象,样本户占比约为 1%。从自身的实践中,塞廖诺夫认为样本调查必须要同分类相结合,才能以较

少的样本观察得到较好的结果。抽样调查实践的先驱者之一——美国人 Wright 在 1873 年利用 400 个志愿者的报告,对当时纽约的失业人数进行了估计。1879 年,英国人 Lawes 选取 5 块地段的小麦收成,对英格兰的小麦产品进行了估计。20 世纪 20 年代初期,美国统计学家 Croxton 等人在俄亥俄州哥伦布市按 10% 的抽样比对该地区的从业人员进行了访问调查。1923 年,英国人 Hilton 从参加社会失业保险工人的个人档案中抽选 10000 个失业救济金申请者进行调查,他认为在一般的选样调查中,只需要 1% 左右的样本基本上就可以了,并指出按照这样的比例确定调查单位,既能达到降低调查费用的目的,同时也能满足行政管理和政策制定的需要。1935 年,美国人 Gallup 创办了美国民意调查研究所,开展了卓有成效的社会舆论抽样调查活动。Gallup 认为,进行抽样调查,对调查对象实行细致的分类是必要的,并且各层应抽取的样本数要与该层占总体单位数目保持某种合适的比例,样本调查总数要考虑调查误差的要求和调查经费的预算。1936 年, Gallup 正式出版了 *Public Opinion Poll*, 书中详细叙述了社会调查的原理和方法。受 Gallup 的影响,此后西方的社会舆论调查基本上都或多或少采用了 Gallup 民意测验的做法,并且还总结出一套调查样本确定的经验规则,即保证调查相对误差不超过 3% 时,至少需要 2000 左右单位的调查样本,如果调查允许误差为 10%, 则只需 200 个左右样本单位。

在样本推断理论发展过程中,一大批数理统计学家做出了重要贡献。法国数学家 Laplace 不仅是古典概率研究的集大成者,同时也在推动概率论与统计研究相结合方面取得了重要的成就,另外,在统计调查领域也开拓了新的思路。1786 年, Laplace 抽选了法国 30 个地区,根据这些地区的人口数和出生率,估算全法国的人口总数,并对估算误差进行了分析。挪威统计局长 Kiaer 是抽样调查理论和实践的代表人物,1895 年,他和 Hanssen 等人分别在城市抽取 1.6%、乡村抽取 3.3% 的成年男性进行调查,以估算挪威成年男性的收入状况。Kiaer 认为:“调查结果的准确性与观察数量的多少没有多大关系,关键是要保证调查样本的代表性。”英国统计学家 Bowley 的卓越研究把 Kiaer 的代表性抽样思想发展到一个新的阶段,他从理论上论证了普查不总是必要的,认为小样本在很多方面就能满足调查研究的需要。1906 年, Bowley 在一次会议上指出:“通过抽样方法可以得到想要的结果,而且只需要很小的样本就可能达到目的,只不过如何保证每个事物都有同等的机会抽取才是问题的关键。”1923 年,苏联统计学家 Tschuprow 系统研究了代表性抽样问题,并给出了层间样本最优分配的计算方法。Neyman 是自 Fisher 之后统计学界的伟大代表,1934 年, Neyman 在《皇家统计学会会刊》上发表了一篇题为 *On the two different aspects of the representative method* 的论文,在这篇文章中, Neyman 认为调查费用和样本调查误差是一对矛盾,在进行样本调查时,需要考虑如何在给定的费用条件下寻求最大的样本估计精度,并给出了层间样本

分配的 Neyman 公式。从 20 世纪 40 年代开始, Neyman 和 Pearson 经过十多年的友好合作,共同完善了频率统计的“区间估计”和“假设检验”理论,同时也为“经典抽样调查”样本容量的确定奠定了坚实的基础。现在,“势函数”规则下的样本容量的确定都是以 Neyman 和 Pearson 的思想为依据的。印度大统计学家 Mahalanobis 也是一位对统计学发展有着众多贡献的人,在抽样调查理论中,他在 20 世纪 40 年代就提出了应该建立调查费用函数,根据调查对象的实际情况,确定最合适、最经济的调查样本。

我们知道,抽样理论主要研究的内容包括样本分布、抽取样本的方式、抽样的估计效果、抽样调查的经济性等。抽样研究的问题虽然很多,但无论如何总有一个环节需要给出样本含量大小的明确回答。所以,人们在发展抽样理论的同时,也在理论和实践上自觉或不自觉的关注着样本容量的确定。随着抽样理论在统计学体系中的地位的确立,有关样本容量确定的研究也随之得到一定的发展。

自 19 世纪末、20 世纪初期开始,建立在 Gosset 的 t 分布“小样本”理论基础上的统计思想逐渐成为现代统计学研究的主流,于是,样本观察的大小问题便被作为统计理论研究的一个附属命题提了出来。在一般的统计学著作中,都会提到“经典”抽样中的样本容量问题,并会给出最基本的基于势函数规则的计算公式。为进一步解决 Gosset“小样本”推断的实现问题,统计学家 Fisher 创立了“方差分析”、“实验设计”这两个统计学研究的分支学科,从理论上探讨和研究实验样本的安排问题,一方面是希望用最小的样本观察规模达到统计研究的目的,另一方面就是在进行观察单位选择的时候避免可能带来的抽样偏差。1947 年, Wald 出版 *Sequential Analysis* 一书,系统阐述了“序贯概率比检验”思想。Wald 认为,在产品抽样检验中,要想保证达到检验的目的,同时又要尽可能减少抽样规模,就不应该在抽样检查之前就事先确定下样本数量;相反,应该根据抽样检查过程中得到的信息,逐次增加抽样检测量,直至能做出判断为止。1950 年, Wald 出版了 *Statistical Decision Functions*,他是想把 Neyman 和 Pearson 的区间估计和假设理论,以及他自己系统发展的序贯分析理论,共同纳入到一个统一的数理逻辑体系中,他提出的损失函数、风险函数等范畴,对样本容量的确定有着重要的指导意义。

以上所述可以说是统计样本确定问题研究的大致历程,这些研究对丰富统计学的理论体系起到了重要的推动作用。抽样调查样本容量确定问题的专门、系统的研究出现在 20 世纪 50~60 年代,其中,最有代表性的成果是 Mace 出版于 1964 年的 *Sample Size Determination*。在这本书中, Mace 介绍和讨论了统计估计、假设检验、统计选择、产品抽样检查,以及序贯抽样时样本容量确定的一般性原理。此后,有关样本容量的系统研究成果鲜有问世。原因可能是样本容量的确定在整个统计理论研究中不占主导地位,所以对从事理论统计研究的人来说,可能挑战性不大,因而没有过多的兴趣,再就是这个问题的研究,正如 Lenth 所讲,“虽然样本

容量的确定是统计研究中的重要环节,但要解决这个问题确实存在着一些障碍不好处理”。

样本容量问题的提出,同统计归纳推断研究方法的形成和抽样调查的兴起有着非常密切的关系。通过对样本容量确定研究的简单回溯,我们可以看出:作为统计归纳推断和抽样调查一个子课题的样本容量的研究,大致经历了从意义不明确的阶段到成为一个专门研究问题的变化过程,其发展历程归纳起来主要有以下 3 个阶段。

(1) 早期阶段。样本观察的大小在早期阶段其基本概念是不清晰的,一般都是在社会问题研究、人口问题调查、经济活动领域为某种目的,结合调查现象本身的特点,在很大程度上是凭经验和感觉选取调查单位的数量。

(2) 经典抽样理论兴起阶段。在这一阶段,样本容量的讨论多半是随着现代统计学的理论发展而不断得到重视。特别是小样本推断理论的确定,直接促进了通过实验设计获取最小的有效样本的研究。到 20 世纪 40~50 年代,伴随着频率统计推断两大内容,即区间估计与假设检验的理论完善,以及 Wald 统计决策思想的诞生,为从势效果和决策角度研究样本容量提供了重要的工具。统计学科体系的丰富,新的统计研究分支的出现,也为样本容量问题的研究开辟了崭新的领域,如多变量统计、非参数统计、现代时间序列分析等的兴起,使不少统计学家在从事主要研究的同时,也不时注意解决抽样规模的问题。

(3) 标准抽样理论兴起阶段。标准抽样理论中的抽样同数理统计学中经常提到的抽样理论是有一定的差别的。数理统计学中讲到的抽样主要有以下几个方面的特征。

① 总是假定样本所来自的总体具有某种特定的分布形式,如二项分布、泊松分布、正态分布等,可是在标准的抽样调查理论中,一般不做这样的假定,或者最多只是知道有关样本来自总体的有限一点的信息,换句话说,数理统计中,抽样是有固定的分布模型的,而标准的抽样调查理论中的抽样是分布自由的。

② 大多指的是简单随机抽样,因为这种抽样方式最符合统计随机性要求,对讨论统计推断的理论方法相对来说比较容易,对比起来,标准抽样调查理论中的抽样采用的抽样组织形式要复杂得多,除简单随机抽样之外,还包括分层抽样、等距抽样、整群抽样以及多阶段抽样等。

③ 数理统计中的抽样总体,既可以是有限总体,也可以是无限总体,因而讨论的统计估计方法具有普遍适用性,可是抽样调查理论面对的总体主要是有限总体。

④ 数理统计中的抽样不要求利用辅助性信息,但在抽样调查理论中,为防止样本发生整体性偏差,同时也是为了调查组织工作的便利,在确定具体的样本单位的时候,往往强调经验性改进。

样本容量确定问题的广泛关注,原则上讲主要得自于标准抽样调查理论和实

实践的发展。因为在抽样调查中,人们不仅关心抽样调查的统计效果,如估计精度、概率把握程度,同时也比较关注抽样的经济效果,包括直接的调查费用预算和潜在的机会损失。这样一来,就涉及最优的样本大小问题。

按照研究主要发生和存在的领域,可以把样本容量确定问题的研究大致划分为以下4个领域。

(1) 社会调查领域。这一领域的样本容量研究可分为前后两个阶段,前期阶段是抽样调查得到确立前的时期,其样本容量一般都不是严格意义上的概念,更多地可能类似于今天所说的“典型调查”、“重点调查”和放大的“个案调查”。后一阶段可以认为是抽样调查在社会问题研究中的应用,如人口抽样调查时的抽样比等。

(2) 数理统计领域。与数理统计理论发展相随的抽样形成了经典的抽样样本容量课题。这些方面的研究主要是回答从总体中应该抽取多大规模的样本,才能保证必要的估计精度和检验效果。当然,也包括其他方面的意图,因为不解决抽样方式和样本大小,可能无助于相应理论研究的完善。

(3) 标准抽样调查领域。样本容量确定在标准抽样调查领域受到的重视无须多说,因为这是开展有效抽样调查方案设计必须要回答的问题。

(4) 生物遗传和医疗卫生领域。在遗传研究和医疗卫生活动中,由于天然不具备大量观察的条件,从而迫使人们更加注意如何通过对实验的安排获取有价值的分析数据。从对样本容量确定问题研究发表的论文数量来看,针对生物遗传和医疗卫生的要远远多于经典抽样和标准抽样领域。

第三节 本书体系

本书共有十二章组成,主要章节安排及其内容简介如下。

第二章着重对样本容量确定的目标以及影响样本容量大小的因素进行了分析。样本容量确定要实现的目标表现在:适度的抽样规模,保证应有的抽样估计精度和检验效果,对样本推断的概率把握程度,尽可能降低机会损失,不超过预算的抽样开支。对样本容量确定的影响因素主要讨论了调查对象内部的变异程度,估计精度的要求,总体分布是否确知,调查费用预算,样本抽取的方式,数据资料的测量水平等。

第三章介绍抽样分布。随机变量的理论概率分布及其衍生出来的样本统计量的抽样分布不仅是统计推断的基础,同时也是确定样本容量的重要条件,特别是在势函数规则下进行样本容量的确定更是如此。为此,本章介绍了由正态分布导出的三大分布,即 χ^2 分布、 t 分布、 F 分布的性质,以及常用的样本统计量的抽样分布,包括样本均值的抽样分布、样本方差的抽样分布等。

第四章讨论了总体均值估计时样本容量的确定。本章假定样本来自于正态总体,并依据势函数规则,对单总体均值估计和两总体均值差估计时的样本容量计算进行了讨论。另外,还对相关条件下均值估计时样本容量的大小给出了相应的查对表,以便实际运用时可以通过查表直接找出必要的样本观察数目。

第五章对单总体均值假设检验和两总体均值差假设检验时的样本容量大小进行了讨论,并且区分了单边假设检验和双边假设检验等情况。最后,也给出了样本容量的查对表。

第六章讨论了总体服从二项分布,样本成数统计量假定在大样本下服从正态分布,然后根据势函数的一般规则,分别导出单总体成数估计和两总体成数差估计时样本容量的计算公式,并根据这些公式给出了总体成数估计时的样本容量的查对表。

第七章依据势函数规则对单总体成数假设检验和两总体成数差假设检验时的样本容量确定进行了分析,并给出了相应的计算公式。根据这些公式,编制了总体成数假设检验时的样本容量的查对表。

第八章根据决策规则对总体均值样本推断时的样本容量进行了讨论,主要包括单总体均值和两总体均值差样本推断时的样本容量确定,单总体成数和两总体成数差样本推断时的样本容量确定。

第九章假定样本来自于正态总体,然后根据势函数方法,对单总体方差估计和两总体方差比估计时的样本容量确定问题进行了讨论,给出了相应的样本容量的计算公式。

第十章着重对正态总体方差假设检验时的样本容量进行了分析,并分别对单总体方差假设检验和两总体方差比假设检验问题给出了各自的样本容量的计算公式。

第十一章对单因素重复试验和两因素重复试验情况下的样本观察次数进行了说明。

第十二章介绍了相关与回归分析样本容量的确定。相关与回归是常用的分析工具,本章对相关分析和回归分析时的样本容量进行了讨论,包括相关系数估计和检验时的样本容量,一元线性回归分析时的样本容量。