新世纪统计与经济优秀著作文库

国家社会科学基金资助项目
教育部新世纪优秀人才支持计划项目
第十届全国统计科学研究优秀成果一等奖

统计调查质量的测度与评估方法

傅德印　黄恒君　著

中国统计出版社
China Statistics Press

新世纪统计与经济优秀著作文库

国家社会科学基金资助项目
教育部新世纪优秀人才支持计划项目
第十届全国统计科学研究优秀成果一等奖

统计调查质量的测度与评估方法

傅德印　黄恒君　著

中国统计出版社
China Statistics Press

(京)新登字 041 号

图书在版编目(CIP)数据

统计调查质量的测度与评估方法 / 傅德印，黄恒君著. —北京 ：中国统计出版社，2011.12

ISBN 978—7—5037—6448—6

Ⅰ. ①统… Ⅱ. ①傅… ②黄… Ⅲ. ①统计调查—方法 Ⅳ. ①C811

中国版本图书馆 CIP 数据核字(2011)第 264937 号

统计调查质量的测度与评估方法

作　　者/傅德印　黄恒君
责任编辑/梁　超　胡文红
装帧设计/杨　超
出版发行/中国统计出版社
通信地址/北京市西城区月坛南街 57 号　邮政编码/100826
办公地址/北京市丰台区西三环南路甲 6 号
网　　址/www. stats. gov. cn/tjshujia
电　　话/邮购(010)63376907　书店(010)68783172
印　　刷/河北天普润印刷厂
经　　销/新华书店
开　　本/710×1000mm　1/16
字　　数/330 千字
印　　张/19.25
版　　别/2011 年 12 月第 1 版
版　　次/2011 年 12 月第 1 次印刷
书　　号/ISBN 978—7—5037—6448—6/C.2601
定　　价/36.00 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。

前言

统计数据是政府科学决策、企业现代管理以及个人投资选择的重要参考依据。统计调查数据质量是统计工作的生命。随着我国经济社会的快速发展，数据用户对统计调查数据的内在质量以及数据产生、提供过程可靠性的要求越来越高，统计调查质量问题越来越引起各界的广泛关注。提高政府统计能力，提高统计数据质量，提高统计公信力，进一步推进统计工作规范统一、改革创新、公开透明，推进统计现代化及科学发展，也是国家统计局当前的重点工作。因此，结合我国政府统计实际，吸收国内外已有的研究成果，系统研究统计调查质量的测定与评估方法，对规范统计数据质量管理，推动统计科学研究及统计事业发展，无疑具有重要的理论意义和现实意义。

截至目前，关于统计调查质量测定与评估的研究，已经历了三个阶段。从20世纪初到30、40年代，研究主要是围绕着调查数据的准确性、样本的代表性进行的；20世纪50年代初到70年代，研究主要是围绕着统计调查误差模型的建立及由此展开的专项研究；20世纪70年代末至今，随着统计数据质量涵义的扩大，对统计调查质量的研究发展到如何对统计调查建立有效的保证、控制、评估体系上来。从当前的研究热点上看，对统计调查质量测定与评估的研究，已包括对调查误差的来源、分类以及误差模型的研究；统计调查设计质量与规范的研究；数据收集质量的研究；数据处理质量的研究；数据质量评估和测定的研究；调查误差对参数估计、统计分析的影响研究；统计调查质量与调查费用及调查效率的研究；统计调查质量改进中的问题与对策等方面。从统计调查管理上看，联合国统计局关于官方统计资料的八项要求、欧洲经济委员会官方统计的十项准则、加拿大统计局的“统计数据质量手册”、美国联邦统计委员会的“企业调查质量测定指标”，以及国际货币基金组织的“统计数据质量评估框架”等，这些都对统计调查质量测定和评估的原则、范围、方法、技术进行了理论和实践探索。

本项研究是在国内外有关统计调查质量研究的基础上，界定统计调查质量

内涵，将统计调查质量分解到统计数据质量和调查过程质量上来。基于此，再对统计调查质量的测度、评估、控制和改进方法进行系统研究，我们做了如下尝试和探索：

1. 按照调查质量维度，讨论了调查质量测度与评估的框架、思路和方法。

2. 按照人们对事物认识规律的顺序（分类、比较、判别、归纳），讨论了调查质量的识别方法。

3. 按照调查过程，讨论了调查质量测度与评估的框架、思路和方法。

4. 按照调查过程，讨论了调查质量管理与改进的框架、思路和方法。

当然，呈现在您面前的这项研究，只是我们对统计调查质量测度与评估这一课题进行探索的阶段性结果，有待理论和实践的进一步检验。本书未涉及的问题以及今后需要进一步研究的问题还有很多。

本书是国家社会科学基金资助项目《统计调查质量的测度与评估方法》（批准号：03BTJ008，获第十届全国统计科学研究优秀成果一等奖）、教育部新世纪优秀人才支持计划项目（2005年度）《政府统计数据质量控制技术方法研究》等研究成果的基础上完成的。课题组主持由傅德印负责，参加成员有：黄恒君博士、梁亚民教授、刘晓梅教授、陶然博士、瞿春玲博士、师国君副研究员。课题组成员发扬团队精神，紧密配合，团结协作，发表和出版了多项研究成果，本书的完成与课题组全体成员的辛勤劳动密不可分！

课题研究得到了国家社科规划办、教育部科学技术司、甘肃省社科规划办、兰州商学院科技处等单位专家的支持和帮助。为完成本项研究，我们参考了国内外同行的大量文献，没有这些同行的前期研究，我们的探索是难以完成的，为便于查阅，我们按章节将主要参考文献列于各章结尾。第十届全国统计科学研究优秀成果评审委员会的专家评委，对该项研究曾给予了高度评价，我们倍受鼓舞，同时也深感还有许多工作有待我们去做。本书的出版也得到了中国统计出版社以及杨映霜女士的支持和帮助，在此，一并向各位专家、教授、同行表示衷心的感谢！

虽然我们为本书的完成付出了很大努力，但限于能力和水平，一定存在很多缺点、不足甚至错误，敬请得到各位同行及读者的批评指正！

傅德印

2011年5月于金城

此为试读，需要完整PDF请访问：www.ertongbook.com

目 录

绪 论

第1章 统计调查质量研究的综述 3

1.1 国外对统计调查质量研究的综述 3
1.2 国内对统计调查质量研究的综述 8
1.3 本书出发点和结构安排 11
参考文献 15

第Ⅰ部分 统计调查质量测度与评估基本理论与方法

第2章 统计调查质量特性研究 21

2.1 统计调查质量的内涵界定 21
2.2 统计调查质量维度分析 24
2.3 统计调查质量维度分解的意义 29
参考文献 30

第3章 统计调查质量与调查误差 32

3.1 统计调查误差分类 32
3.2 统计调查误差的测度 38
3.3 统计调查过程中的误差指标 39

3.4 统计调查误差与信度、效度的关系 41
参考文献 44

第4章 统计调查质量成本与时效性 45

4.1 统计调查质量成本内涵的界定 45
4.2 统计调查质量成本体系结构及要素分解 47
4.3 统计调查的时效性 51
4.4 统计调查时效、成本与误差的关系 53
参考文献 54

第5章 统计调查质量控制评估的其他方法 55

5.1 系统分类评估方法 55
5.2 比较分析与探索性数据分析技术 59
5.3 判别分析模型在汇总数据质量控制的应用 74
5.4 二元选择模型在汇总数据质量分析中的应用 80
5.5 神经网络技术在汇总数据质量分析中的应用 85
参考文献 90

第Ⅱ部分 统计调查阶段质量测度与评估方法

第6章 统计调查设计质量测度与评估 93

6.1 数据收集模式质量评估 93
6.2 调查框质量测度与评估 95
6.3 问卷设计质量测度与评估 106
6.4 抽样设计质量测度与评估 116
6.5 调查设计阶段质量评估方法框架 124
参考文献 125

第7章 统计调查数据收集质量测度与评估 127

7.1 访问员质量因素测度与评估 127
7.2 被调查者质量因素测度与评估 134
7.3 无回答的测度与评估 144
7.4 数据收集阶段质量监控 157

参考文献 161

第 8 章 统计调查数据处理质量测度与评估 164

8.1 数据处理过程描述 165
8.2 数据捕获质量测度与评估 166
8.3 数据编码质量测度与评估 168
8.4 数据编辑和插补质量测度与评估 171
参考文献 178

第 9 章 统计调查数据分析质量测度与评估 180

9.1 估计量的加权调整 180
9.2 估计量的方差估计 192
参考文献 198

第 10 章 统计调查管理质量的评估——兼论统计调查质量保证体系 199

10.1 统计调查质量保证体系的基础 199
10.2 统计调查质量保证体系的原理 202
10.3 统计调查质量保证体系的结构 204
10.4 统计调查质量保证体系作业要素及控制 208
参考文献 213

第 11 章 统计调查质量改进的评估 215

11.1 统计调查质量改进的原理 215
11.2 统计调查持续质量改进评估框架 218
11.3 统计调查质量持续改进的讨论 223
参考文献 224

第 12 章 统计调查质量综合评估 226

12.1 调查误差综合评价 226
12.2 综合质量评估模型 231
12.3 调查过程中的质量文档 234
12.4 调查质量的非技术方面——数据虚假问题分析 240
参考文献 248

第Ⅲ部分　统计调查质量测度和评估应用研究

第 13 章　人口普查质量测度与评估——以我国人口普查质量控制与保证为例　253

13.1　我国人口普查质量控制技术体系　253

13.2　我国人口普查质量保证体系　265

参考文献　270

第 14 章　抽样调查质量测度与评估——以欧盟劳动力调查为例　271

14.1　欧盟劳动力调查概述　271

14.2　欧盟劳动力调查质量评估　272

参考文献　280

第 15 章　插补估算质量测度与评估——以英国年度商业调查(ABI)数据为例　284

15.1　欧盟编辑项目(EUREDIT)概述　284

15.2　英国年度商业调查数据估算质量评估　286

参考文献　290

附表　291

绪　论

此为试读,需要完整PDF请访问: www.ertongbook.com

第1章 统计调查质量研究的综述

质量是统计工作的生命，统计中的质量可归结为统计调查质量。调查(Survey)，在《统计学术语词典》中定义为从目标总体中获得数据[1]。由此，一些统计机构给予调查更清晰的定义[2]：包括普查、抽样调查、报表数据获取以及合成数据获取(模型估计或插补)。而“统计”至少包括统计工作、统计数据两层涵义。因此，统计调查质量可进一步扩展：统计调查作为生产过程，其质量一般的定义是“符合标准”[3]，即调查过程质量；作为统计生产的成果，质量则定义为“满足需求”[4]，即统计数据质量。在统计调查质量中，调查过程质量是必要条件，统计数据质量是最终表现，而准确性是调查质量的最朴素认识，对统计调查质量的研究历史表明，这种认识是从准确性→数据质量→调查质量展开的。本章分国内外两个部分，按照对调查质量认识的顺序，叙述调查质量的研究现状。

1.1 国外对统计调查质量研究的综述

1.1.1 国外关于统计调查质量研究的进程

对统计调查质量的研究是随着社会、经济、科技的发展对大规模的统计数据搜集系统的需求而产生并不断发展的。从20世纪初开始，人们着手研究调查数据误差，并进一步研究统计数据质量和统计调查质量。具体可划分为以下

几个阶段：

1.1.1.1　研究形成期(20 世纪初～20 世纪 40 年代)

该时期统计调查质量研究主要是围绕着调查数据的准确性、样本的代表性，沿着两条思路展开的。

一是对调查数据准确性与抽样误差的研究。20 世纪初期正值抽样调查与普查展开激烈争论的时期。为此围绕统计数据质量，人们讨论的是：①调查结果的准确性是取决于观察数量的多少，还是取决于抽样方法的问题；②采用代表性样本还是随机抽选样本的问题。在 Kiar[5]、Bowley[6]等人的努力下，确立了抽样调查的地位；而 Neyman 的“可测设计”，奠定了调查随机化框架的基础，是统计调查质量研究的一个重要里程碑。

二是围绕着统计数据准确性与非抽样误差。20 世纪初人们就注意到非抽样误差的影响。第一次世界大战后，实际工作者们在许多应用中进一步发现各种非抽样误差比抽样误差的影响更重要、更明显，这就激发了人们对非抽样误差的研究和讨论。而这种研究实际上是随着政府进行大规模统计数据收集而展开的，并在 30、40 年代取得了进展，包括美国普查局进行人口普查时对非抽样误差的研究，以及印度政府进行粮食产量估计时对非抽样误差的研究[7]。

1.1.1.2　误差专项研究期(20 世纪 40 年代初～70 年代)

该时期调查质量的研究是围绕着统计调查误差模型的建立及由此带来的专项研究展开的。

调查误差模型是以整体的方式测定抽样误差和非抽样误差的影响。对调查模型的研究最早始于 20 世纪 40 年代末，1946 年 Hansen 和 Hurwitz 首次提出了无回答问题[8]，50 年代研究取得成果，60 年代形成较为完善的调查误差模型(HHB 模型)[9]。该模型及后来的扩展模型的基本思路为：

均方误差＝测量误差＋抽样误差＋测量误差与抽样离差的协方差＋偏差的平方

当然，HHB 模型虽然仅包含了被调查者和访问员作为误差的来源，但是对调查过程中误差的控制、评估奠定了分析基础，具有现实的重要性。一方面，该模型利用均方误差的概念将总误差分解为各个相对重要的误差组成部分，并使测定各部分在总误差中的影响成为可能；另一方面，在此基础上，人们还可以进一步研究每一分量的分解及测算。如 Kish[10]整合了多数随机误差和偏差的误差来源，Andersen 将这个模型用于公共卫生支出调查[11]，包括了无回答误差，回答偏差，插补偏差。Bailar 和 Biemer 扩展了 HHB 模型，用以显示所有已知的非抽样误差[12]，当然，由于交互效应较为复杂，难以估计模型的参数，Bailar 和 Biemer 的模型在实践中难以应用。因此，Groves 和 Magilavy 的模型

假设交互效应为0，包含了抽样误差，无回答和回答误差[13]。也正因为如此，60、70年代后，围绕调查误差的许多研究是要找出均方误差中占较大份额的组成部分，并研究减少这些成分的方法。如对子样本重复计量、交叉子样本的方法以及二者相结合的方法[14]。同时，这一时期也出现一些专项研究，最典型的是60年代 Warner 提出对敏感性问题调查的随机化方法[15]。

1.1.1.3 研究扩展期(20世纪70年代末～)

该时期关于调查数据准确性的研究仍在发展，同时，随着对统计调查质量内涵认识的深入，研究发展到基于统计调查的多维度质量，建立有效的保证、控制、评估与改进体系。

(1)调查数据的准确性方面

在调查思想理论方面，Royall 于1970年发表的文章[16]形成了模型化推断理论。与随机化理论中总体单位取值固定，抽样程序随机不同，模型化推断法认为，被调查的有限总体是某个“超总体”的具体实现，目标参数的推断建立在“超总体”统计模型的基础上，而抽样不一定要采用随机的程序，只要模型的设定正确，有利于提高估计精度。因此模型化推断给出特定参数估计量是条件估计，其表述与随机化推断往往存在一定差别。后来，在模型化推断的基础上，结合随机化理论，出现模型辅助设计，提出了广义回归估计量、校正估计量，不但有利于提高设计精度[17]，也可以减少非抽样误差(如无回答误差和调查框误差)[18]。Särndal 等人[19]对模型化推断和模型辅助设计的基本思想方法进行了概括。

在方差估计方面，经典(概率)抽样的参数估计及其方差估计在理论上都已解决，但在实际问题中所使用的抽样方法可能是多种抽样的组合，获得的样本较为复杂，估计量也不限于简单的线性估计。复杂样本的方差估计是20世纪70年代以来较多讨论的一种技术，其中主要方法有：①随机组方法，②平衡半样本方法，③刀切法与自助法等再抽样方法，④线性化方法。Wolter 系统地总结上述四种方法在内的复杂样本的方差估计方法[20]。

在非抽样误差的专门研究方面，20世纪80年代开始，无回答问题成为研究热点，一方面，以多重插补法为标志，涌现出众多关于无回答插补的文献[21]；另一方面，以校正估计量为代表[22]，概括了无回答的加权调整[18]。大量先进统计方法在无回答研究领域的应用，带动了这一领域蓬勃发展。关于非抽样误差处理的其他方面全新思想和方法较少，大多数学者或者进行了理论的总结归纳，或者提出了方法的改进、比较研究。包括 Groves 的调查误差与成本的研究[23]，Lessler 和 Kalsbeek 对非抽样误差产生背景、统计影响和处理方法等进行了总结归纳研究[24]。

(2)广义调查质量及管理方面

政府大规模统计数据搜集系统的日益制度化,抽样调查地位的确立及调查模型的建立,统计数据质量的研究趋势是研究调查设计的规范化、调查误差与调查费用的平衡,而且随着调查内容的复杂化,对被调查对象提供资料的保密及个人隐私的保密也已成为统计调查设计质量考虑的重要内容。因此统计数据质量的涵义得到扩展,由单纯的统计数据准确性扩展到更为丰富的内涵。

一是统计数据质量由准确性扩展到更广泛的涵义。

一方面,统计数据质量最广泛的定义是"满足需求",从满足数据使用者需求角度看,调查数据作为统计生产过程的成果,必然表现出多方面的质量要求。Dalenius 提出统计数据质量的"测量向量"[25],包括统计数据的准确性、经济性、保密性、相关性、时效性、数据详细程度等。1980 年联合国统计司出版的《统计组织手册》中,对官方统计资料提出八项要求[26],实际上也远远超出统计数据准确性、及时性的要求,进一步明确了统计数据质量的综合涵义。进而,各个学者和统计机构纷纷提出并完善了包含多个维度的统计数据质量概念,诸如,国际货币基金组织的数据质量评估框架[27],联合国粮农组织的数据质量评估框架[28]等。

另一方面,从统计数据质量扩展为统计调查质量。统计调查质量不但包含统计数据质量,还包含调查过程质量,即表现为"符合标准"。诸如,加拿大统计局关于调查质量的政策指南和数据质量评估框架[29,30],欧盟统计局的调查质量评估框架[31]。对不同框架的调查质量维度比较发现[32],不同框架多维度质量内涵在一定程度上表现出一致性。

二是从调查误差测量到调查质量控制专项研究。

随着调查活动的广泛开展,人们意识到对调查误差模型完全分解既不经济也难以实现。特别是 20 世纪 70 年代"石油危机"后,国外政府非常关心统计数据质量对决策的影响,信息产业的发展,更促使人们开始研究利用现有的各种知识,对统计调查质量进行综合的控制。Dalenius 提出统计调查质量的综合控制应包括调查设计控制和调查实施控制两部分,调查实施控制又分为预防控制和过程控制。20 世纪 80 年代初美国的许多政府统计机构以及大的公司对本单位的统计调查质量都进行了综合质量控制,1988 年 4 月美国田纳西 Oak Ridge 统计调查质量专题研讨会上,提出了统计数据质量环的一般内容并进行讨论。而且就统计数据质量控制达成以下共识[33]:①简单的质量控制应在实施更困难、花费更高的步骤发生之前进行,即强调统计调查质量的事前控制。②数据误差和偏差应在质量控制环最早的可行阶段中发现,即强调统计设计质量。③软件应标准化,即强调统计调查质量管理方法体系的标准化。④进一步

发展统计调查质量的综合测量和对用户提供统计数据使用指南。这实际上是为进一步的规范化的调查质量控制提出了一个管理模式或标准。此后，不仅美国，而且世界上许多其他国家如意大利[34]、瑞典[35]的农业普查中都进行了统计调查质量的综合控制。相应地，如何进行统计调查质量综合控制也成为重要的研究课题。例如，Lyberg 和 Biemer 等在 1995 年国际调查测量与过程质量会议的基础上，组织编写的《调查测量与过程质量》分问卷设计、数据收集、数据处理、质量评估和控制、误差对估计量的影响 5 个专题，对调查质量控制进行了系统的讨论[36]；Biemer 和 Lyberg 的专著[37]则介绍了调查过程控制的原理和研究新进展。

1.1.2 国外统计调查质量研究的前沿面

纵观目前国外对统计调查质量的研究，不难发现，其研究内容是极其广泛的，具体集中于三个方面：

1.1.2.1 关于统计调查质量维度

(1)调查质量的内涵的研究，即统计数据质量、调查过程质量应包含维度的讨论。

(2)准确性方面的研究。包括误差的来源、分类以及如何对误差文档及监测的研究，误差分布以及误差模型的研究，新的估计量的探索研究(如校正估计量、双重稳健估计量)，误差报告指标的研究。

(3)统计调查质量与调查费用及调查效率的研究。

(4)调查质量特性的定量测定研究。

1.1.2.2 关于调查过程质量的测度与控制

(1)统计调查设计质量与规范的研究(如问卷设计、调查规模、分组体系、保密问题)；对统计数据收集目的、用户需求的研究；调查实施计划框架充分性的研究；对调查过程的综合概括以及对统计调查质量环的研究。

(2)统计调查执行质量的研究。如调查参与主体(被调查者、访问员)误差及其交互作用的研究；调查参与主体心理对调查结果影响的研究；数据收集系统的问题及处理；普查与抽样调查有机结合的研究。

(3)统计数据处理及操作质量的研究。如数据捕获、编码、数据编辑、插补质量测定的研究；数据处理质量及改进、数据处理过程的控制系统研究。

(4)数据误差对参数估计，统计分析的影响(分析、评估误差对数据建模的影响及表现方式)。调查误差对数据质量的影响及分析；参数估计的准确性及精确性；探索性数据分析和模型识别技术的应用；奇异值诊断方法和数据编辑技术的应用；缺失数据处理计划和实施。

(5)质量评估和控制的研究。如调查质量评估方法,调查误差模型及分析,误差文档及监测,用户的质量评价,质量管理,统计数据报告内容及结构的研究。

1.1.2.3 关于调查质量的管理与改进

(1)统计调查质量改进中的问题及对策研究;提高统计调查质量的管理步骤;统计理论与实际工作者如何提供支持管理者决策的统计数据;统计数据内部一致性,外部一致性的检验。统计数据文件、统计资料的管理。

(2)统计调查质量控制的成功经验、当前最优统计数据的质量水平、统计调查质量理论研究成果与统计调查质量改进实践相结合的研究。

1.2 国内对统计调查质量研究的综述

在国内,统计调查活动很大程度上表现为借鉴国外的统计理论,用于政府统计活动上来。因此,调查质量多表现为政府统计调查(数据)质量。

1.2.1 国内关于统计调查质量理论研究的历史回顾

1.2.1.1 新中国成立以前,有关统计调查质量的研究

新中国成立以前,由于统计理论水平的落后,以及我国政府统计规模水平的限制,应该说,从理论上对统计调查质量的研究一直很少开展。但也应承认,在实践上,从我国官方统计活动诞生的那一天起,对统计调查质量的控制及保证活动就已经存在。例如秦代有关统计的法规主要是为了保证统计数字的准确,曾规定[38]:各项统计数据,要账账衔接和账账相符;必须做好对有关事项的原始登录工作;要保证统计数字的准确,必须计量无误;度量衡器的误差必须限定在一定的范围以内,否则要受到法律惩处;对度量衡器要由官府经常进行校验,以保证度量准确;对统计簿籍的交接要建立验收手续;对违反统计法规要进行制裁。其中对统计数字上的弄虚作假(同贪污、挪用相联系),秦律把它视为盗窃罪一样。并且对统计数字的不实,要区别两种情况,一种是差错大小,一种是差错的责任,对不同的对象,分别采取不同的惩处办法。继秦代以后,后来的两汉、唐宋以及明清各代都是对秦代的统计法规进行继承和发展。而且自夏商代官方统计活动产生以来,官方统计组织的完善,统计范围的扩大,统计内容的日益丰富等都表明统计数据质量在我国的历史实践中是逐渐提高的。

中华民国时期,受清代以来引入西方统计理论的影响,北洋政府和国民政府统计,推行西方国家的统计组织,依据西方统计理论和方法进行了一系列统

此为试读,需要完整PDF请访问: www.ertongbook.com