

《全国大学生市场调查与分析大赛》指定参考书

# 市场调查理论与方法

---



主编徐映梅

副主编 张海波 孙玉环 中

国商业统计学会组编

主审金勇进

《全国大学生市场调查与分析大赛》指定参考书

## 统计学与数据科学系列教材

国民经济统计学（第三版）

邱东

应用时间序列分析（第二版）

史代敏 谢小燕

多元统计分析及 R 语言建模（第五版）

王斌会

◆市场调查理论与方法

徐映梅



ISBN 978-7-04-050250-3



9 787040 502503 >

定价：56.00 元

# 市场调查理论 与方法

主编徐映梅 副主编张海波

孙玉环 中国商业统计学会

组编

主审金勇进

# 《全国大学生市场调查与分析大赛》指定参考书

## 内容简介

本书结合作者多年教学实践和国内外优秀市场调查教材的成果编写而成。内容包括对市场调查活动整个流程与环节中具体方法、技术与规程的介绍，涉及市场调查方案设计、面向对象的数据采集方法、量表设计技术与问卷设计、抽样设计与样本量确定、数据整理与清洗、抽样估计、单变量与多变量数据分析方法及市场调查报告撰写等。本书将案例与数据分析技术相结合，大部分调查分析方法都给出了 SPSS 的计算过程与结果，书后配有教学和学习辅助资料，方便教师讲授和学生自学。

本书可作为高等院校经济、管理类各专业本科生市场调查课程的教材，也可作为全国大学生市场调查与分析大赛的参考用书。对广大实际工作者也具有参考价值。

## 图书在版编目 (CIP) 数据

市场调查理论与方法 / 徐映梅主编；中国商业统计学会组编。--北京：高等教育出版社，2018.9

ISBN 978 -7 -04 -050250 -3

i. ①市… U. ①徐… ②中… in. ①市场调查-高等学校-教材 ②市场分析-高等学校-教材 IV. , ①F713. 52

中国版本图书馆 CIP 数据核字 (2018) 第 170687 号

Shichangdiaocha Lilun yu Fangfa

策划编辑权利霞

责任编辑吴淑丽

封面设计王鹏

版式设计徐艳妮

插图绘制于博

责任校对刘娟娟

责任印制陈伟光

出版发行 高等教育出版社

网 址 http://www. hep. edu. cn

社 址 邮 北京市西城区德外大街 4 号

http : // www. hep. com. cn

政 编 码 100120

网上订购 http://www. hepmall. com. cn

印 刷 北京佳顺印务有限公司

http: //www. hepmall. com

开 本 787 mm x 1092 mm 1/16

http: //www. hepinall. cn

印 张 25.25

版 次 2018 年 9 月第 1 版

字 数 640 千字

印 次 2018 年 9 月第 1 次印刷

购 书 网址 010 -58581188

定 价 56.00 元

本书如有缺页、倒页、脱页等质量问题，请至所购图书销售部门联系调换 版权所有侵权必究

物料号 50250 -00

市面上关于市场调查或市场调研的教科书很多，为什么还要编写这方面的教材？有两个原因：一是培养学生综合素质与创新能力的需要。市场调查这门课程的知识体系，与一般的工作模式及研究模式具有相似性，都遵循“问题提出—方案设计—证据采集与处理（主要指调查实施与质量控制、数据预处理与分析）—因果解释—报告撰写”的流程，强化这一流程的完整训练，有助于学生建立起自己提出问题、研究问题的路径，提升解决问题的创新能力。市场调查的理论与方法十分丰富，分属不同的课程内容，将其综合在一起，有助于深刻理解调查理论与方法之间的相依关系、理论与实践相结合中处理问题的方式，起到融会贯通、避免误用和滥用的作用。容易引起误用和滥用的知识点都需要在教科书中介绍清楚。第二个原因是试图通俗地将市场调查理论与方法的科学性和严谨性表达出来，以方便人们正确地讲授和使用。在这样的动机驱使下，我们开始了本教材的编写，经过将近两年时间的努力，终于将书稿呈现在各位读者面前。

本教材注重体现以下特色：

注重方法的科学思想。调查是一门技术，阐述清楚技术的重点与难点，选择时的前提与条件、采用后可能的不足与后果，做到心中有数是十分必要的。市场调查方法的内容涉及：方案设计、问卷设计、量表设计、抽样技术、样本量确定、抽样估计、数据分析方法等。本教材与以往教材不同的是更加体现市场调查的专属性与专业性，是在学习统计学原理相应方法后的进一步提升，也是统计技术在市场调查与研究领域的应用。

体现内容的通俗易懂。通俗地阐述市场调查理论与方法中出现的基本概念、基本方法与基本原理，举例说明方法原理中包含的基本思想、操作过程、注意事项，做到清晰地讲解知识点的来龙去脉，便于学生领会与掌握。

贯穿全程的强化实训。市场调查的良好实现需要基于问题，将市场营销、消费者行为、营销管理、统计学、统计分析方法进行综合运用。因而相关知识的讲授与学习需要体现综合训练。该教材体现在章前有导语或引例，有明确的学习目的与要求，有知识结构梳理；章后有小结、关键名词、思考题、练习题和实训作业。章前的设计可以让学生理解学习的目的与要求，而章后的实训可以让学生深刻领会与掌握相关的知识、技巧、方法和技术，从而达到对知识点的融会与贯通，特别要指出的是，每章后配有基础练习题及其参考选项，题型有单项选择题、多项选择题和判断题；读者通过手机扫描二维码即可进行相应练习，方便快捷掌握基础知识，强化学习效果。

全书分为十五章，体现的基本逻辑关系是以调查的流程与环节为主线依次展开，问题的提出—方案的设计—数据的采集技术—调查的实施与过程控制—数据清洗与整理技术—数据的分析技术—调查报告的撰写与展示”

本教材是集体劳动的成果。中南财经政法大学徐映梅教授完成第一章至第三章，第六章至第十章，第十五章的编写；中南财经政法大学张海波副教授完成第四章及第五章的编写；东北财经大学孙玉环教授完成第十一章至第十四章的编写。全书由徐映梅教授总纂，中国人民大学金勇进教授主审…参与案例数据搜集与整理、习题编写和书稿校对等工作的有来自中南财经政法大学的研究生王亚琪、郭晓菲、徐婷婷、高申瑞、陈树德、鄂卓、武治巧、陈华、赵梦凡、牛辑、李坤、陈康、黎倩文、相立博、李健、邱凡、张毅、杨金枝、郭会华、何正平、郭雪雅，来自东北财经大学的研究生刘宁宁、韩月、王建兰、王曦莹、徐晓涵、刘楠、赵明珠、王洁、韩智东、韩雨晴、阿果、林娜。感谢他们的辛苦付出！特别要感谢中国商业统计学

会秘书长江懿和工作人员宋宝生；中国商业统计学会市场调查与教学分会秘书长孙际平，她（他）们为书稿案例及习题的编写提供了很多的支持与帮助！

本书的出版还要感谢高等教育出版社的权利霞编审、责任编辑吴淑丽博士，她们对书稿的很多建议使本书增色很多。

本教材参阅了很多人的成果，我们在书末列出了相应的参考书目与文献，在此特向有关作者表示衷心的感谢！

本教材可以采用选讲与自学相结合的形式。如果是36学时的课程计划，建议选讲章节为第一章至第八章、第十五章，其余内容可通过实训完成相应的指导与学习；如果是48学时的课程计划，建议选讲章节为第一章至第十章、第十五章，其余内容可通过实训完成。

本教材适合于先修了概率论与数理统计学、统计学等课程的大学高年级学生学习，也可以作为全国大学生市场调查与分析大赛的参考用书。

由于作者水平有限，加上时间仓促，尽管已经尽了最大努力，但书中难免存在缺憾与不足，敬请读者不吝指正，我们将在再版时认真修改。

编者

2018年1月于武昌

## 前言

第一章 市场调查概述 .....

    导语 .....

    教学目标与要求 .....

    本章知识结构 .....

    第一节 市场调查的含义、分类与作用 .....

    第二节 市场调查的基本程序 .....

    第三节 市场调查的使用范围 .....

    第四节 市场调查需遵循的道德伦理「 .....

    本章小结 .....

    关键名词 .....

    思考题 .....

    本章实训 .....

第二章 市场调查方案设计 .....

    导语 .....

    教学目标与要求 .....

    本章知识结构 .....

    第一节 市场调查方案设计及其重要性 .....

    第二节 市场调查问题的界定 .....

    第三节 市场调查中的误差及其控制 .....

    第四节 市场调查方案的拟订 .....

    本章小结 .....

    关键名词 .....

    思考题 .....

    本章实训 .....

第三章 数据收集方法 .....

    导语 .....

    教学目标与要求 .....

    本章知识结构 .....

第一节 定性研究的数据收集方法•••••	
第二节 定量研究的数据收集方法•••••	
第三节 各种数据采集方法的比较 .....	61
本章小结 .....	64
关键名词 .....	64
思考题 .....	64
本章实训 .....	65
 第四章 量表测量技术.....	66
导语 .....	66
教学目标与要求 .....	67
本章知识结构 .....	68
第一节 测量与测量工具 .....	69
第二节 量表的尺度与类型 .....	72
第三节 市场调查中的常用量表 .....	78
第四节 量表质量的评价 .....	91
本章小结 .....	97
关键名词 .....	98
思考题 .....	98
本章实训 .....	98
 第五章 问卷设计技术.....	100
导语 .....	100
教学目标与要求 .....	100
本章知识结构 .....	102
第一节 问卷概述_____， ••• .....	103
第二节 问卷设计的原则与基本流程 .....	106
第三节 问卷设计技术 .....	110
第四节 网上创建在线问卷 .....	130
本章小结 .....	134
关键名词 .....	134
思考题 .....	134
本章实训 .....	135
 第六章 抽样方法.....	136
导语 .....	136
教学目标与要求 .....	136

本章知识结构 .....	137
第一节抽样方法概述 .....	138
第二节 非概率抽样技术 .....	142
第三节概率抽样技术 .....	146
第四节 抽样设计中的注意事项 .....	160
本章小结 .....	162
关键名词 .....	162
思考题 .....	163
本章实训 .....	163
 第七章样本量确定.....	164
导语 .....	164
教学目标与要求 .....	164
本章知识结构 .....	165
第一节影响样本量确定的因素 .....	166
第二节不同抽样方式下样本量确定 .....	171
第三节分层抽样样本量的分配 .....	176
本章小结 .....	183
关键名词 .....	183
思考题 .....	184
本章实训 .....	184
 第八章抽样估计.....	185
导语 .....	185
教学目标与要求 .....	185
本章知识结构 .....	186
第一节抽样设计加权与调整 .....	187
第二节 总体均值、比例及总量的点估计 .....	190
第三节辅助信息的使用 .....	197
第四节 调查估计量的抽样误差及区间估计 .....	201
本章小结 .....	210
关键名词 .....	210
思考题 .....	210
本章实训 .....	210
 第九章市场调查的组织与实施.....	212
导语 .....	212

教学目标与要求 .....	212
本章知识结构 .....	213
第一节必要的公关工作 .....	214
第二节调查前的准备 .....	215
第三节数据收集过程 .....	218
第四节调查过程中的质量监控 .....	224
本章小结 .....	227
关键名词 .....	227
思考题 .....	227
本章实训 .....	227
 第十章调查数据处理与分析设计.....	229
导语 .....	229
教学目标与要求 .....	229
本章知识结构 .....	230
第一节 问卷的接收和检查 .....	231
第二节 问卷编码和录入 .....	232
第三节 调查数据的统计预处理 .....	237
第四节 编制数据分析任务书 .....	243
本章小结 .....	248
关键名词 .....	248
思考题 .....	248
本章实训 .....	249
 第十一章单变量数据分析.....	250
导语 .....	250
教学目标与要求 .....	252
本章知识结构 .....	253
第一节 规模、结构与比较分析 .....	254
第二节集中度分析 .....	263
第三节变异性分析 .....	268
第四节参数估计 .....	270
第五节显著性检验 .....	273
本章小结 .....	276
关键名词 .....	277
思考题 .....	277

本章实训	277
第十二章双变量关联性分析	279
导语	279
教学目标与要求	279
本章知识结构	280
第一节双变量列联分析	281
第二节双变量相关分析	288
第三节双变量回归分析	292
本章小结	292
	301
关键名词	301
思考题	302
301	

本章实训 .....	302
第十三章 多变量数据分析（1） .....	303
导语 .....	303
教学目标与要求 .....	306
本章知识结构 .....	308
第一节多变量方差分析 .....	309
第二节 多变量相关与回归分析 .....	319
第三节 分类模型——Logistic 模型 .....	327
本章小结 .....	332
关键名词 .....	333
思考题 .....	333
本章实训 .....	333
第十四章 多变量数据分析（2） .....	335
导语 .....	335
教学目标与要求 .....	335
本章知识结构 .....	336
第一节因子分析 .....	337
第二节聚类分析 .....	343
第三节对应分析 .....	348
第四节多维标度分析 .....	352
第五节联合分析 .....	356
第六节多变量综合评价 .....	362
本章小结 .....	367
关键名词 .....	368
思考题 .....	368
本章实训 .....	369
第十五章市场调查成果报告 .....	371
导语 .....	371
教学目标与要求 .....	371
本章知识结构 .....	372
第一节市场调查报告的功能 .....	373
第二节市场调查书面报告的撰写 .....	374
第三节市场调查口头报告的撰写 .....	383
本章小结 .....	384

关键名词 .....	385
思考题 .....	385
本章实训 .....	385

附录 1 2018 年中国景区旅游消费研究报告 .....	386
附录 2 入户访问项目遵循规范 .....	387
<hr/>	
附录 3 第八届全国大学生市场调查与分析大赛评分标准.....	388
主要参考文献.....	389





# 第一章市场调查概述

## 导语

市场营销管理是以市场为中心、顾客为导向，实施以产品、价格、渠道和促销策略为主要内容的管理过程。管理的重点是决策，而决策的关键则是拥有信息的数量与质量以及如何使用信息。市场调查正是以科学的方法，系统地收集、记录、整理和分析有关市场营销问题所需信息的一种手段，目的是为市场营销策略的制定和实施提供依据。因而加强数据研究，从铺天盖地的海量数据中寻找市场机会、避开潜在威胁、提升企业营销决策的科学性与准确性，以使企业取得良好的经济效益，正是市场调查巨大的价值所在。

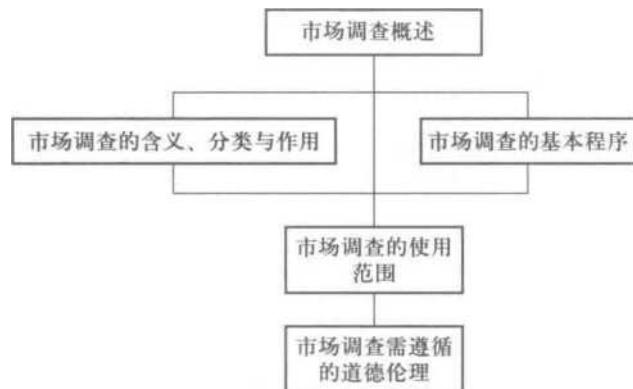
例如 2018 年中国景区旅游消费研究报告就是基于调查而形成（参见附录 1）。

本章将主要介绍市场调查的含义、特点、作用、分类、基本程序、使用范围和市场调查相关行为主体需遵循的道德伦理：

## 教学目标与要求

1. 理解市场调查的含义与特点；
2. 掌握市场调查的程序；
3. 熟悉市场调查的主要内容；
4. 知晓市场调查的道德伦理

## 、本章知识结构



## 第一节

# 市场调查的含义、分类与作用

市场调查的基本问题涉及对市场调查内容的界定、对市场调查活动特点的理解、依据不同 标志对市场调查进行分类和对市场调查成果进行利用等。下面分别简述。

### 一、市场调查的含义与特点

#### (一) 市场调查的含义

由于研究的目的与重心不同，市场调查有广义与狭义之分。广义的市场调查更加强调研究的内涵，因此也称为市场研究、市场调研或市场营销研究，涵盖从认识市场到制定营销决策 的有关市场营销活动的分析与研究。狭义的市场调查更偏重于信息的收集与分析。无论是广义还是狭义的市场调查，其主体或核心的技术内容则是确定的。作为一本教材，由于篇幅所限，难以面面俱到，但必须包括其核心内容。本书侧重于市场调查理论与方法的核心内容，即 在内容上更偏重于调查、研究与分析的基本方法与技术及其具体操作规程。因而在书中不再 区分市场研究、市场调研或市场营销研究，统称为市场调查。本书定义的市场调查是指在市场 活动中，以特定研究目的为指引，有计划地收集、整理和分析市场的信息资料，提出解决问题的 方案及建议的一整套科学方法、技术及规程。本书介绍的方法和技术同样适用于一般的社会 调查。

市场是指在一定的环境下经济活动主体进行公开交易的场所。一定的环境是指当时特定 的政治、经济、社会及生态环境。经济活动主体可以是法人单位或自然人，通常以法人单位即 企业为主体，还包括政府、非营利组织和个人。交易的内容是货物与服务。交易的场所表现为 不同范围的空间。

市场是经济活动主体经营业绩实现的重要途径，是货物与服务交易实现的桥梁。经济活动主体作为货物与服务的提供方与需求方，在市场交易活动中，需要合理的定价、合适的促销、 有效的渠道支撑，需要不断关注市场供需双方收入的变化、心理的变化和偏好的改变，进而能够面向需求者提高产品与服务质量、开发新产品、提升新服务。这样，每个经济活动主体都需 要快速掌握有关市场的准确、可靠的信息，使自己在市场交易中处于优势地位。因此，经济活动主体都需要进行专门的市场调查与分析。

#### (二) 市场调查的特点

市场调查活动目的性强，包括一整套科学的方法与技术。其特点概括如下：

1. 目的性。由于任何调查都受限于一定的人力、物力和财力，要在一定的时空条件下获 取充分的信息，必须目标明确。因此，任何一种调查都有明确的目的，并围绕目的展开具体的 调查。没有目的或多个目的都是不可取的。

2. 系统性。调查活动是一个系统工程，包括问题的定义，计划的编制，方案的设计，样本 的抽取，资料的收集、整理与分析，报告的撰写等。在这一过程中，每个阶段环环相扣，互相联 系、互相依存、互相影响。某一环节的疏忽，都可能会影响整个调查结果的准确性。

3. 科学性。市场调查的科学性体现在三个方面：即科学的方法、科学的技术手段和科学 的分析结论。市场调查工作因为复杂而需要周密的计划，精心的组织和科学的实施。在进行 市场调查时要以科学的方法为指导；在调查设计时要按照科学的程序进行；在研究方法选择 时，需根据具体要求和条件选择最适当的方法；在研究结果报告时，需排除主观因素的影响与 干扰，以科学的态度提供研究报告。

## 二、市场调查的基本分类

划分市场调查的类型，有助于经济活动主体根据需要选择最合适的调查途径

### （一）根据研究的性质分类

按照市场调查的目的和研究的性质，可以分为探索性调查、描述性调查、因果性调查和预测性调查。

#### 1. 探索性调查

探索性调查是指对知之甚少的事物进行初步研究的方法与技术，其目的是提供一些资料以帮助研究者认识和理解所面对的问题、发现想法和洞察内部关系。一般采用小样本研究，用于大规模正式调查之前。探索性调查的常用方法有：专家咨询或调查、预调查、个案研究、二手资料分析、定性研究等。探索性调查得到的结果一般可以通过描述性或因果性调查进行证实

#### 2. 描述性调查

描述性调查是指采用结构性方法描述总体（市场）的特征或功能。使用该方法的前提是假定研究者事先已对所研究的问题具备许多相关的知识。一般以有代表性的大样本为基础，通过事先制定好的具体假设和结构性方案展开研究。常用的方法有：二手资料分析、抽样调查、固定样本连续调查、观察法、模拟法等。

#### 3. 因果性调查

因果性调查是指为了确定研究对象中变量之间的因果关系而进行的调查，是对因素之间作用机制的深刻认识。一般是在因果机制的假定下，对定义好的因变量和自变量进行调查，观察一个因素的改变是否引起另一个因素改变的研究活动。因果性调查要求调研人员对所研究的问题有深刻的理解，能够判断一种情况出现了，另一种情况会接着发生，并能说明其原因所在。因果关系研究的特征为：要处理一个或多个独立变量、要控制其他中间变量或间接变量。常用方法有市场实验法。

#### 4. 预测性调查

预测性调查是指利用已有的市场经验和科学的预测技术对市场未来的发展趋势进行估计和判断的调查方法。预测性调查要求对影响市场未来发展的各种因素进行调查，并且对未来可能出现的各种状况及其概率进行估计和测算。常用方法有专家咨询法、**趋势分析法**、时间序列预测等。

### （二）根据资料来源分类

按照调查资料的来源不同，可以分为文案调查和实地调查，又称为二手资料调查和一手资料调查。文案调查是指通过收集已有的资料、数据、调查报告、已发表的文献等有关的二手信息，加以整理和分析的一种市场调查方法，也叫作二手资料分析或二手数据分析。文案调查经常在探索性的研究阶段中使用。实地调查是指在制定详细的调查方案基础上，由调查员直接向被调查者收集第一手资料，再进行整理和分析，从而写出调查报告的市场调查方法。由于二手数据的信息常常不能很好地满足问题研究的需要，因而实地调查是市场研究中的主流方法。

### （三）根据调查的组织形式分类

按照调查的组织形式，可以分为专项调查、连续性调查和搭车调查。专项调查是指受某个

客户的委托针对某些问题进行一次性的调查，即从给定的总体中一次性地抽取样本进行调查，并且只从样本中获取一次性信息的调查形式。专项调查可以采用定量调查方法，也可以采用定性调查方法。连续性调查是指对一个（或多个）固定的样本进行定期的、重复的调查。样本中的被调查对象（人或单位）一般不随调查时间的变化而变化。例如消费者固定样组（Panel）或其他固定样本调查，连续的跟踪研究和品牌测量，零售扫描研究，连续的媒介研究等，都属于连续性调查。搭车调查是指多个客户共同利用一个样本进行调查。根据各个客户搭车调查问题的个数和类型，来决定客户的费用。一般有搭车调查业务的调查公司，每年实施搭车调查的时间和价格都是固定的，例如每月实施一次或每周实施一次。由于搭车调查的实施一般都是定期的，因此经常将搭车调查归入连续性调查类，但是要注意的是，搭车调查每次所用的样本是可变的。

#### （四）根据调查研究的方法分类

按照调查研究的方法，可以分为定性调查研究与定量调查研究。定性调查研究是指采用专门的方法收集定性的调查数据进行研究的方法，定性研究注重对现象的性质、特性的认识。定性调查研究的方法具体包括：文案调查法、小组座谈会、深层访谈、投影技法和其他定性调查方法。定量调查研究是指采用专门的科学方法收集定量的调查数据进行研究的方法，定量研究注重现象结构性量化特征及数量关系的揭示。定量调查研究方法具体包括：邮寄调查、电话调查、街道或商城拦截面访、中心地调查、入户面访、借助其他电子手段（如传真、互联网等）的调查、神秘购物调查以及其他定量调查方法。

### 三、市场调查的作用

从企业角度理解市场调查的作用，主要表现在以下五个方面：

#### （一）识别机会与问题

机会是指未得到满足的需求或没有得到很好满足的需求。问题是指企业营销中显著的或潜在的不利因素及其影响。通过市场调查，建立起市场营销信息系统，能够方便地检索、查询和利用已有的资料，掌握国际层面、国家层面、行业层面的有关信息，分析预测市场发展总体趋势，为企业战略决策提供依据；通过市场调查，广泛收集整理分析数据资料，建立起监测市场环境变化的预警系统，及时掌握市场可能出现的机会与威胁，从而把握住有利的条件与因素，规避不利的趋势与变化；通过市场调查，可以帮助企业明确所属行业中可能存在的机会，发现对企业现在或将来发展的不利因素；通过市场调查，可以了解企业自身的优势与劣势、机会与威胁，确切感受竞争引起的压力，并转化为企业发展的动力，迫使企业加强经营管理，跟踪最新产品的市场发展趋势，增强产品的适应能力，从而跟上市场的步伐，避免由于市场需求变化给企业带来的机会损失与经济损失。总之，借助市场调查，企业可以明确营销努力的方向，高效率地整合资源，实现企业盈利目标。

#### （二）更好地理解市场

理解市场是指理解市场信息的不断更新中表现出的各种关系的改变：只有通过市场调查，才能充分认识市场的过去和现在，探明市场供求发展变化的规律，了解顾客的需求及变化趋势，掌握竞争对手的数量与策略，为更好地服务于目标顾客提供依据。此外，通过市场调查可以动态了解不同偏好导致的需求差异，认识、分析和解释消费者行为，掌握市场主体及其行为特征，有助于企业改进营销策略。无视市场经济自身发展变化的规律，按照人的主观意志来开发产品、开拓市场，注定会失败。

#### （三）协助制定营销规划

无论是战略性决策还是战术性决策，都需要以大量的信息为依据，特别是在21世纪这样一个信息经济时代，超出企业常规决策以外的发展创新战略不能仅凭经验和猜测进行决策。因为这一类型的决策都具有一定的创新性，存在较大的不确定性和风险性，必须掌握大量的第一手数据和第二手数据，而这些数据资料只有通过文案调查和实地调查才能获得。通过调查可以辨别环境变化所带来的市场机会，发现企业营

销中可能存在的问题，协助制定企业的总体发展战略、目标市场战略和竞争战略——

#### （四）制定营销决策

企业制定营销决策通常需要回答以下问题：顾客关注哪些产品属性？产品设计应集成多少特性？产品应该怎样定价？产品由谁来销售？推广方面的费用多少较经济？科学的营销决策以占有信息为前提，而信息的质和量取决于市场调查的质量。因而，市场调查结果直接影响到决策的准确性。一般来说，决策者掌握与决策问题相关的信息越多，对决策方案选择和制定的余地就越大，对各种方案可能出现结果的估计就越充分，决策的可靠性就越有保证。在市场经济条件下，企业决策者无论是做出目标市场决策，还是产品开发、价格制定、渠道安排及促销策划，都必须对市场状况进行认真的调研，并权衡各种方案的得失。不认真研究市场的实际情况，只凭主观臆断，盲目决策，必然要蒙受损失，付出代价。

#### （五）实施营销控制

控制就是了解目标与实际情况的差距。因此，需要通过市场调查来掌握公司的市场占有率、顾客满意度、退货率、服务质量和服务水平等企业营销指标的实际执行情况。只有通过调查揭示出差距，明确问题的症结所在，才能够“对症下药”，提出相应的矫正方案，改善企业的经营管理，提高企业的经营效益。

作为国民经济微观核心的企业，要在市场经济的海洋中做到勇立潮头，就必须用好市场调查这个有效工具，为企业经营决策服务。

但市场调查结果有其局限性，需要正确利用。市场调查通常是用样本数据推断总体，不可避免存在误差。但其误差可控，因此可以在大概率条件下认同调查的结果。市场调查结果只是找出了企业的病根或部分病因，解决问题还要靠企业对各种资源的运作能力。市场调查结果还需用于决策，准确可靠的调查信息如果不用于决策就不会对营销活动产生影响，调查的所有投入就不会产生任何经济效益。

## 第一 F

### 市场调查的基本程序

市场调查是一个经历一定时期的调查研究过程。这一过程具体包括确定调查问题、设计调查方案、实施调查计划、处理与分析数据和报告研究结果五个前后相连的阶段，每个阶段又可分为若干具体步骤。

#### 一、确定调查问题

市场调查的第一步是明确并定义调查研究的问题，提出必要的研究假设定义调查问题需要背景信息、决策信息等的支撑。一般包括如下几个方面：

##### 1. 理解调查问题的背景

从背景出发有助于理解问题的来龙去脉，了解问题依据的基本信息、侧重方向，并对关联问题有一个初步的研判。其核心是围绕调查费用、调查目的、调查对象、调查内容、资料及数据基础等开展一系列的交流、访谈与摸底。

##### 2. 充分交流问题关联信息

在理解调查问题的背景之后，需要进一步聚焦问题的核心，这时需要与决策者进行充分的交流和讨论，明确决策者所面临的问题的性质，可能采取行动或策略，市场调查可以解决什么问题等。同时向行业专家咨询，通过行业关联分析，加深对问题深度与广度的理解。还需要从更加宽泛的视角收集和分析二手资料，补充从决策者和行业专家中得到的信息。有时可以进行必要的定性研究，如采取小组座谈的方法或小样本的预调查方法等等，从而全面和准确地界定问题的边界。

### 3. 定义调查研究问题

市场调查常常是为了解决管理决策问题而进行的。因此，定义调查问题就需要很好地将 市场调查问题与管理决策问题进行转换，使两者在信息的针对性上实现对接。例如，管理决策 问题是“是否需要研发新产品或新服务”，将其转换为市场调查问题就可能是：“确定消费者的 偏好以及对新产品的购买倾向” “确定对价格的弹性要求以及价格对销售量的影响” “确定各 种价格变化水平的收益”。

研究者在定义调查问题时要避免定义得太宽或太窄。将问题定义得太宽则难于操作；而 定义得太窄又可能限制了研究者的视角，妨碍研究者去关注管理决策问题中的重要部分。因 此，可以先将调查问题用比较宽泛的、一般的术语来陈述，然后再具体规定问题的各个组成部 分、在许多市场调查研究项目中，需要理论的框架和分析模型的支持，研究者才能确定要调查 的变量及变量间的关系，从而帮助研究者明确有关的调查问答题和进一步提出假设。

假设是关于研究者所感兴趣的某个因素或现象的未经证明的陈述或主张。例如，假设可 以是理论框架或分析模型中规定的两个变量或多个变量间关系的一种试验性的描述。假设的 重要作用之一是提醒调查研究方案中应包括哪些变量在内 C

## 二、拟定调查方案

在明确地定义了调查问题和提出了假设之后，研究者就要设计详细的调查研究方案。不 同的调查问题需要有不同的方案，因此在设计中必须具体问题具体分析。但在方案设计中一 般都包括以下内容（详见第二章）：

1. 确定调查的对象和调查范围。它是指明确研究总体及总体单位的特征，明确调查的具体时间与空间范围。
2. 选择合适的研究方法。它是指根据研究目的，选择调查研究性方法中的探索性方法、 描述性方法、 因果性方法，或上述这些方法的组合。
3. 选择数据收集的具体方法。它是指根据研究目的、研究对象特征、调查的时间及经费 限制选择合适的数据收集方法。它可以是定量研究方法中的电话调查、面访调查或邮寄调查； 也可以是定性研究方法中的文案调查、深层访谈或小组座谈会，还可以是它们中多种方法的组 合（详见第三章）。
4. 规定调查内容（所需的详细信息）。调查内容具体体现为问卷中的问题，因研究对象特 点和研究 目的而不同。一般还会包括被调查者的背景信息、被调查者的某种行为或偏好信息、 行为或偏好改变的影 响因素的关联信息等。
5. 规定变量的测量方法和选择适当的量表（详见第四章）。
6. 设计调查问卷和测试问卷（详见第五章）。
7. 设计抽样方案和具体的实施计划（详见第六、七章）。
8. 制定并实施数据分析方案（详见第十章）。
9. 预算调查经费。
10. 安排调查进度。

一般提交给委托方的调查计划书则要求简洁明了。一般是按以下顺序撰写：调查背景和 目的、调查内 容和范围、调查方式和方法、调研经费预算和进度安排、其他说明等。

## 三、 实施调查计划

调查的顺利实施是关系到市场调查成功与否的关键步骤，其重心又在于实施过程中严格 的组织管理和质量控制。具体包括以下内容（详见第九章）：

1. 挑选调查员。

2. 培训调查员。
3. 监督管理调查的实施。
4. 复查验收调查员的工作。
5. 评价调查员的工作。

专业性的市场调查公司对上述每一部分的工作一般都会制定详细的工作手册或工作流程，包括调查员的基本条件、调查员培训手册、督导员工作手册、复查规则、调查员评价标准等。

#### 四、 处理与分析调查数据

市场调查的实施完成后，还必须在数据处理和分析环节严格管理，才有可能得到准确的结果。具体包括（详见第八章、第十至第十四章）：

1. 接收和清点资料。
2. 检查和校订资料。
3. 编码。
4. 录入数据。
5. 查错或数据净化。
6. 处理缺失数据。
7. 统计预处理。
8. 制表、作图。
9. 抽样估计与统计分析。

#### 五、 报告调查结果

市场调查的最后一步是撰写市场调查报告。市场调查报告是调查项目的重要组成部分，是呈交给客户的最终产品。研究者必须花费足够的时间和精力，认真地准备好书面报告和口头报告。市场调查书面报告有不同的类型和格式，但一般情况下都包含以下几个部分（详见 第十五章）：

- 1-报告的摘要。要简短明了，一般包括调查的主要发现、主要结论和建议。
2. 报告的详细目录。包括正文和附录的每一部分的大、小标题。
3. 报告的正文。一般包括调查的基本情况、主要发现、对结果的讨论、总结和建议等。
4. 报告的附录。一般包括问卷、图表、技术细节说明、实施细节说明等。
5. 追踪与反馈：提出了调查的结论和建议，不能认为调查过程就此完结，而应继续了解 其结论是否被重视和采纳、采纳的程度和采纳后的实际效果以及调查结论与市场发展是否一 致等，以便积累经验，不断改进和提高调查工作的质量。

### 第三节

## 市场调查的使用范围

市场调查广泛使用于市场活动的多个方面，如通过市场调查获取相关领域的全面而系统 的定量与定性资料、动态变化规则、行为与活动间的关联规则等，为相关决策提供重要支持。 具体使用领域有以下 7 个方面。

### 一、 市场研究

在市场研究中，需要通过调查和分析，估计某类产品（或服务）市场的现有规模和潜在规模，预测该市场的近期（有时需要预测中期或远期）需求，估计该类产品（或服务）各品牌的市 场占有率；对目标市场进行进一步的细分；了解消费者对该类产品（或服务）的消费形态及变 化趋势；掌握有关该市场的促销和销售渠道方面的信息；等等。

### 二、 消费者行为研究

消费者行为研究主要是指对消费者的购买行为进行调查和分析。一般首先需要了解八个 方面的信息，即所谓 6W 加 2H：购买什么（what）？为什么要购买（why）？购买者是谁（who）？何时购买（when）？何 处购买（where）？采取何种途径（which）？购买多少（how much）？如何决 策购买（how）？同时还需要进一步了解有关消费心理和消费观念方面的信息等。

### 三、 品牌或企业形象研究

品牌或企业形象研究主要包括以下几方面的内容：品牌或企业的知名度；品牌或企业的美 誉度；对品 牌或企业的认知程度及认知途径；品牌或企业的基本形象和具体形象；评价品牌或 企业的指标及指标的相 对重要性；对品牌或企业的名称、标志或商标的联想和印象；品牌的管 理和品牌的忠诚度等。这类研究除了针对品牌形象和企业形象外，有时还会涉及产品类别形 象和品牌使用者形象等。

### 四、 消费者满意度研究

在消费者满意度研究中，需要调查消费者对有关产品或服务的整体满意度；对产品或服务 的各个方面 的具体满意度；满意的 原因或不满意的 原因；对改进产品或服务质量的具体建议； 对各竞争对手的满意度 评价的比较。此外，还往往需要了解有关消费者的生活方式和消费观

念方面的信息等。

## 五、产品研究

产品研究主要是指新产品的开发研究，也叫产品测试，包括对现有产品的开拓或改造。首先需要调查了解消费者对产品的概念理解；对产品的各个属性的重要性的评价；对各种属性水平组合所形成的产品的偏好；在此基础上做进一步的定量分析，以寻找产品属性水平的最佳组合；估计产品的预期市场占有率。产品研究还包括产品的概念测试、定价研究、名称研究、包装研究、家庭产品测试，等等。

## 六、广告研究

广告研究包括为广告创作而做的广告主题调查和广告文案测试，为选择广告媒体而做的广告媒体调查、电视收视率调查、广播收听率调查、报纸或杂志阅读率调查、新媒体浏览及点击率调查；为评价广告效果而做的广告前消费者的态度和行为调查、广告中接触效果和接受效果调查、广告后消费者的态度和行为跟踪调查；为了解同行竞争对手的广告投放情况而做的电视、广播、报纸、杂志、新媒体等的广告媒介监测；等等。

## 七、营销环境研究

营销环境研究一般要结合外在环境研究和内在环境研究的结果进行深入的分析。外在的环境研究主要是指市场研究、消费者研究、竞争对手研究、销售渠道研究、零售店研究（包括零售店的分布、规模、销售的商品种类、数量、型号、品牌、包装、设计、价格）等；内在的环境研究主要是针对企业自身的管理和销售情况（包括商品与服务的数量、质量过程控制、价格、销售网店分布、广告投放、营销人员的能力和素质、企业雇员满意度等）进行分析研究。

实际中的市场调查项目，可以是上述单一问题的调查，也可以是上述多个问题的综合调查，要依据解决问题的需要而定。

# 第四节

## 市场调查需遵循的道德伦理

市场调查道德是调整市场调查所涉及的各行为主体之间关系的行为规范的总和。在市场调查过程中，涉及的行为主体包括市场调查使用者、市场调查提供者和被调查者。对于不同行为主体在市场调查中的道德问题，各国的行业组织都制定了相应的行为准则，如《美国市场调查协会道德准则》《欧洲民意和市场研究协会关于市场和社会研究国际准则》《中国信息协会市场研究业分会管理办法（章程）》等。这些准则或章程在一定程度上规定了市场调查主体的行为，成为市场调查从业者的道德准则，成为维护市场秩序和利益保障的基本规范。下面简要介绍针对不同的市场行为主体的不同要求。

### 一、市场调查提供者的道德规范

市场调查提供者指承担市场调查任务的组织和个人。调查提供者所提供的信息对调查使用者的决策起着至关重要的作用。调查提供者在提供市场调查服务时，必须遵守以下几方面的道德规范。

#### 1. 信守合同

合同是市场调查主体在自愿的原则下，为了维护各自的合法权益，明确各自的权利和义务 关系而签订的契约。调查提供者一旦与使用者签订了调查合同，就应该按照合同的要求按时、 保质完成调查任务。调查提供者不应寻找理由无故拖延调查时间，随意更改或减少调查内容， 以免对调查使用者的决策带来不利影响。

#### 2. 实事求是

调查提供者必须严格控制收集信息的过程，避免带着某种预先设定的目标确定调查范围。在收集信息时应避免使用职业被调查者，不能为了迎合使用者的口味而不顾事实；更不能为了 省事和节约开支而随意减少调查样本数量或人为修改、编造数据资料；在资料处理和分析过程 中，应避免片面追求方法的复杂性而不考虑研究对象的特点。

#### 3. 公平交易

坚持公平交易，首先要坚持公平竞争，反对把同行视为敌人。其次要坚持平等自愿、等价 交换的原则。市场调查的提供者和使用者应按照等价交换的原则，平等协商费用水平 Q 调查 提供者在报价时，应使使用者了解调查项目的难度，并对调查的内容进行合适的界定，避免为 了提高价格而夸大项目的难度和扩大项目的研究范围。同样，对于所需调查费用，提供者不应 以任何理由克扣，以保证调查工作的质量。

#### 4. 为客户保密

市场调查一般都是非标准化的定制服务，不允许同时为处于竞争状态的使用者或有利害 关系的使用者提供同种类型的调查服务。调查提供者不能利用使用者之间的竞争关系达到自 己的目的，更不能搬弄是非，侵害使用者的利益。对于市场调查结果，调查提供者必须为使用 者严格保密，在保密期限内，未征得同意，不得向任何第三方泄漏相关的调查结果。

## 二、市场调查使用者的道德规范

市场调查使用者是指为了解决特定问题而需要了解和掌握相应的市场信息，产生市场调 查需求，提出市场调查任务和要求，并承担市场调查费用，最终使用市场调查结果的政府、企 业、社会组织和个人。市场调查使用者除了同样必须信守合同、坚持公平交易外，还应遵守以 下道德规范。

1. 调查使用者不应以市场调查为由误导公众，不应借助市场调查的方式或借助市场调查 机构和人员实现某种不能公开的目的，不应借助市场调查进行不正当竞争。
2. 调查使用者应与提供者保持一种开诚布公的关系，应该把委托进行调查的真正目的， 所需解决的真实问题，所受到的时间、费用、资源等方面的限制如实地告诉调查提供者。
3. 尊重市场调查提供者和被调查者的意愿并保护其利益。
4. 调查使用者应该恰当地使用市场调查提供的信息资料，不应对市场调查提供的信息随 意进行夸大、断章取义、篡改以及曲解。
5. 调查使用者不应要求调查提供者从事正常市场调查以外的事情，特别是不能要求提供 者收集竞争对手的商业秘密以及政府及相关组织机构尚未公开的情报、信息、资料。
6. 调查使用者应该尊重调查提供者的劳动，不能以委托市场调查为借口，诱使调查提供 者为其提供免费的信息资料、调查方案，或达到获取免费咨询的目的。

7. 调查使用者有权从调查提供者处获得相应的信息资料、调查报告和有关建议。信息资料、调查报告和有关建议仅作为决策的参考，真正的决策由调查使用者做出，决策的后果由调查使用者承担，不能要求调查提供者承担决策责任：

### 三、被调查者的道德规范

被调查者是市场信息的直接提供者。在市场调查过程中，被调查者的权利，如是否接受调查的自主权、隐私权、所花费的时间和精力的补偿权等必须得到尊重。被调查者一旦同意接受调查，则也应遵守相应的道德规范。

1. 被调查者应诚实地提供有关信息，并对所回答的问题和所提供的信息的真实性承担道德责任。
2. 被调查者有责任对在接受调查过程中涉及的调查使用者的商业秘密加以保密，不得随意泄漏给第三方。

### 篇章小结…

市场调查是指在市场活动中，以特定研究目的为指引，有计划地收集、整理和分析市场的信息资料，提出解决问题的方案及建议的一整套科学方法、技术及规程。市场调查可以按照不同的标志进行分类。进行市场调查有着基本的程序规程，市场调查的方法和技术使用范围广泛，几乎涉及营销决策的方方面面。市场调查在发现问题、识别机会、理解市场、制定营销策略、辅助决策和实施管理控制上具有重要作用，但也有其局限性。

市场调查道德伦理是调整市场调查所涉及的各行为主体之间关系的行为规范指南，也是保护市场调查行业正常发展的行规，市场调查道德规范包括调查提供者道德规范、调查使用者道德规范和被调查者道德规范。关键词：…

市场调查探索性调查 描述性调查 因果性调查 定性调查 研究

定量调查研究 一手资料调查 二手资料调查

### 肯思考题



1. 什么是市场？什么是市场调查？市场与市场调查有何关系？
2. 举例说明影响中国某一具体商品与服务市场发展的因素有哪些？
3. 市场调查有哪些类型？
4. 一手资料调查与二手资料调查有何不同？
5. 如何理解企业发展与市场调查之间的关系？
6. 举例说明市场调查的作用，
7. 简述市场调查的基本程序。
8. 如何理解市场调查从业者应该遵循的道德伦理？
9. 到图书馆查找因为进行了市场调查活动而取得成功的实例（个人或企业的均可，但要注意是关于市场调查的实例）；或者相反，查找一个因为没有进行科学的、系统的市场调查而导致失败的例子。组织班级交流。

### 本章实训. …

#### 一、 实训目的

1. 培养团队意识和协作精神；
2. 掌握合理选择市场调查方法的依据；
3. 明确各种类型市场调查的适用场景；
4. 了解目前国内外主要的市场调查机构及其业务范围。

## 二、实训内容

1. 以团队为单位完成实训任务。
2. 学习使用搜索引擎，认知国内外重要调查公司。.

## 三、实训组织

1. 把全班分成多个小组，每组 4~6 人，推选一名组长。组长负责组织本组的讨论，围绕文本撰写、演示稿制作及演讲进行分工。
2. 演讲内容。具体包括：
  - (1) 列举当前国内外 5 家重要调查公司的网址及其名称 Q
  - (2) 各调查公司主要业务类型。
  - (3) 重要案例。
3. 各组交流。围绕演讲内容进行互动交流、质询和提问。理解市场与市场调查的关系。
4. 教师点评并确定实训成绩



## · 导语

、 — — . • • •

无论是为了获取市场调查项目还是接受委托方调查项目后的实施，都需要一份详略得当的市场调查设计方案，以保障市场调查的可行性。

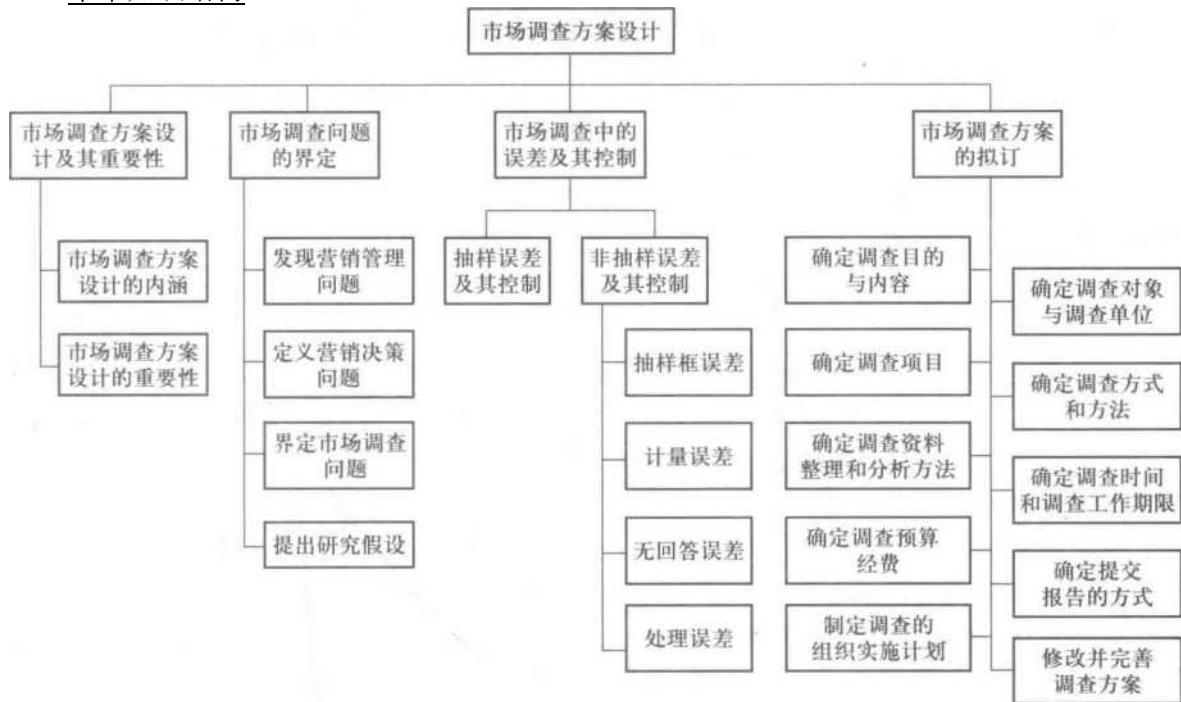
例如，需要了解大学在校学生对 BAT 移动支付使用意愿，如何进行市场调查方案设计？

一份高效的市场调查设计方案，需要根据研究对象的特点确定所需要的信息，包括信息类型、信息来源、信息收集方式、信息代表性、信息可靠性和信息的有效性。本章主要介绍市场调查方案设计的重要性、市场调查问题的界定、调查中的误差及其控制、市场调查方案设计的步骤与内容。精心设计市场调查方案是成功实施市场调查项目的重要保障。

## 、 教学目标与要求

1. 理解市场调查方案设计的含义；
2. 理解并知晓市场调查方案设计的重要性；
3. 掌握确定营销管理问题的方法；
4. 掌握市场调查问题的界定方法与技巧；
5. 掌握提出假设的思路与方法；
6. 掌握市场调查方案设计的过程、内容与技巧。

## 、本章知识结构



## 市场调查方案设计及其重要性

市场调查活动因其目标明确，具有一定的期限，遵循一定的流程与规范，影响因素复杂，所以需要以一种文本形式进行规范与安排，以提高工作效率，这就体现在市场调查方案设计的重要性上。

### 一、 市场调查方案设计的内涵

市场调查方案设计，就是根据调查研究的目的和调查对象的性质，在进行实际调查之前，对调查工作总任务的各个方面和各个阶段进行的通盘考虑和安排，并形成相应的调查实施方案，制定出合理的工作程序。无论是接受委托的市场调查，还是作为企业内部常规工作的一定规模的市场调查，都需要设计市场调查方案。

一定规模的市场调查都会涉及相互联系的各个方面和全部过程。不仅调查全过程的每一个环节相互衔接，不能错位，而且，市场调查内容上广泛而深入，全面而系统，需要进行整体考虑，避免调查内容上出现重复和遗漏整个过程只有事先做出统一考虑和安排，才能保证调查工作有秩序，有步骤地顺利进行，减少差错，保障高质量与高效率。

### 二、 市场调查方案设计的重要性

市场调查是一项复杂、严谨、技术性较强的工作，一定规模的市场调查参加者众多，为了在调查过程中统一认识、统一内容、统一方法、统一步调，圆满完成调查任务，就必须事先制定出一个科学、周密、可行的工作计划和实施措施，以使所有参加调查工作的人员都依此执行。市场调查方案设计的重要性体现在以下方面：

#### （一）调查方案设计是市场调查项目是否立项或能否委托的重要依据

调查方案设计围绕需要解决的调查主题顶答打算做什么，能够做什么，准备怎么做的问题。它是受托方承担能力的一种书面体现。当有多家调查机构竞标时，各机构提供的设计方案成为委托方选择的重要依据

#### （二）调查方案设计是市场研究的第一步

调查方案设计是从定性认识过渡到定量认识的开始阶段任何调查工作都是先从对调查对象的定性认识开始的，没有定性认识就不知道应该调查什么和怎样调查，也不知道要解决什么问题和如何解决问题。例如，要研究某一行业的生产经营状况，就必须先对该行业生产经营活动过程的性质、特点等有详细的了解，设计出相应的调查指标以及搜集、整理调查资料的方法，然后再去实施调查。可见，调查方案设计是定性认识与定量认识的连接点 r

在调查实践中，调查方案设计适应了现代市场调查发展的需要。现代市场调查已由单纯的搜集资料活动发展到把调查对象作为整体来反映的调查活动，与此相适应，市场调查过程也相应扩展为包括调查方案设计、资料搜集、资料整理和资料分析的完整工作过程，调查方案设计正是这个全过程的第一步。

#### （三）调查方案设计是市场调查全过程的计划书与活动指南

在市场调查活动中，调查方案设计起着统筹兼顾、统一协调的作用。现代市场调查是一项

复杂的系统工程，对于较大规模的市场调查，尤其如此。在工作中会遇到很多复杂的矛盾和问题、其中许多问题属于调查本身的问题，也有不少问题并非是调查的技术性问题，而是与调查相关的问题。例如，抽样调查中样本量的确定，按照抽样调查理论，可以根据允许误差和置信度大小，计算出相应的样本量，但这个样本量是否可行，要受到调查经费、调查时间、回答率等多方面条件的限制。同样，只有调查实施过程中许多前期准备工作的安排都一一计划好，才可能取得好的调查效果。因此，只有通过调查设计，设置调查流程，将各种保障条件与影响因素综合处置，才能分清主次，明确责任、任务与分工，根据需要和可能采取相应的方式方法，使调查工作有序进行。这样，调查方案设计实质上成了市场调查全过程的计划书与活动指南。按照计划书内容分步实施的市场调查活动，可以高效率地实现研究目标，并达成预定结果 C，

## 第二节

### 市场调查问题的界定

市场调查的目的是通过对资料的收集、整理和分析，明确企业营销活动中存在的机会或问题，寻找利用机会的可能性及解决问题的可行方案。显然，市场调查方案设计的核心是发现营销管理问题并准确界定市场调查问题。

即从发现营销管理问题开始，通过一系列信息分析定义营销管理问题，进一步确定市场调查问题，并通过研究假设，建立市场调查问题与营销管理问题之间的联系，通过市场调查为营销管理问题的解决提供科学依据。

准确界定市场调查问题，实际就是界定调查所应收集信息的广度与深度，这是市场调查能否取得成效的重要一步：

#### 一、发现营销管理问题

所谓“营销管理问题”，是指决策者无法确定究竟应采取什么措施来应对市场营销活动中出现的各种机会和明显的（或可以预计的）困难。对于一个营销管理问题，如果市场营销者仅有一种选择或根本没有任何选择，那么实际上也就不存在问题，自然而然地形成一种“自觉”的决策。但市场营销中面临着多种状态的决策和可能产生的结果，企业各自的产品与服务竞争优势各有不同，需要决策者在调查了解各种状态的效益和风险的基础上，进行权衡后做出决策。机会与问题的界定通常需要回答以下四个问题：

1. 对应于市场营销活动的机会与困难是什么？（发现营销管理问题）
2. 与问题相关的信息有哪些？基于这些现有的信息能否进行决策？
3. 在信息不充分时，还需要收集哪些信息？（对应市场调查问题）
4. 利用这些信息确实可以解决经营管理中的问题吗？

确定市场调查问题的基本分析框架见图 2-1，这个分析模式阐明了发现与界定市场调查问题的一般过程：

#### 二、定义营销决策问题

营销管理问题是决策者需要做什么的问题，是一种行为导向，关心的是决策者采取什么行

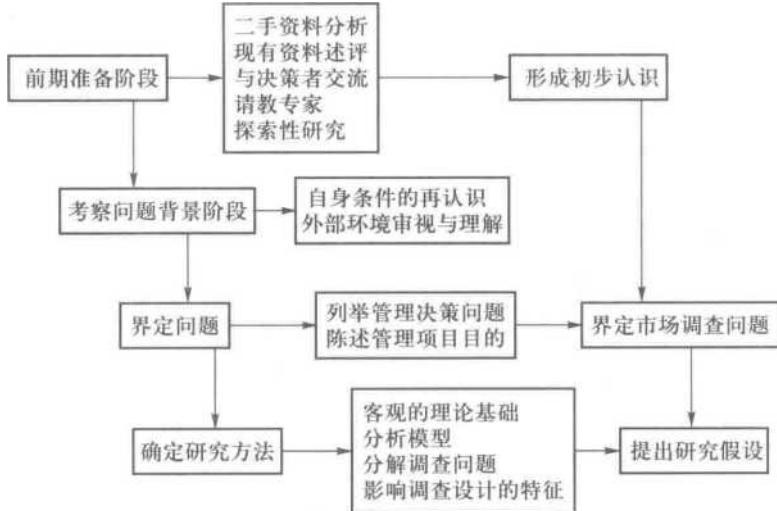


图 2-1 确定市场调查问题的基本分析框架

动步骤才能达到目的。例如，怎样才能收回失去的市场？市场是否可以细分？是否应该投放一种新产品？是否应该增加促销预算？明确营销管理问题的过程，实际上就是对企业市场营销活动出现的机会或威胁进行识别判定的过程。一个企业需要进行决策的问题很多，一定要从企业营销活动的实际出发，在对企业的生产经营状况进行全面分析研究后确定。调查者与客户或决策者进行充分的沟通是成功界定营销管理问题的基础。

明确营销决策问题一般遵循以下七个步骤。

### (一) 熟悉公司、产品和市场的背景

调查者应该熟悉客户的以下情况：

- (1) 长期以来的经营业绩；
- (2) 经营的历史、组织结构、所有权情况、企业目标；
- (3) 产品生产、产品使用、产品特点、产品优势；
- (4) 过去的营销方案，包括产品的价格、分销和促销的具体措施、手段等；
- (5) 顾客的构成及消费行为；
- (6) 竞争对手及市场趋势；
- (7) 相关财务情况。

### (二) 了解决策者的环境、目标及资源

营销管理者都是在一定的环境下做出决策，可能受公司董事会及企业现有各种资源的制约，也可能存在任期内经营目标、时间等条件的限制。例如，广告投放决策面临资金短缺、时段紧张等约束，决策者需要的信息与调查者理解的信息可能会有差异，在此情况下，调查者需与决策者进行交流，理解决策者的意图及决策环境，从而提出有利于搜寻决策者所需信息的调查问题。在预算紧张的情形下，可把迫切需要解决的问题拟订为调查项目。

### (三) 认识问题的征兆

决策者也许只知道公司市场占有率下降了而不知道具体的原因，也许只观察到公司经营不佳的症状但却不知道具体导致这种情况的根源，或许对这些问题的征兆也只是有一种主观的感觉或模糊的概念，调查机构应该帮助决策者“会诊”问题的原因，并分析提出可能的潜在“病因”。问题的征兆，通常采用相关指标的实际完成情况与预定目标之间的差异来测定，它

也可能是指某些市场环境因素的改变对公司带来的机会或造成的不利影响。这些相关指标包括销售量、销售订单、利润、市场占有率、竞争者、顾客投诉等，公司必须监测这些指标的变化、调查机构通过了解这些问题的征兆，确认实际偏离目标的方向和程度。

#### （四）探查问题的可能原因

问题总有其产生的具体原因。也许决策者对此已有所认识，也许只是一种猜测，但他们的看法可能会偏离主要原因，带有强烈的主观色彩。调查机构应该客观地分析，列出各种可能的原因。如果某个可能的原因在征兆出现之前或同时有所改变，则这个原因很有可能就是引起这种征兆的原因；反之，如果某个可能的原因在征兆改变过程中并没有变化，则这个原因未必就是问题产生的原因。例如，某公司的市场占有率下降了，可能的原因是竞争者采取了更高明的促销手段，或者是竞争者采取了先进的生产技术降低了成本获取了价格优势，或者是本公司分销体系不顺畅等调查机构需要尽可能列举出所有可能的原因，找出主要原因，准确地界定所要决策的问题。

#### （五）提出可能的问题解决方案

调查机构应与客户共同研讨，列出一些运用公司现有资源可能解决当前问题的方法。例如，营销策略的调整，可能会从产品改良、价格调整、加强渠道管理及改变促销方式等方面入手来设计新的营销方案，以解决问题或减少损失。

#### （六）预测这些举措的可能结果

通过对可能有助于解决问题的举措的预测结果进行分析，能够使调查机构明确解决方法是否正确。例如，在营销方案设计中调查机构应针对每一种营销行动的可能结果做出分析判断，设想可能产生的所有结果（有利的或不利的）。例如，如果实施某项特定的营销活动，那么它对于当前问题及整个营销计划的影响如何？如果实施了某项用于处理现有问题的计划，那么会产生哪些附加问题？

#### （七）评估已有信息的状况

调查机构必须评估已存在的信息状况，明确决策者已掌握的信息与期望获得的信息，分析二者的信息缺口，根据信息缺口确定市场调查目标。许多客户只知道销售额下降了，市场占有率减少了，但他们并不清楚产生这种现象的具体原因。因而，调查机构需要花更多的时间与客户进行大量的信息沟通，与企业营销经理共同探查“病因”，明确营销决策的关键问题。

### 三、界定市场调查问题

市场调查问题是指确定需要什么样的信息和如何有效地获取这些信息，以帮助决策者解决营销管理问题（避免损失、捕捉机会）。如果说营销管理问题关注管理决策的结果，是行为导向，那么市场调查问题则是信息导向。在了解营销管理问题的基础上，将营销管理决策问题转化为市场调查问题有以下三个过程。

#### （一）阐述构想与操作定义

构想是指对决策者所面临的营销管理问题进行思考的过程，很多时候受到营销管理问题支撑理论如营销学、管理学、行为学等的指引。

而操作层面的定义则是指把构想转化为市场调查问题的设计过程，即用来收集与现有构想有关的特定问题形式，体现为问卷的设计过程。例如，一种行为改变后面对应着一组6W2H的结构化设计。往往一个构想不一定只有一种操作定义，可能会有若干种操作定义，需要进行比较与选择。

## (二) 明确构想间的相关性

由于解决营销问题的复杂性与可选择性，与之对应的市场调查问题就不可能是单一的，很多时候是问题的组合，并体现在问题设计的关联性与结构性上。因此，在问题设计阶段，调查机构必须明确各种构想之间的相关性，并确认相关性所体现的特征、因素、产品特性及假设。这样可以使所需信息更加明确，有助于市场调查问题的正确设计。

## (三) 建立调查问题框架

当调查机构把一系列构想以逻辑关系联系起来，就可构建起调查问题的分析框架。这种框架有时可能是复杂的，但大多数时候它们是简明的，其主要构成要素包括研究的相关理论（如行为学理论等）、模型假设、调查目标分解、影响调查设计的因素分析等。把管理决策问题转换为市场调查问题的举例说明见表 2-1。

表 2-1 管理决策问题转换为市场调查问题举例

管理决策问题	市场调查问题
是否应该引进新产品	针对提议的新产品确定消费者偏好和购买意愿
是否应该改变广告活动	确定现有广告活动的效果
某种品牌的价格是否应该提高	确定需求的价格弹性、确定在不同水平上价格变化对销售额和利润的影响

## 四、提出研究假设

假设是指研究人员对所感兴趣的某个要素或现象未经证实的命题。假设可以是对两个或多个变量之间关系的描述，而变量则由基本理论或分析模型确定。在市场调查中，假设可能是研究问题的一个答案，也可能是实施某种举措后所发生的反应。作为科学研究，不但要了解实施后的结果，而且要知道产生这种结果的可能性。问题出现征兆，分析了可能的原因和对策，设想了执行对策及其结果，但实际情况可能未必如预料的那样。因为还存在很多不确定的因素，决策者不能确定的关键性假设可能会成为研究问题提出的主要原因，调查将有助于消除或减少这种不确定性。当公司的管理层对假设存在较大分歧时，调查将有助于确定何种假设是正确的，从而找出解决问题的真正方法。提出假设有助于调查机构确定研究方向、变量数及其实相关性、设计研究方案和导出研究结果。一个好的假设，应该目的明确、条件清晰；能够指明变量的分布与大小；能够指出需要解释现象的数据资料，进行统计检验；基于样本特征进行的推断，应具有较高的有效性。

提出假设的步骤如下：

第一步，明确需要检验的内容。它可以是单一变量的变化方向与大小；多个变量之间差异的比较；两个变量之间的作用关系；一组变量之间的作用关系（重要程度）；变量分布的特征；等等。

第二步，合理建立零假设与备择假设。把需要充分证据支持的结论（预期结论）放在备择假设上，而其对立面设为零假设。

第三步，构建一个可解释的调查问题分析框架；并有足够的数据支持这一分析。

第四步，掌握假设检验的分析技术。

## 第二章 市场调查方案设计

### 第三节

## 市场调查中的误差及其控制

在调查方案设计阶段，应尽力预测可能发生的调查误差，并采取适当的措施在调查活动的全过程中加以控制，使其达到最小。调查误差根据其误差来源不同可以分为抽样误差和非抽样误差两大类。下面分别介绍。

### 一、 抽样误差及其控制

抽样误差是指完全由抽样引起，根据样本调查数据来估计总体特征所产生的误差。因为样本是随机抽选的，样本的估计值会随着抽选样本的不同而变化，即使调查完全准确，它和总体真值之间也往往存在差异，这种差异又称为代表性误差。

在调查方案设计时必须将抽样误差的大小向调查数据的所有潜在用户提供。对于概率抽样调查，可以计算抽样误差。这将在以后介绍。抽样误差的大小，取决于下列因素：样本量的大小、总体指标的变异程度、抽样设计、估计方法以及总体大小。一个好的抽样设计方案可以在相同的样本量下减少抽样误差。因此，在调查设计阶段，需要充分了解总体特征，如果采用抽样调查，则需要尽可能地设计一个符合研究目的的使抽样误差更小的抽样方案。

### 二、 非抽样误差及其控制

非抽样误差是指除抽样误差以外的，在几乎所有市场调查活动过程所产生的误差。非抽样误差可以分为两类：随机的和非随机的。随机的误差，如果样本足够大，这类误差的影响可以忽略，随机误差只增加估计值的变异性，而非随机误差又称系统误差，这种误差倾向于同一方向，对所有样本单元的误差累加起来导致最终结果的偏差。因此，在调查设计中需要重点关注非随机的非抽样误差。它主要有以下四种来源：

#### （一） 抽样框误差

抽样框误差是指抽样框中的信息存在缺陷或者错误而产生的误差，也称为涵盖误差。当抽样框不完备，总体单元有遗漏时，就出现不完全覆盖误差。而当抽样框包括非目标总体单元时，就出现过覆盖误差。这些都会导致估计的偏差。在调查设计时，应对抽样框带来的误差有一个大致的判断。

#### （二） 计量误差

计量误差是指对一个问题所做的回答记录与真值不同而形成的误差。它可能来源于调查问卷设计、调查员、被调查者、收集数据方式或测量工具。

被调查者和调查员对问题的误解是计量误差产生的主要原因之一，误解的产生可能是因为问卷设计有问题，如使用了专业行话，概念模糊，措辞不当等。

数据收集方式也可以影响计量误差。例如在面访调查中，当使用经过培训的调查员时，会比邮寄调查的计量误差小。在直接测量调查中，调查员通过观察或者使用测量工具收集数据会产生误差。比如在健康调查中测量人的体重时，用于记录被调查者体重的台秤可能因校正有误而显示不准确的数字。计量误差有时也称回答误差，但是两者含义并不完全相同。计量误差包括由于调查员工作中的错误而造成的误差，也包括被调查者有意或无意的错误回答。

在实际操作中往往是很难严格区分各种计量误差的。例如由于问卷设计上的失误造成被调查者对题意的误解，既可以认为是问卷设计误差，也可以认为是被调查者回答误差。计量误差能影响非抽样误差大小并造成估计的偏差。如果计量误差是随机的，则会增大估计值的方差，从而降低调查的精度。另一方面，如果计量误差系统地偏向某个方向或某个类别，就会产生估计的偏差，对调查结果形成误导。例如，在测量学生身

高时，如果调查员要求学生都穿着鞋站在 测量仪器上测量，这样就会产生系统误差，所有的身高都会被高估 c

### （三）无回答误差

无回答是指调查时未能从指定的被调查者处获得有效的回答。无回答有两种类型：一种 是全部无回答，又称单元无回答，它是指被调查单元没有提供任何信息。另一种是部分无回答，又称项目无回答，它是指问卷中某些问题回答的空缺。

无回答的原因有多种，有一部分是属于随机性的，另一部分是属于系统性的。下面是关于 一些无回答情况的原因举例，在每一种情况中，都有随机性和系统性两种可能

#### 1. 地址错误

由于地址错误使调查员无法找到预定的被调查者，从而失去了这一样本可能提供的信息 如果地址错误是一种偶然的情况，则这种误差是随机性的。但如果是某一类被调查者的地址 出现了错误，则这种误差可能是系统性的。例如，由于部分教师集体在某一新区购买了新的住 房，造成抽样框名单上该部分教师的有关信息与其他的总体单元不同，这种误差就是系统 性的。

#### 2. 被调查者不在家

调查员虽然找到了被调查者的地址，但家里没人或者指定的被调查者不在家，造成数据缺失。如果这样的被调查者与其他的单元没有明显的特征，这种无回答是随机性的，但如果是调 查设计的原因使某一类人恰好在调查时不在家，而这类人与其他单元又有明显的特征差异，那 么这种无回答误差是系统性的。例如，设计在周末的休息时间进行调查，而很多较富裕的人喜 欢在假日外出度假。如果调查的内容与收入水平有关，其误差也将是系统性的

#### 3. 无法联系

抽样框中缺乏电话号码和通信地址，电话调查和邮寄调查就无法进行。在面访调查中，由 于调查员的口音或被调查者能力上的缺陷，也会造成联系上的困难。

#### 4. 拒访

拒访是无回答中最常见的一类，包括在调查员上门调查时不开门或拒绝合作。拒访的原 因也是多方面的，有的是由于社会治安原因拒绝接受调查，有的是调查的内容引起被调查者的 反感而不愿意配合。

#### 5. 遗漏

调查员在调查时遗漏了需要调查的某些人，有时虽然取得调查结果，但因调查员的大意丢 失了问卷，这也相当于无回答。

项目无回答主要发生在被调查者不知道答案，拒绝回答某个问题，忘题，忘了回答，或者错 误地跟随问卷的进程漏了回答。

无回答的产生也有调查员的原因，例如，调查时间过长，问卷中某些问题的排列不合逻辑 使被调查者失去耐心，这种情况会使被调查者在没有完成调查之前就终止回答。有些调查员 没有按照调查的规定，或者没有按原文读出。调查员与被调查者没有建立良好的关系，不适当 的追踪回访等都会造成无回答。

#### (四) 处理误差

处理是指将收集的调查结果转化为适合于列表及进一步数据分析需要的形式，它包括数据收集以后、进行估计以前，所有对数据的处理，包括数据的编码、录入、审核及插补等。其中任何一个步骤都有可能发生错误，这种在数据处理环节形成的误差称为处理误差。

数据编码是将回答的结果转化为数值以便于录入及进行一般处理的过程。编码是给每一个回答结果(例如职业)赋予一个代码，或者将回答结果与一组代码相对照，然后挑选一个最接近回答结果的代码。对于封闭式的问题，即对问题回答的所有选项事先已经确定，代码在访问前已经编好；对于开放式问题，即被调查者用自己的语言表述答案，一般采取手工编码。对开放式问题用手工编码时，需要对回答结果进行判断，然后编码，所以容易产生误差。例如对同一个答案可能赋予两个不同的代码。缺乏经验或没有受过正规训练的编码员特别容易犯编码错误。如果采用自动编码，假如程序有误或者程序没能将所有现有信息考虑在内，也会导致误差的产生。

数据录入误差指原始问卷中的数据没能准确地录入到计算机里所引起的误差。这可能是由于字母与数字符号组合造成的数据复杂结构以及所录入信息的含糊不清造成的。调查问卷本身的结构和编码文件也会导致录入误差。

由于审核和插补这两个过程是密切联系的，因此审核误差与插补误差通常同时发生。审核是指对数据记录进行检验以识别出那些有可能导致错误的缺失数据、无效数据和不一致数据的过程；插补是为了解决在审核过程中检查出来的缺失数据、无效数据和不一致数据的修补过程。

审核和插补误差是由于原始数据的质量低劣或者由于数据的结构过于复杂造成的。当审核和插补由计算机自动进行时，未经充分测试的有错误的程序会导致错误，插补方法不恰当也会导致偏差。另外，对数据的错误转换或把本来正确的数据改错当然也会导致错误。

对于随机的非抽样误差可通过控制样本量来减小，而非随机的非抽样误差，在设计时要周密考虑，并在调查的每一环节中计划好质量控制的方法。

## 第四节

### 市场调查方案的拟订

在正式调查项目启动之前，调查机构为争取市场调查项目会撰写一份市场调查提案，它由调查目标确定、调查基本方案、调查费用预算和调查进度安排四个部分构成。这份提案是调查机构提交给客户审定调查目标、调查草案、调查费用和调查时间的重要文件，是委托方的立项依据，也是受托方保护自身合法权益，避免调查方案被窃的依据。市场调查提案与市场调查方案不完全相同，可以认为前者是调查的初步设想，后者是调查的具体实施计划，但在设计流程上具有一致性。本节介绍市场调查方案的设计内容。

市场调查的总体方案设计是对调查工作各个方面和全部过程的通盘考虑，包括了整个调查工作过程的全部内容。市场调查总体方案是否科学、可行，是整个调查成败的关键。其主要内容包括以下十个方面。

#### 一、 确定调查目的和内容

确定调查目的是指明确在调查中要解决哪些问题，通过调查要取得什么样的资料，取得这些资料有什么用途等问题。衡量一个调查设计是否科学的标准，主要就是看方案的设计是否能体现调查目的的要求，是否符合客观实际。

为了进一步明确调查目的，调查设计需要知道谁是数据的使用者，因为调查的目的服务于他们。要把

调查的目的搞清楚，需要不断地与客户进行沟通与交流。因为客户提出的需求，开始往往是含糊的，需要加以提炼和进一步明确。调查目的明确了，调查的内容也就大致确定了。

## 二、确定调查对象和调查单位

明确了调查目的与内容，就需要确定调查对象和调查单位，即解决向谁调查和由谁来具体提供资料的问题。调查对象是根据调查目的、任务确定的调查范围以及所要调查的总体，它由某些性质相同的许多调查单位组成。调查单位或调查单元是所要调查的社会经济现象总体中的个体，即调查对象中的一个个具体单位，它是调查实施中需要具体回答各个调查项目的承担者。

在确定调查对象和调查单位时，应该注意以下三个问题：

第一，准确界定调查对象的内涵与外延。由于市场现象复杂多变，因此，调查对象也是比较复杂的，必须以科学的理论为指导，严格规定调查对象的含义，并指出它与其他有关现象的界限，以免造成调查实施时由于界限不清而发生的差错。例如：以“在校大学生”为调查对象，就应明确“在校大学生”的含义，即明确是普通高等学校类型、只包括本科一至四年级的在读学生等。

第二，要基于调查目的和调查对象明确调查单位。调查目的和对象变化了，调查单位也会随之改变。例如，要调查“家庭户”还是“家庭户中的每个人”的基本情况时，调查单位是不同的，前者为每一户，后者为家庭中的每个人。

第三，基于不同的调查方式确定调查单位。如果采取普查方式，调查总体内所包括的全部单位都是调查单位；如果采取抽样调查方式（绝大多数情况），则用各种抽样方法抽出的样本单位是调查单位，为此，要明确地给出具体的抽样设计思路。

## 三、确定调查项目

明确了整个信息需求以后，就要将调查内容进一步细化为调查项目。调查项目是指对调查单位所要调查的主要内容，确定调查项目就是要明确向被调查者了解什么问题，也是问卷设计的前期工作。调查内容需要通过指标加以反映，统计指标是说明客观现象数量特征的概念和数值，具有确定的统计范围和相应的计量单位，调查的时间范围，形成可操作的概念和定义。在确定调查项目时，除要考虑调查目的和调查对象的特点外，还要注意以下几个问题：

第一，确定的调查项目应当既是调查任务所需，又是能够取得结果的。凡是调查目的需要又可以取得结果的调查项目要充分满足，否则不应列入。

第二，项目的表达必须明确，要使问题备选结果具有确定的表示形式，如数字式、是否式或文字式等。否则，会使被调查者产生不同理解而做出不同的回答，造成汇总时的困难。

第三，确定调查项目应尽可能做到项目之间相互关联，使取得的资料相互对照，以便了解

现象发生变化的原因、条件和后果，便于验证回答的准确性

第四，调查项目的含义要明确，必要时可附加调查细目，

#### 四、确定调查方式和方法

在调查方案中，还要规定采用什么组织方式和方法取得调查资料。调查方式的选择，需要考虑到多种因素：要得到充分而全面的信息，全面调查可能更好或者是必要的，但它的工作量比较大，费用比较高，时间也可能比较长。而选择抽样调查的最主要原因是其能在充分满足客户所需信息质量的前提下，提供一种更快、更节省的方法。相对于全面调查来说，抽样调查规模小，因而更容易进行监控。在抽样调查和全面调查之间做抉择时，既要根据市场调查的目的与研究对象的特点考虑总体大小、小区域（范围）的估计、属性的多寡、调查误差、特殊要求等因素，也要考虑实施调查的费用、时效及其他因素等现实要求；但全面调查和抽样调查并不是完全对立的，根据具体调查的需要，可将这两种调查方法结合起来使用。例如，在工业企业的调查中，对大型企业进行全面调查，而对小型企业进行抽样调查。

在确定选择抽样调查时，要进一步明确目标总体与调查总体的关系；确定合适的抽样框及其类型，抽样框有缺陷时知道如何补救；明确抽样误差的影响因素及其控制方法。

搜集资料的具体调查方法有定性调查研究方法和定量调查研究方法。具体包括文案法、小组座谈法、深层访谈法、投影技法、询问法、观察法和实验法等（详见第三章）。每一种方法都具有各自的优点与不足，都存在合适的应用场景，在调查时，采用何种方法不是固定和统一的，而是取决于调查对象和调查任务的需要，尤其应注意多种调查方法的组合运用。

#### 五、确定调查资料整理和分析方法

采用实地调查方法搜集的原始资料大多是零散的、不系统的，只能反映事物的表象，无法深入研究事物的本质和规律性，这就需要对大量原始资料进行加工汇总，使之系统化、条理化。目前这种资料处理工作一般已可借助计算机进行，这在设计中也应予以考虑，包括确定是采用定性分析还是定量分析的方法；在使用定量分析时需采用何种操作程序以保证必要的运算速度、计算精度及特殊目的。

随着计量分析方法的发展和计算机的运用，越来越多的现代统计分析手段可供我们在定量分析时选择，如回归分析、相关分析、聚类分析等。每种分析技术都有其自身的特点和适用性，有对数据收集的相应要求。因此，应根据调查的需要，各类数据获取的可能，选择最合适的分析方法并在方案中加以规定。

#### 六、确定调查时间和调查工作期限

调查时间是指调查资料所属的时间。如果所要调查的是时期现象，就要明确规定资料所反映的是调查对象从何时起到何时止的资料。如果所要调查的是时点现象，就要明确规定统一的标准调查时点。

调查期限是规定调查工作的开始时间和结束时间。包括从调查方案设计到提交调查报告的整个工作进度，也包括各个阶段的起始时间，其目的是使调查工作能及时开展、按时完成。

详细的调查进度计划要使各个时期的工作大致平衡，避免出现前紧后松或前松后紧的现象，同时也有利于调查机构的管理部门依据进度计划表检查有关部门在各个时期完成进度的情况。调查进度计划的确定通常采用逆推的办法，即从最后的截止日期逆顺序地向前安排完。

成各项任务所需的时间，并推算出起讫的期限，落实承接相应工作的部门或个人。表 2-2 是 安排调查进度计划的一个示例。

表 2-2 xx 项目调查进度安排

序号	工作任务	计划天数	开始日期	结束日期	承接部门	完成情况
1	总体方案论证、设计					
2	抽样方案设计					
3	问卷设计、测试、修改和定稿					
4	调查员的挑选与培训					
5	调查实施					
6	数据的整理、录入和分析					
7	调查报告的撰写					
8	有关鉴定、发布会和出版					

## 七、确定调查预算经费

市场调查费用的多少通常视调查范围和难易程度而定。不管何种调查，费用问题总是十分重要和难以回避的，故对费用的估算也是设计调查方案时必须考虑的内容之一。根据拟定 的调查方案，可将调查项目分解成若干子项目，然后匡算各个子项目所需的时间和费用。调查 项目的费用预算包括直接和间接的人工、材料、交通、管理费用及其他一切费用开支。如果是 专业调查机构承接的调查项目，还应包括服务费——佣金，通常佣金按占总报价的百分比计算。调查中的各种费用预算示例参见表 2-3。

表 2-3 调查费用预算表

序号	预算项目	摘要及说明	预算金额
1	调查方案策划费 与设计费用	二手资料收集、定性研究、实地调查设计	
2	抽样设计费	总体特征分析与历史数据采集、抽样设计	
3	问卷设计费	问卷设计与测试费	
4	问卷印制费	印刷、装订费	
5	调查实施费	预调查费、培训费、交通费、调查员和督导员劳务费、礼品费 和其他费用等	
6	数据编码、录入费	复核、编辑、计算机编码、录入	
7	数据统计分析费	描述统计分析、推断统计分析、假设检验与多元统计分析	
8	调查报告撰写费	报告撰写费用与印制费用	
9	设备使用费	含计算机及其他设备	
10	材料费用	计算机用纸及其他耗材	
11	管理费用	管理人员报酬	

续表

序号	预算项目	摘要及说明	预算金额
12	其他费用	不在以上项目中开支的费用	
13	成本合计	上述成本费用总和	
14	佣金	一般以总报价的 15%—100%测算	
15	总计	加总上述各项	

## 八、确定提交报告的方式

市场调查项目完成的最终成果之一是书面报告。确定提交报告的方式主要包括书面报告书的形式和份数，书面报告的基本内容、图表数量等。

## 九、制定调查的组织实施计划

调查的组织计划是指为确保实施调查的具体工作计划，主要是指调查的组织管理、调查项目组的设置、人员的选择和培训、调查的误差控制和质量保障措施等。其中，对调查员的培训包括解说问卷内容，分配调查对象，掌握访问技巧，明确工作进程及质量要求等。在市场调查中，为保证调查的顺利实施，保障调查质量，在方案确定后和印制调查问卷之前，可先从各类调查对象中抽取少量样本进行预调查。通过预调查，修改问卷，探索针对具体调查对象的访问技巧等，为全面推开调查做好准备。

## 十、修改并完善调查方案

一般在调查方案设计完成后与正式调查实施前会通过预调查的组织实施，检验方案设计的科学性与合理性，并使用其反馈结果修改和完善设计的调查方案。预调查法是指依据调查方案抽取较小规模样本进行实地调查并获取数据，基于调查过程与调查数据进行分析与评估的方法。小规模的样本称为试点，因而预调查法又称为试点调查。

预调查的任务主要有以下两个：

1. 对调查方案内容进行实地检验。调查方案的设计是否切合实际，还要通过预调查进行实地检验，检查目标制定是否恰当，调查指标设计是否正确，哪些需要增加，哪些需要减少，哪些说明和规定需要修改和补充。预调查完成后，要分门别类地提出具体意见和建议，使调查方案的制定既科学合理，又具有可操作性。

2. 作为实战前的演习，可以了解调查工作安排是否合理，哪些是薄弱环节，进而针对实施过程中的不合理和薄弱环节进行修改与完善。

预调查应该注意以下几个问题：

其一，应建立一支精干有力的调查队伍，队伍成员应该包括有关负责人、调查方案设计者和调查骨干，这是搞好试点工作的组织保证。

其二，应选择适当的调查对象。要选择规模较小，代表性较强的试点单位。必要时可采取少数单位先试点，再扩大试点范围、然后全面铺开的做法。

其三，应采取灵活的调查方式和方法。调查方式和方法可以多用几种，经过对比后，从中选择适合的方式和方法。

其四，应做好试点的总结工作。即要认真分析试点的结果，找出影响调查成败的主观原因。不仅要善于发现问题，还要善于结合实际探求解决问题的方法，充实和完善原有调查方案，使之更加科学和易于操作。

一个设计好的市场调查方案在正式采用前还可以进行自我评价。具体内容包括：方案设计是否体现调

查目的和要求；方案设计是否科学，完整和可操作；方案设计能否使调查质量有所提高；调查实效检验，即通过实践检验调查方案的科学性。

### 第五章小结· · ·

市场调查从定义调查问题开始，经过设计调查方案、数据采集、数据分析等中间过程，最后向客户（需求方）提交书面调查报告和演示报告展示调查结果。市场调查方案设计在市场调查项目立项与实施中具有重要作用。它是组织市场调查活动具体实施的行动步骤与进度计划。市场调查的目的是通过对资料的收集、整理和分析，明确企业营销活动中存在的机会和问题，提出利用机会的可能性及解决问题的可行方案。营销管理问题是一种行为导向，关心的是决策者可采取什么行动步骤才能达到目的。市场调查问题是一种信息导向，确定需要■什么样 的信息和如何有效地获取这些信息，以帮助决策者解决营销管理问题（避免损失、捕捉机会）C 在市场调查中，研究假设可能是问题的一个答案，也可能是实施某种举措后所发生的反应。提出假设有助于调查者选择研究方向，确定变量的相关性，合理地进行研究设计和导出研究结果 C 设计市场调查方案包括十个组成部分：确定调查目的和内容、确定调查对象和调查单位、确定调查项目、确定调查方式和方法、确定调查资料整理与分析方法、确定调查时间和调查工作期限、确定调查预算经费、确定提交报告的方式、制定调查的组织实施计划、修改并完善调查 方案。

### 第五章关键词· · ·

调查方案 营销管理问题 市场调查问题 调查目标 研究假设

思考题.....



回%即测即评

1. 举例说明调查机构是如何运作一个市场调查项目的。
2. 解释并说明营销管理问题与市场调查问题的差异。
3. 举例说明行为问题如何转换为信息问题。
4. 举例说明如何界定市场调查问题。
5. 如何分析确认一个公司的营销管理问题？
6. 举例说明怎样陈述研究假设。
7. 市场调查方案包括哪些组成部分？各个部分如何设计？

### 第五章实训· · ·

#### 一、 实训目的

通过实训，使学生了解市场调查的全过程，掌握市场调查方案设计的基本过程，熟悉市场 调查方案的具体框架、内容及设计技巧◎

#### 二、 实训组织

以团队为单位完成实训任务 G

### 三、实训步骤

组建4~6人小组团队，各小组自选调查项目。要求必须选择熟悉的主题，如所在学校食宿条件、学生学习行为、在线教育、就业压力、学生移动支付使用等现实问题，而且还要考虑到后续学习中的抽样调查、问卷设计、数据采集及数据整理分析等环节的难易程度，不建议选择异地和不熟悉主题作为调查项目选题。

1. 各小组（扮演调查公司）拟订调查提案，提交教师（扮演客户）审核。

2. 调查提案审核通过的小组，继续设计调查方案；未获准通过的小组，进一步搜寻背景资料，重新设计调查提案，直至获得通过。教师认定提案后方能进入调查方案设计阶段。

3. 小组在班级研讨交流完成的调查设计方案，修改并完善方案。

# 第三章数据收集方法

## ，导语

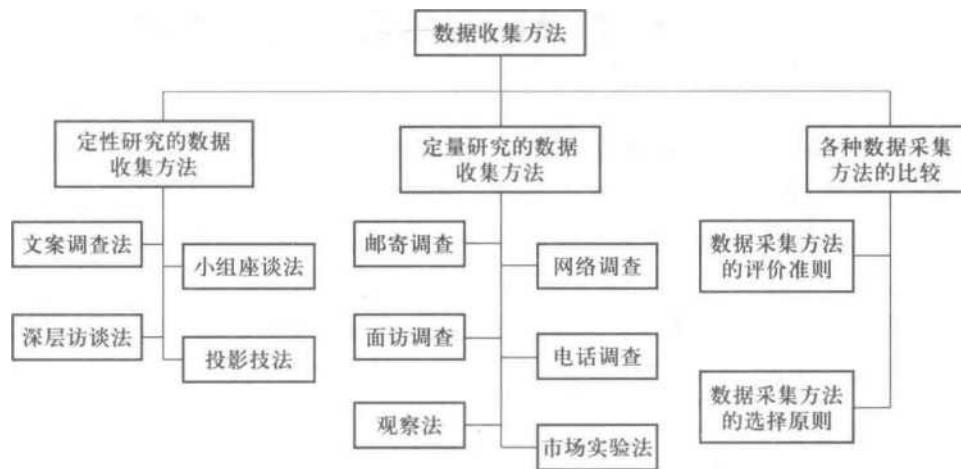


一旦调查项目得以确定，并明确受托方，就需要进一步明确调查目的、了解研究对象特征、进一步细分研究问题，进入相关信息的收集阶段。这一阶段有多种研究方法和调查技术可供选择。本章介绍数据收集的定性调查研究与定量调查研究的主要方法，以及如何合理地选择适当的符合调查对象特点的数据收集方法。

## 、教学目标与要求

1. 了解定性调查研究与定量调查研究方法的特点及其适用对象；
2. 掌握各种定性调查研究方法的特点、作用、实施步骤和适用场景；
3. 掌握各种定量调查研究方法的特点、作用、实施步骤和适用场景；
4. 熟悉调查研究方法的评价准则；
5. 熟悉调查方法的选择原则。

## \本章知识结构



## 第一节

# 定性研究的数据收集方法

定性研究(qualitative research)是指采用非结构化技术，以较小样本量对研究对象做出性质方面的描述与分析的方法。定性研究主要用以识别研究对象的属性特征，方法灵活，形式多样。定性研究可以用于分析消费者的购买行为特征，对产品品牌的态度、感觉及动机等心理层面的问题，寻求新产品的设计创意，也常作为定量研究的前奏或验证、补充定量研究的结果。定性研究的主要数据收集方法有文案调查法、小组座谈法、深层访谈法、投影技法等，下面分别介绍。

### 一、文案调查法

文案调查法是指根据研究目的与任务的需要，通过一定的手段与技术搜集二手资料信息的调查方法。由于所有的研究必须建立在对前期研究成果积累的基础之上，因而，文案调查法的目的是搜集截止到目前，已经形成的研究成果或积累了的数据资料与信息。文案调查法是市场调查前期必选的方法，并贯穿其实施的全过程。

根据市场信息产生的过程不同，资料信息分为原始资料信息和二手资料信息。原始资料信息又称初级信息，它是指在实地调查中从市场活动主体的回答中直接收集到的资料；或是在市场活动中所产生的各种文字、图片、视频和数据资料等。二手资料信息又称次级信息，它是指根据市场活动的需要，对原始信息进行加工、整理和分析后所形成的信息，可以这样理解，第一次形成的市场信息都是原始信息，之后统称为二手资料信息。

对于文献性信息来讲，原始信息为一次文献，对原始信息进行加工、转换，使之有序化和浓缩化，就可形成二次文献，如目录、文摘、索引、统计和会计报表、分类广告等。如果再对二次文献进行综合加工，而得到的有关综述、年鉴、手册、分析报告等，就是通常所说的三次文献。文案调查收集的是二次、三次文献。

#### (一) 文案调查的特点

与实地调查注重现状信息收集相比，文案调查的特点如下：

- (1) 文案调查是收集已经产生并加工过的资料信息，而不是对原始资料的搜集。
- (2) 文案调查以收集文献性信息为主，具体表现为各种文献资料。
- (3) 文案调查所收集的资料包括动态和静态两个方面，尤其偏重于动态资料，即收集服务于研究目的的各种反映相关市场变化的历史与现实资料。

#### (二) 文案调查的作用

文案调查的作用具体表现在以下四个方面：

(1) 文案调查可以发现问题并为市场研究提供重要参考依据。根据市场调查的实践经验，文案调查常被作为市场调查的首选方式。几乎所有的市场调查都始于收集现有资料，只有当现有资料不能为解决问题提供足够的依据时，才进行实地调查。因此，文案调查可以作为一种独立的调查方法加以采用。

在市场研究中，文案调查经常对以下四种情况进行研究：

- ① 市场供求趋势分析 即通过收集各种市场动态资料并加以分析对比，以观察市场发展方向。
- ② 相关和回归分析。即利用一系列相互联系的现有资料进行相关和回归分析，以研究现象之间相互

影响的方向和程度，并在此基础上进行预测。

③ 市场占有率分析。即根据各方面的资料，估算出本企业某种产品的市场销售量占该市 场同种商品总销售量的份额，以了解市场需求及本企业所处的市场地位。

④ 市场覆盖率分析。是用本企业某种商品的投放点与全国该种商品市场销售点总数的 比较，反映企业商品销售的广度和深度。

(2) 文案调查可为实地调查创造条件。文案调查可为实地调查提供经验和大量背景资 料。具体表现 在：

① 通过文案调查，可以初步了解调查对象的性质、范围、内容和重点等，并能提供实地调 查无法或 难以取得的市场环境等宏观资料，便于进一步开展和组织实地调查，取得良好的 效果；

② 文案调查所收集的资料还可用来考证各种调查假设，即可通过对以往类似调查资料的 研究来指导 实地调查的设计，用文案调查资料与实地调查资料进行对比，鉴别和估算实地调查 结果的准确性和可靠性；

③ 利用文案调查资料并结合实地调查，可以推算所需掌握的数据资料；

④ 利用文案调查资料，可以探讨现象发生的各种原因并进行说明。

(3) 文案调查可用于有关部门和企业进行经常性的市场调查。与文案调查相比，实地调 查更费时、费 力，组织起来也比较困难，故不能或不宜经常进行。而文案调查如果经调查员精 心策划，尤其是在建立了 企业及外部文案市场调查体系的情况下，具有较强的机动性和灵活 性，能随时根据企业经营管理的需要， 收集、整理和分析各种市场信息，定期为决策者提供有关 的市场调查报告。

(4) 文案调查不受时空限制。从时间上看，文案调查不仅可以掌握现实资料，还可获得实 地调查所无法取得的历史资料。从空间上看，文案调查既能对企业内部资料进行收集，还可掌 握大量的有关市场环境 方面的资料。尤其是在做国际市场调研时，由于地域遥远、市场条件各 异，采用实地调查，需要更多的时间 和经费，加上语言障碍等原因，将给调查带来许多困难，相 比之下，文案调查更便利。

### (三) 文案调查资料的来源

文案调查需要围绕调查目的，收集一切可以利用的现有资料。从企业经营的角度讲，现有 资料包括企 业内部资料和企业外部资料。因此，文案调查的渠道也主要是这两种来源。

#### 1. 企业内部资料的收集

企业内部资料的收集是收集其经济活动的各种记录，主要包括以下四种：

(1) 业务资料。包括与企业业务经营活动有关的各种资料。如订货单、进货单、发货单、 合同文本、 发票、销售记录、业务员访问报告等。通过对这些资料的了解和分析，可以掌握本企 业所生产和经营商品 的供应情况，分地区、分用户的需求变化情况等。

(2) 统计资料。包括各类统计报表，企业生产、销售、库存等各种数据资料，各类统计分析 资料等。企 业统计资料是研究企业经营活动数量特征及规律的重要定量依据，也是企业进行 预测和决策的基础。

(3) 财务资料。是指由企业财务部门提供的各种财务、会计核算和分析资料，包括生产成 本、销售成 本、各种商品价格及经营利润等。财务资料反映了企业生产要素占用和消耗情况及 所取得的经济效益，通 过对这些资料的研究，可以确定企业的发展前景，考核企业经济效益。

(4) 企业积累的其他资料。如平时剪报、各种调研报告、经验总结、顾客意见和建议、同业 卷宗及有 关照片和录像等。这些资料都对市场研究有着一定的参考作用。

#### 2. 企业外部资料的收集

对于企业外部资料收集，可从以下主要渠道加以收集：

(1) 统计部门与各级各类政府主管部门公布的有关资料。国家统计局和各省及副省级城 市统计部门 都定期发布统计公报等信息，并定期出版各类统计年鉴，内容包括全国人口总数、 国民收入、居民人均可

支配收入、居民人均消费水平等，这些均是很有权威和价值的信息。此外，发改委、财政、工商、税务、银行等各主管部门和职能部门，也都设有各种调查机构，定期或不定期地公布有关政策、法规、价格和市场供求等信息。这些信息都具有综合性强、辐射面广的特点。

(2) 各种经济信息中心、专业信息咨询机构、专业调查机构、各行业协会和联合会提供的市场信息和有关行业情报。这些机构的信息系统资料齐全，信息灵敏度高，为了满足各类用户的需求，它们通常还提供资料的代购、咨询、检索和定制服务，是获取资料的重要来源。

(3) 国内外有关的书籍、报纸、杂志所提供的文献资料，包括各种统计资料、广告资料、市场行情和各种预测资料等。

(4) 有关生产和经营机构提供的商品目录、广告说明书、专利资料及商品价目表等。

(5) 各地电台、电视台提供的有关市场信息。近年来全国各地的电台和电视台为适应市场经济形势发展的需要，都相继开设了市场信息、经济博览等以传播经济、市场信息为主导的栏目、专题节目及各类广告。

(6) 各类国际组织、外国使馆、商会所提供的国际市场信息。

(7) 国内外各种博览会、展销会、交易会、订货会等促销会议以及专业性、学术性经验交流会议所发放的文件和材料。

(8) 互联网与市场信息网络提供的信息。它具有信息量大、速度快、成本低等特点，已成为当前获取市场信息的重要手段。

(9) 其他途径获得的信息。

#### (四) 文案调查的方法

在文案调查中，对于企业内部资料的收集，由于调查费用低，调查的各种障碍少，所以应尽量利用企业的内部资料。对于企业外部资料的收集，可以依不同情况，采取不同的方式：对于具有宣传广告性质的许多资料，如产品目录、使用说明书、图册、会议资料等，是企、事业单位为扩大影响、推销产品、争取客户而免费向社会提供的，可以无偿取得；而对于需要采取经济手段获得的资料，只能通过有偿方式获得，有偿方式取得的资料构成了调查成本，因此，要对其可能产生的各种效益加以考虑；对于公开出版、发行的资料，一般可通过订购、邮购、交换、索取、上网查询等方式直接获得；而对于有一定限制或具有保密性质的资料，则需要通过间接的方式获取。

文案调查方法也就是资料的查寻方法，在此主要介绍文献资料查询的两种方法。

##### 1. 参考文献查找法

它是指利用有关著作、论文的末尾所开列的参考文献目录，或者是文中所提到的某些文献资料，以此为线索追踪、查找有关文献资料的方法。采用这种方法，可以提高资料查找效率。

##### 2. 检索工具查找法

它是指利用已有的检索工具，查找文献资料的方法。依检索工具不同，检索方法主要有手工检索和计算机检索两种。

(1) 手工检索。进行手工检索的前提，是要有检索工具，因收录范围不同、著录形式不同、出版形式不同而有多种多样的检索工具。

以著录方式来分类的主要检索工具有三种：一是目录，它是根据信息资料的题名进行编制的，常见的目录有：产品目录、企业目录、行业目录等；二是索引，它是将信息资料的内容特征和表象特征录出，标明出处，按一定的排检方法组织排列，如按人名、地名、符号等特征进行排列；三是文摘，它是对资料主要内容所做的一种简要介绍，能使人们用较少的时间获得较多的信息。

除了有检索工具外，还要按照一定的检索途径去查寻，才能迅速、准确地找到资料。从表象特征来检

索时，主要有：①资料名途径。它是将有关资料的名称按照一定顺序排列起来的检索系统。②著者途径，著者是指信息资料的个人著者、团体著者、信息提供的单位等，把它们按一定顺序排列起来。在已知著者的情况下，可很快找到所需的信息资料。③资料顺序号途径。每一篇资料都有一个序号，用序号便可组成一个检索系统以方便查找。如果已知要查找的序号即可找到所需的资料。从内容特征来检索时，主要途径有：①分类途径。利用信息资料分类或编码的目录及索引来进行查找。②主题途径。利用信息资料的主题词所组成的检索体系，根据主题词就能查到所需要的资料。手工检索的常用方法有顺查法、倒查法、分段法和追溯法等，在此不再赘述。

(2) 计算机检索。与手工检索相比，计算机检索不仅具有检索速度快、效率高、内容新、范围广、数量大等优点，而且还可打破获取信息资料的地理障碍和时间约束，能向各类用户提供完整、可靠的信息，在计算机广泛普及的今天，文案资料主要依靠计算机来进行检索。

在文案资料中，许多资料是可供长期使用的，对这部分资料就需要加以合理的存储与保管。文案资料存储和管理方式主要有两种：一是档案式的存储和管理方式；二是采用电脑软件进行存储和管理。后者已经十分常用。常见的中文文献管理软件有：NoteExpress、文献之星、医学文献王、PowerRef等。

#### (五) 文案调查的局限性

文案调查法的局限性主要表现在：

首先，文案调查法依据的主要是历史资料，因而过时资料比较多，而且现实中正在发生变化的新情况、新问题也难以得到及时的反映。

其次，所收集、整理的资料和调查目的往往不能很好地吻合，数据对解决问题不能完全适用，收集资料时易有遗漏。

最后，文案调查要求调查员有较广的理论知识、较深的专业知识及技能，因而对调查员的要求较高，而一般的调查员很难达到这样高的要求。

此外，有些资料是由专业水平较高的人员采用科学的方法搜集和加工的，有较高的准确度；而有的资料只是工作人员粗略估算和推测的，准确度较低，对于准确度较低的资料，调查人员较难把握其准确程度。

#### (六) 文案调查资料的评估

由于文案调查的局限性，在运用文案调查资料进行分析研究时，要首先提出问题并对其质量进行评估。应当提出的问题主要有：

资料是由谁收集的？

为何目的收集?

如何收集?

什么内容?

何时收集?

一致性如何?

对文案资料评估,除评估是否满足准确性基本要求之外,还可主要观察文案调查是否满足以下四方面要求:

#### 1. 广泛性

即文案调查对现有资料的收集必须周详,要通过各种信息渠道,利用各种机会,采取各种方式大量收集各方面有价值的资料。一般说来,既要有宏观资料,又要有微观资料;既要有历史资料,又要有现实资料;既要有综合资料,又要有典型资料。如果可能的话,对同种资料应从多种信息源取得,以便相互印证、核实。

#### 2. 针对性

要着重收集与调查主题紧密相关的资料,善于对一般性资料进行摘录、整理和选择,以得到对企业生产经营有参考价值的信息。

#### 3. 时效性

即要考虑所收集资料的时间是否能满足调查的需要。随着信息时代的到来,知识更新速度加快,市场活动的节奏也越来越快,资料适用的时间在缩短,只有反映最新市场活动情况的资料才是价值最高的资料。

#### 4. 连续性

即要注意所收集的资料在时间上是否连续,只有连续性的资料才便于动态比较,便于掌握事物发展变化的特点和规律。

## 二、小组座谈法

小组座谈法,又称焦点访谈法,是指采用小型座谈会的形式,挑选一组具有代表性的消费者或客户,在一个装有单面镜或录音录像设备的房间内(在隔壁的房间里可以观察座谈会的进程),在主持人的组织下,就某个专题进行讨论,从而获得对有关问题的深入了解的调查方法。小组座谈法通常被视为一种最重要的定性研究方法,在国内外得到广泛应用。

### (一) 小组座谈法的特点

小组座谈法的特点在于,它所访问的不是一个一个的被调查者,而是同时访问若干个被调查者,即通过与若干个被调查者的集体座谈来了解市场信息。因此,小组座谈过程是主持人与多个被调查者相互影响、相互作用的过程,要想取得预期效果,不仅要求主持人要做好座谈会的各种准备工作,熟练掌握主持技巧,还要求有驾驭会议的能力:

### (二) 小组座谈法的实施步骤

小组座谈会的实施过程包括三个环节,每个环节有其具体的内容。分述如下:

#### 1. 做好座谈会的会前准备工作

采用座谈会形式,参加人员较多(一般8~12人),会议时间有限(一般1.5~3个小时)。做好会前准备就显得十分必要,需注意以下几个方面:

(1) 确定会议主题,设计详细的座谈提纲。会议的主题应简明、集中,且应是到会者共同关心和了解的问题,这样才能使座谈始终围绕主题进行讨论。提纲通常要在调查者、客户(委托人)与主持人三者之间进行研究。此外,还要注意讨论话题的次序,通常是先提一般的问题,后提特定问

题。当然，在各组座谈会中也要灵活掌握提问方式，从而尽可能得到有用的信息。

(2) 确定会议主持人。主持人对于座谈会的成功与否起关键作用，要求具备丰富的研究经验，掌握与所讨论内容有关的知识，并能左右座谈会的进程和方向。对主持人的素质要求参见表 3-1。

表 3-1 小组座谈法主持人的基本素质要求

素质要求	说明
坚定中的和善	为了促成必要的相互影响，主持人应将训练有素的(不偏不倚)超脱的态度与理解对方并投入感情这两者很好地结合起来
容许	主持人必须容许出现小组的兴奋点或目的不集中的情况，但必须保持警觉性
介入	主持人必须鼓励和促进热情的个人介入
不完全理解	主持人必须通过摆出自己对问题的不完全理解，进而鼓励参加者更具体地阐述其看法
激励	主持人要激励成员积极参与，尤其对不发言者
灵活	在座谈过程出现混乱时，主持人要能随机应变并及时改动计划的座谈提纲
敏感	主持人应有足够的敏感，才能既有感情又有理智地去引导讨论

(3) 选择参加人员。对参加者应做预先筛选，要考虑他们的相似性和对比性。比如参加者都是来自同一目标市场的消费者，其中包括某种产品的使用者和非使用者。最好不要把不同社会层次、不同消费水平、不同生活方式的人放在一组，以免造成沟通障碍，影响讨论气氛。同时，参会人数也要适中，一般在 8~12 人左右。如果参会者过少，难以取得应有的互动效果，参会者过多，发言机会就会减少，观点容易分散；一般来讲，以了解情况为主的座谈会，规模可适当大些；以研究问题为主的座谈会，规模可适当小些。主持人能力强，规模可适当大些；反之，规模可适当缩小。

(4) 选好座谈会的场所和时间。会议场所和时间应对大多数与会者来说是方便和适当的。会场的环境十分重要，应安静，场地布置要营造一种轻松、非正式的气氛，以鼓励大家自由、充分地发表意见。座谈会的时间应比较充裕，时间长度可为 1.5~3 个小时。

(5) 确定座谈会的次数：这主要取决于问题的性质、细分市场的数量、座谈会产生的新想法的数量、时间与经费等。

(6) 准备好座谈会所需的演示和记录用具。如录音、录像设备等。

(7) 在需要同声翻译的情况下（例如有国外客户在观察室监看时），应让翻译熟悉了解讨论的内容，并进行事先的必要练习。

## 2. 组织和控制好座谈会的全过程

(1) 要善于把握座谈会的主题。为避免座谈会讨论离题太远，主持人应善于将与会者的注意力引向讨论的主题，或是围绕主题提出新的问题，使座谈会始终有一个焦点。

(2) 做好与会者之间的协调工作。在座谈会进行过程中，可能会出现冷场、跑题、小组中某个成员控制了谈话的情况，主持人要妥善做好协调、引导工作（如转移话题），以保证座谈会顺利进行。

(3) 做好座谈会记录。座谈会一般由专人负责记录，同时还常常通过录音、录像等方式进行记录。

## 3. 做好座谈会后的各项工作

(1) 及时整理、分析座谈会记录。检查记录是否准确、完整，有没有差错和遗漏。

(2) 回顾和研究座谈会情况。通过反复听录音、录像，回想会议进程是否正常，会上反映情况是否真实可靠，观点是否具有代表性，对讨论结果做出评价、发现疑点和存在的问题。

(3) 做必要的补充调查。对会上反映的一些关键事实和重要数据要进一步查证核实，对于应当出席

而没有出席座谈会的人，或在会上没有能充分发言的人，如有可能也最好进行补充 访问并记录。

(4) 分析和解释结果。讨论结果可以形成供以后进一步检验的假设，在报告书中不要简单重复与会者所说的话，而要着重研究其含义和作用。

### (三) 小组座谈法的应用场景

小组座谈法可以应用于以下场景：

1. 消费者对某类产品的认识、偏好及行为；
2. 对老产品的新想法；
3. 获取对新产品概念的印象；
4. 研究广告创意；
5. 获取消费者对具体市场营销计划的初步反应。

### (四) 小组座谈法的优缺点

#### 1. 小组座谈法优点

(1) 资料收集快、效率高。小组座谈可以同时访问若干个被调查者，与个别面访相比可以 大大节省人力和时间。

(2) 取得的信息比较广泛和深入。由于座谈会由多人参加，在主持人有目的的引导下，大家围绕专题集思广益、互相启发，可以得到更深入、更广泛、有创意的看法和建议。

(3) 方式灵活。有经验的调查主持人可以根据出席座谈会人员的情况，灵活掌握会场，还能将调查与讨论相结合，寻找原因和探讨解决的途径相结合。

(4) 可进行科学监测。小组座谈可以在一个装有单面镜或录音、录像设备的房间内进行，研究人员和客户可亲自观看座谈会现场讨论的情况。

#### 2. 小组座谈法的缺点

(1) 对主持人的要求比较高。主持人不仅要熟悉所讨论主题的业务，而且要头脑清晰、反应灵敏，因而不容易挑选到理想的主持人。

(2) 回答的结果往往比较凌乱，没有统一的格式。对收集到的信息进行汇总、说明和分析 都比较困难。

(3) 有时间和内容方面的局限性。座谈会的时间不能太长，受讨论时间的限制，有时不能 进行深入细致的讨论。有些内容涉及隐私、保密等原因，也不宜在座谈会上讨论。

## 三、深层访谈法

深层访谈法是一种无结构的、直接的、一对一的访问，在访问过程中由掌握高级访谈技巧 的调查员对调查对象进行深入的访谈，用以揭示对某一问题的潜在动机、态度和情感，此方法 最适合于做探测性调查。在市场调查中，常需对某个专题进行全面、深入的了解，同时希望通过访问、交谈发现一些重要情况，要达到此目的，仅靠表面观察和一般的访谈是不可能达到的，这就需要采用深层访谈方法。

### (一) 深层访谈的技术与技巧

#### 1. 深层访谈的技术

常用的深层访谈技术主要有三种：阶梯前进、隐蔽问题探询和象征性分析。

阶梯前进是沿着一定问题的线索进行访谈，使调查员有机会了解被调查者的思想脉络。

隐蔽问题探询是将重点放在个人的“痛点”而不是社会的共同价值观上，以了解与个人密切相关的问

题。

象征性分析是通过反面比较来分析对象的含义。要想知道“是什么”，先要想办法知道“不是什么”。例如在调查某产品时，可先了解某产品的不适用方面以及对立的产品类型。

## 2. 深层访谈的技巧

所谓访谈技巧，是调查员为了获得准确、可靠的调查资料，运用科学的访问方法，引导被调查者提供所需情况的各种方法和策略。访谈技巧在访问调查中占有十分重要的地位，访谈过程大体分为三个阶段，即开始阶段、主要阶段和结束阶段。要取得访谈的成功，必须在访谈过程中各个阶段和各个环节中掌握一定的访谈技巧。

(1) 访谈开始阶段。访谈开始阶段所面临的问题有两个：一是如何做好访谈前的准备工作；二是如何接近访谈对象。在做访谈准备工作时应注意以下三点：

一是准备访谈计划。调查员必须对自己所从事的访谈工作有一定的了解，知道自己在从事什么样的工作，要达到什么访谈目的，准备提哪些问题，重点在哪里等，并预先拟定好访谈提纲，提纲内容一般包括谈话目的、谈话步骤和谈话问题等。

二是预约访谈时间。由于深层访谈一般时间较长，而且访谈的对象常常是身居要职的人员，因此如有可能，最好事先进行电话接触，约定被调查者方便的时间进行访问。

三是准备访谈用品。访谈前，调查员必须准备好能够证明自己身份的证件，如工作证、介绍信等，这对接近访谈对象、取得对方最初信任至关重要。

此外，还要准备访谈必需的物品，如笔、记录本、录音机、摄像机等，如果需要给被调查者一些馈赠物品和宣传资料等，也应准备齐全。

接近被调查者是正式访谈前的序幕，也应掌握一定的技巧”概括地讲，接近被调查者主要有两种方式：

一是正面接近。即开门见山，先介绍自己的身份，直接说明调查的意图，之后就可开始正式访谈。这种方式在访问调查中最常用，它能节省时间，提高效率，但有时显得简单、生硬。

二是侧面接近。即先在某种共同的活动中接近被调查者，如与被调查者一起开会、学习、住宿、娱乐等，但不公开身份，等到与被调查者建立起一定的友谊或有共同语言时，再在一种自然、和谐的气氛中说明来意，进行正式访谈。这种方法有利于消除对方紧张戒备心理，能收集到比较真实、可靠的资料，但调查比较费时、费力。

(2) 访谈的主要阶段。在访谈主要阶段，调查员应该注意以下几点：

调查员要有一个粗略的访谈提纲以防止偏离访谈目标，但访谈的方向需要根据被调查者回答状况而进行适当地调整。

在访谈过程中，对需要引导和追问的问题，调查员要做必要的引导和追询。当被调查者对所提问题不理解或误解时、当被调查者对某一问题的回答有所顾虑时或当被调查者漫无边际地闲谈时，调查员就要有礼貌而巧妙地加以引导。而当被调查者的回答含糊不清时，回答过于笼统，很不准确时，回答残缺不全、不够完整时，则要作适当追问，以使访谈能够顺利进行。

在必要或时间允许的情况下，调查员可从被调查者关心的话题开始，随后逐步缩小访谈范围，最后问及所要提出的问题。

在访谈中，调查员应始终采取公平、中立的立场，使被调查者留下的印象应该是：你对事对人都不带有任何偏见，也不希望左右别人的态度和思想。

调查员应讲文明、有礼貌，用语准确、明了、贴切、恰当。不能以审讯或命令的口吻提问，不能随便打断对方的回答，或对其回答流露出任何鄙视或不耐烦的表情，更不能使用一些对方忌讳、反感的语言。

深层访谈时间比较长，一般都在1个小时左右。在可取得所需资料的前提下，应尽可能缩短时间。

(3) 访谈结束阶段。访谈结束阶段是整个访谈的最后一个环节，这个环节也很重要，不能忽视。首先，在访谈结束时，调查员必须迅速重温一下访谈结果或迅速检查一遍访谈提纲、避免遗漏重要项目，找出回答中出现的各种问题其次，访谈结束时，应再征求一下被调查者的意见，了解他们还有什么想法、要求等，这样可以多掌握一些情况和信息。

## (二) 深层访谈法的应用范围

深层访谈法主要用于获取对问题理解的探索性研究，常适用于以下场景：

- (1) 试图详细地探究被调查者的真实想法，比如购买高档小车的想法；
- (2) 详细地了解一些复杂行为，比如对员工跳槽行为及其原因的调查；
- (3) 讨论一些保密的、敏感的话题，比如个人收入、婚姻状况；
- (4) 访问竞争对手、专业人员或高层领导对事件的应对措施，比如对企业污染物排放的管理手段等；
- (5) 调查的商品比较特殊，通常是那些能够引起人们情感变化的商品，如鲜花等礼品

## (三) 深层访谈法的优缺点

### 1. 深层访谈法的优点

- (1) 能比小组座谈法更深入地了解被调查者的内心想法和态度。
- (2) 能更自由地交换信息，常能取得一些意外资料。
- (3) 便于对一些保密、敏感问题进行调查。
- (4) 能将被调查者的反应与其自身相联系，便于评价所获资料的可信度。

### 2. 深层访谈法的缺点

- (1) 调查的无结构性使这种方法比小组座谈法更受调查员自身素质高低的影响。
- (2) 深层访谈结果的数据常难以解释和分析。
- (3) 由于访问时间长，故所需经费较多，使该法在实际应用中受到一定限制，

## 四、投影技法

投影技法是指采用一种无结构的、非直接的询问方式，激励被调查者将他们所关心话题的潜在动机、态度或情感反映出来的调查方法。前述小组座谈法和深层访谈法都是直接在调查中明显地向被调查者表露调查目的，这些方法在关于动机、原因以及敏感性问题的提问等场合都不太合适，此时，就需要采取那些在很大程度上不依赖研究对象自我意识和情感的新方法。投影技法就是这样的新方法。它有以下四种表现方法，即联想技法、完成技法、结构技法和表现技法。

### (一) 联想技法

联想技法是在被调查者面前设置某一刺激物，然后了解其最初联想事物的一种方法。这类技法中最常用的叫词语联想法，即向被调查者提供一些刺激词，让其说出或写出所联想的东西，调查员通过回答者的不同反应，分析其态度，词语联想法具体分以下三种：

## 1. 自由联想法

自由联想法是不限制联想性质和范围的方法，回答者可充分发挥其想象力。

例如，请您写出（或说出）由下面词所引发的联想

白酒

回答者可能回答：“豪爽”“醉”“浓烈”“营养”“暴力”等。这从不同侧面反映了酒的特点，为改进工艺和市场定位提供有关信息。

## 2. 控制联想法

控制联想法是把联想控制在一定范围内的方法。

例如，请您写出（或说出）由下面词语所联想到的食品。

电视

由电视所联想到的食品，有的是电视广告中出现的食品，有的是看电视时消费的食品，有的兼而有之，有的则什么也不是。对此，研究者在分析结果时可加以区分。

## 3. 引导联想法

引导联想法是在提出刺激词语的同时，也提供相关联想词语的一种方法。

例如，请您就所给的词语按提示写出（或说出）所引发的相关联想。

自行车

联想提示：代步、健身、娱乐、载物、运动、其他。

引导联想法所给出的联想提示带有导向性，如本例的提示，将联想往自行车功能方向引导，回答者的思维也由此向这方面集中。

需要说明的是，在实际调查中运用上述方法时通常会给出一串词语（其中包括一些能隐瞒调查目的的中性词语），让被调查者一一回答。

例如，在下面项目中，看到左边的词，不管联想到什么，请按顺序尽量填写在右边表格中。

① 菜鸟			
② 顺丰			
③ 申通			
④ 圆通			

此法多用于对商品名称、企业名称等做命名调查或名称测试。

采用联想法时，要将被调查者对每个词的回答内容记录下来，再分析它们的含义。对回答的分析可采用以下方式：

- (1) 回答中某词出现的频率。
- (2) 做出回答前用的时间。
- (3) 对某被测词汇在规定的时间内根本没能做出反应的人数。

调查员要仔细记录被调查者对每个被测词汇回答的时间，记录可用秒表计算，也可以由调查员在等待回答时心中默数。被调查者在回答时如果出现犹豫（通常超过三秒），可以认为他确实是在考虑这个词，因而未做出立即回答，这样的回答是可接受的、如果他最终没做出反应，则认为他考虑得太多以致难以回答。每个被调查者的回答方式和回答细节，都可用来分析

其潜在态度或情感。

## (二) 完成技法

完成技法是指设计出不完全的一种刺激情景，要求被调查者来完成的调查方法。常用的方法有句子完成法和故事完成法。

句子完成法是提出一些不完整的句子，让被调查者完成该句子。例如：

(1) 拥有一套住房\_\_\_\_\_ o

(2) 一个家庭必须拥有的交通工具是\_\_\_\_\_ o

(3) 如果我有十万元，我会 \_\_\_\_\_ o

每个人对同一个问题的答案都可能不同，不同的答案表明不同的看法。

句子完成法与词语联想法相比，其优点是具有足够的引导性来使回答者产生一些联想，利用此法进行调查，同样要求被调查者用反应的第一想法回答问题，调查员按原文记录回答并加以分析。

故事完成法是提出一个能引起人们兴趣但未完成的故事，由被调查者来完成，从中看出其态度和情感的方法。例如，某位消费者在一家商场花了很长时间才选中一组价格适宜、造型新颖的家具，在他即将下决心购买时，却遇到售货员的怠慢，这位消费者将做出何种反应？为什么？

## (三) 结构技法

结构技法与完成技法很相近，最常用的是主题统觉法。它是让被调查者看一些内容模糊、意义模棱两可的图画，然后根据图画编一段故事并加以解释。由于这种图画本身没有特定的含义，被调查者往往通过对意义的解释，将其性格和态度反映出来，据此了解其内心活动及潜在需求。

## (四) 表现技法

表现技法是指给被调查者提供一种文字的或形象化的情景，请他们将其他人的态度和情感与该情景联系起来进行调查的方法。具体方法有角色扮演法和第三者技法。

角色扮演法是请被调查者以他人的角色来处理某件事，以间接反映其真实动机和态度。例如，美国某公司在 50 年代就曾用这种方法调查速溶咖啡滞销的原因，他们向被调查者展示两张购物单，让其说出购买速溶咖啡和新鲜咖啡的两个家庭主妇的特点。结果是：被调查者普遍认为，购买速溶咖啡者是懒惰、不会理财、不称职的家庭主妇。这个结果帮助公司了解消费者不愿购买速溶咖啡的真实原因。被调查者在形容购买速溶咖啡的家庭主妇特点时，不知不觉地将自己的看法投射进去了。

第三者技法与角色扮演法有所不同，它是给被调查者提供一种文字的或形象化的情景，让被调查者将第三者的态度与该情景相联系。例如，“这是一套待售的住房。如果您的同事有意购买，那么当他在观看这套住房时，您认为他会如何表现？”研究者可以判断被调查者如何把他自己投影到这个第三者身上，从而揭示出被调查者的真实想法。

## (五) 投影技法的适用场景

投影技法适用于了解以下问题或情形：

1. 某种行为的原因 c
2. 拥有或使用某产品对消费者意味着什么。
3. 当人们不清楚其情感和意见，或不愿意承认对其形象有影响的方面时，或出于礼貌不愿批评他人时。

## (六) 投影技法的优缺点

投影技法的最主要优点是有助于揭示被调查者真实的意见和情感，对那些秘密的、敏感问题的了解尤为有效。

投影技法可能存在的缺点有：需要有训练有素的、专业的调查员，这会增大人员费用；其调查结果可能出现解释偏差；开放式的提问常给分析和研究带来一定困难。

## 第二节

# 定量研究的数据收集方法

定量研究(quantitative research)是指采用结构化问卷，以较大规模样本量对研究对象进行数量特征的统计描述与推断的方法。一般来说，定量调查研究关注研究对象数量特征的测量和描述，需要通过抽样设计、问卷设计、数据分析等手段来收集资料和信息；其调查结果具有客观性和可推断性。定量调查研究采用结构式问卷设计，有着固定的路径和研究程序。定量调查研究广泛应用于市场营销决策，同时还经常是定性调查研究的后续环节。定量调查研究的数据收集方法有邮寄调查、网络调查、面访调查、电话调查、观察法、市场实验法等。下面分别介绍。

### 一、邮寄调查

#### (一) 邮寄调查的特点

邮寄调查是设计好问卷以后，将纸质版问卷邮寄到被调查者手中，由被调查对象自行填写回答的数据收集方法。邮寄调查要求问卷结构严谨，有清楚的说明，让被调查者一看就知道如何完成问卷。与派调查员协助填写问卷相比，邮寄调查问卷应有制作详细、形象友好的说明，这样有助于被调查者完成问卷。在问卷上应有联系电话，如果被调查者需要帮助，可以拨打问卷上列出的联系电话，以便帮助被调查者完成问卷。

#### (二) 邮寄调查的方法

调查实践中，邮寄调查常用的方法有普通邮寄调查、留置问卷调查和固定样本邮寄调查。

##### I. 普通邮寄调查

普通邮寄调查就是将问卷装入信封，通过邮局寄给选定的调查对象，并要求他们按规定的 要求和时间填写问卷，然后寄回调查机构。

邮寄调查的基本步骤如下：

- (1) 根据研究的目的确定调查的总体，收集调查对象的名单、通信地址或电话，抽样确定 调查对象。
- (2) 通过电话、明信片与简短的信件，与调查对象进行事先的接触，说明最近将有一份邮寄的问卷，请求他们协助填写。
- (3) 向调查对象寄出调查邮件。典型的调查邮件一般包括 5 个方面的内容：贴足邮资写清调查对象地址的信封、致调查对象的信、调查问卷、贴足邮资写清调查机构地址的回邮信封、 谢礼或有关谢礼的许诺。
- (4) 通过电话或简短的提示信，与调查对象再次接触，询问是否接收到了问卷，并请求合作早日寄回问卷。
- (5) 对回收的问卷及时登记编码，按回收的日期统计回收的数量。如果回收率还未达到 理想的水平，则再次打电话或寄提示信。
- (6) 如果回收率还达不到研究的要求，则视条件的许可，采取一定的措施修正低回收率所 造成的误差。

## 2. 留置问卷调查

留置问卷调查是介于邮寄调查和面访调查之间的一种方法。它综合了邮寄调查由于匿名而保密性强的优点，又综合了面访调查回收率高的优点。具体做法是，由调查员按面访的方式找到被调查者，说明调查目的和填写要求后，将问卷置于被调查者处，约定几天后再次登门取回填好的问卷。留置问卷调查的关键之一是保证匿名性。

## 3. 固定样本邮寄调查

固定样本邮寄调查是指事先抽取一个地区性或全国性样本，样本中的家庭或个人都已同意参加某方面研究的定期邮寄调查，然后由调查机构向这个固定样本中的成员定期地邮寄调查问卷，样本成员将问卷按要求填写后及时寄回调查机构。

固定样本邮寄调查常用于进行电视收视率、广播收视率、报纸杂志阅读率的调查，机动车燃油消费调查，家计调查等。为防止样本的老化，要定期调整更新样本。

### (三) 邮寄调查的应用范围

目前，市场调查中极少采用邮寄调查的方法来收集数据，原因主要是邮寄调查的局限性。一般来说，在调查的时效性要求不高，调查对象的名单地址都比较清楚，调查经费比较紧缺，而调查的内容又比较多、比较敏感的情况下，采用邮寄调查是比较适合的。其涉及的内容可以是有关日常的消费、日常购物习惯、日常接触媒介习惯等比较具体的方面；也可以是有关消费观念、生活形态、意识、看法、满意度或态度等比较抽象的方面。

### (四) 邮寄调查的优缺点

#### 1. 邮寄调查的优点

(1) 保密性强。由于邮寄调查一般都是匿名的，保密性强，被调查者有安全感，对问题的回答较真实，特别适合于敏感问题的研究。

(2) 调查区域广。从原则上讲，凡是通邮的地方都可以进行邮寄调查，因此该种调查方法能够调查到面访或电话调查难以寻找到的调查对象，如某商品的偏远地区或海外地区的使用者。

(3) 费用较低。一般情况下，邮寄调查的费用比面访和电话调查都低。

(4) 无调查员偏差。面访调查和电话调查的质量与调查员自身的素质有很大的关系，而邮寄调查可以完全避免由于调查员的原因而产生的偏差。

#### 2. 邮寄调查的缺点

(1) 回收率低。在常用调查方法中，邮寄调查的回收率是最低的，因此在邮寄调查中要特别注意采取有效的措施提高回收率，同时对由于回收率低所造成的偏差要进行必要的处理。

(2) 花费时间长。在常用调查方法中，邮寄调查所需的时间是最长的、因此只适用于那些时效性要求不高的项目。

(3) 问卷填答质量难以控制调查对象可能会找他人代为回答，或没有填完全部问题就停止了，这些都将影响数据的质量。

(4) 调查对象受到限制。邮寄调查的最大限制是被调查者必须有较高的文化程度。

## 二、网络调查

随着计算机技术、网络技术的迅猛发展，互联网在社会各个领域的广泛普及，传统的市场调查方式与方法也在发生着变化。作为搜集市场信息的市场调查活动，无论在调查的方式、数据获取的渠道，还是在数据获取的技术与手段方面，呈现出日益多元化、科学化、便捷化的趋势。其中，网络调查已越来越被人们所重视，也越来越多地运用于实际的市场调查活动之中。

### (一) 网络调查的特点

网络调查是指以计算机网络为媒介，以相关专业软件为手段，实现调查问卷或提纲网上制作、发布、填答与回收，或虚拟视频会议空间进行访谈、或直接在虚拟平台实时观察的一种标准化调查方式。网络调查与互联网紧密联系在一起，也被称为互联网调查、在线调查。

网络调查作为调查的一种方式，与传统的面访调查、电话调查、邮寄调查、小组座谈或观察法之间并不存在本质的区别，所不同的是在问卷或提纲制作、发布、填答与回收方面，在虚拟空间利用方面需要依赖计算机网络与有关专业软件。这里所指的专业软件主要包括问卷设计软件、问卷调查软件、数据统计软件、虚拟空间与视频会议软件等。另外，网络调查根据需要，可以实现调查数据的实时汇总、简单描述分析及分析结果的图示化显示。网络调查已成为一种搜集市场信息的主要市场调查方法。

## （二）网络调查的方法

网络调查方法有网上问卷调查、网上讨论法和网上观察法。

### 1. 网上问卷调查

它具体可表现为网页浏览问卷调查、专业网站问卷调查和电子邮件问卷调查三种。

(1) 网页浏览问卷调查。它是指将设计好的问卷放在所选定网站的网页上，由被调查者自愿填答并提交的一种网上调查方法。为了吸引更多的网民参与调查，研究者可以通过海报广告等方式向其所需的研究对象介绍本次调查的目的与主题，并邀请符合调查要求的网民参与调查。感兴趣的网民可以通过打开超级链接或者扫描二维码，在电脑上跳转到相应的网页界面或者在手机设备中下载安装相应的APP后开始填写问卷，并在答题完毕后通过网络将问卷提交到特定的服务器中。这种调查方式能够针对特定所需的研究对象，方便调查对象作答，并具有更好的网络安全特性。另外，这种方式的容错性较高，当发现问卷设计不合理时，可以及时地在服务器中更改问卷，但访问该问卷的超链接不变。

(2) 专业网站问卷调查。它是指借助专业的网上调查网站，从实施问卷设计至调查结果回收的全流程网上调查方法。随着市场调查需求的日益扩大，有不少机构建立了用于网上调查的专业网站，国内的如问卷星、调查派、第一调查网、中国调查网等，国外的如 Socrative、SurveySavvy、Toluna 调查网、TPS 调查网等，为用户提供在线调查服务。用户进入网站需要注册，网站将会提供问卷设计、发布、调查数据汇总与统计分析、调查数据下载等一系列服务。这类服务一般都是有偿的，研究者可根据条件选择合适的专业调查网站实施调查。专业网站问卷调查也是由网民自愿填答问卷，填答者与目标群体往往不一致，将会影响调查数据的质量。

(3) 电子邮件问卷调查。它是指研究者将问卷以电子邮件的方式发放给调查对象，调查对象在接收邮件并填写问卷后提交给研究者的一种网上调查方法。对于这种方式，一方面不论是调查对象还是研究者都需警惕邮件附件中可能存在的网络木马风险。另一方面，基于邮件的调查问卷的容错性较差，当人们发现发布的问卷有错误时，纠正的成本比较大。电子邮件

问卷调查的关键在于要搜集目标群体的电子邮件信箱地址，构建一个比较可靠的抽样框，但这往往难以实现。但是，如果是在局域网中发送电子邮件则比较方便。例如，在调查公司员工满意度时，如果公司拥有局域网且每个员工都分配有个人邮箱，采用电子邮件的方式进行问卷调查就非常方便。

## 2. 网上讨论法

网上讨论法可通过多种途径来实现，如电子公告板(BBS, Bulletin Board System)、新闻组(Newsgroup)、网络实时交谈(IRC, Internet Relay Chatting)、网络会议(net meeting)等。主持人在相应的讨论组中发布调查项目，邀请被调查者参与讨论，发布各自的观点和意见；或是将分散在不同地域的被调查者通过互联网视频会议功能虚拟地组织起来，在主持人的引导下进行讨论。网上讨论法是小组座谈法在互联网上的应用，它的结果需要主持人加以总结和分析，对信息收集和数据处理的模式设计要求很高，难度较大。

## 3. 网上观察法

网上观察法是指利用虚拟平台对网站的访问情况和网民的网上行为进行观察和监测的方法。大量网站都在做这种网上监测，目的是提高网站的点击率。很多可供免费下载的软件，事实上也在做网上行为监测，例如统计下载次数等。

## (三) 网络调查的适用场景

由于网上问卷调查可以设计成自动弹出模式，所以原则上只要有线上业务类型的机构都可以进行相应的网络调查’主要涉及类型如下：

1. 消费者满意度调查；
2. 消费者体验细节调查；
3. 消费者多维评价调查；
4. 线上业务流程合理性调查。

网络调查中的问卷创建详细内容参见本书第四章第四节。

## (四) 网络调查的优缺点

### 1. 网络调查的优点

(1) 成本低、范围广、速度快。由于互联网传播范围的广泛性，利用网络调查可以轻易地实现跨地区甚至跨国界的调查操作，调查成本非常低。相对于传统的邮寄调查，它省去了印刷表格和邮寄等很多环节，使得网络调查的时效性更好。例如，问卷放在网上，被调查者做出的回答能够在瞬间传递到调查员手中.同时也完成了问卷的录入过程.

(2) 交互性好，能够实现问卷多样化设计。有研究表明，问卷中的问题或选项的排列顺序对于调查结果有明显的影响(这个问题将在问卷设计一章中详细讨论)，安排不好会对调查结果产生偏差。例如，在市场调查中对某种汽车的颜色调查，有下面一个问题：

请问您最喜欢下列哪一种颜色的汽车：

- ①红色②绿色③黑色④白色⑤其他

由于红色放在问卷的第一选项，容易当作被调查者的首选，反之，如果把黑色放在问卷的第一选项，其调查结果就可能不同。因此，理想的问卷其选项排列顺序是随机的。而纸介质的问卷不能随意调整结构。网络调查时，设计人员在制作问卷时可以进行一些随机化的设计，包括自动调整问题顺序和随机安排选项顺序。每一名被调查者在下载一份问卷时，由程序产生一份随机化的问卷，这样就解决了上述问题。

- (3) 为各种调查提供了丰富的抽样框 目前，网上站点丰富，涉及许多不同的领域，每一

类站点都有自己特殊的群体，如果对这些群体进行研究，则每一个站点的访问人群，都可以成为一个专业化的抽样框。例如，要进行汽车消费问题的调查，可以将问卷放在汽车类专业的站点上。

(4) 有更好的质量控制。网络调查可以实现自动的逻辑检查，通过程序可以自动将逻辑上有矛盾的选项反馈给被调查者。在问卷中还可以实现动态解释，在传统的调查中，设计者为了避免问卷过于复杂，不能在问卷中设计过多的解释内容，而调查中又往往因为被调查者对问卷误解而出现回答偏差。网络调查可以通过热点提示的方法解决这一难题。

## 2. 网络调查的缺点

(1) 总体代表性不全面。由于网络普及的局限，上网的人多以年轻人和知识分子为主。因此，代表的总体不全面，常常无法推断目标总体，或是一个有偏的估计。所以还只能应用于与网民有关的总体。

(2) 回答率低。目前的网络调查大部分是自愿的，调查员对被调查者并没有什么约束，采用的方法主要是对受访者提供一些礼品(或礼券)或报酬(或打折)以表示感谢。因此，回答率不会很高。

(3) 回答的真实性不能得到保证。网络调查的回答者，仅代表志愿者的总体。大量调查的实践证明，志愿者与非志愿者的态度分布不同，还有一些回答者可能是为了获得礼品和报酬，所以，回答中的偏差无法控制。

(4) 被调查者的身份验证有困难。当有些调查的回答与人群的特征有关系时，如有些化妆品可能对女性有特殊偏好，有些保健品对老年人有特殊偏好，真实的身份就显得重要。由于回答人的身份在网上很难验证，因而影响这类问题的准确分析。

## 三、面访调查

### (一) 面访调查的特点

面访调查是指派调查员与被调查者面对面进行问答的一种数据采集方法。对某些目标总体来说，这是收集数据唯一可行的方法。例如，有些调查无法采用自填的方式；有一些大规模调查的调查对象要求调查员在现场确定被调查对象；很多被调查对象没有电话或者识字率很低。

面访调查可以分为标准式访问和非标准式访问。标准式访问又称结构式访问，它是按照事先设计的、有规定格式的标准化问卷进行的访问。在访谈中调查员按照问卷中的问题逐一提问，得到回答后再将其填入问卷。由于在标准式访问中使用统一的问卷和表格，调查结果便于汇总整理，在面访中通常采用标准式访问。非标准式访问与标准式访问不同，它事先不制订统一的问卷和表格，也没有规定统一的访问程序，而是只给调查员一个题目，由调查员与被调查者就这个题目进行交谈。其特点是具有很大的弹性，访谈中调查员和被调查者可以展开讨论，这就有利于发挥双方的主动性和创造性，获得标准化访问所无法获得的丰富资料。非标准式访问适用于典型调查或个案研究，本章第一节中介绍的小组座谈法和深层访谈法属于非标准式面访。

面访调查中，调查员的个人品质、能力和经验是决定访问调查好坏的关键因素。高质量面访调查的重要前提是有一支良好的调查队伍。这对于保证调查质量，减少调查误差十分重 要。入户面访调查对于调查员的要求，除了熟悉调查内容外，主要是认真负责、诚实可靠、不怕困难和善于交流。

## (二) 面访调查的方法

面访调查根据进行调查的地点不同分为入户面访调查和街头/商城拦截式面访调查，根据所用设备的不同分为普通的面访调查和计算机辅助个人面访调查。

### 1. 入户面访调查

入户面访调查是指调查员按照抽样方案中的要求，到抽中的家庭或单位，按事先规定的方法选取适当的被访者，再依照问卷或调查提纲进行面对面直接的访问。

#### (1) 入户面访调查实施的步骤

进行入户面访调查需要有六个环节的准备和实施：第一环节，确定信息需要，确立调查目标；第二环节，决定调查对象范围（客户或家庭）；第三环节，决定抽样程序，进行抽样；第四环节，设计标准问卷；第五环节，实施预先测试性调查；第六环节，实施和控制正式的入户面访调查。其中，抽样设计和问卷设计至关重要。

#### (2) 入户面访调查中的误差控制

入户面访调查中需要控制的误差包括抽样误差和非抽样误差两大类。

① 抽样误差的控制。入户面访对象由概率抽样方法确定。概率抽样过程中的误差包括随机抽样误差和系统抽样误差。随机抽样误差是由抽样方法本身产生的，即使抽取的样本再合理，当用样本结论推断总体时也可能存在误差，控制这类误差的方法是抽取足够多的面访对象。系统抽样误差是因抽样设计中的错误或问题产生的，包括抽样框误差、调查对象范围误差等。入户面访需要尽可能完善抽样框，准确界定调查对象范围，最大可能避免系统抽样误差。

② 非抽样误差的控制。在入户面访调查中，非抽样误差对调查结果的影响通常比抽样误差更大些。非抽样误差具体包括替代信息误差、替代样本误差、拒访误差、回答误差、调查员误差等。替代信息误差是指由于实际所需的信息与调查员所收集信息之间的差距而产生的误差，这种误差与调查设计的主要问题有关，多数因对一些问题的不恰当定义产生替代样本误差是指由于无法对所抽中的入户对象进行访问而访问其他对象产生的误差。拒访误差是因为需要调查的面访对象拒绝接受调查而产生的误差。回答误差是受访者在回答某一特定问题时有特定倾向而产生的误差，回答误差的产生有两种情况，一是被调查者故意对所提问题不做真实回答；二是回答者希望能够做出真实、准确的回答，但却给了不正确的答案：调查员误差是指因调查员与被调查者之间的相互作用而引起的误差，调查员有时会自觉或不自觉地影响被调查者，使之给出不真实或不准确的回答：上述多种类型的非抽样误差需要在面访调查中有有效控制。

#### (3) 入户面访调查的注意事项

① 访问户的确定。如果抽样方案中已经具体地给出了待访问户的具体地址或名单，调查员必须按方案中的指定地址去访问。但是在许多情况下，抽样方案无法给出具体的待访问户的名单，而只是给出若干个抽样点（如居委会）和如何抽取待访问户的具体规定，这样调查员就有一定的主动权确定访问户。应该注意的是，研究者赋予调查员的抽样主动权应尽量减小，即应尽可能详细地规定抽取访问户的办法。例如，可规定在每个抽样点按等距抽样法抽取5户家庭，还要规定起点的确定方法、抽样间距的计算方法以及行走路线的方向等。抽样方案中还应给出当抽中的家庭户无人或拒访时的处理办法。

② 访问对象的确定：即入户以后确定访问该家庭中的哪一个成员。根据研究目的的不同，访问对象的确定也不同。如果调查的内容涉及整个家庭（住房、耐用品等），一般应访问户主或最具决定权的家庭成员；如果调查的内容主要涉及个人的行为或态度，一般是访问家中所

有某个年龄段的成员（如18岁以上），或是按某种规定选取一位家庭成员进行访问。不管是哪一种情况，抽样方案中都要规定具体的方法。

### 2. 街头/商城拦截式面访调查

街头/商城拦截式面访调查主要有两种方式。

第一种方式是由受到培训的调查员在事先选定的地点（如交通路口、商城），按照一定的程序和要求选取访问对象（如每隔几分钟拦截一位或每隔几个行人拦截一位），征得其同意后，在现场按照问卷进行简短的面访调查。这种方式常用于需要快速完成的小样本的探索性研究，如对某种新上市商品的反应。

第二种方式也叫厅堂调查或中心地调查，是在事先选定的场所按照一定的程序和要求拦截访问对象，征得其同意后，带到该场所附近的事先租借好的房间或厅堂内进行面访调查。这种方式常用于需要进行实物显示或特别要求有现场控制的探索性研究，如广告效果测试等。

### 3. 计算机辅助个人面访调查

计算机辅助个人面访调查（Computer Assisted Personal Interviewing, CAP1），可以是入户的 CAPI，也可以是街头/商城拦截式的 CAPI，该调查法在一些发达国家使用比较广泛。主要有两种方式，一种方式是调查员操作型，即调查员手持普通的笔记本电脑或专用的 APAD，按照屏幕上显示的问题逐题提问，并及时将答案通过鼠标、键盘或专用电脑笔输入计算机。另一种方式是被调查者操作型，即被调查者经过简单的培训或指导后自己操作计算机，通过键盘、鼠标或专用电脑笔，逐题回答电脑屏幕上的问卷，并将答案输入计算机。CAPI 实现了将调查过程、问卷回答审核、数据汇总与数据分析一体化，极大地提高了数据收集和分析的效率和质量。随着计算机及智能手机的普及和广泛使用，计算机辅助个人面访调查正成为市场调查的主流方法。

## （三）面访调查的适用场景

面访调查是目前国内使用最广泛的方法，几乎涉及市场调查的各个应用领域。主要有：

1. 消费者研究。包括消费者的消费行为研究、消费者生活形态研究、消费者满意度研究等。
2. 媒介研究。包括媒介接触行为研究、广告效果研究等。
3. 产品研究。包括产品的使用情况和态度研究、产品的追踪研究、新产品的开发研究等。
4. 市场容量研究：包括对产品当前市场容量和近期市场潜在需求量的估计、竞争品牌市场份额的研究等。

为保障入户访问项目的规范实施，中国市场研究协会于 2016 年制定了相关的行业标准，形成了相应的工作规范，是入户面访的指导文本。详见附录 2。

## （四）面访调查的优缺点

### 1. 面访调查的优点

(1) 入户面访可望取得较有代表性的样本。这是因为入户面访有比较严格的抽样方案，使样本对总体有一定的代表性。另外，利用经过严格培训的丰富经验的调查员，可以大大提高访问的回答率，可望得到质量较高的样本。

(2) 可望获取较多内容、较深问题、较高质量的数据。面访调查的时间一般较长（拦截式调查有时除外），可以采用比较复杂的问卷，可以调查比较深入的问题。面访调查中可以通过调查员现场的解释澄清模糊问题，减少不回答问题或不完全回答的现象。特别是采用 CAP1 进行访问时，在数据收集的灵活性和多样性方面的优势就更为突出了。

(3) 具有激励的效果。由于被访者有机会面对面地向他人发表自己的见解，或与他人讨论某些自己所熟悉的产品或关心的问题，因此有可能达到情绪上的满足或知识上的满足。调查员可以利用这种激励作用，争取被访者的合作，以提高回答率和回答质量。

## 2. 面访调查的缺点

(1) 费用较高。在常用调查方法中，面访调查的费用是最高的，特别是当要求获取全国性或某些区域性的随机样本时。

(2) 时间较长。由于面访调查需要调查员亲自去找到被调查者，大量的时间会花费在路途的奔波和寻找之中，因此调查速度较慢，持续时间较长。

(3) 某些群体的访问成功率低。例如，有些高收入或高地位的特殊阶层，面访调查很难接触；有些居民区由于安全等方面的原因，不允许调查员入户。近年来，大城市入户调查的拒访率不断上升。

(4) 实施质量的控制较困难。在面访调查中，由于调查员分散在许多地点，对调查员的管理监督难度较大，对调查质量的控制也只能通过事先的培训和事后的抽查来实现，这需要较多的人力和经费投入。

## 四、电话调查

### (一) 电话调查的特点

电话调查是调查员利用电话这一通信工具与被调查者进行语言交流，从而获取信息，采集数据的一种调查方法。

电话调查的特点是以电话作为调查工具。由于电话调查中，双方不见面，只能听到对方的声音，这对调查员提出了更高的要求。因此，电话调查员应具备这样的能力，即从对方的声音中能够迅速、准确地判断出对方对于调查的基本态度和感兴趣的程度，要能够熟练地运用访谈技巧完成调查任务。当前电话普及率高，电话调查的前景非常广阔。

### (二) 电话调查的主要方式

电话调查一般包括两种方式，传统的电话调查和计算机辅助电话调查。

#### 1. 传统的电话调查

传统的电话调查使用的工具是普通的电话、普通的印刷问卷和普通的书写用笔。经过培训的调查员在电话室内按照调查设计所规定的随机拨号的方法，确定拨打的电话号码，拨通后按照准备好的问卷和培训的要求，筛选被访对象，然后对合格的调查对象对照问卷逐题逐字地提问，并迅速及时地记录下答案。

随机拨号方法是根据随机抽样的原理设计，有两种常用做法。

(1) 利用现成的电话号码簿作为抽样框，借助随机数字表，随机地选取要拨打的电话号码；或采用等距抽样的方法从电话号码簿中抽取拨打号码。

(2) 按照所调查地区的具体情况和抽样方案先确定拨打号码的前几位，再按照随机的原则确定后几位。

传统的电话调查对于小样本的简单访谈虽然简便易行，但也存在不少问题，如效率低、难以进行统一的监控和管理、难以处理复杂的问卷等。

#### 2. 计算机辅助电话调查

计算机辅助电话调查通常是在一个装备有计算机辅助电话调查设备的中心地点或实验室进行。整套CATI(Computer Assisted Telephone Interview)系统的硬件包括：一台起总控制作用

的计算机主机、若干台与主机相连接的显示器终端、耳机式或耳塞式电话和鼠标、若干台起监视作用的计算机和配套的音响设备。软件包括：问卷设计系统、自动随机拨号系统、自动数据录入系统、简单统计系统和自动访问管理系统。

经过培训的调查员戴上耳机或耳塞，坐在显示器屏幕前，按照屏幕上显示的问答题进行访问，并通过计算机的键盘或鼠标将数据及时录入计算机。与传统的电话调查相比，CATI 收集的数据质量得到了提高，访问的时间大大缩短，也省去了数据编码和录入等繁琐的过程。

督导员可在现场检查和指导调查员的工作，也可以在其他的房间通过监控设备随时了解访问的实况。计算机的主机可以随时提供整个调查的进展、阶段性的调查结果以及每一个调查员完成工作的具体情况。研究人员可以根据阶段性的调查结果及时地调整方案，使调查更有效。对于被调查者不在家需要追访或被调查者没有空需要另约时间的情况，CATI 系统也会自动地储存下次访问的号码和时间，届时该号码会自动地出现在拨号系统中。

### （三）电话调查的适用场景

目前，我国电话调查一般应用于热点问题或突发性问题的快速调查、关于某特殊问题的消费者调查、企业家信心调查和特殊群体的调查等。

### （四）电话调查的优缺点

#### 1. 电话调查的优点

(1) 时效快。要在短时间内完成调查任务，电话调查具有突出优势，尤其是被调查对象分布十分分散，居住在交通不便、边远地区时，电话调查可以大大缩短调查时间。因此，电话调查广泛运用于时间要求很高的调查。如市场应急需求调查、民意调查等。

(2) 费用低。电话调查能以比较合理的费用提供适中的回答率。在常用调查中，面访调查的回答率比较高而费用也比较高，邮寄调查的费用低，但是回答率也低。电话调查的费用和回答率一般处于这两者之间。

(3) 可能访问到不容易接触到的对象。例如，有些高收入或高地位的特殊阶层，面访调查比较难接触到，但是利用电话调查可能访问成功。有些被调查者拒绝陌生人入户访问，有些人工作忙，拒绝面访，但有可能接受短暂的电话调查。

(4) 可能在某些问题上得到更为坦诚的回答。例如，有些关于个人方面的问题，或者是对某些特殊商品的看法，电话调查得到的回答可能要比面访调查时得到的回答真实得多。

(5) 易于控制实施的质量。电话调查员基本上是在同一个地点进行电话调查，督导员或调查组织者可以在调查的现场进行监督，以便纠正调查员不正确的操作。例如，没有严格按问答题提问、说话太快、吐字不清晰、声调不亲切或语气太生硬等可能出现的问题，可以及时予以纠正。

(6) 计算机辅助电话调查更加便利与高效。电话调查一般采用计算机做辅助工具。由于有现成的软件，可以把调查过程和数据录入过程结合在一起，即在实施调查的同时，将调查的结果输入计算机：这样就能随时掌握调查进度和及时得到汇总信息，十分方便。

#### 2. 电话调查的缺点

(1) 调查难以包括目标总体的全部单位。抽样总体实际是全体电话用户，而目标总体包括所有安装有电话和没安装电话的消费者，如果电话普及率未全部覆盖，特别是在边远地区和农村，样本的代表性就会存在问题。所以电话调查只能局限于那些拥有电话的群体。如果有电话那些单元其调查项目的数量特征与总体其他单元的数量特征有显著的差异，电话调查的结果就不能用来说明目标总体。也就是说这种估计是有偏的。

(2) 电话调查的采访时间受到限制，调查内容难以深入。任何一项调查，调查过程都不能太长，以免

占用被调查者过多的时间，引起反感。这个问题在电话调查中显得更为突出。因为 电话调查还占用了对方的电话线路，影响到被调查者及其旁人使用电话。由于时间限制，因而 调查内容也不能太复杂。

(3) 电话调查的回答率比面访调查要低。在面访调查中调查员可以利用与被调查者直接 进行面对面交流的有利条件，提高回答率。而电话调查则不具备这个优势，其回答率受到一定 影响。有些电话调查是采用随机拨号的方式，则常常会遇到空号或错号、占线或被调查对象不 在等，这些都会降低回答率。

(4) 电话调查容易受到语音障碍的影响。邮寄调查和面访调查都是通过文字的问卷进行 调查，因此不存在语音的问题，即使调查员有口音，但比较容易解决。电话调查则不同，如果其 中一方听不懂对方的语言，就很容易产生误差。

## 五、观察法

观察法是由调查员在现场对被调查者的情况直接观察、记录以取得有关市场信息的一种 调查方法。它不是直接向被调查者提出问题要求回答，而是凭借调查员的直观感觉或是利用 有关器械，如照相机、录像机或其他器材，记录和考察被调查者的活动和现场事实，以获得必要 的信息。利用卫星图像的遥感技术也是一种直接观察，一般用于调查农作物的种类和种植面 积的估计。

### (一) 观察法的分类

观察法可从不同角度进行分类。

#### 1. 按照观察结果的标准化程度分为控制观察和无控制观察

控制观察是根据观察目的，预先规定观察范围，在实施观察时，对观察手段、观察技术、观 察程序和记录方式标准化。控制观察一般用于目的性、系统性较强的调查或用于简单观察后， 为使调查更加精确而进行的补充调查或取证。无控制观察比较灵活，对观察项目、程序等不做 严格规定，记录也可采取随意的方式，无控制观察通常用于探索性或有一定深度的专题调查。

#### 2. 按照观察者参与观察活动的程度不同分为完全参与观察、不完全参与观察和非参与 观察

完全参与观察是指观察者较长时期地生活在被观察者之中，甚至改变自己原有的身份。例如，以售货员身份在商场从事销售工作的企业信息员，观察顾客购买本企业产品的情况及同 类竞争性产品的销售情况。不完全参与观察是指观察者不改变身份，只在调查时置身于调查 事项中取得资料。例如，“神秘购物法”就是其中一例。依靠那些经过专门训练的神秘顾客（其本身是普通消费者），由他们进行伪装购物，详细记录下购物或接受服务时发生的一切情 况，以发现商家在经营管理中存在的各种缺陷。非参与观察是指观察者不参与调查活动之中， 而是以局外人的身份去观察事项发生和发展情况。前两种观察法因观察者置身于观察活动之 中，便于了解事物发展状况及产生的原因，取得较为深入的资料;但有时也会不客观，对观察结果带有偏见；非参与观察比较客观、公正，但观察可能停留在表面，无法了解事物发展变化的根本原因。

#### 3. 按照所取得资料的时间特征不同分为纵向观察、横向观察和纵横结合观察

纵向观察又称时间序列观察，是指在不同时期加以观察，取得连续的观察记录。对纵向观 察资料进行分析研究，能了解调查对象发展变化的过程和规律。横向观察又称横断面观察，是

指在某一特定时间内对若干个调查对象所发生的事态同时加以记录。对横向观察资料进行分析研究，能够扩大调查的范围，观察调查对象之间行为上的差异程度。纵横结合观察就是对调查对象的纵横两个侧面进行观察，这样，就能取得更加丰富、相互印证的可靠性高的调查资料，如要开展电视广告对消费者影响力的调查，就可采用这种方法进行□

#### 4. 按照观察的具体形式不同分为人员观察、机器观察和实际痕迹观察

人员观察是由调查员实地观察受访对象以了解情况的一种常用形式。例如，某调查公司为了了解某种品牌的微波炉市场营销情况，可采取以下三种观察方法：

(1) 销售现场观察。调查员到商场、经销店、门市部、各种展销会、交易会等现场，亲自观察和记录顾客的购买情况、购买情绪、踊跃程度、同类产品竞争程度、新产品的设计以及商品的性能式样、价格、包装等。

(2) 使用现场观察。在用户的使用现场，了解用户使用本企业生产或经销商品的情况，观察对象包括产品的用途和使用条件、顾客在使用时遇到的困难、顾客要求等，以发现问题并及时加以改进。

(3) 供应厂家现场观察：派专人到原材料、配套件、协作件生产企业，观察他们的生产条件、技术水平、工艺过程，以了解其所生产和经销的商品能否满足要求，或到商店，观察货源的准备情况与库存状况C

机器观察就是通过机器观察受访对象。这些设备可能需要也可能不需要被调查者直接参与，它们可以连续地记录发生的行为例如，美国一家广告公司，为了进行电视广告收视率调查，经用户同意，在1~200个家庭的电视机上安装了与公司总部相连的电子记录器，当观众收看时，就能把所看电视的频道、节目记录下来，然后再对这些资料加以汇总、分析，这样就可以了解观众收看什么样的电视台和电视节目，确定广告播出的黄金时间。

实际痕迹观察法是指调查员不直接观察受访对象的行为，而是通过一定的途径来了解他们行为的痕迹〔例如：通过对手机售后保修点的观察，了解不同部件的损坏率。损坏率高的部件，说明其设计或质量本身存在某种问题，与广告宣传中所介绍的产品功能也可能存有偏差。〕

### (二) 观察法的记录技术

采用观察法时，记录技术的好坏直接影响调查结果。良好的记录技术，可以减轻观察者的负担，不致因忙于记录而顾此失彼。记录技术主要包括观察卡片、符号、速记、记忆和机械记录五种。

#### 1. 观察卡片

观察卡片或观察表的结构与调查问卷的结构基本相同C在制作观察卡片时，首先根据观察内容，列出所有观察项目；去掉那些非重点的、无关紧要的项目，保留重要的能说明问题的项目；列出每个项目中可能出现的各种情况，合理编排；通过小规模的观察来检验卡片的针对性、合理性和有效性，以修改卡片；最后，定稿付印，制成观察用的卡片。

#### 2. 符号

这是指用符号代表在观察中出现的各种情况，在记录时，只需根据所出现的情况记下相应的符号，或在事先写好的符号上打钩即可，不需要再用文字叙述，这样不仅加快了速度，避免因忙乱而出错，而且便于资料的整理。

#### 3. 速记

这是用一套简便易写的线段、圈点等符号系统来代表文字，进行记录的方法。

#### 4. 记忆

这是在观察调查中，采取事后追忆的方式进行的记录。通常用于调查时间紧迫或不宜现场记录的情况。由于人的大脑不可能准确无误地储存很多信息，因此，必须抓住要点记忆，提纲挈领，事后及时进行整理。事后记忆虽然可以回避被调查者，但常容易遗忘一些重要的信息。

## 5. 机械记录

这是指在观察调查中运用录音、录像、照相、各种专用仪器等手段进行记录。这种记录方法能详尽记录所要观察的事实，免去观察者的负担，但易引起被调查者的顾虑，使调查结果失去真实性。例如，在广告效果调查中，以下四种仪器可以记录消费者对广告的反应：

(1) 眼动仪，又称眼睛照相机。它可以将顾客注视物品时的眼动轨迹记录下来，通过分析记录的数据，如注视顺序、时间、次数、眼跳、距离等，分析观看者的心理活动。可用于探测被调查者对广告的反应。

(2) 瞬间显露器。它可在短暂的时间内显示广告，用以了解广告的各个构成要素所需时间。

(3) 精神电流测定器。它可通过测量脉搏、血压、呼吸等间接测出情感变化和心理反应。

(4) 皮肤电流反射器。利用这种仪器可以根据受测者的精神变化而在皮肤上出现的反应，进行相应的判断。

## (三) 观察法应用范围与注意事项

### 1. 观察法应用范围

观察法在我国市场调查实践中应用较广，主要有以下几个方面：

(1) 在城市集贸市场调查中，对集贸市场上农副产品的上市量、成交量和成交价格等情况进行观察。

(2) 在商品库存调查中，对库存商品直接盘点计数，并观察库存商品冷背残次情况。

(3) 在消费者需求调查中，对消费者购物时对商品品种、规格、牌号、花色、包装、价格等要求进行观察。

(4) 在商场经营环境调查中，对商品陈列、橱窗布置、所临街道的车流、客流量情况进行观察。此外，观察法还可用于产品质量调查、广告调查等领域。

### 2. 观察法的注意事项

(1) 为了使观察结果具有代表性，反映某类事物的一般情况，应设计好抽样方案，以使观察的对象和时段具有较好的代表性。

(2) 在进行实际观察时，最好不让被调查者有所察觉，否则，无法了解被调查者的自然反应、行为和感受。

(3) 在实际观察时，必须实事求是、客观公正，不得带有主观偏见，更不能歪曲事实真相。因此，要对调查员进行有效的培训，提高调查员的业务素质。要求调查员遵守有关法律和道德准则，不能对涉及国家机密和个人隐私的内容进行观察，除非得到了允许。

(4) 调查员的记录用纸和观察项目最好有一定的格式，便于尽可能详细地记录调查内容的有关事项。

(5) 为了观察客观事物的发展变化过程以进行动态对比研究就需要做长期反复的观察。

## (四) 观察法的优缺点

### 1. 观察法的优点

(1) 直观性和可靠性。观察法可以比较客观地搜集第一手资料，直接记录调查的事实和

被调查者在现场的行为，调查结果更接近于实际。这是文案调查等间接调查方法所不可比拟的，也是面访法所不及的。

(2) 少受语言交流与人际交往的影响。特别是非参与观察，它一般不依赖语言交流，不与被调查者进行人际交往，因此，它有利于对无法、无须或难以进行语言交流的市场现象进行调查，有利于排除语言交流或人际交往中可能发生的种种误会和干扰。

(3) 简便、易行，灵活性强，可随时随地进行调查。

## 2. 观察法的缺点

(1) 费用比较高。因为观察法对调查员的要求比较高，所有的调查员都必须经过培训，才能进行观察，有时还需要专业人员参与。例如，请护士测量血压，会录像的人录像等。如果专业人员不够，样本量就受到限制；至于卫星图像的遥感技术，费用就更高了。

(2) 观察对象有局限性这种数据收集方法不适用于所有类型的数据，因为并非所有客观现象都能通过直接观察就能得到。

(3) 观察结果容易受到观察者的影响。观察者往往只能接触到被观察的表面现象，无法接触到隐蔽在背后的原因。特别是如果被调查者想掩盖这些背后的原因，当观察人员在场时，被调查者就会做出另一番表现，使观察人员得到的资料不真实。即便是公开现象，由于观察人员的参与，往往也会在某种程度上影响被观察者的正常活动。此外，观察人员在进行观察时，有时难免带入调查员的主观看法和个人的感情色彩，使得观察结果出现偏差。

(4) 有些现象观察的结果缺乏统计意义。与其他调查方法相比，观察所得到的结果可能缺乏统计意义。观察到的往往是质的方面，观察的结果不宜被量化处理，许多现象不宜被重复验证，因此，很难判断观察结果是否有代表性，也就难以推断总体。

## 六、市场实验法

### (一) 市场实验法的基本概念

实验法广泛应用于自然科学中的因果关系研究中。它是指在受控环境下，使其他因素保持不变，研究所控制的变量（自变量）对某一变量或某些变量（因变量）的影响的方法。

将自然科学研究中的实验法应用于市场研究，称之为市场实验法。例如，实验者控制一个或多个自变量（如价格、包装、广告），研究在其他因素（如质量、服务、销售环境，等等）都不变或相同的情况下，这些自变量对因变量（如销售量）的响应效果。在市场实验中，未能控制而又可能在实验期间有所变动的外来因素太多，例如竞争对手的策略、消费者的偏好等等，这些外来因素都可能对实验的结果有所影响。为此，在进行实验设计时，要特别考虑如何尽可能地降低实验的误差。

市场实验法采用归纳法的逻辑，首先，通过科学设计的实验收集数据，然后进行统计分析和假设检验，以达到利用实验样本对总体的推断。不同于本章前面所述的探索性或描述性的研究方法，它是了解和研究因果关系的主要方法之一。市场实验法常用的技术语如下：

#### 1. 因果关系

它是指某个或某些“起因”变量X影响或引起了某个或某些“效应”变量Y而“X引起Y”包含有以下三方面的意思：一是变量X只是变量Y的可能起因中的一个；二是变量X的出现使变量Y的出现更为可能（X是Y的概率性起因）；三是我们最多只能推断变量X是变量Y的一个起因。变量X和变量Y构成因果关系的三个必要条件是：一是变量X与Y有一种共变的关系；二是变量X在Y之前发生（出现），或变量X与Y同时发生（出现）；三是排除了其他可能的起因。

#### 2. 自变量

它也称为独立变量、外生变量、解释变量。在实验设计中，称为因子，也称处理变量，是指在实验过程中实验者所能控制、处置或操纵的(即实验者可以规定或改变这些变量的水平或 取值)而且其效果可以测量和比较的变量:例如，销售行为研究中产品的价格水平、包装设计、广告主题、促销方法等都属于这类变量。每个因子在实验中需要考虑的不同测量值或种类，称为水平。

### 3. 实验单位

它是指实验的主体，可以是个人、组织或其他实体，它们对自变量的反应(因变量)可以测量。例如消费者、商店、销售区域、分销商等。

### 4. 因变量

它也称为响应变量、内生变量、被解释变量 c 它是测量自变量对实验单位的影响效果的变量。例如，市场研究中消费者的购买量、满意度，商店的销售量、利润、市场占有率等属于这类变量。因变量的取值也叫观察值或实验结果。

### 5. 外来变量

它也叫作无关变量，是指除自变量以外一切能影响因变量取值的其他所有变量。这些外来变量可以影响因变量的测量值，使实验的结果变弱或无效外来变量主要有两类:一类是由实验单位之间的差别造成的影响，如消费者的收入或文化程度方面的差别•商店的位置或规模 的差别等等。这些影响是可以通过设计加以控制的。另一类是不能控制的外来因素，如气候、 竞争对手的策略或行动、消费者的偏好等。通过随机抽样决定实验单位的方法，有可能降低这些外来变量对实验结果的影响

### 6. 实验

它是指研究人员在控制外来变量影响的同时，处理操纵一个或多个自变量，并测量它们对一个或多个因变量的效应。同一实验应该可以在尽可能相同的条件和环境下进行重复

### 7. 实验设计

它是指进行实验时具体规定的一系列方法，包括:

- (1) 规定实验单位，以及如何将这些单位划分为同类或同质的子样本;
- (2) 需要控制或处置的自变量;
- (3) 需要测量的因变量;
- (4) 控制外来变量的方法

### 8. 实验误差

市场实验的因变量并非只受到自变量的影响，还会受到外来变量或测量误差的影响’ 通过实验设计，可以控制或消除部分外来因素的影响，但是还会有一部分未能识别的外来因素的 影响不能消除，此外，还有一些测量上的无法用统计方法消除的随机误差，也会对因变量的变动产生影响。由这些外来因素和随机误差所导致的影响统称为实验误差。

### 9. 内部有效性

它是指通过实验测量的自变量对因变量影响的准确性项控制外来变量是建立内部有效性 的一个必要条件。

### 10. 外部有效性

它是指将实验的结果推广到实验环境以外或更大总体的可能性。如果具体的实验环境没

有现实地考虑到真实世界中其他相关变量的交互影响，外部有效性就会受到很大的影响。

## （二）市场实验法的分类

### 1. 根据实验场所分为实验室实验和现场实验

实验室实验是在人造的环境中进行实验，研究人员可以进行比较严格的高水平的实验控制，比较容易操作，所需的时间比较短，所需的费用也比较低。实验室实验的内部有效性一般比较高，即观察和测量值比较准确，可以比较准确地判断变量间的因果关系。但是由于实验是在实验室的环境中进行的，因此一般情况下外部有效性比较低。

现场实验是在实际的环境中进行实验，研究人员不太容易对实验进行严格的高水平控制，比较难操作，所需的时间比较长，所需的费用比较高。现场实验的内部有效性虽然可能比较低，但是其外部有效性一般都比较高，即其实验结果具有较高的实用意义。

### 2. 根据实验单位随机分组方式不同可分为非随机化实验和随机化实验

非随机化实验，也叫非正式的实验或伪实验，按照是否将实验单位分成处理组和控制组（也叫对照组）以及按照是否只进行事后测量还是同时进行事前事后测量，可以分成四种：事后设计、有控制组的事后设计、事前事后设计、有控制组的事前事后设计。

随机化实验，也叫正式的实验或真正的实验，根据自变量（处理变量或因子）的多少，外来变量的多少，以及是否考虑因子之间的交互作用，可以分为完全随机化设计、随机区组设计、拉丁方设计、正交实验设计等。

## （三）市场实验法的质量控制

实验的结论是否有效，是研究者最为关注的问题，实验质量控制的关键是控制外来变量，以提高实验的内部有效性和外部有效性。

### 1. 控制不同类型外来变量的影响

能否有效控制外来变量不同类型的影响，决定了市场实验设计的成效。除自变量以外其他一切能影响因变量取值的所有变量根据其影响的方式可归纳为选样偏差、时间效应、测试效应、回归效应和失效应。

（1）选样偏差。它是指由于实验单位的选取和分组不当形成的外来变量造成的偏差。研究者如果按照自己的主观判断进行分组，使处理组和对照组内的实验单位在进行实验之前就有明显的差异，选样偏差就形成了。例如，在研究商品的陈列对销量的影响时，处理组和对照组内的商店可能一开始就不等价的。某一组的商店规模可能大一些，而商店规模与销售量是有直接关系的，不管两组商店的商品陈列（自变量）有什么不同，商店规模（外来变量）对销售量（因变量）的可能影响是客观存在的。因此，应当尽量避免这种选样偏差。

（2）时间效应。它是指在实验进行的时间段内发生了其他重大事件或发生了实验单位环境变化而体现的外来变量的影响。以这些事件或环境变化可能会对因变量产生影响。例如，在研究促销手段对空调销售量影响的为期一个月的实验期间，发生连续高温的天气异常事件，这一事件本身会带来空调销售量的增加，这种增加就会混杂在促销的效果中，即销售量的增加除了促销手段外，还有其他因素的影响，实际上会高估促销手段的影响。又例如，如果实验单位是消费者，那么实验期间他们的参与兴趣可能会发生比较大的变化；如果实验单位是商店，那么实验期间商店的摆设或装修也可能有变化，这些变化在实验之前难以预料，因此都可能对实验的结果产生影响。

（3）测试效应。它是指由于实验过程中测试的顺序不同而形成的外来变量的影响。测试效应包括主效应和交互效应两种类型。

主效应是指前期观察对后期观察可能产生的影响。例如，在研究某品牌广告与消费者对该品牌态度的因果关系实验中，先测试消费者对该品牌的态度，然后让消费者接触该品牌的广告，再测试他们对该品牌的

态度。消费者态度的变化除了受广告影响外，还可能更多地受到事先测试的影响；消费者可能会有意识地维持其前后态度的一致性，或正好相反。在前一种情况下，主效应弱化了广告的实际效应，而在后一种情况下，主效应强化了广告的实际效应。在这两种情况下，测试的主效应会影响实验的内部有效性。

交互效应是指前期观察改变实验单位对自变量反应（或回答）的影响效应「例如在上述的广告效果测试中，在询问对某品牌的态度时，参与实验的消费者知道了该品牌，因此，与那些没有参与实验的消费者相比，他们会对该品牌更为敏感，更有可能会对测试的广告表现出关注。那么，所测出的广告效应就无法推广到总体 C 所以，测试的交互效应会影响实验的外部有效性。

(4) 回归效应 它是指第一次测试出现极端值的实验单位，在实验多次测试过程中移至平均值附近所形成的来自外部变量的影响。例如在广告效果的实验中，有些被调查者（实验单位）对某品牌的态度原来是十分极端的（非常喜欢或非常不喜欢），但是在接触广告后再次测量时，他们的态度可能变得比较接近平均值。如果这种情况发生，实验效果的判断就会受到干扰。因为观察到的效应（态度的变化）可能更多是由于统计的回归产生的（极端值对方差的影响更大），而不是真正的实验处置的效应（广告的效应）。统计回归效应的真正原因是偶然因素变化导致的随机误差，以及仅仅根据一次测试结果进行分析的做法。

(5) 丢失效应。它是指由于实验过程中实验单位的“丢失”（退出、拒访、遗失）而形成的外部变量的影响。例如在研究商店的陈列方式对销售量影响的实验中，某处理组（采用某种新的陈列方式）的几个商店可能中途退出实验，在这种情况下，研究者很难判断如果这几个商店还继续参与实验的话，该组的平均销售量会高一些还是低一些

## 2. 控制外部变量的影响

外部变量的存在会影响因变量，因此对市场实验的内部有效性和外部有效性都可能产生严重的威胁，要尽可能地加以控制 常用的控制方法有以下几种。

(1) 随机化分组。随机化分组指的是借助随机数字表或抽签等手段，将实验单位随机地分配到各处理组和对照组，同时，处理条件也是随机地分配到各个组 例如，在研究促销效果的实验中，将实验的商店随机地分成三组，然后再将三个处理条件（促销方式 1、促销方式 2、无促销）随机地分配到这三个组，形成处理组 1、处理组 2 和对照组。随机化分组的方法有利于保证参与实验各组从一开始就是基本平等的。不过，如果实验的样本量比较小、随机化分组的效果可能就会比较差。一般可以通过测量和比较各个实验组外部变量的值，来确定随机化分组的有效性。

(2) 匹配分组。匹配分组指的是按照一些关键的背景变量来比较实验单位，使每个实验组都有相匹配的实验单位。例如，在广告效果测试实验中，可以按性别、年龄、文化程度和家庭收入等变量将参加实验的消费者匹配分组；在商品陈列效果测试的实验中，可以按年销售量、商店规模和位置等变量匹配分组。匹配分组的局限性在于，首先，实验单位只能按照少量的几个变量进行匹配分组，因各组中的实验单位只在所选的变量上是相似的，而在其他方面特征则可能很不相同；其次，匹配变量的选择也是比较困难的，因为如果要使匹配分组形成好的效果，就需要保证所选的匹配变量与因变量之间有较强的关联性，而这在实际操作中很难选择。

(3) 统计控制。统计控制指的是采用统计分析的方法测量外来变量并修正其影响。例如，在研究店内广告对促成消费者冲动购买的效果时，可以对诸如性别、年龄或收入等外来变量加以控制，从而近似地排除这些无关变量的干扰。用于控制外来变量的统计方法很多，例如多维列联表、对数线性模型，多元回归分析、协方差分析等。

(4) 实验设计。实验设计是指通过精心设计的实验方案，使外来变量的影响得到有效的控制。常用的实验设计的例子见下一节。

## (四) 市场实验设计的常用方法

市场实验设计的方法十分丰富，一般分为随机化实验设计的方法和非随机化实验设计的方法，由于非随机化实验设计方法不是标准的实验设计方法，因而这里只给出随机化实验设计常用方法的设计思路。

### 1. 完全随机化设计

完全随机化设计是统计实验中最简单的一种，它清楚地描述了控制和随机化这两项原则。其基本做法是：先把受试对象随机分组，组数和处理数相同，然后对每一组实施一种处理。在完全随机化的实验设计中，所有的实验个体都随机配置给所有的处理。

例如，关于电视广告效果的测试：重复观看同一个广告的效果如何？

以大学生为受试对象，测试广告效果与广告长度、播出频率之间的关系。让所有受试对象都观赏一档40分钟长的电视节目，其中包含某种数码相机的广告。一些广告播出30秒，而其他广告则播出90秒，同一个广告在该电视节目中出现1次、3次和5次。节目结束后，所有受试对象都要回答他们对广告内容的印象，对该相机的好恶，以及是否有购买意愿的问题。其设计过程如表3-2所示。

表3-2 完全随机化实验设计的设计过程列表

实验目的		重复观看同一个广告的效果测试
反应变量		对相机的喜爱程度•或者是对相机的购买意愿的强度
解释变量	广告长度	2种水平：30秒与90秒
	重复次数	3种水平：1次、3次和5次
处理	6种处理	30秒的广告，播放1次、3次、5次 90秒的广告，播放1次、3次、5次

实验时假设招募60个学生（各种背景信息相似），将其随机分为6组，再随机配置在6种处理中。其反应变量的差异来源于处理和实验误差，可采用方差分析方法。

### 2. 适配与随机区组设计

人们发现将受试对象进行适配或区组设计，可以得到比完全随机化设计更高的精确性。结合适配与随机化的常用设计就是配对设计。结合区组与随机化的设计就是随机区组设计。

配对设计只比较两项处理，首先选取“成对”的受试对象，同“一对”中的两个要尽可能接近。然后利用抽签或投掷硬币的方式随机决定把两个处理分别指派给每“一对”当中的两个受试对象；有时候配对设计中的“一对”，实际上只包含一个受试对象，只是分时间先后分别接受两个处理。此时每个受试对象就是他或她自己的控制组，接受处理的顺序可能影响受试对象的反应，所以会再次使用抽签或投掷硬币的方式来随机化每个受试对象接受处理的顺序。配对设计中的随机化不是完全的随机化，只是在每一个配对组中进行了随机化。经过配对，减少了受试对象之间的变异所产生的影响。配对设计是区组设计中的特例。

一个区组就是一组实验个体，在实验之前研究者认为这些个体在影响反应变量的某些方面具有相似性。区组设计中，只在每个区组里面随机指派个体给各处理。

区组设计结合了用适配（选取成对的受试对象）来制造相近的处理组的概念，以及用随机方式产生各处理组的原则，区组是另一种形式的控制，通过把外来变量引进实验里来形成区组，可以控制这些外来变量的影响。

续前例。完全随机化设计把所有受试对象，包括男性和女性，全部放在一起考虑，“随机化”的部分是把受试对象分派到6个处理组，完全不理会他们的性别差异。比较好的设计是将男性和女性分开考虑：将女性随机指派到6个组，每组看一个广告，再将男性也随机指派到6个组。其设计过程如表3-3所示。

表3-3 随机区组设计的设计过程列表

实验目的		重复观看同一个广告的效果测试
反应变量		对相机的喜好程度，或者是对相机的购买意愿的强度
解释变量	广告长度	2 种水平：30 秒与 90 秒
	重复次数	3 种水平：1 次、3 次和 5 次
区组	男性	30 秒的广告，播放 1 次、3 次、5 次 90 秒的广告，播放 1 次、3 次、5 次
	女性	30 秒的广告，播放 1 次、3 次、5 次 90 秒的广告，播放 1 次、3 次、5 次

这样，广告研究中可以分别对男性和女性做结论。分区组也可以把男性和女性之间的系统性的差异去除掉。区组的观念，是统计实验设计中另一个重要原则。经验丰富的实验者会根据实验受试对象之间重要且无法回避的差异来源进行区组设计，然后随机化会把余下的变异平均，使得处理之间能进行无偏的比较，

#### （五）市场实验法的应用范围

市场实验法主要应用于检验有关市场变量间因果关系的假设，研究有关自变量对因变量的影响或效应主要有：

1. 测试各种广告的效果；
2. 研究商品的价格、包装、陈列位置等因素对销售量的影响；
3. 研究品牌对消费者选择商品的影响；
4. 研究颜色、名称对消费者味觉的影响；
5. 测试各种促销方法的效果。

#### （六）市场实验法的优缺点

##### 1. 市场实验法的主要优点

（1） 可以探索不明确的因果关系。通过实验设计，控制一个或几个因子（自变量），尽可能排除外来因素的影响，可以有效地研究各个因子及因子间交互作用对所感兴趣的因变量的效应，并有可能通过适当的统计分析方法找到效应最佳的组合。

（2） 实验的结论有较强的说服力。在实验单位、实验变量、实验设计、实验条件和实验环境都基本相同的情况下，不管是谁来进行实验，也不管是在何时何地进行实验，结果都大致相

同，因此具有较强的说服力

## 2. 市场实验法的主要缺点。

(1) 费时。市场实验一般持续的时间都比较长，特别是正式随机化实验，要求进行事前和事后测量，当所考虑的自变量的个数以及外来变量的个数比较多的时候，实验都可能持续比较长的时间。

(2) 费用高。市场实验在技术上和分析上都有一定的难度，因此一般费用较高，特别是当要求实验组、控制组和多重测量都考虑时，费用又会大大提高。

(3) 管理、控制困难。市场实验比较难于管理，特别是在现场的环境中，既要考虑不影响公司的日常工作，又考虑要取得零售商，批发商或其他人员的合作，还要考虑控制外来变量的影响，这些都是比较困难的事情。

(4) 保密性差。如果是现场实验，那么研究计划是很难不暴露的。竞争对手可能会“故意破坏”现场实验的结果。

g-\* -HR  
第二 T>

# 各种数据采集方法的比较

由于数据采集方法的多样性和适用条件的差异性，需要根据研究目的、研究对象和现实条件的不同进行合理的选择，因而存在基本的评价准则和选择原则。

## 一、数据采集方法的评价准则

一种市场信息收集方法的优劣并不是绝对的。在某种条件下，一种方法可能是相对合适的，但在另一种条件下，这种方法可能就不太合适。因此，对数据采集方法的评价离不开研究的总体特征、问题的性质、经费和时间的限制、项目管理的环境、调查精度的要求等。一般来说，可以遵循以下原则：

### (一) 与研究总体采集的特征信息相一致

一般来说，研究总体能够采集到的特征信息都通过抽样框进行描述。抽样框是联系目标总体与调查单元之间的桥梁。任何抽样调查都必须有一个抽样框。即使一些非概率抽样不需要有具体的抽样框，但也要有一个概念的抽样框。抽样框的信息越是丰富、齐全，选择数据收集方法的余地就越大。而当抽样框的信息不全时，有些收集数据的方法就受到限制。例如，抽样框中缺乏地址，就不能将问卷邮寄给被调查者；抽样框中没有电话号码，电话调查也不能进行。虽然现在有一种随机拨号的电话调查，它不需要具体的抽样框，使用这种电话调查方法时，需要注意目标总体的所有单元是否都有电话，此外，这种随机拨号的空号率比较高。

目标总体的特征信息也影响数据收集方法。如果总体的文化程度不高、识字率较低，就不可能使用邮寄调查方法；而只能选择面访调查或调查员辅助式方法。总体和样本的地理分布也很重要，如果总体和样本散布于全国各地，面访调查的费用就会非常大，也不易管理。除非一些重要的调查必须采用面访调查外，一般的调查采用电话调查或邮寄调查更为合适。

### (二) 与研究问题的性质与要求相吻合

如果所研究的问题是探索性的，那么定性研究的几种方法就比较合适；如果要估计市场的

容量和需求，那么随机抽样调查可能就是比较好的方法；如果要检验因果关系，那么定量研究的市场实验法可能就是最好的选择。调查问题的性质也会影响收集数据的方法，如果涉及一些敏感性问题，采用匿名的收集数据方法比较好，比如用邮寄调查方法可能比较合适；或者如果要采用面访调查，则需要采用随机化设计的调查方法。如果调查的问题比较复杂，一般应该使用调查员进行访问调查，以便于向被调查者解释有关问题和概念。在市场调查中，如果调查的问题需要辅之以向被调查者出示一些材料（如图片、样品），一般以派人访问比较好。

例如，如果要估计某市居民家庭消费的结构、品种、数量和变化特征与趋势，最好的方法可能是固定样本的日记调查法；如果要了解企业对某大型自动化办公系统的购买意向，那么对主管领导的电话调查可能会是针对性较强的方法。民营经济发展景气调查中对企业家的调查，采用电话调查方式比邮寄调查更具针对性 a

### （三）满足调查项目的时间、费用等要求

调查的经费来源、人员、设备和时间等因素都会对数据的收集方法产生重大的影响。如果要采用调查员面访的数据收集方法，就必须有足够的经费来雇佣和培训调查员，还要能够招聘到合适的调查员。如果选择用计算机辅助式方法的话，就需要专门的计算机程序员以及必要的计算机设备。

如果入户面访调查和街头拦截式面访调查都能满足研究的要求，那么后者肯定费用较低，因此是相对较好的方法。

收集数据的时间限制对于选择收集数据方法也是一个重要的因素。邮寄调查明显慢于面访调查，而纸张式调查又慢于计算机辅助调查，需要根据时效性要求进行合理选择 c

### （四）更加易于操作和管理控制

有些收集数据方法要比另外一些方法容易管理。例如，集中的电话调查要比地理上分散的个人面访容易进行组织，整群抽样比随机抽样的面访更加易于管理控制。

如果深层访谈和小组座谈会都能满足研究的要求，例如需要了解现代青年对智能型高科技产品的消费观念，那么在专用座谈会会议室、设施齐备，且拥有经验丰富的主持人的情况下，小组座谈会的方法可能更易于操作和控制，因而可能是更适合的方法。

### （五）尽可能确保精度要求

数据质量要求也是确定数据收集方法的一个重要因素。对于调查中使用的概念，如果调查员接受过有关方面的较好培训，就能减少回答误差。可接受的无回答水平是另一个考虑的因素。一般情况下，邮寄调查回答率低，而面访调查的回答率最高。抽样调查中精度要求是确定样本量和费用的重要因素 [、调查方法需要的单位费用越高，在相同的费用下能负担的样本量就越小。考虑多种因素的常用数据收集方法的评价比较结果见表 3-4。

表 3-4 几种主要数据收集方法的评价比较•

评价准则 (按高、中、低三级)	电话访问	CATI	入户面访	拦截面访	C \1>1	一般邮寄	邮寄 (固定样本)
数据收集的灵活性	中	中高	高	高	中高	低	低
问答题的多样性	低	低	高	高	高	中	中
有形刺激的使用	低	低	中高	高	高	中	中
样本控制	中高	中高	可能高	中	中	低	中高
数据收集环境的控制	中	中	中高	高	高	低	低

续表

评价准则 (按高、中、低三级)	电话访问	CATI	入户面访	拦截面访	<M'1	一般邮寄	邮寄 (固定样本)
实施力的控制	中	中	低	中	中	高	高
数据的数量	低	低	高	中	中	中	高
回答率	中	中	高	高	高	低	中
保密性	中	中	低	低	低	高	高
社会合意性	中	中	高	高	高	低	低
获取敏感信息的可能	高	高	低	低	低中	高	高
可能由调查员造成的偏差	中	中	高	高	低	无	无
速度	高	高	中高	中高	中高	低	低中
费用	中	中	高	中高	中高	低	低中

\* : 本表来自柯惠新、刘红鹰. 民意调查实务. 北京: 中国经济出版社 J996J43

例如, 在新产品属性测试的研究中, 既可以采用直接询问的常规大样本面访调查法, 也可以采用间接询问的使用联合分析法(详见第十四章)的较小样本面访调查, 相比之下, 后者可能得到更合理、精度更高的结果。

## 二、数据采集方法的选择原则

为了满足研究目的和任务的需要, 数据采集方法的选择一般会遵循以下原则:

### (一) 选择最有效的组合式调查方法

市场调查主题一旦确定, 调查对象和研究总体的特征就清晰起来, 需要综合研究问题的性质、研究目的和要求、经费和时间限制等多方面因素进行分析, 选择适当的调查方法的组合, 来完成预定的市场调查项目。例如, 一项完整的市场调查项目, 常常有经费上的限制和时间上的要求, 受委托的调查机构选择调查方法的原则是定性研究和定量研究相结合、以定量研究为主(即在定性的文案调查(有关主题的二手资料的收集和研究)和深层访谈(访谈重要客户)的基础上, 设计定量研究的问卷, 抽取有代表性的随机样本, 进行入户面访调查, 通过对问卷做深入的统计分析, 为下一步的营销决策提供相应的参考资料和建议, 达到市场调查的预期效果)。

### (二) 用户至上原则

在选择调查方法的组合时, 市场研究者常常面临着两难的局面, 是降低研究标准以适应客户在费用和时间上的苛刻要求, 还是坚持研究的高标准, 说服客户提高费用和时间的预算? 后一种做法的结果常常是以失去业务告终。

经验丰富的研究者需要遵循“用户至上”原则, 认真地研究背景资料, 精心地设计研究方案, 选择最有效的调查方法的组合, 使客户的经费和时间尽可能地节约; 同时又要基本达到市场研究的预期效果。做到这一点, 市场调查研究者需要有深厚的理论功底, 敏锐的市场洞察, 以及长期从事市场调查实践的经验积累。

### (四) 本章小结•••

本章主要阐述了定性与定量研究方法在市场调查中的应用。在寻找处理问题的途径时, 定性研究常常用于拟订假设或是确定研究中应包括的变量; 定性研究是以小样本为基础的无预定程序的、探索性的研究方法, 目的是对问题的定位和深入分析提供比较透彻的理解和认识。定性研究方法主要包括文案调查法、

## 小组座谈法、深层访谈法和投影技法 r

定量研究方法是指通过调查、观察和实验的方法收集资料，可运用统计技术分析，并且最后结果可以用具体的数据来表示的调查方法。定量研究方法分为调查法、观察法和市场实验法三大类型。调查法是利用从总体中抽取样本，运用设计好的问卷，从被调查者中获取所需具体信息的定量研究方法。这种结构式的直接调查，是最为常用的数据收集方法。根据调查手段与介质的不同，调查法分为传统自填问卷调查法、网络调查法、面访调查法、电话调查法。观察法是调查者凭借自己的感官和各种技术工具，深入调查现场，直接观察和记录被调查者的行 为以收集市场信息的一种方法。市场实验法是指市场调查者有目的、有意识地改变一个或几个影响因素，观察市场现象在这些因素影响下的变动情况，以认识市场现象的本质特征和发展规律。按照实验的设计方式不同，可分为非随机设计和随机设计实验。实验的有效性由内部有效性与外部有效性来评估。

各种定性研究与定量研究方法各有其利弊，调查员在选择具体的调查方法时，必须综合考虑调查目标、具体问题、时间限制、经费预算、数据精度和被调查者的便利程度等因素。 磁关名词 · · ·

定性研究 文案调查 实地调查 第二手资料 第一手资料

小组座谈法 深层访谈法 投影技法 定量调查法 面访调查法

邮寄调查法 电话调查法 网络调查法 观察法 市场实验法

实验组 控制组 实验误差 内部有效性 外部有效性

完全随机化设计 配对设计 随机区组设计

回喷回

讒

回\*廠即測即評

- 肯思考题.....
1. 定性研究的优点是什么？
  2. 定性研究主要有哪几种方法？
  3. 二手资料与一手资料有何区别？
  4. 收集营销管理决策方面的信息，通常可从哪些途径去寻找？
  5. 任意选择一种搜索引擎，举例说明如何通过网络搜寻二手资料？
  6. 小组座谈法的实施过程大致可分为几个阶段？
  7. 深层访谈法的优点、缺点是什么？
  8. 深层访谈法的实施一般要经过哪几个步骤？
  9. 投影技法有什么特点？比较常用的投影技法有哪几种？
  10. 电话、面访、网络、邮寄调查法各自具有哪些优缺点？试各举出一个适用该调查法的 市场调查项目
  11. 什么是观察法？观察法有哪些分类与观察技巧？

12. 什么是市场实验法？市场实验法有哪些类型？
13. 什么是实验的内部有效性与外部有效性？
14. 如何控制外部变量与外部变量影响的类型？
15. 选择数据采集方法时需要考虑哪些因素？

碧本章实训. . .

### 一、实训目的

1. 熟悉常见的观察调查法(也可以是问卷调查法、实验法等)，掌握如何运用这些观察法 进行市场调查：

2. 能够结合调查实际需要，选择合适的观察方法，选用合适的观察技术，进行观察调查。

### 二、实训内容

1. 观察所在学校附近的购物中心市民购买手机的特点 C

2. 观察所在学校附近共享自行车的消费特点。

3. 观察所在城市消费者拥有私家车的特点。

每一个选题下完成训练的内容包括：

第一步，明确观察目标。观察目标，即通过观察法所能获得的进行某种决策的依据. 明确 采用观察法需要调查哪些信息。

第二步，确定观察方法。

第三步，选定观察对象。

第四步，选用观察技术。

第五步，确定观察时间。

第六步，确定观察地点。

第七步，记录观察结果。在使用观察法进行调查时，良好的记录技术可以及时记下转瞬即逝的宝贵信息，加快调查工作的进程，便于资料的整理及分析。

第八步，观察结果报告。观察活动结束后，需要对记录下来的观察结果进行处理。对于观察到的现象，需经专业人士根据经验进行分析判断，概括、提炼出有用信息；保存好观察调查的所有资料，以便查寻；对有用的信息进行分析，得出进行决策的依据，并形成调查报告。

### 三、实训组织

1. 以团队为单位完成实训任务。

(1) 选派 10~20 名同学到所在城市各大手机专卖店进行观察，了解目前市场上哪些是 主打品牌、主打款式，以及哪些是经销商最看好和最不看好的品牌。跟踪顾客，了解顾客在 哪个品牌、哪个款式前滞留的时间最长，并表现出浓厚的兴趣 Q 观察顾客选购手机的心理。

(2) 选派 10~20 名同学到学校周边进行观察，了解所在城市市民的代步工具中共享自行车的比例，以及目前消费者使用最多的共享自行车品牌。

(3) 选派 10~20 名同学到所在城市各大私家车经销点进行观察，了解目前市场上哪些是 主打品牌、主打款式，以及哪些是经销商最看好和最不看好的品牌与款式。跟踪顾客，了解顾客在哪个品牌、哪个款式前滞留的时间最长，并表现出浓厚的兴趣。观察顾客选购私家车的心理.

2. 各组在不同的场景完成数据采集，并形成书面报告

3. 班级各组交流，教师点评，



## 第四章量表测量技术

### 、导语

随着人工智能技术不断取得新的突破，一系列智能生活硬件逐步走进普通大众的日常生活中。智能手环凭借入门级的售价和产品方案的标准化，市场渗透率在智能可穿戴设备领域中一马当先，远远超过智能手表、智能眼镜等其他产品。智能手环作为新兴产品，其市场潜力还未完全释放，正处于承前启后的发展阶段。现阶段消费者对智能手环的评价如何？智能手环厂商今后应采取什么样的发展策略？针对这些问题，拟进行一次关于智能手环用户群的市场调查。项目组围绕研究目的，经过反复讨论，决定对居民群体就表4-1中变量展开一次抽样调查。

表 4-1 智能手环用户群调查变量

序号	模块	变量
1	个人背景	①性别②年龄③学历④职业
2	智能手环 使用情况 及产品偏好	① 智能手环广告接触情况 ② 智能手环相关信息获取渠道 ③ 购买智能手环考虑的主要因素及其重要程度 ④ 对智能手环功能的预期 ⑤ 智能手环获取途径 ⑥ 所使用的智能手环的品牌及其印象 ⑦ 对所使用的智能手环的外形与材质、功能与服务、使用体验等方面满意度评价 ⑧ 是否打算放弃使用智能手环 ⑨ 放弃使用智能手环的原因 ⑩ 非智能手环用户是否打算购买智能手环及购买动机
3	个人特质	①月收入②最感兴趣的休闲活动③个人行为习惯

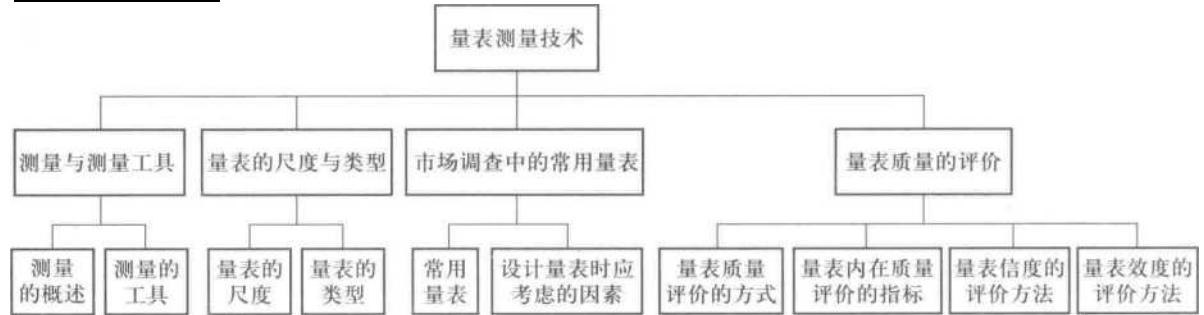
在上述三大模块中，一共涉及 17 个变量，调查过程中需要对每一个被调查者就这些变量逐一进行调查或者测量。在 17 个变量中，有些变量性别、年龄、学历与职业，智能手环广告接触情况、智能手环相关信息获取渠道等属于客观变量；有些变量如对智能手环功能的预期、对智能手环各方面的满意度评价、智能手环购买动机等属于主观变量；有些变量如月收入属于敏感性变量。如何对不同类型或性质的变量进行测量是需要认真分析的。市场调查中的测量工具主要是量表。对于不同的变量，不同的计量水平要求，使用的测量的量表也不同。

本章介绍测量的尺度与类型，常用量表的测量技术，设计量表应考虑的因素，量表质量的评价方式及方法。

### **、教学目标与要求**

1. 了解测量的含义、种类与要求。
2. 了解量表的含义与作用。
3. 掌握量表的计量层次，熟悉量表类型的划分。
4. 系统掌握市场调查中常用量表的表达格式、特点与作用，能够在问卷设计中熟练设计量表。
5. 熟悉量表设计中需要考虑的主要因素。
6. 掌握量表质量评价的一般方式与基本过程。
7. 掌握量表信度与效度的含义及其区别与联系。
8. 掌握量表信度与效度检验的方法。

## 、本章知识结构



市场调查中常常要搜集被调查者心理活动与主观意识方面的数据，心理活动往往比较复杂，被调查者对自我的主观意识往往缺乏知觉而难以准确表达出来，给调查造成一定的困难。此外，有关人们心理活动与主观意识方面的调查结果，多属于定性数据，给调查数据的分析带来一定的难度。所以，在市场调查活动中，如果需要调查被调查者的态度、意见、偏好、动机、认知、评价等心理与主观方面的问题，就必须采用专门的量表测量技术，以使调查结果尽可能准确可靠，且便于数据处理与分析”

## 第一节

### 测量与测量工具

#### 一、测量的概述

##### （一）测量的定义

测量是指按照特定的规则，将数字或符号分配到研究对象的特征上，从而将其特征量化的过程。本质上讲，测量的过程是一个分配数字或符号的过程，用这些数字或符号来反映研究对象所具有的特征。需要说明的是，测量的对象并不是所研究的对象本身，而是研究对象的特征。例如，在消费者调查中，测量的对象并不是某个消费者，而是消费者的诸如态度、动机、品牌忠诚度、收入、年龄等特征变量。

测量需遵循一定的规则。测量的规则是指导调研人员如何实施测量的操作指南和具体方法。测量的规则必须在测量活动开始之前就预先设计出来，测量规则的设计包括了确定测量对象与测量单元、确定测量单元的属性特征、划分属性特征的状态类型、规定数字或符号对相应特征状态的分配方式。例如，在调查消费者对某品牌手机的喜爱程度时，测量对象就是所要调查的消费者群体，测量单元就是消费者个体，测量单元的属性特征就是消费者个体对某品牌手机的喜爱程度，喜欢程度可以划分为非常喜欢、比较喜欢、一般、不太喜欢、很不喜欢五种状态，然后规定将5、4、3、2、1依次分配给五个不同的喜欢程度。这一系列的规定，都属于测量规则设计的范畴。测量规则的设计有难易之别。例如，测量性别就比较容易，性别特征只能划分为两种状态即男性和女性，然后可以规定用“1”表示男性，用“0”表示女性。又如，研究者欲测量消费者对品牌的忠诚度、购买动机等，这些属性特征属于心理活动与主观意识的范畴，并没有明确的表现形态，它们存在于人的大脑内部，也没有既定的状态划分标准，测量起来具有较大的不确定性。此时测量规则的设计就比较困难。

##### （二）测量的种类

测量活动有多种类型，根据不同的标准可以划分出不同的测量类型。市场调查中的测量活动可以做下列划分：

###### 1. 根据测量的内容不同，测量可分为态度测量、行为测量、背景测量

态度测量就是运用测量工具，对被调查者的可能认识或认识态度进行测量。态度是指人们脑海中对某件事物或某个环境的认识、判断以及指导他们的反应、行动的某种状态。态度与个人的基本价值观和基本需要有联系，会造成自我卷入程度上的差异。态度不是直接观察到的，而是通过可见反应显示出来的。可见反应可区别为三类，即认知反应（同意或不同意）、情感反应（喜欢或不喜欢）和行为反应（支持或反对）。对态度进行测量，通过对上述任何一种反应或三种反应的组合的测量来实现。在大多数市场调查中，对态度的测量是其中重要的调查内容。

行为测量就是运用测量工具，对被调查者在某些方面的行为习惯或行为特征进行客观的测量。市场调查中所指行为是消费者行为。消费者行为就是消费者寻找、购买、使用和反馈用以满足需求的商品和服务。

所表现的一切脑体劳动。人们的行为往往取决于他们对待事物的认 识与态度，也取决于他们内心深处的需求。通过对消费者行为的测量，可以分析消费者行为的 特征、影响消费行为的主要因素以及他们的消费态度。

背景测量就是运用测量工具，对被调查者的有关背景信息进行测量”被调查者的个人背景主要包括自然背景和人文背景。自然背景是指被调查者的自然属性，被调查者的属性不同，其自然背景也会不同。如果被调查者是消费者个人，则其年龄、身高、性别等就属于自然背景 信息，如果被调查者是企业，则其所属地区、企业规模等就属于自然背景信息。人文背景是指 被调查者的一系列社会属性。对于消费者个人而言，其学历、职业、收入等就属于人文背景信 息•对于企业而言，其所属行业、企业类型等就属于人文背景信息如果被调查者的背景与调 查研究的课题相关联，此时就需要对被调查者的背景进行测量。

### 2. 根据测量的功能不同，测量可分为检验性测量与研究性测量

检验性测量是指测量结果用于对测量工具的信度与效度的检验。在市场调查活动中，往 往会使用问卷进行调查，问卷中会设计各种测量工具•这些测量工具测量的结果是否具有稳定性与准确性，是需要检验的这种检验一般要抽取一个小样本，对被调查者属性特征进行各种 测量，再根据测量结果，运用统计方法检验测量工具的信度与效度如果测量工具符合要求，则可在大规模正式调查中加以使用。

研究性测量是指测量的结果用于对特定课题的分析与研究研究性测量发生于正式调查 之中，正式调查的结果将用于对特定课题中的各种问题进行分析，得出研究结论研究性测量 所使用的测量工具是经过检骑符合要求的，对于达不到要求的测量工具，必须进行修改或重新 设计。

### 3. 根据测量的标准化程度不同，可分为标准化测量与非标准化测量

标准化测量就是对不同的测量单元按照统一测量工具，对其属性或特征进行测量问卷 中封闭式问题的调查就属于标准化测量-标准化测量所获得的结果具有相同的数据格式和确定的测量结果，便于统计处理与分析

非标准化测量就是不采用统一的测量工具对不同测量单元的属性或特征进行测量。问卷 中开放式问题的调查以及非结构式面访调查，都属于非标准化测量：非标准化测量的结果往 往不具有统一的数据格式，测量的结果具有不确定性，不便于进行统计处理与分析

### （三）测量的要求

任何一次有效的测量，都应该尽可能达到下列要求：

#### 1. 稳定性

稳定性是指测度结果的一致性和可靠性程度。如果在相同条件下，对同一测量单元就同 一测量内容，运用同一测量工具重复多次进行测量的结果是一致的，就表明测量工具具有稳定 性。测量工具的稳定性越高，测量结果的可信度就越高。对于测量工具的稳定性，是需要通过 检验性测量来进行检验的，然而，测量工具具有了稳定性，并不意味着该测量工具就是符合要 求的。这是因为测量工具的稳定性并不能代表该测量工具不存在测量误差如果测量工具设计的不科学，数字或符号分配规则不合理，则测量的结果就不准确。

#### 2. 准确性•

准确性是指测量工具准确地反映客观事物特征和属性的程度。测度总是为了一定的目的 进行的，判断测量工具的准确性是评价其达到测量目的的程度。如果一个测量工具能够正确 地测量出所测对象的属性或特征，则该测量工具就具有有效性。测量工具的准确性越高，测量 结果的有效性也就越高。测量工具的准确性同样需要检验性测量来进行检验。

#### 3. 完备性

完备性是指测量规则中必须包括所测变量的各种状态或变异。由于测量目标并不是被测 对象本身，而是被测对象的某种属性或特征，在设计测量规则时，被测对象的属性状态或特征 状态究竟有哪些，研究者必须事先在有关理论和经验的指导下做出科学的划分 c 划分出的属 性类型或特征类型状态必须完备，不能有任何遗漏。否则可能有部分被测单元的属性或特征 不能被测出或不能被准确地测出。

#### 4. 互斥性

互斥性是指被测单元的属性或特征都能够用一个且只能用一个数字或符号来表示。互斥性要求在设计测量规则时，所划分的所测对象的属性或特征的状态类型不能有任何重叠，一个状态类型必须独立于其他任何一个状态类型，这样才能保证一个状态类型对应唯一一个数字。否则，被测单元的属性或特征就不会是唯一的，会存在多个测量结果。究竟哪个结果是正确的，测量活动本身是无法告诉我们的。这种测量结果的不确定性对测量结果的准确性会造成严重危害。

## 二、测量的工具

测量的对象既可能是有意识的现象，也可能是无意识的现象；测量的内容既可能是物理性测量，也可能非物理性测量。无论是什么对象和什么内容的测量，都离不开测量工具。测量工具按照其测量原理，可分为两大类，一类是物理测量工具，另一类是量表。市场调查活动中，应该根据不同的测量对象和不同的测量内容，选择不同的测量工具。

### （一）物理测量工具

物理测量工具是用来对物理现象的属性或特征进行测量的各种工具的总称。物理测量工具主要包括各种用于测量的仪器、仪表、度量衡等。市场调查中一般不使用物理测量工具进行测量，但在进行特定调查活动中，也可能会使用物理测量工具。例如，人们对外来刺激所引起的心理、情感、情绪、认知的反应是难以用肉眼观察出来的，必须借助物理测量工具来进行观察。如果要测量受试者对某项活动的反应，可采用视向测定器；要测量受试者对广告要素的了解程度，可采用瞬间显示器；要测量受试者情感与心理变化，可采用精神电流测定器；要测量受试者情绪变化，可采用皮肤电流反射器，这些仪器都属于物理测量工具。对于其测量的结果，需要通过比较专业的分析，间接判断受试者心理活动状态。

### （二）量表

量表是指通过一套事先设计的用语、符号和数字，来测量人们心理活动的度量工具。

市场调查中常常要对被调查者的态度、意见、感觉、意向、偏好、印象等心理活动方面的问题进行辨别和测定，这些心理活动存在于人们的大脑之中，不能够直接进行测量。例如，要调查消费者对某品牌手机的喜爱程度，对某项服务的满意程度，购买某类商品时对各种因素的考虑程度等问题时，我们无法用数字来直接测量这类心理活动的状态，此时，就需要设计相应的测量工具，将这些无法用数字反映的属性或特征状态进行量化，从而实现对这些定性问题的调查。人们将这些调查中使用的将定性问题转化为用数字来回答的测量工具或尺度概称为量表。

量表本是心理学上用于测量人们心理活动与人格特征的工具，由于市场调查活动中也常常涉及被调查者有关心理活动方面的问题，所以，量表也是市场调查中一种重要的测量工具。

量表测量技术在问卷调查中具有十分重要的作用。首先，量表能够将不能直接测量的定性问题转化为能够直接测量的变量，提高了测量的可行性；其次，通过量表所调查的数据，具有较高的标准化程度，便于对数据进行统计处理与分析。这两方面的作用，决定了量表在问卷调查中具有重要的地位。

设计一份具有操作意义与准确测量功能的量表，是一项既复杂又有较大难度的工作。心理学家们为我们提供了许多现成的量表格式，但面对特定的调查内容，有些现成的量表格式可能并不适用，需要研究者根据具体的测验对象、测量目的、测量内容自行设计科学合理的量表。

## 第二节

# 量表的尺度与类型

获取不同问题的信息需要采用不同的量表，量表具有不同的尺度。下面介绍常见的量表尺度，以及根据不同标志划分的量表类型

### 一、量表的尺度

量表作为一种测量工具或者计量尺度，其精度越高越好’但事物的属性或特征有各种各样的表现，尤其是一些定性的属性或特征，期望通过量表来加以完全精确的测量是不现实的。人们只能够根据所测事物的属性或特征状态，设计具有相对精度的量表

所谓量表的尺度，也称为量表的计量层次，是指量表对客观事物、现象测度的程度或精确水平。量表根据其计量层次由低级到高级、由粗略到精确分为四类，即定类量表、定序量表、定距量表和定比量表。采用不同的计量尺度可以对不同类型的市场调查问题进行测度，并搜集到相应的数据 C

#### （一）定类量表

定类量表也称为类别量表，或命名量表、名义量表，是用于测量调查对象属性或类别的一种工具或尺度。运用该量表时，首先需要根据研究的目的，对调查对象属性或特征的类别进行划分，并给每一个类别分配一个且只能一个数字符号，一个数字符号代表一个类别。

##### 例 4.1

您的性别:①男②女

上例中”①男、②女”就是测量被调查者性别属性的一个定类量表。

运用定类量表最关键的是对调查对象属性或特征的类别进行划分。有些类别划分具有客观性，如性别类型的划分就不能主观随意，只能划分出男性与女性；又如职业类型的划分应该按照国家职业分类标准来进行，设计者不能随意设定职业类型 Q 有些类别的划分具有一定的主观性，这主要取决于研究的目的。例如，要了解消费者所在的地理区域，如果研究者感兴趣

的是城市与农村市场的差异，就可以将地理区域划分为城市、农村和郊区。如果研究者感兴趣的是消费者不同地理方位的市场差异，则可以将地理区域划分为东北地区、西北地区、东南地区、西南地区和中部地区。调查对象属性或特征类别的划分，一定要遵守穷尽性原则与互斥性原则

定类量表需要对每一个类别赋予一个唯一的数字符号，一个数字符号只能对应于一个特定的类别，这些数字符号并没有数学意义，它们只是某个特定类别的标签或代码，不能作任何算术运算。各类别之间是并列关系，不能比较大小、高低。

定类量表的计量层次最低，所测得的数据称为类别数据，只能运用类别数据的分析方法进行分析：从描述性分析角度、可以进行各类的频次与频率分析，运用众数描述数据的集中趋势。从推断性分析角度，可以进行类别数据相关性或独立性的卡方检验等。

### （二）定序量表

定序量表也称为顺序量表，是用于测量调查对象在某种属性或特征上等级差异的一种测量工具。定序量表较定类量表要高一个计量层次，它不仅将属性或特征的状态进行了分类，而且各类别之间存在类别差异或顺序差异，各类别可以比较属性的优劣、程度的高低等。如产品按质量属性可以划分为优等品、一等品、二等品、等外品，这4个质量等级之间就存在差异，等级越高表明产品质量越好。又如要调查消费者对某款商品功能的喜欢程度，可以将其评价态度划分为非常喜欢、比较喜欢、一般、不太喜欢、很不喜欢5个等级，这5个等级之间存在顺序差异，可以比较态度倾向，也可以比较喜欢程度。定序量表需要将所划分的等级、程度按一定顺序从高到低或从低到高排列，便于被调查者选择。

#### 例 4. 2

您对山地牌自行车价格的评价态度是：

- ①非常贵 ②比较贵 ③一般 ④较便宜 ⑤很便宜

该量表中的5个选项按评价态度由正面到负面及其程度排序，并给每一种态度赋予唯一的一个数字，形成一个定序量表，量表中的1-5这5个数字，也没有数学意义，不能作任何算术运算，每一个数字仅仅代表所对应的态度。比如，“1”表示非常贵，“2”表示比较贵，两个数字之间相差一倍，就不能说“非常贵”比“比较贵”在态度上多一倍，这种解释没有任何意义。

定序量表还有另外一种表述形式。例如，要调查人们对四种不同品牌洗衣粉的喜欢程度，可以按下列方式进行测量。

#### 例 4. 3

下列有四种品牌的洗衣粉，请按你的喜爱程度从1-4进行排序（1表示最喜欢，4表示最不喜欢，数字填入品牌名称后的括号中）：

碧浪【】

佳美【】

白猫【】

奥妙【】

通过被调查者填列的数字，可以测量其对四种不同品牌洗衣粉由高到低的喜爱程度。根据规定的数字来看，上述两个例子中使用了等间隔排序的方式。其实，定序量表本质上就是一个排序过程，只要是大小不同的数字都可以用于上述不同品牌喜欢程度的排序，而且并不会改变顺序。但设计量表时应该尽可能使用等间隔的排序方式，以便更有利于进行统计处理。需要指出的是，定序量表测量的结果只能说明人们对某项事物或行为态度上存在等级差异，但各等级之间究竟有多大的差异，是不能通过该量表测量出来的。

定序量表测量所获得的数据属于顺序数据，类别数据的分析方法也适合于顺序数据。另外，顺序数据还可以采用中位数、百分位数、等级相关系数、肯德尔和谐系数等指标进行分析，也可以进行符号检验、秩检验、秩方差分析等。

### （三）定距量表

定距量表也称为等距量表，是用于测量调查对象在某种属性或特征上的等级差异及差异绝对水平的测

量工具。定距量表较定序量表又高了一个层次，它不仅能够测量等级差异，而且能够测量这种差异绝对水平的大小。更明确地说，就是各等级之间的顺序与各等级之间具体的差异程度都能够被测量出来。因而，定距量表中用于测量的尺度都是刻画事物某方面属性或特征的数值。例如，采用 50 分制打分的方式，测量消费者对甲、乙、丙三个超市的整体满意度，某个被调查者给出的分数是 30 分、20 分、10 分。据此可知，该被调查者对三个超市的满意态度存在三个类别，即最满意、一般、最不满意，而且可按满意态度的程度对三个超市进行排序。显然，定距量表兼有定类量表与定序量表的特点。同时，我们还可以知道，该被调查者对三个超市满意态度上的具体差值。即对甲的评分比对乙高 10 分，对乙比对丙高 10 分。毫无疑问，这种测量结果就更加精确。

定距量表尺度中任意相邻的两个刻度之间的距离必须是相等的。比如，上述评分的计量单位是分数，而两个评分的最小差异值是 1 分。所以，其计量尺度按照 1、2、3 这种离散型数值来标示刻度，1 与 2、2 与 3 之间的距离就是相等的。

虽然定距量表测得的数据可以计算顺序差异的绝对水平，但不能计算顺序差异的相对水平。上述例子中，我们就不能说某个被调查者对甲超市的满意程度是对丙超市的 3 倍。这是因为定距量表上并没有一个用于进行相对比较的绝对零点。所谓绝对零点是指事物本身客观存在的一种数量特征状态。比如，年龄这个变量就存在绝对零点，这个零点是不由研究者来定义的。定距量表上没有这种绝对零点，即便存在一个零点，也只是研究者主观确定的。市场调查中绝大部分变量都属于态度变量，这类变量在客观上就不存在真正的零点。假如上述评价三个超市的整体满意度的打分规则改为 10 分制，测量的结果分别为 70 分、60 分、50 分，显然对三个超市整体满意度的排序及等级差值绝对水平并没有变化，但等级差值的相对水平就不一样了，被调查者对甲超市的满意程度是对丙超市的 1.4 倍，与按 50 分制的打分规则所得的 3 倍比值就不同。正因为定距量表没有一个固定的测量起点即零点，就会导致顺序差值的绝对水平相同但相对水平不同的情况。这说明定距量表的测量结果不适宜进行乘除运算，只能进行加减运算。

由于定距量表的计量层次要高于定类量表与定序量表，所以后两类量表的统计分析方法均适用于定距量表，而且定距量表的测量结果还可以用于计算算术平均值、方差或标准差、相关系数等指标，也可以进行，检验与「检验」。

#### （四）定比量表

定比量表也称为等比量表，是用于测量调查对象在某种属性或特征上的等级差异、差异绝对水平与相对水平的测量工具。由定义可知，定比量表除了具备了定距量表的所有特性外，还具有绝对零点这一特征。等比量表中的绝对零点是客观存在的，量表尺度上的刻度具有类别关系、顺序关系、等距关系与比率关系。

许多变量存在客观上的绝对零点，如被调查者物理特征上的年龄、体重、身高均存在绝对零点，被调查者的收入水平、消费水平、家庭人口数等也存在绝对零点。所以，定比量表的计量起点都是以绝对零点为标准的，所测得的数值就是被测对象特征的实际值，适合于进行相对水平的比较。绝对零点的存在，意味着可以对定比量表的测量结果进行所有的算术运算。定比 第四章量表测量技术

量表的计量层次比定距量表高，任何统计分析方法都是适用的 G

实际调查活动中，究竟采用何种计量层次的量表，首先取决于被测对象自身的特点，其次还取决于测量的目的与研究的要求。例如，性别、职业、居住地等变量本身就是分类属性变量，只能按定类量表进行测量。又如，调查消费者对某品牌商品是否喜欢，就可以采用定类量表，测量出喜欢与不喜欢的人数；如果是调查消费者对某商品的喜爱程度，则要使用定序量表；如果要求消费者按某种打分规则，给某商品的某种属性打分，就需要使用定距量表；如果是要调查被测对象某些特征的实际值，就只能采用定比量表如果对调查的角度要求不高或不可能获得精确的数据，就可以采用计量层次较低的量表。

需要说明的是，能够用较高计量层次的量表类型测量的被测对象，一定可以转化为采用较低计量层次的量表进行测量，但不能反过来。例如调查消费者的消费水平，如果要求的计量层次高，就应该采用定比量表，获得的测量数据就是消费者的实际消费数值；如果研究不需要太高的计量层次，只是需要获得消费水平的等级，可以将量表降低为定序量表，获得的测量数据可以是高消费水平、中等消费水平和低消费水平的人数。

## 二、量表的类型

上述介绍了量表的尺度，也是量表类型的一种划分除此以外，量表还可以根据不同的角度进行划分，形成不同类型的量表。

### （一）根据测定变量数目的多少分类

根据测定变量数目的多少不同，量表可分为一维量表与多维量表。

#### 1. 一维量表

一维量表也称为一元量表，是指用于测量被测对象单一特性的量表。如果一个调查问题只涉及调查对象的某一特征，例如，要了解消费者对某品牌商品价格的评价态度，就适合采用一维量表。有时，对被测对象就某一个概念或属性进行测量，这个概念或属性可能需要分解为多个问题，一个问题对应一个一维量表，就会形成一个一维量表的组合，来共同测量同一个概念或属性。例如要调查消费者对某品牌冰箱质量的评价态度，由于冰箱的质量这个概念可以分解为制冷效果、耗电量、安全性、容量等四个属性，于是就需要分成四个问题，每一个问题对应一个质量特征，需要分别使用四个平行的一维量表来测量。

#### 2. 多维量表

多维量表也称为多元量表，是指用于测量被测对象多个特性的量表。如果调查的内容涉及调查对象多方面的属性或特征，就需要设计一组概念，从不同角度和不同方面对被测对象的特性进行测量，此时就需要使用多维量表。例如，要了解消费者对某品牌冰箱的质量、价格、外观、售后服务的评价态度，就涉及四个不同性质的概念，此时就适合采用多维量表来进行测量。多维量表中对每一个概念进行测量，有时只需要一个一维量表，有时可能需要多个平行的一维量表，形成一个量表体系。

### （二）根据测定的方式不同分类

根据测定的方式不同，量表可以分为分等式量表与排列式量表。

#### 1. 分等式量表

分等式量表是指被调查者在不参照其他事物的条件下，直接确定其对某事物或现象的态度等级。例如，在调查用户对某款手机 APP 使用方便性的评价态度时，可以使用分等式量表。

#### 例 4. 4

您对 xx 手机 APP 使用方便性的评价态度是：

- ①非常满意 ②比较满意 ③一般 ④不太满意 ⑤很不满意

该例采用的量表就是一个分等式量表，数字 1~5 分别代表不同的态度等级。分等式量表属于定序量表，由于定距量表与定比量表均具有定序量表的特性，所以也都属于分等式量表的范畴。分等式量表在市场调

查中被广泛地用于态度与事实的测量。

## 2. 排列式量表

排列式量表是指被调查者首先对若干个受测项目进行比较，然后再按其偏好程度对这些项目排出先后顺序。此类量表主要用于由于其内在的特殊性，无法直接对所有的项目进行比较，只能先做两两比较，然后根据比较的结果进行统一比较分析的问题。下列例子中的量表属于排列式量表。

### 例 4.5

请比较下列每一对不同品牌的洗发液，哪一种您更喜欢？（每一对中只选一个，并在□内画上“V”）

- |      |      |
|------|------|
| □海飞丝 | □潘婷  |
| □潘婷  | □飘柔  |
| □飘柔  | □威娜宝 |

### （三）根据测定中备选答案的对称性分类

在测定被调查者态度时，按照被选答案中有利态度与不利态度的数目是否相等，量表可分为平衡量表与非平衡量表。

#### 1. 平衡量表

平衡量表也称为对称量表，是指有利态度的答案数目与不利态度的答案数目相等的量表。上述例 4.4 中采用的量表就是平衡量表。如果研究者事先并不清楚被调查者的有利态度与不利态度的分布，或欲获得广泛的意见，就应采用平衡量表。平衡量表中由于有利态度与不利态度的答案数目是相等的，因而不会对被调查者的选择产生引导作用，不容易产生回答偏差。采用平衡量表测量的结果，有存在均匀分布的可能，但这种概率是较低的。

#### 2. 非平衡量表

非平衡量表也称为非对称量表，是指有利态度的答案数目与不利态度的答案数目不相等的量表。如果以往的调研或预先研究已表明，大多数的态度都是有利态度或不利态度，则设计量表时就应该考虑这种态度倾向。采用非平衡量表，一是有利于减少答案的分布，简化调查；二是有利于对某种倾向态度的答案设计得更精确、从而进行深入细致的测量；但是，非平衡量表中有利态度答案多于不利态度答案，或不利态度答案多于有利态度答案，都可能使调查结果倾向答案较多的那个态度方向，从而出现回答偏差。

### （四）根据测量语句确定者的不同分类

用于测量的语句可以分为两类，一类是由研究者确定的，另一类是由被调查者来确定的。根据测量语句确定者的不同，量表可以分为直接量表与间接量表。

#### 1. 直接量表

直接量表是由研究者事先设计好有关态度问题的各种语句及答案，直接询问被调查者，由被调查者评定，以反映其态度。大多数场景下，市场调查中用于测量的语句都是由研究者事先设计好的，凡是对事先设计好的语句进行测量的量表，皆属于直接量表。前文中所列举的一系列例子中，用于测量的语句都是事先由研究者设计好的，对这些语句进行测量的量表都是直接量表。

#### 2. 间接量表

间接量表是由被调查者决定测量用的语句，研究者提供答案，由被调查者来评定，以反映其态度。一般情况下，要对被调查者的某些态度特征进行测量，往往是由研究者事先设计好测量语句，但这些语句可能并不是被调查者感兴趣的或所关注的，因此这种测量可能就不会有什么意义。如果测量用的语句由被调查者自己来决定，则测量的内容一定是被调查者所关心的或感兴趣的问题，再使用合理的量表对其进行测量，测量的结果对研究可能更有帮助。Q 后面要介绍的瑟斯顿量表就是典型的间接量表。

### （五）根据测量工具的形式不同分类

测量工具有多种形式，根据测量工具的形式不同，量表可分为文字式量表、图式量表、表式量表。

#### 1. 文字式量表

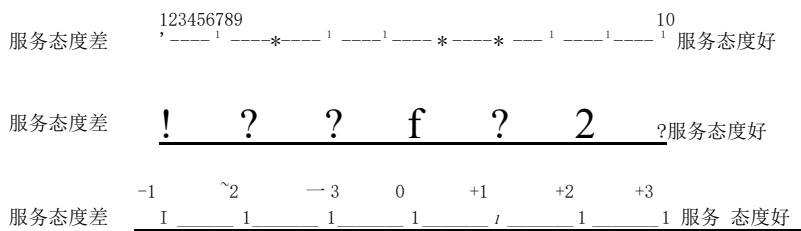
文字式量表是指测量工具上的刻度直接以文字加以表述，由被调查者来选择合适的刻度。用文字来表述的量表有多种具体形式，究竟采用何种形式，主要取决于测量的目的与内容以及设计者的个人偏好。上述例 4.4、例 4.5 所采用的量表均为文字式量表。文字式量表所表达的刻度清楚明了，便于被调查者作答，其不足之处在于，这些刻度往往比较粗略，难以精确地反映被调查者主观态度上的细微差别，有时刻度过于细密，被调查者又难以区分其差别，增加回答难度。但从实践来看，文字式量表使用最为普遍，只要量表的刻度设置合理，一般不会对测量产生太大的负面效应。

## 2. 图式量表

图式量表是指测量工具以图示的形式表现出来，由被调查者在图示上作答。图式量表的设计是有规则的。一般来讲，图示是一段数轴，数轴上最好是标上刻度，以便被调查者分辨。数轴的两端必须标注事先确定的用于评价被测事物某种属性或特征的两个反义词语。要求被调查者在图示上作答，一般是采取标记的方式。下列几个例子中的量表就属于图式量表。

### 例 4.6

调查被调查者对 xx 银行的服务态度做出评价时，可采用下列量表形式。无论采用哪个量表，被调查者都要在数轴上他认为合适的位置标注记号



上述三个量表都可以用来测量被调查者对 xx 银行服务态度的评价。这三个量表本质上讲都属于打分制量表，但这种打分是在标有刻度的数轴上进行，有利于被调查者进行分辨。第一个量表较后面两个量表的刻度更加细密，测量的精度更高。第三个量表采用正负对称的数字标示刻度，便于被调查者分辨有利与不利态度，能更好地避免出现回答错误。此类量表属于定距量表，所得测量结果为态度分数。但是这种分数对态度特征的描述而言并没有太清晰的结论，在数据汇总整理中，可以将这种定距量表转化为定序量表，研究者需要对不同的分值水平定义出不同的态度等级。对于第一个量表，数轴划分为九个间隔，每一个间隔可以定义出一个态度等级，例如分值为 1~2 分时可定义为态度极差，2~3 分时态度很差等等；对于第二和第三个量表刻度为奇数，间隔为偶数，根据对称原理，每一个数字可以定义为一个态度等级。例如在第二个量表上，可以定义 1、2、3、4、5、6、7 分别表示态度非常差、比较差、稍微差、无所谓、稍偏好、比较好、非常好。如果某个被调查者的打分或样本均分为 5.7 分，则按四舍五入的原则计为 6 分，对应的态度等级为“比较好”。当然，在设计量表时，也可以事先将态度等级名称直接标示在量表中，此时量表会显得比较复杂，对回答者的回答也没有多大的帮助，所以本教材并不提倡使用这种方式。

这种图式量表在测量态度时比顺序量表更加准确。这是因为在顺序量表中提供的是态度等级名称，而被调查者很可能难以在相邻的两个名称之间进行分辨，尤其是当态度等级划分过于细致时更是如此。运用图式量表就不要求被调查者分辨态度等级，被调查者只需要根据其心理感知，在量表上自认合适的位置做标记即可，所得测量结果可能更能准确反映其主观态度。不过，使用这种量表将增加数据处理的工作量。

## 3. 表式量表

表式量表是指测量工具以表格的形式表现出来，由被调查者在表格上作答。如果要对多个属性或特征使用相同的量表进行测量，或者测量时用表格的形式更易于被调查者理解与回答时，可以考虑采用表式量表。

### 例 4.7

在调查消费者对某商品的一些属性与特征的陈述所持有的态度时，使用了表式量表，见表4-2。

表4-2 某商品特征测量的表式量表

商品特征描述	完全同意	基本同意	无所谓	不太同意	很不同意
价格合理					
功能齐全					
款式新颖					
服务周到					

上述量表实际上是四个定序量表的组合，采用表格形式不仅有利于节省问卷篇幅，还有利于提高被调查者回答效率。

### 第三节

## 市场调查中的常用量表

在市场调查活动中，测量被测对象的属性或特征时，根据调查研究的目的与研究对象特征的不同，需要采用不同的量表。就测量的内容来看，测量活动主要包括主观性问题的测量和客观性问题的测量。其中，主观性问题诸如态度、愿望、偏好、意见等的测量是比较困难的，心理学家、社会学家、统计学家在此方面进行了许多探索与研究，形成了一系列比较成熟的量表。

## 一、列举评比量表

列举评比量表是市场调查中应用十分普遍的一种分等式量表，研究者在问卷中事先拟定并列举有关问题的答案，由被调查者自由选择回答。量表的两端为极端性答案，中间划分为若干个阶段。阶段的划分可多可少，如果要求测量的精度高，则划分的阶段可以多些。反之，可少些。一般来讲，以5~9个阶段为宜，但最常用的是5个阶段。阶段划分过多，会增加被调查者的分辨难度，影响其回答情绪。列举评比量表可采用文字式列举和表式列举两种方式。

### 例 4.8

对“您对XX牌手表的喜爱程度”这个问题的测量，评比量表可以采用下列文字式列举：

列举一：

- ①喜欢 ②无所谓 ③不喜欢

列举二：

- ①非常喜欢②比较喜欢③无所谓④不太喜欢⑤很不喜欢

上述两种列举中，前一个列举所划分的阶段较少，便于被调查者分辨与回答，但测量结果不太精确。后一个列举所划分的阶段多一些，会相应增加被调查者的分辨难度，但测量结果就要精确一些。

### 例 4.9

如果要对XX手表的一系列属性与特征进行测量，且测量使用的量表是相同的，就可以采用如表4-3所示的表式列举。

表 4-3 XX 手表特征的列举评比量表

属性与特征	极好	非常好	好	一般	差
价格合理性					
品牌知名度					
走时准确性					
款式美观性					

在这个列举评比量表中，一共有五种评价态度，有三个有利态度，只有一个不利态度，采用了非平衡量表。这应该是研究者事先获知了消费者评价中的有利态度居多的信息。如果没有特别的根据和特别的目的，首先要考虑采用平衡量表。

列举评比量表属于定序量表的计量层次，主要用于测量主观性问题，如态度、意愿、偏好、情感的强烈程度。

但列举评比量表容易产生三种误差即仁慈误差、中间倾向误差和晕轮效应。仁慈误差是指有些人对客体评价时，倾向于给予较高的评价，这就产生所谓的仁慈误差；反之，则产生负向的仁慈误差。中间倾向误差是指有些人不愿意给予被评价的客体很高或很低的评价，特别是不了解或难于用适当的方式表达出来时，往往倾向于中间性的评价。晕轮效应是指如果被调查者对被评价的对象有一种整体印象，而导致对该对象的所有属性的评价呈正面或负面倾向，该效应可能会导致系统偏差。

仁慈误差来源于被调查者的性格与心理，研究者一般无法控制。中间倾向误差可以采用一定办法加以控制，如调整叙述性形容词的强度；增加中间评价性语句在整个量表中的比例；使靠近量表两端的各级在语意上的差别加大，使其大于中间各级间的语意差别；增加测量量表的层次。晕轮效应也可以采用一定的方法加以预防，如对所有被评价的对象，每次只评价一个变量或特性；问卷每一页只列一种特性。如果通过实验性调查，晕轮效应的确存在，则上述表式列举评比量表在正式调查中的应用就要慎重。

## 二、等级顺序量表

等级顺序量表是一种定序量表。它是将多个测量项目同时展示给被调查者，并要求他们根据某个标准对这些测试项目进行排序或分成等级。在调查消费者的偏好、意愿以及评价态度时，可采用这一量表。

前面介绍的列举评比量表是一种非比较性的量表，因为被调查者是在没有任何参照物或参照标准的情况下进行判断和选择。等级顺序量表则是比较性的量表，此量表要求被调查者将一个事物或因素与另一个或多个事物或因素进行比较，通过比较来促使被调查者用一种现实的态度来做出评价。

### 例 4.10

您在购买手机时最看重的因素是（选三项并按重要程度由高到低排序，将所选因素的代码填入题干后的方格中）□□□

- ①品牌 ②功能 ③颜色 ④价格 ⑤售后 ⑥其他（请注明）

### 例 4.11

下面列有四种品牌的洗衣液名称，请按您对品牌的喜爱程度分别给予不同的分数（顺序由 1~4, 1 表示最喜欢，4 表示最不喜欢，数字填入品牌名称后的括号中）：

- |       |        |
|-------|--------|
| ①立白【】 | ②蓝月亮【】 |
| ③白猫【】 | ④奥妙【】  |

如果要对这四种不同品牌的洗衣液分品牌要素进行评价，可以采用下列形式的等级顺序量表。

### 例 4.12

请比较下列四种不同品牌洗衣液的特征，并按特征对四种品牌洗衣液进行排序（顺序由 1 至 4, 1 表示最显著，4 表示最不显著，数字填入横线上）

去污强 泡沫少 易漂洗

- |       |     |
|-------|-----|
| ① 立白  | 去污强 |
| ② 蓝月亮 | 泡沫少 |
| ③ 白猫  | 易漂洗 |
| ④ 奥妙  |     |
- 

例 4.12 实际上是三个等级顺序量表的组合，因为四个品牌在三个需要比较的要素上的测量规则是一样的，为了节省问卷篇幅，采用这种形式最合适。

等级顺序量表的优点主要表现在题目容易设计、易于使用；被调查者比较容易掌握回答方法；被评价的事物被排成一定的顺序，彼此成为参照物，被调查者的评价可能更加客观与现实。

等级顺序量表的缺点主要表现在如果被选项中没有包含被调查者的选择项，那么结果就

可能会不准确；要测量的某些因素可能完全超出了个人的选择范围，产生毫无意义的数据；量表仅给研究者提供了顺序信息，人们完全不了解被评价的各客体间有多大的差距，即搞不清为什么被评价的客体按此顺序排列；问卷上列举对象的顺序也可能带来所谓顺序误差；用于排序的对象个数不能太多，否则很容易出现错误、遗漏。

### 三、Q 分类量表

Q 分类量表是根据事先规定的评价类别，将一组同性质的客体进行分类的量表。这一组客体一般是描述被调查者对诸如品牌要素、商品特点的语句。可以说 Q 分类量表是等级顺序量表的一种复杂形式，等级顺序量表中用来进行比较的客体往往是十分有限的，一般不会超过 10 个；但是在实际研究中，用来比较的客体可能非常多，比如有几十个甚至上百个。如果将一组数量庞大的客体由被调查者通过比较而进行排序、分类，几乎是不能完成的。而运用 Q 分类量表则可以实现这一测量目标。

Q 分类量表的实质是按照对称分布（如正态分布）的要求，对一组描述某种价值判断的语句进行分类，然后进行统计分析，例如，为了研究消费者对某种商品的各特性的偏好程度，可以拟定一组描述商品特性的语句，要求被调查者对该组语句按有利态度与不利态度分成不同的等级或类别。待分类完毕，赋予各类卡片或语句一定的数值（评定值），最后进行相关分析。运用 Q 分类表进行测量的步骤如下：

第一步，拟定 60 句～120 句有价值判断的语句。语句需要根据研究的目的来拟定，应该包含有利态度和不利态度，这些语句可以按从有利到不利的方向排列顺序。

第二步，制定一个奇数等级的态度量表，用于测量被调查者对某个语句的态度。以 5 分制为例，态度测量量表可以采用图示量表：

5            4            3            2            1

量表中的数字 5、4、3、2、1 分别代表非常同意、比较同意、无所谓、不太同意、很不同意。第三步，制定语句的分布规则，要求被调查者对全部语句按规则进行分配。语句的分布要求在 5 个计分点上呈对称分布，假如共拟定了 60 条语句，其在 5 个计分点上的分布规则为：

5	4	3	2	1
	1			(
8	12	20	12	8

数轴下面的数字是 60 条语句在各个计分点上的分布频数，频数分布是对称的。每个被调查者在调查员的指导下，将 60 条语句按照上述分布规则，完成语句在 5 个计分点上的分布。

第四步，汇总测量结果，进行相关分析。假如某次调查拟定了关于 A 牌商品特性及消费者态度的 60 条语句，抽取了 80 名被调查者，按上述规则测得如下结果（见表 4-4）：

表 4-4 Q 分类量表调查结果表

语句	被调查者序号						
	1	2	3	4	5		80
1. A 牌品质好	5	2	3	1	5		4
2. 我总是购买 A 牌	4	1	3	1	4		4

语句	被调查者序号						
	1	2	3	4	5		80
3.我喜欢使用 A 牌	4	1	3	2	3		3
...		...				...	
58. A 牌品质较差	1	4	3	4	4		2
59.我不会购买 A 牌	1	5	3	4	4		2
60.我不会使用 A 牌	1	5	3	4	3		3

根据表 4-4 的测量结果, 可以分析消费者在 A 牌上的消费偏好, 消费者群体在 A 牌商品上具有哪些不同的潜在特征, 消费者购买或使用 A 牌商品的行为与 A 牌商品的一些特性之间的关系等。这些都是进行市场细分、研究消费者行为与动机的基础信息。

Q 分类量表的优点主要表现在: 具有较强的逻辑性和实用性, 适用于比较项目众多的场合; 调查数据可用于检测自变量对于复杂因变量的效应; 适宜于进行探索性的研究, 有利于产生新的研究思路和假设; 调查数据还可以用方差分析、相关分析和因素分析等多种方法进行分析。

Q 分类量表的缺点主要表现在: 难以确保样本的代表性; 语句项目不易满足某些统计处理方法的假设, 比如语句项目反映的独立性、项目数据的等距性; 采用了强迫选择和分类, 使得受测者的自由反应受到限制。

#### 四、配对比较量表

配对比较量表是将一组需要比较的项目, 通过两两比较的方式来测量人们态度或偏好的一种量表。它是排列式量表的一种具体应用形式。

例 4.5 中, 提供了四种品牌的洗发液, 成对排列出来, 由被调查者在每一对中选择一个喜欢的品牌, 该量表就是配对比较量表一个简单配比形式。实际上四种不同的品牌所形成的两两配比组合有六种情形, 可以按如下组合进行排列:

- |       |      |
|-------|------|
| ①□海飞丝 | □潘婷  |
| ②□海飞丝 | □飘柔  |
| ③□海飞丝 | □威娜宝 |
| ④□潘婷  | □飘柔  |
| ⑤□潘婷  | □威娜宝 |
| ⑥□飘柔  | □威娜宝 |

针对上述排列, 要求被调查者在每一组中选择一个喜欢的品牌。从严格意义上讲, 配对比较量表要求一组比较项目中的任意两个项目都有一次比较的机会。否则, 不便于进行统计分析。为了尽量避免顺序误差的产生, 上述排列顺序可以打乱。配对比较量表的描述性数据分析可以按下列方式进行。

假如抽取了一个由 100 名消费者组成的样本, 整理的调查结果见表 4-5。

表 4-5 品牌 i 较品牌 j 为佳的人数分布

J	i			
	海飞丝	潘婷	飘柔	威娜宝
海飞丝	—	80	70	40
潘婷	20	—	30	15
飘柔	30	70	—	35
威娜宝	60	85	65	—

为便于分析，可以将此频数分布转化为频率分布。定义各品牌与自己比较的频率为 0.5，按列向计算频率，将每栏的频率相加，就可得出各品牌的态度值(见表 4-6)。

表 4-6 品牌 i 较品牌 j 为佳的频率分布

j	i			
	海飞丝	潘婷	飘柔	威娜宝
海飞丝	0.50	0.80	0.70	0.40
潘婷	0.20	0.50	0.30	0.15
飘柔	0.30	0.70	0.50	0.35
威娜宝	0.60	0.85	0.65	0.50
合计	1.60	2.85	2.15	1.40

四个品牌的态度值相比结果为：潘婷(2.85) > 飘柔(2.15) > 海飞丝(1.60) > 威娜宝(1.40)，说明在四个品牌中，潘婷最受被调查者喜欢。

从计量层次上看，配对比较量表属于一种定序量表，但可以克服一般的定序量表所存在的问题。首先，对被调查者来说，从一组两个成对的比较项目中选择一个要比从多个比较项目中选一个更容易。其次，当成对的比较项目不按项目名称顺序排列，就能克服顺序误差。但是，该量表也存在不足。首先，既然它属于定序量表的范畴，那么上述计算出来的态度值仅能作为排序的符号，而不能用来度量态度之间的绝对差距，更不能进行相对差距的比较。如上例中，我们不能说被调查者对潘婷的喜爱程度是威娜宝的两倍(2.85/1.4)。其次，当需要比较的项目以算术级数增加时，配对数量就会以几何级数增加，配对比较数量过多会让被调查者厌烦。所以，配对比较量表适合于比较项目不多，且被调查者对每一个项目都熟悉或了解的情形。

## 五、固定数目量表

固定数目量表是要求被调查者对一组需要比较的项目，按某种特性的重要程度，将一个固定的总分分配给每一个项目的量表。该量表要求，某一问题下列出需要比较的项目，被调查者根据自己的认识程度给每个项目评分，不管分数如何分配，各个项目分数的总和必须是一个固定值。固定数目量表可以避免次数频繁的配对比较，常用于同类但不同品牌商品之间的比较，以调查消费者对品牌的偏好，也可用于对同一客体的不同因素项目之间重要程度的比较。

### 例 4.13

运用固定数目量表对多种不同品牌的同类商品进行某方面属性或特征的比较，

“下列有甲、乙、丙三个不同品牌的矿泉水，根据您个人的喜好程度，给每个品牌一定分数，其总分必须为 100 分”：

①

甲品牌【】

②乙品牌【】

③丙品牌【】

#### 例 4.14

运用固定数目量表对多种不同品牌的同类商品两两进行某方面属性或特征的比较

“下列有三个不同品牌的矿泉水，分为三个配对比较组，根据您个人的喜好程度，对每一组中的每个品牌各给予一定分数，每组总分均必须为 100 分”：

- ① 甲品牌【】----- 乙品牌【】
- ② 乙品牌【】----- 丙品牌【】
- ③ 甲品牌【】----- 丙品牌【】

#### 例 4.15

运用固定数目量表对同类商品的不同要素进行比较

“您在购买皮鞋时，对下列要素的重视程度如何？请分别打分，总分为 100 分”：

- ① 品牌【】 ② 款式【】 ③ 颜色【】 ④ 价格【】

固定数目量表属于定序量表，所打出的分数只是表示被调查者对某个项目给予的一种等级，也反映了被调查者对待某一项口的态度，分数本身除了用来排序外，不能做其他任何统计运算。但固定数目量表优于一般的定序量表，这是因为采用了打分的方法，对被调查者态度的刻画比较具体，且当两个项目在某种态度上被认为具有相同的价值时，可以被如实地表示出来。固定数目量表的主要缺点是，当比较的项目过多时，可能使被调查者感到难以回答，也就是说为了保持总分数为给定的固定数目，被调查者要在多个项目上小心地分配分数并核对总分。一般认为，使用固定数目量表测量的项目最多不超过 10 个，

## 六、语义差别量表

语义差别量表是用成对反义形容词作为量表的两端，中间划分为若干阶段或等级，来测量人们的态度和个体、群体之间差异的一种测量工具。该量表要求被调查者根据自己的感觉在某一阶段或等级上做上记号，根据其所做记号的位置，来判断被调查者对某一或某一组测试项目的程度。实际应用时，语义差别量表往往由一组子量表所组成，以测量被调查者对一组概念所持有的态度。

语义差别量表由美国心理学家、伊利诺大学 C. E. 奥斯古德于 1957 年提出的一种心理学研究方法，又称 SD 法。奥斯古德等人认为，不论文化和言语的差别有多大，人类对概念或词汇具有颇为广泛的共同的感情意义。因此对“智力高的和言语流利的研究对象”，直接询问一个概念的含义是有效的。

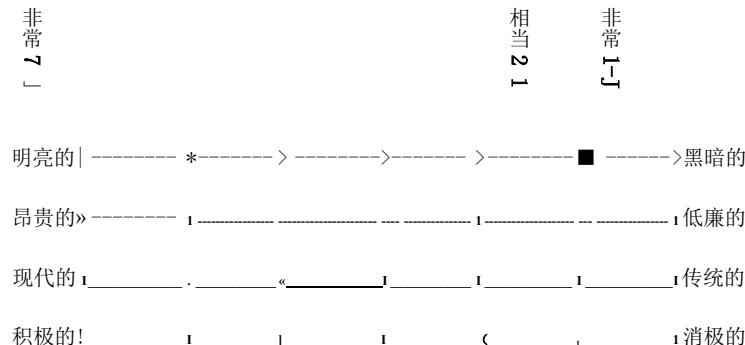
语义差别量表由概念和若干量表构成。这里的“概念”，既包括词、句、段和文章那样的语言符号，也包括像图形、色彩、声音等有感情意义的知觉符号 G

构造语义差别量表，首先要确定测量的概念，如公司形象、品牌形象、商品印象等；其次要挑选一些能够形容这一概念的一组对立的形容词或短语。语义差别量表所包括的形容词或短语所涉及的属性因素可以分为三个类型，即“评价因素（好与坏）”“潜能因素（强与弱）”“活动因素（快与慢）”。评价因素一般是测量概念的性质特征，如商品的品质、等级等；潜能因素一般是测量概念的规模特征，如商品的尺寸、体积等；活动因素一般是测量概念的活动特征，如商品广告、销售网络等。这三个方面的因素构成了一个“语义空间运用这个语义空间，基本

上能够测量出人们对某个概念的印象、感受等。

语义差别量表一般采用数轴的格式加以表现。数轴两端分别为两个反义形容词，中间划分为若干阶段。

例如，假设要调查人们对某商品的印象，那么需要测量的概念就是“商品印象”。商品印象这个概念可以从评价因素的角度，选择形容商品品质特征的形容词“高级的”与“低级的”、“明亮的”与“黑暗的”两组反义词；从潜能因素的角度，选择“昂贵的”与“低廉的”、“传统的”与“现代的”两组反义词；从活动因素的角度，选择“积极的”与“消极的”一组反义词，共同组成一个语义空间。然后按照语义差别量表的格式构造出如下语义差别量表：



上述语义差别量表由 5 个子量表所组成，要求被调查者在每一个子量表上使用标记做出回答，按标记所处的位置，转换出态度分值，并计算每一个子量表上的态度均值，再构造出语义差别量表的“轮廓图”或“形象图”。通过“轮廓图”可以判断被调查者对某种商品的认识与评价态度。

运用语义差别量表还可以比较被调查者在不同项目（如商品）上的认知、态度、偏好等。假如有甲乙两种商品，通过语义差别量表分别测得被调查者在上述五个子量表上的态度均值，将这些态度均值标示在量表上，再将各均值点连成线，就会形成比较轮廓图（见图 4-1，实线表示“甲”、虚线表示“乙”）。

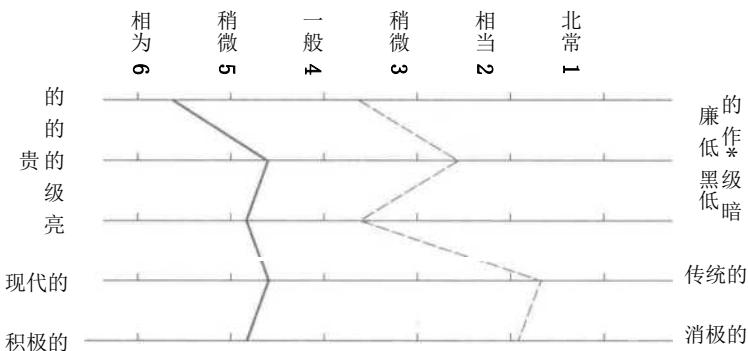


图 4-1 甲乙两种商品语义差别比较轮廓图

从图 4-1 可以看出，被调查者在两种商品上的 5 个语义上存在明显差别，对甲商品均有正面认知态度，而对乙商品多持有负面认知态度。通过这种差异，可以判断人们在此类商品上的属性需求与购买动机。

除了轮廓比较图，还可以使用“距离”来比较人们在不同子量表上对所测概念的印象与态度的相似与差异程度。以商品比较为例，“距离”的计算方法有多种，一般采用欧氏距离即可，计算公式为：

$$d_{ij} = \sqrt{\sum_{k=1}^n (x_{ik} - x_{jk})^2} \quad (4.1)$$

式中， $d_{ij}$  表示态度距离， $i$ 、 $j$  分别表示被调查者对  $i$  商品与  $j$  商品在第  $k$  个量表上的态度均值/表示子量表的个数。距离越小，表明被调查者在两种不同商品上的认知态度越相似。

在市场调查中，语义差别量表可以快速、高效地发现商品、公司形象与竞争对手相比所存在的长处与短处。在营销学与行为科学的研究中发现，语义差别量表在制定决策和进行预测方面具有足够的可靠性与有效

性。当用于公司形象研究时，语义差别量表在统计上具有普遍适用性。语义差别量表适用于测量和比较具有不同背景的被调查者对商品印象、公司形象的看法。

语义差别量表的应用也会存在一些问题。第一，该量表缺乏标准化。具体体现在对一个概念进行测量时，语义空间的构造具有一定的主观性，也具有一定的难度。第二，语义差别量表中评分点数目的确定也是一个难题。评分点数目太少，整个量表过于粗糙，测量结果往往缺乏现实意义；评分点过多，又会超出大多数人的分辨能力。第三，语义差别量表可能会存在“晕轮效应”。对一个特定目标的各个组成要素进行评分，可能会受到被调查者对测试概念总体印象的制约，特别是被调查者对各要素不太清楚时，可能会产生明显的偏差。

## 七、中心量表

中心量表是语义差别量表的一种变通形式，它是将一个形容词置于量表的中间位置，其上下或左右均分为5个等级，由-5至+5共10级，用以测量态度的方向与程度。运用该量表进行测量时，只需要被调查者在量表的数字上画“V”即可。

例如，研究者欲对某银行的服务质量进行评价，采用中心量表对客户的态度进行测量，量表的格式如表4-7所示。

中心量表的主要优点在于，它能使研究者免去设计成组反义形容词的麻烦工作，而且这种量表被认为在态度测量时具有较高的分辨度。其主要缺点是，量表中所使用的描述性形容词可能是中性的，也可能偏向有利态度方向或不利态度方向显然，形容词的选择将会影响测量的结果。中心量表较语义差别量表在实际应用中使用的机会要小些。

表4-7 中心量表示例

+ 5		+5		+5
+ 4		+4		+4
+ 3		+ 3		+3
+2		+2		+2
+ 1		+ 1		+ 1
服务态度友好		等待时间短		理财产品多
-1		-1		-1
-2		-2		-2
-3		-3		-3
-4		-4		-4
-5		-5		-5

## 八、瑟斯顿量表

前面介绍的各种量表，用于测量的问句都是由研究者根据研究的目的与自己的认识来设计的，测量的答案也是事先拟定好的，被调查者只是被动地选择。但是，有时被调查者对问句及其答案的理解与研究者会存在差异，或者问句并不是被调查者感兴趣的，这势必会影响回答质量而产生回答误差。为避免此类问题的发生，就可以考虑由被调查者来决定用于测量的问句，设定回答的规则，评定被调查者的态度。显然，采用这种方式进行测量，并不是直接测量出被调查者的态度，而是通过被调查者选择的问句及其给出的答案来评定被调查者的态度，所以属于一种间接测量技术。在间接测量技术中，瑟斯顿量表、李克特量表与

哥提曼量表都是常用的测量工具。本章重点介绍前两个量表。

瑟斯顿量表是要求被调查者对一系列表明态度的陈述语句加以评判进而测量其态度的工具。具体而言，它是指在调查中设计一系列陈述语句，由被调查者选出其所同意的陈述，再通过统计处理，筛选出具有代表性的陈述语句所表明的态度，进而来测量被调查者的态度。

瑟斯顿量表的设计与测试过程是比较复杂的：其设计过程是通过评判者测试来选择合适的陈述句，再交由被调查者使用n具体步骤如下：

#### 第一步，拟定态度陈述语句。

语句是研究者根据调查的问题拟定的被调查者对该问题一切可能的态度陈述。由于研究者事先并不知道被调查者究竟对所调查问题的哪些属性或特征感兴趣、究竟对所调查的问题存在哪些看法，所以，拟定关于调查问题的有关态度陈述语句要尽可能多，以使被调查者对语句有足够的选择空间。所拟定的陈述语句必须包含有利态度、不利态度、中立态度的描述语句，但这些态度描述性语句以简短为宜。例如•要了解人们对网上购物的看法，研究者可以事先拟定关于网上购物的下列陈述语句（假如有50个语句），见表4-8。

表瑟斯顿量表测试语句

语句编号	语句内容
Q1	网购是人们越来越认可的购物方式
Q2	网购节省购物时间
Q3	网购能够提高购物效率
Q4	网购不受时空限制
.....	...
Q49	网购容易成瘾
Q50	网购商品的质量不可靠

#### 第二步，测定语句。

选择一组评判者，人数一般在20~50名。评判者要具有代表性，对所调查的问题比较熟悉”语句测定之前，研究者还要制定测试规则可以规定语句按照由不利态度到有利态度分为若干组，组数最好为奇数，一般以5~9组为宜，不同组别反映评判者对每一个语句所持有的肯定或否定的态度。若分为7组，不利态度列入1~3组，最不利的列第1组，中立态度的语句列入第4组，有利态度列入5~7组，最有利的列在第7组 在上述规则下，由评判者根据个人

的认知、态度，将各条语句列入他认为合适的组之中。由于评判者在认知、态度上存在差别，同一条语句不同的人可能会将其列入不同的组之中。

### 第三步，统计处理。

所有语句由全部评判者按态度差异进行了分组，每一个组中将会出现一定数量的语句，形成每条语句在各组之间的频数分布。假如针对表 4-8 中的语句，选择了 20 名评判者，按 7 组 进行区分，50 条语句在 7 个组之间的频数分布见表 4-9。

表 4-9 语句态度频数分布

语句	组别							合计
	1	2	3	4	5	6	7	
Q1	6	4	3	2	3	2	0	20
Q2	5	5	5	2	1	1	1	20
Q3	7	5	4	3	1	1	0	20
Q4	4	5	5	3	2	1	0	20
		...				...	...	
Q49	1	2	4	5	3	3	2	20
Q50	0	2	2	3	4	4	5	20

表 4-9 中，中间交叉部分为各句归入相应组别的频数。根据此频数分布数据，计算各语句组别序的均值和标准差：，均值采用加权平均法进行计算，变量为组别序号，权数为各组别所对应的人数各语句组别序的均值按四舍五入的规则取整数。这样按组别序均值，每条语句就归入到某一个组之中语句组别序的标准差用于衡量全体评判者在某条语句上所持有的态度上的差异，这种差异越大，说明评判者对某条语句归入某组的意见分歧越大，那么该语句的分组结果就越没有代表性。另外，这种态度差异也许是语句表意不清或表意不准所导致这样的语句就应该从问卷中剔除。

### 第四步，筛选语句。

经过第三步，可以获得每一个组所归入的语句及其组序标准差。筛选语句的基本规则是，从每一组中选出标准差最小的 2 条语句，作为实际调查中使用的语句——第五步，评价态度。

实际调查中，只需要被调查者对每一条筛选出来的语句表达是否同意的态度调查结果既可以评价被调查者在某条语句上的态度，也可以综合评价被调查者在全部语句上所表现出的对所调查问题所持有的态度，

瑟斯顿量表设计过程复杂，测量过程也比较烦琐，语句评判人员的态度可能与实际调查中被调查者的态度存在差别，且所选择的语句评判人员不同，筛选出来的语句也会有别：、利用语句所测量的态度仅仅只是“同意”或“不同意”，其同意与不同意的程度无法体现。所以，该量表在实际调查活动中使用并不普遍

## 九、李克特量表

李克特量表是由一组能够表达对所研究的概念持有的态度陈述及态度方向与程度的语句 所组成的测量工具，用于测量和评定人们对某事物或现象的态度，是问卷设计中运用非常广泛

的量表。

李克特量表与瑟斯顿量表都是对语句所陈述的概念特征进行态度测量的工具，但两者存在一些区别。瑟斯顿量表中用于实际测量的语句需要经过预先测试、评判、筛选，而李克特量表中的语句既可以按一定方式筛选出来，也可以由研究者主观确定下来；瑟斯顿量表在测量人们对语句所陈述的概念特征时，只需要被调查者回答“同意”还是“不同意”，而李克特量表则要求被调查者回答同意或不同意的程度。因此，在态度测量的角度上，李克特量表要优于瑟斯顿量表，其应用也比瑟斯顿量表更普遍：

严格意义上讲，李克特量表中用于测量的语句，最好要经过预先测试「但这个测试的过程与瑟斯顿量表一样非常复杂。首先，研究者要搜集大量的有关公众对所测概念看法的陈述；其次，将搜集到的这些陈述划分为“正面的”和“反面的”，并询问被测试者对陈述的态度，态度一般划分为5个等级，即完全同意、基本同意、无所谓（不确定）、不太同意、很不同意，并给每一个态度等级分配一个数字，如上述态度等级按顺序分别分配5、4、3、2、1这5个数字。再次，要进行评分处理。评分过程中，态度的方向（从正面到反面的）应与所测内容保持一致。例如，如果给正面项目中的“完全同意”为5分，则反面项目中的“很不同意”也给5分。最后，选择语句。选择语句的依据是测试分数。先计算每个被测试者的总得分，研究人员只选出那些在高的总分与低的总分之间比较好地表现出差别的项目，并比较每个项目的细微差别，最终选出20-25个在平均得分上有最大区别的项目，这些项目都是陈述句，用于测量被调查者对这些描述概念特征陈述的态度。

假设我们要了解人们对某种保健鞋垫的印象，根据选出的20条语句，设计出表4-10所示的李克特量表

表4-10 某保健鞋垫印象测量李克特量表

语句	完全同意	基本同意	无所谓	不太同意	很不同意
1.我很满意产品质量	5	4	3	2	1
2.使用时脚底发热	5	4	3	2	1
3.使用时鞋子变得更紧	5	4	3	2	1
4.可能会继续使用	5	4	3	2	1
18.价格有点贵	5	4	3	2	1
19.我不打算推荐给他人	5	4	3	2	1
20.不会持续使用太久	5	4	3	2	1

实际调查中只需要被调查者在量表中认为适当的格子中画“V”即可。在市场调查实务中，李克特量表非常流行。其设计快捷、简便、易于操作，而且可以通过电话来进行调查。但是，对测量用的语句很少按前面介绍的步骤来进行挑选，往往是由研究者在初步探索的基础上，经过精心设计并由小组访谈后确定下来’

## 十、设计量表时应考虑的因素

设计量表是一项复杂且十分重要的工作，量表设计得是否合理、科学，直接关系到测量结果的质量与研究目的的实现程度。设计量表需要考虑的因素较多，既涉及研究目的、研究对象的特点，也涉及调查方法，还涉及被调查者群体的特征以及不同量表的功能与特点。

### （一）根据研究目的选择量表

研究目的决定研究内容与方法，研究内容与方法决定了所需要搜集的数据的类别，当然也决定了测量中需要使用的量表。例如，如果根据研究的目的需要获取定性数据，就应该选择定类与定序量表，如果需获取定量数据，则应该选择定距与定比量表；如果需要获取被调查者主观层面的数据，定类、定序、定距量表都是可以选择的，前面介绍的一系列量表都可以作为选择对象；如果需要获取被调查者客观层面的事实性数据，大多数情况下首先要考虑使用定比量表。当然，有些事实性数据也只能采用定类量表，如调查性别时就必须采用这种量表。

### （二）根据研究对象的特点选择量表

针对不同的调查对象应该选择不同的量表。一次市场调查活动中，研究的对象可能既有无意识的现象也有有意识的现象。对于无意识的现象，只能采用物理性工具来测量其物理属性，例如对商品的尺寸、重量的测量，就需要采用专用的度量衡工具；对于有意识的现象，既可以采用物理性测量工具，也可以采用非物理性量表。例如，要测量人们对广告的情绪反应，就可以采用诸如视向测定仪进行测量。又如，要了解人们的购买动机，就需要专门设计测量购买动机的量表，具体的量表形式有多种，如列举评比量表、等级顺序量表、配对比较量表都是可以选择的。

### （三）根据调查方法选择量表

依据研究的目的不同，市场调查所采用的调查方法也会不同。面访式调查可以设计复杂一些的量表。例如，在测量被调查者的态度时，一般的量表均可以采用，而且量表的刻度也可以更加细密，量表的表现形式也可以更加灵活；而对于非面访式的各种调查方法，由于被调查者在回答问题时缺少与调查员的沟通机会，设计的量表就不能太复杂，量表上的刻度也不能太细密，以免给被调查者的回答造成障碍。例如，在电话调查中，就不适合采用瑟斯顿量表，对于需要进行测量语句筛选的李克特量表也不能采用。

### （四）根据被调查者群体特征选择量表

这里所指的被调查者群体特征主要是指被调查者的自然属性特征与人文属性特征。就自然属性中的年龄特征来讲，被调查者可以划分为不同年龄段的群体，对特殊年龄段的群体进行调查，量表的设计就可能要采用比较特殊的形式。例如，要调查小学生对网络小游戏的兴趣时，由于这个群体的认知能力十分有限，在量表的形式、量表的表述、量表的用词方面就应该考虑这个特定群体的特点，尽可能做到量表简单易答、生动有趣。比如，量表中可以配卡通图片，这些图片可以与测量对网络小游戏的兴趣程度用词匹配起来，帮助其认知、提高其回答兴趣。但描述兴趣程度的等级不能太细，否则，此年龄段的被调查者难以分辨。

### （五）选择量表时要考虑的其他因素

市场上可使用的量表种类较多，各有其功能与特点，设计量表时还需要考虑以下因素：

从态度答案数目上看，量表有平衡量表与非平衡量表之分。如果研究者想得到广泛的意见，则采用平衡量表比较好。如果以往的调查或初步探索结果已表明，大多数的意见都是肯定的，则设计的量表中就应该有更多的肯定倾向。

从量级层次的个数上看，量表有低层级量表与高层级量表之分。量级层次太少，则量表会显得粗糙和不够全面；显然，一个3层量表无法反映出一个11层量表所能提供的态度的强度。但是，量级层次过多，可能会超出人们的分辨能力。有研究表明，评比量表基本上以5~9层为宜。如果采用电话调查，量级层次个数最多为5层。一般来讲，5个量级层次的量表最为

## 第四章 量表测量技术

参  
普遍。

■ r 八 A \_\_\_\_\_

从量级层次的奇偶性来看，有奇数层级量表与偶数层级量表之分。一般情况下，应使用奇数层级量表。如果使用偶数层级量表，就意味着没有中间答案，量表中没有中间答案，被调查者就会被迫选择一个正向或负向答案，但那些确实持有中立意见的人就无法表明他们的观点。而奇数层级量表给被调查者设立了一个中间答案，当被调查者确实没有明确的态度倾向时，就不必纠结于态度倾向及其程度的选择，而是直接

选择中间答案。

从回答是否具有强制性来看，有强迫性量表与非强迫性量表之分。一般情况下应该采用 非强迫性量表。强迫性量表要求被调查者必须从答案选项中选择一个答案，但是可能被调查 者对问题并不熟悉，或者难以表明态度，本来无法作选又不得不选，就会迫使其任意选择一个 并不能反映其认知与态度的答案，测量结果就不准确在设计量表时，研究者必须对被调查者 群体进行初步分析与判断，如果答案选项对于被调查者来说是完备的，采用非平衡量表可能不 会存在什么问题；反之，就需要在答案选项中列入“其他”或“不知道”选项，这样就不会强迫对 所调查事项一无所知的被调查者也给出某种明确的意见或态度。

## 第四节

### 量表质量的评价

问卷中所设计的量表是用于测量事物或现象属性与特征的工具，自然要求这些测量工具 具有一定的质量 量表的质量属性可分为外在质量与内在质量两个方面，，量表的外在质量主 要从量表是否便于阅读、理解、回答几个方面来评价，而量表的内在质量主要是从量表的信度 与效度两个方面进行评价。问卷中设计的量表首先要进行质量测试与评价，只有符合质量要 求的量表才能够用于实际调查

#### 一、量表质量评价的方式

##### （一） 小组座谈

小组座谈是指研究者组织相关研究人员对所设计出的量表进行集中审核与评估的活动。小组座谈成员应该熟悉被研究的对象与研究的目的，对各种量表的功能与特点均有了解。通 过小组座谈的工作方式，每个小组成员对量表在计量层次选择、类型选择方面的适用性与合理 性进行评估；对量表的表述内容与形式、回答指引等进行检查，评价量表是否便于被调查者阅 读、理解与回答，并指出量表存在的问题，提出修改建议。小组座谈适合于对量表的外在质量 进行评价。

##### （二） 预调查

预调查就是从目标群体中抽取一个小样本，运用经过小组座谈修改后的量表所进行的一 次试验性调查。这时进行预调查的目的有两个方面，一方面是为『评价量表的内在质量而搜 集调查数据，通过对调查数据的分析来评价量表的信度与效度；另一方面是通过预调查，了解 被调查者对量表的适应性，发现量表存在的问题，为量表的进一步修改与完善搜集意见、看法 与回答情况方面的资料。预调查主要用于对量表的内 在质量进行评价，

由于预调查的目的是检验问卷的质量，而不是对所研究总体有关参数进行概率推断，所以 原则上预调 查可以不需要进行概率抽样。但是，如果在评价量表的信度与效度时，要使用一些 具有概率分布的样本统 计量，那么预调查就需要采用概率抽样的方式。

预调查的必要样本量是难以用现成公式计算出来的，这是因为预调查阶段并不涉及对总 体有关参数进 行概率推断时的允许误差问题，如果按照必要样本量的公式进行计算就缺乏已 知条件。实际上，预调查的 样本量一般是按经验主观确定的。经验表明，预调查样本量一般为 问卷中题项数量的 5~10 倍即可。假如某 问卷共有 30 个题项，则预调查样本量可在 150~300 个调查单元之间选择。如果有足够的人力、时间与经费 保障，则调查单元可多一些，反之则可 以少一些。

#### 二、量表内在质量评价的指标

量表内在质量评价指标主要有信度与效度两大类指标。

量表的信度与效度需要从测量的结果进行分析，量表测量的结果即测量值是一个观测值，可能与真实值之间存在一定误差，误差中可能既包含随机误差，也包含系统误差，这些误差对量表的信度与效度产生着不同的影响。观测值（X）、真实值（T）、系统误差（W）、随机误差（E）之间的关系可表示为：

$$X = T + W + E$$

其中，随机误差（E）无法避免，但可以通过多次重复测量或增大样本量进行控制，它影响信度。随机误差越大，信度越低。系统误差（W）是一种偏差-效度受到系统误差和随机误差的综合影响。

### （一）量表的信度

量表的信度是指量表测量结果的一致性、稳定性，也称为量表的可靠性。“如果在相同条件下，运用某量表对某个概念在不同时间上重复多次进行测量，其测量的结果保持不变，就表明该量表是可信的或具有可靠性；否则量表就不可信或不具有可靠性。信度是测量过程中所存在的随机误差大小的反映。信度低，测量随机误差大，测量结果就会与真实数据发生较大偏离。”

例如，用一台电子秤对同一个物品的重量进行了3次测量，结果都是10公斤，就说明这台电子秤在测重功能上是可信的，具有相当的可靠性。又如，如果使用同一个量表测量某个被调查者对某商品的喜爱程度，如果前后调查了3次，三次调查的结果都是“非常喜欢”，则表明该量表是可信的或可靠的。

对于一个具有可靠性的量表而言，短暂性因素和环境因素不会对测量过程产生影响，它能够在不同环境与不同时刻提供稳定的测量结果。所以，量表的可靠性也可以理解为在测量中可以避免随机误差，从而提供前后一致的数据的程度—测量过程中产生的随机误差越小，测量就越可靠。如果在被测概念的属性特征保持稳定的情况下，测量结果保持不变，则量表就具有可靠性。如果被测概念的属性特征变化了，具有可靠性的量表应该能够揭示出这一变化。

量表的信度可分为内在信度和外在信度。外在信度是指对相同的测试者在不同时间测得结果的一致性程度，而内在信度是指对测量同一个概念的一组问题，同时组成量表题项的内在一致性程度，反映测试一个概念的一组问题之间的稳定性和一致性。外在效度与内在信度都可以使用不同的方法进行评价。

量表仅仅具有信度是不够的，因为量表的信度并不能告诉我们测量的结果是否真实、准确。前面曾举了使用电子秤称重的例子，3次称量物品的重量都是10公斤，电子秤具有很高的信度。但是，假如这台电子秤是一台假秤，虽然它测量的结果是稳定的，但却肯定不准确。也许这件物品的实际重量是12公斤，而测量结果却只有10公斤，就存在2公斤的误差。因此，衡量量表的内在质量还要看其是否能够准确测量被测概念的属性与特征，这就是量表的效度。

### （二）量表的效度

量表的效度是指量表准确地反映客观事物属性和特征的程度，也称有效性。市场调查中效度可以理解为调查结果准确地反映调查中所要说明问题的程度。如果一个量表既具有较高的信度，也具有较高的效度，则这个量表就具有较高的内在质量Q。

例如，前述用一台电子秤对一个物品进行称重的例子中，假设这件物品的实际重量为12公斤，而实际测量的结果也为12公斤，就说明这台电子秤在测重功能上是准确的，具有很高的有效性。又如，如果使用一个量表测量某个被调查者对某商品的喜爱程度，假设他确实对该商品持有非常喜欢的态度，而测量结果也是“非常喜欢”，则这个量表就具有较高的效度或有效性。

效度包括两层含义，一是测量的是所要研究的概念而不是其他概念，二是该概念能够被准确地测量。如果我们需要测量的概念是消费者“对某商品价格的评价态度”，而量表却是测量消费者“对某商品的喜爱程度”，显然，对“商品价格评价态度”的测量不可能是准确的。如果量表测量的概念的确是“对某商品价格的评价态度”，且评价态度又被准确地测量出来，毫无疑问这个量表就有效。

如果一个量表具有有效性，那么这个量表能够避免测量过程中所可能产生的系统性误差和随机性误差的程度就高。因此，运用具有较高效度的量表所测量的结果在被调查者之间存在的差异，只反映了所测概念在属性与特征上的真正差异，而并不反映由于系统性因素与随机性因素导致的差异。

量表的效度可以划分为不同的层次，对应采用不同的方法不同层次的效度进行评价。

以上介绍的量表的信度与效度是评价量表内在质量的两个不同的范畴。一份好的量表，应同时具有较高的信度与效度。信度是效度的必要条件，但不是充分条件。一份量表要有效度，首先必须有信度，没有信度就不会有效度，但是有了信度不一定有效度。信度低，效度不可能高，信度高，效度未必高。效度低，信度很可能高；效度高，信度必然高。

### 三、量表信度的评价方法

评价量表信度的方法有多种，主要有再测信度、复本信度、分半信度与内部一致性信度  $\alpha$

#### (一) 再测信度

再测信度是指用同一份量表，对同一组人员进行前后两次测量，然后用两次测量所得结果的相关系数评价量表信度的方法。所计算的相关系数反映两次测量结果有无变动，也就是测量结果的稳定程度，故又称为稳定性系数。其计算公式为：

$$\rho = \frac{\sum_{i=1}^n (x_{1i} - \bar{x}_1)(x_{2i} - \bar{x}_2)}{(\sum_{i=1}^n (x_{1i} - \bar{x}_1)^2)(\sum_{i=1}^n (x_{2i} - \bar{x}_2)^2)} \quad (4-2)$$

式(4.2)中， $\rho$  为再测信度即稳定性系数， $x_{1i}$  为第  $i$  个被调查者第一次与第二次测量结果的得分， $\bar{x}_1$ 、 $\bar{x}_2$  为前后两次测量得分的均值， $s_1$ 、 $s_2$  为前后两次测量得分的标准差， $n$  为被调查者人数；稳定性系数越高，则量表的信度也越高。

稳定性系数的计算必须满足以下几个假设：所测量的特性必须是稳定的；遗忘和练习的效果相同；两次测量期间被调查者对问题的熟悉情况没有差别。如果以上3条不易做到，则不宜采用再测信度。同时，在获得再测信度时，两次测量的时间间隔要适当。时间太短会有记忆性，太长容易受到其他因素变化的影响。同时第二次测量时，需要调动被调查者的积极性，避免被调查者对于重复测量产生不必要的惰性，影响第二次测量效果。

利用稳定性系数评价量表的效度会存在一些问题。首先，安排同一个被调查者群体参加两次测量存在组织上的困难；其次，第一次测量也许会改变被调查者在第二次测量中的行为和态度；最后，环境与个人因素的改变，会使第二次测量结果发生变动。

## （二）复本信度

复本信度是指对于某个被测量的问题，使用两个平行的量表分别进行测量，然后用这两个量表上回答结果之间的相关系数评价量表信度的方法。所计算的相关系数也称为等值系数，等值系数的计算方法就是稳定性系数的算法：等值系数越高，则量表的信度也越高。

利用等值系数评价量表的代表性，需要对所研究的问题设计出两套平行的量表，这两套量表在测量的内容、测量的难度、量表的数量上应该是一致的，测量可以连续进行，也可以间断进行，但间隔时间要适当，最好是在两周以内。然而，在市场调查中，针对同一个问题，设计出两套完全等价的量表是相当困难的，甚至是不可能的。即便是能够设计出两套等价的量表，从成本上讲也是不经济的。

## （三）分半信度

分半信度是指将用来测量某个问题的一组量表分成两半部分，然后根据两半部分的测量结果计算出的相关系数来评价量表信度的方法。一般做法是，先利用该组量表对被调查者进行一次完整的测量，测量结束后再按量表所在题项的奇数编号与偶数编号进行分组，奇数题一个组，偶数题一个组。也可以将量表所在的全部题项随机地划分为两个部分。然后根据两个部分的测量结果，计算它们之间的相关系数。

如果是按题项编序号的奇偶数分组，首先计算两组之间回答结果的相关系数小，由于E只是半个测量的信度，故需要用 Spearman-Brown 公式进行调整，得到整个测量信度估计值。Spearman - Brown 校正公式为：

$$r = \frac{2}{I+E} \quad (4.3)$$

分半信度尽管不需施测两次或编制两份等价量表，但需要假定两半测量具有相等的均值与标准差。此假定不能满足时，可使用 Flanagan（弗朗那根）公式估计信度。Flanagan（弗朗那根）公式如下：

(4.4)

式(4.4)中，毫、竜分别为各半测量结果的内部方差， $\sigma^2$ 为全部测量结果的总方差。

运用上述方法来评价量表的信度，具有不确定性。这是因为将全部量表分成两半部分，有不同的分组方法，而不同的分组方法会得到不同的相关系数。

## （四）内部一致性信度

内部一致性信度是指用测量同一概念的多个量表之间的一致性或同质性来评价量表信度的方法。对某个或某组概念进行测量，往往需要使用一组量表，这些量表测量的是在所研究的问题上具有内在逻辑关系的多个概念，量表测量的结果之间也应该具有某种逻辑关系。这种逻辑关系决定了测量结果具有某种一致性或稳定性。如果量表能够反映这种内在逻辑关系，那么这些量表所测结果在被测个体上应该具有某种共同的方向，其测量结果之间的方差不会太大，量表的一致性或者信度也会高一些，否则，量表的一致性或者信度就不高。内部一致性信度是市场调查中运用最多的评价量表信度方法。

评价内部一致性信度最常用的是 Cronbach 系数，即“系数”，其计算公式如下：

式(4.5)中, 勇为各量表所在题项的方差, 每个方差所包含的信息都是以单个题项为基础 的; $s^2$ ; 为全部测试结果的总方差, 它是所有量表题项方差与协方差之和,  $K$  为量表的个数。上 述。系数也可以采用式(4.6)计算:

$$\alpha = \frac{K \cdot \sigma_{\text{总}}^2}{K + (K-1) \cdot \sigma_{\text{总}}^2} \quad (4.6)$$

式(4.6)中 $\alpha$ 为全部题项所测结果之间的相关系数。

评价量表信度使用的是上述各种信度系数。信度系数表示信度的大小。信度系数越大, 表明量表的可信程度越高。信度水平的评价可参照 DeVellis(1991)提出的标准, 见表 4-1”

表 4-11 信度的评价标准

信度系数	信度水平	量表处理
0.60 ~ 0.65	有些可信	最好不要
0.65-0.70	比较可信	勉强接受, 需要调整
0.70-0.80	相当可信	可以接受, 少量调整
0.80-0.90	非常可信	完全接受, 不需调整
0.90-1.00	十分可信	完全接受, 极为少见

分量表的信度系数最好在 0.70 以上, 在 0.60~0.70 之间也可以接受。若分量表的信度系 数在 0.60 以下, 或者总量表的信度系数在 0.80 以下, 应考虑重新修订量表或增删题项。

在市场调查实践中, 提高问卷信度的措施主要有适当增加测量的长度; 适度增加新项目, 并注意新增加项目必须与原问卷中的所有项目同质; 测量中问题的难度分布接近正态分布, 减 少特别难回答的问题, 适度增加中等难度的问题; 努力提高测量问题的区分度; 选取适当的被 调查样本; 调查员严格按照要求实施调查等。

#### 四、量表效度的评价方法

量表的效度按其层次不同可划分为多种, 目前被广泛采用的是弗兰士(J. W. French) 和米 希尔(B. Michel)提出的分类, 即按层次将效度划分为内容效度、准则效度和结构效度。需要采 用不同的方法对不同层次的效度进行评价。

##### (一) 内容效度

内容效度也称为逻辑效度, 是指量表实际测到的内容与所要测量的内容之间的吻合程度。内容效度关注的是量表是否充分地覆盖了所要研究的主题。例如, 如果调查人们对某品牌酒 店的印象, 一共设计了酒店外观、酒店地理位置、酒店的外在环境、酒店的标志和酒店的声誉等 5 个测量项目, 并采用了语义差别量表, 来度量人们对酒店的印象。只要是曾经路过即便是没

有住过该酒店的人，也能够回答上述测量项目，测量结果很可能既缺乏代表性，也不全面、准确。这是由于测量的内容并没有完全覆盖所要调查问题的全部要素，诸如酒店设施、房间卫生、餐饮品质、服务态度等都没有考虑进去，实际调查的内容与需要调查的内容不相吻合，内容效度自然不高对内容效度主要采用逻辑分析法与统计分析法进行评价。

逻辑分析方法的思路是由研究者、调研人员、有关专家对测量项目与原定调查目的的吻合程度做出判断，评估所选择的项目是否符合测量的目的和要求，所设计的题项能否代表所要测量的内容或主题，所设计的量表是否符合测量的要求。逻辑分析法属于一种主观性的经验判断方法，可以用来对观测结果做大致的评价。为了获得足够的内容效度，要特别注意设计量表时应遵循的程序和规则。

统计分析法主要采用单项目与量表总和相关分析法获得评价结果，计算每个题项得分与题项总分的相关系数，这个相关系数被称为内容效度系数。根据内容效度系数的大小判断量表是否在内容上有效。内容效度系数越高，说明量表实际测到的内容与所要测量的内容之间的吻合程度越高，量表的效度也越高；反之，则越低。需要注意的是，若全部量表中存在态度方向排序不一致时，应将排序方向统一后再计算总分。

## （二）准则效度

准则效度又称为效标效度，是根据已经得到确定的某种理论，选择一种指标或量表作为准则即效标·计算量表得分与准则的相关系数，这种相关系数被称为准则效度系数。根据准则效度系数的大小判断所采用量表与准则之间的一致性程度，一致性程度越高，说明所采用量表的效度越高；反之，则越低。例如，要调查人们对某种商品的购买动机，使用量表A与量表B进行两次测量，假如以量表B为准则，两个量表测量的结果高度相关，即准则效度系数较大，则说明量表A在购买动机的测量上具有较高的准则效度。

准则效度包括预测效度与同时效度两种类型。预测效度是指一个准则变量的未来值可以通过当前量表的测量来预测的程度·如果使用当前量表测量的结果能够对准则变量进行准确的预测，则当前量表具有预测效度「同时效度是指当前量表与准则量表在同一时点上，对所测概念进行测量的结果之间的相关程度。如果相关程度高，则说明当前量表具有同时效度。

在量表的准则效度评价中，一个好的准则需要具备有效性、客观性、可靠性与实用性等特性，选择一个合适的准则往往十分困难，因而这种方法的应用受到一定限制。

## （三）结构效度

结构效度也称为构造效度，是指测量结果体现出来的某种结论与关于所测属性的理论命题之间的对应程度。换言之，结构效度关注的是量表对所研究假设进行证实的程度。由于测量的是概念的属性特征，概念的属性特征之间存在什么关系，往往是调查研究需要回答的问题，为了回答这样的问题，需要以所测概念为基础、以概念所涉及的理论为依据而建立研究假设。通过测量结果形成的结论是经验层面的，而研究假设是理论层面的，如果前者能够证实后的程度越高，就说明量表的结构效度越高。特别需要说明的是，研究假设必须建立在科学、可靠的理论基础之上，不能脱离研究的目的与初步探索所形成的感性认识。如果研究假设是虚伪的，那么通过量表的测量结果就很难证明研究假设的真伪，也就失去了结构效度的判断基础。同时，量表的设计要与研究假设相匹配，测量的变量必须针对研究假设中所涉及的概念，如果没有这种对应，量表的结构效度也不会高。例如，顾客满意度的相关理论表明，顾客的预期质量、感知质量、感知价值、品牌印象对满意度有着较强的相关关系，如果涉及的量表没有脱离这种满意度的理论架构，而且测量结果也能证实理论上的命题，那么这套量表就具有结构效度。

结构效度的评价可以从两个角度进行，一个是收敛效度，另一个是区别效度。收敛效度是指用来测量同一结构的不同量表之间的相关程度，相关程度越高，收敛效度越高；区别效度是指利用相同的量表测量不同的概念或特征之间的相关程度，相关程度越低，区别效度越高。

结构效度的评价可以采用因子分析法和结构方程法。

市场调查中的大多数测量活动往往要验证一组研究假设是否成立，这些研究假设往往要涉及一组概念

‘大多数的概念只是对现象或事物某方面特性的概括、一般是不能直接测量的，属于潜在变量 c 将概念进行操作化定义后，就形成了观测变量，一个概念往往会由多个观测 变量来加以测量，这样，一次测量活动是否实现了对全部研究假设命题中所涉及的概念进行了测量，每一个概念是否由多个变量被准确地测量出来，就需要进行检验。

运用因子分析法检验量表的结构效度，可以从探索性因子分析与验证性因子分析两个角度来进行评价。探索性因子分析是基于数据统计分析基础，只考虑到数据之间的纯数字特征 而没有任何的理论前提，目的在于从数据出发，寻找数据中所蕴藏的规律。因此，探索性因子 分析应该被认为是理论产生的方法而不是理论检验的方法’如果市场调查中的研究假设不确定或根本就没有现成的研究假设，这时可以通过预调查收集数据. 利用探索性因子分析构建研 究假设

如果在量表设计时，理论上已先构想出量表中相关变量由哪些方面或领域所组成，用因子 分析的目的在于确认变量归类的合理性，验证提取的公共因子数目和组成与假设构想是否一 致，或者各个因子是否能够比较好地被解释，即判断该组量表是否具有较好的结构效度，这时 就需要使用验证性因子分析其基本思想是，研究者首先根据先前的理论和已有的知识，经过 推论和假设、形成关于一组变量之间关系的模型 检验目的在于从理论假设出发，检验检验数 据与理论构想是否相符，在确定不相符时可能需要修正理论构想

市场调查实践中，进行量表的结构效度评价时，可以将探索性因子分析与验证性因子分析 结合起来进行交叉证实，但不能采用同一个数据集口可以将调查数据集分成两个部分:一部分 用于探索性因子分析，构建研究假设；另一部分用于验证性因子分析，验证理论构想。

结构方程模型是建立在一定的理论基础之上，用以检验某一先期理论构想或模型存在性 与合理性的方法问卷结构效度的评价，其实就是对观测变量与理论构想吻合程度的检验。 结构方程模型中涉及潜在变量与观测变量，量表所测量的潜在变量就是被测概念或属性，观测 变量就是量表所要测量的可观测的变量。每一次测量都会包含多个潜在变量，每一个潜在变 量都由一组观测变量构成:所设计的潜在变量之间的经验层面上的逻辑关系是否符合理论构 想，某个潜在变量是否被一组观测变量准确地加以测量，就可以通过结构方程模型中的路径分 析结果加以评价 C 由于任何测量都或多或少地存在测量误差，而运用结构方程模型检验量表 的结构效度恰恰包含了测量误差的假定，这就极大地提高了检验的准确性。

市场调查实践中，提高问卷的效度需要精心编制测验量表，避免出现比较大的系统误差； 妥善组织测验，控制随机误差；创设标准的测试环境；选择好的准则，确定好的准则效度。

## 本章小结••••

市场调查中的数据搜集过程是时所测对象的属性或特征进行测量的过程。本章主要介绍了量表测量技术的有关内容口

测量是指按照特定的规则，将数字或符号分配给研究对象的特征上，从而将其特征量化的 过程。测量有多种类型，无论何种测量，都必须达到稳定性、准确性、完备性、互斥性的要求。量表是市场调查中最常用的测量技术。

量表是指通过一套事先设计的用语、符号和数字，来测量人们心理活动的度量工具。量表 按其计量层次可分为定类量表、定序量表、定距量表与定比量表。市场调查中的量表按照不同的标准可以划分出不同的种类，在实际调查中，可以根据调查对象的特点、研究目的对数据的 要求，选择合适的量表搜集调查数据。

市场调查中常用的列举评比量表、等级顺序量表、分类量表、配对比较量表、固定数目量 表、语义差别量表、中心量表、瑟斯顿量表、李克特量表都有其特定的格式、功能、特点及应用场 合，选择量表时需要考虑研究目的、研究对象的特点、调查方法、被调查者群体特征及其他相关 因素。一次测量活动往往是多个不同量表的结合运用。

量表作为一种测量工具，必须具有一定外在与内在质量 C 外在质量一般可以通过小组座 谈的方式进行

评估，而内在质量主要通过预调查，运用信度与效度指标加以评价。信度指标主要包括再测信度、复本信度、分半信度与内部一致性信度，效度指标主要包括内容效度、准则效度与结构效度。一份好的量表应该同时具备较高的信度与效度。

### 渗关键名词•••

测量	量表	量表的尺度	列举评比量表	等级顺序量表
分类量表	配对比较量表	固定数目量表	中心量表	语义差别量表
瑟斯顿量表	李克特量表	量表的信度	量表的效度	再测信度
复本信度	分半信度	内部一致性信度	内容效度	准则效度
结构效度				

### 回澹回

### 肯思考题••

### 回裝即测即评

1. 什么是测量？测量有哪些类型？
2. 测量应达到哪些方面的要求？
3. 什么是量表？它有什么作用？
4. 按照计量层次的不同，量表可分为哪几种？不同计量层次的量表具有什么特点？
5. 量表的类型可以从哪些角度进行划分？不同的量表类型具有什么功能与特点？
6. 市场调查中常用的量表有哪些？每种量表的格式、功能、特点如何？
7. 选择量表时应考虑哪些因素？
8. 什么是量表的信度与效度？两者之间是什么关系？
9. 如何评价量表的信度与效度？

•-

### 碧瘩章实训…

#### 一、实训目的

1. 熟悉市场调查中测量的基本原理与要求。
2. 熟悉量表设计中需要考虑的因素。
3. 熟悉各种量表的格式、功能与作用。
4. 掌握市场调查中常用量表的设计技术。

#### 二、实训内容

在本章导语中，介绍了某次关于智能手环用户群体市场调查项目。该次调查的变量列于 表 4-1 中。

1. 表 4-1 中共列示了 17 个变量，对每一个变量设计一份量表。
2. 利用小样本对量表的信度与效度进行评价。

#### 三、实训组织

1. 根据实训内容组建团队。
2. 各组完成相关问题的量表设计与预调查，并形成书面材料。
3. 各组围绕设计的量表进行互动交流、质询和提问，并进一步完善量表。理解量表设计的基本规则与技巧。
4. 教师点评。
5. 师生共同评定实训成绩。



# 第五章 问卷设计技术

## 、导语

近年来，随着我国经济的高速发展，私家车的数量不断攀升，交通压力越来越大，道路拥堵现象日益严重，大力发展城市绿色交通势在必行。公共自行车作为城市绿色交通工具对于方便人们短途出行，减轻交通拥堵，具有十分重要的作用。从长远来看，公共自行车将拥有广阔的发展前景。然而针对公共自行车的现有研究仅限于现状的基本分析，鲜有对于其更深层次问题的探索与研究。为了了解某市居民对公共自行车的认知、态度、用户满意度，挖掘潜在用户，提高公共自行车的服务水平与管理水平，为公共自行车的运营商及相关管理部门提供决策信息，某研究机构接受委托特地在某市进行一次城市公共自行车用户群体的问卷调查。

问卷设计人员首先通过查阅文献、咨询专家、实地考察，熟悉了城市公共自行车的运营特征、现状以及居民群体对城市公共自行车的认知与使用行为，在此基础上围绕研究目的，提出了多项研究假设，梳理了问卷调查中所涉及的概念，并对概念进行了操作性定义，最终形成了调查变量。问卷一共包括被调查者个人背景、城市公共自行车认知状况、使用意愿、使用行为、评价态度、用户满意度等6个方面共35个问题，针对各个问题的性质与特点，分别设计了符合研究数据要求的量表。35个问题包括33个封闭式问题，2个开放式问题。为了方便被调查者阅读与填答，问卷在排版格式上做了精心的安排。

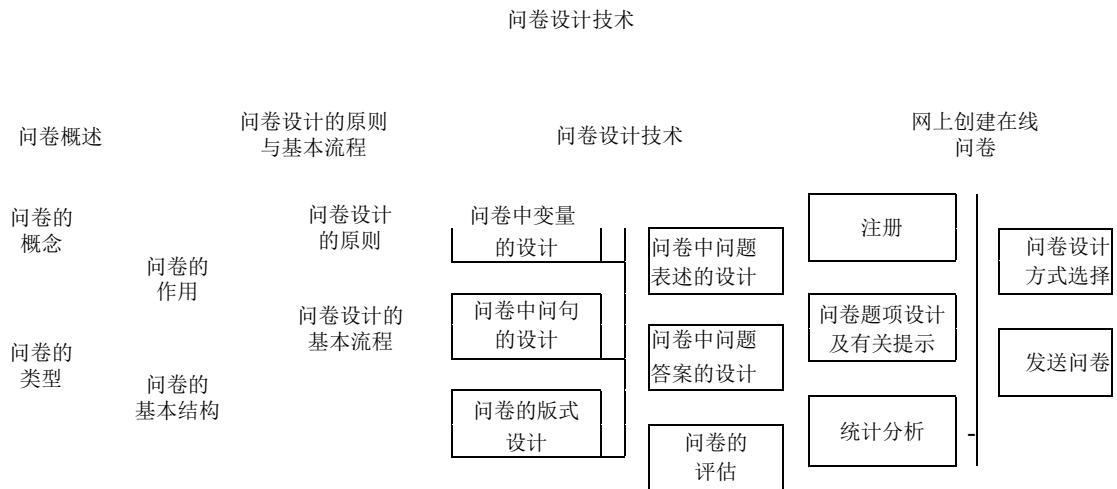
问卷初稿设计完成后，研究人员对问卷中问题的表述与答案的表述进行了初步的评估，认为被调查者能够理解这些表述，且愿意回答并能够回答问卷中的各项问题。通过预调查了解了被调查者填答问题中所存在的问题，并利用预调查的数据对问卷的信度与效度进行了检验。结果发现，问卷中问题与答案的表述还存在个别措辞不易理解、个别问题难以回答或不愿意回答、问卷信度不太高等问题，研究人员针对问卷所存在的问题，对问卷进行了多次修改，最后问卷终于定稿并用于调查的实施。

## 、教学目标与要求

1. 了解问卷的概念、作用与类型；
2. 掌握问卷的基本结构；
3. 熟悉问卷设计的基本原则与基本流程；
4. 掌握问卷中调查问题的确定方法；
5. 掌握问卷中问题表述设计技术及设计问句应注意的问题；

6. 掌握问卷中的设计原则与方法;
7. 熟悉问卷的版式设计要求;
8. 掌握问卷评估的主要内容与方法;
9. 掌握在线问卷创建的基本技能

## 本章知识结构



问卷设计是市场调查流程的重要一环，问卷设计的质量直接影响调查数据的质量，关系到调查研究目标与任务能否实现。问卷设计是一项具有很强的理论性、技术性、艺术性的工作，需要设计者充分熟悉调查研究对象的特点与性质，了解研究课题所涉及领域的相关基础理论，灵活运用统计测量技术。本章重点介绍问卷的特点、类型、结构，问卷设计的一般流程，问卷设计的基本技术以及网络调查中的问卷设计问题。

## 第一节

### 问卷概述

#### 一、问卷的概念

问卷，又称调查表，是指调查研究者根据调查目的和要求，设计出的由一系列问题、备选答案及填答说明等组成的向被调查者搜集调查数据的一种工具。问卷可以是表格式、卡片式或簿记式。

设计问卷，是询问调查的关键。完美的问卷必须具备两个功能，即能将问题清晰、准确地传达给被调查者和使被调查者乐于回答。要完成这两个功能，问卷设计应当遵循一定的原则和程序，运用一定的技巧。

问卷在自然科学、社会科学的许多领域具有广泛的用途。就市场调查活动而言，它是搜集市场信息的基本工具之一。市场研究需要进行规范的调查，标准化的问卷不仅有利于准确、迅速地搜集市场资料和信息，而且便于快速、高效地对这些数据进行处理和分析。

#### 二、问卷的作用

##### 1. 有利于调动被调查者的回答意愿

根据调查研究的目的与任务所设计的调查问卷，能够将需要调查的内容以提问的方式向被调查者提出问题，许多问题还给出了多种可能的备选答案，提供给被调查者进行选择。这种询问方式有助于将调查的内容正确地传递给被调查者，减轻被调查者的回答难度，有利于取得对方的配合，从而提高被调查者的回答意愿，进而为搜集比较准确的数据创造条件。

##### 2. 有利于对调查数据进行统计处理

调查员将设计好的调查问题和答案规范有序地在问卷中列出，并给出相应的计算机编号，可以使调查后期的数据处理变得方便快捷，能够使调查员依据数据处理的结果进行深入的统计分析。

##### 3. 有利于节省调查时间，提高作业效率

由于问卷中已将调查目的、调查内容和填表要求进行了说明。因此，除特殊情况外，一般无须再让调查者就有关问题向被调查者进行说明，只需要对方按照要求完成问卷即可。这样就节省了调查者与被调查者进行沟通的时间，缩短了调查作业流程，提高了作业效率。

#### 三、问卷的类型

问卷的设计必须满足具体的调查目的和任务的要求，需要针对不同的调查研究目的、调查对象、调查内容、调查方法设计不同类型的问卷。

##### (一) 根据使用问卷的方法分类

根据使用问卷的方法可分为自填式问卷和访问式问卷两类。

自填式问卷是指调查者把问卷发给被调查者，由被调查者自己填写问卷，大多数情况下，被调查者在

回答此类问卷时，由于不能与调查者面对面交流，缺乏填答问卷的现场指导，因而从被调查者角度考虑，设计自填式问卷时应尽可能使问卷简单明了、方便填答，无论是问题的表述还是备选答案的表述，都必须能够使被调查者易于准确理解，以确保填答质量。

访问式问卷是由调查者在各种方式的访问中所使用的一种问卷。访问方式包括入户面访、街头拦截访问、小组座谈访问、深层访谈、电话访问等。在访问调查中，由于被调查者与调查者能够面对面交流，有机会获得填答问卷的现场指导，所以访问式问卷较自填式问卷可以设计的稍微复杂一些，以便能够搜集更深层次的数据。但电话访问只是调查者与被调查者的一种语言交流，被调查者不能见到问卷，而且受到时间限制，所以电话访问中的问卷一般比较简单，不能使用比较复杂的问卷。访问式问卷根据需要，可以由调查者向被调查者提问，调查者记录被调查者的回答，也可以由被调查者自行填写。

## （二）根据问卷发放的形式分类

根据问卷发放的形式不同，可分为现场发放式问卷、报刊式问卷、邮寄式问卷、电话访问式问卷、网络问卷。

现场发放式问卷是指在调查现场由调查者将调查问卷直接发给选定好的目标群体，待回答完问题后再进行统一收回，如各类面访问卷。其特点是有确定的传播途径，回收率高，回收时间短，但费用也高。

报刊式问卷是指把问卷刊登在报刊上，随报刊发送到各地，当读者看到报刊后在报刊上填写问卷，然后寄回报刊编辑部。其特点是有稳定的传播途径、保密性好、费用低，回收时间长且回收率不高。

邮寄式问卷是指通过邮局把问卷邮寄给相应的人员，待答完问题后再通过邮局将问卷统一回收。此类问卷适合于特定群体的调查。其特点是有确定的传播途径，保密性好，费用低，回收时间较长且回收率较低。

电话访问式问卷是通过互通电话的形式，向被调查者提问，调查者根据被调查者在电话中的回答情况进行填写。其特点是有确定的传播途径，内容较少，回收率高，回收时间短，费用相对较低。

网络问卷是当前较为普遍的一种形式，是将问卷在网络上发布。其特点是保密措施好，有相对稳定的传播途径，不受时间和空间的限制，可以获得更多的信息。

## （三）根据问卷标准化程度分类

根据问卷标准化程度不同可分为封闭式问卷和开放式问卷。

封闭式问卷是指问卷中事先拟定了问题的备选答案，由被调查者进行选择性回答。封闭式问卷的标准化程度高，提问的方式与备选的答案都是统一的，有利于对调查数据进行统计处理。在调查活动中，若偏重搜集定量数据或可以量化的数据，则提倡采用封闭式问卷。

开放式问卷是指问卷中只是提出了问题而没有提供备选的答案，完全由被调查者根据自己的理解或感受进行填答。开放式问卷由于未提供统一的备选答案，因而其标准化程度低，不便于对调查结果进行统计处理。开放式问卷广泛应用于定性的探索性研究。

# 四、问卷的基本结构

一份完整的问卷通常由标题、说明词、正文、编码、被调查者基本情况、结尾等六部分内容组成。有些特殊的调查问卷可能还包括情景刺激物，如调查中必须向被调查者展示的图片等，可以作为问卷调查的辅助工具。

## （一）标题

问卷的标题是对调查主题的概括说明。标题的拟定应该突出调查的主题，必要时还可以突出调查研究的对象。一个好的标题不仅可以让被调查者了解调查的大致内容，还可以吸引被调查者积极参与。标题要力求简洁明了，切忌模糊冗长。例如大学生网购行为调查问卷，居民保险意识调查问卷，此类问卷标题不仅限定了调查范围，也说明了调查的主题及大致内容。

## （二）说明词

说明词又称封面信，该部分是对调查者身份、调查目的与意义、填答问题的有关事项的说明。说明词往往由两部分组成，一是致被调查者词，二是填答说明。

致被调查者词需要向被调查者说明调查员的身份、调查目的与意义、说明问卷的匿名性和保密性原则，说明对被调查者的希望和要求。该部分的文字表述要力求简练，语气要谦和诚恳、平易近人，给被调查者以好感。要强调调查活动不涉及个人隐私，调查数据仅供研究之用，以消除被调查者的顾虑。这样方能获得被调查者的信任，并激发其参与调查的愿望。特别需要说明的是，调查者身份应尽可能具有较强的权威性与可信度，以获得被调查者的信任与认可。一般而言，研究机构、高等院校、政府机关、社会团体等部门较宜作为调查活动的组织者或委托者。若是企业或行业主管部门所组织的内部调查，就直接以企业或行业主管部门作为调查者的身份。

填答说明需要向被调查者介绍填答问卷的具体方法及有关注意事项，以确保被调查者能够按要求规范地填答问卷，从而有助于保证问卷的有效性及后续的数据处理。填答说明的表述一定要具体、准确、详尽，不能有任何遗漏或让被调查者产生疑惑之处。根据问卷的具体内容，填答说明可以集中安排在问卷前面，也可以分散至各题项之前。为了避免问卷过于啰嗦与凌乱，提倡将填答说明集中安排在问卷的前面。对于个别题项不便于集中说明填答方法时，则在相应题项前说明填答方法。

说明词可以安排在卷首，成为问卷的一部分，也可单独成为问卷的一封附信。

### （三）正文

正文是问卷的主体部分，也是问卷的核心，主要由调查的问题、备选答案、回答方式所组成。调查的问题有封闭式问题与开放式问题两种，但一般以封闭式问题为主。封闭式问题需要给被调查者提供备选答案并规定回答方式，而开放式问题则由被调查者自由作答。正文部分的设计需要调研者在明确调查目的的基础上，充分熟悉被研究现象的基本特征以及与该现象相关的若干学科知识，运用问卷设计中的一些技巧，使所提的问题符合研究的需要。

问卷中所要调查的问题可分为两类，一类是经验事实方面的问题，主要了解市场活动中所发生的客观现象、人们的行为特点和结果；另一类是被调查者的主观判断，如观念、认知、态度、偏好、意见、感受等方面的问题。这两类问题的性质、作用不同，使用的询问方式和询问技术也会有所区别。

### （四）编码

编码就是将被调查者填答的内容，转换成便于识别的数码，以便计算机汇总计算。编码有两方面内容，一是对问题的编码；二是对问题的诸种回答编码。编码一般应用于大规模的问卷调查中。因为在大规模问卷调查中，调查资料的统计汇总工作十分繁重，完整、科学的编码，便于计算机进行调查数据的录入、处理与分析。

编码既可以在问卷设计的同时就设计好，也可以在调查工作完成后再进行。前者称为预编码，后者称为后编码。在实际调查中，对于封闭式问题既可以采用预编码，也可以采用后编码；对于开放式问题则只能采用后编码。在问卷中，编码的方法多种多样，问卷制作者可以根据实际情况选择合适的编码方法。

### （五）被调查者基本情况

被调查者的基本情况是指被调查者的主要自然属性特征与社会属性特征，也称为被调查者背景。在市场调查中，被调查者可以划分为两种，一种是个人，一种是组织单位。这两类被调查者的自然属性特征与社会属性特征是有区别的。如在消费者调查中，消费者自然属性特征包括性别、年龄、身高、体重等，社会属性特征包括婚姻状况、文化程度、职业、收入等；在企业调查中，企业的自然属性特征包括企业名称、所在地区、企业规模、生产能力等，社会属性特征包括企业经济类型、所属行业等。

被调查者的基本情况是多方面的，在一次调查中不必要将其全部基本情况一一加以调查，只需要调查与研究目的有关的少数基本情况。这些少数基本情况的确定往往与研究的假设有关。

## (六) 结尾

问卷结尾部分主要设置一些开放式问题，征询被调查者参与调查的意见与感受；必要时还可以设计调查实施的相关信息，记录调查情况，如调查员的姓名、访问日期、访问时间、访问地点、被调查者的配合情况等；也可以包括对被调查者的感谢语以及有关的补充说明 a

以上六个部分是一份规范、完整的调查问卷所应具有的结构和内容-对于一些简单的调查问卷，如征询意见表、消费调查表等，只需要有标题、问题和答案就够了，无须面面俱到。有时也要考虑问卷调查的方式，如面访式问卷的结尾往往需要有调查情况的记录，而邮寄调查问卷则无须此项内容。

# 第一 T>

## 问卷设计的原则与基本流程

设计问卷是一项技术含量很高的工作。在问卷中设置哪些问题，如何提问，采用什么方式回答，如何使被调查者愿意并准确回答问题，以期获得能够满足研究目的要求的调查数据，都是问卷设计者需要认真思考的。下面主要介绍问卷设计的原则与基本流程。

### 一、问卷设计的原则

为了保证问卷设计的质量，问卷设计中需要遵循一定的原则。

#### (一) 目的性原则

目的性原则是指所设计的问卷必须符合研究目的的要求。具体来讲，问卷的设计必须从调查研究的目的出发，认真分析调查现象的特点与规律，明确调查项目。首先要拟定出与调查研究课题相关的全部变量，并对这些变量进行筛选，保留主要的、关键的变量，然后按变量的特点提出调查的问题。同时，要确保调查问题没有重复与遗漏，使所搜集的数据能够满足调查研究目的的需要。这是问卷设计最基本的要求

## (二) 科学性原则

科学性原则是指设计问卷时必须坚持科学的态度，不能随心所欲。设计问卷是一项复杂且具有较高难度的工作，需要科学理论的指导。例如，在进行消费者动机调查时，就需要了解消费者心理学、行为学中关于动机的有关理论，以及动机的表现、形成因素、关联因素等相关知识。科学性是问卷设计中重要概念、特征、关系合理体现的重要保证。

## (三) 逻辑性原则

逻辑性原则是指问卷设计中问题的生成、问题的排列、问题之间的关系要具有逻辑性。如果调查研究需要验证有关研究假设，就可以从研究假设所涉及的概念出发生成调查研究的具体问题，如果调查研究只是描述市场特征，就应该在相关理论的指导下，通过逻辑分析，推理出需要调查的问题。问题的排列应有一定的逻辑顺序，符合被调查者的思维程序。一般是先易后难、先简后繁、先具体后抽象。同时还要评估问题的编排效应，确保相邻两个问题的回答是独立的，前面问题的回答不影响后面问题的回答。此外，问卷所涉及的调查内容往往区分为多个模块，模块内的问题与问题之间、模块与模块之间也要有一定的逻辑关系。

## (四) 中立性原则

中立性原则是指问卷设计时要保持提问项的中立性。提问不能带有任何诱导、提示的色彩，问句的表述不能掺杂调研者的个人倾向。否则，被调查者在不合理的提问引导下，会出现回答偏差，影响调查数据的质量。尤其是在测度被调查者一些主观方面的问题时，要注意中立性原则。

## (五) 方便性原则

方便性原则是指所设计的问卷必须使被调查者能够方便作答、调研者能够方便地整理数据。调研者在设计问卷时，需要站在被调查者的角度，评估整份问卷是否便于被调查者回答每一个问题。这就要求问卷的问题表述应该通俗、准确、清晰，避免使用过于专业的词语和意义模棱两可的表述，使被调查者能够容易准确理解问题的含义。对于问题的回答方式，问卷中必须给予明确的提示，以方便被调查者回答。另外，设计的问卷还要能够为调研者自身进行数据的整理提供方便。

## 二、问卷设计的基本流程

问卷设计的基本流程包括设计问卷前的探索性研究，设计调查问卷初稿，问卷的试用和修改。一份可实施的问卷需要经过多次讨论、修改与完善。

### (一) 设计问卷前的探索性研究

#### 1. 探索性研究的含义

所谓探索性研究，是调查者对所研究而又不太熟悉的现象或问题，运用一定方法，进行初步熟悉、了解、思考、分析的过程。其基本目的是提供一些资料以帮助调查者认识和理解所面对的问题，使问题定义得更准确、更清晰。探索性研究是问卷设计的基础工作。

设计问卷之前，调查者可能不太熟悉所研究的具体现象所涉及的专业性问题，或对所研究问题的理解与定义把握不准，就需要运用探索性研究。探索性研究的研究过程灵活，无固定结构，所需的信息是不能精确定义的。例如，向行业专家咨询就是一种探索性的研究。探索性研究的样本量一般较小，原始数据一般是定性的。探索性研究的结果一般作为进一步研究的开始。

#### 2. 探索性研究的方法

探索性研究的方法有多种，比较常用的有文献查阅法、咨询访问法、实地考察法、小样本调查法，调查者根据需要可以选择其中的一种或多种方法。

文献查阅法是文案调查法的一种具体应用，是一种比较方便快捷的探索性研究方法，在帮助调查者获

取相关研究背景、熟悉相关专业性问题、了解研究所需的相关理论与方法等方面具有独特的作用。但是，文献查阅法往往不能完全满足探索性研究的需要，还需要采用其他的方法予以补充。

咨询访问法是调查者向专家、知情人士请教、学习的过程，通过查阅文献获得的资料是有限的，并且与实际情况可能存在差距。此时，咨询访问就显得非常必要，咨询访问的对象可以是对本课题有深入研究的专家，掌握第一手资料、具有丰富经验的实际工作者，上级有关领导和主管者。咨询访问可以采取个别访问、集体访谈的方式，咨询前需要列出咨询提纲，明确需要咨询的问题，咨询中要尽可能详细地搜集资料，并做好记录。咨询访问法在帮助调研者了解专业性问题，加深对调研问题的理解，增强问卷中调查问题的针对性方面有较强的作用。

实地考察法是指通过实地调查搜集一手资料的过程。文献查阅的资料是二手资料，咨询访问的资料往往含有主观色彩或不具有普遍的代表性，为了掌握具有说服力的第一手资料，还必须进行实地考察。实地考察不同于正式调查，参加者应该是调查活动的领导者、组织者和调查方案的设计者，调查范围应小，对象应少，可采取解剖麻雀的办法，选择有代表性的对象进行较深入的调查。通过实地考察得到的第一手资料，可以验证通过文献查阅、咨询访问所得材料的真伪，决定取舍。此外，将一手资料与二手资料进行综合，还可能发现课题研究的新角度，形成独特的观点，完善研究假设。

小样本调查法是在大规模正式调查前，通过小样本检验问卷质量的研究方法设计的问卷是否符合课题研究的需要，是否便于被调查者回答，是否具有信度与效度，问卷调查活动的组织方案是否可行等，均需要在正式调查前进行检验。小样本调查是一种预调查，样本规模无须过大，也无需强调样本的随机性。利用小样本调查的结果，可以分析和发现问卷以及整个调查方案可能存在的问题，通过对所发现的问题逐一加以解决，可提高调查方案的可行性及问卷的质量。

## （二）设计调查问卷初稿

设计调查问卷初稿是一项相当复杂的工作。问卷设计者需要根据调查研究的目的，结合探索性研究阶段所掌握的资料，在研究假设的指引下或研究逻辑的推理下，按照问卷的结构与要求，系统地提出调查问题，设计回答方式，形成问卷初稿的书面文件。本阶段工作的核心与关键，是确定调查问题和设计量表。

### 1. 确定调查问题

确定调查问题的根本依据是调查研究的目的，调查研究的目的不同，问卷中设计的调查问题也会不同。同时，也要考虑调查研究的方式与方法，调查的方式与方法不同，问卷中的调查问题也会有差别。在市场调查中，调查研究的目的大致可以分为两大类，一是验证研究假设，二是描述市场特征。

以验证研究假设为目的的调查研究，其问卷设计中调查问题的确定，可根据研究假设的命题拟定调查问题。每一个研究假设均会涉及至少两个概念，这些概念往往就是我们需要调查的问题。如在“消费者的文化差异是导致消费偏好的重要因素”这个假设中，出现了“文化差异”与“消费偏好”两个概念。这两个概念是需要通过问卷加以测度的研究假设命题中出现的概念，有些可以直接对其进行测度，如在“消费者的年龄不同对支付手段的选择也不同”这

个假设命题中，年龄这个概念可以用年龄的问题来测度；支付手段这个概念可以用各种支付手段的选择意愿的问题来测度。对于研究假设命题中出现的不能直接测度的概念，就需要进行变量化处理，即将其转化为可测度的变量。如上述命题中的“文化差异”可以通过“文化程度”或“学历”来进行测度；“消费偏好”则可通过用颜色偏好、款式偏好、支付偏好、环境偏好等变量来具体测度。将需要测度的这些变量按照提问的方式表述出来，则需要调查的问题也就自然形成了。

以描述市场特征为目的的调查研究，往往不需要建立研究假设。对于不需要建立研究假设的调查，其问题的提出往往通过对研究课题的认识，在探索性研究的基础上，运用逻辑推理的方法提出问题。逻辑推理方法主要是明确需要研究的课题需要哪些方面的信息，每一个方面又具体包括哪些调查项目，通过层层推理，形成调查问题。比如，要研究用户满意度，根据满意度的内涵，需要从用户的期望、用户感知的质量、价值、满意度、抱怨等方面进行研究。而这些变量都是一些抽象的概念，属于潜变量，不具备直接测度的条件，还要将其中的每一个潜变量分解为多个可以测度的显变量。如感知的质量可根据研究的具体商品或服务对象，分解为不同的质量属性。对于冰箱而言，感知的质量可能包括冰箱的智能化水平、制冷效果、耗电量、安全性。这些属性就是问卷设计中调查问题设置时需要考虑的。

此外，某些调查研究需要了解被调查者的背景信息，调研者就需要分析哪些背景信息是与调查研究目的相关的，将相关的背景信息筛选出来，也构成问卷中需要调查的问题。假如被调查者的年龄、性别、职业、收入与我们调查研究的课题有关系，则需要将这些变量以提问的方式表述出来，这些问题就构成问卷中问题的一部分。

问卷中问题的确定，也要考虑调查的方式与方法。如果调查方案中规定的是面访式问卷调查，则问卷中的问题可以适当多设置一些，而且问题也可以稍微复杂一些。如果是采用电话调查，则问卷的题量要相对少一些，提问的方式也要简单一些，以便于被调查者理解与记忆。

## 2. 设计量表

问卷中确定了需要调查的问题，只是明确了调查的内容。每一个问题对应一个变量，调查结果的准确性与可靠性，取决于测度这些变量的量表选择。量表的选择要根据问题的特点、属性、复杂程度、敏感程度以及对调查数据的计量层次要求来决定。

量表按其功能与特点可以划分为多种。问卷设计者应该全面了解各种不同量表的功能与特点，在设计量表时科学合理地选择量表。只有将调查问题与相应的量表匹配起来，才可能形成调查问卷。

### （三）问卷的试用和修改

由于受问卷设计人员理论水平、实际经验、对研究对象的了解程度等因素的限制，设计出来的问卷初稿不一定就是完美的。在问卷正式使用前，需要对问卷进行试用和必要的修改，以提高问卷质量。

#### 1. 问卷的试用

问卷的试用就是通过预调查来检验问卷的质量是否符合调查研究的要求、是否能够使被调查者愿意并方便地回答各项问题、是否能够搜集到真实、准确的数据。问卷的试用是问卷设计过程中的必要环节，对调查研究能否实现既定目标有着重要影响。

通过预调查所进行的问卷试用，旨在检验问卷本身的质量，并不是对所研究的对象进行正式的调查。所以，预调查阶段一般采用小样本，而且可以是非随机样本，任何非概率抽样方式都是适用的。根据经验，预调查的样本量以题项数量的5至10倍为宜。调查者可以视经费条

件，确定预调查的样本量。

预调查以面访式问卷调查为宜，这样有利于调查者了解被调查者填答问卷时所遇到的具体问题，并通过与被调查者的交流，明确这些问题产生的原因，以便后续对问卷做有针对性的修改。

预调查除了要搜集规定的问卷数据以对问卷的信度与效度进行检验以外，还要通过与被调查者的交流和对被调查者填答问卷时情绪的观察，着重搜集下列情况：

- (1) 被调查者参与调查的意愿、态度、情绪等主观反应；
- (2) 被调查者对问卷所提问题的理解程度，回答问题是否遇到障碍及具体原因；
- (3) 被调查者对填答问卷的兴趣，以及拒答的情况及其原因；
- (4) 被调查者对问卷题量的反应，对回答难度的评价；
- (5) 被调查者对问卷修改的有关建议。

通过预调查所搜集的问卷数据及相关信息，可以对问卷初稿展开评估，发现问卷初稿所存在的问题与不足，为问卷的修改与完善提供依据。

需要指出的是，问卷的预调查不仅可以用来检验问卷，还可以用来检验整个调查方案的可行性。预调查结束后，研究者可以知道被调查者回答一份问卷一般需要花费多少时间，一般一名调查员一天可以完成多少份问卷的调查，并以此为依据制定出调查进度表；

## 2. 问卷的修改

通过预调查完成问卷的试用后，针对问卷初稿所存在的问题，要分析原因并逐一加以修改。问卷修改的内容可能包括问卷的说明词、填答问卷的要求与注意事项、问题的表述方式、问题的回答方式、问题的编排顺序、问卷的整体结构与逻辑、问卷的排版格式等。对于修改较大的问卷还需要再一次进行试用，以检查问卷是否完全达到质量要求。只有达到质量要求的问卷，才可在正式调查中使用。

心一、-4+  
第二 T>

# 问卷设计技术

问卷设计在很大程度上决定着问卷调查的回收率、效率、回答的真实性，以至整个调查研究的成败。问卷调查的质量取决于问卷设计的质量。因此，科学地设计问卷，在问卷调查中具有重要的意义。问卷设计的核心内容是问卷中变量的设计、问题的设计与答案的设计。

## 一、问卷中变量的设计

问卷中变量的设计是问题设计的基础。需要调查的变量设计好了，就可以将这些调查变量转换成问卷中一个个问题。

调查变量的设计实际上是一个不断将抽象的调查课题降低抽象层次，使之实现操作化的过程。所以，调查变量的设计离不开对调查课题的理解与研究：

调查课题与研究假设是紧密联系的两个范畴。调查课题是市场调查要说明或研究的主题，也就是调查研究的任务。调查课题往往会涉及一系列概念以及概念与概念间的关系，研究假设是对调查任务的一种具体描述，往往由概念或变量及其关系的命题所构成，研究假设所涉

及概念的抽象层次比调查课题要低一些。概念的抽象层次越低，概念越容易被测度。课题操作化主要是概念的操作化与命题的操作化。

概念是科学认识活动最基本的工具，是人们抽象思维活动的产物与基本单位，其实质是综合概括同一类事物或现象的抽象名词，概念一般只能用抽象思维去把握与理解，很难直接加以测度。在市场调查中，通过对抽象概念的定义来确定调查变量，将抽象的概念转化为可测度的变量的过程称为概念的操作化。概念操作化是通过对概念进行界定而实现的。

对概念进行界定有两种方法，即真实界定法与名义界定法。真实界定法就是给概念下抽象定义，抽象定义是用文字来概括说明一个概念的内涵与外延，对某一事物的本质做出陈述。事物的本质是被发现的，而不是人们所赋予的。抽象定义还只是抽象层次上对概念进行界定，而没有说明与概念相对应的各种具体现象，因而不能直接被测度。例如，购买动机是指为了满足一定的需要而引起人们购买行为的愿望或意念。这个概念并没有说明具体的购买对象，也没有说明愿望或意愿的具体内涵。所以，购买动机仍然是一个抽象的概念，不能直接被测度。名义界定法就是给概念下操作定义。所谓操作定义是依据抽象定义所界定的概念内涵与外延提出一些可测度的变量来说明如何度量一个概念，这样就把概念与可测度的特征连接起来。例如购买动机的操作性定义可以概括为：理性动机+感情动机。然后，根据理性动机与感情动机的类型，提出需要测度的具体动机，如求实动机、求新动机、求异动机、趋众动机等。有了这种操作性定义，就能够使调研者设计出具体的调查变量。

通过操作性定义对概念进行界定，可以提高市场调查的客观性，有助于对市场现象进行精确、客观的描述、比较与分析。例如，我们可以通过消费者的收入水平、受教育程度、职业三项可以测度的变量描述和比较不同消费者的“社会地位”这个抽象的概念。操作性定义将抽象概念与经验事实联系起来，使调研者可以对其进行经验研究。有了操作性定义，我们不但可以提高市场调查的可比性，而且可以对调查结论进行客观检验。

对概念进行了名义界定，给出其操作定义后，就可以选择或设计调查变量了。调查变量是概念内涵中某一方面的指示标志，其表现的是经验层次的现象。所以，一个概念的操作定义可以分解为几个调查变量，对同一个概念也可以采用不同的调查变量进行测度。

设计调查变量可以采用三种不同的方法：

1. 采用客观存在的具体事物或数量标准作为概念的调查变量。例如，在测度“企业规模”这个概念时，可以将企业产值、固定资产原值、职工人数、生产能力等这些客观存在的指标作为调查变量，来衡量或判断企业规模。

2. 采用可观测的市场现象或行为作为概念的调查变量。例如，在测度“消费者偏好”这个概念时，由于该概念比较抽象，可以将其分解为品牌选择偏好、购买场所偏好、支付方式偏好等具体的调查变量、来度量消费者的偏好特征。

3. 对于难以用具体客观标准和市场现象来界定的概念，可以采用主观评价的方法设计调查变量。例如，在测度“消费者维权意识”这个概念时，由于没有具体客观的标准变量，也没有相应的能够直接度量的维权意识变量，则只能根据被调查者对一些维权方面的法律法规或规范性文件的了解程度来测量消费者维权意识的强弱程度。设计的每一项关于对维权方面的法律、法规、规范性文件的理解程度都是调查变量。

需要指出的是，上述三种方法并不是相互排斥的，对于同一个概念，可同时采用几种方法设计调查变量。

完成概念操作化后，就可以再进行研究假设的操作化。研究假设是对抽象概念之间关系

所做的一种尝试性的陈述，虽然假设中所使用的概念大多数是结果界定了的，但它们一般还不能与经验事实相联系，还需要进一步地界定和操作化，将抽象假设转为具体假设。假设的操作化可以通过用调查变量替换假设中概念的方法进行。例如，在“手机用户的期望越高满意度越低”这个研究假设中，用户期望与满意度这两个概念都可以分解为具体的调查变量。如满意度可以分解为手机总体满意度、功能满意度、价格满意度、待机时间满意度等调查变量。

问卷中变量设计完成后，需要将变量以提问的方式形成问卷中的问题，即要进行问题设计。

## 二、问卷中问题表述的设计

问卷中问题表述的设计，就是将调查变量具体地转换成问卷中一个个问题的过程。这一过程主要是根据调查变量的属性与特点，运用合适的方式进行问题的表述，问句的表述可以采用问句的形式，也可以采用陈述句的形式”

### （一）问题表述的基本原则

问题的表述，既要考虑调查变量的属性与特点，也要考虑被调查者的理解能力、回答兴趣与回答心理。设计问卷时需要遵循下列基本原则：

#### 1. 准确性原则

准确性原则是指问题的表述能够准确地表达所要调查的内容。c 调查研究课题确定了调查的内容，围绕调查研究的课题而设计的一系列调查变量又将调查内容进一步具体化，问题的表述将调查变量转换成一个个问题。如果问题的表述不能准确地表达调查的内容，所搜集的数据偏离调查研究课题，就会失去调查意义，达不到调查研究的目的。

将调查变量转换成调查问卷中的问题，可能会存在两种情形：一种情形是一个调查变量对应一个问题，另一种情形是一个调查变量对应一组问题。e 如果使用一个问题能够测度所要调查的某个变量，就采用第一种情形，否则采用第二种情形。

#### 例 5.1

调查的变量是“广告的传达效果”，对应的问题表述为：

您是否认为 XX 广告具有较好的效果？

①是                  ②否

针对广告传达效果，这个问题的表述就不准确。首先，变量是要测度广告传达效果，而问题的表述中则抽象化为广告的效果。其次，广告的传达效果需要通过到达度、注意率、记忆率、理解度、喜爱程度等几个方面来进行评价，而上述问题的表述中没能将广告传达效果具体化。所以，针对广告传达效果这个变量，就可以根据上述 5 个测评指标至少可以提出 5 个问题。只有这样才能准确地测评广告的传达效果。

#### 2. 清晰性原则

清晰性原则是指问题的表述要清晰明了，便于理解，不致使被调查者对问题产生歧义。问卷是要给被调查者阅读与回答的，如果问题表述不清、意思模棱两可，容易导致被调查者误解，且易造成回答困难。

#### 例 5.2

您是否认为色彩鲜艳的广告会给人带来积极的视觉效应？

上述表述中的“积极的视觉效应”可能使很多调查者难以理解。这样的问题表述就不符合清晰性原则的要求。

### 3. 客观性原则

客观性原则是指问题的表述应该使用中性词，避免引导和提示，不能体现设计者自己的主观意思。

#### 例 5.3

人们都认为 xx 冰箱的制冷效果不错，您觉得呢？

该问题的表述明显带有诱导性，可能有不少被调查者会被问题诱导，以人们普遍认为的观点作为自己的回答结果，然而这种回答并不一定是自己真实的观点，从而破坏了数据的真实性。

上述原则是问题表述中应遵循的最基本的原则。问句的设计还有许多应该注意的具体细节问题•将在后续内容中介绍。

## (二) 问题表述的一般方式

问题的表述有两类方式，即直接表述法与间接表述法。直接表述法也称为直接提问法，就是将调查意图与所要调查的内容直接用文字表述出来，被调查者能够清楚地了解所要调查的真实意图与内容，直接表述法有许多具体方法，如简单询问法、简单陈述法、释疑法、假定法 等’ 间接表述法也称为间接提问法，就是将调查的真实意图与所要调查的内容采用隐含性文字或图表表述出来，被调查者并不清楚所要调查的真实意图，在此情形下，被调查者更可能就 调查的问题给予真实的回答。直接表述法适合于非敏感性问题，而间接表述法则适合于敏感 性问题。这是因为对被调查者来说，敏感性问题往往具有隐私性、恐惧性、难堪性，如果采取直 接表述的方式，他们往往不愿意给出真实的回答。

### 1. 直接表述法

#### (1) 简单询问法

简单询问法就是将调查内容用一句简短的疑问句直接表述出来。该表述方法是问句设计 中最常见的一种问题表述方式。

#### 例 5.4

您在购买服装时考虑的最主要的因素是什么？

- ①颜色 ②款式 ③质地 ④价格 ⑤品牌 ⑥产地

此类表述方式就是简单询问法，简单询问法的表述方式简单明了，被调查者容易理解题意， 设计也比较简单。但这种表述方式只适合于比较简单的问题，对于比较复杂的问题效果并不好。

#### (2) 简单陈述法

简单陈述法就是将调查内容用简短的陈述句形式表述出来该表述方法也是问句设计中 一种常见的问题表述方式。

#### 例 5.5

请您阅读下列表格中关于本酒店服务的陈述句，根据您自己的感知经历做出判断(在相应的格子中画“。”)：

酒店服务评价陈述	态度				
	非常同意	同意	无评价	不同意	很不同意
入住手续办理速度快					
房间干净整洁					
洗漱用品齐全					

简单陈述的表述方式也比较简单明了，具有简单询问法的特点，但要比简单询问法具有更 高的效率。在测度被调查者的主观态度时，如果测度的项目较多，而测度的量表相同，采用该 种表述方式就比较合适，既能节省问卷篇幅，也能方便被调查者阅读与回答。

#### (3) 释疑法

释疑法就是在问题的主题之前加上一段解释性的文字，对被调查者可能不太熟悉的专业词汇进行解释，帮助其准确理解题意，或消除被调查者的回答顾虑，促使其表露自己真实的想法或态度。

#### 例 5.6

随着社会的进步，有越来越多的男士使用化妆品。您(男士)使用过男士化妆品吗？

- ①使用过      ②未使用过

有部分男士认为自己使用化妆品是一件不好意思的事情，在询问其是否使用过化妆品时，即便有使用经历，也可能不如实回答。此时，在问题前面加上一句解释性文字，有助于消除男士的回答顾虑，以便获得真实回答。所以，如果设计者认为所提的问题可能带有一定的敏感性，就可以考虑使用释疑法来表述问题。但在问卷设计中，这种表述方法不宜过多使用，否则会增加问卷篇幅和被调查者的阅读量，不利于被调查者的回答。

#### (4) 假定法

假定法就是用一个假言判断作为问题的前提，然后再询问被调查者的看法。常用于意愿调查。常用格式是：“假定……您是否会”“如果……您将会……吗？”等。

#### 例 5.7

如果有以下工作，您将会选择哪一项？

- ① 月薪 5 000 元，每天工作 8 小时  
② 月薪 7 000 元，每天工作 10 小时  
③ 月薪 9 000 元，每天工作 12 小时

采用假定法来了解被调查者的意愿和行为倾向是比较有效的，但不宜多用。假定条件毕竟不是事实，被调查者有时也很难把握自己，当假设成为现实时，其态度与行为可能会发生改变。所以，此时我们所搜集的数据不一定十分可靠。

## 2. 间接表述法

#### (1) 转移法

转移法是指由他人直接回答问题，然后再请被调查者对他人的回答做出评价。采用这样的表述方法可以降低问题的敏感性和减轻对被调查者的威胁性。比如，在家庭产品测试中，需要征询用户对产品的评价意见，由于产品是免费试用的，用户往往不好意思对产品给予差评，可能会做出比较中性的评价。此时，若运用转移法提问，可能效果会好些。

#### 例 5.8

对于您试用的这款产品，有的用户认为使用方便，有的用户认为使用不方便，您同意哪种看法？

- ①同意第一种看法      ②同意第二种看法

转移法虽然降低了问题的敏感性，但问题的陈述中一般不能表述出他人过多的观点和比较细致的看法，所以搜集到的数据往往计量层次较低。

#### (2) 情景法

情景法是指设计一个情景，让被调查者设身处地地表露自己的看法或意向，这种方法有利于了解被调查者的真实想法。情景法与假定法既有相同之处，也存在不同特点。两者都是从假定出发，要求被调查者在假设条件的前提下回答问题。两者的区别在于，情景法要设计一个与调查内容相关的情景、而且调查的意图往往隐藏在问题本身之中，被调查者不易察觉设计者的真实意图。

#### 例 5.9

某女工在企业实行的优化组合中失去了工作。您能猜猜是什么原因使她失去了工作吗？

- ①人际关系不好      ②工作不努力

③技术水平不高                          ④身体不好，经常请假

⑤没有什么特殊原因，只因她是女性

情景法将被调查者置于某种情景之中，容易使被调查者将该情景与现实情况联系起来，且不知不觉地根据自己的真实看法或现实生活中的真实感受做出选择，其所选择的答案往往是其内心的真实想法。如果直接询问被调查者，在企业的优化组合中是否存在性别歧视，他们往往按照社会共同价值观而不是根据自己的真实想法来回答问题，选择的答案可能是不存在性别歧视，而其内心深处往往不这么认为。

间接表述法中还可以使用投影技法，参见第三章第一节的相关内容。

### 三、问卷中间句的设计

问卷中的问句是给被调查者阅读的，这就必须站在被调查者的角度来设计问句。设计问句时除了要遵循前述的基本原则外，在用词造句、表述习惯与表述规范方面还要注意以下问题，以便使被调查者能够准确理解所要调查的内容，并给予准确的回答。

#### (一) 问句设计中的用词造句

问句设计中的用词造句是有讲究的，设计者不可以随心所欲地表述问题。问句用词造句最基本的要求就是，被调查者既要能够准确理解问句的含义，也要能够给出尽可能准确的回答。问句设计中的用词造句需要注意以下几个方面的问题：

##### 1. 问句要通俗易懂

设计问句时，要考虑调查对象的特征，问句的表述不能超过被调查者的理解能力。如果所设计的问卷是用于大规模调查，被调查者的文化程度可能参差不齐，问句设计的原则只能是“就低不就高”，也就是问句的表述应尽可能使文化水平较低的被调查者能够准确理解其意义。这就要求在问句的表述中尽可能使用通俗的、一般的词语，避免使用专有名词和专业术语。例如，如果要调查居民对物价变动的感受，问题表述为“您认为 CPI 是否真实地反映了物价水平的综合变动？”，该问句中的“CPI”是英文居民消费价格指数的缩写，是一个专有名词，有些被调查者可能不了解其含义，这个问句就不符合通俗易懂的要求。此时，可以将问句改为“您感觉目前的物价水平怎么样？”，这样提问就能够使大多数被调查者理解问句的含义。

##### 2. 问句要精练简洁

一般而言，被调查者是不乐意阅读表述过长的问句的。表述过长的问句会增加被调查者的阅读量，增加其接受调查的时间，同时也可能造成理解难度，进而影响被调查者的情绪，这对搜集调查数据都是不利的。所以，在意思能够表达清楚的前提下，问句越短越好、越简单越好。例如，“奢侈品消费在一些群体中已成为一种时尚，您是否考虑在适当的时候加入奢侈品消费群体的可能性？”，该问句不仅啰嗦冗长，而且语义不清，不便理解；如果改为“您最近半年是否有购买奢侈品的愿望？”，此时问句就显得简洁明了，易于理解。

##### 3. 问句要清楚明了

问题的表述不能包含概念不清、含义模糊的词语。否则，被调查者难以准确理解问题的含义，或者难以进行准确的回答。例如，“您经常在超市购物吗？”这一问句中，使用了“经常”这个意义比较含糊不清的词语，被调查者很难把握“经常”究竟是指什么频率，不同的被调查者对“经常”的理解或感知是不一样的。两个对此问题回答“是”的被调查者，可能每月到超市购物的次数相差甚远，这样就难以准确度量人们到超市购物的频率。如果将此问题改为“您每月到超市购物有多少次？”，这样提问既便于被调查者准确理解问题的含义，也能使其给予比较准确的回答。又如“您春节是出门旅游，还是休息？”这个问句的概念表述就不清晰，被调查者很难回答。该问句中似乎将出门旅游与休息两个概念对立起来，其实

这是两个有关联意义的概念，出门旅游也是休息的一种形式’如果将此问句改为“您春节是出门旅游，还是在家休息？”就不存在概念不清的问题了，被调查者也能方便作答

#### 4. 问句要一题一问

所谓“一题一问”是指每个问句中只能提一个问题，不能存在一问多答的表述。例如，“您觉得海尔空调的价格与性能怎么样？”这一问句，就存在一问多答的问题，被调查者是难以作答的。可能有被调查者认为海尔空调价格合理但性能一般，或价格稍贵但性能较好，如果只用一个答案来回答相关联的两个问题，被调查者就不好作答。该问句实际上包含了两个问题，应该对海尔空调的价格与性能分别进行提问设计问句时特别要注意，一个问句只能涉及一个单一的事物，不要用两个以上的观念或事件。

#### 5. 问句要尽量避免否定形式提问

被调查者阅读问卷时往往一扫而过，一般不会去仔细阅读，这样就容易把句子中的否定词看漏，误将否定句看成肯定句，造成答案错乱。例如，“您平时对广告不感兴趣吗？”与“您是否赞成电视节目之间不插播广告？”这两个问句均采用了否定式提问。有研究表明，被调查者对否定式提问一是感觉别扭，二是容易忽视否定句，从而给出不准确的答案。而且，采用否定式提问也存在一定的诱导倾向，容易造成被调查者朝着诱导性方向回答问题。上述两个问句分别改为“您平时对广告感兴趣吗？”与“您是否赞成电视节目之间插播广告？”，这样被调查者在阅读时既不感到别扭，也不会被诱导，从而会给出准确的回答。

#### 6. 问句要尽量避免敏感性问题

敏感性问题是指涉及被调查者隐私、被调查者回答可能感到难堪或威胁的一类问题对于敏感性问题，被调查者是不愿意或不敢给出真实回答的。如果问句中涉及敏感性问题，一般难以获得真实的回答。例如，“您公司有过偷税的行为吗？”这一问句，就涉及敏感性问题，被调查者是不愿意或不敢做出真实回答的。

#### 7. 问句要考虑时间性

在调查某些具有时间属性的变量时，不可避免地会采用时间限定词。设计问句时，需要考虑时间限定的合理性。如果时间限定周期太长，容易造成被调查者回忆与回答的难度，从而影响回答结果的准确性。例如，“您去年家庭生活费支出是多少？”这个问句，时间限定周期较长，被调查者是难以进行回忆与准确回答的，如果缩短时间限定周期，改为“您上月家庭生活费支出是多少？”，被调查者的回忆难度就会大大降低，给出比较准确答案的可能性就会提高。问卷设计者在设计问句时，应该根据调查变量的特点，选定合适的时间限定周期。如果问卷设计者认为，问句中所设定的时间限定周期可能会对被调查者造成回忆困难，就应该缩短时间限定周期 C

#### 8. 问句的表述要有礼貌

一般情况下，作为被调查者是没有义务来接受调查的，既然他能够接受调查，就应该对他合作态度表示充分的尊重，其中，在设计问句时就要体现对被调查者的尊重，问句的表述应该礼貌谦逊。例如，“您有刷牙的习惯吗？”这一问句就存在提问不礼貌的问题：讲究个人卫生，应该是人们的一种良好习惯，如果直接询问被调查者是否有刷牙的习惯，就是对被调查者的不尊重，容易引起被调查者的反感或敌意，从而带来回答的偏差，如果将上述问题改为“您每天刷几次牙？”就不存在礼貌问题了。另外，在称谓上最好使用“您”而不使用“你”，给被调查者一种被尊重的感觉，这有利于提高被调查者参与调查的配合程度。

### （二）问句的非倾向性

保持问句的非倾向性是问句设计应遵循的基本原则之一。设计问卷时，要保证问卷中的每一个问句都是中性的，不能带有某种提示性、诱导性、倾向性。如果不注意这一点，往往不能客观地测量被调查者的行为与态度，从而使问卷的效度与信度受到影响。G 问卷设计者往往会有意无意地将自己的观点、看法、态度、习惯显露在问句中，或由于问题的提法不妥，而使问句具有提示性、诱导性这些问题都应该尽力避免。为了保持问句的非倾向性，需要注意以下几个方面的问题◎

### 1. 避免掺杂问卷设计者的主观成分

问句中如果掺杂了问卷设计者的诸如态度、情感、愿望、看法、认识、判断等主观成分，则问句就会存在诱导性，会牵引被调查者的回答，影响调查数据的真实性。例如，“为了不影响收看电视节目情绪，您是否赞成电视节目之间插播广告？”该问句中所出现的“为了不影响收看电视节目情绪”这句话，实际上是问卷设计者个人的一种主观认识与判断，在这句话的引导下，被调查者将会有更大的可能性选择否定的答案。为了保持问句的客观性，完全可以将此句多余的话删除。

### 2. 避免采用特殊语气的措辞

问句中应尽量使用中性词语，避免使用具有价值判断的褒义词与贬义词，也要避免使用具有某种感情色彩的词语，问句的表意不能具有主观上的倾向性，否则会引导被调查者给出具有偏差的回答。例如，“您如何评价购买高档消费品这种奢侈行为？”该问句使用了“奢侈行为”这个贬义词，此时，被调查者将会受到这种贬义表述的影响，一般会对购买高档消费品的行为给予负面评价。此时，可以将上述问句表述为“您对购买高档消费品持什么态度？”，该问句就属于中性提问，可以使被调查者自由表达个人的观点与态度。

### 3. 避免问题的从众效应

从调查心理角度而言，被调查者往往会在回答动机上存在一定程度的从众心态，如果他事先已知或感觉别人普遍对某种现象或行为具有某种评价态度时，他很可能会依照别人的态度来回答问题，从而产生所谓的从众效应，其回答就有可能产生回归偏差。例如，“人们都认为 COOLPAD 手机功能不错，您觉得呢？”，该问句就给出了人们对 COOLPAD 手机功能的普遍评价，被调查者在这种提示下，其回答很可能会有意识地选择大多数人的看法，即便其内心深处并不这么认为。所以，上述问句可以改为“您认为 COOLPAD 手机的功能怎么样？”这种中性提问。

### 4. 避免问题的权威效应

社会对某些问题比较一致的看法，权威人士或权威机构的态度，知名人士的意见，在被调查者的心目中往往具有一定影响力和权威性，容易得到被调查者的认可，这种现象被称为权威效应。权威效应的存在往往会使被调查者放弃自我态度、看法、意见等，其回答将会自觉地向权威效应回归，调查数据的真实性就值得怀疑。因此，在设计问句时应尽量避免权威效应。例如，“世界卫生组织经过二十多年的跟踪调查发现，饮酒有害健康。您是否也这么认为？”，该问句使用了“世界卫生组织”这一权威机构的研究结果，被调查者无论自身的感知如何，很容易被这种具有权威性的研究结果所左右，失去自我认知，从而我们难以获得其真实看法。如果将该问句改为“您是否认为饮酒有害健康？”，就不存在权威效应，调查结果的真实性就会提高。

## 四、问卷中问题答案的设计

问题答案的设计是标准化问卷设计的重要组成部分。问题答案的设计包括两个方面的内容，一是问题答案类型的设计，二是问题答案内容的设计。由于调查的目的、对象、内容不同，问题答案的具体内容千差万别，难以对其一一做出具体规定，但答案的类型是可以进行归纳的。所以，本教材重点介绍问题答案类型的设计方法。

### （一）问题答案设计的基本原则

问题的答案是用来供被调查者选择的项目，答案呈现的方式与答案的具体内容，都需要设计者进行科学的设计，以保证被调查者能够方便阅读、准确理解、正确选择问题的答案项目。设计问题的答案，应遵循下列基本原则。

#### 1. 穷尽性原则

穷尽性原则是指所设计的问题答案应该包括该问题全部可能的潜在答案，不能有任何遗漏。在答案有

遗漏的情况下，有些被调查者难以作选，有可能放弃回答而导致数据的缺失。但是，一个问题的潜在答案可能太多，如果全部将其一一列出往往不可能或没有必要，此时可以将一些重要的答案列示出来，而那些不重要或研究者可能并不关注的潜在答案，可以使用“其他”选项来囊括，这样就保证了答案的穷尽性。例如，如果调查的问题是被调查者的职业，根据《中华人民共和国职业分类大典》，我国职业划分为四个层次，有8个大类、66个中类、413个小类和1838个细类。显然，在设计职业选项时不可能按小类与细类的划分标准列示职业种类，而按照8个大类列示职业又显层次太粗，所以，按中类划分职业就比较合适。但中类也有66种职业类别，问卷中不可能列示出66种职业名称。为了保证答案的穷尽性，可以将研究者重点关注的职业类别列示出来，其余的职业一并归入“其他”选项。顺便指出，答案选项一般控制在8个以内比较合适，以减轻被调查者的阅读量，方便被调查者选择。

## 2. 互斥性原则

互斥性原则是指设计的问题答案彼此之间完全不相关，互不包容。如果答案之间相互关联，不同被调查者则会根据个人的理解或主观认知进行选择，而造成选择的标准与口径不统一，降低了调查数据的标准。在后续对调查数据进行整理与分析时，有可能出现矛盾的结果，不便于对分析结果进行解释。例如，如果调查的问题是被调查者的职业，答案选项中出现了“售货员”和“商业人员”，这就违背了互斥性原则，因为售货员包含在商业人员中。此时，可以将“售货员”一项删除。

## 3. 通俗性原则

通俗性原则是指问题答案的用词与表述应通俗易懂，便于被调查者理解。问题的答案主要由词语、短句或数字来表述，也有用图形、数轴等几何方式呈现的场合。如果是使用词语或短句来表述答案，则不宜采用过于专业、生僻的词语，短句表意要通俗易懂；如果用几何的方式 来表述答案，则不宜采用过于复杂的图形或刻度不明确的数轴。例如，调查问题是询问人们购买某种商品的原因时，答案中出现了“比照集团行为促使购买”这一短句选项，对于大多数被调查者来说，可能并不知晓“比照集团行为”是什么意思，从而难以选择。实际上，“比照集团 行为”是一个社会学概念，其含义是指人们的相互攀比心理。设计答案选项时，要尽量避免出现类似的情况 r

## 4. 准确性原则

准确性原则是指问题答案的内容能够准确反映调查问题的内涵，答案必须与问题相匹配，不能答非所问。问卷设计者应该准确把握调查问题的内涵，根据初步探索所获得的信息，结合研究对象的具体特征，设计出与问题相匹配的答案选项，尤其是在潜在答案较多或者潜在答案不太明确的情况下，应力图筛选出其中主要的潜在答案。答案的设计必须针对问题的内容及调查目的，否则，脱离主题的答案就是不准确的。例如，调查的问题是人们购买某种商品的动机，如果答案中出现了“经常购买”这一选项，就不符合准确性原则，因为“经常购买”是一种行为而不是动机 c 所以，该项答案应该删除。

除了上述基本原则以外，设计问题答案时还要考虑被调查者的阅读兴趣、理解能力、回答能力和回答的方便程度，应注意答案的排列顺序，灵活运用各种不同计量层次和不同形式的量表工具设计比较科学的答案选项。能够使用计量层次较高的量表工具设计问题答案，就不要使用计量层次较低的量表工具 计量层次越高的答案，对问题的度量就越精确，对其进行数据处理与分析的方法也越多，有利于对所调查的问题进行深入的定量研究。

## (二) 封闭式问题的答案设计

封闭式问题是指事先设计好问题的备选答案，由被调查者从备选答案中按照回答要求，选择自己认同的答案。问题的回答被限制在备选答案中，被调查者只能从备选答案中进行选择，不能在答案以外回答。

封闭式问题的答案类型有多种，常用的主要是两项选择法、多项选择法、排序法、等级法、双向列联法、过滤法及比较法。

### 1. 两项选择法

两项选择法也称为真伪法或二分法，是指问题只有两个相互排斥和对立的答案供被调查者选择二答案往往表现为“是”或“否”、“有”或“没有”、“喜欢”或“不喜欢”等类似的答案组合’被调查者只能从中选择一项，做选择的特点是非此即彼。

#### 例 5.10

在本年度中，您是否购买过自己使用的电冰箱？

- ①是 ②否

该方法简单明确，回答容易，便于统计处理和分析，但不能获得深层次的信息。适用于互相排斥的两项择一式问题，及询问较为简单的事实性问题。该方法可以独立使用，也可以作甄别之用。如果是作甄别之用，这种询问方法后面常常紧跟着另外一个问题，目的是对某个特定的群体进行深入的调查。

### 2. 多项选择法

多项选择法是指针对一个问题设计了两个以上的备选答案，被调查者按规定或提示从中作选。为了遵守穷尽原则，答案中往往设置一个“其他”项。

多项选择法根据要求选择的答案多少不同，又有单项选择、多项选择和限制选择三种具体的选择类型。

单项选择型要求被调查者从多项备选答案中只选择其中一项下例就属于单项选择型□

#### 例 5.11

电视播放广告时，您通常的做法是（选一项）

- |                  |               |
|------------------|---------------|
| ①兴趣很浓，从头看到尾      | ②马上换其他频道      |
| ③开着电视去干其他事情      | ④只选择自己感兴趣的广告看 |
| ⑤虽然不感兴趣，但还是耐心地等待 | ⑥其他           |

多项选择型是让被调查者从多项备选答案中，选择自己认为合适的答案，选择数量不限：下例就属于多项选择型。

#### 例 5.12

您对广告感兴趣的原因是（可选多项）：

- |             |               |
|-------------|---------------|
| ①广告带来商品信息   | ②广告带来许多新的生活观念 |
| ③广告有较高的欣赏价值 | ④其他           |

限制选择型是要求被调查者从多项备选答案中，选择自己认为合适的答案，但选择数量有限定。下例就属于限制选择型。

#### 例 5.13

促使您购买西门子冰箱的主要原因是什么（最多选三项）

- |       |       |       |       |
|-------|-------|-------|-------|
| ①名牌产品 | ②性能良好 | ③价格合理 | ④广告宣传 |
| ⑤售后服务 | ⑥他人推荐 | ⑦外观中意 | ⑧其他   |

运用多项选择法时，要注意不要使备选答案的排列顺序存在某种规律，因为有些被调查者习惯选择第一个答案，如果排列有规律性，就会产生偏差。此外，备选答案不能太多，当样本量有限时，容易使调查结果分散，缺乏说服力。如前所述，备选答案的数量不宜超过 8 个。

### 3. 排序法

排序法又称顺序选择法或优先顺位法，是指由被调查者根据自己的经验和认知程度、专业知识以及兴趣需要等，对多个备选答案按重要性（也可以是熟悉程度、重视程度等）排序，排序法有两种方式，一是对全部备选答案排序，二是对部分备选答案排序

对全部备选答案排序是要求被调查者对全部备选答案按重要程度从高到低或从低到高的顺序进行排序。

在例 5.13 中，如果在题干后提示“请按重要程度由高到低对全部选项进行排序”，就属于对全部备选答案进行排序。

对部分备选答案排序是要求被调查者对部分备选答案按重要程度从高到低或从低到高的顺序进行排序。在例 5.13 中，如果在题干后提示“请在下列选项中选出三项，并按重要程度由高到低进行排序”、就属于对部分备选答案进行排序。

排序法不仅能够比较全面地「解被调查者对所调查问题的态度，还能够区别不同态度的重要性，增加了信息量，有利于对调查结果进行深入分析，但备选答案的个数不能太多，否则会增加顺位的难度与准确性。如果根据研究的需要，要求作选答案有先后顺序，就要使用排序法。

#### 4. 等级法

等级法是将备选答案按照强度或程度分成若干等级依次排列，要求被调查者从中选择一个答案的方法。这种答案设计方法适用于了解被调查者意见、态度，感情、情绪等的强烈程度的定序问题。备选答案由表示主观意见、态度等不同等级的词汇组成，如“完全同意”“基本同意”“无所谓”“不太同意”“很不同意”等级一般按奇数次划分，可以是三级、五级甚至七级

但等级划分不能过多，如果等级过多，被调查者难以区分相邻两个等级之间的区别，增加选择难度。从经验来看，七级以内的分等比较易于操作。等级除了可使用文字直接表达外，也可以用数字在量表上表示。

#### 例 5.14

电视广告这种形式适合于食品、化妆品这样的生活用品，而不适用于水泥、农药这样的非生活用品，您的看法是

- ①完全同意      ②基本同意      ③无所谓  
④不太同意      ⑤很不同意

例 5.14 就属于用文字表达的等级法「在设计这种等级答案时，没有特别的先验信息或特殊目的，正反等级数目应该相等，以保持答案分布的均匀性。」

上例也可以使用语义差别量表的方式来设计答案。其方法是使用一个标有刻度的数轴，数轴两端标注意义相反的两个等级词语，如一端标注“完全同意”，另一端标注“很不同意”，要求被调查者在数轴上自己认为恰当的位置做上标记即可”

#### 例 5.15

运用例 5.14 的问题，以数轴的形式设计答案如下：

完全同意 I ----- 1 ----- 1 ----- 1 ----- 1 很不同意

这种以数轴的方式设计的答案，也可以在数轴上标上数字，如果“1”表示“完全同意”，“5”表示“很不同意”，则上述答案形式可以改为：

12                    3                    4                    5  
完全同意 I ----- 1 ----- 1 ----- 1 ----- 1 很不同意

这种答案设计方法，实际上就是让被调查者在完全同意与很不同意两者之间进行选择，如果被调查者在数轴上的标记越靠左，则表示同意的程度越高；反之，越靠右，则表示不同意的程度越高。数据汇总时，分析者可以将标记转化为数字，便于数据的量化处理。

如果在一份问卷中设计有一连串的等级回答方式的问题，最好不要分开选择，可以采用列表的方式设计答案。调查的问题放在列表中的横栏位置，等级词语放在纵栏位置，被调查者只需要在纵横相交所形成的方格中进行勾选即可。这样设计既可以节省问卷篇幅，又便于被调查者阅读与选择。

#### 例 5.16

对下列关于彩虹牌智能手环特征的陈述，您的看法如何？（在相应的方格中画“V”）

彩虹智能手环特征	完全同意	比较同意	一般	不太同意	很不同意
品牌知名度高					
使用方便					
价格合理					
外型轻巧美观					

被调查者在选择每一项陈述的答案时只能作单项选择，不可能出现复选情况，否则就是无效回答。

等级法还可以用于表示事物大小、多少之类的定距与定比问题的答案设计，所设计的答案往往用表示事物数量大小或多少的数值区间来表示，数值区间由低到高排列成序，形成数值等级，反映事物数量上的差异状况。

#### 例 5.17

您的年龄是多少？

- ①21~30岁      ②31~40岁 ③41~50岁  
④51~60岁      ⑤61岁及以上

例 5.17 通过五组数值区间，对目标群体的年龄进行调查。五个年龄区间就是年龄的五个等级，这种等级的划分可以根据研究的具体目的，设置不同的数值区间”

#### 5. 双向列联法

双向列联法是将两类不同的问题综合在一起，通常用表格的形式来表现。表格的横行是一种类型的问题，纵列是另一类型的问题。可以反映两方面因素的综合作用，提供单一类型的问题所无法提供的信息，而已可以节省问卷的篇幅。但被调查者需要较多时间的思考与填答，故在一份问卷中，此类问题不宜过多。

#### 例 5.18

请您对下列各种媒体的广告进行评价（在相应的方格中画“V”）

特征描述	电视广告	报纸广告	广播广告	杂志广告
有较高的欣赏价值				
有较高的可信度				
容易给人留下深刻印象				
不感兴趣				

被调查者在选择答案时，无论在横向还是纵向上，均可以进行多项选择，也可能有弃选的情况。弃选情况的出现，只能说明答案设计存在问题，研究假设的建立缺乏理论与事实依据。研究者在设计问卷前必须做好充分的初步探索工作。

#### 6. 过滤法

过滤法也称条件法，是指通过一个前奏问题的答案，来筛选被调查者的方法。有时，研究者仅对某一类群体的信息感兴趣，就需要对调查对象进行过滤、筛选掉不需要的部分；然后针对特定群体进行提问，这时需要两个或两个以上的问题相连接，第一个问题起过滤作用，从第二个问题开始是具体的调查内容。运用过滤法设计答案时，设计者首先必须明确对于具有何种属性的被调查者需要进一步调查，要将被调查的这种属性设计在答案中。

#### 例 5.19

研究者为了深入研究对广告感兴趣的群体，就可以通过下列的设计来筛选被调查者。

Q1. “您平时对广告感兴趣吗？”

- ①很感兴趣 ②较感兴趣 ③一般 ④不感兴趣 ⑤很不感兴趣

如果选择①或②，请回答下列问题：

Q2. “您对广告感兴趣的原因是”

- ①广告给您带来商品信息      ②广告带来许多新的生活观念  
③广告有较高的欣赏价值      ④其他

上例中第一个问题中的答案①和②，都是用来筛选对广告感兴趣的被调查者的，如果将

答案设计成感兴趣、一般、不感兴趣，也能达到筛选目的。

### 7. 比较法

比较法是把若干可以比较的同类事物或现象，整理成两两对比的形式，要求被调查者进行 两两比较，按要求选择其中一种事物或现象的方法。在市场调查中，该方法适用于对消费者的 偏好进行调查。运用比较法要考虑被调查者对所要比较的事物或现象是否熟悉，否则会导致 选择空项。

#### 例 5.20

请比较下列每一对不同品牌的洗发液，哪一种您更喜欢？（每一对中只选一个，并在口内 画 “W”）：

- |                              |                              |
|------------------------------|------------------------------|
| <input type="checkbox"/> 海飞丝 | <input type="checkbox"/> 潘婷  |
| <input type="checkbox"/> 潘婷  | <input type="checkbox"/> 飘柔  |
| <input type="checkbox"/> 飘柔  | <input type="checkbox"/> 威娜宝 |
| <input type="checkbox"/> 威娜宝 | <input type="checkbox"/> 花王  |

上例中共涉及 5 个品牌的洗发液，如果某个被调查者对其中一个品牌不熟悉，就没法进行 比较，也不能做出有效选择。所以，用来配对比较的事物或现象都必须是被调查者熟悉的。否 则，该方法不能采用。

### （三）开放式问题的答案设计

开放式问题是指对问题的回答不提供任何具体的答案，也不规定回答的范围，由被调查者 根据自己的想法与态度自由填答。开放式问题的这种回答方式被称为自由回答法。虽然开放 式问题一般不需要设计问题的答案，但被调查者必须给出答案。为了使被调查者尽可能准确 回答开放式问题，问卷设计者在设计问题时，要充分考虑被调查者回答的方便性，为其自由回 答提供空间。

开放式问题的答案设计主要是设计预留给被调查者回答问题的书写空间 e 由于开放式问 题的回答属于非标准化回答，回答内容可能千差万别，回答的篇幅可长可短，这将给事后进行 数据的编码、处理与分析带来较大的麻烦。所以，对开放式问题的回答要进行适当的数量上的 控制 Q 其方法是在问题后适当预留用 于填写回答内容的下划线行数。

#### 例 5.21

您喜欢这款商品的原因是：\_\_\_\_\_ 0

#### 例 5.22

您喜欢这款商品的原因是：

上述两例所预留的回答空间都是不合适的。例 5.21 所留下划线行数过少，且一行的下划 线长度也太短，被调查者即便只是回答喜欢这款商品的主要原因，预留空间也不够，将会限制 被调查者的思考与回答。例 5.22 所预留的下划线行数过多，一是给被调查者带来思考与回答 压力■影响其回答情绪与兴趣；二是给研究者带来数据处理的一些困难。一般而言，凡属于了 解被调查者意见、态度、建议、看法、感受等主观层面的资料时，下划线的行数以控制在三行以 内为宜。这样既可以让被调查者比较充分地发表看法，又不至于长篇大论。

在混合型回答方式下，预留的下划线一般只需一行，且不必过长。例如，对多项选择法中的“其他”项，如果研究者对“其他”项的具体内容感兴趣，可以再将“其他”改为“其他（请注 明）\_\_\_\_\_”即可。此时，下划线的长度可参考其他具有明确内容选项的文字长度，一般 不会太长。有时，下划线的长度可能只需要两三个字符的长度。例如，“您的年龄:—”这个开放式问题所预留的下划线有三个字符长度就足够了，

这是由年龄这个数字的位数决定的。

在某些面访调查中，并不采用被调查者自填问卷的方式，而是采用由调查员提问、被调查者回答、调查员记录的方式搜集调查数据。其中，可能有些开放式问题采用回忆法进行调查，了解被调查者对不同事物（如商品质量、品牌）印象的深浅。在问卷设计时，就不仅要考虑回忆的结果与秩序，还要考虑回忆的速度。在这种情况下，问卷中问题最好设计成用于记录回忆过程与结果的表格，以便于调查员进行记录。

**例 5.23**

请您尽可能回忆出最近在电视广告中出现的手机品牌

回忆秩序	1	2	3	...	N
品牌名称					
回忆时间					

针对上述开放式问题，答案要件包括回忆秩序、品牌名称，回忆时间属于作业记录。此时，所搜集的开放式问题的数据具有标准化特点，有利于分析与比较不同手机品牌的广告传达效果。

需要说明的是，在某些特定情况下，开放式问题有时也需要设计答案。例如，运用联想技法进行调查时，就可以给被调查者提供联想的词语供其选择。

开放式问题的回答采用自由回答法，其最大优点是灵活性大、适应性强，可以让被调查者自由地发表自己的看法。若问卷设计者对某些问题不太熟悉，很难设想调查的结果会是什么，也很难事先设计好固定答案时，可以考虑采用自由回答法。另外，在某些情况下，采用自由回答法有助于问卷的简化。例如，当某个问题的潜在答案太多时，如果在问卷上一一列举，就会占用大量版面，也会增加被调查者的阅读量；又如，有些问题比较复杂，若分解成若干个问题并相应设计众多答案，就会导致问卷过于烦琐。此时，若采用自由回答法，就会使问卷变得简洁。

自由回答法也存在明显的缺点，主要表现在两个方面。其一，回答资料的标准化程度低，不易进行统计处理与分析。由于是自由回答，无固定答案，回答内容与文字表达可能千差万别，会大大增加数据归类处理的工作量。有些回答可能答非所问，或不够准确，或不够明确，回答结果可能没有什么价值。其二，无固定答案使被调查者回答问卷的难度增大。除了一些客观性质的开放式问题，对于绝大多数的具有主观性质的开放式问题，采用自由回答法对填答者有较高的文化素养的要求，而且要花费较多的填答时间。由于没有固定答案做参考，有时会使被调查者有不知所措的感觉。在这些情况下，被调查者有可能失去填答兴趣而放弃回答，从而影响了问卷的回收率与有效率。鉴于自由回答法的这些缺陷，不提倡在标准化问卷中过多使用这种回答方式。一般而言，开放式问题的数量控制在三个以内为宜。

需要指出的是，为了便于被调查者准确理解与回答问题，开放式问题的提问内容不能太广泛，太空洞，否则会使被调查者无所适从，不知道该如何回答。

**例 5.24**

问题：您认为现有共享单车的运营模式有哪些经验和教训？

这个问题的内容就太广泛了。不同品牌的共享单车的运营模式不同，经验和教训也不一样；每个品牌的共享单车的运营模式存在多方面的经验和教训，被调查者不通过认真调查、分析与总结，是很难作答的。对于此类问题，被调查者放弃作答可能性较大。设计开放式问题，一定要考虑被调查者的回答能力、回答兴趣，所提问题应该尽可能具体、简单，使被调查者稍加思考即可作答。

## 五、问卷的版式设计

问卷的版式设计是指将已经设计好的问题进行适当的排列组合，并确定问题在问卷版面上的位置和形式。版式设计也是问卷设计的重要内容，应该给予足够的重视。同样内容的问卷，采用不同的版式，会带给被调查者带来不同的感受，对调查过程产生不同的效果，也给研究者进行数据处理与分析带来不同的效率。

### (一) 问卷版式设计的一般要求

问卷版式设计的总体要求是简洁、整齐、美观、清晰、便于阅读与填答。一般而言，问卷版式设计要考虑以下问题。

#### 1. 问卷结构的整体布局

一份问卷由标题、说明词、正文、编码、被调查者基本情况、结尾等六个部分组成。问卷结构的整体布局，原则上根据这六个部分的先后顺序进行排列。

标题一般置于说明词或正文的前面，单独一行居中的位置。市场调查对象及其核心概念、时空特征等应包含其中，表述应简明新颖。

说明词根据其内容的简繁程度，有两种不同的处理方式。如果说说明词内容比较简单，就直接将其排列在问卷卷首；如果说说明词内容比较复杂，就可以考虑将其单独作为问卷的一封附信，独立于问卷的其他部分。

正文部分是问卷的主体，由问题与被选答案所组成。对这一部分编排的总体要求是便于被调查者阅读与填答。问题的排列可以分模块，每一模块内的问题排列要有一定逻辑。必要时需要提示被调查者的回答方式，提示要明确、醒目。

编码包括问题的编码与答案的编码，随同问题与答案一起编排。问题的编码一般使用“Q1、Q2……”；对于封闭式问题，答案的编码一般使用阿拉伯数字标示，如“①、②……”或“1、2……”。对于开放式问题，因为无固定答案，设计问卷时事先不作编码。

被调查者基本情况既可以排列在结尾部分的前面，也可以排列在说明词之后。一般来讲，被调查者基本情况不会涉及任何敏感性问题，大部分被调查者不会介意这些问题。所以，被调查者基本情况排列在前还是在后，并不是一个本质性问题。但是，可能还是有少数被调查者觉得，一开始就询问其个人的一些基本情况，可能心存戒备，产生某种抵触情绪。考虑到这种因素，将被调查者基本情况排列在问卷的结尾之前可能更为妥当。

结尾部分的内容一般是作业记录和致谢语。致谢语采用简明的述语表达对被调查者完成调查问卷填写的感谢与支持。

#### 2. 问卷编排的一般版式

##### (1) 问卷的整体格式

问卷的整体格式要求版面活泼、字体字号选用得当、行间距选择合理、问卷各组成部分之间层次清晰、填答引导与说明醒目、填答方便，设计者在排版时可以根据这些要求，进行灵活处理：比如，问卷的标题、各模块的标题均可以使用加粗字体，行间距可选择1.25倍，问题与答案可以分别采用不同字体与字号加以区别等。问卷的整体格式安排没有严格意义上的统一规定，究竟以什么具体格式来编排问卷，很大程度上还取决于设计者的审美水平。

## (2) 问题的编排版式

问卷中所有的问题排列要整齐，题项前的编码上下对齐。文字可采用五号宋体或其他视觉效果较好的字体。若是封闭式问题，对于统一的回答方式已在填答说明中作了统一规定。但有些问题具有某种特殊性，需要在题干后进行回答方式的提示。例如，使用等级法回答方式时，如果采用了数轴的形式让被调查者在数轴上作标记，题干后就必须说明如何作标记。在题干后面往往要预留填写答案的括号或方框 G 如果预留的是方框，方框的数量应该是所回答问题答案编码的最大位数与应选答案个数的乘积 C

### 例 5.25

对于“您放弃使用智能手环的原因是（可多选）”这个问题，可以采用两种版式：

版式一：Q10. “您放弃使用智能手环的原因是（可多选，请将所选答案的编码填入题干后的‘□’中）□□□□□□

- ①监测数据不准      ②功能不实用      ③用的人太少
- ④与手机关联使用      ⑤不够防水      ⑥其他（请注明）

版式二：Q10. “您放弃使用智能手环的原因是（可多选，请在所选答案前的编码上画“V”）

- ①监测数据不准      ②功能不实用      ③用的人太少
- ④与手机关联使用      ⑤不够防水      ⑥其他（请注明）

由于问题已提示可以多选，答案编码的最大位数是一位，可选答案可能有六个，则版式一的题干后就需要预留六个方框 G 如果该问题限选三项，就只需要留三个方框 G 从方便被调查者填答角度评价，版式二只需要被调查者在选项前的编码上画“V”，较版式一更方便。但版式二不利于数据的录入。究竟采用何种版式，完全由设计者的个人偏好来决定。

## (3) 答案的编排版式

对于封闭式问题，设计有固定的被选答案。答案的排列要整齐，答案与答案之间要保持适当的距离，答案前的编码上下要对齐 G 为了提高视觉效果，答案可以采用与问题不同的字体，字号最好不要大于问题的字号。如果答案较多而采用分行表达，则行间距应不小于 1.25 倍。一般来讲，答案前最好有编码，编码最好用阿拉伯数字标示，这样有利于减轻事后编码的工作量。有些问题中，由于问题的答案并不确定，难以事先编码。例如，采用等级法回答方式中的数轴标示方法，就不需要编码。对于开放式问题，如前所述，所预留的用于回答问题的下划线长度要适当。

### 例 5.26

承例 5.25 的版式一，将问题与答案编排如下：

Q10. “您放弃使用智能手环的原因是（可多选，请将所选答案的编码填入题干后的“□”中）□□□□□□

- ①监测数据不准      ②功能不实用      ③用的人太少
- ④与手机关联使用      ⑤不够防水      ⑥其他（请注明）\_\_\_\_\_

此例中，问题的文字采用小五号宋体，答案的文字采用 11 号楷体，就显得比例 5.25 中的版式一更具有层次感，不会导致被调查者的视觉疲劳。混合回答项中所预留的下划线长度也比较适当，与其他选项的文字长度相匹配，

## (4) 图表的编排版式

问卷中无论是问题还是答案，有时会使用一些图形或表格，例如，投影技法中的图片理解测验法，就需要设计一幅与调查内容有关的漫画，由被调查者根据漫画的理解来回答相关问题。此类漫画往往不要画得过于清晰，画中人物和事物只需要有大致轮廓即可。对于涉及人物的漫画，人物的表情最好不要画得太明确，人物之间的关系可以根据调查目的与内容来设计。总之，此类漫画以不给被调查者带来某种回答暗

示为宜。又如，等级法回答方式中所使用的数轴应该标有刻度，以便于被调查者在数轴上比较准确地标示自己的态度与看法。对于问题与答案中出现的表格，可按照统计表的制作要求进行设计。

## （二）问题的排列顺序

问卷中问题的排列是有一定技巧的。心理学研究表明，调查问题的排列顺序可能影响被调查者的情绪。同样的调查问题，如果排列顺序合理、恰当，可以及时有效地获得调查数据。反之，会影响被调查者作答，影响问卷的回收率及调查数据的质量。设计问卷时，设计者应该始终站在被调查者的角度，考虑被调查者的思维习惯，合理排列问题的顺序。问题的排列应该遵循下列基本原则。

### 1. 问题的排列应具有逻辑性

问题排列顺序的逻辑性是指问题的编排应该与调查内容的逻辑结果相匹配。一份问卷的内容往往会划分成几个模块，每个模块中又包含着若干调查问题。模块之间以及模块内各问题之间是什么逻辑关系，设计者必须清楚。问题的排列顺序，就是基于这种模块之间及模块内各问题之间的逻辑关系来决定的。例如，在调查用户满意度的问卷中，应该包括用户的期望、用户感知的质量、感知的价值、满意度、抱怨及被调查者背景这六个模块。根据满意度内涵及各模块之间的内在关系来看，六个模块就应该按照上述这种顺序进行排列。

### 2. 问题的排列要考虑问题的性质或类别

按照调查内容的不同，所有的问题可分为三大类，即了解被调查者基本资料的问题，了解行为方面的问题和了解态度方面的问题。在每个模块内考虑问题的排列顺序时，最好将同一类问题排在一起。可以考虑先排列行为调查方面的问题，再排列态度调查方面的问题。若采用邮寄、电话调查等方法，则可把被调查者基本资料放在问卷的后面；若采用面访法进行调查，则被调查者基本资料（其中的敏感问题除外）可放在前面。

### 3. 问题的排列应该先易后难

从调动被调查者回答兴趣的角度考虑，同性质的问题应该按先易后难的顺序排列。被调查者首先回答一些比较容易回答的问题，有利于提高其回答问题的积极性，促进其完成整份问卷的填答。如果一开始就回答一些较难回答的问题，容易降低被调查者的回答兴趣。一般而言，有关公开的事实或状态的描述性问题要相对容易一些，被调查者不需要做什么思考即能回答，可将这类问题列在所属模块前面的位置；而有关看法、意见、态度、情感等主观判断性问题，被调查者需要动脑筋思考后才能作答，因此应该放在所属模块后面的位置。

### 4. 问题的排列需要考虑时间要素

从时间角度考虑，一份问卷可能包括好几个时间段的问题，可能既涉及近期的事情（最近一周、最近一个月），也涉及远期的事情（前几个月、上一年度等）。由于近期的事情容易回想，便于作答，因此可以将此类问题置于所属模块靠前的位置。对于远期的事情，由于记忆容易受到干扰，回忆难度较大，可以将此类问题置于所属模块靠后一点的位置。例如，可以先问“您现在使用的是什么品牌的洗发液？”，然后再问“您最近一年以来使用过哪些品牌的洗发液？”有时，为了帮助被调查者回忆过去的事情，问卷设计者可以采用一些提示办法，如运用时间（如重大节日）提示、事件（重要事件）提示、列单提示等方法。

### 5. 敏感性问题与开放性问题可以置于问卷的后面

敏感性问题往往是被调查者不愿回答的问题，应该放在所属模块的后面位置。因为这类问题容易遭到被调查者的拒答，从而影响后继问题的回答。如果将此类问题置于后面，即使这些问题被拒答，对之前问题的回答仍有利用价值。开放式问题一般需要被调查者进行一些思考，需要花费较多的时间作答，这些都是被调查者不太愿意做的事情。如果将开放式问题置于问卷的前面，会对被调查者的回答情绪产生负面影响。开放式问题置于问卷的后面，即便被调查者没有作答，也不影响前面问题的回答及问卷结果的利用价值。

值。需要指出的是，有些关于事 实性的开放式问题比较简单，例如“您的年龄：一”，这类开放式问题即便置于问卷的前面也 不会对被调查者产生回答的负担，不会影响后续问题的作答。

## 6. 要注意问题编排的顺序效应

问卷中可能会出现一些相关联的问题，如果被调查者对前面问题的回答影响了对后序问 题的答案选择，这时就会产生问题编排的顺序效应。顺序效应往往会对被调查者回答的真实 性产生负面影响，因此要尽可能避免。

### 例 5.27

下列两个问题属于关联性问题：

Q1. 您是否认为购买高档消费品是一种奢侈行为？

①是      ②否      ③不知道

Q2. 如果条件允许，您是否有购买高档消费品的打算？

①是      ②否      ③不知道

对于第一个问题，如果被调查者选择了①，那么在回答第二个问题时，即便他在条件允许 时会有购买高档消费品的打算，也很可能选择②或③，因为他并不希望别人认为他是一个有奢 侈动机的人。此时，对第二个问题的回答就失真了。

为了避免这种问题编排的顺序效应，问卷设计者有必要对问卷中出现的一些关联性问题 进行评估，判断是否存在顺序效应：如果确定存在顺序效应，就应考虑将这些关联性问题 错开。

## 六、问卷的评估

问卷初稿设计完毕之后，并不能立即用于际调查，还要对问卷进行评估。由于问卷的设 计者可能对所调查研究的现象或问题并不十分熟悉，对研究假设和相关的研究理论的理解并 不深刻，考虑问题可能不够全面、不够细致，还可能缺乏问卷设计的实践经验，这些原因都可能 使得问卷初稿存在意想不到的一些问题。所以，有必要对问卷初稿进行评估。只有经过评估 且被认为质量合格的问卷，才能用于正式调查。

### （一）问卷评估的主要内容

问卷是用于对被调查者进行各种测量的工具，由被调查回答问卷中的各种问题，以搜集研 究所需的调 查数据。问卷评估的主要内容包括以下几个方面。

#### 1. 问题是否具有必要性

调查研究的目的决定调查的内容，调查的内容决定了需要调查哪些问题。评估时需要对 问卷初稿中的问题逐一进行审查，判断问题存在的必要性。对与研究目的不相关的或难以搜 集数据的问题，应该加以剔除。

#### 2. 问题是否存在遗漏

在设计问卷的过程中，需要根据研究的目的、研究的假设来提出问题一在设计问卷问题 时，如果对研究目的理解不透彻或对研究假设所涉及的概念缺乏系统梳理，可能会导致设计出 来的问题不够全面，存在遗漏现象。此时，有必要增设题项。

#### 3. 问题与答案的表述是否准确、通俗

问卷的主题部分由问题与被选答案构成，问题与答案的表述是否准确、通俗，直接关系到 被调查者能否准确理解问题的内容，能否给出正确的回答’另外，被选答案有时以所采用的量 表所固有的格式表述出来，需要评估量表的复杂程度，被调查者是否具有使用该量表回答问题 的能力。对于表意不清、难于理解的问题及答案，要根据前面所述的问题表述及答案设计的要 求加以修改。

#### 4. 回答导引是否明确、醒目

问卷中所有的问题均具有回答要求，对于统一的要求一般在问卷的说明词中进行了规定，要评估这些规定是否明确、清晰。对于某些存在特殊回答要求的题项，例如跳答题、排序题、数轴测量题（语义差别量表），通常有专门的回答导引 I，这些导引词务必醒目且准确，便于了解。否则，引词需要修改。

#### 5. 问卷的版式设计是否符合要求

问卷的版式设计直接关系到被调查者能否对调查产生兴趣、阅读问卷是否存在困难、回答问题是否方便等。评估问卷初稿还需要根据问卷版式设计的要求对问卷进行评估，对不符合要求的版式设计要加以优化。

#### 6. 问卷的构成要件是否完善

一份完整的问卷一般由标题、说明词、正文、编码、被调查者基本情况、结尾等六部分内容组成。评估问卷初稿还需要逐项审查各个要件，对于遗漏的要件要加以补充。

#### 7. 问卷的长度是否合适

无论采用哪种方法进行调查，问卷的长度都应该合适，篇幅过长的问卷会增加被调查者的填答负担，容易导致厌烦情绪而影响作答。问卷的长度没有固定的标准，但应保证不占用被调查者过多的时间为宜，一面访问卷的长度以被调查者能够在 10 分钟之内完成回答为宜，电话调查问卷则应尽可能使被调查者能够在 2~3 分钟内完成回答为宜，邮寄调查问卷与在线调查问卷的篇幅可以适当地长一些，因为对于此类问卷调查，被调查者可以分多次来完成回答。

### （二）问卷评估的主要方法

#### 1. 专家审查

专家审查就是组织有关的专家、学者或具有丰富问卷调查经验的研究人员，对问卷初稿进行审查与评估。专家审查主要是一种技术性审查，其审查内容主要是审查问卷设计的整体构思是否合理，问题的表述是否准确、清晰；问卷的版式设计是否美观，问卷中量表的使用是否科学等。专家审查可以采用邮寄调查与小组座谈两种组织形式。邮寄调查需要给所邀专家邮寄问卷初稿，既可以邮寄纸质版的问卷，也可以通过电子邮件发送电子版问卷。必要时应该随问卷初稿给专家邮寄或发送调查方案，以便专家审查问卷与调查方案之间的匹配性。

#### 2. 预调查

如果正式调查属于大规模的问卷调查，涉及的范围广、被调查者人数多，此时对问卷的审查必须慎之又慎，稍有疏忽，就会造成难以挽回的损失。除了需要组织专家对问卷初稿进行审查与评估外，还应该组织预调查，通过预调查发现问卷存在的问题，为问卷的修改提供依据。此处预调查与前述第二章第四节介绍的内容具有相似性，只是其关注重点不同。问卷内容的预调查可以单独进行，也可以并入整个调查方案的完善环节。

在问卷测试的预调查过程中，要重点了解被调查者对问题是否完全理解、问卷的措辞和表达方式是否具有诱导性、问卷的长度是否合适、问卷的结构和形式是否合理、问题的编排顺序是否妥当、答案的设计是否合理。另外，通过预调查所获得的调查数据，还可以用于检验问卷的信度与效度。这方面的检验与下一章中所介绍的量表的信度与效度检验是一致的。

通过问卷初稿的审查与评估，发现问卷存在的各种问题，并对问卷进行认真的反复修改后，就形成了正式调查中可使用的问卷终稿。

## 第四节

### 网上创建在线问卷

网上调查问卷设计具有其独特性，因而专设一节进行介绍」

网上调查问卷的创建有两种途径，一是网下自建调查问卷文本，二是网上创建调查问卷 Q 网下自建调查问卷是指在正式实施网上调查之前，预先设计出用于网上调查的问卷文本 在问卷经过预调查以及信度与效度检验后，对于符合质量要求的问卷就可以在网上发布。为了征得更多的网民参与调查，可以同时发送邀请广告，并在其中说明调查目的及有奖填答 事项。

网下自建问卷的基本过程与要求与前面所介绍的问卷设计规范并无差别。但是，因为所设计的问卷是适用于网络调查的，需便于网民填答与数据回收，因而在具体的设计过程中也会有一些特别的要求。例如，需要插入选项按钮，编写问卷接收文件等。但这些设计比较麻烦，建议利用专业调查平台系统实现调查问卷的设计，并通过平台生成问卷链接、邀请邮件、嵌入 到网站等方式发布问卷。

网上创建调查问卷是指通过网上调查平台中的问卷设计功能来实现问卷的设计。有许多 网上调查平台为用户提供了问卷调查系统，其中包括功能强大的问卷设计模块。用户只需要 注册登录至该平台，即可运用其中的问卷设计模块进行问卷的设计。研究者在上网设计问卷 之前，首先要将调查问卷的底稿设计出来，在经过质量评估合格后，再上网制作适合于网上调 查的电子化问卷。

以问卷星为例，该调查网是一个功能强大的问卷调查系统。问卷星是一个专业、无限制的 免费在线问卷调查、测评、投票平台，专注于为用户提供功能强大、人性化的在线设计问卷、不 限题目数，不限答卷数；支持分类统计与交叉分析，可供免费下载报告和原始答卷；完美支持手 机填写、微信群发。该平台具有快捷、易用、低成本的明显优势，已经被大量企业和个人广泛使 用。研究者欲进入该调查网，首先要进行注册，注册成功后即可使用该调查网进行问卷设计 下面介绍利用问卷星平台制作问卷的要领。

#### 一、注册

进入 <https://www.wjx.cn>, 打开问卷星首页，点击注册，成为问卷星的用户。设定用户名、 密码，并输入常用的电子邮件地址。在用户类型一项中选择 “免费版”，当然根据自身需求，也

可以选择支付一定的费用选择“专业版”和“企业版”。填入验证码，并点击创建用户按钮。之后问卷星会向你注册的邮箱发送一封验证邮件，进入你的邮箱确认了邮件内容后，就完成了注册：注意，一个邮件地址只能注册使用一次。

## 二、问卷设计方式选择

首先登录系统，点击页面右上方的“设计新问卷”按钮。问卷星提供了两种设计问卷的方式：一是利用模板创建；二是利用文本创建。

### （一）模板创建问卷

问卷星系统中内置了一些常用的问卷类型，如员工满意度调查、就业情况调查等问卷，如果问卷性质和目标比较一致，就可以利用模板创建问卷。如果没有相似的问卷，需点击“使用空白模板”使用空白模板设计一份新问卷需要进行问卷基本信息的设置：这些设置包括问卷名称、问卷说明、主题、语言、截止日期、分类、问卷公开级别、问卷密码保护、结果公开级别，设置完成后点击“下一步”按钮开始对问卷题项进行设置。

### （二）文本创建问卷

如果已经在Word等文字编辑软件中设计好了问卷，可以采用文本创建问卷的方式，将Word中的内容直接复制到问卷星所提供的文本框中。然后点击“生成问卷”的按钮，即可生成一份问卷星的问卷。生成问卷后，可以修改题目属性，例如为选项增加分值、设置跳题逻辑、引用逻辑、转换题型等。

## 三、问卷题项设计及有关提示

问卷星提供了单选题、多选题、测评与测试、矩阵题、主观题等多种题型。

### （一）单选题设计

鼠标指向单选题，选择具体题型，如列表单选、组合单选、下拉框、投票单选等。我们以列表单选为例，介绍设计单选题的主要流程。单击列表单选，出现本题的编辑框。然后按下列流程操作：

1. 输入问题的题干，例如，在此输入问题“您的学历”。
2. 输入问题的选项，如果有多个选项，请点击选项后的加号。如果该选项允许参与者填空，请在选项后的允许填空一列画“V”。
3. 系统默认所增加的每道题都是“必答题”，可以取消“必答题”的设置，将该题改为“选答题”。
4. 组合单选是默认最后一个选项允许参与者填空；而下拉框则是将选项以下拉框的方式进行呈现。
5. 设计完成后，点击“完成”按钮，完成该题编辑。如果还需要继续编辑，只需将鼠标移动到该题上方，再点击“编辑”按钮即可，如果现有多题，也可以点击“上移”和“下移”按钮改变题的顺序。同时也可以点击“复制”和“删除”按钮执行相应的操作。

### （二）多选题设计

鼠标指向多选题，选择具体题型，如列表多选、组合多选、投票多选。我们以组合多选为例，展示设计多选题的主要流程，单击组合多选，出现本题的编辑框。然后按下列流程操作：

1. 输入问题的题干，例如，在此输入问题“您购买手机时主要考虑哪些因素”。
2. 输入问题的选项，如品牌、价格等，并保留最后一个选项“其他”。由于选择的是“组合”

多选”，默认最后一个选项为填空选项，允许参与者手动输入，如果取消“允许填空”，则该题变成“列表多选如果还有其他选项，请点击选项后的“删掉选项”。

3. 设计完成后，点击“完成”按钮，完成该题编辑。如果还需要继续编辑，只需将鼠标移动到该题上方，再点击“编辑”按钮即可。如果设计多题，也可以点击“上移”和“下移”按钮改变题的顺序。

4. 设计完成的该题，也可以通过题型编辑菜单中的“转换题型”功能快速地转换题型，如转换为列表单选题、下拉框单选题以及问答题等。

### （三）测评与测试设计

鼠标指向测评与测试，选择具体题型，如量表题、评分单选题、评分多选题、测试单选题、测试多选题。该题型适合于态度测量问题，根据采用的量表不同而选择不同的回答方式。例如，如果采用的是语义差别量表可选择“高级题型”中的“滑动条”这一类题型的选项都有可设定分值或设为备选答案。这里重点介绍量表题的设计流程：单击量表题，出现本题的编辑框，然后按下列流程操作：

1. 输入该题的题干。例如，在此输入问题“您对目前所使用手机的满意态度”。

2. 量表题默认本题选项为五个级别，并且选项文字为“很不满意”“不满意”“一般”“满意”“很满意”。可以修改选项的描述文字，或者重新设定每个选项的具体分值。

3. 设计完成后，点击“完成”按钮，完成该题编辑。

4. 问卷星中的多种题型都可以设置跳转逻辑。如果勾选“无条件跳题”，则参与者无论选择哪个选项都将跳转到设定的题作答。

5. 同时也可以设置不同选项跳转到不同的题作答，例如，选择“很不满意”则填写第6题，选择“很满意”则跳转到第7题。做这样的设置需要勾选选项后的“跳题”，同时设置跳转的具体选择。

### （四）矩阵题设计

鼠标指向矩阵题，选择具体题型，如矩阵单选题、矩阵多选题、矩阵量表题、矩阵文本题等，该题型适合于双向列联表的设计。下面我们将以矩阵单选题展示该题型的设计流程。点击矩阵单选题，然后按下列流程操作：

1. 输入该题的题干。例如，在此输入问题“请为下列品牌的手机打分”。

2. 矩阵单选题默认左行标题为两个，也可以通过换行增加标题。该题型默认本题的选项评价有五个等级，并且选项文字为“很不满意”“不满意”“一般”“满意”“很满意”。你可以修改选项描述文字，也可以增加或调整选项内容。

3. 设计完成后，点击“完成”按钮，完成该题编辑。

### （五）主观题设计

鼠标指向主观题，选择具体题型，如单行文本、多行文本、数值、日期、省份、省市区、地址等。单行文本和多行文本只是在答题框的大小设置上有所区别，而其他的选项则是在答题内容的格式上有所区别。下面以单行文本题展示该题型的设计流程：点击单行文本，出现本题的编辑框。然后按下列流程操作：

1. 输入该题的题干，例如，在此输入问题“你的姓名”。

2. 根据答题内容的多少设置答题框的高度，如1行、2行、3行等。

3. 设计完成后，点击“完成”按钮，完成该题编辑。

### （六）其他编辑内容提示

除上述介绍的功能外，问卷星提供了更多的题型和分段设置功能。

1. 分段。共有四种选择：分页栏、信息栏、分割栏、甄别栏等。

2. 更多题型中提供了排序题、表格题、比重题、滑动条、上传文件等功能
3. 常用题库中提供了更为丰富的题型选择。

### （七） 其他注意事项提示

1. 编辑过程中，问卷星每隔几分钟会自动对问卷进行保存，如果需要手动保存，请点击左上角的“保存”按钮’
2. 在编辑过程中，可以点击“预览问卷”按钮预览设计效果。
3. 如果问卷设计完成，则需点击页面上方的“完成编辑”按钮完成问卷编辑工作。

## 四、发送问卷

点击页面上方的“完成编辑”按钮，完成问卷的编辑工作，进入发送问卷环节。

### （一）发布问卷操作设计

在向其他人发送问卷之前，需要进行发布问卷的操作。发布问卷就是将已经编辑完成的问卷发布到网络上，使其有一个唯一的访问地址。

如果发送前需要继续编辑问卷，可以点击“继续编辑问卷”链接，返回到问卷编辑页面。点击“可选操作”链接，展开其他选项，可以有3个选择：

1. 测试填写问卷。通过试填检查问卷逻辑是否正确。
2. 修改问卷设置。设置完成信息、防重复填写、公开级别等属性。
3. 设置问卷外观。可以选择你喜欢的问卷背景与字体样式等。

如果问卷编辑完成，可以点击“立即发布问卷”按钮发布问卷。此时页面会有这样一句提示，点击“发送问卷”链接。

### （二）发送与回收答卷的途径

问卷星提供了多种发送问卷回收答卷的途径，例如直接获得问卷链接、申请推荐、样本服务、互填问卷、邀请邮件、邀请短信以及嵌入到网站等方式。这里重点介绍问卷链接、邀请邮件以及嵌入到网站三种。

1. 问卷链接。点击问卷链接后，页面会出现链接地址信息，任何人通过该地址即可访问到本问卷。
2. 邀请邮件。利用邀请邮件，可以直接通过问卷星向参与问卷调查的人发送问卷的链接信息。此环节重点是需要输入收件人地址，点击“直接输入邮件地址”链接，输入收件人地址和姓名，需要注意的是，地址和姓名之间需要用英文的逗号分开。

3-嵌入到网站。该功能是将本问卷嵌入到自己的网站中，其方法是将需要复制问卷的代码，整合到原有网站的代码中。整合过程中可以修改问卷的宽度和问卷的外观。

## 五、统计分析

一旦有网民开始填写问卷后、问卷星自动统计分析问卷结果，并实时更新结果信息。登录后，找到所发布的问卷，点击“统计分析”，然后选择具体的查看方式。

### （一）单题统计

可以查看每道题的统计数据，也可以直接看到数据图，如饼状图、柱状图等。

### （二）数据分析与报告

可以直接查看所有题的统计结果，默认呈现柱状图，也可以修改图形格式。

### （三）答题来源分析

可以查看用户答题方式，也可以查看时间分布或地域分布情况。

### （四）完成率分析

即查看点击次数和完成情况。

在查看统计结果的同时，也可以下载答卷或下载统计结果。当然也可以点击“查看&下 载”按钮选择下载内容和下载方式。

## 本章小结•…

问卷设计是一项具有很强的理论性、技术性和艺术性的工作，是市场调查流程中的重要一 环，本章比较系统地介绍了问卷设计技术。一份完整的问卷通常由标题、说明词、正文、编码、 被调查者基本情况、结尾等六部分内容组成。问卷按照其使用的方法、发放的方式、标准化程 度的不同，可以划分为不同的种类。

设计问卷应该遵循目的性原则、科学性原则、逻辑性原则、中立性原则和方便性原则。问卷设计是一个比较复杂的过程，问卷设计之前必须进行探索性研究，问卷设计过程中的两个关 键问题是明确调查的问题和设计量表，

问卷设计中必须根据研究的目的，明确调查的主题，梳理与调查主题相关的一系列概念，并将概念转为调查变量，然后针对变量逐一提问，根据要求设计出问句。问句的表述方法有多 种，问句的表述也要遵循诸多规则。针对不同的问题，设计合适的量表，将问题的被选答案表 述出来，答案的类型也有多种。此外，还需要对问卷的版式进行设计，使其简洁、整齐、美观、清 晰、便于阅读与填答。最后还要对问卷进行评估与修改。

网上问卷调查越来越受到人们的重视。网上问卷调查也有多种方式，通过网上专业调查 平台创建在线调查问卷，使问卷的设计实现了半自动化，极大地提高了问卷设计的效率。

设计一份高质量的问卷是一项极其复杂且具有相当难度的工作。问卷设计者需要具备较 为广博的专业知识、对调查现象与调查对象有充分的理解，还需要具备丰富的问卷调查经验 磁关键词•…

问卷封闭式问卷开放式问卷简单询问法简单陈述法释疑法假定法转移法 双向列联法 过滤法 比较法  
顺序效应 两项选择法 多项选择法 排序法 等级法

在线问卷

## 啬思考题

即测即评

1. 什么是调查问卷？它有什么作用？
2. 简述调查问卷的结构。
3. 设计问卷应该遵循的基本原则有哪些？
4. 简述问卷设计的流程。
5. 如何确定问卷中的调查变量？
6. 问卷中问题的表述方式有哪些？各有什么特点？

7. 设计问句应注意哪些问题?
8. 问题答案设计应遵循哪些原则?
9. 封闭式问题的答案设计方法有哪些? 各有什么特点?
10. 问卷的版式设计有哪些要求?
11. 如何对问卷初稿进行审查与评估?
12. 如何创建网上调查问卷?

## 宫本章实训. ...

### 一、 实训目的

1. 熟悉问卷设计的基本原则与基本流程。
2. 训练问卷中调查问题的确定方法。
3. 掌握问卷设计中问题表述与答案表述的设计方法。
4. 熟悉在线问卷的创建方法。
5. 掌握问卷质量的评价方法

### 二、 实训内容

随着电子商务的快速发展，网上购物已经成为许多消费者日常购物的一种重要方式。某研究机构为了研究消费者的网购行为、网购偏好、网购习惯、网购感知与网购评价，拟进行一次 关于消费者网购情况的问卷调查

1. 根据实训材料中提示的研究内容，梳理、归纳问卷调查中涉及的概念。
2. 根据梳理出来的概念，确定问卷调查所涉及的变量。
3. 根据所确定的调查变量，按照问卷设计的规范，设计一份结构完整的纸质版问卷。
4. 根据个人的偏好，选择一个具有问卷设计功能的网上调查平台，借助其中的问卷设计 功能模块，  
创建一份在线调查问卷 G

5. 利用一个样本量不少于 100 个单位的样本进行一次预调查，对问卷的信度与效度进行 评价。

### 三、 实训组织

1. 根据实训内容组建团队。
2. 各组在不同的场景完成问卷设计与预调查-
3. 各组互动交流，完善调查问卷，掌握问卷设计技巧。
4. 教师点评.

# 第六章 抽样方法

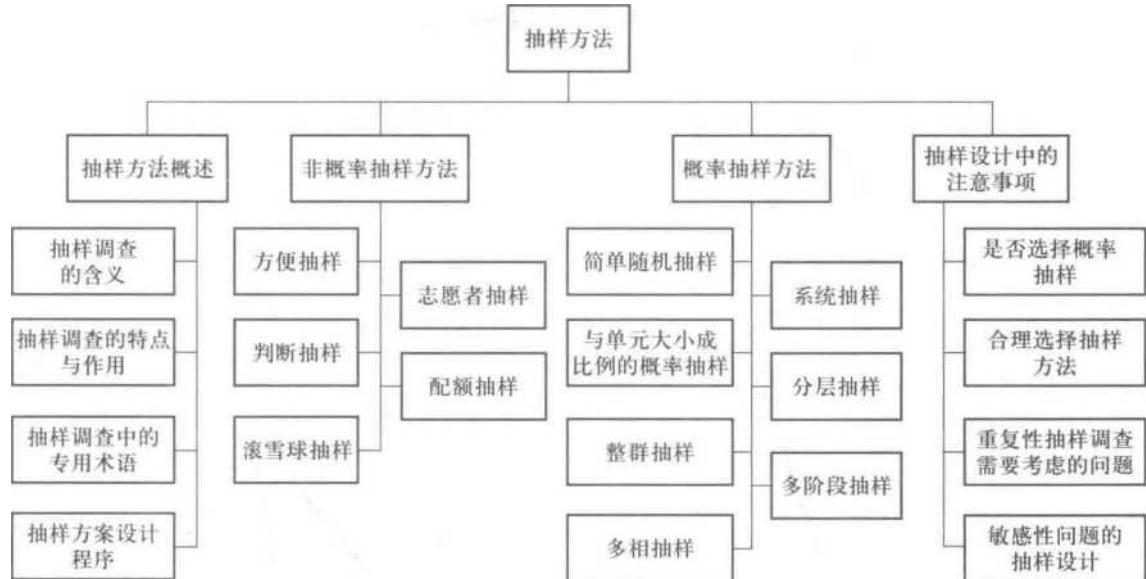


从前面几章我们了解到，在进行一项调查之前，必须考虑许多问题：调查的目标，研究对象的范围与特点，目标总体范围的确定，已有现成数据的使用价值及其局限性等。一旦现有数据不能满足决策管理的需要，就必须采集新数据，这时，就进入调查方式设计阶段。其必须回答的主要问题有：全面调查或抽样调查方式的选择，明确可供利用的调查框，实际调查总体的范围，抽样单元的选取，可能出现的误差，调查结果的评估等。如果决定进行一项抽样调查，下一步是决定如何从总体中抽选样本单元以及抽取多少样本的问题。本章的目的是介绍各种抽样方法，即如何从总体中抽选样本单元。

## 、教学目标与要求

1. 理解抽样调查的含义、特点与作用；
2. 掌握抽样调查重点专用术语；
3. 熟悉抽样方案的设计程序；
4. 理解非概率抽样的含义、优点与缺点；
5. 掌握非概率抽样的专门方法；
6. 理解概率抽样的含义、特点；
7. 掌握各种概率抽样的专门方法及其适用场景；
8. 熟悉抽样设计中需要考虑的相关问题

## 本章知识结构



# 第一节

## 抽样方法概述

在介绍具体的抽样方法之前，需要了解抽样调查的含义、特点与作用，理解广泛使用的专用术语和抽样方案设计的一般程序。

### 一、 抽样调查的含义

抽样调查是指按照一定程序，从所研究对象的总体中抽取一部分总体单元（样本），对其特征进行调查和观察，并在一定的条件下，运用数理统计的原理和方法对总体的数量特征进行估计与推断的方法技术。

依据是否遵循随机原则抽取样本，抽样方法可以分为两类：一类是非概率抽样；一类是概率抽样。选择使用哪种方法主要取决于研究者是否打算对总体进行推断。

非概率抽样，也称非随机抽样，是指调查员依据主观的（非随机的）判断或从方便出发从总体中抽选样本的方法。尽管非概率抽样的实施快速、简易且省钱，由于其主要依赖研究者或调查员的个人经验和判断，不能计算入样概率，难以保证样本对总体的代表性，因而无法估计和控制抽样误差，无法用样本的数据特征来推断总体『

概率抽样，也称随机抽样，它是遵循随机原则从总体中抽取部分作为样本进行观察。与非概率抽样相比，概率抽样较为复杂、费时，费用也更高。然而，由于单元是从总体中随机抽选出来的，而且能计算每一个单元的入样概率，因此能够得到可靠的估计值及其抽样误差的估计值，并对总体进行推断口

非概率抽样和概率抽样都有许多不同的抽样方法（或称抽样设计），选用哪一种方法取决于多种因素，例如，可利用的抽样框形式，总体的分布情况，调查的费用，数据使用者将如何分析数据等。在对概率抽样方法进行选择时，目标是使那些最重要的调查指标估计量抽样误差最小，同时也使调查的时间最短、费用最少

### 二、 抽样调查的特点与作用

#### （一）抽样调查的特点

与全面调查相比，抽样调查具有以下显著特点：

1. 良好的经济性。抽样调查可以根据调查需要，抽取一定数量的样本进行调查，与全面调查相比，样本量大大减少，从而可以显著地节约人力、物力和财力，具有良好的经济性

2. 较高的时效性。由于抽样调查只对总体中的少量单位进行调查，所以能够比较快地得到调查结果。例如，我国居民住户调查，就是采用抽样调查，每月均可取得必要的住户消费和需求资料。

3. 较高的准确性。抽样调查只调查部分总体单位，数目较小，参加调查的人员较精干，登记性调查误差较小，从而能较好地控制非抽样误差，提高调查的质量，具有较高的准确性：

#### （二）抽样调查的作用

1. 适用于不可能或不必要进行全面调查的社会经济现象的调查。例如，进行破坏性质或损耗性质的商品质量检验以及对无限总体的调查（如对森林木材积蓄量调查）等

2. 适用于连续不断的观察。抽样调查可对同一现象在不同时间进行连续不断的调查，可随时了解现象发展变化状况。

3. 适用于对全面调查进行验证。全面调查涉及面广、工作量大，花费时间和经费多，组织起来比较困难。但调查质量如何，需要检查验证，这时，采用抽样调查就比较合适。例如，人口普查，前后要用几年时间才能完成，常用抽样调查进行检查和修正。

4. 适用于统计过程控制。抽样调查还可运用于企业管理，尤其是产品质量过程管理，能更好地提高企业决策的有效性。

### 三、抽样调查中的专用术语

抽样调查是统计学的一个分支，它有一些专门术语，以下对一些常用的术语进行介绍。

#### (一) 总体和样本

总体在数理统计中也称母体，它是指所要研究对象的全体。总体由研究对象中的元素组成，组成总体的元素称为总体单位或总体单元。总体也可以表述为个体的集合，这里的个体可以是指研究对象的实体也可以是指观察到的数据。总体中包括的单元数目，称为总体总量或总体大小。

根据总体单元是否可数，总体可以区分为有限总体和无限总体。有限总体是指总体总量能明确定，并且单位数目有限。例如，2016年年末湖北省人口数，规模以上全国工业企业数等。无限总体则指总体总量不确定，且其单位数目无限：例如海洋中的鱼类资源数量，森林中的树木数量、连续作业的工业产品数量等。在调查中大多数的对象属于有限总体。在抽样中把总体划分为有限总体和无限总体，主要是为了判别在每次抽取样本之间是否相互独立。在无限总体中，每抽取一个单元，并不影响下一次的抽选结果，因此可以看作是相互独立的。在有限总体中抽出一个单元后往往要影响到下一次的抽选结果，因而不独立，且影响到估计量抽样方差的计算。但有两种情况，虽然是有限总体，仍可看作是独立抽取的：一种是有放回的抽选方法；另一种情况是有限总体的总量很大，而抽取的样本单元数比例很小，因此对下一次抽选的影响也很小，可以忽略不计。

样本是指抽样时按照抽样的规则所抽中的那部分单元所组成的集合。这里的总体单元概念与前面的总体单元或个体的定义相一致。由于抽样方式不同，如果样本是随机抽取的，称为随机样本。如果是主观有意选择的，则是主观或判断样本。有时还根据抽选的具体方式加以修饰，如分层样本，表示分层抽样的样本，整群样本表示整群抽样抽取的样本，如此等等。

样本中包括的总体单元数称为样本量。它是决定抽样误差大小的重要因素，同时也是影响调查费用的一个重要因素。从误差角度看，总是希望误差小，这样样本量应该大一些；但从费用角度看，一项调查总是希望调查费用低，这样样本量应该少一些。因此样本量的大小需要在误差和费用之间进行平衡。

#### (二) 参数、估计量和估计值

抽样调查的目的是要得到总体的某种特征，总体的特征广义上我们称作参数。通常参数有总体平均数、总体总量、总体比例等。假设有限总体的单位数为  $M$ ，要调查的目标变量通常用  $y$  来表示。于是其变量值为  $y_1, y_2, \dots, y_M$ ，其参数用符号表示如下：

总体总量  $\mu = \sum_{i=1}^M y_i$

总体平均数  $\bar{y} = \frac{1}{N} \sum_{i=1}^N y_i$

总体比例  $P = \frac{\text{具有属性的单位数}}{总单位数}$

其中：总体比例公式中， $y_i$  为属性变量， $y_i = 1$ ，当该单位具有某属性时为 1，其他则为 0°。总体范围

确定后总体参数客观存在，但却未知，需要通过调查来获得。如果采用抽样方法，需要通过抽样调查，根据样本观察的结果来进行推断估计。

估计量是样本观察值的函数，用它来对总体参数进行估计。由于样本是从总体中抽取的，样本中包含着总体的信息，可以通过它来对总体进行估计。但估计量的取值是随着抽中的样本单元而变动的，是一个随机变量。总体参数的估计量通常也不止一个，我们要选择一个好的估计量。统计理论证明，在简单随机抽样的情况下，以下的估计量是总体参数的好估计量。几表示样本中的单元数量，也称样本量。总体参数、估计量及其计算公式见表 6-1。

表 6-1 总体参数、估计量及其计算公式

参数	估计量	估计公式
$Y$	$\bar{Y}$	$\bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i$
$P$	$\hat{P}$	$\hat{P} = \frac{\text{入样单元数}}{n}$

在不同的抽样方式下，估计量也有不同的形式，同样我们也要选择一个好的估计量来对总体参数进行估计。如何选择好的估计量是抽样理论需要解决的问题。

在实际的抽样调查中，样本只能抽取一次，根据这一具体样本所计算的结果称为估计值，可以根据估计值来对总体参数进行估计。这种估计在统计学中称为点估计

### (三) 样本的抽选概率与抽样比

样本的抽选概率有两重含义。一个是指在一定的抽样方式下某个具体样本被抽中的概率。以简单随机抽样为例，假设有一个总体， $N=5$  (A, B, C, D, E)，从中抽取样本  $n=2$  (无放回)，则共有 10 个可能的样本，分别是：AB, AC, AD, AE, BC, BD, BE, CD, CE, DE。由于每个样本被抽中的概率是相同的，因此每个样本被抽中的概率是  $1/10$ 。另一个含义是指某个总体单元被选入样本的概率，从上面这个例子中，我们可以看到，共有 10 个可能的样本，其中每个总体单元入样的次数是 4 次，因此入样概率为  $4/10 = 2/5$ 。

抽样比是指样本量  $n$  与总体  $N$  之间的比率。即用  $f$  表示，有些介绍抽样理论的教科书上也用  $f$  表示。即  $f=n/N$ 。从这个例子可以看出，在简单随机抽样的情况下，每个单元的入样概率与抽样比是一致的。抽样比在抽样设计中是一个很重要的指标，如果抽样比过高，采用抽

样调查的意义就不大了。抽样比的倒数  $\frac{N}{n}$ ，也称扩张系数，它反映了每个样本单元所代表的总体单元数，在本书中称为权数。

#### (四) 抽样分布、抽样误差和抽样方差

抽样分布是指样本估计量的分布。由于估计量是一个随机变量，如果按照一定的样本量，一定的抽样方式反复抽取样本，每一个样本可以计算一个估计值，这些估计值就会形成一个分布，这一分布就叫作抽样分布。

抽样误差是指抽样调查中由于样本代表性引起的样本估计值与待估的参数真值之间存在的误差，估计量的方差称为抽样方差。它表示该抽样方案抽样结果的离差大小，从直观来看，抽样方差愈小意味着大多数估计值更靠近总体均值，抽到一个好的样本的可能性更大。其平方根称为抽样标准误差，也就是估计量的标准差。它的性质和抽样方差一致。

#### (五) 抽样框和抽样单元

抽样框是指提供给抽样用的所有调查单位的详细名单。一个抽样框一般包含识别资料、联系资料和辅助资料。识别资料是帮助识别抽样框中每个单元的项目，例如姓名、地址、身份证号等。联系资料是指在调查时用来确定总体单元所在位置的项目，例如通信地址或联系电话。辅助资料是指用来分类与维护的资料信息。如果调查的单元是“个人”，则其性别、年龄和文化程度等都是辅助资料。抽样框可以是名录框、区域框，也可以是多重抽样框，还可以是概念框。例如要从 2500 名学生中抽取 700 名学生组成一个样本，则 2500 名学生的名册，就是名录框。抽样框一般可采用现成的名单，如户口、企业名录、企事业单位职工名册、学生名册等，或单元所在的地理区域，或名单与地理区域的结合。在没有现成名单的情况下，可由调查员自己编制。例如某天早上 9 点至晚上 10 点之间进入某购物中心停车场的车辆，就是一个概念名录框。应该注意的是：在利用现有名单作为抽样框时，要先对该名录进行检查，避免有重复、遗漏、分类错误等情况的发生，以提高抽样框对总体的代表性。

抽样单元是指在总体分层的条件下，每层有限且互不重迭又穷尽的每个部分。抽样单元可大可小，例如，在全国居民生活状况调查中，各省就是一级单元，每个省又可分为较小的二级单元，如市、县等，还可按区、街道、家庭户、个人再细分为三级、四级单元等。

#### (六) 精确性和精度

抽样调查中经常用到精确性和精度的概念。精确性这个概念来自于测量问题。抽样调查中的精确性是指在抽样中若反复抽样其结果的差别大小，差别越小其精确性越高。它反映了估计量抽样误差或抽样方差的大小，因此在对同一参数进行估计时，其估计量的抽样标准误或抽样方差愈小，其精确性愈高。

精度通常是一个相对的概念，反映精度的主要指标有相对标准误和相对方差。它是估计量的抽样标准误或抽样方差与待估的参数之比。假如待估参数是总体平均数  $R$ ，它的抽样标准误为，则有：

$$\text{相对抽样标准误: } \frac{\bar{Y}}{\sigma}$$

$$\text{相对方差: } \frac{\bar{Y}^2}{\sigma^2}$$

这两个指标值愈小精度愈高。显然，在抽样设计和选择抽样方法时，在费用一定的情况下，应该选择一个比较精确或精度比较高的方案。

#### (七) 抽样效率与设计效应

抽样效率是指两个抽样方案在样本量相同情况下的抽样方差之比。当某个估计量的抽样

方差与另一个估计量的抽样方差比较时，则称抽样方差小的估计量效率比较高。在抽样调查中，抽样方差的大小和费用的多少成为比较两种不同抽样设计或抽样方案好坏的标准。一个好的抽样方案要求在固定的费用下使抽样方差尽可能小，或者在满足规定的误差限内使费用最省。由于费用与样本量的大小直接相关，因此在费用不易计算的条件下可采用样本量相同的条件下使抽样方差最小，或在抽样方差一定的情况下使样本量最小的原则。

设计效应是用来衡量抽样设计方案好坏的一个评价指标。最常用的一个指标是由美国统计学家 Kish 提出的，它是把设计方案的抽样方差与简单随机抽样的抽样方差进行比较。若设计方案的抽样方差用  $\sigma_{\text{crp}}^2$  表示，简单随机抽样的抽样方差用  $\sigma_{\text{simple}}^2$  表示，则设计效应（呀）可表述为：

$$Deff = \frac{\sigma_{\text{crp}}^2}{\sigma_{\text{simple}}^2} - 1$$

在。呀小于 1 时，表示设计方案的效率高于简单随机抽样，反之则表示设计方案的效率低于简单随机抽样。设计效应指标除用来评估比较复杂抽样方案的效率外，还可以利用它来计算复杂抽样方案的样本量等，因此是一个很有用的指标，

#### 四、抽样方案设计程序

抽样方案设计是指从总体抽取样本资料以前，预先确定抽样程序和方案，在保证所抽选的样本对总体有充分代表性的前提下，力求取得经济、有效的结果。抽样方案设计的难点在于如何在抽样效果和调查成本之间找到平衡点，即抽样效果较好，调查成本较低。依据国内外调查经验，在制订抽样方案时，常将相关人员，即调查资料使用者，抽样专家，实际调查工作组织者和数据处理人员等聚集一起，相互交流意见，协商探讨各种问题，最后提出一个合理可行的抽样方案。

抽样方案设计的基本内容依次为：根据抽样调查的目的、任务和要求定义总体及抽样单位；确定所采用的抽样框；选择抽样技术，确定抽取样本方法；确定必要的样本量；根据主要抽样指标的精度要求，确定总体目标量的估算方法。综合上述内容制订实施方案的办法和步骤

这些内容构成了抽样方案设计的主要程序节点，如图 6T 所示。

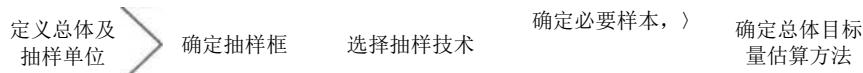


图 6-1 抽样方案设计程序图

## 第二节

### 非概率抽样技术

非概率抽样技术在实践中被广泛采用，具有一定的价值，本节介绍其优缺点与常用的非概率抽样方法。

## 一、非概率抽样含义及其优缺点

### (一) 非概率抽样的含义

非概率抽样没有严格的定义，其抽取样本方法的一个共同特点是用一种主观的（非随机的）方法从总体中抽选单元。由于不需要完整的抽样框，非概率抽样是一种快速、简单且节省费用的获取数据的方法。使用非概率抽样的问题是，我们不清楚能否通过样本对总体进行推断，原因是从总体中抽选单元的方式可能会导致较大的偏差。例如，在非概率抽样中调查员经常主观地决定哪些单元入样。由于调查员往往倾向于选择总体中那些最容易接触到和最友好的单元，使总体中很大一部分单元完全没有被抽中的机会，而这些单元与被抽中的单元之间可能存在系统差异。这样，非概率抽样会使调查结果出现偏差。另一方面调查员也可能有意选择具有平均特征的那些单元，从而有排除极端值的倾向，这将减少总体中明显的变异性。由于非概率抽样抽取样本时存在倾向性与偏差，而且没有抽样框，我们不可能计算出各个单元的入样概率，从而无法得到可靠的估计值及其抽样误差的估计值。此时若能对总体进行推断，必须假定样本对总体有代表性。而这通常需要假定总体指标服从均匀分布或其他分布。由于抽选存在偏差，做这种假定风险较大。

非概率抽样经常被市场研究者作为一种省钱、快速的方法来代替概率抽样。但是由于存在上述问题，它并不是概率抽样的一种有效的替代方法。那么，为什么还要花篇幅来讨论非概率抽样呢？这是因为非概率抽样除了简单易行外，还可以用于下面几个方面的研究：用来形成一种想法；作为设计开发概率抽样调查的初始步骤；在后续步骤中帮助理解概率抽样调查结果。

近年来，依赖网络大数据，人们开始关注利用更多的相关信息对非概率抽样的统计推断问题进行研究。①比如，非概率抽样能在调查的早期就提供非常有价值的信息。它也可用在探索性或诊断性研究中，了解人们的态度、信仰、动机或行为，或分析概率抽样调查的结果。有时，非概率抽样是唯一可行的选择。例如，在医学实验中，采用志愿者样本可能是取得数据的唯一途径。

非概率抽样常被用于抽选一些人来参加小组座谈和深层访谈。发达国家统计机构常用非概率抽样测试在人口普查调查表中所使用的问题，以确保所问的问题和使用的概念对被调查者是清楚的。如果认为问题的内容会引起争议，则在最有可能有争议的地方抽选一些人来试点。如果通过小组座谈证明问题能被这些人接受，则它们在其他地方也能被接受。另一个能较好发挥非概率抽样作用的例子是预研究。如果正在设计一项新的调查，对调查领域又了解不多，则预调查时常采用非概率抽样。例如，我们想对“互联网+书店”这个行业的从业人员进行一项调查，假定开始我们对有多少人在这个行业工作，他们的工作时间、收入状况及其他情况了解不多，也没有现成的抽样框，我们就可以先作一个预调查，初步设计一份问卷放在网上，通过问卷反馈回来的信息设计正式调查。

### (二) 非概率抽样的优点

1. 快速简便。一般而言，非概率样本不需要有完备的抽样框等先决条件，因此能被很快抽出并进行调查。例如，要进行一项市场调查，可以在商店门口或大街上，拦截一些顾客进行调查，这就很容易做到。

① 金勇进，刘展. 大数据背景下非概率抽样的统计推断问题[J]. 统计研究, 2016 (3): 11-17.

2. 费用相对较低。由于非概率样本不受地域的限制，可以使样本比较集中，这样就可以减少调查员的差旅费。

3. 不需要任何抽样框。抽样调查中要取得一个完备的抽样框往往是很费时费力的，而非概率抽样就不需要任何抽样框。

4. 非概率抽样对探索性研究和调查的设计开发很有用。

### (三) 非概率抽样的缺点

1. 为了对总体进行推断，需要对样本的代表性做很强的假定。由于所有非概率抽样的样本抽选都存

在某种倾向性或偏差，做这样的假定通常有很大的风险。如果要对总体进行推断，必须使用概率抽样。

2. 由于不可能确定总体单元的入样概率，故不可能得到可靠的估计值以及抽样误差的估计值。

## 二、非概率抽样方法

非概率抽样方法一般有方便抽样、志愿者抽样、判断抽样、配额抽样和滚雪球抽样 下面 分别介绍。

### (一) 方便抽样

方便抽样又称偶遇抽样，其样本单元的抽选以调查员方便的方式进行，几乎没有或完全没 有计划。方便抽样假定总体是同质的，如果总体各单元相似，那么可以抽选任何单元入样

“街道拦截”访问法是方便抽样的常用方式。调查员可以在街道上或任意便利的地理位置上抽选一个碰巧路过的人，目测其符合被调查者特征后，与其接触进行调查。例如，要调查 图书馆读者的相关情况，可以在图书馆的门口拦截一些读者作为样本进行调查。

该抽样方法简便易行，能及时取得所需的相关信息，省时、省力，节约费用；但容易受到调 查员的倾向性影响；也容易在抽样的时候谁碰巧路过，这个“谁”可能会影响样本的代表性，因 而抽样偏差大。因为其实施比较方便，在总体是同质的情况下有时能取得好的效果。

### (二) 志愿者抽样

志愿者抽样是指被调查者都是志愿者的一种抽样方法。其抽选出来的样本也称为反应性 样本。一般说来，因为作为调查对象的志愿者和非志愿者之间往往存在差别，因而会产生较大的抽选偏差，这时应对志愿人员进行筛选以得到符合调查目标的对象，但这样仍难免存在选择 性偏差。在某些特定情况下，志愿者抽样具有一定的价值。

志愿者抽样的适用场景有以下几种：医学研究中的志愿者招聘。例如，出于道德上的 考虑，我们不得不请求具有特殊病情的人自愿参加某项医疗实验。活动参与者的志愿者招 聘。例如，当电视或广播节目中讨论某一问题时，主持人邀请听众或观众打电话发表他们 的意见，只有那些关心所讨论的主题，且有这样或那样观点的人才会响应，大多数人都不会 响应，这样就会导致较大的抽选偏差。志愿者抽样有时也用于选择参加小组座谈或深层访 谈的人。

### (三) 判断抽样

判断抽样又称目的抽样，它是依据研究者的主观意愿、经验和知识，从总体中选择具有典 型代表性的样本作为调查对象的一种抽样方法，，换句话说，由专家有目的地挑选他们认为 “有代表性”的样本。该方法使用的前提是研究者对调查总体特征十分了解。

判断抽样抽选样本单位的方法有两种:一种是选择与研究目的相关的最具代表特征的调

查对象，可以是“平均型”或“多类型”标准，也可以是“极端型”标准；另一种是利用总体的全 面统计资料，按照一定的标准，主观选取样本。

判断抽样方法在样本量小及样本不易分类挑选时具有优势，但由于其误差大小很大程度 上依赖于研究者对调查对象的了解程度、判断水平以及对结果的解释，所以，该方法的结果容 易遭受质疑

判断抽样方法可应用于探索性研究，如抽选参加小组座谈或深层访谈的人员（例如检验 问卷的某个特定方面）时有用但是它不适用于预调查”

#### （四） 配额抽样

配额抽样类似随机抽样中的分层抽样，首先将总体中的所有单位按一定的标志分为若干 类（组），然后在每个类（组）中用方便抽样或判断抽样的方法选取样本单位。

采用配额抽样，事先要对总体中所有单位按其属性或特征分为若干类型，这些属性、特征 称为“控制特征”，例如，被调查者的年龄、收入、职业、文化程度等，然后，按照各个控制特征分 配样本数额。

按照配额的要求不同，配额抽样可分为“独立控制”和“交叉控制”两种。

独立控制配额抽样是根据调查总体不同特性，对具有某个特性的调查样本分别规定单独 分配数额，不 规定必须同时具有两种或两种以上特性的样本数额。因此，调查员有比较大的自由去选择总体中的样本•例如，调查时关注总体的性别和收入两个特征。抽样时只分别规定 男性与女性各抽取多少人，高收入、中 等收入和低收入各层次分别抽取多少人。

交叉控制配额抽样是对调查对象的特性样本数额进行交叉分配，即复合分组，对每一个控 制特性所需 分配的样本数都做了具体规定，调查员必须按规定在总体中抽取调查单位。由于 各个特性都同时得到了控 制，从而克服了独立控制配额抽样的缺点，提高了样本的代表性。例 如，如果采用交叉控制方法抽取样本， 针对性别和收入两个特征，则需要将性别与收入进行复 合分组，分别得到高收入、中等收入与低收入各组 分别抽取多少男性与女性样本。

配额抽样由于费用相对较低、易于实施并能满足总体比例要求等特点，是广受欢迎的一种 非随机抽样 方法。市场研究者常用它来代替分层抽样（特别在电话调查中），但同时它也掩盖 了潜在的可能显著的偏 差：

#### （五） 滚雪球抽样

滚雪球抽样是指从一个已知属于目标总体对象的名单开始，访问、询问他们所知道的该群 体中其他人的姓名，再通过这些人接触更多其余的人，直到没有新的对象出现为止，或者满足 所需要的样本量为止。

滚雪球抽样适用于那些特征稀少、小的、特殊总体，诸如盲人、聋哑人、同性恋者、流浪儿童 等群体，也适用于音乐家、画家、诗人这类不能很容易地根据其工作场所来识别的群体。

雪球抽样方法的基本步骤为：首先，根据研究目的和相关特征找出少数样本单位；其次，通 过这些样 本单位了解更多的样本单位；如此类推，如同滚雪球，使调查结果愈来愈接近总体。 这种方法的优点是便 于有针对性地找到具有“某种特征”的被调查者，可提高抽样效率。其局 限性是要求样本单位之间必须有 一定的联系并且愿意保持和提供这种关系，否则，将会影响这 种调查方法的进行和实施效果，因为一定会 有些个体，甚至是部分群体，根本没有人知道或被 提及。

## 第三节

# 概率抽样技术

概率抽样技术是市场调查中的主流方法，与非概率抽样技术相比具有明显的优势。本节 主要介绍常用的概率抽样方法及其具体适用场景。

## 一、 概率抽样的含义及其优缺点

### （一） 概率抽样的含义

概率抽样是指遵循随机原则从总体中抽选单元组成样本进行调查的方法。这里的随机并不意味着随意。特别是调查员不能随意选择被调查者，因为如果那样的话，抽样将受到个人意愿的影响。随机抽选一般通过随机数表或由计算机随机生成随机数，来确定哪些单元应该入样。

概率抽样有两条基本准则。

第一，样本单元是按随机原则抽取的。所谓随机原则就是在抽选样本时排除主观上有意识地抽选调查单元，被抽中的单元既不取决于调查研究者的愿望，也不取决于被调查者的态度，而是依赖于客观的机遇。

第二，调查总体中的每一个单元都有一定的概率被抽中。所谓一定的概率被抽中有两层含义：第一层含义是指每一个单元都应有机会被抽中，用数学的语言来表述，就是每一个单元都有一个非零的入样概率。这意味着每一个单元被抽中的概率可大可小，但一定要有机会被抽中，它可以被指定被抽选的概率为1，即一定要抽中，但不能指定被抽选的概率为0，即一定不抽中，因为这样就意味着将该单元排除在总体之外。第二层含义是每个单元入样的概率应该已知，能计算出这些概率。总体中所有单元不一定都要有相同的入样概率。实际上在大多数复杂抽样中，总体单元入样概率差异很大，这在后面介绍各种抽样方法时将会看到。

### （二） 概率抽样的优缺点

1. 概率抽样的主要优点。由于每个样本单元都随机抽取，且能计算各个单元的入样概率，所以能得到总体的可靠估计值并能计算每个估计值的抽样误差，因而能对总体进行推断。实际上，采用概率抽样方法，用一个相对小的样本，就能对一个大的总体进行推断。因而能节约资源并提高效率。

2. 概率抽样的主要缺点。与非概率抽样相比，概率抽样比较复杂、费时，通常也更费钱。一般而言，创建和维护一个高质量的抽样框的费用很高。另外，由于概率样本在总体中的分布通常比非概率样本广，因此数据收集的费用比非概率样本高，且更难实施。尽管有这些缺点，但就能从概率样本进行推断而言，其利远大于弊。

## 二、 概率抽样方法

概率抽样设计有多种不同类型，最基本的是简单随机抽样，较为复杂的包括系统抽样、与单元大小成比例的概率抽样、整群抽样、多阶抽样、多相抽样等。这些抽样技术在不同情况下都很有用，需要根据调查目的、调查要求、调查费用支持等不同条件进行选择。下面分别介绍主要抽样设计的类型。

## (一) 简单随机抽样

### 1. 简单随机抽样的含义

简单随机抽样又称纯随机抽样，它是最基本的抽样方式。所有概率抽样的出发点都是简单随机抽样。

简单随机抽样是一步抽样法，假设调查总体共有  $V$  个单元，从中随机抽取  $n$  个单元作为样本，每一个单元都有同样的机会被抽中，也即有相同的入样概率，这个概率为  $P$ ，它等于这种抽样就称作简单随机抽样，所抽到的样本称为简单随机样本。

$/V$

简单随机抽样按抽样方式的不同，分为有放回的抽样和无放回的抽样。有放回抽样是指允许一个单元被抽中的次数多于一次，又称为重复抽样；无放回抽样意味着一个单元一旦被抽中，就不能被再抽中，又称为不重复抽样。一般情况下，无放回抽样得到的估计结果更精确，实际操作也方便一些。在本章中，除非特别指明，简单随机抽样都是指无放回抽样。

### 2. 简单随机样本抽取方法

简单随机样本抽取的方法有：掷硬币、掷骰子、抽签或抓阄、随机数表等方法。在市场调查中一般总体单位较多，主要采用后两种方法。

(1) 抽签或抓阄法。它是指将总体的每个单元编号，并做成号签，把号签混合之后从中抽取所需单位数，然后，按照抽中的号码，查对调查单位并进行记录。抽中号签对应的单位就是需要调查的样本。

(2) 随机数表法。随机数表是将 0~9 十个数字用完全随机顺序排列编制而成的表。随机数表各有不同，下面举例说明如下。

表 6-2 随机数字表（截取）

行序	列序				
	75-79	80-84	85-89	90-94	95-99
00	82848	04190	96579	90464	29065
01	96621	43918	01896	83991	51141
02	61891	27101	37855	06235	33316
03	59798	32803	67708	15297	28612
04	94335	34936	02566	80972	08188

为了更好地具体说明随机数表抽样方法，以一项商业零售网点调查为例。该项调查的目的是估计在一个给定年度经营一个商业零售网点所需的费用。假定有一个适当的现成商业零售网点名录，或可以从现有材料中建立，这样的名录就可以用作抽样框。假定总体名录中有  $V=1530$  个商业零售网点，要从中抽选  $n=90$  个网点的样本。下一步则是决定如何抽选这 90 个商业零售网点。

样本的抽取可以通过使用随机数表进行（见表 6-2）。抽取过程如下：

第一步，抽取一个四位数（抽四位数是因为  $V=1530$  是四位数）。我们可以在表中任意一个位置选一个四位数；

第二步，任选一个方向连续抽下去。如果决定顺着列向下抽，前 90 个不大于 1530 且不同的四位数入选。

假定第 01 行和第 85-88 列被选为起始点，顺着这个起始点按列方向向下，满足条件的相

应的随机数依次是：189, 256 等，直到得到 90 个不同的数为止。这样得到的样本由总体中以这些数编号的那些商业零售网点组成，要注意的是，这里讨论的是无放回简单随机抽样，当相同的数字重复出现时，则应该略去。此外，由于 N=1530，第一个数字 1 比较小，因此大量的四位数将抛弃不用，使随机数表利用不充分。因此对随机数表的应用也有一些技巧，我们可以把 2001 至 4000 之间的数字减去 2000, 4001 至 6000 之间的数字减去 4000，依次类推。这样随机数表的数字可以比较充分利用，既做到等概率抽选，又可以加快抽选速度。上例就可以补充相应的调查编号：77()

上面介绍的随机数表的使用是以手工抽取进行的，但从实际情况出发，这样长的商业零售网点名录很有可能是以计算机文件的形式存在的，这种情况可以用计算机程序来生成样本。

利用 SPSS 实现简单随机抽样的步骤如下：

在已经建立的抽样框数据编辑窗口中，单击“数据”-“个案”“个案样本”-“样本”，出现“选择个案：随机样本”对话框。如果是近似抽样，根据抽样的比率，输入百分比；如果是精确抽样，则输入样本量和总体大小，单击“继续”-\*“过滤掉未选定个案”-“确定”可得到抽取的样本编号。

再点击“删除未选个案”-“确定”，保存数据，得到简单随机样本数据集。

### 3. 简单随机抽样的优点与缺点

(1) 简单随机抽样的优点。  
①与其他抽样技术相比，它是最简单、直观的抽样技术，比较容易理解和掌握。  
②不需要其他(辅助)信息就能对抽样框进行抽样，唯一需要的只是一个关于调查总体所有单元的一个完整的清单和与其联系的信息。  
③它已建立了成熟的理论，关于样本量的确定、总体估计与方差估计都有标准的现成公式可以利用，而且这些公式很容易掌握。

(2) 简单随机抽样的缺点。  
①抽样框中即使有现成的辅助信息也不加利用，使得估计的统计效率较其他利用辅助信息的抽样设计要低；  
②由于样本在总体中的地理分布范围比较广，如果采用面访调查，费用较高；  
③有可能抽到一个不好的简单随机样本，这是因为这种抽样不加任何限制，所有样本量为 n 的样本，都有相同的机会被抽中，因此抽出的样本有可能不能很好地代表总体。  
④如果不使用计算机，而用随机数表抽选一个大样本，将十分单调而又费时。

## (二) 系统抽样

### 1. 系统抽样的含义

系统抽样又称等距抽样，是指将研究总体按一定的标志排序后，每隔一定的间隔抽取一个单元直至抽到所需样本量的抽样方法。进行系统抽样需要一个抽样间距和一个随机起点。当使用名录框时，设总体单元数为 N，要从中抽取 n 个单元做样本，先要计算出抽样的间距  $k, k = N/n$ 。随机起点 r 是 1 到 k 范围内的一个随机数，被抽中单元的顺序位置分别是：

$$r, \dots, r + (n - 1)k$$

与简单随机抽样一样，每个单元的入样概率 P 都等于  $n/N$ ；不同的是并非几个单元的每一种组合都有相同的机会被抽中。系统抽样只能抽选距离为 k 的单元组成样本：所以从总体中只能抽出左 n 个可能的系统样本。

### 2. 系统抽样样本抽取方法

#### (1) 等间隔随机起点抽取样本

等间隔随机起点抽取样本是指将总体单位排序后，划分为间隔相等的若干部分，在第一个部分按照随机原则抽取样本，接着等间隔陆续抽取样本，直至抽到所需的样本数。

例如，假定我们有一个含有单元数为 V=54 的总体，并打算抽取一个 n = 9 的样本量，如何实现系统抽样？

第一步，确定抽样间隔如  $k = N/n = 54/9 = 6$

第二步，确定随机起点。抽一个 1 到 6 范围内的随机数，比如说 2，则编号为 2、8、14、20、26、32、38、44 和 50 的单元即为入样单元。在  $V = 54$ ,  $n = 9$   $J = 6$  的情形下，只有 6 个可能的系统样本而对样本量为 9 的简单随机抽样，其可能的样本数有：54 个中随机抽取 9 个的组合” / 54\，数，即  $h$  个，要超过 53 亿一

使用系统抽样可能会遇到的一个问题是有时  $V$  不能被  $n$  整除，这时抽样间距不是一个整数。这种情况下  $J$  取与其最接近的整数。不幸的是，如果这样做，实际抽到的样本大小就会不一样。例如，假定  $A = 55$ ,  $n = 9$ ，则  $J = 6.1$ 。假如取  $J = 6$ ，且如果  $A = 2$ ，则样本中的单元是编号为 2、8、14、20、26、32、38、44 和 50 的单元，共 9 个。如果随机起点是 1，则样本是编号为 1、7、13、19、25、31、37、43、49 和 55 的单元。这种情况下，实际样本量为 10。

如果  $V$  不能被  $n$  整除，则可以使用圆形系统抽样法来避免出现可能样本量不一致的情况。我们把总体单元假想排列在一个圆上，并采用模余数法取  $J$  为与  $V$  几最接近的整数，现在随机起点  $r$  在 1 到  $M$  而不是 1 到  $n$  的范围内抽取（即第一个样本单元可以是名录中的任何单元）。与前面一样，被抽中的单元是：

$$r, r+(n-1)k$$

如果第  $r$  个单元使得  $r+(J-1)k > N$ ，则取单元  $r+(A-V)$ 。也就是说，在达到名录的末尾后转回到名录的开始。圆形系统抽样法的优点是每一个单元都有一个相同的被抽入样本的机会。例如，在前面的例子中，假定  $A = 55$ ,  $n = 9$ ，取  $J = 6$ ，在 1 到 55 的范围内抽取一个随机起点，例如， $r = 42$ ，则总体中被抽中的单元分别是：42, 48, 54, 5, 11, 17, 23, 29 和 35。

使用系统抽样的另一个问题是，抽样间距  $k$  有可能正好碰上总体中变化的某种周期性。例如，假定我们正在某个地区进行交通流量调查，且我们只能选一周中的某一天抽样。换句话说，每隔七天取一次样，容易想象，都选星期日取样和都选星期二取样估计的流量就会有很大的差别，在大多数情况下，我们根本就不知道总体是否有周期性变化。

## (2) 对称等距抽取样本

一般系统抽样中第一个单元是从 1 到  $M$  中随机抽选的，当总体单元是按有关标志排列时，就会产生一个新的问题：由于有关标志与调查变量之间相关，调查变量也是由高到低或由低到高形成一种趋势。因此第一个单元的偏高或偏低就会影响到以后各单元，使整个样本可能会产生偏差。为了减少这种偏差，有人提出了对称等距抽样的方法，以消除由于第一个单元位置引起的系统偏差。对称等距抽样方法也有多种，一种是由统计学家塞蒂 (Sethi) 提出的。当总体容量为  $M$  要抽取样本量为  $n$  时，其抽样间隔为  $k = N/n$ ，这时实际上把  $V$  个单元分成  $\lceil \frac{V}{n} \rceil$  组，每组有  $k$  个单元，现在把临近的两组加以合并，合并成  $\lceil \frac{n}{2} \rceil$  组，每组  $2k$  个单元。当组内前一个抽中的单元位置偏低时，后一个单元的位置使它偏高，使它两两对称，偏低和偏高相互抵消；具体的做法按照公式抽取。

首先在 1 到  $M$  中抽取一个随机数，假设为  $r$ ，抽样间隔为  $k$ ，则各抽中单元的位置按下面的公式计算可得：

$$\text{官+2 必, } 2(j+1) \leq r+1] \quad j = -1$$

下面用一个例子来说明这种抽样方法。已知某总体单元数为  $V=300$ ，要从中抽取  $n=20$  个单元为样本，先计算抽样间隔  $k = 300/20 = 15$ 。在 1 到 15 之间抽选一个随机起点，假定随机起点  $r=3$ ，按公式计算出应抽取的单元号码为：

$$\begin{array}{lll} 7 = 0 & r+2jk = 3 & 2(J+1)k-r+1 = 28 \\ j=1 & r+2jk = 33 & 2(j+1)/c-r+1 = 58 \\ \hline j=9 & r+2/7c = 273 & 2^{\circ} +1)4r+1 = 298 \end{array}$$

这样抽中的单元号码为：3, 28, 33, 58, …, 273, 298。

当以为奇数时，仍按上式进行，但只取到 $(n-1)/2$ 为止，并要增加靠近终端的一个样本单元。也用一个例子来说明。

假设 $n=35$ ,  $k=5$ ,

这里抽选的样本单元数为7，是一个奇数。先在1到5之间取一个随机数，假定为 $r=4$ 。  
则应抽的单元号码为：

$$\begin{array}{ll} r+2jk = 4 & 2(j+1)J+1=7 \\ r+2jk = 14 & 2(j+1)A;-r+1 = \\ \text{产 } 2 & r+24 = 2(j+1)4-r+1=27 \\ J = 3 & r+2jk = 34 & 2(j+1)A: -r+1 \end{array}$$

最后应加一个单元 $\lceil +(\square-1)\rceil = 34$ 。因此由编号为4, 7, 14, 17, 24, 27, 34。这7个单元组成样本。

另一种对称等距抽样方法是由统计学家辛(Singh)提出，也称修正的对称等距抽样，其方法是抽选对总体两端等距离的成对单元。当样本量几为偶数时，其样本都是对称的，抽样间隔的计算方法与前面相同。  
假设随机起点为，则抽选样本的公式为：

$$n \\ [,+必, 3-北)-r+1] \\ 1$$

当 $n$ 为奇数时，用辛的方法要增加靠近中间的第8个单元。

为了便于和塞蒂方法进行比较，仍用前面的例子。用修正的对称等距抽样方法抽取的样本结果及其比较见表6-3：

### 3. 系统抽样的优点与缺点

(1) 系统抽样的优点。①在没有抽样框时，可以构造一个概念抽样框，可代替简单随机抽样。②与简单随机抽样一样，系统抽样不需要辅助的抽样框信息。③与简单随机抽样相比，系统抽样样本的分布比较好，当然这还取决于抽样间隔及名录是如何排列的。④与简单随机抽样一样，系统抽样有较好的理论支持，估计值容易计算。⑤由于只需要一个随机起点，系统抽样比简单随机抽样简单。

(2) 系统抽样的缺点。①如果抽样间隔正好碰上总体变化的某种未知周期性，就会得到一个“差的”系统样本，形成系统偏误。②与简单随机抽样相同，由于不使用抽样框中的辅助信息，抽样策略的效率不高。③在使用概念框时，不能预先知道最终样本量。④系统抽样的抽样方差，没有一个无偏的估计量。为了进行方差估计，必须把系统样本看作简单随机样本，

而这一点只有在名录中的单元是随机排列时才是合适的。⑤在总体大小/V 不能被样本量 n 整除且不使用圆形抽样法时，会得到不同样本量大小的样本

表 6-3 两种对称等距抽样的结果比较

方法名称	塞蒂方法		辛方法		
抽样公式	$\lceil \frac{1+2(\lfloor \frac{n}{b} \rfloor -1)}{2} \rceil = 0, 1, \dots, \lfloor \frac{n}{b} \rfloor$		$lr^*jk, (N-jk) - r+1 \rceil = 1, 2, \dots, \lfloor \frac{n}{b} \rfloor - 1$		
抽样结果	样本量为偶数 $n=300$ $n=20$ $k=15$ $r=3$ $j=0,1,2, \dots, 8,9$	3	28	3	298
		33	58	18	283
		63	88	33	268
		93	118	48	253
		123	148	63	238
		153	178	78	223
		183	208	93	208
		213	238	108	193
		243	268	123	178
		273	298	138	163
样本量为奇数 $V=35$ $n=7$ $k=5$ $r=4$ $\sigma \approx 0 \pm 2$		4	7	4	32
		14	17	9	27
		24	27	14	22
		34		19	

### (三) 与单元大小成比例的概率抽样

#### I. PPS 抽样的含义

与单元大小成比例的概率抽样 (Proportional Probability Sampling , PPS) 是指采用抽选的概率与单元规模大小成比例用英语单词的第一个字母表示，简称 PPS 抽样。如果抽样单元大小的辅助信息是已知的，这样的信息就可以用在抽样中，以提高统计效率。如果大小度量是准确的，而且所研究的变量与单元的大小相关，PPS 抽样就能极大地提高精度。如果大小的度量不大准确，最好按单元规模大小分组并使用分层抽样。

PPS 抽样不同于简单随机抽样与系统抽样的等概率抽样，它是一种不等概率抽样。

在实际的抽样调查中，经常会遇到总体单元的大小不同，例如学校、社区、工厂、农场作为抽样单元时，它们的大小相差非常悬殊，这种情况下采用等概率抽样的效果就不会好，因为一个大的单元相当几个或几十个小单元，因此很自然会想到采用不等概率来抽选。

农场调查经常使用 PPS 抽样，有关单元大小变量的测量指标是面积，例如，以公顷衡量的农场规模大小。当然，农场的规模每年会有些变化，但对大多数农场来说，农场的规模在各个年份都是常量。此外，农场调查中的典型问题，如收入、农作物产量、牲畜拥有量和农场的支出，都与农场规模有关，在工商调查中可用的其他规模的度量，包括雇员数、年销售额、经营场所数（含线上与线下数量）等变量，在不同年份发生变化的可能性要大一些。

下面用一个抽选农场的例子来说明 PPS 抽样。以土地面积作为农场的规模，那么一个面积为 200 公顷的农场被抽入样本的概率，是面积为 100 公顷农场的两倍。现假定一个总体有 六个农场，我们要估计这个农场总体的总支出。

假定我们已知每个农场的规模（以公顷计的农场大小），为便于说明，进一步假定我们已 知它们的支出。显然，在现实生活中，如果我们已知各农场支出的话，就不需要再做这个调查 了，这里是想说明 PPS 抽样优于简单随机抽样

先考虑这个总体的名录，如表 6-4 所示。

表 6-4 农场的抽样框名录

抽样单元:农场	农场的规模/公顷	调查变量: 支出/元
1	50	26 000
2	1 000	470 000
3	125	63 800
4	300	145 000
5	500	230 000
6	25	12 500
总计	2 000	947 300

对这六个农场的总体，真正的总支出为 947 300 元。假定我们没有对这六个农场进行普 查的经费，预算只够对一个农场的样本进行调查。因此把样本量取为 1, 这里只是为了说明方 便, 在实际调查中，只抽一个单元的情况极少。如果我们采用简单随机抽样，每个样本只含一 个单元，每个单元的入选概率为  $1/6$ , 那么大小为  $n=1$  的样本有六种可能，如表 6-5 所示。

表 6-5 所有可能的  $n=1$  的简单随机样本

样本	入选概率	样本权数	农场的支出	估计总支出
农场 1	$1/6$	6	26 000	136 000
农场 2	$1/6$	6	470 000	2 820 000
农场 3	$1/6$	6	63 800	382 800
农场 4	$1/6$	6	145 000	870 000
农场 5	$1/6$	6	230 000	1 380 000
农场 6	$1/6$	6	12 500	75 000
总计		—	—	947 300

下面我们来分析简单随机抽样的结果。为此，必须先引入一些估计的概念，（这将在第八 章讨论），也就是说根据样本的结果，对总体做出估计。关于大小为 1 的样本对总体总支出的 估计, 就是由抽中样本单元的支出乘以该单元的权数。而权数则由样本单元所代表的总体单 元个数得到。从表 6-5 中可以看出，用简单随机抽样时，总体的估计值随样本的不同而其变 化的范围从 75 000 元到 2 820 000 元，这说明总体估计值差别很大：对差别很大的总体，用简 单随机抽样的结果是不好的。如果采用 PPS 抽样，同样用样本量为 1, 其结果如表 6-6 所示。

表 6-6 所有可能的  $n=1$  的 PPS 样本

样本	农场的规模 (公顷)	入样概率	样本权数	农场的支出	估计总支出
农场 1	50	50/2 000	2 000/50	26 000	1 040 000
农场 2	1 000	1 000/2 000	2 000/1 000	470 000	940 000
农场 3	125	125/2 000	2 000/125	63 800	1 020 800
农场 4	300	300/2 000	2 000/300	145 000	966 667
农场 5	500	500/2 000	2 000/500	230 000	920 000
农场 6	25	25/2 000	2 000/25	12 500	1 000 000
总计	2 000	—	—	—	947 300

从表 6-6 可以看出, 对 PPS 抽样, 估计结果的变异性要小得多。六个可能样本得到的估计值的变化范围, 从最低的 920 000 元到最高的 1 040 000 元, 比简单随机抽样要好得多, 与实际的支出额 947 300 元都比较接近 (PPS 抽样的入样概率是按农场规模除以所有农场的总规模计算得到)。

本例中假定农场的支出与其规模大小有关, 这一假定在这里显然是成立的, 否则 PPS 抽样就不会这样有效。如果调查变量与规模大小之间不相关, PPS 抽样可能不比简单随机抽样好, 甚至可能更差。

## 2. PPS 抽样的方法

如何抽取一个 PPS 样本呢? 有三种主要方法, 它们是随机法、系统 PPS 抽样方法和随机系统抽样法。随机法包括随机放回的 PPS 抽样和随机不放回的 PPS 抽样, 由于随机不放回的 PPS 抽样比较复杂, 下面只介绍随机放回的 PPS 抽样。在讨论中, 假定单元大小皆为整数。

### (I) 随机放回的 PPS 抽样

首先, 对总体中每一个单元的规模大小进行累计, 然后确定总体每个单元与它相对应的代码范围; 在 1 到总的累计大小 (即总体所有单元规模大小的总和) 范围内, 抽取一个随机数, 该随机数相应的代码范围中的单元即为入样单元; 重复上述步骤, 直到抽得  $n$  个单元为止。下面以农场调查为例进行具体说明, 如表 6-7 所示。

表 6-7 使用随机放回的 PPS 抽样

农场	规模大小	累计	代码范围
1	50	50	1~50
2	1 000	1 050	51-1050
3	125	1 175	1 051-1 175
4	300	1 475	1 176-1 475
5	500	1 975	1 476 - 1 975
6	25	2 000	1 976-2 000

假设现在需要抽取一个由三个农场组成的样本，前两步已经在上面的表中完成，现在必须在 1 至 2 000 之间抽三个随机数，以得到三个样本农场。假定抽到的随机数是 1 697, 624 和 1 109，I 697 位于第 5 个农场的代码 1 476-1 975 之间，因此农场 5 是第一个抽中的样本单元，以次类推 624 和 1 109 抽中的农场是农场 2 和农场 3。很显然，单元的规模越大，用数字表示的代码范围越宽，被抽到的概率也越大。由于是有放回的，因此规模较大的单元，就有可能再次被抽中，该单元入样概率会增加，影响评估的结果。在不放回抽样的情形，如果抽选的单元多于一个，要使人样概率与单元大小成正比以及对调查估计值的误差进行估计都比较复杂。可参考有关抽样的专著。

### (2) 系统 PPS 抽样方法

系统 PPS 抽样方法前面两步的做法与随机放回的 PPS 抽样方法相同，都是在总体单元按一定顺序排列的基础上，先将单元的规模大小进行累计，对总体中每个单元确定与它相对应的代码范围。其做法不同的是，系统 PPS 抽样方法要计算一个抽样间隔  $R$ ，设累计的规模大小为将累计的规模大小除以样本量  $n$ ，则抽样间隔  $k = M/n$ 。然后在 I 到左的范围内抽取一个随机数  $r$ ，则与代码  $dr+2, \dots, 5(nT)k$  所在范围相对应的单元入样。前例中，假如仍要抽取三个农场为样本，先计算抽样间隔  $k = 2 000/3 = 666$ ，然后在 1 到 666 之间抽取一个随机数，假定为 526，则与  $526, 526+666 = 1 192, 526+2(666) = 1 858$  所在代码范围对应的农场即为样本，它们是农场 2、农场 4 和农场 5。

这里需要注意的是这种方法的单元有可能被重复抽中，当单元的规模大于两倍的抽样间隔时，就肯定会被重复抽中；此外，本例中抽样间隔如果严格计算并不是一个整数，因此要严格与规模大小成比例，最好用前面介绍的圆形系统抽样。

### (3) 随机系统抽样法

在这种抽样方法中，名录在系统抽样前要进行随机排序。如果完全按原来的排序，会把一些可能的样本排除在外。在上面介绍的系统抽样的例子中，不可能把单元 3 和 4 抽入同一个样本。把名录重新随机排列后会增加潜在的可能抽到的样本数量。

这种抽样方法也同样存在大于抽样间隔的单元有重复被抽中的可能。只有把这些大的单元放入单独的一个层，并独立地对它们进行抽样，才能克服这个问题。另一个问题是估计值抽样方差估计比较困难。

## 3. PPS 抽样的优缺点

(1) PPS 抽样的优点 主要是它使用了辅助信息，从而提高了抽样效率，与简单随机抽样相比能显著地减少抽样误差。

(2) PPS 抽样的缺点 ①抽样框中的所有单元，都要有高质量的、能使用大小度量的辅助信息；②抽样框的创建比简单随机抽样和系统抽样成本高，更复杂，因为总体中每一个单元的大小都需要度量、记录和存储下来；③并非在任何情况下都能使用，因为并不是每一个总体都有稳定且与主要调查变量相关的有关大小或规模的度量；④对于那些与大小变量不相关的调查变量，PPS 抽样可能会导致抽样效率比简单随机抽样还低；⑤估计量抽样方差的计算比较复杂；⑥当单元大小度量不准确或不稳定时 PPS 抽样不适用，此时最好的办法是将单元按大小分组并使用分层抽样：

### (四) 分层抽样

#### 1. 分层抽样的含义

分层抽样也称分类抽样或类型抽样，这种抽样方法是在抽样之前将总体分为同质的、互不重叠的若干子总体，也称为层。然后在每一个层中独立地抽取样本，在各层中的抽样方法可以选用比较简单的简单随机抽样、系统抽样，也可以选用较复杂的与规模大小成比例的概率抽样、整群抽样等，总之是随机抽选的。因此分层抽样也称为分层随机抽样。

采用分层抽样的前提是需要有各层的抽样框。假如在调查前总体单元的抽样框中所有单元都具有辅助

信息，就可以利用这些信息来对总体进行分层。若这种已知信息只是总体单元 所处的地理位置，我们就可以按地理位置分层，或按行政区域分层。如果有收入的信息，就可 以按收入高低对总体分层。通常用于分层的辅助变量还有：年龄、性别、家庭规模、企业类型、 资产总额、员工数等。

## 2. 采用分层抽样的主要理由

(1) 为了提高抽样效率。即在同样的样本量条件下提高估计的精确度，或在一定精确度 的条件下减少样本量以节约调查费用。要使分层抽样的效率高于简单随机抽样，要求层内各 单元必须有高度的同质性，即同一层内每一个单元的调查指标接近。而不同层之间差异应尽可能大。一般而言，如果分层变量和调查指标相关，以上两点就能做到。正如科克伦在《抽样 技术》一书中曾经解释了分层抽样能使估计值的精度比简单随机抽样高的理由：如果每一层 都是同质的，即单元的指标变化很小，则每一层的均值都可以通过该层的一个小样本得到精确 的估计，而将这些均值估计综合在一起，可以得到整个总体的一个精确估计。

分层还受到总体偏斜程度的影响。例如，工商业和农场调查，总体分布常有较大的偏斜，有一些工商业单位和农场，其指标值(如收入、支出、员工人数)特别大。因此这些单元对总体 的任何一个估计值都会有很大影响。如果对这样的单元在抽样方案中不规定特殊的处理办法，估计的抽样方差就会加大。因此这种情形的处理方法是：特别大的单位自成一层，仅仅代 表自己，不代表总体中其他的、可能小得多的单位。为此，在工商业和农场的调查中，常用一个 规模变量作为分层变量，可以是资产总额或员工人数，据此划分为大型、中型和小型。对大型 采用全面调查，对中型和小型采用抽样调查。这样做可以保证对估计值有较大影响的大单元 只代表它们自身，而不代表总体中的其他单元。如果工商业单元在大、中、小型之间销售额差 别大，而大、中、小型内部差别小，这样的分层就可以提高抽样效率。类似地，对于一个用地域 框的抽样方案，可能要将大城市组成单独的层，然后对中小城市进行抽样，以确保样本的代 表性。

上例中，员工人数是企业规模的一种度量，又与销售额高度相关，用其进行分层比较合理。如果一项调查指标是员工的年龄，按员工人数分层就没有任何意义，因为他们之间不相关。另 外，分层若对某个调查变量是有效的，对另一个调查变量则未必同样有效。通常，分层变量的选 择是按它们与最重要的调查变量间 的相关程度进行确定的。这意味着对那些重要程度较低，与 分层变量不相关的调查变量，分层样本估计值的效率有可能比简单随机抽样低。

(2) 需要对子总体进行估计。这些子总体通常称为“域”，在对 “域” 进行估计时，应确保 对需要估计的每个 “域” 都有足够的样本量。在做一项调查设计时，总的目标常是对总数进行 估计:如上月失业人数是多少？上个月的销售额是多少？但除了总数外，客户还常需要对某些 子总体进行估计。如客户可能希望了解男性的失业人数有多少？并把该数和女性的失业人数 相比较。或者，想知道上月服装店的销售额，甚至某省所有零售店的销售额。对某些子总体进 行估计称为“域估计”。如要进行 “域估计”，每个域中样本单元是否足够，就应该在抽样设计 时加以考虑。如果抽样框中有可用的信息，最简单的做法就是在每个 “域” 中建立一个单独 的“层”。

(3)通过分层可以避免得到很差的样本。在简单随机抽样的情况下，样本的抽选完全取 决于偶然的机会，分层抽样则试图确保总体中的所有部分在样本中都有代表，从而避免出现缺 乏代表性的极端样本。

除了这些原因外，选择分层也是为了操作或管理上的方便，使调查机构能控制现场调查工作在不同地区机构间的分布。例如，如果数据收集是按省进行的，则按省分层比较好。在这种 情况下，省级机构只须承担本省范围内的工作。

一旦总体被分了层，我们需要确定从各层中抽选多少个单元，这一步骤称为样本分配，这 将在第七章 中介绍。各个层的人样概率可能不相同，要计算入样概率，必须考虑各个层的样本 量和层的大小。

## 3. 分层抽样的优点与缺点

(1) 分层抽样的优点。①能提高对总体估计值的精度，从而使抽样效率更高；②能保证 样本对被定义

为层的那些子总体的代表性，从而得到有效的域估计；③操作与管理方便；④能避免得到一个“差的”样本；⑤在不同的层中可以使用不同的抽样框和不同的抽样方法（例如，在一个层中用简单随机抽样，在另一个层中采用 PPS 抽样）。

(2) 分层抽样的缺点。①要求抽样框中的所有单元，而不仅仅是抽入样本的那些单元都必须有高质量的、能用于分层的辅助变量；②由于需要辅助信息，对抽样框的要求更高，较简单随机抽样和系统抽样需要更多的费用；③如果调查变量与分层变量不相关，抽样效率可能比简单随机抽样还低；④估计值的计算比简单随机抽样和系统抽样稍为复杂。

## (五) 整群抽样

### 1. 整群抽样的含义

整群抽样是指将总体划分为若干个有联系的基本单元所组成的群，以群为抽样单元抽取样本的方法。它不同于前面介绍的简单随机抽样、系统抽样、PPS 抽样和分层抽样方法，它们在抽取样本单元时，都是直接抽取组成总体的基本单元，也就是说抽样单元与基本单元是一致的。而整群抽样中抽样单元与基本单元不一致。

### 2. 应用整群抽样的原因

(1) 抽样框对选用抽样方法有影响。前面曾经提到抽样调查需要有一个抽样框，它应包括所有的总体单元的名录或地图。抽样时需要编上号码，这样才能应用随机数表或其他方式从中抽取所需的样本。然而，有时候总体很大没有现成的名录，而要着手编制一个抽样框也十分费时、费工。如果总体单元是自然分成组或群的，创建一个这种群的抽样框并对它们进行抽样比创建总体中所有单元的名录框更为容易。例如，我们需要调查武汉市中学生中近视眼的比例有多大？就需要全武汉市中学生的名录，这是不太容易办到的。如果我们以中学作为抽样单元，编制一个中学的名录就方便得多。

(2) 抽选群能大大降低数据收集的费用。有时即使具备了基本单元的抽样框，但由于总体很大，分布很广，若采用简单随机抽样，势必使样本的地域分布很分散，调查所需要花费的人力和费用也比较大。如上例中，假如具有全武汉市的中学生名录，要从中抽 2 000 名中学生进行调查，抽中的样本很可能分布在全市的许多中学，调查起来会费时、费力。若能抽选若干中学，对抽中中学的全部学生进行调查，这样，样本单元比较集中，调查会更方便，费用也会更节省。

### 3. 整群抽样的具体方法

整群抽样过程如下：首先将总体划分为若干群，群内单元可以相等，也可以不等；然后抽选

群的样本；再调查样本群中的所有单元。在整群抽样中，总体被看作由不同级的单元组成。例如，人们居住在相应的小区，许多小区组成一个街区，所有的街区构成城市。

整群抽样的“群”大致可分为两类：一类是根据行政、地域及其他自然因素划分的群体，如学校、工厂、城市、街区、农场等；抽取这一类群主要是为了方便和节约费用。而另一类群是一个连续总体，调查设计者可以根据需要来划分群的大小。例如，一大块土地的面积可以划分成许多小面积的地块群，连续作业生产的产品可以以每5分钟生产量为一个群。

整群抽样的效率与群内各单元的相似程度有关。它要求群内的相似程度较低，也就是说群内有各种各样的单元，对总体的代表性就比较高，抽样效率也就比较高。不幸的是，在同一个群内的单元通常非常相似、从而导致整群抽样的效率通常低于简单随机抽样。例如，假定一个城市有100 000个人、我们要抽两个样本，第一个样本用整群抽样，随机抽取一个有400个人的街区，第二个样本用简单随机抽样，从全市居民的名录中抽400个居民。简单随机样本中的400人的收入、年龄、工作及受教育程度的差别，可能要比整群样本中都住在同一街区的400个居民的差别大得多，这是因为邻近的居民倾向具有相近的特征。

为了提高整群抽样的统计效率，抽取数量较多的小群比抽取少量较大的群的效率高。如果采用面访，样本的分布越广，调查的费用越高。“一般做法是选择最优群的数量和大小，同时使费用节约。”

整群抽样可能有一些具体困难，如果用地图作为地域框，可能很难决定某个居住地是否在一个特定的地域群内：必须建立一个基本原则以决定哪些单元在一个群内。例如，规定如“居住地属于它的正门所在的群”这么一条原则，就能解决大多数问题。（通常，所有的居住地不是在群内就是在群外）。如果一个居住地似乎平等地跨越两个地域群，则采用抛硬币的方法来确定，以避免产生任何抽选的偏差。例如，在加拿大的劳动力调查中，群是通过街中心画线决定的。这使得决定一个居住地是否在样本中变得很容易，但使群的形状有点怪。

如果总体单元既可以划为群作整群抽样，也可以划为层作分层抽样的情形，当进行整群抽样时，抽取的是群样本，被抽中的群中的每一个单元都要进行调查，而分层抽样则须在每个层中抽取样本。一般而言，相对于简单随机抽样，分层抽样通常能提高精度，而整群抽样则会降低精度。但从费用来看，整群抽样能节约费用。另外还可以注意到，整群抽样和分层抽样之间的关系，整群抽样是在全部群中抽取若干群，再对抽中群的全部单元进行调查；分层抽样则是对所有的层都要进行抽样，但在每一层中并不对全部单元进行调查，而是抽选部分单元进行调查。整群抽样要求群内各单元相似程度低，才能提高抽样效率；而分层抽样则相反，它要求层内相似程度高，才能提高抽样效率。

#### 4. 整群抽样的优点与缺点

(1) 整群抽样的优点。  
①费用较低。由于样本不像简单随机样本那样分散，整群抽样能大大降低收集数据的费用。这一点在总体分布很广而且要使用面访时特别重要，因为整群抽样能大量节省调查员在路上的旅行时间，这对于分布在农村的总体尤为明显。  
②创建自然形成群的名录较容易。当总体单元为自然形成的群（例如住户、学校）时，整群抽样比简单随机抽样或系统抽样更容易。因为对后两者都需要创建一个基本单元的名录，这样的名录建立很困难且费用高。  
③一定条件下有较高的抽样效率。当调查变量群内单元差异较大，而群与群之间的差异较小时，整群抽样的效率将高于简单随机抽样。

(2) 整群抽样的缺点。  
①对于调查变量，若群内单元有趋同性，则整群抽样的统计效率比简单随机抽样低，而这正是通常遇到的情况，但对此项效率的损失可通过增加群的抽取数量来弥补。  
②通常无法提前知道调查总样本量，因为在进行调查前，我们通常不知道一个群内有多少个单元。  
③调查的组织比其他方法复杂。  
④方差估计往往比简单随机抽样更为复杂。

### (六) 多阶段抽样

#### 1. 多阶段抽样的含义

•多阶段抽样也称多级抽样，它是用两个或更多个连续的阶段抽取样本的过程。第一阶段 抽取的单元称为初级或一级抽样单元，用英语单词的第一个字母表示为 PSU, 第二阶段抽取的 单元称为次级或二级抽样单元，表示为 SSU, 依次类推。每个阶段抽取的单元在结构上是不同的，与整群抽样中一样，单元是分级的。在两阶段抽样中，次级抽样单元通常就是总体的基本 单元。

## 2. 多阶段抽样的方法

(1) 常规方法。常见的多阶段抽样是二阶段抽样，第一阶段抽样用地域框抽取小的地理 小区作为 PSU, 第二阶段用系统抽样抽取小区内的居住地为 SSU。在两阶段抽样中，从每个抽 中的群(一级)单元中再抽一部分(二级)单元进行调查。它不同于前面介绍的一阶整群抽样 中，样本群中的所有单元都入样的选样方法。

多阶段抽样常常与地域框结合使用，作为解决整群抽样效率低的一种方法。实际上，很少 使用纯粹的整群抽样，如果群内的邻近单元相似，那么在抽取同样数的基本单元时，要多抽初 级抽样单元，在每个初级抽样单元中抽取较少的基本单元。这样，比少抽初级抽样单元，而在 每个初级抽样单元中多抽基本单元的抽样效率高。

(2) 阶数的选择。多阶段抽样中的阶数可以是任意的，但因为阶数越多，设计就越复杂，估计也更复杂。通常我们只用二阶段或三阶段抽样。要注意的是，第一阶段的抽样框通常是 非常稳定的，例如，覆盖大地理区域的地域框，一般不会随时间很快发生变化，最后一阶段的抽 样框通常不够稳定，这些抽样框常常是在实地调查中现场建立的名录框。在多阶段抽样中，要 提高抽样精度，第一阶段的样本量起着十分重要的作用，但是多抽取第一阶段的样本，会使样 本的分布比较分散，这就会增加费用，因此要在精度和费用之间适当平衡。

(3) 多阶段中的抽样方法选择。在多阶段抽样的各阶段中可以使用任何一种抽样方法。因此，多阶段抽样的主要优点之一是灵活。例如，对某个第一阶段抽样中的初级抽样单元，第 二阶段采用简单随机抽样，而对另一个样本初级单元，如果有与主要调查变量相关的大小度量 可以利用，则第二阶段抽样可以采用 PPS 抽样。中国国家统计局进行的许多全国性调查都是 多阶段抽样。例如，全国的农产量调查、居民的住户调查、中小工商企业调查等由于样本单元 遍布全国各地，显然不可能直接一次抽到所需要的样本，只能分几个阶段来逐级抽取。但是每 增加一级抽样，就会增加抽样设计的复杂性，同时还会增加抽样误差。因此，要在满足信息需 要的前提下尽可能减少阶段数。

加拿大劳动力调查所用的样本就是一个多阶段样本的实例。整个国家被划分为 1 100 多 个层，每个层由一组普查小区组成。普查小区是由人口普查时所确定的地理区域，普查区根据 一个普查员能力所及条件下的地域大小和人口密度划分，普查员能对它所覆盖的区域进行详 细调查。劳动力调查的第一阶段抽样是从每个层中抽取普查小区或普查小区组的样本。在第 二阶段抽样中，将这些小区绘图，列出它们中的所有居住地，并从中根据系统抽样抽取居住地 C 然后，调查抽中居住地内的所有人。

最后要注意的是，在上面的举例中，第一阶段抽样所用的都是地域框，但这并不是多阶段

抽样的一个必要条件。多阶段抽样可以用其他形式的抽样框。例如，在机场进行的旅客调查，它使用的初级抽样单元是时间单元，是一个月中的各天，第二级抽样单元是旅客对更复杂的调查，第二级抽样单元可以是到达的客机，第三级抽样单元则是飞机上的实际座位。

### 3. 多阶段抽样的优点与缺点

(1) 多阶段抽样的优点。①以调查变量衡量的群特征差异不大时，多阶段抽样比单阶段整群抽样的效率更高。②整群样本的分布比简单随机样本集中，因此采用面访调查的时间和费用都大为降低。③不需要有整个总体的名录框，所要的只是在每个阶段都有一个好的抽样框。换句话说，只要有关于初级抽样单元的好抽样框以及在以后各阶段抽样中有被抽中单元的完整名录即可。

(2) 多阶段抽样的缺点。①虽然多阶段抽样的效率比单阶段整群抽样要高，但它没有简单随机抽样的效率高。②通常不能提前知道多阶段抽样的样本量，因为在具体调查前，我们不知道在每个大单元中到底包含多少个下级单元(如果固定每个大单元中的抽样数目，则样本总量也可控制)。③与单阶段整群抽样相比，调查组织较为复杂。④估计值与抽样方差的计算较为复杂。

### (七) 多相抽样

#### 1. 多相抽样的含义

多相抽样是指在同一抽样框中，先抽一个含有很多单元的大样本，收集基本的信息，然后在这个大样本中抽一个子样本，收集更详细信息的抽样方法。多相抽样与多阶段抽样虽然名称相近，但内涵却完全不同。虽然多相抽样也要抽两个或两个以上的样本，但所有的样本都是从同一个抽样框中抽出的，而且每一相的单元在结构上是相同的。多相抽样最常见的形式是二相抽样，或称二重抽样，但三相或更多相也是可以的。然而，使用多相抽样时，相数越多，抽样设计与估计就越复杂，这一点和多阶段抽样类似。

#### 2. 使用多相抽样的原因

(1) 为了获得分层信息。应用多相抽样的原因是当抽样框缺乏辅助信息，而又想对总体进行分层或筛选部分总体的情况。例如我们前面曾经介绍分层抽样，把总体单元中相似的单元分在一起，缩小层内各单元之间的差异，可以提高抽样效率，但这需要有关分层的信息。然而有时候并没有现成的总体辅助信息可供使用。一种补充的方法是采用先抽取一个较大的样本以获取有关总体的信息，然后在此基础上再抽一个较小的样本，利用大样本得到的辅助信息进行分层，可以改进估计，这就是二相抽样。又如，假定对农场进行调查，要了解养牛的情况，而抽样框只列出了农场，没有所需的辅助信息。我们可以进行一个简单的调查，只问一个问题：“你的农场养牛吗？”由于只有一个问题，该调查的单元调查成本会很低(如用电话调查)，因此，我们能抽一个大样本，一旦这个大样本被抽取，就能从养牛的农场中抽取第二个较小的样本，并询问更详细的问题。这种方法能避免调查那些不在范围内(即不是养牛场)的单元，节省调查费用。显然，在这里需要指出，在抽取一个较大样本以获取总体信息时需要增加一定的费用，但利用这些辅助信息又可以提高效率或节省费用而有所得益，因此，只有当得益大于费用时采用多相抽样才合算。

(2) 经费有限条件下获得更详细信息。当没有充足的预算经费收集整个样本的信息时，也可以用多相抽样抽取子样本的方法来收集更详细的信息。加拿大零售商品季度调查就是这样一个例子。此调查的第一相是商业的批发零售的月度调查。每个月，商业的批发零售调查要求批发零售基层单位只提供两个数据：月度销售额和库存量。零售商品季度调查再从这些零售基层单位抽取一个子样本，在调查时要求它们把销售额分解到各类商品中，如服装、电器、食品等。

类似地，多相抽样可以用于在一项调查中不同问题的数据收集费用差异很大的情形。例如，一个健康调查，它要问很多基本问题，涉及饮食、吸烟、运动、饮酒。此外，该调查还要求被调查者能参加一些直接测量，例如在踏车上跑步并测量血压和胆固醇指标。相对而言，问几个问题不需花很多费用，但医学测量却需要占用一个经过良好训练的保健工作者的时间，还要使用实验室的相应设备，因而相对来说费用较高。

这个调查就可以使用二相抽样，对第一相样本 单元只问基本问题，而对较小的第二相样本单元则需进行直接测量。

### 3. 多相抽样的优点与缺点

(1) 多相抽样的优点. ①与简单随机抽样相比，能显著提高估计值精度。②能用来获得抽样框中所没有的辅助信息(特别是分层信息)。③适用于某些调查指标的数据收集费用 特别高，或会给被调查者带来较重负担的情况。

(2) 多相抽样的缺点。①如果需要根据第一相的结果来进行第二相调查，得到整个调查 结果的时间比单相调查的时间要长。②由于对某些样本单元访问次数超过一次，故所需费用 比一相调查要多。③调查的组织比较复杂。④估计值和抽样误差的计算相当复杂。

## 第四节

### 抽样设计中的注意事项

#### 一、首先要考虑是使用概率抽样还是非概率抽样

在考虑采用概率抽样还是非概率抽样时，首先要根据调查目的，判断是否要以样本的结果 为基础对总体进行推断。如果是要对总体进行推断，则应选用概率抽样，如果只是了解一些概 况或者是正式调查前的预调查，则可用非概率抽样。

#### 二、考虑概率抽样时需要合理选择抽样方法

1-在考虑选择哪一种概率抽样方法时，要结合抽样框的情况和收集数据的方法来决定。例如，是否具备或有可能建立一个包括所有总体单元的名录框。如果具备这样的名录框，而且 总体单元的分布并不很广，总体单元之间的差别也不很大，则可考虑采用简单随机抽样。如果 上述条件并不具备，则考虑其他抽样方法。

2. 调查是否要求对一个相当大的地理区域进行面访？如果是，则不宜采用简单随机抽 样，而应考虑使样本单元相对集中，例如采用一阶段或二阶段整群抽样。

3. 总体单元是否自然形成群体或是否有此类群体的名录，如住户、机关、学校等，或者是否打算使用地域抽样框？若总体单元自然形成群体，有关群体的名录也有现成的，或建立这样 的名录比建立一个总体单元的名录要便宜得多，或者可以用地域抽样框，则应考虑使用一阶段 整群抽样或多阶段抽样。

4. 是否具有辅助信息的抽样框？如抽样框中包括年龄、性别、省份、收入、农场规模 和企业规模等。如果有，而且这些变量与调查变量之间相关，则可以考虑采用分层抽样，

如果有准确的且与主要调查指标相关的大小度量，则可以考虑使用 PPS 抽样。如果大小 度量不太准确，或与主要调查指标的相关性不太大，则考虑按度量大小分组并以此分层。

5. 是否需要作“域估计”，“域”能否在抽样框中确定？如果是，则考虑将“域”作为层进行 分层抽样，若不是，考虑采用多相抽样。

6. 如果抽样框缺乏分层的信息也缺乏对总体单元进行筛选的辅助变量，而又打算分层，则考虑二相或多相抽样。先抽取一个相当大的第一相样本，用以收集第二相样本中所需要的 分层信息 Q

7. 如果在搜集的信息中有部分信息收集起来费用比较高, 如需要专门的技术人员当面进行直接测量, 这种情况可以考虑采用多相抽样, 只用较小的第二相样本来收集那些费用较高的 数据。

### 三、重复性抽样调查需要考虑的问题

在抽样设计中, 一次性调查和重复性调查在很多方面是不同的。重复性调查也称连续性 调查, 其目的通常是研究调查指标在一段时间内的趋势或变化。重复性调查的一种特殊类型 是固定样本连续调查, 即对同一样本单元多次收集数据。这样的调查通常要度量总体指标的 变化, 在相同样本量的条件下, 固定样本连续调查比抽取一系列独立的样本调查更为精确。

#### 1. 固定样本调查的优点

如果一项调查要重复进行, 与一系列独立的抽样调查相比, 固定样本调查的优点如下:

(1) 减少对变化(即 $\bar{\gamma}-\bar{\gamma}$ )估计的抽样误差。这里 $\bar{\gamma}$ 是时间 1 的总量,  $\bar{\gamma}$ 是时间 2 的总 量、例如我们估计的可能是两个月失业人数的变化。

(2) 能得到被调查者随时间变化的信息。

(3) 能减少回答误差, 因为连续调查, 被调查者对问卷有更好的理解 Q

(4) 费用会随时间而减少。因为调查的设计开发、计算机编程、工作人员培训等费用, 一 次性投入, 较长时间使用。

#### 2. 固定样本调查的缺点

与独立样本相比, 固定样本调查的缺点是:

(1) 估计与无回答的处理更复杂。

(2) 需要在较长时间内保证费用。因为连续性调查只要样本存在就必须保证其预算。

(3) 使样本保持在长时间内的代表性比较困难。因为总体将随时间发生变化, 如新单元 的产生(诞生)和某些老单元的退出(消亡), 或分类的变化。换句话说, 在第一时段具有很好 代表性的样本, 在以后的时段不一定能同样好地代表总体。

(4) 可能会增加回答误差。例如, 由于被调查者对问卷有一定的了解, 他们中的一部分人 可能会为了加快面访的进度而不正确地回答问题。

(5) 回答负担比较重。同一个问题被调查者在不同时间被多次调查, 随着时间的推延, 导 致被调查者的疲劳而可能增加无回答。

(6) 组织工作比一次性调查复杂得多。

(7) 容易产生诱导性, 如多次问一个被调查者看医生的事可能使他开始看医生。

(8) 会造成某些概念上的模糊 例如, 随着时间的推移, 住户构成会发生变化, 因此给连 续观察下如何定义一个住户带来不确定性’ ,

(9) 如果一开始抽到的就是一个 “差” 的样本, 也不得不继续使用下去。

在重复性调查设计中, 一般通过样本轮换处理诸如新生、死亡及大小度量变化这样的问题。详细内容可参考相关书籍。

### 四、敏感性问题的抽样设计

在调查中, 常会涉及一些诸如收入、吸毒、作弊等敏感性问题。如果对被调查者直接询问, 可能会招致拒答或不真实回答。因此, 如何采用适当的方法, 既能保护被调查者隐私, 又能使 其讲真话, 这也是抽样设计中应考虑的问题。现介绍两种有效的处理方法。

### 1. 随机化的回答

这个方法由沃纳 1965 年提出。它向被调查者提出两个问题。假如要调查对改革的看法，这两个问题可以是：

A：您赞成改革吗？1 一是，2 一不是

B：您不赞成改革吗？1 一是，2 一不是

被调查者随机抽一个问题回答。调查员不知道每人回答哪个问题，但回答 A 类问题的人 占的比例 P 是他事先确定的。例如向 100 个人作调查，取  $P = 0.7$ （注意 P 不能取 0.5）。可以 制作 100 张卡片，其中 70 张上印上问题 A, 30 张印上问题 B, 让被调查者随意抽取卡片。当被 抽到的问题与自己的情况一致时，回答“是”；相反时，回答“不是”。对上面的问题：如被调查 者赞成改革且抽到 A 或不赞成改革且抽到 B 都应回答“是”，如被调查者不赞成改革且抽到 A 或赞成改革且抽到 B 都应回答“不是”。显然，这个方法要取得成功，关键在于让被调查者确 信：调查员不知道他（或她）回答的是哪一个问题。

### 2. 提出无关的第二个问题，进行随机化回答

沃纳的方法虽然比直接提问敏感性问题好，但所提的两个问题都具有敏感性，被调查者可能仍存有戒心，不予配合。而且，这个方法中回答问题 A 的人占的比例不能等于  $1/2$ 。西蒙斯 作了改进，问题 A 仍为原来的敏感性问题，把问题 B 换成与问题 A 无关的、毫无敏感性的问题。例如：“您是五月份出生的吗？您的住址中最后一个数字是奇数吗？”等等。这样，被调 查者的合作态度可能会有所改进。更具体的介绍参见相关文献。

## 恒] 痞章小結•…

抽样调查是指从总体中抽取一部分个体作为样本进行观察，由样本调查结果来推断总体 的一种调查方式。抽样调查的目的就是用样本推断总体，只需调查总体的一小部分（即样 本），具有费用低、效率高、准确和有代表性等特征。抽样设计程序包括五个步骤：界定目标总 体，确定抽样框，选定抽样方法，确定样 本量，制定抽样计划。抽样技术有两类：非随机抽样和 随机抽样。非随机抽样是指抽样不按照随机原则，而是根据调研者的主观判断或方便来抽取 样本，包括方便抽样、志愿者抽样、判断抽样、配额抽样以及滚雪球抽样等。随机抽样是按照随 机原则抽取样本，包含简单随机抽样、系统抽样、PPS 抽样、分层抽样、整群抽样、多阶段抽样和 多相抽样等形式。每一种方法都有其各自的优点和缺点，需要根据研究目的、调查 对象的特 点、辅助信息多少、经费与时间限制、调查的精度要求等进行合理选择。

## 礲关梶名词••••

抽样调查 随机抽样 简单随机抽 系统抽样 PPS 抽样 判断抽样 分层抽样

整群抽样非随机抽样方便抽样志愿者抽样配额抽样滚雪球抽样

肯思考题 .....



即测即评

1. 如何区分普查与抽样调查?
2. 抽样方案设计程序是什么?
3. 随机抽样技术包括哪几种? 各有何优缺点?
4. 非随机抽样技术包括哪几种? 各有何优缺点?
5. 抽样设计时需要考虑哪些因素?

## 碧本章实训. ...

### 一、实训目的

1. 要求学生掌握简单随机抽样、系统抽样、分层抽样和整群抽样的实施方法。
2. 通过实训掌握抽样技术中重要的概念、技巧与设计思路。

### 二、实训内容

#### 1. 调查对象

以所在大学的本科生与研究生为调查对象，收集学生在学校学习、生活及电子产品消费等 情况 或者是前一章自选调查项目中确定的调查对象。

#### 2. 具体任务

首先，到相关机构或学校相关部门(如教务处或招生就业处、学生处、研究生部)掌握调查 对象的总体规模及构成情况。

其次，按照简单随机抽样、系统抽样、分层抽样和整群抽样的要求，对调查对象的抽样方案 进行设计 和实施

#### 3. 具体要求

- (1) 通过有关部门获取在校学生详细资料，并按学院、专业、班级三级建立抽样框，即学校 抽取出需要调查的学院，学院中抽出需要调查的班级，抽中的班级则进行全面调查，
- (2) 根据学生的学号分布，对学生逐一编号，建立学生抽样框。
- (3) 根据简单随机抽样、系统抽样、分层抽样和整群抽样的抽样要求对学生进行调查。

### 三、实训组织

1. 将教学班级学生按 4~6 人的标准分成若干小组，每个小组采用组长负责制。每组选择 一种抽样方法，按照实训目的与要求，合理分工，团结协作，建立抽样框。
2. 完成抽样方案设计的文本文件。
3. 各小组充分讨论设计完成的抽样方案，认真听取全体组员的意见，分析、修改和调整， 在此基础 上形成小组的 PPT 实训报告文稿。
4. 各小组在班级进行实训汇报和展示。
5. 指导教师进行综合评定和总结



## 第七章样本量确定

### 、导语

、 … .. . “.. “- “ “ — • . • •

需要多大的样本量？这是在调查设计过程中经常碰到的一个问题。由于样本量直接影响到调查的费用、完成调查所需要的时间、确定所需调查员的数量以及其他一些重要的现场操作限制条件，所以调查管理人员急于得到调查设计中这个基本问题的答案。

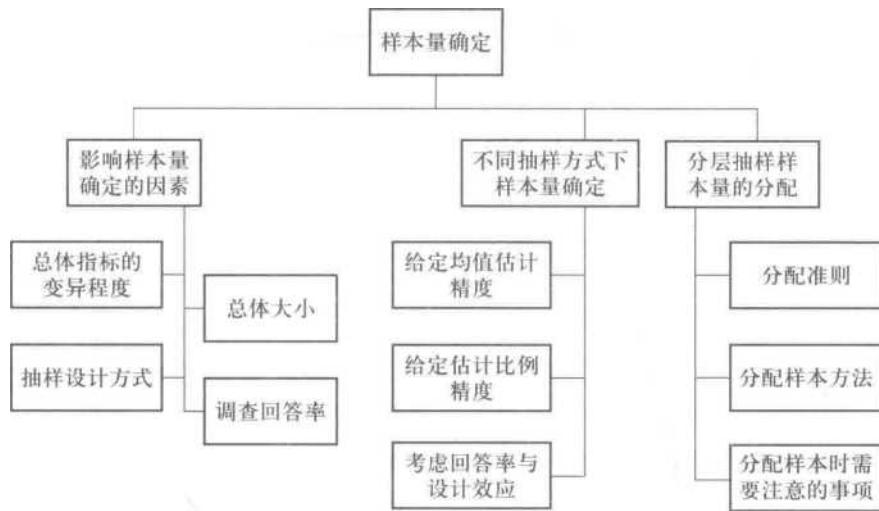
样本量的确定是在多种约束条件下进行折中的过程。影响调查样本量的因素包括总体指标的变异程度；总体的大小；样本设计、估计量的精度以及回答率的高低，实际调查运作的限制也是影响样本量的主要因素，虽然样本量计算公式没有涉及这些限制，但是在确定调查最终样本量时必须加以考虑，

确定样本量是为控制抽样误差而不是非抽样误差。为了确保抽样结果的准确，应尽可能使非抽样误差引起的偏差最小。由于只有在概率抽样条件下，才能对精度进行计量，所以本章的内容仅适用于概率抽样的样本量确定方法。

### 、教学目标与要求

1. 熟悉精度的表示与计算；
2. 掌握确定精度需要考虑的因素；
3. 理解影响样本量的各个因素；
4. 掌握简单随机抽样样本量的计算方法；
5. 掌握其他随机抽样样本量的计算方法；
6. 掌握分层抽样样本量的计算及其分配方法。

## 、本章知识结构



## 第一节

### 影响样本量确定的因素

市场调查的目的是要保证主要指标的估计精度，实际上调查指标的估计值精度就决定了调查所需要的样本量。因此，影响样本量确定的因素就是影响调查估计值精度的因素。

#### 一、精度的表示方法

在第六章中，我们介绍了精度的概念及其测度方法。无论估计量形式如何，精度都与其抽样方差或抽样误差有关。调查估计值的精度与样本量密切相关，随着样本量的增加，调查估计值的精度也会不断提高。换句话说，样本量取决于调查估计值所要求的精度。

但精度还有另外的表示方法，它也可以通过其他的误差指标，如标准差、误差限和变异系数来表示。最常用的是一定置信概率下的置信区间。因此有必要复习一下区间估计的概念。估计量的标准差(即标准误差)是估计量抽样方差的平方根。误差限是标准差的倍数，倍数因子取决于我们在调查估计中所希望达到的置信水平。对于具有正态分布的估计量，置信度为95%的置信区间意味着在同样的条件下，如果重复多次调查，100次中有95次置信区间会覆盖总体真值。

例如，我们用 $\hat{\theta}$ 表示一个任意的估计量，它可以是样本均值、样本比例或样本方差。在给定标准差估计挤出的情况下，估计量的置信区间可表示为：

其中， $\delta$ 是误差限， $Z$ 是对应于某一置信水平标准正态分布的分位点值，该值可从标准正态分布表中查得。大多数统计学教材中都有这样的统计表。常用的 $Z$ 值包括：

90%的置信度，对应的 $Z$ 值为1.65

95%的置信度，对应的 $Z$ 值为1.96

99%的置信度，对应的 $Z$ 值为2.58

置信度或置信水平通常叫可靠性，精度与可靠性的关系有绝对和相对两种表示方法，即：

$$P(|\hat{\theta} - \theta| < za) = 1-\alpha$$

$$\text{F号}(*)=1$$

其中， $1-\alpha$ 为置信度或称可靠性， $\alpha$ 为显著性水平，取值一般为0.05, 0.01或0.1。通常也有人把 $\alpha$ 称为置信度。

把 $3-\alpha$ 称作可靠性为的估计精度，由于 $\theta$ 是待估的参数，是未知的，因此可用 $\hat{\theta}$ 来代替

8，从而得到近似的估计精度估计量的误差限越小，精度越高。

2

#### 二、确定精度需要考虑的因素

确定调查需要达到的合适精度水平时，调查机构需要考虑以下相关的问题：

(一) 允许多大的误差限

对于调查估计值来说，多大的抽样方差可以接受？

在调查估计时，客户能容忍多大的不确定性？例如，常用的 95% 置信度±6% 的误差限对客户的调查目标是否适宜，估计值是否需要更高（或更低）的精度。

如果调查结果将用于一项有重大影响或有较大风险的决策，那么客户对估计值就需要有较高的精度；如果他们只是简单地希望取得对所研究总体某个指标的感性认识，那么低一点的精度就可以满足要求了。

### （二）是否需要对调查总体中的子总体（域）进行估计

调查结果可能需要包括一些细化的数据，这些数据称为域估计值，曾经在第六章讨论过。例如，除了需要对全国进行估计外，可能还需要省或地市的估计值，或者还需要在调查总体中按性别、年龄或受教育程度等因素细分成子总体（域），并对这些子总体进行估计。为满足这些数据的估计，需要确定合适的精度。当然不同的域可能精度有不同的要求。例如对一项全国性调查，调查的主办者可能要求对全国估计的精度，其误差限为±3%，但对于省一级的估计值，其精度要求误差限为±5% 就可以满足。而对于地、市级的估计值，±10% 的误差限就足够了。在这种情况下，通常每个域作为层来处理。

### （三）调查估计值有关的抽样方差大小

相对于调查结果要求的一定精度，最小的调查估计值多大？设一项调查的目的是多个指标的比例估计，其中有些指标的比例估计可能是  $p = 0.5$  或更高，但另外一些指标的比例估计则可能较低，如  $P = 0.05$  或  $p = 0.1$  在确定调查估计值所需的精度时，调查机构应该考虑当达到某个既定精度时可能的最小估计值。如果最小的估计值是  $P = 0.5$ ，那么调查机构要求的误差限就应该小于 0.05c

举例说，政府要决定是否实行“为某种使用人数较少的语言提供服务”的新政策，如果至少有  $P = 0.05$ （或 5%）的人群对这一语种存在需求，政府就应为这种语言提供服务。因此计划对当地居民进行一项调查、以估计他们对这种语言的需求一显然，对于  $P = 0.05$  水平左右的调查估计值，±0.05 的误差限似乎过高。在这种情况下，应规定更小的误差限，如不大于±0.01、±0.02 等，这时候，置信区间应该是  $0.05 \pm 0.01$  或  $0.05 \pm 0.02$ 。

### （四）需要在精度和费用间平衡

随着样本量的增加，估计值的精度也将提高。但是，精度的得益并不与样本量的增加成正比，考虑只有两个选项“A”和的定性变量，客户对总体中性质“A”的个体所占的比例感兴趣。表 7-1 描述了 95.45% 置信度下，对不同样本量的简单随机样本进行比例估计得到的误差限，其中选择 A 的个体占总体的真实比例值是  $P = 0.5$  或 50%。

表 7-1 简单随机抽样估计比例  $\pi$  的样本量与误差限的关系 ( $P=0.5$ )

样本量 ( $n$ )	误差限 ( $2 \times \sqrt{\frac{p(1-p)}{n}}$ )
50	0.14
100	0.10
500	0.045
1 000	0.032

续表

样本量 ( $n$ )	误差限 ( $2 \times \sqrt{\frac{p(1-p)}{n}}$ )
100	0.10
500	0.045
1 000	0.032

表 7-1 中，当样本量从 50 增加一倍到 100 时，比例估计值的误差限有所降低，从±0.14 减少到±0.10，

并没有如我们可能想象的那样，减半至 $\pm 0.07$ 。同样，对于 500 的样本量，或者 5 倍于 100 的样本量，误差限也只是减少到 $\pm 0.045$ ，也没有与样本量增加 4 倍相对应，减少到原来的  $1/5$ ，即从 $\pm 0.1$  减少到 $\pm 0.02$ 。最后，样本量从 500 增加到 1 000 时误差限也没有减少一半。由此可见，样本量和误差限之间并不是线性关系。

上述例子表明，调查机构和客户必须判断花费更多人力和物力调查 1 000 人而不是 500 人，以使误差限从 $\pm 0.045$  减少到 $\pm 0.032$  是否值得。

为得到最小的误差限而选择最大可能的样本量并不总是最佳的解决办法。有时我们可以 在有效利用现有资源的基础上，获得相对精确的估计结果，接受一个较大的误差限，采用一个 较小的样本而不是一个 较大的样本而节省下来的费用，可用来调整其他影响调查结果精度的因素，例如减少无回答率（如对无回答者进行追踪回访、测试问卷、培训调查员等等），这样做 也许更有效。一般来说，为确定调查估计值所要求达到的精度，调查机构除了需要考虑基于估 计值将做的决策之外，还需要考察数据汇总与分析对调查估计的影响」

### 三、影响调查估计值精度的因素

调查所需样本量取决于调查估计值要求的精度，而精度又受到以下因素的影响，即总体指 标的变异程度、总体大小、抽样设计方式和调查回答率。下面分别介绍。

#### （一）总体指标的变异程度

在调查总体中，我们所研究的变量或指标会随着个人、住户、企业或农场等的不同而不同 虽然我们不能控制这种变异性，但它的大小却影响给定精度水平下对研究指标估计所必需的 样本量。

假设首次进行一项调查，目的是估计顾客对某企业所提供的服务的满意程度。对“顾客满 意度”这一指标，设置了两个可能的选项：满意或者不满意表 7-2 中列出了持满意和不满意 态度的顾客可能占的比例

表 7-2 中的第 1 种和第 11 种的变异程度完全相同，即“顾客满意度”指标没有任何变 异——所有顾客都持相同的态度。同样，第 2 种和第 10 种这两组反映了“顾客满意度”的 变异性很小，90%的顾客持一种意见，另外 10%的顾客持相反的意见。类似地，下列的每一 组内，即（3 和 9）、（4 和 8）、（5 和 7）呈现了相同的变异程度. 从第 1 到第 6 或者从第 11 到 第 6，“顾客满意度”指标的变异程度在不断递增。数字 6 这一组代表有 50%的顾客表示满 意，而另外 50%的顾客表示不满意，这是总体中有关“顾客满意度”的变异程度最大的情况。 如果所有的顾客对所提供的服务都表示满意，这就意味着“顾客满意度”没有变异，此时只要有一个顾客的样本就可以提供关于“顾客满意度”的完全可靠的估计值：随着调查总体 中所研究指标的实际变异程度的增加，样本量也必须随之增大，这样才能使所研究指标的 估计值具有较高的精度。

表 7-2 对服务表示满意和不满意的顾客所占比例及其变异程度

序号	满意比例/%	不满意比例/%	变异程度（比例方差）/%
1	100	0	0
2	90	10	9.00
3	80	20	16.00
4	70	30	21.00
5	60	40	24.00
6	50	50	25.00
7	40	60	24.00

8	30	70	21.00
9	20	80	16.00
10	10	90	9.00
11	0	100	0

若总体指标变异性很大,或具有所研究特征的单元数量很少,对这样的总体,要求精确估计就需要较大的样本量。当研究指标具有最大的变异程度时,调查需要的样本量也最大。对于只取两个值的指标,即二元变量,则当这两个值在总体中以 50% 和 50% 的比例出现时,总体指标的变异程度最大因此,为确定调查所需的样本量,首先需要得到目标总体的研究指标变异程度的估计值。在通常情况下,变异的真值未知,需要从过去相关主题的调查,或从预调查中得到它的估计值。一旦实施调查,调查机构就会意识到,如果所研究指标的实际变异程度大于确定样本量时估计的变异程度,那么,调查估计值的实际精度就会低于期望的精度。相反,如果所研究指标的实际变异程度比所估计的变异程度小,调查所得到的估计值会比预计的更精确。

为确保达到调查要求的精度,在计算样本量时,建议对某一指标的总体变异程度采用较为保守的估计。换句话说,在实际中如果事先不知道调查中要测量指标变异程度的数据,那么最好的办法是假定研究指标具有最大的变异程度。例如对于二元变量,应该假定总体中该变量的变异程度为 50%,即假定  $P = 0.5^{\circ}$

抽样调查时,调查指标通常不止一个,每个指标的变异程度可能互不相同。对某一指标来说足够大的样本,对变异程度更大的另一个指标来说可能就有些偏小。因此,为确保样本量对所有的研究指标都足够大,应该根据最大变异程度或被认为最重要的指标来确定样本量。

## (二) 总体大小

在样本量确定过程中,人们常常以为总体越大,应抽的样本就应该大,其实总体所起的作用因它的大小而有所差异。对于小规模总体,它起着重要作用;对于中等规模的总体,其作用也是中等;而对大总体其作用最小。例如,在顾客的满意度调查中,假定持满意态度顾客的真实比例是  $P = 0.5$ 。调查中,使用简单随机抽样抽取样本,并以 95.45% 置信度下 0.05 的误差限(即置信区间为  $0.5 \pm 0.05$ ),对  $P$  进行估计,表 7-3 显示了不同大小总体所需的样本量。

表 7-3 简单随机抽样条件下不同大小总体所需的样本量

总体大小	所需的样本量
50	45
100	80
500	223
1 000	286
5 000	371
10 000	385
100 000	399
1 000 000	400
10 000 000	400

由此可知,为满足要求的精度水平,随着总体大小的增加,样本量增加的比率逐渐减小到零。对于单元总数为 50 的调查总体,需要 45 个样本,对于两倍于此的调查总体,我们并不需要将样本量翻倍,即 90

个有效样本。对于  $N = 5\,000$  或更多的调查总体，所需的样本量快速地逼近  $n = 400$ 。因此对于简单随机抽样，在真实总体比例是  $P = 0.5$  的情况下，400 份有效问卷对于大于 5 000 的总体，已足以满足给定的精度要求对于很小规模的总体，通常必须调查较大比例的样本，以取得所期望的精度。因此在实际操作中，对小规模总体经常采用普查

### （三）抽样设计方式

计算样本量时，通常假定采用的抽样为简单随机抽样。当使用复杂的抽样设计时，估计值可能比简单随机抽样精确，也可能没有简单随机抽样精确。当估计值更精确时，称所采用的样本设计更为有效’如果在确定样本量时是按简单随机抽样的公式计算的，那么需要使用实际抽样设计的效率对它进行调整。

在分层抽样中，当分层的变量与调查指标相关时，所得的估计值通常比相同样本量的简单随机抽样更精确，或者至少一样精确。另一方面，整群抽样的估计精度通常低于使用同一估计量进行估计时简单随机抽样估计值的精度（因为群内相邻单元通常比较相似），

一般说来，当样本量采用简单随机抽样的计算公式，而实际使用的是更复杂的抽样方式时，为达到给定精度所需的样本量，应该在此基础上乘以设计效应因子。我们知道，简单随机抽样的设计效应等于 1；分层抽样设计效应小于等于 1；整群抽样设计和多阶抽样的设计效应一般大于等于 1。

那么如何获得设计效应因子呢？如果过去相同或相似主题的调查所用的抽样设计与我们计划实施的抽样设计相同或相似，就可以将过去相关调查的设计效应因子作为当前调查主要变量设计效应的估计值。或者，通过预调查获得设计效应的估计值。

如果调查机构计划使用分层抽样设计，但不能从过去相关的调查中得到设计效应的估计值，可以将设计效应设定为“1”，以此来计算所需的样本量  $C$ 。这时，所得的样本量与简单随机抽样的样本量相同。调查估计值最终的精度，应该不会低于根据简单随机抽样设计取得的精度。如果分层是有效的，那么实际得到的精度会更高。如果计划使用整群抽样设计，同时事先又不了解整群抽样对抽样方差带来的影响，对设计效应进行估计通常很困难，在这种情况下，设计效应至少应取为 2，也许它实际可能高达 6 或 7：

#### (四) 调查回答率

调查回答率是指回收到的有效问卷数占计划样本量的百分比，表示在所有发放的问卷中回收的有效问卷的比例。无效的调查问卷可能是因为：样本单元超出调查范围；在住户调查中，住宅是可居住的却无人居住（是空的）；样本单元没有回答等。

为了达到调查估计值要求的精度，调查机构需要根据预计的回答率调整样本量的大小。预计的回答率可依据对同一总体的小范围预调查或者过去类似的调查得到。例如根据计算，初始样本量是400，预计回答率为75%，那么样本量就应该定为：

$$n = \frac{400}{0.75} = 534$$

一旦调查机构和客户确定了某一回答率，就必须尽最大的努力，保证这一回答率的实现。

如果不能达到所预期的回答率，就会影响调查结果的精度。实际回答率偏低会导致有效样本数小于精度所需要的样本数；若实际的回答率比预期的高，结果正好相反。

务必注意，为了妥善处理无回答，简单地增加样本量是不够的。因为如果调查中，拒绝回答与回答的被调查者在所研究的指标上存在显著的差异，简单地增加样本会产生估计的偏差。

## 第二节

### 不同抽样方式下样本量确定

因为不同的抽样方式会得到不同的样本量，各自存在必要样本量的计算公式。如果以简单随机抽样为基础对样本量进行调整，则需要考虑的因素包括抽样方式下的设计效应和调查回答率。

#### 一、简单随机抽样方式下必要样本量确定

##### (一) 简单随机抽样，给定均值估计初的精度（不考虑回答率）

简单随机抽样下，通常使用误差限和估计量的标准差来确定所需的样本量。

1. 有限总体或不重复抽样情形。在不重复的简单随机抽样情形下，总体均值估计量V（或顶）标准差的表达式为：

$$\text{cr} = \sqrt{\frac{nS^2}{N}} \quad (7.1)$$

其中， $\sigma$ 是总体标准差的估计值。假设允许误差为 $\epsilon$ 那么：

$$\epsilon = z\sigma = z \sqrt{\frac{nS^2}{N}} \quad (7.2)$$

其中 $z$ 依赖于置信水平。从式(7.2)中解得：

$$n = \frac{z^2 \sigma^2}{\epsilon^2} \quad (7.3)$$



2. 无限总体或重复抽样情形。在重复的简单随机抽样情形下，总体均值估计量  $\bar{Y}$ （或子）标准差的表达式为：

$$\frac{\sigma}{\sqrt{n}} \quad (7.4)$$

其中，“ $S$ ”是总体标准差的估计值”假设允许误差为  $e$ ，那么：

$$e = z \frac{\sigma}{\sqrt{n}} \quad (7.5)$$

其中  $z$  依赖于置信水平。从式 (7.5) 中解得：

$$n = \frac{z^2 S^2}{e^2} \quad (7.6)$$

因此，为确定必要样本量  $n$ ，需要知道允许的误差限与给定置信水平相对应的  $z$  以及总体大小  $N$  和总体方差估计值。其中，总体方差估计不容易得到，通常需要根据过去对类似总体所做的研究确定一个近似值。

## （二）简单随机抽样，给定估计比例/的精度（不考虑回答率）

在确定总体比例  $P$  的样本量时，需要知道用误差限表示的精度。对于大样本，比例估计量后近似正态分布，而二元指标总体的方差可由式 (7.7) 估计：

$$S^2 = P(1-P) \quad (7.7)$$

于是计算必要样本量的公式为：

$$n = \frac{z^2 P(1-P)}{e^2} \quad (7.8)$$

或者

$$n = \frac{z^2 P(1-P)}{e^2} + \frac{A}{U(i)} \quad (7.8)$$

$$n = \frac{z^2 P(1-P)}{e^2} \quad (无限总体或重复抽样) \quad (7.9)$$

如果能够获得以往调查中总体比例的一个较好的估计  $A$ ，那么直接将它代入上面的公式，就可以得到所需的样本量。否则，可以用  $P = 0.5$ ，因为这时的总体方差最大。

如果不是简单随机抽样，那么在计算样本量时，还需要抽样设计的设计效应。盼的估计值，若回答率小于 100%，还需要一个回答率的估计值  $U(i)$

## 二、考虑回答率、设计效应的最终样本量确定

多数抽样方式都以简单随机抽样为基础，因此，为简化起见，其他抽样方式下样本量的确 定也可以以简单随机抽样的样本量为基础进行适当的调整。具体思路为在计算必要样本量基 础上，根据总体大小、设计效应和回答率等分别对其进行调整，最后求得最终样本量。下面以 比例估计计算样本量为例介绍详细步 骤。

第一步，计算初始样本量。

当总体是无限总体或抽样是有放回的抽样时，初始样本量  $S$  应按式 (7.9) 计算。如果  $e$



和『都不用比例表示，而用百分数表示，式（7.9）同样成立。

第二步，若总体是有限总体，而且是无放回抽样，那么要根据总体大小对所需样本量进行调整。调整的样本量化计算如下：

或直接采用式（7.8）计算样本量。

第三步，如果抽样设计不是简单随机抽样，则相应调整的样本量计算如下：

$$n_3 = Deff \times n_2$$

其中 $Deff$ 是设计效应，可从过去相同或相似主体的调查中或预调查中取得，通常，不同抽样方式的设计效应如下：

简单随机抽样，其  $Deff = 1$ ；

分层抽样设计，其

整群抽样或多阶段抽样设计，其  $Deff > 1$ 。

第四步，根据无回答再次进行调整，以确定最终样本量 $n$ ：

$$n = \frac{n}{r}$$

其中， $r$ 为预计回答率。

#### 例 7.1

某杂志出版商希望得到读者对该杂志综合满意度的估计值。通过邮寄调查，出版商可以联系到所有 2500 个订户，但由于时间的限制，出版商决定使用简单随机抽样进行电话调查。请问应调查多少个订户？

我们假定：

如果真实的总体比例落在总体比例的样本估计值 $\pm 0.1$ 范围内，该出版商就感到满意。换句话说，误差限 $e$ 设定为 0.1；

出版商希望调查估计值的置信度为 95%，即在 100 次抽样中只有 5 次所得的估计值确定的置信区间 $P \pm 0.1$ 不包含总体真值 $P$ 。此时  $z=1.96$ ；

使用简单随机抽样；

预计回答率为 65%，即  $r=0.65$ ；

由于事先没有关于顾客满意度的估计因此为留有余地，方差应取最大，即  $P = 0.5$  样本量的计算步骤如下：

第一步，计算初始样本量(3)：

$$\frac{z^2 P(1-P)}{(0.1)^2} = \frac{1.962(0.5)(1-0.5)}{(0.1)^2} = 97$$

第二步，根据总体大小调整样本量（这一步只需对中小规模的总体进行）：

$$n_2 = n \frac{N}{N+1} = 97 \times \frac{2500}{2500+97} = 94$$

第三步，根据抽样的设计效应来调整样本量：

在这个例子中，由于假定采用简单随机抽样设计，所以  $Deff = 1$ ：

$$n_3 = Deff \times n_2 = 1.94 = 94$$

第四步，根据无回答情况进行调整，确定最终的样本量”：

$$n = \frac{94}{0.65} r = 145$$

注意，在实施调查时，如果实际回答率不到 65%，最终的有效样本数将小于预期的样本量，所得的估计值的精度将比要求的精度低；如果实际回答率高于 65%，有效样本量将大于预期的样本量，调查得到的估计值的精度也会更高。

以上步骤告诉我们，为估计读者对杂志的综合满意度，在置信度为 95%，误差限为 0.1 且假定回答率为 65% 的条件下，出版商需从 2 500 个订户中随机抽取 145 个订户进行电话调查。

### 例 7.2

准备实施一项民意调查，以测定赞成建立一个市级公园的居民的比例。总体由所有在两个城市和一个农村地区居住的成年人组成。通过从每个城市及农村中各抽取一个简单随机样本，得到一个分层随机样本。问每一层需要多大的样本量？总体大小即总人口数为 657 500 人，总体在各层的分布如表 7-4：

表 7-4 总体及其分层

<i>h</i>	层	各层人口总数
1	城市 1	400 000
2	城市 2	250 000
3	农村地区	7 500
合计		657 500

所需要的样本量取决于调查对数据的具体要求，为此考虑以下两个方案：

方案 1：不需要得到各层的精确估计

假设不需要得到各层的精确估计，对整个地区的估计值在 95% 的置信度下，±0.05 的误差限，就认为估计值足够可靠了。由于没有整个地区赞成建立市级公园居民比例的可靠估计，所以假设  $P = 0.5$ ，预计回答率为 50%。

第一步，计算初始样本量<sup>⑧</sup>：

$$\hat{e} = \sqrt{\frac{0.5(1-0.5)}{657500}} = 385$$

第二步，计算初步修正的样本量  $n_2$

$$n_2 = \frac{657500}{657500+385} = 657.5$$

71] 这里，将而忽略不计，故  $n_2 = n$ ，

第三步，根据设计效应，再次进行调整：

$$n_3 = 1 \times 385 = 385$$

对于分层随机抽样，提出  $Deff < 1$  但由于没有可利用的估计值，因此，取  $Deff = 1$ ，以得到保守的样本量，即比需要的样本量更大一些。

第四步，对无回答进行调整，确定最终的样本量

$$n = \frac{3}{e^2} \cdot \frac{385}{N} = \frac{3}{0.05} \cdot \frac{385}{7500}$$

根据该方案，调查所需的样本量为 770。至于如何将这 770 个样本单元在 3 个不同的层 进行分配，将在第三节进一步介绍。

### 方案 2：要求得到各层的精确估计

假设对每一层，都要求  $P$  在置信度为 95% 下，误差限为 0.05 的估计结果，那么就需要单独 计算各层的样本量，即将每一层作为一个总体，估计调查所需要的样本量。计算之前，考虑到 城市 1 和城市 2 是大总体，因而可以认为，总体大小对样本量的确定没有影响。由此，如果第 1、2 层中  $P$  的取值与方案 1 相同，这两层需要的样本量都应是 770。然而对于农村地区，由于 总体较小，因此总体大小对该层样本量的确定会产生一定的影响。

农村地区需要的样本量计算为：

$$\begin{aligned} n &= \frac{z^2 P(1-P)}{e^2} \cdot \frac{N}{N-n} \\ &= \frac{1.96^2 (0.5)(1-0.5)}{(0.05)^2} \cdot \frac{7500}{7500-367} \\ &= \frac{385}{7500-367} = 367 \\ n_3 &= Deffx n_2 - 1 \times 367 = 367 \\ n &= \frac{3}{0.5} \cdot \frac{367}{7500-367} = 734 \end{aligned}$$

因此，方案 2 所需的总样本量为 770（城市 1）+770（城市 2）+734（农村地区）= 2 274。

方案 2 的总样本量为 2 274，几乎是方案 1 的样本量 770 的 3 倍，换句话说如果仅仅需要 得到包含所有层的整个总体的估计值，那么要求的样本量将大大小于对各层分别进行估计时 所需的样本量。因为对各层分别进行估计时，需要确保使得每一层的样本量都足够大。

这个例子清楚地说明了对每一个不同子总体审查精度的重要性。但如果调查涉及许多研 究域，这一要求可能会使样本量显著增大，并可能导致样本量超出客户的预算和现有资源的承 受能力。一般来说，要求估计的域越多，需要的样本量就越大。因此，需要在精度与费用之间 进行折中，以保证估计的误差在可接受的范围之内。通过增大每一层估计值的允许误差，或合 并其中某些域，使精度和费用达到平衡。

至此，我们所涉及的只是确定样本量过程的一个方面，即仅考虑最重要的指标估计值的精 度要求。但在实际中，在确定样本时还需考虑时间和费用这两个极为重要的因素。即最终确 定的样本量必须与可获得的经费预算和允许的时限保持一致。

对于许多调查，在确定实施调查的细节之前，就已经确定了调查经费预算，并限定了完成 调查的最后期限。实际中经常会出现这种情况，即实施调查所需要的样本量大于现有经费所 能支撑的样本量，在这种情况下，如果不能得到更多的经费，就得削减样本量，从而降低精度要 求。对于时间因素也一样，如果时间不充裕，也需要限制样本量，以保证按时完成调查。除了 时间和费用之外，其他一些现场操作因素如：数据收集的方法，能否招聘到合适的现场调查员， 以及处理数据的设备等等，都会对样本量的确定产生一定 的影响。

最终样本量的确定需要在精度、费用、时限和操作的可行性等相互冲突的限制条件之间进 行协调。它还可能需要重新审查初始的样本量、数据需求、精度水平、调查计划的要素和现场 操作因素，并做必要的调整。通常调查机构和客户寻求在最有效使用费用的基础上（例如缩 短访问时间），使用户能对所需的样本量提供经费支持。

### 第三节

## 分层抽样样本量的分配

决定分层抽样效率的一个很重要的因素是全部  $n$  个样本单元在各层间的分配方式。首先 回顾有关分层抽样设计的内容：包含  $V$  个单元的总体，被分为  $L$  个层，这  $L$  个层应该包括全部 总体单元，而各层之间应互不重复，其大小分别为  $N_1, \dots, N_L$ ，而且有  $V = N_1 + N_2 + \dots + N_L$ 。从 每一层分别抽取一个样本，各层的抽样过程是独立进行的，各层的样本量分别用  $n_1, n_2, \dots, n_L$  表示。且有  $n = n_1 + n_2 + \dots + n_L$

将  $n$  个样本单元分配到  $L$  层中去有两种方式：一是先用本章前面介绍的方法确定总的样本量，然后再在各层间进行分配，即总样本量固定的情况；另一种情况是先根据各层预定的精度要求，确定每一层需要的样本量，然后将各层的样本量相加得到总的样本量。（若精度用变异系数表示，即为给定变异系数的情况）

### 一、分配准则

本节将详细介绍总样本量固定和变异系数给定两种情况之间的差异。

#### (一) 总样本量固定

这种准则是先确定总的样本量  $n$ ，然后再将确定的总样本量以某种方式分配到各层。分配给第  $h$  层的样本量确定一个分配系数  $a_h$ ，每个  $a_h$  都在  $0$  和  $1$  之间取值，包括  $0$  和  $1$ （即  $0 \leq a_h \leq 1$ ），同时， $\sum a_h = 1$ 。  
 $a_h = \frac{n_h}{n}$

由此，对于第  $h$  层，其样本量  $n_h$  等于该层的样本分配系数  $a_h$  乘以总的样本量  $n$ 。即

$$n_h = n \cdot a_h \quad (7.10)$$

例如，如果某一层的分配系数为  $a_h$ ，那么样本中的一半，就会被分配到这一层。在总样本量固定这一分配原则下，由于样本量  $n$  已事先确定，每一层的叫值一旦确定，各层所需的样本量  $n_h$  也就可以求得。确定有多种方法，通常是取使总体估计值的方差最小的  $a_h$  值。

#### (二) 给定变异系数

确定总样本量的另一个准则是在给定总体估计值精度水平的情况下，计算确定每一层所需的样本量  $n_h$ ，使估计值的变异系数不超过事先估计的值  $\sigma_h$ 。

例如，考虑分层随机抽样的总量估计  $\hat{Y}_h$  将前面确定样本量的计算公式作某些变换，就可以得到在给定估计值的变异系数条件下计算总样本量  $n$  的表达式：

$$\begin{aligned} n &= \frac{\hat{Y}_h^2}{C_h^2 S_h^2}, \\ &\text{----- } (7.11) \end{aligned}$$

其中： $M$  是第  $h$  层单元的总数；

$S_2$  是第  $l$  层的总体方差估计值；

(3) 是分配给第  $l$  层的比例，即样本分配系数；

$c$  是  $y$  的变异系数的估计值；

$\bar{Y}$  是总体总值的估计值。

第  $h$  层总体方差的估计值  $S_h$  的计算公式如下：

$$\frac{\sum_{i=1}^{n_h} (x_i - \bar{x}_h)^2}{n_h T}$$

其中， $\bar{x}_h$  是第  $h$  层样本的均值：

$\bar{x}_h$

$\bar{Y}$

注意，如果  $\bar{Y}$  是一个二元变量，那么层的均值就是层的比例。即  $\bar{Y} = \frac{\sum_{i=1}^n Y_i}{n}$  且  $S_h = \frac{\sum_{i=1}^{n_h} (Y_i - \bar{Y})^2}{n_h}$ 。

将  $n_h = n a_h$  代入式 (7.11)，就可以得到冒的公式：

$$\begin{aligned} n_h &\sim n a_h \sim a_h \\ C^2 Y^2 + 2 \sum_{i=1}^{n_h} (Y_i - \bar{Y})^2 &+ 2 N_h S_h \end{aligned} \quad (7.12)$$

因此，在确定了每一层叫的值之后，就可以计算各层的样本单元数如同本章前面介绍过的样本量确定方法一样，为了确定哪，还应该进行无回答调整。

注意，进行样本分配时，用总样本量固定方法要比给定变异系数法简单，因而在下面样本分配的例子中只用总样本量固定的方法。

## 二、分配样本方法

式 (7.10) 和式 (7.12) 是分层抽样样本分配的基本公式，一旦各个(3)值确定，我们就可以 使用其中任何一个公式计算所需的样本量。每层七的选择方法可划分为两类：按比例分配或 不按比例分配。这些方法与各层总体单元数、表示层总体大小的其他一些度量，层总体方差以 及各层的调查费用等因素有关。

### (一) 按比例分配

按比例分配是指根据总体大小或总体规模指标的比例分配样本的方法，具体包括有与  $V$  成比例分配、与  $\bar{Y}$  成比例分配、与  $a_h$  成比例分配等。

#### 1. 与 $V$ 成比例分配

按比例分配即是与  $N$  成比例分配，就是每一层的样本量叫与该层的总体大小叫成比例的分配。因此，层的规模越大，分配到的样本单元数也就越多。即各层的抽样比  $f_h$  是相同  $\frac{n_h}{N_h}$

的，等于总体的抽样比

由于在比例分配的分层随机抽样下，各层抽样比都相同，并等于总体抽样比，因此可以得到下面的等式：

$$h_i = \frac{n_i}{N} h \quad (7.13)$$

因此，对于与  $N$  成比例的分配——换句话说，每一层的分配系数叫等于各层的  $n_i / N$  总体大小与整个总体大小之比。

当层的总体方差未知时，通常使用比例分配。由于在使用给定变异系数准则计算各层样本量时，需要知道每一层的方差。因此在使用这种准则时，通常不使用比例分配的方法。比例分配的分层随机抽样还是一个自加权的样本设计。

如果各层总体均值又的差异非常显著，那么比例分配的分层随机抽样（简称比例分层随机抽样）比对整个总体的简单随机抽样更有效率。但是如果分层后，各层的总体均值相差不大，那么比例分层随机抽样减少抽样方差的程度就不大。当然，比例分层随机抽样的效果一般不会比简单随机抽样的效果差，因为它“设计效应”通常不会大于 10。

下面举例说明总样本量固定下与  $N$  成比例的样本分配过程。

#### 例 7.2（续 1）

在例 7.2 的方案 1 中，计算得到样本量“为 770。对于总样本量固定的情形，下面用与  $N$  成比例分配的方法将这 770 人分配到各层中去。

第一步，根据与  $N$  成比例分配的原则，计算各层的分配系数(3)：

$$\text{城市 1: } \frac{M}{nV} = \frac{400000}{657} = 0.608$$

$$\text{城市 2: } a_2 = \frac{250000}{657} = 0.380$$

… 地 7500 农村地区  $a_3 = 0.0114$  第二步，计算各层的样本量叫：

$$\text{城市 1: } n_1 = net |na_1| = 770 \times 0.6084 = 468$$

$$\text{城市 2: } n_2 = na_2 = 770 \times 0.3802 = 293$$

$$\text{农村地区: } n_3 = na_3 = 770 \times 0.0114 = 9$$

正如我们所预料的那样，样本中大多数单元都被分配到总体较大的层，在城市 1 和城市 2，应分别抽取 468 人和 293 人，而总体最小的层即农村地区，却只得到整个样本的很少一部分，仅抽取 9 个单元。计算结果见表 7-5。

表 7-5 分层抽样的分配系数及其层样本分配

$h$	层		$^a h$	$^n h$	九一瓦
1	城市 1	400 000	0.608 4	468	0.001 2
2	城市 2	250 000	0.380 2	293	0.001 2
$h$	层	%	%	s	

3	农村地区	7 500	0.011 4	9	0.001 2
合计		657 500	1	770	0.001 2

表 7-5 表明，3 个层的抽样比都等于 0.001 2，这表明比例分配是一个自加权的抽样设计。

注意，上述分配方法和例 7.2 方案 2 中样本量的确定过程有所不同：这里的分配方法是满足整个估计值精度的要求，而例 7.2 方案 2 满足的却是各个层所要求的精度。在采用不按比例分配时，各层之间的抽样比是不相同的。下面将讨论几种不按比例分配的方法。

### 2. 与 K 成比例分配

在名称上与 Y 成比例分配和与 V 成比例分配相差不大，容易和按比例分配相混淆，但请读者注意，是否是按比例分配，主要是看各层的抽样比是否相同。如果相同，则是按比例分配，否则就不是按比例分配。调查中如果变量  $\cdot \cdot$  (第  $h$  层的第  $i$  个指标) 是关于大小 (或规模) 的一种度量，那么第  $h$  层的样本量  $n_h$  就可以是第  $h$  层大小 (或规模) 度量  $N_h$  的一个比例，这种分配方法称为与  $V$  成比例分配。这时

$$n_h = S_h N_h \quad (7.14)$$

即每层的分配系数  $S_h$  等于该层总值大小度量与总体总值大小度量的比。

与  $V$  成比例的分配方法在市场调查中有着广泛的应用，因为这类调查中，“经常呈偏态分布。典型的例子是制造业的就业调查，以及零售业的销售量调查。在这些调查中，少量的企业就能说明该行业总就业情况或该产品总销售情况的较大的百分比，而其余规模小、数量多的企业却只代表该行业总就业情况或该产品总销售情况的较小部分。

市场调查通常在现有大小或规模度量（例如，就业人数、公司总收入、销售量）的基础上构造层。例如，将公司划分为小规模、中等规模和大规模三个层。包含规模最大单元的层，变异程度也往往最大。极端的情况下，与  $V$  成比例分配意味着对偏态总体中最大规模的层进行普查，而对其余层进行抽样。当调查估计值与  $V$  相关程度大于它与  $M$  的相关程度时，用与  $V$  成比例分配方法得到的估计值的精度，要高于用与  $/V$  成比例分配得到的精度。需要说明的是在实际调查中  $V$  与  $N_h$  都是未知的，因此只能参照过去的历史数据或与指标值高度相关的指标获取信息。

### 3. 与 $J$ 成比例分配

到目前为止，所介绍的所有分配方法都只考虑了总体估计值  $V$  的精度。但是通常客户还希望层估计值  $R_h$  也具有较高的精度。例如，如果以省为层，那么省和全国的估计值可能都很重要。使用与  $V$  成比例的分配，将样本单元分配至各层，就能改善层估计值的精度。在这种分配方法下，分配系数  $a_h$  的计算公式为：

$$a_h = \frac{\sqrt{N_h}}{\sum_{h=1}^L \sqrt{N_h}}$$

换句话说，分配系数叫等于该层总体大小的平方根与所有层总体大小平方根的总和之比。与用成比例分配可以提供精度更高的层估计值。实际上，经常借助它在最优分配和满足所有

域估计要求的分配之间进行折衷。例如,对于全国估计值来说,最优的分配会导致研究域(例如省)估计值的抽样方差变大,而选择满足所有域的要求(如例 7.2 方案 2)的样本分配方法却可能会使得总体估计值的效率非常低与川成比例的分配则可以在总体分配与域分配之间进行折中

下面举例说明当总样本量  $n$  固定时与  $V$  的平方根成比例分配的过程。

#### 例 7.2 (续 2)

我们在前面已使用比例分配方法将总样本量 770 人分配到 3 个层。现在我们使用与用 成比例分配法重新分配样本。表 7-6 列出了已知信息。

城市 1	$\frac{5}{\$}$	—顷 9.06	5 = 0588
	$h = I$	何	
	500		0.4102
—, —1	219.06	—	

$h = I$	£何
	$I$

农村地区: ———

£何

$h = I$

第二步, 按比例分配各层的样本叫 城市

$1 = na, n = 770 \times 0.5188 - 399$  城市 2:  $n_2$

表 7-6 与亦成比例分配法的分配系数

$h$	层		用
1	城市 1	400 000	632.46
2	城市 2	250 000	500
3	农村地区	7 500	86.60
合计		657 500	1 219.06

第一步, 根据与  $N$  的平方根成比例分配计算各层分配系数%:

表 7-7 对上述计算结果作了小结, 同时比较了与  $V$  成比例分配和与用成比例分配的计算结果。

表 7-7 与成比例分配和与用成比例分配的比较

$h$	层	与可成比例分配			与加成比例分配			
		$n_h$	$a_h$	A	用	$a_h$	$n_h$	$f_h$
1	城市 1	400 000	0.608 4	468	0.001 2	632.46	0.518 8	399
2	城市 2	250 000	0.380 2	293	0.001 2	500	0.410 2	316
3	农村地区	7 500	0.011 4	9	0.001 2	86.60	0.071 0	55
	合计	657 500	1	770	0.001 2	1 219.06	1	770
								0.001 2

由表 7-7 可知, 与比例分配相比, 按与用成比例分配对城市 I 赋予了较小的样本量。但是, 它对城市 2 和农村地区却赋予了较大的样本量。因此, 对于与  $V$  成比例分配, 城市 2 和农村地区得到的估计值的精度将高于比例分配得到的精度, 因为在这种分配方法下, 这两个地区 分配到的样本单元数较多。按与  $N$  成比

例分配农村地区只有 9 个样本单元，仅仅根据 9 个单 元，很难得到关于农村地区可靠的估计值。对于城市 1, 样本量的减少不会对估计值的精度产 生显著影响。而对于城市 2, 样本量的增加仅对其结果的精度产生轻微的正面影响。但对于 农村地区，增加的样本量却使其估计值的精度有了显著提高，这样农村地区精度的提高大大 弥补了城市 1 在精度上的损失。

#### 4. 与户成比例分配

确保层估计值和总估计值都有较高精度的另一种方法，是使用与旧成比例分配。这里的  $\bar{h}$  也是对大小（规模）的一种度量。与疗成比例分配是与  $V$  成比例分配的一种替代方法。在调查估计值与大小变量  $\bar{h}$  的相关度高于与层的大小  $M$  相关的情况下，与户成比例分配法 比与用成比例分配法更精确。在与拒成比例分配中，分配系数%的计算公式如下：

(7.16)

即分配系数  $a_h$  等于层大小度量的平方根与所有层大小度量平方根的总和之比。

同与  $V$  成比例一样，若以总的样本量最小为最优分配的标准，那么，使用与拒成比例分 配来计算  $S$  的值（以及随后的也），也不如其他分配方法更有效。但是，它可以提供精度更高 的层估计值。

与何成比例分配和与拒成比例分配都是与“幕”有关的分配，即与  $\bar{Y}$  的某个幂（乘方）有 关，更一般的定义为：

(7.17)

其中， $P$  常取分数，如  $1/2$ ，即为平方根幕。

#### （二）不按比例分配

不按比例分配是指考虑各层总体单元数、表示层总体规模大小的度量，层总体方差以及各 层的调查费用等因素的样本分配方法，常见的有最优分配、奈曼分配和方差相等时的最优 分配。

##### 1. 最优分配

当各层单元的调查费用不全相同，以及层间方差  $S^2$  变化很大的情况下，可以考虑使用以 下称为最优分配的不按比例分配的方法. 这是考虑费用的分配方法。

为了用最优分配，需要假设一个函数，用于模拟调查费用。最简单的费用函数具有下面的 形式：

其中， $c$  是总费用且是第  $i$  层  $i=1, 2, \dots, L$  每个单元的调查费用， $\alpha_i$  表示该调查固定的 管理费用。当费用中主要开支是访谈费用，或者在每个单元上测定数据时，这一费用函数是较 为合理的。

最优分配的分配系数  $S$  按下式计算：

N&

(7.18)

X A

最优分配法可以在规定的费用下，使估计值的方差最小，或者在给定的总抽样方差下，使 费用最小。为达到这一目标，内部变异程度较大或规模较大的层，抽取的样本单元数应该相应 增多，而对于那些调查费

用昂贵的层，则应减少取样的数目。也即最优分配的一般规则是对以下情况的层需抽更多的样本单元：层的规模较大；层的内部变异程度较大；层内每个单元的调查费用比较低。

为了运用最优分配法，需要知道每个层的层内方差和层内单元的调查费用。但在实际中，可能并不知道层内方差和层内每个单元调查费用的有关数据。解决这一问题的途径之一是根据初始样本或从以前类似调查来估计方差和费用。但不足之处是，这种估计值的稳定性较差。因此，对于重复性调查，经过一段时间后，设计往往就不是最优的。

当所有层的方差和费用都相等时，最优分配就简化为比例分配。在这种情况下，与“成比例分配”是使估计值方差最小的分配形式。当所有层的层内抽样调查费用都相等时，最优分配就简化为下面要讨论的奈曼分配。

## 2. 奈曼分配

各层中每个单元的调查费用都相同的最优分配就称为奈曼分配  $Q$ 。这一分配方法是在总样本量固定的情况下，使估计值的总抽样方差最小的一种分配方法。在奈曼分配下，规模或方差较大的层将分配到更多的单元。由于它是最优分配的特例，所以它也必须由初始或以前的样本本来估计层内方差。

奈曼分配下，分配系数(3)表示为：

(7.19)

即分配系数  $S$  等于第  $h$  层的层规模和其标准差的乘积与所有层规模和其标准差乘积之比。

当得不到层内方差的精确数值时，可以认为各层的层内方差与该层总体均值 ( $=$ ) 的比率  $\pi_h$  不变。在这一假设下，奈曼分配就还原为与  $K$  成比例分配。运用奈曼分配的困难和最优分配一样，就是层内方差可能不大稳定。

## 3. 方差相等时的最优分配

最优分配的另一个特例是所有层内方差都相等的情况。由于方差在层间通常是变化的，所以这一分配方法较少使用。但是，当不知道有关层方差的任何信息，或假设这些方差在层间是相等的，或者当费用是影响分配系数的主要因素时，可以使用这一分配方法。该方法将更多

的样本单元分配到规模大的层和单元费用较小的层分配系数(3)计算如下:

$$a_h = \frac{\frac{N_h}{\sqrt{c_h}}}{\sum_{k=1}^L \frac{N_k}{\sqrt{c_k}}} \quad (7.20)$$

### 三、分配样本时需要注意的事项

在进行样本分配时，应考虑以下问题：

#### (一) 利用好辅助信息

在分层抽样设计中，理论上可以用与  $V$  或与  $\bar{Y}$  的平方根成比例分配法分配样本单元。但实际上我们根本不可能知道总体中所有单元的指标值，如果我们知道了这一数值，那么就不需要对这一指标进行调查了因此通常借助于与指标值高度相关的辅助信息来实施这些分配。这些辅助信息来自过去的调查或行政统计资料。但使用辅助信息，可能得不到调查变量的变异系数使用辅助信息进行分层的效率以及用于样本分配的精确性取决于调查变量和所用的辅助变量之间的相关程度。

#### (二) 超额分配的处理

在分层抽样设计中，使用最优分配、奈曼分配、与  $K$  成比例分配或与  $\bar{Y}$  的平方根成比例分配，都有可能发生对某些层，分配的七大于层总体大小  $M$  的情况，这种情况称作超额分配。当出现超额分配时，应该对发生超额分配的层进行全部调查。由这种超额分配得到的有效总样本量将小于初始确定的样本量，因此估计值可能达不到规定的精度。解决的办法是增加叫小的那些层的样本量。

#### (三) 多个变量的分配

对一个变量来说是最优的分配，对另一个调查变量来说未必就是最优的分配。如果我们希望基于多个变量分配样本，就必须达成某些折中的分配方案。为此专家们研究了多变量的分配方法来解决此类问题，其中有些方法使用了线性规划的方法，已超出了本书的范围。

应该记住的是，分配样本的目的是满足主要调查变量的精度要求。因此，对那些不太重要变量的精度可能比主要调查变量的精度稍差一些，

### 篇章小结••••

置信度或置信水平通俗的叫法也叫可靠性，精度与可靠性的关系有绝对和相对两种表示方法。确定调查达到的合适精度水平时，需要考虑允许误差限的大小、调查总体中子总体的估计、调查估计值抽样误差的大小和时间、经费的限制。影响样本量大小的因素有总体指标的变异程度、总体大小、抽样设计方式、回答率。这些因素在确定样本量时都必须认真考虑。在简单随机抽样方式下给定均值估计量的精度和给定估计比例的精度都可以按照公式计算必要的样本量：其他抽样方式的样本量可以在简单随机抽样样本量的计算基础上考虑设计效应和回答率进行调整。分层抽样样本量在层间的分配方法很多，可以有等比例和不等比例分配法。磁关推名词••••

精度 允许误差限 回答率 最优分配 奈曼分配 超额分配 分配系数

思考题.....



1. 什么是精度？如何表示？
2. 什么是允许误差？
3. 确定精度时需要考虑哪些因素？
4. 影响样本量大小的因素有哪些？
5. 在给定均值估计精度条件下，简单随机抽样样本量如何计算？
6. 在给定估计比例精度条件下，简单随机抽样样本量如何计算？
7. 如何考虑回答率和设计效应对样本量进行调整？
8. 分层抽样的分配准则有哪些？
9. 分层抽样的样本分配方法有哪些？各有何特点与条件？
10. 分配样本时需要注意哪些事项？

## 可 •本章实训…

### 一、 实训目的

1. 要求学生掌握简单随机抽样、分层抽样的样本量计算方法。
2. 通过实训掌握样本量确定中的重要因素、技巧与设计思路二

### 二、 实训内容

首先，完成前述调查设计项目中样本量的确定任务以

其次，按照简单随机抽样、分层抽样和其他方式抽样的要求，确定合适的样本量。

### 三、 实训组织

1. 将教学班级学生按 4~6 人的标准分成若干小组，每个小组采用组长负责制。每组选择 一种抽样设计中确定样本量的方法，按照实训目的与要求，合理分工，团结协作。
2. 收集相关资料和数据。
3. 完成抽样设计中样本量确定及其分配的文本文件。
4. 各小组充分讨论，认真听取全体组员的意见，分析、修改和调整，在此基础上形成小组 的 PPT 实训报告文稿。
5. 各小组在班级进行实训汇报和展示
6. 教师进行成绩评定与总结。

# 第八章抽样估计

## 、导语

调查数据收集完成后的重要任务之一是对主要调查指标进行估计。抽样估计是根据样本的统计特征来估计推断总体相应特征的方法，是统计学中参数估计的原理在抽样估计中的具体应用。在进行参数估计时同样需要选择一个好的估计量来推断总体指标。评价估计量的标准依然遵循一致性、无偏性和有效性原则。

但是抽样估计与传统统计学中的参数估计不尽相同，第一，抽样调查在现实中通常是有限 总体，而且大多数社会经济现象，无法重复进行；而在传统统计学中往往假设被研究总体是无限 总体，建立在可实验观察的基础上，可以无限进行。第二，抽样调查的观察值通常是在有限总体 中进行不重复抽样后得到，因而观察值之间不独立，这使得一些抽样方差的计算比较复杂，而在 传统统计学中，假定样本观察值独立同分布，样本方差的计算较简单；第三，抽样调查中研究对象 的总体分布常常未知，使用比较多的是大样本情况下估计量的近似正态分布，而参数估计理论中 通常假定总体分布的形态是已知的正态分布，从而在理论上比较侧重于讨论小样本的精确分布； 第四，抽样调查中经常由于抽样单元的大小不同或分层抽样等原因而实施不等概率抽样和多种 方式的抽样，而参数估计中讨论的样本，通常是等概率的随机抽样；第五，抽样调查除了考察抽样 误差外，还要注意非抽样误差，而传统参数估计主要是考察抽样误差.

本章主要介绍通过加权处理提高抽样估计精度的方法，不同数据类型下的点估计，利用辅助信息进行分层、加权与估计，方差估计及调查指标的区间估计方法。

## ’教学目标与要求

、————，—————

1. 理解加权在抽样估计中的作用；
2. 掌握不同情形下的加权处理技巧；
3. 掌握总体均值、比例和总量的点估计方法；
4. 熟悉抽样估计中辅助信息的用途；
5. 掌握抽样方差及其估计的方法；
6. 掌握区间估计方法的具体应用。

## 、本章知识结构



# 第一节

## 抽样设计加权与调整

估计的第一步，就是给每一个样本单元或样本的回答单元赋予一个权数，这一权数成为样本指标估计总体相应指标的乘数因子。一般情况下，在概率抽样设计时抽样估计的权数就已经确定，但经常由于调查实施中数据采集的无回答，需要进行设计权数的调整。

### 一、等概率抽样加权

如果所有样本单元的设计权数都相同，那么称这样的抽样设计为自加权设计，这种情况发生在当每个单元都有相同的入样概率时。对于自加权的抽样设计，如果没有随后的权数调整，（例如，出于处理无回答或使用辅助变量的需要），那么在计算诸如总值、均值等估计量时可以将权数忽略。对总值的估计仅仅需要将样本总值乘上某个倍数即可。

那么，哪些设计是自加权的呢？简单随机抽样是一种自加权设计，其每个单元都有相等的入样概率。同样，系统抽样也是一种自加权设计。对于分层抽样，如果按各层的大小等比例分配样本，而且每层内都采用简单随机抽样，那么它也是自加权设计。这就是说，如果各层的抽样比相同，总体中每个单元就具有相同的入样概率，这样的分层抽样设计就是自加权设计。我们称分层抽样的这种样本分配方法为与各层大小/V成比例的分配。

下面以一个简单分层抽样为例，对这种自加权设计进行说明。

#### 例 8.1

一个由/V=10 000 个人构成的总体被划分成两个层：第一层由 "5 200 名男性组成；第二层由 % = 4 800 名女性组成。从中抽取一个样本量为  $n = 250$  的样本，将样本等比例地分配给各层，使得两层的抽样比都等于（二洁 3=1）。

男性层的样本量是：

$$n_1 = \frac{n}{N} = \frac{1}{2} = \frac{250}{10000} = \frac{1}{40}$$

女性层的样本量是：

$$n_2 = \frac{n}{N} = \frac{1}{2} = \frac{250}{10000} = \frac{1}{40}$$

此时，设计权数是多少？

对于男性层，入样概率  $P_i$  为：

$$P_i = \frac{n_1}{N_1} = \frac{1}{40} = \frac{1}{5200}$$

对于女性层，入样概率  $P_i$  为：

$$P_i = \frac{n_2}{N_2} = \frac{1}{40} = \frac{1}{4800}$$

这样，每个人都有相同的入样概率  $1/40$ ，设计权数都等于  $40^\circ$

为设计一个自加权的多阶段抽样，可以采用除最后一阶段抽样外，其他各阶段都按照与单元大小成比例的概率抽样方法抽选样本，而在最后一阶段用等概率抽样方法抽选单元多阶 抽样经常使用这种与单元大小成比例的概率抽样方法，因为这样可以得到自加权的样本，并能 控制样本量的大小。

## 二、不等概率抽样设计的加权

自加权设计因为简单与使用便利而常被采用，但它有时并不可行 G 例如，进行一项全国性 调查时，采用不等比例分层抽样，对较小地区进行估计所需的样本量，可能比采用比例分配方 法得到的样本量要大，从而避免了对较小地区使用按比例分层抽样所得的样本量不够大，而对 于较大地区得到的样本量又太大的情形。

当采用的抽样方法不是等概率抽样时，正确地使用权数就显得尤为重要。下面的例子说 明当抽样总体被划分为两层，而样本不是按比例分配时的权数计算问题

### 例 8.2

对一项在线教育利用效果进行调查，总体由 1 100 人组成，考虑到在线教学模式可能存在 明显的教育层次差别，所以采用了分层抽样设计，将总体分为本科生与研究生两个层。

本科生层的大小为此 $=1 000$ ，研究生层的大小为以  $2=1$ 。。。打算抽取一个 $n = 250$  人的样 本，本科生层抽取 $n_1 = 200$ ，研究生层抽取  $n_2 = 50$  人（按比例抽取样本仅为 23，样本量偏小）。 具体信息见表 8-1 o

对于这项调查，被调查者的设计权数是多少？

表 8-1 调查的分层总体及其样本量

层	总体大小	样本量
本科生	$M = 1 000$	$n = 200$
研究生	$N = 100$	$n_2 = 50$

各层的权数叫是单元入样概率的倒数：

本科生层：叫. 广 上些卯  
 $n, 200$

研究生层：町 2 =--- =  
 $n_2 50$

因此，对本科生层的每个被调查者的权数为 5，而对研究生层的每个被调查者的权数为 2。

## 三、对无回答的权数调整

所有调查都会受无回答的困扰。所谓无回答是指由于某些原因，从抽中的样本单元无法 获得所需要的信息。它可以是样本单元没有提供或者没有完全提供所需的信息，也可以是提 供的信息中有一部分无法使用。一般情况下，无回答的数量既取决于调查的内容，也取决于数 据收集和数据处理方法。

无回答有两种主要类型：项目无回答和单元无回答。项目无回答是指被调查者虽然接受 了调查，但是仅仅回答了问卷的一部分，因而不能得到某些项目的信息，在这种情况下，最常用 的调查方法就是对缺失数据进行插补（详见第十章）。

单元无回答主要是指一个样本单元所有或几乎所有的数据都缺失。单元无回答的产生，

可能是由于被调查者拒绝合作或者无法找到被调查者的住所，或者调查结果无效。处理单元 无回答最简单的办法就是忽略它 在一些特殊的情况下对单元无回答进行调整后所得的均值 或比例的估计值，与未做任何调整的估计值相比并没有任何改进。然而对总值的估计则不同， 此时如果不对单元无回答进行弥补，则会导致对总值的低估。例如， 总体的大小会被低估， 总 收入或总产量等总体指标也会被低估。

### 例 8.3

从一个  $N= 100$  人的总体中抽取一个  $n = 25$  人的简单随机样本，结果只有 20 个人提供了 所需要的信息，其中有 5 个人为单元无回答。如果我们不考虑无回答，其抽选的概率  $P$  和设计的权数分别为：

$$\frac{25}{100} = \frac{1}{4}$$

$$W_d = \frac{N}{n} = \frac{100}{25} = 4$$

假设我们要调查这一总体的总收入，现在只有 20 个人做了回答，其每人平均收入 1 000 元，样本的总收入为 20 000 元。若要估计总体的总收入，就要将样本的总收入乘以权数，得到  $20 000 \times 4 = 80 000$  元。从这个例子可以看出，如果无回答的人每人平均收入也是 1 000 元的话，这个 100 人的总体的总收入应为 100 000 元，由于无回答的存在，总收入被低估了。

由于无回答所导致的这种低估称为无回答偏差。如果发现忽略单元无回答是不适当的， 最常用的办法是对权数进行调整”这样处理的理论假定是回答单元不仅能代表回答单元，而 且能够代表无回答单元。当回答单元和无回答单元具有相似的调查指标时，这是一个合理的 假定。此时无回答的设计权数需在回答单元之间进行重新分配，通常的做法是将一个设计权 数乘以一个无回答调整因子，由此得到无回答的调整权数。

无回答调整因子是原样本单元的权数之和与给出回答的单元权数之和的比值。对于自加权设计，这个比值也可以用原样本的单元数与给出回答的单元数比值来表示。由于后者更容易演示，所以下面就用一个自加权样本的例子来说明无回答调整因子的计算。

### 例 8.4(续例 8.3)

由例 8.3 已经计算出每个样本单元的设计权数为 4。下面进一步计算无回答的调整因子。 由于样本单元为  $n=25$ ,回答单元用  $n_r$  表示，则  $n_r = 20$ ，他们提供了所需的信息。假定回答单 元不仅能代表回答单元，而且能代表无回答单元，这个例子又是等概率的自加权抽样设计，此 时可计算无回答的调整因子为：

$$\text{无回答调整因子} = \frac{77}{25} = 1.25$$

$$\frac{n_r}{n} = \frac{20}{25} = 1.25$$

接着计算无回答的调整权数。无回答的调整权数用星号表示，它等于设计权数与无回答 调整因子的乘积：

$$* = 4 \times 1.25 = 5$$

所以做出回答的 20 人中，每个人都代表被调查总体中的 5 个人，把它们合在一起就代表 整个总体共  $5 \times 20 = 100$  人(即  $n = JF_{nr}X_n$ )。此时，对数据文件中的每一条记录赋予权数 5。

如果可以假定所有的无回答单元与回答单元具有相似的调查指标，那么可以对所有的回答单元使用相同的无回答调整因子。但是，不同的子总体往往具有不同的回答率，在这种情况下，对所有的无回答单元使用同样的调整将导致结果的偏差. 例如，与多个成员组成的家庭相 比，只有一个成员的家庭通常具有更低的回答率；高收入的人通常也比其他人具有更低的回答 率。这种情况下，由于不同子总体具有不同的调查指标，应对这些子总体分别进行无回答 调整。

在下面的例子中，对本科生和研究生两个层使用了不同的无回答调整因子，这样做是因为 发现本科生层和研究生层具有不同的回答率，而且本科生和研究生在调查指标上也存在着 差异。

### 例 8.5 (续例 8.2)

数据收集后发现，本科生层中只有  $n_{r1} = 150$  人，研究生中只有  $n_{r2} = 40$  人提供了所需的信息，见表 8-2° 那么，此时回答者的权数是多少？

表 8-2 调查后的分层及其回答者数量

层	总体大小	样本量	回答者数量
本科生	$M = 1000$	$= 200$	$\% = 150$
研究生	$/v_2 = 100$	$n_2 = 50$	$s = 40$

步骤 1：各层的设计权数为：

本科生层 驴产 5 研究生层 町=2

步骤 2：进行调整以弥补无回答 c，本例中，各层的无回答调整因子计算如下：

$$\begin{array}{l} \text{本科生层: } \frac{n_1}{M} = \frac{200}{1000} = 0.2 \\ \text{研究生层: } \frac{n_2}{v_2} = \frac{50}{100} = 0.5 \end{array}$$

$$\begin{array}{l} \text{本科生层: } \frac{n_1}{M} = \frac{200}{1000} = 0.2 \\ \text{研究生层: } \frac{n_2}{v_2} = \frac{50}{100} = 0.5 \end{array}$$

步骤 3：无回答的调整权数等于设计权数与无回答调整因子的乘积：

$$\begin{array}{l} \text{本科生层: } \frac{n_1}{M} \times 0.2 = \frac{200}{1000} \times 0.2 = 0.04 \\ \text{研究生层: } \frac{n_2}{v_2} \times 0.5 = \frac{50}{100} \times 0.5 = 0.25 \end{array}$$

在样本数据文件中，本科生层的每条回答记录应赋予权数 0.04，研究生层的每条回答记录 应赋予权数 0.25。

无回答调整只适用于调查范围之内的单元，如果发现一些样本单元不在调查范围之内，则需要另外的修正。

## 第二节

### 总体均值、比例及总量的点估计

市场调查中对主要调查指标的总体估计是通过调查样本量获得样本的主要特征指标，然后根据设计的抽样方式进行总体均值、比例和总量的点估计。下面分别介绍不同数据类型下 和等概率与不等概率抽样设计下的变量估计。

## 一、不同数据类型下的变量估计

多数市场调查都要求收集很多指标，这些被称作变量的指标可以是定性或定量数据类型。对于不同类型的变量，需要使用不同类型的估计。通常对定性变量计算比例与总数；对定量变量计算平均数与总量。本节将分别介绍定性变量和定量变量估计值的计算方法。

在估计时，需要考虑针对的总体范围。估计既可以是针对整个调查总体，也可以针对特定的子总体或特定的域 G 例如除需要计算全国的估计值外，可能还需要计算省及地、市的估计值。同时，还可能需要计算诸如性别、年龄、受教育水平等这样的域估计值。

下面介绍对定性与定量数据，使用权数估计总体及总体中域的总量、均值及比例的方法。

### （一）总体总量的点估计

无论对定性还是定量数据，调查总体中单元总数的估计都是对回答单元的最终调整权数之和：

$$N = \sum_{i \in S_r} w_i$$

其中 Q 是样本中回答单元的编号， $w_i$  是它的最终调整权数， $S_r$  是所有由回答单元组成的集合 Q。

对定量数据，总体总值的估计，是每个回答单元的权数  $w_i$  与其值  $y_i$  的乘积之和：

$$\sum_{i \in S_r} w_i y_i$$

### （二）总体比例的点估计

对定性数据，对调查总体中具有某特定特征的单元所占比例进行估计，是将具有该特征单元的权数进行加总，然后将它除以所有回答单元的权数之和。换句话说，总体比例的估计，是具有某特征的单元总数的估计值除以总体中单元总数的估计值。即：

$$P = \frac{\sum_{i \in S_r} w_i}{\sum_{i \in S_r} w_i}$$

其中 c 是具有特定特征的单元集合。

### （三）总体平均数的点估计

对定量数据，总体平均数的估计是回答单元的样本值与权数的乘积之和，除以这些单元权数之和。换句话说，总体平均数的估计值，是定量数据总值的估计值除以总体单元总数的估计值，即：

$$\bar{Y} = \frac{\sum_{i \in S_r} w_i y_i}{\sum_{i \in S_r} w_i}$$

### （四）总体域的点估计

有时候还需要估计总体中一些特定部分或称子组的估计值，例如按年龄组、住所类型、家庭规模、收入等级等划分的子组的估计值。这些需要进行估计的子组通常称作域。如果把域作为整个总体处理，那么可以将上面讨论的关于计算总量、比例或平均数的估计值方法用于域的估计。

无论是定量数据或定性数据，域的大小都可以用下式估计：

$$\sum_{i \in S_r} w_i$$

对定量数据，域总值的估计值为：

$$\bar{M} = \frac{\sum_{i \in S_r} w_i y_i}{\sum_{i \in S_r} w_i}$$

对定量数据，域均值的估计值为：

对于定性和定量数据，域比例的估计值为：

\*      ie.s.n 域 nc  
P 域二—F

Le s.n 域

这些估计方法的具体说明见下面各节。

## 二、等概率抽样设计的变量估计

理论上自加权抽样设计，所有样本单元的最终权数都相等。因此计算比例或均值的估计值时，就可以忽略权数。对总体总量的估计，只需要用同一个设计权数将样本总值放大即可。但是在实际中这种情况很少发生，这是因为通常需要根据无回答或利用辅助信息进行调整。

下面的例 8.6a 说明了在自加权设计的情况下，如何计算定性和定量数据的比例，平均数 和总量的估计值。当计算比例和均值时，忽略了权数。例 8.6b 列举的是同一个例子，但使用了设计权数。这两个例子说明两种不同的方法得到相同的结果。然而，因为估计总体总量时 要用到权数，也因为由于无回答或辅助信息需要对权数进行调整，导致即使是自加权设计，最 终权数也不同，所以使用第二种方法更为稳妥。

### 例 8.6a

下面所列是例 8.3 所用的调查数据文件。在本例中，所有回答单元都有相同的最终权数 5。数据中有一个定量变量（年龄）和两个定性变量（性别和婚姻状况），见表 8-3。在数据文件中，对两个定性变量取值的编码如下：

性别：1 二男性, 0 二女性

婚姻状况：1 =已婚, 0 =单身

表 8-3 回答样本的部分人口统计特征与权数

回答者	年龄	性别	婚姻状况	权数
1	29	1	1	5
2	32	1	1	5
3	45	0	0	5
4	16	1	0	5
5	30	0	1	5
6	38	0	1	5
7	43	1	1	5

回答者	年龄	性别	婚姻状况	权数
8	15	0	0	5
9	22	1	0	5
10	49	1	0	5
11	39	0	0	5
12	42	0	1	5
13	53	0	1	5
14	18	1	1	5
15	26	0	1	5
16	76	0	1	5
17	48	1	1	5
18	15	0	0	5
19	24	0	1	5
20	65	0	1	5

### (一) 对总体或总体中某个域的总量估计

总体中男性总数的估计值为样本中男性的总数乘以无回答调整权数  $w_{nr}$ :

$$* \text{男性} = \text{吃次} * \text{男性} = 8 \times 5 = 40$$

总体中单身女性总数的估计值为样本中单身女性的总数乘以无回答调整权数  $W_{nr}$ :

$$\text{。女性} = \text{女性} = 4 \times 5 = 2。$$

### (二) 忽略权数对总体的比例和平均数的估计

总体中男性所占比例的估计值即为样本中男性所占的比例:

$$\frac{8}{20} = 0.4$$

总体中平均年龄的估计值即是样本的平均年龄:

$$\frac{29 + 32 + \dots + 24 + 65}{20} = 36.25$$

### (三) 对总体中的子组或域的比例和均值的估计

女性总体中单身女性所占比例的估计值, 即样本中婚姻状况回答为“单身”的女性在所有女性样本中所占的比例:

$$\frac{\text{女性}}{\text{女性}} = \frac{4}{20} = 0.33$$

总体中男性平均年龄的估计值即样本中男性的平均年龄:

$$\frac{\text{男性}}{\text{男性}} = \frac{29 + 32 + 16 + 43 + 22 + 49 + 18 + 48}{8} = 32.18$$

### 例 8.6b

使用权数，重复前例所做的各项估计：

### 1. 对总体或总体中域的总量的估计。

总体总人数的估计值为：

$$\text{应} \Rightarrow \text{吧} = \frac{\$ + 5 + 5 + \dots + 5 + 5 + 5}{20} = 100$$

总体中男性总数的估计值，为样本中所有男性的权数之和：

$$N = \text{f. } W. = \frac{5+5 + 5 + 5 + 5 + 5 + 5 + ?}{\text{i 住 n 男性}} = 40,$$

类似地，女性总数的估计值为：

$$N = V \quad W_t = \frac{?+5+5 + 5 + 5+ 5+ 5+ 5+ 5+ 5+ 5+ 5+ ?}{苦} = 60$$

<5:D 女性

## 2. 对总体的比例和平均数的估计。

总体中男性所占比例的估计值为总体中男性总数的估计值与总体总人数估计值的比值：

$$P = \frac{iws}{yw} = 0.4$$

总体平均年龄的估计值，为总体所有年龄和的估计值与总体人数估计值的比值（它等于 样本平均年龄）：

$$Y = \frac{mW}{20} = \frac{5+5+5+\dots+5+5+5}{20} = 36.25$$

### 3. 对总体中域的比例和平均数的估计

女性中单身所占比例的估计值为单身女性总数的估计值与女性总数估计值的比值：

$$\frac{20}{60} \ll 0.33$$

总体中男性平均年龄的估计值为男性年龄总和的估计值与男性总数估计值的比值。

$$Y = \frac{5 \times 29 + 5 \times 32 + \dots + 5 \times 18 + 5 \times 48}{8} = 32 ]$$

例 8.6a 中忽略权数和例 8.6b 中使用权数，估计平均数和比例时得到了相同的结果，这是因为本例是简单随机抽样的自加权设计，而且也没有无回答的调整。

### 三、不等概率抽样设计的变量估计

在估计过程中，正确使用最终权数不仅对计算不等概率抽样设计的估计值非常重要，而且对于当最终（经调整后的）权数不全相等时，计算自加权设计的估计值也非常重要。

### 例 8.7

进行一项关于农村扶贫农民收入水平改善的调查，假设某地区共有  $10$  个村，分为平原和山区两层，平原有  $34$  个村，即  $N=34$ ；山区有  $66$  个村，即  $M=66$ 。现在从中抽取  $10$  个村为样

本其中，平原抽取6个村，即 $\frac{6}{10} = 0.6$ ，山区抽取4个村，即 $\frac{4}{10} = 0.4$ 。因此，第一层的抽选概率为

n. 6	34	几 2 4
/二】，其抽样权数为抽选概率的倒数 $2=\frac{1}{0.6}=1.67$ 。第二层的抽选概率为 $\frac{4}{10}=0.4$ ，其抽样权数为		
M 34	6	$N_2 = 66$

丝=16.5。所有的10个村都给出了回答，因此无须对无回答作调整。现在得到表8-4所示的4

数据文件，要求估计有关总体的估计值。（注：村类型中的C表示以种植业为主的村，L表示以养殖业为主的村）。

表 8-4 10个调查村的相关指标与权数

层	权数	村类型	收入/元
1	5.67	C	75 000
1	5.67	L	15 000
1	5.67	C	125 000
1	5.67	C	67 000
1	5.67	L	80 000
1	5.67	C	40 000
2	16.5	C	30 000
2	16.5	C	14 000
2	16.5	L	48 000
2	16.5	C	22 000

总体农村总数的估计值为：

$$V = v \cdot \sum_{i=1}^{10} w_i = 5.67 \times 6 + 16.5 \times 4 = 100$$

这和原来是一致的。养殖村总数的估计值为：

$$W_t = v \cdot \sum_{i=1}^{10} w_i = 5.67 + 5.67 + 16.5 = 27.84 \approx 28 \text{ 户养殖村}$$

养殖村所占比例的估计值为：

$$\begin{aligned} & V \text{ 吧} \\ & \frac{\text{户养殖村}}{\text{户养殖村} + \text{户种植村}} = \frac{28}{100} = 0.2784 = 27.84\% \\ & i.e.S \end{aligned}$$

总体中所有村总收入的估计值为：

$$V \text{ JF.} = 5.67 \times 75 000 + 5.67 \times 15 000 + \dots + 16.5 \times 22 000 = 4 160 340 \text{ (元)}$$

1

I  $\in S$

总体中所有村的平均收入的估计值为：

$$Y = \frac{V \text{ JF.}}{100} = \frac{4 160 340}{100} = 41 600 \text{ (元)}$$

养殖村总收入的估计为：

$$P_{\text{养殖户}} = \frac{\sum_{\text{养殖户}} \text{收入}}{\sum_{\text{所有村}}} = \frac{5.67 \times 15000 + 5.67 \times 80000 + 16.5 \times 48000}{27.8} = 1330650 \text{ (元)}$$

养殖户平均收入的估计值为：

$$\bar{y}_{\text{养殖户}} = \frac{1330650}{27.8} = 47769.34 \text{ (元)}$$

域估计值的计算可以定义一个新变量，当单元属于该域时，它等于指标的原值，其他情况下则取值为0。在上述例子中，如果要计算养殖户总收入的估计值，可以设置一个新变量，当样本村是养殖户时，就取该村的收入值，若是其他类型的村则取值为0。对整个样本计算这个新变量的估计值，即得到与前面相同的结果。

在实际中样本单元的数量和所要估计的指标数量可能非常多。目前已发展了一些有效的计算技术可进行大规模的计算。例如对定性数据的估计，可以先对数据按类别检索，然后加总各类的权数。

一些常用的定量数据估计方法也可用于计算定性数据的估计值方法是对定性变量的每一个类别，定义一个特殊变量：当单元属于该类别时取值为1，否则取值为0。于是，属于该类别的也即具有相应特征的单元总数的估计，即是对所有回答单元，这个新变量的值与权数乘积的总和。通过这种变换，定性变量与定量变量的估计方法就完全相同。在前面的例子中，我们可以定义一个变量，当一个村是养殖户时取值为1，否则取值为0。将这个新变量看作定量变量，对它的总值估计就是养殖户总数的估计。

以上介绍的都是利用计算机来进行汇总估计的方法。如果是用手工汇总，分层随机抽样对总体总量或总体平均数进行估计时，也可以先计算各层的估计值，再汇总或加权。应该注意的是当各层的抽选概率不等时，必须以总体单元的总数或比重来加权。

例如在上面的例子中，当要计算农村地区总体的总收入和每个村的平均收入时，我们也可以先分别算出平原和山区各层每个村的平均收入，然后加以汇总。

平原层村的平均收入：因为在层内是等概率抽选的，可以直接计算平均数

$$\bar{y}_{\text{平原}} = \frac{75000 + 15000 + 125000 + 67000 + 80000 + 40000}{6} = 67000 \text{ (元)}$$

山区层村的平均收入：

$$\bar{y}_{\text{山区}} = \frac{30000 + 14000 + 48000 + 22000}{4} = 28500 \text{ (元)}$$

农村总体总收入的估计值为：

$$y = ME = 34 \times 67000 + 66 \times 28500 = 4159000 \text{ (元)}$$

每个村的平均收入的估计值为：

$$\bar{y} = \frac{\sum_{\text{所有村}} \text{收入}}{N} = \frac{34 \times 67000 + 66 \times 28500}{100} = 41590 \text{ (元)}$$

其结果和前面的计算结果是一致的，尾数稍有出入，主要是在计算权数时，对尾数实施了四舍五入的缘故。

## 辅助信息的使用

原则上，设计权数与无回答调整因子的乘积可以用来估计所需的总体指标。但是，有时候我们也可以从其他渠道，比如从以前的普查中获得有关调查总体的某些信息，将它们合并到权数中。

使用辅助信息主要有两个原因，一. 首先，为使调查估计值与已知的总体总值，或者从另一调查所得的估计值相匹配，，例如，许多社会调查使用最新的人口普查数据来调整估计值，以确保这些估计值（如年龄、性别分布等）的一致性。辅助信息也可以从行政数据中获得，或从另一个被认为更加可靠的调查中获得，这些调查或者具有较大的样本量，或者其调查结果被认为更有权威性。其次，是为了提高估计值的精度。前面讲过抽样方差是度量抽样误差的一个指标，估计值的抽样方差越小，精度就越高。在计划一项调查时，评估用于提高精度的所有外部信息非常重要，将辅助信息与抽样设计相结合，是提高估计精度的一个重要途径。

要想在调查的设计阶段使用辅助信息，抽样框中的所有单元都必须具有这类辅助信息。否则，只有在获得所有被调查单元的辅助变量的值后，才能在估计阶段利用辅助信息提高估计值的精度。

例如，抽样框中没有年龄信息，那么就不能利用年龄对样本进行分层。但是如果通过调查获得了这个信息，那么就可以对调查的估计值进行调整，使之与人口普查的年龄分布相匹配。另外，如果年龄与调查所得的其他变量（例如表示健康状况的变量）相关，那么就可以利用它来提高这些估计值的精度。

辅助变量还可以用来对不同子总体所对应的不同无回答率进行修正。它还可以用来调整由于不完善的抽样框，或称涵盖误差导致的调查总体与目标总体之间的差异所造成的影响。若想在估计阶段成功地使用辅助信息，应具备以下三个基本条件：

- ① 有关总体的外部信息必须是准确的；
- ② 应收集所有样本的辅助信息；
- ③ 辅助信息与调查变量之间存在相关性。

辅助信息的准确与否非常重要。辅助信息不仅要可靠，而且外部数据应来自同一个总体，并且应建立在与该调查可比的概念、定义、参照期等基础上。为了修正抽样框的不完全涵盖，辅助资料必须能涵盖调查的目标总体。下面介绍如何在估计中使用辅助信息。

### 一、事后分层

事后分层是指利用数据收集之后的辅助信息对变量进行分层，进而调整样本权数的方法。在抽样方法设计时，分层抽样可以提高抽样效率。但在数据收集之前，可能因得不到适合分层的抽样框信息（如年龄与性别）而无法进行分层。但如果可能从其他方面获得分层信息，或者调查样本后得到了更新、更可靠的分层信息，那么可以利用数据收集的结果，选用其中合适的变量对收集到的样本进行分层，这就是通常所说的事后分层。事后分层常用于计数形式的辅助数据，例如，总体中性别变量。所调查变量的总体均值在各层之间的差异越大，事后分层的效果就越好。下面举例说明如何使用事后分层，对某公司吸烟人数的估计进行改进。

### 例 8.8

为得到某公司职员是否有吸烟习惯的信息，进行了一项调查，调查前不清楚公司职员分性别信息。该公司共有 156 人，即  $N = 156$ ，在调查设计阶段，并没有可用于分层的辅助信息，故用简单随机抽样抽取了  $n = 50$  人的样本。在收集关于吸烟习惯信息的同时，收集了每个回答者的年龄和性别信息。共有  $n_r = 30$  人进行了回答，由此得到样本数据的分布如表 8-5 的示。

表 8-5 某公司职员吸烟习惯信息

回答者数量	男性	女性	总计
总人数	6	24	30( $n_r$ )
其中：吸烟者人数	2	10	12

1. 每个回答者的权数以及对无回答所做的调整计算如下：设计权数是入样概率的倒数：

假定样本中每个被调查者回答的概率都相同，用调查的总人数除以回答者的总人数，就得到无回答调整因子：

无回答调整因子  $1 - \frac{1}{67}$

$$n_f = 30$$

无回答调整权数即为设计权数与无回答调整因子的乘积：

$$W_{nt} = W_{tl} = \frac{n}{n_r} = \frac{30}{12} \times 1.67 = 5.21$$

2. 利用无回答的调整权数，即可得到如表 8-6 所示的调查估计值。

表 8-6 某公司职员吸烟行为信息的调查估计值

调查的估计值	男性	女性	总计
总人数	31.2(6x5.2)	124.8(24x5.2)	156.0(30x5.2)
吸烟者人数	10.4(2x5.2)	52.0(10x5.2)	62.4(12x5.2)
吸烟者比例	0.33	0.42	0.40

注：括号内为计算过程

通过无回答调整权数，我们估计出该公司约有 31 名男性职员和 125 名女性职员，而且男女吸烟的比例不同。

假定调查完成后我们得到了如下的辅助信息：该公司共有 84 名男性职员和 72 名女性职员。这说明通过调查所得的估计值与真值之间存在明显差异。这时如何修正？

一般而言，总是希望调查所得的估计值与已知的男性和女性数量相一致，并认为一个人是否吸烟与其性别之间可能存在相关。这时可以利用事后分层提高估计的精度：

3. 可以在抽样后对样本进行分层，计算事后分层权数用于估计。事后分层权数是事后层的辅助变量总和与该层回答单元数量之比。在本例中，男性的事后分层权数为：

$$\text{中}= \frac{1}{84} = 0.0119$$

女性的事后分层权数为：

当使用这些新权数时，估计值就变为表 8-7 所示的结果。

表 8-7 某公司职员吸烟行为信息经事后分层权数调整后的估计值

事后分层估计值	男性	女性	总计
总人数	84(14x6)	72(3x24)	156
吸烟者人数	28(14x2)	30(3x10)	58
吸烟者比例	0.33	0.42	0.37

注：括号内为计算过程

此时，男性和女性人数的估计值与该公司男性和女性的实际人数一致。而且，由于吸烟者 人数及比例与性别之间存在相关，估计的精度得到了显著的提高：注意到在每个事后分层的 层中吸烟者的比例并没有改变，但总体吸烟者比例的估计发生了改变。

在实际中，对每个事后层都计算一个因无回答引起的调整因子，这个因子即总体中各事后层的单元数量 N 与其事后分层所得的估计值应之比。（后者用无回答的调整权数进行估计）。本例中的事后分层调整因子为：

$$\frac{84}{31.2} = 2.69$$

% 72

女性：  $\frac{72}{124.8} = 0.58$

将这个调整因子应用到无回答调整权数时，就可以得到与前面相同的最终事后分层权数：

$Ng$ 、

男性：  $W_{\text{网}} \curvearrowleft \text{产町} \cdot 1 = 5.2 \times 2.69 = 14 Ng$

女性 %  $5.2 \times 0.58 = 3$

## 二、比率估计

比率估计是指当目标变量与已知的辅助变量之间存在相关关系时，利用辅助变量值与同类 样本估计值的比率作为乘数因子对各类权数进行调整的方法。采用比率估计方法，可以用来改 进调查估计量的精度。例如，如果调查的目的是估计某地区的产品销售总额，那么这一地区的居 民可支配收入就是一个很有用的辅助信息。由于该地区的产品销售收入与居民可支配收入相 关，那么辅助信息就可以用来改进对地区产品销售收入的估计。下面举例说明。

### 例 8.9

某县农业部门，需要估计该县年末生猪存栏头数 该县共有 276 个村，以村为抽样单元，采取简单随机抽样，抽取 24 个村为样本，考虑到村的大小差别比较大，因此饲养生猪数量的差 别也比较大，采用简单估计方法的误差比较大，也即抽样效率不高。需要利用一个辅助变量采 用比率估计方法。根据分析每个村的户数与生猪存栏数之间有较强的相关关系，而且资料也 容易取得，所以将户数作为辅助变量。取得如表 8-8 所示的数据

表样本村户数与生猪存栏数数据

样本村	户数(药) /户	生猪存栏数(yi) /头	样本村	户数(xi) /户	生猪存栏数(yi) /头
1	158	266	13	320	346
2	930	842	14	289	320
3	462	540	15	76	96
4	734	690	16	246	230
5	294	420	17	166	188
6	42	38	18	154	202
7	650	1 020	19	230	198
8	532	840	20	144	154
9	76	169	21	69	70
10	168	154	22	102	110
11	84	90	23	36	30
12	330	408	24	90	86

本例可以有两种估计方法 一种是前面介绍的简单随机抽样的简单估计方法，即以样本 的均值来估计总体均值，也就是先求出平均每村生猪存栏数，再用总村数乘以每村生猪存栏 数。即：

$$\bar{y} = \frac{\sum y_i}{n} = \frac{1}{n} \sum y_i = \frac{1}{24} (266 + 842 + \dots + 408) = \frac{1}{24} \times 330.5 = 13.77 (头)$$

另一种就是比率估计方法，以户数作辅助变量，在获取生猪存栏数的同时，要取得户数的 资料。根据样本求得生猪头数与户数之间的比率，用  $r$  表示，则有：

$$r = \frac{\sum y_i}{\sum x_i} = \frac{1}{24} (266 + 842 + \dots + 408) / \frac{1}{24} (158 + 930 + \dots + 330) = 1.1759$$

这个比率的直观含义就是样本中平均每户的生猪存栏数，因此要估计全县的生猪存栏数就需要乘以全县的总户数。假设已知全县农村共有 75 000 户，则估计全县的生猪存栏数为：

$$\hat{y} = \bar{x}_r = \bar{x} \cdot r = 75000 \times 1.1759 = 88182.5 (头)$$

这两种方法的计算结果不同，从理论上说，用比率估计方法的估计量是有偏的，但随着样 本量的增大，其偏差会缩小。此外如果辅助变量与目标变量之间存在较高的相关性，比率估计 量的效率比简单估计的效率要高，抽样误差较小。但它计算误差的公式比较复杂，有兴趣的读 者可参阅有关抽样的书籍：

### 三、更复杂的权数调整：广义回归和校正

比率估计基于这样一个假定，即调查的指标与辅助变量之间存在一个简单的倍数关系。例如，总人数扩大两倍意味着吸烟的人数也相应增加两倍，耕地面积增加五倍，意味着小麦的 播种面积也相应扩大五倍 c 然而这个假定过于简单，现实情况要复杂得多。首先，使用一个辅 助变量有时对改进估计还不够；其次，调查指标与辅助变量之间的关系有时比简单的倍数关系

更为复杂：在假定不成立的情况下可以采用回归估计，回归估计是一种更为复杂的估计方法。它使得我们可以考虑更高级的模型■

对于存在多个调查指标需要估计的情形下，要求保证它们的加权估计总值与所参照的总体总值保持一致，此时就应将事后分层和比率估计的思想加以推广，以产生更加复杂的估计方法 C 例如，总体可以分别按照年龄或性别进行分组，但是每个年龄组中男性和女性的人数未知，这时就不能同时对两个指标使用事后分层。

对这种多个指标的情形，已开发了若干广义估计方法，这些方法能同时保证多个指标的估计值与其参照值的一致性。这些广义估计方法中，使估计值与其参照值一致的过程被称为校正，所计算的权数调整因子称为校正因子。与事后分层和比率估计一样，这些方法的结果也是对权数进行调整。用于估计的最终权数是最初的设计权数与校正因子的乘积。感兴趣者可参阅广义估计方法的相关内容。

## 第四节

### 调查估计量的抽样误差及区间估计

%

通过调查所得的估计值都会有误差，前面我们提到有两种类型的误差：抽样误差和非抽样误差 非抽样误差很难进行度量，但是抽样误差通过调查数据本身就可以进行估计。这一节 将专门讨论抽样误差的度量。

#### 一、抽样方差及其估计

在进行抽样设计时，希望估计量无偏，且估计量的抽样分布在估计的参数真值周围，变化越小越好「估计量的抽样方差是指使用相同的抽样设计时对所有可能的样本计算的估计值与其平均值离差的平方的平均。抽样方差的大小取决于：样本量与总体的大小、总体中所研究指标的变异程度、抽样设计、使用的估计量以及调查的无回答情况。具体表现为：

1. 如果其他因素不变，抽样方差的大小一般随样本量的增大而减少。然而，为获得给定的方差，所需样本量并不随总体大小的增大而等比例增加。实际上，在确定给定抽样方差所需的样本量时，对中等规模的总体来说，总体大小仅起有限的影响；而对大的总体则几乎没有影响 e

2-抽样方差会随着总体指标的变异程度变化而正向变化 在样本量相同的情况下，那些变异程度较大的指标，相比那些变异程度较小的指标具有更大的抽样方差。

3. 在样本量给定的情况下，抽样方差还依赖于抽样设计因素。一种抽样设计可能比另一种设计更为有效，因其估计量具有更小的方差。例如分层抽样通常比简单随机抽样更精确。抽样设计的其他方面，如样本在层间的分配、抽样单元和抽选方法的选择等都会影响抽样方差 的大脱

4. 一些估计量可能比另一些估计量具有更小的变异性，因而抽样方差更小。例如，我们已经知道事后分层和比率估计都能提高估计的精度° 因此在开发抽样设计和估计程序时，应该利用现有的信息制定有效的抽样策略、以降低抽样方差□

5. 无回答将使有效的样本量减少，从而导致抽样方差的增加，这是因为样本越小，方差会增大。

下面举例说明抽样方差的估计。

#### 例 8.10

对于一个由四个学生社团组成的总体，他们花费在音像制品上的支出如表 8-9 所示 ‘

表 8-9 四个学生社团音像制品支出相关数据

学生社团序号/个	音像制品支出/百元
1	10
2	20
3	30
4	40

如果从中抽取一个样本量为 2 的简单随机样本，对音像制品的平均支出进行估计，那么这个估计值的方差是多少？

总体平均支出的真值是 25（百元），即

$$\bar{y} = \frac{1}{N} \sum_{i=1}^N y_i = \frac{10 + 20 + 30 + 40}{4} = 25 \text{ (百元)}$$

为计算一个样本量为 2 的简单随机样本估计值的抽样方差，可以通过观察总体中所有可能样本量为 2 的简单随机样本得到，具体结果见表 8-10。

表 8-10 四个社团中抽取 2 个所有可能样本的抽样方差计算表

样本	样本单元 $k$	样本观测值	样本估计值 $\bar{y}$	样本估计值与均值之差	样本估计值与均值之差的平方 $(\bar{y} - 25)^2$
(1)	(2)	(3)	(4)	(5)	(6)
1	(1,2)	10,20	15	-10	100
2	(1,3)	10,30	20	-5	25
3	(2,3)	20,30	25	0	0
4	(1,4)	10,40	25	0	0
5	(2,4)	20,40	30	5	25
6	(3,4)	30,40	35	10	100
	平均数	—	25	—	41.7

抽样方差的计算步骤如下：

1. 计算所有可能的样本均值的平均数。

$$\bar{\bar{y}} = \frac{1}{6} (15 + 20 + 25 + 25 + 30 + 35) = \frac{15+20+25+25+30+35}{6} = 25 \text{ 百元}$$

注意：计算结果等于总体均值的真值。这是预料之中的事，因为简单随机抽样得到的样本均值是总体均值的一个无偏估计。

2. 计算所有可能的样本估计值与它们的平均数之间的差，称为离差。如对第  $k$  个样本，其 离差为  $\bar{y}_k - \bar{y}$ ，具体结果见表 8-10 中的第 (5) 列。

3. 计算这些离差的平方，即  $(Y_i - \bar{Y})^2$ ，见表 8-10 的第(6)列口

4. 计算这些离差平方的平均数:

$$\begin{array}{r} \text{£ } (\text{p(A)} - \bar{x}) 2 \\ 6 \end{array} \quad \begin{array}{r} 100 + 25 + 0 + 0 + 25 + 100 \\ 6 \end{array} \quad \begin{array}{r} -41.7 (\bar{x}_1) \end{array}$$

因此，从该总体抽取的样本量为 2 的简单随机样本，其总体均值估计的抽样方差为 41.7 元。

实际中，我们不可能把所有的样本都抽出来，以计算估计值的方差。但是如果使用概率抽样方法，根据单个样本就能估计抽样方差。对于那些标准的抽样设计，如简单随机抽样、分层抽样等，抽样方差的计算公式可以在任何一本抽样理论的教材中找到。如果采用的抽样设计或估计程序较为复杂，那么就需要使用本章后面讲述的其他方法。

如前所述，估计量的抽样方差是总体变异的函数。也就是说与总体方差有关。总体方差的计算公式如下：

$$\frac{f(K, -F)^2}{N-1}$$

在本例中，总体方差等于：

$$\text{痔一}'' \quad (10 - 25)^2 + (20 - 25)^2 + (30 - 25)^2 + (40 - 25)^2 = 500$$

£ (丸 2  
，其中，—<sub>3</sub> 才是样本均值。

N - 1

4-1

我们知道，对简单随机抽样，总体均值

对简单随机样本，可以证明， $\bar{Y}$  的一个无偏估计量是： $\bar{Y}$  的估计量为样本均值，即  $\bar{y} = \frac{1}{n} \sum y_i$  对不重复简单随机抽样，可以证明  $y$  的抽样方差公式为：

$$\frac{n}{N}$$

当总体方差未知时，它可以用下式来估计：

$$\frac{n}{N}$$

前面介绍过号称为抽样比,通常用 $\lambda$ 来表示,  $\lambda^2$ 是样本方差。从这个公式我们可以看出, 随着样本量 $n$ 的增加, 均值估计的抽样方差将减少。当样本量大到是全面调查总量时, 即 $n = N$ 时, 均值估计就完全没有抽样方差。当抽样比很小时, 因子 $(1 - \lambda)$ 就可以忽略, 科克伦在《抽样技术》一书中提出, 当抽样比不超过5%, 这个因子可以被忽略。另外应注意上述公式仅应用于当回答率为100%, 且没有使用任何辅助信息对样本权数进行调整的情形。

除了上面的抽样方差以外，估计量抽样误差的大小还可以用标准差和变异系数来衡量。

估计量的标准差，也称标准误它是抽样方差的平方根，这种度量方式比较容易理解，因为它的量纲与估计值相同，而抽样方差的量纲是指标单位的平方。在很多情况下，结合估计值本身的大小，用标准差来衡量误差大小可能更加直观有用但是有时候即使是标准差也很难解决“多大的标准差是可以接受的？”这个问题。标准差的大小取决于估计值的大小，例如标准差为100公斤，对于每亩地的平均产量来说可能很大，对一个乡的总产量来说就显得，不是很大。

变异系数是估计值的标准差与估计值本身之比，它解决了标准差不能直接对比的问题，它把误差与估计值联系在一起，通常用百分比来表示：当比较不同规模、不同量纲单元的样本估计值的精度时，变异系数非常有用。

## 二、复杂抽样设计的方差估计

到目前为止，我们在估计抽样方差的讨论中都局限于简单随机抽样，关于其他典型的抽样方式的抽样方差估计，也可以在大多数抽样理论书籍中找到。但是在实际中，抽样设计经常无法满足公式应用所需的各项条件，或者会让公式变得很复杂。复杂抽样设计经常包括分层、多阶抽样、不等概率抽样、多相抽样等多种方式的综合应用，在估计方法上也包括采用比率估计、回归估计等非线性估计量因此无现成的公式可以套用。怎样对估计量的方差进行估计？这是一个相当困难的问题目前已有一些其他方法可用于抽样方差的估计，而不管它的抽样设计或估计程序有多复杂，复杂抽样设计方差估计的一种方法是自举法。

自举法是近年发展起来的解决上述问题的一种较为有效的方法，它是从1个样本的资料产生“新”的能够代表被研究总体的多个样本，从而模拟出参数的分布，并得到其平均数和标准差。依赖于电子计算机的大容量和高效率，自举抽样程序已非常容易实现。

设 $I$ 个样本具有变量 $Y_1, Y_2, \dots, Y_I$ ，用某种方法得到其对总体参数的估计值 $\hat{\theta}_I$ ，则在抽样时，每轮都要从这几个变量中抽取 $n$ 个容量也是 $n$ 的随机样本，并保证每一变量在每轮每次抽样中被抽取的概率都是 $1/n$ 。于经典方法的重复抽样此过程一般可用随机数表或 $n$

由计算机输出 $n$ 个 $0 \sim 1$ 之间的随机数实现。即先将样本变量编码为 $1, 2, \dots, I$ ，然后取分组距离 $1/n$ ，将区间 $[0, 1]$ 分成 $n$ 个互斥的分隔：

$n$

$$\begin{array}{c} 1 \\ \hline n & n & n & n & n \end{array} \quad (8.1)$$

出现的随机数，凡在 $0 \sim 1/n$ （下限用 $W$ 号，上限用 $E$ 号，即 $0WR$ （上；下同）之间的取编码 $n$

$2$   $1$

“1”变量，凡在 $n/n \sim 1/n$ 之间的取编码“2”变量，……，凡在 $(n-1)/n \sim 1/n$ 之间的取编码变量。这就

称为自举抽样，由这 $n$ 个“新”变量组成的样本则称为自举样本(bootstrap sample)当对原始样本变量重复进行 $s$ 轮自举抽样，就得到 $n^s$ 个容量均为 $n$ 的自举样本，再对每一样本都按照在原始样本中所用方法估计 $\hat{\theta}_s$ ，得到 $s$ 个次数为 $n$ 的分布。这称为自举统计参数分布。根据此分布，我们就可以研究。「的基本性质，其最主要特征数是自举平均数 $O_s$ ：( bootstrap

average) 和自举标准误  $\text{SE}$  (bootstrap standard error)，定义为：

$$e = \bar{y} - \bar{y}_0; \quad (8.2)$$

$$\begin{aligned} & I_m \quad -] \quad 1/2 \\ & \text{粉} = r \quad (\text{露}''V) \quad (8.3) \\ & L_m J &= i \end{aligned}$$

不论行是什么样的统计指标，自举抽样都能稳健地提供它的平均数和标准误估计。这是自举法的一个重要特征。

如果能能够描述有关随机变量的取值中心（即通常所说的“无偏估计”），则和之间的差异应在误差（标准误）所许可的范围内；如果  $\theta$  的标准误为未知，则将是其最佳替代。

一般为估计  $\mu$  和  $s$  补， $n=100$  即可（这在露和标相对稳定的意义上而言），不必大于 200。但是如果要了解露的较精确分布和对于的 100 ( $1-a$ ) % 置信区间，一般应选  $m=1000$ 。这时，在露值从小到大的顺序排列中，“6, 26 和 51 的露值依次对应于。的置信度为 99%, 95% 和 90% 的区间下限；而  $\alpha = 995, 975$  和 950 的”：值则依次为相应的区间上限。它们与。的分布形状（如是否左右对称）无关。

#### 例 8.11

在一项关于病患的调查中获得 10 位被调查者的年龄，列于表 8-11 列 2，计算的统计指标有：年龄的算术平均数  $\bar{y}$ ，标准差  $S$ ，平均数的标准误或，标准差的标准误  $\text{SE}$ 。其计算方法为：

$$\begin{aligned} \bar{y} &= \frac{1}{n} \sum_{i=1}^n y_i = \frac{260}{10} = 26.0 \\ S &= \sqrt{\frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2} = \sqrt{\frac{1}{9} \sum_{i=1}^n (y_i - 26)^2} = \sqrt{\frac{1}{9} (100 + 49 + 16 + 16 + 16 + 16 + 16 + 16 + 16 + 16)} = \sqrt{\frac{1}{9} \cdot 180} = \sqrt{20} = 4.47 \end{aligned}$$

$$18.5$$

$$\begin{aligned} \text{SE} &= \frac{S}{\sqrt{n}} = \frac{4.47}{\sqrt{10}} = 1.41 \\ \text{SE}_{\bar{y}} &= \frac{S}{\sqrt{n}} = \frac{4.47}{\sqrt{10}} = 1.41 \\ \text{SE}_{\bar{y}} &= \frac{S}{\sqrt{n}} = \frac{4.47}{\sqrt{10}} = 1.41 \end{aligned}$$

此例  $n=10$ ，故根据式 (8.1)，自举随机号的分组数列为  $0 \sim 0.1; 0.1 \sim 0.2; 0.2 \sim 0.3; 0.3 \sim 0.4; 0.4 \sim 0.5; 0.5 \sim 0.6; 0.6 \sim 0.7; 0.7 \sim 0.8; 0.8 \sim 0.9; 0.9 \sim 1.0$ 。其第  $I \sim 3$  个自举样本及其计算结果均列于表 8-11。经 100 次自举抽样后，各自举统计指标的变化区间及根据式 (8.2) 及 (8.3) 计算的平均数、标准误则列于表 8-12。

表 8-11 原始样本与 3 个自举样本数据表及指标计算表

原始样本		第 1 个自举样本			第 2 个自举样本			第 3 个自举样本		
变1#	变量Y	随机号	变量号	5>	随机号	变量号	变量I	随机号	变最号	变量*
1	6	0.627	7	29	0.240	3	13	0.302	4	18
2	9	0.255	3	13	0.827	9	39	0.468	5	20
3	13	0.918	10	69	0.950	10	69	0.614	7	29

续表

原始样本		第1个自举样本			第2个自举样本			第3个自举样本		
变量号	变量 $Y$	随机号		「)	随机号	牡号	J)	随机?	2 吊	变量 y
4	18	0.355	4	18	0.240	3	13	0.945	10	69
5	20	0.369	4	18	0.066	1	6	0.548	6	i2
6	22	0.082	1	6	0.855	9	39	0.804	9	39
7	29	0.677	7	29	0.239	3	13	0.808	9	39
8	35	(0.538	6	22	0.416	5	20	(1.592	6	22
9	39	0.411	5	20	0.207	3	13	0.402	5	20
10	69	0.836	9	39	0.239	3	13	0.241	3	13
离=	26.0			26.3		们:	23.8	的	29.1	
$s_{\bar{e}} = s$	18.5			17.6		円:	19.5		16.4	
或二 s;	5.85			5.56			6.16		5.20	
	4.14			3.92			4.36		3.68	

表 8-12 100 次自举抽样各统计指标的平均数和标准误 ( $/\ll = 100$ )

统计指标	自举样本变化区间	平均数 $e$ :	标准误%:-
$e = \bar{y}$	16.0-36.7	25.01	4.77
$s_{\bar{e}} = s$	7.4-25.6	16.46	4.83
	23.8-1	5.20	1.53
	L6-5.7	3.68	1.08

表 8-11 和表 8T2 的结果主要说明:

(1)  $\bar{e}$ : (表 8-12) 和。 「(表 8-11) 都有一定差异, 但其最大差异都在  $\pm 0.5$  个标准误差范围内。例如对于”是  $(26.0-25.01)/4.77 = 0.208$ , 对于  $s$  是  $(18.5-16.46)/4.83 = 0.422$ 。这表明各个  $Q^*$  都是无偏估计指标。

(2) 三个样本平均数的标准误 ( $s_{\bar{e}}$ ) 来自各不相同的定义: 5.85 是基于观察样本随机抽自 正态总体的假设; 4.77 是自举样本平均数分布的标准误; 5.20 是自举样本平均数标准误的平均数, 其最大差异达到  $(5.85 - 4.77)/1.53 = 0.706$  个标准误。而具有相应定义的三个样本 标准差的标准误 ( $s_e$ ) 4.14, 4.83 和 3.68, 其最大差异达到  $(4.83 - 3.68)/1.08 = 1.065$  个标准 误。它们虽然都未达到显著水平, 但与和  $Q^*$  相比, 差异有扩大趋势。这表明标准差有较 大的抽样变异。

根据原始样本计算的标准差为 18.5, 而根据自举法估计的标准差为 16.46

### 三、不同抽样方式下调查指标的区间估计

构造置信区间的理论依据是中心极限定理, 对于一个大样本, 总体平均数的估计值是 R 它的标准差的估计值是那么根据中心极限定理和正态分布的理论, 我们得到如下的结果:

绝对误差  $|V-W|$  超过 1.65 倍标准差的概率为 0.1 (对应于 90% 的置信区间)；

绝对误差  $|Y-\bar{Y}|$  超过 1.96 倍标准差的概率为 0.05 (对应于 95% 的置信区间)；

绝对误差  $|U-\bar{U}|$  超过 2.58 倍标准差的概率为 0.01 (对应于 99% 的置信区间)，具体见表 8-13。

表 8-13 Z 值与置信度的关系

Z 值	概率值 (置信度, %)
1	68.27
1.645	90
1.960	95
2	95.45
2.575	99
3	99.73

有关置信区间理论的详细内容，可以参阅任何一本抽样理论方面的书或统计学的书。

### (一) 简单随机抽样指标的区间估计

下面举例说明简单随机样本调查指标的估计。假定所有的单元都做出了回答，而且没有利用辅助信息做任何调整。

#### 例 8.12

从一个 500 人的总体中，采用简单随机抽样抽取一个 10 人的样本，每个入样单元的年龄 见表 8-14。求平均年龄估计值的抽样方差、标准差与变异系数，并计算在 95% 的置信水平下的误差限与相应的置信区间。

表 8-14 10 人样本的年龄分布及其抽样方差计算表

样本单元	单元的年龄	估计值与平均数之差	离差的平方
1	21	-13.4	179.56
2	26	-8.4	70.56
3	27	-7.4	54.76
4	32	-2.4	5.76
5	34	-0.4	0.16
6	37	2.6	6.76
7	38	3.6	12.96
8	40	5.6	31.36

续表

样本单元	单元的年龄	估计值与平均数之差	离差的平方
9	42	7.6	57.76
10	47	12.6	158.76
合计	—	—	578.40

1. 计算样本的平均年龄

2. 计算抽样方差的估计值

计算每个人的年龄与其平均数的差（见表 8-14 中第三列）；计算离差的平方（见表 8-14 中第四

$$\frac{21 + 26 + \dots + 42 + 47}{10} = 34.4$$

列）；

将离差平方之和除以样本量减 1，即

$$\text{再受沪 } 578.4 \\ \text{---} = \text{-----} = 64.27 \\ 10 - 1$$

抽样方差的估计值为：注意到抽样比斤十冒很小，可以忽略不计。

$$\text{CT} = \frac{\text{A}_2}{n} = \frac{64.27}{10} = 6.4$$

3. 计算其他相关指标 标准差的估计为

变异系数的估计为：

$$cp = \frac{2.5}{34.4} = 0.073 = 7.3\%$$

在 95% 的置信水平下的误差限和置信区间为：

误差限 =  $1.96a = 1.96 \times 2.5 = 4.9$

置信区间 =  $y \pm 1.96a = 34.4 \pm 4.9 = [29.5, 39.3]$

因此，我们可以说，在 95% 的置信水平下，总体平均年龄的真值在 29.5 和 39.3 之间。严格地表述应为：如果重复抽样多次，置信区间有 95% 的机会覆盖总体真值。

类似的方法可以推广至总体比例的估计。包括计算比例的方差、标准差和在一定置信水平下的区间估计。计算公式如下：

$$Sp = \sqrt{\frac{p(1-p)}{n}}$$

$$S_p$$

$$P - ZS \leq P \leq P + ZSp$$

系统抽样与简单随机抽样的区别在于具体实施的路径不同，除了其精度略高一点外，其估计理论方法与简单随机抽样相似 所以，不展开介绍。

## （二）分层抽样指标的区间估计

与简单随机抽样相比，分层抽样估计是在各层中先按简单随机抽样估计，然后依据各层的 权数进行加

权计算。由于在每层中取样，样本分布比较均匀，所以分层抽样估计的精度高于简单随机抽样。用叫表示样本分层；“表示第人层的第，个个体的 $y$ 值； $A$ 表示第 $A$ 层的样本平均数； $n$ 表示各层权数； $s^2$ 表示各层抽样方差； $r$ 表示各层抽样比例；“ $\alpha$ ”表示显著性水平； $z$ ”表示置信水平对应的双侧分位点。其具体统计指标的估计方法如下：

1. 计算各层样本平均数，采用简单算术平均方法。

$$\bar{y}_h = \frac{\sum y_{hi}}{n_h}$$

2. 计算各层抽样方差。

$$s_i^2 = \frac{\sum (y_{hi} - \bar{y}_h)^2}{n_h - 1}$$

3. 计算总平均数。以各层权数进行加权平均。

$$L$$

$$A = 1$$

4. 计算层间方差与标准差。

$$s^2 = \frac{\sum (L - \bar{y}_h)^2}{n_h - 1}$$

### 摘要

5. 进行样本平均数的区间估计。

$$P\{y - \bar{y}_h \leq \bar{y} \leq y + \bar{s}_h\} = 1 - \alpha$$

### (三) 整群抽样的区间估计

整群抽样的计数、加总、比例、分布四类统计指标的计算与简单随机抽样完全一样。但整群抽样的效率要低于简单随机抽样。用 $V$ 表示总体群数； $n$ 表示样本群数； $M$ 表示每群的平均规模，即每群所包含个体数目的平均数； $y_{ij}$ 表示第 $i$ 群的第 $j$ 个个体的 $y$ 值； $\bar{y}_i$ 表示第 $i$ 群的样本平均数； $\bar{y}$ 表示调查群的总平均数； $r$ 表示抽样比； $z$ ”表示置信水平对应的分位点；“ $\alpha$ ”表示显著性水平。则整群抽样统计指标的估计方法如下：

1. 计算样本平均数

2. 计算样本群间方差

3. 计算总体平均数抽样方差

$$\frac{nM}{n-1} (\bar{y} - \bar{y}_i)^2$$

#### 4. 计算总体平均数的置信区间

$$= 2 [ ( > ) ] , W \approx Wy + z_s [ \sigma ]^{1/2}$$

#### 5. 总体总量的置信区间。

$$NM y - z_{\alpha/2} NM [ ( ; ) ]^{1/2} \leq VW/VM y + z_{\alpha/2} NM [ p(y) ]^{1/2}$$

其他复杂抽样调查指标的估计可以采用各自复杂的公式进行计算，或采用自举法进行估计。

### (四) 本章小结 · · ·

估计的第一步需要考虑权数：设计权数是指由抽样设计所决定的每个样本单元所代表的调查总体的单元数，表示为样本单元入样概率的倒数。自加权设计是指所有样本单元设计权数都相同的抽样设计。这类抽样设计在计算诸如总值、均值等估计量时不需要进行权数调整。自加权设计包括简单随机抽样，系统抽样，按各层大小等比例分配样本的分层抽样不等概率抽样不属于自加权设计，需要进行加权调整。对单元无回答进行总值估计时需要进行加权调整。利用调查所得到的辅助信息进行事后分层和比率估计，可以提高估计的精度。对于标准的抽样设计，抽样方差可以采用常规方法进行估计，而复杂抽样需要采用自举法进行相关指标的估计，在此基础上再根据抽样方式的特点进行区间估计。

### 磁关键名词 · · ·

加权权数调整点估计区间估计事后分层比率估计抽样方差自举法

### 回諸回

#### 肯思考题

即测即评

1. 抽样估计中为什么需要使用权数？什么是设计权数？等概率抽样设计中如何进行加权调整？不
2. 等概率抽样设计中如何进行加权调整？如何利用辅助信息进行事后分层？如何利用辅助信息进
3. 行比率估计？简单随机抽样的抽样方差如何计算？
- 4.
- 5.
- 6.
7. 复杂抽样设计中如何利用自举法估计统计指标？
8. 如何针对不同抽样方式进行区间估计？

### 碧瘠章实训 · · ·

#### 一、 实训目的

1. 要求学生掌握简单随机抽样、分层抽样、整群抽样和复杂抽样方式下调查指标估计中的点估计与区间估计的计算方法。

2. 通过实训掌握加权的重要性与辅助信息利用的价值。

#### 二、 实训内容

首先，完成前述调查项目中调查指标的抽样估计，包括总量、比例、平均数的估计；

其次，按照简单随机抽样、分层抽样、整群抽样和其他复杂抽样的要求，确定合适的调查指标的估计。

第三，完成加权的处理、辅助信息的具体使用及自举法的应用。

#### 三、 实训组织

1. 将教学班级学生按4~6人的标准分成若干小组，每个小组采用组长负责制。每组选择一种抽样设

计中调查指标的估计方法，按照实训目的与要求，合理分工，团结协作。

2. 收集相关资料和数据。
3. 完成抽样估计的文本文件。
4. 各小组充分讨论，认真听取全体组员的意见，分析、修改和调整，在此基础上形成小组的 PPT 实训报告文稿以
5. 各小组在班级进行实训汇报和展示。
6. 教师进行成绩评定与总结口

# 第九章 市场调查的组织与实施



## \导语 \_\_\_\_\_

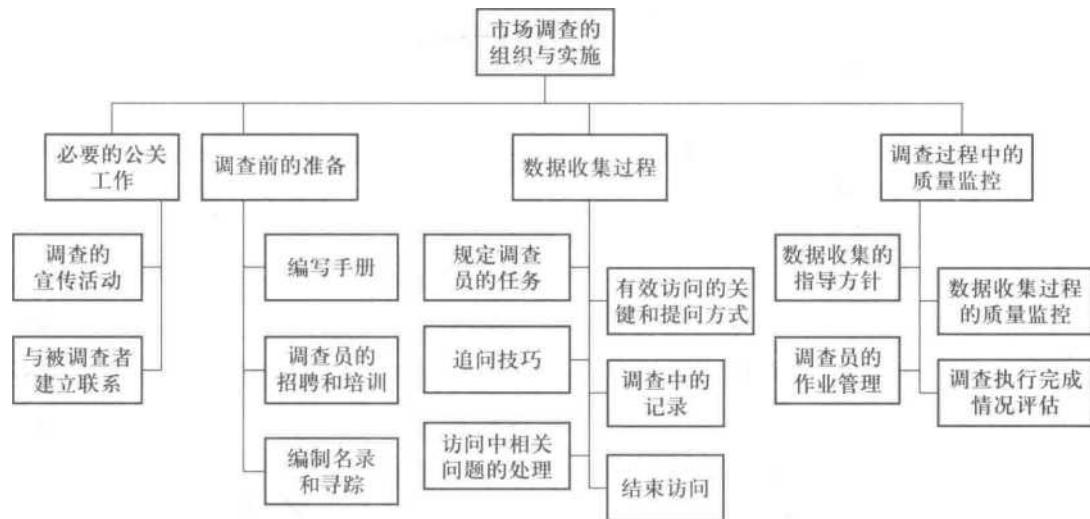
数据采集是调查中收集每一个被抽中的单元（住户、个人、农场、企业、学校等）个体数据 的过程。这一过程看起来很容易，但它是一项具体而复杂的工作，它可能会遇到各种意想不到 的情况。需要在这一过程中做好的事项包括：与被调查者建立密切融洽的关系；熟悉并使用调 查问卷；处理尴尬和敏感情况；控制文件（即数据收集过程中跟踪表格，例如寄出、返回、完成、 需要回访等的表格数目）；对数据的预审核（检查信息是否完整、正确）；对预审核时发现的错 误和无回答进行追踪回访；控制好调查质量

本章的目的是介绍在数据收集期间发生的各种活动和应该如何去组织、实施这些活动 应该尽可能高效地组织数据收集工作，同时保持各调查员在调查活动中的言行一致性。

## 、教学目标与要求

1. 熟悉市场调查实施的全过程；
2. 明确公关工作的主要内容；
3. 熟悉调查前的准备工作内容；
4. 掌握数据收集过程中的管理工作要点；
5. 掌握调查过程及质量控制过程的要点。

## 、本章知识结构



## 第一节

### 必要的公关工作

对任何调查机构而言，最好的公关策略是确保数据的可靠性、多渠道性、广泛适用性和高度受欢迎性，并切实尊重被调查者的个人隐私，从而建立和维持良好的专业信誉。良好的信誉有助于调查员赢得被调查者的合作与支持。公关工作既包括一些大型调查的宣传活动，也包括在小规模调查中使用的致被调查者的信、调查手册和一般机构介绍等各种宣传资料。为使被调查者了解调查的意义并能够与调查员积极合作，一项调查应当开展广泛的宣传活动。这种活动通常要达成两个目的：第一，增加本项调查活动的知名度，从而减少被调查者对调查员的怀疑；第二，增加本项调查对公众利益重要性的重视，以此激发被调查者的兴趣与合作意愿。

#### 一、 调查的宣传活动

根据调查的重要性、调查主题、经费预算和总体目标的不同，宣传活动选择方式的不同，大型调查通常包括以下部分：

1. 通过报纸、电台等媒体发布公告。
2. 在邮局和图书馆等公共场所张贴海报。
3. 积极处理好与媒体的关系（为报纸、电台和电视台等媒介提供报道和发言人）。
4. 给有关部门或与调查有关的协会单位和机构负责人写信，以获得他们的支持。

#### 二、 与被调查者建立联系

为增进被调查者的理解与合作，避免在公众中引起混乱，调查机构通常致力于与被调查者进行交流。最常用的交流材料包括：

##### （一）致被调查者的信

致被调查者的信提供调查的目的和解释调查的重要性，同时还提供数据收集的方法和日期。在邮寄调查时，包括调查手册与问卷信件的同时寄送。在面访调查时，包括调查手册的信件通常在与被调查者首次接触之前送出，并且在访问前一星期内送达被调查者。

##### （二）调查宣传手册

调查宣传手册可以是简单散页，描述调查对于公众利益的重要性。调查手册中应该举例说明数据的应用和数据的来源。手册可以送给所有的被调查者，或者仅仅在被调查者不愿接受调查的情况下使用。

##### （三）调查机构介绍手册

调查机构介绍手册是用于一般目的的小册子，说明调查机构收集和发布数据的多样性，以增加调查机构的可信度，它也经常用作与被调查者联系的手段。

## 第二节

### 调查前的准备

数据收集之前，有许多准备工作必须完成，具体包括编写手册、职员和调查员的招聘和培训、编制名

录和寻踪。

## 一、编写手册

调查现场实施成功的关键是调查员必须具有较强的能力和良好的个人素质。数据收集的质量依赖于他们工作的质量和一致性。条理清楚的指导手册和对现场工作人员进行专业化培训有助于保证调查的一致性和高质量。现场指导手册列出了标准的访问程序并给出了对调查中出现常见问题的处理方法。一般由调查项目领导小组设计手册，包括调查员手册、督导手册等，如果有必要，还要提出区域内的单元名录编制方法。

### （一）调查员手册

调查员手册是调查员的工作指南，调查员根据调查员手册开展工作。调查员手册通常包括以下内容：

1. 一般信息。这部分陈述调查的目的、重要性、数据的用途及调查机构收集数据的原则，并陈述为被调查者保密，忠实于被调查者的回答以及调查机构或调查员的职责等，同时还附上致被调查者的信和样本抽取方法的基本材料。

2. 简介。这部分解释怎样与被调查者进行第一次接触，怎样确保所接触到的是正确的样本单元，怎样验证或纠正抽样框信息，如电话号码等。

3. 问卷说明。这部分向调查员提供问卷调查中及问卷中所用的概念和术语的定义。对每一个问题的含义和目的给出清楚的说明。最后，对调查中最可能出现的问题及疑难提供解决办法。

4. 问卷的审核与整理。通常要求调查员在访问期间或访问结束之后立即对问卷进行现场审核，其作用在于识别缺失、无效或前后不一致的问卷。这些问卷表明有关数据存在潜在错误。这项工作需遵循一定的规则，务必给调查员阐明审核规则。

5. 单个样本单元的管理，这是处理对无回答的追踪回访，明确调查员需要回访的次数及其规则。此部分还包括如何对每个样本单元的最终回答进行编码。

6. 作业管理。这部分涵盖一些管理细节，如工作人员怎样报告调查过程，怎样分发和回收问卷及调查所需的物资设备，调查员如何报销他们调查中的各项费用等。

7. 问题与解决方案“这是调查员手册的最后一部分，列出了调查员将会遇到的一些常见问题及其正确解决办法。一般访问技巧和技术也可以包括在调查员手册中，特殊或复杂的调查还要给出具体的例子。

### （二）督导手册

督导必须非常熟悉调查员手册的内容。此外，还必须提供一份督导员专用的督导手册，为调查的管理提供指导。督导手册通常包括下列内容：

1. 招聘和培训调查员。
2. 向调查员分配任务。
3. 根据预定的质量目标、经费开支和时间要求监控调查过程。
4. 后勤服务。包括分发和回收调查物资，发放调查员报酬，回收并提交问卷以便数据录入等。
5. 被调查者的安全和隐私保护。
6. 特殊情况下替代调查员进行数据收集。比如在访问调查中遇到一些特殊情况，像聋哑残疾人、只会说方言的人如何沟通；或者外籍人员或游客是否应包括在目标总体中进行调查等情况下有效处理。

## 二、调查员的招聘和培训

任何一次调查都需要有人去执行，市场调查活动的执行者主要是调查员。因而招聘和培训调查员具有

重要作用。这里主要介绍调查员的招聘和培训

### (一) 调查员的招聘

对一个调查机构来说，除了具有自己的工作人员外，掌握一份训练有素的调查员名单非常有益，名单上记载着在以往的调查中录用过的调查员或其他人员。这样一旦需要人员就可以按图索骥。如果需要的调查员数量很大，就应该招聘一些新的调查员。招聘新的调查员的方法很多，可以发通告，可以在地方报纸上做招聘广告，也可以使用一些其他手段来吸收合适人员（比如聘请教师、学生或有关人员）。

在任何情况下，应当明确聘用调查员的标准。当招聘调查员时，教育程度，社交技能，当地语言的流利程度、组织能力以及人品都是需要考虑的重要因素。对于面访式调查，调查员的地理知识和定位能力也是很重要的。通常还要注意以下的素质要求：

1. 诚实与认真。这是调查员必须具备的基本品质，诚实与认真具体表现在两个方面，一是准确地遵守工作细则，按规则操作；二是要忠于访问的事实，客观地反映被调查者的意见。在调查中切忌歪曲调查结果。例如没有经过访问而私自填答问卷，把主观意志强加于人。此外，对被调查者要一视同仁，不能有亲有疏，必须认真进行访问记录，不能敷衍了事。

2. 责任心要强。由于被调查者差别大，调查中有的比较顺利，有的不太顺利。例如，有些人经常不在家，要求调查员做到耐心地多次访问；有些人不愿意合作，要求调查员查明原因，有针对性地再做说明、解释工作。如果调查员没有责任心，遇到困难就会打退堂鼓，完不成调查任务。

3. 要有调查的兴趣和信心。如果没有调查的兴趣和信心，经过几次调查访问以后，就会感到调查工作枯燥乏味，也就不会钻研调查内容和访问技巧。如果对调查内容不熟悉，被调查者提出问题后无法解释，如果不掌握一些基本的访问技巧，也难以使访问达到预期的效果。

4. 对调查对象有一定的了解。例如在进行居民入户调查中，可以选择一些街道居委会的成员作为调查员，他们对被调查者的情况比较了解，又容易消除对方的疑虑，从而获得对方的支持与合作。此外最好选择与被调查者背景相近（如职业、社会地位、地区、民族等）的人作为调查员，这样易融洽双方的关系。

5. 要有健康的体魄。访问调查是一项比较艰苦的工作，为了及时完成调查任务，难免要牺牲个人的休息时间。有时还需要到交通不便的地区进行调查，不但要跑腿，有时赶不上吃饭等。因此调查员应具备良好的身体素质以适应艰苦的工作环境，能在困境中完成调查任务。

### (二) 调查员的培训内容

调查进行之前需对调查员进行培训，使他们深刻理解调查内容，掌握访问技巧。调查员培

训工作必须仔细规划，因为它是保证正确理解调查概念和统一行动的唯一办法。大型调查（如设立督导）通常先培训督导，然后由督导培训调查员。关于培训内容因调查内容和要求不同也有所不同。一般说来应包括以下四方面的内容：

1. 进行思想教育和素质教育。要根据调查的目的和要求，针对调查员的思想情况，有的放矢地组织学习和讨论，激发调查员参与调查的积极性和主动性，增强调查员的事业心和责任感。
2. 调查设计者作简要介绍。培训中需要向调查员介绍调查目的、意义以及调查的范围，调查样本的抽选和调查对象的数量，调查的步骤及需要注意的问题，调查时间和每人的工作量，使调查员做到心中有数。在市场调查中，有时会涉及专业知识的介绍，如果是有关新产品 的调查，调查员应当亲自观看新产品的实物和实际的使用过程。
3. 仔细阅读问卷。调查员要仔细阅读调查问卷，或由督导向调查员介绍问卷的结构和调查项目，明确每个项目的内容，回答类别及记录方式。明确每个概念的具体含义。调查员要吃透问卷才能保证调查的质量。
4. 其他知识的准备。包括讲授与调查内容有关的政策法规，向调查员介绍调查对象的情况以及出现问题后应当采用的处理方法。

### （三）调查员的培训方式

1. 自学。要求调查员认真学习“调查员手册”并完成书面作业。
2. 课堂培训。在课堂上，督导和调查员学习如何与被调查者建立良好关系，以促使被调查者如实回答。一般督导比较有调查经验，可以由督导向调查员讲述调查中的技巧。课堂上还可以安排模拟面访，传授面访的技巧和技术（特别是有许多新的调查员时）。督导可对调查员的作业进行讲评，调查员对整个问卷反复学习。对问卷中的特殊之处和容易出现问题的地方要格外注意。培训过程中督导应该留出充分的时间，允许调查员提出问题并给予正确的解答。
3. 模拟访问。每个调查员都做模拟访问练习，由督导或另一位调查员来充当被调查者，在模拟的过程中既要有典型的情景，又要有关困难的情景。模拟调查可以使调查员对调查过程留下深刻的印象，使其对调查过程有切身的体会。在此过程中，还可要求旁观者（其他调查员）就模拟情况发表评论，并与其他调查员探讨调查的技巧与技术。
4. 实习访问。如果条件允许，调查员可以对调查样本之外的其他对象进行实习调查访问。这种做法可以使调查员感受到实际调查中的经历，使他们取得调查经验，并在出现意外情况时能进行正确的调整培训结束时，督导或调查的其他组织者要给每位调查员布置任务，如分配问卷和样本单元的控制表，对调查员提出要求，如建立及时汇报的制度等。

## 三、编制名录和寻踪

### （一）编制名录

编制名录的目的是为了对位于一个具体地理区域内的所有单元建立一个名录。因为当进行多级抽样时，如果第一级抽选的初级单元（PSU）是区域框，如地理小区，在第二级要抽选被抽中的地理小区中的住户，为了抽取住户就必须编制所抽中的地理小区内所有住户的名录清单，编制名录就是创建这个清单的过程。在我国有比较严格的户籍制度，因此名录清单也可以从街道办事处或小区居委会取得。如果第二级抽样单元是商店、学校等单位，也可以从有关部门获得。符合条件（或称合格）单元的定义和需要的信息取决于调查的目的与要求。

### （二）寻踪

实际调查中很多情况下都需对调查的单元进行寻踪，即在现场操作时仅根据抽样框的信息还不能对样本单元定位，因此需要进一步寻找。对于邮寄调查，如果地址不详或地址错误、公司停业或被调查者搬迁等情况，问卷就会被邮局退还，尤其当抽样框已经老化，在正式调查之前更应对样本单元进行寻踪。

寻踪的目的包括：对样本单元进行定位、更新抽样框中单元的基本识别信息（例如，对名录框中住户的姓名、住址、电话号码，联系人姓名等信息）、确认样本单元是否仍在调查范围内（例如，公司是否已经停业？被调查者是否已经迁移？）。

有些寻踪可以由调查员来做，更有效的办法是在调查办公室内组建一支有经验的熟悉各种文件和资料的专职寻踪队伍。寻踪一旦成功，抽样框的信息就可以得到更新，如果寻踪是在数据收集开始后进行的，那么后来寻找到的样本单元可以由最初决定的调查员进行访问，也可以分配给专门负责“清理”的调查员。

寻踪的工作量取决于原始抽样框中联系信息的质量。例如，在对近期获得学位的大学毕业生进行的调查中，抽样框来自大学，只有学生的最后一次已知的通讯地址和电话号码。因为刚毕业的学生流动性较大，调查时许多人的联系方式发生改变了。如果联系的资料中也包括了他们父母的名字和地址，调查员就能通过他们的父母找到学生；或者通过学生建立的QQ群、微信群更新联系方式。

## 第三节

# 数据收集过程

数据收集过程由调查员主导，因而首先必须规定调查员的任务，培训其如何进行有效的访问和正确提问，掌握追问技巧，客观真实地记录被调查者的回答，面临突发情况时能应急处理以及愉快地结束访问。

### 一、规定调查员的任务

#### （一）调查员的职责

对调查员来说，在访问一开始就与被调查者建立良好的关系非常重要。这样可以赢得被调查者的充分合作。调查员应该在调查中表现出敬业的、和蔼可亲的态度，因为这有助于建立良好的关系，促使被调查者提供完整和准确的回答。调查员的职责如下：

1. 确保信息是从正确的样本单元处收集的。
2. 确保样本涵盖的完整性，使无回答的偏差减到最小。
3. 准确、完整地收集信息。
4. 尽量减少记录和数据录入的错误。
5. 遵守安全程序、保证数据的机密性。

#### （二）调查前和被调查者接触

无论是电话调查还是面访，调查员都必须做好充分准备。当调查员计划好工作日程、安排了打电话或进行面访的时间并对调查表格和材料充分了解之后，他们工作的效率就会更高。

下面是对调查员安排工作的一些指南。

1. 安排时间。调查员对他一天中所要打的电话个数或进行面访的次数了如指掌，细心 安排并在连续两次调查之间留出足够的时间，补全调查中未完成的记录以及完成问卷所必需 的审核工作。

2. 调查情况记载 c 调查员应该在问卷的空白处记载有关调查情况的说明性文字。这些 说明性文字的内容包括：给被调查者打电话的最佳时间，答应接受回访的访问形式及时间安 排等。

3. 预约。调查员应该准备一本日历或一个笔记本，记录约会的安排。

### （三）安排面访注意事项

安排面访时，调查员应该牢记下列事项：

1. 如果是面访，调查员要掌握好时间，避免过早或过晚。因为许多人在天黑之后就不愿 再开门，或者对晚上敲门非常反感。调查员可以预先将致被调查者的信留置其邮箱中，同时附 上一张个人便条，告知被调查者下次返回的时间。或留下电话号码，以便被调查者打电话确认 或重新安排时间。如果碰巧打扰了被调查者进餐，调查员应该诚恳地向被调查者道歉以取得 其积极配合。

2. 首次面访时，经常会遇到被调查者工作繁忙的情况，因此调查员需要与其商量另外安 排一个方便的时间，或者留一个电话号码，以便被调查者有空时给调查员打电话，约定调查时 间。调查员提供的电话号码应该随时有人接听（或标明有人接听的时段）。

3. 当调查员首次面访时，如果预先选定的被调查者恰好不在，调查员也应该和接待自己 的人建立友好的关系，并且定下与被调查者联系的最佳时间。在无法联系到被调查者时，调查 员应该尝试在其他时间再次进行访问。对于面访，如果调查员两次访问仍不能找到被调查者， 可以向其邻居打听，以确定找到被调查者的最佳时间。行业规范一般要求，调查员对大多数面 访调查至少进行 3 次尝试，而对电话调查则可能需要多达 10 次的尝试。

### （四）调查员自我介绍

调查员和被调查者建立良好关系的基础是恰当的自我介绍。恰当的自我介绍对于塑造出 一个既专业又友好的形象是非常重要的。自我介绍应该简洁（特别是电话调查）而诚恳，其内 容包括：

1. 调查员的姓名和调查机构名称。
2. 调查名称和目的。
3. 数据的用途（表明调查的重要性）。
4. 数据收集的权威性。
5. 对数据安全及保密性的保证。

对于面访，分发调查的介绍材料或调查手册是介绍调查的一个较好的办法。这样在进行 访问的时候，被调查者会记起他们曾经收到过介绍调查的一些资料，从而不会对调查的合法性 产生疑问。调查员必须随 身携带身份证件，并在每次访问时都主动向被调查者出示证件。

### （五）消除被调查者顾虑

从调查一开始就要与被调查者建立良好的关系，这对访问的成功具有决定性作用。访问 过程中，调查 员必须倾听，并且随时准备回答被调查者提出的问题（调查员手册应该包括最常 见问题的答案）。如果被 调查者不愿意继续进行调查，调查员应该努力消除被调查者的顾虑。 被调查者的顾虑及应对方法主要有以 下三种。

1. 你为什么挑选我？调查员应该对随机抽样的方法进行说明，让被调查者明白他是被随机抽取的，他实际上代表了总体中的个体，所以他的回答对于整个调查非常重要。
2. 谁将看到我的资料？你们将怎样使用我的回答？这里的主要顾虑是机密性。调查员应该告诉被调查者，他的回答只用于和其他人回答的汇总，根据汇总结果编制统计表格或计算出汇总结果。政策制定者和决策者根据调查结果，考虑调查反映出来的情况是否令人满意，并决定是否需要采取相应措施。
3. 我现在没有时间。调查员必须如实告诉被调查者访问所需要的时间。调查员与被调查者接触时，应该争取立即开始调查，不要因被调查者声称非常忙而有所迟疑。如果被调查者真的不能立即接受访问，也应该建议另选时间，对重新访问做一个确定的安排。有时，被调查者声称没有时间是由于其实际上没能清楚地理解调查的重要性和数据的用途，调查员必须在介绍时对此进行清楚的说明。

## 二、有效访问的关键和提问方式

### (一) 有效访问的关键

1. 信心。调查员必须相信自己的能力，只有这样才能心中有数，对自己充满信心。调查员必须理解自己在调查中所起的作用。
2. 倾听技巧。调查员应当等被调查者把话讲完。调查员可以通过适时插入类似“对，是这样”等简短的话语，以表示自己正在倾听。然而调查员不应该猜测被调查者将要说的话，也不应该随便打断被调查者的话。如果调查员认为被调查者没有把握好某个问题，最好把这个问题重复一遍。
3. 投入。进行电话调查或面访时，调查员应该时刻注意被调查者的表情。如果被调查者谈到与调查无关的私事时，应该表示对此很感兴趣，但是不要妄加评论。然后自然而然地把被调查者引回主题。
4. 语言的表达。口头表达非常重要，在电话调查中尤其如此。调查员应该吐字清楚，并保持中等的语速。因为如果调查员说得太快，被调查者就可能会漏听问题的某些部分；相反，如果调查员说话太慢，则会使被调查者没有听完问题就急于回答，因而也会出现偏差。访问过程中，调查员务必要降低语调，这在电话调查中应该更加引起注意，低调的声音有助于言语的表达，并可以使话语说得更清晰。
5. 洞悉问卷。调查员必须熟悉与调查相关的材料，特别是问卷。因为访问期间，调查员根本没有时间去查看手册中对问题的定义或答案，没有什么能比长时间地停顿更能损害调查员和被调查者之间业已建立的良好关系了，这在电话调查中更是如此。

### (二) 提问方式

由于调查机构需要对所有的调查数据进行汇总，因此数据的收集方式必须保持一致。这意味着必须以同样的方式向所有被调查者提供同样的问题。当使用问卷来收集一致性的数据时需遵循以下准则：

1. 严格按问卷的措辞提问。研究表明，如果调查员没有完全准确地按问卷措辞去提问，那么很细微的措辞变化都可能会使被调查者的回答发生改变，可能会歪曲设计的原意而造成偏差。因此在实施调查中，调查员必须严格按照问卷的要求来提问，绝对不能按自己的理解来修改问卷中问题的提法。
2. 按问题给定的顺序依次提问。问卷中问题的排列是经过通盘考虑的，而且这种排列也保证了前面的问题不会对后面问题的回答产生不利的影响。
3. 每个问题都应该被问及。有时被调查者在回答一个问题的时候，同时也回答了后面要提问的另外的问题。这时，调查员在后面的问题出现时，仍然需要提问。在这种情况下，调查员可以说，“您已经告诉了我关于这个问题的一些东西，但是……”这表示调查员知道前面的回答，但现在要求被调查者再次回答这个问题。
4. 按正面方式提问。有时调查员可能会对某些类型问题的提问方式不太适应，因此常常改用委婉的方式，例如，“你可能不想回答这个问题，但是……”“这个问题可能对你没有什么意义……”这样提问

开头往往会对被调查者的回答产生负面影响，并且有引导被调查者回答的倾向。相反，如果使用正面或中性的语气提问，答案就不会受到干扰了。

5. 对问题之间的停顿加以解释。这在电话调查中显得更为重要。相比之下，某些答案的记录可能要花费更长的时间。这时调查员可以说，“请原谅我稍停一下，我正在记录您给我的信息从而得到被调查者的理解”

6. 重复被误解或曲解的问题：问卷中问题的措辞是按照普通的被调查者的理解能力来设计的，因此大多数被调查者都能理解。然而，偶尔某个被调查者也可能会误解或曲解其中的一些问题。当这种情况发生时，调查员应该把问题按原文再重复一遍。如果答案仍然不符合问题的意图c，调查员可能不得不进行追问，追问的技巧随后再做说明。

7. 特别关注跳转模式。“过滤”问题（有时也称为“跳转”问题）决定随后的问题是否适用，而且还决定了问卷的流程。因此调查员必须对跳转模式给予特别注意，并且详细阅读问卷上的调查员指南。计算机辅助调查可以通过程序解决跳转问题，但是调查员仍然应该熟悉问卷的流程。

8. 对被调查者提供信息持中立态度。对于被调查者来说，如果他们觉得访问像是审问，那么他们可能会将自己的真实想法隐瞒起来，而向调查员提供一般被社会认可的答案。因此，调查员应该尽量使被调查者明确，他们提供的答案没有正确与否或好坏之分。对于被调查者的回答，调查员都应该持中立态度，而不应有不满、吃惊、赞成或不赞成、同意或不同意等的言语或表情。如果被调查者的回答在答案的可接受域内（符合逻辑），调查员就必须承认被调查者的回答。

### 三、追问技巧

追问是当调查员认识到答案没能满足问题的意图时使用的一种技巧。被调查者可能不知道该如何回答，或者可能误解或曲解了某个问题，从而导致对该问题的回答不完整、不清楚或与其他问题的答案不一致。当这种情况发生时，调查员必须持中立态度去探求所需要的信息。问卷中也可能会有一些开放式的问题需要被调查者深入地回答，调查员要采取有效的追问技巧，使被调查者能够进一步扩展、阐明他们自己的回答。追问技巧的关键是既深入、客观又不至于诱导产生偏差。常用的追问方法有以下五种：

1. 重复提问，在追门前，调查员应先按照已有的措辞重复这个问题，以防被调查者只是因为没有听清楚这个问题而导致回答不合题意：当需要解释这个问题时，必须使用手册提供的定义。

2. 重复被调查者的回答或用中性的语言追问。如果重复提问后，回答还不合题意，那么可以使用中性的追问以获取更多的信息，例如：

(I) “我不能确切领会你说……的意思”

- (2) “你能多告诉我一点吗？”
- (3) “还有吗？”
- (4) “其他理由呢？”
- (5) “对此您还能谈谈您的想法吗？”

3. 启发性地帮助被调查者回忆。调查员也可以通过围绕合乎题意的回答来帮助被调查者回忆。如：

- (1) “您认为最接近哪个数？”
- (2) “比…大还是比…小？”（对数字型的回答）
- (3) “是春天、夏天、秋天、还是冬天？”

问题也可以改述，但必须非常小心谨慎。不能把问题改述成答案的提示。例如，如果问题是“去年你工作了多少星期？”，调查员就不应该将它改述为“你全年都工作了吗？”：但可以这样问：“你去年工作了吗？”若回答“是”，接着再问“工作了多少星期？ ”

4. 利用适当的停顿或沉默。通过停顿、沉默或注视，都可以暗示被调查者提供一个更完全的答案。不过，一定要适当，不要让沉默或注视使被调查者感到难堪。

5. 鼓励被调查者放心回答。如果被调查者在回答中有犹豫，调查员应让他们放心。例如可以说，回答没有什么对与错，我们只是想要知道您的看法。

#### 四、 调查中的记录

记录被调查者的回答看起来很简单，但调查员常常容易在这方面犯错误，应该培训调查员在调查过程中如何进行记录，记录应该使用相同的格式、符号和修改方法。对于结构式的问题，主要是要求调查员注意划圈或打钩地方不要搞错，只要注意力集中，一般不会出错。对于开放式问题，要逐字逐句地按被调查者的原话进行记录。以下是美国密歇根大学社会研究调查中心制订的调查员针对开放式问题进行记录的指南：

- 1. 在访问期间随时记录（不要过后补记）；
- 2. 使用被调查者自己的语言（记录原话）；
- 3. 不要对被调查者的回答进行归纳总结或解释；
- 4. 记录与问答题有关的全部内容；
- 5. 记录所有的追问语和对应的回答；
- 6. 边记录边重复所记录的答案。

#### 五、 访问中相关问题的处理

调查员在进行访问的过程中经常会遇到一些棘手问题，需要认真仔细地处理。这些问题的处理同样有一些基本规则，分述如下：

##### （一）拒访的处理方式

拒访是数据收集过程中经常遇到的情况，下列方法可以帮助调查员适当地处理拒访。

1. 被调查者拒访。通常是因为他们对调查或调查机构不太了解，或者是调查时间选择不当。在这种情况下，调查员应该见机行事，对时间的安排不当，应向被调查者表示诚恳的歉意（而不能断定是被调查者拒绝访问）并建议再约回访的时间。

2. 被调查者不情愿接受访问。在此情形下，调查员可以建议先试几个问题，而访问一旦开始，被调查者就可能会认识到调查对他没有什么妨害。这样调查员就可以继续调查后面的问题。出于礼貌，调查员要让被调查者知道可以不回答某些问题。

3. 当被调查者断然拒绝回答时，调查员应该礼貌地退出并报告给督导，以便督导追踪回访。调查员如果坚持访问，可能会损害督导赢得被调查者合作的机会。

## （二）代理访问的处理

代理访问，又称代理回答，是指由于被调查者不在，有关信息可以从另一位有文化的家庭成员那里获得的过程。代理回答仅适用于所有家庭成员都知道的信息调查，这种方法不适用于收集个人数据或个人意向的调查。在调查培训中就要让调查员知道该调查是否允许代理访问，什么类型的被调查者可以接受代理访问，或是否所有的访问都不能通过代理进行。若督导不做特殊说明，调查员应该认为该调查是不能代理的。在不能代理的情况下，若遇到语言障碍，调查员仍然可以请其他家庭成员充当翻译。

非代理访问一般回答率比较低，而且比代理访问需要花费更多的精力。非代理访问中，调查员应该牢记需要预约和再访的情况，以提高非代理访问的回答率。同时调查员还应该仔细地安排，以免给被调查者造成负担，并控制调查费用。

## （三）其他棘手情况的处理

1. 调查涉及机密信息的处理。当调查涉及机密信息时，调查员不应该要求被调查者当着其他人回答。如果被调查者有客人在场，调查员应该与被调查者另约时间再访问，或者另选私人场所进行访问。寄宿者、租宿者和未成年孩子，可能会要求避开其他家庭成员单独接受访问。

2. 与被调查者语言沟通不畅的处理。当与被调查者语言不通时，调查员可以将被调查者移交给与之语言相通的调查员或者被调查者的其他家庭成员做翻译。

3. 绝不允许调查员直接或含蓄地与被调查者发生争执，或者威胁被调查者。完整可靠的信息需要被调查者自愿合作。因此访问过程中调查员不应该与被调查者讨论敏感的政治话题或其他容易引起争执的问题。

4. 访问期间如果被调查者正好处于患有重病或丧亲等困境，调查员必须根据情况决定是否继续进行访问，或者与被调查者协商，约定一个方便的回访时间。当然，没有继续访问或回访的可能性，就应终止访问。

5. 采取合适的方式进入高级住宅区或公寓区。与生活在高级住宅区或公寓里的被调查者联系，有时是一个难题，因为这些地方时刻有保安守卫，外人出入都需出示证件并得到被调查者的允许，楼内住户的电话一般不容易得到。遇到这种情况时，调查员可以尝试与本地居委会负责人或被调查者的所在单位联系，解释访问的目的并允许进入调查。

6. 调查员不得使用不正当的方法进行访问。例如，访问时若被调查者恰好不在家，这时调查员可以向他的邻居询问，但所问的问题不能超出如何和什么时候能与被调查者取得联系等内容。

7. 调查员要保护自己的人身权利。如果受到辱骂、口头恐吓、身体骚扰或暴力等威胁，调查员应该立即撤退并向督导和有关领导报告。

I

## 六、结束访问

访问过程的最后一步是保证所有必需的信息都已收集完整，并且已经做了清楚的记录。每次访问结束时，调查员都要全面检查问卷并做必要的审核。当然，由于时间的限制，调查员可能无法当着被调查者的面来做这些事。因此访问结束时，礼貌地对被调查者提供信息表示

感谢非常重要，这也可以方便以后的回访。同时，调查员也应该尽量回答被调查者仍然可能存在的关于调查的疑问。重要的是让被调查者感到时间没有白花，通过这次调查，他们做出了重要的有价值的贡献。有的市场调查公司采取向被调查者赠送一些小礼品以表示对被调查者合作的诚挚谢意。

如果访问是通过电话进行的，上述访问技巧和注意事项也仍然适用。但是由于调查员与被调查者分别在电话线的两端，调查员只能完全依靠自己的声音来控制访问的过程，因此在某种意义上对调查员的要求更高。调查员的声音要亲切有礼貌，而且吐字必须清楚，语速要适当。

## 第四节

### 调查过程中的质量监控

数据收集的目标是尽可能完整和准确地获取数据。由于在整个调查活动期间每一方面都可能发生误差，所以必须小心地计划数据收集以使误差减到最小。调查过程中的质量监控包括遵循数据收集的指导方针、数据收集过程的质量监控、调查员的作业管理和调查执行完成情况的评估等多个方面”下面分别介绍。

#### 一、数据收集的指导方针

为了保证较高的数据质量，应该遵守以下指导方针：

1. 要确保聘用具备调查能力和良好个人素质的人当调查员。并要对他们进行适当的培训，配备访问所需要的工具。
2. 应该对所有的数据收集规定适当的样本控制程序。这种程序应该从数据收集开始，到完成收集数据和数据录入，自始至终地对样本单元的状态进行跟踪。
3. 调查项目负责人应该建立有效的控制系统以保证数据传输和处理的安全，以维持被调查者的信任和统计安全性。
4. 调查项目主管应该在组织机构中设立合适的联系人，以便被调查者在适当的时间询问有关信息，从而提高回答率和改善来自经营机构及公共机构的信息质量。如果可能，应该允许被调查者使用他们便于操作的方法和方式提供数据。
5. 所有的调查员都应该接受相同的培训，使用相同的手册，所有的编码人员也都应该使用相同的编码规则等，从而保证对所有被抽中的样本单元都使用一致的数据收集程序。
6. 数据收集过程中应该贯彻一套管理信息系统来跟踪质量的度量、现场费用的开支和其他各种执行方法。现场费用包括邮资、电话费、差旅费和员工的日常消费等。质量度量包括回答率、追踪回访率和各种原因导致的无回答的数量等。从数据收集工作一开始，就应该对他们进行测量和监控。具体操作时，操作人员可以把质量和产出的度量标准结合起来使用，以决定数据收集在达到什么程度时就应予以结束（例如，当回答率达到了预定的目标或者进一步改进回答率需要花费太多）当然，这些方法也适用于其他调查和数据质量的评估。

7. 调查项目主管应及时追踪调查数据的变化。这时审核检出错误的频率和数据修正的数量及类型可以提供关于数据质量的有用信息；同时，它们也可以为改进今后调查中的数据收

集过程或数据收集方法提供重要的参考。

## 二、数据收集过程的质量监控

数据收集过程中，调查员的工作好坏直接影响到数据的质量，但这又往往是比较薄弱的环节。在大多数调查中，设计人员和数据处理人员都只有少数几个人，而调查员人数比较多，在一些大型的调查中，还可能临时聘用大量业余的调查员，对于这个人数众多的群体，调查技术水平参差不齐，要做到全都全心全意地投入到调查中去是不现实的。同时，由于调查活动大都是在远离管理人员的现场单独工作，监督工作很难展开。因此，有些调查员就会不负责任，敷衍了事。即使有些调查员主观上愿意做好调查工作，但客观上的疲劳、能力上的缺陷，或者调查中艰苦的环境，都会给调查的质量造成影响。因此对数据收集质量进行监控非常必要。数据质量的监控方法如下：

### 1. 建立督导机制，严格管理调查员

调查中许多因素可能诱使调查员作弊。一种比较恶劣的作弊是调查员根本不到现场，对所有问卷都根据自己的理解来填写，或者请自己的朋友代为填写。此外还有一些比较隐蔽的作弊方式，如当被调查者不在时更换调查对象；有时为了减少工作量，在询问时可能跳过一些问题，这些问题往往是问卷中难度比较大的问题，受访者不愿意回答的问题。还有当调查内容与调查员有利害关系时，调查员可能会有意或无意地采取诱导的方式促使被调查者按自己的意愿来回答。因此建立督导机制，在调查过程中对调查员的严格管理十分必要。督导应按一定的比例采取公开或隐蔽的方法监督调查员的工作，以保证调查员正确使用问卷和确保访问技术的有效性。

### 2. 检查（或现场检查）已完成的问卷

由督导回收并审核问卷，这是督导的核心工作内容之一，由于调查员人数较多，调查后的问卷需要进行审核，工作量非常大。一般方法就是几个调查员编为一个小组，安排一名督导负责，该组调查员的问卷都由督导回收后逐份审核，发现问题及时责令调查员返工。这可以确保调查员使用正确的方法进行问卷的调查，而且一旦发现缺失数据，调查员能够及时地从被调查者处获得遗漏的信息。督导的审核工作也能为以后计算机审核阶段可能发生的错误提供预警。

### 3. 严格的文档控制

必须对每个样本单元的最终状态进行编码。有关人员将会对每份问卷进行编码。编码时，问卷在现场操作中的每个阶段都必须有所说明，如“调查员提交问卷的日期”“寻踪的日期”“录入的日期”等。督导的样本控制表格（与具体作业控制表连接）是获得有效样本和进行文档控制的关键。如果可能的话，当调查员书面报告他们的访问进程时，推荐使用自动样本控制方法，以简化督导的工作。

### 4. 对调查员询问

当数据收集结束时，调查负责人应该向调查员询问访问的情况，这可以揭示问卷中存在的许多问题。这些问题在问卷设计阶段可能并没有被注意，但却是许多调查员在调查过程中切身体会到的。将这些信息反映到数据处理和数据审核过程中，对整个调查来说非常重要。这些信息也有助于对调查手册的修订和为重复性调查的再设计提供参考。

### 三、调查员的作业管理

在整个数据收集过程中，调查员要将全部作业的状况写在作业控制表上，定期(通常每周一次)向督导报告。那些还有待于联系或尚处在联系过程中的样本单元，要标上“未解决”的记号(还要进一步细分为“还没有尝试”和“已经尝试”两种情况)。对于那些已经尽了很大的努力去处理的样本单元，要标上“已解决”的字样，并注明已解决的程度，具体包括以下3种情况：

1. 全部完成。表明整个调查过程已经全部结束。
2. 部分完成。指被调查者没有完成整份问卷，但是回答了关键问题。关键问题是事先确定的，是形成一份有效问卷所必需的问题。
3. 没有完成。应该列出原因，其原因包括：
  - (1) 拒访(指需要督导回访的情况)
  - (2) 被调查者在调查期间不在家；
  - (3) 无法寻找(相对于寻踪而言)；
  - (4) 被调查者不在样本范围内，如死亡、停业、房屋拆迁；
  - (5) 被调查者暂时不在样本范围内，例如住宅空置；
  - (6) 没有接触(无人在家)。

### 四、调查执行完成情况评估

执行任务和质量目标是在现场调查开始时就确定的，督导应该及时监控和管理调查员的工作，在调查执行过程中计算解决率、回答率等指标，以保证达到既定的目标。

首先，执行任务完成情况可以用解决率来反映，即

$$\text{解决率} = \frac{\text{已解决的样本单元数}}{\text{样本单元总数}} \times 100\%$$

这个比率是衡量调查员计划完成情况的一个指标。将每周完成的比率与目标的比率相比较，并以此进行督促，保证调查按时完成。同时还需要把费用与这一指标结合起来进行评估，以确定调查费用是否已经超出预算范围。执行费用的评估指标也包括两项，即“每个已解决单元的平均费用”和“每个未解决单元的余下预算气

其次，要计算无回答率。无回答是非抽样误差(以偏差的形式出现)和抽样误差(导致精度的下降)的一个重要来源。因此质量目标主要根据回答率来确定。回答率的计算公式如下：

$$\text{回答率} = \frac{\text{被单数数量}}{\text{在范围内已解决单元数+未解决单元数}} \times 100\%$$

例如，若抽取了1 000个单元，开始收集数据一个星期后有800个得到解决，它包括完成、拒访和超出范围的单元。其中有700个在规定的调查范围之内，而在这范围内，又有550个作了回答(完全回答或部分回答)。则调查第一周的回答率是：

$$\text{第一周的回答率} = \frac{550}{700+200} \times 100\% = 61.1\%$$

然而考虑到未解决单元一旦得到解决，可能有一部分超出范围，因而需对未解决单元用一个调整因子(700/800)来修正，在上述例子中，200个未解决单元中大约有175(200×7/8)个在

调查范围内’因此回答率应是：

$$\frac{550}{700+175} \times 100\% = 62.9\%$$

督导也应该对其他能预示质量问题的指标进行监控。个别调查员作业中反映的高拒访率或无法接触率预示着问题可能出在调查员。但如果出现整个样本的高拒访率，则显示了公众的抗拒心理，这种情况的出现使得在既定的时间和经费下，质量目标很难实现。若被调查者中“超范围”的数量高于期望的数量，说明抽样框存在问题。如果某个代码组（特别是“空缺”组）在某些作业中出现的频率较高，则可能预示着调查员对单元分类不当（例如，住宅是否真正空置，或居民是否只是暂时不在）

### 本章小结…

市场调查的组织与实施是一个完整的过程，需要在这一过程中做好的事项包括：与被调查者建立密切融洽的关系；熟悉并使用调查问卷；处理尴尬和敏感情况；控制数据收集过程中的各类文件；对数据的进行预审核；对预审核时发现的错误和无回答进行追踪回访；控制好调查质量。在组织队伍对数据进行采集过程中，调查员是数据采集的关键，其素质和条件与调查质量的高低息息相关，进而影响到调查项目能否获得成功。调查员应该具备良好的职业道德水准、沟通能力和宽广的知识视野，依据上述标准选拔出来的调查员仍需针对所开展调查项目的特定情形进行系统和专业化的培训。在调查实施过程中，调查员要掌握与被调查者接触、提问、追问及记录答案等访问技巧。在调查实施过程中贯彻提高数据收集质量的指导方针。

### 关键名词•••

公关宣传 编制名录 寻踪 模拟访问 拒访 追问 代理回答

人员培训 质量控制

直思考题..... 盐  
回奕黄 即测即评

1. 调查实施中的公关工作内容有哪些？
2. 调查前的准备工作内容有哪些？
3. 调查员应具备哪些基本素质？
4. 如何对调查员进行培训？
5. 追问的技巧有哪些？
6. 如何对调查员进行管理？
7. 如何有效防止调查员作弊现象的产生？
8. 如何对调查过程中数据收集质量进行监控？

### 本章实训. …

#### 一、实训目的

针对自选调查项目进行具体实施或参加一次市场调查公司或者某个项目的调查员应聘，并实际从事调查访问

1. 通过具体实施项目，掌握项目实施中如何对调查员进行培训。

2. 应聘调查员，了解项目实施中对调查员进行培训的情况，结合所学知识分析其中的长处与不足。

## 二、实训内容

完成一份书面报告，具体要求：

如果是实施一个调查项目：

\*

1. 准确说明项目实施的过程安排。
2. 清晰描述招聘人员的场景和接受培训的情况。
3. 描述自身对调查访问的感受和体会。

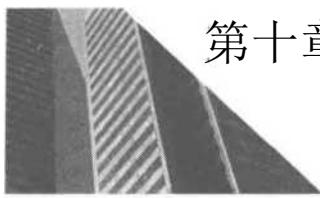
如果是应聘调查员：

1. 准确说明应聘的时间、地点等基本信息。
2. 清晰描述应聘时的场景和接受培训的情况。
3. 对该公司或项目的调查员选择和培训进行评价，并提出改进意见。
4. 描述自身对调查访问的感受和体会。

## 三、实训组织

1. 以项目组为单位组织实施。
2. 以学生个体为单位完成相应活动。
3. 教师对交流情况进行总结和点评。

# 第十章 调查数据处理与分析设计



## 导语

、”■ ■，..•.

数据处理就是把调查中收集到的数据转换为适合汇总制表和数据分析的形式。它包括在数据收集完毕之后、进行估计之前对数据进行加工处理的所有活动。处理的整个过程既有自动完成的，也有手工完成的。整个过程费时、费力，但对数据的最终质量和成本却有很大的影响。因此对处理工作进行周密的计划、实施质量监控，并在必要时进行调整就显得不可或缺。经过预处理后的“干净”数据进行的统计分析还需要在原有调查设计方案基础上进一步修改与完善。

本章将介绍数据处理的主要工作：即问卷接收与检查，编码、录入与数据净化，统计预处理与数据库的创建，在“干净”数据基础上编制统计分析任务书。

## 教学目标与要求

1. 了解市场调查数据整理的全过程及其主要内容；
2. 掌握问卷编码的规则与方法；
3. 熟悉数据录入的环节及差错处理方法；
4. 熟悉数据净化的含义；
5. 掌握缺失数据插补与离群值处理的方法；
6. 熟悉如何创建数据库；
7. 熟悉调查数据的统计预处理流程；
8. 掌握编制数据分析任务书的基本原则。

## 、本章知识结构

### 调查数据处理 与分析设计



# 第一节

## 问卷的接收和检查

问卷的接收和检查自市场调查项目实施开始，到整个调查活动完成时才结束。全过程需要认真细致，规范操作。

### 一、问卷接收

调查数据的预处理应该在研究设计阶段就制定好，但真正着手整理是从仍在实施的现场中回收的第一份问卷开始的。因此，如果一旦发现问题，还可以及时地纠正或改进实施的工作。

调查机构通常需要非常仔细地控制数据收集和整理的过程。从实施一开始，就要通过实施主管，每天一次或至少每周两至三次从每一个调查员或督导那里获取工作进度的报告。

完成的问卷就是获取调查数据的原始文件，需要设计一套工作系统来处理原始文件，并在接收资料的过程中自始至终地坚持这个工作系统。这个工作系统的第一步通常就是打开所收到的材料，记录收到的日期和交付人的姓名。问卷接收工作系统样表见表 10-U

表 10-U 问卷接收工作系统

日期	问卷编号	完成情况	接收情况	存放处或保管人
10月1日				

按问卷编号（唯一识别码）的顺序记录所有接收的问卷，问卷编号不但记录在原始文件上，也同时记录在数据库的数据列表中。因此，如有必要进行查错时，研究者可以随时找到原始的资料。I 规定并坚持一套系统的记录制度，明确原始文件的责任人，原始文件不得错放或丢失，责任人负有保管的责任。如果事后发现任何问题或错误，都能找到责任人。

### 二、问卷检查

问卷检查是指对回收问卷的完整性和访问质量的检查。目的是要确定哪些问卷可以接收，哪些问卷要作废。这些检查常常是在调查还在进行的过程中就开始了。如果调查是委托某个数据收集机构去做的，那么研究者在调查工作结束后还要进行独立的检查。其重点是：规定若干规则，使检查人员明确问卷完整到什么程度才可以接收，包括完成问题数量和每一题完成的回答情况。

对于每份看似完成了的问卷都必须彻底检查，要检查每一页和每一题及其回答，以确认调查员是按照指导语进行了访问并将答案记录在了恰当的位置上。G

一般不能接收的问卷如下：所回收的问卷明显不完整，例如，缺了一页或多页；问卷从整体上回答不完全；问卷的几个部分回答不完全；问卷只有开头的部分回答完全；问卷中的问题没有按要求跳答；问题答案几乎没有什么变化；问卷的回收是在规定的截止日期以后；问卷由不合要求的被调查者回答。

一般情况下会存在检查人员难于判断的问卷，这些问卷应该先放在一边，通知研究人员来检查以决定取舍。因此，通常最好建议检查人员将原始文件（问卷）分成三部分：可以接收的问卷；明显要作废的问卷；

是否接收还存有疑问的问卷。

如果有配额的规定或对某些子样本有具体的规定，那么应将可以接收的问卷分类并数出其数量。如果没有满足抽样的要求，就要采取相应的行动，例如在资料的校订之前对不足份额的类别再做一些补充访问

C

### 三、问卷校订

问卷的校订包括两个环节：即找到回答不满意的问卷，处理不满意的答案，

#### （一）找到回答不满意的问卷

通过检查，找到回答不满意的问卷。不满意的问卷主要有：①字迹模糊的回答问卷；②回答不完全的问卷；③回答问题前后不一致的问卷；④模棱两可或分类错误的问卷。校订人员应该用红笔将这些问卷中的回答圈出来或写出来，使之与问卷中用于记录数据的方式有明显的区别。同时对这些不满意的回答进行适当的处理

#### （二）处理不满意的答卷

处理不满意的答卷，通常有三种处理办法：

1. 退回实施现场以获取更好的数据把有不满意答卷的问卷退回实施现场，让调查员再次去接触被调查者。在市场调查中，有时候样本量比较小，而且被调查者比较容易识别不过由于访问时间和所采用方法的变化，第二次得到的数据和第一次的可能会有些差别。

2. 按缺失值处理。如果将答卷退回实施现场的做法无法实现，校订人员可能就要把不满意的答案按缺失数据来处理。在满足以下条件的前提下，这种方法是可行的：有不满意答案的问卷数量很小；每份有这种情况的答卷中，不满意答案的比例很小；有不满意答案的变量不是关键变量。

3. 整个答卷作废。当存在以下情况时问卷可作废处理。不满意答卷的比例很小（小于10%）；样本量很大；不满意的答卷和满意的答卷之间没有明显的差别（例如，人口背景资料，产品适用特征，等等）；每份不满意的答卷中，不满意答案的比例很大；关键变量的答案缺失。

由于不满意的答卷与满意的答卷之间一般都会有差异，而且将某份答卷指定为不满意的答卷也可能是主观的。按缺失值处理或将整个答卷作废，都可能会使数据产生偏差。如果研究者决定要扔掉不满意的答卷，应该向客户报告识别这些答卷的方法和作废的数量。

## 第二节

### 问卷编码和录入

在问卷接收和检查完成之后，进入问卷编码和录入阶段Q其目的是为建立方便计算机处理的可使用数据库做准备问卷编码包括事前和事后编码，数据录入必须准确无误，

#### 一、事前编码

问卷编码就是给每一个问答题的每一个可能答案分配一个代号，通常是一个数字。编码

第十章 调查数据处理与分析设计

可以在设计问卷时进行，也可以在数据收集结束以后进行，分别叫事前编码和事后编码。事前编码的问卷通常是将每个答案的对应值印在问卷上，数据文件用的记录格式常常放在最右边或放在某处的括弧内

大多数编码工作都几乎可以在调查实施的同时进行。在许多情况下，为了简化，表示列位置的括弧内

的数字是省略的。因为列位置可以在数据录入时才规定，或采用自由格式录入。表 10-2 为第三次农业普查农户普查表的事前编码表（部分），其中 H000-H015 都为事前编码。例如，“H002”这个问题（变量）的回答编码“1”（变量值）表示户主的户籍在本乡镇，“2”（变量值）表示户主的户籍不在本乡镇。

表 10-2 第三次农业普查农户普查表的事前编码表

普查区代码：口□□□□□□□□□□□□

普查区名称：

普查小区代码：□□

普查小区名称\*

户编码：□□□

#### H0 ( ) 0 您家的基本情况

H001 户主姓名：\_\_\_\_\_

H002 户主的户籍是否在本乡镇？（单选，在选择处划“V”） 1. 是 2. 否

H003 住户成员（与本户经济、生活连为一体的人）共有几人？\_\_\_\_\_人

请将您家各成员基本情况依次填写（序号 I 为户主）：

本户成员	1	2	3	4	5	6	7
H004 性别 1. 男 2. 女							
H005 年龄（周岁）							
H006 婚姻状况 1. 未婚 2. 有配偶 3. 离婚 4. 丧偶							
H007 受教育程度 1. 未上过学 2. 小学 3. 初中 4. 高中或中专 5. 大专及以上							
H008 是否在校学生？ 1. 是 2. 否							
H009 是否离开本乡镇 6 个月及以上？ 1. 是 2. 否							
H010 2016 年从事农业生产和管理时间有多少天？（请选择） ①没有（跳至 H014）②1~14 天 ③15~29 天 ④30 天及以上							
H011 从事农业的行业? R 种植业 2. 林业 3. 畜牧业 4. 渔业 5. 农林牧渔服务业 6. 无	主要口 次要口	次要口	主要口 次要口	主要口 次要口	主要口 次要口	主要口 次要口	主要口 次要口
H012 2016 年是否在本户以外从事农业 30 天及以上？ 1. 是 2. 否							

续表

本户成员	1	2	3	4	5	6	7
H013 是否受过农业专业技术培训? 1. 是 2. 否							
H014 是否从事非农行业? 从业方式是什么? 1. 没有 2. 雇主 3. 自营 4. 务工 5. 公职 6. 其他							
H015 年龄 60 周岁及以上者, 请问: 您的子女是否全部外出本乡镇 6 个月及以上? 1. 是 2. 否							

## 二、事后编码

事后编码是指给某个没有事先编码的答案分配一个代码。通常需要事后编码的有: 封闭式问答题的“其他”项; 开放式(含填空)问答题 0

封闭式问答题可能有几个供选择的答案。再加上需要被调查者具体说明的“其他”类别。由于“其他”的答案没有事先规定的代码, 因此在数据录入前编码员要完成事后编码的工作。

对于开放式的问答题, 事后编码的工作量就更大。这是因为研究人员一般无法事先告诉 编码员会出现多少新的代码和答案, 而且还有一些答案是类似的, 必须决定是将它们合并为一类, 还是分成几类。

事后编码通常可参照以下要点:

1. 提供编码员一份空白的“参照问卷丁”
2. 提供每一个需要事后编码的项目一份编码表或编码名单。
3. 对每一个项目做一份编码本, 内含一页或几张单页。
4. 让所有的编码员都在同一地点, 使用同一编码本进行工作。
5. 提供编码指南, 说明什么时候以及怎样设立一个新的代码或合并答案。
6. 设立较多较窄的类别要优于设立较少较宽的类别。
7. 保持编码册的整洁和清晰。

事前编码和事后编码所用的编码本最后将合并为一个编码本。一般来说, 编码本不但是 编码人员的工作指南; 也提供了数据集中变量的必要信息, 编码本一般包含变量的以下方面 信息:

1. 所在列的位置(列数);
2. 变量的顺序编号;
3. 变量名称及变量说明(变量及变量标识);
4. 问答题编号;
5. 编码说明(变量值及变量值标识)。

表 10-3 给出了 200 份问卷的一个编码格式。

表 10-3 问卷的编码格式

列数	变量编号	变量名称及变量说明 对应问卷中 的问题	问答题编号	编码说明 对应事先编码 选项很多时代码说明 缺失值 代码
----	------	------------------------	-------	---------------------------------

1-3	I	被调查者编号		从 001-200
4	2	调查城市编号		1 北京 2 武汉 3 广州
5-7	3	调查员编号		第 1 位代表城市, 后 2 位代表调查员编号
8-9	4	QH 第一重要	Q1	按问卷上的事先规定编码; 99-不知道
10-11	5	Q111 第一重要的其他问题	Q1	事后编码 10-Jl-,12-,13-,14-

### 三、数据录入

数据录入是指将问卷编码表的内容逐项转化成可机读的数据形式。由于数据采集方式可以是纸质问卷方式和计算机辅助收集数据方式，两者在数据录入的进程上有差别。采用纸质问卷方式收集数据时，数据录入应该在数据收集完成之后尽快进行，这样可以尽早地将情况反馈给数据收集人员，并通过反馈，不断改善数据收集过程。采用计算机辅助数据收集方法时，数据录入是在数据收集的同时完成的。只有准确的数据录入，才能在数据分析阶段提高效率。数据录入需要完成的工作内容包括录入员的培训、数据录入方式的选择和录入审核及差错处理，全部数据录入完成后，还需要进行数据净化」

#### (一) 录入员的培训

录入员培训的主要内容包括：

1. 明确任务的具体要求和注意事项；
2. 向每一个录入员提供一份清晰的记录格式说明件，用于明确每个记录包含的变量及其相对位置（在数据库中对应的列）；
3. 录入员必须熟悉使用录入设备及相关设施；
4. 说明数据或研究的最终目的，进一步重视数据录入的准确性要求；
5. 开始阶段的实战检验，开始录入阶段、研究人员必须在场，用于指导和展示准确的数据录入。
6. 录入完成一部分后需对录入的数据进行全面核查或抽样核查，发现问题及时修正“

#### (二) 数据录入方式的选择

数据录入的方式决定了录入的质量和速度。根据采集数据的方式不同，数据录入有对纸质问卷通过键盘人工录入、对纸质问卷进行光电扫描录入以及采用计算机辅助数据录入三种方式。

通过键盘录入方式是传统的数据录入方式，面向终端，利用人工一个一个敲入数据。不仅速度慢，且容易出差错。它只适合小规模调查，大规模调查的数据录入需要其他方式的支持。

光电扫描录入是指采用光电扫描设备对纸质材料进行快速扫描录入数据的方式。如果调

查机构决定采用纸质问卷数据收集方法，光电扫描可以大大提高数据录入的速度。光电扫描特别适用于封闭式选择题，即打钩式的回答（此时，一般将答案框画成圆形，在选中某项时，将相应的圆圈涂黑），对文字叙述式（开放式问题）的回答，则可靠性较低，需要辅之以手工录入，可以减少录入差错。一般扫描结束后要做检查工作，以控制录入差错。

光电扫描录入方式需要对设备使用进行培训，提供出现问题时的解决方法说明。该方法曾在中国经济普查中使用。但是由于扫描数据只能呈现纸质问卷的原始表格，要建立便于分析的结构式数据库，还需要进一步转换，因而其适用领域受到一定的影响。

计算机辅助数据收集方法的主要优点是数据收集和数据录入可以同时进行，从而比纸质问卷数据收集方法速度更快，效率也更高。其核心工作是需要专门的技术人员进行软件的开发和测试，并在测试阶段解决数据录入过程中可能发生的各种问题，它是现在通用的大规模数据调查的录入方式。2017年完成的中国第三次农业普查就采用了pda软件、调查员利用统计局专属的平板下载pda.app，就可以根据6张普查表的内容逐项完成调查，提交后软件系统即可在后台完成编码、录入、审核、预处理，并最后形成数据库。

### （三）录入审核及差错处理方法

1. 人工审核方法。在录入结束后，由审核人员将数据库中的记录与问卷资料进行核对，确定是否存在错录的情况。这种方法能发现差错的比例，但不能确定具体问卷录入中错误的所在，也就无法提供修改依据。

2. 双机录入。双机录入是指将同一份问卷上的内容在计算机中录入两次，两次使用不同的录入员。从理论上讲，两个录入员在同一份问卷的同一个地方按同样的方式发生错误的概率非常小，因此在录入完成后，研究设计人员可以将两个数据库进行比较，找出其中对应量之差不为0的变量，根据问卷号码查找，以便发现变量的真实数值。

3. 计算机辅助审核。随着计算机辅助调查的大力推广，利用设计好的软件直接对数据进行自动审核，大大提高了效率，节约了成本。

### （四）数据净化

全部数据录入完成后需要进行数据净化。数据净化是传统意义上的数据审核，是完美分析的前提。如果数据“不干净”，会发生两方面的严重问题。一方面，很有可能无法适当地执行下一步的数据分析，因而将延迟呈交报告的时间；另一方面，可能数据分析和报告已经完成，但是研究人员并没有意识到里面的许多错误。数据净化主要是尽可能地处理错误的或不合理的数据以及进行一致性检查。尽管在数据的校订阶段已经进行了初步的检查，但那只是局部的，分阶段的。而在完整数据的整理阶段，采用计算机进行的检查会更彻底、更广泛。

数据净化通常可采用统计软件进行，例如，用SAS、SPSS、BMDP等软件，可以很简便地寻找超出范围、有极端值或逻辑上不一致的数据。

数据净化通常的做法分为以下阶段：

首先通过描述统计分析进行审核完成一张所有非连续变量的频数表，以及计算连续变量的均值、标准差、最小值、最大值等统计量，那么超出范围的数据或极端值就可以检查出来。例如，假定“受教育程度”的编码应该是从1至4，分别对应4种不同学历层次的被调查者。假定用9表示缺失的数据，那么频数表中出现的不在1~4范围内的数据就是超出范围的。将超出范围（该变量取值为1~4和9之外的值）的样本筛选出来，根据对应的被调查者编号、变量编码、记录号码、列号码以及超范围的变量值等，就可以找到原始的问卷和数据文件的对应位置，进行必要的修改。

其次利用统计表与统计图检查逻辑上的一致性。通过统计软件，编制交叉表进行交互分析，从中可以很方便地发现逻辑上不合理的数据。也可以通过统计图分布的特点，发现逻辑上不合理的数据。然后根据

异常数据对应的被调查者编号、变量编码、记录号码、列号码以及不合理的变量值等，进行必要的修改。

## S 第二市

# 调查数据的统计预处理

数据收集完毕后，经过净化处理，还需要根据不同的要求进行进一步的统计预处理。主要包括缺失数据的处理、离群值检验和处理、加权、原始数据或变量的转换以及数据库创建。

## 一、缺失数据的处理

在本章第一节问卷的校订环节就提出了不满意的答卷按缺失数据处理的条件。在多数情况下，少量的缺失回答是可以容忍的。但是如果缺失值的比例超过了 10%，就可能出现严重的问题。处理缺失值主要有四种方法：用一个样本统计量的值去代替缺失值；用从一个统计模型计算出来的值去代替缺失值；将有缺失值的个案整个删除（list-wise）；将有缺失值的个案保留，仅在相应的分析中作必要的排除（pair-wise）。下面分别介绍。

### （一）缺失值插补的常规方法

#### 1. 样本统计量的值代替缺失值 Q

缺失值可以用一个样本统计量代替，最典型的做法是使用变量的平均值，被称为平均值插补。这样，由于该变量的平均值会保持不变，那么其他的统计量，例如标准差和相关系数等也不会受很大的影响。例如，一个被调查者没有回答其收入，那么就用整个样本的平均收入或用该被调查者所在的子样本（比方说是属于社会地位比较高的那个阶层）的平均收入去代替。用平均值替代缺失数据，相当于对同一插补类中所有的被调查者使用相同的无回答权数进行调整。平均值插补还假定无回答是一致的，且无回答的被调查者与提供回答的被调查者具有相似的特征。平均值插补可能会得到较好的点估计（对总量估计和均值估计等），但插补后在均值这一点会形成一个人为的“峰值”，从而破坏了分布形态和变量之间的关系；而用常规的抽样方差公式进行计算会低估最终的方差。因此，平均值插补通常在没有辅助信息可用或只有少量记录需要作插补处理时才最后被采用。平均值插补不会对数据的分布产生实质性的影响。

#### 2. 统计模型计算结果代替缺失值。

利用由某些统计模型计算得到的比较合理的值来代替缺失值也是一种常用方法，例如利用回归模型、判别分析模型等。

利用回归模型的方法称为比率或回归插补。它是使用辅助信息及其他记录中的有效回答建立一个比率或回归模型，该模型表明了两个或多个变量之间的关系。例如，比率插补所使用的模型为：

其中： $y_i$  表示  $y$  的第  $i$  个单元值；

的表示该单元值，其与  $y$  相关；

$R$  是直线的斜率（即  $x$  每变动一个单位， $y$  平均变动的数值）；

$\epsilon$  是模型的随机误差项，均值为 0、方差为  $\sigma^2$ 。

换句话说，模型假定  $y$  服从近似的线性关系，且观察值  $K$  围绕这条直线上下波动，波动的幅度是随机项  $\epsilon$ 。

这时  $y_i$  的插补值按如下公式计算：

$$\hat{y}_i = \bar{y} + R(\bar{x} - \bar{X})$$

式中： $\hat{y}_i$  是变量的第  $i$  个记录的插补值；

$\bar{x}$  是插补类中记录的  $X$  值的均值；

$\bar{y}$  是插补类中记录的  $y$  值的均值。

例如，一份关于公司人员情况的调查问卷，问卷中有两个问题分别是公司职员月收入总额  $K$  和职工人数  $n$ 。但其中一份问卷中月收入总额项缺失，而职工人数已正确填报，同时该公司所属行业也已知。这样就可以用全部问卷中属于这一行业的其他有效问卷（即这些问卷中职工收入总额和职工人数都有效）求得每个职工的平均收入，再根据这个比值（平均收入总额与平均职工人数之比  $A$ ）与已知的该公司的职工人数，就可以确定该公司的月收入总额。

比率或回归插补值的精度很大程度上取决于要插补的变量与已知的变量之间是否存在密切的关系，取决于所运用的计算是严格限制在一个插补类中还是在全部数据集中。比率或回归插补方法的一个优点是可以保持变量之间的关系。如果使用常规的抽样方差公式计算，会低估最终的抽样方差。

一般而言，比率和回归估计产生的插补值比简单的均值法产生的插补值更加稳定。这种方法常用于市场调查中定量变量数据的插补。

判别模型常用于分类变量的插值处理。例如，在选举预测中，如果问在下次选举中会投谁的票时，有许多被调查者常常会给出“还没有决定”的回答。如果只是简单地删除掉这一部分的回答（有时可能高达 30% 左右），那么肯定会引起严重的预测偏差。处理这一问题的统计方法之一是寻找一个判别函数，使其能够区分那些已经决定投票选 A（假定只有两个候选人 A 和 B）的群体和选 B 的群体。这个函数可能由一些独立特征变量来解释，比如被调查者的社会地位、职业、党派、受教育程度、生活形态，等等。假定某位说“还没有决定”的被调查者给出了上述变量的答案，那么就可以通过其特征将他（她）划入“已经决定选 A”或“已经决定选 B”的群体中。这样，选举预测的成功率就会大大提高。

缺失值的插补方法还有热平台插补、冷平台插补、最近邻插补以及 EM 算法等，读者可查阅相关教材。

## （二）插补后数据的方差估计

前面讲到所有的插补方法都能为某一个缺失或不一致的值生成一个单一的插补值，这些方法也都会在一定程度上扭曲插补变量的原始分布，并导致使用标准方差估计公式得出不适当的方差估计。分布扭曲很大程度上取决于插补量的大小和使用插补的方法。

使用插补方法时有一点特别重要，即应该尽量避免因插补而引起估计量的方差估计被人为低估。插补会导致方差估计的置信区间过窄，从而引起检验中出现虚假的显著性。

经过插补操作后，估计量的方差包括两个分量：抽样引起的方差分量（抽样方差）和插补引起的分量（插补方差）。有插补数据时，抽样方差通常会低估，因为常规的抽样方差计算公式假定回答率为 100%。随机插补方法（是指插补定量数据时，用确定性方法得出一个插补值，加上从某个合适分布或模型生成的一个残差作为最后的插补值  $Q$ ）的一个好处是，它能给插补后的完整数据集加进一些噪声。因此，使用随机

插补方法时，用常规方法在大多数情况下 能正确地估计某一估计量的抽样方差。但是，如果要得到估计量的总方差的话，还必须估计插 补方差。

估计总方差中的抽样方差和插补方差两个分量都非常重要，因为它们能如实地告知数据 用户调查数据的质量，有助于数据用户做出正确的推断 C 了解抽样方差和插补方差的相对重 要性，有助于资源在样本量和插补过程之间进行合理的分配。

### （三）删除缺失值个案方法

将有缺失值的个案整个删除 (list-wise) 的方法，结果可能会导致很小的样本，因为很多被 调查者都多多少少会有一些项目没有回答。删除大量数据并不可取，因为数据的收集花费了 大量的经费和时间。而且，有缺失回答的被调查者与那些全部回答的被调查者之间可能会有 显著的差异。如果真是如此，这种整个删除个案的方法会导致有严重偏差的结果，需要谨慎 使用。

### （四）排除缺失项目的方法

将有缺失值的个案保留，仅在相应的分析中将缺失的项目做必要的排除 (pair-wise) 的方法，会使分析中不同的计算将根据不同的样本量进行，这也有可能导致不适宜的结果。如果一 项调查中，样本量很大、缺失值很少、存有缺失值的变量之间不是高度相关，还是可以选择这一 方法对数据进行处理) 因而在实践中这种方法常被研究人员所采用心

不同的缺失值处理方法可能会产生不同的结果，特别是在回答的缺失不是随机的以及变 量之间存在高度相关的情况下。因此，应当使缺失数据保持在最低的水平，在选择一种处理 缺失数据的方法之前，研究人员应该仔细地考虑各种方法可能产生的后果。如果对缺失数据 进行了处理，应该有文件描述，并应向客户报告。

## 二、离群值检测和处理

离群值也称异常值。离群值检测主要是用来发现和确认可疑的记录：离群值被定义为看 起来与数据集中的其他观察值不一致的一个观察值或一组观察值。在离群值中要注意区分极 值和影响值。极值是指一组排序观察序列中的最大值或最小值。如果一个记录值和抽样权数 的组合对估计有较大的影响，我们就称这样的观察值为影响值：极值不一定是影响值，同样 地，影响值也不一定是极值。

离群值可以是单变量，也可以是多变量：如果一个离群值对应一个变量，该观察值就是 一个单变量离群值；如果一个离群值对应两个或多个变量，我们就说该观察值是一个多变量离群 值。例如，在一个关于身体素质的调查数据中，身高是 2 米或体重是 45 公斤的人会大有人在， 而符合身高 2 米且体重只有 45 公斤的人，就是一个多变量离群值的例子。

在调查中，有许多原因会导致离群值的产生。这些原因包括：数据中有错误，例如，数据录 入错误；可能混杂了来自另一模型或分布的数据，例如，大多数数据服从某个正态分布，但所推 测的离群值很可能是来自一个指数分布；也可能是由于数据固有的变异性，它可能来自于某一 分布的尾部，但却是完全可能的观察值，例如，按公司规模划分的销售额就服从典型的偏态分 布，少数大型公司常占整个销售额的绝大部分。

### (-) 离群值检验方法

最常见的离群值检验方法是单变量检验，因为单变量检验比多变量检验方法更简单 G 初 学者多倾向于随意舍弃离群值，以获得精度较好的分析结果。对于离群值，首先应从技术上设 法判断其出现的原因，如果查明确由技术上的失误引起的，不管这样的观测值是否为异常值，都应舍弃，而不必进行统计检验。但是，有时由于各种原因未必能从技术上找出它出现的原 因，在这种情况下应对其进行统计检验，以便从统计上判明离群值是否应该保留或舍弃

离群值的检验可分为两大类：一类是标准差预先已知的场合；另一类是标准差未知的场合。只能利用待检验的一组分析数据本身来检验其中的离群值是否为异常值 G 这里只对同一组分析数据中异常值的取舍予以介绍，不涉及同一总体中不同样本间的离群值的问题。

1. 标准差预先已知。检验时使用统计量为：

$$T = \frac{x - \bar{X}}{s}$$

式中， $x$  是被检验的离群值， $\bar{X}$  是一组观测值的算术平均值是由不包括异常值在内的 其他观测值求得标准差。如果计算的  $T$  值查表大于舍弃界限中相应置信度下的临界值，则将  $x$  作为异常值舍弃。

2. 标准差未知。在更多的情况下标准差是未知的，只能利用待检验的一组分析数据本身 来检验其中的离群值是否应该保留或舍弃。常用方法如下：

(1) 拉依达法。根据拉依达法，如果离群值  $x$  与观测值的平均值之差的绝对值大于 3 倍 的标准差 ( $s$ )，即

$$|x - \bar{X}| > 3s$$

则可以认为该离群值  $x$  为异常值。应从该组分析数据中舍弃；否则，应予保留。

(2) Q 检验法。亦称狄克松法，该法由迪安(Dean) 和狄克松(Dixon) 在 1951 年提出。具体步骤如下：

将分析数据由小至大按顺序排列： $x_1, x_2, \dots, x_n$ ，其中可疑值为  $x_1$  或  $x_n$  求出可疑值与其最邻近值之差 ( $x_2 - x_1$ ) 或然后用它除以极差 ( $x_n - x_1$ ) 计算出统计量：

$$Q = \frac{x_2 - x_1}{x_n - x_1} \text{ 或 } Q = \frac{x_n - x_1}{x_n - x_1}$$

$Q$  值越大，说明  $x_1$  或  $x_n$  “离群越远”。根据观测次数和所要求的置信度查表，若。大于所查 数值，则相对于。的  $x_1$  或  $x_n$  “为异常值，应予舍弃；否则，应予保留。

(3) 肖维特法。用肖维特法检验离群值，使用统计量：

根据上式算出的  $Q$  值，如果大于肖维特系数表中相应观测次数时的值，则可以认为该离 群值  $x_1$  或  $x_n$  为异常值，应从该组分析数据中舍弃；否则，应予保留，

(4) 格鲁布斯法。将分析数据由小至大按顺序排列： $x_1, x_2, \dots, x_n$ ，其中可疑值为  $x_1$  或  $x_n$ 。先计算出该组数据的平均值  $X$  和标准差  $s$ ，再计算统计量

$$\frac{|x - X|}{s} \quad | \quad \frac{x_n - x_1}{x_n - x_1}$$

根据事先确定的置信度和观测次数查表，如果  $G$  大于所查数值，则相对于  $G$  的  $M$  或  $X''$  为异常值，应予舍弃；否则，应予保留。

(5) 「检验法。用  $I$  检验法检验离群值，使用统计量

$$K = \text{-----}$$

和肖维特法不同的是，式中  $\bar{x}$  和  $s$  由不包括离群值在内的  $n-1$  个数据计算。根据事先确定的置信度和观测次数查表，如果计算出的  $K$  值大于表中相应的值，则可以认为该离群值  $X$  为异常值，应从该组分析数据中舍弃；否则，应予保留。

(6) 极差法’利用极差检验一组分析数据中的离群值时，使用统计量：

式中  $K$  为极差 根据事先确定的置信度和观测次数查表，如果计算出的门值大于表中相 应的值，则可以认为该离群值\*/为异常值，应从该组分析数据中舍弃；否则，应予保留。

在利用统计软件进行数据分析时，离群值检验一般通过残差分析部分来完成。

## (二) 离群值处理方法

对于在调查过程中检验出来的离群值，可以用多种方法来处理在手工审核系统中，经常 要对离群值进行插补处理 有些情况下，如果认为离群值无妨大碍，可以不对离群值做任何处 理；因此，主观判断就非常 重要，因为忽略或纠正离群值都会对数据的质量有很大的影响。

在审核时没有进行处理的离群值可以在估计的时候进行处理。简单地忽略未经处理的离 群值会影响估 计的效果，并导致估计时方差增大，给离群值赋予 1 或。的权数会使估计结果发 生偏差 离群值处理的目的就是在不引入较大偏差的前提下，尽量减少离群值对估计量抽样 误差的影响

估计时有三种方法可以处理离群值，即：改变数值、调整权数、使用稳健估计。

### 1. 改变数值

处理极值的一种方法是缩尾化，这种方法首先要将样本数据从小到大依次排序，如果认为 处于尾部的  $k$  个值为离群值，则采用余下的  $n-k$  个样本数据中的最大值代替这些离群值再进 行估计的方法。

例如，在简单随机抽样中，总体总量  $\bar{Y}$  的无偏估计公式如下：

$$\hat{y} = \frac{N}{n} \sum_{i \in s} y_i$$

其中， $y_i$  是指样本中第  $i$  个单位， $S$  是所有样本单元的集合(假定回答率为 100%)。

假设…，“是将样本观察值按从小到大的顺序重新排列得到的有序样本，若样 本数据中前 4 个最大值 (即有序样本值中的被认为是离群值，则单侧人次缩尾 估计量就可以通过采用前  $T$  个样本数据中的最大值 代替所有离群值来定义。即：

$$R^{\text{NI}} = (\text{Ex} < \text{---})$$

注意，缩尾化方法适合于处理单个变量的情况，在多变量抽样调查中很少应用。

### 2. 调整权数

处理离群值的另一种方法是，降低离群值的权数使它们的影响变小。若赋予离群值的抽

样权数为 1 或 0, 会对估计的结果产生太大的影响, 特别是对偏态总体的估计, 这样做会产生严重的偏差, 所以实际中很少采用。例如, 如果两个大公司的零售额占整个行业零售额的绝大部分, 而其中一个公司的零售额被确定为离群值, 若从估计中剔除这个离群值, 显然就会严重低估整个行业的总零售额。有些学者提出了一些降低离群值权数的估计量, 由于超出了本书的范围, 有兴趣的读者可参见 Rao (1970), Hidiroglou 与 Srinath (1981)。

### 3. 使用稳健估计量

经典的估计理论中, 总体参数的估计是基于某种分布的假设。通常假定估计量服从正态分布, 样本均值和样本方差估计量在正态分布的假定下是最优的。但是这些估计量对离群值非常敏感。稳健估计量则能克服这种局限性, 因为它对分布的假定不太敏感, 比如中位数比均值更稳健; 四分位数间距比通常的方差估计量更稳健。过去几年中已经提出了许多复杂的稳健估计量, 关于稳健估计量和离群值检测的详细内容, 请参见有关专著。

## 三、加权处理

在市场研究中, 加权处理比较常用。对完全无回答(即单元无回答)的情况, 常用的方法是调整被调查者的抽样权数。

加权处理的两个主要目的分别是增强部分样本的代表性和强调部分被调查者群体的重要性。在加权处理时, 要给数据库中的每一个个案或被调查者以一个权重, 用于反映其相对于别的个案或被调查者的重要性的差异。加权处理的效果是在具有某些特征的样本中增加(权重)1)或减少(权重<1)个案的数量, 使这些样本对目标总体更具代表性或减弱代表性。

如果研究的目的是要确定如何对现有的某种产品进行改造, 那么研究人员就可能想要赋予该产品的使用大户以更大的权重, 因为客户会更加重视这种大户的意见。例如可以给使用大户赋予权重3, 使用中户赋予权重2, 而对只是少量使用或不使用的消费者则赋予权重1。这样的加权处理就增强了大客户群体的重要性。

进行加权处理应该慎重, 加权的过程应该有文件描述, 并应向客户报告。

## 四、原始数据或变量的转换

在进行统计分析之前, 常常需要将原始数据重新分类、重新编码, 或需要重新定义变量、修改变量。

### (一) 原始数据的重新分类或重新编码

将原始数据重新分类是因为有时需要将数据分成更有意义的类别; 或者是需要将数据合并成更少的几个大类别。将原始数据重新分类时, 要注意重新构成的类别必须满足以下三点:

1. 所有的情况都已包括在新的类别之中;
2. 各个类别之间没有交叉或重叠;
3. 类别间的差异大于类别内的差异。

### (二) 变量的重新定义或修改

一种在分类中常用的重要变量类型叫做“哑变量”或“二值变量”“二分变量”“指示变量”“虚拟变量”, 等等。这种变量只有两个取值, 通常是0和1。为了方便数据处理, 经常需要重新定义变量或修改变量。对于分类变量定义的一般规则是, 如果一个定类变量有K个类别, 那么需要用KT个哑变量来表示。为什么是KT而不是K, 原因是在K类中, 只有K-1类是独立的, 因为第K类的信息可以通过K-I类推得。例如, 变量“性别”有两个类别, 因此只需要

用一个哑变量 X 来表示。

$y = r_1$ , 如果是男性

$\sim 1_0$ , 如果是女性

如果变量“所居住的城市”共有三个类别：北京、上海、武汉，那么需要两个哑变量 X<sub>1</sub> 和 X<sub>2</sub> 来表示 Q

$X_1 = 1$ , 如果是北京

其他

$X_2 = r_1$ , 如果是上海

$\sim 1_0$ , 其他

显然，当 X<sub>1</sub> 和 X<sub>2</sub> 都等于 0 时，表示的居住地就是武汉。

### (三) 数据标准化

重新定义变量的另一种情况是标准化，目的是为了使不同单位或不同量表的变量在分析中具有可比性。例如，对品牌 A 的评价使用的是 7 级量表（%），品牌 B 使用的是 5 级量表（分），等等，直接比较这些品牌的得分没有意义。为此，需要采用一般标准化方法，可将得分 …，L 转换成标准化得分 Z：即

$$Z = \frac{X - \bar{X}}{s}$$

其中， $\bar{X}$  和  $s$  分别表示变量  $X$  的平均值和标准差。

## 五、数据库创建

编码、数据录入、审核、插补和离群值的检验完成之后，就可以着手建立数据库，以方便后续阶段数据分析的计算机处理。数据储存的两种主要格式是数据库和平面文件数据。库是存储在一起的相关数据的集合。平面文件就是计算机化后的二维记录格式和它的相应值，它容易从一种平台转到另一种平台，并能用不同的表格软件或统计软件读取。平面文件的主要不足是，大多数的统计软件都要求数据按一种特定的格式存储，以便加快处理的速度，而用平面文件的时候，特定的格式还要重新生成，这就会造成效率低下。如果数据以一种适当的数据库格式存储，大多数统计和数据库软件就可以不需要重新生成文件而直接处理，查询操作也可以在数据库上直接进行。但数据库格式的选择可能会限制统计软件或数据库软件的选择。因此，较好的选择是对调查结果生成一个平面文件，同时生成几种不同的数据库文件。

一旦储存数据的格式选好了，就可以计算估计的最终权数，同时生成计划好的汇总表格。通常，需要编制计算机程序来完成这些工作。

## 第四节

### 编制数据分析任务书

在市场调查总体方案设计书或计划书中，可能已经给出了数据处理和分析的计划。但一般来说，总体方案中给出的计划只是原则性的或粗线条的，不具有操作性，项目主管或分析人员应该在数据采集与预处理完成后将数据分析的具体要求详细地列出，以使后续工作人员能明确其所需完成的全部统计工作。本节介绍编制数据分析任务书的主要内容以及选择统计方法时需要遵循的原则。

## 一、数据分析任务书的内容

一份详细地列出全部必需的统计工作的可操作的清单被称为数据分析任务书) 数据分析 任务书一般应包括：变量标识和变量值标识的编制说明、基本统计表、统计图的形式与数量、基础统计分析以及高级统计分析或模型五个方面的内容。具体分述如下。

### (一) 变量标识和变量值标识的编制

在数据录入阶段，一般只用字母和数字对变量进行编码’但在进行统计分析之前，还需要 对这些变量或由这些变量派生出来在进一步分析中使用的新变量的名称（简称变标）和取值 （简称值标）进行说明，并将这些说明制作成可运行的文件. 有经验的工作人员一般都会按照 问卷中回答题的内容来编制变标和值标，但是对于派生出来的新变量，需在任务书上加以说明。表 10-4 为完成值标和变标的编制任务后给出的截取结果。

表 10-4 部分变量标识和变量值标识的编制结果

变量序号	变量名	变量说明	对应题号	变量类型	变量取值标识
1	A1	性别	1	虚拟变量	1 =男； 0 =女
2	A2	学历	2	定序变量或虚拟变量	•专科及以下; 2=大学本科; 3=硕士研究生; 4=博士研究生
3	A3	月均可支配 生活费	3	定序变量	1 = 600 元以下； 2 = 600~1 200 元； 3=1 200-1 800 元; 4=1 800-2 400 元； 5 = 2 400 元以下
4	A41	目前使用的 BAT 移动支付软件	4	定类变量或 虚拟变量	1 = 百度钱包; 0 = 不选
5	A42				1 = 支付宝; 0 = 不选
6	A43				1 = 微信支付; 0 = 不选
7	A44				1 = 其他; 0 = 不选
8	A5	每月平均使用 BAT 移动支付的频率	5	定类或定序 变量	1 = 1-2 次； 2=3~5 次； 3=6~10 次； 4=11-15 次； 5=16~20 次； 6= 20 次以上

注：来自案例“关于 BAT 移动支付使用意愿影响因素分析——武汉大学生的调查”

### (二) 基本统计表

在市场调查中，向客户提供的最基本的统计表一般是两部分：第一部分是所有问答题（变量）的频数表。如果是连续型变量则可以分组后做出频数表，或计算其均值和标准差（到底是 分组还是直接计算均值，需要在任务书上给予具体的说明）。第二部分是反映被调查者背景 资料（如性别、年龄段、文化程度、收入段等）的变量与所有其他问答题（变量）的交互分析表。 如果是连续型变量则需要具体说明是分组后做出交互分析表，还是按背景资料变量的不同类别计算其分类均值

这一部分的说明可以很详细，也可以很概括，主要取决于项目主管与分析员的配合习惯。变量类型与统计表的关系见表 10-5。

表 10-5 变量类型与统计表的关系

变量类型		统计表
单变量	定类、定序变量	简单频数（含百分数）表 分组频数（含百分数）表
	定距变量	单变量简单分组表
	定比变量	单变量并列分组表
两变量	定类、定序、定距与定比变量两交 $>1$	二维列联表、两变量并列分组、两变量复合分组
三个及以上变量	引进一个或多个控制变量	多维列联表、多变量并列分组、多变量复合分组

### （三）统计图的形式与数量

在最终呈交的报告中，最好多用直观的统计图，以代替比较烦琐的统计表。数据分析任务书中需要说明：需要作图回答题及其数量，各种类型的图形分布，样本数据作图与背景资料分 类作图具体说明，等等。变量类型与图形形式的匹配表见表 10-6。

表 10-6 变量类型与图形形式的匹配表

变量类型	图形形式
定类变量	条形图、饼图
定序变量	条形图、饼图、折线图、态度对比图、轮廓形象图
定距变量	条形图、饼图、折线图、直方图、箱线图、雷达图、趋势图、散点图
定比变量（空间变量）	条形图、饼图、折线图、直方图、箱线图、雷达图、瀑布图、漏斗图、地图、趋势图、散点图

### （四）基础统计分析

基础统计分析主要指完成一般的调查报告所需的最低要求的分析一大部分是指单变量统计分析，如描述统计、主要指标的误差计算、统计推断、假设检验（t 检验、卡方检验、W 检验）、单因素方差分析、相关分析与回归分析等。这一部分的说明一般都是比较概括的，只需提出具体的要求，不必详述所采用的具体统计方法，因为一般有经验的分析员都掌握这些基础的知识，基础统计分析方法相关内容见表 10-7。

表 10-7 基础统计分析方法列示

基础统计方法	内容
描述统计	规模分析、结构分析、比较分析：频数、百分数
	集中度分析：均值（众数、中位数）、分位数
	变异性分析：极差（最大值、最小值）、方差或标准差、变异系数
	分布形态：偏态、峰度
样本误差计算	抽样方差、抽样标准差、变异系数

基础统计方法	内容
统计推断指标	总数、总量、平均数、比例、方差（点估计与区间估计）
正态分布检验 随机性检验 独立性检验 相关性检验	正态分布（JB 统计量） K-S 统计量 卡方统计量 积距相关系数、等级相关系数、一致性系数
平均数差异分析	单因素方差分析
回归分析	"统计量、，统计量、参数估计

## （五）高级统计分析或模型

在一般情况下没有必要使用高级统计分析或模型，因为利用任务书中的（一）~（四）步已经完全可以满足绝大多数客户的要求。但是对于有些比较专门的或深入的研究，高级统计分析或模型是必要的’或是客户虽然没有具体要求（大多数情况下是因为客户没有这方面的专门知识），但研究人员希望能呈交给客户一份更具吸引力、更能揭示内在规律的报告，在时间和经费许可的情况下，可以适当地使用一些高级统计分析方法或模型（例如因子分析、聚类分析、对应分析，等等）。

## 二、选择统计方法的原则

任务书的制定离不开一系列的统计方法。目前数据分析一般都采用现成的统计软件，软件中可供选择的统计方法十分丰富。选择时一般要考虑以下四个方面：

### （一）根据研究目的和研究的性质合理选择

在制定数据分析任务书之前，必须再次认真地参考调查项目总体方案设计的计划书，明确该项研究的主要目的，客户所面临的决策问题，本项目研究的主要问题，研究设计中有待检验的基本假设，等等。数据分析必须紧紧围绕着这些方面去设计和安排。尽量避免出现如下情况，例如只是按照问卷的顺序，逐题计算频数或均值、标准差，或者再增加一些交互分析表，或者只是运用高级统计方法做了大量的分析，但是并不能帮助解决客户所关心的问题。数据分析必须始终围绕所研究的问题，采用合适的统计方法，提供正确决策的有用信息。必须根据研究问题的不同性质，选择对应的方法。无论是学术性的研究，还是商业性的研究；描述性的研究还是预测性的研究，在任务书中都应该体现适用方法的差异与特点。

### （二）研究方法与数据的属性要匹配

数据的性质不同，所采用的分析方法就不同。例如对随机的大样本数据和对小样本的实验数据、对按不等概率抽取的样本数据和对等概率抽取的样本数据，选择的分析方法都会有所差异。此外，问卷中使用的量表和本章前面部分所述的数据整理方法都将影响数据的性质，进而影响到统计方法的选择。例如，对于定类变量，只能进行简单的分类或频数分析，对于定距或定比数据，才有可能使用高级的统计方法。因此，需要根据变量所属的不同类别，选用各自适用的统计方法。

### （三）正确辨别统计方法的作用和性质

在选择统计方法时还应充分理解各种方法的适用场合、应用目的、解决问题的性质、必须满足的假定或前提，等等。例如有些统计方法适用于考察不同群体间是否存在显著的差异，有些则适用于研究变量间的联系紧密程度，有些用于分类，有些用于简化数据，还有些特别适用于进行预测。有些方法需要假定变量服从正态分布，有些要求样本是独立同分布的，也有些方法并不需要任何的前提或假定。如果对统计方法

作用与性质没有基本的认识,将可能导致滥用统计软件中的各种方法,也许造成了严重的错误但并无觉察,后果的影响会很恶劣。

#### (四) 充分考虑研究者和客户的背景

项目研究主管的专业背景、工作经历和处理问题的习惯也会影响到统计方法的选择、有些研究者会对所研究的变量或总体做一些合理的假定;但有些研究者可能不愿意这样,他们也许宁愿保守一些,因此可能选择不需要对分布有前提假定的方法,例如非参数方法等。有些客户喜欢简单明了和直截了当的方法,但也有些客户希望做深入的分析甚至要求做模型。因此,在选择统计方法时也应充分理解客户的喜好和要求

### 三、常用统计方法的分类

常用的统计方法可以按单变量技术和多变量技术分成两大类。单变量技术也称之为基础统计分析法,适用于样本中的一个单位(单元)只有一个度量值,或虽有几个度量值,但每一个度量都是按一个变量单独地进行分析的情况。例如在一项市场调查中,样本中的每一个被调查者都回答了他们对某产品品牌的认知(乙)、使用情况(乙),满意程度(X<sub>0</sub>等项目,但在分析时,这些变量A<sub>1</sub>、A<sub>2</sub>和A<sub>3</sub>都是单独地进行分析。多变量技术也称之为多元统计分析法,适用于样本中的一个单位(单元)有多个度量值,而且对应的多个变量是同时地进行分析的情况。这两大类方法将在相关章节中进行介绍。基础统计分析法又按照所分析变量的性质(定类、定序、定距、定比)以及所涉及样本的情况(一个或多个)可以进行更细致的分类。同样,多元统计分析法可按照所涉及的变量中是否有因变量以及因变量多少进一步地分类。详细的分类如表10-8和表10-9所示。

表10-8 常用基础统计方法(单变量分析技术)

定类变量			定距、定比、定序变量		
单个样本	两个或以上样本		单个样本	两个或以上样本	
	独立样本	配对样本		独立样本	配对样本
频数分析	卡方检验	符号检验	Z检验	Z检验	Z检验
卡方检验	W检验	马克纽摩检验	t检验	t检验	单因素方差分析
	K-S检验				

表10-9 常用多元统计分析方法(多变量分析技术)

对称性技术(不含因变量)		非对称性技术(含因变量)	
变量间的相互依存关系	变量间的相似性	-一个因变量	两个及以上因变量
主成分分析	聚类分析	多维列联表	多元方差协方差分析
因子分析	联合分析	方差分析	典型相关分析
综合评价	对应分析	多重回归分析	多组判别分析
	多维标度分析	两组判别分析…	结构方程模型..

续表

对称性技术（不含因变量）	非对称性技术（含因变量）	
变量间的相互依存关系	变量间的相似性	一个因变量
		对数线性模型.
		方差协方差分析
		多元回归分析
		Logistic 回归

注：\*\*表示对应的方法超出了本书的要求，因此没有在书本中进行介绍；\*表示仅进行了简单介绍，没有详细说明

### 本章小结••••

实地调查完成后，市场调查工作就进入了调查数据处理与分析设计阶段。数据处理包括 问卷的接收、检查和校订，问卷的编码、录入与数据净化等主要的环节，随后进入统计预处理与 数据分析设计阶段。问卷接收与检查是研究者对回收到的问卷按收集时间、覆盖范围、调查 员、收集顺序等要素进行系统登记和记录的过程。问卷校订是对问卷进行验收、审查、校正的 过程，目的是消除数据中错误或含糊不清之处，使数据更加准确，校验工作的重点是数据资料 的易读性、完整性、一致性和准确性，编码是指在分类的基础上用数字代表类别，便于问卷查 找与数据分析。对于封闭式问卷，并有明确备选答案的问题，多采用事前编码；采用填空式问 题或开放式问题时，只能用事后编码。数据录入是指将问卷编码表的内容逐项转化成可机读 的数据形式。数据录入方式有人工录入、光电扫描录入和计算机辅助数据录入三种方式 每 一种录入方式都需要专门培训或专门程序设计。数据录入之后要依靠软件或人工对数据实施 净化，即通过逻辑性和一致性检查，保证数据的准确性、合理性和一致性。当完成数据的净化 工作后，通常进入数据预处理阶段，它包括缺失数据的处理、离群值的检测与处理、加权处理和 原始数据或变量的转换，并创建适合数据分析的数据库文件。基于完整数据库的分析需要在 原有调查设计基础上进一步完善，编制数据分析任务书，使之成为后续分析的指南。数据分析 任务书包括变量的确定、基本统计表、基本统计图、基础统计分析和需要的高级统计分析 统 计分析方法的使用以合适为原则，充分考虑研究的目的、对象特征、数据的属性与可获取 性，还 有研究者的需要与客户的背景 通常可以利用统计软件对原始问卷数据进行处理 磁关键名词••••

问卷审查 缺失值一致性检查 逻辑检查 事前编码

事后编码 编码手册 描述性统计 数据分析设计 离群值

均值插补回归插补加权数据库

直思考题 ..... .



即测即评

1. 数据整理包括哪些主要内容？
2. 什么是问卷编码？事前与事后编码如何进行？针对一份设计好的问卷，如何对其进行 编码？
- 3 如何保证数据录入的准确度？
4. 什么是数据净化？什么是统计预处理？
5. 缺失数据有哪些主要的处理方法？
6. 离群值检验与处理的方法有哪些？
7. 常用的统计分析方法有哪些？分别适用于什么场合？
8. 数据分析任务书的主要内容有哪些？

## 碧本章实训. . .

### 一、 实训目的

1 . 以前述完成的市场调查项目为基础，进行以下内容的实训：

理解和掌握调查资料的预处理和初步分析方法，包括问卷发放回收后所需完成的一系列 问卷编码、数据录入、审核、缺失数据插补、创建数据库、调查数据的初步描述分析等理论知识 与规范流程，练习运用 SPSS 软件进行数据处理与图表制作。

2. 可以收集二手微观调查数据，进行上述内容的实训。

### 二、 实训内容

1. 分组完成收集不少于 100 份的有效问卷，进行资料的预处理，并编写分析任务书。

2. 各小组在一周内完成数据整理和分析报告

### 三、 实训组织

1. 以小组为单位完成实训，每组人数控制在 4~6 人，各组做好数据的收集、整理、录入、分 析等分工安排。

2. 占用 2 课时，分组交流完成实训报告。

3. 在教师的指导下，针对资料整理、问卷编码与录入、数据分析和图表制作过程中遇到的 问题进行讨论，主要包括：

(1) 问卷的整理与编码。

(2) 运用 SPSS 软件进行数据录入，考虑变量的设置、变量赋值、缺失值的处理、变量尺度 类型等细节。

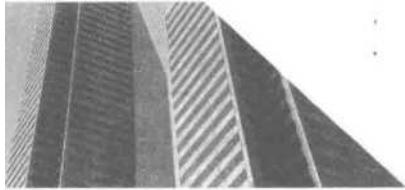
(3) 运用 SPSS 软件进行数据分析，考虑描述性统计分析、多选题的分析，以及交叉列表中 双变量的卡方检验。

(4) 量表题型如何录入与分析。

(5) 开放式问题如何编码、录入与分析。

4. 教师根据各小组实训报告和课堂交流情况进行总结和点评，并记录成绩 Q

# 第十一章单变量数据分析



## 、导语

二孩政策的全面放开，给婴幼儿奶粉市场带来了新的销售高峰。然而，随着母婴用品渠道渗透率的提高以及消费方式的升级，在移动互联网的推动下，各品牌婴幼儿奶粉的销售竞争激烈。对其推广提出了新的挑战。面对新的市场机遇和激烈的市场竞争，某婴幼儿奶粉品牌为了规划品牌结构，确定更加精准的品牌发展方向，委托某市场咨询公司开展婴幼儿奶粉市场研究项目，以期为该品牌提供精准品牌战略支持和精准营销沟通支持。

该研究项目采取预约上门访问的方式访问了2255位0~3岁宝宝的妈妈，样本筛选条件包括在目标城市连续居住三年以上、年龄在20~45岁的女性、婴幼儿奶粉用户以及婴幼儿奶粉品牌购买决策者等。访问内容包括宝宝养育、品类知识与态度、宝宝遇到过的健康问题、奶粉利益及特性、品牌认知与评价、品牌形象、购买模式、育儿理念与个性以及背景资料（居住城市、年龄、学历、家庭收入、宝宝月均支出等）。

本书截取该研究项目中的购买模式部分以及部分背景资料的信息，随机抽取全部调查样本中来自北京、上海、成都、合肥、南京和武汉六个城市的370个样本数据①（见数据集Example11-1和图1J-1SPSS格式截图），讲解数据分析方法并演示相应的实现过程，具体变量信息如表11-1所示。

①出于商业案例数据的保密性原则，本案例的原始数据已经过编辑和处理，不再代表实际市场情况。其分析结果仅用作方法演示和学习参考。

表 11-1 数据集 Examplell-1 变量信息

变量分类	变量名称	变量赋值及含义
背景资料	居住城市(city)	1=北京;2=上海;3=成都;4=合肥; 5=南京;6=武汉
	年龄(age)	以周岁计，连续变量
	平均每个月宝宝的总支出(「expend）	单位：元，连续变量
	平均每个月宝宝的奶粉支出(m.expend)	单位：元，连续变量
	平均每个月的家庭总收入(income)	单位：元，连续变量
购买模式	过去三个月，购买奶粉的总次数(num)	单位：次，取值范围0~9之间
	奶粉购买渠道•(channel)	1=连锁类母婴店；2=非连锁类的单体母婴店；3=大卖场/大超市；4=中小超市；……；15=保税区购买；16=代购；17=其他
	奶粉购买频率(frequency)	1=每周1次；2=每月2~3次；3=每月1次；4=每2~3个月1次；5=每4~5个月1次；6=说不清

\*购买渠道为多选题，在数据集中设置成 channel】-channel 17 共 17 个二分类变量，'1' 代表使用过该渠道购买婴幼儿奶粉，'0' 代表未使用过该渠道购买婴幼儿奶粉-

QMbs Ifarform 伽版 graphs 部《 为瞄 A 碰成：疆答•德：																
诺	city	虹	[expand ! m expand]	income j	<	channel!	channel j	channel? I	cbann	滿:	channeE	i	channel? chan	喃'	channels i cl	:
6096	北京	31	2797	1314	15677	0	0	0	0	Q	i	0	0	0	d	
6105	北布	31	2798	1235	16725	1	1	1	0	1	tr	0	0	0	0	
60	6107	北京	的	2923	1039	16894	0	1	0	.I	0	0	0	0	Q	
61	6134	32	2595	1293	17350	7	0	0	1	0	0	0	0	0	Q	
62	6149	北卓	30	2894	1118	18086	7	1	G	1	0	0	1	1	0	0
	6165	北京	32	2634	1534	18639	8	0	0	t	:	0	p	0	0	fi -
	6157	北点	31	2913	103J	19429	3	1	0	0	I	0	0	0	0	,0
K	6169	北京	37	2963	1406	19731	5	Q	0	w	J	0	0	1	0	0
56	6185	北奉	36	2505	1701	20114	3	1	0	1	j	Q	0	D	B	
伊	6218	北京	35	6712	2520	61235	7	EJ	Q	1	0	0	0	0	D	0
68	1009	上青	22	437	186	2761	6	0	0	1	0	Q	0	1:	0	0
69	1016	±3	35	234	156	3502	9	0	0	1	0	0	b	0	0	0
70	1027	上海	21	689	142	4348	1	0	D	1	0	0	0	1	0	0
71	1047	上海	20	917	119	4515	4	D	0	1	T	1	i	a	Q	0
72	1049	上萼	24	035	179	4524	8	0	0	0	i	0	i	T	D	0
73	1CB8	上萼	21	684	211	5824	5	1	8	I	0	Q	0	D	0	
74	2009	上海	20	566	277	6989	2	u	Q	i	0	0	0	1	0	-0
	2017	上海	23	862	357	5958	5	0	0	O'	0	0	0	1	D	D
	2G21	上海	20	939	398	5713	8	0	fI	i	0	0	1	f	0	0
77	WED	上海	20	815	264	6412	8	0	0	0	in	1	tt	0	□	
	2J710	上萼	21	866	225	6331	B	0;	.a	i	j	Q	0	0	B	
79	2083	28	764	368	6511	6	0	0	i	J	0	0	0	G	2.	
80	3030	上導	26	757	391	7660	7	0	0	i	1	0	G	0	0	Q
							—TM					*****				
												海	3Y			
	0<tvView															

图 11-1 婴幼儿奶粉市场研究案例数据集 Examplell-1 SPSS 格式截图

数据经过收集整理后，进入数据分析阶段。分析方法的选择是研究者面临的首要问题。根据调查目的确定研究变量，是选择分析方法的第一步。根据研究变量的个数，数据分析方法

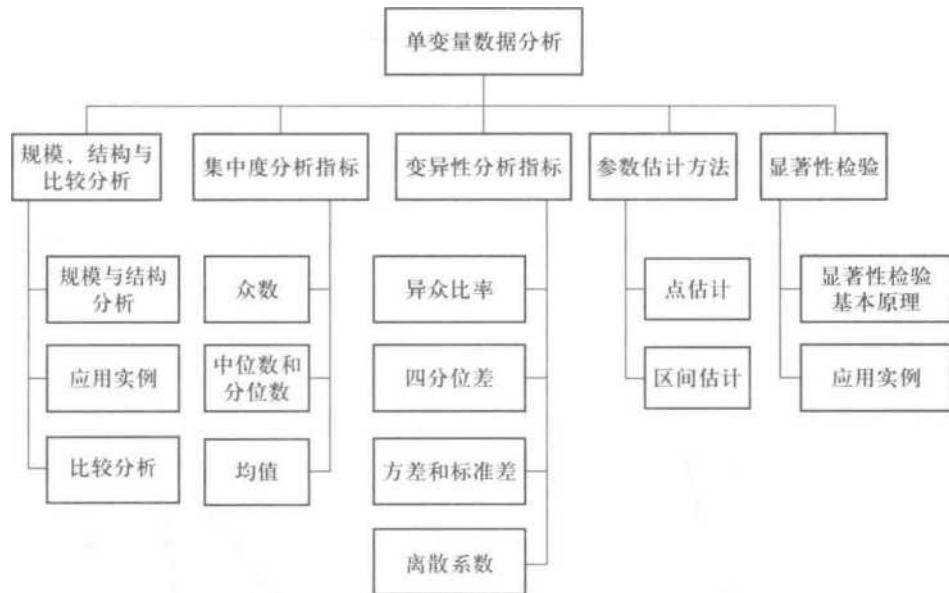
可分为单变量分析、双变量分析和多变量分析。在实际调查中，应根据不同的研究目的和不同的变量尺度，选择不同的分析方法。本章将具体介绍单变量分析方法。

单变量数据分析的目的在于对样本所有元素在某一方面（变量）的观察值进行概括性的描述或统计推断。单变量数据的描述性分析包括规模、结构与比较分析，集中度分析，差异性分析；推断性分析包括参数估计和显著性检验。在本书第四章第二节中，我们提到变量可以划分为定类、定序、定距和定比四种尺度，针对不同尺度的变量，分析方法往往也有所不同。本章将从引例出发，讲解上述分析方法的基本原理和在实际应用中应该注意的问题，同时介绍其中常用方法的SPSS软件实现过程。

### 、教学目标与要求

1. 了解与掌握单变量规模、结构与比较分析的常用方法；
2. 掌握不同尺度数据的集中度测度方法；
3. 掌握不同尺度数据的差异性测度方法；
4. 掌握参数估计的两种形式；
5. 了解显著性检验的基本原理，掌握显著性检验的步骤；

## 本章知识结构



## 第一节

# 规模、结构与比较分析

市场调查进入数据分析阶段后，最基础的分析内容是从整体上把握现象的规模和结构，并大体判断其在同行业、同区域或者同类现象中所处的水平。以引例中的宝宝平均每月奶粉支出变量为例，分析中首先要了解宝宝平均每月奶粉支出的整体规模和分布结构，获取该变量最基本的信息。

### 一、规模与结构分析

规模分析通常以频数或频率来描述。频数是指变量值中代表某种特征的观测值出现的次数，频率是指频数与全体观测值个数的比值，也可用来反映数据的结构。频数的表现方式主要有频数分布表和频数分布图。

#### (一) 频数分布表

把总体按某一标志分组，并按一定顺序列出每个组的单位数，所形成的总体单位在各组间的分布就是频数分布；把总体中各个类别及其相应的频数、频率及累计频率等指标用汇总表格的形式展示出来，就形成了频数分布表。

品质型数据（包括定类与定序型数据）的频数分布表只需汇总观测值的个数并计算其相应的频率即可绘制。如表 11-2 所示。

表 11-2 婴幼儿奶粉市场研究项目样本居住城市频数分布表		
城市	频数/个	频率/%
北京	67	18.1
上海	60	16.2
成都	55	14.9
合肥	60	16.2
南京	71	19.2
武汉	57	15.4
合计	370	100.0

例如，引例中共包含 370 个样本，居住城市包括北京、上海、成都、合肥、南京和武汉，对应的样本个数依次为 67、60、55、60、71 和 57 个，对应的频率依次为 18.1%、16.2%、14.9%、16.2%、19.2% 和 15.4%，绘制的频数分布表如表 11-2 所示。

数值型数据需要绘制组距式频数分布表，绘制的关键在于组数、组距和组限的确定。

1. 确定组数时，不宜过多也不宜过少，应以能够充分显示频数分布的整体特征为原则；
2. 确定的组距，既要能区分各组的性质差异，也要能反映总体资料的分布特征，应尽可能为 5 或 10 的整数倍；
3. 关于组限的确定，有了组距之后，只要确定了最小组的下限，其余各组的组限也将随之确定，最小

组的下限必须包含数据中最小的观测值。此外，实践中所遇到的数值型变量多为连续型，为避免汇总过程中遗漏观测，相邻两组中较小一组的上限应当与较大一组的下限相重合；若数据中存在极大或极小的异常值，为了避免出现空白组，首末两组也可以使用“xx 以下”及“xx 以上”的开口组的表示形式。

组距式频数分布表通常为等距分组，比如，对于人均月收入变量，可以划分为“1 000 元以下 1 000—3 000 元 3 000—5 000 元”“5 000—7 000 元”“7 000—9 000 元”和“9 000 元及以上”六组，这里需要注意统计分组应遵循“上限不计入”原则。不过，在实际应用中，也可以根据特定现象进行不等距分组，比如，若调查所得到的人均月收入变量数据差距过大，也可以根据实际情况划分为“1 000 元以下 000—2 000 元 2 000—3 000 元 3 000—5 000 元”“5 000—7 000 元 7 000—10 000 元”“10 000—15 000 元”和“15 000 元以上”八组。

## （二）频数分布图

为了更为直观地显示数据的频数分布状况，可以在频数分布表的基础上绘制频数分布图，主要有条形图、饼图、直方图、折线图等形式。适用于描述品质型数据的频数分布图主要是条形图和饼图；而数值型数据的频数分布状态，可以通过直方图更为直观地显示，为强调频数分布的整体特征，还可以在直方图的基础上进一步加工制作出频数分布折线图或曲线图。

频数分布图可以直观地显示数据的结构。品质型变量通过饼图反映该变量的数据结构；数值型变量通过直方图或曲线图来显示数据分布的类型和特征，反映数据的结构。数据的分布特征不同，形成的频数分布图也表现出各种不同的类型，常见的类型主要包括：钟形分布（又分正态分布、右偏分布、左偏分布）、J 形分布（又分正 J 形分布和反 J 形分布）、U 形分布等。应根据研究现象的具体情况来判断分布的类型、特点及其原因。

钟形分布具体表现为中间隆起、两侧逐渐降低的曲线（见图 11-2）。钟形分布表明数据具有集中的趋势，多数数据集中在中间；越往两端，数据越少；在远离中心的位置，只有极少数的数据。钟形分布的中间隆起部分称为峰，两侧称为尾。若较长的尾指向左方，则称为负偏（左偏）分布；若较长的尾指向右方，则称为正偏（右偏）分布；若左右对称，则称为正态分布。

J 形分布通常是累积分布的表现形式，在图形上表现为一条从下向上或从上向下单调变化的曲线（见图 11-3）。若曲线呈单调递增的趋势，则称为正 J 形分布，一般用于描述向下累积的现象；若曲线呈单调递减的趋势，则称为反 J 形分布，一般用于描述向上累积的现象。

U 形分布是指中间凹陷、两端凸起的分布形式（见图 11-4）。U 形曲线通常用于描述具有生命或者质量特征的现象，例如产品的故障率数据，产品使用初期和老化期的故障率都比较高，中间阶段则故障率比较低。

此外，还可以通过计算偏度和峰度来进一步了解数据分布的形态特征。

偏度（skewness）是指数据分布的不对称性，它是反映变量偏斜程度的指标，记为 SK。偏度不仅能判断数据的偏斜方向，还能够测度数据偏斜的具体程度，计算公式为：

(a)右偏分布

(b)左偏分布

(c)正态分布

图 11-2 钟形分布图

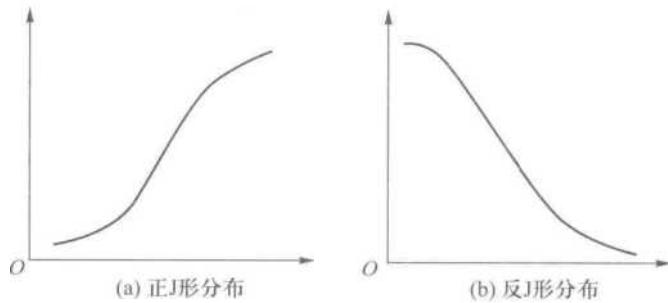


图 11-3 J 形分布图

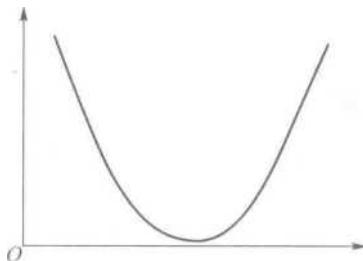


图 11-4 U 形分布图

$z$  (总\_力<sup>3</sup>)  
j1F' 对于未分组数据

(11.1 )

$z$  (州 - &)  $y$   
——，对于分组数据  
 $ns$

式中，\*为算术平均数，羽为第  $i$  个变量值  $m$  为样本单位总数， $s$  为样本标准差， $\Delta$  为组数， $M$  为第  $i$  组组中值  $J$  为第  $i$  组频数。 $SK = 0$  时，为对称分布； $SK < 0$  时，为左偏分布； $SK > 0$  时，

为右偏分布。且偏度越接近 0, 偏斜程度就越低:

峰度 (kurtosis) 是指数据分布的平峰或尖峰程度, 也就是反映频数分布曲线顶端尖峭程度 的综合指标, 记为 K'', 计算公式为:

$$KU = \frac{n(n+1) \sum (x_i - \bar{x})^4 - 3 \left[ \sqrt{\sum (x_i - \bar{x})^2} \right]^4 (n-1)}{(n-1)(n-2)(n-3)s^4},$$

$\bar{x}$  对于未分组数据  
 $s$  对于分组数据  
 $n$

11.2)

$KU = 0$  时, 为正态分布;  $K'' > 0$  时, 为平顶峰态;  $KU < 0$  时, 为尖顶峰态。

## 二、应用实例

### 应用实例 H-1

在婴幼儿奶粉市场研究项目中, 研究团队获取了受访者平均每月用于宝宝奶粉支出的数据, 根据宝宝平均每月奶粉支出的情况, 可以从消费者的角度初步了解婴幼儿奶粉市场销售额的总体规模和结构。

以引例中的数据集 Example1-1 提供的 370 个样本数据为例, 演示规模与结构分析在 SPSS16.0 (英文版) 软件中的实现过程。

分析: 宝宝平均每月奶粉支出为连续的数值型变量, 分析该变量的总体规模, 需要编制组距式频数分布表, 首先结合变量含义和数据的分布特点, 将宝宝平均每月奶粉支出分成等距的 9 组, 即: “200 元以下” “200 ~ 400 元” “400 ~ 600 元” “600 ~ 800 元” “800 ~ 1 000 元” “1 000 ~ 1 200 元” “1 200 ~ 1 400 元” “1 400 ~ 1 600 元” 和“1 600 元及以上”; 然后, 汇总出各组的频数、频率及累计频率等指标。宝宝平均每月奶粉支出的结构特征通过绘制频数分布直方图来展示。具体过程利用 SPSS 16.0 (英文版) 软件中的“ Frequencies (频数) ” 和“ Graphs (图形) ” 菜单实现。

#### 1. 操作步骤

(1) 打开数据集 Example 1-1-1。

(2) 对宝宝平均每月奶粉支出变量进行分组。分组后的变量命名为“G\_m\_expand”, 包含“200 元以下” “200 ~ 400 元” “400 ~ 600 元” “600 ~ 800 元” “800 ~ 1 000 元” “1 000 ~ 1 200 元” “1 200 ~ 1 400 元” “1 400 ~ 1 600 元” 和“1 600 元及以上”九组。分组过程通过“Transform (转换)”菜单中的“Recode into different variables (编码为不同变量)”过程实现。请读者注意, 运用 SPSS 中的 Recode 过程确定各组组限的归属时, 遵循“先定义的区间优先”的原则, 为实现“上组限不计入”, 应按照变量值从高到低依次定义各组区间 (如图 11-5 所示重新编码后的“G\_m\_expand”变量为数值型, 再依次为各个值设置值标签 (如图 11-6 所示), 即可得到如图 11-7 所示的分组后的数据集。具体操作步骤 读者可查阅 SPSS 相关教程, 这里不做具体讲解。

(3) 绘制频数分布表 依次选择 Analyze (分析) → Descriptive Statistics (描述性统计) → Frequencies (频数) , 展开 Frequencies (频数) 对话框, 如图 11-8 所示将变量“G\_m\_expand”

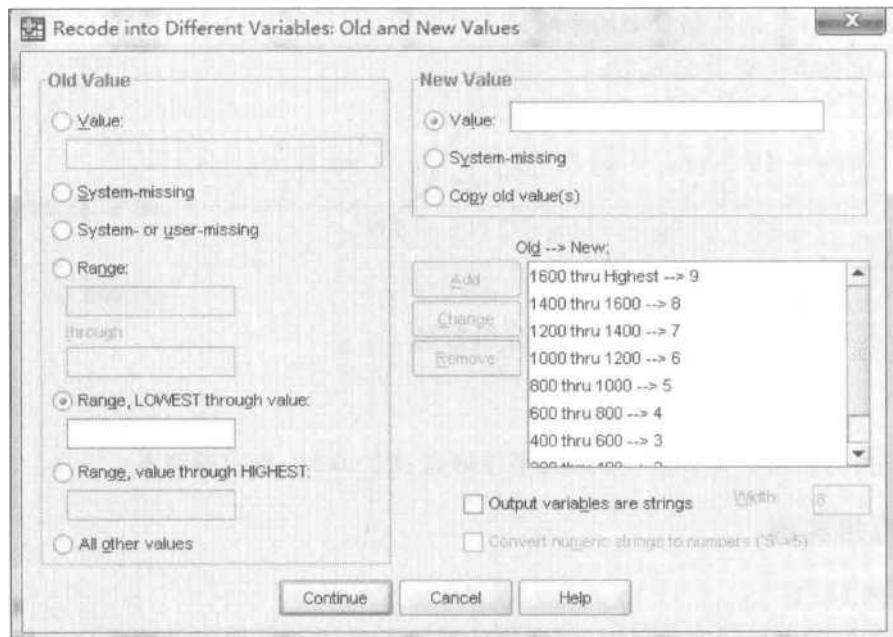


图 11-5 Recode 过程中各组区间定义顺序（由高到低）

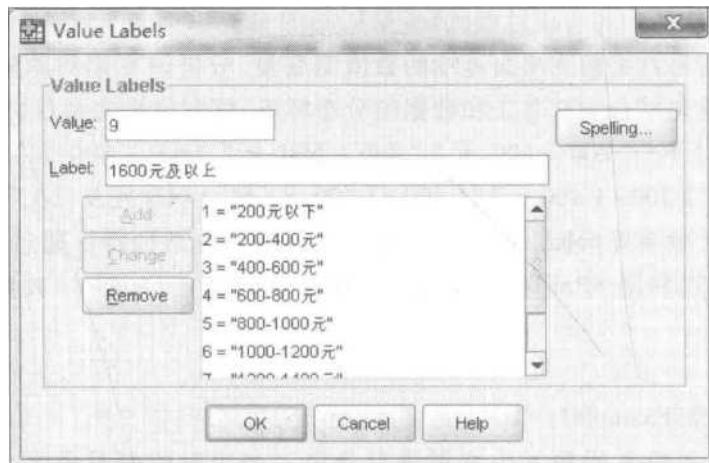


图 11-6 分组后的数值型变量设置值标签

（分组后的奶粉支出）”选入 Variable (s)（变量）框后，单击 OK 按钮，即可得到频数分布表输出结果（见表 H-3）。

（4）绘制频数分布直方图。依次选择“Graphs（图形）→ Legacy Dialogs（传统对话框）→ Histogram（直方图）”，在展开的“Histogram（直方图）”对话框中（见图 11-9），将变量“m\_expand”选入“Variable（变量）”栏，单击 OK 按钮，即可得到直方图输出结果（见图 11-10）。

宝宝每月奶粉支出分组后的数据集																	
姓名		性别		年龄		居住城市		收入		购买渠道		购买方式		购买量		购买金额	
S2	张	女	1岁	上海	13785	1	0	~U	1	channell	0	0	0	0	0	0	
ms	李	男	2岁	北京	15890	4	0	ab&tooCn	0	channeC	0	0	0	0	0	0	
5<	王	女	3岁	上海	912	a>io@n	1944	8	1	ch@rme0	0	0	0	0	0	0	
55	赵	男	4岁	北京	16045	0	0	aXMCOH	0	channeM	0	0	0	0	0	0	
57	钱	女	5岁	上海	1029	IQ09	6	0	0	channels	0	0	0	0	D	0	
61	孙	男	6岁	北京	1527	16930	2	0	0	j channeE	0	0	0	0	0	0	
J g	周	女	7岁	上海	13U	imuo 訂	0	0	1	chaKie(7 charmdB	0	0	0	0	0	0	
®	吴	男	8岁	北京	1235	I200-1400n	15725	1	1	channel	0	0	0	0	e	0	
SO	冯	女	9岁	北京	1039	100CH200	15994	0	1	channel	0	0	0	0	0	0	
6t	陈	男	10岁	北京	259	183	17350	7	0	channel	0	0	0	0	0	0	
62	胡	女	11岁	北京	289	180%	180%	7	1	channel	0	0	Q	1	i	0	
63	林	男	12岁	北京	26.3	1534	10639	8	0	channel	0	g	0	0	0	0	
64	秦	女	13岁	北京	291	140Q-	19429	3	1	channel	0	0	0	0	0	0	
65	徐	男	14岁	北京	296	14W60Q	19731	5	0	channel	0	0	0	1	0	0	
66	朱	女	15岁	北京	36	17DI	20114	3	1	channel	0	0	a	0	0	0	
87	蒋	男	16岁	北京	35	252D	61235	7	0	channel	0	0	0	0	0	0	
®	谢	女	17岁	北京	49	1%	2761	6	0	channel	0	0	t	0	0	0	
35	高	男	18岁	北京	23	15B	3582	9	0	channel	0	0	0	0	0	0	
上詢	王	女	19岁	北京	21	2D 元以	4348	1	0	channel	1	t	c	0	1	0	
71	范	男	20岁	北京	91	3B 无虾	<16	4	0	channel	1	1	1	0	0	0	

图 11-7 宝宝每月奶粉支出分组后的数据集截图

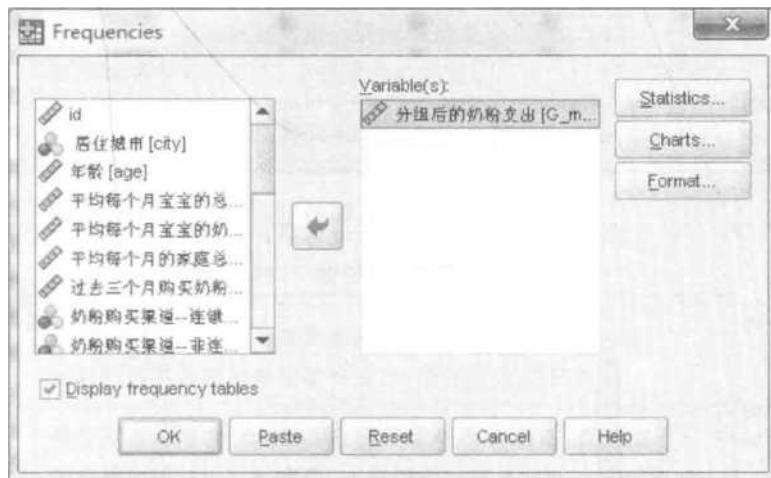


图 11-8 Frequencies 对话框

表 11-1 宝宝平均每月奶粉支出的频数分布表

	Frequency	Percent	Valid Percent	Cumulative Percent
200 元以下	27	7.3	7.3	7.3
200-400 元	68	18.4	18.4	25.7
400-600 元	88	23.8	23.8	49.5
600-800 元	56	15.1	15.1	64.6
800-1 000 元	67	18.1	18.1	82.7
1 000-1 200 元	17	4.6	4.6	87.3
1 200~1 400 元	14	3.8	3.8	91.1
1 400-1 600 元	13	3.5	3.5	94.6
1 600 元及以上	20	5.4	5.4	100.0
Total	370	100.0	100.0	—

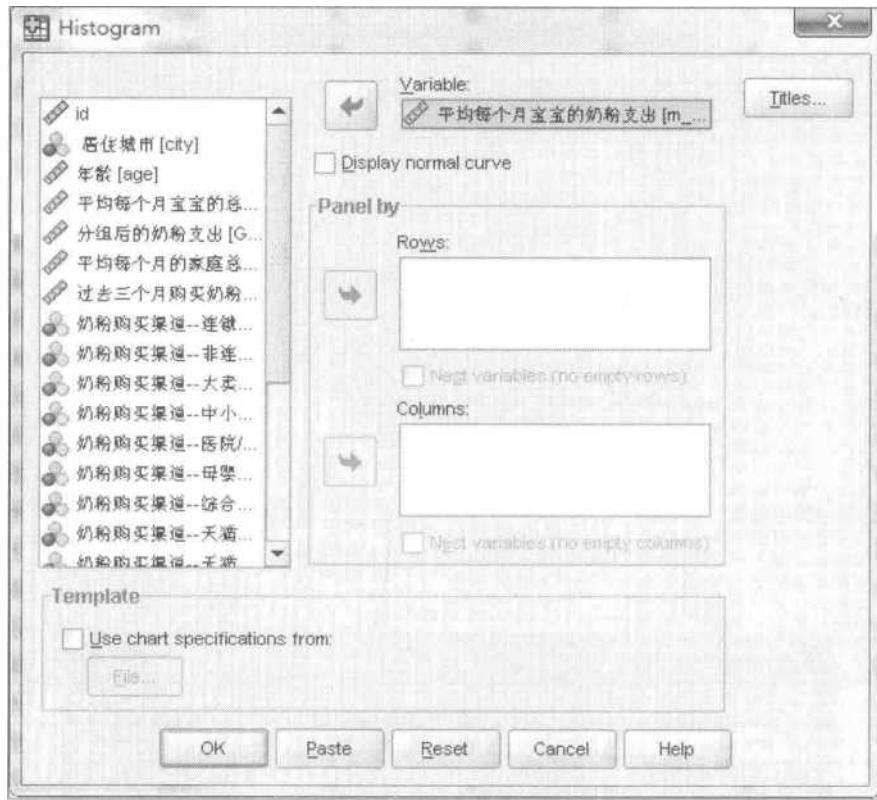


图 11-9 定义直方图的对话框

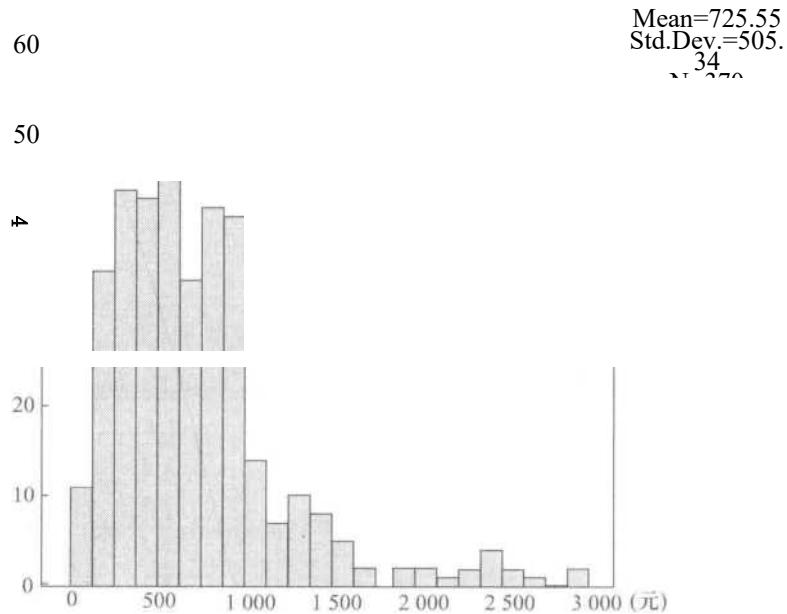


图 11-10 宝宝平均每月奶粉支出的频数分布直方图

## 2. 解读输出结果

在宝宝平均每月奶粉支出的频数分布表输出结果(见表 11-3)中, 从左到右依次为: 分组 区间、频数(Frequency)、频率(Percent)、有效频率(Valid Percent)和累计频率(Cumulative Per-

cent). 可以看出, 宝宝平均每月奶粉支出变量不存在缺失值, 进入分析过程的有效值为 370 个; 奶粉支出在 400–600 元区间的客户最多, 有 88 个, 占样本总量的 23.8%。从频数分布直方图(见图 1—10)可以看出, 宝宝平均每月奶粉支出为右偏分布。

### 三、比较分析

在解现象的总体规模和结构后, 若要了解其在不同时期所处水平的变化, 或评判其在同行业、同区域、同类现象中所处的水平, 需要进行比较分析。前者属于时间维度上的比较, 后者则是空间维度上的比较。进行比较分析时, 要注意统计口径和时期的一致性。

时间维度上的比较分析, 通常是指将同一指标不同时期的数据进行相互比较, 也称为纵向比较。常用的比较指标包括逐期增减量、累积增减量、环比发展速度、定基发展速度、环比增减速度和定基增减速度。比如, 知道某地区某批发商近十年的奶制品销售额, 就可以在时间维度上分析销售额的增长量和增长速度。各指标的具体含义和计算方法如表 11-4 所示。

表 11-4 时间维度上的比较分析指标

指标	含义	计算方法
逐期增减量	报告期观测值与报告期前一期观测值之差, 反映观测值逐期波动幅度的大小, 一般用绝对数表示。	
累积增减量	报告期观测值与固定基期观测值之差, 反映观测值在给定时期内累计波动幅度的大小, 一般用绝对数表示	
环比发展速度	报告期观测值与报告期前一期观测值之比, 反映观测值逐期波动速度的快慢, 一般用百分数表示。	$\frac{y_t}{y_{t-1}} \times 100\%$
定基发展速度	报告期观测值与固定基期观测值之比, 反映观测值在给定时期内累计波动速度的快慢, 一般用百分数表示	$\frac{y_t}{y_0} \times 100\%$
环比增减速度	报告期逐期增减量与报告期前一期观测值之比, 反映观测值逐期增减程度的高低, 一般用百分数表示。	$\frac{y_t - y_{t-1}}{y_{t-1}} \times 100\%$
定基增减速度	报告期累积增减量与固定基期观测值之比, 反映观测值在给定时期内增减程度的高低, 一般用百分数表示,	$\frac{y_t - y_0}{y_0} \times 100\%$

进行空间维度上的比较分析, 常用统计指标是相对数。相对数是指两个有联系的总量指标对比计算的比率, 根据相比较的总量指标之间的关系不同, 相对数又可以划分为 结构相对数、比例相对数、强度相对数、比较相对数和动态相对数。各指标的计算方法如表 11-5 所示。

表 11-5 空间维度上的比较分析指标

指标	计算方法
结构相对数	$\frac{\text{总体的某一部分}}{\text{总体的整体}}$
比例相对数	$\frac{\text{总体的甲部分}}{\text{总体的乙部分}}$

指标	计算方法
强度相对数	<u>总体的甲方面</u> / <u>总体的乙方面</u>
比较相对数	<u>甲总体某时期某指标</u> / <u>乙总体同时期同指标</u>
动态相对数	<u>总体的甲时期指标</u> / <u>总体的乙时期指标</u>

此外，基于数据结构进行比较分析时，还可以采用的统计量是指数。用各种指数来比较一些特定的市场问题，既方便又直观。

例如，根据武汉市城市公共自行车使用情况调查数据，通过计算“受众指数”，反映目标群体的倾向性。城市公共自行车受众指数，是指使用城市公共自行车的某一目标群体（指按照年龄、职业、性别等某种特征定义的群体）相对于调查总体的特征指数。以年龄为特征将目标群体划分为“16~24岁”“25~29岁”“30~39岁”“40~49岁”以及“50~65岁”五个群体，则城市公共自行车受众指数的计算公式如下：

$$\text{城市公共自行车受众指数} = \frac{\sum_{i=1}^n p_i d_i}{\sum_{i=1}^n p_i} \quad (11.3)$$

$p_i$  表示第  $i$  个群体中使用城市自行车的比例  
 $d_i$  表示第  $i$  个群体的受众指数

这样定义的指数可以对不同目标群体做横向比较。例如，城市公共自行车在“25~29岁”人群中的“受众指数”为 130.5，在“40~49岁”人群中的“受众指数”为 84.7，通过比较可以发现，不同年龄段的群体使用城市公共自行车的比例存在差异，“25~29岁”群体的使用比例高于总体的使用比例，“40~49岁”群体的使用比例低于总体的使用比例①。需要指出的是，一般来说，不同年度调查获得的受众指数不宜进行纵向比较。

此外，市场调查中常用的指数还有满意度指数。满意度指数是指通过对评价分值的加权计算，得到测量满意程度的一种指数概念。例如，在某项大连市医护人员工作状况满意度的调查中，采用李克特 5 分量表，即“非常满意”为 5 分、“比较满意”为 4 分、“一般”为 3 分、“比较不满意”为 2 分、“非常不满意”为 1 分。按百分制计算，通常将上述各个得分段依次赋值为 100、80、60、30、0 分，则满意度的计算公式为：

$$\text{满意度} = \frac{\sum_{i=1}^n s_i p_i}{n} \quad (11.4)$$

其中， $s_i$  为对某项目的满意度， $n$  为对该项目进行评价的样本总量， $p_i$ 、外、%、几 5 分别为对项目表示“非常满意”“比较满意”“一般”“比较不满意”和“非常不满意”的样本单位数。经计算，受访医护人员对工作状况中各项目的满意度分别为：工作量（47.30 分）、工作环境（52.42 分）、工资福利（41.57 分）、职称评定（50.68 分）、管理制度（58.85 分）、职业发展（53.54 分）、社会地位（40.73 分），可以看出，各项目的满意度均低于“一般”水平对应的 60 分，其中管

①该案例来源于第五届（2015 年）全国大学生市场调查与分析大赛获奖作品《城市公共自行车使用现状及发展分析 报告一基于武汉六大城市抽样调查》，团队成员包括中南财经政法大学统计与数学学院孔静、尹璐、杜雨桐及林成范。指导教师：张海波，

管理制度的满意度相对最高，社会地位的满意度最低，仅有 40.73 分①。

指数的主要作用是用于比较，其特点是方便和直观因此，在使用中应注意以下几点：自定义指数时，除了应该考虑定义是否科学合理之外，还应考虑其计算是否简便可行；引用已有 的指数时，应明确其定义和算法，以便正确地使用；基数的确定也要与研究目的紧密结合。

## 第二节

# 集中度分析

在第一节中，我们以宝宝平均每月奶粉支出变量为例，在对其进行统计分组的基础上，通过制作频数分布表和条形图等形式，直观地反映了宝宝平均每月奶粉支出的规模和结构 在此基础上，为进一步了解宝宝平均每月奶粉支出的分布特征和规律，寻找代表值来衡量奶粉支出的一般水平，则需要对其进行集中度分析，目的是寻找一组数据一般水平的代表值或中心值。

### 一、集中度分析指标

集中度分析，即测度一组数据的集中趋势，反映各调查数据向其中心值靠拢或聚集的程度。测定集中度的指标有多种，不同测量尺度变量适用的集中度指标有所不同。一般来说，低层次数据的集中趋势测度指标适用于高层次的数据度量，但反之，高层次数据的集中趋势测度指标并不适用于低层次的数据度量。本节将按照测量尺度从低到高的顺序，依次介绍众数、中位数、分位数和均值等集中度测度指标。

#### （一）定类数据：众数

众数（mode）是指数据中出现次数最多的变量值，记为 $\text{e}$ 。主要用于测度定类数据的集中趋势，若更高层次的数据具有集中趋势，也可计算众数。

定类数据或单项式分组数据的众数可用直接观察法得到，即出现次数最多的变量值就是众数。比如，在某项关于武汉高校大学生最常使用移动支付消费服务的调查中，选择各项服务的人数分别为：缴费充值 260 人，网 E 购物 298 人，转账汇款 223 人，信用卡还款 71 人，打车 93 人，发红包 137 人，投资理财 77 人，票务支付 83 人，团购订餐 181 人，商店超市付款 61 人，在这组数据中出现次数最多的是网上购物，因此，网上购物为这组定类数据的众数，即 $\text{e} = \text{网上购物}$ ②。

定距变量分组数据的众数，除可以用众数所在组的组中值近似代替外，还可以用插值公式来确定，即：

$$\frac{\text{d}}{\text{f}_0 + \text{d}} \cdot \frac{\text{A}_1 - \text{A}_0}{\text{A}_1 - \text{A}_2} \quad (115)$$

① 该案例来源于第四届（2014 年）全国大学生市场调查与分析大赛获奖作品《医患关系缘何紧张——基于大连市医疗现状的调查》，团队成员包括东北财经大学统计学院裴晨、张梦茹、刘宁宁、崔惠及冷倩文，指导教师：孙玉环<sup>1</sup>

② 该案例数据来源于 2015 年中南财经政法大学专业硕士研究生案例分析大赛参赛作品《BAT 移动支付使用意愿影响因素分析——以武汉高校大学生为例》。团队成员：王亚琪、郭晓菲、徐婷婷、田晓春，指导教师：徐映梅

式中 $\text{A}_0$ 为众数所在组的下限值， $\text{d}$ 表示众数所在组变量值的次数与下一组变量值的次数之差， $\text{f}_0$ 表示众数所在组变量值的次数与上一组变量值的次数之差 $\text{f}_1$ 为众数所在组的组距。

有时出现频数最大的数据值可能会有两个或更多，在这种情况下，就存在不止一个众数。不过，在多众数的情况下，也就等于没有众数，因为对于三个或更多个众数来说，对于描述数据的位置不能起多大作用，也就没有什么统计意义。

## (二) 定序数据：中位数和分位数

### 1. 中位数

中位数 (median) 是一组数据按照大小顺序排列后，位置居中的变量值，记为中位数。将全部调查数据分为两部分，其中一半的数值小于中位数，另一半的数值大于中位数。

根据未分组数据计算中位数时，要先对数据进行排序，然后确定中位数的位置，最后确定中位数的具体数值。中位数位置的确定公式为：

$$\text{中位数位置} (11.6)$$

式中  $m$  为数据个数。

一组数据按照从小到大的顺序排列，设排序结果为  $w_1 w_2 \dots w_m$ ，则中位数为：

$$M = \begin{cases} [m+1]/2, & m \text{ 为奇数} \\ [(m+2)/2] + 1, & m \text{ 为偶数} \end{cases}, \quad (11.7)$$

由于分组数据中位数的确定方法相对复杂，这里不再介绍，有兴趣的读者可以参考相关统计学教材。

### 2. 分位数

将一组排序后的数据进行不同的等分后，各分位点上的值就是对应的分位数，可以用来衡量数据的位置特征。

中位数是从中间点将全部数据等分为两部分，与中位数相类似的还有四分位数、十分位数和百分位数等。

把数列中各单位的变量值按照由小到大排序，然后通过三个点将全部数据等分为四部分，其中每部分包含 25% 的数据，处于三个分割点位置上的数值就是四分位数 (quartile)。很显然，中间的四分位数就是中位数，因此，通常所说的四分位数是指第一个四分位数 (也称作下四分位数) 和第三个四分位数 (也称作上四分位数)。与中位数的计算方法类似，根据未分组数据计算四分位数时，首先对数据进行排序，然后再确定四分位数所在的位置。设下四分位数为  $Q_1$ ，中间的四分位数为  $Q_2$ ，上四分位数为  $Q_3$ ，则各个四分位数的位置分别为：

$$Q_1 \text{ 的位置} = \frac{m}{4} \quad (H.8)$$

4

$$Q_2 \text{ 的位置} = \frac{3(m+1)}{4} \quad (H.9)$$

$$Q_3 \text{ 的位置} = \frac{m+3}{4} \quad (H.10)$$

当四分位数的位置不在某一个数值上时，可根据四分位数的位置，通过按比例分摊四分位数位置两侧数值的差值来估算四分位数。

百分位数也是用于衡量数据位置的指标，其含义是，对于第  $P$  百分位数来说，它使得至少有  $P\%$  的数据项小于或等于这个值，且至少有  $(100 - P)\%$  的数据项大于或等于这个值。也就是说，第  $P$  百分位数将数据分为两部分，大约  $P\%$  的数据项的值比第  $P$  百分位数小，而约  $(100 - P)\%$  的数据项的值比第  $P$  百分位数大。

例如，用百分位数的形式报告某批发商奶制品销售额的分布状况，如果批发商对于某客户的奶制品销售额为 8 259 元，那么我们无法知道批发商对于该客户的销售额在所有客户中处于什么位置。不过该销售额恰好对应的是第 80 百分位数，据此可以知道大约 80% 的客户奶制品销售额要低于该客户，而约 20% 的客户奶制品销售额高于该客户。

### （三）定距和定比数据：均值

均值（mean）又称平均数，主要有算术平均数、调和平均数和几何平均数三种计算方法。其中，算术平均数是最常用的计算方法，根据所依据的数据不同，又可以分为简单算术平均数 和加权算术平均数两种形式。具体计算公式为：

①简单算术平均数——根据原始数据计算

$$X = \frac{\sum_{i=1}^n x_i}{n}$$

式中序号为算术平均数”为第  $i$  个变量值  $M$  为单位总数， $\Sigma$  为求和符号。

②加权算术平均数——根据分组数据计算

式中上为第  $i$  组的组中值或变量值  $J$  为第  $i$  组的单位数上为组数  $f_i$

特别地，由于均值受极端值的影响较大，因此可以考虑将数据排序后，按照一定的比例截去两端的数据，对余下的数据计算平均数，这样得到的平均数称为截尾平均数（trimmed mean）。如果截尾平均数和原平均数相差不大，则说明数据不存在极端值，或者两侧极端值的影响正好抵消；常用的截尾比例为 5%，即两端各去掉 5% 的数据，在 SPSS 中的“Explore（探索）”过程可以自动计算 5% 截尾平均数。

## 二、应用实例

### 应用实例 11-2

在应用实例 H-1 中，我们采用图表形式对宝宝平均每月奶粉支出的规模和结构进行了基本的刻画，在此基础上，可以通过计算集中度测度指标进一步分析奶粉支出的分布特征和规律。

分析：首先，判断宝宝平均每月奶粉支出为定比数据，可以采用中位数、分位数和均值等指标测度其集中趋势。其次，受访者主要来自北京、上海、成都、合肥、南京和武汉六个城市，为更全面地分析宝宝平均每月奶粉支出情况，可以分别计算不同城市宝宝平均每月奶粉支出的集中程度，即需要先按居住城市变量分组，然后再分别计算各组奶粉支出的中位数和平均数等指标。具体计算过程利用 SPSS16.0（英文版）软件中的“Explore（探索）”菜单实现。

#### 1. 操作步骤

- (1) 打开数据集 Example1-U
- (2) 依次选择“Analyze (分析) —> Descriptive Statistics (描述统计) —> Explore (探索性分析)”，展开“Explore (探索)”对话框，如图 11-11 所示。然后，将变量“m\_expand (宝宝平均每月奶粉支出)”作为分析变量选入“Dependent list (因变量列表)”框；再将变量“city (居住城市)”选入“Factor List (因子列表)”框，用以指定按不同城市进行分组分析。最后，单击 OK 按钮，即可得到不同城市的宝宝平均每月奶粉支出的描述统计结果（见图 H-12 和图 11-13）



图 11-11 Explore 对话框

		Case Processing Summary					
		Valid		Missing		Total	
		N	Percent	N	Percent	N	Percent
平均每个月宝宝的奶粉支出	北京	67	100.0%	0	.0%	67	100.0%
	上海	60	100.0%	0	0%	60	100.0%
	成都	55	100.0%	0	.0%	55	100.0%
	合肥	60	100.0%	0	.0%	60	100.0%
	南京	71	100.0%	0	0%	71	100.0%
	武汉	57	100.0%	0	.0%	57	100.0%

图 11-12 数据整理汇总结果截图

注：图中第 5 列数据个位数默认为 0，后续截图中小于 1 的个位数据表达同此

如果不区分居住城市，直接计算对于全部受访者的宝宝平均每月奶粉支出的集中度指标，则无须在“Factor List (因子列表)”框中指定变量。

## 2. 输出结果的解读

输出结果由数据整理汇总表（见图 11-12）和按城市分组计算的宝宝平均每月奶粉支出描述统计结果两部分组成（见图 11-13）。可以看出，北京、上海、成都、合肥、南京和武汉的样本数分别为 67、60、55、60、71 和 57 个，数据集中没有缺失值。描述统计结果中给出的集中度测度指标有均值（Mean）、去掉 5% 极端值后的截尾平均数（5% trimmed Mean）和中位数（Median）。由计算结果可知，北京市宝宝平均每月奶粉支出的均值为 769.36 元，去掉 5% 极端值后的截尾平均数为 742.09 元，中位数为 786 元；上海市宝宝平均每月奶粉支出的均值为

Descriptives				Statistic	Sid Error
盖均每个月宝宝的奶粉支 北京	Mean			769.36	49.900
	95% Confidence Interval for Mean	Lower Bound	Upper Bound	669.73 868.99	
	5% Trimmed Mean			742.09	
	Median			786.00	
	Variance			1 668E5	
	Std. Deviation			408451	
	Minimum			109	
	Maximum			2520	
	Range			2411	
	Interquartile Range			431	
	Skewness			1 369	.293
	Kurtosis			4 139	.578
上有 上海	Mean			632.97	50.704
	95% Confidence Interval for Mean	Lower Bound	Upper Bound	531.51 734.42	
	5% Trimmed Mean			601.06	
	Median			535.00	
	Variance			1 543E5	
	Std. Deviation			392.749	
	Minimum			119	
	Maximum			2033	
	Range			1914	
	Interquartile Range			369	
	Skewness			1.421	.309
	Kurtosis			2.237	.608
成都	Mean			732.38	63.841
	95% Confidence Interval for Mean	Lower Bound		604.39	

图 11-13 不同类型客户奶制品销售额的描述统计结果截图

632.97 元，去掉 5% 极端值后的截尾平均数为 601.06 元，中位数为 535 元。限于篇幅，其他四个城市的计算结果不做解读。根据不同城市的集中度计算结果，可以比较各城市宝宝平均每 月奶粉支出的平均水平所存在的差异。

另外，描述统计结果中给出的其他指标属于差异性分析的内容，将在本章第三节中讲解。

### 三、进行集中度分析时应注意的问题

分析数据的集中程度，应结合统计数据的实际分布状况和数据类型，选用恰当的指标形式，以克服不同形式的测度指标在使用范围上的局限性。一般来说，平均数充分利用了数据的全部信息，是实践中应用最广泛的测度指标，但其只适用于定距和定比数据；另外，平均数还容易受极端值的影响，如果数据中存在明显的极端值，则平均数的代表性就可能较差，此时可以选用截尾平均数来测度数据的平均水平。中位数和众数都是位置代表值，不受极端数值的影响，虽然它们的敏感性要弱于平均数，但当数据中存在极端数值或分布的偏斜程度很大时，选用它们描述数据的集中趋势，代表性要好于平均数。

## 变异性分析

在第二节中，我们通过计算中位数和均值等集中度测度指标描述了宝宝平均每月奶粉支出的集中趋势，而该变量各观测值之间的差异状况如何，还需要考察其离散程度，这是宝宝平均每月奶粉支出分布的另一个重要特征。另外，集中趋势测度指标值反映了奶粉支出的平均水平，那么这一平均水平的代表性如何，也应结合奶粉支出的离散程度加以判断。

### 一、变异性分析指标

变异性分析即测度数据的离散趋势，反映各调查数据远离其中心值的程度根据数据类型的不同，测定离散程度的指标主要包括：异众比率、四分位差、方差、标准差等绝对量指标，以及测度相对离散程度的离散系数等。

#### (一) 定类数据：异众比率

异众比率(variation ratio)是指非众数组的频数占总频数的比率，用  $V_r$  表示，其计算公式为：

$$V_r = \frac{f}{n} \quad (11.13)$$

式中， $f$  表示变量值的总频数， $n$  表示众数组的频数。

异众比率主要用于衡量众数对一组数据的代表程度。异众比率越大，说明非众数组的频数占总频数的比重越大，众数的代表性就越差；异众比率越小，说明非众数组的频数占总频数的比重越小，众数的代表性越好。异众比率主要用于测度定类数据的离散程度，当然对于定序数据和定比数据也可以计算异众比率。此外，还可以利用异众比率对不同总体或样本的离散程度进行比较。

#### (二) 定序数据：四分位差

四分位差(quartile deviation)是上四分位数和下四分位数之差，记为  $Q_d$ ，用公式表示为：

$$Q_d = Q_3 - Q_1 \quad (11.14)$$

四分位差反映了中间 50% 数据的离散程度，其数值越小，说明中间的数据越集中，数值越大，说明中间的数据越分散。四分位差不受极端数值的影响。此外，由于中位数处于数据的中间位置，四分位差的大小在一定程度上也可以说明中位数对一组数据的代表程度。四分位差主要用于测度定序数据的离散程度，对于定距和定比数据也可以计算四分位差，但对于定类数据并不适合计算四分位差。

#### (三) 定距和定比数据：方差和标准差

方差(variance)是指调查数据中各变量值与其算术平均数离差平方和的算术平均数，记为  $S^2$ ；标准差(standard deviation)是方差的算术平方根，记为  $S$ 。方差和标准差反映了所有变量值对平均数的离散程度，其数值大小与平均数代表性的大小呈反向变化，即方差或标准差越大，平均数的代表性越差；反之，方差或标准差越小，平均数的代表性越好。方差及标准差的计算

公式为：

① 根据原始数据计算方差

$$s^2 = \frac{2 \sum (x_i - \bar{x})^2}{n-1} \quad (11.15)$$

② 根据分组数据计算方差

$$s^2 = \frac{\sum g_i (x_i - \bar{x})^2 f_i}{k-1} \quad (11.16)$$

$$\bar{x} = \frac{\sum x_i f_i}{\sum f_i}$$

对方差进行开方，即可得到标准差。

方差和标准差是测定离散程度的最重要、最常用的指标，不受极端值的影响，能综合反映全部单位标志值对其平均数的实际差异情况；同时，用平方的方法消除各标志值与其算术平均数离差的正负值问题，在数学运算中处理起来也比较方便。

#### (四) 离散系数

离散系数是指调查数据的标准差与其算术平均数的比值，也称为变异系数，记为  $CV$ 。离散系数主要用于比较不同类别数据的离散程度，用公式表示为：

$$CV = \frac{s}{\bar{x}} \times 100\% \quad (11.17)$$

四分位差、方差和标准差等都是反映数据离散程度的绝对指标，离散系数则是反映数据离散程度的相对指标。离散系数越大，说明数据的离散程度越大，数据间的差异越大，平均指标对总体一般水平的代表性越差；反之，离散系数越小，说明数据的离散程度越小，数据间的差异越小，平均指标对总体一般水平的代表性越好。

## 二、应用实例

### 应用实例 11-3

在应用实例 11-2 中，我们分别测度了六个城市宝宝平均每月奶粉支出的集中趋势，得到了宝宝平均每月奶粉支出的平均水平，也在一定程度上反映了奶粉支出的分布特征。不过，奶粉支出平均水平的代表性如何，还需考察各观测值之间的变异状况，测度数据的离散程度。

应用实例 H-2 中的操作方法，同时计算出集中度指标和离散程度指标。图 H-13 输出结果中给出的离散程度测度指标有方差(Variance)、标准差(Std. Deviation)和四分位差(Interquartile Range)。可以看出，北京市宝宝平均每月奶粉支出的方差为  $1.668 \times 10^{-10}$ ，标准差为 408.45 元，四分位差为 431 元；上海市宝宝平均每月奶粉支出的方差为  $1.543 \times 10^{-10}$ ，标准差为 392.75 元，四分位差为 369 元。限于篇幅，其他四个城市的差异性分析结果不做解读。

## 三、进行变异性分析时应注意的问题

比较不同类别数据的离散程度时，应使用变异系数指标，而不宜使用标准差、方差等反映数据离散程度的绝对量指标。一方面，采用不同计量单位计量的变量值，其离散程度的测度值也会不同；另一方面，由于标准差、方差的数值大小受数据本身水平高低的影响，绝对水平高的数据，其标准差、方差的数值会偏

大，绝对水平低的数据，其标准差、方差的数值又会偏小，因此，对于平均水平不同或计量单位不同的不同组别的变量值，不宜使用上述绝对量指标来比较数据离散程度的大小，而变异系数反映的是数据的相对变异程度，可以准确地反映不同类别数据离散程度的差异。

## 第四节

### 参数估计

市场调查通常采用抽样调查的方法，即抽取部分个体或单位进行调查，抽样调查的目的是通过样本信息判断调查对象全体的特征。假设导语中给出的370个样本单位是通过简单随机抽样方法在各个城市的全部婴幼儿奶粉消费者中抽取的，在本章前三节中，我们基于样本数据，从规模和结构、集中趋势、离散程度三个方面描述了宝宝平均每月奶粉支出的分布特征和规律。在一定条件下，还可以通过这些样本数据来推断各个城市宝宝平均每月奶粉支出的平均水平，即进行参数估计。

参数估计是指在满足一定精度和把握程度的条件下，利用样本信息来估计总体特征的统计分析方法。由样本统计量的值推断总体参数的基础和前提是样本统计量作为随机变量的概率分布①，因此，抽样调查中只有采用概率抽样方法获取的样本数据才能用来推断总体信息，非概率抽样方法获取的数据不能用于推断总体。

由于在抽样调查中，参数估计涉及的理论知识很多，因此在本书第八章中已单列“抽样估计”一章加以介绍，本节将直接结合实际案例，给出参数估计方法的具体应用和基于SPSS软件的实现过程。

#### 一、参数估计方法

用样本统计量推断总体参数有点估计和区间估计两种形式。下面以简单随机抽样为例，介绍点估计和区间估计的计算公式。关于其他复杂抽样形式的计算公式，可以参阅本书第八章以及其他相关教材。

##### (一) 点估计

点估计是直接用样本估计量作为总体未知参数的估计量。例如，导语中通过样本计算的北京市宝宝平均每月奶粉支出的均值为769.36元，如果以769.36元直接作为北京市全部婴幼儿奶粉消费者平均每月用于宝宝奶粉支出的估计值，这种估计方法就是点估计。点估计方法具有简便、直观等优点，但由于点估计仅以一次抽样的结果来估计总体参数值，因此无法提供误差程度的准确信息，无法知道做出估计的把握程度有多大。

##### (二) 区间估计

与点估计不同，区间估计方法是以区间的形式给出总体参数的取值范围和推断的把握程度。对于定距和定比变量，区间估计方法主要是对总体的均值和方差进行估计；对于定类和定序变量，主要是对估计总体的比例进行估计。区间估计方法弥补了点估计不能给出推断把握程度的不足。

①概率分布的相关知识，读者可参阅统计学教材。

区间估计的基本原理是，根据给定的置信度构造出被估计的总体参数的上限和下限。设被估计的总体参数  $e$  落在区间  $(\bar{e} - a, \bar{e} + a)$  内的概率为  $P(\bar{e} - a < e < \bar{e} + a) = 1 - \alpha$ ，则称  $(\bar{e} - a, \bar{e} + a)$  为总体参数的估计区间， $a$  为下限估计值， $\bar{e} - a$  为上限估计值， $1 - \alpha$  为估计置信度， $\alpha$  为显著性水平。

对于单个总体，常用的置信区间的形式有：

(1) 对于正态总体，当总体方差已知时，总体均值  $\mu$  的置信度为的置信区间为：

$$\bar{M} - Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \bar{M} + Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \quad (11.18)$$

式中  $\bar{M}$  为样本均值， $n$  为样本量， $Z_{\frac{\alpha}{2}}$  为总体标准差又  $\alpha$  为显著性水平  $\alpha$  所对应的双侧分位数。

(2) 对于正态总体，当总体方差未知时，总体均值  $\mu$  的置信度为的置信区间为： $\bar{M} \pm t_{\frac{\alpha}{2}, n-1} \frac{s}{\sqrt{n}}$ ，如

$$\bar{M} \pm t_{\frac{\alpha}{2}, n-1} \frac{s}{\sqrt{n}}$$

式中  $t_{\frac{\alpha}{2}, n-1}$  为显著性水私所对 应的自由度为  $(n-1)$  的，值。

## 二、应用实例

### 应用实例 11-4

假设导语中的样本是采用简单随机抽样方法在各城市全部婴幼儿奶粉消费者中抽取的，可以根据各城市的样本数据，在 90% 的置信度下，推断各城市婴幼儿奶粉消费者平均每月用于宝宝奶粉的支出情况。如果通过其他辅助信息获取到各城市  $\sim 3$  岁婴幼儿总人数，还可以进一步推断各城市每月该年龄段婴幼儿的奶粉消费支出总量。

分析：本例是在总体方差未知的情况下，求总体均值  $\mu$  的区间估计问题，选取的置信度为 90%。由于简单随机抽样是一种自加权设计，其每个单元都有相等的入样概率，可以不考虑加权问题、具体计算过程利用 SPSS16.0 软件中的“Descriptive Statistics（描述统计）”菜单实现。

#### 1. 操作步骤

- (1) 打开数据集 Example" -1"。
- (2) 依次选择 "Analyze (分析) → Descriptive Statistics (描述统计) → Explore (探索性分析)"，展开 "Explore (探索)" 对话框，将变量 "m\_expand (宝宝平均每月奶粉支出)" 作为分析变量选入 "Dependent list (因变量列表)" 框；再将变量 "city (居住城市)" 选入 "Factor List (因子列表)" 框。该过程与集中度分析的操作过程一致，见应用实例 11-2 中的图 11-1"

如果不区分居住城市，则无需在 "Factor List (因子列表)" 框中选入变量。

- (3) 继续单击 "Statistics (统计量)" 按钮，在展开的 "Explore statistics (探索性统计量)" 对话框中，选中 "Descriptives (描述性)" 项，并将 "Confidence interval for mean (均值的置信区间)" 框中的数字改成 90（默认值为 95，读者可根据研究的具体情况设置不同的置信度）。再单击 "ContinueC (继续)" 按钮，返回到主对话框（见图 11-14）。最后，单击 OK 按钮，即可得到区间估计结果。

## 2. 解读输出结果

如图 11-15 所示的输出结果，分别给出了六个城市在 90% 置信度下均值的置信区间 (90% Confidence Interval for Mean)。其中，北京市宝宝平均每月奶粉支出估计区间的置信下限 (Lower Bound) 为 686.11 元，置信上限 (Upper Bound) 为 852.61 元，即北京市宝宝平均每月奶粉支出的置信区间为 686.11 元 ~ 852.61 元，上海市宝宝平均每月奶粉支出的置信区间为 548.24 元 ~ 717.70 元，做出上述估计的把握程度为 90%。

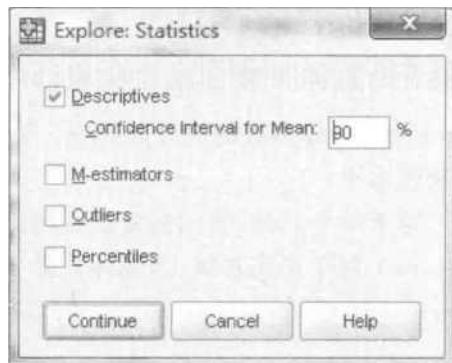


图 11-14 参数估计调用过程中

Statistics 对话框

Descriptives		Statistic	Std. Error
平均每个月宝宝的奶粉支	北京 Mean	769.36	49.900
	90% Confidence Interval Lower Bound for Mean	686.11	
	Upper Bound	852.61	
5%	Trimmed Mean	742.09	
	Median	786.00	
	Variance	1.668E5	
	Std. Deviation	408.451	
	Minimum	109	
	Maximum	2520	
	Range	2411	
	Interquartile Range	431	
	Skewness	1.369	.293
	Kurtosis	4.139	578
上海	Mean	632.97	50.704
	90% Confidence Interval Lower Bound for Mean	548.24	
	Upper Bound	717.70	
5%	Trimmed Mean	601.06	
	Median	535.00	
	Variance	1.543E5	
	Std. Deviation	392.749	
	Minimum	119	
	Maximum	2033	
	Range	1914	
	Interquartile Range	369	
	Skewness	1.421	.309
	Kurtosis	2.237	608
成都	Mean	732.38	63.841

图 11-15 按城市分组的宝宝平均每个月奶粉支出的区间估计结果截图

在应用实例 11-2 的输出结果中，给出了该变量在 95% 置信度下的区间估计结果，北京市宝宝平均每月奶粉支出的置信区间为 669.73 元 ~ 868.99 元，上海市宝宝平均每月奶粉支出的置信区间为 531.51 元 ~ 734.42 元，不同置信度下的区间估计结果是不同的。

限于篇幅，其他四个城市的区间估计结果不做解读。另外，为进一步推断各城市每月该年龄段婴幼儿的奶粉消费支出总量，如果通过其他辅助信息获取到各城市 0~3 岁婴幼儿总人数，用所得到的各个城市宝宝平均每月奶粉支出的估计区间上下限，分别乘以各城市 0~3 岁

婴幼儿总人数即可得到各城市每月该年龄段婴幼儿的奶粉消费支出总量的估计区间。

### 三、运用参数估计方法时应注意的问题

第一，应合理确定允许误差的范围。允许误差是样本统计量与总体参数离差绝对值的可允许变动范围，离差的绝对值越小，表明抽样估计的准确度越高；离差的绝对值越大，表明抽样估计的准确度越差。由于样本统计量为随机变量，所以估计误差难以避免，但若估计误差超过一定限度，参数估计本身就会失去价值。另一方面，估计误差也不是越小越好，因为减少误差势必会增加调查费用，甚至会失去抽样调查的意义。因此进行参数估计时，应根据所研究对象的变异程度及分析任务的要求，确定一个合理的误差范围，只要估计值与被估计值之间的离差不超过该允许范围，均可以认为这种估计是有效的。

第二，应合理确定置信度的大小。置信度也就是估计推断的概率保证程度。因为样本统计量存在随机性，所以实际进行抽样推断时，不可能有100%的把握做到被估计的总体参数落在允许误差的范围之内，这就产生要承担多大风险相信所做出估计的问题。如果估计的可信度很低，意味着风险很大，估计也就没有价值；另一方面，增大估计的可信度又会增加相应的调查费用，如果费用过高，也同样会失去抽样调查的意义。

此外，置信度和估计精度是相反的，因此进行参数估计时，应根据所研究问题的性质和工作需要，确定合理的误差范围和置信度。

## 第五节

### 显著性检验

显著性检验是推断统计中的另一个重要内容，应用范围非常广泛。比如，在导语中，不同城市的宝宝平均每月奶粉支出是否存在显著差异呢？在某公司研发的减肥茶效果的测试中，测试者饮用减肥茶前后的体重是否存在显著差异呢？解决上述问题，需要采用恰当的方法进行显著性检验。

显著性检验(significance test)是指事先对总体分布形式或总体(随机变量)的参数做出一个假设，然后利用样本信息来判断这个假设是否合理，即判断总体的真实情况与原假设是否有显著性差异。显著性检验可以分为参数检验和非参数检验两种，参数检验是在总体分布已知的情况下，先对总体参数提出假设，然后再利用样本信息去检验该假设是否成立；非参数检验是在总体分布未知的情况下，先对总体提出假设，然后再根据样本资料对假设的正确性进行判断。参数检验是显著性检验的核心内容，本节只介绍参数检验的基本原理和应用。

#### 一、显著性检验基本原理

显著性检验的基本思想是“小概率事件”原理，即发生概率很小的随机事件在一次试验中几乎不可能发生。根据这一原理，基于样本数据判断原假设 $H_0$ 是否成立，但这种判断总有可能是错误的。可能 $H_0$ 为真时，拒绝了 $H_0$ ，这类“弃真”错误称为第I类错误；也可能 $H_0$ 为假时，接受了 $H_0$ ，这类“取伪”错误称为第II类错误。显著性检验则是指在给定样本量的情况下，只限定犯第I类错误的最大概率不控制犯第II类错误的概率 $\alpha$ 。

显著性检验中涉及统计假设、统计量、显著性水平及显著性检验的  $p$  值等要素。统计假设是指对总体参数的假设值，一般分为原假设和备择假设，原假设是接受检验的假设，记作  $H_0$ ，备择假设是当原假设被否定时，另一种可能成立的假设，记作  $H_1$ 。在检验原假设是否成立时，对不同的问题需要构造不同的检验统计量，如：统计量、 $\chi^2$  统计量、 $P$  统计量等。显著性水平是指在  $H_0$  成立的情况下，统计量落在拒绝域的条件概率，记为“”，它是判断样本估计值与  $H_0$  统计假设值之间是否有显著差异的临界概率，一般取  $\alpha = 0.05$ 。显著性检验的  $p$  值是实际样本估计值与  $H_0$  下统计假设值之间不一致程度的概率值， $p$  值越小，说明实际样本估计值与  $H_0$  下统计假设值之间不一致的程度越大，表明检验结果越显著。

显著性检验的一般步骤包括：

- (1) 根据实际问题，对总体提出原假设  $H_0$  和备择假设出；
- (2) 确定显著性水平。；
- (3) 确定适当的统计量和拒绝域的形式；
- (4) 计算检验统计量的值和拒绝域；
- (5) 根据样本观察值，判断接受还是拒绝原假设

总体分布形式的正态性检验，一般用来检验总体是否服从或近似服从正态分布：服从正态分布的统计量具有一系列优良性，是基本统计方法应用的前提。常用的检验方法有：

- ① 偏度和峰度系数判定法，即分别计算样本的偏度系数和峰度系数，若两个系数都接近于 0，则可以认为近似服从正态分布；
- ② Q-Q 图检验法，以样本的分位数作为横坐标，以按照正态分布计算的相应分位点作为纵坐标，将样本表现为直角坐标系的散点，若样本点呈一条接近第一象限对角线的直线，则总体服从正态分布。

在实际调查中，大部分数据都不会是完全服从正态分布，只要数据样本量足够大，数据接近正态分布形态即可。

总体参数的显著性检验包括，检验、方差分析、 $t$  检验等，接下来只介绍单变量数据分析中常用的/检验，其他检验方法在后续章节中介绍。 $Z$  检验适用于数值型数据、正态分布、方差相等的两组小样本比较情况，包括两独立样本，检验和两配对样本，检验。其中，两独立样本，检验是根据分别来自两个独立总体的两组样本数据，检验这两个独立总体的均值之间是否存在显著性差异的检验方法；两配对样本，检验是根据分别来自两个配对总体并且具有一对一应关系的两组样本数据，检验这两个配对总体的均值之间是否存在显著性差异的检验方法。

## 二、应用实例

### 应用实例 11-5

假设导语中各城市宝宝平均每月奶粉支出均近似服从正态分布，可以进一步分析任意两个城市的宝宝平均每月奶粉支出是否存在显著性差异。

分析：由于每一个婴幼儿奶粉消费者是相互独立的，那么不同城市的婴幼儿奶粉消费者群体也相互独立，因此可以利用两独立样本，检验过程来检验任意两个城市的宝宝平均每月奶粉支出是否存在显著性差异。待检验的原假设是：城市 1 和城市 2 的宝宝平均每月奶粉支出不存在显著性差异。计算过程利用 SPSS 软件中的 Independent-Samples T test(独立样本/检验)”菜单实现。

#### I. 操作步骤

- (1) 打开数据集 Example 11 -1。
- (2) 依次选择<sup>44</sup> Analyze (分析) — Compare Means (比较均值) → Independent-Samples T test (独立样本 Z 检验)，展开独立样本，检验过程的对话框，如图 11-16 所示。然后，将变量“m\_expand (平均每个月宝宝的奶粉支出)”作为待检验变量，选入“Test Variable(s) (检验变量)”框中；再将变量“city (居住城市)”作为分组变量，选入“Grouping variable (分组变量)”框中，并单击“Define Groups (定义分组)”按钮，展开“Define Groups (定义分组)”对话框（见图 11-17），分别在“Group1”和“Group2”框中填入待检验的两个城市所对应值标签，比如，若检验北京（值标签为“1”）和上海（值标签为“2”），则需要在“Group”和“Group?”框中填入“1”和“2”，若检验上海和成都（值标签为“3”），则需要在“Group”和“Group?”框中填入“2”和“3”；继续单击“Continue (继续)”按钮，返回到主对话框。

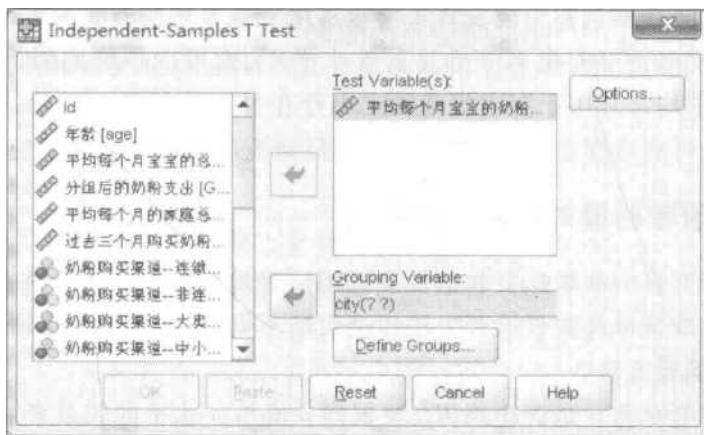


图 H-16 独立样本, 检验过程对话框

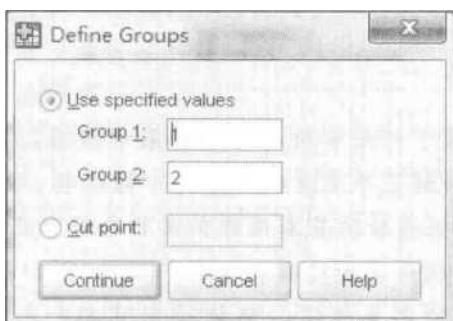


图 11T7 定义分组过程对话框

- (3) 单击 OK 按钮，即可得到检验结果（见图 11-18）。

## 2. 解读输出结果

输出结果由分组描述分析表和独立样本[检验表两部分组成（见图 11-18）以对北京和上海的检验结果为例，独立样本 z 检验表的左半部分，是判断两总体方差是否相等的“检验过程，由于  $p = 0.858 > 0.05$ （表中  $p$  值显示为 Sig.），可以认为北京和上海的宝宝平均每个月的奶粉支出的方差不存在显著差异，因此需要看该表中第三行“假定方差相等 (Equal variances assumed)”时的 t 检验结果：有  $t = 1.913$ ,  $p = 0.058 > 0.05$ ，据此可以得出如下统计结论：在 5% 的显著性水平下，没有理由拒绝“北京和上海宝宝平均每月奶粉支出不存在显著性差异”的原假

Group Statistics				
	N	Mean	Sid D 的 a, M	Sto Error
平均每个月宝宝的奶粉支出 北京	67	7B938	4Q8451	49900
平均每个月宝宝的奶粉支出 上海	60	63297	392 749	50 704

Independent Samples Test									
		Levene's Test for Equality of Variances		Mestfor Equality of Means					
		F	8ta	!	df	3ta (2 tailed)	Mean Difference	95% Confidence interval otthe Difference	SW. Enor Difference
平均每个月宝气硼棚支 出	Equal variances assumed	032	85B	1.913	126	.058	136 392	71 298	*4 710
	Equal variances not assumed			1317	124 3 公	.058	136.392	71140	-4 411
									277 493
									277104

图 H-18 独立样本, 检验结果截图 (北京和上海)

设, 即认为北京和上海婴幼儿奶粉消费者平均每月用于宝宝奶费的支出不存在显著性差异

但是, 按照上述分析方法, 在 10% 的显著性水平下, 则可以拒绝原假设, 认为北京和上海 婴幼儿奶粉消费者平均每月用于宝宝奶费的支出存在显著性差异 可见, 在不同显著性水平 下, 假设检验的结论可能会存在差异, 需要根据实际研究问题的特点来确定

### 三、进行显著性检验时应注意的问题

第一, 应根据研究目的和数据资料的性质选用恰当的检验方法一资料性质不同, 样本大小 不同, 所适用的检验统计量及假设检验方法也会有所不同, 因此需要根据研究目的和数据资料 的性质选用恰当的检验方法。

第二, 结论不能绝对化。是否拒绝要根据 P 值与。水平的关系来判断, 而 a 值又往往 需要根据所研究问题的特点来确定。当  $p < a$  时, 拒绝仇, 接受当  $p > a$  时, 不拒绝仇。实 际工作中, a 的确定往往具有灵活性, 有时在  $a = 0.05$  的水平下拒绝%. 而按  $a = 0.01$  时则可能 不拒绝仇。

## 必读! h 站

单变量数据分析的目的在于对样本所有元素在某一方面 (变量) 的观察值进行概括性的 描述或进行统计推断。本章讲解了单变量数据分析中的规模、结构与比较分析、集中度分析、 差异性分析、参数估计和显著性检验的基本原理和实际应用中应该注意的问题, 同时基于引例 介绍了其中常用方法的 SPSS 软件实现过程。

规模分析通常以频数或频率来描述。频数的表现方式主要有频数分布表和频数分布 图 o 数据的结构分析, 主要通过频数分布图以及偏度、峰度等统计指标来显示数据分布 的类型和特征, 反映数据的结构与特点。常见的分布类型包括: 钟形分布 (又分正态分 布、右偏分布和左偏分布) 、J 形分布 (又分正 J 形分布 和反 J 形分布) 、u 形分布等。比较 分析, 用来评判某变量在不同时期或在同行业、同区域或者同类现象中所处的水平, 分为 时间和空间两个维度。此外, 还可以采用各种指数来比较一些特定的市场问题, 既方便又直观, 比如市场调查中常用的受众指数和满意度指数等 Q

集中度分析, 即测度一组数据的集中趋势, 其目的是寻找一组数据的代表值或中心值, 以 此衡量全体数据的一般水平。测度集中趋势的指标主要有众数、中位数、分位数和平均数, 不 同类型的数据适用不同的指标。平均数能够充分利用全部数据信息, 但易受极端值的影响; 众 数、中位数和分位数是位置代表值, 不受极端值的影响, 但代表性弱于平均数

变异性分析，即反映各变量值之间的变异状况，需要测度数据的离散趋势，测度离散程度的指标主要有众数比率、四分位差、方差、标准差和变异系数等，比较不同类别数据的离散程度时，应使用反映相对差异的变异系数指标，而不宜使用标准差、方差等反映数据离散程度的绝对量指标。调查数据的离散程度越高，变量值之间的差异越大，数据集中度指标的代表性就越差，使用这些代表性指标进行统计分析的效果也就越差。

参数估计指在满足一定精度和把握程度的条件下，利用样本信息来估计总体特征的统计分析方法，主要有点估计和区间估计两种形式。抽样调查中只有采用概率抽样方法获取的样本数据才能用来推断总体信息，非概率抽样则不能用于推断总体。进行参数估计时，应合理确定允许误差的范围以及置信度的大小。

显著性检验是指事先对总体参数提出一个假设，然后利用样本信息判断该假设是否成立，即判断总体的真实情况与原假设是否有显著性差异。显著性检验可以分为参数检验和非参数检验两种：进行显著性检验时，应根据研究目的和数据资料的性质选用恰当的检验方法，同时注意结论不能绝对化。

### 多关键名词•••

频数 分布表 分布图 平均数 中位数 众数

方差 标准差 变异系数 区间估计 显著性检验



即测即评

### 肯思考题

1. 简述组距式频数分布表的编制步骤，以及编制过程中应当注意哪些问题。
2. 描述集中趋势的指标有哪些？有何异同之处？
3. 描述离散程度的指标有哪些？有何异同之处？
4. 应用参数估计方法时应注意哪些问题？
5. 简述显著性检验的一般步骤。
6. 进行显著性检验时应注意哪些问题？ 完本章实训•••

#### 一、实训目的

1. 明确单变量数据分析方法的基本原理。
2. 掌握应用 SPSS 软件进行单变量数据分析的操作方法。

#### 二、实训内容

1. 数据准备

葡萄牙某货物批发商经营生鲜产品、日用品、冷冻食品、洗涤剂和纸制品、熟食品以及奶制品，其客户主要包括酒店餐厅和零售商两类。某调查公司收集到该批发商 2014 年全年的销售数据（见数据集 Example 11-2 和图 11-19“SPSS 格式截图”），要求通过这些数据分析该商家的销售情况。

#### 2. 具体要求

应用数据集 Example11-2，完成对该批发商 2014 年全年各类商品销售情况的整体分析，包括：

- (1) 分析一类商品的销售情况。

	Num	Channel	Region	Fresh	Grocery	Frozen	Detergents Paper!	Delicassen	Milk
18/	137	酒店餐厅	其他	3009	854	347	949	727	521
188	188	酒店餐厅	其他	2438	9819	626	3459	3	8002
	189	零售商	其他	8040	11687	275	6839	404	7639
190	190	零售商	其他	834	11522	27	4027	1856	11577
191	191	酒店餐厅	其他	16936	1981	733	118	64	6250
192	192	酒店餐厅	其他	13624	1381	890	43	84	295
193	193	酒店餐厅	其他	5509	2251	547	187	409	1461
194	194	零售商	其他	180	20292	959	5618	666	3485
195	195	酒店餐厅	其他	7107	2974	806	355	1142	1012
196	196	酒店餐厅	其他	17023	5230	788	330	1755	5139
197	197	酒店餐厅	里斯本	30624	4897	18711	763	2876	7209
198	198	零售商	里斯本	2427	10391	112	4314	1468	7097
199	199	酒店餐厅	里斯本	11686	6824	352	592	697	2154
200	200	酒店餐厅	里斯本	9670	2112	520	402	347	2280
201	201	零售商	里斯本	3067	23127	394	9959	731	13240
	202	零售商	里斯本	4484	24708	354	14235	1681	14399
203	203	酒店餐厅	里斯本	25203	9490	506	284	6854	11487
204	204	酒店餐厅	里斯本	583	2216	469	954	18	685
205	205	酒店餐厅	里斯本	1956	5226	138	5	1328	891
206	206	零售商	里斯本	1107	23596	955	9265	710	11711
	207	酒店餐厅	里斯本	6373	950	878	288	285	780
208	208	零售商	里斯本	2541	6089	294	5316	120	4737

图 H-19 某批发商 2014 年全年销售额数据集 SPSS 格式截图

资料和数据来源：UCI 机器学习数据库①(<http://archive.ics.uci.edu/ml/datasets.html>)中的 Wholesale customers 数据集，该数据集共包含 9 个变量、440 个观测，其中，变量“Num”表示序号，“Channel”和“Region”分别表示客户类型和所在地区，变量“Fresh”“Grocery”“Frozen”“Detergents-Paper”“Delicassen”和“Milk”分别表示生鲜产品、日杂用品、冷冻食品、洗涤剂和纸制品、熟食品以及奶制品的销售额。

- (2) 绘制频数分布表，描述各类商品销售额的频数分布情况。
- (3) 计算集中趋势和离散程度指标，分别对该批发商对于酒店餐厅和零售商两类客户的各类商品销售情况进行描述性分析口
- (4) 在 90% 的置信度下，根据 440 个客户样本数据，推断该商家对于每个客户的各类商品平均销售额。
- (5) 分析该批发商对于酒店餐厅和零售商两类客户的各类商品销售情况是否存在显著性差异
- (6) 汇总各类商品销售情况的分析结果，并进行综合对比分析，完成该批发商 2014 年全年销售情况分析报告

### 三、实训组织

1. 以个人为单位，每人负责分析一类商品的销售情况。
2. 提交数据分析书面报告，
3. 教师批改报告，点评与总结一，

① UCI 机器学习数据库(The UCI Machine Learning Repository)，由加州大学欧文分校(UC Irvine)于 1987 年创建，截至 2018 年 1 月共包含 399 个基础测试数据集，被全球的研究者广泛使用。

# 第十二章 双变量关联性分析

## 、导语

、 ■ · ■ - · · ·

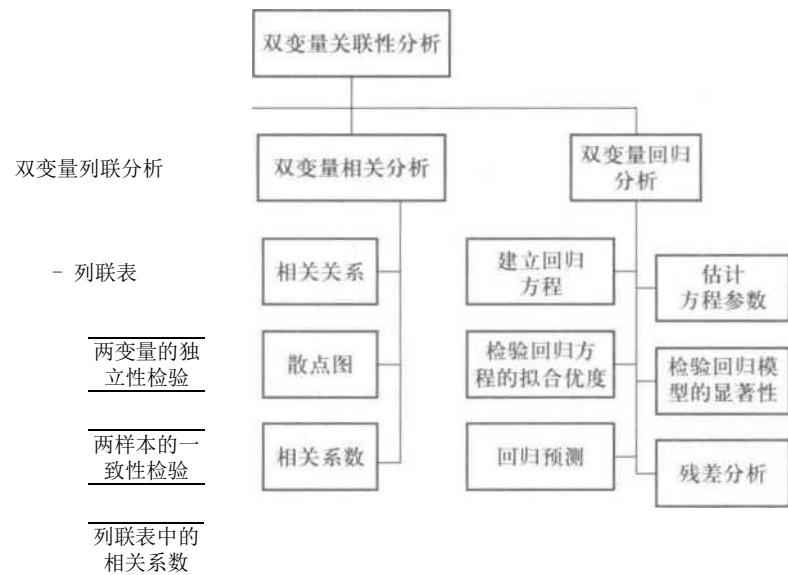
第十一章的引例中介绍了某调查公司进行的婴幼儿奶粉市场研究项目，研究团队计划通过调查分析为某婴幼儿奶粉品牌提供精准的营销沟通支持。他们在研究中设计了与消费基本情况和购买模式相关的调查问题，收集了婴幼儿奶粉消费者用于宝宝奶粉的支出数据和购买频率数据，那么，不同的消费者群体特征与购买频率是否存在某种联系？家庭的基本收入状况是否会影响宝宝的奶粉支出？如果有影响，那么具体是如何影响的？为了解决这些问题，研究团队计划选择适合的双变量分析方法进行探究。

在市场调查中，研究者感兴趣的常常是两个或以上变量之间的关系。比如品牌知名度和销售量有何关系，商店附近的车流量与商店的销售额有无关系，顾客对于某品牌的态度和对于其他品牌的态度有何不同，等等。要回答这一类问题，就需要对调查数据进行双变量统计分析。双变量统计分析方法通常用来探究两个变量之间的关联性，按照变量尺度的不同，需要采用不同的统计分析方法。本章将继续结合“婴幼儿奶粉市场研究项目”案例，讲解常用的列联分析、相关分析和回归分析的基本原理以及实际应用中应该注意的问题，同时介绍上述方法的SPSS软件实现过程。

## ’教学目标与要求

1. 了解与掌握列联表的编制；
2. 掌握列联表中 $\chi^2$ 检验的作用和方法；
3. 掌握两个变量间线性相关关系的测度；
4. 掌握一元线性（包括可线性化的非线性关系）回归分析的步骤，

## 、本章知识结构



# 第一节

## 双变量列联分析

列联分析，是通过原始数据结构，揭示品质型变量（定类变量和定序变量）之间以及品质型变量各种状态之间相关关系的统计分析方法。在引例中，研究团队想要探究不同的消费者群体特征与购买频率是否存在某种联系，例如消费者的年龄特征与购买频率是否有关联？不同年龄阶段消费者的购买频率是否有差异？消费者的受教育程度特征与购买频率是否有关联？这种对于年龄段、受教育程度、购买频率等品质型变量之间的关联分析，需要采用列联分析方法加以探究。

### 一、列联表

列联表（contingency table）是列联分析的必要数据基础，是对两个或两个以上的分类变量进行交叉分类所形成的复合频数分布表。最常用的列联表形式是二维  $r \times c$  列联表，如表 12-1 所示。表中，变量 X 是列变量，类别数为一般表示影响原因的自变量；变量 Y 是行变量，类别数为八一般表示变动结果的因变量；表示在列变量 X 取第 j 个类别的条件下，行变量 Y 取第 i 个类别的观测频数。表中最后一行是每列的总频数，称为行边缘频数；表中最后一列是每行的总频数，称为列边缘频数。列联表可以直观地反映出在变量 Y 的条件下，变量 X 的频数分布情况，以及在变量 X 的条件下，变量 Y 的频数分布情况，每个具体的观察值称为条件频数：

表 12-1 二维  $r \times c$  列联表的一般形式

	X、	1	...	$X_j$		X、	合计
Y		$f_{1j}$	...	A			$\sum_i f_{ij}$
$Y_2$	$h$	$f_{2j}$		A		A	$ZA$
...			...	...	•	...	...
Y	Z.	$f_{aj}$		A			
...		...	...	...	...	...	...
$Y_r$	$z_r$	$f_{rj}$	...	...		L	$\sum_j f_{rj}$
合计	$\sum_i f_{ij}$	$\sum_i f_{i-1j}$	...	$\sum_i f_{i1j}$	...	$\sum_i f_{ij}$	$\sum_i \sum_j f_{ij}$

如果把表 12-1 中的观测频数转换成百分比形式，就得到了频率分布形式的列联表。根据研究问题的需要，可以计算以下三种形式的百分比：

行百分比：指单元格的频数占所在行总观测频数的比重（ $f_{ij}/\sum_j f_{ij}$ ）；

列百分比：指单元格的频数占所在列总观测频数的比重（ $f_{ij}/\sum_i f_{ij}$ ）；

总百分比：指单元格的频数占全部观测总频数的比重（/； / ”  $\sum f_{ij}$  ） 0

基于列联表，可以进行两变量的独立性检验和两样本的一致性检验，两者的检验目的不同，前者用于判断两个分类变量之间是否相互关联，后者是检验不同类别的目标变量之间是否存在显著性差异。两者均采用  $\chi^2$  检验的方法，关于  $\chi^2$  检验的基本原理，请参阅相关数理统计教材。

## 二、两变量的独立性检验

独立性检验（test of independence）是利用样本资料来检验两个变量的独立性，即样本资料所显示的两个分类变量的相关性在总体中是否存在。如果相关性不存在就称为独立。例如对消费品的选择是否与受教育程度或性别有关，企业的创新能力是否与企业规模有关。

独立性检验也需要遵循假设检验的基本步骤，具体包括：

1. 建立原假设和备择假设。原假设为真：两个分类变量之间相互独立（不相关联）；备择假设为行变量和列变量之间不相互独立（相关联）。

2. 计算检验统计量，即  $\chi^2$  值。若用  $\chi^2$  表示观测值频数，根据列联表中的观测值分布计算期望值%的分布，在各单元的期望值均大于 5 的情况下，计算  $\chi^2$  统计量的值：

足，% (12.1)

3. 确定自由度和显著性水平。自由度  $\nu = (r-1) \times (c-1)$ ，其中  $r$  和  $c$  分别为行变量和列变量的分类类别个数。显著性水平  $\alpha$ ，是根据样本列联表对总体规律进行检验时可接受的犯第一类错误的概率大小（原假设为真而拒绝的概率）。

4. 依据检验规则拒绝或接受原假设。当  $\chi^2 > \chi^2_{\alpha/2}$  时，拒绝原假设，说明两个分类变量存在相关性；当时，不拒绝原假设，说明两个分类变量之间相互独立，调查数据显示的变量之间的相关性不显著。

## 三、两样本的一致性检验

一致性检验（test of homogeneity）是对多项总体，即单个总体划分为三个或更多类别后，为了判断样本资料所显示的不同类别的比例差异在总体中是否也存在而进行的检验。例如检验不同受教育程度的消费者对某商品的购买兴趣是否一致。

一致性检验也采用  $\chi^2$  检验的方法，建立的原假设为总体各类别中的比例没有差异，备择假设为总体各类别中的比例存在差异；检验统计量  $\chi^2$  值的计算方法和检验规则与上文讲解的两变量独立性检验相同，不再赘述；若经过检验拒绝原假设，则说明多项总体各类别中的比例存在差异；若不能拒绝原假设，则说明多项总体各类别中的比例没有差异。

## 四、列联表中的相关系数

列联表中的  $\chi^2$  检验已经对两个分类变量之间的相关性给出了统计推断当我们拒绝两

变量间相互独立的原假设时，就表明两个变量间是相关的。不过，列联表中的 V 检验只能说明两个变量之间是否相关，却不能说明相关程度的高低。所以，需要一种尺度，用以测量两个变量是否相关以及相关程度的高低。常用的测量尺度有三种： $\phi$  相关系数、C 相关系数和  $\gamma$  相关系数。

以表 12-2 所示的列联表为例，假设变量 X 与变量 Y 相关，可以给出上述三个相关系数的计算公式。

表 12-2 X 与 Y 相关的列联表

	%	X)	合计
Y	a	b	a+b
	c	d	c+d
合计	a+c	b+d	n

列联表中的相关系数与 V 统计量之间具有确定的联系，因此，可以通过 V 统计量的值来计算列联表的相关系数。由表 12-2 得到 V 统计量的计算公式为：

$$(a+b)(c+d)(a+c)(b+d)$$

$\phi$  相关系数、C 相关系数和 V 相关系数的计算公式、取值范围和适用条件如表 12-3 所示。采用三个相关系数进行分析时，应注意的问题有：

第一，对于  $2 \times 2$  列联表\*相关系数取值范围为  $[0, 1]$ ，但当列联表的行数或列数大于 2 时淬相关系数会随着行数或列数增加而变大，并且没有上限，伊相关系数的含义不易解释。

第二，由 C 相关系数的计算公式可知，当两个变量独立时，C = 0，但即使两个变量完全相关，列联系数也不可能等于 1，对 C 列联系数含义的解释就不够方便。

表 12-3 列联表中相关系数的计算和应用

相关系数	计算公式	取值范围	适用条件	相关程度判断规则
相关系数		Ow   0 1 W 1	适用于 $2 \times 2$ 列联表	
C 相关系数		OwC<1	适用于具有相同行数 和列数的大于 $2 \times 2$ 的列联表	相关系数为 0 时完全不相关； 相关系数为 1 时，完全相关； 相关系数越接近于 1，相关程度越高； 相关系数越接近于 0，相关程度越低。
V 相关系数	0W   W1	任何情况		<p 相关系数的正负 没有意义上的不同，以   见进行判断

## 五、应用实例

### 应用实例 12-1

在引例中提到的婴幼儿奶粉市场研究项目中，研究者计划探究不同的婴幼儿奶粉消费群体特征与购买频率存在何种联系。基于数据集 Example11-1，以年龄特征为例，分析消费者年龄与婴幼儿奶粉购买频率之间的交叉频数分布情况，在此基础上，进一步分析不同年龄阶段的消费者群体在购买频率上是否存在差异。

分析：由于年龄为连续变量，且数据的不同取值个数较多，经常会出现稀疏单元（单元格频数小于5），所以为构建能反映年龄和购买频率之间交叉频数分布情况的列联表，首先需要对年龄进行分组整理，划分为“20-24岁”“25-29岁”“30-34岁”以及“35-45岁”四组；然后，再对分组后的年龄变量和购买频率变量进行交叉分组形成列联表，最后利用 $\chi^2$ 统计量对变量的独立性进行检验。具体实现过程利用SPSS16.0（英文版）软件中的“Crosstabs（交叉表）”菜单实现。

### 1. 操作步骤

(1) 打开数据集 Example 11 - 1，依次选择 Transform（转换）→ Recode into Different Variables（转换到不同的变量），展开变量转换的对话框，如图 12-1 所示。然后，将变量“age（年龄）”选入“Input Variable→ Output Variable（输入变量→输出变量）”框中，在“Name（名称）”和“Label（标签）”框中输入衍生变量的名称和标签，并单击“Change（改变）”按钮。进一步单击“Old and New Values（新旧值）”按钮设置转换规则，如图 12-2 所示。然后，单击“Continue（继续）”按钮，返回到主对话框。单击“OK（继续）”按钮，得到分组后的年龄变量 age”。

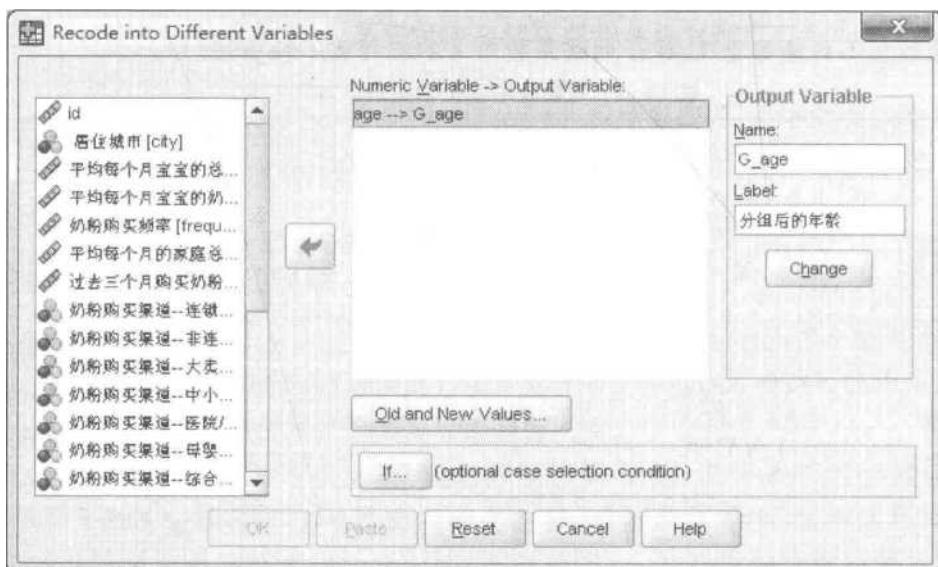


图 12-1 变量转换的对话框

(2) 依次选择“Analyze（分析）→ Descriptive Statistics（描述统计）→ Crosstabs（交叉表）”，展开编制列联表对话框，如图 12-3 所示；然后，将变量“frequency（购买频率）”作为行变量选

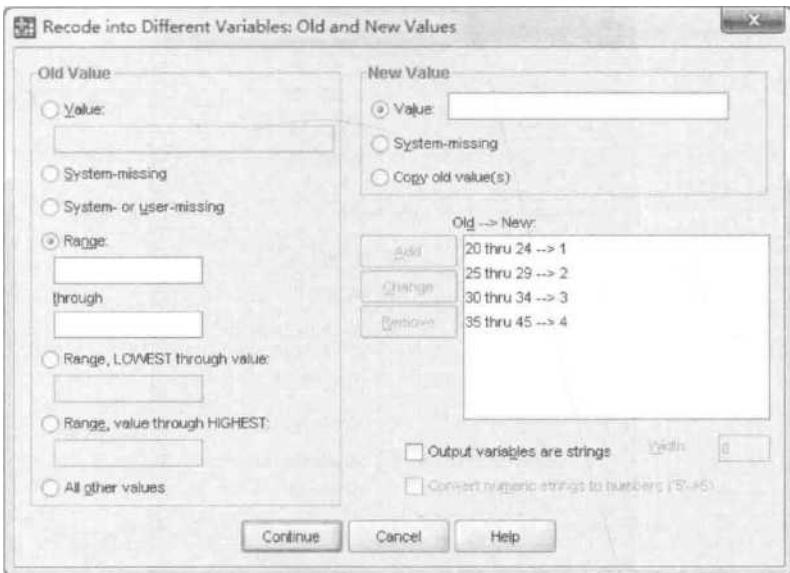


图 12-2 设置变量转换规则的对话框

入 "Row (s) (行)" 框中，将变量 "G\_age (分组后的年龄)" 作为列变量选入 "Column(s) (列)" 框中。



图 12-3 生成列联表的对话框

(3) 单击“Cells (单元格)”项，打开“Crosstabs: Cell Display (交叉表: 单元格显示)”对话框，如图 12-4 所示。在“Percentages (百分比)”框中选择“Row (行)”、“Column (列)”、“Total (总)”复选框，用于指定输出三种形式的百分比；然后，单击“Continue (继续)”按钮，返回到主对话框。

(4) 单击“Statistics (统计量)”项，打开“Crosstabs: Statistics (交叉表: 统计量)”对话框，如

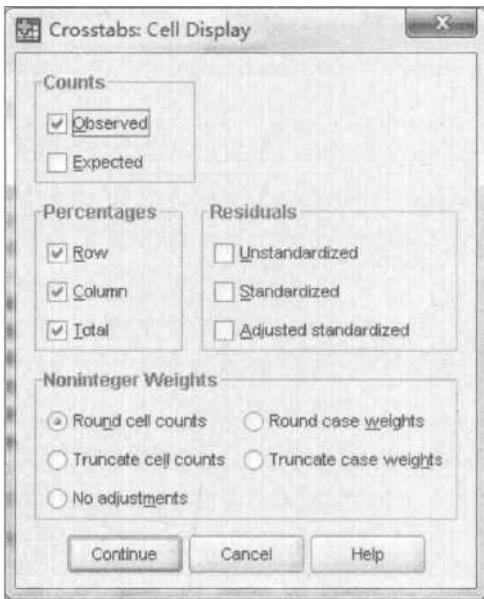
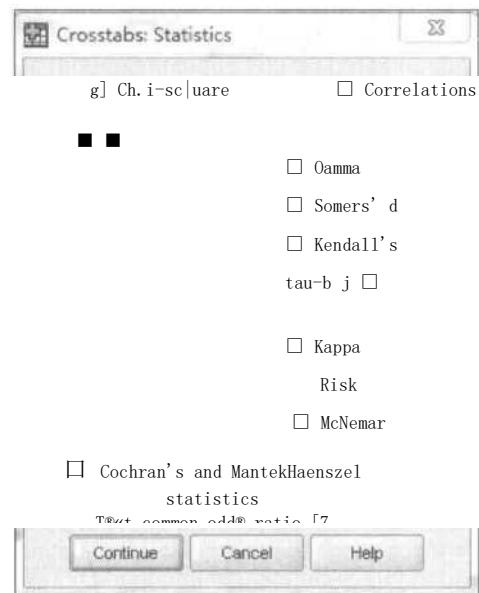


图 12-4 Cells 对话框

图 12-5 所示。选择“Chi-Square（卡方）”复选框，用于指定进行尸检验；在“Nominal（名义）”框中选择 Contingency coefficient（相依系数）和“Phi and Cramer's V（Phi 和 Cramer 变量）”复选框，用于指定输出三种相关系数的计算结果；再单击“Continue（继续）”按钮，返回到主对话框。



Q Contingency coefficient

O Phi and Cramer's V

Lambda

Uncertainty coefficient

r Nominal Interval

Eta

图 12-5 Statistics 对话框

(5) 单击 OK 按钮，即可得到列联表输出结果（见图 12-6）. 列联表的/检验结果（见图 12-7）和列联表的相关系数计算结果（见图 12-8）。

## 2. 解读输出结果

输出结果共包括三部分，如图 12-6、图 12-7 和图 12-8 所示 ‘第一部分是分组后的年龄

奶粉购买频率, 分组后的年龄 Crosstabulation

		分组后的年龄				Total	
		20-24岁	25-29岁	30*34岁	35-45岁		
奶粉购买频率	每周一次	Count	11	45	20	6	82
		% within 奶粉购买频率	13.4%	54.9%	24.4%	7.3%	100.0%
		% within 分组后的年龄	17.5%	23.0%	25.3%	18.8%	22.2%
		% of Total	3.0%	12.2%	5.4%	1.6%	22.2%
	每月 2-3 次	Count	20	29	21	6	76
		% within 奶粉购买频率	26.3%	38.2%	27.6%	7.9%	100.0%
		% within 分组后的年龄	31.7%	14.8%	26.6%	18.8%	20.5%
		% of Total	5.4%	7.8%	5.7%	1.6%	20.5%
	每月 1 次	Count	11	29	13	7	60
		% within 奶粉购买频率	18.3%	48.3%	21.7%	11.7%	100.0%
		% within 分组后的年龄	17.5%	14.8%	16.5%	21.9%	16.2%
		% of Total	3.0%	7.8%	3.5%	1.9%	16.2%
	每 2-3 个月一次	Count	11	50	14	7	82
		% wrthm 奶粉购买频率	13.4%	61.0%	17.1%	8.5%	100.0%
		% within 分组后的年龄	17.6%	25.5%	17.7%	21.9%	22.2%
		% of Total	3.0%	13.5%	3.8%	1.9%	22.2%
	每 4-5 个月一次	Count	10	43	11	6	70
		% within 奶粉购买频率	14.3%	61.4%	15.7%	8.6%	100.0%
		% within 分组后的年龄	15.9%	21.9%	13.9%	18.8%	18.9%
		% of Total	2.7%	11.6%	3.0%	1.6%	18.9%
Total	Count	63	196	79	32	370	
	% within 奶粉购买频率	17.0%	53.0%	21.4%	8.6%	100.0%	
	% within 分组后的年龄	100.0%	100.0%	100.0%	100.0%	100.0%	
	% of Total	17.0%	53.0%	21.4%	8.6%	100.0%	

图 12-6 分组后的年龄与婴幼儿奶粉购买频率列联表结果截图

Chi-Square Tests

	Value	df	Asymp Sig (2-sided)
Pearson Chi-Square	15.220»	12	.230
Likelihood Ratio	15.008	12	.241
Linear-by-Linear Association	125	1	.723
N of Valid Cases	370		

a. 0 cells (.0%) have expected counts less than 5. The minimum expected count is 5.19

图 12-7 列联表的 V 检验结果截图

Symmetric Measures

	Value	Approx. Sig
Nominal by Nominal Phi	.203	.230
Cramer's V	.117	.230
Contingency	.199	.230
N of Valid Cases	370	

图 12-8 列联表的相关系数计算结果截图

和婴幼儿奶粉购买频率的列联表, 行变量是“奶粉购买频率”变量, 列变量是“分组后的年龄”变量。这里以第 1 行第 2 个单元格(即“每周一次”和“25-29岁”交叉的单元格)为例, 说明列联表中各数值的含义: 45 为该单元格的实际频数, 54.9% 为行百分比, 23.0% 为列百分比, 12.2% 为总百分比, 表明每周购买一次婴幼儿奶粉的“25-29岁”消费者共有 45 人, 占全部每周购买一次奶粉的消费者的 54.9%, 占全部“25-29岁”消费者人数的 23.0%, 占全部婴幼儿奶粉消费者的 12.2%。

输出结果的第二部分是变量独立性的/检验结果。可以看出  $15.220/\approx 0.23$  [表中 显示为 Sig. (2-sided)], 因此, 在 0.05 显著性水平下, 没有理由拒绝原假设, 即可以认为年龄和·奶粉购买频率两个变量

具有独立性，即两个变量间没有显著关联。

输出结果的第三个部分是列联表相关系数的计算结果。其中，“Phi”是伊相关系数 Cramer<sup>1</sup> “V”是 V 相关系数，“Contingency Coefficient”是 C 相关系数。由于本例中变量独立性的 / 检验结果显示年龄和购买频率两个变量间没有显著关联，因此无须计算相关系数，从相关系数计算结果的 P 值也可以看出相关系数是不显著的

## 六、进行列联分析时应注意的问题

第一，编制列联表时，如果两个变量之间不存在因果关系，行变量和列变量可以随意指定。如果两个变量之间存在因果关系，一般需要把表示影响原因的自变量作为列变量，列入表的纵列；把表示变动结果的因变量作为行变量，列入表的横行。

第二，列联表只是检验变量之间是否有关联，而并非检验变量之间是否具有因果关系。

第三，进行列联分析的变量必须是取值个数有限的离散变量，如果是连续变量，则要先对变量进行分组整理，将其转换成有限多个取值的离散变量后，再进行列联分析。

第四，使用卡方分布进行独立性检验和一致性检验时，一般要求样本量必须足够大(如 n>50)，每个单元格中的期望频次也不能过少，通常要求每个单元格中的期望频数大于等于 5，若小于 5，则需要将相邻单元格合并。

# 第一 P

## 双变量相关分析

在本章第一节中，我们采用列联分析方法分析了年龄特征和购买频率两个品质型变量之间的关联性，对于宝宝平均每月奶粉支出和家庭平均每月总收入这类数值型变量之间的相关关系却不适用，需要采用相关分析方法和回归分析方法。

两个变量之间的数量关系可以分为函数关系和相关关系两大类 其中，函数关系指变量之间所具有的确定性的对应关系，相关关系指变量之间存在的不能用精确的函数表达式来表示的关系。把握函数关系和相关关系两种内涵的关键在于，函数关系具有确定性，相关关系具有不确定性。

### 一、相关关系

相关分析(correlation)是研究两个变量之间相关关系密切程度的统计方法。在相关关系中，当一个变量的取值发生变化时，另一个变量的取值是不确定的，但是总会遵循某种规律在一定范围内发生变化。如果一个变量的取值整体上是沿着一条直线，随着另一个变量取值的变化而变化，则称两者之间具有线性相关关系(linear correlation)；如果是沿着一条曲线变化，则称两者之间是非线性相关关系(nonlinear correlation)，或称曲线相关关系(curvilinear correlation)。在线性相关关系中，如果两个变量的变动方向相同，即一个变量的数值增大或减少，另一个变量的数值也随之增大或减少，则称之为正线性相关；如果两个变量的变动方向相反，则

称之为负线性相关，如图 12-9 所示。

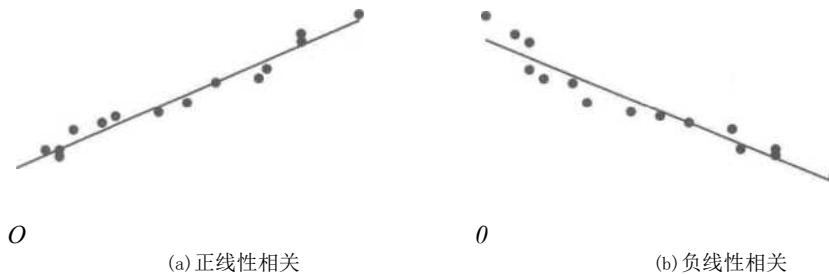


图 12-9 线性相关关系的两种类型

例如，仔细观察导语中的样本数据，宝宝平均每月奶粉支出和家庭平均每月总收入两个变量之间存在着某种关联，即随着收入的增加，奶粉支出相应地也在增加。但是这种关联的具体状态和密切程度等，都无法直接从数字中直接得到答案。通常用来测度变量之间相关关系的工具有散点图和相关系数。

## 二、散点图

散点图(scatter diagram)是以直角坐标系的横轴和纵轴分别代表两个变量，将两个变量间相对应的变量值用坐标点的形式描绘在坐标平面上所形成的图形。利用散点图可以直观地看出两变量之间相关关系的形式、方向及密切程度。散点图是从样本数据上直观判断两个数值型变量之间是否具有相关关系的重要工具。

如图 12-10 所示的散点图，横轴代表家庭平均每月总收入，纵轴代表宝宝平均每月奶粉支出，图中的各个点，由每一个观测中的奶粉支出和家庭总收入的一对观测值所决定。该图直观地显示了数据中 30 个样本的宝宝平均每月奶粉支出会随着家庭平均每月总收入增加而增加的正相关关系。

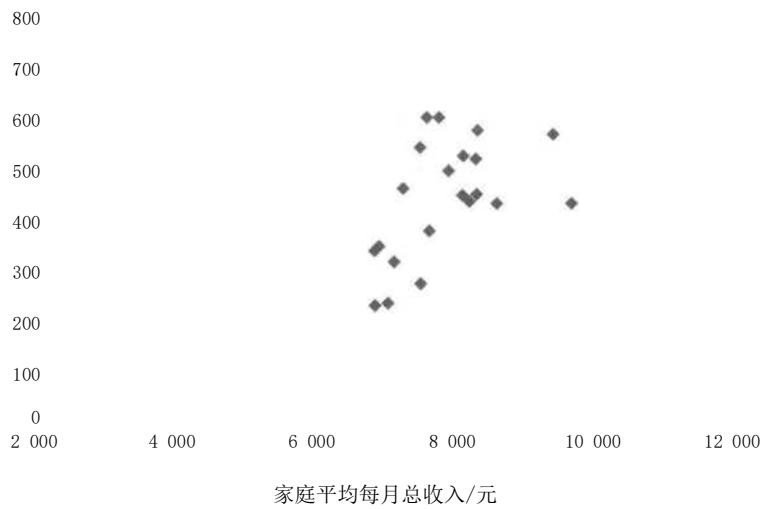


图 12-10 宝宝平均每月奶粉支出与家庭平均每月总收入样本数据散点图



### 三、相关系数

散点图有助于观察和判断两个变量之间相关关系的类型及其密切程度，但判断结果往往 带有很强的主观性。在线性相关的条件下，相关系数① (correlation coefficient) 是度量两个变量 之间相关关系的类型及其密切程度的统计指标，记为  $r$ ，其计算公式为：

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{s_x^2 s_y^2}} \quad (12.3)$$

式中,  $\text{cov}(x, y)$  为协方差,  $s_x$  和  $s_y$  分别为变量  $X$  和变量  $y$  的标准差。

相关系数的取值范围为 “ $-1 < r < 1$ ” 若  $0 < r < 1$ , 表明  $x$  和  $y$  正相关; 若  $-1 < r < 0$ , 表明  $x$  和  $y$  负 相关; 若  $r = 0$ , 表明  $x$  和  $y$  之间不存在任何线性相关关系; 若  $|r| = 1$ , 表明  $x$  和  $y$  完全相关。 $r$  的 绝对值越接近于 1, 表明两变量之间线性相关的程度越大, 当  $|r| > 0.75$  时, 一般可以认为两变 量之间存在高度相关关系。

在现实中, 两变量之间总体上的相关系数的理论真值通常是未知的, 只能根据样本数据计 算出样本相 关系数, 再由样本相关系数在一定把握程度下推断出总体相关系数, 即检验相关系 数的显著性。

对相关系数的显著性进行检验, 需要遵循第十一章第五节中所介绍的显著性检验的基本 步骤, 即: 提出原假设, 设定显著性水平计算检验统计量, 并判断是否拒绝原假设, 进而做出 统计决策。检验相关系数显著性的统计量采用自由度为  $n-2$  的  $t$  统计量:

$$t = \frac{r \sqrt{n-2}}{\sqrt{1-r^2}} \quad (12.4)$$

在显著性水平  $\alpha$  下, 临界值为  $t_{\alpha/2}$ , 若  $|t| > t_{\alpha/2}$ , 表明相关系数  $r$  在统计上是显著的, 也就是说在  $1-\alpha$  的把握程度下, 可以认为总体上两个变量之间是 线性相关的; 否则, 不能说两个变量 之间线性相关。或直接比较  $P$  值与显著性水平  $\alpha$ , 若表明不拒绝原假设; 若则拒绝 原假设。

### 四、应用实例

#### 应用实例 12-2

在婴幼儿奶粉市场研究项目中, 研究者考虑到支出往往随着收入的增加而增加, 同时通过 观察样本数据, 认为宝宝平均每月的奶粉支出会随着家庭平均每月总收入的增加而增加, 即二 者之间存在正线性相关关系。但是这种关联的具体状态和密切程度却无法通过观察数据直接 得到答案, 可以通过计算相关系数来确定。

分析: 由于宝宝平均每月奶粉支出变量和平均每月的家庭总收入变量均为定距变量, 因此 应该选择皮 尔逊相关系数来测度二者之间的相关程度。具体实现过程利用 SPSS16.0 (英文 版) 软件中的 “ Correlate (相关) ” 菜单实现。

①本节所讲的相关系数是用于衡量两个定距变量之间相关程度的统计量, 也称简单相关系数, 或皮尔逊相关系数 (Pearson 相关系数) 此外, 还有适用于衡量两个定序变量之间相关程度的斯皮尔曼相关系数 (Spearman 相关系数), 读者 可参阅相关统计学教材

## I. 操作步骤

- (1) 打开数据集 Example1 - 1。
- (2) 依次选择“Analyze (分析) → Correlate (相关) → Bivariate (双变量)”，展开双变量相关分析对话框，如图 12-11 所示。首先，将变量“m\_expand (平均每个月宝宝的奶粉支出)”和 income (平均每个月的家庭总收入) 同时选入“Variables (变量)”框中。然后，依次勾选“Correlation Coefficients (相关系数)”复选框中的“Pearson”Test of Significance (显著性检验) 复选框中的“Two-tailed (双侧)”和“Flag significant correlations (标记相关的显著性)”。最后，单击 OK 按钮，即可得到宝宝平均每月奶粉支出和平均每月家庭总收入的相关系数 (见图 12-12)。

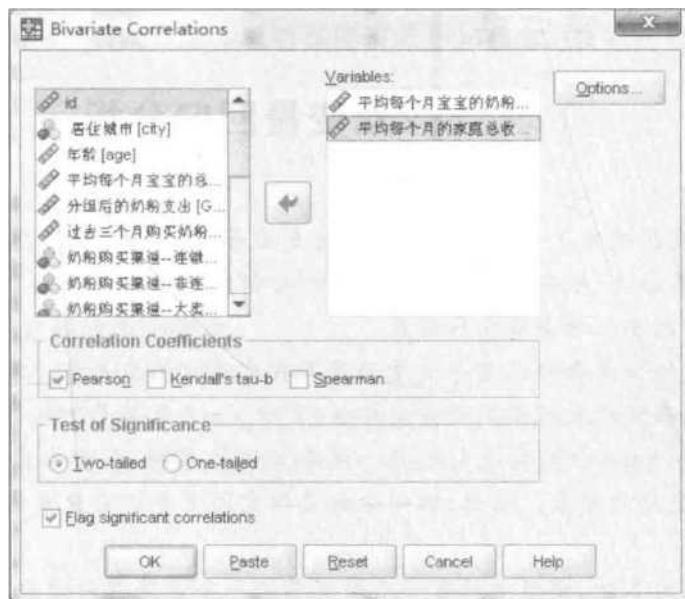


图 12-11 Bivariate Correlations 对话框

Correlations			
平均每个月宝宝的奶粉支...	Pearson Correlation Sig. (2-tailed) N	1 370	.913** .000 370
平均每个月的家庭总收入	Pearson Correlation Sig. (2-tailed) N	.913** .000 370	1 370

\*\*. Correlation is significant at the 0.01 level (2-tailed)

图 12-12 相关系数输出结果截图

## 2. 解读输出结果

相关系数的计算结果以相关系数矩阵的形式输出，并给出了显著性检验的结果。对角线的上下所显示的内容是相同的：可以看出，370 个样本的宝宝平均每月奶粉支出和平均每月家庭总收入的相关系数达 0.913，且  $p = 0.000$  [表中显示为 Sig. (2-sided)]，远小于设定的显著性水平 0.05，可以认为宝宝平均每月奶粉支出和平均每月家庭总收入存在高度正相关关系。

## 五、进行相关分析时应注意的问题

第一，相关分析的目的是测定变量之间相关关系的方向和程度，进行相关分析时，可以不必事先确定哪个是自变量，哪个是因变量，所涉及的变量可以都是随机变量。

第二，进行相关分析时，不能仅凭相关系数  $r$  的大小来解释变量之间的相关程度，否则有可能会得出不切实际的结论。例如， $r = 0$  只表示  $x$  和  $y$  之间不存在线性相关关系，并不能说明  $x$  和  $y$  之间也不存在其他形式的相关关系；用样本数据计算的  $r$  值具有一定的随机性，在实际应用时，需要对其进行显著性检验。

### 第三节

## 双变量回归分析

通过相关分析，我们确定了宝宝平均每月奶粉支出和家庭平均每月总收入之间具有高度线性相关关系，在此基础上，研究者还可以采用回归分析进一步探究家庭平均每月总收入对宝宝平均每月奶粉支出的具体影响方向和程度。

相关分析的目的仅仅是要确认两个变量间是否具有线性相关关系。回归分析的目的则是通过自变量的给定值来估计或预测因变量的均值，其中，一个或若干个起着影响作用的变量称为自变量（independent variable），它是引起另一现象变化的原因；而受自变量影响的变量称为因变量，它是自变量变化的结果。因此，回归分析是研究因变量对自变量依赖关系的一种统计分析方法。

以线性相关关系为基础，模型中只有一个因变量与一个自变量的回归分析，称为简单线性回归分析（或一元线性回归分析）；若模型中有一个因变量和多个自变量，则称为多重线性回归分析（或多元线性回归分析）。本节以一元线性回归分析为例，介绍回归模型的理论基础和回归分析的基本步骤，这在多元线性回归分析中同样适用。

### 一、双变量回归分析的步骤

一元线性回归是最简单的回归形式，它反映的是因变量与一个自变量之间的回归关系。假定自变量为  $x$  与因变量  $y$  在总体上存在着线性相关关系，就可以用下面的等式来模拟  $y$  与  $x$  之间的这种线性相关关系：

$$y = \beta_0 + \beta_1 x + \varepsilon \quad (12.5)$$

该等式即为  $x$  与  $y$  的一元线性回归模型。其中  $\beta_0$  为因变量  $y$  的第  $i$  个取值次为自变量  $x$  的第  $i$  个取值； $\beta_1$  和  $\varepsilon$  分别为线性函数的截距和斜率，称作模型参数； $\varepsilon$  是误差项，反映除  $x$  与  $y$  之间线性关系之外的对  $y$  影响的随机因素。

一元线性回归分析的基本步骤为：“建立回归方程—估计方程参数—检验回归方程的拟合优度—检验回归模型和参数的显著性—>回归预测—残差分析”。其具体步骤如下：

#### （一）建立回归方程

由于存在误差项随机变量  $\varepsilon$ ，给定一个自变量  $x$  的取值，仍然无法确切地估计出因变量  $y$  的相应取值。因此，需要对误差项随机变量  $\varepsilon$  的概率分布做出如下假定： $\varepsilon$  的期望值为零； $\varepsilon$  的方差存在且相等，记为  $\sigma^2$ ； $\varepsilon$  服从正态分布，即  $\varepsilon$  之间相互独立。基于上述假定，将回归模型的等号两边同时取数学期望，即可得到一元线性回归方程：

$$\hat{y} = \beta_0 + \beta_1 x \quad (12.6)$$

其中， $\beta_0$  和  $\beta_1$  均是未知参数， $\beta_0$  为直线在  $y$  轴上的截距， $\beta_1$  为直线的斜率，也称作回归系数，它表示在假定所有其他可能引起因变量变动的变量保持不变的情况下，自变量  $x$  每增加一个单位时因变量  $y$  的平均变动量。

#### （二）估计方程参数

如果们和用是已知的，回归方程就代表一条确定的直线，只要给定自变量、的一个取值 吼，就可依此方程得出对应的义的期望值回归分析可以采用最小二乘法对上述回归 方程的参数伍和 R 进行估计，其基本思想是使误差平方和最小。在此规则下可以得到具体 的参数估计公式为：

$$\hat{a} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (12.7)$$

### (三) 检验回归方程的拟合优度

确定估计的回归方程的目的是要用自变量的变动来解释因变量的变动。显然，所确定的 估计的回归方程与样本数据拟合得越紧密，自变量对因变量的解释能力就越强 c 判定系数就 是用来检验回归方程对样本数据拟合紧密程度的统计量，记作 b，计算公式为：

$$R^2 = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (12.8)$$

其中,SSR 为回归平方和，表示样本中第，次观测的因变量的估计值初与因变量均值 y 之 间的离差平方和； SST 为总平方和，表示样本中第，次观测的因变量的实际值咒与因变量均值 ”之间的离差平方和。

判定系数的取值范围是 0 ~ 1, 当 k=1 时，说明估计的回归方程可以完全拟合所有样 本；如果没有完 全拟合，则朋越接近于 1, 表明估计的回归方程对样本数据的拟合效果越好，越接近于 0 表明拟合效果越差。

### (四) 检验回归模型的显著性

回归分析中的显著性检验包含回归方程的显著性检验和回归系数显著性检验两个部分。在一元线性回 归分析中，回归方程的显著性检验和回归系数显著性检验是等价的，即若回归方 程显著，则在同一显著性 水平下，回归系数也显著，反之结论也成立。

回归方程的显著性检验，采用”检验的方法。在一元线性回归方程中，若斜率用为 0, 则 回归方程代表 一条水平直线，这表明 y, 的期望值 E (力) 不依赖于此的变化而呈线性变化，意 味着总体中尤与 y 之间不 存在线性相关关系。因此，回归方程显著性检验的根本目的是检验 回归方程的斜率，是否显著为 0, 即设定原假设 H<sub>0</sub>: β<sub>1</sub> = 0, 备择假设出：，关 0。检验统计量 为尸，计算公式为：

$$F = \frac{\frac{SSR/l_z}{Y(y_i - \bar{y})^2/1}}{\frac{SSE/(n-2)}{\sum(y_i - \hat{y}_i)^2/(n-2)}} \quad (12.9)$$

其中,  $SSR$  为回归平方和, 表示样本中第  $l_z$  次观测的因变量的估计值与因变量均值之差的离差平方和;  $SSE$  为残差平方和, 表示样本中第  $i$  次观测的因变量的实际值与估计值  $\hat{y}_i$  之间的离差平方和;  $1$  为分子的自由度,  $n-2$  为分母的自由度/样本量。

在显著性水平  $\alpha$  上, 若  $F > F_{\alpha}$ , 则拒绝  $H_0$  的原假设, 这表明可以在  $1-\alpha$  的把握程度上推断总体中  $x$  与  $y$  两个变量之间存在线性相关关系, 即回归方程是显著的; 否则, 没有理由拒绝原假设, 回归方程未通过显著性检验。

需要注意的是, 在多元线性回归分析中, 回归方程的显著性检验和回归系数的显著性检验不是等价的, 需要分别进行检验, 可参看第十三章第二节。

### (五) 回归预测

基于模型假定, 采用最小二乘法由样本数据得到估计的回归方程, 若回归方程在一定的显著性水平下通过  $F$  检验, 接下来即可以运用该回归方程进行统计预测(或称估计)。预测包括点估计和区间估计两种, 点估计即通过自变量的值来预测因变量的均值; 区间估计即在一定的把握程度下, 通过自变量的值估计因变量的均值的置信区间。对于给定的计划, 平均值  $E(y_0)$  在  $1-\alpha$  置信水平下的置信区间为:

$$\overline{\hat{y}_0} \pm \left( \frac{1}{\sqrt{n}} \right) \quad (12.10)$$

$i=1$

### (六) 残差分析

一个估计的回归方程可能拥有令人满意的判定系数, 也可能顺利地通过显著性检验, 但即使如此, 也仍然不能据此直接认定就是一个好的模型。因为无论是估计的回归方程的得出, 还是判定系数的计算, 或是显著性检验, 以及回归预测中的点估计和区间估计, 都是建立在模型遵循假定条件的基础之上, 如果最初的几个模型假定中任意一个是不真实的, 就有理由怀疑该估计回归方程的适用性。因此, 需要证实模型假定的真实性。残差分析是证实模型假定的基本方法, 主要包括残差图、标准化残差图和正态概率图三种方法。

若残差图中的样本点分布在一条水平带中间, 或标准化残差图中大约有 95% 的标准化残差介于  $-2$  和  $+2$  之间, 标准化残差分布直方图呈现一个大体上以 0 为中心的对称的钟形分布, 或正态概率图中有较多的样本点密集地分布在  $45^\circ$  线附近, 则说明该模型满足关于残差的基本假定①。

以上即为进行一元线性回归分析的完整过程。然而, 在实际问题中, 自变量和因变量之间的关系并非都是线性的。当因变量随着自变量的变化而变化, 并且呈现出某种曲线形态时, 称二者之间为非线性关系。对于变量之间的非线性关系, 可以用适当类型的曲线来描述, 也可以通过变量转换将非线性回归方程转换为线性方程后, 再进行求解。表 12-4 列举了对部分非线性回归方程进行线性转换的常用形式 D

① 度差图、标准化残差图和正态概率图的具体绘制方法, 读者可查阅相关的统计学教材, 本节不做详细讲解。

表 12-4 可线性化的非线性模型的常用形式

函数名称	非线性函数	线性化模型	变换规则
指数函数	$y_t = ae^{pt}$	$(\ln y_t) = (\ln a) + pt$	两边取对数
幂函数		$(\ln y_t) = (\ln a) + \beta(\ln x_t)$	两边取对数
对数函数	$y_t = a + \beta \ln x_t$	$y_t = a + p(\ln x_t) + u_t$	变量代换
双曲线函数	$y_t = a + b/x_t$	$y_t = a + b(\ln x_t) + u_t$	变量代换
二次多项式		$y_t = a_0 + a_1 t + a_2 t^2 + u_t$	变量代换
生长曲线函数	$y_t = a e^{kt}$	$\ln(y_t) = (\ln a) + kt + u_t$	等式变换后两边取对数
龚伯斯曲线		$\ln(\ln y_t) = (\ln a) - kt + u_t$	等式变换后两次取对数

## 二、应用实例

### 应用实例 12-3

在应用实例 12-2 中，计算了宝宝平均每月奶粉支出和平均每月家庭总收入之间的相关系数，计算结果显示这两个变量之间存在显著的线性相关关系，因此可以进一步建立回归方程，探究平均每月家庭总收入对宝宝平均每月奶粉支出的影响。

分析：探究平均每月家庭总收入对宝宝平均每月奶粉支出的影响，以宝宝平均每月奶粉支出为因变量，平均每月家庭总收入为自变量，建立一元线性回归方程。具体实现过程利用 SPSS 16.0（英文版）软件中的“Regression（回归）”菜单实现。

#### 1. 操作步骤

(1) 打开数据集 Example1 1-1

(2) 依次选择“Analyze（分析）—Regression（回归）Linear（线性）”，展开回归过程对话框，如图 12-13 所示。然后，将变量“income（平均每个月的家庭总收入）”作为自变量，选入“Independent（s）”框中，将变量“m\_expand（平均每个月宝宝的奶粉支出）”作为因变量，选入“Dependent（s）”框中。由于只有一个自变量的简单线性回归分析，在“Method（方法）”复选框中，选择默认项“Enter（进入）”选项即可。

(3) 点击“Statistics（统计量）”按钮，在如图 12-14 所示的“Linear Regression：Statistics（线性回归统计量）”对话框中，分别勾选“Estimate（估计）”和“Model fit（模型契合度）”选项，用于计算回归模型并对其进行显著性检验；为进行样本异常值检验，可以在该对话框中选择“Casewise diagnostics（个案诊断）”，并在“Outliers outside（离群值）”的参数框中键入 2，设置标准化残差的绝对值大于等于 2 时识别为异常值（也可根据研究需要设置为 1 或 3）点击 Continue（继续），回到主对话框。

(4) 绘制残差图，继续点击“Plots（绘制）”按钮，在如图 12-15 所示的“Linear Regression：Plots（线性回归：图）”对话框。选取变量“\*ZRESID”进入 y 框内，选取变量“DEPENDNT”进入 x 框内，用于输出关于因变量的标准化残差图；选择“Histogram（直方图）”选项，用于输出标准化残差的直方图。



图 12-13 Linear Regression 对话框

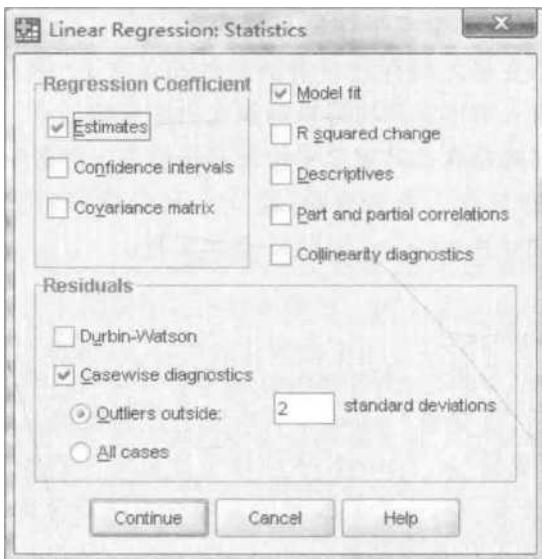


图 12-14 Linear Regression : Statistics 对话框

(5) 进行点估计和区间估计的预测。可以在“Linear Regression（线性回归）”主对话框中点击“SAVE（保存）”按钮，打开“Linear Regression : Save（线性回归：保存）”对话框，如图 12-16 所示。然后在“Predicted Values（预测值）”选项框中，选择“Un standardized（未标准化）”，用于输出未标准化的预测值；在“Prediction Intervals（预测区间）”选项框中，选择“Mean（均值未或 Individual（单值）”选项，并设定置信度，比如 95%，实现对总体均值或个别值的区间估计；在 Residuals（残差）选项框中，选择“Unstandized（未标准化）”或“Standized（标准化）”选项，即可得到残差或标准化残差的计算结果。点击“Conlinue（继续）”按钮，返回到主对话框-同时，上述操作的系统输出结果将在 SPSS 主窗口的数据浏览界面中以新生成的变量形式显示，一

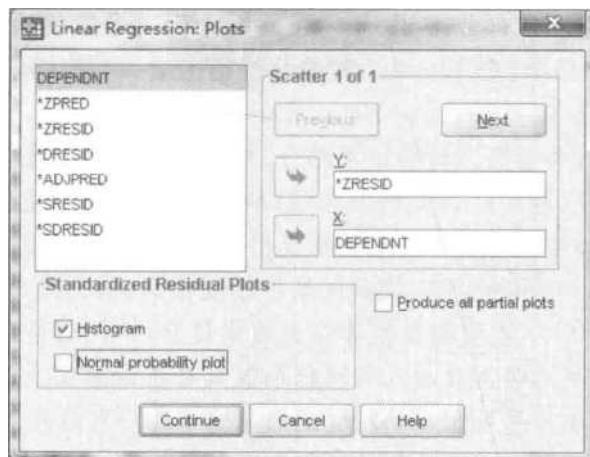


图 12-15 Linear Regression ; Plots 对话框

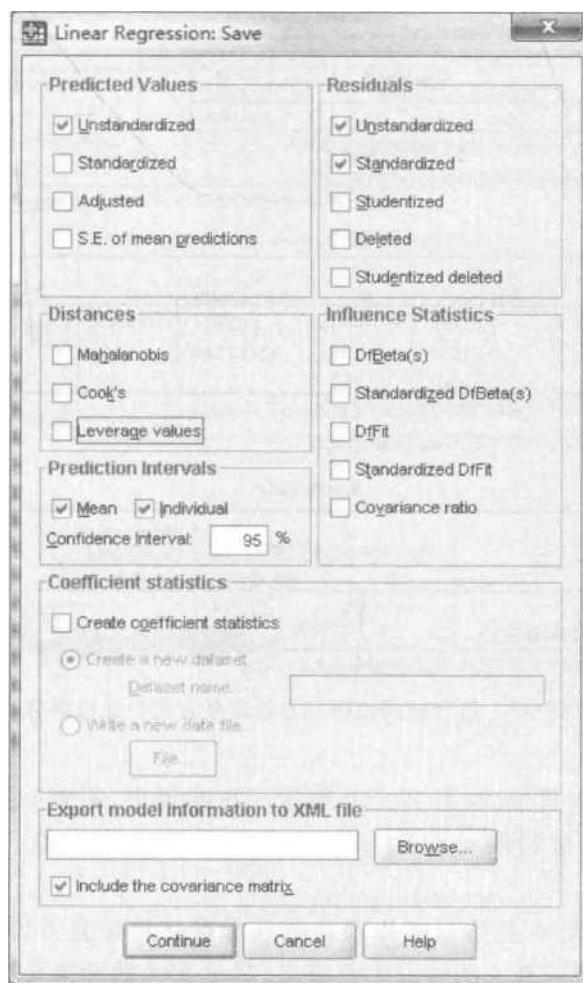


图 12-16 \*\* Linear Regression:Save\*, 对话框

(6)最后，单击OK按钮，即可得到回归过程的全部运行结果(见图 12-17. 至图 12-21)。

## 2. 解读输出结果

回归过程的输出结果由四部分组成，分别是回归方程的参数估计和显著性检验结果(图 12-17). 异常值检测输出结果(图 12-18). 残差分析输出结果(图 12-19 和图 12-20)以及回 归方程的预测结果(图 12-21)。

回归方程的参数估计和显著性检验结果部分，包括回归模型总结、方差分析表以及回归系数和检验结果，如表 12-17 所示。回归模型总结(Model Summary)，也就是对方程拟合情况的 简单描述，从中可以看出，宝宝平均每月的奶粉支出和平均每月家庭的总收入的相关系数 $R^2$ 和 判定系数①? 分别为 91.3% 和 83.3%，模型的解释程度很高；由方差分析表(ANOVA)可知 $F = 1842$ , p 值(Sig.) 小于 0.01, 表明模型整体拟合效果良好；由回归系数及其检验结果(Coefficients) 可知，自变量家庭平均每月总收入的回归系数具有统计意义上的显著性( $t = 42.915$ ,  $p < 0.01$  )。据此可以构造出宝宝平均每月奶粉支出对家庭平均每月总收入的回归方程： $m_{expand} = -94.797 + 0.074 \times income$

式中, $m_{expand}$  为宝宝平均每月奶粉支出 ,  $income$  为家庭平均每月总收入。该方程表明， 家庭平均每月总收入增加 1 元，宝宝平均每月奶粉支出增加 0.074 元。

Model Summary<sup>1\*</sup>

Mode	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.913 <sup>a</sup>	833	833	206.505

a. Predictors: (Constant),平均每个月的家庭总收入 b Dependent Variable,平均每个月宝宝的奶粉支出

ANOVA<sup>1\*</sup>

	Sum of Squares	df	Mean Square	F	Sig.
1 Regression	7854E7	1	7854E7	1842E3	000 <sup>b</sup>
Residual	1.569E7	368	42644.364		
Total	9.423E7	369			

a. Predictors: (Constant),平均每个月的家庭总收入 b Dependent Variable 平均每个月宝宝的奶粉支出

Coefficients<sup>1\*</sup>

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
1 (Constan. 平均每个月的家庭总收入)	-94.797 .074	21.924 .002	.913	-4.324 42.915	000 000

a Dependent Variable 平均每个月宝宝的奶粉支出

图 12-17 估计的线性回归方程及显著性检验结果截图

异常值检测输出结果部分，如图 12-18 所示。在本例中，设置标准化残差的绝对值大于 等于 2 时识别为异常值，检测结果显示共有 15 个异常值， 分别为第 57、67、120、121、181、…、 363 . 367 条观测。

残差分析输出结果部分，包括标准化残差图和标准化残差直方图，分别如图 12-19 和图 12-20 所示。从标准化残差直方图可以大致看出，超过 95% 的标准化残差介于-2 和+2 之间， 结合异常值检测结果可以确定只有 15 个标准化残差的绝对值大于 2, 占总数的 4%；此外，标 准化残差分布直方图呈现为大体上以 0 为 中心的对称的钟形分布。因此，建立的一元线性回

①判定系数表示自变量所能解释的方差占总方差的百分比，判定系数越大，说明模型的拟合效果越好

### Casewise Diagnostics

Case Nu... 57 67 120 121 181 182 235 236 237 238 303 311 312 363 367	Std. Residual 2.119 -9382 2.075 2.519 5.032 4.310 -2.619 3.394 2.633 4.484 2.001 3.957 3.280 2.050 4.385		Predicted Value 1527 2520 1501 1570 2487 2497 1028 2210 1975 2467 1441 2341 2082 1446 2324	Residual 437.530 -1.938E3 428.406 520.229 1 039E3 889 961 -540.901 700.944 543.631 925.977 413.160 817.150 677.320 423.363 905 567
--	---	--	---	---

a Dependent Variable: 平均每个月宝宝的奶粉支出

### Residuals Statistics<sup>1</sup>

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	53.89	4457.53	725.55	461.346	370
Residual	-1.938E3	1039 202	.000	206.225	370
Std Predicted Value	-1.456	8.089	.000	1.000	370
Std Residual	-9.382	5.032	.000	.999	370

a Dependent Variable: 平均每个月宝宝的奶粉支出 图 12-18 异常值检测输出结果截图

归模型满足残差的基本假定。

### Scatterplot

Dependent Variable: 平均每个月宝宝的奶粉支出

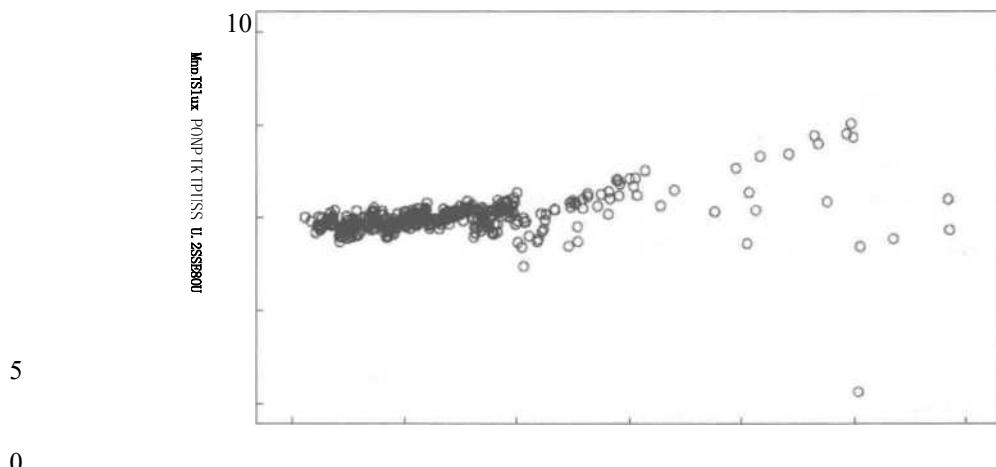


图 12-19 标准化残差图

回归方程的预测结果部分,如图 12-21 所示。在数据集中新生成的变量 PRE\_1、RES\_1、ZRE\_1、LMCI\_1、UMCI\_1、LICI\_1 和 UICI\_1 依次代表点估计预测值、未标准化的残差、标准化的残差、总体均值的区间估计下限、总体均值的区间估计上限、个别值的区间估计下限和个别值的区间估计上限。若需要利用建立的一元线性回归模型预测平均每月家庭总收入为 27 500。

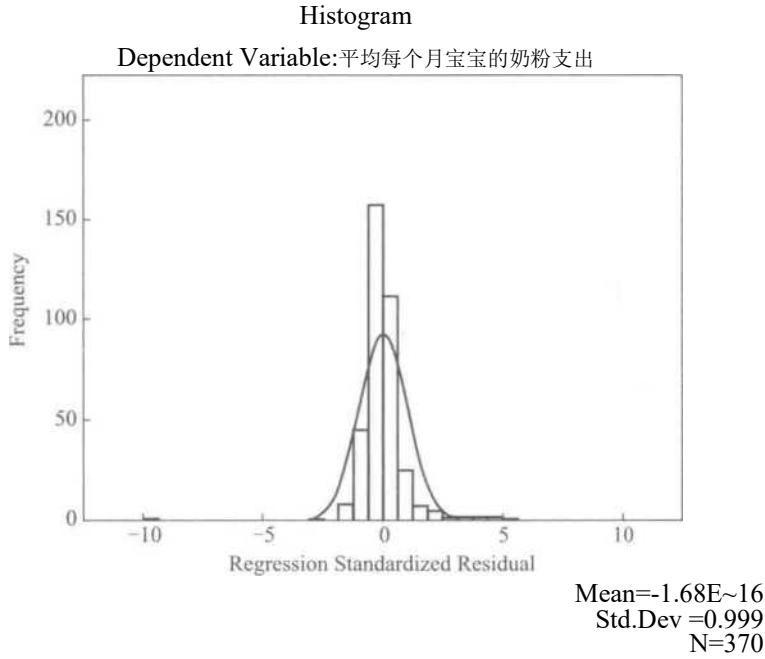


图 12-20 标准化残差直方图

元和 18 950 元的两条新增观测对应的宝宝平均每月奶粉支出，只需在原数据集中新增两条观测，且将平均每月家庭总收入的值填入对应变量“income”中，重新运行该回归过程，即可得到对应的预测值，如图 12-21 所示。

段啓	View	anstorm	Add-2f^					
焰 13 A		、m S M	砒蠱 E 虫或	就 M				
	jIOIW							
359	13465	660	906 21700	-46 21700	-0 22381	883 54CB1	928 89319	499 50623
360	13899	81B	938 46141	-120 48141	-058343	915.22472	961 73810	531 73786
361	14643	877	993 79183	-116 79183	-0 56555	969 36329	1018 22037	586 97959
362	16623	688	105921275	-171,21275	#0.8292	103334685	108527866	652 29691
363	15031	1446	1022 63651	423 35349	2 05014	997 51696	1047 75606	615 78220
364	16390	1409	1123.66723	2B533277	130172	109576626	1151 56820	716S3172
365	17063	1519	117369937	345 30063	167212	114424833	1203 15041	766 55467
368	1B517	1003	12B1.79258	-278 79258	-1 35005	1248 69702	131498614	874 36804
367	20365	2324	1418 43310	905 56690	4 30520	1380 30591	1456 56029	101066900
386	31000	2381	220980326	171 19674	0 82902	2138 59117	2281 01535	1797 62835
369	39652	2528	2953 009®	-325 00989	-1 57366	27S3 26654	2952 75325	2434 66138
37Q	42563	2929	3069 41934	-140 41934	-0 67998	2959 96430	3178 87430	2648 64848
371	27500		1949,60640			1889 67666	2009 53615	1539 12983
372	18950		1313 9B265			1279 73821	1348 22708	906 46317

图 12-21 回归方程的预测结果截图

### 三、进行回归分析时应注意的问题

第一，与相关分析不同，进行回归分析必须事先确定自变量和因变量，其中，自变量是事先 给定的非随机变量，因变量是随机变量。

第二，在进行一元线性回归分析前，应注意通过散点图检验是否存在异常值。所谓异常 值，是指数据集中过大或过小的观测值。异常值的存在，对于回归直线方程的拟合、判定系数 及显著性检验的结果都将产生很大影响。所以，进行回归分析时，首先需要检测数据集中是否 存在异常值。

第三，在实际操作中，建立回归模型的过程非常复杂，变量之间的关系很难确切估计，有时

甚至都无法找到哪些变量可以对因变量的变动形成统计意义上的显著性影响，而且实际被广泛应用的回归分析技术也要比这里介绍的内容复杂得多，应用时必须结合具体情况进 行探讨 和分析。

## 本章小结••••

双变量数据分析的目的在于探索两个变量之间的关系，按照变量尺度的不同，需要采用不同的统计分析方法。研究两个品质型变量（即定类变量和定序变量）之间的关系，可以采用列联分析；研究两个数值型变量（即定距变量和定比变量）之间的关系，可以采用相关分析和回归分析因此，在选择分析方法时，应首先确定变量的尺度。

列联分析是通过原始数据结构，揭示品质型变量（定类变量和定序变量）之间以及品质型变量各种状态之间的相关关系的统计分析方法列联表是列联分析的必要数据基础，它是由两个分类变量进行交叉分类的频数分布表。列联表的  $\chi^2$  检验主要包括独立性检验和一致性检验，前者是利用样本资料来检验其所显示的两个分类变量的相关性在总体中是否存在；后者是对多项总体，即单个总体划分为三个或更多类别后，检验样本资料所显示的不同类别的比例差异在总体中是否也存在。如果独立性检验结果显示两个分类变量之间存在相关性，为了进一步测度两个分类变量之间的相关程度，通常要计算相关系数，具体包括  $w$  相关系数、 $C$  相关系数和  $V$  相关系数。列联表的类型不同，适合采用的相关系数也有所不同。

相关分析是研究两个变量之间相关关系密切程度的统计方法。相关关系按表现形式不同，可以分为线性相关和非线性相关（也称曲线相关）；按相关的方向不同，可以分为正相关和负相关。通常用来测度变量之间相关关系的工具有散点图和相关系数。散点图是以直角坐标系的横轴和纵轴分别代表两个变量，将两个变量间相对应的变量值用坐标点的形式描绘在坐标平面上所形成的图形。利用散点图可以直观地看出两变量之间相关关系的形式、方向及密切程度。散点图的判断结果往往带有很强的主观性，在线性相关的条件下，相关系数是度量两个变量之间相关关系的类型及其密切程度的统计指标，在现实中，若需要根据样本相关系数在一定把握程度下推断出总体相关系数，则需要检验相关系数的显著性。

相关分析的目的仅仅是要确认两个变量间是否具有线性相关关系。回归分析是研究因变量对自变量依赖关系的一种统计分析方法，其目的则是通过自变量的给定值来估计或预测因变量的均值。以线性相关关系为基础，模型中只有一个因变量与一个自变量的回归分析，称为简单线性回归分析（或一元线性回归分析）。一元线性回归分析的基本步骤为：“建立回归方程  $\rightarrow$  估计方程参数  $\rightarrow$  检验回归方程的拟合优度  $\rightarrow$  检验回归模型和参数的显著性  $\rightarrow$  回归预测  $\rightarrow$  残差分析”。在实际问题中，自变量和因变量之间的关系并非都是线性的。当因变量与自变量之间为非线性关系时，可以用适当类型的曲线来描述，也可以通过变量转换将非线性回归方程转换为线性方程后，再进行求解 磁关键名词…

列联表 独立性检验 一致性检验 夕相关系数 C 相关系数 / 相关系数 散点图 相关系数 回归分析 一元线性回归

## 肯思考题

1. 列联表中的行变量与列变量位置可以互换吗？为什么？
2. 简述独立性检验和一致性检验的意义。
3. 列联表中的相关系数包括几种？各自的计算方法、取值范围和适用条件是什么？
4. 如何判断两个变量之间是否具有相关关系？
5. 一元线性回归分析的基本假定有哪些？
6. 简述一元线性回归分析的步骤，，



即测即评

## 可 •本章实训…

### 一、 实训目的

1. 明确双变量相关分析和回归分析的基本原理和分析步骤
2. 掌握应用 SPSS 软件进行双变量相关分析和回归分析的操作方法。

### 二、 实训内容

#### 1. 数据准备

StopGo 公司最近致力于测算特殊地段车流量对坐落于该地的商店年销售量的影响。为了控制其他变量，调查员选定了 20 家商店。这些商店的共同之处是，显著影响商店销售量 的其他变量（如商店面积、停车量、周围居民点的人口统计特征等）是完全相同的。该项目是 Stop'NG。公司致力于识别和量化影响商店销售量因素的整体项目中的一部分。公司的最终 目的是构建出一个用来筛选潜在店址的模型，考虑实际购买情况与店铺建设，从中选择能够带 来最多销售额的店址。

在确定了 20 个商店以后，Stop'NG 公司在长达 30 天的时间内，每天到每个观测地点做 车流量记录。同时，公司通过内部记录获得了这 20 家店前 12 个月的全部销售数据（见数据集 Example 12-1）。（资料来源：（美）小卡尔•麦克丹尼尔（Carl McDaniel, Jr.），罗杰盖茨（Roger Gates）.当代市场调研[M].李桂华等译.8 版.北京：机械工业出版社, 2015.）

#### 2. 具体要求

应用数据集 Example12-1, 选择合适的方法完成车流量和商店年销售量之间的关联性分 析。（注意：确保分析过程的完整性。）

### 三、 实训组织

1. 以个人为单位完成实训要求。
2. 提交数据分析书面报告 ■
3. 安排部分个人进行分析报告演示。
4. 教师批改报告，点评与总结。

# 第十三章 多变量数据分析 (1)

## 、导语

随着打车软件的兴盛以及微信红包的横空出世，网上支付不再是阿里巴巴一家独大，BAT（百度、阿里巴巴和腾讯）三足鼎立的格局已经形成，特别是在线上移动支付方面，阿里巴巴有支付宝，腾讯有微信支付，百度则有百度钱包。为了解客户对BAT移动支付的认知过程，对武汉市高校中使用移动支付软件的人群进行问卷调查①。

问卷分为两个部分，第一部分为基本信息及使用情况调查，包括被调查者的个人信息，如性别、受教育程度、月可支配金额，以及被调查者对BAT移动支付的基本使用情况，如最常用哪种移动支付软件、使用频率、使用场景等。第二部分采集的是用户对BAT移动支付使用意愿，包括BAT认知、感知利益、感知风险、个人创新力、方便性、安全性、娱乐性、使用意愿、是否投资理财等信息。

### BAT移动支付的用户使用意愿调查问卷

亲爱的朋友：

您好！我们是中南财经政法大学的学生，正在对高校大学生使用BAT移动支付的情况进行研究。您的个人信息仅用于汇总数据处理，绝不对外公开，请放心填写，非常感谢您的支持与合作！

.注：本问卷中，BAT移动支付主要是指百度、阿里巴巴和腾讯的用户使用移动终端设备（通常是手机）通过移动支付平台（百度钱包、支付宝和微信支付），实现货币支付与资金转移的行为。

问卷填写方式：请在您认为合适的选项下打钩（V）或直接填写。

#### 基本信息

1. 您的性别\_\_\_\_\_o

A.男 B.女

2. 您的学历\_\_\_\_\_o

A.专科及以下 B.大学本科 C.硕士研究生 D.博士研究生

3. 您每月平均可支配生活费金额\_\_\_\_\_。

A. 600 元以下 B. 600~1 200 元 C. 1200~1 800 元 D. 1 800~2 400 元 E. 2 400 元以上

① 资料来源于：2015年中南财经政法大学专业硕士研究生案例大赛参赛作品《BAT移动支付使用意愿影响因素分析——以武汉高校大学生为例》，小组成员：王亚琪、郭晓菲、徐婷婷、田晓春，指导教师：徐映梅。

4. 您目前使用的 BAT 移动支付软件是 \_\_\_\_\_ (可多选)。  
A. 百度钱包      B. 支付宝  
C. 微信支付      D. 以上三类均未使用过 (选 D 请跳至第 8 题)
5. 您每月平均使用 BAT 移动支付的频率为 \_\_\_\_\_ o  
A. 1~2 次      B. 3~5 次      C. 6~10 次      D. 11~15 次以上  
E. 16~20 次      F. 20 次以上
6. 每笔消费超过多少金额时, 您不会选择 BAT 移动支付 \_\_\_\_\_  
A. 500 元以下      B. 500~1 000 元      C. 1 000~1 500 元      D. 1 500~2 000 元      E. 2 000 元以上
7. 您使用 BAT 移动支付的场景和用途包括 \_\_\_\_\_ (可多选)。  
A. 缴费充值      B. 网上购物      C. 转账汇款      D. 信用卡还款  
E. 打车      F. 红包      G. 投资理财      H. 票务支付  
I. 团购订餐      J. 商店超市付款      K. 其他 \_\_\_\_\_
8. 个人创新性
- 8.1 我乐于接受新观点和新事物 \_\_\_\_\_ Q  
A. 强烈不认同      B. 不认同      C. 不确定      D. 认同      E. 强烈认同
- 8.2 我通常比周围人更先使用新产品或新技术 \_\_\_\_\_ 0  
A. 强烈不认同      B. 不认同      C. 不确定      D. 认同      E. 强烈认同
- 8.3 我通常比周围人对新产品或新技术有更多的了解 \_\_\_\_\_ —  
A. 强烈不认同      B. 不认同      C. 不确定      D. 认同      E. 强烈认同
9. 移动支付认知
- 9.1 我经常使用第三方移动支付平台 \_\_\_\_\_ O  
A. 强烈不认同      B. 不认同      C. 不确定      D. 认同      E. 强烈认同
- 9.2 在购买产品或服务时我总是优先考虑 BAT 移动支付 \_\_\_\_\_ .  
A. 强烈不认同      B. 不认同      C. 不确定      D. 认同      E. 强烈认同
- 9.3 我可以很快熟悉并掌握 BAT 移动支付的相关用途—.  
A. 强烈不认同      B. 不认同      C. 不确定      D. 认同      E. 强烈认同
10. 方便性
- 10.1 我觉得 BAT 移动支付不受地点的约束 \_\_\_\_\_ o  
A. 强烈不认同      B. 不认同      C. 不确定      D. 认同      E. 强烈认同
- 10.2 我觉得使用手机进行支付不会影响手机提供通讯等其他服务 \_\_\_\_\_  
A. 强烈不认同      B. 不认同      C. 不确定      D. 认同      E. 强烈认同
- 10.3 我觉得 BAT 移动支付操作步骤简单 \_\_\_\_\_ □  
A. 强烈不认同      B. 不认同      C. 不确定      D. 认同      E. 强烈认同
11. 安全性
- 11.1 支付系统显示的支付金额与交易数据通常很准确—  
A. 强烈不认同      B. 不认同      C. 不确定      D. 认同      E. 强烈认同
- 11.2 BAT 移动支付系统经常通过一些途径来进行安全阐述 \_\_\_\_\_  
A. 强烈不认同      B. 不认同      C. 不确定      D. 认同      E. 强烈认同
- 11.3 使用 BAT 移动支付时很少会发生意外状况或者错误—  
A. 强烈不认同      B. 不认同      C. 不确定      D. 认同      E. 强烈认同
12. 娱乐性
- 12.1 我觉得使用 BAT 移动支付的过程是愉悦的一  
A. 强烈不认同      B. 不认同      C. 不确定      D. 认同      E. 强烈认同
- 12.2 BAT 移动支付的支付场景是有趣的 \_\_\_\_\_ 0

A.强烈不认同	B.不认同	C.不确定	D.认同	E.强烈认同
---------	-------	-------	------	--------

13. 感知利益

13.1 我觉得 BAT 移动支付使购买产品或服务更加便捷。A.强烈不认同\_\_\_\_\_o\_\_\_\_\_

B.不认同	C.不确定	D.认同	E.强烈认同
-------	-------	------	--------

13.2 使用 BAT 移动支付能提高我的支付效率\_\_\_\_\_o\_\_\_\_\_。

A.强烈不认同	B.不认同	C.不确定	D.认同	E.强烈认同
---------	-------	-------	------	--------

13.3 我觉得使用 BAT 移动支付可以带来一定的优惠\_\_\_\_\_o\_\_\_\_\_。

A.强烈不认同	B.不认同	C.不确定	D.认同	E.强烈认同
---------	-------	-------	------	--------

14. 感知风险

14.1 我觉得使用 BAT 移动支付不会有很大的风险\_\_\_\_\_o\_\_\_\_\_ A.\_\_\_\_\_o\_\_\_\_\_

强烈不认同	B.不认同	C.不确定	D.认同	E.强烈认同
-------	-------	-------	------	--------

14.2 我觉得使用 BAT 移动支付不会有很大的潜在损失\_\_\_\_\_o\_\_\_\_\_。

A.强烈不认同	B.不认同	C.不确定	D.认同	E.强烈认同
---------	-------	-------	------	--------

15. 使用意愿

15.1 我打算开始使用或继续使用 BAT 移动支付\_\_\_\_\_o\_\_\_\_\_

A.强烈不认同	B.不认同	C.不确定	D.认同	E.强烈认同
---------	-------	-------	------	--------

15.2 我愿意向身边的朋友推荐 BAT 移动支付\_\_\_\_\_o\_\_\_\_\_。

A.强烈不认同	B.不认同	C.不确定	D.认同	E.强烈认同
---------	-------	-------	------	--------

在问卷调查中，若是对某种特质、态度、行为或心理知觉等潜在概念的调查，在分析时不能逐题分析，因为单一题项所要测量的信息不足以代表某一潜在特质或概念，此种潜在特质或概念通常包含数个题项，这些属性相似的题项所要测量的共同特质为建构效度，建构效度中的层面测量值是数个题项的加总分数①。在本导语中，问卷各层面信度和效度通过检验，将涉及的各影响因子变量的题项进行简单加总得分，得分越高，表明此特质水平越高②。

根据问卷收集的数据，整理出相关变量信息（见数据集 Example 13-1）。图 13-1 给出 SPSS 格式数据部分截图：

	sex	degree	bat	pbe	prs	inn	con	sec	ent	wil	tzlc	consu	freq
1	女	专科及以下	13	13	11	9	10	9	11	8	进行投资理财	5	2
2	女	大学本科	15	15	15	10	15	7	15	10	进行投资理财	4	2
3	男	硕士研究生	8	10	10	6	9	11	9	7	没有投资理财	2	2
4	男	大学本科	14	13	14	9	12	8	14	10	没有投资理财	4	3
5	男	大学本科	14	14	13	10	13	6	14	10	没有投资理财	4	3
6	男	硕士研究生	14	14	13	9	13	7	13	9	进行投资理财	4	3
7	女	大学本科	13	12	13	10	13	8	13	8	进行投资理财	5	2
8	男	专科及以下	13	13	11	8	7	8	10	8	没有投资理财	4	3
9	男	大学本科	13	13	14	9	13	7	13	9	没有投资理财	3	4
10	女	大学本科	10	11	10	9	11	9	11	7	进行投资理财	3	3
11	女	硕士研究生	14	14	13	9	14	7	14	10	进行投资理财	3	5
12	男	硕士研究生	6	7	5	5	4	8	6	5	没有投资理财	3	3
13	男	大学本科	14	15	13	8	13	7	13	10	没有投资理财	4	4
14	女	大学本科	15	9	12	10	15	11	15	8	没有投资理财	5	4
15	女	大学本科	13	14	13	9	14	8	14	9	没有投资理财	4	2
16	女	大学本科	12	12	11	8	10	9	10	9	进行投资理财	3	5
17	男	大学本科	13	12	13	7	14	9	13	9	没有投资理财	2	2
18	男	大学本科	13	14	14	9	14	7	13	10	进行投资理财	2	2
19	男	大学本科	13	14	12	9	13	6	13	10	没有投资理财	3	2

图 13-1 武汉市高校学生对 BAT 认知及使用意愿数据集 Example 13-1 截图

① 吴明隆, 问卷统计分析实务: SPSS 操作与应用[M]. 重庆大学出版社, 2015.

② 因为得分的数值范围较大, 故本书在实际分析过程中将其作为连续变量处理。但是此做法存在争议, 本书只是为了方便就此数据, 展现相关方法的分析过程,

该数据集包含 13 个变量，343 条观测，其中，背景变量包括“性别(sex)”“受教育程度(degree)”以及“消费水平等级”。变量说明如表 13-1 所示。

表 13-1 BAT 移动支付案例变量说明表

变量说明	意义
性别(sex)	1 =男; 2=女
受教育程度(degree)	1=专科及以下; 2 =大学本科;3 =硕士研究生;4 =博士研究生
消费水平等级(consu)	1 = 600 元以下; 2 = 600~1 200 元; 3=1 200- 1 800 元;4=1 800~2 400 元; 5 = 2 400 元以上
BAT 认知 (bat)	数值越大，表明认知程度越高
感知利益(pbe)	数值越大，表明对利益的感知越高
感知风险(prs)	数值越大，表明对风险的感知越高
个人创新(inn)	数值越大，表明个人创新能力越强
方便性(con)	数值越大，表明受访者认为 BAT 的方便性越好
安全性(sec)	数值越大，表明受访者认为 BAT 的安全性越高
娱乐性(ent)	数值越大，表明受访者认为 BAT 的娱乐性越强
使用意愿(wil)	数值越大，表明受访者使用意愿越强
投资理财(tzlc)	1=进行投资理财;0 =没有投资理财
使用频率(freq)	1 = 1-2 次; 2=3~5 次; 3 = 6~10 次; 4=11~15 次; 5 = 16~20 次; 6 =月均 20 次以上

第十一章和第十二章分别介绍了关于单变量和双变量的统计分析方法，但在实际问题中，为反映现象的全貌，涉及的变量往往不止一个，而是多个变量，并且这些变量之间又存在一定的联系，那么涉及多个变量的情况又该如何进行分析呢？结合本章导语，如我们想了解不同受教育程度学生在 BAT 移动支付“安全性”“便捷性”与“娱乐性”功能层面的知觉是否存在显著差异，显然，问题中涉及的自变量“受教育程度”为四分类别变量，其水平数值在三个以上，不能采用独立样本 t 检验方法，此时，应采用方差分析方法；又如，影响用户使用 BAT 支付意愿的因素不止一个，如何分析多个变量对支付意愿的影响呢？上述问题的解决需要引入多变量统计分析方法。多变量统计是分析多个变量在不同取值情况下的数据分布状况以及不同变量之间密切程度的一类统计分析方法。常用的多变量统计方法有方差分析、相关与回归分析、聚类分析、因子分析等。本章将从实例出发，主要介绍多变量方差分析、相关与回归分析、分类模型—Logistic 模型的基本原理以及实际应用中应注意的问题，并同时给出上述分析技术中常用方法的 SPSS 软件实现过程。

## 、教学目标与要求

1. 了解与掌握多变量方差分析的适用条件、原理及方法；
2. 了解相关分析与回归分析的联系和区别；

3. 掌握线性回归分析的适用条件、应用过程与结果解释；
4. 理解分类模型的适用条件；
5. 熟悉掌握 Logistic 模型的原理与应用；

## 、本章知识结构



# Afr 多变量方差分析 4- 第一 T

在方差分析(Analysis of Variance, ANONA)中，一般把因变量称为观测变量，把自变量或 控制变量称为因素(factor)。方差分析是推断各种因素状态对观测变量影响效应的一种统计 分析方法，它通过分析数据中不同来源的变异对总变异的贡献大小，来确定可控因素对研究结 果影响力的大小。在方差分析中，按照观测变量的个数，可分为单变量方差分析和多变量方差 分析；按照涉及的因素个数划分，可分为单因 素方差分析和多因素方差分析。限于篇幅，本节 将只介绍单因素方差分析方法的基本思路及其应用。

## 一、单因素单变量方差分析

单因素方差分析是在控制变量的不同水平下，检验各总体分布是否存在显著性差异，进而判断 控制变 量是否对观测变量具有显著性影响的分析方法。单因素方差分析的一般步骤如下。

(1) 明确因变量和自变量，建立原假设。设因变量为 V, 自变量为  $X_i$  有 r 个水平，在  $X_i$  的第 i 个水 平下， $r_i$  的观测值的个数为 “ “则：

$$\text{总样本量: } n = \sum_{i=1}^r n_i$$

$$\text{原假设 } H_0: \mu_1 = \mu_2 = \dots = \mu_r$$

(2) 分别计算总方差 (SS7)、组间方差 (SS4)、组内方差 (SSE) 及 F 统计量，具体公式 如下：

$$SST = \sum_{i=1}^r \sum_{j=1}^{n_i} (y_{ij} - \bar{y})^2 \quad (13.1)$$

$$SSA = \sum_{i=1}^r n_i (\bar{y}_i - \bar{y})^2 \quad (13.2)$$

$$SSE = \sum_{i=1}^r \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2 \quad (13.3)$$

$$\begin{aligned} S5A &= (r-1) \\ &\sim SSE/(n-r) \end{aligned} \quad (13.4)$$

(3) 建立单因素单变量方差分析表(见表 13-2)，并查临界值  $F_a(r-1, n-r)$ 。

表 13-2 单因素方差分析表

方差来源	平方和	自由度	方差	F 值	P 值
组间(因素影响)	SSA	r-1	SSA/(r-1)	F	P
组内(误差)	SSE	n-r	SSE/(n-r)		
总和	SST	n-1			

(4) 分析结果。如果  $F > F_a(r-1, n-r)$ ，或 P 值的对应概率 p 小于给定的显著性水平”，则 拒绝原假设，这 时可以认为不同水平下各总体均值之间存在显著差异；如果尸值较小，或其对 应概率”大于显著性水平。， 则不能拒绝原假设，即认为不同水平下各总体均值之间没有显著

差异。

## 二、单因素多变量方差分析

单因素多变量方差分析适用于一个因素、两个以上观测变量的检验，其中观测变量应为连续变量，控制变量为类别变量（间断变量）。在单变量方差分析（univariate analysis of variance）中只检验因素的各个水平在单一观测变量测量值平均数上的差异，使用的检验为 F 检验，而多变量方差分析（multivariate analysis of variance；简称 MANOVA）则同时检验因素的各个水平在两个以上观测变量是否有显著差异，常用的 MANOVA 显著性检验方法为 Wilks A 值，其值介于 0 至 1 之间，Wilks A 越小，表明控制变量对观测变量的影响越显著①。

单因素多变量方差分析的一般步骤如下：

(1) 明确因变量和自变量，建立原假设。

设有左个 P 元正态整体  $N(p_1, \dots, p_k, x; \Sigma)$  ( $\ell = 1, 2, \dots, n_1, n_2, \dots, n_Q$ ) 是来自  $M_S$  ……肉的样本，记第  $\ell$  个 P 元总体矩阵为：

$$\begin{bmatrix} & & & \cdots \\ & & & \ddots \\ & & & \Sigma_{\ell} \\ & & (4) & x_{n,2} \\ x_{n,j} & & & (1) \\ & & & \vdots \end{bmatrix}$$

其中  $Q=1, 2, \dots, n$

原假设  $H_0$  (2) 二 … 小(1)。

(2) 分别计算总离差矩阵组内离差矩阵组间离差矩阵 B 及 Wilks A 统计量，具体公式如下：

如时 (13.5)

$r(i) = X_i - \bar{X}$ , (13.6)

$$B = \sum_{i=1}^n (r(i) - \bar{X})(r(i) - \bar{X})' \quad (13.7)$$

HI

(13.8)

成

(3) 建立多变量方差分析表（见表 13-3），并查临界值  $A_{Ap,m}$  ( $A$  服从威尔克斯分布，可用 F 分布或/分布近似二关于其关系，如果读者感兴趣可参阅统计学相关教材）。

表 13-3 单因素多变量方差分析表

方差来源	平方和	自由度	A 值	P 值
组内（误差）	A	$n-k$	A	P
组间（因素影响）	B	$k-l$		
总和	T	$l-1$		

① 读者应该注意：在单变量方差分析中 A 值越大表示越容易达到显著水平，关于单变量方差分析以及多变量方差分析中推导的具体介绍，读者可参阅：韩明。应用多元统计分析 [M] 上海：同济大学出版社，2013：133-139。

(4) 分析结果。如果  $A \ll Wp, H$ , 或 A 值的对应概率 p 小于给定的显著性水平  $\alpha$ , 则拒绝原假设, 这时可以认为不同水平下多总体均值向量之间存在显著差异; 反之, 则不能 拒绝原假设, 即认为不同水平下多总体均值向量之间没有显著差异。

### 三、应用实例

#### 应用实例 13-1

在本章导语中提到, 如何推断受教育程度不同状态下对 BAT 移动支付功能层面的三个观测变量“安全性 (sec)” “方便性 (con)” “娱乐性 (ent)” 感知的影响效应? 上述研究问题中, 分组变量 (自变量) 为受教育程度, 共有四个水平, 1 为“专科及以下”, 2 为“大学本科”, 3 为“硕士研究生”, 4 为“博士研究生”。同时, 检验的因变量包括“安全性” “方便性” “娱乐性”。

#### 1. 操作步骤

(1) 打开数据集 Example 13-1, 执行工具栏<sup>44</sup> Analyze (分析), 依次选择“General Linear Model (一般线性模型) — Multivariate (多变量)”, 在打开的“Multivariate (多变量) 对话框中 (见图 3-2), 将学生对 BAT 移动支付功能层面的三个感知指标“安全性 (sec)” “方便性 (con)” “娱乐性 (ent)” 选入“Dependent Variables (因变量) 方格中, 将“受教育程度 (degree) ” 选入“Fixed Factor(s) 方格中。

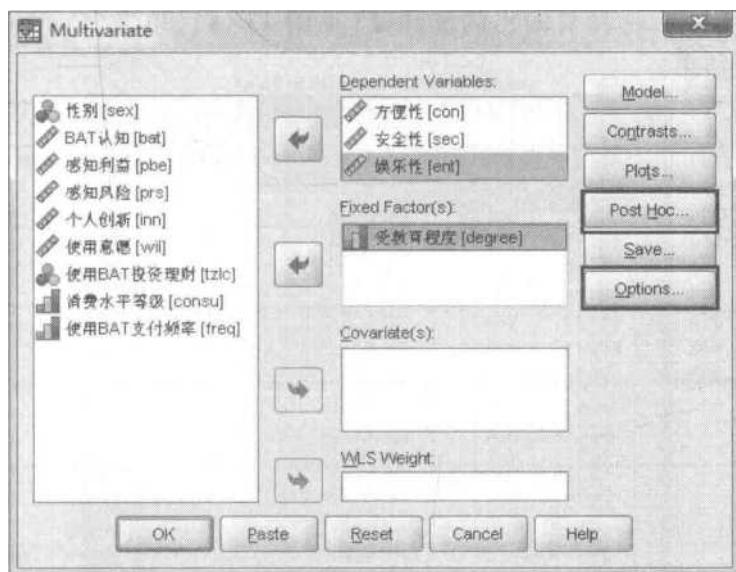


图 3-2 Multivariate 对话框

【注意】若“Fixed Factor (s) ^方格中若包含多个因子, 则变为多因素多变量方差分析。

(2) 按“Post Hoc . . .”按钮, 开启 Post Hoc 多重比较对话框 (见图 13-3), 选取进行事后比较的方法, 点击“Continue (继续) ”返回多变量方差分析对话框。

【注意】在方差分析中, 如果分析结果表明整体通过显著性检验, 表明至少有两个组别的平均数的差异达到显著性水平, 至于具体是哪几组的均值有差异需进一步使用事后比较进行判断。事后比较方法分为两种情况:

- ① 假设样本方差具有同质性, 常用方法有 Scheffé 法. Tukey 法等;
- ② 样本方差不满足同质性假设, 此时, 若选择 Tamhane's T2 检验法、Dunnett's T3 检验法

等，不必对数据进行转换

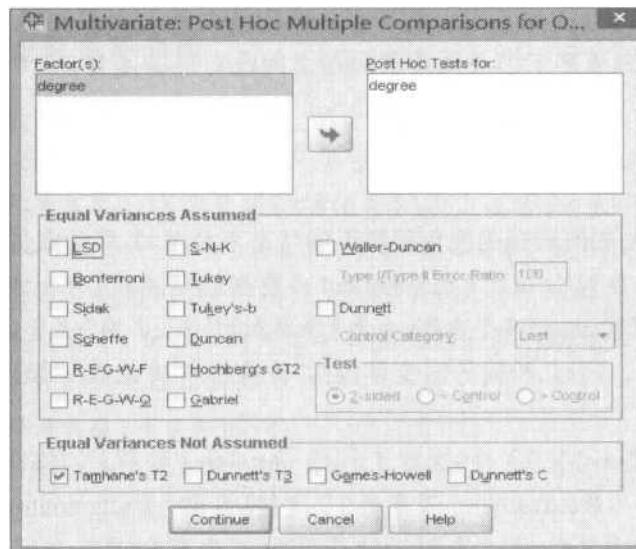


图 13-3 Post Hoc 多重比较对话框

(3)点击“Options”，选择要输出的统计量(见图 13-4)，单击“Continue（继续）”，单击“OK”，即得到分析结果。

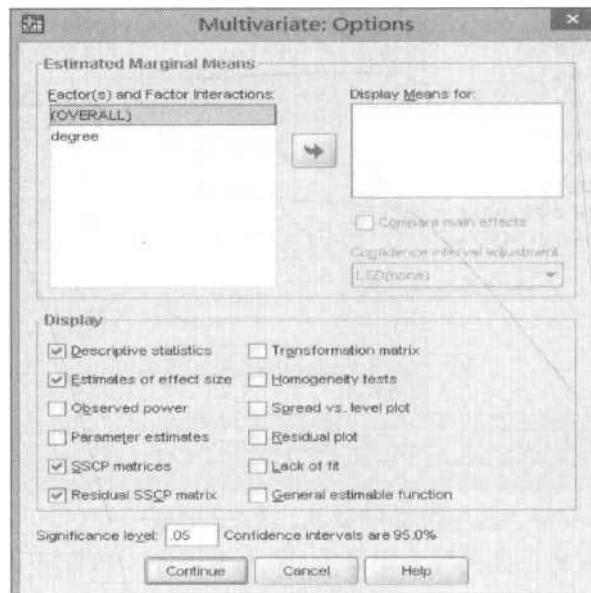


图 13-4 Options 对话框

## 2. 结果解读

受教育程度在 BAT 移动支付“功能认知”三个层面的差异比较结果如图 13-5 所示：

图 13-5 为单因子多变量检验的统计量，四个统计量在 0.05 的显著性水平下通过检验，表明学生对 BAT 移动支付“安全性(sec)”“方便性(con)”“娱乐性(ent)”感知在受教育程度的差异检验上达到显著。Wilks A 值为 0.772，对应概率  $p$  值为  $0.000 < 0.05$ ，通过显著性检验，这表明在“安全性(sec)”“方便性(con)”“娱乐性(ent)”中至少有一个在自变量“受教育程度”

Multivariate Tests*						
Effect		Value	F	Hwolhesia (i)	Error df	Si a
Intercept Pillafs Trace Wilks*	Pillai's Trace	951	2.16SE9*	3 000	337 000	000
	Lambda Hotelling's	049	2 166E3 <sup>b</sup>	3 000	337 000	000
	Trace Roy's	19 283	2 166E3,	3000	337 000	000
	Largest Root	19 283	2 166E3-	3000	337 000	000
degree	Pillai's Trace	244	10 005	9000	1 017E3	000
	Wilks'Lambda	772	10J25	9000	820.320	000
	Hotelling's Trace	275	10 256	9 000	1 007E3	000
	Roy's Largest Root	176	19 833 *	3 000	339 000	000

a. Exact statistic  
b The statistic is an upper bound on F that yields a lower bound on the significance level c Design intercept + degree

图 13-5 多变量检验结果截图

的平均数差异达到显著，具体是哪几组因变量有显著差异，需使用单因素单变量方差分析以及 事后比较进行判断。

图 13-6 为受教育程度变量与三个因变量的单变量方差分析的结果。由图可知，受教育 程度与 “安全性” “方便性” “娱乐性” 三个变量的单变量方差分析的 F 值分别为 1 1.859、 11.169, 9.622, 对应概率 p 值为 0.000<0.05, 通过显著性检验。表明不同受教育程度组别在三 个变量上均存在显著性差异。

Tests of Between-Significance Effects						
	Depen	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	方便性	125 882-	3	41 961	11 169	000
	安全性	106 253*	3	35 418	11 859	000
	娱乐性	122 156*	3	42 385	9622	000
intercept	方便性	10332 003	1	10332 003	2750E3	000
	安全性	6707 793	1	6707 793	2246E3	000
	娱乐性	9633 914	1	9633 914	2187E3	000
degree	零知■	125 882	3	41 961	11 169	000
	安全性	106 253	3	35 418	11 859	000
	娱乐性	122 156				095
Error	方便性	1273 576	339	3 757		
	安全性	1012 459	339	2 987		
	娱乐性	1493 229	339	4 405		
Total	方便性	51346 000	343			
	安全性	28009 000	343			
	娱乐性	<5127 000	343			
Corroded Total	方便性	1 399 458	342			
	安全性	1118 711	342			
	娱乐性	1620 385	342			

a R Squared = 090 (Adjusted R Squared = 082) b R Squared = 095 (Adjusted R Squared = .087) c R Squared = 078 (Adjusted R Squared = 07tr>

图 13-6 受教育程度变量与三个因变量的单变量方差分析的结果截图

图 13-7 为使用 Tamhane\* ST2 方法下的多重比较表。多重比较表(multiple comparisons) 是对影响因素的各组均值之间做两两比较，其中标记“ \* ”的对比组，表示对应两组的均值之 间存在显著性差异①。由图可知：

①读者只需关注标记“ \* ”的数值为正数的对比组别

**Multiple Comparisons**

Dependent Variable		Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
方便性	专科及以下大学本科	-2.25 <sup>†</sup>	.526	.001	-3.73	-.76
	硕士研究生	-2.17 <sup>†</sup>	.537	.002	-3.68	-.66
	博士研究生	-3.1 <sup>‡</sup> *	.941	.045	-6.27	-.07
	大学本科 专科及以下	2.25 <sup>†</sup>	.526	.001	.76	3.73
	硕士研究生	.07	.220	1.000	-.51	.66
	博士研究生	-.92	.804	.883	-4.20	2.35
	硕士研究生专科及以下	2.1 <sup>‡</sup>	.537	.002	.66	3.68
	大学本科	-.07	.220	1.000	-.66	.51
	博士研究生	-1.00	.811	.848	-4.24	2.25
安全性	专科及以下大学本科	-.29	.334	.952	-1.22	.65
	硕士研究生	-.96	.358	.060	-1.95	.03
	博士研究生	-3.86 <sup>†</sup>	.770	.009	-6.63	-1.09
	大学本科 专科及以下	.29	.334	.952	-.65	1.22
	硕士研究生	-.68 <sup>†</sup>	.211	.010	-1.24	-.11
	博士研究生	-3.57 <sup>†</sup>	.713	.021	-6.48	-.66
	硕士研究生专科及以下	.96	.358	.060	-.03	1.95
	大学本科	-.68 <sup>†</sup>	.211	.010	.11	1.24
	博士研究生	-2.90 <sup>†</sup>	.725	.048	-5.76	-.03
娱乐性	专科及以下大学本科	-1.78 <sup>†</sup>	.450	.002	-3.05	-.52
	硕士研究生	-1.08	.473	.161	-2.39	.24
	博士研究生	-3.86 <sup>†</sup>	1.045	.044	-7.62	-.10
	大学本科 专科及以下	1.78 <sup>†</sup>	.450	.002	.52	3.05
	硕士研究生	.71 <sup>†</sup>	.251	.031	.04	1.37
	博士研究生	-2.08	.965	.400	-6.03	188
	硕士研究生专科及以下	1.08	.473	.161	-.24	2.39
	大学本科	-.71 <sup>†</sup>	.251	.031	-1.37	-.04
	博士研究生	-2.78	.976	.179	-6.70	1.13
	博士研究生专科及以下	3.86 <sup>†</sup>	1.045	.044	10	7.62
	大学本科	2.08	.965	.400	-1.88	6.03
	硕士研究生	2.78	.978	.179	-1.13	6.70

Based on observed means.

The error term is Mean Square(Error) = 4.405.

\* The mean difference is significant at the .05 level.

图 13-7 多重比较结果截图

- ① 在感知方便性方面，“本科”“硕士研究生”“博士研究生”组均显著高于“专科及以下”组，同时，“博士研究生”组显著高于其他三组；
- ② 在感知安全性方面，“本科”高于“专科及以下”，“硕士研究生”高于“本科”和“专科及以下”，“博士研究生”高于其他三组；
- ③ 在感知娱乐性方面，“大学本科”高于“专科及以下”和“硕士研究生”，“硕士研究生”和“博士研究生”均高于“专科及以下”。

对上述结果进行归纳可以发现，‘专科及以下’组别在 BAT 支付安全性、方便性、娱乐性三个维度上感知均为最低；‘博士研究生’组在安全性和方便性感知上高于其他两组。由此可推

断, 受教育程度高低对三个观测变量具有影响效应。

【注意】在多变量方差分析中, SPSS 并没有直接给出其方差分析表, 在实际应用中, 读者最好对 SPSS 给出的分析结果进行整理, 得到更加直观的多变量方差分析表, 表格范例如下 (由于篇幅原因, 表 13-4 中组内与组间 SSCP 计算结果截图本文没有给出) :

表 13-4 单因素多变量方差分析摘要表

变异来源 df	SSCP	多变量 A	单变量 F 值		
			方便性	安全性	娱乐性
组内 339	( 1 274 -125.46 817.65 - 125.46 1 012 -270.83 < 817.65 - 270.83 1 <125.8 46.81 101.32 ) 组间 3 J01.32 33.94 127.16	0.77**	11.17***	11.86 ...	9.62 ...

注: \*\*\*为在 0.001 的显著性水平下显著

由于进行多变量方差分析需要满足以下前提条件:各因变量为正态分布且方差相等, 各因变量之间为多元正态分布, 故对该检验过程补充因变量正态性和方差同质性检验, 步骤如下:

(1) 进行因变量正态性检验。依次选择“Analyze (分析) ^ Descriptive Statistics (描述性统计) ^ Explore (探索)”, 在打开的“Explore (探索)”对话框中(见图 13-8), 将学生对 BAT 移动支付功能层面的三个感知指标“安全性 (sec)”“方便性 (con)”“娱乐性 (ent)”—选入“Dependent List (因变量列表)”框中, 单击“Plots (图形)”按钮, 选中“Nomality plots with tests (正态性图形检验)”复选框, 单击“Continue (继续)”按钮返回, 单击“OK (确定)”按钮进行正态性检验。

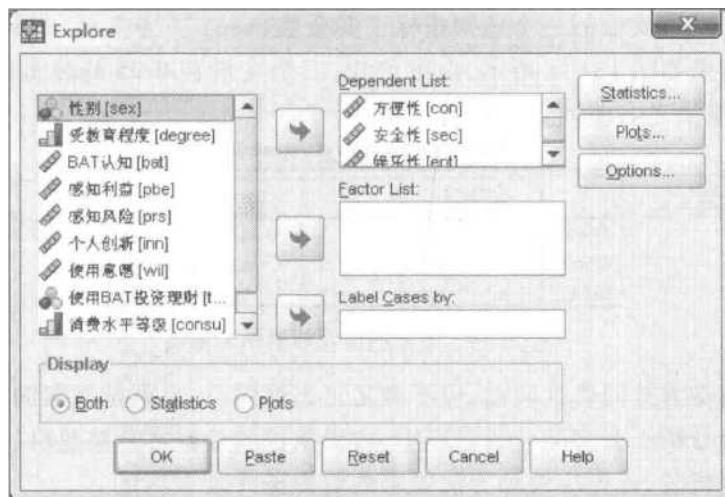


图 13-8 Explore 对话框

BAT 移动支付功能层面的三个感知指标“安全性 (sec)”“方便性 (con)”“娱乐性 (ent)”的正态性检验结果如图 13-9 所示, 由图可知, 三个变量在 0.05 的显著性水平下拒绝原假设, 即说明各因变量不满足正态性假设。

(2) 进行因变量方差同质性检验。依次选择“Analyze (分析) — Compare Means (均值比

	Tests of Normality			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
方便性	.192	343	.000	.918	343	.000
安全感	.136	343	.000	.964	343	.000
娱乐性	.143	343	.000	.959	343	.000

a Lilliefors Significance Correction

图 13-9 方差正态性检验结果截图

较) ^One-Way ANOVA (单因素方差分析)”，在打开的“单因素方差分析”对话框中(见图 13-10)，将学生对 BAT 移动支付功能层面的三个感知指标“安全性(sec)”“方便性(con)”“娱乐性(ent)”选入<sup>44</sup> Dependent List (因变量列表) 框中，将“受教育程度(degree)”。变量选入“Factor (因素)”框中。单击“Post Hoc (多重比较)”按钮，选中“LSD”复选框，单击“Continue (继续)”按钮返回。单击“Options (选项返按钮，选中“Homogeneity of variance test (方差齐性检验)”复选框，单击“Continue (继续)”按钮返回：单击“OK (确定)”按钮进行方差齐性检验。

图 One-Way ANOVA

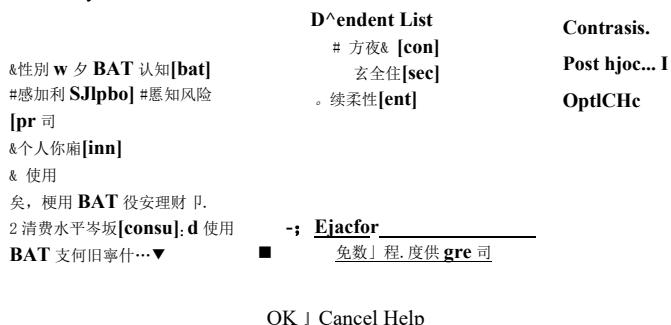


图 13-10 One-Way ANOVA 对话框

BAT 移动支付功能层面的三个感知指标“安全性(sec)”“方便性(con)”“娱乐性(ent)”的方差齐性检验结果如图 13-11 所示，由图可知，三个变量在 0.05 的显著性水平下不拒绝原假设，即说明各变量满足方差同质性假设。

Test of Homogeneity of Variances

	Levene Statistic	df1	df2	Sig.
方便性	2284	3	339	.079
安全感	980	3	339	.402
娱乐性	176	3	339	.913

图 13-11 方差同质性检验结果截图

故各因变量满足方差同质性假设，但不满足正态性假设，也就是该案例其实并不适合进行单因素多变量方差分析。在实际分析过程中，前提条件的检验常常被忽略，进而导致分析方法的误用。在进行分析时，首先应检验方法要求的前提条件是否满足。

例如，根据某连锁超市不同门店销售数据，判断不同地区三种商品(啤酒、肉类、米类)的销售情况是否有差异，数据详见数据集 Example 13-2①。

①出于商业案例数据的保密性原则。本案例的原始数据已经过编辑和处理，不再代表实际市场情况，其分析结果仅用作方法演示和学习参考。

首先进行正态性检验，由检验结果可见（图 13-12），三个变量均满足正态性假设；然后进行方差同质性检验，检验结果见图 13-13，三个变量均满足方差同质性假设。故可进行单因素 多变量方差分析，分析结果见图 13-14。由结果可知，四个地区三种商品的销量不存在显著性差异，对结果的详细解读请参照上例，此处不再重复介绍。

Tests of Normality						
	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
啤酒销重	.048	207	.200 <sup>b</sup>	994	207	.580
肉类销重	.070	207	.016	.992	207	.314
米类销重	.046	207	.200 <sup>b</sup>	993	207	.498

a Lilliefors Significance Correction

\* This is a lower bound of the true significance.

图 13-12 因变量方差正态性检验结果截图

ANOVA					
		Sum of Sauares	df	Mean Sauare	F
啤酒销重	Between Groups	86.475	3	28.825	.279
	Within Groups	20965.970	203	103.281	
	Total	21052.444	206		
肉类 i 肉重	Between Groups	47.008	3	15.669	.564
	Within Groups	5640.770	203	27.787	
	Total	5687.778	206		
米类销量	Between Groups	222.157	3	74.052	.754
	Within Groups	19936.693	203	98.210	
	Total	20158850	206		

图 13-13 方差同质性检验结果截图

Multivariate Tests <sup>a</sup>						
Effect	Value	F	Hypothesis df	Error df	SIQ.	Partial Eta Sauared
intercept	Pillai's Trace	.996	1.628E4-	3.000	.201.000	.000
	Wilks' Lambda	.004	1.626E4-	3.000	.201.000	.000
	Hotelling <sup>*</sup> ® Trace	242 B80	1.626E4 <sup>3</sup>	3; 000	.201.000	.000
	Ro/s Largest Root	242 630	1.626E4 <sup>3</sup>	3.000	.201.000	.000
area	Pillai's Trace	.023	.512	9.000	.609.000	.866
	Wilks'Lambda	.978	.509	9.000	.489.332	.868
	Hotelling's Trace	.023	.506	9.000	.599.000	.871
	Roys Largest Root	.014	948 <sup>k</sup>	3.000	.203.000	.418

Tests of Between-Subjects Effects						
Source	Depend ent ***	Type III Sum of Sauares	df	Mean Sauare	F	Sia.
Corrected Model	啤酒销重	88.475-	3	28.825	.279	.840
	肉类销重	47.008 <sup>b</sup>	3	15.669	.564	.639
	米类销量	222.157=	3	74.052	.754	.521
Intercept	啤酒销重	M47B91 524	1	1147691 524	1.111E4	.000
	肉类销量	516270.013	1	518270.C13	1.858E4	.000
	米类销重	2031182.476	1	2031182.476	2068E4	.000
area	啤酒销量	86.475	3	28.825	.279	.640
	肉类销重	47.008	3	15.669	.564	.639
	米类销重	222.157	3	74.052	.754	.521
Error	啤酒销重	20965 970	203	103281		
	肉类销重	5640.770	203	27.787		
	米类销量	19936.693	203	98.210		
Total	啤葛龍	'181980.000	207			
	肉类 f 曜	52779S.000	207			
	米类销量	2078974.000	207			
Corrected Total	啤酒销重	21052.444	206			
	肉类销重	5687.778	206			
	米类销重	20158.850	206			

a. R Squared = .004 (Adjusted R Squared = -.011)

b. R Squared = .008 (Adjusted R Squared = -.006) c R Squared = .011 (Adjusted R Squared = -.004)

		Multiple Comparisons				
Dependent Variable	間地	Mean Difference (kJ)	Std Error	Sio	95% Confidence Interval	
					Lower Bound	Upper Bound
啤酒銷量	A	48	2169	1.000	-5.36	6.31
	B	.90	2122	.999	-4.60	0.60
	C	1.77	1,974	.939	-3.55	7.09
	D					
B	A	-48	2169	1.000	-6.31	5.36
	C	.42	2 058	1.000	-5.10	5.95
	D	1.29	1 905	.984	-3.83	6.41
C	A	-90	2122	.999	-6.60	4.80
	B	-42	2 058	1.000	-5.95	5.10
	D	.87	1 851	.998	-4.09	5.83
D	A	-1.77	1,974	.939	-7.09	3.55
	B	-1.29	1 905	.984	-6.41	3.83
	C	-.87	1.851	.998	-5.83	4.09
肉类銷量	A	-38	1.070	1.000	-3.26	2.50
	B	48	.983	.997	-2.16	3.11
	C	89	.999	.942	-1.80	3.57
	D					
B	A	38	1 070	1.000	-2.50	3.26
	C	86	1 072	.964	-Z02	374
	D	1.27	1 087	.816	-1.65	4.19
C	A	-48	.983	.997	-3.11	2.16
	B	-86	1 072	.964	-3.74	2.02
	D	41	1 001	.999	-2.27	3.09
D	A	*89	.999	.942	-3.57	1.80
	B	-1.27	1 087	.816	-4.19	1.65
	C	*41	1 001	.999	-3.09	2.27
米类請里	A	-2.84	2.243	.754	-8.87	319
	B	-2.06	1.953	.877	-7.33	3.21
	C	-.93	2113	.998	-6.62	475
	D					
B	A	2.84	2 243	.754	-3.19	887
	C	78	1 880	.999	-4.28	5.84
	D	1.91	2 046	.927	-3.59	7.40
C	A	2.06	1.953	.877	-3.21	7.33
	B	-.78	1.880	.999	-5.84	4.28
	D	113	1 724	.987	-3.49	5.75
D	A	93	2113	.998	-4.75	6.62
	B	-1.91	2.046	.927	-7.40	3.59
	C	-1.13	1 724	.987	-5.75	3.49

Based on observed means.

The error term is Mean Square(Error) = 98.210

图 13-14 不同地区商品销量差异单因素多变量方差分析结果截图

#### 四、多变量方差分析应注意的问题

第一，进行方差分析需要满足一定的假定前提，如每个样本独立、服从正态分布、总体方差相同或为常数等。对于多变量方差分析而言，需要满足的假定前提包括：数据来源于随机抽样，样本观察值之间相互独立；各因变量为正态分布且方差相等；各因变量之间为多元正态分布。但在实际调查中，完全符合这些条件是不现实的，在调查样本数较大的情况下，能近似满足上述要求，并不影响对调查结果的判断。

方差分析对变量的数据类型要求为因变量是间距测度等级变量，自变量为定类测度等级变量。

方差分析对因变量之间关系的要求是因变量之间需要存在一定程度的线性相关，如果是非线性关系，需要转换为线性关系，相关程度要达到一定的强度。

方差分析对样本量的要求是总量足够大，分类样本足够大。

第二，在实际应用时，由于观测次数有限，影响因素可能只有有限的几个状态水平，这时一般只考察影响因素有限集中各状态的不同搭配对所观测变量的影响。

第三，在实际应用两因子方差分析方法时，影响因素之间是否存在交互作用，除了可以依据有关专业理论及实践经验做出初步判断外，最终还应在对试验数据进行方差分析后才能得到可靠的结论。

## 第二节

# 多变量相关与回归分析

相关与回归分析也是一种有效分析变量间相互关系的统计方法。多变量方差分析得到的是因素对观测变量是否具有显著影响的整体判断，即因素与观测变量间是否相互独立，而多变量回归分析得到的是在不独立的情况下自变量与因变量之间的更加精确的回归函数式，也即判断相关关系的类型。在第十二章中，对双变量间的相关与回归分析理论已经做过相关介绍，本节将基于导语中分析多个因素对用户 BAT 支付使用意愿影响情况的需求出发，重点介绍在涉及多变量时如何具体应用该方法。

### 一、 多变量相关分析

相关分析的方法很多，比较直接和常用的是绘制散点图。但是散点图并不能精确地展示变量间的相关关系大小，还需要进一步计算相关系数。多变量间两两相关系数可用相关系数矩阵来表示  $G$

记互/2, 的, …成为左个变量, 任意角与易之间的相关系数记为七, 其中,

$k$

注  $g_{10} = \text{值} : = (13.9)$

3, -島) 2 言 (气-£) 2

则上个变量之间的相关系数矩阵为:

…5、

$R = \begin{matrix} & & & \\ & & & \\ & & & \\ & & & \\ & & & \end{matrix}$

(13.10)

、 「妇」 \*2 \* \* \*  $_{kk}$

### 二、 应用实例

#### 应用实例 13-2

在本章导语中，数据集（Example 13-1）中包含使用意愿、BAT 认知、个人创新、感知风险、感知利益等变量。为了解 BAT 移动支付使用意愿与 BAT 认知、个人创新、感知风险、感知利益的相关性，可使用相关分析方法分析各变量之间的相关关系。

##### 1. 操作步骤

打开数据集 Example 13-1, 执行工具栏“Analyze（分析）”，依次选择“Correlate（相关）—► Bivariate（双变量）”，打开双变量相关分析窗口（如图 13-15 所示）。将“使用意愿(wil)”“BAT

认知(bat)”“个人创新(inn)”“感知风险(prs)”与“感知利益(pbe)”五个变量选入“Variables”方格中，点击“OK”，得到输出结果（见图 13-16）。

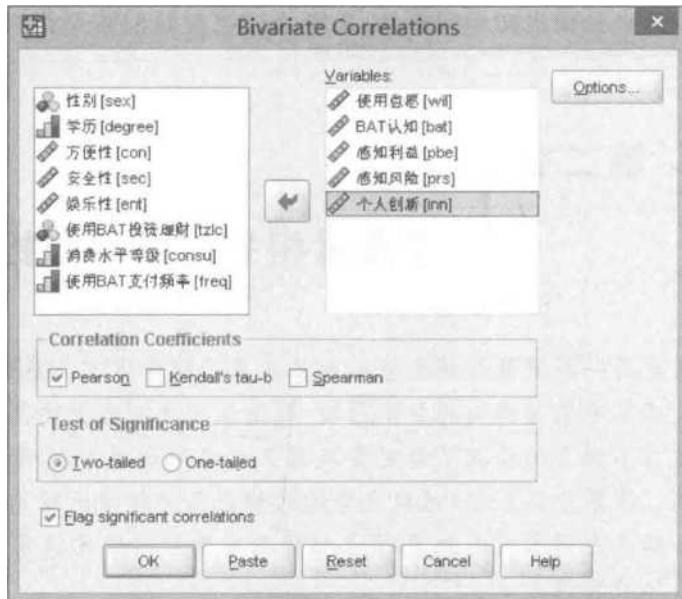


图 13-15 双变量相关分析窗口

## 2. 结果分析

输出结果如图 13-16 所示：

Correlations						
		伸用意愿	BATH 知	感知利緒	感知風盼	个人创新
使用意愿	Pearson Correlation	1	.715**	.657**	.705**	.643**
	Sig. (2-tailed)		.000	.000	.000	.000
	N	343	343	343	343	343
BAT 认知	Pearson Correlation	.715**	1	.763**	.735**	.646**
	Sig. (2-tailed)	.000		.000	.000	.000
	N	343	343	343	343	343
感知利益	Pearson Correlation	.657**	.763**	1	.722**	.643**
	Sig. (2-tailed)	.000	.000		.000	.000
	N	343	343	343	343	343
感知风险	Pearson Correlation	.705**	.735**	.722**	1	.672**
	Sig. (2-tailed)	.000	.000	.000		.000
	N	343	343	343	343	343
个人创新	Pearson Correlation	.643**	.646**	.643**	.672**	1
	Sig. (2-tailed)	.000	.000	.000	.000	
	N	343	343	343	343	343

\*\* Correlation is significant at the 0.01 level (2-tailed).

图 13-16 相关分析输出结果截图

由图 13-16 相关分析输出结果可知，使用意愿与其他四个变量均存在显著的相关性（ $p$  值 小于 0.01），其中，使用意愿与 BAT 认知、感知利益、感知风险、个人创新的相关系数大小分别为 0.715、0.657、0.705、0.643，呈现中高度的相关性，同时，BAT 认知、感知利益、感知风险、个人 创新两两之间也存在显著的相关关系（ $P$  值小于 0.01），且感知利益与 BAT 认知，感知风险与 BAT 认知，感知利益与感知风险之间的相关系数都大于 0.70，表明如果进行多元回归，这些变

量之间可能存在共线性问题，但这只是初步判断，具体判断情况需要在多元回归分析的过程中 做多重共线

性检验（在案例 13-2 中将讨论此问题）。

### 三、多变量回归分析

#### （一）多变量回归分析基本原理

在实际问题中，影响因变量的因素往往不止一个，如影响学生 BAT 移动支付使用意愿的因素不仅包括个人对 BAT 认知，还包括个体创新力、感知风险以及感知利益等因素。在实际应用中很难分辨哪一个因素重要，哪一个因素不重要，各个因素对因变量的影响都是必不可少的，这时就需要进行一个因变量与多个自变量之间的回归，即多元回归。多元回归与相关分析两者之间的联系和区别见表 13-5。

表 13-5 相关分析与回归分析的联系和区别

分析方法	分析目的	变量处理	分析原理	解释
相关分析	依存关系	对自变量和因变量不作区分	测度变量的联系方向和大小	相关系数表明变量间的联系大小
回归分析	因果关系	对自变量和因变量作区分	用自变量对因变量的解释度	根据回归模型可进行总体推断及预测

多元回归分为多元线性回归和多元非线性回归两种形式。在实际问题中，自变量和因变量之间的关系并非都是线性的。当因变量随着自变量的不同而变化，并且呈现出某种曲线形态时，称二者之间为非线性关系。对于变量之间的非线性关系，可以用适当类型的曲线来描述，也可以通过变量转换将非线性回归方程转换为线性方程后，再进行求解。本节重点介绍多元线性回归。

假设影响因变量  $y$  的自变量共有上个，对两个自变量的多元线性回归模型为：

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \epsilon$$

它的估计方程为：

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \cdots + \hat{\beta}_k x_k \quad (13.11)$$

式中， $\hat{\beta}_0$  为直线的截距， $\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_k$  为直线的偏回归系数。

#### （二）多变量回归分析的基本假定

应用多元线性回归时，所分析的数据必须符合下列基本假定：

- 随机误差项零均值： $E(\epsilon_i) = 0$  ( $i=1, 2, \dots, n$ ，这里  $n$  是样本数量)。
- 随机误差项服从正态分布： $\epsilon_i \sim N(0, \sigma^2)$  ( $i=1, 2, \dots, n$ )。
- 随机误差项同方差和无自相关：随机误差项前后期不相关假定：

$$\text{cov}(\epsilon_i, \epsilon_j) = E[(\epsilon_i - \bar{\epsilon})(\epsilon_j - \bar{\epsilon})] = 0 \quad (i \neq j, i, j = 1, 2, \dots, n)$$

- 随机误差项和自变量不相关。
- 自变量彼此间无多重共线性。

多元线性回归方程的基本原理和步骤与简单线性回归类似，在相关系数和判定系数指标上仅存在较小差异。

多元回归分析中，简单相关系数可能不能够真实地反映出因变量和自变量之间的相关性，由于变量之间关系较复杂，它们可能受到不止一个变量的影响，此时可考虑使用偏相关系数。偏相

关系数是指在消除其他变量影响的条件下，所计算的某两个变量之间的相关系数

在多元线性回归中，自变量个数会影响因变量中被估计回归方程所解释的变差数量。当增加自变量时，会使预测误差变小，从而减少残差平方和 SSE。当 SSE 变小时，SSR 就会变大，从而使  $R^2$  变大。如果模型中增加一个自变量，即使这个自变量在统计上并不显著， $R^2$  也会变大。因此，为避免增加自变量而高估 k，统计学家提出用样本量 n 和自变量的个数 k 去调整  $R^2$ ，计算出调整的多重判定系数，记为其计算公式为：

(13.12)

$$n - 1$$

$R^2$  的解释与  $R^2$  类似，不同的是

同时考虑了样本量 n 和模型中自变量个数 k 的影响，这就使得  $R^2$  的值始终小于 1，而且  $R^2$  的值不会由于模型中自变量个数的增加而越来越接近 1。因此，在多元回归分析中，通常用调整的多重判定系数来评价回归方程的拟合优度。

$R^2$  的平方根称为多重相关系数，也称为复相关系数，它度量了因变量同多个自变量的总体相关程度。

在多元线性回归中，要判断每个自变量对因变量的影响是否都显著，需要对各回归系数  $\beta_j$  分别进行检验。首先，提出假设： $H_0: \beta_j = 0$ ， $H_A: \beta_j \neq 0$ ，进一步计算检验的统计量：

(13.13)  $Z_j = \frac{\hat{\beta}_j}{\text{se}_{\beta_j}}$

式中， $\text{se}_{\beta_j}$  是回归系数  $\beta_j$  的抽样分布的标准差。最后根据，统计量做出决策，给定显著性水平  $\alpha$  根据自由度 H 计算出统计量的  $t$  值。若  $|Z_j| > t_{\alpha/2, H-1}$  拒绝原假设，表明回归系数  $\beta_j$  显著。

关于多元线性模型的建模过程以及检验等问题，将结合案例进行具体介绍。

## 四、应用实例

### 应用实例 13-3

实例 13-2 分析了使用意愿与 BAT 认知、个人创新、感知风险、感知利益的相关程度，结果表明使用意愿与其他四个变量均存在显著的相关性。为有效了解四个变量对使用意愿的影响效应的大小，可选择多元线性回归进行进一步分析。

在建立模型之前，首先对被解释变量与解释变量之间的理论关系进行简单的说明：

1. 认知是行为控制的决定因素之一，考虑将 BAT 认知纳入模型；
2. 用户在决策是否使用移动支付工具时会权衡行为所带来的收益与风险；
3. 个人创新性通常表现为对新鲜事物有浓厚的兴趣并敢于尝试，因而个人创新性在新兴的 BAT 移动支付使用意愿中发挥着一定作用，在移动商务的新环境下，将个人创新性作为一个影响变量是具有合理性的。

综上，初步考虑将 BAT 认知 (bat)、个人创新 (inn)、感知风险 (prs) 与感知利益 (pbe) 四个变量作为自变量，使用意愿 (wil) 作为因变量。

### 操作步骤

(1) 打开数据集 Example 13-1，执行工具栏“Analyze（分析）”，依次选择“Regression（回归方法）”→“Linear（线性）”，开启线性回归对话窗口（图 13-17）。在开启的“Linear Regression（线性回归）”对话框中，将“使用意愿 (wil)”点送至“Dependent Variables（因变量）”方格中，将“BAT 认知 (bat)”、“个人创新 (inn)”、“感知风险 (prs)”与“感知利益 (pbe)”四个变量点选至

Independent s) ”方格中。在”Method (方法) ”右侧的下拉式选项中选择”Enter (直接进入

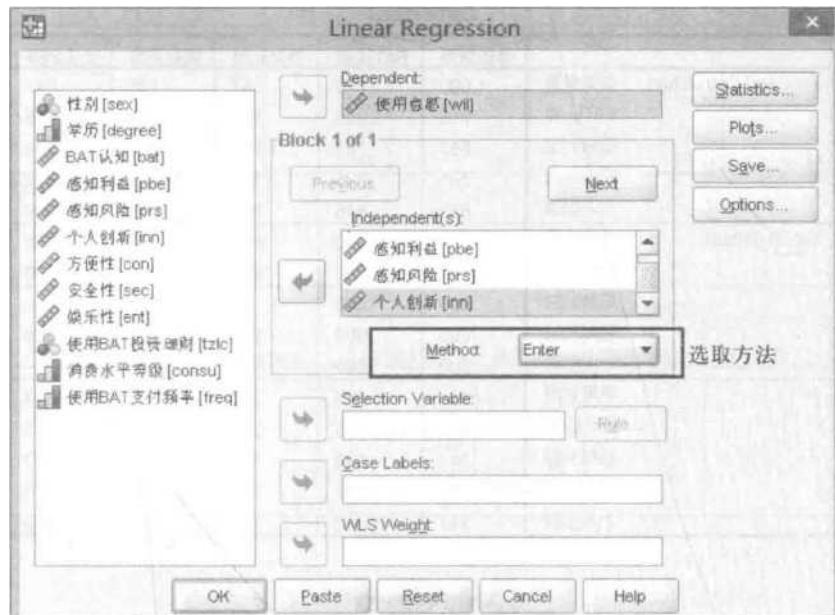


图 13-17 线性回归对话框

【注意】在多元回归分析过程中，SPSS 共有 5 种进行变量选择的方法：直接进入法（Enter）、逐步回归法（Stepwise）、向前法（Forward）、向后法（Backward）、删除法（Remove）。其中，最为常见的为直接进入法（该选项为 SPSS 系统默认）和逐步回归法。直接进入法将所有的解释变量同时纳入模型中，而不论个别解释变量对被解释变量的影响是否显著。逐步回归只包含对被解释变量有显著影响的解释变量，其余没有达到显著性水平的变量则会被排除在回归模型之外。

(2) 按<sup>44</sup> Statistics ... 钮，打开 \*\* Linear Regression : Statistics (统计量) ”对话框（见图 13 - 18），选取 “ Estimate (估计值) ” “ Model fit (模型拟合度) ” “ Descriptive (描述性统计量) ” “ Collinearity diagnostics (共线性诊断) ”以及 “ Durbin-Watson (杜宾统计量) ”，点击”Continue “返回 线性回归对话窗口。最后，单击 OK 按钮，即可得到回归过程的运行结果)

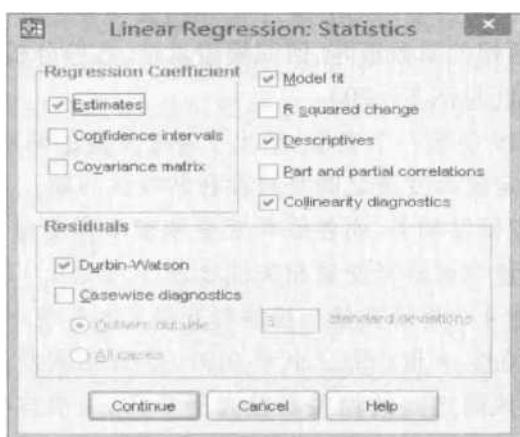


图 13-18 线性回归统计量对话框

## 1. 结果解读

多元线性回归的输出结果如图 13-19 所示：

Correlations						
		使用意愿	BAT 知	感知利害	感知 RI 路	个人创 I 新
Pearson Correlation	使用意愿	1.000	.715	.657	.705	,643
	BAT 认知	.715	1.000	.763	.735	,646
	感知利益	.657	.763	1.000	.722	643
	感知风险	,705	.735	.722	1.000	,672
	个人创新	,643	.646	.643	,672	1.000
Sig. (1-tailed)	使用意愿		,000	,000	,000	,000
	BAT 认知		,000	,000	,000	,000
	感知利益		,000	,000	,000	,000
	感知风险		,000	,000	,000	,000
	个人创新		,000	,000	,000	,000
N	使用意愿	343	343	343	343	343
	BAT 认知	343	343	343	343	343
	感知利益	343	343	343	343	343
	感知风险	343	343	343	343	343
	个人创新	343	343	343	343	343
Model Summary **						
Model I	R	R Square	Adjusted R Square	Std Error of the Estimate	Durbin-Watson	
1	,778=	,606	,601	,812	2.056	

a. Predictors: (Constant), 个人创 I 新, 感知利益, 感知风险, BAT 认知

b. Dependent variable 使用意愿

ANOVA <sup>b</sup>					
Model	Sum of Squares	df	Mean Square	F	Sig.
1	342.103	4	85.526	129.787	,000-
	Regression 222731	338	,659		
	Residual Total 564.834	342			

a. Predictors: (Constant), 个人创新, 感知利益, 感知风险, BAT 认知

b. Dependent Variable 使用意愿

图 13-19 多元线性回归输出结果截图（1）

回归过程输出结果包括相关系数矩阵、回归模型总结、方差分析表（见图 13-19）和回归系数数、共线性检验、残差统计量（见图 13-20）。

相关系数矩阵（图 13-19 中第一个表格）给出了各个变量的两两相关性的大小和方向，根据相关系数矩阵，可初步判定解释变量之间是否存在共线性问题。一般来讲，变量间的最佳状态是解释变量间不相关或中低度相关，而各解释变量和被解释变量之间高度相关。从相关系数矩阵可以看出，各解释变量与被解释变量相关性显著，P 值（Sig.）小于 0.01，且相关系数大小在 0.643 到 0.715 之间，表明 4 个解释变量与被解释变量之间存在中高度相关性。四个解释变量之间也两两具有显著相关性，P 值（Sig.）小于 0.01，且感知利益与 BAT 认知、感知风险与 BAT 认知、感知利益与感知风险之间的相关系数大于 0.70，表明这些变量之间可能存在共线性问题。

回归模型总结部分（图 13-19 中第二个表格），也就是对方程拟合情况的简单描述，从中

Model	Coefficients <sup>18</sup>						
	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
	B	Std. Error	Beta			Tolerance	VIF
1 (Constant) BAT 认知 益 个人创 <sub>諸</sub>	1.734	.294		5.902	.000		
	183	.034	.317	5.378	.000	.335	2.981
	.059	.037	.092	1.588	.113	.349	2.868
	185	.038	.274	4.817	.000	.361	2.772
	174	.044	.195	3.962	.000	.480	2.083

a Dependent variable: 使用意愿

Mode	Dime nsio l n	Eigenvalue	Condition Index	Variance Proportions				
				(Constant)	BAT 认知	利益	感知风哈	个人革新
1	1	4.953	1.000	.00	.00	.00	.00	.00
	2	.020	15.913	.86	.08	.00	.01	.07
	3	.013	19.376	.00	.19	.06	.01	.86
	4	.008	25.602	.06	.26	.03	.96	.06
	5	.007	26.939	.08	.46	.91	.02	.00

a Dependent variable: 使用意愿

Residuals Statistics <sup>*1</sup>					
	Minimum	Maximum	Mean	Std Deviation	N
Predicted Value	4.81	9.88	8.06	1.000	343
Residual	-2.591	2347	.000	807	343
Std Predicted Value	-3.248	1.823	.000	1.000	343
Std. Residual	-3.192	2.891	.000	994	343

a. Dependent Variable: 使用意愿

图 13-20 多元线性回归输出结果截图（2）

可以看出，四个解释变量与被解释变量的多元相关系数  $R$  和判定系数① $R^2$  分别为 77.8% 和 60.6%，模型的解释程度较高；。“统计量可检验模型中的序列相关现象，即回归方程中的各个误差项是否存在自相关，本例中  $\text{DW}=2.056$ ，接近于 2，表明相关系数接近于 0，残差间无自相关性。

【注意】。值介于 0 到 4 之间，当  $DW$  值越接近于 2 时，表明相关系数越接近于 0；当  $DW$  值越接近于 0 时，表明相关系数越接近于 1；当  $DW$  值越接近于 4 时，表明相关系数越接近于 -1。在解释变量数目确定以及给定的显著性水平下，为准确判断，读者可由分布表查出其上限临界值和下限临界值《，然后根据以下规则进行判断：

当  $d_u < DW < 4 - d_u$  时，不存在序列相关；

当  $DW < d_l$  时，为正序列相关；

当  $0 < DW - d_l$  时，为负序列相关；

当  $d_t < DW < d_u$  以及  $4 - d_l < DW < 4 - d_u$  时无结论。

由方差分析表（图 13-19 中第三个表格）可知， $F=129.787$ ,  $p$  值 (Sig.) 小于 0.01，表明模型整体拟合效果良好，回归方程中至少有一个回归系数不为 0，系数显著性的具体情况需要观察回归系数及其检验结果。

①判定系数表示自变量所能解释的方差占总方差的百分比，判定系数越大，说明模型的拟合效果越好。

由回归系数及其检验结果（图 13-20 中第一个表格）可知，除感知利益的回归系数未达到统计学意义上的显著性外 ( $Q1.588, p = 0.113 > 0.1$ )，其他三个自变量回归系数均达到显著性水平 ( $p < 0.01$ )。在回归分析中，当某个变量的回归系数未达到显著性水平时，不能认为其与被解释变量之间没有关系，如本例中的感知利益变量，其与使用意愿的相关系数大小为 0.657，之所以出现回归系数不显著的现象，可能是由于解释变量之间存在高度的相关性（感知利益与 BAT 认知、感知风险之间的相关系数大小分别为 0.763、0.722）。回归系数及其检验结果还给出了容忍度 (Tolerance；TOL 值) 和方差膨胀因子 (Variance inflation factor；VIF)，两者可有效度量多元共线性问题。其中，TOL 值越接近于 0，多元共线性问题越严重；VIF 大于 10 时，则表示有多重共线性问题。本例中四个解释变量的 TOL 值在 0.335 至 0.480 之间，VIF 值在 2.083 到 2.891 之间，多重共线性不明显。

输出结果中还给出了单独的共线性检验表（图 13-20 中第二个表格），包括特征值 (Eigenvalue)、条件指标 (Condition Index；CI 值)，在回归分析中，当特征值小于 0.01，条件指标 (CI 值) 大于 30 时，认为解释变量间存在共线性。由图 13-20 可知，有两个特征值小于 0.01，但是，条件指标 (CI 值) 均小于 30，故认为各解释变量之间多重共线性问题不严重。

## 2. 模型整理

根据 SPSS 的多元回归输出结果，整理得到如表 13-6 所示的结果摘要，

表 13-6 多元回归输出结果摘要表 (1)

变量	回归系数	标准误	t 值	P 值
截距项	1.734	0.294	5.902	0.000...
BAT 认知 (bat)	0.183	0.034	5.378	0.000 ...
感知利益 (A)	0.059	0.037	1.588	0.113
感知风险 3 小)	0.185	0.038	4.817	0.000 ...
个人创新 (访 '')	0.174	0.044	3.962	0.000 ...

$$R = 0.778 \quad \text{调整后 } R^2 = 0.601 \quad \text{F} = 129.787 \dots \quad \text{数 } F = 2.056$$

注：\*\*\* 为在 0.001 的显著性水平下显著

据此，得到影响学生 BAT 支付使用意愿的多元线性模型：

$$wil = 1.734 + 0.183x_{bat} + 0.059x_A + 0.185x_{risk} + 0.174x_{inn}$$

式中，加为使用意愿 J 小为 BAT 认知，p 为感知利益，prs 为感知风险 Q”为个人创新。根据该回归模型可知，高校学生群体 BAT 使用意愿受 BAT 认知、感知利益、感知风险和个人创新等因素影响。在控制其他变量不变的情况下，BAT 认知水平每提高一个单位，使用意愿增加 0.183 个单位；用户感知利益对使用 BAT 移动支付的意愿作用效应为正，但在统计学意义上效果并不显著；感知风险回归系数为 0.185，即感知风险每提高一个单位，使用意愿提升 0.185 个单位，与常理相悖，故决定舍弃该变量，出现这种现象的原因可能是由于多重共线性、数据质量问题或样本量较小等。个人创新每提高一个单位，使用意愿增加 0.174 个单位。

舍弃感知风险变量后重新拟合模型，结果见表 13-7 据此，得到最终影响学生 BAT 支付使用意愿的多元线性模型：

$$wil = 2.112 + 0.238x_{bat} + 0.111x_A + 0.238x_{inn}$$

对比两个回归模型可以看出，去除感知风险变量后，其余三个变量的系数均有所增大，且

感知利益变量变得显著。在控制其他变量不变的情况下, BAT 认知水平每提高一个单位, 使用意愿增加 0.283 个单位; 感知利益水平每提高一个单位, 使用意愿增加 0.110 个单位; 个人创新水平每提高一个单位, 使用意愿增加 0.283 个单位。

表 13-7 多元回归输出结果摘要表 (2)

变量	回归系数	标准误	t值	P 值
截距项	2.112	0.292	7.230	0.000…
BAT 认知( <i>hat</i> )	0.238	0.033	7.164	0.000…
感知利益 (冲 e)	0.110	0.037	2.992	0.003
个人创新 3")	0.238	0.043	5.516	0.000 …

R = 0.761 /?^2 = 0.579 调整后广 = 0.575 F = 155.157 \*•, 。昨 2.083

## 五、应用多元相关与回归分析应注意的问题

第一, 在多元回归分析时, 要注意多重共线性问题。如果自变量间存在完全共线性或者高度共线性, 会直接影响模型参数估计与统计推断的准确性。如: 在参数显著性检验中, 多重共线性使得错误拒绝原假设的概率增大; 回归得到的参数的正负号与实际情况不符合; 模型回归结果整体显著, 而大部分参数不显著等。要最大限度地避免以上问题, 在进行构建模型之前, 可先进行各解释变量之间的相关分析, 初步判定解释变量之间有无共线性问题。

第二, 在实际操作中, 建立回归模型的过程非常复杂, 变量之间的关系很难确切估计, 有时我们甚至都无法找到哪些变量可以对因变量的变动形成统计意义上的显著性影响, 而且实际被广泛应用的回归分析技术也要比我们这里介绍的内容复杂得多, 应用时必须结合具体情况进行探讨和分析。

## 第三节

### 分类模型一 Logistic 模型

在第二节中介绍的多元线性回归模型, 只适用于被解释变量是连续变量的情形(在案例 13-2 中, 使用意愿为连续变量), 但在实际应用中, 也会遇到因变量是分类变量的情况。例如, 要了解客户使用 BAT 进行投资理财的影响因素(投资理财变量为二分变量), 此时基于同方差、线性和正态性等假定的线性回归不再适用, 可以考虑采用 Logistic 回归。当 Logistic 回归涉及的因变量为二分变量时, 称为二分 Logistic 回归; 当所涉及的因变量为两个及以上类别时, 称为多分类 Logistic 回归, 限于篇幅, 本节只介绍二分 Logistic 回归模型。

#### 一、Logistic 回归模型的基本原理

设 y 为因变量, 且 y 服从于二项分布(Binomial), 其取值为 0 和 1, 不适合直接作为因变量, 记 p = P(y = 1), q = P(y=0), </= 1-p, 则 p == 1 即是我们所要研究的对象。在实际应用中, 如果有很多因素影响 y 的取值, 那么将这些因素当作自变量, 自变量既可以是连续变量也

可以是分类变量。值得注意的是, Logistic 回归与一般的线性回归不同, 它解释了自变量和因变量概率取值  $P = P(y=D)$  之间的关系, 一般用 Logistic 曲线表示  $p = P(y=1)$  与自变量之间的关系:

$$p = P(y=1) = \frac{\exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k)}{1 + \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k)} \quad (13.14)$$

对上式作 Logit 变换, 则可得到 Logistic 回归模型的形式:

$$\ln(p/(1-p)) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k \quad (13.15)$$

其中角,  $\beta_0, \beta_1, \dots, \beta_k$  为待估参数。已称为优势比率 (odds ratio; OR), 即事件发生与不发生概率比, 优势比率的自然对数称为 y 的 Logit, 记作  $\text{Logit}$  优势比率也可以表示为:

$$(13.16)$$

估计的系数反映了自变量的变化情况。可以证明,  $\text{Logit}(y)$  的取值范围为  $(-\infty, +\infty)$ , 在这种情况下, 因为无穷大或无穷小数值的存在, 普通最小二乘法 (OLS) 难以对参数进行估计, 故 Logistic 回归模型的参数估计使用极大似然估计法。模型拟合好坏通过似然值来测度 (通常统计软件中将似然值取自然对数并乘以 -2, 记为  $-2\ln L$ ), 当似然值为 1 时, 此时  $-2\ln L = 0$ , 达到最小值, 模型完全拟合。对于系数的显著性检验, Logistic 回归模型常用的统计量为 Wald 统计量①, Wald 检验原假设为: 待检验的系数为零, Wald 统计量计算公式见式 (13-17), 单个解释变量的 Wald 统计量服从自由度为 1 的  $\chi^2$  分布。

$$(13.17)$$

## 二、应用实例

### 应用实例 13-4

为研究武汉高校学生的感知利益、安全感、消费水平、BAT 移动支付使用频率四个变量是否可有效解释以及预测他们的投资理财选择, 将感知利益、安全感、消费水平、使用频率作为自变量, 投资理财作为因变量, 构建二分 Logistic 回归模型。其中, 感知利益与安全感为连续变量; 消费水平和使用频率为等级变量, 共有 5 个等级, 取值越高, 表明消费水平和使用频率越高; 因变量投资理财为二分变量, 1 表示使用 BAT 进行投资理财, 0 表示没有使用 BAT 进行投资理财。

#### 1. 操作步骤

打开数据集 Example 13-1, 执行工具栏“Analyze (分析)”, 依次选择“Regression (回归方法) ^ Binary Logistic (二元 Logistic)”, 在打开的“Logistic Regression (Logistic 回归)”对话框中

① 关于 Logistic 回归参数估计以及假设检验介绍, 读者可参阅: (美) 梅纳德 (Menard,S.) . 应用 Logistic 回归分析 [M]. 李俊秀译. 2 版. 上海: 格致出版社. 上海人民出版社, 2012.

(如图 13-21 所示), 将“投资理财 (tzlc ) ”选入“Dependent Variables (因变量) ”方格中, 将“感知利益 (pbe) ” “安全感 (sec) ” “消费水平等级 (consu) ”与“使用频率 (freq) ”四个变量选入“Covariates”方格中。在“Method (方法) ”右侧的下拉式选项中选择“Enter”, 点击 OK, 即可得到回归结果。

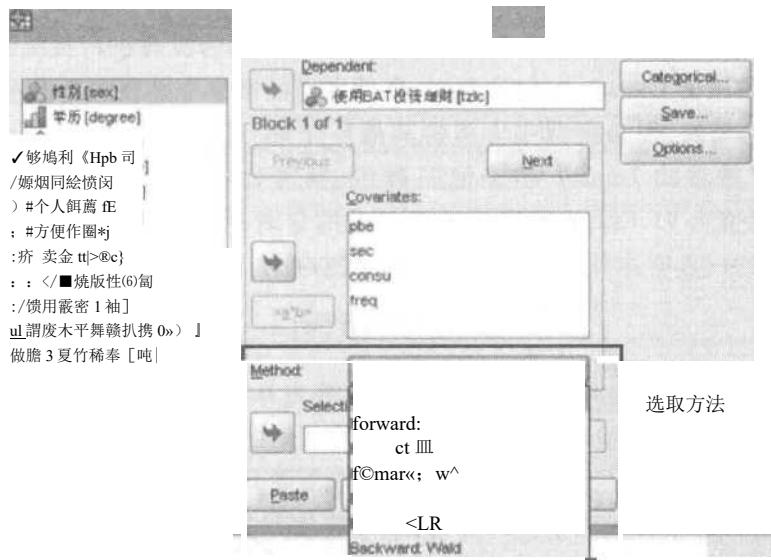


图 13-21 Logistic 回归对话窗口

**【注意】**在 SPSS 的“Method”下拉框中提供了“Enter”、“Forward: Conditional”、“Forward: LR”、“Forward : Wald”、“Backward : Conditional”、“Backward : LR”和“Backward : Wald”七种 Logistic 回归方法。系统默认的选项为 Enter 法，即将所有的自变量全部纳入 Logistic 回归模型中，而不论其显著与否，相当于在多元线性回归分析中提到的直接进入法 forward 即向前逐步选择，表示根据一定准则逐步选择模型中显著的自变量 backward 即向后逐步选择，表示根据一定准则逐步剔除模型中不显著的自变量。

## 2. 结果解读

Logistic 回归分析的输出结果如图 13-22 所示：

Case Processing Summary		
Unweighted Cases	N	Percent
Selected Cases Included in Analysis	342	99.4
Missing Cases	2	.6
Total	344	100.0
Unselected Cases	0	.0
Total	344	100.0

a. If weight is in effect, see classification table for the total number of cases

## Dependent Variable Encoding

Oriainal ...	Internal Value
没有投资理财	0
进行投资理财	1

图 13-22 样本基本信息表输出截图

图 13-22 给出了样本信息摘要表以及因变量投资理财的编码表，从图中可知，有效观测 值共有 342 个，缺失值为 2，样本单位总数为 344。投资理财的数值编码为 0 时，表示没有投资 理财；数值编码为 1 时，表示进行投资理财。

图 13-23 为 Logistic 初步回归的信息，包括变量分类表、只包含常数项的 Logistic 初步回 归模型。由图中的分类表可知，未分类前没有投资理财的样本单位数为 265，进行了投资理财 的样本单位数为 77，没有进行投资理财的样本占 77.5%。因为没有进行投资理财的样本数较 多，当用只包含常数项的 Logistic 回归预测时，事件发生的概率全部预测为没有投资理财，其预 测正确性为 77.5%，显然，因为模型中未包含其 他自变量，只有常数项，这种预测的意义并不 大。观察只包含常数项的 Logistic 初步回归模型，其常数项 的估计值为 -1.236，标准误为 0.129，Wald 统计量值为 91.137，达到统计学意义上的显著性水平(Sig. = 0.000)，Exp( b) 表示 优势比率，值为 0.291，表明在只包含常数项进行预测时，有：

$$\hat{p} = \text{exp}(-1.236) / (1 + \text{exp}(-1.236)) = 0.291$$

即“进行投资理财”和“没有投资理财”相比，胜算为 0.291。

图 13-23 step = 0 的结果输出中还给出了不在模型中的变量的 Score 检验结果，在进行 Logistic 回归分析时，如果 Score 检验结果显示所有的自变量均在统计意义上不显著，则表明所有 的自变量对因变量没有解释以及预测的意义，此时应放弃 Logistic 回归分析。当至少有一个自 变量达到显著性水平时，就 可以将该变量纳入模型进行进一步估计。本例中的所有自变量均 达到显著性水平。

### Block 0: Beginning

Classification Table<sup>a,b</sup>

	Predicted		Percentage Correct	
	没有投资理财	进行投资理财		
Step 0 使用 BAT 投资理财	265	0	100.0	
没有投资理财	77	0	.0	
Overall Percentage			77.5	

a. Constant is included in the model.

b. The cut value is .500

Variables in the Equation

	B	S.E.	Wald	df	Sia.	Exp(B)
Step 0 Constant	-1.236	.129	91.137	1	.000	291

Variables not in the Equation

	Score	df	Sid.
Step 0 Variables pbe	20.270	1	.000
sec	9.715	1	.002
consu	19.978	1	.000
freq	16.478	1	.000
Overall Statistics	41.983	4	.000

图 13-23 开始区块(step = 0)的结果截图

图 13-24 为所有自变量加入模型后的分析输出结果，输出结果包括四个部分，即模型的

Omnibus 系数检验表、模型摘要表、分类表、包含在模型中变量的参数估计及显著性检验。

**Omnibus Tests of Model Coefficients**

	Chi-square	df	Sig.
Step 1 Step	46.099	4	.000
Block	46.099	4	.000
Model	46.099	4	.000

**Model Summary**

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	318.709	.126	.192

a Estimation terminated at iteration number 5 because parameter estimates changed by less than .001

**Classification Table<sup>a</sup>**

		Predicted			Percentage Correct	
		使用 BAT	阴资理	讲行将资理财		
Step 1	使用 BAT 投资理财没有投资理财 进行理财 Overall Percentage	254	11	11	95.8	
		65	12	12	15.6	
					77.8	

a. The cut value is .500

**Variables in the Equation**

Step		B	S.E.	Wald	df	Sig.	Exp(B)
1 <sup>b</sup>	pbe	.276	.088	9.876	1	.002	1.318
	sec	-.094	.083	1.302	1	.254	.910
	consu	405	.125	10.564	1	.001	1.500
	freq	.267	.107	6.237	1	.013	1.305
	Constant	-5.915	1.555	14.466	1	.000	.003

a. Variable(s) entered on step 1: pbe, sec, consu, freq.

图 13-24 区块 1 (step=1) 的结果截图

首先，观察 Omnibus 系数检验表，纳入四个变量的 Logistic 回归模型的卡方值为 46.099, P = 0.000，在 0.01 的显著性水平下显著，表明在感知利益、安全感、消费水平、使用频率这四个变量中，至少有一个变量可以有效解释以及预测客户是否投资理财的概率。要具体分析哪些变量具有影响显著，需观察包含在模型中变量的参数估计及显著性检验输出结果。

模型摘要输出结果包含了 -2LL, Cox-Snell  $R^2$  以及 Nagelkerke  $R^2$  值。-2LL 用来评价模型拟合的好坏， $\sim 2LL$  值越小，拟合效果越好；Cox-Snell  $R^2$  和 Nagelkerke  $R^2$  评价自变量与因变量之间的关联性（类似于多元回归分析中的正，但不能说明自变量解释因变量变异的比例）。Cox-Snell  $R^2$  和 Nagelkerke  $R^2$  值分别是 0.126、0.192，表明自变量与因变量之间有关联，但关联度较小。

再来看 Logistic 回归模型的分类预测结果，原来有 265 个没有投资理财的样本单位，根据加入 4 个自变量的预测结果，有 254 个被归类为没有投资理财的组别（分类正确），11 个被归类为进行投资理财的组别（分类错误）；原来有 77 个进行投资理财的样本单位，根据 Logistic

回归模型预测结果，有 12 个被归类为进行投资理财的组别（分类正确），65 个被归类为没有进行投资理财的组别（分类错误）。整体预测归类正确率为 77.8%。一般来讲，整体预测归类正确率越高，表明模型拟合越好，自变量对因变量影响越大。

参数估计及显著性检验结果表明，除了安全感 (sm) 的 Wald 检验值未达到显著性水平外 (sig. = 0.254)，其他三个变量均在 0.05 的显著性水平下显著。

根据 SPSS 的回归输出结果，整理得到如表 13-8 所示的摘要表。

表 13-8 Logistic 回归分析摘要表

变量	回归系数	标准误	Wald 值	Df 值
常数项	-5.915	1.555	14.466	1
感知利益 (pbe)	0.276	0.088	9.876*	1
安全感 (sec)	-0.094	0.083	1.302	1
消费水平等级 (consu)	0.405	0.125	10.564*	1
使用频率 (freq)	0.267	0.107	6.237,	1
关联强度 整体拟合度	Cox-Snell 广义 = 0.126 -2LL = 318.709	Nagelkerke $\chi^2 = 0.192$ $x^* = 46.099 \dots$		

注：\*表示在 0.05 的显著性水平下显著，\*\*表示在 0.001 的显著性水平下显著

据此，得到最终影响学生使用 BAT 进行投资理财的 Logistic 回归模型：

$$\log (j) = -5.915 + 0.276x_{pbe} - 0.094x_{sec} + 0.405x_{consu} + 0.267x_{freq}$$

$$p = \frac{\exp(-5.915 + 0.276x_{pbe} - 0.094x_{sec} + 0.405x_{consu} + 0.267x_{freq})}{1 + \exp(-5.915 + 0.276x_{pbe} - 0.094x_{sec} + 0.405x_{consu} + 0.267x_{freq})}$$

式中， $p$  恪为感知利益，腿 ( $:=$  为安全感,  $consu$  为消费水平等级，斤舛为使用频率) 如果预测值  $p$  的概率大于 0.5，则样本被归类为进行投资理财组别；如果预测值的概率小于 0.5，则样本被归类为没有投资理财组别。

### 三、应用 Logistic 回归模型应注意的问题

第一，当变量之间的关系呈非线性时，可以对因变量或自变量进行转换，虽然此时因变量与自变量的实质关系仍为非线性的，但在表现形式上为线性，即参数间的关系是线性的。在数学中，概率的 Logit 能有效地修正二分因变量与自变量之间的非线性关系。

第二，进行线性回归分析，关注的是所有自变量对因变量的解释程度以及是否能更好地预测因变量。对于 Logistic 回归模型而言，由于因变量为分类变量（通常为二分变量），更多关注的是正确预测的频率以及模型能否有效减少误差，而不是预测值接近观测值（0 或 1）的程度。

### 本章小结•••

多变量数据分析用于分析多个变量在不同取值下的数据分布状况以及不同变量之间的密切程度。本章重点介绍了多变量方差分析、多变量回归分析、分类 Logistic 回归模型的基本原理以及在实际中应该注意的问题，并结合本章导语介绍了常用方法的 SPSS 软件实现过程。

在实际问题中，影响一个事物的因素很多，方差分析是研究一种或多种因素的变化对试验结果的观测值是否有影响的统计方法，它通过分析数据中不同来源的变异对总变异的贡献大小，来确定可控因素对研究结果影响力大小。在方差分析中，按照观测变量的个数，可分为

单变量方差分析和多变量方差分析；按照涉及的因素个数划分，可分为单因素方差分析和多因素方差分析。

相关与回归分析也是一种可有效分析变量间相互关系的统计方法。两变量间的相关分析方法包括制作散点图和计算相关系数  $r$ 。相关系数是用于描述样本两个变量之间联系紧密程度的统计量，多变量间的两两相关系数可用相关系数矩阵来表示。单纯的相关分析可以不必区分自变量与因变量，但在回归分析中则必须首先对自变量与因变量做出区分。多元回归分为多元线性回归和多元非线性回归。多元线性回归分析的关键在于合理地选择变量， $t$  检验是筛选自变量的有效方法：

当因变量是分类变量时，基于同方差、线性和正态性等假定的线性回归不再适用，此时可采用 Logistic 回归模型。当 Logistic 回归涉及的因变量为二分变量时，称为二分 Logistic 回归；当所涉及的因变量为两个及以上类别时，称为多分类 Logistic 回归。Logistic 回归与一般的线性回归不同，它解释了自变量和因变量概率取值  $p = P(y=1)$  之间的关系，一般用 Logistic 曲线表示，对于 Logistic 回归而言，更多地关注正确与错误预测的频率以及模型能否有效地减少误差  $a$ 。

## 磁关撻名词

方差分析 单因素多变量方差分析相关系数 回归分析判定系数调整的判定系数

多元线性回归偏相关系数分类模型 Logistic 回归优势比率

## 肯思考题



即测即评

1. 简述多变量方差分析的基本原理，并举例解释。
2. 简述多变量相关分析与回归分析的联系和区别。
3. 简述多元线性回归分析和简单线性回归的联系与区别。
4. 多元线性回归分析的基本假定有哪些？
5. 多元线性回归分析如何选择自变量？应注意哪些方面的问题？
6. 简述 Logistic 回归的基本思路以及应用中应注意的问题。

可，本章实训. . .

### 一、实训目的

1. 掌握方差分析的基本操作。
2. 了解多元线性回归模型的基本思路以及操作过程，并对结果进行解读。
3. 熟练应用回归模型，掌握分类模型 Logistic 模型的适用情形，并对结果进行解释。

### 二、实训内容

消费者信心指数是反映消费者信心强弱的指标，综合反映并量化了消费者对当前经济形势评价和对经济前景、收入水平、收入预期以及消费心理状态的主观感受。20世纪40年代，美国密歇根大学的调查研究中心为了研究消费需求对经济周期的影响，首先编制了消费者信心指数，随后欧洲一些国家也先后开始建立和编制消费者信心指数。

在2007年，联恒市场研究启动中国消费者信心调查项目，该项目由联恒的张文彤博士与美国密歇根大学社会研究所消费者信心调查课题组负责人Richard Curtin博士共同开发完成。由于数据保密需要，本书中所附的仅为部分数据（见数据集Example13-3），且包含的消费者信心指数与实际值存在一定偏差（计算消费者信心指数所需的权重为保密数据，不对外提供），故本书所附数据只供读者进行练习使用。

本实训中，所需要的变量信息说明如表13-9：

表 13-9 消费者信心指数案例 Example13-3 变量说明表

变量说明	意义
性别 (sex)	1 =男; 2 =女
受教育程度 (degree)	1 =初中/技校以下; 2 =高中/中专; 3 =大专; 4 =本科; 5 =硕士及以上
年龄 (age)	年龄变量
月收入 (income)	为连续变量, 数值越高, 收入越高
婚姻状况 (s7)	1 =已婚; 2 =未婚; 3 =分居/离婚/丧偶
总指数 (index)	反映消费者信心强弱

实训资料来源：张文彤，钟云飞. IBM SPSS 数据分析与挖掘实战案例精粹[M]. 清华大学出版社, 2013.

应用数据集 Example13-3, 分析受访者的性别、受教育程度、年龄、月收入、婚姻状况等背景资料对消费者信心指数的影响情况, 具体要求如下:

1. 分析不同性别、受教育程度状态下受访者的消费者信心指数是否存在显著差异。
2. 建立回归模型, 分析月收入与年龄对消费者信心指数的影响方向以及影响程度
3. 将总指数按照 100 的临界值分为两类, 大于 100 表示与基期(本例中基期为 2007 年 4 月)相比, 受访者更乐观; 小于 100 表示与基期相比, 受访者更为悲观「请读者探究影响受访者是否乐观的因素。」

### 三、实训组织

1. 以个人为单位负责完成数据分析任务。
2. 提交数据分析书面报告。
3. 组织部分个人课堂交流书面报告。
4. 教师批改报告, 点评与总结。



### 导语

在第十三章的引例中介绍了“BAT 移动支付用户的使用意愿调查”案例，其中包含 BAT 认知、个人创新、感知利益、方便性等影响 BAT 支付使用意愿的因素以及受教育程度、性别、消费水平等基本特征信息。如何基于这些数据对用户群体进行细分，进而了解各群体的使用意愿偏好，是研究者关注的一个重要问题，那么，选择哪些统计分析技术可以实现上述目的呢？

在第十三章中，结合“BAT 移动支付用户的使用意愿调查”案例，介绍了方差分析、多元线性回归分析以及 Logistic 回归分析等多变量描述性统计分析方法的基本原理，并给出了在不同受教育程度受访者对 BAT 移动支付功能感知的影响效应问题，以及在分析使用意愿与 BAT 认知、个人创新、感知风险、感知利益的相关关系问题等方面的具体应用。在此基础上，还可以通过因子分析对众多的使用意愿的影响因素进行降维，进而利用聚类分析对 BAT 移动支付用户进行群体细分。此外，在产品营销案例的分析过程中，若想了解客户的消费水平差异和需求倾向，可对商品销售额进行对应分析和综合评价；为分析商品属性对消费者决策的影响，可对消费者对该商品不同属性组合的评分数据进行联合分析。在实际调查分析中，应根据不同的研究目的，选择不同的分析方法。

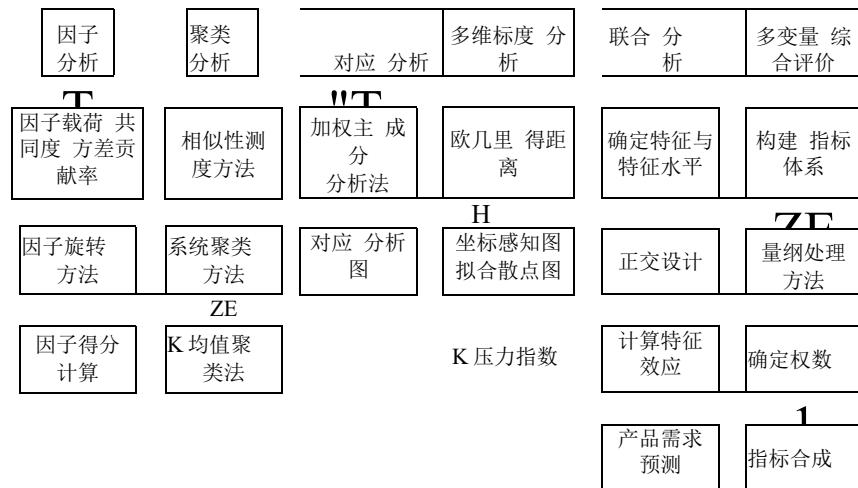
本章将具体介绍因子分析、聚类分析、对应分析、多维标度分析、联合分析以及多变量综合评价等多变量统计分析方法的基本原理，同时结合实际应用案例，演示上述方法在 SPSS 等统计分析软件中的实现过程。

### 教学目标与要求

1. 掌握因子分析的意义和主要步骤；
2. 掌握聚类分析的意义和操作方法；
3. 了解对应分析的概念和基本思想，掌握对应分析的实现过程；
4. 了解多维标度分析的基本理论；
5. 掌握联合分析的基本理论和方法；
6. 了解综合评价中指标体系的构建方法

## 、本章知识结构

多变量数据分析  
(2)



## 因子分析

进行市场研究时，人们往往希望尽可能多地收集与研究目的相关的数据信息，以避免由于信息缺失所造成的决策失误。这种做法虽然有助于全面、详尽地把握现象，但同时也会导致指标间出现信息重叠，使指标间存在包含关系或因果关系。这不仅会降低指标的解释效力，也给实际统计分析工作带来较大困难。

例如，在BAT移动支付用户使用意愿调查中，设计了七类可能影响用户使用意愿的指标，包括BAT认知、感知利益、感知风险、个人创新、方便性、安全性和娱乐性。可以利用这七个指标来预测用户的使用意愿或投资理财的选择等问题，不过这些指标之间或多或少会存在一定相关性，即存在信息重叠。这时便可采用因子分析方法，解决变量间的信息重叠问题，以进行客观准确的评价或预测。

### 一、基本原理

因子分析(Factor Analysis)是一种降维、数据简化技术，是把具有错综复杂关系的变量综合为数量较少的几个因子，以再现原始变量与因子之间的相互关系。在因子分析过程中，既要保证被描述的原始变量所损失的信息尽可能地少，又要求所提取的因子变量具有较强的可解释性。从显在变量提炼潜在因子，判定和消除指标间的信息重叠，是因子分析的出发点。

因子分析的基本思想是：通过研究众多变量之间的内部依赖关系，探求观测数据中的基本结构，并将其用少数几个抽象的变量来表示。这几个抽象的变量被称作公共因子，能反映原来众多变量的主要信息。原始的变量是可观测的显在变量，而因子一般是不可观测的潜在变量。在因子载荷矩阵中，各行元素的平方和(即共同度)表示每个指标被各个因素所解释的方差总和，各列元素的平方和表示每个因子对方差的解释能力，即方差贡献率。方差贡献率越大，说明该因子所包含的信息越多。

因子分析的假设前提是观测变量能够转换为一系列潜在变量(因子)的线性组合。因子可以与两个以上的变量相关联，也可以只与其中的一个变量相关联。

因子分析中的公共因子是不可直接观测但又客观存在的共同影响因素，每一个变量都可以表示成公共因子的线性函数与特殊因子之和，即：

$$X_i = \alpha_{i1}F_1 + \alpha_{i2}F_2 + \dots + \alpha_{im}F_m + \epsilon_i \quad (14.1)$$

这就是因子分析的数学模型，用矩阵形式表示为：

$$X = AF + s \quad (14.2)$$

1 因子分析中涉及的因子载荷、共同度以及方差贡献率等概念，读者可参阅相关多元统计分析教材，这里不做赘述。

其中，自变量几，「2, … /」，称为公共因子，是不可观测的潜在变量，参数(3)3=1, 2, … %；户 1, 2, … F) 称为因子载荷系数，其意义是第 j 个观测变量与第 j 个公共因子的相关系数，即  $X_i$  依赖  $F_m$  的比重。

对原始数据进行因子分析前，首先要对数据进行标准化处理，并对数据是否适合因子分析进行适用性

检验。

在对多个具有不同量纲的变量进行处理时，常常需要对各变量的数据进行标准化处理，也就是把一组数据转化成具有平均数为 0、标准差为 1 的新数据。实际上，标准化只是将原始数据进行线性变换，并没有改变某个数值在该组数据中的位置，也没有改变该组数据分布的性状。统计上常用的标准化公式为：

$$Z = \frac{X - \bar{X}}{S} \quad (14.3)$$

因子分析常用的检验方法包括 KMO 检验法和巴特莱特球形度检验法。具体来说，KMO 检验值大于 0.5 或巴特莱特球形度检验尹值小于 0.05 即通过检验，可以进行因子分析。因子分析过程一般分为三个步骤：确定因子个数—对因子载荷求解和旋转—计算因子得分。

1. 确定因子个数。提取公共因子的一般原则是，特征根  $\lambda > 1$  或者累计贡献率达到 85% 以上，同时也可以使用碎石图作为辅助判断。为了能够更清晰、合理地解释因子含义并给因子命名，往往还需要结合因子载荷矩阵（特别是旋转后的因子载荷矩阵）进行分析和判断。

2. 因子载荷的求解和旋转。因子载荷的求解方法很多，比如主成分法、主轴因子法、极大似然法等。由于得到的因子载荷矩阵是不唯一的，可能有些变量在多个公共因子上都有较大的载荷，有些公共因子也可能同时对多个变量具有较大的载荷。这时为得到含义清晰的公共因子，需要通过因子旋转的方法，使每个变量仅在一个公共因子上有较大的载荷，而在其余的公共因子上的载荷较小。最常用的因子旋转方法是方差最大化正交旋转法。

3. 计算因子得分。即计算公因子  $U_1, U_2, \dots, U_k$  在每一个样本点上的得分。首先需要用回归的思想给出公共因子用原始变量表示的线性表达式：

$$U_i = b_{1i}X_1 + b_{2i}X_2 + \dots + b_{ni}X_n \quad (14.4)$$

式中， $b_{1i}, b_{2i}, \dots, b_{ni}$  称为因子得分系数。基于上述表达式，利用原始变量计算因子得分；也可以用因子得分做进一步的分析，如样本间的比较分析和样本的聚类分析等。

## 二、应用实例

### 应用实例 14-1

在第十三章的应用实例 13-2 中，基于“BAT 移动支付用户的使用意愿调查”项目，探究了 BAT 认知、感知利益、感知风险和个人创新等因素对使用意愿的影响。在分析过程中，相关分析结果显示上述因素存在较高的相关性，多元线性回归分析中的共线性检验结果也说明这些因素之间存在一定的共线性。可见这些影响因素之间存在信息重叠，可以采用因子分析方法进行降维，提炼潜在因子。

分析：该调查项目中设计的影响用户使用意愿的因素共有七个，包括 BAT 认知、感知利益、感知风险、个人创新、方便性、安全性和娱乐性，这七个因素的取值均为得分，是连续变量，可以应用因子分析方法提炼潜在因子，进行降维。具体实现过程利用 SPSS16.0（英文版）软件中的“Crosstabs（交叉表）”菜单实现。

#### 1. 操作步骤

- (1) 打开数据集 Example13-1。
- (2) 依次选择“Analyze（分析）—Data Reduction（降维）—Factor Analysis（因子分析）”，展开“Factor Analysis（因子分析）”对话框，如图 14T 所示。从对话框左侧的变量列表中选择“bat（BAT 认知）”“pbe（感知利益）”“prs（感知风险）”“inn（个人创新）”“con（方便性）”“sec（安全性）”和 5（娱乐性）七个变量进入 Variables（变量）框。

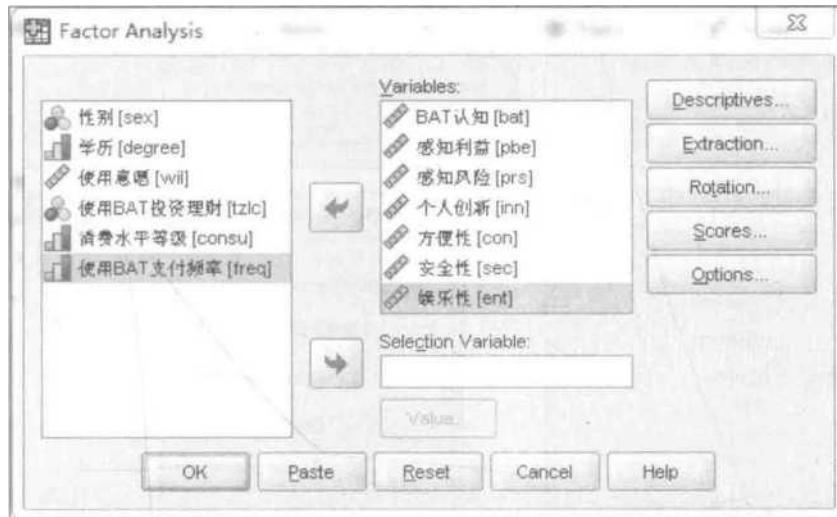


图 14-1 Factor Analysis 对话框

(3) 单击“Descriptive (描述)”按钮，选中“KMO and Bartlett \* s test of sphericity (KMO 和 Bartlett 的球形度检验)”复选框；返回主对话框后，单击“Exraction (抽取)”按钮，选中“Scree Plot (碎石图)”复选框；返回主对话框后，单击“Rotation (旋转)”按钮，选中“Varimax (最大方差法)”；返回主对话框后，单击“Scores (得分)”按钮，选中“Save as variables (保存为变量)”和“Display factor score coefficient matrix (显示因子得分系数矩阵)”复选框，再返回主对话框上述过程如图 14-2 所示。

(4) 最后，单击“OK (确定)”按钮，即可得到因子分析结果，如图 14-3、图 14-4 和图 14-5 所示。

## 2. 解读输出结果

图 14-3 所示的 KMO 和 Bartlett 球形度检验结果显示，KMO 值为 0.886，说明数据适合做因子分析。

图 14-4 为方差贡献率汇总表，包括初始的因子特征根及方差贡献率。第一主成分的特征根为 4.27，对应的累积方差贡献率为 60.996%；第二主成分的特征根为 0.996，对应的累积方差贡献率为 75.231%。利用特征根选择公共因子时，通常要求特征根 N1；利用主成分分析法选择公共因子时，一般要求因子的累计方差贡献率大于 80%。本例中符合特征根 N1 的公共因子只有一个，其对应的累积方差贡献率只有 60.996%，而且一个公共因子的实际含义不好解释，综合考虑本例将提取 2 个公共因子进行分析。

在上述 SPSS 操作步骤中，单击“Exraction (抽取)”按钮时，选中“Number of factors (因子数量)”复选框，并设置固定因子个数为 2（如图 14-6 所示），继续运行因子分析过程，输出结果。

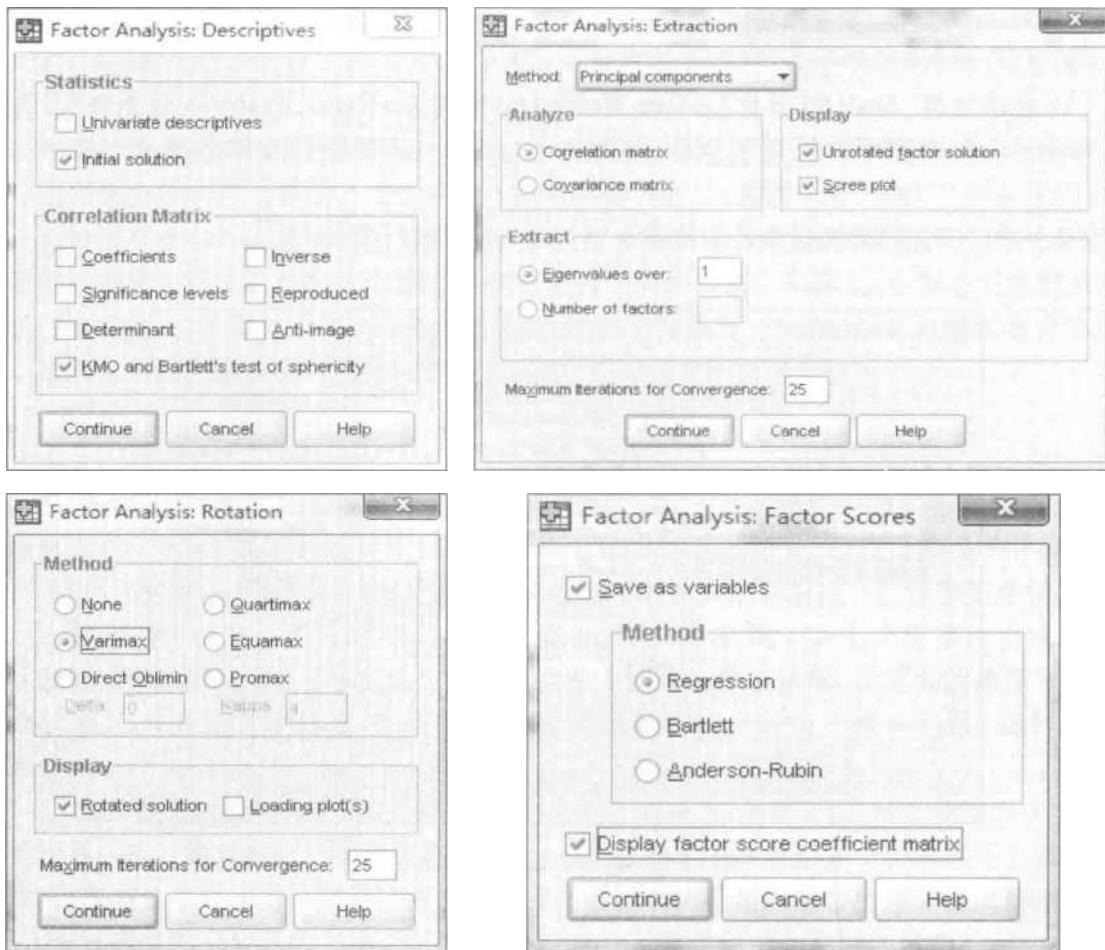


图 14-2 Descriptive, Extraction, Rotation 和 Scores 对话框

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		
Bartlett's Test of Sphericity	Approx Chi-Square	1.387E3
Sig.		.21000

图 14-3 KMO 和 Bartlett 的球形度检验结果截图

如图 14-7 所示。

提取两个公因子的输出结果显示，前两个因子的累计方差贡献率为 75.231%，在可接受范围内。<sup>M</sup> Component Matrix<sup>M</sup> 和 "Rotated Component Matrix" 分别为旋转前、后的因子载荷矩阵，因子载荷反映的是因子与变量之间的相关程度，可以看出，旋转后的两个公共因子与原始变量之间的关系变得清晰化了，因子 1 对除“安全性”变量外的六个变量的因子载荷较大，因子 2 对“安全性”变量的因子载荷较大，因此可以分别将两个因子定义为其他因子和安全性因子，这样较好地区分了原来 7 个评价指标的信息。

进一步，为评价各用户的使用意愿程度，还可以进一步利用“Compute”菜单，以旋转后的

Commonen t	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	4.270	60.996	60.996	4.270	60.996	60.996
2	.996	14.234	75.231			
3	.526	7.514	82.745			
4	.373	5.335	88.080			
5	.357	5.099	93.179			
6	.271	3.873	97.052			
7	.206	2.948	100.000			

Extraction Method: Principal Component Analysis.

图 14-4 “公因子方差”及“解释的总方差”结果截图

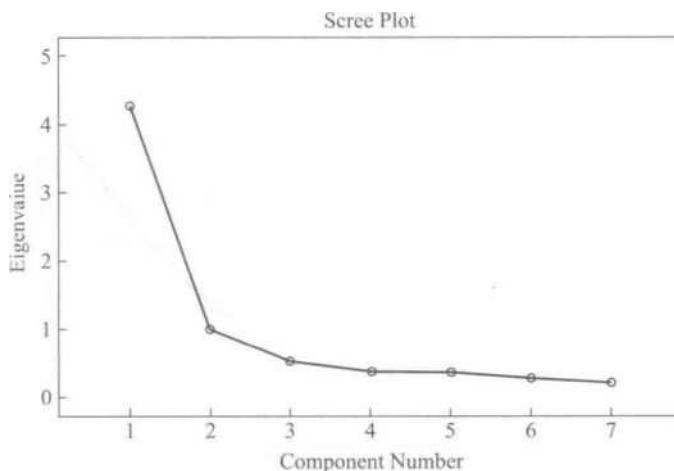


图 14-5 碎石图

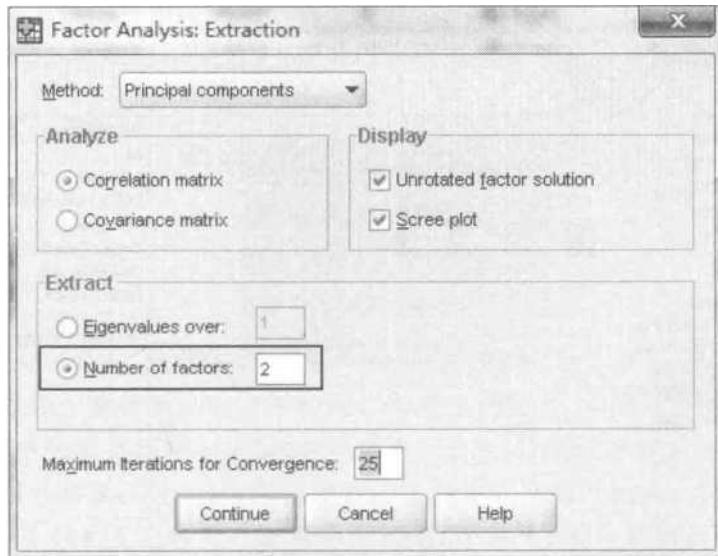


图 14-6 Extraction 子对话框（设置因子个数）

公共因子的方差贡献率为权数，进行归一化后，以因子得分为变量值，计算出各用户的综合得

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% "Variance	Cumulative %	Total	% of Variance	Cumulative %
1	4.270	60.996	60.996	4270	60.996	60.996	4.005	57.218	57.218
2	.996	14.234	75.231	.996	14.234	75.231	1.261	18.013	75.231
3	.526	7.514	82.745						
4	.373	5.335	88.080						
5	.357	5.099	93.179						
6	.271	3.873	97.052						
7	.206	2.948	100.000						

Extraction Method: Principal Component Analysis.

图 14-7 提取两个公共因子的分析结果截图

	Component		Rotated Component	
	1	2	1	2
BAT 认知	.888	.053	.867	.202
感知利益	.854	.053	.834	.191
感知风险	.872	-.054	.821	.299
个人创新	.817	-.169	.735	.394
方便性	.785	.352	.853	-.114
安全性	-.364	.895	-.094	-.962
娱乐性	.758	.186	.780	.037

Extraction Method: Principal Component

Analysis a 2 components extracted

Extraction Method: Principal Component  
Analysis Rotation Method: Varimax with Kaiser

a. Rotation converged in 3 iterations.

分。调用“Compute”过程的对话框如图 14-8 所示，运行结果如图 14-9 所示，其中“FAC1\_1”和“FAC2\_1”分别代表每一条观测在第一个公共因子和第二个公共因子上的得分，“FAC”即为各用户的综合得分。

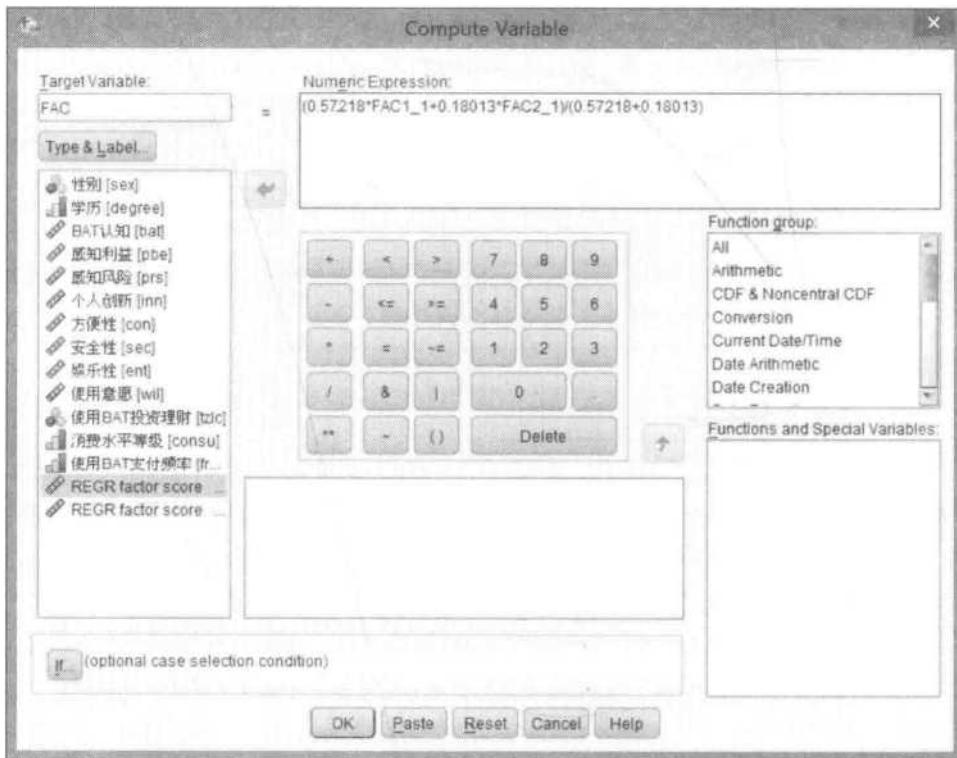


图 14-8 Compute Variable 对话框



	sex	degree	bat	pbe prs	■	see	砌	wil	consu freq	FAC1.1	Fg	FAC
1	女	专科及以下	13	13 11		9	11	8 进行殷理財	5	2	-OOS65	42166 .09
2	女		15	t: 5 15	10	15	7	10 进行對理財	4	2	176696	85080 155
3	男	硕士研究生	8	10 10	&	9	11	9 7 没齊投資理	2	2	132803	-83217 -1 21
4			u	13 14	g	12	a	W 奶投賈理財	4	3	85759	62775 80
5		大学	14	14 13	10	13	s	14 to	4	3	90505	1 560S9 106
6	男	硕士研究生	U	14 1J	9	13	7	13 9 进行投O 财	4	3	80467	98002 86
	女		13	12 13	10	13	8	V) 8 进行頗理財	5	2	69061	.63168 68
8		专科及以下	n	13 11	3	7	8	10 8 没有投资理	4	3	*71130	1 24442 *24
9			13	13 14	9	13	7	13 9 没有投资理	3	4	69930	1 03136 78
10	女		to	11 10	9	11	g	11 7 制技炽財	3	3	46332	20637 -.30
11	女	硕士研究生	14	U 13	9	U	7	14 W 行投貸理	3	5	1 04405	77954 98
12		硕士研究生	S	7 5	5	4	8	6 6 没有投资理	3	3	3 67956	1 06716 2 54
13	具		M	15 13	8	13	7	13 10 没齊投資理	4	4	81459	8325S 82
U	女		15	g 12	10	15	11	15 8 没有投貿理	5	4	1 24436	-1 29666 64
16	女		t3	t4 13	g	14	8	14 g 没削頗理財	4	2	1 04363	28636 86
			12	12 11	8	10	9	10 9 行投漁財	3	5	-40512	32427 23
18			ta	t2 13	7	14	9	13 9 没有投资理	2	2	64201	*44986 38
IS			13	U 14	g	14	7	13 W 送行投貨理	2	2	93980	89264 93
20	女	硕士研交生	11	12 11	8	12	9	10 T w 没有頗理財	3	2	51742	14083S 73
21	女	大学本叫	M	12 12	8	14	7	t2 8 没隽投資瓊	4	3	-22801	03782 16
											44281 <sup>j</sup>	66446 50

图 14-9 综合得分计算结果

### 三、进行因子分析时应注意的问题

第一，由于因子分析对变量之间的相关程度很敏感，为保证分析质量，需要事先进行稳健性比较。另外，由于异常值、缺失值和不规则分布均会对变量间的相关性产生影响，进行因子分析时，还需要事先对数据进行整理，比如剔除异常值、插补缺失值等。

第二，由于小样本数据相关系数的可靠性较差，因子分析要求观测数至少应是变量数的 5 倍，最好是 10 倍以上。

第三，参与因子分析的变量必须是数值型变量，即要求数据是由定距或者定比测量尺度得到的。对于非数值型或非测量型变量，一般可以利用虚拟变量先将其转化成数值型变量，然后再进行因子分析。

## 第一 F

### 聚类分析

在多变量数据分析阶段，通常可以对样本（或变量）进行量化分类，对不同属性的对象进行市场细分，并从整体上把握数据的内在结构。在市场研究中，需要发现不同的客户群，并且通过消费模式刻画不同客户群的特征；在电子商务细分市场中，通过对具有相似浏览行为的客户进行分组，并分析各组客户的特征，可以更好地了解客户的需求倾向，为客户提供更适合的服务。比如，在“BAT 移动支付的用户使用意愿调查”项目中，可以根据影响使用意愿的因素对用户进行群体细分，刻画不同客户群的特征，以便更好地采取有针对性的营销策略。聚类分析就是解决这类问题的有效工具，它在市场细分、消费者行为研究、产品研究、竞争对手研究等市场研究领域均有着广泛的应用。

聚类分析常常需要在因子分析的基础上进行。比如，进行目标市场划分时，由于同一阶层

在价值观念、自我意识、生活方式、消费需求等方面均表现出相似的心理特征，因此可以在因子分析的基础上，通过对消费者的特征进行聚类分析，实现市场细分，从而界定各个目标市场

---

## 一、基本原理

聚类分析（Cluster Analysis）又称群分析或类分析，它是依据某种准则对个体（样本或变量）进行分类的一种多元统计分析方法。聚类分析能够在没有先验知识的情况下，将一批样本数据（或变量）按照它们在性质上的亲疏程度自动进行分类。

聚类分析的基本思想是：认为所研究的样本单位或变量之间存在程度不同的相似性，根据一批样本单位的多个观测指标，具体找出能够度量样本或指标之间相似程度的统计量，以此作为划分类型的依据，把一些相似程度较大的样本（或变量）聚合为一类，把另外一些彼此之间相似程度较大的样本（或变量）聚合为另一类，直到把所有的样本（或变量）聚合完毕。聚类分析是依据分析数据本身所具有的定量特征来对大量的数据进行分组归类，以了解数据的内在结构，这样的分类更具客观性，并能反映出事物的内在必然联系。好的聚类结果是类内相关性尽量大，类间相关性尽量小。

聚类分析分为Q型聚类和R型聚类。Q型聚类是对样本进行分类处理，R型聚类是对变量进行分类处理。R型聚类本质上也是一种降维技术，实际应用中较为常见的是对样本进行聚类。因此，本节将重点讲解Q型聚类，关于R型聚类的相关内容请读者参阅相应的多元统计分析教材。

聚类分析实质上是寻找一种能客观反映研究对象之间亲疏关系的统计量，相似性测度的三种主要方法：相关测度、距离测度、关联测度，它们从不同角度测度了研究对象的相似性，其中相关测度和距离测度适用于定距或定比变量，关联测度适用于定类或定序变量。

聚类方法有很多种，不同的方法适合解决不同类型的问题。实际应用中常用的聚类方法有系统聚类法（Hierarchical Cluster）和K均值聚类法（K-Means Cluster）。

系统聚类法的基本思想是：距离相近的研究对象先聚成类，距离较远的后聚类，过程一直进行下去，每个研究对象总能聚到合适的类中。应用系统聚类法时，首先需要对数据进行变换，变换的方法有平移变换、极差变换、标准化变换、对数变换等；然后再选取聚类方法，主要有最短距离法、最长距离法、中间距离法、重心法等。这里需要注意的是，不同的聚类方法结果不一定完全相同，一般只是大致相似，如果差异很大，则应仔细考查，找到问题所在。

K均值聚类法是一种快速聚类法，适用于大样本并且均为连续型变量的情形。K均值聚类的基本思想是将每一个样本分配给最近重心（均值）的类中。其实质是分步聚类，即先选定一批凝聚点，让研究对象向最近的凝聚点靠拢，形成初步分类；然后再对凝聚点的选点进行调整，重复前一步骤，直到得到比较合理的结果为止。

K均值聚类法和系统聚类法一样，都是以距离的远近亲疏为标准进行聚类。两者的不同之处在于系统聚类对不同的类数产生一系列的聚类结果，而K均值法只能产生指定类数的聚类结果。具体类数的确定，离不开实践经验的积累。系统聚类法以一部分样本为对象进行聚类的结果可作为K均值法确定类数的参考。

## 二、应用实例

### 应用实例 14-2

在应用实例 14-I 中，我们对影响用户使用意愿的七个因素进行了因子分析，提取了两个公因子，分别是安全性因子和其他因子。在此基础上，结合消费水平等级和 BAT 支付使用频率两个变量，对用户群体进行划分，以发现各群体的使用倾向特征。

分析：消费水平等级和 BAT 支付使用频率两个变量均为定序变量，在本案例中认为这两个变量对应的取值为得分，按照连续变量进行分析。由于四个变量的取值存在数量级的差异，在聚类分析前需要进行标

准化处理。此外，该数据集共有 340 条观测，样本容量较大，可以采用 K 均值聚类法。具体过程利用 SPSS16.0（英文版）软件中的“K-Means Cluster Analysis（K 均值聚类分析）”菜单实现。系统聚类法的 SPSS 软件实现过程与 K 均值聚类法相似，这里不再介绍，感兴趣的学可参照相关 SPSS 软件数据分析书籍。

### 1. 操作步骤

- (1) 打开数据集 Example13-1。
- (2) 首先利用“descriptives”菜单对变量进行标准化，标准化后的变量分别为“Zconsu（标准化后的消费水平等级）”“Zfreq（标准化后的使用频率）”“ZFAC1\_1（标准化后的其他因子得分）”以及“ZFAC2\_1（标准化后的安全性因子得分）”。
- (3) 依次选择“Analyze（分析）→ Classify（分类）→ K-Means Cluster Analysis（K 均值聚类分析）”，展开“K-Means Cluster Analysis（K 均值聚类分析）”对话框，如图 14-10 所示。从变量列表中选择“Zconsu”“Zfreq”“ZFAC1\_1”以及“ZFAC2\_1”变量进入“Variables（变量）”框，在“Number of Clusters】聚类数】文本框输入需要划分的组数。本例经过不断尝试，最终选择的合适组数为 3 组。

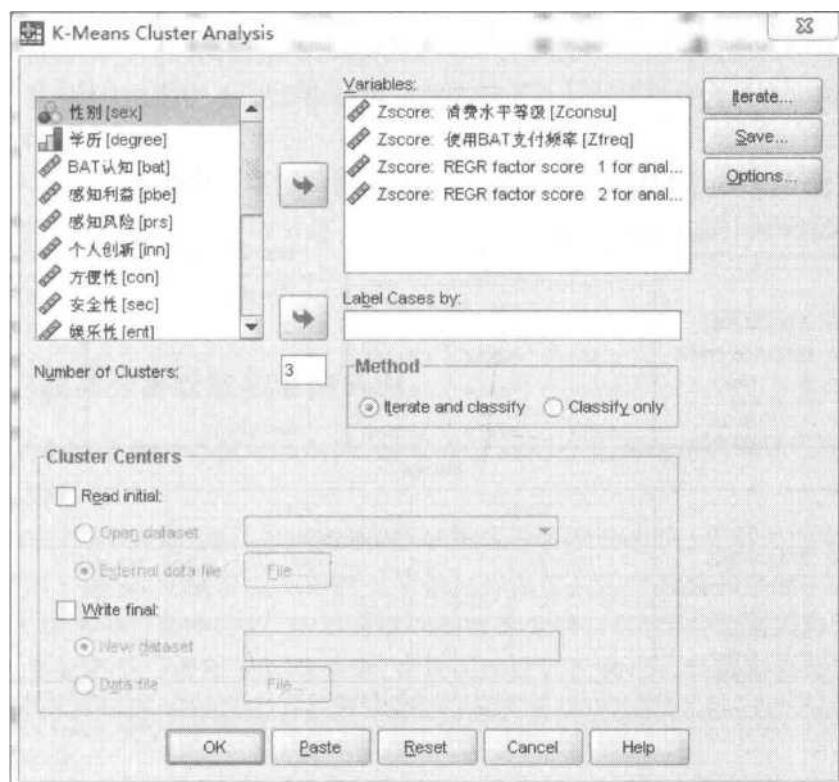


图 14-10 K-Means Cluster Analysis 对话框

- (4) 单击“Save（保存）”按钮，弹出“K-Means Cluster Analysis: Save New Variable（K 均值

聚类：保存新变量”对话框，选中“Cluster Membership（聚类成员）”以及<sup>44</sup>“Distance from cluster center（与聚类中心的距离）”复选框（如图 14-11 所示），在原始数据文件中逐一显示分类结果以及观测与所属类中心的距离，单击“Continue（继续）”按钮，返回“K-Means Cluster Analysis（K 均值聚类分析）”对话框。

(5) 继续单击“Options（选项）”按钮弹出“K-Means Cluster Analysis: Options（K 均值聚类：选项）”对话框，在“Statistics（统计量）”选项组中选中“ANOVA table（ANOVA 表）”复选框（如图 14-12 所示），对聚类结果进行方差分析，单击“Continue（继续）”按钮返回“K-Means Cluster Analysis（K 均值聚类分析）”对话框。

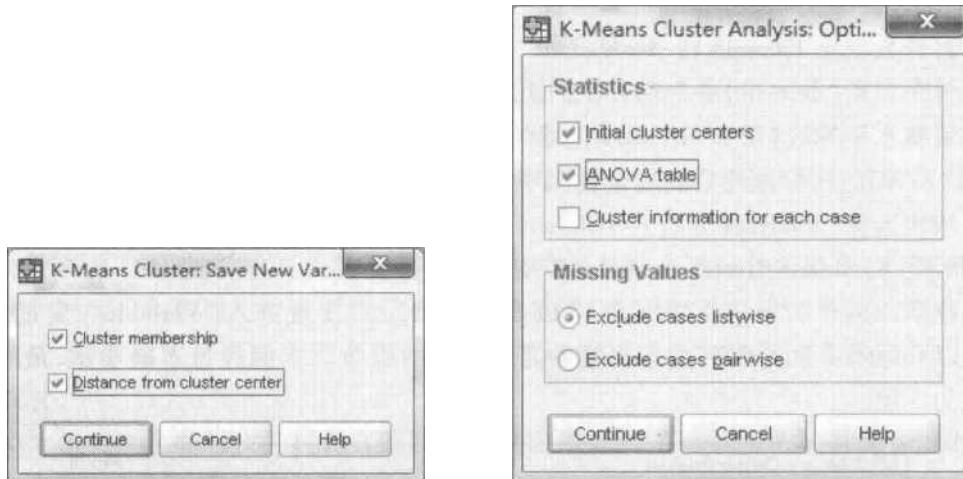


图 14-11 Save 子对话框

图 14-12 Option 子对话框

(6) 最后，单击“OK（确定）”按钮，即可得到 K 均值聚类分析结果（如图 14-13 所示）。

Final Cluster Centers				Number of Cases in each Cluster		
	Cluster					
	1	2	3	Cluster 1	2	3
Zscore 消费水平等级	-.67769	1.13890	.06901	175.000		
Zscore: 使用 BAT 支付频率	-51504	-04824	1 41724		100.000	
Zscore: REGR factor score 1 for analysis 1	-.36819	.19464	.67716		67.000	
Zscore: REGR factor score 2 for analysis 1	-.32814	.62379	-.06620		342.000	
ANOVA						
	Cluster		Error		F	Sig.
	Mean Square	df	Mean Square	df		
Zscore: 消费水平等级	105.199	2	.388	339	271.003	.000
Zscore: 使用 BAT 支付频率	90.613	2	.471	339	192.259	.000
Zscore: REGR factor score 1 for analysis 1	29.117	2	.837	339	34.804	.000
Zscore: REGR factor score 2 for analysis 1	29.023	2	.837	339	34.603	.000

The F tests should be used only for descriptive purposes because the clusters have been chosen to maximize the differences among cases in different clusters. The observed significance levels are not corrected for this and thus cannot be interpreted as tests of the hypothesis that the cluster means are equal.

图 14-13 K 均值聚类分析输出结果截图

## 2. 解读输出结果

输出结果由四部分组成，如图 14-13 和图 14-14 所示。第一部分是最终的聚类中心，根

AS XJRT ORBS Xrootdf UMict **									
-■鼻必 M “e 懶雨關”@ ·									
1—									
1	.FAC3.1	F*C	Zwu	列 a !	—ZFAC2.1	或 2			
	*oocess 17G696	G07 0 85060	1 84305 0 96277	4)5365 OSK385	-0OOOB55 1	0 4216E aasosc	2 五	0 92999 1 68061	
3	.32B03	4H32V	-091	0 79B46	*0533ES	-1 32003 0 86759	1 0 6277!	1 09075 0 7272B	
4	0 05759	Q 62776	060	096377	0 20080	0 90505 1 56D6S	2 四	1 21742 0 76523	
5	0 90605	1 66069	oce	096277	0 20080	0 8D4G7 0 9000:	2 四	1 13519 1 018Z*	
	0 8CM87	0 9800?	0 64	0 96277	0 20080	0 8D4G7 0 9000:	2 四	1 13519 1 018Z*	
7	O6806t	0 63158	asi	1 84339	43.55365	0.69061 0.6316E	2 四	0 76523 1 018Z*	
8	-071130	1 24442	-018	0 gk	0.20060	■071130 1 24442	2 四	1 13519 1 018Z*	
9	0 69930	t 0313B	069	0 08236	0 96644 0 930	1 03136 1 201 中	3 三	1 13519 1 018Z*	
ID	*046332	0 20637	-023	0 08236	0 XD60 0 46332	0 20837 1	1 18445 0 96661		
11	104405	0 77954	074	0 06216	17100B 1U4406	0 9594	3 三	374757	
12	-3 67956	t 05716	-1.91	0 08216	0 20060	■.B7966 1 0671E	1 一		
13	0 81459	053256	0 62	0 96377	0 95544	0 81 <59 0 8375e	2 三	1 20711 0 96661	
14	1 24436	1 29666	0 48	1 64339	0 96644	1 24436 1 29666	3 三	Z260D8	
15	1 0063	0 2B636	0&	0 96277	0 55385	1 04363 0 26636	2 二	1 05973 1 18311	
16	*0 40512	0 32427	-0.17	0 08218	1 71000	■0 40519 0	i 1 02177		
17	0 64201	<0 44988	0 29	0 79B46	05S3B5	0B42Q1 0 44986	i 1 02177		
IB	0 93960	0 89264	070	4179646	-0 56385	0 93960 D9SU64	T 1 79493		
19	0.51742	1 40B35	a 55	0 08216	-0 56385	031742 1 4CB35	? 1 43690		
20	*0 22801	003782	*012	*1 62we	*05365	0ZWH 0 03782	1 1 07324		
21	0.44291	0GB446	0 38	0 96277	02CX360 0.44281	0 6844c 0 03782	R 0 38961		
22	*0 26287	0 651^	QOS	1 B4389	0 65385	■0262B7 0 55159	W 0 99626		
23	0 07B13	0 47322	0 19	0 96277	L7UXB	0 17613 0 47322	1 19133		

图 14-14 K 均值聚类分析原数据集显示结果截图

据每一类的聚类中心，可以判断该类群体的特征。以第二类为例，该类用户群体具有“高消费 水平、中等支付频率、非常注重安全性、一般关注其他”的特征。

第二部分是方差分析表，F 检验仅用于描述性分析目的，无法将其解释为是对聚类均值相 等这一假设的检验。通过方差分析表可以看出，4 个变量均对分类贡献显著。

第三部分是每一类包含的样本数，结果显示，第一类包含 175 个样本，第二类包含 100 个 样本，第三类包含 67 个样本。此外，有两个样本存在缺失值，没有纳入分析。

第四部分是聚类分析后的数据集(见图 14-14)，“QCL\_1”和“QCL\_2”变量表示每个观测 的分类情况。其中，“QCL\_1”变量表示该观测属于的类别，“QCL\_2”表示该观测与其所属类别 中心之间的欧氏距离。

扩展：进一步还可以利用每个样本所属的类别变量与划分等级后的使用意愿变量进行列 联分析，以探究每一类群体的使用意愿是否存在差异，以便采取更有针对性的提升使用率的 措施。

### 三、进行聚类分析时应注意的问题

第一，如果聚类变量的计量单位不同，应事先对其进行标准化后再进行聚类分析，否则可能会 得出错误的结论。

第二，选择不同的变量组合、聚类方法、测度方法及标准化方法，所得出的聚类过程及结果 可能会有 所不同。其中，变量组合的差异对聚类过程及结果的可能性影响最大。

第三，对大样本进行聚类分析，当要聚成的类数确定时，可使用快速聚类方法，其特点是处 理速度快。但快速聚类只适用于连续变量，并且采用欧式距离平方法进行聚类；如果聚类变量 是离散变量，或采用其 他距离测度方法进行聚类，则需要使用系统聚类方法(Hierarchical Cluster)进行聚类分析 c，

## 第三节

### 对应分析

在社会、经济以及其他领域中，进行数据分析时不但要处理因素与因素、样本与样本之间的 关系，往

往还需要处理因素与样本之间的相互关系。例如，在“BAT 移动支付的用户使用意愿调查”项目中，我们想探究什么消费水平的群体具有更强的 BAT 支付使用意愿。这类问题可以利用对应分析的方法，将样本和变量放在一起进行分类、作图，从而更好地描述两者之间的关系。

## 一、基本原理

对应分析 (Correspondence Analysis)，又称相应分析，是利用加权主成分分析法来描述两个或多个分类变量各水平间相关性的分析方法。对应分析方法通过分析由定性变量构成的交互汇总表，来揭示变量间、样本间以及变量与样本间的对应关系，因此可以弥补因子分析、聚类分析等统计方法只能单独研究变量间或者样本间关系的不足。对应分析在品牌形象研究、市场细分和市场定位研究等领域都有着非常广泛的应用。

对应分析的基本思想是将一个列联表的行和列各元素的比例结构以点的形式在较低维的空间中表示出来，其特点是能把众多的样品和众多的变量同时作在同一张图上，将样品的大类及其属性在图上直观展现出来。对应分析方法不用进行因子选择和因子轴旋转等复杂的数学运算及中间过程，而是从因子载荷图上对样品进行直观分类，并能够标示出分类的主要参数（主因子）以及分类的依据。

对应分析图是对应分析方法的主要应用形式。在对应分析图中，每一个散点代表某个变量的一个类别，该类别所包含的信息由两个维度 (dim1, dim2) 坐标值反映。根据对应分析图中各个散点分布距离的远近，可以直观解释不同变量之间以及不同变量的不同类别之间的对应关系。如果同一变量的两个类别在某个维度上分布较近，表明这两个类别在该维度上区别不大；如果不同变量的两个类别在某个维度上分布较近，表明这两个类别彼此联系，相关性较大。在以“dim1”和“dim2”分别作为横轴和纵轴的直角坐标系内，每个变量的一个类别就是一个点，各变量间的相关距离越近，表示相关性越大；相关距离越远，表示相关性越小。因此相关距离相当于相关系数。两变量之间距离的计算公式如下：

$$D(A-B) = (\sqrt{(x_A - x_B)^2 + (y_A - y_B)^2})^2 \quad (14.5)$$

其中为  $A$ 、 $B$  两点间的距离， $x_A$ 、 $y_A$  分别为  $A$  点的横坐标和纵坐标， $x_B$ 、 $y_B$  分别为  $B$  点的横坐标和纵坐标。

通过对对应分析，可以把变量和样本同时反映到相同坐标轴的一张图形上，这样便于对变量和样本一起进行分类、作图和解释。如果将其中的样本也用变量代替，对应分析就可以用于分析两个变量之间的内在关系。

## 二、应用实例

### 应用实例 14-3

在“BAT 移动支付用户的使用意愿调查”项目中，除了应用相关分析和多元线性回归模型探究一些因素对使用意愿的影响，还可以采用对应分析方法将其中某个重点关注的因素和使

用意愿变量放在一起进行分类、作图，以比较不同群体的 BAT 支付使用意愿是否存在差异。本例中以消费水平等级因素为例，分析不同消费水平群体的使用意愿有何差异。

分析：在本例中，只关注不同群体使用意愿的等级，将原始数据集 Example13-1 中的使用意愿得分划分为三组，分别是低使用意愿（“5 分以下（含 5 分）”）、中等使用意愿（“6-8 分”）和高使用意愿（“9-10 分”），分组后的变量为“G\_will（使用意愿等级）”，分组的具体实现过程在应用实例 11-1 中有所涉及，不再赘述。将消费水平等级变量和分组后的使用意愿等级变量看作定类变量，采用对应分析方法探究不同消费水平群体的使用意愿有何差异。具体实现过程利用 SPSS16.0（英文版）软件的“Correspondence Analysis\*\*菜单实现。

### 1. 操作步骤

- (1) 打开数据集 Example13-1
- (2) 依次选择“Analyze（分析）—Data Reduction（降维）—>Correspondence Analysis（对应分析）”，展开“Correspondence Analysis（对应分析）”对话框，然后将“（consu）消费水平等级”选入“Row（行）”，再点击“Define Range（定义范围）”来定义范围 1（最小值）到 5（最大值）；将“（G\_will）使用意愿等级”选入“Column（列）”，并定义其范围 1 到 3，如图 14-15 和图 14-16 所示。

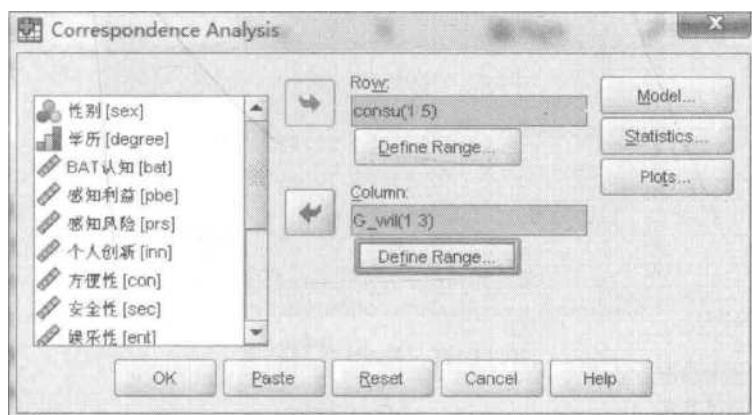


图 14-15 Correspondence Analysis 对话框

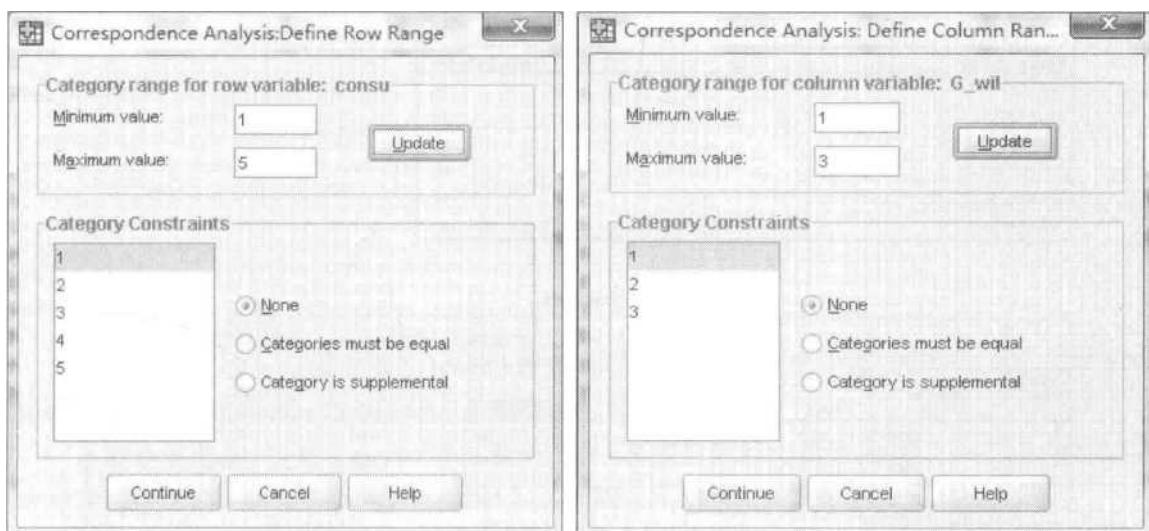


图 14-16 定义分组子对话框

(3) 继续单击“Model (模型)”按钮，在出现的对话框中选择数据标准化方法，如图 14-17 所示，由于本例中的两个变量均为分类变量，因此“距离度量”点选“Chi square (卡方)”单选按钮，下面“Standardization Method (标准化方法)”选项被激活，本例选择默认的“Row and column means are removed (删除行列均值)”，读者也可尝试使用其他方法，点击“Continue (继续)”按钮返回主对话框。

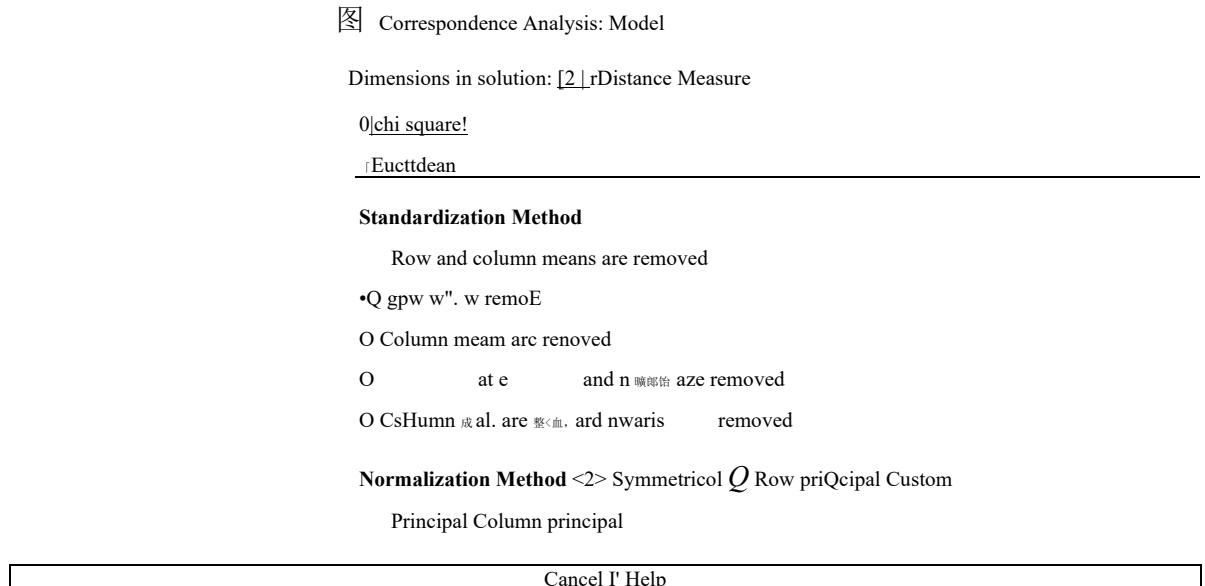


图 14-17 Model 子对话框

(4) 最后点击“OK (确定)”按钮，运行后即可得到对应分析结果（如图 14-18、图 14-19、图 14-20 所示）。

Correspondence TWe		使用意匾等级				
		福			Active Marain	
• 甘蓄水平盖加	600 元以下	2	14	2	16	
	600-1200zt	4	98	36	138	
	1200-180 阮	5	62	24	91	
	1800-2400^	0	20	30	50	
	24。阮及以上	0	29	17	46	
	Active Margin	11	223	109	343	

Summary									
Dime nsio n	Singular Value	Inertia	Chi Sauare	Sia.	Proportion of Inertia		Standard Deviation	Confidence Singular Value	
					Accounted for	Cumulative		Correlatlon	2
1	.289	.083			897	897	.050		270
2	.098	.010			103	1.000	.058		
Total		.093	31 884	000 <sup>1</sup>	1 000	1.000			

a 6 degrees of freedom

图 14-18 对应分析的对应分析表和模型概况截图

Overview Row Points <sup>a</sup>									
	Mass	Score in Dimension		inertia	Contribution				
		1	2		Of Point to Inertia of Dimension		Of Dimension to Inertia of Point		
					1	2	1	2	
6。既以下	.052	-1.025	.721	.019	.191	.278	.856	.144	1.000
600-1200 元	.402	-.194	-.230	.006	.052	.217	.678	.322	1.000
1200-1800 <sup>7</sup> 元	.265	-.276	.226	.007	.070	.139	.814	.166	1.000
1800-2400 元	.146	1.130	.323	.055	.645	.155	.973	.027	1.000
2400 元及以上	.134	.300	-.392	.006	.042	.210	.634	.366	1.000
Active Total	1 000			.093	1 000	1 000			

a. Symmetrical normalization

Overview column Points <sup>b</sup>									
	Mass	Score in Dimension		inertia	Contribution				
		1	2		Of Point to Inertia of Dimension		Of Dimension to Inertia of Point		
					1	2	1	2	
來用意慮等	.032	-1.324	1.536		.195	.773	.687	.313	1 000
弱	.650	-.297	-151	.024 .018 .05	.199	.151	.920	.080	1.000
一般	.318	.742	153		.606	.076	.986	.014	1 000
強 Active	Total	1 000		.093	1 000	1 000			

a Symmetrical normalization

图 14-19 对应分析结果的行变量及列变量信息截图

### Row and Column Points Symmetrical Normalization

。弱

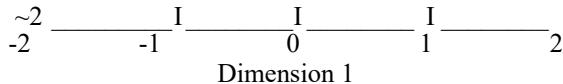
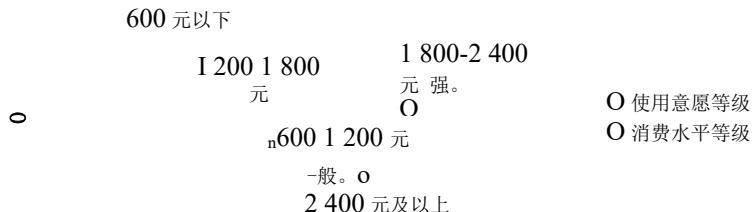


图 14-20 对应分析图

## 2. 解读输出结果

对应分析的输出结果主要由对应分析表、模型概况、行变量信息、列变量信息及对应分析 图等部分组成（见图 14-18. 图 14-19 和图 14-20）。

由模型概况表（见图 14-18）可以看出，第一维度的解释比例达到 89.7%，因此可以选择一 维度进行分析。

从行变量信息表（见图 14-19 中的 Overview Row Points 表）来看，各消费水平等级群体所 占比例有所差别 600-1 200 元”群体所占比例最高, 为 40.2%, “600 元以下”群体所占比例 最低, 为 5.2%, 其他消费水平群体所占比例差别不大。根据每一维度中所得的分数（score in dimension），可以看出行变量值对每一维度特征根的贡献。从列变量信息表（见图 14-19 的 Overview Column Points 表）来看，使用

意愿等级为“一般”的被调查者所占比重相对较高、为 65.0%。

从对应分析图（见图 M-20）可以看出，消费水平在“600—1 200 元”和“1 200—1 800 元”等 级的用户的使用意愿偏向于“一般”，消费水平在“1 800—2 400 元”等级的用户具有较强的使 用意愿，而消费水平在“600 元以下”和“2 400 元及以上”的用户群体使用意愿则比较分散， 同时，也可以看出，在不考虑极低消费水平和极高消费水平的用户群体时，消费水平越高，BAT 支付的使用意愿越强。

### 三、进行对应分析时应注意的问题

第一，由于对应分析只是一种描述性的统计分析方法，它虽然可以揭示变量间的对应关 系，但不能用于相关关系的假设检验。

第二，对应分析是一种减少维度的方法，在分析过程中，最终用几维进行分析需要由研究 者自己决定。

第三，对应分析对极端值比较敏感，极端值的存在会对分析结果造成较大影响。因此进行 对应分析时，需要事先对数据进行整理，比如剔除异常值、插补缺失值等

## 第四节

### 多维标度分析

聚类分析是根据预设维度上数据间的相似性进行对象分组，多维标度分析（Multidimensional Scaling）则在一定程度上是相反的过程。多维标度分析是一种将多维空间的研究对象（样本或变量）简化到低维空间进行定位、分析和归类，同时又保留对象间原始关 系的数据分析方法。在市场营销调研中，多维标度法的用途十分广泛。比如，被用于确定空间 的级数（变量或指标），以反映消费者对不同品牌的认知，并标明各品牌在消费者心目中的位 置。本节将在介绍多维标度分析基本理论的基础上，采用此方法分析消费者在 10 种不同饮料 感知上的相似程度 Q

#### 一、基本原理

多维标度分析的基本思想是，在一个确定维数的空间中估计一组样品的坐标，其基本数据 是配对样品间的距离◎多维标度模型包括可以用来计算距离，并且使距离与实际数据相关联 的两因子和三因子的度量以及非度量模型。多维标度过程的数据包括样品间的一个或者多个 对称或者不对称方阵，也称为相似性数据，该方阵名义上是距离阵，但并不一定是 “个点的距 离，即不是通常所理解的距离阵。于是首先需要将距离阵的概念加以拓展。

一个几矩阵  $D = (d_{ij})$ ，若满足  $D' = D, d_{ij} = O, d_{ii} \neq O \quad (i, j = 1, \dots, n)$ ，则称  $D$  为距离阵°这样定义的距离并不一定满足通常距离的三角不等式。

对于距离阵。二 (4)，多维标度法的目的是要寻找  $k$  和  $R^*$  中 “个点 [, …次 “，用品表示

与  $X$  )的欧氏距离,  $D = (d_{ij})$ ，使得方与  $o$  在某种意义上相近。在实际中，常取  $s 1, 2$  或  $3$ ° 令  $J = (x_1, x_2, \dots, x_n)$ ，为方便叙述，称  $X$  为  $D$  的拟合构造点。当  $D = D'$  时， $X$  称为  $D$  的构造点。

需要指出的是，多维标度法的解并不唯一。若  $X$  是解，令

$$Y = FX + a \quad (14.6)$$

式中，「为正交阵，。为常数向量，则  $Y = (y_1, y_2, \dots, y_n)$  也是解，因为平移和正交变换不改 变欧

氏距离。

由于多维标度法的求解原理相对复杂，这里不再介绍，只结合实例介绍如何采用 SPSS 软件进行多维标度分析，有兴趣的读者可参考相关多元统计分析教材。

## 二、应用实例

### 应用实例 14-4

在一次调查中，收集了多位消费者对 10 种不同饮料 (A~J) 的差异性评分数据，1 分表示 差异较小，9 分表示差异较大，从 1 分到 9 分的差异程度逐渐增加。图 14-21 为某位消费者对 10 种不同饮料的差异性评分数据（见数据集 Example14-1）。采用多维标度法分析消费者在 10 种不同饮料感知上的相似程度，具体过程利用 SPSS16.0（英文版）软件中的“Multidimensional Scaling（多维尺度）”菜单实现。

	V1	A	B	C	D	E	F	G	H	I	J
1	A	0									
2	B	4	0								
3	C	7	5	0							
4	D	8	3	7	0						
5	E	7	3	5	5	0					
6	F	6	4	3	9	3	0				
7	G	5	8	7	5	7	7	0			
8	H	6	8	7	8	6	5	6	0		
9	I	6	2	6	6	2	3	7	7	0	
10	J	2	6	7	5	6	5	3	5		0

图 14-21 消费者对 10 种不同饮料的差异性评分数据 SPSS 格式截图

### 1. 操作步骤

(1) 打开数据集 Example 14-I。

(2) 依次选择“Analyze(分析)→Scale(度量)→Multidimensional Scaling( ALSCAL ) (多维尺度)”，展开“Multidimensional Scaling(多维尺度)”对话框。将 A~J 共 10 个变量选入“Variases(变量)”框（见图 14-22）。

(3) 继续单击“Model(模型)”按钮，弹出“Multidimensional Scaling: Model (多维尺度：模型)”对话框，在“Level of Measurement(度量水平)”选项组中选中“Ordinal(序数)”单选按钮，并选中下面“Untie tied observations(打开结观察值)”复选框，返回主对话框。继续单击“Options(选项)”按钮、弹出“Multidimensional Scaling: Options(多维尺度：选项)”对话框，在“Display(输出)”选项组中选中“Group plots(组图)”复选框，再返回主对话框（如图 14-23 所示）。

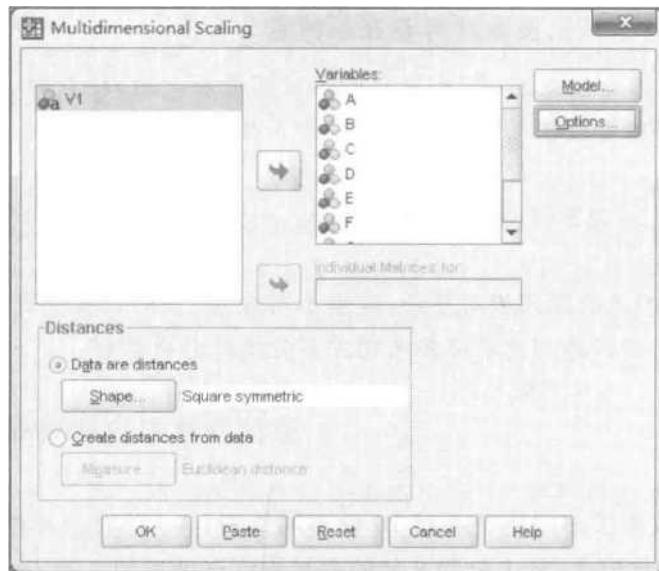


图 14-22 Multidimensional Scaling 对话框

**Multidimensional Scaling: Model**

**Level of Measurement**
 **Ordinal**  
 **Untie tied observations**  
 **Interval**  
 **Ratio**

**Conditionality**
 **Matrix**  
 **Row**  
 **Unconditional**

**Dimensions**  
 **Minimum:** 2    **Maximum:** 2

**Scaling Model**
 **Euclidean distance**  
 **Individual differences Euclidean distance**  
 **Allow negative subject weights**

**Multidimensional Scaling:**

Display  
 **Brouppiois**  
 Individual street plots  
 Data matrix  
 Model and options summary

rCriteria –  
 S-stress convergence: 0.001  
 Minimum s-stress value: 0.005  
 Maximum iterations 130  
 Treat distances less than 0 as missing

**Continue** | **Cancel** | **Help**

图 14-23 Model 子对话框和 Option 子对话框

(4) 最后, 单击“OK (确定)”按钮, 即可得到多维标度分析结果(见图 14-24 至图 14-27)。

## 2. 解读输出结果

图 14-24 为统计量 Stress 和 RSQ 的具体解释和计算结果。RSQ 即决定系数, 表示总变异 中能够被相对空间距离解释的比例; Stress 是 K 压力指数: 这里 Stress 值为 0.04509, 低于 5%, RSQ 的值为 0.99142, 非常接近 1, 可以认为关于 10 种饮料的多维标度模型的拟合效果非常好。

图 14-25 显示的是 10 种饮料在二维空间上的一种坐标值, 根据多维标度法解的概念和有

Stress and squared correlation (RSQ) in distances  
 RSQ values are the proportion of variance of the scaled data (disparities)  
 in the part it ion (row, matrix, or entire data) which  
 is accounted for by their corresponding distances.

Stress values are Kruskal\* s stress formula 1.

For matrix  
 Stress .04509      RSQ = .99142

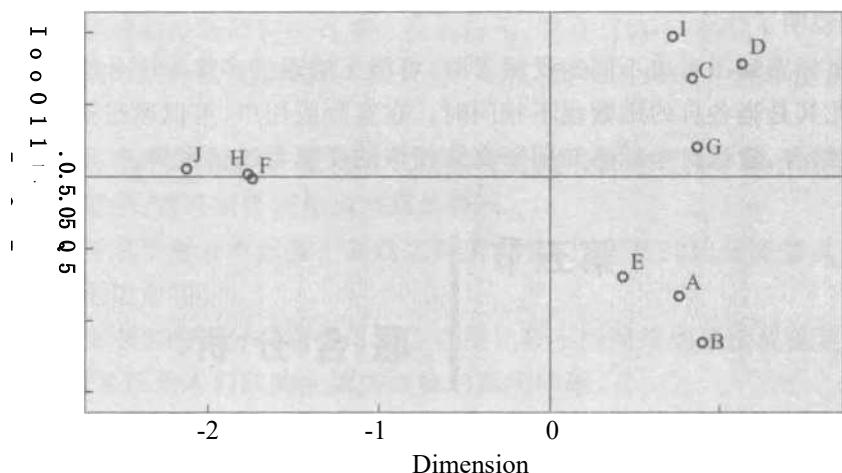
图 14-24 压力指标检验结果截图

		Dimension	
Stimulus Number	Stimulus Name	1	2
1	A	.7560	-.8309
2	B	.8930	-1.1536
3	C	.8345	.6770
4	D	1.1298	.7792
5	E	.4320	6946
6	F	-1.7410	0204
7	G	.8621	.2023
8	H	-1.7651	.0164
9	I	.7210	.9718
10	J	-2.1222	.0529

图 14-25 拟合点在二维标度中的坐标结果截图

图 14-26 坐标感知图

Euclidean distance



关性质，它的解不是唯一的 图 14-26 是多维尺度中输出的 10 种饮料二维空间匹配图，反映 的是各饮料根据实际评分数据计算得到的相应距离在空间中的排列结果。可以看出 I、C、D、 G 四种饮料相对较接近，A、B、E 三种饮料较相近，J、H、F 三种饮料较相近。需要注意的是，图 14-26 仅是源于一个消费者感知的 10 种饮料之间的相关性，因此并不一定反映其他消费者的

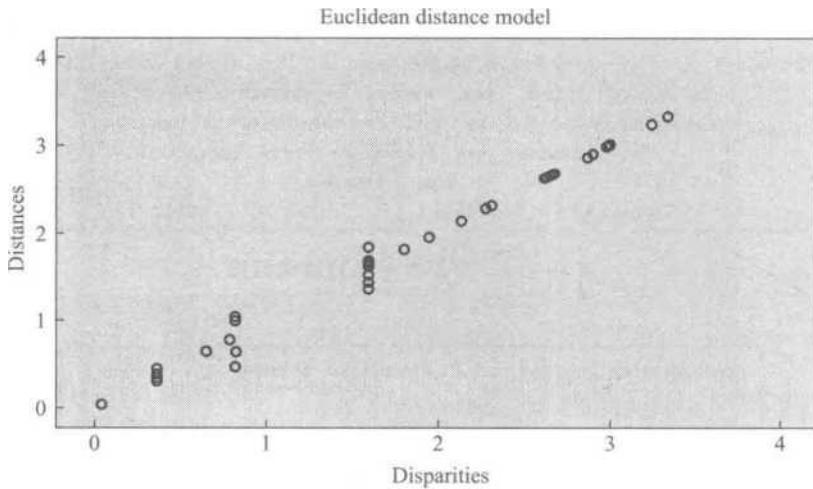


图 14-27 拟合散点图

观点。分析其他消费者的消费评价感知特征，可以依据该案例的分析过程逐一进行。

图 14-27 是欧氏距离模型线性拟合散点图，提供的是原始数据的不一致程度和用线性模型计算出来的欧氏距离间的散点图。如果模型的拟合程度较高，则所有散点应在一条直线上。由图可见，模型对变换后数据的解释程度是非常高的。

### 三、进行多维标度分析时应注意的问题

第一，多维标度分析得到的坐标感知图，显示了每两个对象间的距离和排序，其所蕴含的实际意义，取决于两个坐标轴所代表的含义，即需要给坐标轴命名。而标识坐标轴具有主观性，需要沿着每一个维度检查研究对象的相对位置，并且根据对研究对象先前的认识，推断这些维度最可能说明了什么。

第二，当坐标感知图对应不同的受测者时，可能无法通过多维标度分析的结果得到一个普适性的推论，尤其是当各自的维数也不相同时。在实际应用中，可以通过把聚类分析方法和多维标度分析相结合，尝试对坐标感知图形高度相似的受测者进行聚类。

## 第五节

### 联合分析

在市场调查中，有时需要分析消费者对商品不同属性的偏好结构，这时往往需要用到联合分析方法。联合分析是对人们购买决策的一种现实模拟，因为人们在实际决策过程中，由于价格等原因，需要综合考虑产品的多个特征，往往要在满足一些要求的前提下，牺牲其他部分特性，是一种对特征的权衡与折中。

通过联合分析，可以模拟人们的抉择行为，并预测不同类型人群抉择的结果。在现代市场研究的各个方面，如新产品的概念筛选和开发、竞争分析、产品定价、市场细分等领域，都可以见到联合分析的应用。本节将结合餐厅调查的案例，讲解联合分析的基本原理以及利用 SPSS

## 一、联合分析基本原理

联合分析(Conjoint Analysis)是在已知受测者对某一受测对象整体评估结果的情形下，通过分解受测者的整体反应，来估计其偏好结构的一种分析方法，其中，受测对象由研究人员事先根据某种因子结构加以设计。联合分析首先是通过假定产品具有某些特征，对现实产品进行模拟，然后让消费者根据自己的喜好评价这些虚拟产品，并采用数理统计方法将这些特性与特征水平的效用分离，从而对每一特征以及特征水平的重要程度做出量化评价的方法。

进行联合分析的一般步骤如下：

1. 确定产品特征与特征水平进行联合分析首先要识别能够显著影响消费者购买行为的产品或服务的特征，一个典型的联合分析一般可以包含6~7个显著因素。确定了特征之后，还应该确定这些特征恰当的水平，例如内存容量是台式电脑的一个重要特征，目前市场上台式电脑的内存容量主要有2G、4G、8G和16G四种，这四种内存容量即为台式电脑内存的特征水平。特征与特征水平的个数决定了分析过程中需要进行估计的参数个数。

2. 产品模拟。联合分析采用正交设计方法，将所考虑产品的所有特征与特征水平进行组合，生成一系列虚拟产品。

3. 数据收集，由受访者对虚拟产品进行评价，通过打分、排序等方法调查受访者对虚拟产品的喜好、购买的可能性等。

4. 计算特征的效用从收集的信息中分离出消费者对每一特征以及特征水平的偏好值，即特征的“效用L对于任一特性的某一水平，效用值越高，消费者对包含该水平的描述性形式的偏好越大。

5. 市场预测。利用效用值预测消费者将如何在不同产品之间进行选择，从而决定应采取的措施

联合分析在各领域的市场研究中有着广泛的应用，它具有以下优点：

(1) 联合分析既可以分析度量属性(如价格)的重要性，又可以分析非度量属性(如品牌名称)的重要性；

(2) 收集资料的程序简单易行，受测者只需要对受测对象进行排序或者评分，联合分析对受测者只做很少的要求，就可以得到相当可靠的资料；

(3) 联合分析要求受测者考虑各个属性之间的兑换，比直接询问受测者其理想点的属性水平及属性重要性更切合实际；

(4) 联合分析所求出的成分效用值可用于直接比较不同尺度的属性或是更基本的非度量属性，而这些比较因素正是人们选购决策所面临的真实问题。

## 二、应用实例

### 应用实例 14-5

为了分析消费者对餐厅就餐的偏好结构，某调研公司开展了一项餐厅就餐调查研究，选择了6个属性，每个属性有3种水平，如表14-1所示。应用联合分析方法探究消费者对餐厅就餐的偏好结构，具体过程利用SPSS16.0(英文版)软件中的“Orthogonal Design(正交设计)”菜单及“Syntax(语法)”窗口实现。

表 14-1 餐厅的属性和各属性的水平描述

属性	水平		
	1	2	3
1. 从住处抵达餐厅的时间	1 ( ) 分钟以内	10-15 分钟	15 分钟以上
2. 餐厅的整洁程度	非常整洁	一般	脏乱
3. 餐厅提供食物的种类	较丰富	一般	稀少
4. 在餐厅就餐的价格水平	较高	正常	较低
5. 在餐厅就餐是否拥挤	不拥挤	一般	较拥挤
6. 餐厅工作人员服务态度	良好	一般	较差

## 1. 操作步骤

### (1) 正交设计生成方案

① 在 SPSS 中选择“Data (数据)”→“Orthogonal Design (正交设计)”“Generate (生成)”，展开“Generate Orthogonal Design (生成正交设计)”对话框。在“Factor Name (因子名称)”和“Factor Label (因子标签)”框中分别输入对应因子变量和标签，如图 14-28 所示。



图 14-28 Generate Orthogonal Design 对话框

② 单击“Define Values (定义水平值)”按钮弹出“Generate Design : Define Values (生成设计: 定义水平值)”对话框，在“Value (值)”和“Label (标签)”框中输入对应的水平值及标签，返回主对话框（见图 14-29），B-F 因子设计过程与 A 相同。

③ 继续单击“Options (选项)”按钮，在“Minimum number of cases to generate (生成的最小个案数)”框中输入 18，再返回主对话框「」

④ 最后，单击“OK (确定)”按钮，即可得出一种正交设计方案。（结果见数据集 Example 14—2 及图 14-30）。

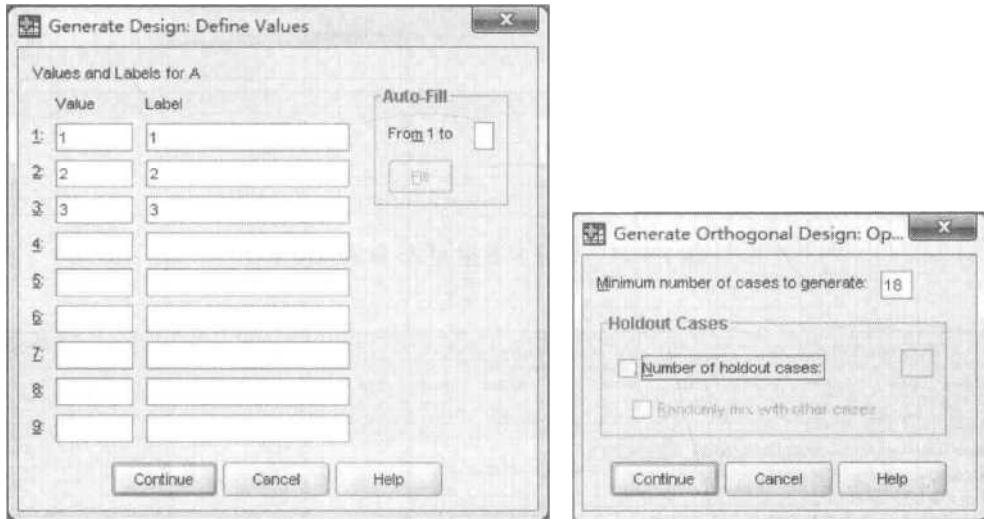


图 14-29 Generate Design : Define Values 子对话框和 Options 子对话框

A								
	A	L	B	:J....	c	D	E	F
2	3,00	3,00	3,00			3,00	3,00	3,00
	2,00	2,00	2,00			2,00	2,00	2,00
3	1,00	1,00	1,00			1,00	1,00	1,00
	3,00	3,00	2,00			2,00	1,00	0
5	2,00	2,00	1,00			3,00	1,00	0
6	1,00	1,00	3,00			2,00	2,00	0
7	3,00	2,00	3,00			1,00	2,00	0
8	2,00	1,00	2,00			3,00	1,00	0
9	1,00	3,00	1,00			2,00	3,00	0
10	3,00	1,00	1,00			3,00	2,00	0
11	2,00	3,00	3,00			2,00	1,00	0
12	1,00	2,00	2,00			1,00	3,00	0
13	3,00	2,00	1,00			2,00	3,00	0
	2,00	1,00	3,00			1,00	2,00	■ 0
15	1,00	3,00	2,00			3,00	1,00	0
16	3,00	1,00	2,00			2,00	3,00	0
17	2,00	3,00	1,00			1,00	3,00	0
18	1,00	2,00	3,00			3,00	2,00	0

图 14-30 正交转换后的数据 SPSS 格式截图

(2) 根据调查结果生成评分卡数据文件(见数据集 Example14-3 及图 14-31)。

(3) 进行联合分析。在 SPSS 中选择“File (文件)”→“New (新建 Syntax (语法))”，展开“SPSS Syntax Editor (SPSS 语法编辑器)”对话框，输入以下程序(如图 14-32 所示)，后点击“Run Current (运行)”按钮，即可得到联合分析结果(见图 14-33 和图 14-34)。

程序：conjoint plan = ' C : \Users\ Administrator\Desktop\ 10\_3.sav'

/data = ' C :\Users \Administrator\Desktop \dafen.sav'

/score = x 1 to x 1 8 /plot = summary.

	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10	x11	x12	x13	x14	x15	x16	x17	x18
1	1	4	18	3	14	12	2	11	17	16	5	1	9	13	7	10	6	15
2																		
3																		
4																		

图 14-31 评分卡数据 SPSS 格式截图

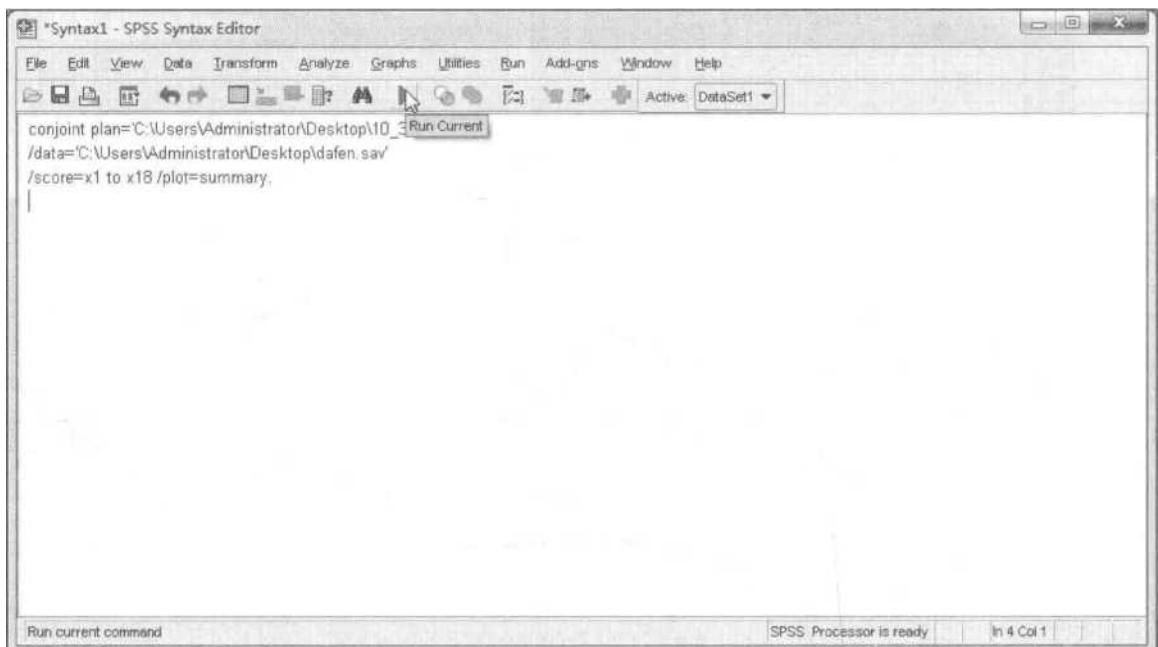


图 14-32 SPSS Syntax Editor 窗口

## 2. 解读输出结果

联合分析的输出结果包含：联合分析的拟合优度检验（见图 14-33 中的“Correlations”表），各属性每一个因子的效用值（见图 14-33 中的“Utilities”表），各属性重要性取值（见图 14-33 中的“Importance Values”表）以及各属性重要性条形图（见图 14-34）。

联合分析的拟合优度检验使用皮尔逊（Pearson）的 R 检验和肯德尔（Kendall）的 tau 检验，相应的斯皮尔曼和肯德尔秩检验的结果如下：皮尔逊 R 检验的预测排序值与实际排序值的相关系数为 0.987，双尾检验显著性水平为 0.000。肯德尔的 tau 检验的实际排序值与预测排序值的相关系数为 0.967，双尾检验显著性水平为 0.000。由此可见，两个相关系数的检验结果均为高度显著，模型拟合精度很高，可以认为联合分析模型所做出的假设和得出的成分效用值是合理的，可以说明受测者 1 在选择餐厅时的偏好结构。

接下来，分析各属性每一个水平的效用值以及重要性。由效用值计算结果可以看出，A 属性（从住处抵达餐厅的时间）的三个水平所对应的效用值依次为 2.833、-0.167 和 -2.667，水平 1（10 分钟以内）对应的效用值最高，从住处抵达餐厅需要的时间越长消费者在该属性上的效用值越低；同理可以分析 B、C、D、E、F 属性各水平所对应的效用值。

此外，还可以计算各属性的效用范围，比如 A 属性的效用范围为 5.5（= 2.833 -

Utilities			
		Utility Estimate	Std. Error
A	1	2.633	.527
	2	-.167	.527
	3	-2.667	.527
B	1	2.167	.527
	2	-1.167	.627
	3	-1.000	.527
C	1	5.000	.527
	2	-2.333	.527
	3	-3.667	.527
D	1	-.500	.527
	2	.000	.527
	3	.500	.527
E	1	-.333	.527
	2	-.333	.527
	3	.667	.527
F	1	.167	.527
	2	.333	.527
	3	-.500	.527
(Constant)		9.500	.373

Importance Values

A	25.781
B	15.625
C	45.312
D	4.688
E	4.688
F	3.906

Averaged Importance Score

Correlations<sup>3</sup>

	Value	Sia.
Pearson's R	.987	000
Kendall's tau	.967	.000

a Correlations between observed and estimated preferences

图 14-33 “各属性各水平的效用值”“重要性值”及“相关关系”结果截图

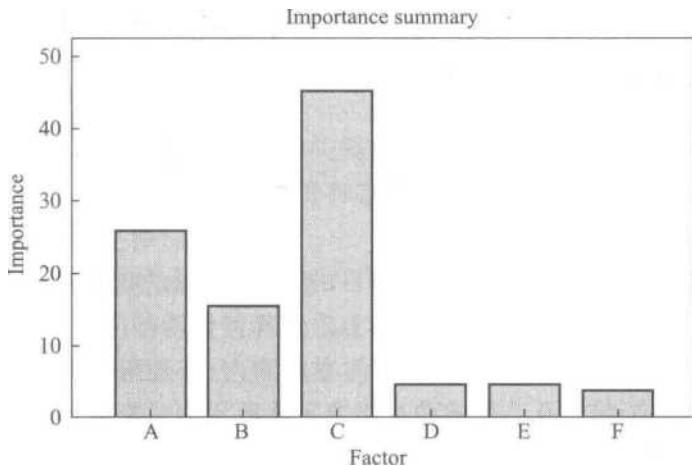


图 14-34 联合分析结果——“重要性总结”条形图

(-2.667) ), B、C、D、E、F 属性对应的效用范围依次是 3.167、9.667、1、1 和 0.667, 效用范围越宽 意味着消费者偏好对属性水平的变化越敏感 也说明该属性在影响消费者对该产品的偏好方 面是最不重要的。结合各属性的重要性可以看出, 属性 C(餐厅提供的食物种类)的效用范围 最宽, 重要性取值最大, 说明食物种类对该消费者选择餐厅的影响最大。同时, 该消费者对于 距离情况(从住处抵达餐厅的时间)也比较关注, 整洁程度次之, 而对另外三种属性即价格水 平、是否拥挤和工作人员服务态度则不是很重视。

根据各属性每一种水平的效用值, 还可以得出各属性水平之间的兑换关系, 比如对于属性 A, 由水平 1 改为水平 2, 效用值将下降 3 (=2.833 - (-0.167)), 而将属性 B ,由水平 2 上升至水 平 1, 效用值将提高 3.334 ( =2.167 - (-1.167)), 两者大致相当, 可以互相弥补。

### 三、进行联合分析时应注意的问题

与许多其他调查技术一样，联合分析在一定程度上也受到人为因素的影响。受访者在调研过程中可能比在真实情境下更深思熟虑，他们也可能在调查中比在实际市场环境中获得更多的产品信息。此外，任何新产品或服务的广告和促销都能够导致消费者感觉的形成，这种感觉与通过用于调查的产品描述所产生的感觉是不同的。

## 第六节

# 多变量综合评价

在第一节中，我们采用因子分析方法从显在变量中提炼潜在因子，在此基础上可以对受访者的 BAT 支付使用意愿进行综合评价；在调查分析过程中，也可采用构建指标评价体系的方法，从整体的角度对客户进行综合评价。多指标综合评价是根据研究的目的，首先建立综合评价指标体系，然后基于评价指标体系，对现象的各个方面进行定量分析，得出一个整体性的评价结果。

### 一、多变量综合评价基本原理

多指标综合评价的基本步骤一般包括：选择并构建综合评价指标体系、确定观测指标的量纲方法、选择综合评价指标的合成方法、确定各评价指标的权重等。

#### （一）选择并构建综合评价指标体系

在明确了评价目标之后，需要根据研究的目的，选择合适的统计指标，建立一个能够从不同角度、不同侧面反映评价对象的评价指标体系。选择评价指标应遵循如下原则：①客观性。评价指标体系要能够准确把握所要研究问题的本质和内涵，能够客观地反映事物的特征；②全面性。各个评价指标要能够从不同的角度综合反映所研究对象的全貌；③相互独立性。在评价指标体系中，应尽可能选择相关程度低的指标，这是因为如果指标之间相关程度过高，用一项指标能说明问题，却采用多项指标，事实上相当于加大了这类指标的权重；④可比性。评价指标要含义明确，计算口径一致，达到动态可比、横向可比；⑤可操作性。对于选择的评价指标，要考虑其数据资料收集的可能性，尽可能利用现有的统计资料。

#### （二）确定观测指标的量纲方法

根据综合评价指标计算过程的不同特点，确定观测指标的量纲方法大致可分为两类：一类为有量纲指标评价方法；另一类为无量纲指标评价方法。

##### 1. 有量纲指标评价方法

对于有量纲指标进行综合评价，主要采用总分评定法，或称综合积分法，其基本步骤可以归纳如下：①根据评价的目的和评价对象的特点，选择若干个评价项目或评价指标，构成评价指标体系；②确定各项目或各指标的评价标准和计分方法，常用的评分法有等级量化处理；③综合评判结果，把各指标（或各项目）的得分相加，即得该评价对象的总分 C

##### 2. 无量纲指标评价方法

评价指标的无量纲化是通过某种变换方式消除各个观测指标的计量单位，使其转化为统

一、可比的变换过程。常用的无量纲化的处理方法包括标准化变换方法、归一化变换方法、功效系数变换方法以及指数化变换方法。

在实际变换中，人们习惯于按百分制对所评价总体的各个观测单位进行变换，常将上述变换方法的公式乘以 100。此外，有时为使综合评价指标不出现 0 值和负值，常在变换公式后加上一个常数项。改进的无量纲指标评价方法如下：

#### (1) 标准化变换方法

$$z_i = \frac{x_i - \bar{x}}{s} \quad (14.7)$$

式中， $\bar{x}$  是观测值， $s$  是均值， $s'$  是标准差。

#### (2) 归一化变换方法

$$\tilde{x}_i = \frac{x_i}{\max(x)} \quad (14.8)$$

式中小是观测值， $s''$  是第  $i$  个指标的最小观测值， $s'''$  是第  $i$  个指标的最大观测值。 (3) 功效系数变换方法

$$z_i = \frac{x_i - a}{b - a} \quad (14.9)$$

式中， $a$  是观测值， $b$  是评价指标的不允许值， $c$  是评价指标的满意值，变换后的指标称为功效系数。

#### (4) 指数化变换方法

$$z_i = \frac{x_i}{\max(x)} \quad (14.10)$$

式中， $x$  是观测值， $x^*$  是评价标准值。

### (三) 选择综合评价指标的合成方法

评价指标的合成方法是指将经过无量纲化变换的各个指标按照某种方法进行综合，得出一个可用于评价比较的综合指标。合成方法主要有总和合成法、乘积合成法和混合合成法三种。其中较为常用的是总和合成法，其公式有如下两种：

#### 1. 简单算术平均法

$$\bar{x} = \frac{\sum_{i=1}^m x_i}{m}$$

其中  $\bar{x}$  是评价总体中第  $i$  个观测单位的综合评价值， $m$  表示指标个数。

#### 2. 加权算术平均法

$$\bar{x} = \frac{\sum_{j=1}^t w_j z_j}{\sum_{j=1}^t w_j} \quad (14.12)$$

其中， $w_j$  是第  $j$  个指标  $Z_j$  的权数。

简单算术平均法同等看待不同评价指标的重要性，但是在现实中，综合评价指标体系中各指标的重要性往往是不同的，故应赋予不同分量的权数，才能准确反映综合指标的合成值。

#### (四) 确定评价指标的权数

评价指标的权数是指在评价指标体系中每个指标在综合评价中的重要程度，因为各指标在指标体系中的重要性不同，不能等量齐观，必须客观地确定各指标的权数。确定评价指标的

权数的方法有：德尔菲法（又称专家评估法）、层次分析法、强制打分法、主成分分析法、因子分析法和相关系数构权法等，其中最常用的是德尔菲法和层次分析法。

### 1. 德尔菲法

德尔菲法（Delphi Method），又称专家规定程序调查法。该方法主要是由调查者拟定调查表，按照既定程序，以函件的方式分别向专家组成员进行征询；而专家组成员又以匿名的方式（函件）提交意见；经过几次反复征询和反馈，专家组成员的意见逐步趋于集中，最后获得具有很高准确率的集体判断结果。利用德尔菲法确定各个评价指标权重的一般程序是，首先由评价对象所属领域内的多名专家单独对各个评价指标的重要程度进行评定，然后由综合评价人员综合各个专家的评定结果，计算出平均数后再反馈给各位专家，如此反复进行几轮，使各位专家的意见趋于一致，从而确定出各个评价指标的权数。

### 2. 层次分析法

层次分析法（Analytic Hierarchy Process, 简称 AHP），是将与决策总是有关的元素分解成目标、准则、方案等层次，在此基础之上进行定性和定量分析的决策方法。层次分析法计算过程的核心问题是权数的构造。层次分析法的基本思路为：建立评价对象的综合评价指标体系，通过指标之间的两两比较确定出各自的相对重要程度，然后通过特征值法、最小二乘法、对数最小二乘法、上三角元素法等的客观运算来确定各评价指标权数。其中，特征值法是层次分析法中最早提出的、也是使用最广泛的权数构造方法，其基本步骤包括：构造判断矩阵、计算各指标的权数以及检验判断矩阵的一致性。

## 二、应用实例

### 应用实例 14-6

在第十一章的实训材料中，给出了调研公司调研得到某批发商主要经营的六类商品的销售情况数据（见数据集 Example11-2），分别采用综合评价法和层次分析法对客户消费水平进行全面评价。评价过程可采用 Excel、R 语言和 MATLAB 等软件实现，考虑到在处理该类问题时，Excel 相比其他软件容易掌握，故本节采用 Excel 软件演示综合评价的操作过程。

#### （一）德尔菲法的实现

##### 1. 观测指标的无量纲化

因为各个指标的量纲或数量级通常是不同的，所以需要对各个指标数据进行无量纲化。这里采用归一化变换方法，并引进功效系数，计算得出各个指标单向评价分数：

$$x = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}}$$

式中：

和表示第  $i$  个地区第  $j$  项指标的实际数值；

表示第  $j$  项指标的最大值；

表示第  $j$  项指标的最小值；

表示第  $i$  个地区第  $j$  项指标的无量纲值。

这种无量纲化方法的好处是，不仅纵向上可以消除不同指标的不同数量级的影响，还能使横向上各地区的得分包含在 1 至 100 之间，易于比较。数据无量纲化运用 Excel 软件实现，计算过程及结果见图 14-35。

	C	D	E	F	G	H	A	J	K	L	M	N			
/	Fresh	Milk	Grocery	Frozen	Detergent	Delicassen	ZFresh	ZKilk	ZGrocery	ZFrozen	ZDetergents	Paper	ZDelicassen		
1	12669	9656	7561	214	2674	13381	46.7764147.	84363	44.	88785	40.18638	43.	92563198		
2	7057	9810	9568	1762	3293	1776	43.	77394	47.	96945	46.1858	41.71291	44.	83539095	
3	6353	8808	7684	2405	3516	7844	43.3973	47.	15085	44.96739	42.	34699	45.16313933		
4	13265	1196	4221	6404	507	1788	47.	09527	40.	93215	42.	72783	46.29051		
5	22615	5410	7198	3915	1777	5185	52.	09759	44.	37482	44.	65309	43.83604		
6	9413	8259	5126	666	1795	1451	45.03442	46.	70234	43.	31311	40.63211	42.	63374486	
7	12126	3199	6975	480	3140	545	46.48589	42.	56852	44.	50888	40.44869	44.	61052322	
8	7579	4956	9426	1669	3321	2566	44.	05322	44.	00392	46.	09397	41.6212	44.	87654321
9	5963	3648	6192	425	1716	750	43.	18864	42.	93534	44.00225	40.39445	42.51763668	40.93491865	
10	6006	11093	18881	1159	7425	2098	43.	21165	49.	01761	52.	20863	41.11827	50.90828924	42.62202753
11	3366	5403	12974	4400	5977	1744	41.	79923	44.	3691	48.	3885	44.31431	48.	78012934
12	13146	1124	4523	1420	549	497	47.	0316	40.	87333	42.	92314	41.37565	40.80246914	40.61827284
13	31714	12319	11757	287	3881	2931	56.	96562	50.	0192	47.	60145	40.25837	45.69958848	43.66458073
14	21217	6208	14982	3095	6707	602	51.	34965	45.	02676	49.	687143.	02741	49.	85302763
15	24653	9465	12091	294	5058	2168	53.	18793	47.	68759	47.	81745	40.26257	47.42945326	42.70963705
16	10253	1114	3821	397	964	412	45.	48382	40.	86516	42.	46915	40.36684	41.41240447	40.51188986
17	1020	8816	12121	134	4508	1080	40.	5441	47.	15739	47.	83686	40.10749	46.	62110523
18	5876	6157	2933	839	370	4478	43.	1421	44.	98509	41.	89487	40.80271	40.	53938859
19	18601	6327	10099	2205	2767	3181	49.	95007	45.	12397	46.	5292	42.14976	44.06231628	43.97747184

图 14-35 数据无最纲化计算过程及结果

## 2. 计算综合得分

m

I=I

假设采用德尔菲法确定各类商品的权重分别为：生鲜产品 0.2、日杂用品 0.1、冷冻食品 0.2、洗涤剂和纸制品 0.1、熟食品 0.2、奶制品 0.2，进一步根据公式计算得到综合得分，计算结果见图 14-37 中“综合评价法得分”列。

### (二) 层次分析法的实现

#### 1. 确定权重

构建 6 类商品的判断矩阵，如表 14-2 所示：

表 14-2 6 类商品的判断矩阵

	Fresh	Milk	Grocery	Frozen	Delergents_Paper	Delicassen
Fresh	1	2	4	2	4	2
Milk	1/2	1	2	2	3	1/2
Grocery	1/4	1/2	1	1/2	1	1/2
Frozen	1/2	1/2	2	1	2	1
Detergents_Paper	1/4	1/3	1	1/2	1	1/3
Delicassen	1/2	2	2	1	3	1

将判断矩阵输入 Excel 表格中，进行一致性检验，如图 14-36 所示。

其中，H4 = PRODUCT(B4: G4)，表示 B4、C4、D4、E4、F4、G4 各单元格连乘，复制公式至 H9 单元格。

I4 = POWER(H4, 1/6)，表示将 H4 单元格的值开 6 次方，复制公式至 19 单元格。

I10 = SUM(14: 19)，表示求和。

J4 = I4/\$I\$10，复制公式至 J9 单元格。

K4 = B4 \* J\$4 + C4 \* J\$5 + D4 \* J\$6 + E4 \* J\$7 + F4 \* J\$8 + G4 \* J\$9，复制公式至 K9 单元格令

	A	B	C	D	E											
	Fresh	Milk	Grocery	Frozen	Detergents	Paper	Delicassen	按行相乘	开n	放方	相重”	Awi	Awi/li	CI	CR	
4	Fresh	1	2	4		4		2	128.2	244924	0.327947	1.979271	6.035329			
5	Xlk	1/2	1	2	2		3	1/2	31.	200937	0.175438	1.108886	320653			
6	Grocery	1/4	1/2	1	1/2		1	1/2	0.031250	561231	0.081987	0.494818	6.035329			
7	Frozen	1/2	1/2	2	1		2	1	1	1	0.146084	0.901916	6.17396			
8	Detergents.Paper	1/4	1/3	1	1/2		1	1/3	0.013889	0.49028	0.071622	0.432758	6.042236			
9	Delicassen	1/2	2.	2	1		3	1	61.348006	0.196922	1.236695	6.280124	0.029588	0.023682		
									6.845379				6.147938			

图 14-36 判断结果一致性检验结果截图

$$L4 = K4/J4$$

$$J = \text{AVERAGE}(L4: L9)$$

$$CI = (L10-6) / (6-1), CR = CI / 1.2494 = 0.023681523 < 0.21, \text{即通过一致性检验。}$$

各变量指标权重依次为：

$$A\_W = (0.3279, 0.1754, 0.0820, 0.1461, 0.0716, 0.1969)$$

## 2. 计算综合得分

得到各指标的权重后，把各项指标的无量纲化后指标值(同综合评价法的无量纲化处理)，采用线性综合评价法即把各项指标的评价分值乘以相应权重，即可得到各个地区的综合得分值，结果见图 14-37 中“层次分析法得分”列。

	Fresh	Milk	Grocery	Frozen	Detergent	Delicasse	ZFresh	Zlilk	ZGrocery	ZFrozen	说 t 痘 Wei 云岳合评价瓣济歛分根确				
2	12669	9656	7561	214	2674	1338	46.7764	47.84363	44.88785	40.18638	43.92563	41.67084	44.17679879	44.63652531	
3...	7057	9810	9568	1762	3293	1776	43.77394	47.96945	46.1858	41.71291	48.83539	42.21902	44.23718237	44.17647279	
4	6353	8808	7684	2405	3516	7844	43.3973	47.15085	44.96739	42.34699	45.16314	49.81352	45.55478357	45.42107394	
5	13265	1196	4221	6404	507	1788	47.09527	40.93215	42.72783	46.29051	40.74074	42.23404	43.65725209	44.12598382	
6	22615	5410	7198	3915	1777	5185	52.09759	44.37482	44.65309	43.83604	42.60729	46.48561	46.08484894	47.14065486	
7	9413	8259	5126	666	1795	1451	45.03442	46.70234	43.31311	40.63211	42.63374	41.81227	43.43091166	43.73634859	
—	12126	3199	6975	480	3140	545	46.48589	42.56852	44.50888	40.44869	44.61052	40.67835	42.94823044	43.4766592	
9	7579	4956	9426	1669	3321	2566	44.05322	44.00392	46.09397	41.6212	44.87654	43.20776	43.67426938	43.74908264	
—	5963	3648	6192	425	1716	750	43.18864	42.93534	44.0025	40.39445	42.51764	40.93492	42.14268397	42.31087599	
..U..	6006	11093	18881	1159	7425	2098	43.21165	49.01761	52.20863	41.11827	50.90829	42.62203	45.50560217	45.09719721	
12,	3366	5403	12974	4400	5977	1744	41.79923	44.3691	48.3885	44.31431	48.78013	42.17897	44.24918666	43.73249558	
13	13146	1124	4523	1420	549	497	47.0316	40.87333	42.92314	41.37565	40.80247	40.61827	42.35233144	43.07905552	
X	31714	12319	11757	287	3881	2931	56.96562	50.0192	47.60145	40.25837	45.69959	43.66458	47.5116565	49.11238981	
15	21217	6208	14982	3095	6707	602	51.34965	45.02676	49.6871	43.02741	49.85303	40.74969	45.98471309	46.6937645	
16	24653	9465	12091	294	5058	2168	53.18793	47.67859	47.81745	40.26527	47.42945	42.70964	46.29477686	47.41902104	
I2	10253	1114	3821	397	964	412	45.48382	40.86516	42.46915	40.36684	41.4124	40.51189	41.83369816	42.40817658	
18	1020	8816	12121	134	4508	1080	40.5441	47.15739	47.83686	40.10749	46.62111	41.34793	43.27717868	42.83198906	
19	5876	6157	2933	839	370	4478	43.1421	44.98509	41.89487	40.80271	40.53939	45.60075	43.14955506	43.31917582	
20	18601	6327	10099	2205	2767	3181	49.95007	45.12397	46.5292	42.14976	44.06233	43.97747	45.29940638	46.08559396	
21	7780	2495	9464	669	2518	501	44.16075	41.99338	46.11854	40.63507	43.69635	40.62328	42.4639858	42.69611325	
22	17546	4519	4602	1066	2259	2124	49.38563	43.64691	42.97423	41.02656	43.3157	42.65*57	43.97172673	44.87183094	

图 14-37 综合评价结果截图

由图 14-37 可见，两种计算结果还是有一定的差异，这主要是由于综合评价法和层次分析法对各评价指标给定的权重不同。根据综合评价的结果，可了解客户对特定权重下各类商品的综合消费水平。

## 三、进行综合评价时应注意的问题

第一，确定评价指标体系时，要在系统全面的基础上，尽量选择具有代表性的综合指标，要避免选择含义相近的指标；指标体系的详略也必须适宜，指标体系的设置应具有一定的科学性，做到简明科学。

第二，评价指标权重的确定是多变量综合评价的关键，选用不同的确定指标权重的方法，所得到的评价指标的权数会有差异，进而综合评价的结果自然也会有些差异。

## 篇章小结 · · ·

多变量数据分析的目的在于确定两个以上变量间的关系，进而探索和描述研究现象的特征，广泛应用于各类市场调查研究中。本章重点介绍了因子分析、聚类分析、对应分析、多维标度分析、联合分析以及多变量综合评价六种常用的多变量统计分析技术。

因子分析是一种降维、简化数据的技术。它是研究如何以最少的信息丢失将众多原始变量简化成少数几个代表变量间关系的因子，并使因子具有一定的命名解释性的分析方法。即因子分析是一种通过显在变量测评潜在变量，通过具体指标测评抽象因子的统计分析方法。

聚类分析是按照某种相似性或差异性指标，定量地确定样本（或变量）之间的亲疏关系，并按这种亲疏关系程度对样本进行分类的多元统计方法。聚类分析分为Q型聚类和R型聚类，Q型聚类是对样本进行分类处理，R型聚类是对变量进行分类处理。常用的聚类方法有系统聚类法和K均值聚类法，不同的方法适合解决不同类型的问题。

对应分析是指对两个或多个定性变量的多种水平进行相应性研究，由定性变量构成的交互汇总表来揭示变量间的联系，揭示同一变量的各个类别之间的差异，以及不同变量各个类别之间的对应关系。对应分析的基本思想是将列联表的行和列各元素的比例结构以点的形式在较低维的空间中表示出来。

多维标度分析是一种将多维空间的研究对象（样本或变量）简化到低维空间进行定位、分析和归类，同时又保留对象间原始关系的数据分析方法。它是通过一系列技巧，使研究者识别构成受测者对样品进行评价基础的关键维数。多维标度法的实质是根据配对样品间的距离数据，在一个确定维数的空间中估计样品的坐标。

联合分析是在已知受测者对某一受测体集合整体评估结果的情形下，通过分解的方法去估计其偏好结构的一种分析方法。联合分析的目的在于将受测者的整体反应加以分解，从受测者对受测体的整体评估结果中估计每一受测体成分的效用。

多指标综合评价是根据研究的目的，首先建立综合评价指标体系，然后基于评价指标体系，对现象的各个方面进行定量分析，得出一个整体性的评价结果。多指标综合评价的基本步骤一般包括：构建综合评价指标体系、确定观测指标的量纲方法、选择综合评价指标的合成方法、确定各评价指标的权重等。

### 常用数据分析方法概览

类型	分析方法	主要目的
单变量 数据 分析	规模与结构分析	应用频数分布表和频数分布图，了解数据的总量规模和分布结构
	比较分析	从时间和空间两个维度比较数据的变化，应用指数比较不同群体在同一问题上的差异
	集中度分析	测度数据的集中趋势
	变异性分析	测度数据的离散程度，衡量集中度指标的代表性
	参数估计	在满足一定精度和把握程度的条件下，利用样本信息来估计总体特征
	显著性检验	事先对总体分布形式或总体（随机变量）的参数做出一个假设，然后利用样本信息来判断这个假设是否合理

类型	分析方法	主要目的
双变量 数据 分析	列联分析	揭示品质型变量（定类变量和定序变量）之间以及品质型变量各种状态之间的相关关系
	双变量相关分析	测度两个数值型变量之间相关关系的密切程度
	双变量回归分析	分析因变量对自变量的依赖关系，同时在给定自变量值时预测因变量，模型中只包含一个自变量，因变量为数值型
多变量 数据 分析	方差分析	验证不同的自变量水平是否会对因变量产生显著的影响
	多变量回归分析	同双变量回归分析，差别在于多变量回归分析包含多个自变量
	Logistic 回归分析	分析因变量与自变量的概率取值之间的关系，因变量为分类变量
	因子分析	以最少的信息损失将众多原始变量简化成少数几个代表变量间关系的因子，并使因子具有一定的命名解释性
	聚类分析	按照某种相似性指标，定量地确定样本（或变量）之间的亲疏关系，并按这种亲疏关系程度对样本进行分类
多变量 数据 分析	对应分析	揭示同一变量的各个类别之间的差异，以及不同变量各个类别之间的对应关系
	多维标度分析	反映研究对象的感知和偏好，用于确认消费者对不同品牌感知维度的数目与性质、现有品牌在这些维度上的定位、消费者的理想品牌在这些维度上的定位
	联合分析	确定消费者选择过程中不同属性的相对重要性；确定最受欢迎的品牌的属性构成；根据不同属性水平偏好的相似程度进行市场细分
	综合评价分析	在众多影响评价对象的因素中选出主要因素构建成综合评价指标体系，用一个总指标来说明被评价对象的总体水平

## 磁关璀璨名词••••

因子分析 聚类分析 对应分析

多维标度法联合分析

多变量综合评价 数据归一化 德尔菲法 层次分析法

肯思考题 .....



即测即评

- 简述因子分析的基本思想和步骤。
- 简述聚类分析的基本思想。
- 常用的聚类方法有哪些？分别简述每种方法的基本思想和适用条件。
- 什么是对应分析？应用对应分析方法时应注意哪些问题？
- 简述多维标度法的作用。

6. 简述联合分析在市场研究中的应用。
7. 概括多变量综合评价的步骤，，

## 嘗本章实训. . .

### 一、 实训目的

1. 明确各种多变量统计分析方法的基本原理和作用。
2. 熟练掌握应用 SPSS 软件进行多变量数据分析的操作方法。

### 二、 实训内容

在某项保健产品研究中，客户开发了一种针对女性的新保健产品，经过初步测试，该产品具有多种保健功能。然而，鉴于保健品市场的激烈竞争，客户希望通过调研，对该类保健品市场结构进行细分，然后根据公司的营销能力选择 1~2 个细分市场。研究中，客户计划采用 3 个步骤：市场细分和市场机会评估—> 目标人群确定—> 产品/品牌定位，以确定产品的市场定位，确保产品上市成功。

为此，客户委托某研究公司进行了系列市场调查，研究共分为三个阶段。其中，第二阶段即为定量研究阶段，通过随机抽样方式了解目标市场消费者的基本状况，并对市场进行细分。经过比较筛选，研究者发现购买动机细分是比较有效的细分方式。因此，研究者确定在问卷中对购买动机数据进行采集。问卷采用李克特五分量表测试消费者的态度，其中 1 分表示非常不同意，5 分表示非常同意，量表问题见数据集 Example14-4 以及图 14-38 “数据集的 SPSS 格式截图”。除消费者购买动机变量外，数据中还包括婚姻状况（已婚、独身）、教育水平、个人收入级别以及主要使用的保健产品（太太\*\*、乌鸡\*\*、排毒\*\*、阿\*\*、红桃\*\*、脑白\*\*、朵\*\*、其他）四个消费者特征变量。

				Values	Missing		Align	Measure
010203040505070B0901213456703920	Ques Bansferm fmatyxe	(jaties Adorns andaw tjepl	wiiM 面細昭曲 M 嘅憶：土前 ti@<r					
13	Numenc	令而色紅掏 使吱		None	None	,	■ Right	H Ordinal
14	Numenc	驗有光蚤 fil 神狀		None	None	5	塞 Right	d Ordinal
15	Numenc	态好去		None	None	5	■ Right	d Ordinal
16	Numeric	去病		None	None	5	■ Right	jI Ordinal
17	Numenc	去皱		None	None	5	看 Right	d Ordinal
18	Numenc	够决反缺干換		None	None	5	■ Right	d Ordinal
19	Numenc	调节内分全 藏缓		None	None	5	■ Right	d Ordinal
20	Numenc	疫敏衰老 去隊盤		None	None	5	M Right	d Ordinal
21	Numenc	劳		None	None	5	S Right	d Ordinal
22	Numenc	摄葛 It 脂馬更		None	None	8	S Rigttr	d Ordinal
23	Numeric	直便词%		None	None	5	■ Right	d Ordinal
24	Numenc	抗豪走		None	None	5	羊 Right	d Ordinal
25	Numenc	援高免疫力		None	None	5	3 Right	H Ordinal
26	Numenc	应该对自己好点		None	None	5	31 Right	HI Ordinal
27	Numenc	希管自己		None	None	5	寧 Right	d Ordinal
28	Numeric	同部人邮在曜用		None	None	5	■ Right	Ordinal
29	Numenc	年轻入对裁桢嬈		None	None	5	■ Right	d Ordinal
30	Numenc	RB 用 女为快己		None	None	5	3 Right	d Ordinal
31	Numenc	在客 主更股用的		None	None	5	首 Right	d Ordinal
32	Numeric	产品		None	None	8	Mi Right	JH Nominal
33	Numenc			”, 未婚)	None	6	9 Right	JH Nominal
34	Numenc			None	None	3	■ Right	異 Nominal
35	Numenc	戰 S3		None	99	2	■ Right	冬 Nominal

图 14-38 数据集 Example 14-4 的 SPSS 格式截图

研究目标是，探究保健品市场上可被直接测量的购买动机指标的背后所代表的消费者内在购买需求。要求基于数据集 Example14-4，选择恰当的统计分析方法（不局限于多变量统计

分析），尽可能多地挖掘数据中包含的信息。分析过程中，请注意方法选择的合理性和分析过程的完整性。

提示：1. 不同保健品品牌的使用者在购买动机上有无明显差异？

2. 基于内在购买需求，是否可将消费者分为若干群体？

3. 若将消费者划分为若干群体，各群体的偏好如何？

实训资料来源：张文彤，钟云飞. IBM SPSS 数据分析与挖掘实战案例精粹 [M]. 清华大学出版社，2013。

### 三、实训组织

1. 以个人为单位负责完成数据分析任务。

2. 提交数据分析书面报告。

3. 组织部分个人课堂交流书面报告。

4. 教师批改报告，点评与总结 Q

# 第十五章 市场调查成果报告



## 、导语

、—……

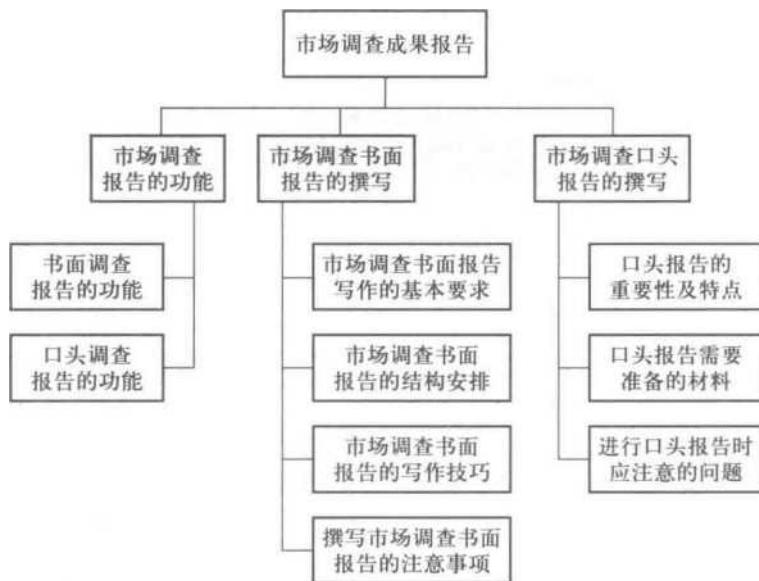
在经过精心设计调查方案、严格的过程控制、详细的统计分析等一系列的努力后，最终到了调查报告的撰写阶段，调查报告是整个研究过程所付出的所有劳动的最终结果。在调查报告中，既需要表明前面所有环节的工作，更需要突出最终的结果。对委托方而言，他们需要的是一个有过程保证的最终结果。在调查活动结束后，受委托方应当向调查的委托方汇报调查的过程与结果，这是一个双方相互沟通交流的过程。委托方一方面迫切想知道研究的最终结果有什么新发现，有哪些有价值的建议等；另一方面，也关心得出结果与结论的过程是否能有效地支持研究结论，有什么新证据，采用了哪些科学的方法。由于研究结果会对企业的市场营销决策行为产生影响，因此需要受委托方精心准备，并与委托方进行深度沟通。

市场调查成果的报告，需要根据调查的要求与目的进行设计。报告方式可以根据不同的场景进行选择。本章主要介绍市场调查结果的书面报告与口头报告。

## 、教学目标与要求

1. 了解展示市场调查结果的多种方式及其选择规则。
2. 熟悉市场调查书面报告与口头报告的基本要求。
3. 掌握市场调查书面报告和口头报告的写作步骤。
4. 掌握书面报告和口头报告的优点与缺点。

## 、本章知识结构



## 第一节

# 市场调查报告的功能

市场调查报告是指根据调查目的，将调查资料通过文字、图、表格组织而成的主题鲜明、逻辑清晰的有观点和建议的文字材料。书面报告和口头报告存在表现形式上的差异。前者详细而具体，主要用于市场调查最终成果的汇报交流；后者简明而概述，主要用于口头交流。由于书面调查报告与口头报告的功能有所侧重，其写作要求也相差较大。因此，这里将两者分开讨论。

## 一、书面调查报告的功能

### (一) 书面调查报告给谁看

从委托调查公司方面看，书面调查报告主要有两个层次的读者：一是公司总经理一级的决策者；二是公司各部门的负责人，如市场部、产品部、广告策划部、市场销售部等部门经理。总经理作为公司的主要决策者，其需要的是主要的结果，以帮助决策，而各部门经理是主要的执行负责人，因此他们往往需要更多详细的信息，供执行参考。

### (二) 书面调查报告的功能

书面调查报告的功能有三个：证明调查过程的科学性与真实性；阐述调查结果，提出建议；文献功能。

#### 1. 证明调查过程的科学性与真实性

书面报告通过对市场调查背景、目的、内容、方法、过程控制、结果及结论各部分的详细描述，阐释一个围绕特定目的的调查活动的科学性与真实性。科学性与真实性表现在调查体系设计的科学性与操作过程的真实有效性两个方面。

首先是调查设计的科学性。这表现在市场调查选题下研究问题的合理界定、相关概念的可操作化、搜集资料方法与抽样方法的合理选择、统计分析方法的正确运用、结论与建议的推理过程严谨等。

其次是操作过程的真实有效性。即对市场调查项目从方案设计、调查实施到数据分析全过程进行严格的操作、有效地管理与执行，从而得到最终有据可依的可操作性结果。

#### 2. 阐述市场调查结果，提出建议

书面报告的第二个功能就是阐述市场调查结果，分析得出调查结论，从而在发现问题的基础上，提出解决问题的建设性意见。这是市场调查报告的核心部分。**市场调查报告的内容必须详细、涵盖调查的全部主要细节**。市场调查报告的科学性和真实性是结果、结论与建议的保证，是得到有价值的参考意见与建议的手段，而结果、结论、建议才是企业需求的内容。

#### 3. 文献功能

书面报告具有正式文献的权威性作用，可作为信息资料文件以备查阅和参考。书面报告既是本次市场调查项目结果阶段所需查阅的正式文本，又是以后相似调查的重要的二手资料来源，并可能成为市场调查行业的重要案例资料。

## 二、口头调查报告的功能

口头调查报告是书面调查报告的浓缩版本，其主要作用是作为一个正式的交流汇报文本，在向调查委托方进行正式汇报交流时使用。一般而言，在一次调查活动结束后，受托方都会向委托方做一个正式的汇报交流，向其汇报调查过程的主要情况和主要结果、结论及建议。对委托方而言，正式的口头汇报非常重

要，他们可以借此向公司相关部门通报有关企业产品或服务的问题，给企业下一步的工作提供重要的参考。同时，各部门人员也可以就市场调查中的结果、结论及建议等方面内容，与市场研究人员进行交流。

## 第二节

# 市场调查书面报告的撰写

一份好的市场调查书面报告是市场调查项目结项的重要文字材料，不仅要满足报告撰写的基本要求，符合内容结构的合理设计，而且还需要掌握专门的写作技巧和相关撰写规范。

### 一、市场调查书面报告写作的基本要求

调查报告不同于一般的学术论文，也不同于一般的工作总结，它的基本特点是以实际的调查资料为基础，对调查获得的数据资料进行分析研究后得出结论。调查报告应该是用事实说话，实事求是地反映客观实际，通过对实际资料的分析得出客观的结论。调查报告切忌先确定出某些结论，然后去寻找证据来验证这一结论的正确性。因此，市场调查书面报告撰写的基本要求如下：

#### （一）实事求是

调查报告作为调查研究的成果，最基本的特点就是尊重客观实际，用事实说话。因此引用的材料、数据必须真实可靠，不能弄虚作假、先入为主或迎合他人意图选择性地进行编写；在文体上一般采用第三人称或非人称代词，如“资料表明……”等语句；行文时以向读者报告的中性语气撰写，表达调查的结果和发现，而不是个人的主观看法；在数据分析部分要充分和准确地利用图、表信息，切记不能误导；报告应当具体而合理地给出市场调查项目的研究方法、调查结论等。

#### （二）目的明确

撰写市场调查书面报告，必须目的明确。任何市场调查都是为了解决某一问题，或说明某一问题而开展，因而其最终成果必须围绕其目的进行论述。要中心突出，条理清楚。由于市场调查中会使用多种方法，得到的材料量大、庞杂，要善于根据调查目的对材料进行严格的鉴别与筛选，适当归类，并分清材料的主次轻重，按照一定的逻辑主线，在报告中充分展示利用材料进行归纳推理的全过程。

#### （三）简明易读

市场调查报告的读者一般是管理人员或一般工作人员，大多不喜欢冗长、乏味、呆板的语言，大部分人也不精通市场调查的专业术语。因此，撰写市场调查书面报告时语言要力求简单准确、通俗易懂，切勿陷入“报告越长，质量越高”的误区，不要试图告诉读者你所知道的一切，而是要以客观、简洁和有效的表述吸引读者的注意力。书面报告逻辑结构要清晰，不能简单地罗列数据和堆砌资料，避免使用生硬、晦涩的措辞和术语，有明确的提纲与重点。提倡使用简短、直接、清楚的句子把事情说清楚，使报告易读易懂。

## （四）观点和数据统一

市场调查书面报告的特点是以调查资料为依据，而其中数据资料显得尤为重要。数据具有很强的概括力，往往比长篇大论更能使人信服。在报告撰写过程中，要善于用资料说明观点，用观点概括资料，切忌资料与观点相分离。

### 二、市场调查书面报告的结构安排

从内容构成上看，书面调查报告一般应当包括整个市场调查与分析过程的各个环节，即市场调查背景、目标、内容、方法、搜集资料的过程、统计分析结果、结论、建议、调查问卷等。但是，由于书面调查报告的特殊性，它又不能按部就班地一一将整个过程不分轻重缓急地写出来。

书面调查报告必须既充分展示调查的全貌，又考虑报告阅读者的心理特征与阅读特征。一次调查活动结束后，调查委托方的负责人最迫切想要阅读的部分是结果、结论与建议部分，其次才是其他部分。作为决策者，他们很难从头到尾读完一份市场调查报告的所有内容，他们最关心的是结果。因而书面调查报告在内容的安排上就应当充分照顾到这种阅读需求的心理特点，将企业负责人最想阅读的内容放在突出的位置。

根据报告阅读者关心报告内容的顺序组织框架，一份规范的市场调查书面报告一般应该包含以下八大部分：

#### （一）封面、标题与目录

1. 封面。封面是读者看到的一份市场调查书面报告的第一印象。封面的内容一般应当包括市场调查报告名称、委托方、调查公司名称及项目负责人、完成报告撰写日期。封面的设计要简洁大方，具有品质感。在封面之后，就是标题与目录。

#### 2. 标题。

市场调查书面报告名称就是标题。标题应是画龙点睛之笔，它必须准确揭示调查报告的主题思想。标题要简单明了，高度概括，具有较强的吸引力。标题一般有两种：单标题和双标题。单标题就是调查报告只有一个标题。一般是通过标题把被调查单位、调查内容等明确而具体地表示出来。例如全国大学生市场调查大赛部分调查报告标题如下：

- ① 济南市自然风光类景区游客满意度的影响因素研究；
- ② 关于育龄妇女对于生育健康基因检测的意愿及其影响因素的调查。

双标题就是调查报告有两行标题，采用正副标题形式。一般正标题表达调查的主题，副标题用于补充说明调查对象或主要内容。例如，

- ① “天生好米”有机胚芽大米市场需求调查与分析——以沈阳市为例；
- ② “生”在何方，“鲜”去哪寻——生鲜农产品购买渠道的调查。

标题的形式大致可以分为三类，即“直叙式”标题、“表明观点式”标题和“提出问题式”标题。“直叙式”标题是反映调查意向和说明调查范围、调查对象、时间的一种标题，这种标题简明、客观，使人一目了然。例如：

#### 2018年中国景区旅游消费调查分析报告

“表明观点式”标题是直接阐明作者的观点、看法，或对事物的判断、评价的标题。例如：

- ① 新能源汽车“热”的冷思考；
- ② 领“鲜”享受——消费者网购生鲜意愿调查分析报告。

“提出问题式”标题是以设问、反问等形式，突出问题的焦点吸引读者阅读，促使读者思考。例如：

“信”可贷？——P2P投资者意愿调查及发展前景分析

以上几种标题的形式各有所长，特别是后面两种形式的标题，它们既表明了作者的态度，又揭示了主

题，具有很强的吸引力。但是这类标题不容易看出调查的范围和调查的对象。因此这种形式的标题可用双标题的形式，用一个副标题补充说明调查内容和调查对象。例如：

不求洋品但求最好——对大学生品牌消费观的调查

3. 目录。目录是指书面市场调查报告中各章节、目的标题，具体包括：

- ① 章、节、目标题与副标题，并附页码；
- ② 表格目录：标题与页码；
- ③ 图形目录：标题与页码；
- ④ 附录：标题与页码。

一般而言，内容目录只列出各部分的标题名称及页码。目录的篇幅以不超过一页为宜。有时如果报告中图、表比较多，也可再列一张图表目录。在现代办公系统软件中，目录一般根据标题层次自动生成，常见的是三级目录形式，即包括章、节、目三级标题的目录形式。

## （二） 内容摘要（或摘要）

阅读市场调查书面报告的人往往对调查过程的复杂细节没有什么知识或兴趣，他们只想知道调查所得的主要结果和主要结论以及他们如何根据调查结果行事。因此摘要可以说是调查报告中极其重要的一节，它也许是读者唯一阅读的部分。由于这一部分如此重要，所以它应当用清楚、简洁而概括的手法，扼要地说明调查的主要结果。一般而言，摘要包括如下方面的内容：

1. 市场调查背景与目标简述。在摘要中，首先要阐明调查背景与目的，即为什么公司要在这方面花费时间和金钱，它想得到些什么。这部分不仅为报告的其余部分规定了切实的方向，同时也使得管理者在评审调查的结果与建议时有了个大致的参考框架，即所陈述的问题。

2. 市场调查方法及调查过程简述。这部分说明资料来源，调查方法及调查过程的可靠性、科学性与合理性。

3. 主要调查结果简述。它包括来自市场的真实情况是什么，有什么发现等。

4. 主要的结论与建议简述■即从结果中能归纳总结出哪些结论，从而提出哪些有意义的建议，企业应当针对这些结论采取什么样的措施。

5. 市场调查的局限性简述。说明受到时效性及经费等现实因素影响下调查结果中存在可能偏差。

除上述内容外，在摘要的篇幅上有如下建议可供参考：第一，调查背景与目标应尽量明确，简明扼要，最多用1~2页的文字进行阐释；第二，摘要的主体部分，即主要结果、结论及建议部分要比其他部分多几页，长短视具体情况而定。原则上，该部分也应当简短，一般认为最多不要超过报告内容的1/5。其中，主要调查结果的陈述应限制在为数不多的几页之内，结论和建议要言简意赅，必要时还应加上简短的解释：

## （三） 市场调查背景与目的

1. 市场调查背景。调查背景是整个研究的缘由与整个调查方案的依据。研究者要对调查的由来或受委托进行该项调查的原因做出说明。调查背景中需要着重说明的问题有两个：一是企业的产品或服务当时面临的问题或所处的状态；二是在当时的情况下，想通过调查解决什么问题。一般而言，需要根据委托企业所在的行业特点进行调查背景的说明。具体而言，调查背景可能会涉及以下一些方面的问题。

(1) 委托企业所在的行业特点及企业的整体管销战略 D 它主要包括：产品或服务在一定时期内的销售情况，产品价格、包装情况，企业的销售渠道及销售方法，广告、促销策略与实施情况，消费者的基本情况及消费者对企业营销执行情况的反应。

(2) 竞争对手的情况。主要包括：竞争对手的市场占有情况，对手的营销策略，对手的营销手段组合及市场反应情况，对手的新动向。背景资料的介绍不仅可作为调查目的提出的铺垫，还可以作为调查结论

和建议的佐证，并可与调查结果相结合来说明问题。背景资料的介绍 不一定要面面俱到，但必须与调查主题有关。

2. 市场调查目的。调查目的通常针对背景分析认为存在的问题而提出。它一般是为了 探究市场上的某些情况或论证某些假设。例如，产品销售下滑的原因是什么。该问题可能有 两个方面：一方面，如果不清楚产品下滑的原因，那就需要进行探究性的调查，找到影响产品销 售下滑的主要原因，从而提出解决方案；另一方面，如果企业已经感觉到产品下滑的一些原 因，但不能肯定或者不知道这些影响因素到底有多大，以及在众多影响因素中，重要的因素是 哪些。针对这样的问题就需要对假设进行详细的检验与论证。在调查目的中，研究者应当对 本项调查的预期结果列出一张清单，主要包括事实性的描述、主要变量之间的关系分析、主要 假设及其验证方法与结果。

#### (四) 正文

正文是市场调查书面报告的核心部分，它决定着整个调查报告质量的高低和作用的大小。它使用客观 中性的语言阐述通过调查了解到的事实，分析说明被调查对象发生、发展和变化 过程，调查的结果及存在的问题，提出具体的意见和建议 D 正文一般涉及很多内容，通常需分 成若干章节。

##### 1. 正文文字表述要求。

首先，应提供与本次市场调查有关的概念及重要指标的定义(更多细节可列在附录中)， 对调查数据资料及背景资料和调查使用的方法进行说明。其次，论述调查结果时，通常是以文 字为主，插入适当的图、表。数据汇总表一般附在报告的后面。最后所使用的语言应尽可能简 洁明了。对报告主体的每个章节来说，常用且有效的结构是针对一个个引人关注的问题，陈述 该问题的各种数据及材料的合理结构和问题答案。每一个用于陈述结果的章节都应该先概述 主要的调查结果及研究成果，紧接着给出更详细的内容’ 应当按照各种论点的重要程度合理 地安排它们的逻辑顺序。可以是按照主报告、分报告逐一展开；也可以是按重要性排序的调查 结果一一展开。既要层次分明，又要重点突出。

关于正文论述部分的层次段落，一般有四种展开形式：

第一，层层深入形式。各层内容之间是一层深入一层，层层剖析。

第二，先后顺序形式，按事物发展的先后顺序安排层次，各层内容之间有密切联系。

第三，综合展开形式。先说明总的情况，然后分段展开，或先分段展开，然后综合说明。 第四，并列形式。各层之间是并列的关系。

当然，也不限于以上的形式，撰写市场调查书面报告时应结合主题的需要，采取适当的形式，使之能充分反映主题。

##### 2. 统计表编制的注意事项。

市场调查书面报告涉及大量统计表的编制与使用「合理使用统计表可以增强报告的清晰 与简明性以下是编制表格的注意要点：

(1) 表格应该尽量简单并且只显示主要信息。通常情况下一个表格只说明一个要点。使用两张简单的表格比使用一张过于复杂的表格更好。

(2) 整体格式、间距和标题位置等表格布局应该合理，不要使表格过长或过短。

(3) 总标题应清楚、简洁、醒目，应避免缩写。

(4) 行标题、列标题有时可以互换，但应当有助于表中数据的清晰表达，避免产生误解。

(5) 各标题应该按一定顺序合理排列，用于支持分析内容。一般应将最显著的放在前面，如果强调的是时间，可按时间顺序排列，如果强调的是大小，就按大小顺序排列，当然也可以按其他顺序排列。

(6) 分组要适当，不可过细，以免烦琐。尤其是频数分布，若分组过细，就会使频数太少，不能看出分布的规律，但分组太粗也会掩盖差别，不易看出分布的规律。

(7) 所有计量单位都必须表述清楚，如果表中的数据都是同一计量单位，可放在表的右上角标明，若各指标的计量单位不同，则应在每个指标后或单独列出一列标明。

(8) 表格的设计应尽可能简便易读，通常表的左右两边不封口，印刷要清晰。

(9) 表中的数据，个位数、十位数等应上下对齐，一般是右对齐，如果有小数点时应以小数点对齐，而且小数点的位数应统一。

(10) 对于没有数字的表格单元，一般用“—”表示。一张填好的统计表不应出现空白单元格。

(11) 应保证数据的舍入不会影响读者对估计精度的理解。

(12) 当数据来自抽样调查时，应当使用最终的样本权数计算估计值和精度。

(13) 在引用别人的统计表时，可在表的下方加注释，要注意注明来源，以表示对他人劳动成果的尊重，也方便读者查阅使用，了解指标的口径。

### 3. 统计图编制中的注意事项。

统计图可以直观地显示数据，它能集中强调数据的相对模式，反映数据的分布形状、结构和数量的级别等。好的统计图是对正文和表格的补充，而不仅仅是重复的信息。适用于统计表的基本规则一般也同样适用于统计图。统计图编制中的注意事项如下：

(1) 保持形式上的简单明了，一个图一般只说明一个要点。

(2) 每张图都要有标号与标题；标题简明扼要。

(3) 项目较多时按照一定顺序排列，使其一目了然。

(4) 尽可能将图标的含义和表示的数量标记在对应的位置上，避免使用附加的图标说明。

(5) 保持变量数值与作图刻度之间的比例关系适当，表达清晰又简明。

(6) 适当选择度量单位，使得图形表现均衡，变量间的差异可见且可解释。

(7) 作图时使用线型、图标及文字综合说明。

(8) 在颜色和纹理选择时突出重要内容的表达，并具有逻辑性。

(9) 图形的排列符合人们的阅读习惯。

(10) 图形下面注明数据来源。

(11) 一般情况下应避免使用三维图形。

(12) 当用样本数据制图来说明总体情况时，必须使用最终样本权数对抽样调查获得的数据进行加权，以防止估计中的偏差。

### (五) 结论

结论部分是调查报告的结束语，好的结尾，可使读者明确主题，加深对问题的认识，并能进一步启发思考和联想。结尾通常有以下几种形式可供参考：

第一，概括全文。经过层层剖析后，综合说明调查报告的主要观点，深化报告的主题。

第二，形成结论。在对收集到的资料进行深入细致的科学分析的基础上，得出报告的 结论 6

第三，提出看法和建议。通过分析，形成对事物的看法，在此基础上，提出建议和可行性 方案。

第四，展望未来，说明意义 通过调查分析，展望未来前景 c

#### . (六) 建议

该部分是可选的 “因为前面的结尾中已经对调查结果提出建议，这里可侧重于调查方法 技术方面，在技术报告中可以就任何调查活动中遇到的问题提供建议。通常这样做将会有益 于其他调查项目的管理者，因为他们可能会遇到同样的情况，而且对以后的调查也会有帮助。

#### (七) 参考书目与附录

所有的参考文献都应该清楚地标明

附录是指调查报告的主题包含不了或没有提及，但与主题有关必须附加说明的部分。它 是对主报告的补充和更详细的说明这样做可以使报告的主题更突出，也避免了正文过于冗 长附录可以包括以下的内容：

第一，调查目的和研究总体，借助问卷收集的数据类型；抽样设计的主要特点，以及建立抽 样框或抽选样本中遇到的问题

第二，现场采集数据的方法，实地调查员的培训和监督的程序：

第三，调查中所使用的计算机硬件和软件，以及计算机系统的开发。

第四，数据的汇总表。

附录还可以包括其他重要文档，例如术语索引、各种数据表、关于数据质量的技术细节、方 法论、调 查问卷以及对统计检验的描述等 在报告适当地方，还可以给出相关的数学公式，例 如方差估计公式。

附录中要包括承担项目方的联系方式:这一部分也是必需的。对任何调查，必须提供给 用户一个明确的直接联络人，这个人应十分 r 解这个项目。应该写明联络人的电话、传真、电 子邮件地址及通讯地址等 网址的提供也越来越受重视，因为从网上可以查到很多信息。

#### (八) 数据质量报告

数据质量报告是指对整个市场调查项目完成全过程的质量评估的书面材料一在任何市场 调查中，数据质量评估都是一个重要的步骤，并且应该记录在主报告中，或撰写一份单独的数据质量报告 数据质量评估指根据统计活动的最初目的，从数据的精确性和可靠性等方面对 调查的最终产品进行评估的过程这种信息可以帮助数据用户更好地对调查结果进行解释和 使用 用户必须学会评估数据的误差会在何种程度上限制他们对数据的使用。但是，很少有 用户能够独自估算调查所得数据的精确性，因此市场调查机构有责任进行必要的数据质量评 估，并及时地以易于使用的方式向数据用户发布评估的结果数据质量评估对统计机构同样 有利对误差的追踪应具体到调查过程的每一个特定的步骤，这样的质量评估可以用于改进 调查后续阶段以及其他同类调查的数据质量控制和质量保证活动适用于调查的每一步骤。 影响数据质量的每个问题都应予以报告和记录，并且尽快找到解决办法。分析期间，分析人员 应该验证所得的估计值和表格，结果是否内在一致？这意味着不同表格中关于同一个变量的 边际总和应该相等，子总体的数量相加应该等于总体的总数量，等等还有总量与其他数据来 源的结果是否一致？最后，还应该考虑使用验证法。验证涉及很多数据分析步骤，包括对数据 资源与历史数据的大量比较研究。

根据发达国家的经验，大规模的市场调查数据都需经过严格的质量评估和近乎苛刻的评 审，以判断在特定的报告地区水平下，向公众发布这些调查数据的适当性和可靠性。所谓适当 性指的是数据符合特定的质量标准和保密标准的程度。数据质量报告中应将调查的方法论与 数据质量指标结合起来。对方法论的描述应使用户能够判断数据是否适合用来估计他们想要 测量的指标，以及所得的估计对预期目标来说是否可 接受，

数据质量报告在内容上的详细程度依赖于下列各组成部分：

第一，收集的数据类型、数据来源类型（普查、抽样调查、行政数据等）和数据分析的类型。 第二，发布媒体。

第三，数据使用的范围和影响，

第四，项目的总预算 G

第五，数据的用途（对政策、社会及经济计划的影响等）。

第六，潜在的错误和应用这些数据的潜在危险。

第七，质量随时间的变化（指重复性调查）。

第八，相对于总费用的质量评估费用。

第九，在以后调查中提高质量的潜力 G

第十，提供给用户的各种度量的解释要通俗易懂。

下面是可供参考的一份不完全的数据质量报告的项目清单：

第一，给用户的注释（如果适用的话），它应该包括信息的重要部分或需要特别提出来加以说明和用户应该注意的问题。

第二，数据的来源和有关的方法。如一般的方法问题，目标总体、对数据来源及收集方法的说明、数据处理和估计方法的说明、数据的标准期、数据的修订及调整情况（如果有的话）。

第三，概念和变量的测量。这部分应包括主要的概念、变量、所用的分类及关键指标“

第四，数据的准确程度。应该有一个关于数据准确程度的说明，包括分地区、分特征的抽样误差和非抽样误差的类型和来源（例如：无回答率，插补率等）。

第五，调查数据与其他相关资料的可比性 Q 如果存在这样的问题，应该说明这些数据是否可以和过去进行比较或不能进行比较，并说明原因（如问题的问法有改变）。

第六，其他质量指标和评介。（如果有的话）例如，给出了分析结果，应概述所用的分析方法和对统计的显著性进行讨论。

第七，附录（如果有必要的话）。

### 三、市场调查书面报告的写作技巧

市场调查书面报告的撰写过程包括解析调查材料，确定主题思想，构建文体结构，起草、修改与定稿。除了需要多次交流与修改外，写作中还需要掌握以下技巧：

#### （一）开头部分的写作技巧

书面报告开头部分的撰写一般有以下几种形式：

##### 1. 开门见山，揭示主题

一开始就交代市场调查的目的或动机，揭示主题。例如，“我公司受华为集团的委托，对消费者进行一项有关智能手机市场需求状况的调查，预测未来消费者对智能手机的需求量和需求的种类，使华为集团能根据市场需求及时调整其产量及种类，确定今后的发展方向”

##### 2. 结论先行，逐步论证

先将调查的结论写出来，然后逐步论证。许多大型的市场调查报告均采用这种形式。特点是观点明确，使人一目了然。例如，“我们通过对玉露茶叶（湖北恩施富硒茶品牌）在武汉市的销售情况和购买意向的调查，认为它在市场竞争力上存在不足，原因主要从以下几方面阐述……”

##### 3. 交代情况，逐步分析

先交代背景情况，调查数据，然后逐步分析，得出结论。例如，“本次关于德系车的销售情况的调查

主要集中在北京、上海、广州、深圳，调查对象集中于中青年消费者……”

#### 4. 提出问题，引入正题

用这种方式提出人们所关注的问题，引导读者进入正题。例如，“奢侈品在中国的消费额不断增加，目前中国已经成为全球第三大奢侈品市场。中国人为什么热衷于购买奢侈品呢？……”

#### (二) 论述部分的写作技巧

论述部分必须准确阐明全部有关论据，根据推断所得的结论，建议有关部门采取的措施，最终形成的解决问题方案。论述部分主要包括基本情况和分析两部分。基本情况部分是对调查数据资料及背景进行客观地介绍说明、提出问题并积极回应问题；分析部分主要包括原因分析、利弊分析和预测分析三种类型，写作时注重推理的依据、方法的科学与合理、论证的严谨与充分。

#### (三) 写作部分的相关技巧

市场调查书面报告的写作技巧主要包括文字表达、表格和图形表现等技巧。表达技巧主要包括叙述、说明、议论、语言运用四个方面的技巧。

##### 1. 叙述的技巧

市场调查的叙述，主要用于开头部分。叙述事情的来龙去脉，表明市场调查的目的和根据，以及过程和结果。市场调查报告常用的叙述技巧有：

(1) 概括叙述。叙述有概括叙述和详细叙述之分。市场调查报告主要用概括叙述，将调查过程和情况概略地陈述，不需要对事件的细枝末节详加铺陈。这是一种浓缩型的快节奏叙述，文字简约，一带而过，给人以整体、全面的认识，以适合市场调查报告快速、及时反映市场变化的需要。

(2) 按时间顺序叙述。市场调查报告在交代调查的目的、对象、过程时，往往按时间顺序叙述，前后连贯。如开头部分叙述事情的前因后果，主体部分叙述市场的历史及现状，就体现为按时间顺序叙述。

(3) 叙述主体的省略。市场调查报告的叙述主体是写报告的单位，叙述中用“我们”称呼。为行文简便，叙述主体一般在开头部分出现后，在后面的各部分即可省略，并不会因此而令人误解。

##### 2. 说明的技巧

市场调查报告常用的说明技巧有：

(1) 数字说明。市场主体活动的过程与结果都可表现为数字，反映市场发展变化情况的市场调查报告要运用大量数据，以增强报告的精确性和可信度。

(2) 分类说明。市场调查中所获材料杂乱无章，根据主题表达的需要，可将材料按一定标

准分为多种类别，进而分别说明，例如，根据调查的基本情况•按问题性质归纳成几类，或按不同层次分为几类。每类前冠以小标题，按关键句的形式表述。

(3) 对比说明。市场调查报告中有关情况、数字说明，往往采用对比形式，以便全面、深入地反映市场变化情况。对比说明要清楚事物的可比性，在同标准的前提下做切合实际的比较-

(4) 举例说明。为说明市场发展变化情况，举出具体典型事例是常用的方法。市场调查中会遇到大量事例，应从中选取有代表性的例子。

### 3. 议论的技巧

市场调查报告常用的议论技巧有：

(1) 归纳论证。市场调查报告是在获得大量材料之后，进行分析研究，得出结论，从而形成论证过程。这一过程主要运用议论文方式，所得结论是从具体事实中归纳出来的。

(2) 局部论证。市场调查报告不同于议论文，不可能形成全面论证，只是在情况分析、对未来预测中做局部论证。如对市场情况从几个方面做分析，每一方面形成一个论证过程，用数据、情况等做论据去证明其结论，形成局部论证。

### 4. 语言运用的技巧

语言运用的技巧包括用词方面和句式方面的技巧'

(1) 用词方面。市场调查报告中数量词用得较多，因为市场调查离不开数字，很多问题要用数字说明。可以说，数量词在市场调查报告中以其特有的优势，越来越显示出其重要作用：市场调查报告中介词用得也很多，主要用于交代调查目的、对象、根据等方面，如用“为、对、根据、从、在”等介词。此外，还多用专业词，以反映市场发展变化，如“企业战略”“流通机制”“市场竞争”等词。为使语言表达准确，撰写者还需熟悉市场研究领域有关专业术语。

(2) 句式方面。市场调查报告多用陈述句，陈述调查过程以及调查到的市场情况，表示肯定或否定判断。祈使句多用在提议部分，表示某种期望，但提议并非皆用祈使句，也可用陈述句。

## 四、撰写市场调查书面报告的注意事项

### (一) 报告的针对性

针对性是市场调查报告的灵魂，主要包括两方面：

第一，撰写市场调查报告必须明确调查的目的，撰写报告时必须做到目的明确、有的放矢，围绕主题展开论述。

第二，市场调查报告必须明确阅读对象•阅读对象不同，他们的要求和所关心问题的侧重点也不同。如果既不明确解决什么问题，又不明确读者对象，针对性不强，撰写出来的调查报告就是盲目和毫无意义的。例如针对市场研究人员而写的调查报告，可以篇幅长些，突出一些技术性细节，有助于其判明市场调查结果的科学性，以决定其对报告有关资料的取舍及应用程度，针对企业决策者而写的调查报告则要简明扼要，突出结论和重点，并注意运用图表加以表现，调查过程和细节可以略写。

### (二) 市场调查报告的新颖性

市场调查报告应紧紧抓住市场活动的新动向、新问题，引用一些人们未知的通过调查研究得到的新发现，提出新观点，形成新结论。特别是能紧密结合市场发展的新形势、新变化，为科学决策提供依据。这样的调查报告，才更有使用的价值。避免只把众所周知的，常识性的或陈旧的观点和结论写进去。

### (三) 市场调查报告的可读性

调查报告作为一种应用型文体，应重视其可读性。首先是观点鲜明，突出；其次是内容的组织安排有序；再次是行文流畅、通俗易懂。

### (四) 市场调查报告的公正性

调查报告要完整报告调查的结果，不应略去或故意隐藏所知事实。如果调查实施中出现严重问题（如回收率过低），研究者应有勇气承认，同时不能随便报道结果，以免误导读者。即使便是成功的调查，在调查报告中也不应只选择那些对自己有利的结果进行报道，其他则避而不提。

在具体的写作过程中有下面一些指导性的建议供参考：

第一，措辞简洁清楚：一个好的调查报告要传达很多详细而精确的信息。别让不必要的细枝末节弄乱了整篇文章。语言需要简洁，但并不总是越短越好，表述清楚永远是第一位的。

第二，少用缩略语。不要设想读者一定知道缩略语的含义，当有可能引起歧义时就要进行详细解释。首次提到某个术语时，应该写完整，并在其后的括号中给出它的缩写形式。如：消费者物价指数（CPI），按规模大小成比例的概率抽样（PPS），在下文中可单独用该缩略语。术语首次出现后其缩略语的使用在整篇文章中应前后一致，不要在某些地方使用全称，而在其他地方使用缩略语。一旦你选定了一种缩写方式，在同一文档中就不要使用其他形式。

第三，使用术语要一致。报告各部分所使用的术语必须保持一致。如果标题和正文提到的是“资产与债务”，那么表和图中就不要标记为“资产负债表”。

第四，谨慎使用“其他”分类。在分类中，有些情况不好分类而列入其他类，尽管其他类别经常大量存在并且值得注意，然而“其他”只是一个含义不明、模棱两可的术语。所以在分类时只要有可能，最好精确地解释或识别其各种成分，而不是笼统地称其为“其他”。

第五，使用清晰的时间概念。不要使用模糊不清的日期，比方说上一年或上个月等。在首次出现时应该使用明确的年和月。在每次发布调查数据以及报告的主标题下应立即给出相应标准时期。

## 第二 T>

### 市场调查口头报告的撰写

市场调查口头报告很多时候是其书面报告的缩略版，但仍要从其重要性和特点出发，精心准备和高质量完成。本节主要介绍市场调查口头报告的重要性及特点、需要准备的材料和报告时的注意事项。

#### 一、口头报告的重要性及特点

在很多情况下，需要将市场调查报告的结果向管理层或委托方进行口头报告。口头报告可以帮助管理部门或委托方简明扼要地理解书面调查报告的内容并接纳书面报告，同时，可以针对委托方提出的问题及时做出解答。口头报告对于有关人士迅速掌握和理解报告内容具有重要的作用。

与书面报告相比，口头报告具有以下特点：

1. 口头报告能用较短的时间说明所需调查的问题。
2. 口头报告生动、具有感染力，容易给对方留下深刻印象。
3. 口头报告能与听者直接交流，便于增强双方的沟通。
4. 口头报告具有一定的灵活性，一般可根据具体情况对报告内容、时间做出必要的调整。

## 二、口头报告需要准备的材料

一份能发挥其效果的市场调查口头报告需要进行精心的设计和充分的准备。具体归纳为以下三点：

1. 演讲提纲及汇报提要。按照书面调查报告的格式提纲挈领准备好详细的演讲提纲。采用演示报告方式并不意味着可以随心所欲、信口开河。它同样需要有一份经过精心准备的提纲，包括报告的基本框架和内容。为了取得好的效果，使报告的内容和风格与听众的知识水平相吻合，需要在报告前了解听众的兴趣、专业技术水平和文化程度。必要时可为听众准备一份汇报提要，便于听众记录或述评。

2. 视觉辅助工具。为了使报告更生动灵活，富有吸引力，提高报告效果，在条件许可的情况下，应尽量借助现代技术包括胶片和投影仪、录像片和电视以及电脑投影等作为辅助手段，以保持与会者的注意力，增强记忆，同时也可以促使讲解者以一种规范方式去组织思维。演示文稿 PPT (PowerPoint) 是其最常见的表达方式。

3. 最终报告简明版。受托方主要负责人在口头汇报中省略了报告中的很多细节，作为补充，在口头报告结束时，应该准备一份最终报告的复印件，以备需要者索取。在有些情况下，需要将最终书面报告在做口头报告之前呈现给听众。

## 三、进行口头报告时应注意的问题

1. 十分熟悉报告的内容。按照书面调查报告的格式准备好演讲提纲后，需要加强练习，直至十分熟悉报告的内容。

2. 采用通俗易懂的语言。口头报告要求语言简洁明了，通俗易懂，要有趣味性和说服力。如果问题比较复杂，可先做一简要、概括的介绍，并运用声音、眼神和手势等变化来加深听众的印象。

3. 运用简明清晰的图表。口头报告时，要善于使用计算机生成的图表来辅助和支持演讲的内容。增强陈述的生动性、丰富性与形象性。注意图表的清晰、简明和规范。

4. 充满自信地进行报告着装得体，仪态端庄；开头结尾语句尊重有礼；口头报告时面对听众，与其保持目光接触，随时判断听众的喜爱与理解程度。

5. 把握回答问题的时机。在报告开始前可告之听众，你会在报告后回答问题并进行个别交流。在报告过程中最好不要回答问题（有关演讲清晰性问题除外），保证演讲思路顺畅、主题连贯，避免出现时间不够等现象。

6. 掌握报告时间。口头报告常有一定的时间限制，在有限的时间内讲完报告是最基本的要求。时间利用不充分，往往不能表达完整的思想；超过报告时间，不仅是对听众的不尊重，影响后续事项的安排，而且经常因为强制中断报告，会极大地影响报告的效果。

### 痞章小结…

市场调查报告的撰写与呈递是市场调查工作的最后阶段。市场调查报告是为了某个特定目的提供给调查委托者的调查结果，一份好的调查报告是决策者进行决策的重要依据。Q 调查

报告按照不同分类标准可以归并到不同的类型中。调查报告按呈递方式有书面调查报告和口头报告两种。书面报告的写作要按照基本要求进行。写作要求包括实事求是、目的明确、简明易读、观点和数据统一。其写作步骤包括解析调查材料，确定主题思想，构建文体结构，起草、修改与定稿。一份规范的书面报告的文体结构包括封面、标题、目录、摘要、正文、结论、建议、参考书目与附录、数据质量报告。撰写一份好的市场调查书面报告需要掌握一定的写作技巧，主要包括开头部分、论述部分和写作中的相关技巧，并谨记注意事项。在多数情况下，客户可能会要求进行调查结果的口头报告，口头报告既可以作为书面报告的补

充也可以作为独立的 报告形式。一份口头报告需要进行充分的准备，包括演示文稿的设计制作、演讲场景的结构布置、文本材料的准备及演讲人的反复练习等，目的是取得汇报的成功 关键词•••••

市场调查书面报告 市场调查口头报告 统计表 统计图

肯思考题 .....

回！即测即评

1. 市场调查书面报告的作用是什么？
2. 市场调查书面报告的基本结构包括哪些？
3. 市场调查书面报告的撰写原则是什么？
4. 市场调查书面报告的写作步骤有哪些？
5. 在展示调查结果时如何利用好统计表和统计图？
6. 如何准备一份高质量的口头报告？

碧本章实训. . .

### 一、 实训目的

培养学生实际撰写市场调查书面报告和制作口头报告的能力

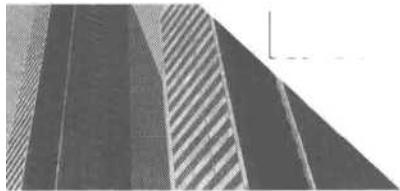
### 二、 实训内容

从前面章节实训要求所提出并实施后的某一市场调查项目为基础，以其结果为核心，撰写 一份规范的市场调查书面报告和相关的口头报告

### 三、 实训组织

1. 按前述项目小组成员的团队构成，每组成员不少于 4 人，采用组长负责制。
2. 指导教师说明实训项目的内容，才旨出要点和注意事项。
3. 小组成员讨论形成核心思想，归纳要点，形成讨论稿，完成市场调查书面报告和口头 报告
4. 各小组在班级展示市场调查书面报告和口头报告，并由各组互评成绩，（可参见附录 3：第八届全国大学生市场调查与分析大赛评分标准）
5. 指导教师进行综合评定和总结

# 附录 1 2018 年中国景区旅游消费研究报告



《2018 年中国景区旅游消费研究报告》是通过艾瑞 iClick 在线调研社区、高星景区（4A/5A）工作人员进行面访或电话访谈，对在线景区门票市场销售情况、企业品牌、用户满意度、用户出行习惯及预订习惯进行在线调查及研究。以期为关注在线景区门票市场的企业及从业者提供参考依据”（资料来源：艾瑞网, <http://report.iresearch.cn/>）

（报告为 PDF 文档，通过二维码扫描可以下载）

回院回

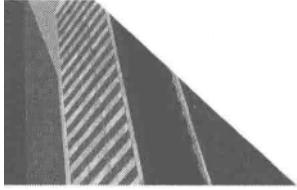


## 附录 2 入户访问项目遵循规范



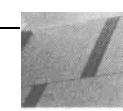
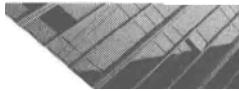
由中国市场研究协会制定的“入户访问项目”行业执行标准，于 2016 年提出，成为市场调查与研究行业执行入户访问项目时的遵循规范。资料来源参见：<http://www.cmra.org.cn/>  
(报告为 PDF 文档，通过二维码扫描可以下载)





# 附录 3 第八届全国大学生市场调查与分析大赛评分标准

---



一. 曹.

序号	分类	分项	小项	分项得分	得分	
1	书面报告 (40%)	研究选题	选题新颖	10%		
			结合实际			
2		方案设计	调查方案设计合理	10%		
			问卷设计科学			
			调查方法设计合理			
3		调查实施	调查组织合理	10%		
			调查程序完整			
			调查控制严格			
4		分析与结论	选择方法正确	10%		
			数据运用充分			
			分析支持结论			
5	演示报告 现场答辩 (60%)	展示	展示文件水平	10%		
6		陈述	逻辑严谨	25%		
			层次清楚			
			语言表达准确清晰			
7		答辩	理解问题	25%		
			思路清晰			
			回答正确			
		总计		100		

特别注意：报告陈述展示 10 分钟，现场答辩 5 分钟。

- [1] 柯惠新, 丁立宏. 市场调查[M]. 北京: 高等教育出版社, 2008.
- [2] 倪加勋. 调查概论[M]. 北京: 中国财政经济出版社, 2004.
- [3] 蒋萍. 市场调查[M]. 2 版. 上海: 格致出版社, 上海人民出版社, 2013.
- [4] 王旭. 市场调研[M]. 北京: 高等教育出版社, 2012.
- [5] (美) 戴维 S 穆尔. 统计学的世界[M]. 5 版. 北京: 中信出版社, 2003.
- [6] (美) Douglas C. Montgomery. 实验设计与分析[M]. 傅珏生, 等译·北京: 人民邮电出版社, 2009.
- [7] 简明, 金勇进, 蒋妍. 市场调查方法与技术[M]. 北京: 中国人民大学出版社, 2012.
- [8] 陈殿阁. 市场调查与预测[M]. 北京: 清华大学出版社, 2004.
- [9] 郭星华, 谭国清. 问卷调查技术与实例[M]. 北京: 中国人民大学出版社, 1997.
- [10] 张灿鹏, 郭砚常. 市场调查与分析预测[M]. 北京: 北京交通大学出版社, 2013.
- [11] 鄢奋. 市场调查[M]. 北京: 经济管理出版社, 2015.
- [12] 宋文光. 市场调查与分析[M]. 北京: 高等教育出版社, 2015.
- [13] 陆跃祥. 消费者行为学[M]. 北京: 中国统计出版社, 2005.
- [14] 马连福、张慧敏. 现代市场调查与预测[M]. 北京: 首都经济贸易大学出版社, 2016.
- [15] 庄贵军. 市场调查与预测[M]. 北京: 北京大学出版社, 2014.
- [16] 罗洪群, 王青华. 市场调查与预测[M]. 北京: 清华大学出版社, 2011.
- [17] 胡咏雪. 市场调查[M]. 上海: 华东师范大学出版社, 2016.
- [18] 杜子芳. 市场调查实务[M]. 北京: 中国财政经济出版社, 2004.
- [19] 徐井岗. 市场调研与预测[M]. 北京: 科学出版社, 2004.
- [20] CarlMcDaniel Jr., Roger Gates. 市场调研精要[M]. 范秀成等, 译. 北京: 电子工业出版社, 2002.
- [21] 金勇进, 杜子芳, 蒋妍. 抽样技术[M]. 4 版. 北京: 中国人民大学出版社, 2015.
- [22] 贾俊平, 何晓群, 金勇进. 统计学[M]. 5 版. 北京: 中国人民大学出版社, 2012.
- [23] 徐映梅. 市场分析方法[M]. 北京: 中国财政经济出版社, 2006 .
- [24] 雷培莉, 张英奎, 秦颖. 市场调查与预测[M]. 2 版. 北京: 经济管理出版社, 2014.
- [25] 孙玉环. 统计学[M]. 北京: 中国统计出版社, 2015.
- [26] 孙玉环, 宋马林. 调研与精选案例[M]. 北京: 科学出版社, 2015.
- [27] 薛薇. SPSS 统计分析方法及应用[M]. 3 版. 北京: 电子工业出版社, 2013.
- [28] 闫秀荣. 市场调查与预测(第二版)[M]. 上海: 上海财经大学出版社, 2013.
- [29] (美) 达莫达尔 N. 古扎拉蒂 (DamodarN. Gujarati), 道恩 C. 波特 (Dawn C. Porter). 经济计量学精要[M]. 张涛, 译. 4 版. 北京: 机械工业出版社, 2011.
- [30] 徐建邦, 冯立. 统计学[M]. 大连: 东北财经大学出版社, 2013.
- [31] 韩明. 应用多元统计分析[M]. 上海: 同济大学出版社, 2013.
- [32] 吴明隆. 问卷统计分析实务:SPSS 操作与应用[M]. 重庆: 重庆大学出版社, 2015.
- [33] 卫海英等. 市场调研[M]. 北京: 高等教育出版社, 2016.
- [34] 张文彤, 钟云飞. IBM SPSS 数据分析与挖掘实战案例精粹[M]. 北京: 清华大学出版社, 2013.
- [35] (美) 梅纳德著. 应用 Logistic 回归分析[M]. 李俊秀, 译. 2 版. 上海: 格致出版社, 上海人民出版社, 2012.
- [36] 王力宾. 多元统计分析: 模型、案例及 SPSS 应用[M]. 3 版. 北京: 经济科学出版社, 2010.

- [37] 何晓群. 应用多元统计分析[M]. 北京: 中国统计出版社, 2014.
- [38] 李春林, 陈旭红. 应用多元统计分析[M]. 北京: 清华大学出版社, 2013.
- [39] 王斌会. 多元统计分析及 R 语言建模[M]. 4 版. 广州: 暨南大学出版社, 2014.
- [40] 朱建平. 应用多元统计分析[M]. 2 版. 北京: 科学出版社, 2012.
- [41] 黄良文. 统计学[M]. 3 版. 北京: 中国统计出版社, 2012.

## 郑重声明

高等教育出版社依法对本书享有专有出版权任何未经许可的复 制、销售行为均违反《中华人民共和国著作权法》，其行为人将承担相应的民事责任和行政责任；构成犯罪的，将被依法追究刑事责任。为了维护市场秩序，保护读者的合法权益，避免读者误用盗版书造成不良后果，我社将配合行政执法部门和司法机关对违法犯罪的单位和个人进行严厉打击。社会各界人士如发现上述侵权行为，希望及时举报。本社将奖励举报有功人员。

反盗版举报电话

(010)58581999 58582371 58582488

反盗版举报传真

(010)82086060

反盗版举报邮箱

dd@ hep.com.cn

通信地址 北京市西城区德外大街 4 号

高等教育出版社法律事务与版权管理部

邮政编码 100120

### 防伪查询说明

用户购书后刮开封底防伪涂层•利用手机微信等软件扫描二维码，会跳转至防伪查询网页。获得所购图书详细信息，用户也可将防伪二维码下的 20 位密码按从左到右、从上到下的顺序发送短信至 106695881280，免费查询所购图书真伪。

### 反盗版短信举报

编辑短信“JB,图书名称，出版社，购买地点”发送至 10669588128 防伪客服电话

(010)58582300

[General Information] 书名=14485428

SS 号=14485428