

调查方案设计

暨南大学 统计学系

陈光慧 教授

统计学研究的基本过程

实际问题

统计调查

(收集数据)

得出结论

(指导实践)

数据整理

(处理数据)

数据分析

(研究数据)



基本定义

- **统计调查：**根据统计研究目的的需要，有组织、有计划地搜集统计数据资料的过程。
- **统计调查方案设计：**关于统计调查的一份切实可行、周密细致的工作计划。



调查方案设计的基本内容

- (一) 确定研究问题和调查目的
- (二) 明确调查对象和调查单位
- (三) 统计调查方法
- (四) 统计数据搜集方法
- (五) 统计调查误差及控制
- (六) 其它问题

(一) 确定研究问题和调查目的

- 在社会科学、自然科学领域，遇到需要定性，特别是定量地研究解决某些问题时，均可考虑应用统计调查方法。
- 科学合理确定统计研究问题和调查目的是有效开展统计调查活动最为关键的一步。
- 调查目的不同，调查对象、范围、内容、方法就不同。目的不明确就无法确定向谁调查、调查什么、用什么方法取得数据。其结果就是调查混乱和调查结果错误。

举例说明:

研究背景:

- 近年来，某网络公司的用户增长速度持续放慢，离网客户大量增加。

研究问题:

- 评估该公司在竞争市场的优势、劣势，分析具体原因，寻找解决对策。

调查目的:

- 掌握目前竞争对手的各个品牌状况，包括市场占有率。
- 分析是什么因素导致用户转网？
- 用户转到了哪里？这些用户有什么特征？

（二）明确调查对象和调查单位

- **调查对象**：根据调查目的明确调查的内涵和外延，即统计调查中的总体。
- **调查单位**：构成调查总体的各个单位，如住户、学生、企业或其它事物。
- **报告单位**：具体负责报告调查数据的单位，只能是人或者机构。

举例说明：

研究背景：

- 某企业计划开发一款新的宠物食品。

研究问题及调查目的：

- 通过市场调查获取关于价格、包装、食量、周期、口味、配料等方面信息，为开发产品提供对策。

调查对象及调查单位：

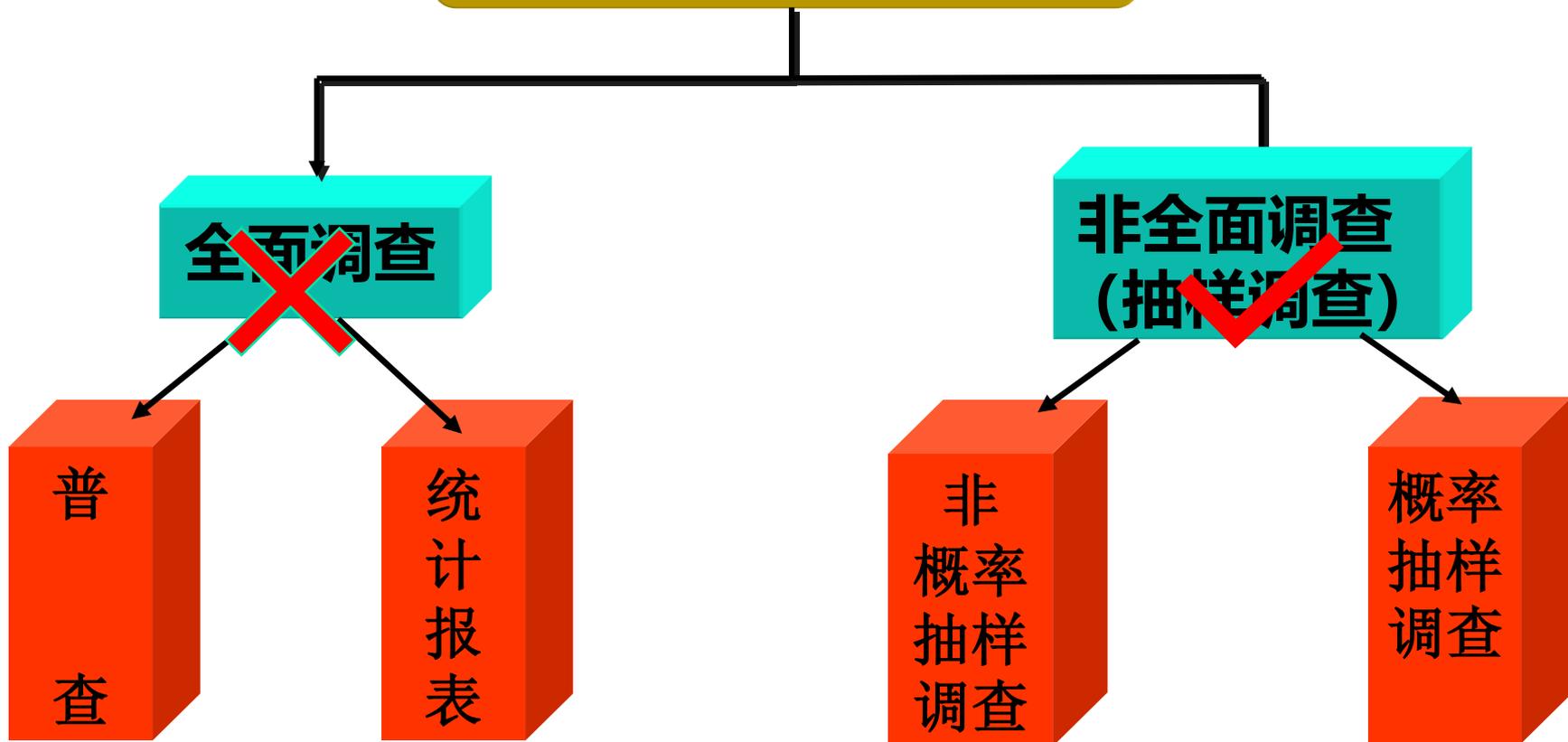
- 某地养宠物的人？宠物？

报告单位：

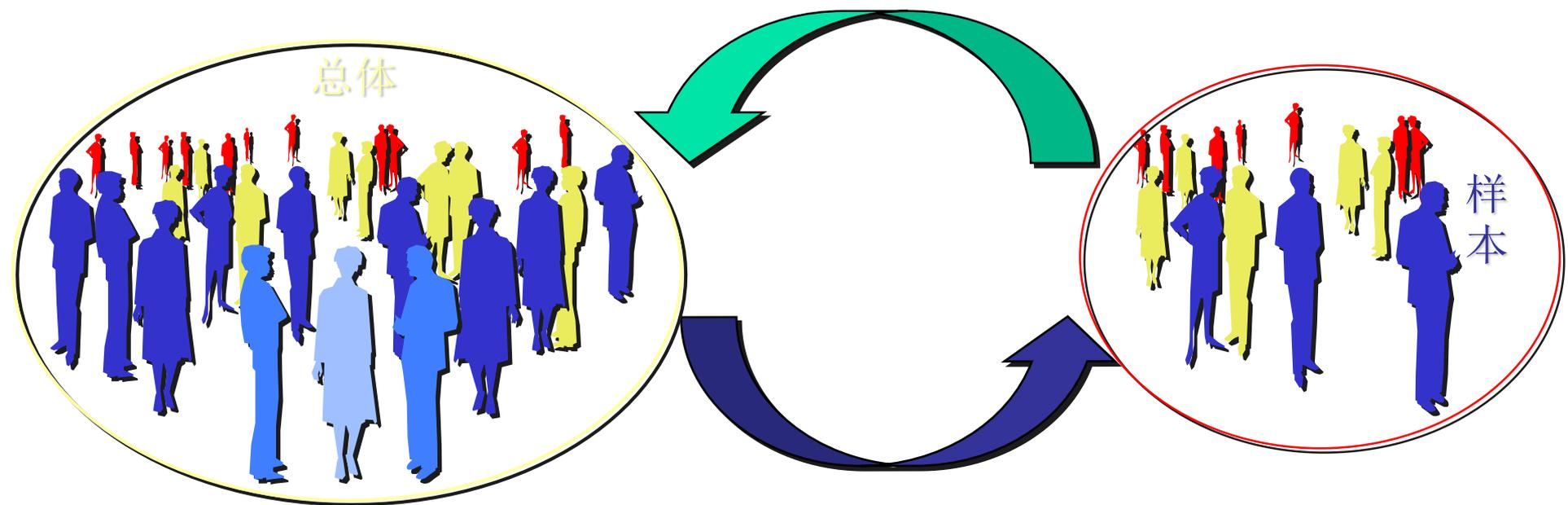
- 养宠物的人。

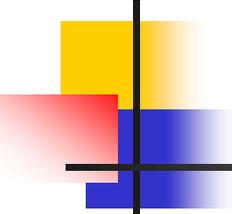
(三) 统计调查方法

统计调查方法分类



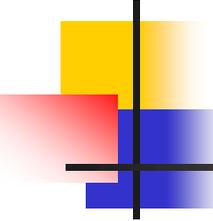
抽样调查的基本过程





抽样调查的优点

- (1) 节约调查费用
- (2) 时效性强
- (3) 适应面广
- (4) 准确性高



几种常用的抽样调查方法

非概率抽样调查

- 自愿样本
- 方便抽样
- 判断抽样
(包括重点调查和典型调查)
- 配额抽样
- 滚雪球抽样
-

概率抽样调查

- 简单随机抽样
- 分层抽样
- 系统抽样
- 不等概率抽样
- 整群抽样
- 多阶段抽样
-



1.非概率抽样调查

- 非概率抽样：以方便为出发点或根据研究者主观的判断来抽取样本。非概率抽样主要依赖研究者个人的经验和判断。
- **优点：**抽取样本方便、省时省力。
- **缺点：**难以计算和控制抽样误差，无法用样本定量化推断总体，也无法保证推断的准确性和可靠性。
- **应用场合：**用于定性的、探索性研究，调查目的主要是发现问题，寻找解决途径；还可用于概率抽样调查之前的预调查。

举例：非概率抽样调查

自愿样本：

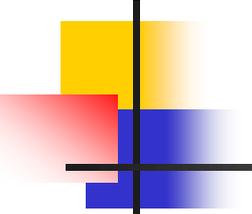
- 在微信朋友圈转发调查问卷。

方便抽样：

- 地铁口、交通路口、饭堂等拦截调查。

判断抽样：

- **重点调查**：组织某省产量最大的几个口罩企业开展调查；
- **典型调查**：辅导员通知几名有代表性的同学召开座谈会。



2. 概率抽样调查

- 概率抽样也称随机抽样，是一种以概率论和随机原则为依据来抽取样本的调查方法，要求使总体中的每个调查单位都有一个事先已知的、非零概率被抽中。
- 优点：可以计算并加以控制抽样误差；可以运用概率估计的方法对总体数量特征进行推断估计。
- 缺点：需构建抽样框，并保证随机抽取样本。
- 应用场合：推断和估计总体数量特征（总体均值、总值、比例等），从而可以定量化地进行统计分析研究。

举例：多阶抽样，结合其它抽样方法

以全国为总体的四阶抽样：

- 全国抽城市、城市抽街道、街道抽居委会、居委会抽住户。

以某省为总体的三阶抽样：

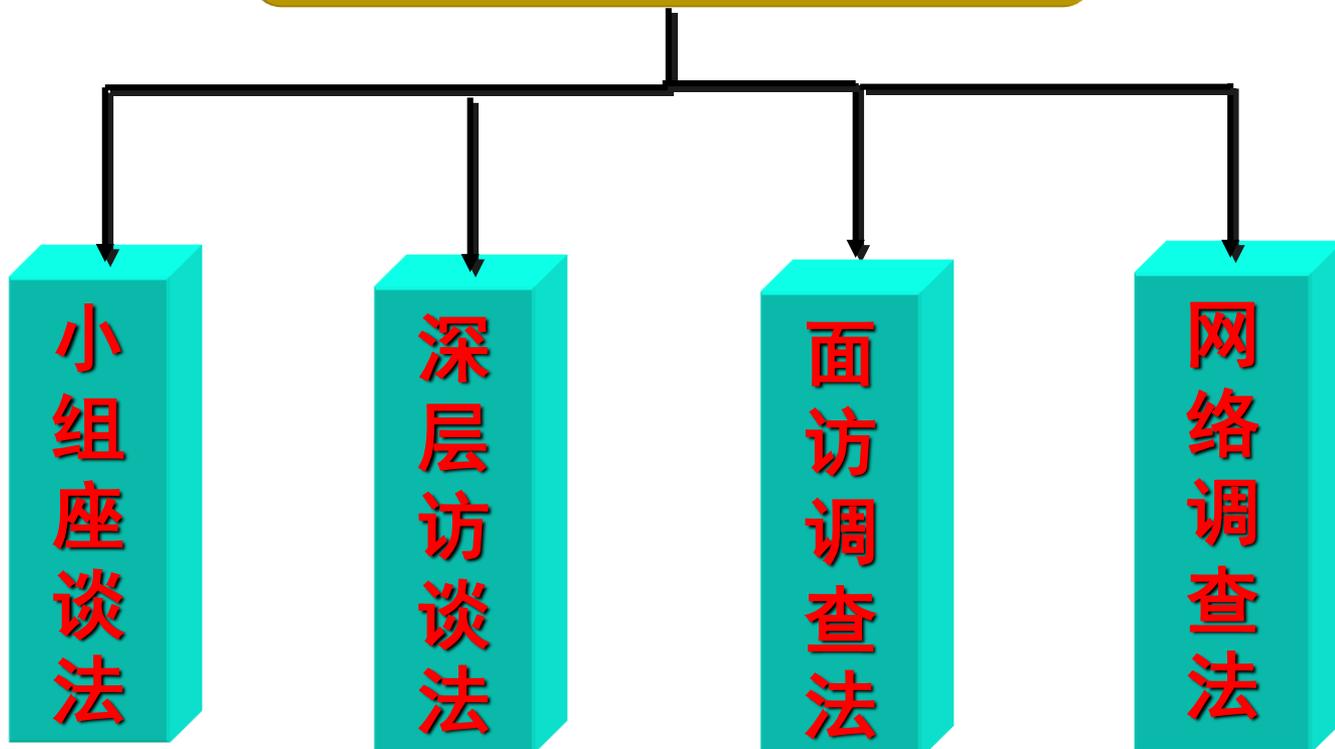
- 省抽县、县抽村、村抽住户或地块。

以某高校为总体的多阶抽样：

- 二阶抽样：全校抽班级、班级抽学生；
- 三阶抽样：全校抽专业、专业抽班级、班级抽学生；
- 三阶抽样：全校抽宿舍楼、宿舍楼抽宿舍、宿舍抽学生。

(四) 统计数据搜集方法

数据搜集方法分类



1.小组座谈会

- 也称集体访谈调研。
- 将一组被调查者集中在调查现场，让他们对调查的主题发表意见以获得调查信息。
- 参加座谈会的人数不宜过多，一般为6~10人，如学生座谈会、企业座谈会等。
- 侧重于定性研究，案例分析，发现问题，难以定量推断。



2. 深层访谈法

- 一次仅一名受访者参加、针对问题进行深入交流访谈。
- 适合于心理学、医学等领域较详细、隐秘的问题，如个人隐私、社会热点问题。
- 如是敏感性问题，通常需要与随机化回答方法结合使用。



3. 面访调查法

- 调查者与被调查者通过面对面地交谈，获得调查信息，是最常用的数据搜集方法。
- 有标准式访问和非标准式访问
 - 标准式访问通常按事先设计好的问卷进行
 - 非标准式访问事先一般不制作问卷
- 例如：入户调查访问、街头拦截访问等形式。





4.网络访问法

- 当前最为流行的数据搜集方法。
- 应用场合：通过问卷星设计问卷，使用微信等社交媒体发布、传播问卷；通过电子邮件发送问卷等。
- **优点：**线上操作完成、省时省力。
- **缺点：**难以保证样本的随机性，难以控制各类误差。这是值得继续探讨的难题。

(五) 统计调查误差及控制

调查总误差

```
graph TD; A[调查总误差] --> B[非抽样误差]; A --> C[抽样误差]; B --> D[抽样框误差]; B --> E[无回答误差]; B --> F[工作性误差];
```

非抽样误差

抽样误差

抽样框误差

无回答误差

工作性误差



1.非抽样误差

- 在抽样方案设计阶段，非抽样误差的来源：
 - (1) 抽样框的编制与准备不够充分完善；
 - (2) 问卷设计不够科学合理。
- 在数据收集阶段，非抽样误差的来源：
 - (1) 无回答误差；
 - (2) 数据计量误差。
- 在数据处理阶段，非抽样误差存在于对调查资料的整理、分组、计算、编码和计算机录入等过程中，是一种工作上的差错。

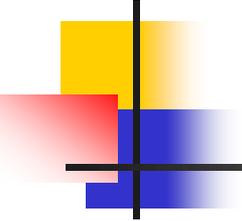


2. 抽样误差

- 抽样误差：用随机样本推断总体特征时所造成的样本推断值与总体真实值之间的差异。
- 抽样误差虽然不可避免、无法消除，但可以计算、控制，并逐渐减少。
- **影响因素**：样本量、抽样设计方法、抽样估计方法等。

(六) 其它问题

- 调查问卷设计
- 调查经费预算
- 组织实施计划：调查时间、地点、人员安排、工作分配.....



谢谢！ 欢迎多联系交流！