

“正大杯”第十二届全国市场调查 与分析大赛公益培训

1.5 抽样方法

1.5 抽样方法

主要内容

1

基本概念

2

简单随机抽样

3

分层随机抽样

4

整群抽样与系统抽样

5

非概率抽样

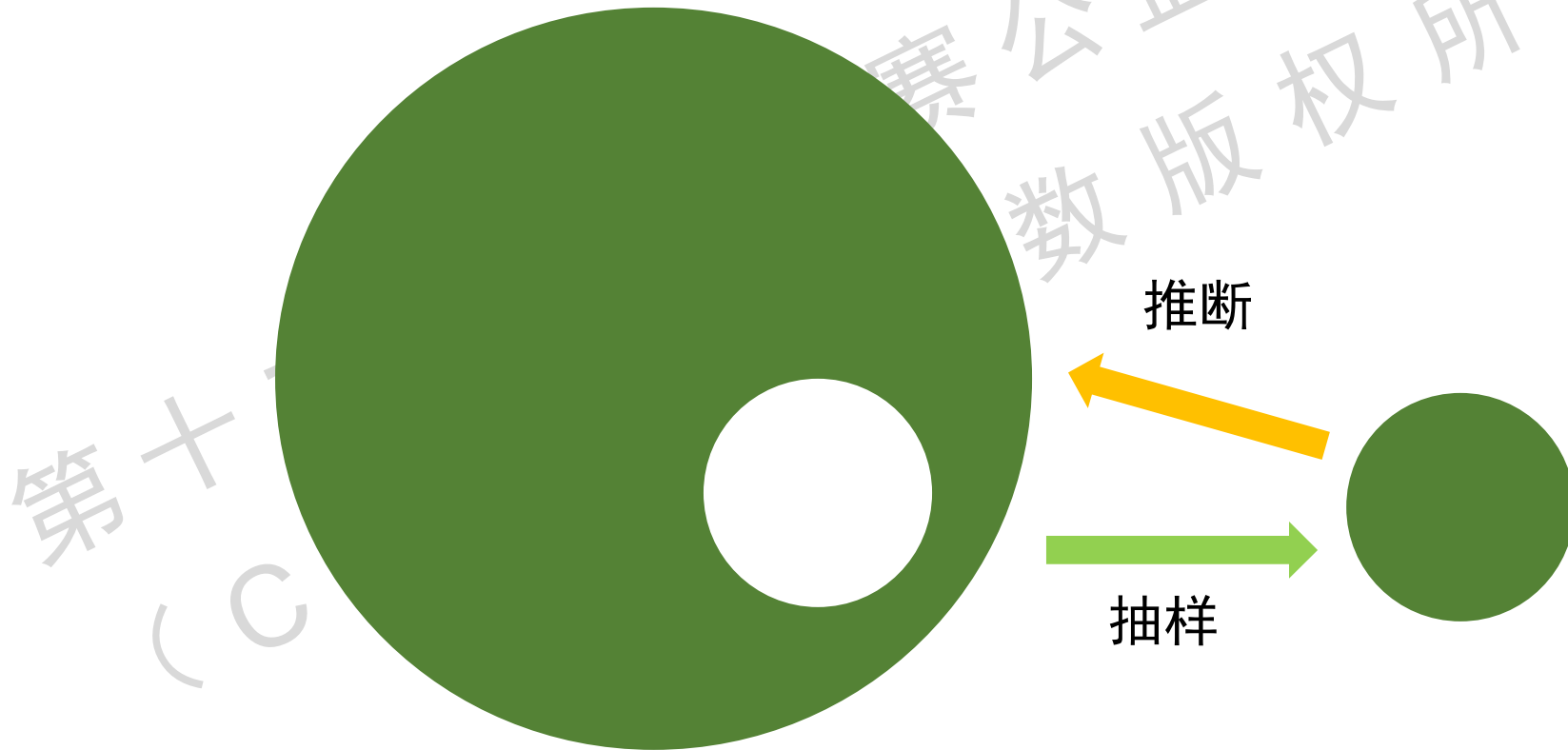


基本概念

第十二届市桥杯大赛公益培训课件
(Credamo 见数 版权所有)

基本概念

- **抽样调查**: 从构成总体的所有个体中按照一定的程序选择一部分个体, 并根据这部分个体的特征估计或推断总体的特征



基本概念

- **总体与个体**：总体与个体是抽样中最基本的一对概念，对于一次调查活动，所有调查客体的集合称为调查总体，相应的构成总体的这些调查客体则被泛称为个体。
- **抽样与样本**：从总体中选择部分个体的活动称为抽样，而抽样的结果，即从总体中所选取的部分个体，称为样本。
- **估计量**：根据样本个体的变量值计算出的对总体特征值的估计。

基本概念

按照抽样的过程与方法，可将抽样分为概率抽样与非概率抽样。

- **概率抽样**：按照一定概率从构成总体的所有个体中随机选取一部分个体作为样本的抽样方法，包括简单随机抽样、分层抽样、整群抽样、系统抽样等。
- **非概率抽样**：指样本不是按照一定的概率抽出，而是由研究者主观抽出或调查对象资源进入样本的抽样方法，包括便利抽样、判断抽样、滚雪球抽样、配额抽样等。

2

简单随机抽样

第十二届市桥杯大赛公益培训课件
(Credamo 见数版权所有)

简单随机抽样

- **简单随机抽样**：从一个个体数为 N 的总体中等概率逐个抽取个体并且无放回，每次都在所有尚未进入样本的个体中等概率的抽取，直到抽完样本所需的 n 个个体。
- 因为总体中的每个个体都有均等的机会被抽入样本，因此避免了任何会带来系统性或偏向性影响的认为干预。
- 但如果总体 N 很大时（例如中国全部大学生），则很难获取包含全部个体信息的完整抽样框，这时候简单随机抽样就遇到了问题。

3

分层随机抽样

第十二届市桥杯大赛公益培训课件
(Credamo 见券版权所有)

分层随机抽样

- **分层抽样**：先按照某种规则把总体划分为不同的层，然后在层内再进行抽样，各层的抽样之间是独立进行的。
- 分层抽样的估计是先在各层内进行，然后各层的估计量进行**加权平均或求和**，从而得出总体的估计量。
- 如果各层内进行的是简单随机抽样，则该过程称为**分层随机抽样**。
- 例如，在全国性调查中，通常按照各省市的经济发达程度分层；在住户调查中，按照户主的职业类型分层；在企业调查中，按照企业的行业或规模分层等。

分层随机抽样

各层样本量的分配

- **比例分配**：在分层抽样中，将每层的样本量按照该层在总体中所占的比例进行分配，这种分配方式称为比例分配。
- **最优分配**：在分层随机抽样中，若考虑费用问题，应当尽可能使估计量的方差达到最小，或者对于给定的估计量的方差，使得总费用达到最小。
- **指标分配**：根据调研目的选取关键指标（如性别、收入水平等），然后根据所掌握的经验资料，对每层的样本量进行划分

4

整群抽样与系统抽样

第十二届市桥杯大赛公益培训课件
(Credamo 版权所有)

整群抽样与系统抽样

- **整群抽样**：将总体划分为K个子样本，然后随机抽取其中的一个或几个子样本，作为最后的调查样本
- **群的划分方法**：一类是按照行政或地域进行划分，如班级、企业、乡镇等；一类是由调研人员认为确定，以保证群与群之间的差异小，而群内差异则越大越好，以保证每个群都具有足够好的代表性。
- 整群抽样的方法相对简单高效、方便经济，在一定情况下具有较高的精度。
- 但如果群与群之间的差异较大，则会带来较大的抽样误差，对整体代表性也会较差。

整群抽样与系统抽样

- **系统抽样**：先将总体中的个体按照某种次序排列，再规定范围内随机抽取一个初始个体，然后按照事先规定的格则抽取其他样本个体。
- 例如，在抽取初始个体后，按照相等的间距抽取其余样本个体。
- 其特点是实施简单，只需要随机抽取初始个体，后面按照规则和顺序将其他个体纳入样本即可，缺点是该抽样方法受到个体排序方式的影响，可能会带来偏差。

5

非概率抽样

第十二届市桥大赛公益培训课件
(Credamo 见数版权所有)

非概率抽样

- 在很多情况下，严格的概率抽样几乎无法进行，例如无法获取总体中所有个体的信息，或者总体边界不清等，在这种情况下都无法进行概率抽样。
- 非概率抽样不是按照随机的原则抽取样本，而是按照研究人员的主观经验或其他条件来抽取样本的抽样方式。
- 非概率抽样一般具有操作方便、省时省力的特点，通常实施起来比概率抽样要简单，而且也可以对调查总体有较好的了解，因此被广泛使用，具体包括便利抽样、判断抽样、滚雪球抽样等。

非概率抽样

- **便利抽样**：又称就近抽样、偶遇抽样等，是指调查人员根据自己的方便，自行确定进入样本的个体。
- 研究人员通常抽取哪些离得最近，或是最容易找到的人作为样本进行调查，例如街头拦访、校园调查，教师以自己所教班级的学生作为样本调查等。
- 其优点是简单易行、省时省力，能为非正式的探索性研究提供良好的数据源，缺点是样本获取的偶然性大，很难保证样本对于总体的代表性。

非概率抽样

- **判断抽样**：又称目的抽样、主观抽样、专家抽样等，是一种按照研究人员的主观判断、经验和知识，从总体中选择具有代表性的个体构成样本的抽样过程。
- 有时调查人员为了探讨某一现象产生的原因，会选择“极端”的个体作为样本进行调查，因为该类个体可能会表现出更为突出或明显的特征，例如重点调查与典型调查都属于判断调查的应用。
- 由于判断抽样方法中样本的选取具有主观性，因此估计精度严重依赖于研究人员的自身素质。

非概率抽样

- **滚雪球抽样**：又称链式抽样、辐射抽样等，即通过一些“种子”样本个体去获取更多样本的信息，以不断扩大样本量的抽样过程。
- 该抽样方法适用于对象不容易获取的情况，例如某些疾病的患者（HIV），某类职业的从业者（流动劳工、无家可归者）等。
- 滚雪球抽样的特点是可以有针对性的选取样本，可以高效的获取调查对象的信息，但其对初始样本点（种子）的依赖程度较高，因此需要慎重选择。

应用举例

第十二届市调大赛公益培训课件
(Credamo 版权所有)

应用举例

• 通过多次发布进行样本纠偏

案例：本次调研所收集的样本，计划男生样本占比50%，女生样本占比50%

Step1：通过平台随机发放180份问卷；

Step2：查看样本被试性别组成，发现男生100名，女生80名

Step3：再次发放20份问卷，通过设定被试性别特征为女性，对样本比例进行纠偏

样本特征设置 ? (1)最多选择两项, 限 [高级用户](#) (2)更多特征要求, 请 [联系我们](#)

人口统计	<input checked="" type="checkbox"/> 性别	<input type="checkbox"/> 年龄段	<input type="checkbox"/> 当前所在城市	
	<input type="checkbox"/> 行业	<input type="checkbox"/> 职业	<input type="checkbox"/> 户口	
家庭情况	<input type="checkbox"/> 收入	<input type="checkbox"/> 婚姻状况	<input type="checkbox"/> 子女状况	<input type="checkbox"/> 家用车状况
收起 ▲				
教育情况	<input type="checkbox"/> 最高学历	<input type="checkbox"/> 常用外语		
设备信息	<input type="checkbox"/> 手机品牌			
个人信息	<input type="checkbox"/> 常用投资理财方式			

谢谢大家

关注右边公众号



及时了解大赛资讯和进程

随时学习大赛公益培训

中国商业统计学会

官方网址：<http://www.china-cssc.org>

公众号：



Credamo见数

官方网址：www.credamo.com

公众号：

