

定量研究方法（二）

实验法

林景 讲师

南京财经大学法学院

课程主页

- law2019.netlify.com 参考书解压密码：2019



Social Science Methodology

Contents

2019-2020学年社会科学研究方法课程大纲

October 2019 · 1 minute read

授课教师：林景 讲师

南京财经大学法学院

[linjing\(AT\)nufe.edu.cn](mailto:linjing(AT)nufe.edu.cn)

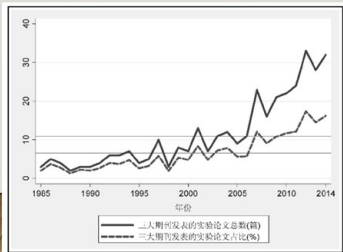
一、课程概要

(一) 课程宗旨与预期目标

这是一门为人文社会科学专业学生开设的研究方法课程，旨在教授学生掌握一套完整的社会科学研究方法技巧，内容包括何为社会科学研究、研究方法的类型与适用范围、如何进行研究设计、如何采集和分析数据以及如何基于既有数据进行科学推论等。

为何需要实验法？

- 问卷调查法的局限
 - 遗漏变量偏误 (omitted variable bias)
 - 样本选择偏误 (selection bias)
- 实验法的应对：用随机化替代控制变量



思考题：南财学生的补习与学业成绩

- 研究假设
 - 理论假设是什么？
 - 操作假设是什么？
- 操作化与测量
 - 补习开支与学业成绩怎么操作化？用什么来测量？
- 抽样与采集数据
 - 总体是什么？
 - 样本是什么？
 - 怎么抽样？
 - 如何设计问卷？

思考题：南财学生的补习与学业成绩

- 如果用实验法进行研究，有何异同？

思考题：南财学生的补习与学业成绩

- 如果用实验法进行研究，有何异同？
- 相同：
 - 都必须建立变量之间因果关系的假设
 - 都是从自变量出发，研究其对因变量的影响
- 差异：
 - 自变量能否被独立地操纵和控制
 - 用于比较的组别，是否是随机分配产生

常用随机化方法

- 按抛硬币或掷色子进行分配的方法
- 按研究对象序列的单双号进行分配的方法
- 按研究对象排列顺序的先后次序进行分配的方法
- 按随机数表或随机数生成器进行分配的方法
 - 美国著名智库兰德公司（RAND Corporation）整理了上百万的随机数并出版为著作*A Million Random Digits with 100,000 Normal Deviates*。



随机化何以管用？

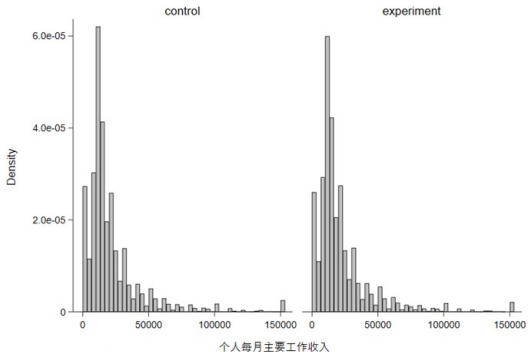
- 女士品茶：茶加奶 V.S. 奶加茶？
 - 统计学家Ronald Fisher与植物学家Muriel Bristol的赌局。
 - 从8杯奶茶，其中4杯为奶加茶，另外4杯为茶加奶。Bristol女士只要品出其中的4杯即胜出。
 - Bristol女士盲品猜中的概率为 $\frac{1}{\binom{8}{4}} = \frac{1}{70} \approx 0.014$ 。
 - Fisher如何最大限度提高自己的胜算？



随机化的效果：基于香港人口数据的模拟

- 2016年香港人口中期统计数据（1%样本）
- 筛选出16-55岁之间拥有工作的本地劳动人口
 - 样本规模：24561名
 - 随机分成两组，实验组为17281名，控制组为17281名
 - 如果随机化分配（randomize assignment）得以实现，实验组与控制组劳动人口的年龄、性别比例、教育程度、住房面积、工作收入水平、工作时间甚至家庭距离工作地点的距离在统计意义上无明显差异。

随机化的效果：基于香港人口数据的模拟



Graphs by RECODE of t

随机化的效果：基于香港人口数据的模拟

	样本量	均值	标准误	标准差	95%置信 区间下限	95%置信 区间上限
控制组	17281	21749.58	175.24	23037.59	21406.08	22093.08
实验组	17281	21898.66	170.98	22476.57	21563.52	22233.8
全体	34562	21824.12	122.41	22758.6	21584.18	22064.06
控制组-实验组		-149.08	244.83		-628.97	330.80

两组均值t检验：

H0：两组无差异；自由度=34560；t值=-0.6089

Ha：两组若有差异，p值=0.5426

实验法的类型：实验室实验与自然实验

- 实验室实验：一般在有专门设备的实验室中进行，并对实验的条件、控制以及实验设计都有严格的规定。
 - 日照强度与公鸡打鸣
 - 斯坦福监狱实验与“路法西效应”
 - 仿制药的药效
- 自然实验：又称实地实验，一般在实际场所中进行，**借助已存在的自然情境**实现对实验条件、实验控制和实验设计的执行。
 - 室外禁烟是否影响服务业经济
 - 房产税是否影响房价

实验法的基本概念：干预、前测与后测

- 干预 (treatment)：在实验过程中施加给实验对象的一系列刺激因素，并预期这类因素将引起实验对象的变化。
- 前测 (pretest)：在实验之前对所有实验对象进行测量。
- 后测 (posttest)：在实验之后对所有实验对象进行测量。
 - 后测 - 前测 = 实验效果？



实验法的基本概念：实验组、控制组、双盲法与安慰剂

- 实验组 (experimental group)：指在实验过程中，实际接受干预或刺激的一组实验对象；
- 控制组 (control group)：又称为对照组，指在实验过程中不接受干预或刺激的一组实验对象。
 - 设置控制组的目的，是为了将研究本身对实验对象的影响，与实验干预或刺激对实验对象的影响区别开来。
- 双盲法 (double blinding)：在实验过程中，实验对象和实验人员均不知道谁属于实验组，谁属于控制组；换言之，实验对象和实验人员均只知道谁被施以真正的干预或刺激，谁被施以安慰剂。
 - 安慰剂 (placebo)：指没有实际效用但与真正实验相似的干预或刺激。

实验法的关键概念：随机化

- 随机化 (randomize)：以随机化分配 (randomize assignment) 的方式将实验对象分到不同的组别中，其中一些组别将成为实验组，另一些将成为控制组。
 - 在自然实验中，研究者通常借助已存在的自然情境引入随机分配。
- 随机化的效果：
 - 实验前，实验对象的条件和特征是相等的，如年龄、性别比例和教育程度等；
 - 实验后，实验对象由于自身因素所导致的条件和特征之变化是相等的，如体貌、情绪和经验等。

实验设计的类型

- 古典实验设计

实验前	实验过程	实验后
对实验组前测	实施干预或刺激	对实验组后测
对控制组前测	不实施干预或刺激	对控制组后测

- 实验效果=实验组后测-控制组后测

- 所罗门四组设计

实验前	实验过程	实验后
对实验组1前测	实施干预或刺激	对实验组1后测
对控制组1前测	不实施干预或刺激	对控制组1后测
不对实验组2前测	实施干预或刺激	对实验组2后测
不对控制组2前测	不实施干预或刺激	对控制组2后测

实验法的实施步骤

- 准备阶段
 - 确定研究问题和研究目的
 - 提出理论假设和操作假设
 - 进行实验设计
- 实施阶段
 - 选取实验对象：随机法或指派法
 - 实施实验
- 数据处理阶段
 - 整理数据和撰写研究报告

实验法案例：入伍参军经历对人们自杀倾向的影响

- 研究背景
 - 美国国立卫生研究院（National Institutes of Health）在一项全国性调查中发现，拥有入伍参军经历的人群呈现较高的自杀率。相关研究者认为，入伍参军经历可能给某些成人造成严重的创伤后压力症候群（PTSD），使其长期处于精神上或身体上的不适和紧张，进而做出剧烈自我伤害的应激反应。
- 如何排除入伍参军以外的因素对人们自杀倾向的影响？



《全金属外壳》



《猎鹿人》

实验法案例：入伍参军经历对人们自杀倾向的影响

- *The New England Journal of Medicine*, 1986
- 作者：诺曼·赫斯特 (Norman Hearst)，加州大学旧金山分校
- 数据：加州和宾夕法尼亚州1970~1972年的全体居民数据
 - 两州在1970~1972年的征兵政策规定，适龄成年居民如果被征召，必须无条件接受指派按时入伍
 - 征兵中心采取随机抽签的方式，先从1~365的数字中随机抽取若干个数字，代表一年内从第1天到最后1天的日期，然后将出身日期符合该数字的居民悉数征召入伍。
- 发现：拥有入伍参军经历的成年人相比其他成年人，自杀率高出13%。



实验法案例：入伍参军经历对人们职业生涯的影响

- *American Economic Review*, 1990
- 作者：约书亚·安格里斯特 (Joshua Angrist) , MIT
- 数据：加州和宾夕法尼亚州1970~1972年的全体居民数据
- 发现：拥有入伍参军经历的成年人相比其他成年人，职业收入低15%。



实验法案例：入伍参军经历对人们职业生涯的影响

- *MIT Technology Review*, 2013
- 1986年在普林斯顿大学，计量经济学家奥利·阿申费尔特（Orley Ashenfelter）给研究生上方法课时提到，旧金山加州大学的研究人员正在设计一个自然实验，评估在入伍从军对成人自杀率的影响。
 - “这是个极好的研究设计——早就应该有人把这个方法推广到退役士兵的收入研究上。就在当天下午，我立即投入了工作。”——Joshua Angrist
 - “很少有人能在这样的研究中坚持下去，包括我自己。这就像在边境下挖隧道什么的，尽是没完没了的、乏味的工作”。——Orley Ashenfelter



实验法案例：性别偏向型职业对地区性别失衡的影响

- 研究背景

- 消失的女孩（missing girls）：指亚洲发展中国家的女性人口比例相比西方发达国家的赤字。如印度和中国女性在总人口中的比率常年维持在48%左右，而欧洲女性在总人口中的比率长期维持在50%左右。
- 在亚洲贫困地区，大量家庭采取选择性堕胎、遗弃甚至溺婴的手段来进行生育的性别选择。

- 研究问题

- 生育的性别选择是文化偏好，抑或是经济理性？

实验法案例：性别偏向型职业对地区性别失衡的影响

- *The Quarterly Journal of Economics*, 2012
- 作者：钱楠筠，布朗大学
- 数据：中国农业普查1%抽样数据+人口普查1%抽样数据
 - 如果女性的劳动价值得到提高，是否会改善该地区的性别失衡？
 - 在中国9000个茶园中，95%的劳动者是女性。而茶园的茶叶品质变化主要受地区降雨量影响，是完全随机事件。茶叶品质会进一步影响茶叶价格。
- 发现：成年女性平均每提高7.7美元，将使得该地区女孩的存活率提高1%。



实验法案例：单性学校对学生学业成绩的影响

- *Economics of Education Review*, 2018
- 作者：朴贤俊，宾夕法尼亚大学
- 数据：2004-2011年韩国教育与劳动雇佣追踪调查（KEEP）
- 发现：男女同校会同时提高男学生和女学生的STEM成绩，男子学校也可以提高男学生的STEM成绩，但女子学校无助于提高女学生的STEM成绩。
- 拓展研究：私立学校/公立学校对学生学业成绩的影响；上学距离对学生学业成绩的影响；学校规模对学生学业成绩的影响。



实验法案例：小额信贷的奇迹

- Prize motivation: "for their **experimental approach** to alleviating global poverty."
- *American Economic Journal: Applied Economics*
- 作者：阿比吉特·班纳吉（Abhijit Banerjee）与艾斯特·杜芙若（Esther Duflo），MIT
- 数据：印度南部城市海德拉巴（Hyderabad）104个贫民窟的信贷、工作与消费
- 发现：接受Spandana公司小额信贷资助的家庭明显增加耐用品消费，减少非耐用品消费；增加对经营性活动的投资，创业可能性更高；劳动时间更长。



实验法是否会测不准？

- 内部效度问题
 - 实验对象未能完全随机分组
 - 实验对象不完全遵从实验设计
 - 实验对象中途退出实验
 - 实验对象的成长或成熟
 - 实验对象受到外在刺激的扰动
 - 经历重大社会变故
 - 霍桑效应

实验法是否会测不准？

- 外部效度问题
 - 实验对象的样本代表性不足
 - 实验条件与现实条件不同
 - 实验中无法模拟现实社会的系统聚现（emergence）效应
 - 实验忽视现实情境的自我选择效应