定量研究方法  
**Quantitative Research Methods**

**授课教师**彭泽宇 青年副研究员（[zpeng@fudan.edu.cn](mailto:zpeng@fudan.edu.cn)）

**助教**蒋屹阳、周督竣、肖伟林、张睿明、林佳怡

**目录**

[第一讲 绪论：定量研究方法在社科研究中的应用 2](#_Toc178352409)

[一、观察性研究 2](#_Toc178352410)

[（一）观察性研究的案例 2](#_Toc178352411)

[（二）观察性研究的局限性 3](#_Toc178352412)

[二、随机对照试验 3](#_Toc178352413)

[（一）随机对照试验的案例 3](#_Toc178352414)

[（二）随机对照试验的局限性 3](#_Toc178352415)

[第二讲 R语言 3](#_Toc178352416)

[第三讲 概率与条件概率 4](#_Toc178352417)

[一、概率的基本概念 4](#_Toc178352418)

[（一）概率的定义 4](#_Toc178352419)

[（二）概率模型的基本概念 4](#_Toc178352420)

[（三）概率公理 4](#_Toc178352421)

第一讲 绪论：定量研究方法在社科研究中的应用

2024.9.6

社会科学量化研究通常分为观察性研究（observational studies）和随机对照试验（randomized controlled trials, RCTs）。

* **观察性研究：**研究人员无法在现实世界中随机分配干预措施，只能观察和收集自然发生的事件中的数据；
* **随机对照试验：**研究人员随机分配干预给研究对象（观察值，observations）。

一、观察性研究

（一）观察性研究的案例

1. 案例背景：无差别炮击是否会引起更强烈的反抗

无差别炮击、轰炸是军事冲突双方常见的策略，同时也是国际安全、军事理论研究中常见的话题。由此催生了一个社会科学问题：在内战中政府军无差别炮击是否会引起民众更强烈的反抗？

这个问题的潜在答案有两个：会引起强烈反抗；或，不会引起强烈反抗，并有效遏制镇压了反抗势力。

我们首先运用推理，提出导向上述答案的可能原因：

* **会引起强烈反抗**
  + 加剧对政府军的仇恨
  + 向叛军/抵抗组织寻求保护
* **有效遏制镇压反抗势力**
  + 减少叛军可获得的资源
  + 威慑恐吓当地民众
  + 打击叛军的后勤网络

我们可以运用定量研究方法验证上述假说。当然，对于这种社会科学问题（尤其是与国际经济军事相关的问题），观察性研究的数据采集难度是很大的。

2. 案例研究：自变量、因变量与干扰因素

杰森·力欧（Jason Lyall, 2009）收集了2000年至2005年俄罗斯联邦军队对车臣村庄实施炮击的数据。数据中存在自变量和因变量之分：

* **自变量（解释变量、预测变量）：**相对稳定且能够决定其他变量的数值的变量；
* **因变量（被解释变量、响应变量）：**数值取决于自变量或其他变量的变量。

我们需要以实证方式验证自变量与因变量之间的关系。在该案例中，自变量是该村是否受到过俄军的无差别炮击；因变量是炮击前后俄军和亲政府武装在该村遭受车臣反政府武装袭击的次数。

当然，存在影响自变量的干扰因素，如人口、贫困率、地形高度（海拔）、是否有俄军进驻、控制该村的车臣领导人派系等。干扰因素又称控制变量。

（二）观察性研究的局限性

观察性研究可能存在选择性偏误（selection bias），其内在效度（internal validity）可能会受到影响，难以证实因果关系。例如，在上面的例子中，研究者可能没有观测到一个同时影响了俄军炮击某一村庄的决定以及当地袭击俄军及亲俄武装的频率的干扰因素。

因此，我们可以在随机对照试验当中研究者可以将干预变量进行随机分配，从而解决选择性偏误问题。

二、随机对照试验

（一）随机对照试验的案例

1. 案例背景：劳动力市场是否存在种族歧视

种族歧视是以美国为代表的西方国家中普遍存在的社会问题。西方国家的劳动力市场是否真的存在种族歧视是社会科学界常年争论的学术话题——例如，不同种族之间的失业率差异是否是由于其他因素（比如教育程度差异）造成的？

在实验中，研究者很难操纵或改变一个人的种族。因此，一个重要的问题就是：如何设计一个随机试验来检视雇主是否会区别对待不同种族的求职者？

2. 案例研究：随机对照试验的设计

经济学家玛丽安娜·伯特兰德（Mariannne Bertrand）和赛德希尔·穆来纳森（Sendhil Muillainathan）进行了以下试验来考察美国劳动力市场的种族歧视问题：研究人员向在报纸上登广告的潜在雇主发去虚构的求职简历；只改变求职者的名字，而将其他信息保持不变。一些求职者用了非常典型的非洲裔美国人的名字（如Lakisha），而其他求职者用了典型的白人名字（如Emily）；比较这两个群体之间的电话回复率。

（二）随机对照试验的局限性

随机对照试验的内在效度普遍高于观察性研究，通常被认为是建立许多科学学科的因果关系的黄金标准。然而，出于道德和操作的种种因素，在很多情况下研究者无法在现实世界随机分配干预措施。例如，在上文提及的“无差别炮击”案例中，社会科学家就不可能将每次炮击的目标随机分配；在本案例中，寄出大量虚构的求职简历也可能会扰乱用人单位人事部门的正常运转。

随机对照试验可能缺乏“外在效度（external validity）”。如在本案例中，研究对象是否包含一个国家/社会中所有所有族群？在美国，简历上一般不贴照片；那么这个实验是否符合其他国家的社会文化背景？

第二讲 R语言

2024.9.13 / 2024.9.20 / 2024.9.27

见“R语言语法速查”。

第三讲 概率与条件概率

2024.9.27

一、概率的基本概念

概率论是解决社会科学研究中不确定性的重要工具。社会科学研究一般由两步构成：

* 构建理论：在所学的基础上建立一个概率模型；
* 运用数据和其他经验性证据证明概率模型从而支持你的理论。

（一）概率的定义

**概率**是一系列在现实世界中用来测量随机性并对随机性建模的数学工具。

频率统计学派（frequentist）认为，概率是对应频率的极限，是事件发生次数与同样条件下重复试验次数之比，即当在相同条件下反复进行的实验次数接近无穷次是某事件发生频率的极限。

贝叶斯统计学派（Bayesian）认为，概率是用来测量一个人认为某件事发生可能的主观信念。

（二）概率模型的基本概念

**试验（experiment）**是一个或一组产生与某一问题相关的随机事件的行动。

**样本空间（sample space）**是试验所有可能的一组结果，通常用Ω表示。

**事件（event）**是样本空间的一个子集。

（三）概率公理

第一，任一事件A的概率非负，即：

第二，样本空间中所有结果中任一发生的概率是1，即：

第三，概率遵循加法法则，即：

对于任意一给定的事件A和B：

如果事件A和B是互斥的：