



THE WORLD BANK



# 技术路线 讲座 2: 随机试验

中国，北京  
2009

# 随机试验

---

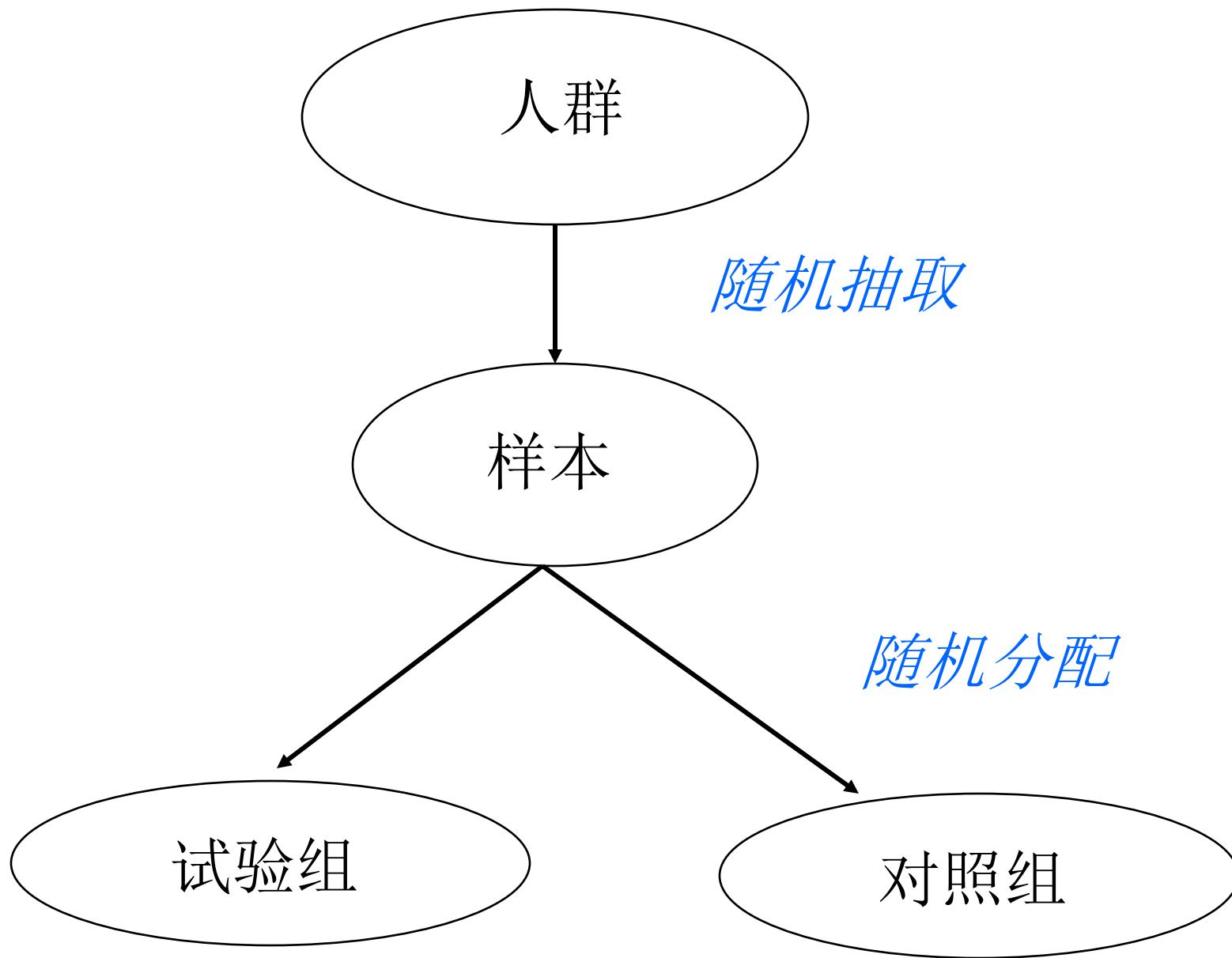
- 研究者们如何在现实世界的实际操作中了解虚拟条件？
- 在很多领域，尤其是医学研究领域，虚拟条件的证据是由随机试验得出的。
- 在一定条件下，随机试验可以确保对照组的试验结果即是试验组真实的虚拟条件。

# 因果推断的随机化

---

统计学家们建议使用规范的两步随机模型：

- **第一步：**从确定的人群(总体)中随机抽取一个样本。
- **第二步：**把这个样本随机分配到试验组和对照组中。



# 为什么两步都是随机的？

---

## □ 第一步:为了确保外部有效性

- 即确保在一定的抽样误差水平内，样本的结果能够代表总体的结果。

## □ 第二步:为了确保内部有效性

- 即确保因变量的观测结果是干预手段产生的，而不是其他混杂因素造成的。

# 两步随机试验

---

在大样本中，两阶段随机试验能够确保：

$$[\bar{Y}_1 | D=1] = [\bar{Y}_1 | D=0] \text{ 并 } [\bar{Y}_0 | D=1] = [\bar{Y}_0 | D=0]$$

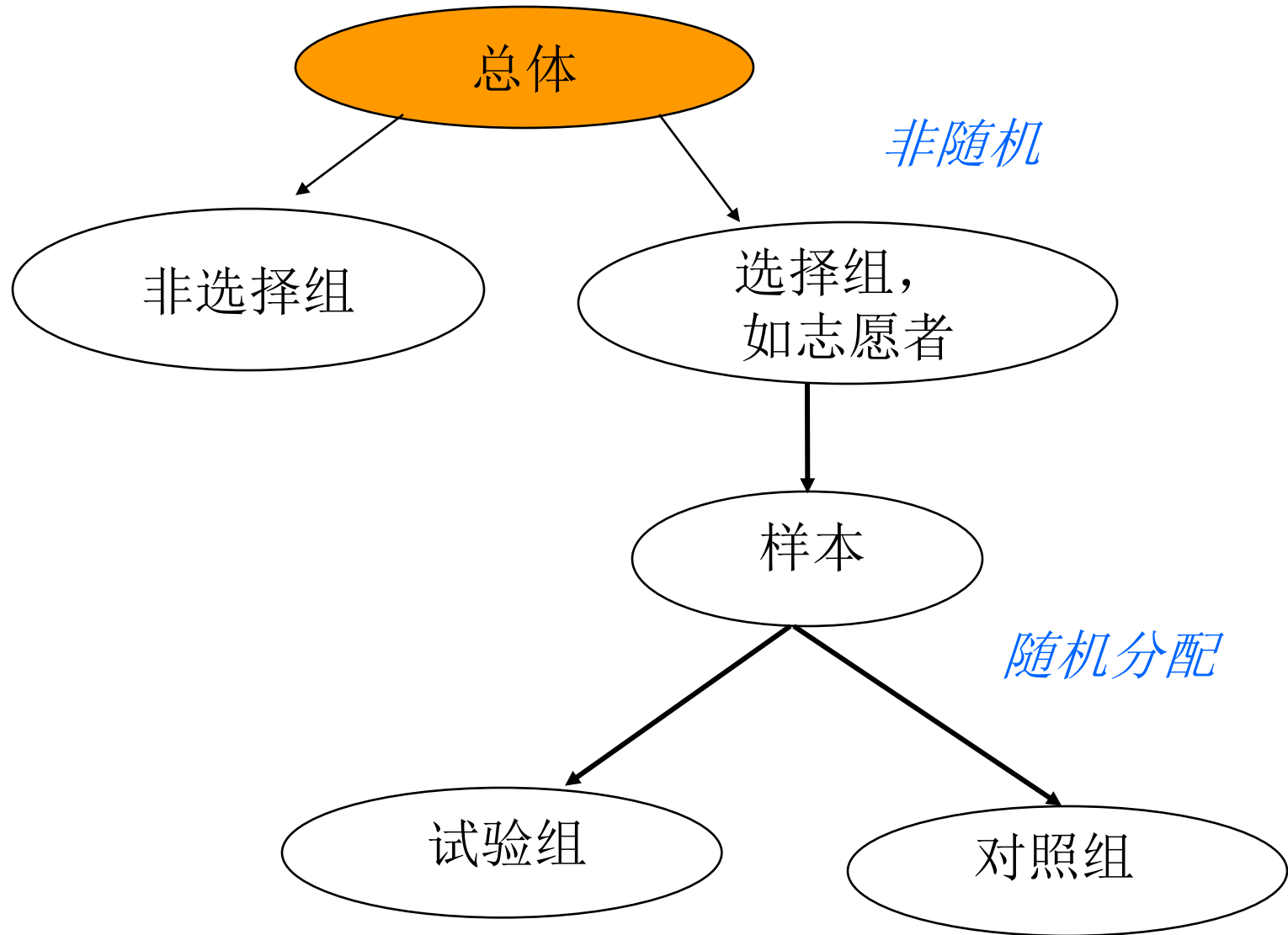
为什么此式成立...

因此，估计值

$$\hat{\delta} = [\hat{Y}_1 | D=1] - [\hat{Y}_0 | D=0]$$

可以一致地 估计平均试验效果 (*ATE*) 。

# 总体 vs. 选择组？



# 总体 vs. 选择组？

---

- 如果随机化只针对选择组
- 我们的估计将会是？ ？ ？ ？
  
- 试验效果是针对选择组的！



# 随机试验：一些需要注意的地方

---

- 非依从
  - 并非试验组中所有的个体都会接受干预（非依从）
  - 对照组中的一些个体可能会接受干预（不依从）
- 损耗：我们也不可能观察到所有个体发生的情况
- 霍索恩效应：观测本身也会使个体的行为发生改变
- 约翰·亨利效应：对照组个体因知道自己在对照组中而发生行为改变,努力表现去抵消差距

# 随机试验 vs. 非随机试验

---

## □ 随机试验

- 假设发挥的作用很小
- 或者当检验“无效”假设时，假设根本没有作用
- 不存在不依从、丢失等困难的情况下.....

## □ 非随机试验

- 为了验证虚拟状态，需要有强假设
- 同随机试验一样，损耗也是一个问题

# 例1：哥伦比亚关于“电脑对教育的影响”的随机试验

---

- 项目活动：
  - 对私人公司捐赠的电脑进行整修，并在公立学校安装这些电脑。
  - 在当地大学的帮助下，培训公立学校的教师学会在教学中使用电脑。
- 2006年：随机选择了100所学校
  - 其中的50所配置了电脑
  - 其余50所没有配置电脑

附表B: 哥伦比亚基线数据, 一些选定的统计指标

	平均数		差异	平均数	差异
	城市	乡村	城市/乡村	样本	试验组/对照组
<b>学校信息</b>					
教师人数	16 (12.25)	7.55 (5.86)	8.45*** (1.84)	9.64 (8.67)	0.26 (1.34)
学生/教师数之比	14.1 (13.53)	13.53 (11.26)	0.57 (2.87)	13.67 (11.79)	-0.73 (1.77)
10年以上教龄的教师数	0.88 (0.13)	0.64 (0.25)	0.24*** (0.02)	0.7 (0.25)	0.03 (0.03)
学生总数	200.08 (217.21)	105.48 (140.87)	94.60*** (24.21)	128.89 (166.94)	-26.17 (21.17)
留级人数	14.08 (24.88)	4.96 (6.22)	9.12*** (3.04)	7.22 (13.89)	5.14*** (0.59)
辍学人数	17.21 (24.37)	12.3 (15.15)	4.91** (2.42)	13.52 (17.86)	-1.23 (1.47)
教室数	10.83 (7.04)	6.48 (3.54)	4.35*** (1.10)	7.56 (4.99)	0.38 (0.90)
图书馆数	0.54 (0.51)	0.6 (0.55)	-0.06 (0.06)	0.59 (0.54)	-0.01 (0.06)
学校的主要部门	0.83 (0.38)	0.86 (0.35)	-0.03 (0.06)	0.86 (0.35)	-0.17** (0.08)
有用性	1.08 (0.28)	1.04 (0.26)	0.04 (0.05)	1.05 (0.27)	-0.02 (0.03)

## 个体信息

性别	0.55 (0.0067)	0.52 (0.0056)	0.03 (0.05)	0.54 (0.0043)	-0.04 (0.04)
年龄	11 (0.0330)	12.54 (0.0311)	-1.54*** (0.54)	11.91 (0.0237)	0.17 (0.51)
兄弟姐妹数量	3.2 (0.0337)	4.16 (0.0322)	-0.96*** (0.25)	3.77 (0.0239)	-0.12 (0.28)
工作	0.16 (0.0050)	0.23 (0.0047)	-0.07*** (0.02)	0.2 (0.0035)	-0.01 (0.02)
上一年度上学	0.97 (0.1757)	0.98 (0.1544)	-0.0075 (0.0049)	0.97 (0.1635)	-0.0013 (0.0048)
上一年度留级	0.29 (0.4542)	0.38 (0.4861)	-0.0920*** (0.0251)	0.35 (0.4755)	0.0011 (0.0276)
上周没有上学	0.21 (0.4104)	0.24 (0.4250)	-0.0222 (0.0231)	0.23 (0.4193)	-0.0097 (0.0264)
天数	2.15 (2.9484)	1.9 (1.6404)	0.2571* (0.1500)	2 (2.2373)	0.1003 (0.1523)
喜欢学校	0.98 (0.1313)	0.98 (0.1427)	0.0033 (0.0038)	0.98 (0.1382)	-0.0022 (0.0039)
知道因特网	0.48 (0.4996)	0.35 (0.4758)	0.1314** (0.0517)	0.4 (0.4900)	-0.0009 (0.0592)
使用因特网（在知道的人中）	0.8 (0.3993)	0.65 (0.4756)	0.1464*** (0.0422)	0.73 (0.4459)	-0.0077 (0.0514)
校外学习的小时数	1.47 (1.0064)	1.31 (0.9610)	0.1565*** (0.0559)	1.38 (0.9826)	0.0812 (0.0542)
考试成绩：语言类	0.45 (0.2667)	0.4 (0.2579)	0.0520** (0.0207)	0.42 (0.2627)	0.0072 (0.0229)
考试成绩：数学类	0.33 (0.2574)	0.31 (0.2437)	0.0239 (0.0235)	0.31 (0.2495)	-0.0077 (0.0231)



## 例2: 缩小班级规模能否提高小学教育质量?

---

- STAR计划（学生-教师成绩比）
- 为评估班级规模大小对学习的影响，进行了一项自幼儿班至小学3年級的为期4年的试验。
- 试验标准：
  1. 正常的班级规模：每班22-25名学生，1名教师。
  2. 较小的班级规模：每班13-17名学生，1名教师。
  3. 助教班：班级规模正常，但增加1名助教。

## 例2: 缩小班级规模能否提高小学教育质量?

---

- 每所学校至少各有一个不同类型的班级。
- 在参与试验的学校中，进入幼儿班的学生被随机分配到这三组班级中的一个。
- 随机配备教师。

## 试验措施对标准化考试成绩的影响效果估计

$$Y_i = \beta_0 + \beta_1 I_i^{\text{小班}} + \beta_2 I_i^{\text{助教班}} + u_i$$

	年级			
回归量	幼儿班	一年级	二年级	三年级
截 距	918.0*** (1.6)	1,039.4*** (1.8)	1,157.8*** (1.8)	1,228.5*** (1.7)
小 班	13.9*** (2.5)	29.8*** (2.8)	19.4** (2.7)	15.6*** (2.4)
助教班	0.3 (2.3)	12.0*** (2.7)	3.5 (2.5)	-0.3 (2.3)
观测样本数	5,786	6,379	6,049	5,967



## 例2: 缩小班级规模能否提高小学教育质量?

---

### □ 发现:

- 缩小班级规模对考试成绩有影响,
- 给正常规模的班级增加助教对考试成绩的影响极小或没有影响。

### □ 需要注意之处:

- 这些估计同时忽略了损耗和不依从的影响。
- 上述两项有害因素的影响较大 => 试验结果也许存在偏差。

# 非依从及损耗：解决办法？

---

## □ 非依从：

1. 参与试验的意愿分析
2. 工具变量分析（当地的平均试验效果）

## □ 损耗（隐蔽偏差）

1. 确保试验组和对照组的损耗没有差异
2. 使用工具变量以及匹配法

---

## 例3:

哥伦比亚的私立学校教育券：来自一项  
随机自然试验的证据

**Angrist等人（2002年）**  
**AER**

# 中学教育覆盖面计划（PACES）

---

## □ 项目有利于

- 支持就读私立中学的学校教育券计划
- 教育券支付一半以上的私立中学费用
- 如果学生们保持令人满意的学习成绩，教育券可以继续使用

## □ 受益人

- 125,000 名接受试验的学生
- 受益人是通过抽签法从大批符合条件的申请者中选出的

# PACES评价设计

---

- 1998年，面试了1,600名PACES计划的申请者
  - 用抽签法选取受益人，中选者和落选者的数量相当
  - 1995年和1997年的申请者来自波哥大
  - 1993年的申请者来自贾穆迪
  
- 面试方法：电话交流
  - 应答率： +/- 60%
  - 应答情况并不受中选情况的影响（即中选和落选的机会是一样的）

## 波哥大的个体特征及教育券情况，1995年

因变量	落选者平均值	获得教育券 估计影响
调查时年龄	15.0 (1.4)	-0.013 (0.078)
男性	0.501	0.004 (0.029)
母亲受教育的最高年限	5.9 (2.7)	-0.079 (0.166)
母亲年龄	40.7 (7.3)	-0.027 (0.426)
父亲年龄	44.4 (8.1)	0.567 (0.533)
父亲工资 (>2倍最低工资)	0.100	0.005 (0.021)
N	583	1,176

注：本表提供了未获教育券的落选者的平均值，以及获得教育券的估计影响。在“落选者平均值”一列，括号中的数字为标准差；在“获得教育券估计影响”一列，括号中的数字为标准误。

## 波哥大的教育产出以及教育券状况，1995年

因变量	落选者 平均值 (1)	无对照 (2)	基本 对照 (3)	完全 对照 (4)
读完6年级	0.943 (0.232)	0.026** (0.012)	0.023* (0.012)	0.021* (0.011)
读完8年级	0.632 (0.483)	0.112*** (0.027)	0.100*** (0.027)	0.094*** (0.027)
6年级的复读率	0.194 (0.454)	-0.066*** (0.024)	-0.059** (0.024)	-0.059** (0.024)
中选后的复读率	0.224 (0.417)	-0.060*** (0.023)	-0.055** (0.023)	-0.051** (0.023)
中选后的受教育年限	3.7 (0.951)	0.058 (0.052)	0.034 (0.050)	0.031 (0.050)
样本量	562		1,147	

注：本表提供了未获教育券的落选者的平均值，以及获得教育券的估计影响。

在“落选者平均值”一列，括号中的数字为标准差；在“获得教育券估计影响”一列，括号中的数字为标准误差。“\*\*\*”表示显著性水平为1%，“\*\*”表示显著性水平为5%，“\*”表示显著性水平为10%。

# 参考文献

---

- ❑ Rosenbaum, Paul (2002): Observational Studies, Springer. Chapter 2.
- ❑ Cochran, W. G. (1965): “The planning of observational studies of human populations”, *Journal of the Royal Statistics Association Series A* 128, pp. 134-155, with discussion.
- ❑ Angrist, J., E. Bettinger, E. Bloom, E. King and M. Kremer (2002): “Vouchers for Private Schooling in Colombia: Evidence from a Randomized Natural Experiment”, *American Economic Review*, 92, pp. 1535-58.
- ❑ Angrist, J. and V. Lavy (2002): “The Effect of High School Matriculation Awards: Evidence from Randomized Trials”, NBER Working Paper.