Drawing Statistical Conclusions

RANDOMIZED EXPERIMENTS V. OBSERVATIONAL STUDIES

RANDOM SAMPLES V. SELF-SELECTION

Two motivating examples
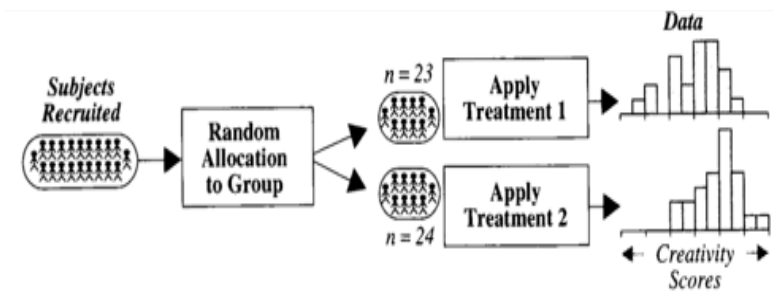
Creativity Scores:
Intrinsic vs. Extrinsic Motivation

Creativity scores in two motivation groups, and their summary statistics

Intrinsic group		Extrinsic group	
12.0	20.5	5.0	17.4
12.0	20.6	5.4	17.5
12.9	21.3	6.1	18.5
13.6	21.6	10.9	18.7
16.6	22.1	11.8	18.7
17.2	22.2	12.0	19.2
17.5	22.6	12.3	19.5
18.2	23.1	14.8	20.7
19.1	24.0	15.0	21.2
19.3	24.3	16.8	22.1
19.8	26.7	17.2	24.0
20.3	29.7	17.2	

	Intrinsic	Extrinsic
Sample Size:	24	23
Average:	19.88	15.74
Sample Standard Deviation:	4.44	5.25

Subjects Volunteered for the Study.

A 95% confidence interval for a difference in score due to having extrinsic motivation rather than intrinsic motivation is between 1.3 and 7.0 points.

Subjects Recruited → Random Allocation to Group → $n = 23$ Apply Treatment 1 → Data

→ $n = 24$ Apply Treatment 2 → Creativity Scores

Starting Salaries: Female vs. Male

Display 1.3 Starting salaries ($U.S.) for 32 male and 61 female clerical hires at a bank

| Males			Females					
4,620	5,700	6,000	3,900	4,500	4,800	5,220	5,400	5,640
5,040	6,000	6,000	4,020	4,620	4,800	5,220	5,400	5,700
5,100	6,000	6,000	4,290	4,800	4,980	5,280	5,400	5,700
5,100	6,000	6,300	4,380	4,800	5,100	5,280	5,400	5,700
5,220	6,000	6,600	4,380	4,800	5,100	5,280	5,400	5,700
5,400	6,000	6,600	4,380	4,800	5,100	5,400	5,400	5,700
5,400	6,000	6,600	4,380	4,800	5,100	5,400	5,400	6,000
5,400	6,000	6,840	4,380	4,800	5,100	5,400	5,520	6,000
5,400	6,000	6,900	4,440	4,800	5,100	5,400	5,520	6,120
5,400	6,000	6,900	4,500	4,800	5,160	5,400	5,580	6,300
	6,000	8,100						6,300

n = 32
males

n = 61
females

The mean starting salary for males is estimated to be $560 to $1080 larger than the mean starting salary for females (95% confidence). Such a large difference is unlikely to occur by chance (p-value < .00001 from a two sample t-test of means).
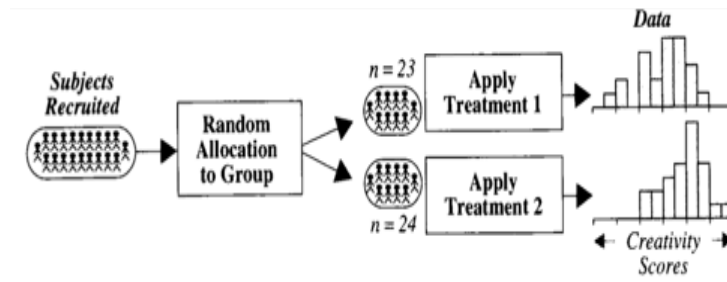
Random sampling study with two populations

Population 1 → Random Sampling → n_1 → Data

Population 2 → Random Sampling → n_2 → Data
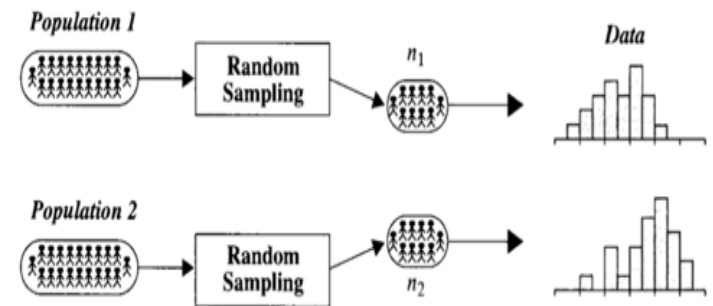
Cause, Effect, and Inference to Populations

Types of Studies

Creativity Study

Randomized Experiment

Salary Study

Observational Study

Causal Inference: Randomized vs. Observational Study

- Causal inferences **can** be drawn from randomized experiments

- Causal inferences **cannot** be drawn from observational studies due to **CONFOUNDING**
  (*Technical point: there are causal inference techniques for observational studies*)

CONFOUNDING VARIABLE: Related to both group membership and to the outcome

Example: Since 2000 the U.S. median wage…

- has overall increased about 1%

- Has decreased for high school (or below) dropouts and high school graduates

- Is this a paradox? No, more people are graduating high school

Causal Inference: Randomized vs. Observational Study

- Causal inferences **can** be drawn from randomized experiments

- Causal inferences **cannot** be drawn from observational studies due to **CONFOUNDING**

Maybe males have
- more education
- more seniority
- more age (older)

Display 1.3 Starting salaries ($U.S.) for 32 male and 61 female clerical hires at a bank

Older

4,620	5,700	6,000
5,040	6,000	6,000
5,100	6,000	6,000
5,100	6,000	6,300
5,220	6,000	6,600
5,400	6,000	6,600
5,400	6,000	6,600
5,400	6,000	6,840
5,400	6,000	6,900
5,400	6,000	6,900
5,400	6,000	
	6,000	8,100

Younger

3,900	4,500	4,800	5,220	5,400	5,640
4,020	4,620	4,800	5,220	5,400	5,700
4,290	4,800	4,980	5,280	5,400	5,700
4,380	4,800	5,100	5,280	5,400	5,700
4,380	4,800	5,100	5,280	5,400	5,700
4,380	4,800	5,100	5,400	5,400	5,700
4,380	4,800	5,100	5,400	5,400	6,000
4,380	4,800	5,100	5,400	5,520	6,000
4,440	4,800	5,100	5,400	5,520	6,120
4,500	4,800	5,160	5,400	5,580	6,300
					6,300

Intrinsic group		Extrinsic group	
12.0 o	20.5 o	5.0 o	17.4 o
12.0 y	20.6 o	5.4 y	17.5 o
12.9 y	21.3 o	6.1 y	18.5 o
13.6 o	21.6 y	10.9 o	18.7 y
16.6 y	22.1 y	11.8 y	18.7 y
17.2 y	22.2 y	12.0 y	19.2 y
17.5 y	22.6 o	12.3 y	19.5 o
18.2 o	23.1 y	14.8 o	20.7 y
19.1 y	24.0 o	15.0 y	21.2 o
19.3 o	24.3 y	16.8 o	22.1 y
19.8 o	26.7 y	17.2 o	24.0
20.3 o	29.7	17.2 y	
24		23	
19.88		15.74	
4.44		5.25	

Why do an observational study?

- Establishing causation not always the goal

  - Predict whether or not an email is spam

- Randomization may not be ethical

  - Assign clinical trial subjects to treatment or placebo

- May be arguable scientifically that a confounder is "unlikely"

  - 6 month smoking ban in Helena, MT coinciding with 40% reduction in heart attacks

- Might have an incidentally observed dataset

  - Walmart collects petabytes of data/day. Should this data be discarded because it is observational?

Inference to Populations: Random Sample vs. Self-Selection

- Inference to populations **can** be drawn from a R<small>ANDOM SAMPLE</small>

- Inference to populations **cannot** be drawn if units are self-selected

R<small>ANDOM SAMPLE</small>**:** Experimental units selected via a "chance mechanism" from a well defined population

- Example: call randomly selected phone numbers for a survey

S<small>IMPLE RANDOM SAMPLE</small>**:** Every subset of size n is equally likely

- Example: I'll assign everyone in this class a random integer 17, 200, -3, 472, … and survey the n people (units) with smallest numbers

Inference to Populations: Random Sample vs. Self-Selection

- Inference to populations **can** be drawn from a R̲A̲N̲D̲O̲M̲ ̲S̲A̲M̲P̲L̲E̲

- Inference to populations **cannot** be drawn if units are self-selected
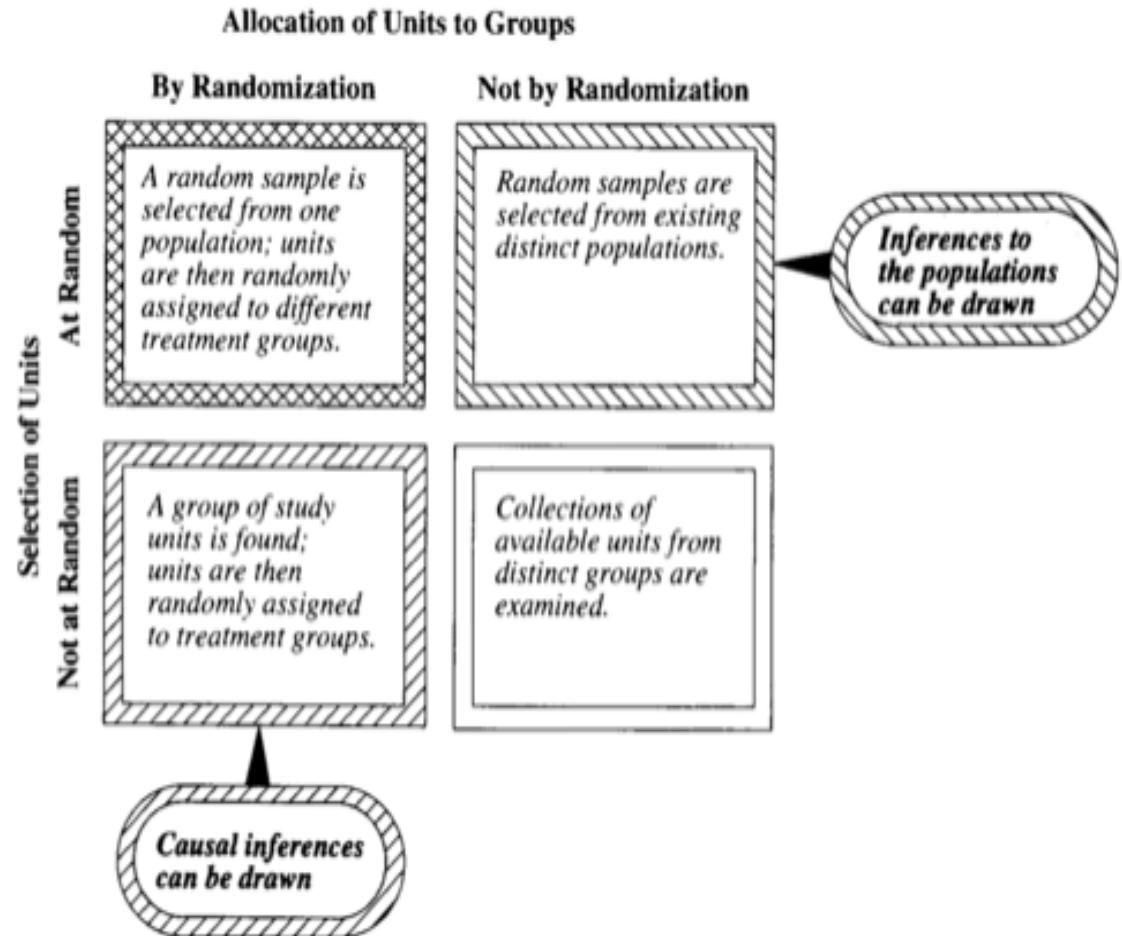
- Neither study uses random sampling
    - <u>Creativity study</u>: units are volunteers
    - <u>Bank study</u>: units are the entire staff
- No inference about a larger population is possible
- Does not mean the results are not interesting or compelling!

Display 1.3 Starting salaries ($U.S.) for 32 male and 61 female clerical hires at a bank

Males			Females					
4,620	5,700	6,000	3,900	4,500	4,800	5,220	5,400	5,640
5,040	6,000	6,000	4,020	4,620	4,800	5,220	5,400	5,700
5,100	6,000	6,000	4,290	4,800	4,980	5,280	5,400	5,700
5,100	6,000	6,300	4,380	4,800	5,100	5,280	5,400	5,700
5,220	6,000	6,600	4,380	4,800	5,100	5,280	5,400	5,700
5,400	6,000	6,600	4,380	4,800	5,100	5,400	5,400	5,700
5,400	6,000	6,600	4,380	4,800	5,100	5,400	5,400	6,000
5,400	6,000	6,840	4,380	4,800	5,100	5,400	5,520	6,000
5,400	6,000	6,900	4,440	4,800	5,100	5,400	5,520	6,120
5,400	6,000	6,900	4,500	4,800	5,160	5,400	5,580	6,300
	6,000	8,100						6,300

Intrinsic group		Extrinsic group	
12.0	20.5	5.0	17.4
12.0	20.6	5.4	17.5
12.9	21.3	6.1	18.5
13.6	21.6	10.9	18.7
16.6	22.1	11.8	18.7
17.2	22.2	12.0	19.2
17.5	22.6	12.3	19.5
18.2	23.1	14.8	20.7
19.1	24.0	15.0	21.2
19.3	24.3	16.8	22.1
19.8	26.7	17.2	24.0
20.3	29.7	17.2	
24		23	
19.88		15.74	
4.44		5.25	

Statistical Inferences Permitted by Study Design

Allocation of Units to Groups

Digitus Tertius