Stat 411/511

Statistical Inference

Jan 18 2012

Charlotte Wickham

stat511.cwick.co.nz

Today

- Scopes of statistical inference

  Study design

- Quantifying evidence

  Study design + data

Inference

A conclusion that patterns in the data are present in some broader context

Inference

A conclusion that patterns in the data are present in some broader context

Statistical Inference

An inference justified by a probability model linking the data with the broader context

Creativity scores in two motivation groups, and their summary statistics

	Motivation Group			
	Intrinsic		***Extrinsic***	
	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	
Sample Size:	24		23	
Average:	19.88	–	15.74	= 4.1
Sample Standard Deviation:	4.44		5.25	

Read: 1.1.1 in Sleuth

Creativity scores in two motivation groups, and their summary statistics

Does intrinsic motivation improve creativity?

	Motivation Group			
	Intrinsic		***Extrinsic***	
	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	
Sample Size:	24		23	
Average:	19.88	–	15.74	= 4.1
Sample Standard Deviation:	4.44		5.25	

Read: 1.1.1 in Sleuth

Creativity scores in two motivation groups, and their summary statistics

Does intrinsic motivation improve creativity?

	Motivation Group: Intrinsic		Motivation Group: Extrinsic	
	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	
Sample Size:	24		23	
Average:	19.88	–	15.74	= 4.1
Sample Standard Deviation:	4.44		5.25	

Motivation Group
Assigned randomly by researcher

Read: 1.1.1 in Sleuth

Creativity scores in two motivation groups, and their summary statistics

Does intrinsic motivation improve creativity?

Motivation Group			
Assigned randomly by researcher			
Intrinsic		***Extrinsic***	
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
20.3	29.7	17.2	

The intrinsic group has a creativity score 4.1 points higher than the extrinsic group

Sample Size:	24		23	
Average:	19.88	–	15.74	= 4.1
Sample Standard Deviation:	4.44		5.25	

Read: 1.1.1 in Sleuth

Statistical Inference

Statistical Inference

Statistical Inference

Question of interest

Data  Broader context

Statistical Inference

Question of interest

Data Broader context

Mechanism of chance

Statistical Inference

Question of interest

Data → Broader context

Mechanism of chance

Does intrinsic motivation improve creativity?

Statistical Inference

Question of interest

Data → Broader context

Mechanism of chance

Does intrinsic motivation improve creativity?

Creativity scores for an intrinsic and extrinsic group

Statistical Inference

Question of interest

Data → Broader context

Mechanism of chance

Does intrinsic motivation improve creativity?

Creativity scores for an intrinsic and extrinsic group → Intrinsic motivation increases creativity

(in this set of people)

Statistical Inference

Question of interest

Data → Broader context

Mechanism of chance

Does intrinsic motivation improve creativity?

Creativity scores for an intrinsic and extrinsic group → Intrinsic motivation increases creativity (in this set of people)

Random allocation to groups

Two scopes of inference

Two scopes of inference

Causal inference

The difference seen in the data was due to the group assignment. Cause and effect relationship.

Two scopes of inference

Causal inference

The difference seen in the data was due to the group assignment. Cause and effect relationship.

Population inference

The difference seen in the data can be inferred to a wider population.

Two mechanisms of chance

Two mechanisms of chance

Random sampling

Subjects are selected from a population at random.

Two mechanisms of chance

Random sampling

Subjects are selected from a population at random.

Random allocation to groups

Subjects are assigned a group (or treatment) at random.

Two mechanisms of chance

Random sampling

Subjects are selected from a population at random.

Random allocation to groups

Subjects are assigned a group (or treatment) at random.

Two mechanisms of chance

Random sampling

Subjects are selected from a population at random.

Random allocation to groups

Subjects are assigned a group (or treatment) at random.

controlled experiment

Two mechanisms of chance

Random sampling

Subjects are selected from a population at random.

Random allocation to groups

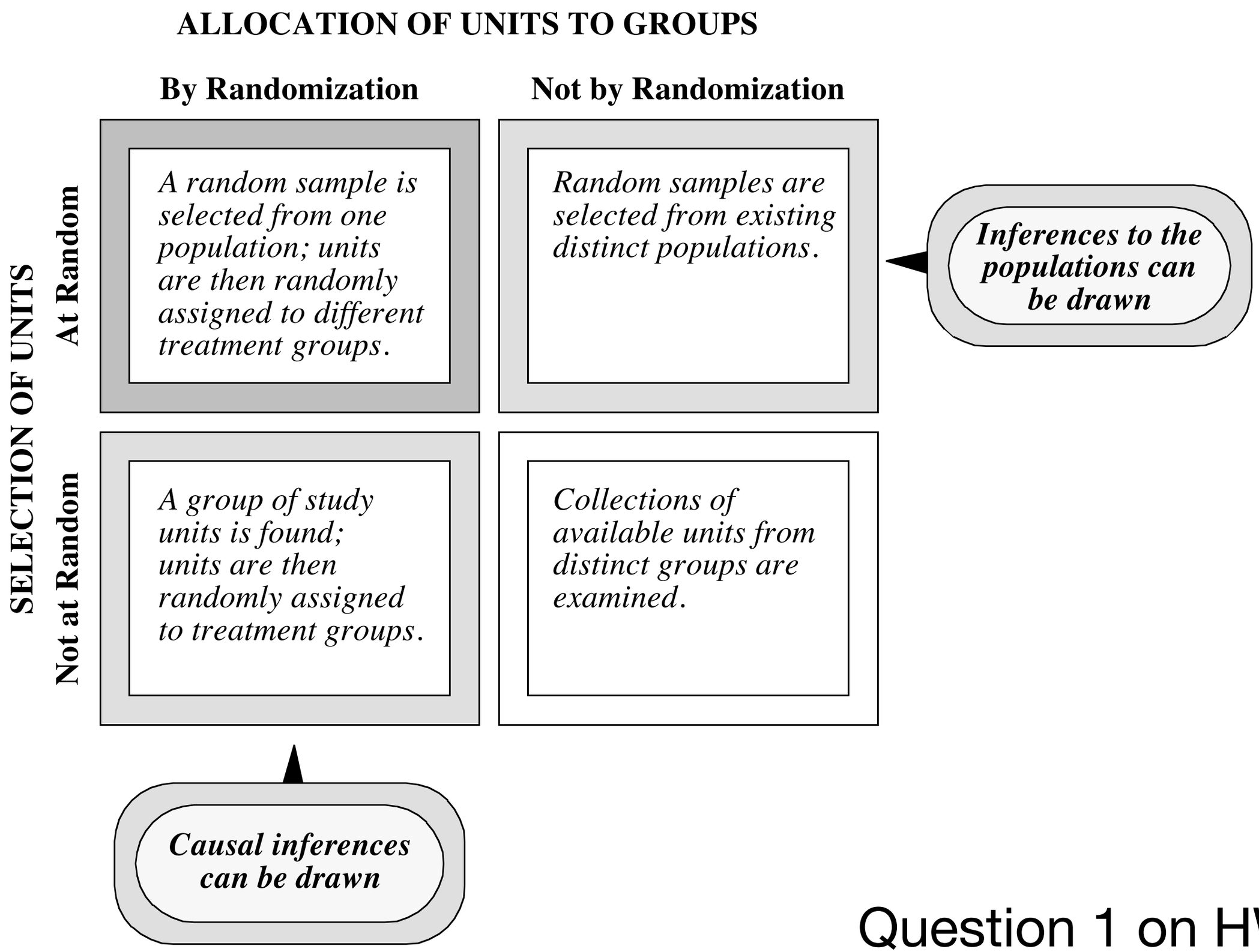
Subjects are assigned a group (or treatment) at random.

controlled experiment

If group assignment is not controlled

observational study

Statistical inferences permitted by study designs



Question 1 on HW#1

Your turn

For each of the following studies decide on the scope of the inferences that could be made:

I compare the heights between male and female students in ST411/511.

I assign you randomly to labs and compare the average midterm grades between the 8am and noon lab classes.

I randomly sample OSU students with a GPA >= 3 and randomly sample OSU students with GPA <3. I compare the number of facebook friends of the high and low GPA samples.

I randomly sample 1000 people living in Corvallis and ask whether they commute by bike.

I pick 20 students in class and give them a extra review lecture every week and compare their final grades to the rest of the class.

Remember

For **population inference** the subjects must be randomly sampled from the population of interest.

For **causal inference** the group membership must be randomly assigned (must be a controlled experiment).

Why do groups need to be randomly assigned for cause and effect?

A **confounding variable** is related to both group assignment and outcome, and makes it hard to determine cause and effect.

The difference in the outcome might be due to the confounder not the group membership.

Example: people who live in houses with > 2 bathrooms have higher blood pressure, wealth might be a confounder.

Randomly assigning subjects to groups breaks the relationship.

Your turn

Come up with potential confounders for these (fictitious) cause and effect relationships:

- Eating oatmeal increases your jogging stamina.
- Breastfeeding babies increases infant mortality.
- Cold weather makes Oregonians spend more money.
- Commuting by bike decreases your commute time.

Which could you do a controlled experiment on?

The mechanism of chance:

- determines how our findings can be extended to a broader context.

Study design

- allows us quantify the evidence the data gives us.

Study design + data