Lecture 3: Chapter 3, Section 4 Designing Studies (Experiments); Begin 4.1: Single Categorical Vars.

- Definitions
- Randomization
- Control
- Blind Experiment
- Pitfalls
- Specific Experimental Designs
- Displaying/Summarizing Single Categorical Variables

Looking Back: *Review*

- **4 Stages of Statistics**
  - Data Production
  - Displaying and Summarizing
  - Probability
  - Statistical Inference

Looking Back: *Review*

- 2 Types of Study Design
  - **Observational study:** record variables' values as they naturally occur
    - **Drawback:** confounding variables due to self-assignment to explanatory values
    - **Example:** *Men who drink beer are more prone to lung cancer than those who drink red wine*

      (what is the confounding variable here?)
  - **Experiment:** researchers control values of explanatory variable
    - ***If well-designed, provides more convincing evidence of causation.***

Definitions

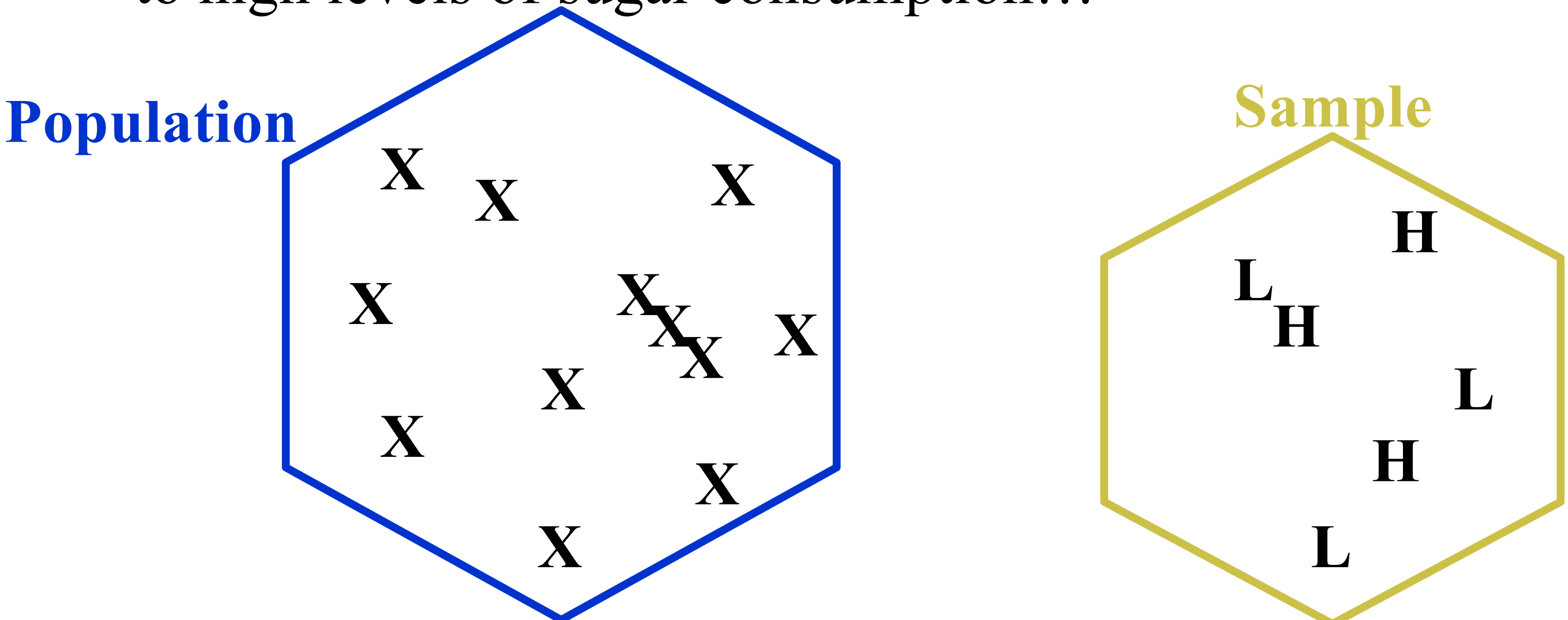
- **Factor**: an explanatory variable in an experiment.
- **Treatment:** value of explanatory variable imposed by researchers in an experiment.

A control group (individuals receiving no treatment or base-line treatment) may be included for comparison.

If individuals are human, we call them subjects.

Example: *Randomized Controlled Experiment*

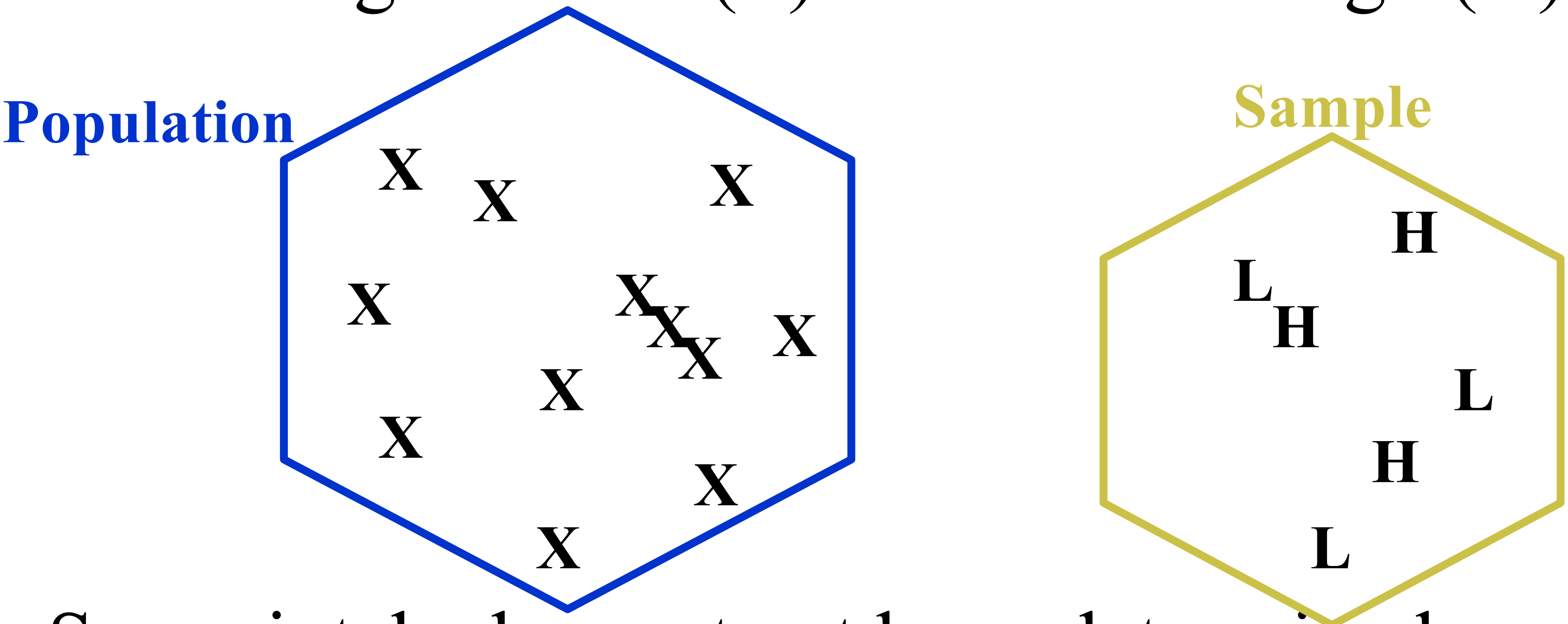
- **Background**: To test if sugar causes hyperactivity, researchers randomly assign some children to low and others to high levels of sugar consumption…



- **Question:** Why are random assignments best?
- **Response:** Randomization helps rule out

Experiment vs. Observational Study

In an experiment, researchers decide who has low sugar intake (L) and who has high (H).

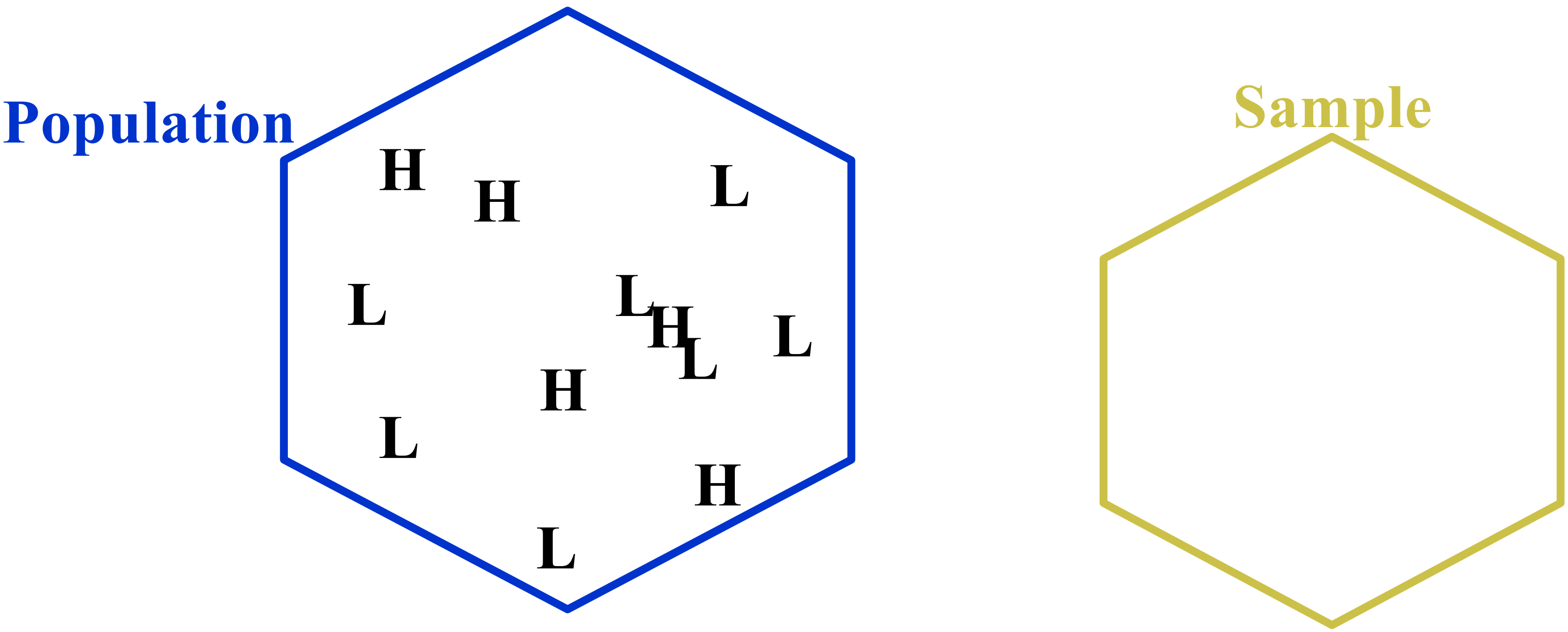


Sugar intake has not yet been determined.

Researchers assign sugar intake L or H.

Experiment vs. Observational Study

In observational study, individuals have already chosen low (L) or high (H) sugar intake.



Researchers make no changes to sugar intake.

Example: Randomize at 1st or 2nd Stage?

- **Background:** Consider two selection issues in our sugar-hyperactivity experiment:
  - What individuals are included in the study?
  - Who consumes low and high amounts of sugar?
- **Question:** At which stage is randomization important?
- **Response:**
  - **1st stage:** Individuals studied ____________________ (Otherwise_____________________ may be an issue.)
  - **2nd stage:**Assignment to sugar (L or H) ______________ Volunteering which treatment to get is ___________

Practice: 3.48c p.61

Must an experiment have a control group?

Recall our definition:

- **Experiment:** researchers manipulate explanatory variable, observe response

Thus, experiment may have no control group…

- if all subjects must be treated
- if simulated treatment is risky
- if the experiment is poorly designed

As long as researchers have taken control of the explanatory variable, it is an experiment.

Definitions: Three Meanings of "Control"

- We **control** for a confounding variable in an observational study by separating it out.
- Researchers **control** who gets what treatment in an experiment by making the assignment themselves, ideally at random.
- The **control** group in an experiment consists of individuals who do not receive a treatment per se, or who are assigned a baseline value of the explanatory variable.

Double-blind Experiments

Two pitfalls may prevent us from drawing a conclusion of causation when results of an experiment show a relationship between the so-called explanatory and response variables.

- If subjects are aware of treatment assignment
- If researchers are aware of treatment assignment

Definitions

- The **placebo effect** is when subjects respond to the *idea* of treatment, not the treatment itself.
- A **placebo** is a “dummy” treatment.
- A **blind** subject is unaware of which treatment he/she is receiving.
- The **experimenter effect** is biased assessment of (or attempt to influence) response due to knowledge of treatment assignment.
- A **blind** experimenter is unaware of which treatment a subject has received.

Example: *Subjects Not Blind*

- **Background**: Suppose after children are randomly assigned to consume either low or high amounts of sugar, researchers find proportion hyperactive is greater for those who consumed higher amounts.
- **Question:** Can we conclude sugar causes hyperactivity?
- **Response:**

Improvement:

Practice: 3.36a p.59

Example: *Experimenters Not Blind*

- **Background**: Suppose after children are randomly assigned to diets sweetened either artificially or with sugar, researchers find proportion hyperactive is greater for those who consumed sugar.
- **Question:** Can we conclude sugar causes hyperactivity?
- **Response:**

More problematic if responses are assessed

____

Improvement:

Practice: 3.37 p.59

Best Evidence of Causation

In general, conclusions of causation are most convincing if a relationship has been established in a randomized controlled double-blind experiment.

A Closer Look: *In the original studies reporting a relationship between sugar and hyperactivity, conducted in the 1970's, experimenters may have been aware of the children's diet when they assessed behavior (randomized controlled **single**-blind). Many studies since then have failed to establish a relationship.*

Other Pitfalls in Experimentation

- Lack of realism (lack of ecological validity)
- Hawthorne effect (people's performance is improved due to awareness of being observed)
- Non-compliance
- Treatments unethical
- Treatments impractical/impossible to impose

Example: *Hawthorne Effect, Lack of Realism*

- **Background**: Suppose researchers want to determine if TV makes people snack more. *While study participants are presumably waiting to be interviewed, half are assigned to a room with a TV on (and snacks), the other half to a room with no TV (and snacks). See if those in the room with TV consume more snacks.*
- **Question:** If participants in the room with TV snack more, can we conclude that, in general, people snack more when they watch TV?
- **Response:** No: ________________
(TV & snacking habits different in contrived setting);
________________ (if people suspect they're observed).

Example: *Non-Compliance in Experiment*

- **Background**: To test if sugar causes hyperactivity, researchers randomly assign 50 children to low and 50 to high levels of sugar consumption; 20 drop out of each group. For remaining children (30 in each group), suppose proportion hyperactive is substantially greater in the high-sugar group.
- **Question:** Can we conclude sugar causes hyperactivity?
- **Response:** _____ __________________makes treatment and control groups different in ways that may affect response.

Practice: 3.39 p.60

Example: *Another Flawed Experiment*

- **Background**: To test if stuttering is a learned (rather than inborn) trait, a researcher in Iowa in 1939 randomly assigned subjects to…
  - Control: 11 orphans in ordinary speech therapy
  - Treatment: 11 orphans badgered and interrupted in sessions with speech therapist

  Of the 11 in treatment group, 8 became stutterers.
- **Question:** What's wrong with this experiment?
- **Response:** ______________

Practice: 3.40e p.60

Examples: *Treatments Impossible/Impractical*

- Taller men get married sooner, promoted quicker, and earn higher wages…
- There is a link between obesity and low socio-economic status in women…

Height is impossible to control.

Weight is difficult to control.

Socio-economic status is too costly to control.

Modifications to randomized experiment

- Blocking: Divide first into groups of individuals who are similar with respect to an outside variable that may be important in relationship studied.
- Paired design: Randomly assign one of each pair to receive treatment, the other control. (*Before-and-after* is a common paired design.)

Looking Back: *blocking is to experimentation as stratification is to sampling.*

Example: *Blocked Experiment*

- **Background**: Study tested theory that use of stronger sunscreen *causes* more time in sun. Before vacation, 40+ students given weak sunscreen, 40+ given strong. Students recorded time spent in sun each day.
- **Question:** How to incorporate blocking, if researchers suspect location (beach or mountains) plays a role in relationship between type of sunscreen and amount of time spent in sun?
- **Response:**

Example: *Paired Experiment*

- **Background**: Study tested theory that use of stronger sunscreen *causes* more time in sun. Before vacation, 40+ students given weak sunscreen, 40+ given strong. Students recorded time spent in sun each day.
- **Question:** How to incorporate paired design, if researchers suspect location plays a role in relationship between type of sunscreen and amount of time spent in sun?
- **Response:**

Practice: 3.40g p.60

Advantage of Paired Design

The paired design helps to ensure that treatment and control groups are as similar as possible in all other respects, so that if their responses differ, we have evidence that the treatment is responsible.

Discussion Question: *Why do not just twins, but also researchers, flock to the annual festival in Twinsburg, Ohio?*

Example: *Combining Paired and Two-Sample Designs*

- **Background**: Studies often randomly assign one group to a placebo and the other to a drug. Responses to the variable of interest are assessed before and after a period of time, then compared to see benefits or side effects.
- **Question:** What aspect of the design is two-sample, and what aspect is paired?
- **Response:**

  two-sample:

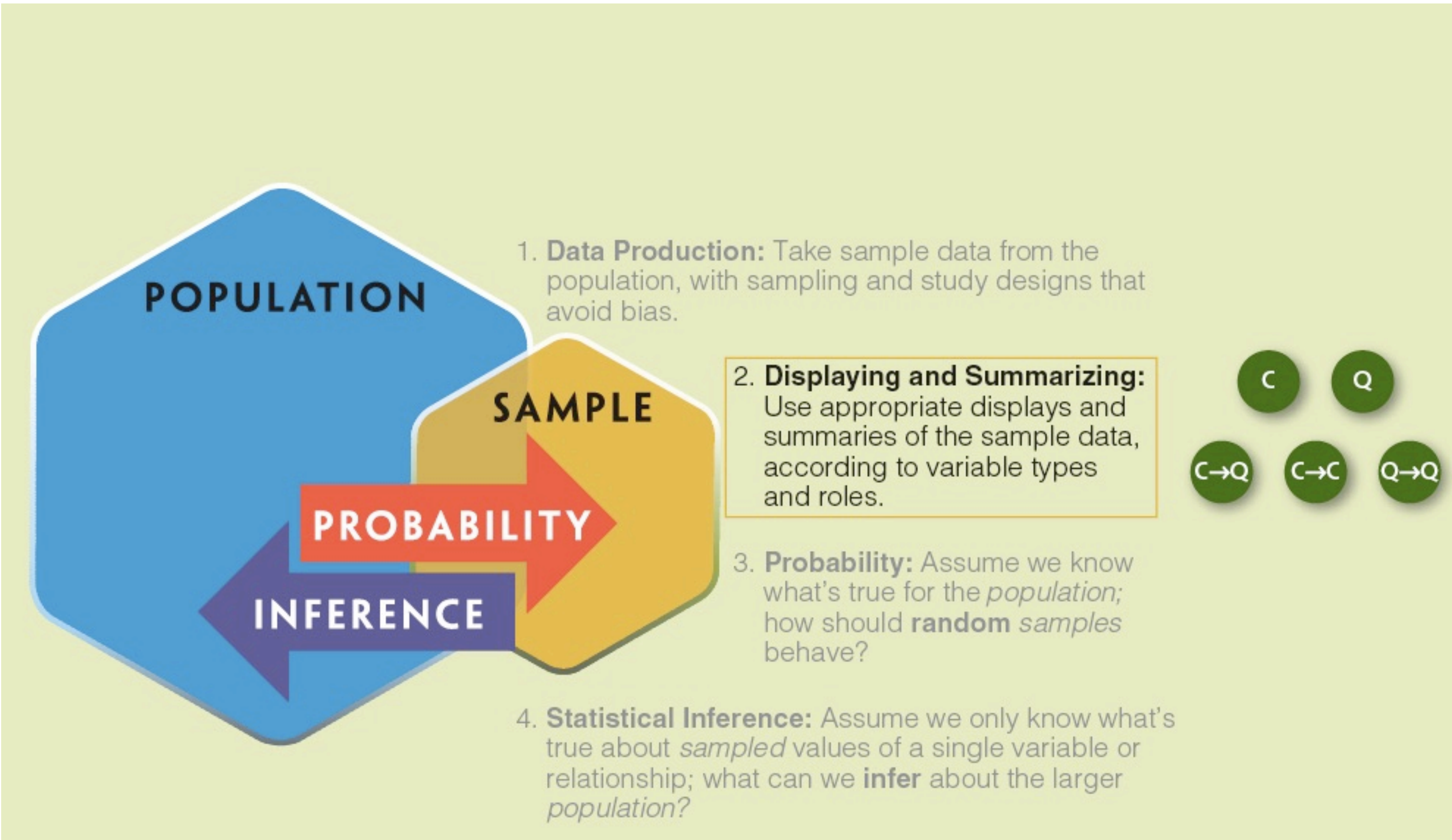
  paired:

Practice: 3.47 p.61

Looking Back: *Review*

- **4 Stages of Statistics**
  - Data Production (discussed in Lectures 1-4)
  - Displaying and Summarizing
    - Single variables: 1 categorical, 1 quantitative
    - Relationships between 2 variables
  - Probability
  - Statistical Inference

Focus on Displaying and Summarizing



Handling Single Categorical Variables

- **Display:**
  - Pie chart
  - Bar graph
- **Summary:**
  - Count
  - Percent
  - Proportion

Definitions and Notation

- **Statistic:** number summarizing sample
- **Parameter:** number summarizing population

- $\hat{p}$: sample proportion (a statistic) [“p-hat”]
- p**:** population proportion (a parameter)

Example: *Issues to Consider*

- **Background**: 246 of 446 students at a certain university had eaten breakfast on survey day.
- **Questions:**
  - Are intro stat students representative of all students at that university?
  - Would they respond without bias?
- **Responses:**
  - ______________________________________
  - ______________________________________

***Looking Back**: these are data production issues.*

Practice: 4.4e p.79

Example: *More Issues to Consider*

- **Background**: 246 of 446 students at a certain university had eaten breakfast on survey day.
- **Questions:**
  - How do we display and summarize the info?
  - Can we conclude that a majority of *all* students at that university eat breakfast?
- **Responses:**
  - Display: ___________

  Summary: ________________________
  - ____

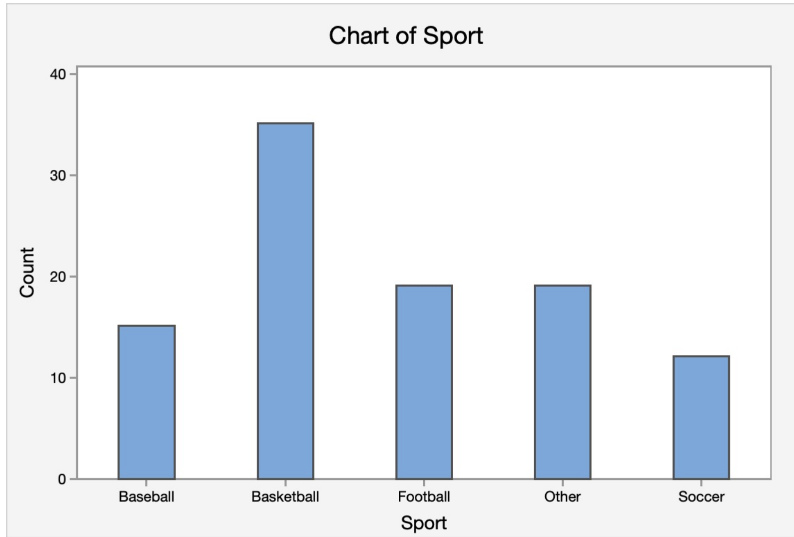
Example: *Statistics vs. Parameters*

- **Background**: 246 of 446 students at a certain university had eaten breakfast on survey day.
- **Questions:**
  - Is 246/446=0.55 a statistic or a parameter? How do we denote it?
  - Is the proportion of all students eating breakfast a statistic or a parameter? How do we denote it?
- **Responses:**
  - 246/446=0.55 is a ___________denoted ____.
  - Proportion of all students eating breakfast is a___________ denoted ____.

Practice: 4.3c p.79

Example: *Summary Issues*

- **Background**: Sports played by world's top (highest paid) 100 athletes:
  - 35 basketball
  - 19 football
  - 15 baseball
  - 12 soccer
  - 19 other



- **Question:**
  - Why is it not appropriate to use this info to draw conclusions about the larger population of athletes?
- **Response:**
  - ____

Example: *Notation*

- **Background**: In study of 20 antarctic prions (birds), 17 correctly chose the one of two bags that had contained their mate.
- **Questions:** How do we denote sample and population proportions? Are they statistics or parameters?
- **Responses:**
  - sample proportion ________________ is a ________________.
  - population proportion____ is a ____.

Definitions

- **Mode:** most common value
- **Majority:** more common of two possible values (same as mode)
- **Minority:** less common of two possible values

Example: *Role of Sample Size*

- **Background**: In study of 20 antarctic prions (birds), 17 correctly chose the one of two bags that had contained their mate.
- **Question:** Would we be more convinced that a *majority* of all prions would choose correctly, if 170 out of 200 were correct?
- **Response:**

Practice: 4.3i p.79

Example: *Sampling Design*

- **Background**: In study of 20 antarctic prions (birds), 17 correctly chose the one of two bags that had contained their mate.
- **Question:** Is the sample biased?
- **Response:**

Example: *Study Design*

- **Background**: Antarctic prions presented with Y-shaped maze, a bag at the end of each arm. One bag had contained mate, the other not.
- **Question:**
  - What were researchers attempting to show?
- **Response:**

Example: *Study Design*

- **Background**: Antarctic prions presented with Y-shaped maze, a bag at the end of each arm. One bag had contained mate, the other not.
- **Question:**
  - Why use bags and not birds themselves?
- **Response:**

Example: *Study Design*

- **Background**: Antarctic prions presented with Y-shaped maze, a bag at the end of each arm. One bag had contained mate, the other not.
- **Question:**
  - Why *"had"* contained (bird no longer in bag)?
- **Response:**

Example: *Study Design*

- **Background**: Antarctic prions presented with Y-shaped maze, a bag at the end of each arm. One bag had contained mate, the other not.
- **Question:**
  - OK to always place correct bag on right?
- **Response:**

Example: *Study Design*

- **Background**: Antarctic prions presented with Y-shaped maze, a bag at the end of each arm. One bag had contained mate, the other not.
- **Question:** Should the other be just any empty bag?
- **Response:**

Looking Ahead: *Researchers were careful to avoid bias in their study design. A success rate of 85% is impressive but we need inference methods to quantify claims that prions in general can recognize their mate by smell.*

Lecture Summary *(Experiments)*

- Definitions
- Randomization; 2 stages of selection
- Control group
- "Blind" study design
  - Subjects blind to avoid placebo effect
  - Researchers blind to avoid experimenter effect
- Other pitfalls of experiments: lack of realism, Hawthorne effect, non-compliance, unethical or impractical treatment
- Specific experimental designs
  - Blocked
  - Paired or two-sample

Lecture Summary *(Categorical Variables)*

- **Display:** pie chart, bar graph
- **Summarize:** count, percent, proportion
- **Sampling:** data unbiased (representative)?
- **Design:** produced unbiased summary of data?
- **Inference:** will we ultimately draw conclusion about population based on sample?
- **Mode, Majority:** most common values
- **Larger samples:** provide more info