Lecture 1: Chapter 1

C C Moxley

UAB Mathematics

29 August 16

§1.2 Basic Concepts

§1.2 Basic Concepts

Definition (Population)

A **population** is the entire collection of all measurements/data that are under consideration.

§1.2 Basic Concepts

Definition (Population)

A **population** is the entire collection of all measurements/data that are under consideration.

Definition (Sample)

A **sample** is a subcollection of a population.

§1.2 Basic Concepts

Definition (Population)

A **population** is the entire collection of all measurements/data that are under consideration.

Definition (Sample)

A **sample** is a subcollection of a population.

Can our class be a population?

§1.2 Basic Concepts

Definition (Population)

A **population** is the entire collection of all measurements/data that are under consideration.

Definition (Sample)

A **sample** is a subcollection of a population.

Can our class be a population? Can it be a sample?

§1.2 Basic Concepts

§1.2 Basic Concepts

Definition (Parameter)

A **parameter** is an aggregated measurement depending on all members of a population.

A **census** is required for determining a parameter!

§1.2 Basic Concepts

Definition (Parameter)

A **parameter** is an aggregated measurement depending on all members of a population.

A **census** is required for determining a parameter!

Definition (Statistic)

A **statistic** is an aggregated measurement depending on only the members of a sample.

§1.2 Basic Concepts

Definition (Parameter)

A **parameter** is an aggregated measurement depending on all members of a population.

A **census** is required for determining a parameter!

Definition (Statistic)

A **statistic** is an aggregated measurement depending on only the members of a sample.

If we considered the mean age of students in this class, in what cases would this aggregated measurement be a statistic?

§1.2 Basic Concepts

Definition (Parameter)

A **parameter** is an aggregated measurement depending on all members of a population.

A **census** is required for determining a parameter!

Definition (Statistic)

A **statistic** is an aggregated measurement depending on only the members of a sample.

If we considered the mean age of students in this class, in what cases would this aggregated measurement be a statistic? A parameter?

§1.2 Common Pitfalls in Statistics

Statistics is the process of planning studies/experiments, collecting data, and analyzing & interpreting this data. Here are some rookie mistakes often made in the process of doing statistics.

§1.2 Common Pitfalls in Statistics

Statistics is the process of planning studies/experiments, collecting data, and analyzing & interpreting this data. Here are some rookie mistakes often made in the process of doing statistics.

- **bias**:

§1.2 Common Pitfalls in Statistics

Statistics is the process of planning studies/experiments, collecting data, and analyzing & interpreting this data. Here are some rookie mistakes often made in the process of doing statistics.

- **bias**: *a possible interest in the outcome of a survey, experiment, etc.*

§1.2 Common Pitfalls in Statistics

Statistics is the process of planning studies/experiments, collecting data, and analyzing & interpreting this data. Here are some rookie mistakes often made in the process of doing statistics.

- **bias**: *a possible interest in the outcome of a survey, experiment, etc.* Two possible sources: subject bias and investigator bias.

§1.2 Common Pitfalls in Statistics

Statistics is the process of planning studies/experiments, collecting data, and analyzing & interpreting this data. Here are some rookie mistakes often made in the process of doing statistics.

- **bias**: *a possible interest in the outcome of a survey, experiment, etc.* Two possible sources: subject bias and investigator bias.
- **confusing correlation & causation**

§1.2 Common Pitfalls in Statistics

Statistics is the process of planning studies/experiments, collecting data, and analyzing & interpreting this data. Here are some rookie mistakes often made in the process of doing statistics.

- **bias**: *a possible interest in the outcome of a survey, experiment, etc.* Two possible sources: subject bias and investigator bias.
- **confusing correlation & causation**
- **pitfalls of data analysis and collection**: self-reported data, small samples, loaded questions, question order, nonresponse, missing data, precise numbers, percentages.

Concept Mastery:

In a survey of 109 randomly selected cell phone users, subjects were asked to tell how many texts they sent each month and how many they received. What are the potential pitfalls in this survey?

Concept Mastery:

Is anything wrong with this statement?

Concept Mastery:

Is anything wrong with this statement? In the last 50 years, a low crime rate in Hartford, CT, was highly correlated with having a republican mayor in office. Thus, republican mayors are more effective at preventing crime.

Concept Mastery:

Is it possible for a quantity to increase by 200%?

Concept Mastery:

Is it possible for a quantity to increase by 200%? Can a quantity decrease by 200%?

Concept Mastery:

Is anything wrong with this statement?

Concept Mastery:

Is anything wrong with this statement? The number of flu deaths in the US is remarkably high this year: 53,000 people have died this year compared to 45,000 last year.

Concept Mastery

Identify the parameter and the statistic. A survey of 2500 people determined that 40% of the 241,472,385 adults in America are employed in their desired field.

§1.3 Types of Data: Qualitative vs. Quantitative

Qualitative/categorical/attribute data vs. quantitative/numerical data:

§1.3 Types of Data: Qualitative vs. Quantitative

Qualitative/categorical/attribute data vs. quantitative/numerical data: qualitative is non-measurement data whereas quantitative is measurement data.

§1.3 Types of Data: Qualitative vs. Quantitative

Qualitative/categorical/attribute data vs. quantitative/numerical data: qualitative is non-measurement data whereas quantitative is measurement data.
Can numbers be qualitative? Can letters be quantitative?

§1.3 Types of Data: Discrete vs. Continuous

discrete: countable - either infinitely countable or finitely countable

§1.3 Types of Data: Discrete vs. Continuous

discrete: countable - either infinitely countable or finitely countable
continuous: not countable

§1.3 Types of Data: Discrete vs. Continuous

discrete: countable - either infinitely countable or finitely countable
continuous: not countable
The speed of a car can be between 0 and 120mph. Is this data discrete or continuous?

§1.3 Types of Data: Levels of Data

Level of Measurement	Description	Example

§1.3 Types of Data: Levels of Data

Level of Measurement	**Description**	**Example**
nominal	not orderable, labels only	hair color

§1.3 Types of Data: Levels of Measurement

Level	**Description**	**Example**
nominal	not orderable, labels only	hair color
ordinal	orderable, differences meaningless	college rankings

§1.3 Types of Data: Levels of Measurement

Level	**Description**	**Example**
nominal	not orderable, labels only	hair color
ordinal	orderable, differences meaningless	college rankings
interval	no natural zero, differences useful	body temperature

§1.3 Types of Data: Levels of Measurement

Level	**Description**	**Example**
nominal	not orderable, labels only	hair color
ordinal	orderable, differences meaningless	college rankings
interval	no natural zero, useful differences	body temperature
ratio	natural zero, useful differences	speed

Concept Mastery

Identify the level of measurement and determine if the calculation is meaningful.

Concept Mastery

Identify the level of measurement and determine if the calculation is meaningful. The average flight number of 55 different flights is 451.2.

Concept Mastery

Identify the level of measurement and determine if the calculation is meaningful. The average flight number of 55 different flights is 451.2.
Is this a statistic or a parameter?

Concept Mastery

Give the level of measurement of the following.

- The star-rating on Yelp!

Concept Mastery

Give the level of measurement of the following.

- The star-rating on Yelp! **Ordinal.**

Concept Mastery

Give the level of measurement of the following.

- The star-rating on Yelp! **Ordinal.**
- Letter grades.

Concept Mastery

Give the level of measurement of the following.

- The star-rating on Yelp! **Ordinal.**
- Letter grades. **Ordinal.**

Concept Mastery

Give the level of measurement of the following.

- The star-rating on Yelp! **Ordinal.**
- Letter grades. **Ordinal.**
- Length of a movie.

Concept Mastery

Give the level of measurement of the following.

- The star-rating on Yelp! **Ordinal.**
- Letter grades. **Ordinal.**
- Length of a movie. **Ratio.**

Concept Mastery

Give the level of measurement of the following.

- The star-rating on Yelp! **Ordinal.**
- Letter grades. **Ordinal.**
- Length of a movie. **Ratio.**
- Years in which the economy shrunk.

Concept Mastery

Give the level of measurement of the following.

- The star-rating on Yelp! **Ordinal.**
- Letter grades. **Ordinal.**
- Length of a movie. **Ratio.**
- Years in which the economy shrunk. **Interval.**

§1.4 Collecting Sample Data: Observation vs. Experiment

An **observational** study does not modify the subjects of the study - it just observes a particular characteristic.

§1.4 Collecting Sample Data: Observation vs. Experiment

An **observational** study does not modify the subjects of the study - it just observes a particular characteristic.
An **experimental** study is one which seeks to modify its subjects. Often, what is sought is an *understanding of the change* which occurs due to this modification.

§1.4 Collecting Sample Data: Simple Random Sample

A **simple random sample** has a fixed size (n) and is such that any collection of n subjects in the population has an equal likelihood of being selected.

§1.4 Collecting Sample Data: Simple Random Sample

A **simple random sample** has a fixed size (n) and is such that any collection of n subjects in the population has an equal likelihood of being selected.
What's the difference between this and a **random sample**?

§1.4 Collecting Sample Data: Types of Samples

Type	**Description**
Random	each subject equally likely to be chosen

§1.4 Collecting Sample Data: Types of Samples

Type	**Description**
random	each subject equally likely to be chosen
simple random	each group of n subjects equally likely

§1.4 Collecting Sample Data: Types of Samples

Type	**Description**
random	each subject equally likely to be chosen
simple random	each group of n subjects equally likely
systematic	select each kth subject

§1.4 Collecting Sample Data: Types of Samples

Type	**Description**
random	each subject equally likely to be chosen
simple random	each group of n subjects equally likely
systematic	select each kth subject
convenience	select an easy obtained sample

§1.4 Collecting Sample Data: Types of Samples

Type	**Description**
random	each subject equally likely to be chosen
simple random	each group of n subjects equally likely
systematic	select each kth subject
convenience	select an easy obtained sample
stratified	subdivide and select sample from each subgroup

§1.4 Collecting Sample Data: Types of Samples

Type	**Description**
random	each subject equally likely to be chosen
simple random	each group of n subjects equally likely
systematic	select each kth subject
convenience	select an easy obtained sample
stratified	subdivide and select sample from each subgroup
cluster	subdivide, randomly choose collection of subgroups

§1.4 Collecting Sample Data: Types of Samples

Type	**Description**
random	each subject equally likely to be chosen
simple random	each group of n subjects equally likely
systematic	select each kth subject
convenience	select an easy obtained sample
stratified	subdivide and select sample from each subgroup
cluster	subdivide, randomly choose collection of subgroups

These types of sampling can be combined!

§1.4 Collecting Sample Data: Types of Studies

Cross-sectional studies are those which involve data collected at a single point in time.

§1.4 Collecting Sample Data: Types of Studies

Cross-sectional studies are those which involve data collected at a single point in time.
Retrospective studies involve data collected in the past.

§1.4 Collecting Sample Data: Types of Studies

Cross-sectional studies are those which involve data collected at a single point in time.
Retrospective studies involve data collected in the past.
Prospective studies involve data collected at a future point in time.

§1.4 Collecting Sample Data: Experimental Design

Important characteristics to consider in experimental design: randomness, replication, blinding, double-blinding, and placebo effects.

§1.4 Collecting Sample Data: Experimental Design

Important characteristics to consider in experimental design: randomness, replication, blinding, double-blinding, and placebo effects. Avoid

§1.4 Collecting Sample Data: Experimental Design

Important characteristics to consider in experimental design: randomness, replication, blinding, double-blinding, and placebo effects. Avoid **confounding**!

§1.4 Collecting Sample Data: Experimental Design

Important characteristics to consider in experimental design: randomness, replication, blinding, double-blinding, and placebo effects.
Avoid **confounding**!
Also consider randomized experimenting and a randomized block design.

§1.4 Collecting Sample Data: Errors

Sampling errors happen when there is a difference between the measurement in the sample and its corresponding parameter, despite good sampling methods.

§1.4 Collecting Sample Data: Errors

Sampling errors happen when there is a difference between the measurement in the sample and its corresponding parameter, despite good sampling methods. **Non-sampling errors** occur are the result of human error or poor judgment.

§1.4 Collecting Sample Data: Errors

Sampling errors happen when there is a difference between the measurement in the sample and its corresponding parameter, despite good sampling methods. **Non-sampling errors** occur are the result of human error or poor judgment. **Non-random sampling errors** happen when there is a difference between the measurement in the sample and its corresponding parameter and when the sampling is non-random.

Concept Mastery

Is the following an observational or experimental study?

Concept Mastery

Is the following an observational or experimental study?
A study wants to know if calling at different times of day affects responses. They call a randomly selected sample of people at noon and another randomly selected sample of people at 6PM and compare the results.

Concept Mastery

Is the following an observational or experimental study?
A study wants to know if calling at different times of day affects responses. They call a randomly selected sample of people at noon and another randomly selected sample of people at 6PM and compare the results.
A study wants to know the effect of laughter on respondents' answers. They ask people political questions after they've watched a stand-up comic.

Concept Mastery

What kind of sampling is used?

Concept Mastery

What kind of sampling is used?
An exit poll asks questions of all voters at several polling stations which were randomly selected.

Concept Mastery

What kind of sampling is used?
An exit poll asks questions of all voters at several polling stations which were randomly selected.
An exit poll asks questions of every 5th voter who leaves at several polling stations which were randomly selected.

Concept Mastery

What kind of sampling is used?
An exit poll asks questions of all voters at several polling stations which were randomly selected.
An exit poll asks questions of every 5th voter who leaves at several polling stations which were randomly selected.
CBS polled neighbors at a house fire to determine the level of concern they had about fires in their home.

Concept Mastery

How might you "control" to see if a subject's sex was a determining factor in the the effectiveness of a drug?