

Stratified and Cluster Sampling

Lecture 8 Sections 2.6, 2.8

Robb T. Koether

Hampden-Sydney College

Tue, Jan 27, 2008

Outline

Stratified and
Cluster
Sampling

Robb T.
Koether

Introduction

Stratified
Random
Samples
Example

Estimating
Parameters

Cluster
Samples
Example

Stratified vs.
Cluster

Assignment

- 1 Introduction
- 2 Stratified Random Samples
 - Example
- 3 Estimating Parameters
- 4 Cluster Samples
 - Example
- 5 Stratified vs. Cluster
- 6 Assignment

Introduction

Stratified and Cluster Sampling

Robb T.
Koether

Introduction

Stratified
Random
Samples
Example

Estimating
Parameters

Cluster
Samples
Example

Stratified vs.
Cluster

Assignment

- Suppose we know that our population is 60% male and 40% female.
- Wouldn't it make sense to ensure that our sample is also 60% male and 40% female.
- After all, our goal is *representativeness*, not *randomness*, right?
- If our sample size were $n = 100$, we could
 - Select a simple random sample of 60 males.
 - Select a simple random sample of 40 females.
- Is this a good idea?
- If we left it to chance, we might end up with 50 males and 50 females.

Introduction

Stratified and Cluster Sampling

Robb T.
Koether

Introduction

Stratified Random Samples Example

Estimating Parameters

Cluster Samples Example

Stratified vs. Cluster

Assignment

- Suppose we know that our population is 60% male and 40% female.
- Wouldn't it make sense to ensure that our sample is also 60% male and 40% female.
- After all, our goal is *representativeness*, not *randomness*, right?
- If our sample size were $n = 100$, we could
 - Select a simple random sample of 60 males.
 - Select a simple random sample of 40 females.
- Is this a good idea?
- If we left it to chance, we might end up with 50 males and 50 females.

Introduction

Stratified and
Cluster
Sampling

Robb T.
Koether

Introduction

Stratified
Random
Samples
Example

Estimating
Parameters

Cluster
Samples
Example

Stratified vs.
Cluster

Assignment

- Suppose we know that our population is 60% male and 40% female.
- Wouldn't it make sense to ensure that our sample is also 60% male and 40% female.
- After all, our goal is *representativeness*, not *randomness*, right?
- If our sample size were $n = 100$, we could
 - Select a simple random sample of 60 males.
 - Select a simple random sample of 40 females.
- Is this a good idea?
- If we left it to chance, we might end up with 50 males and 50 females.

Introduction

Stratified and Cluster Sampling

Robb T.
Koether

Introduction

Stratified Random Samples Example

Estimating Parameters

Cluster Samples Example

Stratified vs. Cluster

Assignment

- Suppose we know that our population is 60% male and 40% female.
- Wouldn't it make sense to ensure that our sample is also 60% male and 40% female.
- After all, our goal is *representativeness*, not *randomness*, right?
- If our sample size were $n = 100$, we could
 - Select a simple random sample of 60 males.
 - Select a simple random sample of 40 females.
- Is this a good idea?
- If we left it to chance, we might end up with 50 males and 50 females.

Introduction

Stratified and Cluster Sampling

Robb T.
Koether

Introduction

Stratified Random Samples Example

Estimating Parameters

Cluster Samples Example

Stratified vs. Cluster

Assignment

- Suppose we know that our population is 60% male and 40% female.
- Wouldn't it make sense to ensure that our sample is also 60% male and 40% female.
- After all, our goal is *representativeness*, not *randomness*, right?
- If our sample size were $n = 100$, we could
 - Select a simple random sample of 60 males.
 - Select a simple random sample of 40 females.
- Is this a good idea?
- If we left it to chance, we might end up with 50 males and 50 females.

Introduction

Stratified and
Cluster
Sampling

Robb T.
Koether

Introduction

Stratified
Random
Samples
Example

Estimating
Parameters

Cluster
Samples
Example

Stratified vs.
Cluster

Assignment

- Suppose we know that our population is 60% male and 40% female.
- Wouldn't it make sense to ensure that our sample is also 60% male and 40% female.
- After all, our goal is *representativeness*, not *randomness*, right?
- If our sample size were $n = 100$, we could
 - Select a simple random sample of 60 males.
 - Select a simple random sample of 40 females.
- Is this a good idea?
- If we left it to chance, we might end up with 50 males and 50 females.

Introduction

Stratified and Cluster Sampling

Robb T.
Koether

Introduction

Stratified Random Samples Example

Estimating Parameters

Cluster Samples Example

Stratified vs. Cluster

Assignment

- Suppose we know that our population is 60% male and 40% female.
- Wouldn't it make sense to ensure that our sample is also 60% male and 40% female.
- After all, our goal is *representativeness*, not *randomness*, right?
- If our sample size were $n = 100$, we could
 - Select a simple random sample of 60 males.
 - Select a simple random sample of 40 females.
- Is this a good idea?
- If we left it to chance, we might end up with 50 males and 50 females.

Introduction

Stratified and
Cluster
Sampling

Robb T.
Koether

Introduction

Stratified
Random
Samples
Example

Estimating
Parameters

Cluster
Samples
Example

Stratified vs.
Cluster

Assignment

- Suppose we know that our population is 60% male and 40% female.
- Wouldn't it make sense to ensure that our sample is also 60% male and 40% female.
- After all, our goal is *representativeness*, not *randomness*, right?
- If our sample size were $n = 100$, we could
 - Select a simple random sample of 60 males.
 - Select a simple random sample of 40 females.
- Is this a good idea?
- If we left it to chance, we might end up with 50 males and 50 females.

Introduction

Stratified and
Cluster
Sampling

Robb T.
Koether

Introduction

Stratified
Random
Samples
Example

Estimating
Parameters

Cluster
Samples
Example

Stratified vs.
Cluster

Assignment

- On the other hand, is this necessary?
- What is we did end up with 50 males and 50 females?
- Suppose male opinion in our sample was divided 20 yes, 30 no and female opinion was divided 10 yes, 40 no.
- How would we come up with an overall estimate for the proportion of yeses in the population?

Introduction

Stratified and
Cluster
Sampling

Robb T.
Koether

Introduction

Stratified
Random
Samples
Example

Estimating
Parameters

Cluster
Samples
Example

Stratified vs.
Cluster

Assignment

- On the other hand, is this necessary?
- What is we did end up with 50 males and 50 females?
- Suppose male opinion in our sample was divided 20 yes, 30 no and female opinion was divided 10 yes, 40 no.
- How would we come up with an overall estimate for the proportion of yeses in the population?

Introduction

Stratified and
Cluster
Sampling

Robb T.
Koether

Introduction

Stratified
Random
Samples
Example

Estimating
Parameters

Cluster
Samples
Example

Stratified vs.
Cluster

Assignment

- On the other hand, is this necessary?
- What is we did end up with 50 males and 50 females?
- Suppose male opinion in our sample was divided 20 yes, 30 no and female opinion was divided 10 yes, 40 no.
- How would we come up with an overall estimate for the proportion of yeses in the population?

Introduction

Stratified and
Cluster
Sampling

Robb T.
Koether

Introduction

Stratified
Random
Samples
Example

Estimating
Parameters

Cluster
Samples
Example

Stratified vs.
Cluster

Assignment

- On the other hand, is this necessary?
- What is we did end up with 50 males and 50 females?
- Suppose male opinion in our sample was divided 20 yes, 30 no and female opinion was divided 10 yes, 40 no.
- How would we come up with an overall estimate for the proportion of yeses in the population?

Stratified Random Sample

Stratified and
Cluster
Sampling

Robb T.
Koether

Introduction

Stratified
Random
Samples

Example

Estimating
Parameters

Cluster
Samples

Example

Stratified vs.
Cluster

Assignment

Definition (Homogeneous)

A group is **homogeneous** if its member all have similar characteristics with regard to the variables of interest.

Definition (Stratum)

A **stratum** is a homogeneous subset of the population.

Definition (Stratified random sampling)

Stratified random sampling is a sampling method in which the population is first divided into strata. Then a simple random sample is taken from each stratum. The results constitute the sample.

Examples

Stratified and
Cluster
Sampling

Robb T.
Koether

Introduction

Stratified
Random
Samples

Example

Estimating
Parameters

Cluster
Samples

Example

Stratified vs.
Cluster

Assignment

- Possible strata:
 - Male and female strata.
 - Resident and non-resident strata.
 - White, black, Hispanic, and Asian strata.
 - Protestant, Catholic, Jewish, Muslim, etc., strata.

Stratified Random Sampling

Stratified and
Cluster
Sampling

Robb T.
Koether

Introduction

**Stratified
Random
Samples**

Example

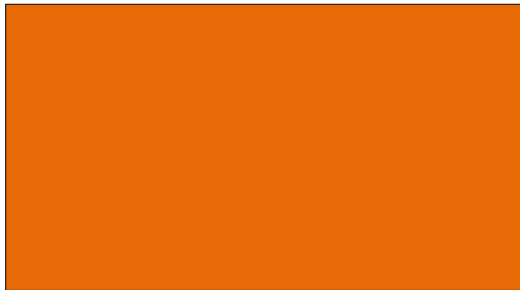
Estimating
Parameters

Cluster
Samples

Example

Stratified vs.
Cluster

Assignment



The population

Stratified Random Sampling

Stratified and Cluster Sampling

Robb T. Koether

Introduction

Stratified Random Samples

Example

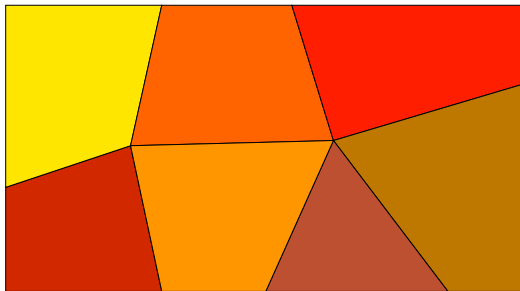
Estimating Parameters

Cluster Samples

Example

Stratified vs. Cluster

Assignment



The strata

Stratified Random Sampling

Stratified and
Cluster
Sampling

Robb T.
Koether

Introduction

Stratified
Random
Samples

Example

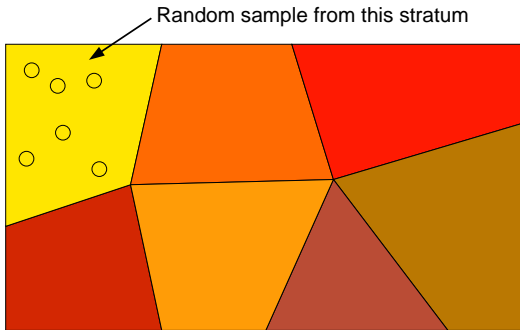
Estimating
Parameters

Cluster
Samples

Example

Stratified vs.
Cluster

Assignment



The strata

Stratified Random Sampling

Stratified and
Cluster
Sampling

Robb T.
Koether

Introduction

Stratified
Random
Samples

Example

Estimating
Parameters

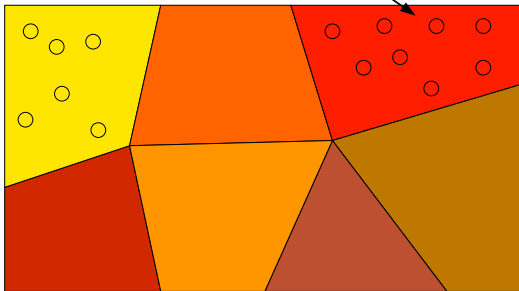
Cluster
Samples

Example

Stratified vs.
Cluster

Assignment

Random sample from this stratum



The strata

Stratified Random Sampling

Stratified and Cluster Sampling

Robb T. Koether

Introduction

Stratified Random Samples

Example

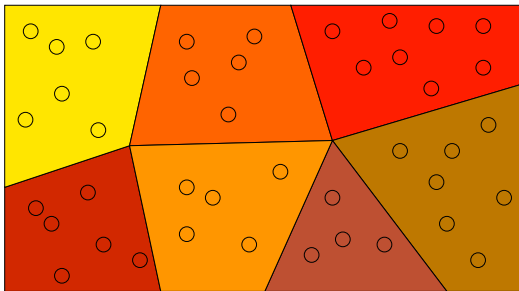
Estimating Parameters

Cluster Samples

Example

Stratified vs. Cluster

Assignment



Random samples from all strata

Stratified Random Sampling

Stratified and
Cluster
Sampling

Robb T.
Koether

Introduction

Stratified
Random
Samples

Example

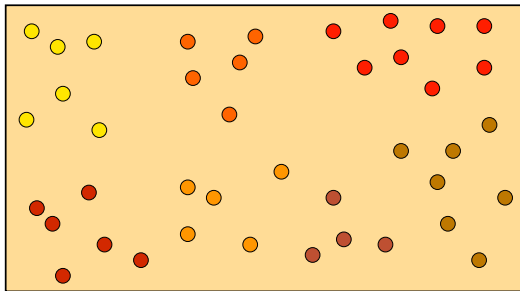
Estimating
Parameters

Cluster
Samples

Example

Stratified vs.
Cluster

Assignment



The stratified sample

Example

Stratified and Cluster Sampling

Robb T.
Koether

Introduction

Stratified
Random
Samples
Example

Estimating
Parameters

Cluster
Samples
Example

Stratified vs.
Cluster

Assignment

Example (Stratified random sample)

- Let the population consist of males Anthony, Benjamin, Christopher, Daniel, Ethan, Francisco, Gabriel, and Hunter and females Isabella, Jasmine, Kayla, Lily, Madison, Natalie, Olivia, and Paige.
- Choose a stratified sample of size $n = 8$, where the strata are the two sexes.
- Is the sample representative with regard to sex?

Estimating Parameters

Stratified and Cluster Sampling

Robb T.
Koether

Introduction

Stratified Random Samples

Example

Estimating Parameters

Cluster Samples

Example

Stratified vs. Cluster

Assignment

- Suppose that
 - The population is 60% male and 40% female.
 - Our sample is 50% male and 50% female.
 - Our variable has an average value of 10 for the males and 15 for the females.
- What is our best estimate for the variable's average for the population?

Estimating Parameters

Stratified and
Cluster
Sampling

Robb T.
Koether

Introduction

Stratified
Random
Samples
Example

Estimating
Parameters

Cluster
Samples
Example

Stratified vs.
Cluster

Assignment

- We need to compute a *weighted average*.

$$\begin{aligned}\text{average} &= (0.60)(10) + (0.40)(15) \\ &= 6 + 6 \\ &= 12\end{aligned}$$

Cluster Sampling

Stratified and
Cluster
Sampling

Robb T.
Koether

Introduction

Stratified
Random
Samples

Example

Estimating
Parameters

Cluster
Samples

Example

Stratified vs.
Cluster

Assignment

Definition (Heterogeneous)

A group is **heterogeneous** if its members vary in regard to the variables of interest in the same way that the population varies.

Definition (Cluster)

A **cluster** is a heterogeneous subset of the population.

Definition (Cluster random sampling)

Cluster random sampling is a sampling method in which the population is first divided into clusters. Then a simple random sample of clusters is taken. All the members of the selected clusters together constitute the sample.

Cluster Sampling

Stratified and Cluster Sampling

Robb T.
Koether

Introduction

Stratified Random Samples Example

Estimating Parameters

Cluster Samples

Example

Stratified vs. Cluster

Assignment

- Note that it is the clusters that are selected at random, not the individuals.
- It is hoped that each cluster by itself is representative of the population, i.e., each cluster is heterogeneous.

Cluster Random Sampling

Stratified and Cluster Sampling

Robb T.
Koether

Introduction

Stratified
Random
Samples

Example

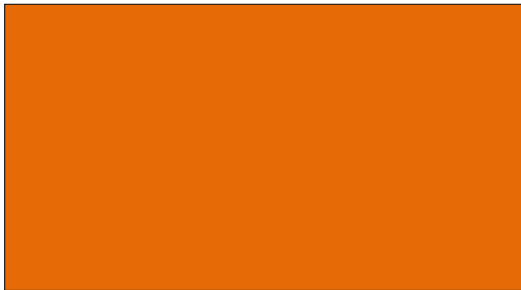
Estimating
Parameters

**Cluster
Samples**

Example

Stratified vs.
Cluster

Assignment



The population

Cluster Random Sampling

Stratified and
Cluster
Sampling

Robb T.
Koether

Introduction

Stratified
Random
Samples

Example

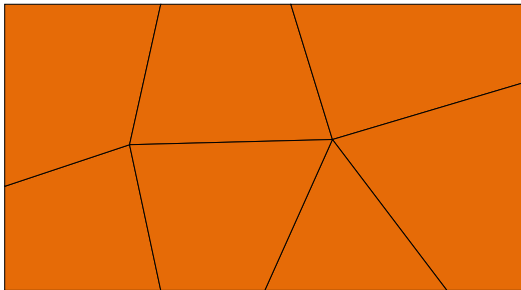
Estimating
Parameters

**Cluster
Samples**

Example

Stratified vs.
Cluster

Assignment



The clusters

Cluster Random Sampling

Stratified and Cluster Sampling

Robb T. Koether

Introduction

Stratified Random Samples

Example

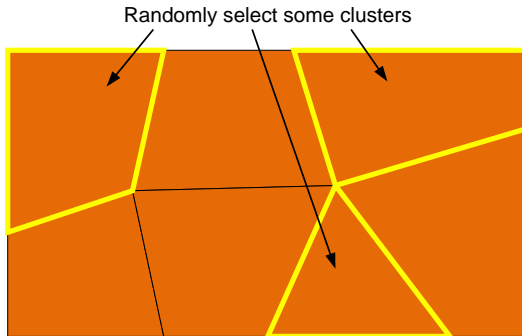
Estimating Parameters

Cluster Samples

Example

Stratified vs. Cluster

Assignment



The clusters

Cluster Random Sampling

Stratified and Cluster Sampling

Robb T. Koether

Introduction

Stratified Random Samples

Example

Estimating Parameters

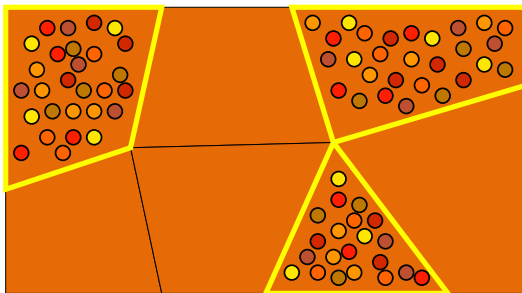
Cluster Samples

Example

Stratified vs. Cluster

Assignment

Selected all member in the clusters



The clusters

Cluster Random Sampling

Stratified and Cluster Sampling

Robb T. Koether

Introduction

Stratified Random Samples

Example

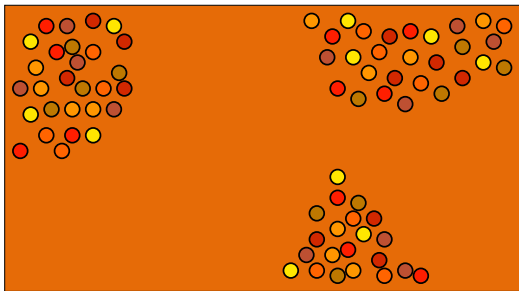
Estimating Parameters

Cluster Samples

Example

Stratified vs. Cluster

Assignment



The cluster sample

Example

Stratified and
Cluster
Sampling

Robb T.
Koether

Introduction

Stratified
Random
Samples
Example

Estimating
Parameters

Cluster
Samples
Example

Stratified vs.
Cluster

Assignment

Example (Cluster sample)

- Now suppose that
 - Anthony, Benjamin, Isabella, and Jasmine live in Fredericksburg.
 - Christopher, Daniel, Kayla, and Lily live in Richmond.
 - Ethan, Francisco, Madison, and Natalie live in Charlottesville.
 - Gabriel, Hunter, Olivia, and Paige live in Roanoke.

Example

Stratified and
Cluster
Sampling

Robb T.
Koether

Introduction

Stratified
Random
Samples

Example

Estimating
Parameters

Cluster
Samples
Example

Stratified vs.
Cluster

Assignment

Example (Cluster sample)

- Use cluster sampling to choose a sample of size $n = 8$, where the clusters are the cities.
- Is the sample representative with regard to sex?

Stratified Sampling vs. Cluster Sampling

Stratified and Cluster Sampling

Robb T.
Koether

Introduction

Stratified Random Samples

Example

Estimating Parameters

Cluster Samples

Example

Stratified vs. Cluster

Assignment

- In stratified sampling
 - From **all of the strata** we take **randomly selected individuals**.
- In cluster sampling
 - From **randomly selected clusters** we take **all of the individuals**.

Assignment

Stratified and
Cluster
Sampling

Robb T.
Koether

Introduction

Stratified
Random
Samples
Example

Estimating
Parameters

Cluster
Samples
Example

Stratified vs.
Cluster

Assignment

Homework

- Read Sections 2.6, 2.8, pages 108 - 115, 122 - 126.
- Let's Do It! 2.6, 2.8.
- Page 115, exercises 19 - 23, 25.
- Page 126, exercises 35 - 38.