

Stratified and Cluster Sampling

Lecture 8 Sections 2.6, 2.8

Robb T. Koether

Hampden-Sydney College

Tue, Sep 9, 2008

Outline

Stratified and
Cluster
Sampling

Robb T.
Koether

Homework
Review

Introduction

Stratified
Random
Samples

Example

Estimating
Parameters

Cluster
Samples

Example

Stratified vs.
Cluster

Assignment

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

Homework Review

Stratified and
Cluster
Sampling

Robb T.
Koether

Homework
Review

Introduction

Stratified
Random
Samples

Example

Estimating
Parameters

Cluster
Samples

Example

Stratified vs.
Cluster

Assignment

Exercise 1.17(a, b), p. 107

In the past, about 10% of customers were dissatisfied with the service received from local automobile dealers. This year, the manufacturer has reason to believe that the percentage of dissatisfied customers has decreased. He decides to test this theory.

- (a) State the competing hypotheses H_0 and H_1 clearly.
- (b) Out of a total of 40,000 customers, he decides to select a simple random sample of size 120. Use your calculator with a seed value of 32 to give the labels of the first four selected customers.

Homework Review

Stratified and
Cluster
Sampling

Robb T.
Koether

Homework
Review

Introduction

Stratified
Random
Samples
Example

Estimating
Parameters

Cluster
Samples
Example

Stratified vs.
Cluster

Assignment

Solution (a, b)

- (a)
- H_0 : The percentage of dissatisfied customers is the same as in the past.
 - H_1 : The percentage of dissatisfied customers is less than in the past.
- (b) The population size is $N = 40,000$ and the sample size is $n = 120$. So we must choose 120 random numbers from 1 to 40,000. First enter `32` → `rand` to set the seed. Then enter `randInt(1, 40000)` in the TI-83 and press ENTER 5 times. The numbers

34318, 15553, 8461, 614, 38964

appear.

Homework Review

Stratified and
Cluster
Sampling

Robb T.
Koether

Homework
Review

Introduction

Stratified
Random
Samples

Example

Estimating
Parameters

Cluster
Samples

Example

Stratified vs.
Cluster

Assignment

Exercise 1.17(c, d), p. 107

- (c) The data were found to be not statistically significant at $\alpha = 0.10$ level. What type of error could have been made?
- (d) Give a possible p -value for this test.

Solution (c, d)

- (c) “Not statistically significant” means that the null hypothesis was accepted. Therefore, it is possible that a Type II error was made.
- (d) Since the null hypothesis was accepted, the p -value must have been greater than α . So a possible p -value is 0.11. Another possible value is 0.12.

Introduction

Stratified and Cluster Sampling

Robb T. Koether

Homework Review

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

Homework Review

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

Homework
Review

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

Homework
Review

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

Homework
Review

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

Homework
Review

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

Homework
Review

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

Homework
Review

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.

Stratified Random Sample

Stratified and
Cluster
Sampling

Robb T.
Koether

Homework
Review

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

Homework Review

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

Homework
Review

Introduction

**Stratified
Random
Samples**

Example

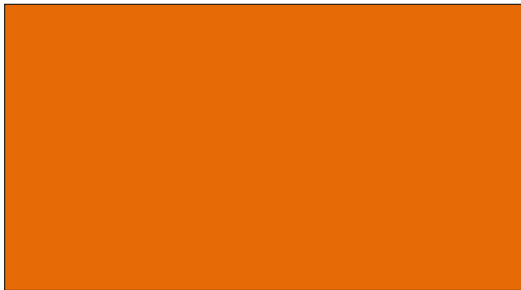
Estimating
Parameters

Cluster
Samples

Example

Stratified vs.
Cluster

Assignment



The population

Stratified Random Sampling

Stratified and
Cluster
Sampling

Robb T.
Koether

Homework
Review

Introduction

**Stratified
Random
Samples**

Example

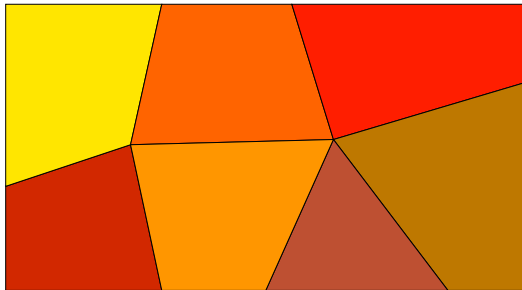
Estimating
Parameters

Cluster
Samples

Example

Stratified vs.
Cluster

Assignment



The strata

Stratified Random Sampling

Stratified and
Cluster
Sampling

Robb T.
Koether

Homework
Review

Introduction

**Stratified
Random
Samples**

Example

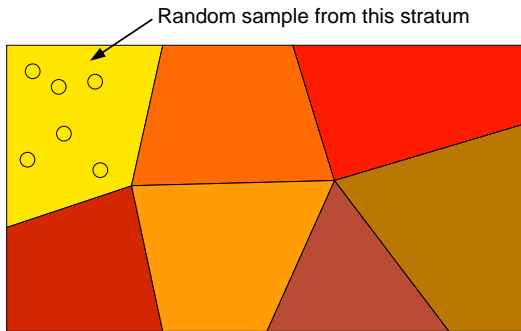
Estimating
Parameters

Cluster
Samples

Example

Stratified vs.
Cluster

Assignment



The strata

Stratified Random Sampling

Stratified and
Cluster
Sampling

Robb T.
Koether

Homework
Review

Introduction

**Stratified
Random
Samples**

Example

Estimating
Parameters

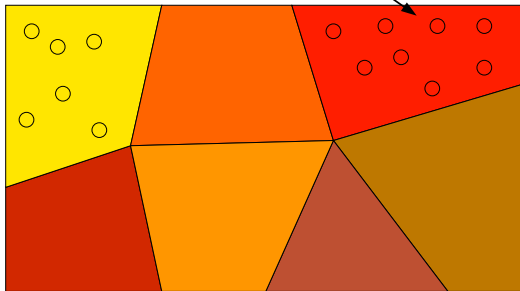
Cluster
Samples

Example

Stratified vs.
Cluster

Assignment

Random sample from this stratum



The strata

Stratified Random Sampling

Stratified and
Cluster
Sampling

Robb T.
Koether

Homework
Review

Introduction

**Stratified
Random
Samples**

Example

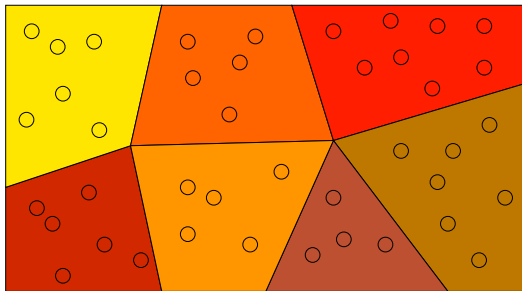
Estimating
Parameters

Cluster
Samples

Example

Stratified vs.
Cluster

Assignment



Random samples from all strata

Stratified Random Sampling

Stratified and
Cluster
Sampling

Robb T.
Koether

Homework
Review

Introduction

**Stratified
Random
Samples**

Example

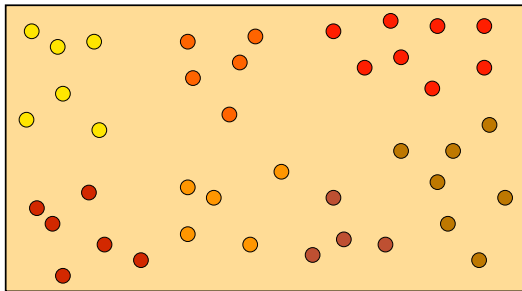
Estimating
Parameters

Cluster
Samples

Example

Stratified vs.
Cluster

Assignment



The stratified sample

Example

Stratified and
Cluster
Sampling

Robb T.
Koether

Homework
Review

Introduction

Stratified
Random
Samples
Example

Estimating
Parameters

Cluster
Samples
Example

Stratified vs.
Cluster

Assignment

Example (Stratified random sample)

- Let the population consist of males Andy, Bob, Charlie, Don, Ed, Fred, Greg, and Hank and females Pam, Queeny, Rachel, Susie, Terri, Ursula, Valerie, and Wendy.
- 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

Homework
Review

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

Homework
Review

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

Homework
Review

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

Homework
Review

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

Homework
Review

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

Homework
Review

Introduction

Stratified
Random
Samples

Example

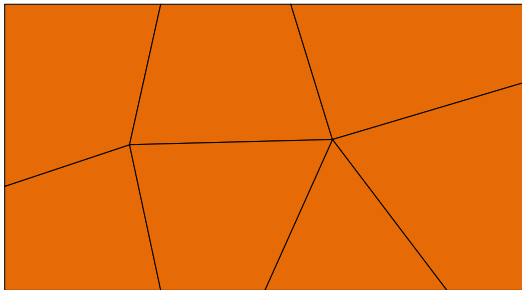
Estimating
Parameters

**Cluster
Samples**

Example

Stratified vs.
Cluster

Assignment



The clusters

Cluster Random Sampling

Stratified and Cluster Sampling

Robb T. Koether

Homework Review

Introduction

Stratified Random Samples

Example

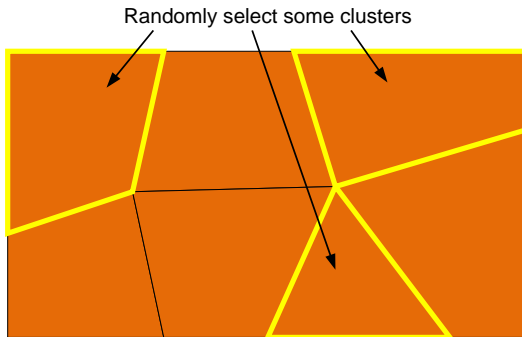
Estimating Parameters

Cluster Samples

Example

Stratified vs. Cluster

Assignment



The clusters

Cluster Random Sampling

Stratified and Cluster Sampling

Robb T. Koether

Homework Review

Introduction

Stratified Random Samples

Example

Estimating Parameters

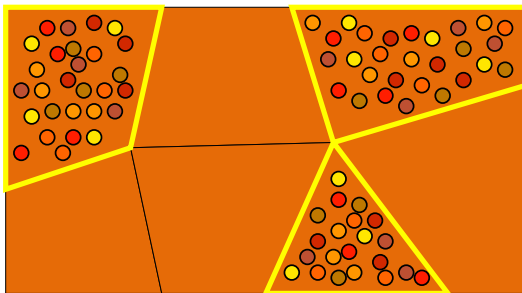
Cluster Samples

Example

Stratified vs. Cluster

Assignment

Selected all member in the clusters



The clusters

Cluster Random Sampling

Stratified and
Cluster
Sampling

Robb T.
Koether

Homework
Review

Introduction

Stratified
Random
Samples

Example

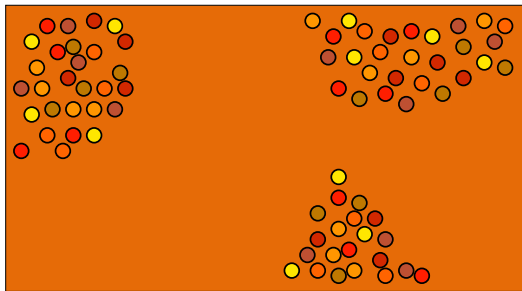
Estimating
Parameters

**Cluster
Samples**

Example

Stratified vs.
Cluster

Assignment



The cluster sample

Example

Stratified and
Cluster
Sampling

Robb T.
Koether

Homework
Review

Introduction

Stratified
Random
Samples
Example

Estimating
Parameters

Cluster
Samples
Example

Stratified vs.
Cluster

Assignment

Example (Cluster sample)

- Now suppose that
 - Andy, Bob, Pam, and Queeny live in Fredericksburg.
 - Charlie, Don, Rachel, and Susie live in Richmond.
 - Ed, Fred, Terri, and Uzi live in Charlottesville.
 - Greg, Hank, Valerie, and Wendy live in Roanoke.

Example

Stratified and
Cluster
Sampling

Robb T.
Koether

Homework
Review

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

Homework
Review

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

Homework
Review

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.