Stratified and Cluster Sampling

Lecture 8

Sections 2.6, 2.8

Robb T. Koether

Hampden-Sydney College

Tue, Jan 31, 2012

Outline

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

Outline

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

Introduction

- Suppose we want to measure support for the president in Virginia.

Introduction

- Suppose we want to measure support for the president in Virginia.
- Suppose further that we know that the Virginia population is 36% Democrat, 24% Republican, and 40% Independent.

Introduction

- Suppose we want to measure support for the president in Virginia.
- Suppose further that we know that the Virginia population is 36% Democrat, 24% Republican, and 40% Independent.
- Suppose even further that we suspect that party affiliation is a relevant variable.

Introduction

- Suppose we want to measure support for the president in Virginia.
- Suppose further that we know that the Virginia population is 36% Democrat, 24% Republican, and 40% Independent.
- Suppose even further that we suspect that party affiliation is a relevant variable.
- We plan to take a sample of 100 individuals.

Introduction

- Suppose we want to measure support for the president in Virginia.
- Suppose further that we know that the Virginia population is 36% Democrat, 24% Republican, and 40% Independent.
- Suppose even further that we suspect that party affiliation is a relevant variable.
- We plan to take a sample of 100 individuals.
- What might go wrong if we take a simple random sample?

Introduction

- Suppose we want to measure support for the president in Virginia.
- Suppose further that we know that the Virginia population is 36% Democrat, 24% Republican, and 40% Independent.
- Suppose even further that we suspect that party affiliation is a relevant variable.
- We plan to take a sample of 100 individuals.
- What might go wrong if we take a simple random sample?
- How can we be assured that each party affiliation will be sufficiently represented in our sample?

Introduction

- We could choose 36 Democrats, 24 Republicans, and 40 Independents.

Introduction

- We could choose 36 Democrats, 24 Republicans, and 40 Independents.
- What if we chose to survey 25 Democrats, 25 Republicans, and 50 Independents?

Introduction

- We could choose 36 Democrats, 24 Republicans, and 40 Independents.
- What if we chose to survey 25 Democrats, 25 Republicans, and 50 Independents?
- Could we accurately estimate the population proportion?

Outline

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

Stratified Random Sample

Definition (Homogeneous)

A group is homogeneous if its member all have similar characteristics with regard to a variable 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 combined results constitute the sample.

Examples

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

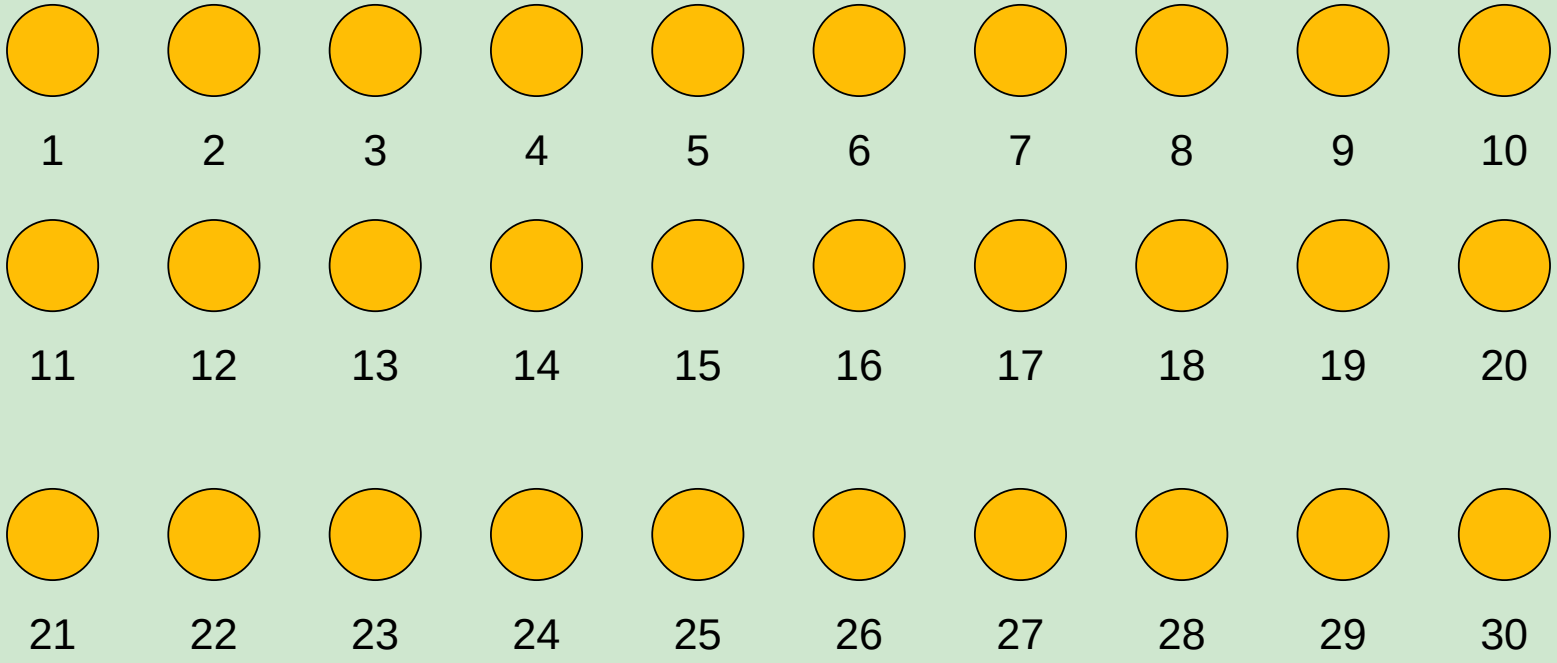
An Example

Example (Sunday Hunting)

- Virginia is currently considering a law to allow Sunday hunting. (Sunday hunting is currently illegal.)
- Suppose we have a population of 30 students, some hunters and some not hunters.
- We wish to survey them to estimate overall support for Sunday hunting.
- We will take a sample of 10 students.

An Example

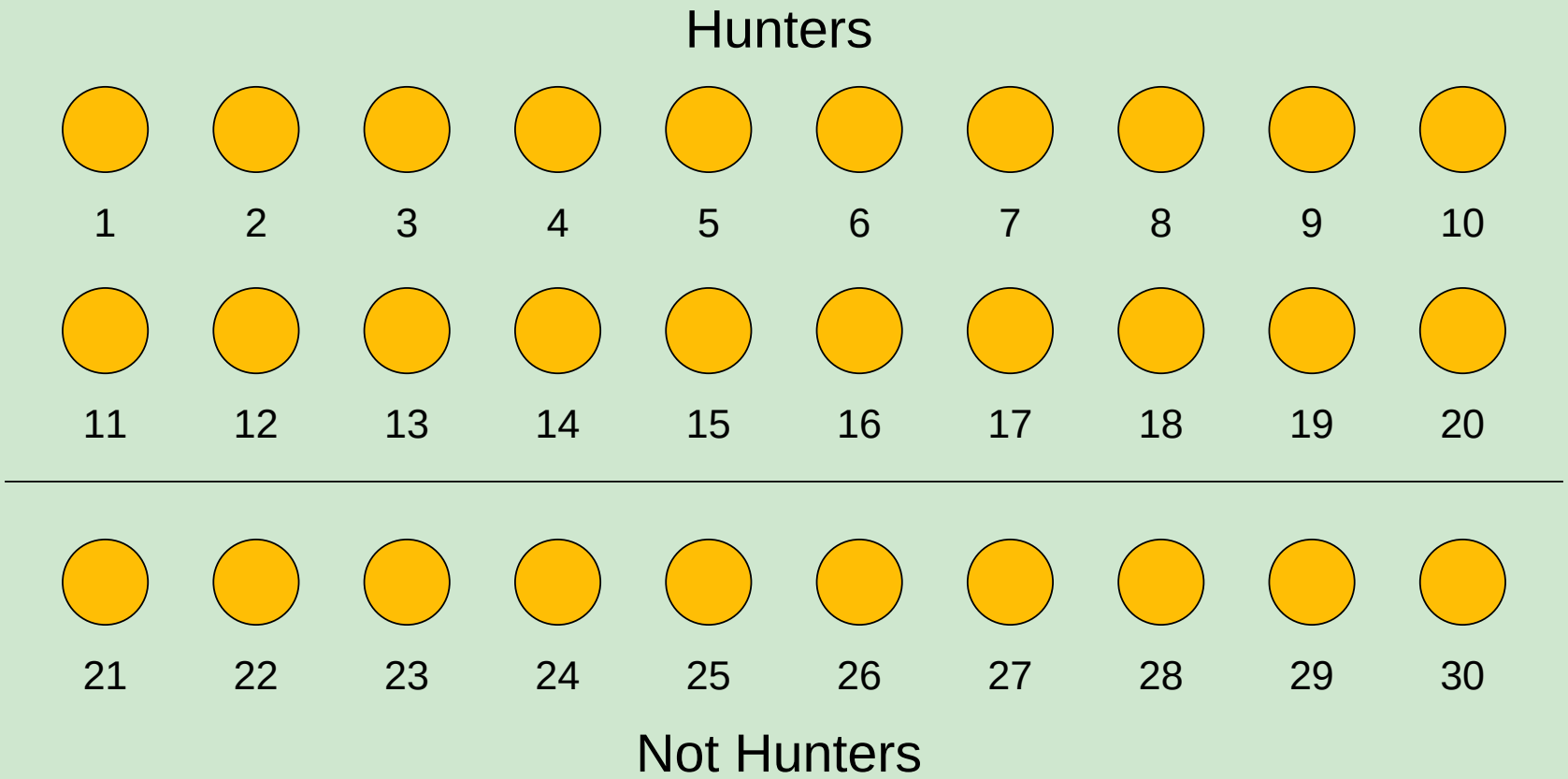
Example (Sunday Hunting)



The population of 30 students

An Example

Example (Sunday Hunting)



20 hunters and 10 non hunters

An Example

Example (Sunday Hunting)

Hunters

N	Y	Y	Y	Y	Y	Y	Y	N	Y
1	2	3	4	5	6	7	8	9	10
Y	Y	Y	N	Y	Y	Y	Y	Y	Y
11	12	13	14	15	16	17	18	19	20

N	Y	N	N	N	N	Y	N	Y	N
21	22	23	24	25	26	27	28	29	30

Not Hunters

Their opinions about Sunday hunting

An Example

Example (Sunday Hunting)

Hunters

N	Y	Y	Y	Y	Y	Y	Y	N	Y
1	2	3	4	5	6	7	8	9	10
Y	Y	Y	N	Y	Y	Y	Y	Y	Y
11	12	13	14	15	16	17	18	19	20

N	Y	N	N	N	N	Y	N	Y	N
21	22	23	24	25	26	27	28	29	30

Not Hunters

Take a simple random sample of 10 from the population

Outline

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

Estimating Parameters

Example (Sunday Hunting)

- What sample proportion did we observe?

Estimating Parameters

Example (Sunday Hunting)

- What sample proportion did we observe?
- How does it compare to the population proportion?

Estimating Parameters

Example (Sunday Hunting)

- What sample proportion did we observe?
- How does it compare to the population proportion?
- Were hunters and non-hunters properly represented?

An Example

Example (Sunday Hunting)

Hunters

N	Y	Y	Y	Y	Y	Y	Y	N	Y
1	2	3	4	5	6	7	8	9	10
Y	Y	Y	N	Y	Y	Y	Y	Y	Y
11	12	13	14	15	16	17	18	19	20

N	Y	N	N	N	N	Y	N	Y	N
21	22	23	24	25	26	27	28	29	30

Not Hunters

The population of 30 students

An Example

Example (Sunday Hunting)

Hunters

N	Y	Y	Y	Y	Y	Y	Y	N	Y
1	2	3	4	5	6	7	8	9	10
Y	Y	Y	N	Y	Y	Y	Y	Y	Y
11	12	13	14	15	16	17	18	19	20

N	Y	N	N	N	N	Y	N	Y	N
1	2	3	4	5	6	7	8	9	10

Not Hunters

Take a stratified sample of 5 hunters and 5 non hunters

Estimating Parameters

Example (Sunday Hunting)

- What sample proportions did we observe in each stratum?

Estimating Parameters

Example (Sunday Hunting)

- What sample proportions did we observe in each stratum?
- What is the overall estimate?

Estimating Parameters

Example (Sunday Hunting)

- What sample proportions did we observe in each stratum?
- What is the overall estimate?
- How does it compare to the population proportion?

Estimating Parameters

Example (Sunday Hunting)

- What sample proportions did we observe in each stratum?
- What is the overall estimate?
- How does it compare to the population proportion?
- Were hunters and non-hunters properly represented?

Outline

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

Cluster Sampling

Definition (Heterogeneous)

A group is heterogeneous if is 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

- 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.

Estimating Parameters

Example (Sunday Hunting)

- Consider again the population of 30 students, some hunters and some non-hunters.

An Example

Example (Sunday Hunting)

Hunters

N	Y	Y	Y	Y	Y	Y	Y	N	Y
1	2	3	4	5	6	7	8	9	10
Y	Y	Y	N	Y	Y	Y	Y	Y	Y
11	12	13	14	15	16	17	18	19	20

N	Y	N	N	N	N	Y	N	Y	N
21	22	23	24	25	26	27	28	29	30

Not Hunters

The population of 30 students

An Example

Example (Sunday Hunting)

Norfolk		Richmond					No. VA		
N	Y	Y	Y	Y	Y	Y	Y	N	Y
1	2	1	2	3	4	5	1	2	3
Y	Y	Y	N	Y	Y	Y	Y	Y	Y
3	4	6	7	8	9	10	4	5	6
N	Y	N	N	N	N	Y	N	Y	N
5	6	11	12	13	14	15	7	8	9

Create clusters based on place of residence

An Example

Example (Sunday Hunting)

Norfolk		Richmond					No. VA		
N	Y	Y	Y	Y	Y	Y	Y	N	Y
1	2	1	2	3	4	5	1	2	3
Y	Y	Y	N	Y	Y	Y	Y	Y	Y
3	4	6	7	8	9	10	4	5	6
N	Y	N	N	N	N	Y	N	Y	N
5	6	11	12	13	14	15	7	8	9

Take a cluster sample of 2 clusters

Estimating Parameters

Example (Sunday Hunting)

- What sample proportions did we observe in each cluster?

Estimating Parameters

Example (Sunday Hunting)

- What sample proportions did we observe in each cluster?
- What is the overall estimate?

Estimating Parameters

Example (Sunday Hunting)

- What sample proportions did we observe in each cluster?
- What is the overall estimate?
- How does it compare to the population proportion?

Estimating Parameters

Example (Sunday Hunting)

- What sample proportions did we observe in each cluster?
- What is the overall estimate?
- How does it compare to the population proportion?
- What is the difference between these clusters and the strata of the previous example?

Outline

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

Stratified Sampling vs. Cluster Sampling

- 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.

Outline

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

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.