PLSC 308: Introduction to Political Research

Christopher Zorn

February 25, 2016

The Margin of Error (MOE)

Sampling error is the (random) difference between the value you want to know in the population and its respective value in the sample.

Characteristics:

- Intuition: "Repeated samples"
- A function of:
    - The sample size,
    - The sampling design, and
    - The size of the population.

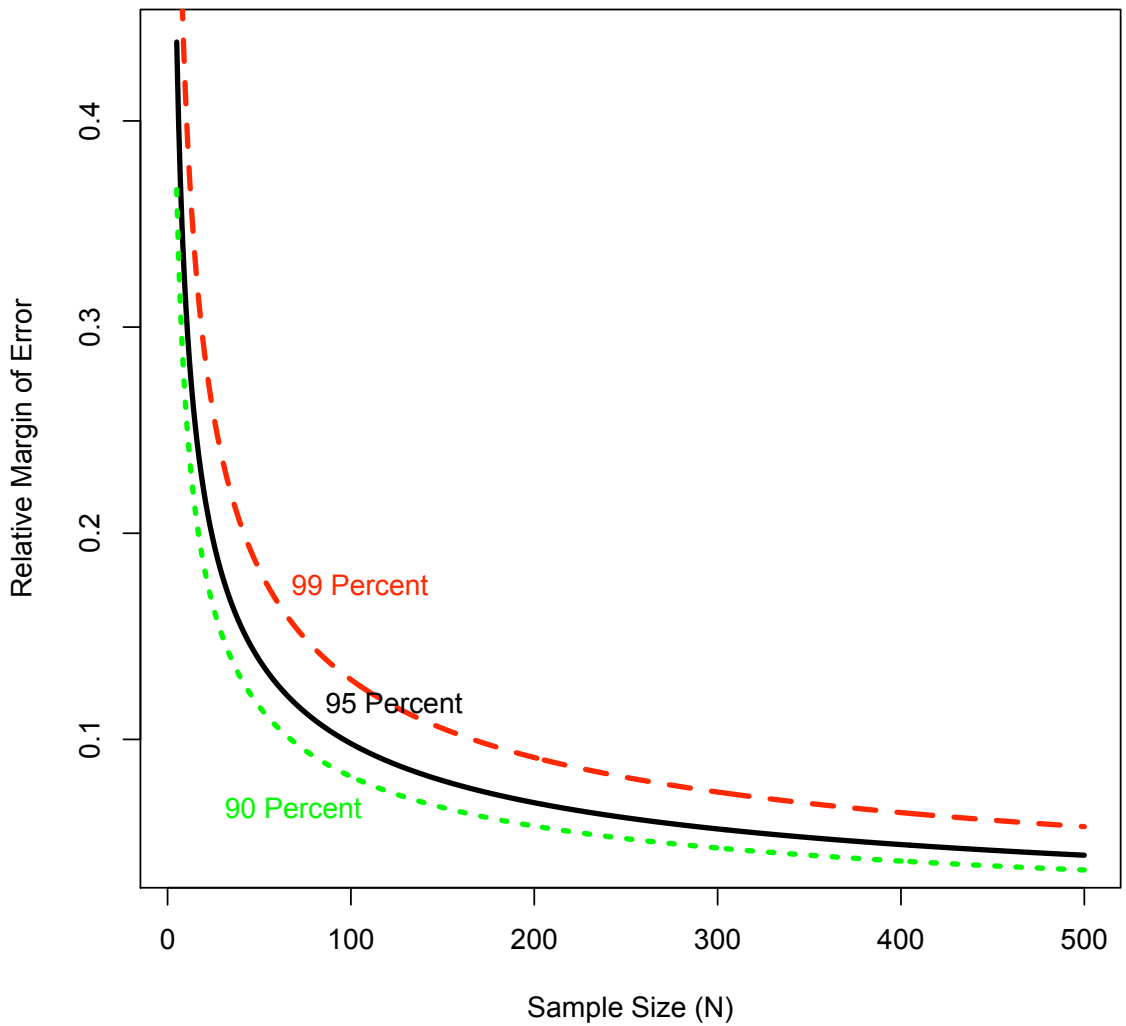
MOE Example

Consider the proportion Q of observations in the population that have some (binary) trait. For a simple random sample of size N, the margin of error (sampling error) for the sample proportion q is:

$$\text{Standard error} = \sqrt{\frac{q(1-q)}{N}}$$

We typically calculate relative sampling error for a given *level of confidence*...

MOE and Sample Size



XXX

Sampling: An Example

	Name	sex	namegroup	score
1	ARNOLD, EMILY	F	AB	31
2	BERNICKER, BRENDAN	M	AB	62
3	CHEN, TIANGE	M	CD	61
4	DANIEL, ASHLEY	F	CD	82
5	KLARIC, ROMAN	M	KL	86
6	KLAVANS, JORDAN	M	KL	64
7	KNABB, CHRISTOPHER	M	KL	1
8	LAMAS, MATTHEW	M	KL	23
9	LEVIN, LOUIS	M	KL	67
10	LI, ZIYU	M	KL	51
11	MCCABE, MICHAEL	M	MN	69
12	MCDONNELL, HANNAH	F	MN	74
13	NAHVI, SUSAN	F	MN	48
14	PARIKH, YASH	M	OP	92
15	PELOSI, OLIVER	M	OP	29
16	RANKIN-WAHLERS, JADEN	F	QR	104
17	SALVESEN, BRIAN	M	ST	29
18	SHAFFER, JOHN	M	ST	27
19	SHIPMAN, MATTHEW	M	ST	19
20	THOMAS, CORY	M	ST	23
21	WAGNER, SEAMUS	M	WX	32
22	WESTERFER, JOSEPH	M	WX	30

Sampling Example (continued)

Population characteristics:

- **Overall mean = 50.2** ($N = 22$)
- Mean for males $= 45$ ($N_M = 17$)
- Mean for females $= 67.8$ ($N_F = 5$)
- No differences by last name

Simple Random Sampling via Random Numbers

	Name	sex	namegroup	score	rando
2	BERNICKER, BRENDAN	M	AB	62	0.03999592
8	LAMAS, MATTHEW	M	KL	23	0.04577026
1	ARNOLD, EMILY	F	AB	31	0.15904600
13	NAHVI, SUSAN	F	MN	48	0.18109621
15	PELOSI, OLIVER	M	OP	29	0.20124804
3	CHEN, TIANGE	M	CD	61	0.21879954
16	RANKIN-WAHLERS, JADEN	F	QR	104	0.25880982
10	LI, ZIYU	M	KL	51	0.26518667
11	MCCABE, MICHAEL	M	MN	69	0.30467220
21	WAGNER, SEAMUS	M	WX	32	0.31182431
9	LEVIN, LOUIS	M	KL	67	0.45609148
12	MCDONNELL, HANNAH	F	MN	74	0.50730687
5	KLARIC, ROMAN	M	KL	86	0.52569755
19	SHIPMAN, MATTHEW	M	ST	19	0.55333359
22	WESTERFER, JOSEPH	M	WX	30	0.62181920
20	THOMAS, CORY	M	ST	23	0.64640609
14	PARIKH, YASH	M	OP	92	0.75967064
18	SHAFFER, JOHN	M	ST	27	0.80735234
4	DANIEL, ASHLEY	F	CD	82	0.81059855
7	KNABB, CHRISTOPHER	M	KL	1	0.83134505
6	KLAVANS, JORDAN	M	KL	64	0.91465817
17	SALVESEN, BRIAN	M	ST	29	0.99215042

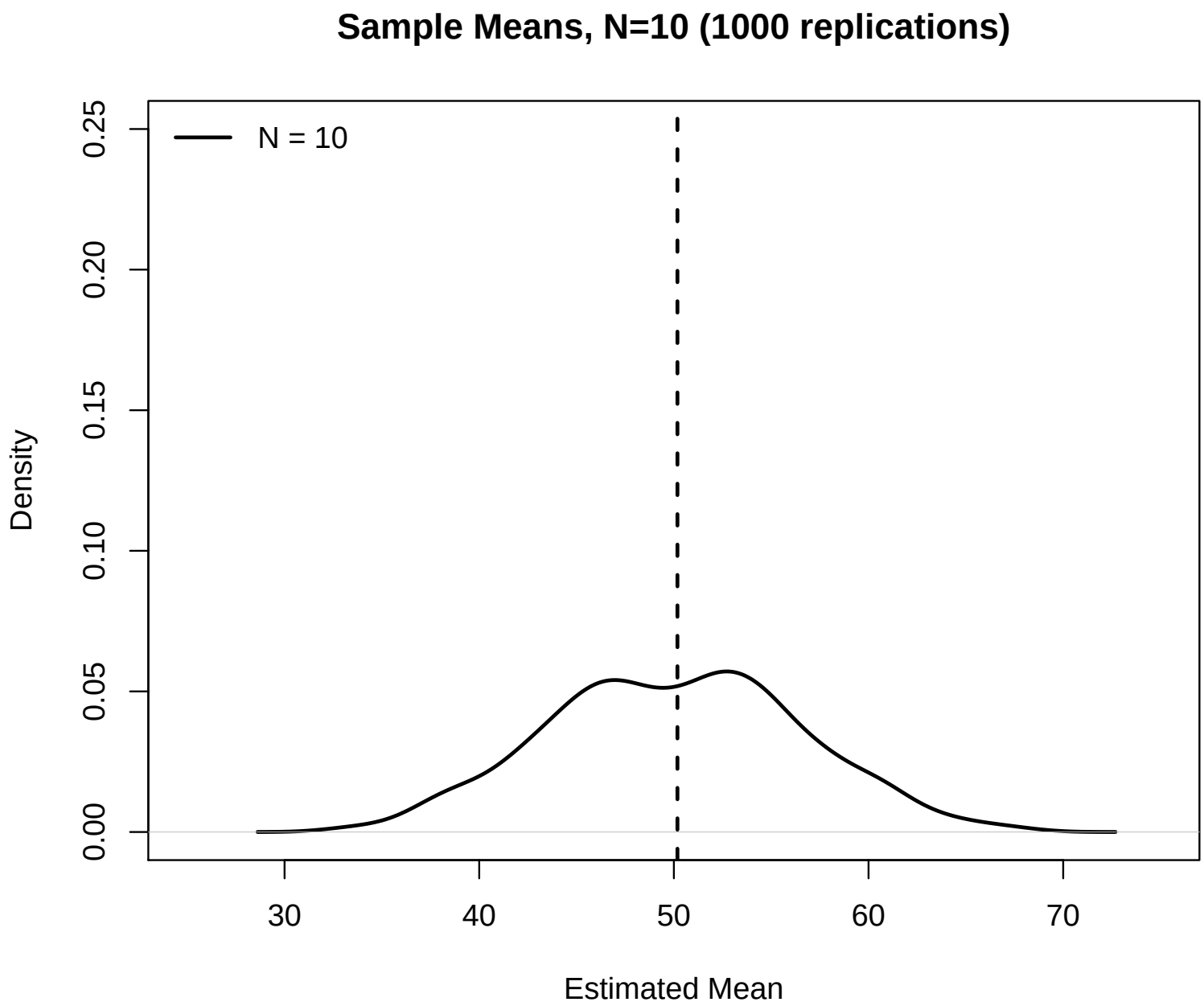
Simple Random Sample ($N = 10$)

Sample characteristics:

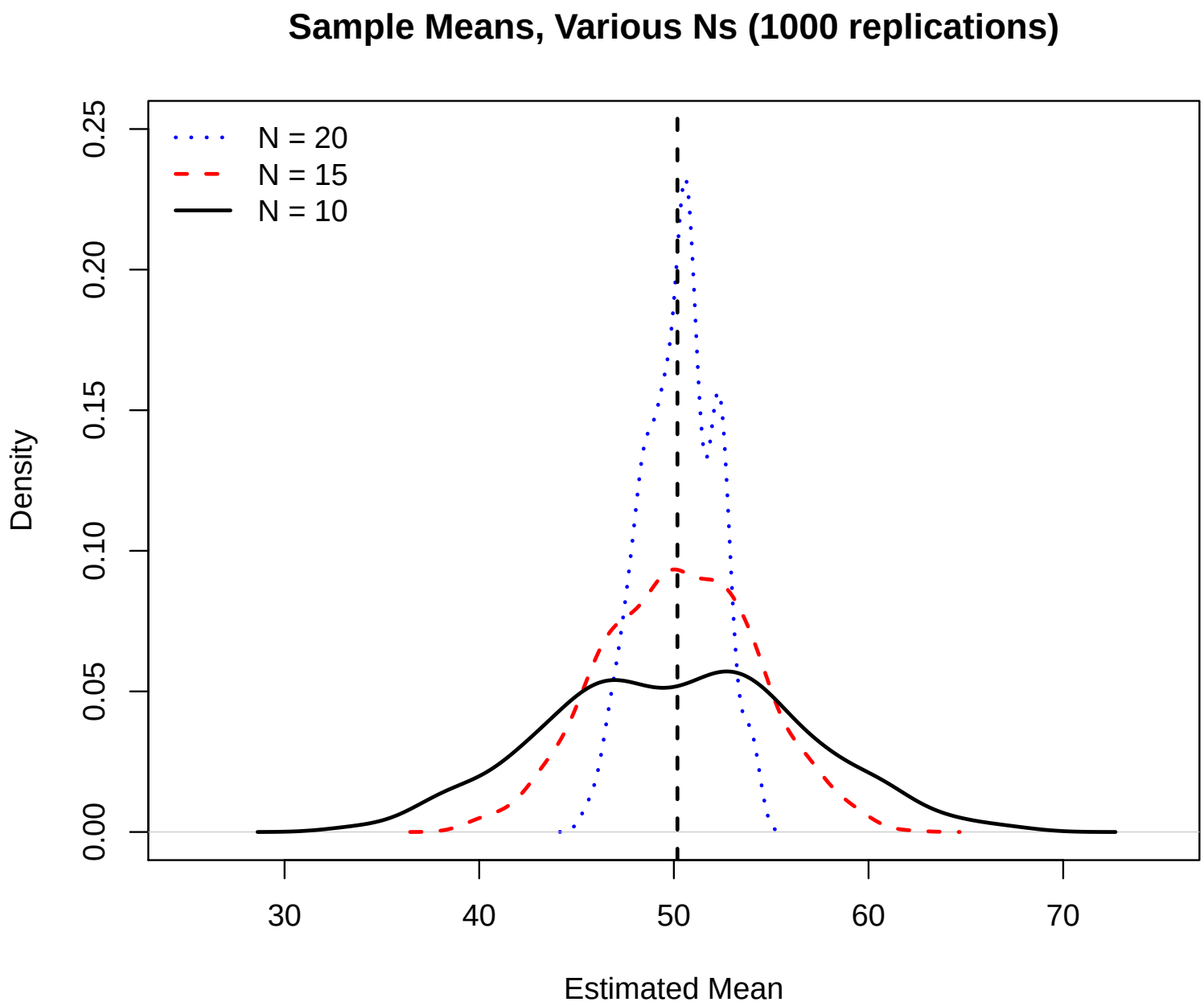
- **Overall mean = 51** ($N = 10$)
- Mean for males $= 46.7$ ($N_M = 7$)
- Mean for females $= 61$ ($N_F = 3$)
- No differences by last name

$$\text{Overall MOE} = \sqrt{\frac{0.51(1 - 0.51)}{10}} = 0.158 \text{ (15.8 percent)}$$

Simple Random Sampling, repeated



Vary Sample Sizes



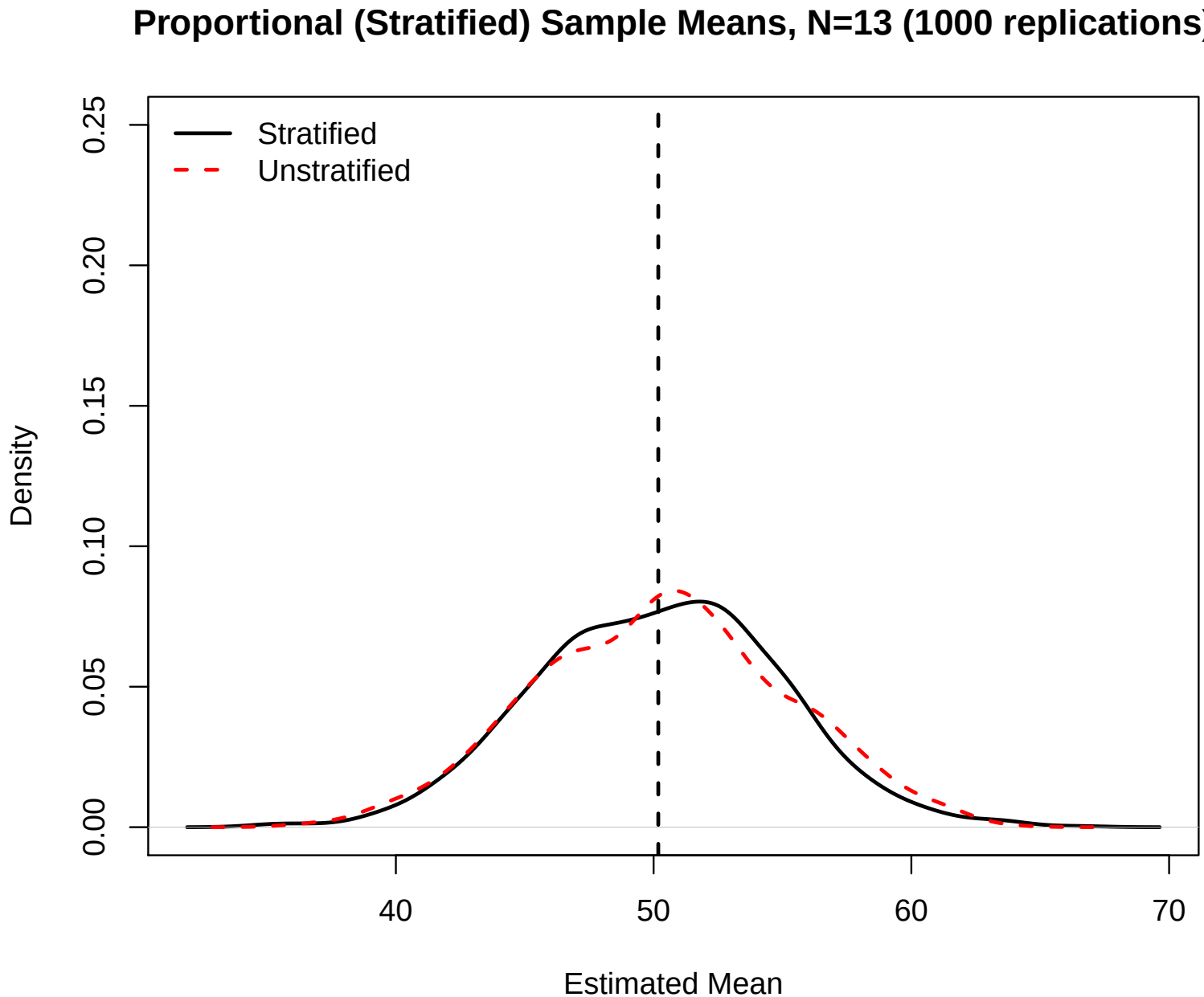
Stratified Sampling

Women are 22.7 percent of the population...

Proportional stratified sampling (say, select 3 women and 10 men randomly):

	Name	sex	namegroup	score	rando
1	ARNOLD, EMILY	F	AB	31	0.26093924
13	NAHVI, SUSAN	F	MN	48	0.79497190
16	RANKIN-WAHLERS, JADEN	F	QR	104	0.72192125
10	LI, ZIYU	M	KL	51	0.64051016
2	BERNICKER, BRENDAN	M	AB	62	0.30613168
3	CHEN, TIANGE	M	CD	61	0.43401093
5	KLARIC, ROMAN	M	KL	86	0.50771100
17	SALVESEN, BRIAN	M	ST	29	0.14978794
8	LAMAS, MATTHEW	M	KL	23	0.12362789
9	LEVIN, LOUIS	M	KL	67	0.62641954
6	KLAVANS, JORDAN	M	KL	64	0.05987933
21	WAGNER, SEAMUS	M	WX	32	0.08098174
22	WESTERFER, JOSEPH	M	WX	30	0.79409840

Proportional Stratification: Results

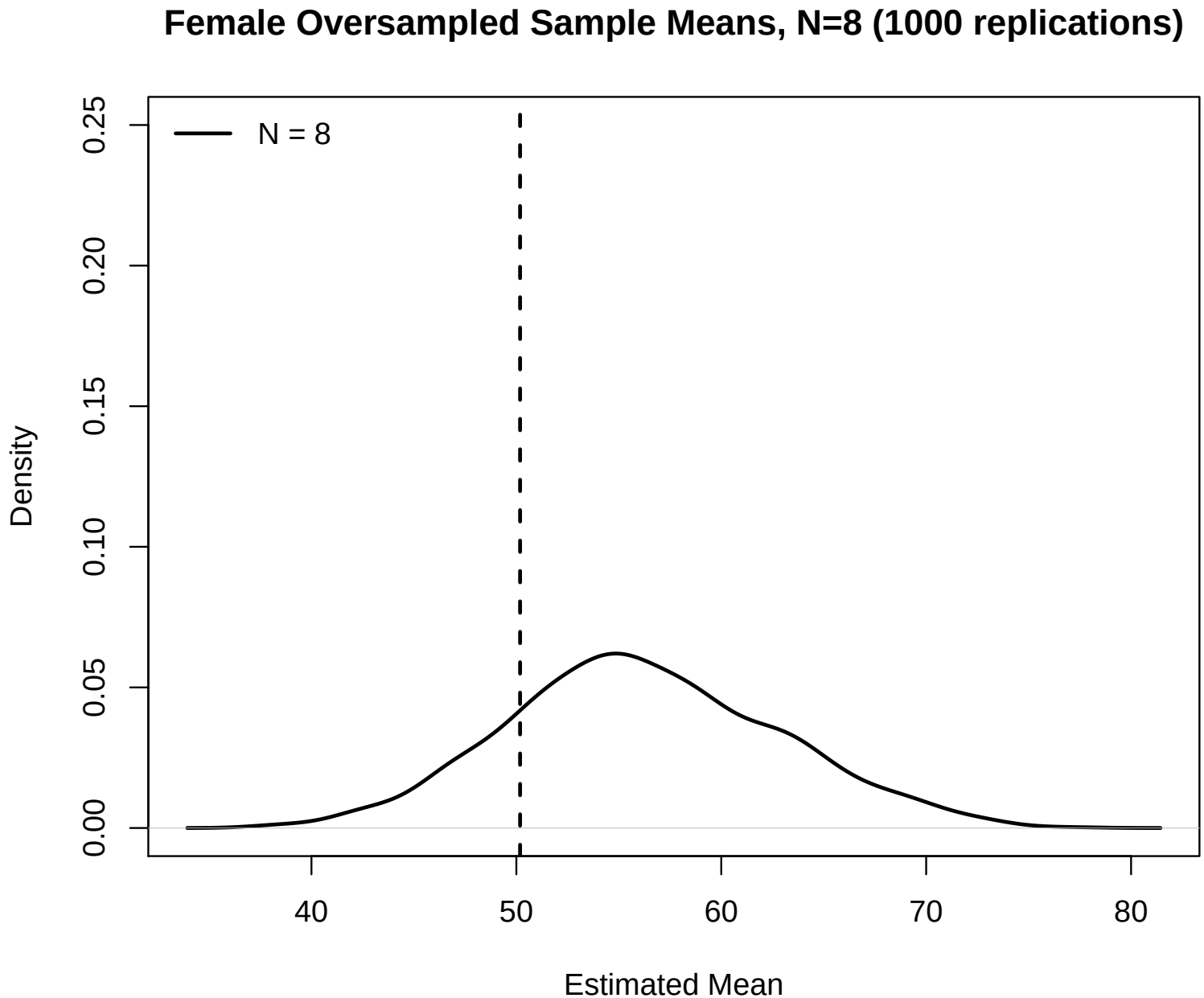


Stratified Sampling

Oversampling women (say, a 50-50 male-female split):

	Name	sex	namegroup	score	rando
1	ARNOLD, EMILY	F	AB	31	0.0939031
13	NAHVI, SUSAN	F	MN	48	0.2346178
16	RANKIN-WAHLERS, JADEN	F	QR	104	0.4263274
4	DANIEL, ASHLEY	F	CD	82	0.5336345
10	LI, ZIYU	M	KL	51	0.1426833
2	BERNICKER, BRENDAN	M	AB	62	0.1178203
3	CHEN, TIANGE	M	CD	61	0.5439105
5	KLARIC, ROMAN	M	KL	86	0.7212281

Oversampling: Results



Cluster Sampling

Sample on `namegroup` with $N_c = 3$

	Name	sex	namegroup	score
3	CHEN, TIANGE	M	CD	61
4	DANIEL, ASHLEY	F	CD	82
5	KLARIC, ROMAN	M	KL	86
6	KLAVANS, JORDAN	M	KL	64
7	KNABB, CHRISTOPHER	M	KL	1
8	LAMAS, MATTHEW	M	KL	23
9	LEVIN, LOUIS	M	KL	67
10	LI, ZIYU	M	KL	51
17	SALVESEN, BRIAN	M	ST	29
18	SHAFFER, JOHN	M	ST	27
19	SHIPMAN, MATTHEW	M	ST	19g
20	THOMAS, CORY	M	ST	23

Mean score = 44.1...

Midterm Matters...

- In-class (75 minutes), 3/1/2016
- Closed-book / closed-notes / etc.
- Materials:
  - Chapters 1 - 7 of the text
  - Class lecture material
- Format:
  - Definitions
  - Short answer questions
  - A short "essay"
- **Worth 20% of course grade (200 points)**