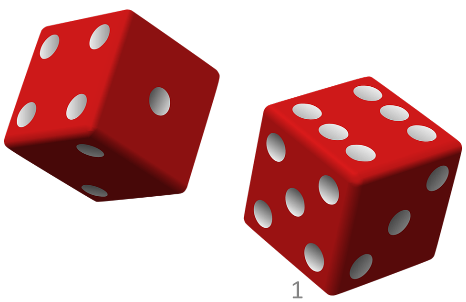
Sampling People, Records, & Networks

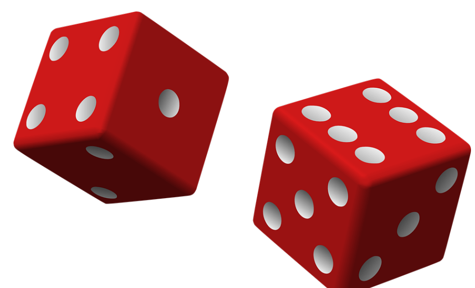
Jim Lepkowski, PhD
Professor & Research Professor *Emeritus*
Institute for Social Research, University of Michigan
Research Professor,
Joint Program in Survey Methodology, University of Maryland



Unit 2

- 1 Simple random sampling
- 2 History
- 3 Sampling distribution
- 4 Sample size
- 5 Margin of error
- 6 Sample & population size

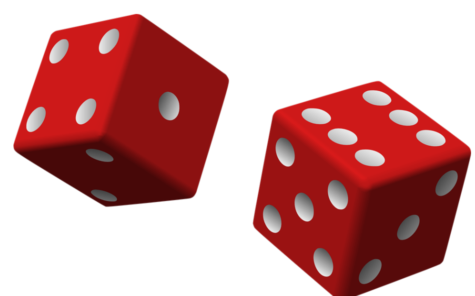
- Unit 1: Sampling as a research tool
- Unit 2: Mere randomization
  - Lecture 1: Simple Random Sampling (SRS)
  - Lecture 2: A short history
  - Lecture 3: The SRS sampling distribution
  - Lecture 4: Sample size
  - Lecture 5: Margin of error
  - Lecture 6: Sample size & population size
- Unit 3: Saving money
- Unit 4: Being more efficient
- Unit 5: Simplifying sampling
- Unit 6: Some extensions & applications



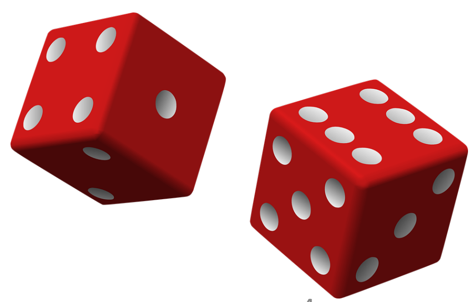
Unit 2

- 1 Simple random sampling
- 2 History
- 3 Sampling distribution
- 4 Sample size
- **5 Margin of error**
- 6 Sample & population size

- Unit 1: Sampling as a research tool
- Unit 2: Mere randomization
  - Lecture 1: Simple Random Sampling (SRS)
  - Lecture 2: A short history
  - Lecture 3: The SRS sampling distribution
  - Lecture 4: Sample size
  - **Lecture 5: Margin of error**
  - Lecture 6: Sample size & population size
- Unit 3: Saving money
- Unit 4: Being more efficient
- Unit 5: Simplifying sampling
- Unit 6: Some extensions & applications



- Using desired standard errors
- Margin of error
- An alternative formula



- **Using desired standard errors**
- Margin of error
- An alternative formula

- **Recall that we got the necessary sample size n' from**

$$n' = \frac{S^2}{V_d}$$

- **And the we could calculate the actual *n* needed for a population of a particular size by**

$$n = \frac{n'}{1 + \frac{n'}{N}}$$

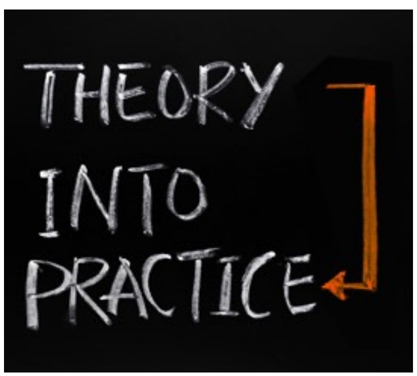
- **Using desired standard errors**
- Margin of error
- An alternative formula

- **The example developed a desired level of precision from the width of a confidence interval:**

$$\left(Lower\ limit, Upper\ limit \right) = \left(p - z \times se\left(p \right), p + z \times se\left(p \right) \right)$$

- **We set upper and lower limits for a 95% confidence interval, where *z* = 2 (approximately – 1.96 exactly for large samples):**

$$\left(Lower\ 95\%\ limit, Upper\ 95\%\ limit \right) = \left(p - 2 \times se\left(p \right), p + 2 \times se\left(p \right) \right)$$



- Using desired standard errors
- **Margin of error**
- An alternative formula

- **Suppose, like before, we want**

$$\left(Lower\,95\%\,limit, Upper\,95\%\,limit \right) = \left(0.58, 0.62 \right)$$

- **Then some will refer to the "margin of error" *e* as the distance from the upper limit to the middle, or the lower limit to the middle.**
  - **In most practice, "margin of error" is about proportions or percentages, as here.**
- **In some areas of application of probability sampling, this distance is referred to as the "precision"**

MARGIN OF ERR+/-R

- Using desired standard errors
- **Margin of error**
- An alternative formula

- **Suppose, like before, we want**

$$\left(Lower\,95\%\,limit, Upper\,95\%\,limit\right) = \left(0.58, 0.62\right)$$

- **Then the "margin of error" *e* is the distance from the upper limit to the middle, or the lower limit to the middle**
- **Calculate then**

$$e = 2 \times se\left(p\right) = \left(\frac{U - L}{2}\right)$$

MARGIN OF ERROR

- Using desired standard errors
- **Margin of error**
- An alternative formula

- **Suppose, like before, we want**

$$\left(Lower\,95\%\,limit, Upper\,95\%\,limit\right) = \left(0.58, 0.62\right)$$

- **Then the "margin of error" *e* is the distance from the upper limit to the middle, or the lower limit to the middle**
- **Calculate then**

$$e = 2 \times se\left(p\right) = \left(\frac{0.62 - 0.58}{2}\right) = 0.02$$

MARGIN OF ERR+/-R

 9

- Using desired standard errors
- Margin of error
- An alternative formula

- **In a newspaper report, you might see then the "margin of error" reported, but never the standard error …**

"President Obama's approval rating now stands at 60% (plus or minus 2%)"

- **The public has gotten used to forming the 95% confidence interval from this statement**

- **It's only one step to get the desired standard error and sampling variance:**

$$\sqrt{V_d} = \frac{e}{2} = \frac{0.02}{2} = 0.01$$

$$V_d = 0.0001$$

- Using desired standard errors
- Margin of error
- An alternative formula

- **But some trained are to use *e* directly in calculating sample size**
  - **You may see sample size formulas that are based on *e***
  - **These alternative formulas yield the same result as what we do here**
  - **But it can be confusing, especially if one has learned one way rather than the other**

 12

- Using desired standard errors
- Margin of error
- **An alternative formula**

- **The necessary sample size formula using *e* is**

$$n' = \frac{S^2}{\left(\frac{e}{2}\right)^2}$$

- **And this can be then 'adjusted' to obtain the final sample size as**

$$n = \frac{n'}{1 + \frac{n'}{N}}$$

13

- Using desired standard errors
- Margin of error
- **An alternative formula**

- **The necessary sample size formula using *e* is**

$$n' = \frac{S^2}{\left(\frac{e}{2}\right)^2}$$

- **And this can be then 'adjusted' to obtain the final sample size as**

$$n = \frac{n'}{1 + \frac{n'}{N}}$$

- Using desired standard errors
- Margin of error
- **An alternative formula**

- **The necessary sample size formula using *e* is**

$$n' = \frac{S^2}{\left(\frac{e}{2}\right)^2} \; or \; \frac{4S^2}{e^2}$$

- **And this can be then 'adjusted' to obtain the final sample size as**

$$n = \frac{n'}{1 + \frac{n'}{N}}$$

- Using desired standard errors
- Margin of error
- **An alternative formula**

- **The necessary sample size formula using *e* is**

$$n' = \frac{S^2}{\left(\frac{e}{2}\right)^2} \; or \; \frac{4S^2}{e^2} \; or \; \frac{z^2 S^2}{e^2}$$

- **And this can be then 'adjusted' to obtain the final sample size as**

$$n = \frac{n'}{1 + \frac{n'}{N}}$$

- Using desired standard errors
- Margin of error
- **An alternative formula**

- **The necessary sample size formula using *e* is**

$$n' = \frac{S^2}{\left(\frac{e}{2}\right)^2} \; or \; \frac{4S^2}{e^2} \; or \; \frac{z^2 S^2}{e^2} \; or \; \frac{z^2 p(1-p)}{e^2}$$

- **And this can be then 'adjusted' to obtain the final sample size as**

$$n = \frac{n'}{1 + \frac{n'}{N}}$$

- Using desired standard errors
- Margin of error
- **An alternative formula**

- **The necessary sample size formula using *e* is**

$$n' = \frac{S^2}{\left(\frac{e}{2}\right)^2} \ or \ \frac{4S^2}{e^2} \ or \ \frac{z^2 S^2}{e^2} \ or \ \frac{z^2 p(1-p)}{e^2} \ or \ \frac{z^2_{1-\alpha/2} p(1-p)}{e^2}$$

- **And this can be then 'adjusted' to obtain the final sample size as**

$$n = \frac{n'}{1 + \frac{n'}{N}}$$

 18

- Using desired standard errors
- Margin of error
- **An alternative formula**

- **And finally, the calculation can also be done in one step, rather than two:**

$$n = \frac{S^2}{\left(\frac{e}{2}\right)^2 + \frac{S^2}{N}}$$

19

- Using desired standard errors
- Margin of error
- **An alternative formula**

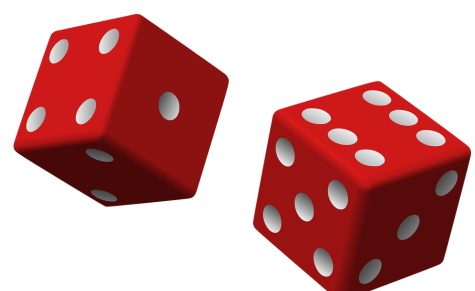
- **And finally, the calculation can also be done in one step, rather than two:**

$$n = \frac{S^2}{\left(\frac{e}{2}\right)^2 + \frac{S^2}{N}}$$

- **For our example, then, where *e* = 0.02,**

$$n = \frac{0.24}{\left(\frac{0.02}{2}\right)^2 + \frac{0.24}{250{,}000{,}000}} = 2{,}399.97 = 2{,}400$$

Unit 2



21