Chapter 5
Sampling Distribution of the Proportion

J.C. Wang

Western Michigan University
Department of Statistics

Goal and Objectives
of Chapter 5

To learn about the sampling distribution of the proportion

iClicker Question 5.1 Pre-lecture

iClicker Question 5.1 Pre-lecture

Outline

Sampling Distribution of the Proportion
- TV+More Example
- Theory
- Law of Large Numbers for Sample Proportions

Estimating the Population Proportion
- Questions Asked About Population Proportion
- Examples

Sampling Distribution of the Proportion

TV+More Example

- Suppose TV+More sells 60 extended warranties with 300 TV sets sold. The warranty sales rate is $\frac{60}{300} = 0.20$.
- Therefore, let X denote the number of successes out of a sample of n observations. Then X is a binomial random variable with parameters n and p.
- The proportion of successes, $\hat{p} = \frac{X}{n}$ in a sample is also a random variable.

Sampling Distribution

of the Proportion

- $\hat{p} = \frac{X}{n}$ = (number of successes) / (sample size)
- For the binomial, X is expected to be around np give or take $\sqrt{npq}$, where $q = 1 - p$.
- For the proportion, $\hat{p}$ is expected to be $p = \frac{np}{n}$ give or take

$$\sqrt{\frac{pq}{n}} = \frac{\sqrt{npq}}{n}$$

TV+More Example

revisited

- The number of warranties sold is expected to be around 60 ± 7
- The proportion of warranties sold is expected to be around

$$\frac{60}{300} \pm \frac{7}{300} \text{ or } 0.2 \pm 0.02.$$

Law of Large Numbers

for sample proportions

The sample proportion tends to get closer to the true proportion as sample size increases.
For TV+More Example:

- Recall if TV+More sold 300 TV sets then $\hat{p} = .2$ and $sd = 0.02$.
- If TV+More sold 1200 TV sets and $\hat{p} = .2$ and now

$$sd = \sqrt{\frac{0.2 \times (1 - 0.2)}{1200}} = 0.0115$$

Sampling Distribution of Sample Proportion

is approximately normal

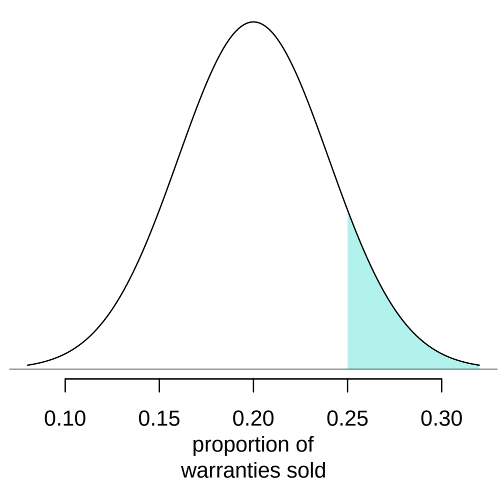
If TV+More sold 100 TV sets last year, the percentage of sets sold with extended warranties is expected to be around 20% give or take 4%.
Estimate the likelihood that more than 25% of TV sets sold last year had sold with extended warranties, in other words,

$$P(\hat{p} > 0.25) = ?$$

Sample Proportion is approximately normal

continued

Given: $n = 100$ and $p = .2$
$P(\hat{p} > .25) =$
$\text{normalCDF}(.25, 9999, .2, \sqrt{\frac{.2\times(1-.2)}{100}})$
$= 0.1056$



Questions Asked

about the population proportion

- The population proportion p are generally unknown and are estimated from the data.
- Suppose we want to estimate the number of students planning to attend graduate school.
  1. Will the sample proportion equal the population proportion? Yes or No.
  2. If not, by how much will we miss it?

Estimating the population proportion p

- $\hat{p}$ is an estimate of the population proportion, i.e.,

$$E[\hat{p}] = p$$

- We will miss it by the standard error of the proportion

$$= \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}}$$

- Consider our example: $n = 40$ graduating seniors, $X = 6$ plan to attend graduate school.
  1. What is the proportion of graduating seniors planning to attend graduate school?
  2. By how much will we miss the true population proportion?

Estimating the population proportion
continued

- $\hat{p} = \frac{X}{n} = \frac{6}{40} = 0.15$.
- $$SE_{\hat{p}} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{0.15 \times (1-0.15)}{40}} = 0.05646.$$

iClicker Question 5.2

iClicker Question 5.2

Estimating the population proportion p

Consider this example: $n = 40$ graduating seniors, $X = 6$ is the number of graduating seniors planning to attend graduate school.

- If we take another random sample of size 40 graduating seniors, what is the probability that 17 percent or less of the graduating seniors will attend graduate school?
- $P(\hat{p} \leq .17) = \text{normalCDF}(-99, .17, .15, .05646) = 0.6384$
- If we take another random sample of size 40 graduating seniors, what is the probability that 20 percent or more of the graduating seniors will attend graduate school?
- $P(\hat{p} \geq .2) = \text{normalCDF}(.2, 99, .15, .05646) = 0.1879$

Estimating the population proportion

theory

- The population proportion p is estimated using the sample proportion $\hat{p}$, i.e., $E[\hat{p}] = p$. This estimate tends to miss by an amount called the $SE_{\hat{p}}$.
- The $SE_{\hat{p}}$ is calculated as

$$SE_{\hat{p}} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}.$$

- As sample size increases, the $SE_{\hat{p}}$ decreases.

iClicker Question 5.3

iClicker Question 5.3

MS Example

Only five percent of US families have a net worth in excess of 1 million dollars, and thus can be called millionaires. However, 30 percent of MSs 31,000 employees are millionaires (Harvard Business Review, July-August, 2000). If a random samples of 100 MS employees are selected at random, what proportion of the samples will have

- (a) more than 36% millionaires,
- (b) less than 29% millionaires, and
- (c) between 25 and 35% millionaires?

MS Example

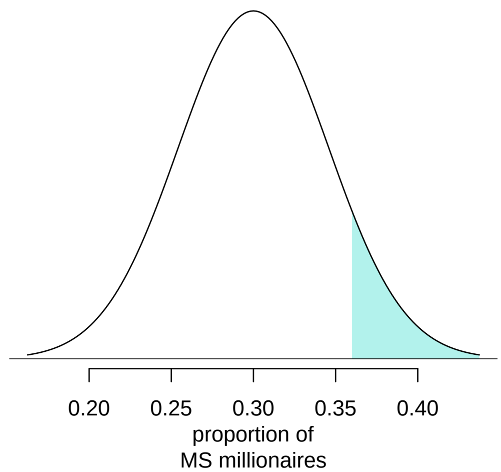
answer to part (a)

Given: $n = 100$ and $p = .3$

$P(\hat{p} > .36) =$

$\text{normalCDF}(.36, 9999, .3, \sqrt{\frac{.3\times(1-.3)}{100}})$

$= 0.0952$



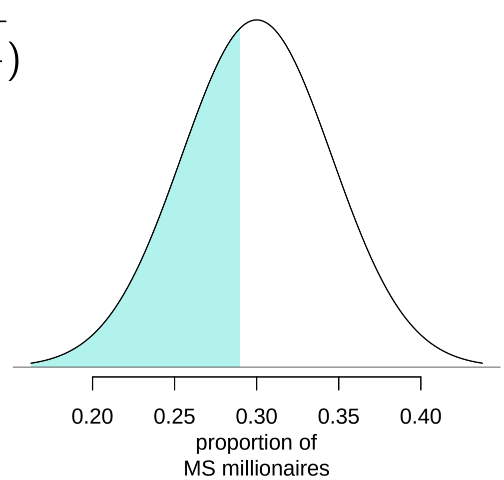
MS Example

answer to part (b)

$P(\hat{p} < .29) =$

$\text{normalCDF}(-9999, .29, .3, \sqrt{\frac{.3\times(1-.3)}{100}})$

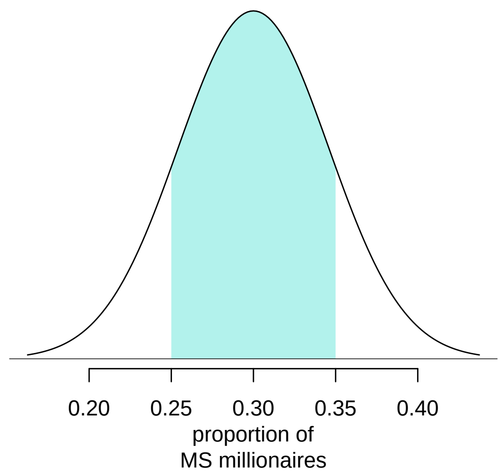
$= 0.4136$



MS Example

answer to part (c)

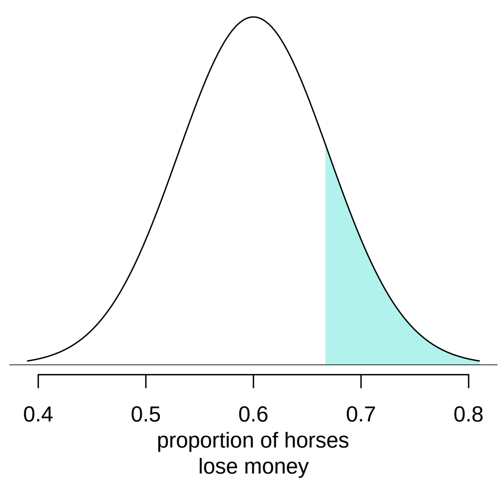
$P(.25 < \hat{p} < .35) =$
$\text{normalCDF}(.25, .35, .3, \sqrt{\frac{.3\times(1-.3)}{100}})$
$= 0.7248$



Problem 1, Page 68

answer

Given: $n = 49$ and $p = \frac{3}{5} = .6$
$P(\hat{p} > \frac{2}{3}) =$
$\text{normalCDF}(\frac{2}{3}, 9999, .6, \sqrt{\frac{.6\times(1-.6)}{49}})$
$= 0.1704$



iClicker Question 5.1 Post-lecture

iClicker Question 5.1 Post-lecture