S160 #11
Sampling Distribution of the Proportion, Part 1

JC Wang

February 23, 2016
Outline

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

2. Estimating Proportion
   - Questions Asked About Population Proportion
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 $p$. Note that $p$ is the (population) proportion of successes. The (sample) proportion of successes, $\hat{p} = \frac{X}{n}$, in a sample is also a random variable.
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 \( p \). Note that \( p \) is the (population) proportion of successes.

The (sample) proportion of successes, \( \hat{p} = \frac{X}{n} \), in a sample is also a random variable.
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 \( p \). Note that \( p \) is the (population) proportion of successes.

The (sample) proportion of successes, \( \hat{p} = \frac{X}{n} \), in a sample is also a random variable.
Sampling Distribution of the Proportion

- \( \hat{p} = \frac{X}{n} = \text{(number of successes)} / \text{(sample size)} \)
- For the binomial, \( X \), the number of successes, is expected to be around \( np \) give or take \( \sqrt{npq} \).
- For the proportion, \( \hat{p} \) is expected to be \( p = \frac{np}{n} \) give or take \( \sqrt{pq} n \) gives or take \( \sqrt{npq} n \).
Sampling Distribution of the Proportion

- \( \hat{p} = \frac{X}{n} = \frac{\text{number of successes}}{\text{sample size}} \)

- For the binomial, \( X \), the number of successes, is expected to be around \( np \) give or take \( \sqrt{npq} \).

- For the proportion, \( \hat{p} \) is expected to be \( p = \frac{np}{n} \) give or take \( \sqrt{npq} \).
Sampling Distribution of the Proportion

- \( \hat{p} = \frac{X}{n} = \frac{\text{(number of successes)}}{\text{(sample size)}} \)
- For the binomial, \( X \), the number of successes, is expected to be around \( np \) give or take \( \sqrt{npq} \).
- 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 $60 \pm 7$
- The proportion of warranties sold is expected to be around $\frac{60}{300} \pm \frac{7}{300}$ or $0.2 \pm 0.02$.
- The percentage of warranties sold is expected to be around 20% give or take 2% (Note: percentage = proportion $\times$ 100%)
TV+More Example
revisited

- The number of warranties sold is expected to be around $60 \pm 7$
- The *proportion* of warranties sold is expected to be around

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

- The *percentage* of warranties sold is expected to be around 20% give or take 2% (Note: percentage = proportion $\times$ 100%)
The number of warranties sold is expected to be around $60 \pm 7$

The *proportion* of warranties sold is expected to be around

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

The *percentage* of warranties sold is expected to be around 20% give or take 2% (Note: percentage = proportion $\times$ 100%)
Campus Video Rental Example, revisited

Historically, 5% of rentals from Campus Video are returned late.

- Campus Video rented out 100 videos yesterday. The percentage that will be returned late should be around 5%, give or take

\[
100\% \times \sqrt{\frac{0.05 \times 0.95}{100}} \approx 2.2\% \quad \text{(note: } q = 1 - p = 1 - 0.05 = 0.95)\]

- Campus Video rented out 700 videos yesterday. The percentage that will be returned late should be around 5%, give or take

\[
100\% \times \sqrt{\frac{0.05 \times 0.95}{700}} \approx 0.8\%\]
A study surveyed 100 students who took a standardized test. Among these students, 43 said they would like math help. What is the sample percentage of students needing math help?

A. 100%
B. 43%
C. 0.43%
D. 1%
E. cannot determine
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 $sd = 0.02$. Note that $p = 0.2$.
- If TV+More sold 1200 TV sets then now

$$sd = \sqrt{\frac{0.2 \times (1 - 0.2)}{1200}} = 0.0115$$
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 $sd = 0.02$. Note that $p = 0.2$.
- If TV+More sold 1200 TV sets then 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 it sold warranties with each TV is more than 25% of those sets, in other words,

\[ P(\hat{p} > 0.25) = ? \]
Sample Proportion is approximately normal

continued

Given: \( n = 100 \) and \( p = .2 \)

\[ P(\hat{p} > .25) = ? \]

Note that

\[ \hat{p} \approx N\left(0.2, \sqrt{\frac{0.2 \times (1 - 0.2)}{100}} = 0.04 \right) \]

\[ z = \frac{0.25 - 0.2}{0.04} = 1.25 \text{ and hence} \]

\[ P(\hat{p} > .25) \approx P(Z > 1.25) \]

\[ = 1 - P(Z \leq 1.25) = 1 - 0.8944 = 0.1056 \]
iClicker Question 11.2

Recall that 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%. What is the chance that the percentage of sets sold with extended warranties is between 12% (= 20% − 2 × 4%) and 28% (= 20% + 2 × 4%)?

A. 99.7%
B. 95%
C. 68%
D. 75%
E. cannot determine
Outline

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

2. Estimating Proportion
   - Questions Asked About Population Proportion
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.
  - Will the sample proportion equal the population proportion? Yes or No.
  - If not, by how much will it miss?
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 it miss?
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 it miss?
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 it miss?
Estimating the population proportion $p$

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

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

- Our estimate misses 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 it miss the true population proportion?
Estimating the population proportion $p$

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

- Our estimate misses it by the *standard error* of the proportion
  \[ = \sqrt{\hat{p}(1 - \hat{p})} \]

Consider our example: $n = 40$ graduating seniors, $X = 6$ plan to attend graduate school.

- What is the proportion of graduating seniors planning to attend graduate school?
- By how much will it miss the true population proportion?
Estimating the population proportion $p$

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

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

- Our estimate misses 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 it miss the true population proportion?
Estimating the population proportion $p$

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

- Our estimate misses it by the *standard error* of the proportion
  \[ = \sqrt{\hat{p}(1 - \hat{p})} \]

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 it miss the true population proportion?
Estimating the population proportion $p$

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

- Our estimate misses it by the *standard error* of the proportion
  
  $$\sqrt{\hat{p}(1 - \hat{p}) \frac{1}{n}}$$

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 it miss the true population proportion?
Estimating the population proportion

Continued

\[ \hat{p} = \frac{X}{n} = \frac{6}{40} = 0.15. \]

\[ SE_{\hat{p}} = \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = \sqrt{\frac{0.15 \times 0.85}{40}} = 0.056. \]

What if 54 out of 360 students plan to go to graduate school. The proportion of all students who plan to go to graduate school is estimated as

\[ \hat{p} = \frac{54}{360} = 0.15 \text{ with } SE_{\hat{p}} = \sqrt{\frac{0.15 \times 0.85}{360}} = 0.0188. \]
Estimating the population proportion

\[ \hat{p} = \frac{X}{n} = \frac{6}{40} = 0.15. \]

\[ SE_{\hat{p}} = \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = \sqrt{\frac{0.15 \times 0.85}{40}} = 0.056. \]

What if 54 out of 360 students plan to go to graduate school. The proportion of all students who plan to go to graduate school is estimated as

\[ \hat{p} = \frac{54}{360} = 0.15 \text{ with } SE_{\hat{p}} = \sqrt{\frac{0.15 \times 0.85}{360}} = 0.0188. \]
Estimating the population proportion
continued

\[ \hat{p} = \frac{X}{n} = \frac{6}{40} = 0.15. \]

\[ SE_{\hat{p}} = \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = \sqrt{\frac{0.15 \times 0.85}{40}} = 0.056. \]

What if 54 out of 360 students plan to go to graduate school. The proportion of all students who plan to go to graduate school is estimated as

\[ \hat{p} = \frac{54}{360} = 0.15 \text{ with } SE_{\hat{p}} = \sqrt{\frac{0.15 \times 0.85}{360}} = 0.0188. \]
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.
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.
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 11.3

Which of the following statements is true about the standard error of the sample proportion?

A. The standard error increases when sample size increases.

B. The standard error decreases when sample size decreases.

C. The increase/decrease of sample size has no effect on the value of the standard error.

D. The standard error decreases when sample size increases.

E. None of the previous.