Inference for a Single Proportion

Nate Wells

Math 141, 4/5/21

Outline

In this lecture, we will. . .

Outline

In this lecture, we will. . .

- Use theory to find the standard error for one sample proportions

Outline

In this lecture, we will. . .

- Use theory to find the standard error for one sample proportions
- Calculate confidence intervals and perform hypothesis tests for proportions using the theory-based method

The Sampling Distribution for Sample Proportion

- Consider a population variable that takes only two levels, success S and failure F. Let p be the proportion of success in the population.

The Sampling Distribution for Sample Proportion

- Consider a population variable that takes only two levels, success S and failure F. Let p be the proportion of success in the population.
- Suppose we randomly choose a single observation from a population, and denote the result as 1 if observation is S and 0 if it is F.

The Sampling Distribution for Sample Proportion

- Consider a population variable that takes only two levels, success S and failure F. Let p be the proportion of success in the population.
- Suppose we randomly choose a single observation from a population, and denote the result as 1 if observation is S and 0 if it is F.
- The *mean* of this variable is p, and the *standard deviation* is $\sqrt{p(1-p)}$

The Sampling Distribution for Sample Proportion

- Consider a population variable that takes only two levels, success S and failure F. Let p be the proportion of success in the population.
- Suppose we randomly choose a single observation from a population, and denote the result as 1 if observation is S and 0 if it is F.
- The *mean* of this variable is p, and the *standard deviation* is $\sqrt{p(1-p)}$
- If we instead take an SRS of size n from the population, we can view the sample proportion $\hat{p}$ as a sample mean:

The Sampling Distribution for Sample Proportion

- Consider a population variable that takes only two levels, success S and failure F. Let p be the proportion of success in the population.
- Suppose we randomly choose a single observation from a population, and denote the result as 1 if observation is S and 0 if it is F.
- The *mean* of this variable is p, and the *standard deviation* is $\sqrt{p(1-p)}$
- If we instead take an SRS of size n from the population, we can view the sample proportion $\hat{p}$ as a sample mean:
  - We are averaging across each person in the sample the variable that takes the value 1 if the individual is a success and 0 otherwise.

The Sampling Distribution for Sample Proportion

- Consider a population variable that takes only two levels, success S and failure F. Let p be the proportion of success in the population.
- Suppose we randomly choose a single observation from a population, and denote the result as 1 if observation is S and 0 if it is F.
- The *mean* of this variable is p, and the *standard deviation* is $\sqrt{p(1-p)}$
- If we instead take an SRS of size n from the population, we can view the sample proportion $\hat{p}$ as a sample mean:
  - We are averaging across each person in the sample the variable that takes the value 1 if the individual is a success and 0 otherwise.
- By the central limit theorem, if n is large, then $\hat{p}$ is approximately Normal, with mean p and standard deviation $\sqrt{\frac{p(1-p)}{n}}$

Examples

Using data from the gss General Social Survey. . .

- 56% identified as female
- 27.2% obtained a college degree
- 96.7% were 21 or older

Examples

Using data from the gss General Social Survey...

- 56% identified as female
- 27.2% obtained a college degree
- 96.7% were 21 or older

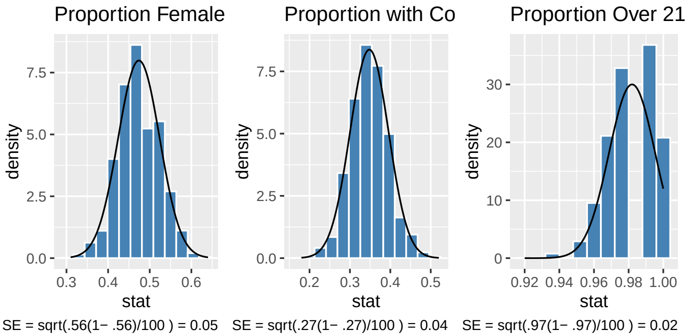
If we draw samples of size 100 from the GSS, the sampling distributions look like...

Examples

Using data from the gss General Social Survey. . .

- 56% identified as female
- 27.2% obtained a college degree
- 96.7% were 21 or older

If we draw samples of size 100 from the GSS, the sampling distributions look like. . .

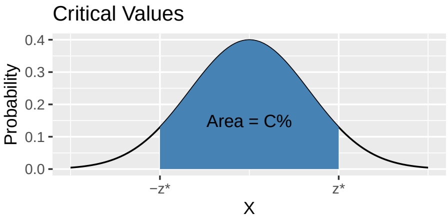


Critical Values

- The **critical value** z^* for a $C\%$ confidence interval is the value so that $C\%$ of area is between $-z^*$ and z^* in the standard Normal distribution.

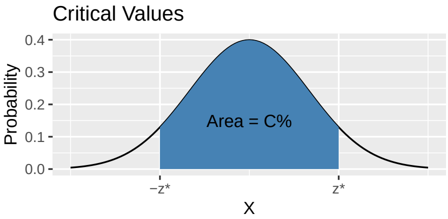
Critical Values

- The **critical value** z^* for a $C\%$ confidence interval is the value so that $C\%$ of area is between $-z^*$ and z^* in the standard Normal distribution.



Critical Values

- The **critical value** z^* for a $C\%$ confidence interval is the value so that $C\%$ of area is between $-z^*$ and z^* in the standard Normal distribution.



- Previously, we saw that for Normal distributions, 95% of observations are within 2 standard deviations of the mean. So the critical value for 95% confidence is

$$z^* = 2$$

Confidence Intervals

When a sample statistic is approximately Normally distribution, the $C\%$ confidence interval is

$$\text{statistic} \pm z^* \cdot SE$$

where z^* is the critical value for $C\%$ confidence and SE is the standard error for the statistic.

Confidence Intervals

When a sample statistic is approximately Normally distribution, the $C\%$ confidence interval is

$$\text{statistic} \pm z^* \cdot SE$$

where z^* is the critical value for $C\%$ confidence and SE is the standard error for the statistic.

- The standard error for a sample proportion $\hat{p}$ is $SE = \sqrt{\frac{p(1-p)}{n}}$. Since we don't know p, we estimate it in the SE formula with $\hat{p}$.

Confidence Intervals

When a sample statistic is approximately Normally distribution, the $C\%$ confidence interval is

$$\text{statistic} \pm z^* \cdot SE$$

where z^* is the critical value for $C\%$ confidence and SE is the standard error for the statistic.

- The standard error for a sample proportion $\hat{p}$ is $SE = \sqrt{\frac{p(1-p)}{n}}$. Since we don't know p, we estimate it in the SE formula with $\hat{p}$.

Theorem

Suppose an SRS of size n is collected from a population with parameter p. If n is large enough so that both $n\hat{p}$ and $n(1-\hat{p})$ are at least 10, then the confidence interval for p is

$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

An Example

On June 28, 2012 the U.S. Supreme Court upheld the much debated 2010 healthcare law, declaring it constitutional. A Gallup poll released the day after this decision indicates that 47% of 1,012 Americans agreed with this decision. Use the theory-based method at 99% confidence to estimate the true proportion of Americans that agreed with this decision.

An Example

On June 28, 2012 the U.S. Supreme Court upheld the much debated 2010 healthcare law, declaring it constitutional. A Gallup poll released the day after this decision indicates that 47% of 1,012 Americans agreed with this decision. Use the theory-based method at 99% confidence to estimate the true proportion of Americans that agreed with this decision.

- Our sample statistic is $\hat{p} = 0.47$

```
p_hat<-0.47
p_hat
```

```
## [1] 0.47
```

An Example

On June 28, 2012 the U.S. Supreme Court upheld the much debated 2010 healthcare law, declaring it constitutional. A Gallup poll released the day after this decision indicates that 47% of 1,012 Americans agreed with this decision. Use the theory-based method at 99% confidence to estimate the true proportion of Americans that agreed with this decision.

- Our sample statistic is $\hat{p} = 0.47$

```
p_hat<-0.47
p_hat
```

```
## [1] 0.47
```

- The critical value z^* for 99% confidence is $z^* = 2.58$

```
z<-qnorm(.995, 0 , 1)
z
```

```
## [1] 2.575829
```

An Example

On June 28, 2012 the U.S. Supreme Court upheld the much debated 2010 healthcare law, declaring it constitutional. A Gallup poll released the day after this decision indicates that 47% of 1,012 Americans agreed with this decision. Use the theory-based method at 99% confidence to estimate the true proportion of Americans that agreed with this decision.

- Our sample statistic is $\hat{p} = 0.47$

```
p_hat<-0.47
p_hat
```

```
## [1] 0.47
```

- The critical value z^* for 99% confidence is $z^* = 2.58$

```
z<-qnorm(.995, 0 , 1)
z
```

```
## [1] 2.575829
```

- The standard error for $\hat{p}$ is $SE = 0.016$

```
SE<-sqrt(p_hat*(1- p_hat)/1012)
SE
```

```
## [1] 0.01568905
```

An Example

- The theory-based confidence interval is $(0.43, 0.51)$

```
CI_low<-p_hat-z*SE
CI_high<-p_hat+z*SE
```

```
##      CI_low   CI_high
## 1 0.4295877 0.5104123
```

An Example

- The theory-based confidence interval is $(0.43, 0.51)$

```
CI_low<-p_hat-z*SE
CI_high<-p_hat+z*SE
```

```
##      CI_low   CI_high
## 1 0.4295877 0.5104123
```

- How does this compare to the bootstrap method?

An Example

- The theory-based confidence interval is $(0.43, 0.51)$

```
CI_low<-p_hat-z*SE
CI_high<-p_hat+z*SE
```

```
##       CI_low   CI_high
## 1 0.4295877 0.5104123
```

- How does this compare to the bootstrap method?

```
health %>% specify(response = agree, success = "yes") %>%
  generate(reps=10000, type = "bootstrap") %>%
  calculate(stat = "prop") %>%
  get_ci(level = .99, type = "se", point_estimate = p_hat)
```

```
## # A tibble: 1 x 2
##   lower_ci upper_ci
##      <dbl>    <dbl>
## 1    0.429    0.511
```

z-Scores

- The **z-score** for a test statistic x with standard error SE and mean μ under the Null hypothesis is

$$z = \frac{x - \mu}{SE}$$

z-Scores

- The **z-score** for a test statistic x with standard error SE and mean μ under the Null hypothesis is

$$z = \frac{x - \mu}{SE}$$

- Suppose X is approximately Normal with mean μ and standard deviation σ. Then

$$Z = \frac{X - \mu}{\sigma}$$

is approximately standard Normal.

z-Scores

- The **z-score** for a test statistic x with standard error SE and mean μ under the Null hypothesis is

$$z = \frac{x - \mu}{SE}$$

- Suppose X is approximately Normal with mean μ and standard deviation σ. Then

$$Z = \frac{X - \mu}{\sigma}$$

  is approximately standard Normal.

- By location-scale invariance,

$$P(X > x) = P\left(Z > \frac{x - \mu}{\sigma}\right)$$

z-Scores

- The **z-score** for a test statistic x with standard error SE and mean μ under the Null hypothesis is

$$z = \frac{x - \mu}{SE}$$

- Suppose X is approximately Normal with mean μ and standard deviation σ. Then

$$Z = \frac{X - \mu}{\sigma}$$

  is approximately standard Normal.

- By location-scale invariance,

$$P(X > x) = P\left(Z > \frac{x - \mu}{\sigma}\right)$$

- If we want to compute a P-Value for test statistic x, we can instead compute a P-value for its z-score z:

$$\begin{aligned} \text{P-value} &= P(Z > z) && \text{if } H_a \text{ is one-sided right} \\ \text{P-value} &= P(Z < z) && \text{if } H_a \text{ is one-sided left} \\ \text{P-value} &= 2 \cdot P(Z > |z|) && \text{if } H_a \text{ is two-sided} \end{aligned}$$

Hypothesis Tests

By the central limit theorem, if $H_0 : p = p_0$ is true, then for large n, the standard error for the sample statistic $\hat{p}$ is

$$SE = \sqrt{\frac{p_0(1 - p_0)}{n}}$$

Hypothesis Tests

By the central limit theorem, if $H_0 : p = p_0$ is true, then for large n, the standard error for the sample statistic $\hat{p}$ is

$$SE = \sqrt{\frac{p_0(1-p_0)}{n}}$$

Theorem

To test $H_0 : p = p_0$ against $H_a : p \neq p_0$ (or the one-sided alternative) we use the standardized test statistic

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

If n is large enough so that both $n\hat{p}$ and $n(1-\hat{p})$ are at least 10, then the p-value for the test is computed using the standard Normal distribution.

Rock-Paper-Scissors

In Rock-Paper-Scissors, each player chooses one of 3 symbols (Rock, Paper, Scissors). Are all three options chosen with equal frequency?

Rock-Paper-Scissors

In Rock-Paper-Scissors, each player chooses one of 3 symbols (Rock, Paper, Scissors). Are all three options chosen with equal frequency?

A study observed that for 119 people playing Rock-Paper-Scissors in an official tournament, 66 players selected Rock on their first turn.

Rock-Paper-Scissors

In Rock-Paper-Scissors, each player chooses one of 3 symbols (Rock, Paper, Scissors). Are all three options chosen with equal frequency?

A study observed that for 119 people playing Rock-Paper-Scissors in an official tournament, 66 players selected Rock on their first turn.

- We test $H_0 : p_0 = 1/3$ against $H_a : p_0 \neq 1/3$, where p_0 is the theoretical frequency a player chooses rock on the first turn.

Rock-Paper-Scissors

In Rock-Paper-Scissors, each player chooses one of 3 symbols (Rock, Paper, Scissors). Are all three options chosen with equal frequency?

A study observed that for 119 people playing Rock-Paper-Scissors in an official tournament, 66 players selected Rock on their first turn.

- We test $H_0 : p_0 = 1/3$ against $H_a : p_0 \neq 1/3$, where p_0 is the theoretical frequency a player chooses rock on the first turn.
- The sample statistic is $\hat{p} = 0.55$

```
p_hat<-66/119
p_hat
```

```
## [1] 0.5546218
```

Rock-Paper-Scissors

In Rock-Paper-Scissors, each player chooses one of 3 symbols (Rock, Paper, Scissors). Are all three options chosen with equal frequency?

A study observed that for 119 people playing Rock-Paper-Scissors in an official tournament, 66 players selected Rock on their first turn.

- We test $H_0 : p_0 = 1/3$ against $H_a : p_0 \neq 1/3$, where p_0 is the theoretical frequency a player chooses rock on the first turn.
- The sample statistic is $\hat{p} = 0.55$

```
p_hat<-66/119
p_hat
```

```
## [1] 0.5546218
```

- The standard error is $SE = 0.04$

```
SE<- sqrt((1/3)*(1-(1/3))/119)
SE
```

```
## [1] 0.04321358
```

Rock-Paper-Scissors

In Rock-Paper-Scissors, each player chooses one of 3 symbols (Rock, Paper, Scissors). Are all three options chosen with equal frequency?

A study observed that for 119 people playing Rock-Paper-Scissors in an official tournament, 66 players selected Rock on their first turn.

- We test $H_0 : p_0 = 1/3$ against $H_a : p_0 \neq 1/3$, where p_0 is the theoretical frequency a player chooses rock on the first turn.
- The sample statistic is $\hat{p} = 0.55$

```
p_hat<-66/119
p_hat
```

```
## [1] 0.5546218
```

- The standard error is $SE = 0.04$

```
SE<- sqrt((1/3)*(1-(1/3))/119)
SE
```

```
## [1] 0.04321358
```

- The test statistic is $z = 5.12$

```
z<- (p_hat - 1/3)/ SE
z
```

```
## [1] 5.120809
```

Rock-Paper-Scissors

- The P-Value (probability of observing a sample proportion as extreme as 66/119) is 0.0000003

```
Pval<- 2*pnorm(-z, mean = 0, sd = 1)
Pval
```

```
## [1] 3.04227e-07
```

Rock-Paper-Scissors

- The P-Value (probability of observing a sample proportion as extreme as 66/119) is 0.0000003

```
Pval<- 2*pnorm(-z, mean = 0, sd = 1)
Pval
```

```
## [1] 3.04227e-07
```

- We reject the null hypothesis in favor of the alternative at significance $\alpha = 0.05$.

Rock-Paper-Scissors

- The P-Value (probability of observing a sample proportion as extreme as 66/119) is 0.0000003

```
Pval<- 2*pnorm(-z, mean = 0, sd = 1)
Pval
```

```
## [1] 3.04227e-07
```

- We reject the null hypothesis in favor of the alternative at significance $\alpha = 0.05$.

How does this compare to the simulation based test?

Rock-Paper-Scissors

- The P-Value (probability of observing a sample proportion as extreme as 66/119) is 0.0000003

```
Pval<- 2*pnorm(-z, mean = 0, sd = 1)
Pval
```

```
## [1] 3.04227e-07
```

- We reject the null hypothesis in favor of the alternative at significance $\alpha = 0.05$.

How does this compare to the simulation based test?

```
rps %>% specify(response = choice, success = "rock") %>%
  hypothesize(null = "point", p = 1/3) %>%
  generate(reps = 5000, type = "simulate") %>%
  calculate(stat = "prop") %>%
  get_p_value(obs_stat = p_hat, direction = "both")
```

```
## # A tibble: 1 x 1
##   p_value
##     <dbl>
## 1       0
```