Section 8.4 — Estimating Population Proportions

Chris Godbout

Outline

Introduction

Interpretation of Confidence Intervals

Finding Confidence Intervals

Estimating Population Proportions

Minimum Sample Size

Introduction

Definition (Point Estimate)

A point estimate is a single value (or point) used to approximate a population parameter.

$\hat{p}$ is the best *point*	p is the population proportion

Definitions

Definition (Confidence Interval)

A confidence interval (or interval estimate) is a range (or an interval) of values used to estimate the true value of the population parameter. A confidence interval is sometimes abbreviated as CI.

Definitions

Definition (Confidence Interval)

A confidence interval (or interval estimate) is a range (or an interval) of values used to estimate the true value of the population parameter. A confidence interval is sometimes abbreviated as CI.

Confidence Level

The confidence level is the probability $1 - \alpha$ (such as 0.95 or 95%) that the confidence interval actually does contain the population parameter, assuming that the estimation process is repeated a large number of times. It is also sometimes called the degree of confidence or the confidence coefficient.

Confidence Levels and α

Table 1: Most common confidence intervals

Confidence Level	α
90% or 0.90	0.10
95% or 0.95	0.05
99% or 0.99	0.01

Interpretation of Confidence Intervals

Example

If a 95% confidence interval of the population proportion p is $0.499 < p < 0.561$, what does it mean?

Example

If a 95% confidence interval of the population proportion p is $0.499 < p < 0.561$, what does it mean?

Interpretation

If we were to select many different samples of the same size and construct the corresponding confidence intervals, 95% of them would actually contain the value of the population proportion p.

Finding Confidence Intervals

Critical values

Definition (Critical Value)

A critical value is the number on the borderline separating sample statistics that are likely to occur from those that are unlikely.

Critical values

Definition (Critical Value)

A critical value is the number on the borderline separating sample statistics that are likely to occur from those that are unlikely.

Notation

$z_{\alpha/2}$ is a critical value that is a z score with the property that it separates an area of $\alpha/2$ in the right tail of the standard normal distribution.

Critical values

Definition (Critical Value)

A critical value is the number on the borderline separating sample statistics that are likely to occur from those that are unlikely.

Notation

$z_{\alpha/2}$ is a critical value that is a z score with the property that it separates an area of $\alpha/2$ in the right tail of the standard normal distribution.

Example

Find the critical value value $z_{\alpha/2}$ corresponding to a 95% confidence level.

Common Critical Values

Confidence Level	α	Critical Value
90%	0.10	1.645
95%	0.05	1.96
99%	0.01	2.575

Margin of Error

Definition (Margin of Error)

When data from a simple random sample are used to estimate a population proportion p, the margin of error, denote E, is the maximum likely difference between the observed sample proportion $\hat{p}$ and the true value of the population p.

Estimating Population Proportions

Requirements

1. The sample is a simple random sample.
2. The conditions for the binomial distribution are satisfied.
3. There are least 5 success and at least 5 failures.

Margin of Error

If $\hat{p}$ is the sample proportion, then the margin of error is

$$E = z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

Notation

If $\hat{p}$ is the sample proportion and E is the margin of error, then the confidence interval is expressed as one of

Notation

If $\hat{p}$ is the sample proportion and E is the margin of error, then the confidence interval is expressed as one of

$$\hat{p} - E < p < \hat{p} + E$$

Notation

If $\hat{p}$ is the sample proportion and E is the margin of error, then the confidence interval is expressed as one of

$$\hat{p} - E < p < \hat{p} + E$$

$$\hat{p} \pm E$$

Notation

If $\hat{p}$ is the sample proportion and E is the margin of error, then the confidence interval is expressed as one of

$$\hat{p} - E < p < \hat{p} + E$$

$$\hat{p} \pm E$$

$$(\hat{p} - E, \hat{p} + E)$$

Identity Theft

From a KRC Research poll in which respondents were asked if they felt vulnerable to identity theft: $n = 1002$, $x = 531$ who said "yes". Use a 95% confidence interval

Identity Theft

From a KRC Research poll in which respondents were asked if they felt vulnerable to identity theft: $n = 1002$, $x = 531$ who said "yes". Use a 95% confidence interval

1. Find the best point estimate of the population

Identity Theft

From a KRC Research poll in which respondents were asked if they felt vulnerable to identity theft: $n = 1002$, $x = 531$ who said ”yes”. Use a 95% confidence interval

1. Find the best point estimate of the population
2. Identify the margin of error E

Identity Theft

From a KRC Research poll in which respondents were asked if they felt vulnerable to identity theft: $n = 1002$, $x = 531$ who said ”yes”. Use a 95% confidence interval

1. Find the best point estimate of the population
2. Identify the margin of error E
3. Find the confidence interval.

Identity Theft

From a KRC Research poll in which respondents were asked if they felt vulnerable to identity theft: $n = 1002$, $x = 531$ who said ”yes”. Use a 95% confidence interval

1. Find the best point estimate of the population
2. Identify the margin of error E
3. Find the confidence interval.
4. What does the confidence interval mean?

Minimum Sample Size

Minimum sample size

To determine how large a sample a sample should be in order to estimate the population proportion with a confidence level of $1 - \alpha$ and a margin of error E, use one of the following

p is known

$$n = p(1 - p) \left(\frac{z_{\alpha/2}}{E} \right)^2$$

p is not known

$$n = 0.25 \left(\frac{z_{\alpha/2}}{E} \right)^2$$

Fortune tellers

Find the sample size needed to estimate the percentage of adults who have consulted fortune tellers. We want the confidence level to be 98% and the error to be within 3 percentage points. Use results from a prior Pew Research Center poll suggesting that 15% of adults have consulted fortune tellers.