Confidence Intervals for a Proportion

Colby Community College

Example 1

In a Gallup poll of 1487 adults, 43% of them said that they have a Facebook page.

Based on this result, what is the best point estimate of the proportion of all adults who have a Facebook page.

Example 1

In a Gallup poll of 1487 adults, 43% of them said that they have a Facebook page.

Based on this result, what is the best point estimate of the proportion of all adults who have a Facebook page.

The sample proportion, 0.43, is the best point estimate of the population proportion.

Example 1

In a Gallup poll of 1487 adults, 43% of them said that they have a Facebook page.

Based on this result, what is the best point estimate of the proportion of all adults who have a Facebook page.

The sample proportion, 0.43, is the best point estimate of the population proportion.

Note

We have no indication of how *good* of an estimate 0.43 is, just that it is the best of the available options.

Definition

A **confidence interval** is a range of values around the point estimate used to estimate the true value of a population parameter.

$$[\text{point estimate} - \text{some value}, \text{point estimate} + \text{some value}]$$

A confidence interval is sometimes abbreviated as CI.

Definition

A **confidence interval** is a range of values around the point estimate used to estimate the true value of a population parameter.

$$[\text{point estimate} - \text{some value}, \text{point estimate} + \text{some value}]$$

A confidence interval is sometimes abbreviated as CI.

Definition

The **confidence level** is the probability that the confidence interval actually does contain the population parameter, assuming that the estimation process is repeated a large number of times.

Definition

A **confidence interval** is a range of values around the point estimate used to estimate the true value of a population parameter.

$$[\text{point estimate} - \text{some value}, \text{point estimate} + \text{some value}]$$

A confidence interval is sometimes abbreviated as CI.

Definition

The **confidence level** is the probability that the confidence interval actually does contain the population parameter, assuming that the estimation process is repeated a large number of times.

Note

Round the confidence interval limits to three significant digits.

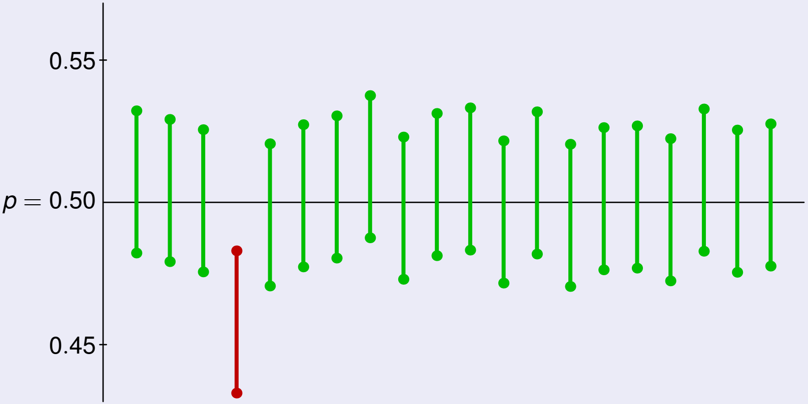
The Process Success Rate

A confidence level of 95% tells us that the process we use should, given enough iterations, result in a confidence interval that contains the true population proportion 95% of the time.

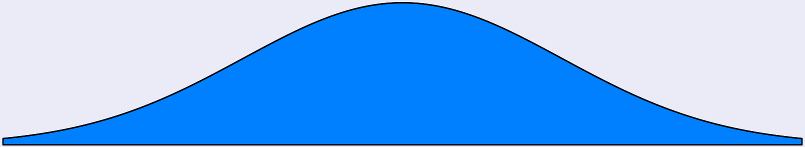
The Process Success Rate

A confidence level of 95% tells us that the process we use should, given enough iterations, result in a confidence interval that contains the true population proportion 95% of the time.

If the true population proportion is $p = 0.5$, then we expect around 19 of 20 confidence intervals to contain the true value of p.

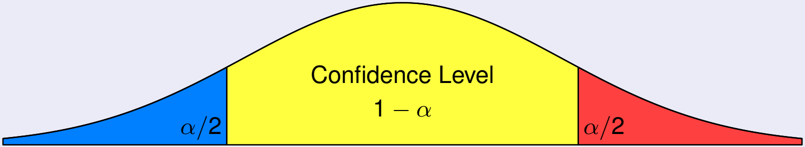


A Few Observations



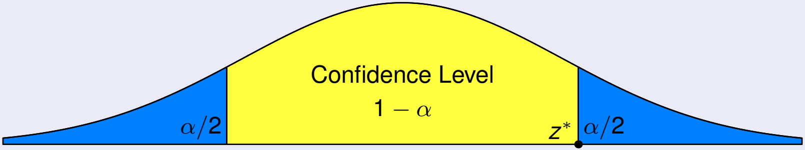
- When the requirements of the Central Limit Theorem are met, the sampling distribution of sample proportions can be approximated by a normal distribution.

A Few Observations



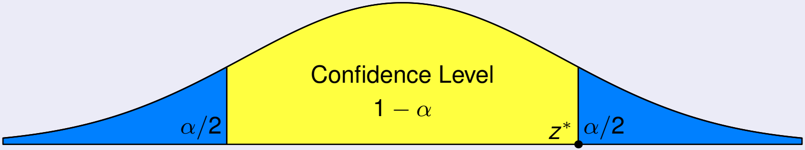
- When the requirements of the Central Limit Theorem are met, the sampling distribution of sample proportions can be approximated by a normal distribution.
- A z score associated with a sample proportion has a probability $\alpha/2$ of falling in the right tail portion.

A Few Observations



- When the requirements of the Central Limit Theorem are met, the sampling distribution of sample proportions can be approximated by a normal distribution.
- A z score associated with a sample proportion has a probability $\alpha/2$ of falling in the right tail portion.
- The z score at the boundary of the right-tail region is commonly denoted by z^*.

A Few Observations



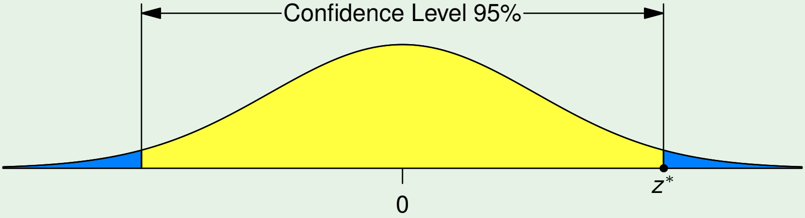
- When the requirements of the Central Limit Theorem are met, the sampling distribution of sample proportions can be approximated by a normal distribution.
- A z score associated with a sample proportion has a probability $\alpha/2$ of falling in the right tail portion.
- The z score at the boundary of the right-tail region is commonly denoted by z^*.

Definition

The value z^* is called a **critical value**.

Example 2

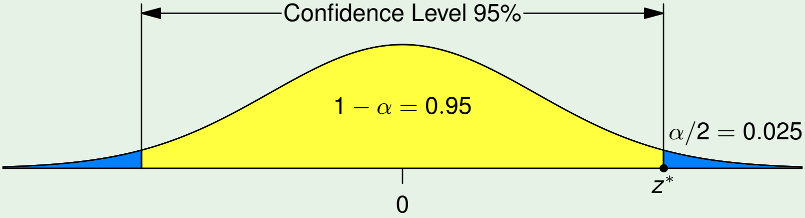
Let us find the critical value z^* corresponding to a 95% confidence level.



Example 2

Let us find the critical value z^* corresponding to a 95% confidence level.

A 95% confidence interval gives $\alpha = 0.05$ and $\alpha/2 = 0.025$.

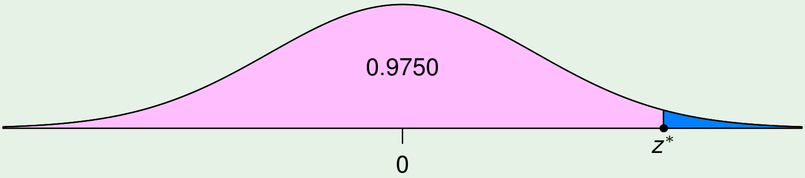


Example 2

Let us find the critical value z^* corresponding to a 95% confidence level.

A 95% confidence interval gives $\alpha = 0.05$ and $\alpha/2 = 0.025$.

To find the z value using the inverse normal distribution, we need to know the cumulative area to the left of the right tail, $0.025 + 0.95 = 0.9750$.



Example 2

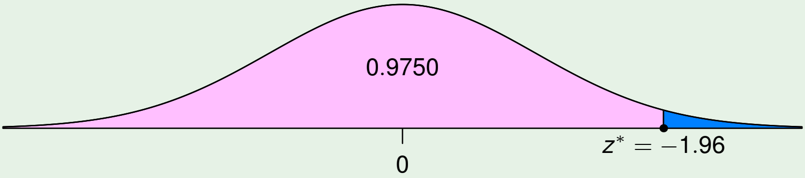
Let us find the critical value z^* corresponding to a 95% confidence level.

A 95% confidence interval gives $\alpha = 0.05$ and $\alpha/2 = 0.025$.

To find the z value using the inverse normal distribution, we need to know the cumulative area to the left of the right tail, $0.025 + 0.95 = 0.9750$.

Using technology we get

$$z^* = 1.96$$



Common Confidence Levels

Confidence Level	α	Critical Value
90%	0.10	1.645
95%	0.05	1.960
99%	0.01	2.575