Introduction to confidence intervals

Millie Harris

Summary

In statistics, to estimate an unknown parameter you can construct a confidence interval. This is the range of values you expect the true estimate to fall between if you were to repeat the study several times, with a certain level of confidence. They are a vital tool used in economics, medicine, and they measure uncertainties in everyday life. For example, weather forecasting or outcomes in sports. This study guide introduces confidence intervals, confidence levels, and Z values using the Normal Distribution.

Before reading this guide, it is recommended that you read [Guide: Intervals], [Guide: The normal distribution].

What is a confidence interval?

If you were conducting a study and took several different samples of data, the mean for that data would be slightly different each time. So, when estimating population means, instead of providing one value, you can specify a **range of values** which is likely to contain the true mean. This is called a confidence interval (CI).

This guide will focus on how to construct and interpret a confidence interval using the **normal distribution**.

For more information on constructing CIs using other distributions see Guide: [insert].

> **i** Definition of confidence interval using the normal distribution
>
> A **confidence interval** (CI) is a range on values, consisting of an **lower** and **upper** bound which show the sample mean plus or minus the standard error.
> The CI is given by the interval [sample mean $\pm$ $standard\ error$] $= [\bar{x} \pm Z_{\frac{\alpha}{2}} \frac{\sigma}{n}]$, where:
>
> - $\bar{x}$ is the sample mean
> - $Z_{\frac{\alpha}{2}}$ is the Z-value
> - σ is the standard deviation
> - $n = sample\ size$

So your CI would appear $[\bar{x} - Z_{\frac{\alpha}{2}}\frac{\sigma}{n}, \bar{x} + Z_{\frac{\alpha}{2}}\frac{\sigma}{n}]$.

- $\bar{x} - Z_{\frac{\alpha}{2}}\frac{\sigma}{n}$ is your **lower** bound
- $\bar{x} + Z_{\frac{\alpha}{2}}\frac{\sigma}{n}$ is your **upper** bound

Check your work!

The average of your confidence interval should equal your sample mean. For example, for $\bar{x} = 50$. If your CI is $[48, 52]$, take $48 + 52 = 100$. Then $100/2 = 50 = \bar{x}$.

Each CI has a specific confidence level which is stated at the beginning of the study. For example, you could have a $90\% CI$ or a $95\% CI$.

What is a confidence level?

Definition of a confidence level

A **confidence level** (CL) suggests that if you were to repeat the study many times, you would expect the true estimate to fall within CL$\%$ of the results.

A CL is typically represented using a percentage or decimal. For example, a $95\% CL$ can be represented as 0.95.

Caution

This does not mean there is a 95% chance the range of values contains the true estimate. Instead, if you were to repeat the study many times, with a CL of 95% - you would expect 95% of the CIs to **contain the true estimate.**

What is a Z value (Z score)?

Using the normal distribution, a Z value (sometimes called Z score or standard score) is a known test statistic. It shows how many $\sigma's$ above or below the mean an observed data point is.

Definition of the z value using the normal distribution

A Z value is written $Z_{\frac{\alpha}{2}}$, where:

- $\alpha = 1 - CL$

$\frac{\alpha}{2}$ explains the two tails of the normal distribution.
These values are **known** and you can use the calculator below to find specific z values.

This graph shows the standard normal distribution, where:

- $P(x \leq -z$ or $x \geq z) = \alpha = 1 - CL$ (if your $\alpha = 0.05$, your $CL = 0.95 = 95$)
- $Z = Z\ value$

Try!

Slide the $Z\ value$ and notice these commonly used CLs:

CL	$\frac{\alpha}{2}$	$Z_{\frac{\alpha}{2}}$
$0.80 = 80\%$	0.10	1.282
$0.85 = 85\%$	0.075	1.440
$0.90 = 90\%$	0.05	1.645
$0.95 = 95\%$	0.025	1.960
$0.99 = 99\%$	0.005	2.576

Example 1

On a random weekday, Cantor's Confectionery sells 50 bags of sweets. On average, the bags weigh $300g$ with a \$standard deviation \$ of $10g$. Construct a 95% confidence interval for the true estimate of the weight of a bag of sweets sold by Cantor's Confectionery.

Step 1: Identify what you need from the text.

- $\bar{x} = 300$
- $CL = 95\% = 0.95$
- $\sigma = 10$
- $n = 50$

Step 2: Use Z Table to find Z Value [Insert Table]
Your $CL = 0.95$ so $\frac{\alpha}{2} = 0.025$, from Z calculator your $Z_{\frac{\alpha}{2}} = 1.960$.
Step 3: Construct CI
$95\%CI = [\bar{x} \pm Z_{0.025}\frac{\sigma}{n}]$

$= [300 \pm (1.960\frac{10}{50})]$
$= [299.608, 300.392]$
Step 4: Check
$299.608 + 300.392 = 600$

$$600/2 = 300 = \bar{x}$$

Example 2

The number of eggs laid a year by a farmhouse chicken is Normally distributed with mean μ and standard deviation $\sigma = 45$. A random sample of 65 birds gives a mean of 178. Build a 90% confidence interval for μ.
Step 1: Identify what you need from the text.

- $\bar{x} = 178$
- Confidence Level = 90% = 0.90
- $\sigma = 45$
- n = 65

Step 2: Use Z Table to find Z Value [Insert Table]
Your CL = 0.90 so $\frac{\alpha}{2} = 0.05$, from Z Table your $Z_{\frac{\alpha}{2}} = 1.645$
Step 3: Construct CI
90% CI $= [\bar{x} \pm Z_{0.05}\frac{\sigma}{n}]$
$= [178 \pm (1.645\times \frac{45}{65})]$
$= [176.861, 179.136]$

Example 3: Summary Table

A study on 210 sunflowers in one field follows the Normal Distribution and produced the following table of results:

Average Height	**1.68ft**
Average Weight	**0.9kg**
Variance in Height	**0.37ft**
Variance in Weight	**0.045kg**

Construct a 90% CI for both the height (ft) and the weight (kg) of all of the farmers sunflowers.
Height:

- $\bar{x}$ = 1.68
- $\sigma = \sqrt{Variance} = \sqrt{0.37}$ = 0.608
- n = 210
- CL = 90% = 0.90, $Z_{\frac{\alpha}{2}}$ = 1.645.
- So a 90% CI for height = $[\bar{x} \pm Z_{0.05}\frac{\sigma}{n}]$

= [1.68 ± (1.645 × $\frac{0.608}{210}$)]
= [1.675, 1.684]
Weight:

- $\bar{x}$ = 0.9
- $\sigma = \sqrt{Variance} = \sqrt{0.045}$ = 0.212
- n = 210
- CL = 90% = 0.90, $Z_{\frac{\alpha}{2}}$ = 1.645.
- So a 90% CI for height = $[\bar{x} \pm Z_{0.05}\frac{\sigma}{n}]$

= [0.9 ± (1.645 × $\frac{0.212}{210}$)]
= [0.898, 0.902]

Further reading

For more questions on the subject, please go to Questions: Introduction toconfidence intervals

For information on the normal distribution and hypothesis testing, please see Guide: [insert] and Guide: [insert]

v1.0: initial version created 04/11 by mh.