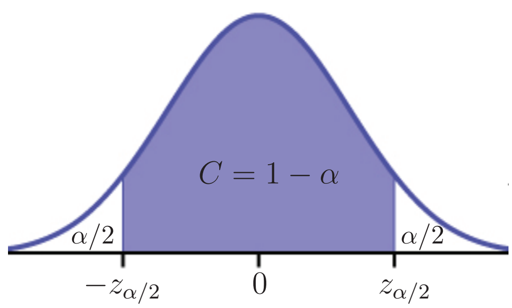
9 Confidence Intervals

In this section, we want to use the size, mean and SD of a *sample* to find an interval estimate for the mean of the *population* that the sample came from. All interval estimates are calculated by first selecting a *confidence level*, which measures the degree of certainty that the interval estimate produced using a normal distribution will contain the true population mean. The most common values for the confidence level are 90%, 95%, 98% and 99%. A confidence level of 95% means that, of all possible samples of size n taken from a population, 95% of them will give an interval estimate that contains the true population mean and 5% will not. It does not mean that the probability that the population mean is in a certain interval is 95%.

9.1 Large Samples

For random samples of size $n \geq 30$ taken from any population, the Central Limit Theorem tells us that $\bar{x}$-values are normally distributed. As a result, given a confidence level, we can use a reverse look-up on the *Standard Normal Distribution Table* to find an interval estimate for the population mean.



Since confidence levels are usually given as 90%, 95%, 98% or 99%, it is quicker to use this reverse look-up table:

$1-\alpha$	0.9	0.95	0.98	0.99
$z_{\alpha/2}$	1.645	1.960	2.326	2.576
z_α	1.282	1.645	2.054	2.326

Definition: Given a confidence level $C = 1 - \alpha$ and a random sample of size $n \geq 30$ from a population with standard deviation σ, a confidence interval (CI) for the population mean, μ, is given by

$$\bar{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}},$$

where σ may be approximated using the sample standard deviation s.
The quantity $z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ is called the margin of error (ME).

Example 9.1. A random sample of 60 cans of Coke had an average volume of 355.3 mL and a standard deviation of 2.5 mL.

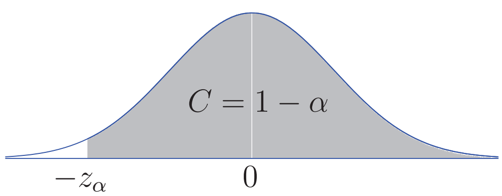
(a) Find a 95% confidence interval for the average volume among all cans of Coke.

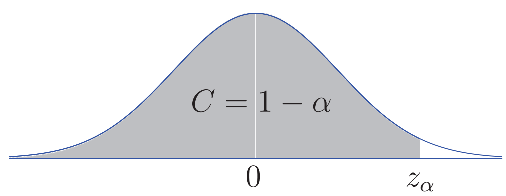
(b) Would a 99% confidence interval be wider or narrower than the 95% confidence interval in part (a)?

Example 9.2. [2, p. 233] 80 readings of daily emission (in tons) of sulfur oxides from an industrial plant had an average of 18.85 tons and a standard deviation of 5.55 tons. Use this data to construct a 99% confidence interval for the plant's true average daily emission of sulfur oxides.

Example 9.3. [3, p. 638] The thickness of a certain type of sheet metal has a known standard deviation of 0.27 mm. We want to estimate μ with a 95% margin of error of less than 0.01 mm. What is the minimum sample size n required?

Sometimes we are interested only in the lower or the upper bound on an estimate of the population mean, rather than an interval.





Definition: Given a confidence level $C = 1-\alpha$ and a random sample of size $n \geq 30$ from a population with standard deviation σ, a <u>lower confidence bound</u> (LCB) for the population mean, μ, is given as

$$\mu > \bar{x} - z_\alpha \frac{\sigma}{\sqrt{n}}$$

and an <u>upper confidence bound</u> (UCB) for the population mean, μ, is given as

$$\mu < \bar{x} + z_\alpha \frac{\sigma}{\sqrt{n}}$$

where σ may be approximated using the sample standard deviation s.

As with confidence intervals, we use the handy table given at the beginning of this section to find the z_α-values for the common confidence levels rather than doing a reverse look-up on the *Standard Normal Distribution Table*.

Example 9.4. 30 randomly selected water samples have a mean pollution concentration of 48.1 ppm with a standard deviation of 6.2 ppm. Find a 99% UCB for the mean pollution concentration in the body of water.

Example 9.5. In a large class, test marks have a SD of 10.3. A random sample of 40 tests has an average mark of 69.1. Find a 98% LCB for the class average.

9.2 Small Samples

If the sample size is $n < 30$, we need to know that the population has a normal distribution to be able to calculate confidence intervals. Moreover, the sample size affects the shape of the bell curve for small values of n, so a new parameter called the *degrees of freedom*,

$$df = n - 1,$$

is needed as well. As a result, the probabilities are not associated with the standard normal distribution, and we use the *t-Distribution Table* instead.

Definition: Given a confidence level $C = 1 - \alpha$ and a random sample with size $n < 30$ and standard deviation s from a normally distributed population, a confidence interval (CI) for the population mean, μ, is given by

$$\bar{x} \pm t_{\alpha/2}\frac{s}{\sqrt{n}}.$$

The quantity $t_{\alpha/2}\frac{s}{\sqrt{n}}$ is called the margin of error (ME).
A lower confidence bound (LCB) for the population mean, μ, is given as

$$\mu > \bar{x} - t_{\alpha}\frac{s}{\sqrt{n}}$$

and an upper confidence bound (UCB) for the population mean, μ, is given as

$$\mu < \bar{x} + t_{\alpha}\frac{s}{\sqrt{n}}.$$

Demonstration: Confidence Intervals

Example 9.6. [2, p. 237] Ten bearings made by a certain process have a mean diameter of 0.5060 cm and a standard deviation of 0.0040 cm. Assuming that the data is a random sample from a normal population, construct a 95% confidence interval for the actual average diameter of bearings made by this process.

Example 9.7. [3, p. 640] A sample of 15 washing machines of a certain brand had a mean replacement time of 9.1 years, with a standard deviation of 2.7 years. Find a 90% lower confidence bound for the mean replacement time of all washing machines of this brand.

Example 9.8. [3, p. 640] A test station measured the loudness of a random sample of 22 jets taking off from a certain airport. The mean was found to be 107.2 dB, with a standard deviation of 9.2 dB. Find a 95% upper confidence bound for the mean loudness of all jets taking off from this airport.

Additional Notes

Additional Notes