**Interval Estimates**

Table of Contents

[Mean of Normal Population with Known 3](#_Toc64281801)

[One-Sided Confidence Interval 7](#_Toc64281802)

[Mean of Normal Population with Unknown 8](#_Toc64281803)

[Variance of a Normal Population 9](#_Toc64281804)

[Difference of Means of Two Normal Populations 10](#_Toc64281805)

[Sample Sizes 11](#_Toc64281806)

We have already learnt about estimating parameters using sample means and sample variances, and even with MLE. The problem with these was that they were point estimates, meaning they gave a single value. This estimated value will most likely not be exactly the same as the true value of the parameter.

In some situations, instead of giving a single estimated value, if we could state an interval for the estimated value in which we think the true value of the parameter will fall, we will be more confident about our results. For example, if we are asked the estimate price of a laptop, it is nearly impossible to give an accurate answer, since the prices vary so much. It would be much more useful to state an interval in which the values may lie. Such estimates are called interval estimates.

When discussing interval estimates, we deal with two values. Firstly, we have the interval, which is the range for the estimate. Secondly, we have the confidence level, which tells us the level of certainty we have that the true value lies in this range.

To calculate the interval, we first find a point estimate and we state our interval. There can be three types of intervals. The first type is where we state limits on both sides of the point estimate, essentially giving a range. This is called a two-sided confidence interval. The second type is where we state an upper limit, but not a lower limit. This is called a lower confidence interval. The third type is where we state a lower limit, but not an upper limit. This is called an upper confidence interval.

The other issue is the confidence level. Essentially, the wider the interval, the higher our confidence level, since it is more likely that the true value of the parameter falls into that range.

One of the variables, either the confidence level or the interval, is usually given to us and we must find the other.

## Mean of Normal Population with Known

If we take a random sample of size , then the sample mean will be given by

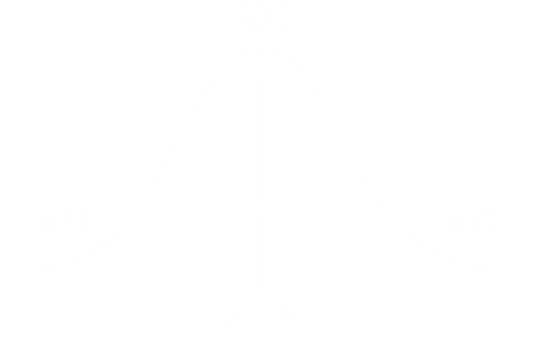
This can be represented using a random variable.

Using the Central Limit Theorem, we know that summing up a large number of random variables that have similar distributions will give us a random variable that has approximately a normal distribution. However, since we are summing up random variables that are normally distributed in this case, as we have previously seen, the result will be a random variable that is exactly normally distributed.

If we standardize this,

We want to find an interval estimation of .

Since we have standardized the random variable, we can use the graph for the PDF of a standard normal random variable. From here, we will find the interval. For now, let’s assume we want a two-sided interval. Now the question is, what is the confidence level. For some value of , we want to ensure that of the values are within the interval.

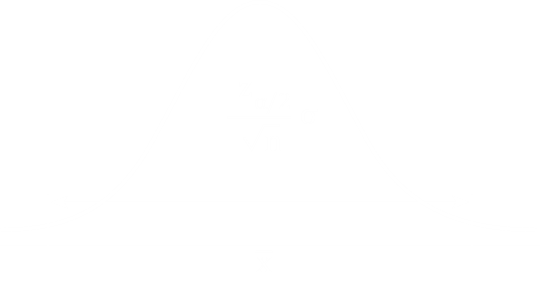


We know that

Say we want to have a confidence interval. Thus, .

From here,

If we draw the standard normal curve for this, it will look like this:



Now, if someone were to ask us to find the interval estimate of the mean packet delay from a normal population having , where we collected pieces of data and found , we can easily do so.

, so

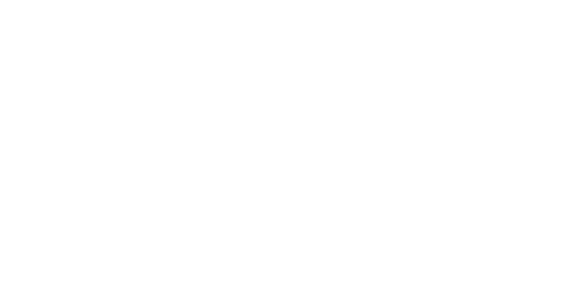
There is one issue. We said that

However, from the frequentist perspective, is a parameter with some fixed value that we do not know. As such, there cannot be a probability that has a value within this range. For there to be a probability, would have to be a random variable. What this actually tells us is that if we take random samples a very large number of times, of those times, we will find that the value of is within this range. It is still entirely possible that the value of is sometimes outside of this range.

## One-Sided Confidence Interval

The upper confidence interval starts at a lower bound and has no upper bound, while a lower confidence interval starts at an upper bound and has no lower bound.

First, let us consider the upper confidence interval. We need to find the lower bound.



From the diagram, we can tell that of the time, the value will be within our interval.

Thus, the one-sided upper CI is

Similarly, we could have had a one-sided lower CI,

## Mean of Normal Population with Unknown

We know that the sample variance can be calculated from the sample data as

We also know that standardizing this gives us a distribution with degrees of freedom.

If we follow the same pattern to find the limits, we will see that

Of course, we can also have upper and lower confidence levels.

## Variance of a Normal Population

We know that

We can solve this in a similar manner to the previous ones. However, one problem here is that the PDF curve for a chi-square distribution is not symmetrical. Due to this, for two-sided intervals, we will need to calculate two separate values, and , since we cannot use the symmetric property as we do with normal distributions.



Thus, the CI for variance is

## Difference of Means of Two Normal Populations

We have seen the details of how this problem works while studying hypothesis testing, so we will jump directly into the solution now.

We know that

Thus,

We can calculate the confidence intervals from here.

One issue is that, here, we are assuming that and are known. As we saw with hypothesis testing, depending on whether or not these two values are known, we may need to calculate the sample variance and the distribution will be either a standard normal distribution or a distribution depending on that.

## Sample Sizes

We know that, for a specific value of , we will get some confidence level. The interval is found from the equation

Say we are given a specific confidence level. As seen above, we will find an interval for this confidence level. However, say we are asked to increase our confidence level without changing the size of the interval. This is where the size of the sample, , becomes important. We can increase the confidence level without changing the interval by changing the value of .

Think about why this makes sense. Say we are collecting data about the heights of people in a group. If we collect all the data, we can say with certainty what the mean height is. This is no interval involved at all. We increased the value of so much that we included the entire group. The more data we collect, the more accurate our estimation.