# MTL108
# Interval Estimation and Confidence Intervals

Rahul Singh

While a point estimate ($\hat{\theta}$) provides our single "best guess" for an unknown population parameter ($\theta$), it is almost certainly wrong. It provides no information about the precision (variance or concentration) of the estimate.

Informally, a confidence interval (CI) is a range of estimates for an unknown population parameter, computed from sample data. Instead of providing a single point estimate, it provides an interval of plausible values to formally account for sampling error and uncertainty. The method of obtaining CI is known as interval estimation.

**Definition 1** (Confidence Interval)**.** Let $X_1, X_2, \ldots, X_n$ be a random sample drawn from a probability distribution that depends on an unknown population parameter $\theta$. A $100(1-\alpha)\%$ confidence interval for $\theta$ consists of two statistics, a lower bound $L(X_1, \ldots, X_n)$ and an upper bound $U(X_1, \ldots, X_n)$, such that the probability of the random interval $[L, U]$ containing the true parameter $\theta$ is exactly $1-\alpha$ prior to observing the data:

$$P(L \leq \theta \leq U) = 1 - \alpha$$

Here, the value $1-\alpha$ is called the *confidence level* (typically set to 0.90, 0.95, or 0.99), and $\alpha$ represents the significance level.

**The Frequentist Interpretation:**

It is a common and fundamental mathematical error to say, "There is a 95% probability that the true parameter $\theta$ falls within our calculated interval." In classical (frequentist) statistics, the parameter $\theta$ is a fixed, deterministic constant—it does not move, and it is not a random variable. Therefore, it is either 100% inside your specific interval or 100% outside of it.

The probability $1-\alpha$ refers to the *reliability of the procedure itself*, not the specific numbers calculated from one sample. The correct statistical interpretation is

> "If we were to repeat this exact sampling procedure an infinite number of times, constructing a new confidence interval from each new sample, exactly $100(1-\alpha)\%$ of those constructed intervals would successfully cover the true, fixed population parameter $\theta$."

## Pivotal quantity

**Definition 2.** In statistical inference, a **pivotal quantity** is a function of both the sample observations $(X_1, X_2, \ldots, X_n)$ and the unknown population parameter $(\theta)$, such that its probability distribution is **completely known** and does not depend on $\theta$ or any other unknown parameters.

Mathematically, if $Q = g(X_1, \ldots, X_n; \theta)$, then $Q$ is a pivotal quantity if the distribution of $Q$ is identical for all valid values of $\theta$ in the parameter space.

**Example 1** (Normal distribution with known Variance)**.** Let $X_1, \ldots, X_n$ be a random sample from a Normal distribution $N(\mu, \sigma^2)$, where the true mean $\mu$ is unknown, but the variance $\sigma^2$ is miraculously known.
The sample mean $\bar{X}$ has the distribution $\bar{X} \sim N(\mu, \sigma^2/n)$. Notice that the distribution of $\bar{X}$ depends heavily on the unknown parameter $\mu$. Therefore, $\bar{X}$ itself is *not* a pivotal quantity.
However, if we standardize $\bar{X}$, define

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}.$$

This function $Z$ contains both our data ($\bar{X}$) and our unknown parameter ($\mu$). By definition, its distribution is Standard Normal: $Z \sim N(0, 1)$. Because the $N(0, 1)$ distribution has a fixed mean of 0 and a variance of 1, its probability curve is entirely independent of the unknown $\mu$. Therefore, $Z$ is a perfect pivotal quantity.

**Example 2** (Normal distribution with unknown Variance)**.** In the real world, the population variance $\sigma^2$ is almost always unknown. If we don't know $\sigma$, the $Z$-statistic above fails to be a pivotal quantity because we cannot calculate its value without knowing all the underlying parameters.
Instead, we substitute the sample standard deviation $S$ for $\sigma$. Define

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

This function $T$ relies only on sample data ($\bar{X}$ and $S$) and the target parameter ($\mu$). By statistical theory, its distribution is exactly the Student's $t$-distribution with $n - 1$ degrees of freedom: $T \sim t_{n-1}$. Because the $t_{n-1}$ distribution depends strictly on the sample size $n$ and is completely free of both $\mu$ and $\sigma^2$, $T$ acts as a highly useful pivotal quantity when variance is unknown.

**Remark 1.** Pivotal quantity are the fundamental algebraic engines used to construct confidence intervals. Because the probability distribution of $Q$ is completely independent of the unknown parameter, we can easily find exact critical values (from the known distribution). We then algebraically invert the inequality to isolate the unknown parameter $\theta$ in the center.

## CI for a Single Population Mean ($\mu$)

## Case A: Population Variance ($\sigma^2$) is Known

If the underlying population is Normal, or if the sample size is large ($n \geq 30$) allowing the Central Limit Theorem to apply, the sample mean $\bar{X}$ is normally distributed. Our pivotal quantity is the standard Normal $Z$-statistic,

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$$

To find a $100(1-\alpha)\%$ CI, we bound this pivotal quantity between the critical values $-z_{\alpha/2}$ and $z_{\alpha/2}$,

$$P\left(-z_{\alpha/2} \le \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \le z_{\alpha/2}\right) = 1-\alpha$$

Rearranging the inequality to isolate $\mu$ in the center, we have

$$P\left(\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$$

**Conclusion:**

$$\text{CI is } \left(\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right).$$

This is also presented as

$$\text{CI} = \bar{X} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}.$$

**Example 3** (High-Precision CNC Machining)**.** A factory produces titanium shafts for aerospace engines. The CNC machines are rigorously calibrated such that the standard deviation of the shaft diameters is known to be exactly $\sigma = 0.02$ mm. An engineer takes a random sample of $n = 16$ shafts and measures a sample mean diameter of $\bar{x} = 15.40$ mm.

**Task:** Construct a 95% CI for the true mean diameter ($\alpha = 0.05 \implies z_{0.025} = 1.96$).
Using the preceding result, we have

$$\text{CI} = 15.40 \pm 1.96\left(\frac{0.02}{\sqrt{16}}\right) = 15.40 \pm 1.96(0.005) = 15.40 \pm 0.0098$$

*Result:* So, the CI is $[15.3902, 15.4098]$ mm.

## Case B: Population Variance ($\sigma^2$) is Unknown

In reality, $\sigma$ is rarely known. We must estimate it using the sample standard deviation, which introduces additional uncertainty.
**Setup:**
Let $X_1, X_2, \ldots, X_n$ be a random sample of size $n$ drawn from a Normal population $N(\mu, \sigma^2)$, where both the population mean $\mu$ and the population variance $\sigma^2$ are unknown. Our objective is to construct a $100(1-\alpha)\%$ confidence interval for $\mu$.
**The Sample Statistics**
Because the population parameters are unknown, motivated by the $t$-distribution definition, we consider the following estimators from the sample data.

- Sample Mean: $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$
- Sample Variance: $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$

**The Pivotal Quantity**
As $\sigma^2$ is unknown, the quantity $Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$ is no more pivotal.

Recall, the ratio of a Standard Normal random variable to the square root of an independent Chi-Square random variable (divided by its degrees of freedom) follows a Student's $t$-distribution. Therefore, our pivotal quantity is

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t_{n-1}$$

Notice that this is a valid pivotal quantity because its distribution ($t$ with $n-1$ degrees of freedom) is completely known and does not depend on the unknown parameters $\mu$ or $\sigma^2$.

**The Probability Statement**

Because the $t$-distribution is perfectly symmetric around zero, we select the critical values $-t_{\alpha/2,n-1}$ and $t_{\alpha/2,n-1}$, such that

$$P\left(-t_{\alpha/2,n-1} \leq T = \frac{\bar{X} - \mu}{S/\sqrt{n}} \leq t_{\alpha/2,n-1}\right) = 1 - \alpha$$

**Algebraic Inversion**

We now use basic algebra to manipulate the inequality inside the probability statement to isolate the unknown parameter $\mu$ strictly in the center. First, multiply all parts of the inequality by the standard error of the mean, $S/\sqrt{n}$, we get

$$P\left(-t_{\alpha/2,n-1}\frac{S}{\sqrt{n}} \leq \bar{X} - \mu \leq t_{\alpha/2,n-1}\frac{S}{\sqrt{n}}\right) = 1 - \alpha$$

Next, subtracting the sample mean $\bar{X}$ from all parts of the inequality, we have

$$P\left(-\bar{X} - t_{\alpha/2,n-1}\frac{S}{\sqrt{n}} \leq -\mu \leq -\bar{X} + t_{\alpha/2,n-1}\frac{S}{\sqrt{n}}\right) = 1 - \alpha$$

Finally, multiply the entire inequality by $-1$. Remember that multiplying an inequality by a negative number reverses the direction of the inequality signs,

$$P\left(\bar{X} + t_{\alpha/2,n-1}\frac{S}{\sqrt{n}} \geq \mu \geq \bar{X} - t_{\alpha/2,n-1}\frac{S}{\sqrt{n}}\right) = 1 - \alpha$$

Rearranging it to read naturally from left to right (smallest to largest),

$$P\left(\bar{X} - t_{\alpha/2,n-1}\frac{S}{\sqrt{n}} \leq \mu \leq \bar{X} + t_{\alpha/2,n-1}\frac{S}{\sqrt{n}}\right) = 1 - \alpha$$

**The Final Confidence Interval**

The algebraic manipulation shows that the procedure will successfully bracket the true mean $\mu$ exactly $100(1-\alpha)\%$ of the time. The calculated bounds form our confidence interval,

$$\text{CI} = \left[\bar{X} - t_{\alpha/2,n-1}\frac{S}{\sqrt{n}}, \quad \bar{X} + t_{\alpha/2,n-1}\frac{S}{\sqrt{n}}\right]$$

Which is commonly written in the shorthand notation,

$$\bar{X} \pm t_{\alpha/2,n-1}\frac{S}{\sqrt{n}}.$$

**Example 4** (Environmental Toxicity Testing). *Scenario:* An environmental scientist is testing the lead concentration in a local municipal water supply. They take a small sample of $n = 10$ vials. The sample mean is $\bar{x} = 14.5$ ppb (parts per billion), with a sample standard deviation of $s = 2.1$ ppb.

**Task:** Construct a 99% CI for the true mean lead concentration ($\alpha = 0.01$, $df = 9 \implies t_{0.005,9} = 3.250$).
So using above mentioned approach we get

$$\text{CI} = 14.5 \pm 3.250 \left( \frac{2.1}{\sqrt{10}} \right) = 14.5 \pm 3.250(0.664) = 14.5 \pm 2.158$$

*Result:* The CI is $[12.34, 16.66]$ ppb.

## Confidence Interval for a Population Proportion ($p$)

In many practical scenarios, we are not measuring a continuous variable (like height or weight) but rather a categorical outcome (like success/failure, defective/non-defective, or voting preferences). In these cases, the parameter of interest is the **true population proportion**, denoted by $p$.

### 1. The Point Estimator and its Distribution

Given a random sample of size $n$, let $X$ be the number of "successes" in the sample. The random variable $X$ follows a Binomial distribution: $X \sim \text{Binomial}(n, p)$.

The best unbiased point estimator for the population proportion $p$ is the sample proportion, denoted by $\hat{p}$ (read as "p-hat"):

$$\hat{p} = \frac{X}{n}.$$

We have

- **Expected Value:** $E[\hat{p}] = E\left[\frac{X}{n}\right] = \frac{1}{n}E[X] = \frac{np}{n} = p$. (This proves $\hat{p}$ is an unbiased estimator).
- **Variance:** $\text{Var}(\hat{p}) = \text{Var}\left(\frac{X}{n}\right) = \frac{1}{n^2}\text{Var}(X) = \frac{np(1-p)}{n^2} = \frac{p(1-p)}{n}$.

The **Standard Error** of the proportion is therefore $\text{SE}(\hat{p}) = \sqrt{\frac{p(1-p)}{n}}$.

### 2. The Normal Approximation and Pivotal Quantity

To construct a confidence interval, we need a pivotal quantity. While the exact distribution of $X$ is Binomial, working with Binomial sums for large sample sizes is computationally difficult.

Instead, we rely on the **Central Limit Theorem (CLT)**. If the sample size is sufficiently large, the sampling distribution of $\hat{p}$ is approximately Normal.
*A Rule of Thumb for "Large Enough":* We require at least 5 expected successes and 5 expected failures:

$$n\hat{p} \geq 5 \quad \text{and} \quad n(1 - \hat{p}) \geq 5$$

If these conditions are met, we can construct the following approximate Standard Normal quantity,

$$Z = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} \approx N(0, 1).$$

However, the denominator relies on the unknown parameter $p$, which is exactly what we are trying to estimate! To solve this, we substitute our point estimate $\hat{p}$ into the standard error formula. This yields an approximate CI known as the **Wald Confidence Interval**.

### 3. Deriving the Wald Confidence Interval

We set up the probability bound using the standard normal critical values, $z_{\alpha/2}$,

$$P\left(-z_{\alpha/2} \le \frac{\hat{p} - p}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}} \le z_{\alpha/2}\right) \approx 1 - \alpha$$

Rearranging the inequality to isolate the unknown parameter $p$ in the center yields,

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

Where the term $z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ is defined as the **Margin of Error (E)**.

### 4. Determining Required Sample Size

A common question in experimental design is: *"How large of a sample do I need to achieve a specific margin of error E?"* We take the Margin of Error formula and solve algebraically for $n$,

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

Because we do not know the true $p$ before conducting the study, we must use a planning value.

- **Historical Estimate:** Use a previous study's $\hat{p}$.
- **Conservative Approach:** If we have no idea, we use $p = 0.5$. The function $p(1-p)$ reaches its absolute mathematical maximum at $p = 0.5$. Using this value guarantees that our calculated sample size $n$ will be large enough regardless of what the true proportion actually is.

**Example 5** (Election Polling)**.** A news agency wants to predict the outcome of an upcoming mayoral election. They survey a random sample of $n = 800$ likely voters. In the sample, $X = 432$ voters state they will vote for Candidate A.
**Task 1: Calculate the 95% Confidence Interval** First, calculate the sample proportion ($\hat{p}$):

$$\hat{p} = \frac{432}{800} = 0.54 \text{ (or 54\%)}$$

Verify the conditions: $n\hat{p} = 432 \geq 5$ and $n(1-\hat{p}) = 368 \geq 5$. The normal approximation is valid. For a 95% confidence level, $\alpha = 0.05$, so $z_{0.025} = 1.96$.

$$\text{SE}(\hat{p}) = \sqrt{\frac{0.54(1-0.54)}{800}} = \sqrt{\frac{0.2484}{800}} = \sqrt{0.0003105} \approx 0.0176$$

$$\text{Margin of Error (E)} = 1.96 \times 0.0176 \approx 0.0345 \text{ (or 3.45\%)}$$

$$\text{CI} = 0.54 \pm 0.0345 = [0.5055, 0.5745]$$

**Conclusion:** The news agency is 95% confident that Candidate A's true share of the vote is between 50.55% and 57.45%. Because the entire interval is strictly greater than 50%, they can confidently predict Candidate A will win the election.
**Task 2: Reducing the Margin of Error** The news agency thinks a 3.45% margin of error is too wide. For their next poll, they want a strict Margin of Error of exactly 2% ($E = 0.02$) at the 99% confidence level ($z_{0.005} = 2.576$). Using the conservative estimate ($p = 0.5$), how many voters must they survey?

$$n = \left(\frac{2.576}{0.02}\right)^2 (0.5)(1-0.5) = (128.8)^2(0.25) = 16589.44 \times 0.25 \approx 4147.36$$

They must survey at least 4,148 voters to achieve this highly precise margin of error.

**Example 6** (Server Uptime Guarantees)**.** A cloud computing company wants to estimate the proportion of time its servers experience latency exceeding 50ms. They want to be 95% confident that their estimate is within a Margin of Error of $E = 2\%$ (0.02). How many ping tests must they run?

$$n = \left(\frac{1.96}{0.02}\right)^2 (0.5)(1-0.5) = (98)^2(0.25) = 9604 \times 0.25 = 2401$$

*Result:* The automated testing script must ping the servers at least 2,401 times to guarantee this level of statistical precision.

## Confidence Interval for a Population Variance ($\sigma^2$)

In many applied fields, minimizing variability is just as important as hitting the target mean. For example, in manufacturing, parts must be strictly uniform; in finance, variance represents risk; and in pharmaceuticals, inconsistent dosages can be dangerous. To quantify the precision of a process, we construct a confidence interval for the true population variance, $\sigma^2$.

### 1. Crucial Assumption: Normality

Unlike confidence intervals for the mean ($\mu$), where the Central Limit Theorem (CLT) bails us out for large sample sizes even if the data is skewed, the confidence interval for variance is **highly sensitive to departures from normality**. We must strictly assume that the underlying population is Normally distributed: $X \sim N(\mu, \sigma^2)$.

**2. The Pivotal Quantity**

To construct the interval, we need a pivotal quantity that relates our sample variance ($S^2$) to the population variance ($\sigma^2$). Statistical theory proves that if we scale the sample variance appropriately, it follows a Chi-Square ($\chi^2$) distribution with $n-1$ degrees of freedom. Therefore the pivotal quantity is

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}.$$

**3. Understanding the Asymmetry**

Unlike the Normal ($Z$) or Student's ($t$) distributions, the Chi-Square distribution is strictly positive (since variances cannot be negative) and heavily right-skewed. Because it is not symmetric around zero, we cannot use a simple "±" formula. We must find two distinct critical values from the $\chi^2$ table that trap the middle $(1-\alpha)$ probability,

- **Lower Critical Value:** $\chi^2_{1-\alpha/2,n-1}$ (The value with $1-\alpha/2$ area to its right).
- **Upper Critical Value:** $\chi^2_{\alpha/2,n-1}$ (The value with $\alpha/2$ area to its right).

**4. The CI**

We begin with the probability statement bounding our pivotal quantity,

$$P\left(\chi^2_{1-\alpha/2,n-1} \leq \frac{(n-1)S^2}{\sigma^2} \leq \chi^2_{\alpha/2,n-1}\right) = 1-\alpha$$

To isolate $\sigma^2$ in the middle, we first take the reciprocal of all three terms. *(Note: Taking the reciprocal flips the direction of the inequality signs)*,

$$P\left(\frac{1}{\chi^2_{1-\alpha/2,n-1}} \geq \frac{\sigma^2}{(n-1)S^2} \geq \frac{1}{\chi^2_{\alpha/2,n-1}}\right) = 1-\alpha$$

Next, we multiply all terms by $(n-1)S^2$ to completely isolate $\sigma^2$,

$$P\left(\frac{(n-1)S^2}{\chi^2_{1-\alpha/2,n-1}} \geq \sigma^2 \geq \frac{(n-1)S^2}{\chi^2_{\alpha/2,n-1}}\right) = 1-\alpha$$

Finally, we rewrite the inequality in the standard left-to-right (smallest to largest) format. **Notice that dividing by the larger upper-tail critical value yields the lower bound of the interval.**

**Final Formula for the $100(1-\alpha)\%$ CI for $\sigma^2$:**

$$\text{CI} = \left(\frac{(n-1)S^2}{\chi^2_{\alpha/2,n-1}},\quad \frac{(n-1)S^2}{\chi^2_{1-\alpha/2,n-1}}\right)$$

**Remark 2.** Note: To find the confidence interval for the standard deviation $\sigma$, simply take the square root of both the lower and upper bounds of this interval.

**Example 7** ( Drone Motor Manufacturing)**.** An aerospace startup is manufacturing high-speed brushless motors for racing drones. For the drone to remain stable in flight, the RPM output across all four motors must be nearly identical. The engineering specifications state that the variance in motor RPMs must not exceed 150 $\text{RPM}^2$.
A quality control engineer tests a random sample of $n = 15$ motors from the assembly line. The sample yields a mean of $\bar{x} = 24,000$ RPM and a sample variance of $s^2 = 95$ $\text{RPM}^2$.
**Task:**
Construct a 95% confidence interval for the true population variance ($\sigma^2$) to determine if the manufacturing process is reliably meeting the stability specification.
**Solution:**

1. **Identify Parameters:** $n = 15$, $s^2 = 95$, $\alpha = 0.05$.
2. **Degrees of Freedom:** $df = n - 1 = 14$.
3. **Find Critical Values:** Using a Chi-Square table for $df = 14$:
   - Upper tail ($\alpha/2 = 0.025$): $\chi^2_{0.025,14} = 26.119$
   - Lower tail ($1 - \alpha/2 = 0.975$): $\chi^2_{0.975,14} = 5.629$
4. **Calculate Bounds:**

$$\text{Lower Bound} = \frac{(15-1)(95)}{26.119} = \frac{1330}{26.119} \approx 50.92$$

$$\text{Upper Bound} = \frac{(15-1)(95)}{5.629} = \frac{1330}{5.629} \approx 236.28$$

**Result and Interpretation:**
The 95% confidence interval for the true population variance is $[50.92, 236.28]$ $\text{RPM}^2$.
Even though the sample variance (95) is well below the maximum allowable limit (150), the quality engineer **should not** certify the batch. Because the upper bound of the confidence interval (236.28) far exceeds the 150 $\text{RPM}^2$ limit, there is not enough statistical evidence to confidently guarantee the motors meet the strict stability requirements. The sample size must be increased, or the manufacturing tolerances must be tightened.

**Example 8** (Pharmaceutical Quality Assurance)**.** The FDA requires that the variance of the active ingredient in a specific blood pressure medication not exceed 1.5 $\text{mg}^2$ to prevent accidental overdoses. A QA technician tests $n = 20$ pills and finds a sample variance of $s^2 = 1.1$ $\text{mg}^2$.
*Task:* Construct a 90% CI for the true variance ($\alpha = 0.10$, $df = 19$). We look up $\chi^2_{0.05,19} = 30.144$ and $\chi^2_{0.95,19} = 10.117$.

$$\text{CI} = \left( \frac{(19)(1.1)}{30.144}, \quad \frac{(19)(1.1)}{10.117} \right) = \left( \frac{20.9}{30.144}, \quad \frac{20.9}{10.117} \right)$$

*Result:* $[0.693, 2.066]$ $\text{mg}^2$. Even though the sample variance (1.1) is well below the FDA limit (1.5), the upper bound of the confidence interval (2.066) exceeds it. The batch fails QA because there is not enough statistical confidence that the true population variance is safe.

## CI for the Difference Between Two Means ($\mu_1 - \mu_2$)

When conducting comparative studies (e.g., treatment versus control, or comparing two manufacturing processes), our goal is to estimate the true difference between two population means, $\mu_1 - \mu_2$.

Let $X_1, \ldots, X_{n_1}$ be a random sample from a population with mean $\mu_1$ and variance $\sigma_1^2$. Let $Y_1, \ldots, Y_{n_2}$ be an independent random sample from a population with mean $\mu_2$ and variance $\sigma_2^2$.

The most logical point estimator for the difference in population means is the difference in sample means,

$$\bar{X} - \bar{Y}.$$

By the linearity of expectation and the properties of independent variances, the sampling distribution of this point estimator has the following properties:

$$E[\bar{X} - \bar{Y}] = E[\bar{X}] - E[\bar{Y}] = \mu_1 - \mu_2$$

$$\text{Var}(\bar{X} - \bar{Y}) = \text{Var}(\bar{X}) + \text{Var}(\bar{Y}) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

### Case 1: Both Population Variances ($\sigma_1^2, \sigma_2^2$) are Known

If both populations are normally distributed (or if $n_1, n_2 \geq 30$ so the Central Limit Theorem applies), the difference $\bar{X} - \bar{Y}$ is normally distributed. Because the true variances are known, we use the Standard Normal ($Z$) distribution to construct our pivotal quantity,

$$Z = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0, 1)$$

By bounding this pivot between $-z_{\alpha/2}$ and $z_{\alpha/2}$ and solving for $(\mu_1 - \mu_2)$, we obtain the confidence interval.

**Formula:**

$$\text{CI} = (\bar{X} - \bar{Y}) \pm z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

### Case 2: Variances Unknown but Assumed Equal ($\sigma_1^2 = \sigma_2^2 = \sigma^2$)

This is a very common scenario in controlled experiments where a treatment might shift the mean but is not expected to change the inherent variability of the subjects.

Since the two unknown variances are mathematically identical, calculating separate sample variances ($S_1^2$ and $S_2^2$) is inefficient. Instead, we "pool" the two sample variances together to create a single, highly efficient unbiased estimator of the common variance $\sigma^2$. We weight each sample variance by its degrees of freedom,

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

Because we are estimating the variance from the data, the pivotal quantity shifts from the $Z$-distribution to the Student's $t$-distribution, utilizing the combined degrees of freedom from both samples $(n_1 + n_2 - 2)$,

$$T = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{n_1+n_2-2}$$

**Formula:**

$$\text{CI} = (\bar{X} - \bar{Y}) \pm t_{\alpha/2,\, n_1+n_2-2}\, S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

**Example 9** (Materials Engineering). A civil engineer is testing two different concrete curing methods to see which yields a higher compressive strength (in MPa).

- Method 1 (Standard): $n_1 = 12, \bar{x}_1 = 32.5, s_1^2 = 4.2$
- Method 2 (New Formula): $n_2 = 10, \bar{x}_2 = 35.8, s_2^2 = 3.8$

First, calculate the pooled variance:

$$S_p^2 = \frac{(11)(4.2) + (9)(3.8)}{12 + 10 - 2} = \frac{46.2 + 34.2}{20} = \frac{80.4}{20} = 4.02$$

Calculate the 95% CI for $\mu_1 - \mu_2$ ($df = 20, t_{0.025,20} = 2.086$):

$$\text{CI} = (32.5 - 35.8) \pm 2.086\sqrt{4.02\left(\frac{1}{12} + \frac{1}{10}\right)}$$

$$\text{CI} = -3.3 \pm 2.086\sqrt{4.02(0.1833)} = -3.3 \pm 2.086(0.858) = -3.3 \pm 1.79$$

*Result:* $[-5.09, -1.51]$ MPa. Because the entire interval is strictly negative, we can conclude with 95% confidence that the New Formula (Method 2) is definitively stronger than the standard method by somewhere between 1.51 and 5.09 MPa.

**Example 10** (Clinical Trial for Treatment Efficacy (Hypertension)). A pharmaceutical company conducts a randomized clinical trial to test the efficacy of a newly developed hypertension medication (Drug X) against the current standard-of-care ACE inhibitor (Control). The primary endpoint is the reduction in systolic blood pressure (measured in mmHg) after 8 weeks of daily treatment.
The researchers assume that while the new drug might lower blood pressure more effectively, the biological variation (variance) in patient responses will be similar for both drugs ($\sigma_1^2 = \sigma_2^2$).

- **Treatment Group (Drug X):** $n_1 = 40$ patients, mean reduction $\bar{x}_1 = 18.5$ mmHg, sample standard deviation $s_1 = 4.2$ mmHg.
- **Control Group (Standard):** $n_2 = 40$ patients, mean reduction $\bar{x}_2 = 13.2$ mmHg, sample standard deviation $s_2 = 4.6$ mmHg.

*Task:* Construct a 95% Confidence Interval for the true difference in mean blood pressure reduction ($\mu_1 - \mu_2$).

**Step 1: Calculate the Pooled Variance ($S_p^2$)** Because the sample sizes are strictly equal ($n_1 = n_2$), the pooled variance will simply be the exact average of the two sample variances. We calculate it formally to demonstrate the method:

$$S_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} = \frac{(39)(4.2)^2 + (39)(4.6)^2}{40 + 40 - 2}$$

$$S_p^2 = \frac{39(17.64) + 39(21.16)}{78} = \frac{687.96 + 825.24}{78} = \frac{1513.2}{78} = 19.4$$

The pooled standard deviation is $S_p = \sqrt{19.4} \approx 4.405$.

**Step 2: Find the Critical $t$-Value** The total degrees of freedom is $df = n_1 + n_2 - 2 = 78$. For a 95% confidence level ($\alpha = 0.05$), we look up the two-tailed critical value: $t_{0.025,\,78} \approx 1.99$.

**Step 3: Calculate the Confidence Interval** We plug our values into the pooled variance CI formula:

$$\text{CI} = (\bar{x}_1 - \bar{x}_2) \pm t_{0.025,\,78}\, S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$\text{CI} = (18.5 - 13.2) \pm 1.99(4.405)\sqrt{\frac{1}{40} + \frac{1}{40}}$$

$$\text{CI} = 5.3 \pm 1.99(4.405)\sqrt{0.05}$$

$$\text{CI} = 5.3 \pm 1.99(4.405)(0.2236) = 5.3 \pm 1.96$$

**Final Result:** $[3.34, 7.26]$ mmHg.

**Clinical Interpretation:** The FDA regulators evaluate this interval and note that it is strictly positive—the value 0 is nowhere near the interval. Therefore, the researchers can conclude with 95% statistical confidence that Drug X is more efficacious than the current standard of care. Specifically, Drug X reduces systolic blood pressure by an additional 3.34 to 7.26 mmHg on average compared to the standard treatment.

# References

[1] Blitzstein, J. K., & Hwang, J. (2019). *Introduction to probability.* Chapman and Hall/CRC.

[2] Ross, Sheldon M. (2020). *Introduction to probability and statistics for engineers and scientists.* Academic press.

### Disclaimer