Confidence intervals for unknown means

Chrysafis Vogiatzis

Department of Industrial and Enterprise Systems Engineering
University of Illinois at Urbana-Champaign

Lecture 20

ILLINOIS
ISE | Industrial & Enterprise Systems Engineering
GRAINGER COLLEGE OF ENGINEERING

Motivation

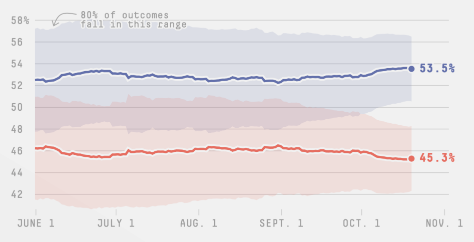
During these past weeks, we introduced **point estimation** and **point estimators**.

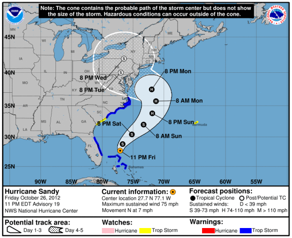
- Looking for an unknown mean? Report the sample average.
  - The bigger the sample, the smaller its mean square error.
- Looking for an unknown distribution parameter θ? Use an estimation method and report its point estimate $\hat{\theta}$.

All these items have one thing in common: to address uncertainty, we resort to reporting a single value.

Motivation

Typically, we prefer to report **intervals** of estimation:





In these next lectures, we will have a full discussion about one such mean of interval estimation, referred to as a **confidence interval**.

ILLINOIS

Point vs. interval estimation

- **Point** estimation: a *single* estimate with our best guess at what the unknown parameter is.
- **Interval** estimation: an interval of values where the unknown parameter is believed to belong in.

Advantage of interval estimation:

- A point estimate reveals a single point, and we have no idea of how close the actual parameter may or may not be:

Point vs. interval estimation

- **Point** estimation: a *single* estimate with our best guess at what the unknown parameter is.
- **Interval** estimation: an interval of values where the unknown parameter is believed to belong in.

Advantage of interval estimation:

1. A point estimate reveals a single point, and we have no idea of how close the actual parameter may or may not be:

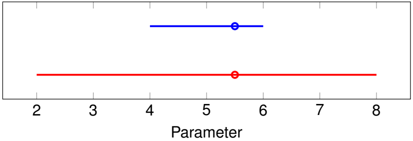
2. An interval estimate reveals a "margin of error" as a measure of accuracy for our parameter.

Point vs. interval estimation

- **Point** estimation: a *single* estimate with our best guess at what the unknown parameter is.
- **Interval** estimation: an interval of values where the unknown parameter is believed to belong in.

Advantage of interval estimation:

1. A point estimate reveals a single point, and we have no idea of how close the actual parameter may or may not be:



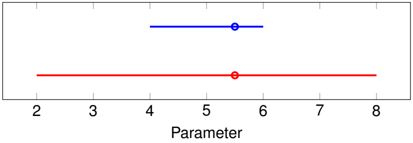
2. An interval estimate reveals a “margin of error” as a measure of accuracy for our parameter.

Point vs. interval estimation

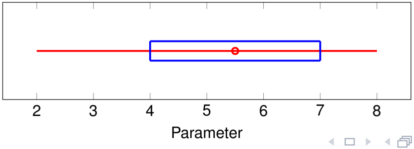
- **Point** estimation: a *single* estimate with our best guess at what the unknown parameter is.
- **Interval** estimation: an interval of values where the unknown parameter is believed to belong in.

Advantage of interval estimation:

1. A point estimate reveals a single point, and we have no idea of how close the actual parameter may or may not be:



2. An interval estimate reveals a "margin of error" as a measure of accuracy for our parameter.



Confidence intervals

Assume we want to estimate the value of an unknown parameter θ.

- A **confidence interval**, presented in the form of $[L, U]$, contains the most "believable" values for the estimated parameter.
- Every confidence interval is associated with a **confidence level**, which represents the probability that the true parameter value falls in that interval:

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

- Combining, we write that $[L, U]$ is a $100 \cdot (1 - \alpha)\,\%$ confidence interval for parameter θ.
  - Most commonly used confidence intervals are 95% confidence intervals, i.e., $\alpha = 5\%$.

The confidence interval, much like point estimates, is finally calculated based on the random sample selected.

Confidence intervals

Assume we want to estimate the value of an unknown parameter θ.

- A **confidence interval**, presented in the form of $[L, U]$, contains the most "believable" values for the estimated parameter.
- Every confidence interval is associated with a **confidence level**, which represents the probability that the true parameter value falls in that interval:

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

- Combining, we write that $[L, U]$ is a $100 \cdot (1 - \alpha)$ % confidence interval for parameter θ.
  - Most commonly used confidence intervals are 95% confidence intervals, i.e., $\alpha = 5\%$.

The confidence interval, much like point estimates, is finally calculated based on the random sample selected.

Confidence intervals

Assume we want to estimate the value of an unknown parameter θ.

- A **confidence interval**, presented in the form of $[L, U]$, contains the most "believable" values for the estimated parameter.
- Every confidence interval is associated with a **confidence level**, which represents the probability that the true parameter value falls in that interval:

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

- Combining, we write that $[L, U]$ is a $100 \cdot (1 - \alpha)\,\%$ confidence interval for parameter θ.
  - Most commonly used confidence intervals are 95% confidence intervals, i.e., $\alpha = 5\%$.

The confidence interval, much like point estimates, is finally calculated based on the random sample selected.

ILLINOIS

Confidence intervals

Assume we want to estimate the value of an unknown parameter θ.

- A **confidence interval**, presented in the form of $[L, U]$, contains the most "believable" values for the estimated parameter.
- Every confidence interval is associated with a **confidence level**, which represents the probability that the true parameter value falls in that interval:

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

- Combining, we write that $[L, U]$ is a $100 \cdot (1 - \alpha)\,\%$ confidence interval for parameter θ.
  - Most commonly used confidence intervals are 95% confidence intervals, i.e., $\alpha = 5\%$.

The confidence interval, much like point estimates, is finally calculated based on the random sample selected.

Confidence intervals

Assume we want to estimate the value of an unknown parameter θ.

- A **confidence interval**, presented in the form of $[L, U]$, contains the most "believable" values for the estimated parameter.
- Every confidence interval is associated with a **confidence level**, which represents the probability that the true parameter value falls in that interval:

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

- Combining, we write that $[L, U]$ is a $100 \cdot (1 - \alpha)\,\%$ confidence interval for parameter θ.
  - Most commonly used confidence intervals are 95% confidence intervals, i.e., $\alpha = 5\%$.

The confidence interval, much like point estimates, is finally calculated based on the random sample selected.

ILLINOIS

Representing confidence intervals

We can also represent a confidence interval as

$$\boxed{\text{Point estimate} \pm \text{Margin}}$$

- For example, here are 10 samples and their confidence intervals (with confidence level at 95%).

- The same experiment with a confidence level of 50%.

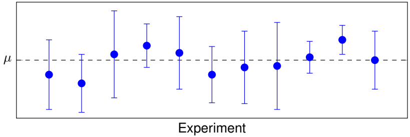
Observe how almost all intervals contain the true mean in the first example; and about half of them do in the second example.

Representing confidence intervals

We can also represent a confidence interval as

$$\boxed{\text{Point estimate} \pm \text{Margin}}$$

- For example, here are 10 samples and their confidence intervals (with confidence level at 95%).



- The same experiment with a confidence level of 50%.

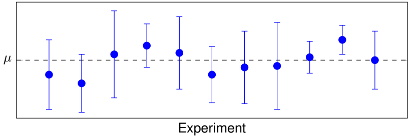
Observe how almost all intervals contain the true mean in the first example; and about half of them do in the second example.

Representing confidence intervals

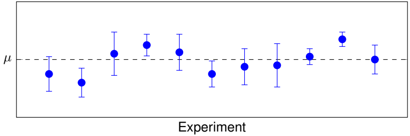
We can also represent a confidence interval as

$$\boxed{\text{Point estimate} \pm \text{Margin}}$$

- For example, here are 10 samples and their confidence intervals (with confidence level at 95%).



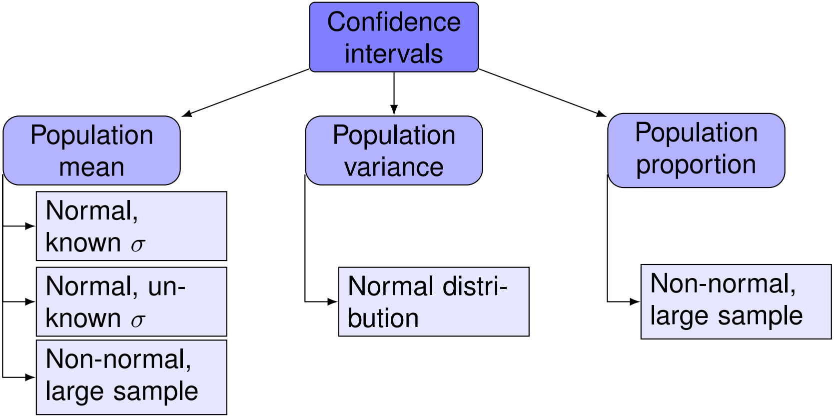
- The same experiment with a confidence level of 50%.



Observe how almost all intervals contain the true mean in the first example; and about half of them do in the second example.

Overview of single population confidence intervals

In the next few lectures, we will discuss:



Population mean: normal, known σ

The simplest of cases assumes a population that is normally distributed with unknown μ but known σ.

- Suppose we have observed a sample $X_1, X_2, \ldots, X_n$.
- We know that $\overline{X} \sim \mathcal{N}(\mu, \sigma^2/n)$.
- We also know, then, that $Z = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1)$.
- We are then looking for the values $-z_{\alpha/2}$ and $z_{\alpha/2}$ such that

$$P(-z_{\alpha/2} \leq Z \leq z_{\alpha/2}) = 1 - \alpha,$$

- where $z_{\alpha/2}$ is called the **critical** z value and we have that $P(Z > z_{\alpha/2}) = \alpha/2$.

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

- where $L = \overline{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ and
- $U = \overline{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$

Population mean: normal, known σ

The simplest of cases assumes a population that is normally distributed with unknown μ but known σ.

- Suppose we have observed a sample $X_1, X_2, \ldots, X_n$.
- We know that $\overline{X} \sim \mathcal{N}(\mu, \sigma^2/n)$.
- We also know, then, that $Z = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1)$.
- We are then looking for the values $-z_{\alpha/2}$ and $z_{\alpha/2}$ such that

$$P(-z_{\alpha/2} \leq Z \leq z_{\alpha/2}) = 1 - \alpha,$$

- where $z_{\alpha/2}$ is called the **critical** z value and we have that $P(Z > z_{\alpha/2}) = \alpha/2$.

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

- where $L = \overline{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ and
- $U = \overline{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$

Population mean: normal, known σ

The simplest of cases assumes a population that is normally distributed with unknown μ but known σ.

- Suppose we have observed a sample $X_1, X_2, \ldots, X_n$.
- We know that $\overline{X} \sim \mathcal{N}(\mu, \sigma^2/n)$.
- We also know, then, that $Z = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0, 1)$.
- We are then looking for the values $-z_{\alpha/2}$ and $z_{\alpha/2}$ such that

$$P(-z_{\alpha/2} \leq Z \leq z_{\alpha/2}) = 1 - \alpha,$$

- where $z_{\alpha/2}$ is called the **critical** z value and we have that $P(Z > z_{\alpha/2}) = \alpha/2$.

$$\boxed{P(\overline{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \overline{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}) = 1 - \alpha}$$

- where $L = \overline{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ and
- $U = \overline{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$

Population mean: normal, known σ

The simplest of cases assumes a population that is normally distributed with unknown μ but known σ.

- Suppose we have observed a sample $X_1, X_2, \ldots, X_n$.
- We know that $\overline{X} \sim \mathcal{N}(\mu, \sigma^2/n)$.
- We also know, then, that $Z = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0, 1)$.
- We are then looking for the values $-z_{\alpha/2}$ and $z_{\alpha/2}$ such that

$$P(-z_{\alpha/2} \leq Z \leq z_{\alpha/2}) = 1 - \alpha,$$

- where $z_{\alpha/2}$ is called the **critical** z value and we have that $P(Z > z_{\alpha/2}) = \alpha/2$.

$$\boxed{P(\overline{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \overline{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}) = 1 - \alpha}$$

- where $L = \overline{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ and
- $U = \overline{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$

Population mean: normal, known σ

The simplest of cases assumes a population that is normally distributed with unknown μ but known σ.

- Suppose we have observed a sample $X_1, X_2, \ldots, X_n$.
- We know that $\overline{X} \sim \mathcal{N}(\mu, \sigma^2/n)$.
- We also know, then, that $Z = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0, 1)$.
- We are then looking for the values $-z_{\alpha/2}$ and $z_{\alpha/2}$ such that

$$P(-z_{\alpha/2} \leq Z \leq z_{\alpha/2}) = 1 - \alpha,$$

- where $z_{\alpha/2}$ is called the **critical** z value and we have that $P(Z > z_{\alpha/2}) = \alpha/2$.

$$\boxed{P(\overline{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \overline{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}) = 1 - \alpha}$$

- where $L = \overline{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ and
- $U = \overline{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$

ILLINOIS

Critical z values

Recall that $z_{\alpha/2}$ is a **critical** z value such that we have:

$$P(Z > z_{\alpha/2}) = \alpha/2$$

Some common critical values:

- $\alpha = 10\% \implies z_{0.05} = 1.645$ as $\Phi(1.645) = 95\% = 1 - \alpha/2$.
- $\alpha = 5\% \implies z_{0.025} = 1.96$ as $\Phi(1.96) = 97.5\% = 1 - \alpha/2$.
- $\alpha = 1\% \implies z_{0.005} = 2.576$ as $\Phi(2.576) = 99.5\% = 1 - \alpha/2$.

You can verify these using a z-table!

Population mean: normal, unknown σ

Assuming the knowledge of the actual population variance is a little unrealistic.

- When assuming that σ is known, then $\overline{X} \sim \mathcal{N}(\mu, \sigma^2/n)$ and $Z = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0,1)$.
- When we do not know σ, we can *estimate* it by s, based on the sample selected.
- $T = \frac{\overline{X}-\mu}{s/\sqrt{n}}$ is **not normally distributed**. Instead:
  - It is distributed following a Student's T distribution,
  - with $n-1$ degrees of freedom.

Population mean: normal, unknown σ

Assuming the knowledge of the actual population variance is a little unrealistic.

- When assuming that σ is known, then $\overline{X} \sim \mathcal{N}(\mu, \sigma^2/n)$ and $Z = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0, 1)$.
- When we do not know σ, we can *estimate* it by s, based on the sample selected.
- $T = \frac{\overline{X}-\mu}{s/\sqrt{n}}$ is **not normally distributed**. Instead:
  - it is distributed following a **Student's T distribution**;
  - with $n-1$ **degrees of freedom**.

Population mean: normal, unknown σ

Assuming the knowledge of the actual population variance is a little unrealistic.

- When assuming that σ is known, then $\overline{X} \sim \mathcal{N}(\mu, \sigma^2/n)$ and $Z = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0, 1)$.
- When we do not know σ, we can *estimate* it by s, based on the sample selected.
- $T = \frac{\overline{X}-\mu}{s/\sqrt{n}}$ is **not normally distributed**. Instead:
  - it is distributed following a **Student's T distribution**;
  - with $n-1$ **degrees of freedom**.

Population mean: normal, unknown σ

Assuming the knowledge of the actual population variance is a little unrealistic.

- When assuming that σ is known, then $\overline{X} \sim \mathcal{N}(\mu, \sigma^2/n)$ and $Z = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0, 1)$.
- When we do not know σ, we can *estimate* it by s, based on the sample selected.
- $T = \frac{\overline{X}-\mu}{s/\sqrt{n}}$ is **not normally distributed**. Instead:
  - it is distributed following a **Student's T distribution**;
  - with $n-1$ **degrees of freedom**.

Critical t values

- For the T distribution we again have a table of values.
- It is read a little differently: the rows now specify the degrees of freedom.
- We again need to track down critical t values. For example:
  - $t_{0.025,15} =$
  - $t_{0.05,10} =$
  - $t_{0.05,25} =$
  - $t_{0.90,5} =$

STUDENT'S t PERCENTAGE POINTS

ν	0.4	0.33	0.25	0.2	0.125	0.1	0.05	0.025	0.01	0.005	0.001
1	0.325	0.577	1.000	1.376	2.414	3.078	6.314	12.706	31.821	63.657	318.31
5	0.267	0.457	0.727	0.920	1.301	1.476	2.015	2.571	3.365	4.032	5.893
10	0.260	0.444	0.700	0.879	1.221	1.372	1.812	2.228	2.764	3.169	4.144
15	0.258	0.439	0.691	0.866	1.197	1.341	1.753	2.131	2.602	2.947	3.733
25	0.256	0.436	0.684	0.856	1.178	1.316	1.708	2.060	2.485	2.787	3.450
∞	0.253	0.431	0.674	0.842	1.150	1.282	1.645	1.960	2.326	2.576	3.090

Critical t values

- For the T distribution we again have a table of values.
- It is read a little differently: the rows now specify the degrees of freedom.
- We again need to track down critical t values. For example:
  - $t_{0.025,15} = 2.131$
  - $t_{0.05,10} = 1.812$
  - $t_{0.05,25} = 1.708$
  - $t_{0.90,5} = -1.476$

STUDENT'S t PERCENTAGE POINTS

ν	0.4	0.33	0.25	0.2	0.125	0.1	0.05	0.025	0.01	0.005	0.001
1	0.325	0.577	1.000	1.376	2.414	3.078	6.314	12.706	31.821	63.657	318.31
5	0.267	0.457	0.727	0.920	1.301	1.476	2.015	2.571	3.365	4.032	5.893
10	0.260	0.444	0.700	0.879	1.221	1.372	1.812	2.228	2.764	3.169	4.144
15	0.258	0.439	0.691	0.866	1.197	1.341	1.753	2.131	2.602	2.947	3.733
25	0.256	0.436	0.684	0.856	1.178	1.316	1.708	2.060	2.485	2.787	3.450
∞	0.253	0.431	0.674	0.842	1.150	1.282	1.645	1.960	2.326	2.576	3.090

Critical t values

- For the T distribution we again have a table of values.
- It is read a little differently: the rows now specify the degrees of freedom.
- We again need to track down critical t values. For example:
  - $t_{0.025,15} = 2.131$
  - $t_{0.05,10} = 1.812$
  - $t_{0.05,25} = 1.708$
  - $t_{0.90,5} = -1.476$

STUDENT'S t PERCENTAGE POINTS

ν	0.4	0.33	0.25	0.2	0.125	0.1	0.05	0.025	0.01	0.005	0.001
1	0.325	0.577	1.000	1.376	2.414	3.078	6.314	12.706	31.821	63.657	318.31
5	0.267	0.457	0.727	0.920	1.301	1.476	2.015	2.571	3.365	4.032	5.893
10	0.260	0.444	0.700	0.879	1.221	1.372	1.812	2.228	2.764	3.169	4.144
15	0.258	0.439	0.691	0.866	1.197	1.341	1.753	2.131	2.602	2.947	3.733
25	0.256	0.436	0.684	0.856	1.178	1.316	1.708	2.060	2.485	2.787	3.450
∞	0.253	0.431	0.674	0.842	1.150	1.282	1.645	1.960	2.326	2.576	3.090

Critical t values

- For the T distribution we again have a table of values.
- It is read a little differently: the rows now specify the degrees of freedom.
- We again need to track down critical t values. For example:
  - $t_{0.025,15} = 2.131$
  - $t_{0.05,10} = 1.812$
  - $t_{0.05,25} = 1.708$
  - $t_{0.90,5} = -1.476$

STUDENT'S t PERCENTAGE POINTS

ν	0.4	0.33	0.25	0.2	0.125	0.1	0.05	0.025	0.01	0.005	0.001
1	0.325	0.577	1.000	1.376	2.414	3.078	6.314	12.706	31.821	63.657	318.31
5	0.267	0.457	0.727	0.920	1.301	1.476	2.015	2.571	3.365	4.032	5.893
10	0.260	0.444	0.700	0.879	1.221	1.372	1.812	2.228	2.764	3.169	4.144
15	0.258	0.439	0.691	0.866	1.197	1.341	1.753	2.131	2.602	2.947	3.733
25	0.256	0.436	0.684	0.856	1.178	1.316	1.708	2.060	2.485	2.787	3.450
∞	0.253	0.431	0.674	0.842	1.150	1.282	1.645	1.960	2.326	2.576	3.090

Population mean: normal, unknown σ

Since $T = \frac{\overline{X}-\mu}{s/\sqrt{n}}$ is distributed with Student's T-distribution with $n-1$ degrees of freedom, we are then looking for the values $-t_{\alpha/2,n-1}$ and $t_{\alpha/2,n-1}$ such that

$$P(-t_{\alpha/2,n-1} \leq T \leq t_{\alpha/2,n-1}) = 1-\alpha,$$

which leads to:

$$\boxed{P(\overline{X} - t_{\alpha/2,n-1}\frac{s}{\sqrt{n}} \leq \mu \leq \overline{X} + t_{\alpha/2,n-1}\frac{s}{\sqrt{n}}) = 1-\alpha}$$

- where $L = \overline{X} - t_{\alpha/2,n-1}\frac{s}{\sqrt{n}}$ and
- $U = \overline{X} + t_{\alpha/2,n-1}\frac{s}{\sqrt{n}}$

ILLINOIS

Population mean: normal, unknown σ

Since $T = \frac{\overline{X}-\mu}{s/\sqrt{n}}$ is distributed with Student's T-distribution with $n-1$ degrees of freedom, we are then looking for the values $-t_{\alpha/2,n-1}$ and $t_{\alpha/2,n-1}$ such that

$$P(-t_{\alpha/2,n-1} \leq T \leq t_{\alpha/2,n-1}) = 1-\alpha,$$

which leads to:

$$\boxed{P(\overline{X} - t_{\alpha/2,n-1}\frac{s}{\sqrt{n}} \leq \mu \leq \overline{X} + t_{\alpha/2,n-1}\frac{s}{\sqrt{n}}) = 1-\alpha}$$

- where $L = \overline{X} - t_{\alpha/2,n-1}\frac{s}{\sqrt{n}}$ and
- $U = \overline{X} + t_{\alpha/2,n-1}\frac{s}{\sqrt{n}}$

ILLINOIS

Population mean: non-normal, large sample

When we do not deal with normally distributed populations, we **need to** have a large sample (let us assume $n \geq 30$).

- Due to the *central limit theorem* $Z = \frac{\overline{X}-\mu}{s/\sqrt{n}} \sim \mathcal{N}(0,1)$.
- The only difference between the first and the third case then, is that in the third case we are allowed to use the sample standard deviation if the true standard deviation σ is not known.

Overall:

$$\boxed{P(\overline{X} - z_{\alpha/2}\frac{s}{\sqrt{n}} \leq \mu \leq \overline{X} + z_{\alpha/2}\frac{s}{\sqrt{n}}) = 1 - \alpha}$$

- where $L = \overline{X} - z_{\alpha/2}\frac{s}{\sqrt{n}}$ and
- $U = \overline{X} + z_{\alpha/2}\frac{s}{\sqrt{n}}$

Population mean: non-normal, large sample

When we do not deal with normally distributed populations, we **need to** have a large sample (let us assume $n \geq 30$).

- Due to the *central limit theorem* $Z = \frac{\overline{X}-\mu}{s/\sqrt{n}} \sim \mathcal{N}(0,1)$.
- The only difference between the first and the third case then, is that in the third case we are allowed to use the sample standard deviation if the true standard deviation σ is not known.

Overall:

$$\boxed{P(\overline{X} - z_{\alpha/2}\frac{s}{\sqrt{n}} \leq \mu \leq \overline{X} + z_{\alpha/2}\frac{s}{\sqrt{n}}) = 1 - \alpha}$$

- where $L = \overline{X} - z_{\alpha/2}\frac{s}{\sqrt{n}}$ and
- $U = \overline{X} + z_{\alpha/2}\frac{s}{\sqrt{n}}$

Population mean: non-normal, large sample

When we do not deal with normally distributed populations, we **need to** have a large sample (let us assume $n \geq 30$).

- Due to the *central limit theorem* $Z = \frac{\overline{X}-\mu}{s/\sqrt{n}} \sim \mathcal{N}(0,1)$.
- The only difference between the first and the third case then, is that in the third case we are allowed to use the sample standard deviation if the true standard deviation σ is not known.

Overall:

$$\boxed{P(\overline{X} - z_{\alpha/2}\frac{s}{\sqrt{n}} \leq \mu \leq \overline{X} + z_{\alpha/2}\frac{s}{\sqrt{n}}) = 1 - \alpha}$$

- where $L = \overline{X} - z_{\alpha/2}\frac{s}{\sqrt{n}}$ and
- $U = \overline{X} + z_{\alpha/2}\frac{s}{\sqrt{n}}$

ILLINOIS

Population mean: non-normal, large sample

When we do not deal with normally distributed populations, we **need to** have a large sample (let us assume $n \geq 30$).

- Due to the *central limit theorem* $Z = \frac{\overline{X}-\mu}{s/\sqrt{n}} \sim \mathcal{N}(0, 1)$.
- The only difference between the first and the third case then, is that in the third case we are allowed to use the sample standard deviation if the true standard deviation σ is not known.

Overall:

$$\boxed{P(\overline{X} - z_{\alpha/2}\frac{s}{\sqrt{n}} \leq \mu \leq \overline{X} + z_{\alpha/2}\frac{s}{\sqrt{n}}) = 1 - \alpha}$$

- where $L = \overline{X} - z_{\alpha/2}\frac{s}{\sqrt{n}}$ and
- $U = \overline{X} + z_{\alpha/2}\frac{s}{\sqrt{n}}$

Quick check

1. We are told that the mean battery lifetime of a cell phone on a single charge is between 13 and 16 hours with a confidence level of 95%. That means:
   a. There is a 5% chance that the mean cell phone battery is less than 13 hours or more than 16 hours.
   b. If we buy 100 phones, we should expect 95 to have battery lifetimes between 13 and 16 hours.
   c. We are 95% certain that the true mean battery lifetime is between 13 and 16 hours.
   d. All of the above.

2. The same company claims that the battery lifetime is between 14 and 15 hours with a confidence level of 99%. This statement is:
   a. True.
   b. False.
   c. We can't tell – we need more information.

Quick check

1. We are told that the mean battery lifetime of a cell phone on a single charge is between 13 and 16 hours with a confidence level of 95%. That means:
   - a. There is a 5% chance that the mean cell phone battery is less than 13 hours or more than 16 hours.
   - b. If we buy 100 phones, we should expect 95 to have battery lifetimes between 13 and 16 hours.
   - c. We are 95% certain that the true mean battery lifetime is between 13 and 16 hours.
   - d. All of the above.

2. The same company claims that the battery lifetime is between 14 and 15 hours with a confidence level of 99%. This statement is:
   - a. True.
   - b. False.
   - c. We can't tell – we need more information.

ILLINOIS

Quick check

1. We are told that the mean battery lifetime of a cell phone on a single charge is between 13 and 16 hours with a confidence level of 95%. That means:
   a. There is a 5% chance that the mean cell phone battery is less than 13 hours or more than 16 hours.
   b. If we buy 100 phones, we should expect 95 to have battery lifetimes between 13 and 16 hours.
   c. We are 95% certain that the true mean battery lifetime is between 13 and 16 hours.
   d. All of the above.

2. The same company claims that the battery lifetime is between 14 and 15 hours with a confidence level of 99%. This statement is:
   a. True.
   b. False.
   c. We can't tell – we need more information.

Quick check

1. We are told that the mean battery lifetime of a cell phone on a single charge is between 13 and 16 hours with a confidence level of 95%. That means:
   a. There is a 5% chance that the mean cell phone battery is less than 13 hours or more than 16 hours.
   b. If we buy 100 phones, we should expect 95 to have battery lifetimes between 13 and 16 hours.
   c. We are 95% certain that the true mean battery lifetime is between 13 and 16 hours.
   d. All of the above.

2. The same company claims that the battery lifetime is between 14 and 15 hours with a confidence level of 99%. This statement is:
   a. True.
   b. False.
   c. We can't tell – we need more information.

Quick check

1. We are told that the mean battery lifetime of a cell phone on a single charge is between 13 and 16 hours with a confidence level of 95%. That means:
   - **a.** There is a 5% chance that the mean cell phone battery is less than 13 hours or more than 16 hours.
   - **b.** If we buy 100 phones, we should expect 95 to have battery lifetimes between 13 and 16 hours.
   - **c.** We are 95% certain that the true mean battery lifetime is between 13 and 16 hours.
   - **d.** All of the above.

2. The same company claims that the battery lifetime is between 14 and 15 hours with a confidence level of 99%. This statement is:
   - **a.** True.
   - **b.** False.
   - **c.** We can't tell – we need more information.

ILLINOIS