**Confidence Intervals:**

Based on a sample, calculate the interval of values in which the population parameter (mean, proportion, standard deviation) may lie with a degree of certainty. This degree of certainty (level of confidence) is given as a percentage or the equivalent decimal.
The most used confidence intervals are $80\%,\ 90\%,\ 95\%,\ 98\%\ \textit{and}\ 99\%$.

The population parameter lies in between the sample point estimate plus or minus an error:

$$Sample\ point\ estimate\ \pm Error$$

The three confidence intervals in STA2023:

1. Confidence Interval for the population mean, population standard deviation σ, known.

$\bar{x} \pm Error$

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

2. Confidence Interval for the population mean, population standard deviation σ, unknown. Use the sample standard deviation (for a normal distributed variable), use the $t\ distribution$.

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

3. Confidence Interval for the population proportion:

$\hat{p} \pm z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}\cdot\hat{q}}{n}}$ where $\hat{q} = 1 - \hat{p}$

| Sample size for means: | Sample size por proportions: |
|---|---|
| $n = \left(\frac{z_{\alpha/2} \cdot \sigma}{E}\right)^2$ | $n = \hat{p} \cdot \hat{q} \left(\frac{z_{\alpha/2}}{E}\right)^2$<br>If $\hat{p}$ unknown, use 0.5 |

**Critical values, $Z_{\alpha/2}$:**

90%, $Z_{\alpha/2} = 1.645$ 95%, $Z_{\alpha/2} = 1.960$ 98%, $Z_{\alpha/2} = 2.326$ 99%, $Z_{\alpha/2} = 2.576$

T-Critical values:
See complete table at the end of this document.
Degrees of freedom, $df = n - 1$ where $n$ is the sample size.

| d.f. | Confidence intervals | 80% | 90% | 95% | 98% | 99% |
|---|---|---|---|---|---|---|
| | One tail, $\alpha$ | 0.10 | 0.05 | 0.025 | 0.01 | 0.005 |
| | Two tails, $\alpha$ | 0.20 | 0.10 | 0.05 | 0.02 | 0.01 |
| 1 | | 3.078 | 6.314 | 12.706 | 31.821 | 63.657 |
| 2 | | 1.886 | 2.920 | 4.303 | 6.965 | 9.925 |
| 3 | | 1.638 | 2.353 | 3.182 | 4.541 | 5.841 |
| 4 | | 1.533 | 2.132 | 2.776 | 3.747 | 4.604 |
| 5 | | 1.476 | 2.015 | 2.571 | 3.365 | 4.032 |
| 6 | | 1.440 | 1.943 | 2.447 | 3.143 | 3.707 |
| 7 | | 1.415 | 1.895 | 2.365 | 2.998 | 3.499 |
| 8 | | 1.397 | 1.860 | 2.306 | 2.896 | 3.355 |
| 9 | | 1.383 | 1.833 | 2.262 | 2.821 | 3.250 |
| 10 | | 1.372 | 1.812 | 2.228 | 2.764 | 3.169 |
| 11 | | 1.363 | 1.796 | 2.201 | 2.718 | 3.106 |
| 12 | | 1.356 | 1.782 | 2.179 | 2.681 | 3.055 |
| 13 | | 1.350 | 1.771 | 2.160 | 2.650 | 3.012 |
| 14 | | 1.345 | 1.761 | 2.145 | 2.624 | 2.977 |
| 15 | | 1.341 | 1.753 | 2.131 | 2.602 | 2.947 |
| 16 | | 1.337 | 1.746 | 2.120 | 2.583 | 2.921 |
| 17 | | 1.333 | 1.740 | 2.110 | 2.567 | 2.898 |
| 18 | | 1.330 | 1.734 | 2.101 | 2.552 | 2.878 |
| 19 | | 1.328 | 1.729 | 2.093 | 2.539 | 2.861 |
| 20 | | 1.325 | 1.725 | 2.086 | 2.528 | 2.845 |

1. **Examples 1: Z-Test for mean (sigma known)**

The lengths, in inches, of adult corn snakes are normally distributed with a population standard deviation of 8 inches and an unknown population mean. A random sample of 25 snakes is taken and results in a sample mean of 58 inches. Find a 99% confidence interval estimate for the population mean.

$\bar{x} \pm z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$ $58 \pm 2.576 \cdot \frac{8}{\sqrt{25}}$ $58 \pm 4.1216$

by subtracting and adding $E = 4.1216$, yields the 99% Conf Interval: $(53.88, 62.12)$

Interpretation:

We can estimate with 99% confidence that the true population mean length of adult corn snakes is between 53.88 and 62.12 inches.

2. **Example 2: T-Test for mean, sample standard deviation known:**

   The commute times for workers in a city are normally distributed with an unknown population mean and standard deviation.

   If a random sample of 20 workers is taken and results in a sample mean of 21 minutes and sample standard deviation of 6 minutes, find a 95% confidence interval estimate for the population mean using the Student's t-distribution.

| | Confidence intervals | 80% | 90% | 95% |
|---|---|---|---|---|
| | One tail, $\alpha$ | 0.10 | 0.05 | 0.025 |
| d.f. | Two tails, $\alpha$ | 0.20 | 0.10 | 0.05 |
| 1 | | 3.078 | 6.314 | 12.706 |
| 2 | | 1.886 | 2.920 | 4.303 |
| 3 | | 1.638 | 2.353 | 3.182 |
| 4 | | 1.533 | 2.132 | 2.776 |
| 5 | | 1.476 | 2.015 | 2.571 |
| 6 | | 1.440 | 1.943 | 2.447 |
| 7 | | 1.415 | 1.895 | 2.365 |
| 8 | | 1.397 | 1.860 | 2.306 |
| 9 | | 1.383 | 1.833 | 2.262 |
| 10 | | 1.372 | 1.812 | 2.228 |
| 11 | | 1.363 | 1.796 | 2.201 |
| 12 | | 1.356 | 1.782 | 2.179 |
| 13 | | 1.350 | 1.771 | 2.160 |
| 14 | | 1.345 | 1.761 | 2.145 |
| 15 | | 1.341 | 1.753 | 2.131 |
| 16 | | 1.337 | 1.746 | 2.120 |
| 17 | | 1.333 | 1.740 | 2.110 |
| 18 | | 1.330 | 1.734 | 2.101 |
| 19 | | 1.328 | 1.729 | 2.093 |
| 20 | | 1.325 | 1.725 | 2.086 |

   Find the margin of error, for a 95% confidence interval estimate for the population mean using the student's t-distribution. By the table, for $n = 20$, $df = 19$: $t_{\alpha/2} = 2.093$

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

   $21 \pm 2.81$

   by subtracting and adding E, yields the 95% Confidence interval: $(18.19, 23.81)$ *or* $18.19 < \mu < 23.81$

3. **Example 3: 1-Proportion Z-Test.**

   Emma wants to estimate the percentage of people who use public transportation in a city. She surveys 140 individuals and finds that 62 use public transportation. Find the confidence interval for the population proportion with a 99% confidence level.

   Answer: This a proportion interval. Find $\hat{p}$ first:

   $\hat{p} = \frac{x}{n} = \frac{62}{140} = 0.4429$ and $\hat{q} = 1 - \hat{p}$ Therefore, $\hat{q} = 1 - 0.4429 = 0.5571$

   $\hat{p} \pm z_{\alpha/2} \cdot \sqrt{\frac{\hat{p} \cdot \hat{q}}{n}}$ substituting values, becomes:

$0.4429 \pm 2.576 \cdot \sqrt{\frac{0.4429 \text{ x } 0.5571}{140}}$

$0.4429 \ \pm 0.1081$

$0.335 \ < \text{p} < \ \ 0.551$

4. **Example 4: Sample size for proportion:**

   Suppose a clothing store wants to determine the current percentage of customers who are over the age of forty.

   How many customers should the company survey in order to be 90% confident that the estimated (sample) proportion is within 4 percentage points of the true population proportion of customers who are over the age of forty?

   Answer: Sample size por proportions:

$n = \hat{p} \cdot \hat{q} \left(\frac{z_{\alpha/2}}{E}\right)^2$ since $\hat{p}$ is unknown (some times researchers used a previous study result, that is the case in which $\hat{p}$ is known; otherwise, like in this case, use $0.5$; the $z_{\alpha/2}$ = 1.645 for a 90% CI, 4% as decimal, $E \ = \ 0.04$; substituting values:

$$n = 0.5 \cdot 0.5 \left(\frac{1.645}{0.04}\right)^2 \ = \ 422.81 \ \ = \ 423$$