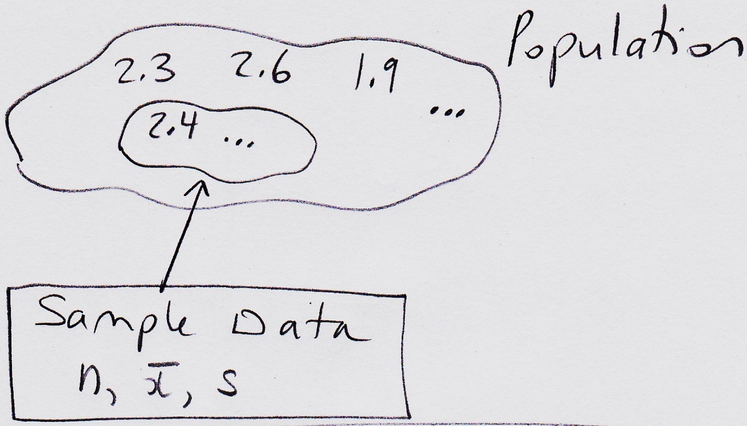
9. Confidence Intervals



Goal: Estimate population mean μ based on sample data

e.g. We're 95% confident that $1.8 \le \mu \le 2.7$

Common confidence levels: 90%, 95%, 98%, 99%

Confidence level is written $1-\alpha$

Interpretation of a 95% confidence interval for μ:

95% of possible samples of size n lead to a confidence interval that contains μ

$$\boxed{\begin{array}{l}\underline{\text{Large}}\text{ Sample Confidence Interval for } \mu \\ \qquad \bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}\end{array}}$$

Assumption: Random sample with $n \geq 30$

Note: σ can be replaced with s since $\sigma \approx s$ when $n \geq 30$

Values of $z_{\alpha/2}$ are on formula sheet

<u>Ex</u>: Volumes in cans of Coke have a SD of 2.5 mL. A random sample of 60 cans had an average volume of 355.3 mL. Find a 95% CI for the average volume among all cans of Coke.

$\sigma = 2.5\text{mL}$ $\quad n = 60$ $\quad \bar{x} = 355.3$

$(n \geq 30 \checkmark)$ $\quad$ Confidence level $1-\alpha = 0.95$

$$z_{\alpha/2} = 1.960$$

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

$$355.3 \pm 1.960 \left(\frac{2.5}{\sqrt{60}}\right)$$

$$354.7 \leq \mu \leq 355.9$$

We're 95% confident that $354.7\text{mL} \leq \mu \leq 355.9\text{mL}$

Ex: Consider a 99% CI with the same σ, n and $\bar{x}$. Is it a wider or narrower interval?

WIDER

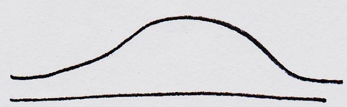
Algebra: $z_{\alpha/2}$ would be bigger

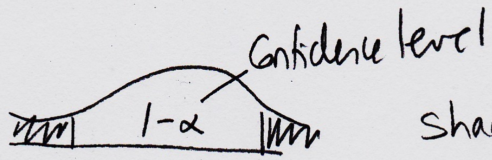
Conceptual: A wider interval gives more confidence that we've captured μ

> There is a trade-off between confidence and precision
> A 100% CI would be infinite

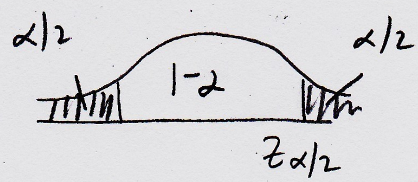
Why we use $z_{\alpha/2}$:

$\bar{x}$-values are normal by Central Limit Theorem





Shaded area = probability of error



Ex: 50 random Camosun students were polled on their # of googles per week. The average was 23.4 with a SD of 3.7. Find a 90% CI for the average # of googles among all Camosun students

$n = 50 \quad \bar{x} = 23.4 \quad s = 3.7$

$(n \geq 30 \checkmark)$

$1 - \alpha = 0.9$

$z_{\alpha/2} = 1.645$

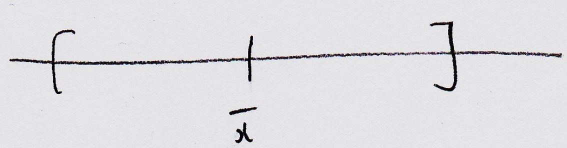
$$\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

$$\approx \bar{x} \pm z_{\alpha/2} \frac{s}{\sqrt{n}}$$

$$\approx 23.4 \pm 1.645 \left(\frac{3.7}{\sqrt{50}} \right)$$

$$22.5 \leq \mu \leq 24.3$$

CI:



margin of error or ME

$$\boxed{ME = z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \text{ or } z_{\alpha/2} \frac{s}{\sqrt{n}}}$$

Ex: A sample of lightbulb lifetimes had a SD of 8.9 months. Want to estimate μ with a 90% ME < 0.1 months. What is the minimum sample size?

Idea: Want to be confident and precise
Need lots of data (large n)

$s = 8.9$ $\quad 1-\alpha = 0.9$

$z_{\alpha/2} = 1.645$

$$ME < 0.1$$

$$z_{\alpha/2} \frac{s}{\sqrt{n}} < 0.1$$

$$1.645\left(\frac{8.9}{\sqrt{n}}\right) < 0.1$$

$$\frac{1.645(8.9)}{0.1} < \sqrt{n}$$

$$\left[\frac{1.645(8.9)}{0.1}\right]^2 < n$$

$$n > 21434.4$$

Minimim size is $n = 21435$

Upper Confidence Bound (UCB) : $\mu \le \#$

Lower " (LCB) : $\mu \ge \#$

Large Sample UCB/LCB for μ:

UCB: $\mu \le \bar{x} + z_\alpha \frac{\sigma}{\sqrt{n}}$

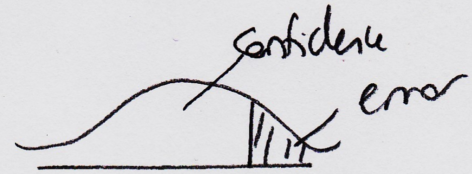
LCB: $\mu \ge \bar{x} - z_\alpha \frac{\sigma}{\sqrt{n}}$

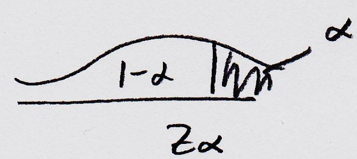
Assumption: Random sample with $n \ge 30$

Note: σ can be replaced by s

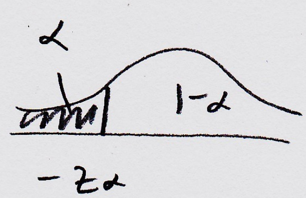
Why we use z_α:

UCB:





LCB:



Ex: 30 random water samples have a mean pollution concentration of 48.1 ppm with a SD of 6.2 ppm. Find a 99% UCB for the mean pollution concentration in the body of water.

$n = 30 \quad \bar{x} = 48.1 \quad s = 6.2$

$(n \geq 30 \checkmark) \quad 1 - \alpha = 0.99$

$z_\alpha = 2.326$

$$\mu \leq \bar{x} + z_\alpha \frac{s}{\sqrt{n}}$$

$$\mu \leq 48.1 + 2.326 \left(\frac{6.2}{\sqrt{30}}\right)$$

$$\mu \leq 50.7 \text{ ppm}$$

Ex: 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.

$\sigma = 10.3 \quad n = 40 \quad \bar{x} = 69.1$

$(n \geq 30 \checkmark) \quad 1 - \alpha = 0.98$

$z_\alpha = 2.054$

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

$$\mu \geq 69.1 - 2.054 \left(\frac{10.3}{\sqrt{40}} \right)$$

$$\mu \geq 65.8$$

SMALL SAMPLES

When $n < 30$, we use the t-table rather than the z-table.

The t-curve is flatter than the z-curve

Degrees of freedom $df = n-1$ describes the flatness of the t-curve.

$df=1$

$df=30$

$\boxed{\begin{array}{lll} \underline{\text{Small}}\ \text{Sample CI for } \mu: & & \bar{x} \pm t_{\alpha/2} \frac{s}{\sqrt{n}} \\ \quad " \quad \text{UCB} \ " & & \mu \le \bar{x} + t_\alpha \frac{s}{\sqrt{n}} \\ \quad " \quad \text{LCB} \ " & & \mu \ge \bar{x} - t_\alpha \frac{s}{\sqrt{n}} \\ df = n-1 & & \end{array}}$

Assumptions: Random sample, $n < 30$, normal population and σ unknown

<u>Ex</u>: The radius of ball bearings is normally distributed. A random sample of 10 ball bearings has a mean radius of 4.9 cm with a SD of 0.9 cm. Find a 95% CI for the mean radius.

Normal pop.

$n = 10 \qquad \bar{x} = 4.9 \qquad s = 0.9$

$\boxed{\begin{array}{l} n \ge 30 \rightarrow \text{use } z \\ n < 30, \text{ normal pop}, \ \sigma \text{ unknown} \rightarrow \text{use } t \end{array}}$

$$\left.\begin{array}{l} df = n-1 = 9 \\ 1-\alpha = 0.95 \\ \alpha = 0.05 \\ \frac{\alpha}{2} = 0.025 \end{array}\right\} \quad t_{\frac{\alpha}{2}} = t_{0.025} = 2.262$$

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

$$4.9 \pm 2.262 \left(\frac{0.9}{\sqrt{10}} \right)$$

$$4.3\text{cm} \le \mu \le 5.5\text{cm}$$

Ex: The mass of chocolate bars are normal. A random sample of 18 bars had a mean mass of 84.7g with a SD of 2.6g. Find a 99% UCB for the mean mass among all bars.

Normal pop.

$n = 18 \qquad \bar{x} = 84.7 \qquad s = 2.6$

$\boxed{\text{normal pop.},\ n < 30,\ \sigma \text{ unknown} \rightarrow \text{use } t}$

$$\left.\begin{array}{l} df = n-1 = 17 \\ 1-\alpha = 0.99 \\ \alpha = 0.01 \end{array}\right\} \quad t_\alpha = t_{0.01} = 2.567$$

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

$$\mu \le 84.7 + 2.567 \left(\frac{2.6}{\sqrt{18}} \right)$$

$$\mu \le 86.3\text{ g}$$

Ex: The breaking strengths of a brand of rope are normal. A random sample of 8 ropes had a mean of 62.1 lbs with a SD of 2.5 lbs. Find a 90% LCB for the mean breaking strength.

Normal pop. $n=8$ $\bar{x}=62.1$ $s=2.5$

$\boxed{\text{Normal pop., } n<30,\ \sigma \text{ unknown} \rightarrow \text{use } t}$

$$\left.\begin{array}{l} df = n-1 = 7 \\ 1-\alpha = 0.9 \\ \alpha = 0.1 \end{array}\right\} t_\alpha = t_{0.1} = 1.415$$

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

$$\mu \geq 62.1 - 1.415\left(\frac{2.5}{\sqrt{8}}\right)$$

$$\mu \geq 60.8 \text{ lbs}$$

Overview

Large Sample	Small Sample
CI	CI
UCB	UCB
LCB	LCB
Minimum Sample Size	
Use z	Use t