CEE 260/MIE 273: Probability and Statistics in Civil Engineering

Lecture M4b: Confidence Intervals for a Proportion

Prof. Oke

UMass Amherst

College of Engineering

October 20, 2025

Outline

Recap: parameter estimation

- A point estimator is used to find the estimate $\hat{\theta}$ for a given parameter θ.
- An ideal estimator is: unbiased, consistent, efficient and sufficient
- The bias of a parameter estimate is the difference between its expected value and the true value of the parameter:

$$\text{Bias}_{\hat{\theta}} = \mathbb{E}(\hat{\theta}) - \theta \tag{1}$$

Today's objectives

- Understand confidence intervals and their importance
- Compute confidence intervals for a proportion
  - find the critical z^* for a corresponding confidence level
  - compute the standard error
  - compute the margin of error
- Learn how to use the success-failure condition to determine whether CLT holds for a proportion

Confidence intervals

Definition

A confidence interval (CI) defines the range within which a population parameter θ lies with a given probability $1-\alpha$

$$\langle\theta\rangle_{1-\alpha} = \left(\hat{\theta} + z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}, \hat{\theta} + z_{\left(1-\frac{\alpha}{2}\right)}\frac{\sigma}{\sqrt{n}}\right) \tag{2}$$

where:

- $\hat{\theta}$ is the point estimate
- $\frac{\sigma}{\sqrt{n}}$ is the **standard error** (SE)
- $z_{\left(1-\frac{\alpha}{2}\right)}\frac{\sigma}{\sqrt{n}}$ is the **margin of error** (ME)
- $z_{\frac{\alpha}{2}}$ and $z_{\left(1-\frac{\alpha}{2}\right)}$ are the **critical z-scores**

$$z_{\frac{\alpha}{2}} = -z_{\left(1-\frac{\alpha}{2}\right)} \equiv -z^* \tag{3}$$

Thus, we can express the CI as:

$$\langle\theta\rangle_{1-\alpha} = \hat{\theta} \pm z^* SE = \hat{\theta} \pm ME \tag{4}$$

Confidence intervals (cont.)

We define:

- Confidence level: $1 - \alpha$
- Significance level: α
- Find critical z-score (standardized) values: $z_{\alpha/2}$ and $z_{(1-\alpha/2)}$
- Convert these to same scale as original variable X

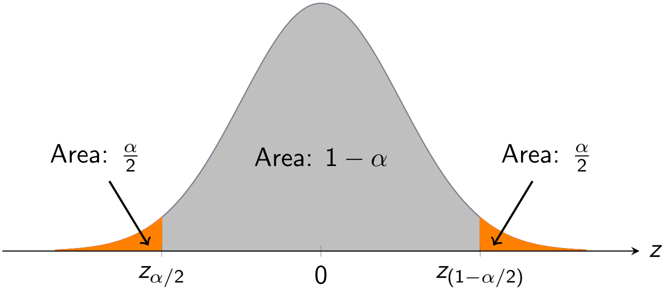


Figure: Standard normal distribution of the mean

Confidence interval of proportion

Given a proportion (occurrence probability) p, then

$$\mathbb{E}(\hat{P}) = \hat{p} \quad (5)$$

$$\mathbb{V}(\hat{P}) = \frac{\hat{p}(1-\hat{p})}{n} \quad \text{(according to CLT)} \quad (6)$$

The confidence interval is given by:

$$\boxed{\langle p \rangle_{1-\alpha} = \left(\hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}; \hat{p} + z_{(1-\alpha/2)}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \right)} \quad (7)$$

Thus:

- $SE = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$
- $ME = z_{(1-\alpha/2)}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = z^* \times SE$

Must check that success-failure condition holds: $np \geq 10$ and $n(1-p) \geq 10$

Working with confidence intervals

Example 1: Identifying confidence levels

Given a normal population distribution with known variance:

(a) What is the confidence level for the interval $\overline{x} \pm 2.81\sigma/\sqrt{n}$?

(b) What is the confidence level for the interval $\overline{x} \pm 1.44\sigma/\sqrt{n}$?

(c) What value of $z_{\alpha/2}$ results in a confidence level of 90%?

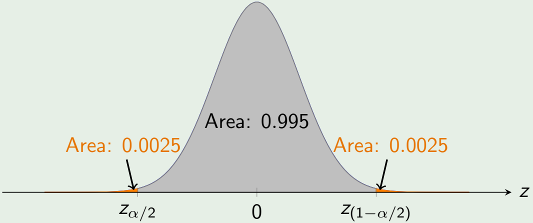
Working with confidence intervals

Example 1: Identifying confidence levels (cont.)

(a) What is the confidence level for the interval $\overline{x} \pm 2.81\sigma/\sqrt{n}$?

$$
\begin{aligned}
z_{(1-\alpha/2)} &= +2.81 \\
1-\alpha/2 &= \Phi(2.81) = 99.75\% \\
\alpha/2 &= 0.25\% \\
\alpha &= 0.5\%
\end{aligned}
$$

The confidence level is $= 1-\alpha = \boxed{99.5\%}$.



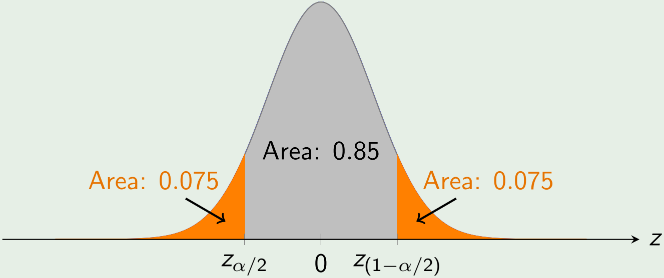
Working with confidence intervals

Example 1: Identifying confidence levels (cont.)

(b) What is the confidence level for the interval $\overline{x} \pm 1.44\sigma/\sqrt{n}$?

$$\begin{aligned} z_{(1-\alpha/2)} &= +1.44 \\ 1-\alpha/2 &= \Phi(1.44) = 92.5\% \\ \alpha &= 15\% \end{aligned}$$

The confidence level is $= 1 - \alpha = \boxed{85\%}$.

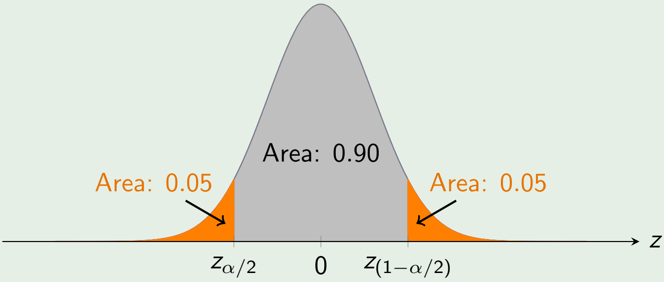


Working with confidence intervals

Example 1: Identifying confidence levels (cont.)

(c) What value of $z_{\alpha/2}$ results in a confidence level of 90%?

$$\begin{aligned} z_{\alpha/2} &= \Phi^{-1}(0.05) \\ &= -\Phi^{-1}(0.95) \\ &= -1.64 \end{aligned}$$



Example 2: Solar power expansion

A poll indicates that 88.7% of a random sample of 1000 American adults supported solar power expansion. Compute and interpret a 95% CI for the population propotion p, given that $SE_{\hat{p}} = 0.01$

Solution

The corresponding critical z-score z^* for a 95% confidence level is `norminv`(1 - .025) $= 1.96$. Thus, the CI is given by

$$
\begin{aligned}
\langle p \rangle_{90\%} &= \hat{p} \pm z^* \times SE_{\hat{p}} \\
&= .887 \pm 1.96(.01) \\
&= (0.8674, .9066)
\end{aligned}
$$

Example 3: Mental health

The General Social Survey asked the question: "For how many days during the past 30 days was your mental health, which includes stress, depression, and problems with emotions, not good?" Based on responses from 1,151 US residents, the survey reported a 95% confidence interval of 3.40 to 4.24 days in 2010.

(a) Interpret this interval in context of the data.

(b) What does "95% confident" mean? Explain in the context of the application.

(c) Suppose the researchers think a 99% confidence level would be more appropriate for this interval. Will this new interval be smaller or wider than the 95% confidence interval?

(d) If a new survey were to be done with 500 Americans, do you think the standard error of the estimate be larger, smaller, or about the same.

Recap

Steps to find the CI of a proportion

- Compute n and $\hat{p}$
- Check that the success-failure condition holds
- Find the critical z-score for the confidence level $1 - \alpha$
- Compute the $SE = \sqrt{p(1-p)/n}$
- Find the ME ($z^* \times SE$)
- Find the CI as $\hat{p} \pm ME$