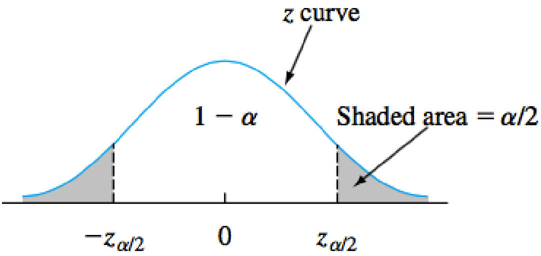
MATH 205: Statistical methods

Lecture 26: Large-sample confidence intervals

$100(1-\alpha)\%$ confidence interval



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

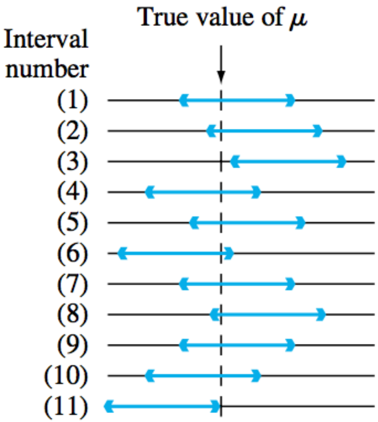
Assumption: Normal distribution with known σ

A $100(1-\alpha)\%$ confidence interval for the mean μ of a normal population when the value of σ is known is given by

$$\left(\bar{x} - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}\right) \tag{8.5}$$

or, equivalently, by $\bar{x} \pm z_{\alpha/2} \cdot \sigma/\sqrt{n}$.

Interpreting confidence intervals



95% confidence interval: If we repeat the experiment many times, the interval contains μ about 95% of the time

Assumptions

- So far
  - Normal distribution
  - σ is known
- Large-sample setting
  - ~~Normal distribution~~
    $\rightarrow$ use Central Limit Theorem $\rightarrow$ needs $n > 30$
  - ~~σ is known~~
    $\rightarrow$ replace σ by s $\rightarrow$ needs $n > 40$

Sample standard deviation

- Chapter 1: Given a data set $x_1, x_2, \ldots, x_n$, the sample standard deviation, denoted by s, is given by

$$s = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n}}$$

- As a random variable:

$$S = \sqrt{\frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n}}$$

Note: when n is small, it is recommended that we use

$$S = \sqrt{\frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n-1}}, \quad s = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}}$$

Principles

- Central Limit Theorem
$$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$$
is approximately normal when $n > 30$
- Moreover, when n is sufficiently large $s \approx \sigma$
- Conclusion:
$$\frac{\bar{X}-\mu}{S/\sqrt{n}}$$
is approximately normal when n is sufficiently large

If $n > 40$, we can ignore the normal assumption and replace σ by s

95% confidence interval

If after observing $X_1 = x_1$, $X_2 = x_2, \ldots$, $X_n = x_n$ $(n > 40)$, we compute the observed sample mean $\bar{x}$ and sample standard deviation s. Then

$$\left(\bar{x} - 1.96\frac{s}{\sqrt{n}}, \bar{x} + 1.96\frac{s}{\sqrt{n}}\right)$$

is a 95% confidence interval of μ

$100(1-\alpha)\%$ confidence interval

If after observing $X_1 = x_1$, $X_2 = x_2, \ldots, X_n = x_n$ $(n > 40)$, we compute the observed sample mean $\bar{x}$ and sample standard deviation s. Then

$$\left(\bar{x} - z_{\alpha/2}\frac{s}{\sqrt{n}}, \bar{x} + z_{\alpha/2}\frac{s}{\sqrt{n}}\right)$$

is a 95% confidence interval of μ

One-sided CIs

A **large-sample upper confidence bound for μ** is

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

and a **large-sample lower confidence bound for μ** is

$$\mu > \bar{x} - z_{\alpha} \cdot \frac{s}{\sqrt{n}}$$

Example

Example

A random sample of 50 patients who had been seen at an outpatient clinic was selected, and the waiting time to see a physician was determined for each one, resulting in a sample mean time of 40.3 min and a sample standard deviation of 28.0 min.

- Construct an 95% confidence interval of the true average waiting time.
- Assuming it is known that the true standard deviation of the waiting time is 27 min, construct an 95% confidence interval of the true average waiting time.

Example

Example

A sample of 66 obese adults was put on a low-carbohydrate diet for a year. The (sample) average weight loss was 7.7 lb and the sample standard deviation was 19.1 lb. Calculate a 99% lower confidence bound for the true average weight loss.

Question: From the result, do you think that the diet is effective?