

M5S1 - Confidence Intervals

Professor Jarad Niemi

STAT 226 - Iowa State University

October 9, 2018

Outline

- Confidence intervals for the population mean when the population standard deviation is known
 - Relation to Central Limit Theorem
 - Based on the Empirical Rule
 - Finding z critical values
 - significance level
 - confidence level
 - margin of error

Sample mean as an estimator for the population mean

Recall that due to the CLT, $\bar{X} \sim N(\mu, \sigma^2/n)$ where

- $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ is the (random) sample mean,
- μ is the population mean,
- σ^2 is the population variance, and
- n is the sample size.

Suppose μ is unknown. Then \bar{X} is an unbiased estimator for μ , since

$$E[\bar{X}] = \mu,$$

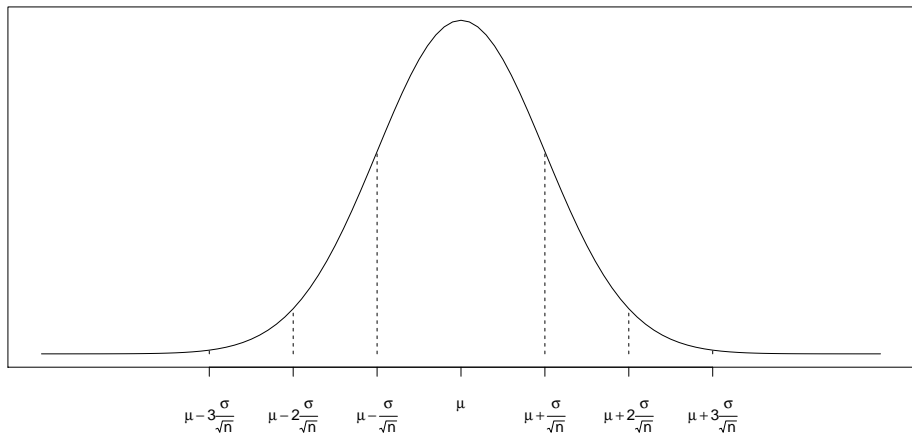
and its variability decreases with increased sample size since

$$SD[\bar{X}] = \sqrt{Var[\bar{X}]} = \sigma/\sqrt{n}.$$

How can we use this knowledge to describe our uncertainty in μ ?

How close is \bar{X} to μ ?

Sampling distribution for sample mean



Empirical Rule Confidence Intervals

From the Central Limit Theorem, we can write

$$\begin{aligned}
 P\left(\mu - \frac{\sigma}{\sqrt{n}} < \bar{X} < \mu + \frac{\sigma}{\sqrt{n}}\right) &\approx 0.68 \\
 P\left(\mu - 2\frac{\sigma}{\sqrt{n}} < \bar{X} < \mu + 2\frac{\sigma}{\sqrt{n}}\right) &\approx 0.95 \\
 P\left(\mu - 3\frac{\sigma}{\sqrt{n}} < \bar{X} < \mu + 3\frac{\sigma}{\sqrt{n}}\right) &\approx 0.997
 \end{aligned}$$

We can rewrite these inequalities by subtracting \bar{X} , subtracting μ , and multiplying by -1:

$$\begin{aligned}
 P\left(\bar{X} - \frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + \frac{\sigma}{\sqrt{n}}\right) &\approx 0.68 \\
 P\left(\bar{X} - 2\frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + 2\frac{\sigma}{\sqrt{n}}\right) &\approx 0.95 \\
 P\left(\bar{X} - 3\frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + 3\frac{\sigma}{\sqrt{n}}\right) &\approx 0.997
 \end{aligned}$$

We will call these intervals, e.g. $\left(\bar{X} - \frac{\sigma}{\sqrt{n}}, \bar{X} + \frac{\sigma}{\sqrt{n}}\right)$, **confidence intervals** and their **confidence level** is the probability (usually written as a percentage).

Example

US Bank provides students with savings accounts having no monthly maintenance fee and a low minimum monthly transfer. US Bank is interested in knowing the mean monthly balance of all its student savings accounts. They know the standard deviation of balances is \$20. They take a random sample of 64 student savings accounts and record that at the end of the month the sample mean savings was \$105. Construct a 68% confidence interval for the mean monthly balance.

Let X_i be the end of the month balance for student i . Then $E[X_i] = \mu$, the mean monthly balance, is unknown, but $SD[X_i] = \sigma = \$20$ is known. We obtained a sample of size $n = 64$ with a sample mean $\bar{x} = \$105$. To obtain the 68% confidence interval for μ , we calculate

$$\begin{aligned}\bar{x} \pm \frac{\sigma}{\sqrt{n}} &= \left(\bar{x} - \frac{\sigma}{\sqrt{n}}, \bar{x} + \frac{\sigma}{\sqrt{n}} \right) \\ &= \left(\$105 - \frac{\$20}{\sqrt{64}}, \$105 + \frac{\$20}{\sqrt{64}} \right) \\ &= (\$102.5, \$107.5)\end{aligned}$$

Confidence Intervals for μ when σ is known

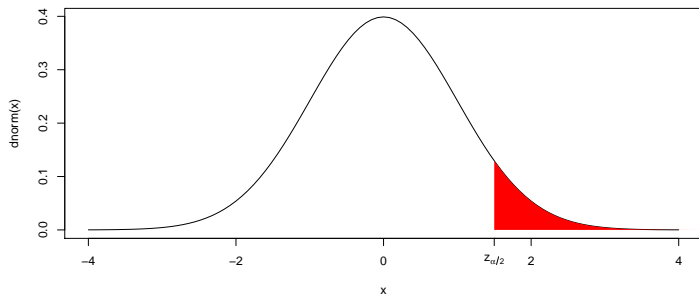
Definition

Let μ be the population mean and σ be the **known** population standard deviation. Choose a **significance level** α which you can convert to a **confidence level** $C = 100(1 - \alpha)\%$ and a **z critical value** $z_{\alpha/2}$ where $P(Z > z_{\alpha/2}) = \alpha/2$. You obtain a **random sample** of observations from the population and calculate the sample mean \bar{X} . Then a $C = 100(1 - \alpha)\%$ **confidence interval for μ** is

$$\bar{X} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}} = \left(\bar{X} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}}, \bar{X} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right)$$

where $z_{\alpha/2} \cdot \sigma / \sqrt{n}$ is called the **margin of error**.

Finding z critical values



Recall that

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

Check that

C	α	$\alpha/2$	$z_{\alpha/2}$
68%	0.32	0.16	≈ 1
95%	0.05	0.025	≈ 2
99.7%	0.003	0.0015	≈ 3

Example

US Bank provides students with savings accounts having no monthly maintenance fee and a low minimum monthly transfer. US Bank is interested in knowing the mean monthly balance of all its student savings accounts. They know the standard deviation of balances is \$20. They take a random sample of 64 student savings accounts and record that at the end of the month the sample mean savings was \$105. Construct a 80% confidence interval for the mean monthly balance.

Let X_i be the end of the month balance for student i . Then $E[X_i] = \mu$, the mean monthly balance, is unknown, but $SD[X_i] = \sigma = \$20$ is known. We obtained a sample of size $n = 64$ with a sample mean $\bar{x} = \$105$. For a confidence level of 80%, we have $\alpha = 0.2$, $\alpha/2 = 0.1$ and $z_{\alpha/2} \approx 1.28$. Then we calculate

$$\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}} = \$105 \pm 1.28 \frac{\$20}{\sqrt{64}} = (\$101.8, \$108.2)$$

which is an 80% confidence interval for μ