Confidence Interval Estimation

James H. Steiger

Department of Psychology and Human Development
Vanderbilt University

Confidence Interval Estimation

1. Introduction
2. The Sampling Distribution of M Revisited
3. Taking a Walk with Mr. Mu
4. Constructing a Confidence Interval
5. The Advantages of Confidence Intervals
   - Relationship to the 2-Tailed Hypothesis Test
   - Relationship to the 1-Tailed Hypothesis Test
6. Confidence Intervals with the t Distribution
7. Let's Practice!

Introduction

- In several previous lectures, we studied, in detail, the characteristics of the Z-statistic and the Student t test for comparing a mean with a null-hypothesized value.
- In recent years, a number of authorities on statistical methods have expressed the notion that another form of inference, *confidence interval estimation*, offers all the advantages of hypothesis tests, and more.
- While this view is still somewhat controversial, it seems to be gaining wide credence.
- In this module, we'll discuss, from the ground up,
  1. What a confidence interval is, and
  2. How to compute and evaluate it

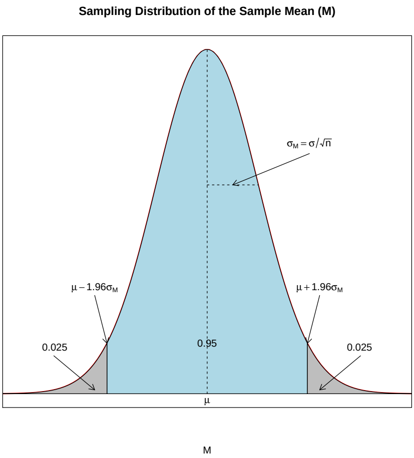
The Sampling Distribution of M Revisited

- In earlier lectures, we discussed in detail the fact that, because of the Central Limit Theorem effect, over a wide variety of circumstances, the sample mean M based on n *independent* observations from a standard normal distribution with mean μ and finite standard deviation σ has an approximately normal distribution with mean μ and standard deviation $\sigma_M = \sigma/\sqrt{n}$.

The Sampling Distribution of M Revisited

- As with any normal distribution, 95% of the time values of M will be within ± 1.96 standard deviations of the mean.
- The plot on the next page shows the basic situation. *Regardless of the value of* μ, in the long run over repeated samples, M will be within a "critical distance" of $\pm 1.96\sigma/\sqrt{n}$ from μ 95% of the time.
- This zone is diagrammed in light blue.

The Sampling Distribution of M Revisited



Taking a Walk with Mr. Mu

- In the previous slide, we demonstrated that 95% of the time, M will be within a certain "critical distance" of μ. We can therefore state that, in the long run,

$$\Pr\left(\mu - 1.96\frac{\sigma}{\sqrt{n}} \leq M \leq \mu + 1.96\frac{\sigma}{\sqrt{n}}\right) = .95 \qquad (1)$$

- After applying some standard manipulations of inequalities, we can manipulate the μ to the inside of the equality and the M to the outside, obtaining

$$\Pr\left(M - 1.96\frac{\sigma}{\sqrt{n}} \leq \mu \leq M + 1.96\frac{\sigma}{\sqrt{N}}\right) = .95 \qquad (2)$$

- Equation 2 implies that, if we take M and add and subtract the "critical distance" $1.96\sigma/\sqrt{n}$,we obtain an interval that contains the true μ, in the long run, 95% of the time.
- Such an interval is called a 95% *confidence interval* for μ.

Taking a Walk with Mr. Mu

- Even if you are not familiar with the manipulation of inequalities, there is a way of seeing how Equation 2 follows from Equation 1.
- The first inequality states that there is a critical distance, $1.96\sigma/\sqrt{n}$. and M is within that distance of μ 95% of the time, over repeated samples.
- Imagine that you had a friend named Mr. Mu, and you went for a stroll with him. After a certain length of time, he turned to you and said, "You know, about 95% of the time, you've been walking within 2 feet of me."
- You could, of course, reply that 95% of the time, he has also been within 2 feet of you!
- The point is, if M is within a certain distance of μ 95% of the time, it must also be the case (because distances are symmetric) that μ is within the same distance of M 95% of the time.

Constructing a Confidence Interval

- Equation 1 leads directly to the formula for constructing a confidence interval on a single mean when σ is known.
- We need to digress and develop some more notation.
  1. A confidence interval that includes the true parameter 95% of the time is referred to as a 0.95 confidence interval or a 95% confidence interval.
  2. More generally, we refer to a $1 - \alpha$ confidence interval (*or* a $100(1 - \alpha)\%$ confidence interval), with α being the proportion of the time the confidence interval does not include the true parameter.
  3. The proportion $1 - \alpha$ is sometimes called the *coverage rate* for the confidence interval.
- To construct the $1 - \alpha$ confidence interval, simply take

$$M \pm Z_{1-\alpha/2} \times \frac{\sigma}{\sqrt{n}} \tag{3}$$

- In the above equation, $Z_{1-\alpha/2}$ stands for the critical value of the standard normal distribution at the $1 - \alpha/2$ quantile. Note that the critical Z for the $1 - \alpha$ confidence interval is the same as the critical Z for the 2-tailed t-test that $\mu = \mu_0$ at the α significance level.

Constructing a Confidence Interval
An Example

Example (Confidence Interval on a Single Mean with σ Known)

Suppose you take a sample of size 49, and observe a sample mean of 104.88 and it is known that $\sigma = 15$. Construct a 95% confidence interval for μ.

(Answer on next slide...)

Constructing a Confidence Interval

An Example

Example (Confidence Interval on a Single Mean with σ Known)

Suppose you take a sample of size 49, and observe a sample mean of 104.88 and it is known that $\sigma = 15$. Construct a 95% confidence interval for μ.

Answer. Using Equation 3 we get

$$\begin{aligned} 104.88 &\pm 1.96 \times \frac{15}{\sqrt{49}} \\ 104.88 &\pm 1.96 \times \frac{15}{7} \\ 104.88 &\pm 4.2 \end{aligned}$$

We can compute the endpoints of the interval in R as

```
> 104.88 - 1.96*15/7
[1] 100.68
> 104.88 + 1.96*15/7
[1] 109.08
```

We might say that we are "95% confident that μ is between 100.68 and 109.08."

The Advantages of Confidence Intervals

- Confidence intervals provide all the information in a hypothesis test, and more. You can use a confidence interval to perform a hypothesis test.
- When plotted, confidence intervals provide a visual display of the relative precision of several experiments, and may help to explain why some hypothesis tests fail to replicate.
- In many exploratory studies, there is no null hypothesis. The confidence interval shows what we have learned about the value of μ, without requiring that we state any hypothesis about it. It that sense, confidence intervals may match the goals of modern psychological and educational experimentation more closely than hypothesis tests.

Relationship to the 2-Tailed Hypothesis Test

- First, we see how to use a confidence interval to perform a hypothesis test.
- Suppose a hypothesis test of the 2-tailed hypothesis $\mu = \mu_0$ is performed with $\alpha = 0.05$. *The hypothesis will be rejected if and only if the confidence interval does not include* μ_0 *within its limits.*
- The implication is clear:
  1. Compute the $1 - \alpha$ confidence interval.
  2. See if the confidence interval includes the value μ_0.
  3. If the confidence interval includes μ_0, do not reject H_0.
  4. If the confidence interval excludes μ_0, reject H_0.

Relationship to the 1-Tailed Hypothesis Test

- A confidence interval may also be used to perform a 1-tailed test.
- Suppose a hypothesis test of the 1-tailed hypothesis $\mu \leq \mu_0$ is performed with $\alpha = 0.05$. *The hypothesis will be rejected if and only if the* $1 - 2\alpha$ *confidence interval lies entirely above* μ_0.
- The implication is clear:
  1. Compute the $1 - 2\alpha$ confidence interval.
  2. See if the confidence interval lies above the value μ_0.
  3. If the confidence interval does not lie above μ_0, do not reject H_0.
  4. If the confidence interval is completely above μ_0, reject H_0.
- A similar rule holds for a hypothesis of the form $\mu \geq \mu_0$, except that the entire $1 - 2\alpha$ confidence interval must lie below μ_0 to indicate rejection of the 1-tailed hypothesis at the α significance level.

Confidence Intervals with the t Distribution

- In a previous module, we discussed how, when σ is not known, we modify the Z statistic by using the sample standard deviation s and use the Student t distribution.
- The confidence interval for μ is modified in a similar manner.
- To construct the $1 - \alpha$ confidence interval, simply take

$$M \pm t_{n-1,1-\alpha/2} \times \frac{s}{\sqrt{n}} \tag{4}$$

- In the above equation, $t_{n-1,1-\alpha/2}$ stands for the critical value of t with $n-1$ degrees of freedom at the $1 - \alpha/2$ quantile. Note that the critical t for the $1 - \alpha$ confidence interval is the same as the critical t for the 2-tailed t-test that $\mu = \mu_0$ at the α significance level.
- Compared to a hypothesis test, confidence intervals using the t distribution have the same benefits as those using the normal distribution. They also may be used to perform 1-tailed hypothesis tests using the methods discussed above.

Let's Practice!

Example (01: Computing a Confidence Interval)

Suppose you observe a sample mean of $M = 44.65$ and a sample standard deviatio of $s = 10.10$ based on a sample of $n = 34$ independent observations from a normal population with a normal distribution. Construct a 0.99 confidence interval for the population mean μ.

(Answer on next slide...)

Let's Practice!

Example (01: Computing a Confidence Interval continued . . .)

Suppose you observe a sample mean of $M = 44.65$ and a sample standard deviation of $s = 10.10$ based on a sample of $n = 34$ independent observations from a normal population with a normal distribution. Construct a 0.99 confidence interval for the population mean μ.

Answer. Since this is a $1 - \alpha$ confidence interval, $\alpha = 0.01$, the critical t value is at the $1 - \alpha/2 = 1 - .01/2 = 0.995$ quantile. This can be obtained using R as

```
> n <- 34
> alpha <- 0.01
> t.crit <- qt(1 - alpha/2,n-1)
> t.crit
[1] 2.733277
```

The confidence interval is constructed as

$$\begin{aligned} M &\pm t_{crit} \times \frac{s}{\sqrt{n}} \\ 44.65 &\pm 2.7333 \times \frac{10.1}{\sqrt{34}} \\ 44.65 &\pm 4.7344 \end{aligned}$$

Let's Practice!

Example (02: Testing a Hypothesis with a Confidence Interval)

Suppose you wish to test a hypothesis that $\mu = 10$ at the 0.05 significance level (i.e., with $\alpha = 0.05$). You decide to compute the confidence interval on μ and use the confidence interval to perform the hypothesis test. You have a sample size of $n = 47$.

1. What confidence level, 0.90, 0.95, 0.99 should you use in constructing the confidence interval?
2. What critical value of the t distribution will you use to construct the confidence interval?

(Answer on the next slide ...)

Let's Practice!

Example (02: Testing a Hypothesis with a Confidence Interval)

Answer. In this case, you need a $1 - \alpha = 1 - 0.05 = 0.95$ confidence interval, because your hypothesis test is 2-tailed. The critical value of t that you need is the $0.975 = 1 - \alpha/2$ quantile of the t distribution with $n - 1 = 47 - 1 = 46$ degrees of freedom

```
> qt(0.975,46)
[1] 2.012896
```

Let's Practice!

Example (03: Testing a Hypothesis with a Confidence Interval)

You have a null hypothesis that $\mu \leq 0$. You construct a 0.90 confidence interval, and find that the confidence interval has endpoints 1.76 and 12.44. Which of the following statements definitely true?

a. The null hypothesis is not rejected.
b. The null hypothesis is rejected at the 0.01 level.
c. The null hypothesis is rejected at the 0.05 level.

Let's Practice!

Example (03: Testing a Hypothesis with a Confidence Interval)

Answer. The correct answer is (c).

- This is a 1-tailed hypothesis, with the rejection region in the upper tail.
- A 0.90 confidence interval corresponds to a 1-tailed test at the 0.05 significance level. The null hypothesis is rejected if and only if the confidence interval *is entirely above* μ_0.
- In this case, $\mu_0 = 0$, and the confidence interval is entirely above 0, so the null hypothesis is rejected at the 0.05 level.
- Would the null hypothesis be rejected at the 0.01 level? This is a complex question, and at first glance it might appear that there is no way we can answer it. However, with a little statistical detective work, we can in fact determine that, if an appropriate confidence interval were constructed (a 0.98 confidence interval) it would not exclude 0, and the null hypotehsis would not be rejected at the 0.01 level with the same data.