§8.4–Confidence Intervals for One Population Mean When σ is Unknown

Tom Lewis

Fall Term 2009

Outline

If σ is known

If we draw a random sample of size n from a population with mean μ and standard deviation σ, then the distribution of the sample mean, $\overline{x}$, is normal with mean μ and standard deviation $\sigma/\sqrt{n}$. In other words,

$$z = \frac{\overline{x} - \mu}{\sigma/\sqrt{n}}$$

has a standard normal distribution.

If σ is unknown

If σ is unknown

- In sampling from a population to estimate μ, it is more often the case that we do not know σ. In such cases, we substitute the sample standard deviation, s, for the population standard deviation. This has important and unexpected consequences.

If σ is unknown

- In sampling from a population to estimate μ, it is more often the case that we do not know σ. In such cases, we substitute the sample standard deviation, s, for the population standard deviation. This has important and unexpected consequences.
- View the simulation.

If σ is unknown

- In sampling from a population to estimate μ, it is more often the case that we do not know σ. In such cases, we substitute the sample standard deviation, s, for the population standard deviation. This has important and unexpected consequences.
- View the simulation.
- If we draw a random sample of size n from a population with mean μ and unknown standard deviation, then the statistic

$$t = \frac{\overline{x} - \mu}{s/\sqrt{n}}$$

called the studentized version of $\overline{x}$ is not normally distributed.

If σ is unknown

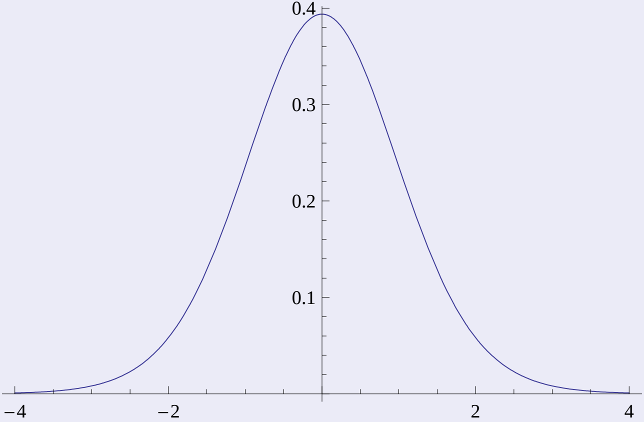
- In sampling from a population to estimate μ, it is more often the case that we do not know σ. In such cases, we substitute the sample standard deviation, s, for the population standard deviation. This has important and unexpected consequences.
- View the simulation.
- If we draw a random sample of size n from a population with mean μ and unknown standard deviation, then the statistic

$$t = \frac{\overline{x} - \mu}{s/\sqrt{n}}$$

  called the studentized version of $\overline{x}$ is not normally distributed.
- The distribution of t is called a t-distribution with $n - 1$ degrees of freedom.

A *t*-distribution curve

Here is a picture of a *t*-distribution curve with 19 degrees of freedom.



Basic properties of t-curves

Basic properties of *t*-curves

- The total area under a t-curves is 1.

Basic properties of t-curves

- The total area under a t-curves is 1.
- A t-curve is always positive, but it tends toward 0 in both the positive and negative directions.

Basic properties of *t*-curves

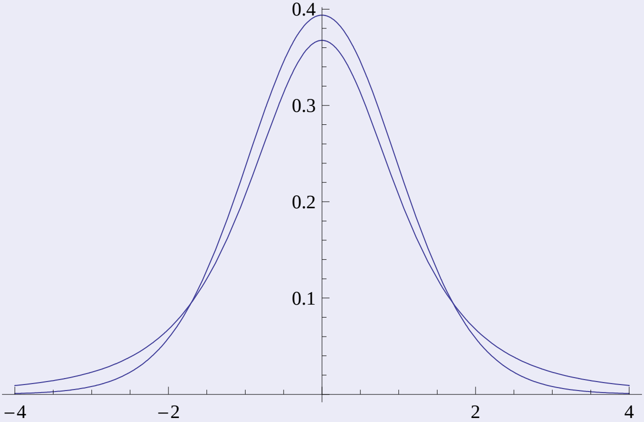
- The total area under a t-curves is 1.
- A *t*-curve is always positive, but it tends toward 0 in both the positive and negative directions.
- A *t*-curve is symmetric about the *y*-axis.

Basic properties of *t*-curves

- The total area under a t-curves is 1.
- A *t*-curve is always positive, but it tends toward 0 in both the positive and negative directions.
- A *t*-curve is symmetric about the *y*-axis.
- As the number of degrees of freedom becomes larger, *t*-curves look increasingly like the standard normal curve.

A comparison of *t*-distribution curves

Here are pictures of the graphs of two *t*-distribution curves. The squat curve has 3 degrees of freedom and the other has 19 degrees of freedom.



Area under the t-curve

Our text does not have a t-table to give the area under a specified region of a t-curve. This can be done by **R Commander** using

Distributions → Continuous distributions
→ t distribution → t probabilities ...

Upon filling in the dialog box, the desired probability will be issued.

Area under the *t*-curve

Our text does not have a *t*-table to give the area under a specified region of a *t*-curve. This can be done by **R Commander** using

Distributions → Continuous distributions

→ *t* distribution → *t* probabilities . . .

Upon filling in the dialog box, the desired probability will be issued.

Problem

Area under the t-curve

Our text does not have a t-table to give the area under a specified region of a t-curve. This can be done by **R Commander** using

Distributions → Continuous distributions
→ t distribution → t probabilities . . .

Upon filling in the dialog box, the desired probability will be issued.

Problem

- *Find the area under the t-curve with 19 degrees of freedom to the left of* 1.28.

Area under the t-curve

Our text does not have a t-table to give the area under a specified region of a t-curve. This can be done by **R Commander** using

Distributions → Continuous distributions

→ t distribution → t probabilities ...

Upon filling in the dialog box, the desired probability will be issued.

Problem

- *Find the area under the t-curve with 19 degrees of freedom to the left of* 1.28.
- *Find the area under the t-curve with 26 degrees of freedom to the right of* 2.25.

Area under the t-curve

Our text does not have a t-table to give the area under a specified region of a t-curve. This can be done by **R Commander** using

Distributions → Continuous distributions → t distribution → t probabilities ...

Upon filling in the dialog box, the desired probability will be issued.

Problem

- *Find the area under the t-curve with 19 degrees of freedom to the left of* 1.28.
- *Find the area under the t-curve with 26 degrees of freedom to the right of* 2.25.
- *Find the area under the t-curve with 11 degrees of freedom between* -1.35 *and* 1.86.

The quantity t_α

- Let $0 < \alpha < 1$. Given a number of degrees of freedom, t_α denotes the number such that the area under the corresponding t-curve to the right of t_α is α.
- This can be done by **R Commander** using

  Distributions $\rightarrow$ Continuous distributions

  $\rightarrow$ t distribution $\rightarrow$ t quantiles ...

  Upon filling in the dialog box, the desired value of t_α will be issued.

Problem

Find the value of $t_{.05}$ for a t-distribution with 5 degrees of freedom. Use **R Commander** *and the table in the text.*

Finding a confidence interval for the mean

Finding a confidence interval for the mean

- Let $0 < \alpha < 1$ be specified. We wish to find a $1 - \alpha$ confidence interval for the mean, μ, of a normal population whose standard deviation, σ, is unknown.

Finding a confidence interval for the mean

- Let $0 < \alpha < 1$ be specified. We wish to find a $1 - \alpha$ confidence interval for the mean, μ, of a normal population whose standard deviation, σ, is unknown.
- Let a random sample of size n be drawn from the population. Calculate $\overline{x}$. s, and $t_{\alpha/2}$ for $n - 1$ degrees of freedom.

Finding a confidence interval for the mean

- Let $0 < \alpha < 1$ be specified. We wish to find a $1 - \alpha$ confidence interval for the mean, μ, of a normal population whose standard deviation, σ, is unknown.
- Let a random sample of size n be drawn from the population. Calculate $\overline{x}$. s, and $t_{\alpha/2}$ for $n - 1$ degrees of freedom.
- With probability $1 - \alpha$, μ will land in the interval

$$\left[\overline{x} - t_{\alpha/2}\frac{s}{\sqrt{n}}, \overline{x} + t_{\alpha/2}\frac{s}{\sqrt{n}}\right]$$

Finding a confidence interval for the mean

- Let $0 < \alpha < 1$ be specified. We wish to find a $1 - \alpha$ confidence interval for the mean, μ, of a normal population whose standard deviation, σ, is unknown.
- Let a random sample of size n be drawn from the population. Calculate $\overline{x}$. s, and $t_{\alpha/2}$ for $n - 1$ degrees of freedom.
- With probability $1 - \alpha$, μ will land in the interval

$$\left[\overline{x} - t_{\alpha/2}\frac{s}{\sqrt{n}}, \overline{x} + t_{\alpha/2}\frac{s}{\sqrt{n}}\right]$$

- In other words, the chance that the actual mean falls outside of this interval is α; this represents our risk in using the stated confidence interval to estimate μ.

Newcomb's "speed of light" data

In 1878, Simon Newcomb worked out an ingenious method for measuring the speed of light. He measured the time (in nanoseconds) for light to travel from the United States Naval Observatory to the Washington Monument and back, a distance of some 7400 meters. His measurements are tabulated below. These measurements represent deviations from 24,800 nanoseconds.

28	26	33	24	34	-44	27	16	40	-2
29	22	24	21	25	30	23	29	31	19
24	20	36	32	36	28	25	21	28	29
37	25	28	26	30	32	36	26	30	22
36	23	27	27	28	27	31	27	26	33
26	32	32	24	39	28	24	25	32	25
29	27	28	29	16	23				

Find a 90% confidence interval for the mean of this data.