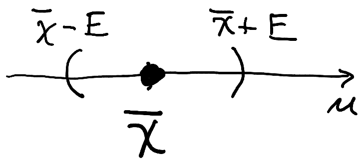
03/20/2023

Chapter 7 (continued)

A CI for a Population Mean (μ)



or $\bar{x} \pm E$

Suppose $x_1, x_2, \ldots, x_n$ is a random sample from a population where σ is unknown. Instead of a Z-variable, the resulting standard variable is

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

03/22/2023

Remarks

1. t above is not exactly the same as Z. Distribution of t is centered at zero but its standard deviation is greater than 1.

2. The distribution of t depends on the sample size. As n increases, the t-curve approaches the Z-curve.

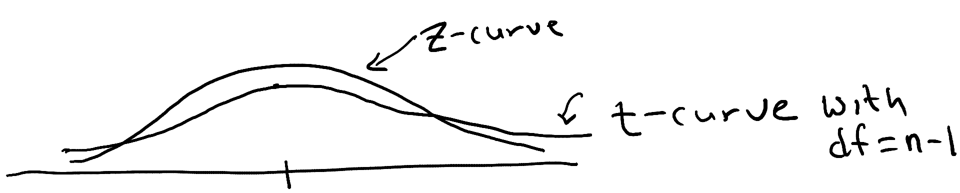
Notation: $Z \sim N(0,1)$

$t \sim$ t-distribution with $n-1$ degrees of freedom

In short, $t \sim t(df)$, where

df = degrees of freedom $= n-1$

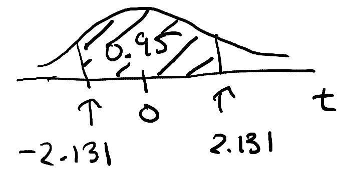
Ex:



$df = 15 \quad (n = 16)$

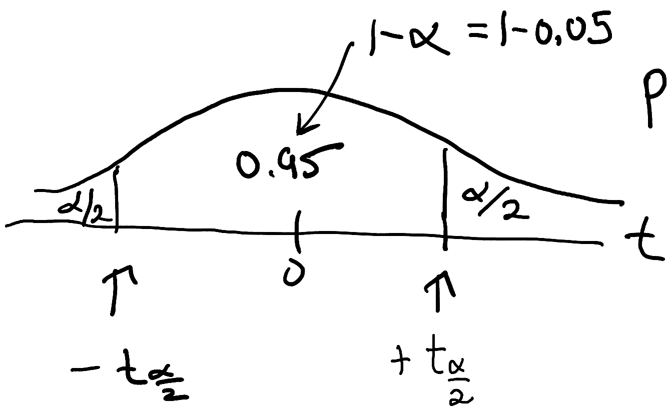
$$P(-2.131 < t < 2.131) = 0.95$$

df=15



$df = 20 \quad P(-2.845 < t < 2.845) = 0.99$

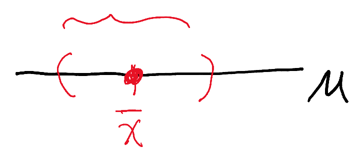
Ex: $df = 60 \quad (n = 61)$



$$P(-2.000 < t < 2.000) = 0.95$$

Result: If $x_1, x_2, \ldots, x_n$ is a random sample of size n from population where σ is unknown, then a t-based CI for μ is

$$\bar{x} \pm t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$$



Formally, a $100(1-\alpha)\%$ t-based CI for μ is

$\bar{x} \pm t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}}$, where $df = n - 1$.

Ex: $df = 15 \quad (n = 16)$

$$0.95 = P(-2.131 < t < 2.131) = P\left(-2.131 < \frac{\bar{x} - \mu}{S/\sqrt{n}} \leq 2.131\right)$$

$\vdots$

$= P(\bar{x} - 2.131\, S \ldots < \mu < \bar{x} + 2.131\, S \ldots)$

$$\vdots$$

$$= P\left(\underbrace{\bar{x} - 2.131\frac{s}{\sqrt{n}}} < \mu < \underbrace{\bar{x} + 2.131\frac{s}{\sqrt{n}}}\right)$$

Endpoints vary from sample-to-sample

<u>Remark</u>: See p. 326 for a discussion for more information and cases where σ is known or population is not normally distributed and $n \leq 30$.

<u>Ex</u>: x = daily iron intake (mg) of an adult living in the U.S.

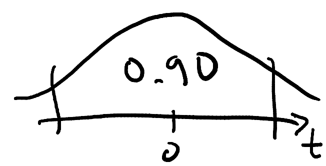
μ = mean daily iron intake of all U.S. adults

Construct a 90% CI for μ.

$n = 100$ $\qquad x_1, x_2, \ldots, x_{100}$

$\bar{x} = 16.50$ mg

$s = \sqrt{s^2} = \sqrt{\dfrac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}} = 3.08$ mg



$t_{\frac{\alpha}{2}} = t_{0.05} = 1.660$ $\qquad$ df = 99

$$\bar{x} \pm t_{0.05}\frac{s}{\sqrt{n}}$$

$$16.50 \pm 1.660 \times \frac{3.08}{\sqrt{100}}$$

16.50 ± 0.511

so a 90% CI for μ is $(15.99 \text{ mg}, 17.01 \text{ mg})$

$\underset{16.50 - 0.511}{\parallel}$

Comments:

Point estimate μ is 16.50mg

Center of the CI is 16.50mg

Width of the CI is $17.01 - 15.99 = 1.02$ mg

Is μ contained in the interval $(15.99, 17.01)$?

Answer: Cannot tell.

Is a 95% CI narrower or wider than a 90% CI?

Answer: Wider.

Remarks

1. An interval having a high confidence which is also narrow provides a lot of information about μ.

2. See Handouts 5 and 6 for additional information.

Remark:

$$\begin{aligned} E &= \text{half width of CI for } \mu \\ &= t_{\frac{\alpha}{2}} \frac{S}{\sqrt{n}} \\ &= \text{margin of error} \end{aligned}$$

Solving for n yields

Solving for n yields

$$n = \left(\frac{t_{\alpha/2} S}{E}\right)^2$$

For instance, the sample size required to estimate μ to within the distance E with 95% confidence is

$$n = \left(\frac{1.96\, S}{E}\right)^2$$

when n is "large".

A value of S is determined by using past information. One could use $\frac{\text{Range}}{4}$ if population is not too skewed.

"range rule of thumb"

Ex: X = weight of newborn manatee

Estimate μ.

How many manatees must be sampled to be 90% certain that the margin of error does not exceed 0.5 lb?

$E = 0.5$

$t_{0.05} = 1.645$ assuming n is "large" $\quad 0.05 = \frac{\alpha}{2}$

Suppose we have $s = 4$ lbs from a previous study.

$$n = \left(\frac{1.645 \times 4}{0.5} \right)^2 = 173.19 \quad \text{so use } n = 174.$$

Omit sections 7.3, 7.4