ZOL851
Standard errors, degrees of freedom, Z, t and Cis.
Part 1 & 2

Tuesday Sept 4 2014

Readings for this lecture

- Since this is "review" you probably already have older material to go through.
- Vasishth and Broe: Chapter 3 (make sure to run through their R code as well).
- I have put a draft of this text on ANGEL.

Screencasts

- Please go through the remaining three screencasts
- Control flow
- Apply like functions
- R for simulations (script in its own folder)

Readings for thursday

- We will be discussing effect size. There are three papers in the angel folder to read.

Concepts (and terms) to be familiar with by the end of class (with the help of your readings)

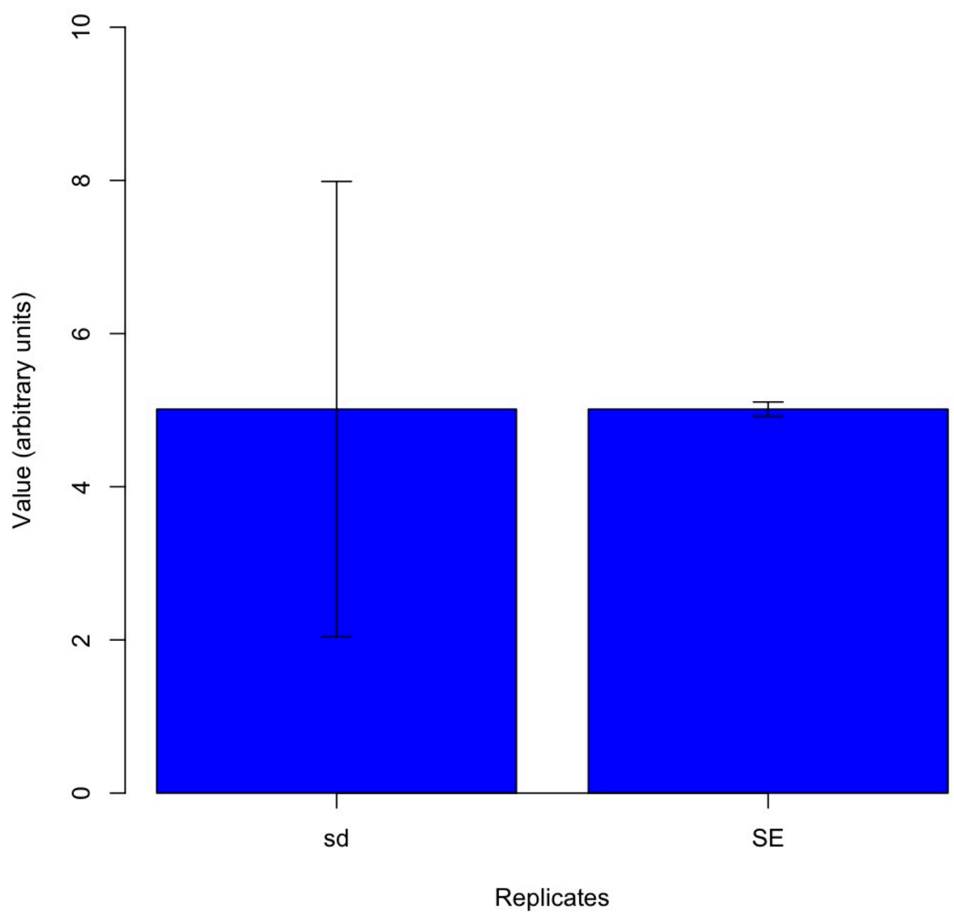
- Variance & Standard deviation
- Standard error
- Degrees of Freedom

Standard deviation, Standard Error, which one to use, and when?

- Some times people will report a figure with the standard error, sometimes with a standard deviation.

- Which (if any, or both) are appropriate? When?

Example



ppl will erroneously report SE so that their error bars look small. The confidence intervals would be somewhere in between these two.

SD is nicer than variance because SD is in the same units (variance is in squared units)

Variance, SD, SE

Why do we have all of these measures that seem to be based on the same expectation namely

$$SS_X = \sum (X_i - \bar{X})^2$$ Sum of squares

$$\bar{X} = \frac{1}{n} \sum X_i$$ Sample mean

$$s_X^2 = \frac{1}{n-1} \sum (X_i - \bar{X})^2$$ Sample variance (also sometimes called mean square)

$$s_X = \sqrt{\frac{1}{n-1} \sum (X_i - \bar{X})^2}$$ sample standard deviation

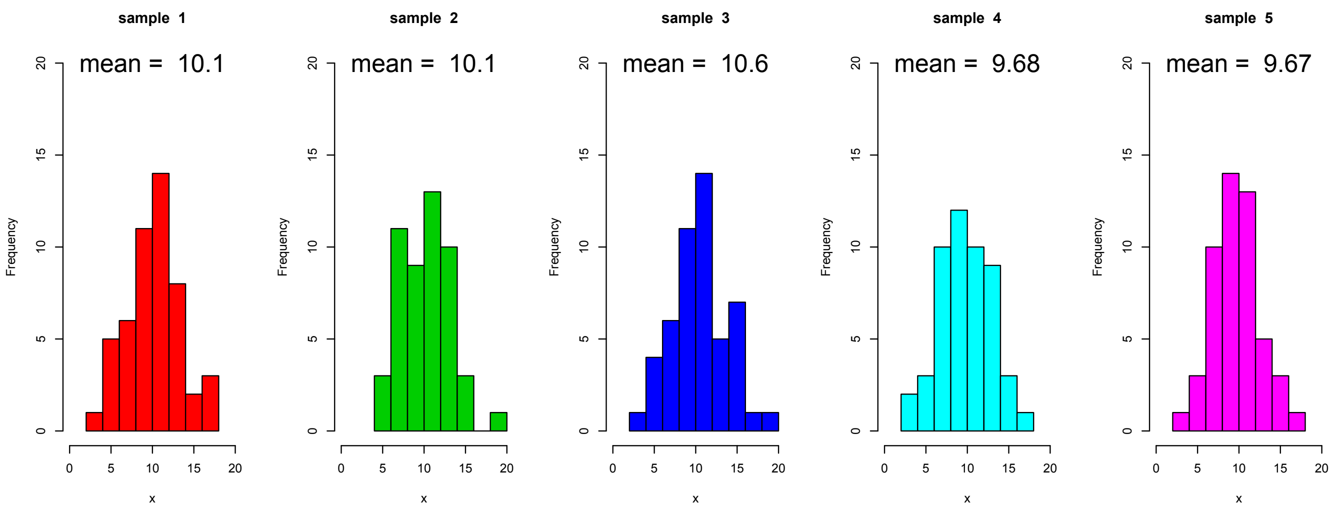
Standard Error

$$s_{\overline{X}} = \frac{s_X}{\sqrt{n}} = \sqrt{\frac{\frac{1}{n-1}\sum(X_i - \bar{X})^2}{n}}$$

Standard error of the "mean"

Let's think about what the above really means

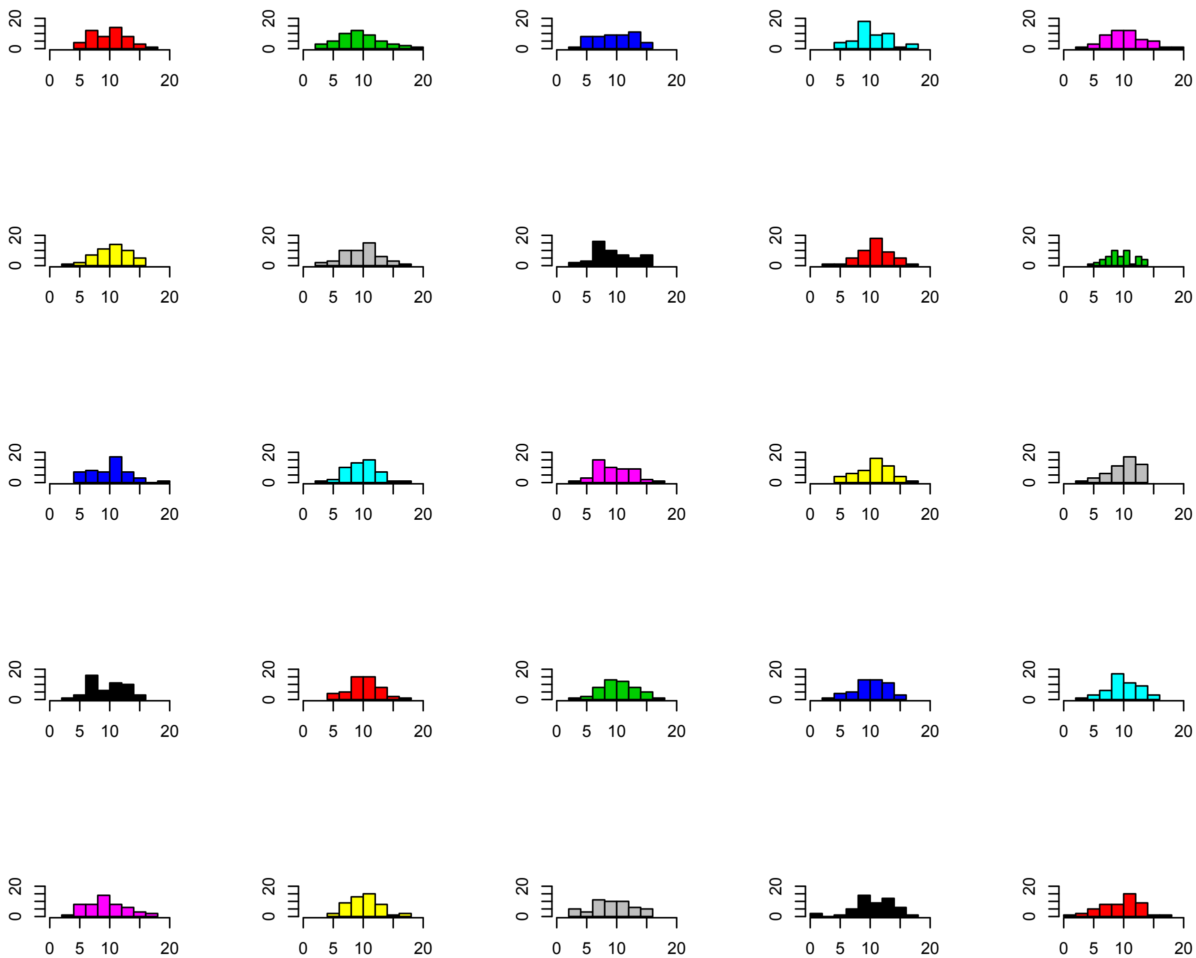
Standard Error



We now have a new vector of data composed of the mean

sample.means <- c(10.1, 10.1, 10.6, 9.68, 9.67)

Standard Error



Degrees of Freedom

- Why all of the n-1? For variance, and sd?
- We call this quantity, the degrees of freedom.
- What on earth is this?

What are we trying to capture with the ideas of “Degrees of Freedom”

R tutorial on standard error (via simulation).

- We will go through this on Thursday….

R tutorial

- Use the same data from last week
- Write functions to compute the coefficient of variation and the standard error of the mean.
- Place these in a file called MyFunctions.R
- Source these functions source("MyFunctions.R")
- Use these functions to compute the SE and CV of the femur, tibia, tarsus and SCT
- Compute the SE and CV for SCT for each "line" using an apply family function
- Do the same with a for loop.

So how can we start to use all of this in statistics?

- Let's say we have two data sets....
- How do we make sense of how different are they?

CSE891: t, Z, and CI: More than letters in the alphabet.

Warning: I am a Canadian. I proudly pronounce "Z" as "Zed", not "Zee"… So do members of all other English speaking countries outside of the US… Get the snickering out of your system now…… You have already converted my daughter to your insidious ways, but you will not get me…..

The utility of the sampling distribution

- Allows us to develop criteria to evaluate the estimators.
- Often we are interested in what the sampling distribution looks like at the extremes of the distribution (which is how we calculate p values).
- i.e. to make an inference: is the estimated value based on our observation extremely unlikely based on the sampling distribution?

The t-distribution

- So far we have only used the standard error with respect to the "**rule (of thumb) of 2**".

- However we can start to use it in a more rigorous fashion, for inference and to develop the ALL IMPORTANT concept of confidence intervals.

How do I know if my mean is different from 0 (or any other number)

- Say we measured the height of everyone in the class. The average height was 173.5cm with a standard error of 1.9
- How do we evaluate whether this is different than the (known) mean of 169.7 cm for all MSU students?
- Let’s formulate this as a hypothesis…

The t

height of sample

$$\bar{X} - \mu = 0 \quad \text{(No difference, the null hypothesis)}$$

height of population

$$\bar{X} - \mu \neq 0 \quad \text{(values are different)}$$

$$t = \frac{\bar{X} - \mu}{S_{\bar{X}}}$$

Xbar is the mean we estimated
Mu is the value (hypothesis we are comparing to)
S[Xbar] is the standard error of the mean.

value is scaled by our uncertainty. if uncertainty is low, t will be high (and vice versa)

The t

$$t = \frac{\overline{X} - \mu}{S_{\overline{X}}}$$

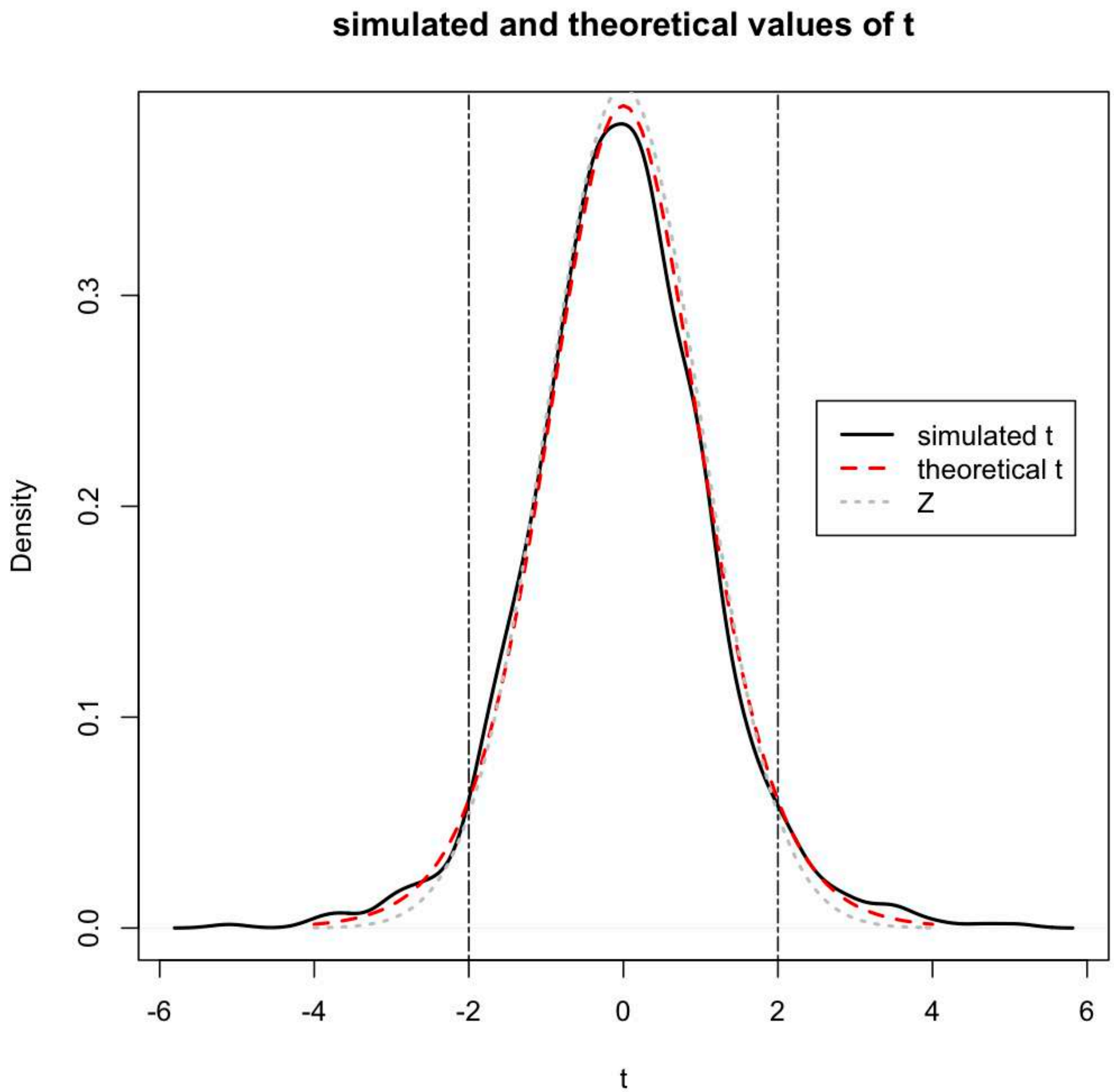
Let’s think about what this means…..

How do we know if the value of |t| is large enough to reject the null hypothesis?

$$t = \frac{\overline{X} - \mu}{S_{\overline{X}}}?$$

As we have seen with the mean and the standard error, we can think about the distribution of possible values of t.

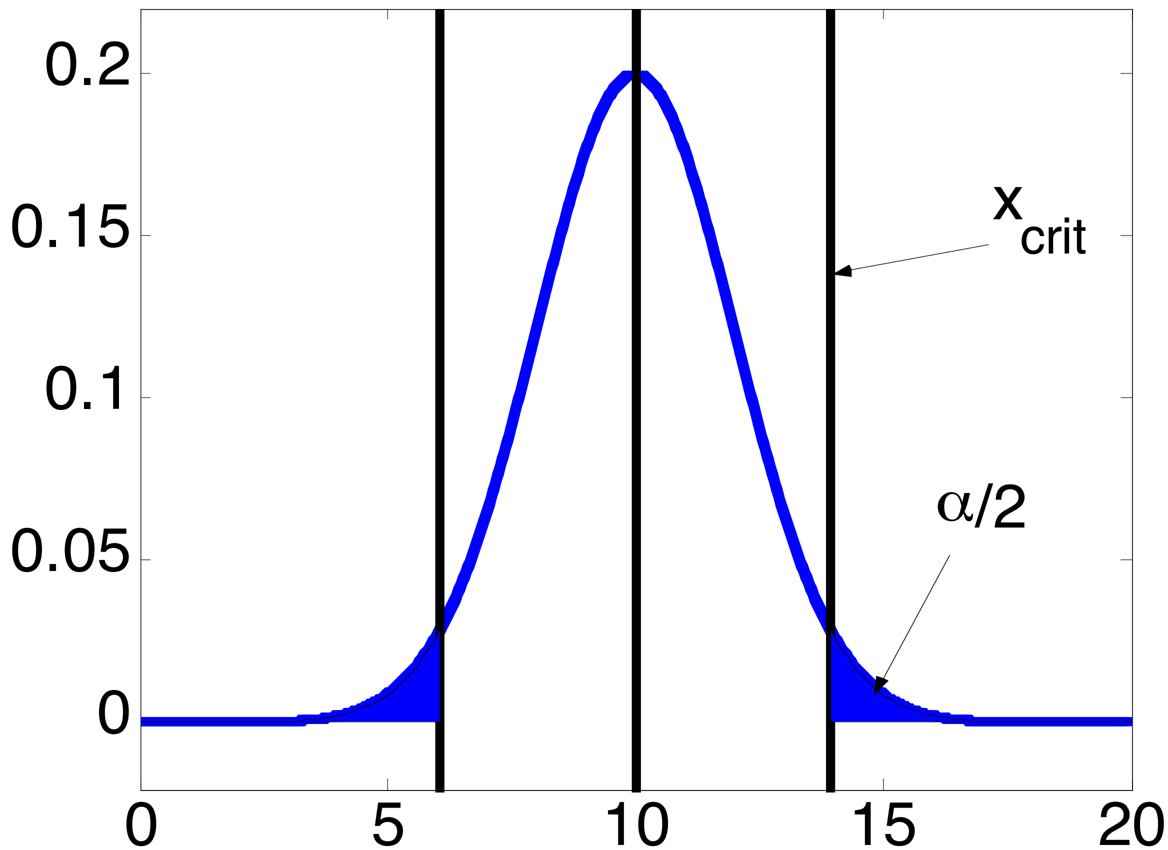
The t-distribution



The t distribution is a distribution of the t test statistic under the null model of no difference (t=0) given repeated sampling.

Back to R....

Critical values



t-distribution

Confidence Intervals

$$s_{\overline{X}} = \frac{\sigma_X}{\sqrt{n}} = \sqrt{\frac{\frac{1}{n-1}\sum(X_i - \bar{X})^2}{n}}$$

Standard error of the "mean"

95% Confidence interval

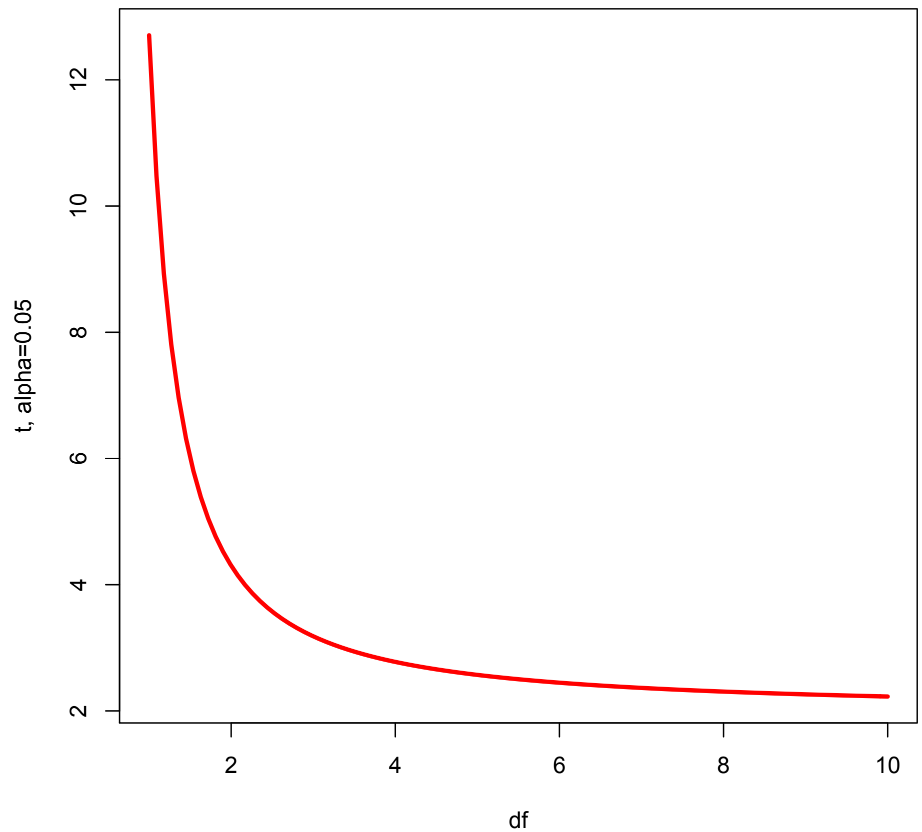
$$P(\overline{X} - 1.96 * S_{\overline{X}} \leq \theta \leq \overline{X} + 1.96 * S_{\overline{X}}) = 0.95$$

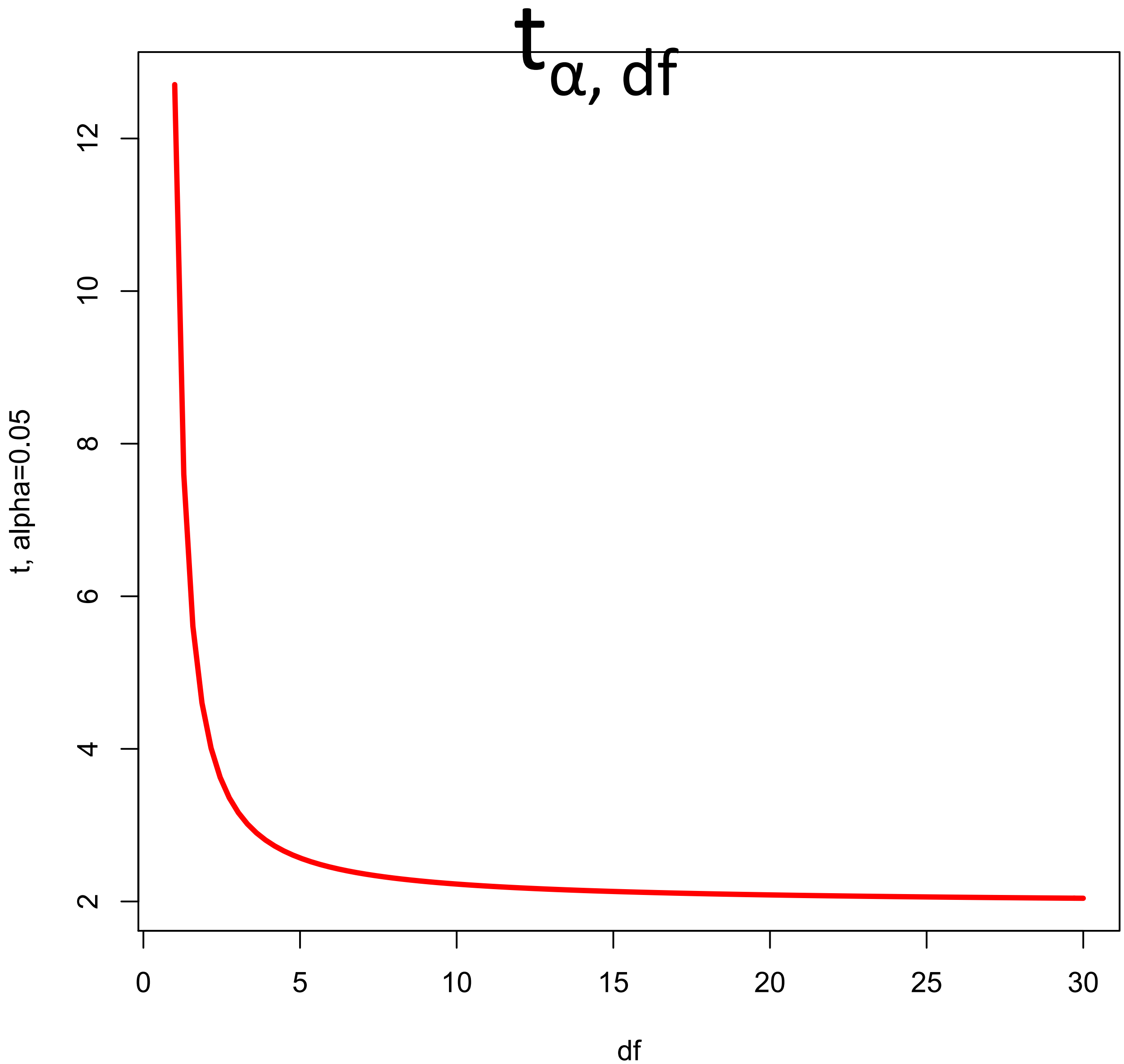
Rule of 2 version of CIs
t = 1.96 ~ 2 SDs
t corresponding to confidence multiplied by SE

$$P(\overline{X} - t_{\alpha,df} * S_{\overline{X}} \leq \theta \leq \overline{X} + t_{\alpha,df} * S_{\overline{X}}) = (1-\alpha)$$

Generalized confidence intervals

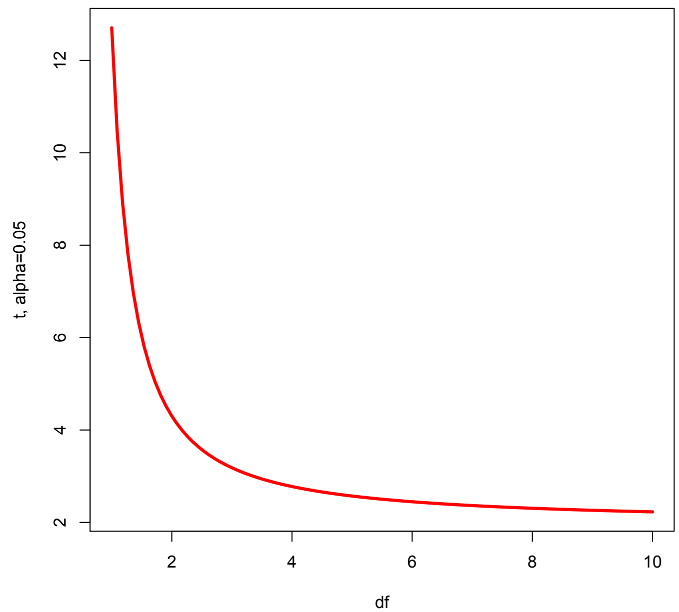
$t_{\alpha,\,df}$





$t_{\alpha,\,df}$
t, alpha=0.05
2
4
6
8
10
12
0
5
10
15
20
25
30
df

$t_{\alpha,\,df}$



Take home message: For very low sample sizes the 2*SE ~ for CI is far too narrow. This only works for sample sizes greater than ~20.

Confidence Intervals

$$s_{\overline{X}} = \frac{\sigma_X}{\sqrt{n}} = \sqrt{\frac{\frac{1}{n-1}\sum(X_i - \bar{X})^2}{n}}$$

Standard error of the "mean"

95% Confidence interval

$$P(\overline{X} - 1.96 * S_{\overline{X}} \leq \theta \leq \overline{X} + 1.96 * S_{\overline{X}}) = 0.95$$

$$P(\overline{X} - t_{\alpha,df} * S_{\overline{X}} \leq \theta \leq \overline{X} + t_{\alpha,df} * S_{\overline{X}}) = (1-\alpha)$$

Generalized confidence intervals

What is a confidence interval (verbal description)?

1 - alpha (for us, 95%) certain that the true value (true mean) lies within these values

Credible intervals vs Confidence intervals

CI do not reflect true value, but more with sampling. It’s 1.96 or (Xbar-T alpha,df) times the standard error of the population.

What is a 95(1-α)% confidence interval? Does this seem right?

"there is a 95(1-α)% chance that the true population mean μ occurs within the interval"

What is a 95(1-α)% confidence interval?

~~“there is a 95(1-α)% [illegible] that the true population me[illegible] occu[illegible]hin the interval”~~

Why not?

What is a 95(1-α)% confidence interval?

~~“there is a 95(1-α)% [illegible] hat the true population me[illegible] occu[illegible] hin the interval”~~

Why not?
Because under the frequentist/classical approach to statistics, **μ is fixed**! It can not be both within and outside of this interval!!!!

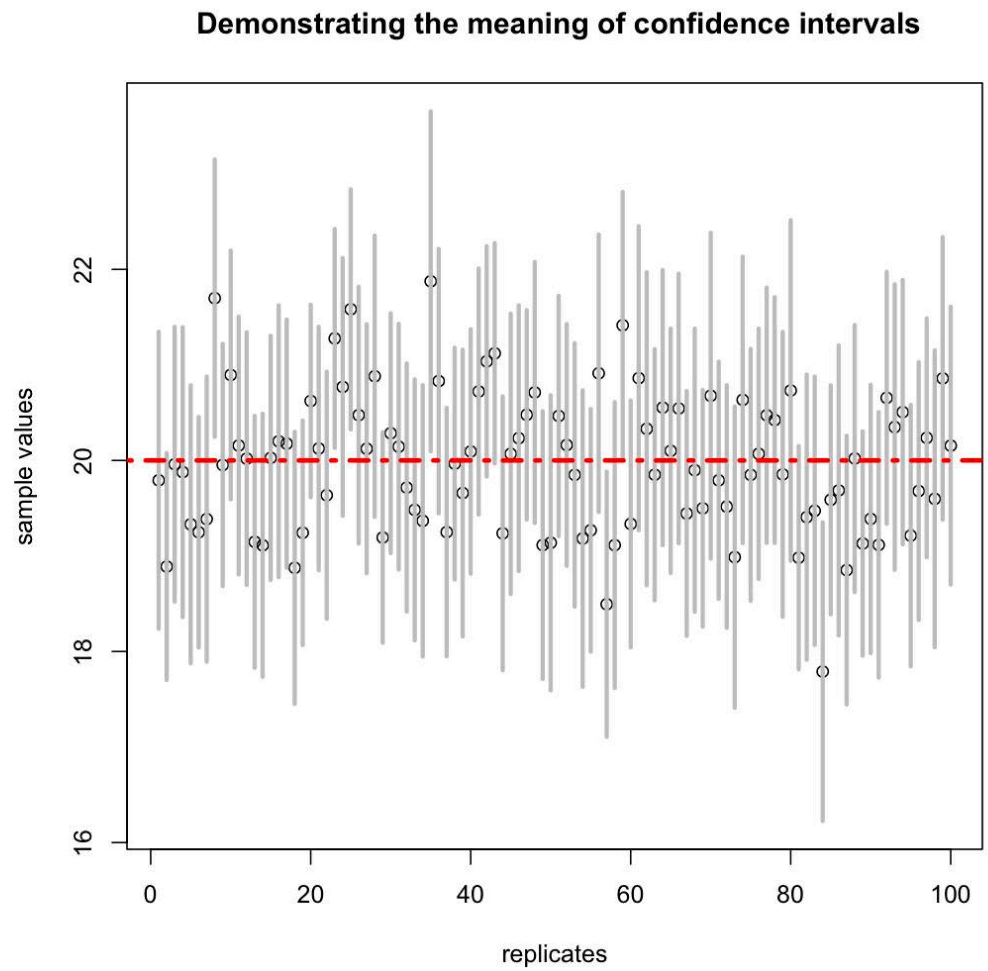
Cool! major difference between Bayesian and frequentist

Ok, we know what a confidence interval isn't. What is it?

Ok, we know what a confidence interval isn't. What is it?

- What can be said is that 95/(1-α)% of the time (upon repeated sampling), an interval calculated this way will overlap with the true value of μ.

…(95)% of the time (upon repeated sampling), an interval calculated this way will overlap with the true value of μ.



Back to R

notice the couple of bars towards the center that do not overlap with μ

Confidence intervals

- I will let you in on a little secret….
- Despite the fact that you can not interpret CIs in the manner you would like, they do approximate that interpretation pretty well…..
- In ZOL851, we will learn about some alternative approaches that do have a direct and intuitive probability interpretation.