

ENV 710: Lecture 5

inference, confidence intervals, p-values

statistical inference

confidence intervals

learning goals

- what do confidence intervals signify?
- how are confidence intervals calculated for large and small samples?
- what is the t-distribution?

stuff you
should
know

confidence intervals

- $\bar{x} = 307 \text{ Mg ha}^{-1}$
- 95% CI = [287, 327]

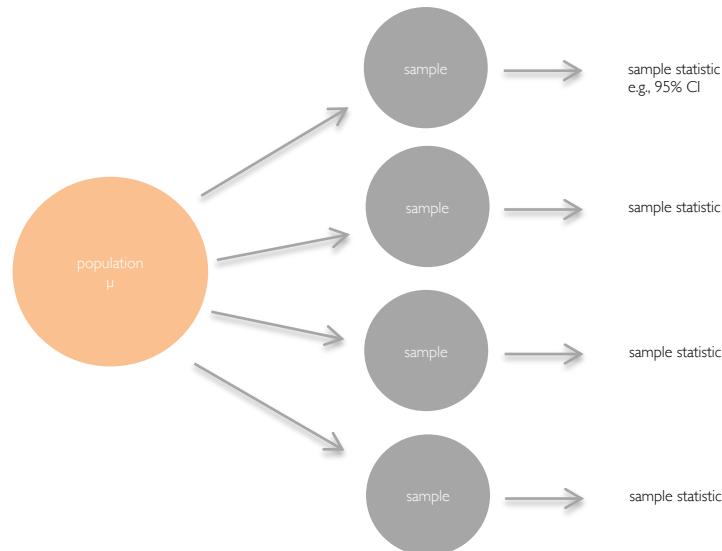
what do 95% confidence intervals signify?



confidence intervals

confidence intervals define the chances that an interval will contain the fixed value, μ

because of variability,
each sample mean, \bar{x} ,
it may be smaller or
larger than μ

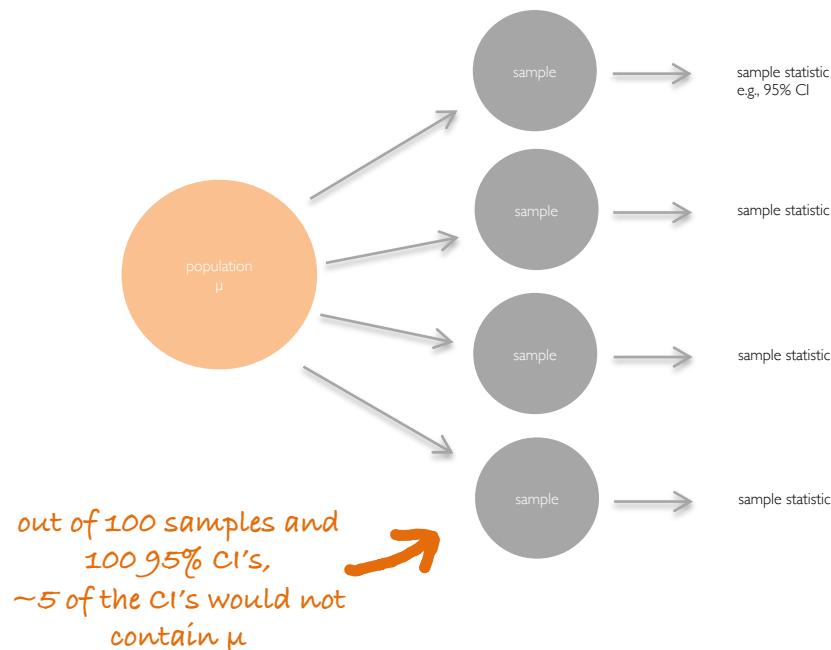


confidence intervals

confidence intervals define the chances that an interval will contain the fixed value, μ



if we were to repeatedly sample the population, we would expect the true μ would lie outside of the 95% confidence interval 5% of the time

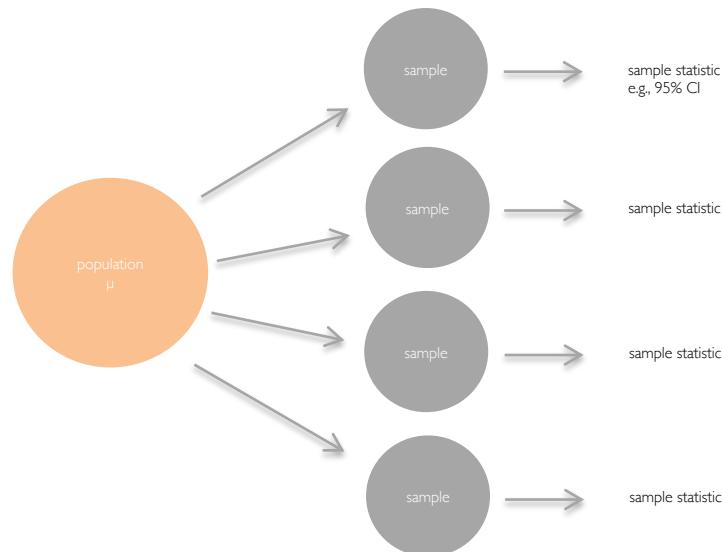


confidence intervals

confidence intervals define the chances that an interval will contain the fixed value, μ

if we were to repeatedly sample the population, we would expect the true μ would lie outside of the 95% confidence interval 5% of the time

the confidence interval varies, not the population mean



confidence interval simulation code

```
ci.sim.fun <- function(mu=mu, std=std, n=n, runs=runs) {  
  
  ci <- matrix(nrow=runs, ncol=2)  
  for (i in 1:runs){  
    samp <- rnorm(n=n, mean=mu, sd=std)  
    sx <- mean(samp)  
    conf <- qt(.975, n)  
    if(n >30){conf <- qnorm(0.975)}  
    ci[i,] <- c(sx-conf*(sd(samp)/sqrt(n)),  
    sx+conf*(sd(samp)/sqrt(n)))  
  }  
  
  par(mfrow=c(1,1))  
  plot(0,0, xlim=c(min(ci[,1])-abs(0.5*min(ci[,1])),  
  max(ci[,2])+0.5*max(ci[,2])),  
  ylim=c(0,100), type="n",  
  xlab="CI's of Samples", ylab="Number of Runs", las=1)  
  abline(v=mu, lwd=2)  
  
  cnt <- 0  
  for(i in 1:runs){  
    clr <- 1  
    if(ci[i,1]>mu)clr=2  
    if(ci[i,2]<mu)clr=2  
    if(clr == 2)cnt=cnt+1  
    segments(ci[i,1], i, ci[i,2], i, col=clr)  
  }  
  text(min(ci[,1]), runs*0.95, paste("Count = ", cnt), cex=0.9)  
}  
ci.sim.fun(mu = 0, std = 1, n = 50, runs = 100)
```

confidence intervals

sample mean \pm margin of error

$$\bar{x} \pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

α is the significance level: probability of rejecting a null hypothesis in a statistical test when it is true

Z is the statistic representing the probability of $\alpha/2$ on a standard normal distribution

```
abs(qnorm(p = 0.05/2, mean = 0, sd = 1))
```

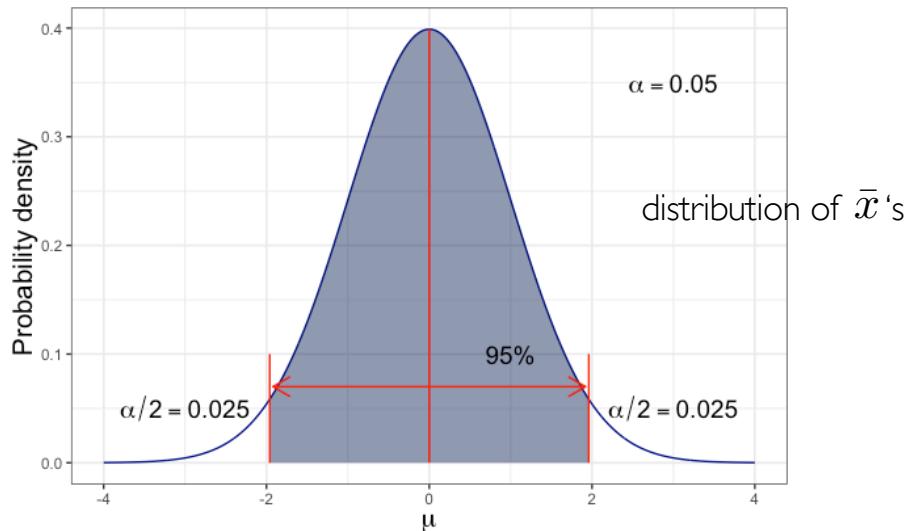
```
qnorm(p = 1-0.05/2, mean = 0, sd = 1)
```

confidence intervals

confidence
interval for a
population
mean

sample mean \pm margin of error

$$\bar{x} \pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$



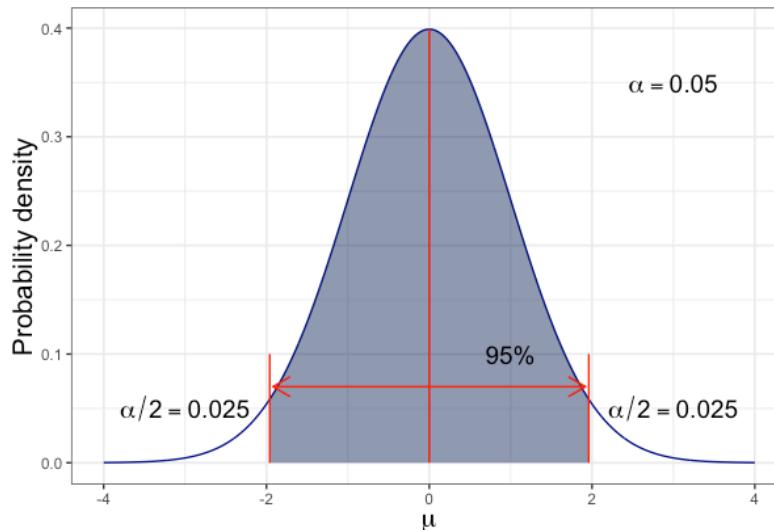
$$P(\bar{x} - 1.96 \cdot \sigma / \sqrt{n} \leq \mu \leq \bar{x} + 1.96 \cdot \sigma / \sqrt{n}) = 0.95$$

confidence intervals

confidence
interval from a
sample (no
knowledge of μ
or σ)

sample mean \pm margin of error

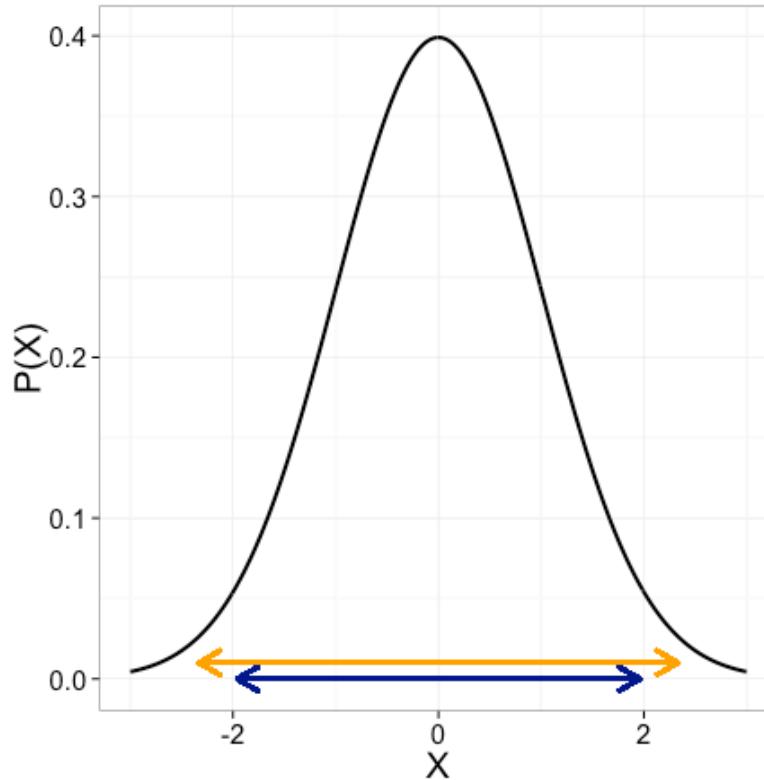
$$\bar{x} \pm Z_{\alpha/2} \frac{s}{\sqrt{n}}$$



$$P(\bar{x} - 1.96se \leq \mu \leq \bar{x} + 1.96se) = 0.95$$

confidence intervals

- samples with more variability generate wider CI's
- small sample sizes = wider CI's
- to cut your margin of error in half, need to ~ quadruple the sample size



large sample sizes ($n \geq 30$)

when the σ of the population is unknown and the sample size is large ($n \geq 30$), the standard normal distribution is used

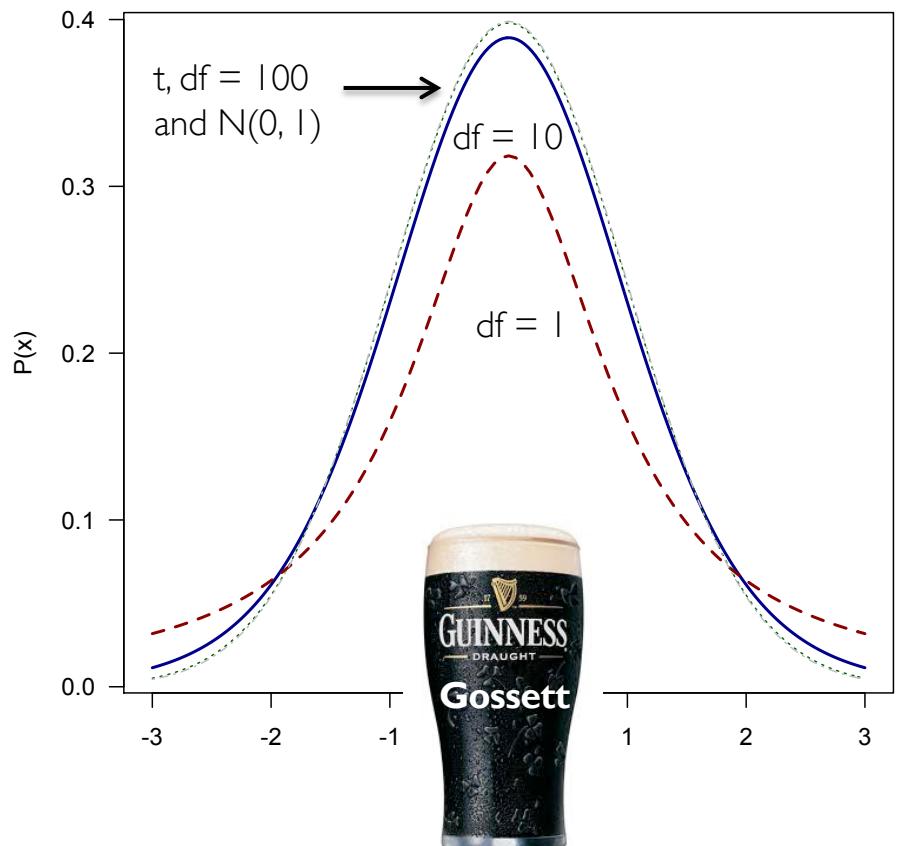
$$\bar{x} \pm Z_{\alpha/2} \frac{s}{\sqrt{n}}$$

small sample sizes ($n < 30$)

when the sample size is small ($n < 30$), Student's t distribution is used instead of the normal distribution

$$\bar{x} \pm t_{[n-1]\alpha/2} \left(\frac{s}{\sqrt{n}} \right)$$

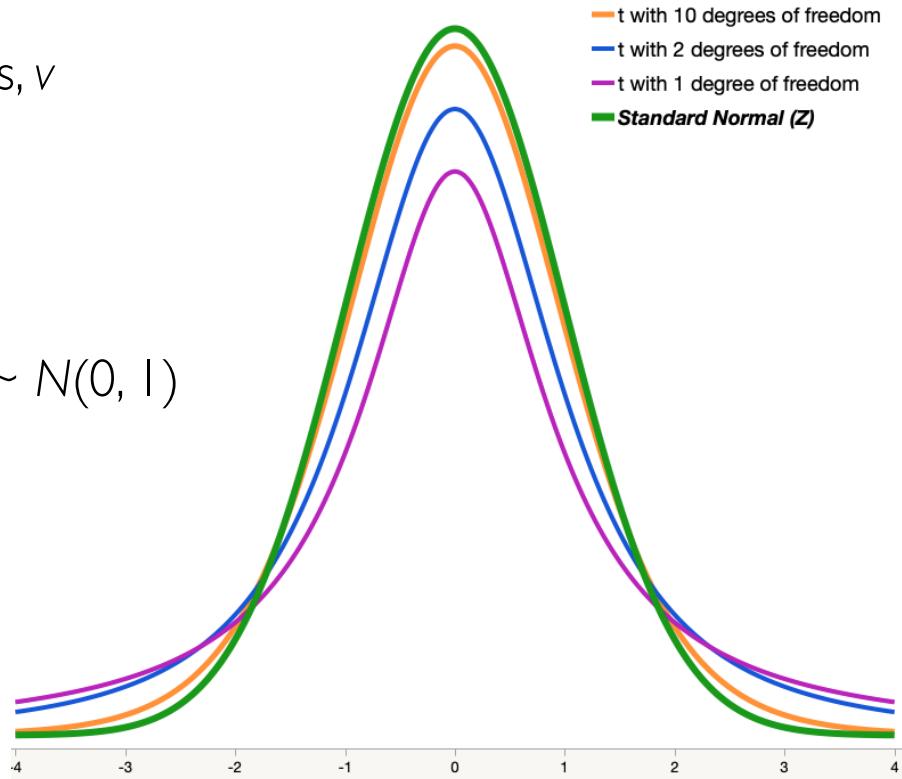




Student's t distribution widens the CI to account for the fact that we are using s (sample standard deviation) to estimate σ .

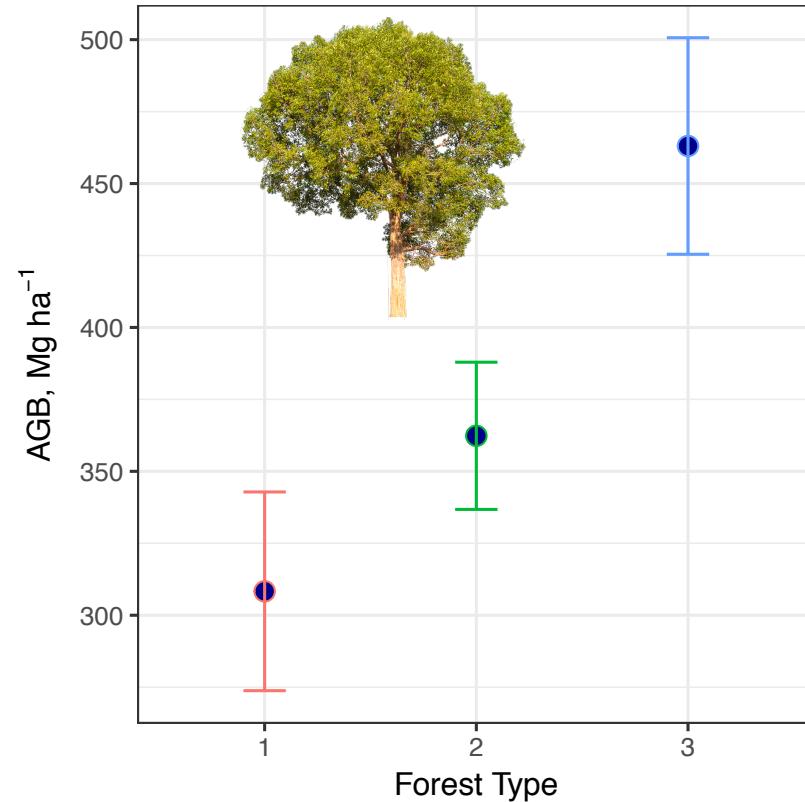
t distribution

- t distribution is determined by its df's, ν
- mean of the distribution = 0
- variance is equal to $\frac{\nu}{\nu - 2}$
- variance is always greater than 1
- with infinite df's the t distribution is $\sim N(0, 1)$

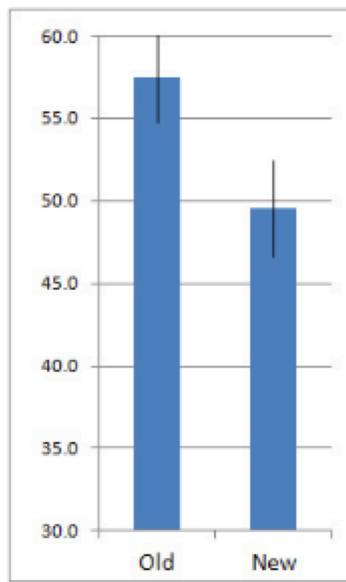


example

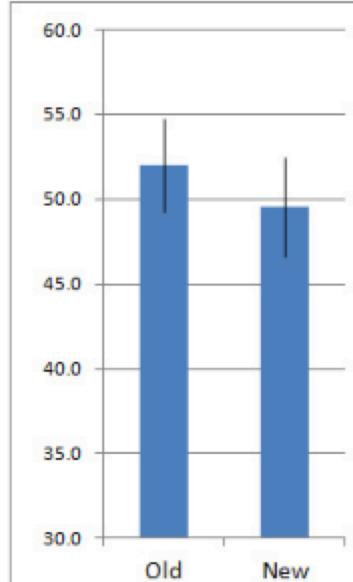
- researchers measured the biomass of 30 1-ha tree plots.
 $\bar{x} = 371 \text{ Mg ha}^{-1}$
- plots disturbed by different logging, hunting or neither
 - #1 - 10 logged, hunted plots
 - #2 - 10 logged plots
 - #3 - 10 undisturbed plots
- is mean biomass for the three types of plots the same?



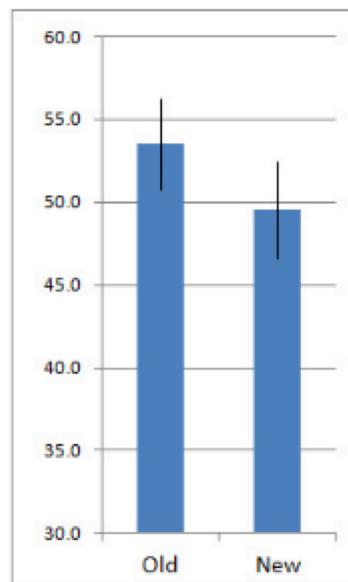
CI's & statistical significance



No Overlap: Significantly different ($p < .05$)



Lots of Overlap: Not Significantly different ($p > .05$)



Some Overlap: Can't tell, run the 2-sample t-test ($p = .04$, go publish!)

statistical inference

**hypothesis testing and
p-values**

learning goals

- how does hypothesis testing relate to p-values?
 - null vs. alternative hypothesis
- what are p-values and how do we use them?
 - calculate p-values
 - one-sided vs. two-sided tests.

stuff you
should
know

hypothesis testing

null - (H_0): usually no difference between the population mean and a given value (=)

alternative - (H_a): a difference between the population mean and a given value (<>)

- abandon H_0 only if the evidence in favor of H_a is so strong as to **reject** H_0 in favor of H_a
- statistical test reveals whether pattern is likely or unlikely, **given that the H_0 is true**
- p-value is guide to rejecting the H_0

null hypothesis

H_0 : Defendant is innocent

alternative hypothesis

H_A : Defendant is guilty

present the evidence

collect data

burden
of proof

judge the evidence

“Could these data plausibly have happened by chance if the null hypothesis were true?”

yes

Fail to reject H_0

no

Reject H_0

p-value



$P(\text{observed or more extreme outcome} | H_0 \text{ true})$

use test statistic to calculate the p-value...

the probability of viewing data at least as favorable to the alternative hypothesis as our data, if the null hypothesis was true

p-value

$P(\text{observed or more extreme outcome} | H_0 \text{ true})$

use test statistic to calculate the p-value, the probability of viewing data at least as favorable to the alternative hypothesis as our data, if the null hypothesis was true

- if the p-value is low ($< \alpha$), we are very unlikely to observe the data if the null hypothesis were true → reject H_0
- if the p-value is high ($> \alpha$), we are likely to observe the data if the null hypothesis were true → do not reject H_0

α is the significance level

if the p-value is low, the null hypothesis must go!

p-value depends on...

- number of observations in the sample(s), N
- difference between the means of the samples

$$\bar{X}_i - \bar{X}_j$$

- degree of variation among individuals (s^2)

two-tailed test

testing for possibility of relationship in both directions

- allots half of α to test statistical significance in one direction, and half to the other

H_0 : mean = y

H_A : tests both whether the mean is significantly greater than y or significantly smaller than y

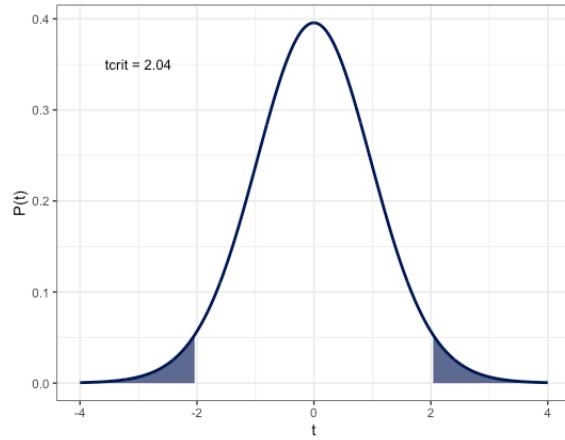
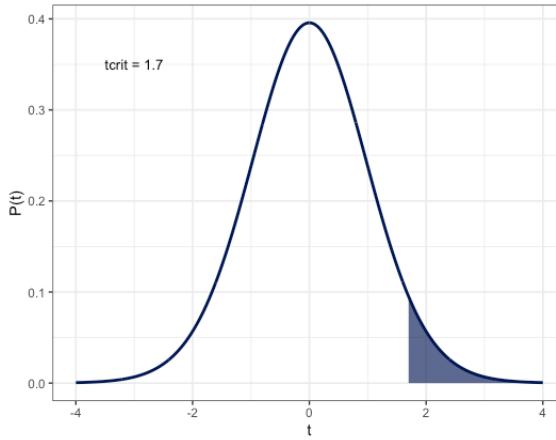
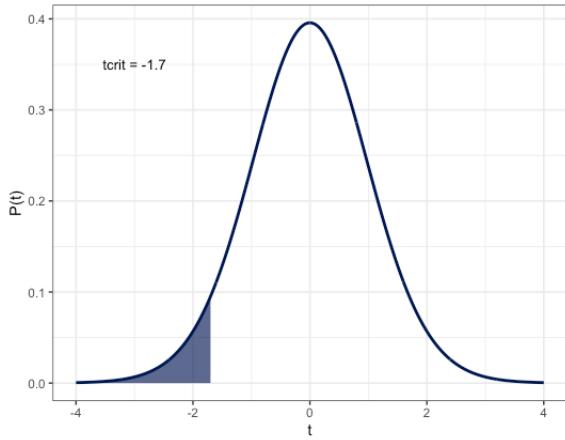
one-tailed test

testing for possibility of relationship in one direction

- allots α to test statistical significance in one tail or the other

one-tailed test provides you with more power – probability of detecting an effect, given the effect exists

one and two-tailed tests



critical value - threshold to which the value of the test statistic in a sample is compared to determine whether the null hypothesis is rejected

critical region or rejection region - set of values of the test statistic for which the null hypothesis is rejected in a hypothesis test

Post your questions to be
answered during lecture