R version 3.3.0 (2016-05-03) -- "Supposedly Educational"

Copyright (C) 2016 The R Foundation for Statistical Computing

Platform: x86\_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.

You are welcome to redistribute it under certain conditions.

Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.

Type 'contributors()' for more information and

'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or

'help.start()' for an HTML browser interface to help.

Type 'q()' to quit R.

[Previously saved workspace restored]

> library(swirl)

| Hi! I see that you have some variables saved in your workspace. To keep

| things running smoothly, I recommend you clean up before starting swirl.

| Type ls() to see a list of the variables in your workspace. Then, type

| rm(list=ls()) to clear your workspace.

| Type swirl() when you are ready to begin.

Warning message:

package ‘swirl’ was built under R version 3.3.1

> swirl()

| Welcome to swirl! Please sign in. If you've been here before, use the same

| name as you did then. If you are new, call yourself something unique.

What shall I call you? SY

| Please choose a course, or type 0 to exit swirl.

1: Statistical Inference

2: Take me to the swirl course repository!

Selection: 1

| Please choose a lesson, or type 0 to return to course menu.

1: Introduction 2: Probability1

3: Probability2 4: ConditionalProbability

5: Expectations 6: Variance

7: CommonDistros 8: Asymptotics

9: T Confidence Intervals 10: Hypothesis Testing

11: P Values 12: Power

13: Multiple Testing 14: Resampling

Selection: 12

| Attempting to load lesson dependencies...

| Package ‘reshape2’ loaded correctly!

| Package ‘ggplot2’ loaded correctly!

| | | 0%

| Power. (Slides for this and other Data Science courses may be found at github

| https://github.com/DataScienceSpecialization/courses/. If you care to use

| them, they must be downloaded as a zip file and viewed locally. This lesson

| corresponds to 06\_Statistical\_Inference/11\_Power.)

...

| |= | 1%

| In this lesson, as the name suggests, we'll discuss POWER, which is the probability of

| rejecting the null hypothesis when it is false, which is good and proper.

...

| |== | 2%

| Hence you want more POWER.

...

| |== | 3%

| Power comes into play when you're designing an experiment, and in particular, if you're

| trying to determine if a null result (failing to reject a null hypothesis) is

| meaningful. For instance, you might have to determine if your sample size was big

| enough to yield a meaningful, rather than random, result.

...

| |=== | 4%

| Power gives you the opportunity to detect if your ALTERNATIVE hypothesis is true.

...

| |==== | 5%

| Do you recall the definition of a Type II error? Remember, errors are bad.

1: Miscalculating a t score

2: Misspelling the word hypothesis

3: Rejecting a true null hypothesis

4: Accepting a false null hypothesis

Selection: 4

| All that practice is paying off!

| |===== | 7%

| Beta is the probability of a Type II error, accepting a false null hypothesis; the

| complement of this is obviously (1 - beta) which represents the probability of

| rejecting a false null hypothesis. This is good and this is POWER!

...

| |===== | 8%

| Recall our previous example involving the Respiratory Distress Index and sleep

| disturbances. Our null hypothesis H\_0 was that mu = 30 and our alternative hypothesis

| H\_a was that mu > 30.

...

| |====== | 9%

| Which of the following expressions represents our test statistic under this null

| hypothesis? Here X' represents the sample mean, s is the sample std deviation, and n is

| the sample size. Assume X' follows a t distribution.

1: (X'-30)/(s^2/n)

2: X'/(s^2/n)

3: (X'-30)/(s/sqrt(n))

4: 30/(s/sqrt(n))

Selection: 3

| You are doing so well!

| |======= | 10%

| In the expression for the test statistic (X'-30)/(s/sqrt(n)) what does (s/sqrt(n))

| represent?

1: a standard variance

2: a standard sample

3: a standard bearer

4: a standard error

5: a standard measure

Selection: 4

| Excellent work!

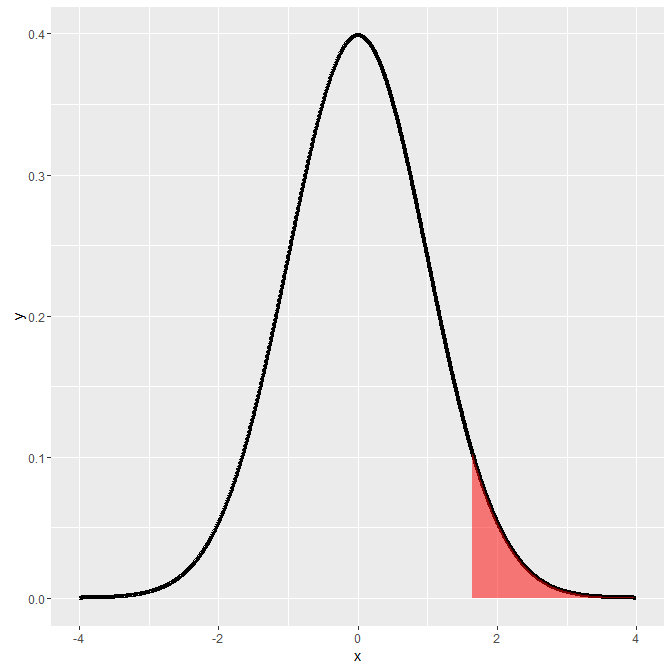
| |======== | 11%

| Suppose we're testing a null hypothesis H\_0 with an alpha level of .05. Since H\_a

| proposes that mu > 30 (the mean hypothesized by H\_0), power is the probability that the

| true mean mu is greater than the (1-alpha) quantile or qnorm(.95). For simplicity,

| assume we're working with normal distributions of which we know the variances.

... 

| |======== | 12%

| Here's the picture we've used a lot in these lessons. As you know, the shaded portion

| represents 5% of the area under the curve. If a test statistic fell in this shaded

| portion we would reject H\_0 because the sample mean is too far from the mean (center)

| of the distribution hypothesized by H\_0. Instead we would favor H\_a, that mu > 30. This

| happens with probability .05.

...

| |========= | 13%

| You might well ask, "What does this have to do with POWER?" Good question. We'll look

| at some pictures to show you.

...

| |========== | 14%

| First we have to emphasize a key point. The two hypotheses, H\_0 and H\_a, actually

| represent two distributions since they're talking about means or centers of

| distributions. H\_0 says that the mean is mu\_0 (30 in our example) and H\_a says that the

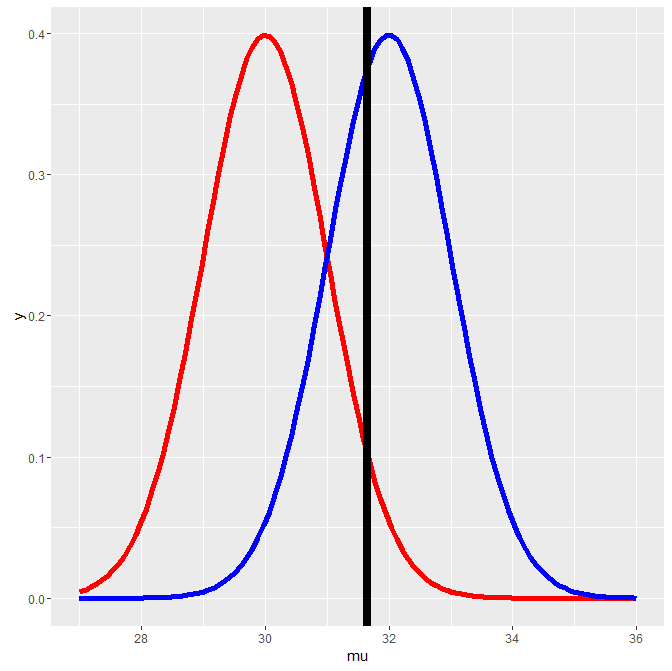
| mean is mu\_a.

...

| |=========== | 15%

| We're assuming normality and equal variance, say sigma^2/n, for both hypotheses, so

| under H\_0, X'~ N(mu\_0, sigma^2/n) and under H\_a, X'~ N(mu\_a, sigma^2/n).

... 

| |=========== | 16%

| Here's a picture with the two distributions. We've drawn a vertical line at our

| favorite spot, at the 95th percentile of the red distribution. To the right of the line

| lies 5% of the red distribution.

...

| |============ | 17%

| Quick quiz! Which distribution represents H\_0?

1: the blue

2: the red

Selection: 2

| Perseverance, that's the answer.

| |============= | 18%

| Which distribution represents H\_a?

1: the blue

2: the red

Selection: 1

| Great job!

| |============== | 20%

| From the picture, what is the mean proposed by H\_a?

1: 30

2: 28

3: 32

4: 36

Selection: 3

| All that hard work is paying off!

| |============== | 21%

| See how much of the blue distribution lies to the right of that big vertical line?

...

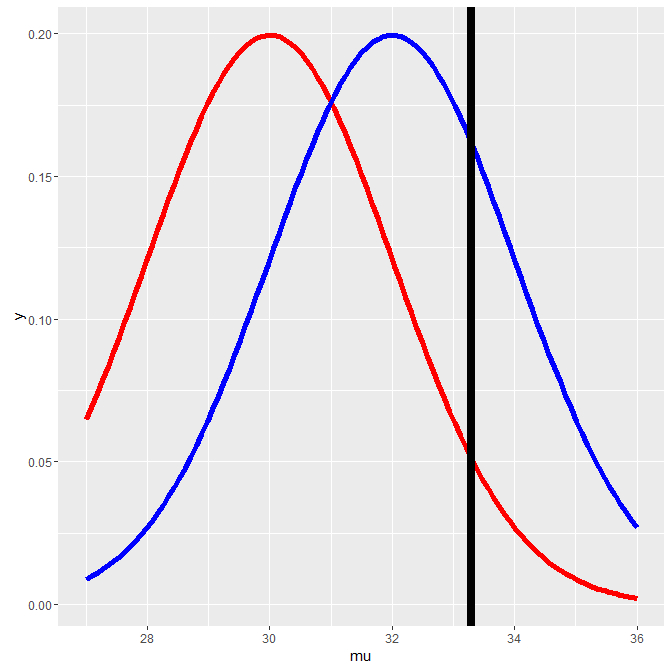
| |=============== | 22%

| That, my friend, is POWER!

...

| |================ | 23%

| It's the area under the blue curve (H\_a) to the right of the vertical line.

... 

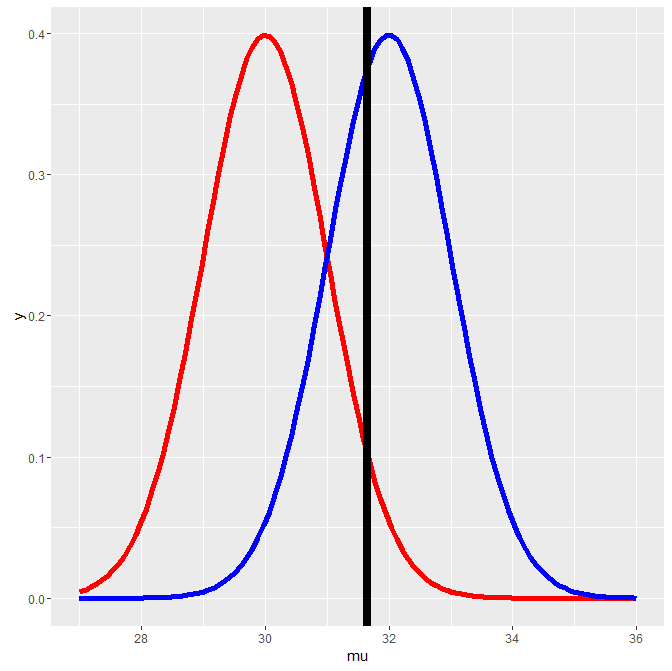
| |================= | 24%

| Note that the placement of the vertical line depends on the null distribution. Here's

| another picture with fatter distributions. The vertical line is still at the 95th

| percentile of the null (red) distribution and 5% of the distribution still lies to its

| right. The line is calibrated to mu\_0 and the variance.



...

| |================== | 25%

| Back to our original picture.

...

| |================== | 26%

| We've shamelessly stolen plotting code from the slides so you can see H\_a in action.

| Let's look at pictures before we delve into numbers. We've fixed mu\_0 at 30, sigma

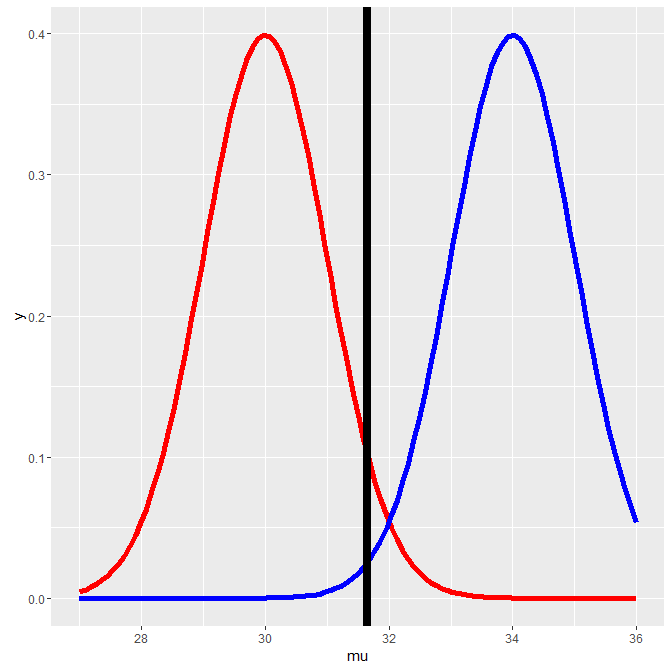
| (standard deviation) at 4 and n (sample size) at 16. The function myplot just needs an

| alternative mean, mu\_a, as argument. Run myplot now with an argument of 34 to see what

| it does.

> myplot(34)

| Keep working like that and you'll get there!



| |=================== | 27%

| The distribution represented by H\_a moved to the right, so almost all (100%) of the

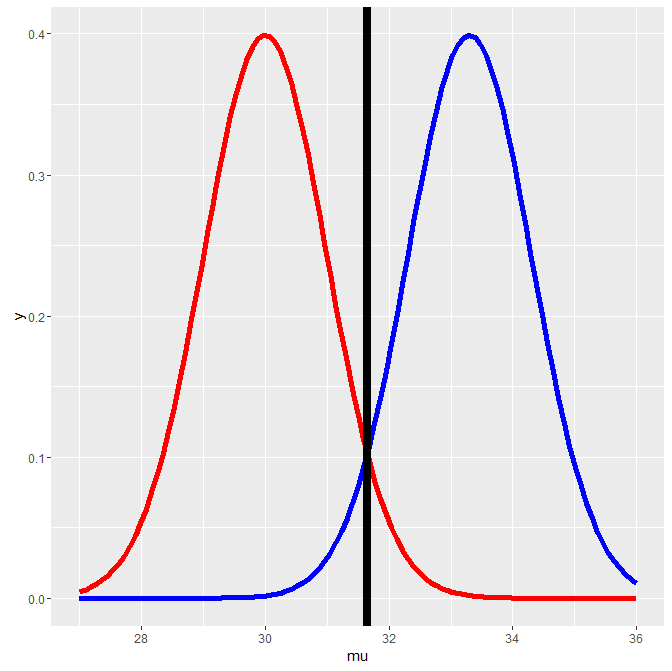
| blue curve is to the right of the vertical line, indicating that with mu\_a=34, the test

| is more powerful, i.e., there's a higher probability that it's correct to reject the

| null hypothesis since it appears false. Now try myplot with an argument of 33.3.

> myplot(33.3)

| Excellent work!



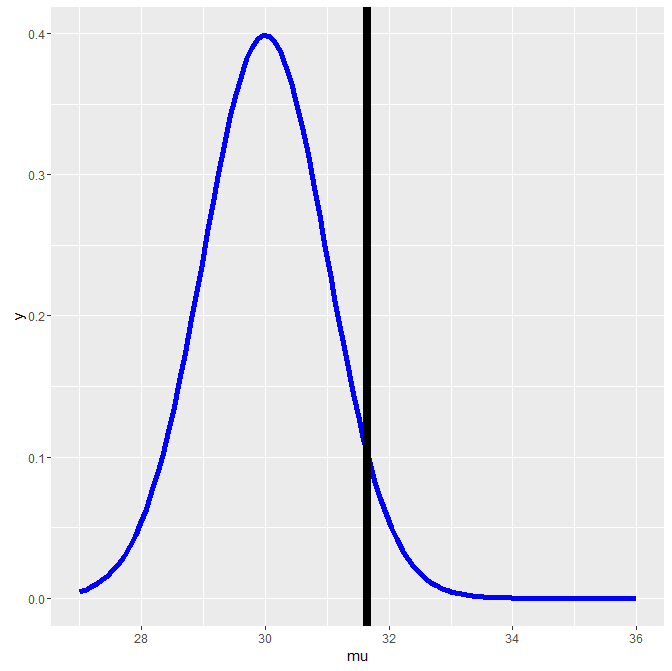
| |==================== | 28%

| This isn't as powerful as the test with mu\_a=34 but it makes a pretty picture. Now try

| myplot with an argument of 30.

> myplot(30)

| You are amazing!



| |===================== | 29%

| Uh Oh! Did the red curve disappear? No. it's just under the blue curve. The power now,

| the area under the blue curve to the right of the line, is exactly 5% or alpha!

...

| |===================== | 30%

| So what did we learn?

...

| |====================== | 32%

| First, power is a function that depends on a specific value of an alternative mean,

| mu\_a, which is any value greater than mu\_0, the mean hypothesized by H\_0. (Recall that

| H\_a specified mu>30.)

...

| |======================= | 33%

| Second, if mu\_a is much bigger than mu\_0=30 then the power (probability) is bigger than

| if mu\_a is close to 30. As mu\_a approaches 30, the mean under H\_0, the power approaches

| alpha.

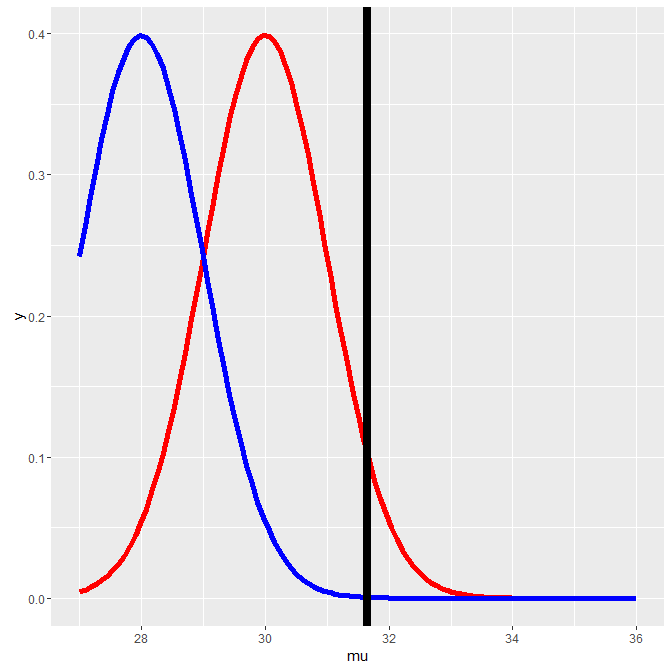
...

| |======================== | 34%

| Just for fun try myplot with an argument of 28.

> myplot(28)

| You are amazing!



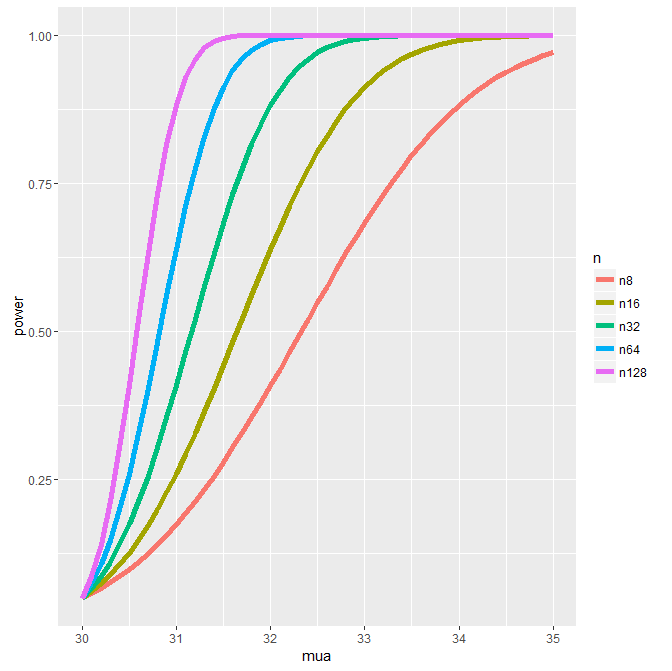
| |======================== | 35%

| We see that the blue curve has moved to the left of the red, so the area under it, to

| the right of the line, is less than the 5% under the red curve. This then is even less

| powerful and contradicts H\_a so it's not worth looking at.

...

 | |========================= | 36%

| Here's a picture of the power curves for different sample sizes. Again, this uses code

| "borrowed" from the slides. The alternative means, the mu\_a's, are plotted along the

| horizontal axis and power along the vertical.

...

| |========================== | 37%

| What does the graph show us about mu\_a?

1: power is independent of mu\_a

2: as it gets bigger, it gets less powerful

3: as it gets bigger, it gets more powerful

Selection: 3

| Excellent job!

| |=========================== | 38%

| What does the graph show us about sample size?

1: as it gets bigger, it gets more powerful

2: power is independent of sample size

3: as it gets bigger, it gets less powerful

Selection: 1

| Excellent job!

| |=========================== | 39%

| Now back to numbers. Our test for determining rejection of H\_0 involved comparing a

| test statistic, namely Z=(X'-30)/(sigma/sqrt(n)), against some quantile, say Z\_95,

| which depended on our level size alpha (.05 in this case). H\_a proposed that mu > mu\_0,

| so we tested if Z>Z\_95. This is equivalent to X' > Z\_95 \* (sigma/sqrt(n)) + 30, right?

...

| |============================ | 40%

| Recall that nifty R function pnorm, which gives us the probability that a value drawn

| from a normal distribution is greater or less than/equal to a specified quantile

| argument depending on the flag lower.tail. The function also takes a mean and standard

| deviation as arguments.

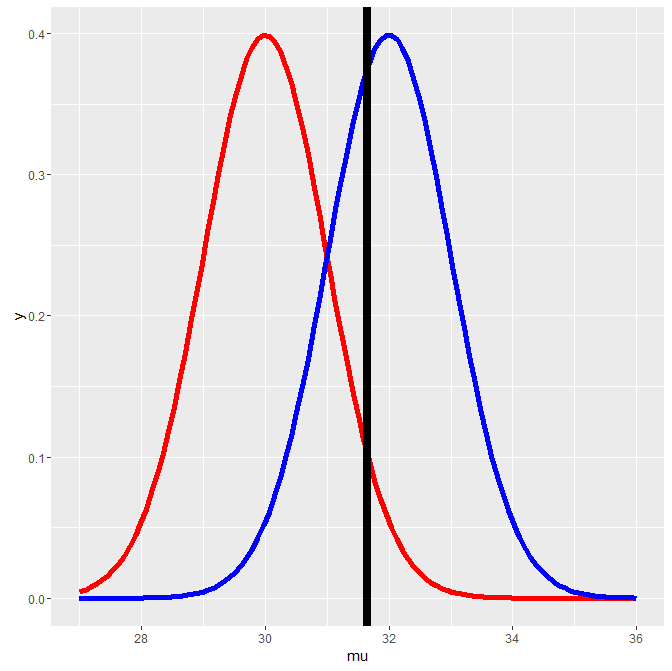
...

| |============================= | 41%

| Suppose we call pnorm with the quantile 30 + Z\_95 \* (sigma/sqrt(n)) and specify mu\_a as

| our mean argument. This would return a probability which we can interpret as POWER.

| Why?



...

| |============================== | 42%

| Recall our picture of two distributions. 30 + Z\_95 \* (sigma/sqrt(n)) represents the

| point at which our vertical line falls. It's the point on the null distribution at the

| (1-alpha) level.

...

| |============================== | 43%

| Study this picture. Calling pnorm with 30 + Z\_95 \* (sigma/sqrt(n)) as the quantile and

| mu\_a, say 32, as the mean and lower.tail=FALSE does what?

1: returns the area under the blue curve to the left of the line

2: returns the area under the blue curve to the right of the line

3: returns the area under the red curve to the right of the line

4: returns the area under the red curve to the left of the line

Selection: 2

| You got it!

| |=============================== | 45%

| Let's try some examples now. Before we do, what do we know pnorm will return if we

| specify a quantile less than the mean?

1: an answer less than .50

2: an answer dependent on beta

3: an answer dependent on alpha

4: an answer greater than 1

Selection: 1

| Keep up the great work!

| |================================ | 46%

| First, define a variable z as qnorm(.95)

> z <- qnorm(.95)

| Keep working like that and you'll get there!

| |================================= | 47%

| Run pnorm now with the quantile 30+z, mean=30, and lower.tail=FALSE. We've specified

| sigma and n so that the standard deviation of the sample mean is 1.

> pnorm(30+z,mean=30, lower.tail=FALSE)

[1] 0.05

| That's a job well done!

| |================================= | 48%

| That's not surprising, is it? With the mean set to mu\_0 the two distributions, null and

| alternative, are the same and power=alpha. Now run pnorm now with the quantile 30+z,

| mean=32, and lower.tail=FALSE.

> pnorm(30+z,mean=32, lower.tail=FALSE)

[1] 0.63876

| All that hard work is paying off!

| |================================== | 49%

| See how this is much more powerful? 64% as opposed to 5%. When the sample mean is quite

| different from (many standard errors greater than) the mean hypothesized by the null

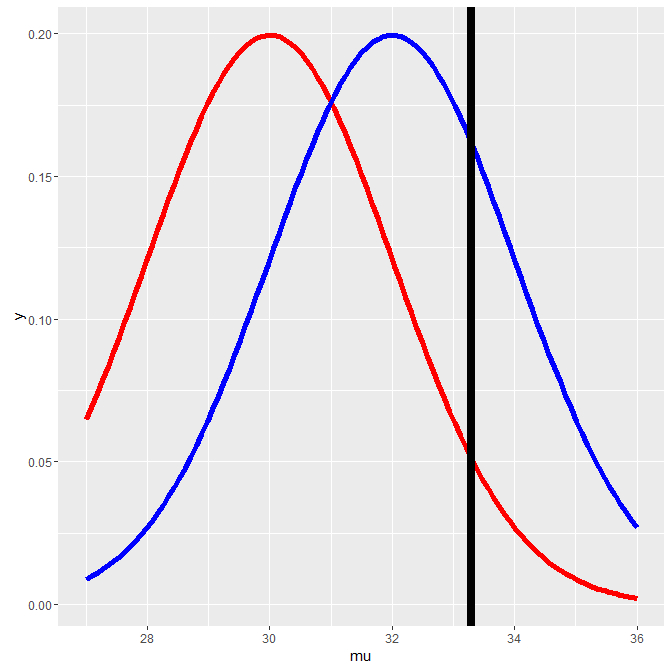
| hypothesis, the probability of rejecting H\_0 when it is false is much higher. That is

| power!

...

| |=================================== | 50%

| Let's look again at the portly distributions.



...

| |==================================== | 51%

| With this standard deviation=2 (fatter distribution) will power be greater or less than

| with the standard deviation=1?

1: the same

2: less than

3: greater

Selection: 2

| All that practice is paying off!

| |===================================== | 52%

| To see this, run pnorm now with the quantile 30+z, mean=32 and sd=1. Don't forget to

| set lower.tail=FALSE so you get the right tail.

> pnorm(30+z,mean=32, sd=1, lower.tail=FALSE)

[1] 0.63876

| You are really on a roll!

| |===================================== | 53%

| Now run pnorm now with the quantile 30+z\*2, mean=32 and sd=2. Don't forget to set

| lower.tail=FALSE so you get the right tail.

> pnorm(30+z\*2,mean=32, sd=2, lower.tail=FALSE)

[1] 0.259511

| You are really on a roll!

| |====================================== | 54%

| See the power drain from 64% to 26% ? Let's review some basic facts about power. We saw

| before in our pictures that the power of the test depends on mu\_a. When H\_a specifies

| that mu > mu\_0, then as mu\_a grows and exceeds mu\_0 increasingly, what happens to

| power?

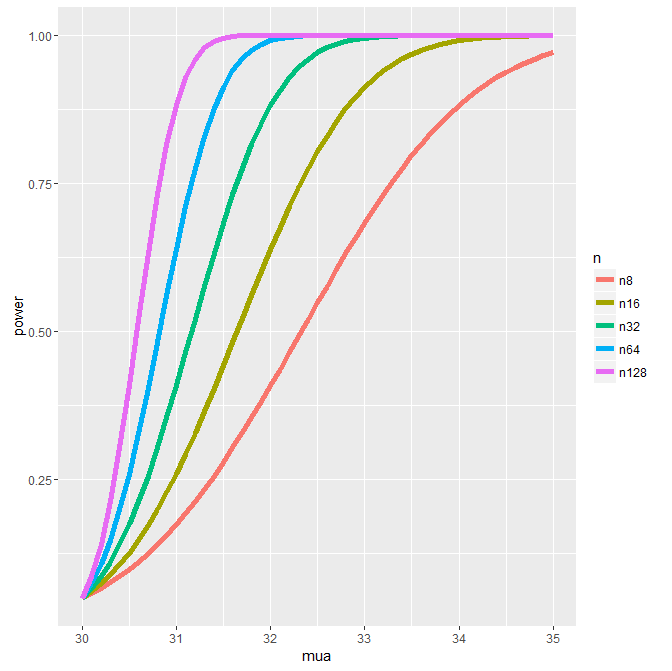
1: it increases

2: it doesn't change

3: it decreases

Selection: 1

| Your dedication is inspiring!



| |======================================= | 55%

| Here's another question. Recall our power curves from before.

...

| |======================================== | 57%

| As the sample size increases, what happens to power?

1: it increases

2: it decreases

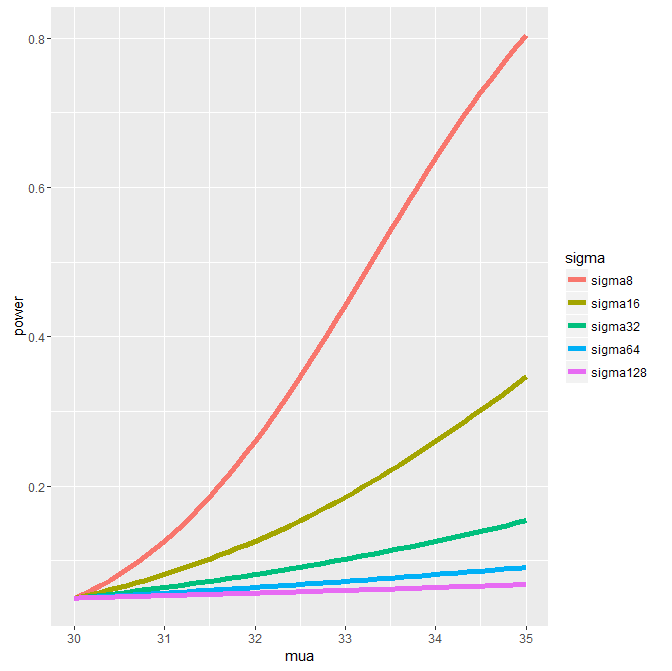
3: it doesn't change

Selection: 1

| Keep up the great work!

| |======================================== | 58%

| Here's another one. More power curves.



...

| |========================================= | 59%

| As variance increases, what happens to power?

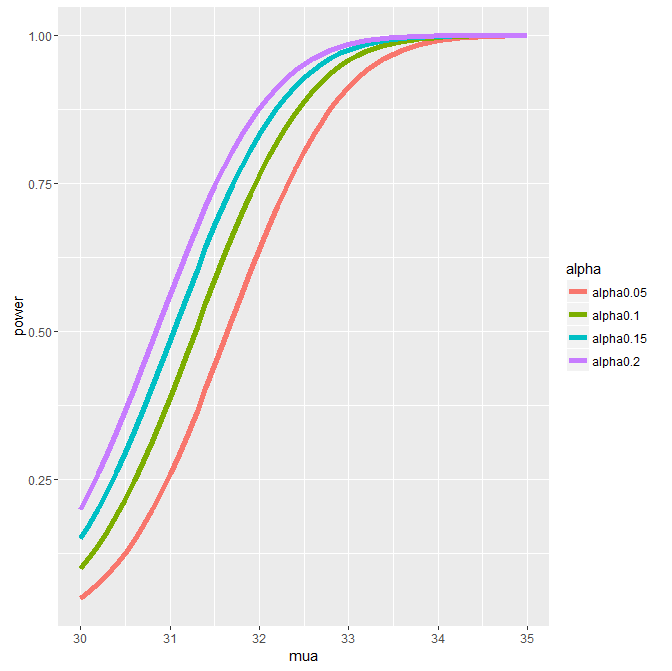
1: it doesn't change

2: it increases

3: it decreases

Selection: 3

| Great job!



| |========================================== | 60%

| Here's another one. And even more power curves.

...

| |=========================================== | 61%

| As alpha increases, what happens to power?

1: it decreases

2: it doesn't change

3: it increases

Selection: 3

| Excellent work!

| |=========================================== | 62%

| If H\_a proposed that mu != mu\_0 we would calculate the one sided power using alpha / 2

| in the direction of mu\_a (either less than or greater than mu\_0). (This is only

| approximately right, it excludes the probability of getting a large test statistic in

| the opposite direction of the truth.

...

| |============================================ | 63%

| Since power goes up as alpha gets larger would the power of a one-sided test be greater

| or less than the power of the associated two sided test?

1: they're the same

2: greater

3: less than

Selection: 2

| You got it right!

| |============================================= | 64%

| Finally, if H\_a specified that mu < mu\_0 could we still do the same kind of power

| calculations?

1: Yes

2: No

Selection: 1

| Excellent work!

| |============================================== | 65%

| Suppose H\_a says that mu > mu\_0. Then power = 1 - beta = Prob ( X' > mu\_0 + z\_(1-alpha)

| \* sigma/sqrt(n)) assuming that X'~ N(mu\_a,sigma^2/n). Which quantities do we know in

| this statement, given the context of the problem? Let's work through this.

...

| |============================================== | 66%

| What does the null hypothesis H\_0 tell us that the population mean equals?

1: mu\_a

2: beta

3: mu\_0

4: alpha

Selection: 3

| You are amazing!

| |=============================================== | 67%

| After the null mean mu\_0 is proposed what does the designer of the hypothesis test

| specify in order to reject or fail-to-reject H\_0? In other words, what is the level

| size of the test?

1: mu\_0

2: mu\_a

3: alpha

4: beta

Selection: 3

| Nice work!

| |================================================ | 68%

| So we know that the quantities mu\_0 and alpha are specified by the test designer. In

| the statement 1 - beta = Prob( X' > mu\_0 + z\_(1-alpha) \* sigma/sqrt(n)) given mu\_a >

| mu\_0, mu\_0 and alpha are specified, and X' depends on the data. The other four

| quantities, (beta, sigma, n, and mu\_a), are all unknown.

...

| |================================================= | 70%

| It should be obvious that specifying any three of these unknowns will allow us to solve

| for the missing fourth. Usually, you only try to solve for power (1-beta) or the sample

| size n.

...

| |================================================= | 71%

| An interesting point is that power doesn't need mu\_a, sigma and n individually.

| Instead only sqrt(n)\*(mu\_a - mu\_0) /sigma is needed. The quantity (mu\_a - mu\_0) / sigma

| is called the EFFECT SIZE. This is the difference in the means in standard deviation

| units. It is unit free so it can be interpreted in different settings.

...

| |================================================== | 72%

| We'll work through some examples of this now. However, instead of assuming that we're

| working with normal distributions let's work with t distributions. Remember, they're

| pretty close to normal with large enough sample sizes.

...

| |=================================================== | 73%

| Power is still a probability, namely P( (X' - mu\_0)/(S /sqrt(n)) > t\_(1-alpha, n-1)

| given H\_a that mu > mu\_a ). Notice we use the t quantile instead of the z. Also, since

| the proposed distribution is not centered at mu\_0, we have to use the non-central t

| distribution.

...

| |==================================================== | 74%

| R comes to the rescue again with the function power.t.test. We can omit one of the

| arguments and the function solves for it. Let's first use it to solve for power.

...

| |==================================================== | 75%

| We'll run it three times with the same values for n (16) and alpha (.05) but different

| delta and standard deviation values. We'll show that if delta (difference in means)

| divided by the standard deviation is the same, the power returned will also be the

| same. In other words, the effect size is constant for all three of our tests.

...

| |===================================================== | 76%

| We'll specify a positive delta; this tells power.t.test that H\_a proposes that mu >

| mu\_0 and so we'll need a one-sided test. First run power.t.test(n = 16, delta = 2 / 4,

| sd=1, type = "one.sample", alt = "one.sided")$power .

> power.t.test(n = 16, delta = 2 / 4,sd=1, type = "one.sample", alt = "one.sided")$power

[1] 0.6040329

| You nailed it! Good job!

| |====================================================== | 77%

| Now change delta to 2 and sd to 4. Keep everything else the same.

> power.t.test(n = 16, delta = 2,sd=4, type = "one.sample", alt = "one.sided")$power

[1] 0.6040329

| Excellent work!

| |======================================================= | 78%

| Same answer, right? Now change delta to 100 and sd to 200. Keep everything else the

| same.

> power.t.test(n = 16, delta = 100,sd=200, type = "one.sample", alt = "one.sided")$power

[1] 0.6040329

| You nailed it! Good job!

| |======================================================== | 79%

| So keeping the effect size (the ratio delta/sd) constant preserved the power. Let's try

| a similar experiment except now we'll specify a power we want and solve for the sample

| size n.

...

| |======================================================== | 80%

| First run power.t.test(power = .8, delta = 2 / 4, sd=1, type = "one.sample", alt =

| "one.sided")$n .

> power.t.test(power = .8, delta = 2 / 4, sd=1, type = "one.sample", alt ="one.sided")$n

[1] 26.13751

| You got it right!

| |========================================================= | 82%

| Now change delta to 2 and sd to 4. Keep everything else the same.

> power.t.test(power = .8, delta = 2, sd=4, type = "one.sample", alt ="one.sided")$n

[1] 26.13751

| Excellent work!

| |========================================================== | 83%

| Same answer, right? Now change delta to 100 and sd to 200. Keep everything else the

| same.

> power.t.test(power = .8, delta = 100, sd=200, type = "one.sample", alt ="one.sided")$n

[1] 26.13751

| Nice work!

| |=========================================================== | 84%

| Now use power.t.test to find delta for a power=.8 and n=26 and sd=1

> power.t.test(power = .8, n= 26, sd=1, type = "one.sample", alt ="one.sided")$delta

[1] 0.5013986

| You're the best!

| |=========================================================== | 85%

| Not a surprising result, is it? It told you before that with an effect size of .5 and

| power .8, you need a sample size a little more than 26. Now run it with n=27.

> power.t.test(power = .8, n= 27, sd=1, type = "one.sample", alt ="one.sided")$delta

[1] 0.4914855

| You are quite good my friend!

| |============================================================ | 86%

| What do you think will happen if you doubled sd to 2 and ran the same test?

1: delta will halve

2: delta won't change

3: delta will double

Selection: 3

| All that hard work is paying off!

| |============================================================= | 87%

| Now for a quick review. We call this the power.u.test since it comes after the

| power.t.test. LOL.

...

| |============================================================== | 88%

| 1. The level of a test is specified by what?

1: delta

2: alpha

3: beta

4: gamma

5: None of the others

Selection: 2

| Your dedication is inspiring!

| |============================================================== | 89%

| 2. What is a Type II error?

1: rejecting a false hypothesis

2: rejecting a true hypothesis

3: accepting a false hypothesis

4: accepting a true hypothesis

Selection: 3

| Perseverance, that's the answer.

| |=============================================================== | 90%

| 3. What is power?

1: thrilling

2: None of the others

3: gamma

4: beta

5: delta

6: alpha

Selection: 2

| Your dedication is inspiring!

| |================================================================ | 91%

| 4. You're a perfectionist designing an experiment and you want both alpha and beta to

| be small. Can they both be 0 for this single test?

1: Yes

2: No

Selection: 2

| All that practice is paying off!

| |================================================================= | 92%

| 5. Suppose H\_0 proposes mu = mu\_0 and H\_a proposes that mu < mu\_0. You'll test a series

| of mu\_a with power != alpha. Which of the following is NOT true?

1: mu\_a-mu\_0 < 0

2: mu\_a-mu\_0=0

3: huh?

4: mu\_0-mu\_a > 0

Selection: 2

| You got it right!

| |================================================================= | 93%

| 6. Suppose H\_0 proposes mu = mu\_0 and H\_a proposes that mu < mu\_0. Which of the

| following is true?

1: mu\_0=mu\_a maximizes the power

2: the smaller mu\_a-mu\_0 the more powerful the test

3: the smaller mu\_0-mu\_a the more powerful the test

Selection: 2

| That's correct!

| |================================================================== | 95%

| 7. Which expression represents the size effect?

1: (mu\_a - mu\_0) / sqrt(n)

2: (mu\_a - mu\_0) / sqrt(sigma)

3: (mu\_a - mu\_0) / sigma

4: (mu\_a - mu\_0) / n

Selection: 3

| Perseverance, that's the answer.

| |=================================================================== | 96%

| 8. True or False? More power is better than less power.

1: True

2: False

Selection: 1

| You got it right!

| |==================================================================== | 97%

| 9. True or False? A larger beta (call it beta\_max) is more powerful than a smaller

| beta.

1: False

2: True

Selection: 1

| You got it right!

| |==================================================================== | 98%

| 10. True or False? The larger the sample size the less powerful the test.

1: True

2: False

Selection: 2

| That's correct!

| |===================================================================== | 99%

| Congrats! You finished this powerful lesson. We hope you feel emPOWERED.

...

| |======================================================================| 100%

| Would you like to receive credit for completing this course on Coursera.org?

1: Yes

2: No

Selection: 1

What is your email address? sweeyean@gmail.com

What is your assignment token? CDS92hsLfrQO018u

Grade submission succeeded!

| Nice work!

| You've reached the end of this lesson! Returning to the main menu...

| Please choose a course, or type 0 to exit swirl.

1: Statistical Inference

2: Take me to the swirl course repository!

Selection: