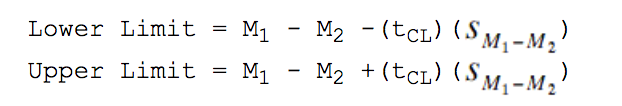
**INTRODUCTION TO STATISTICAL CONCEPTS**

**ASSIGNMENT 3B**

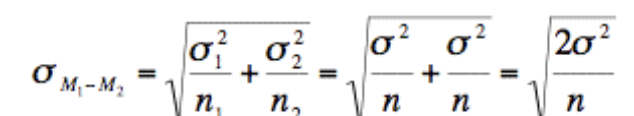
**Chapter 10: Probability**

**The following questions are from the**[**Angry Moods**](http://onlinestatbook.com/2/case_studies/angry_moods.html)**(AM) case study.**

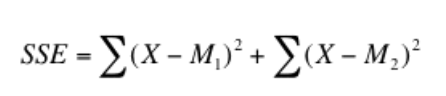
**Question 24: (AM#6c) Is there a difference in how much males and females use aggressive behavior to improve an angry mood? For the "Anger-Out" scores, compute a 99% confidence interval on the difference between gender means. (**[**relevant section**](http://onlinestatbook.com/2/estimation/difference_means.html)The difference between gender means for anger out scores is 0.795. To find the 99% confidence interval for men and women, we need to use the below formula, to get a confidence interval of (-1.803, 3.339).



First, we compute the estimate of the standard error between means, or find . To estimate this quantity, we need to estimate the variance and use the estimator in place of the actual. The SE of the difference in means in the population is written as follows:

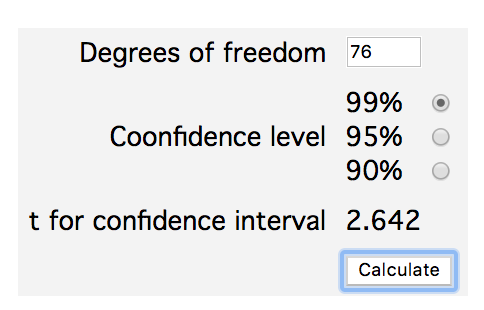


To estimate the variance, we average between the two sample variances of men and women to get the formula for SSE:

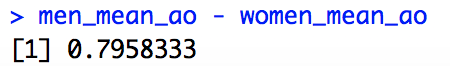


Next, we want to find t that we can use for the confidence interval tCL. To do this, we need to know the degrees of freedom. This is equal to (n1 - 1) + (n2 - 1) where n1is the sample size of the first group and n2 is the sample size of the second group.

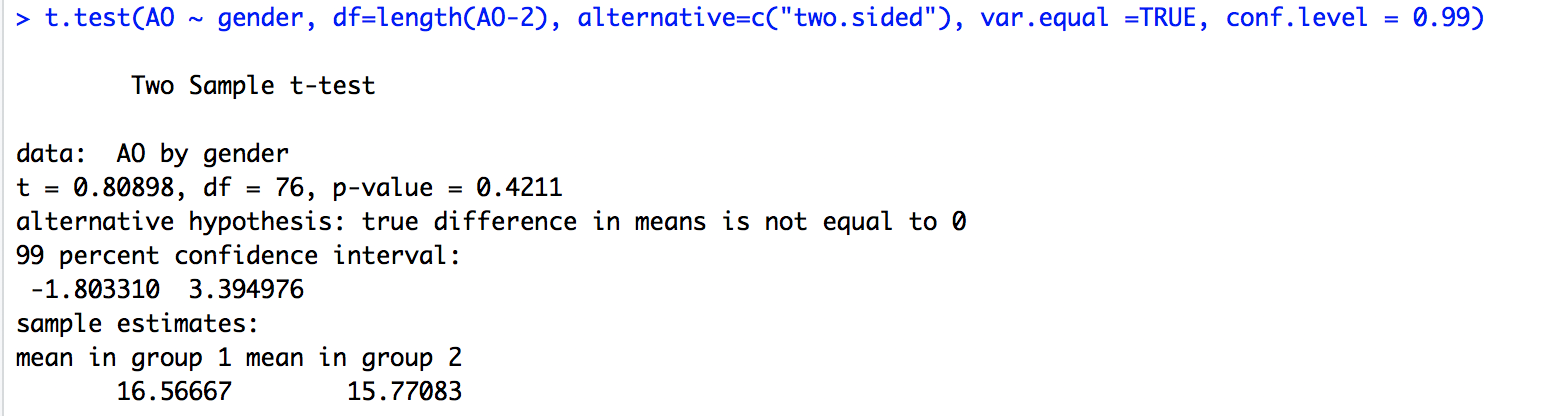
We can use the t-table or the inverse t calculator to find the t for a 99% confidence interval with df =76.



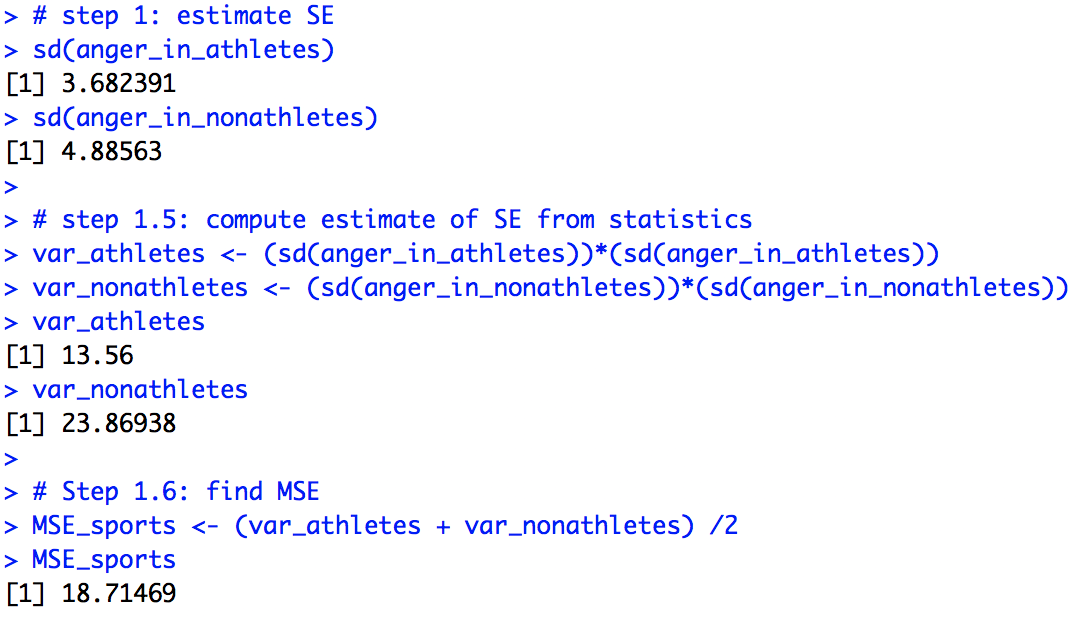
From last week’s exercise, we also already have the difference in means:

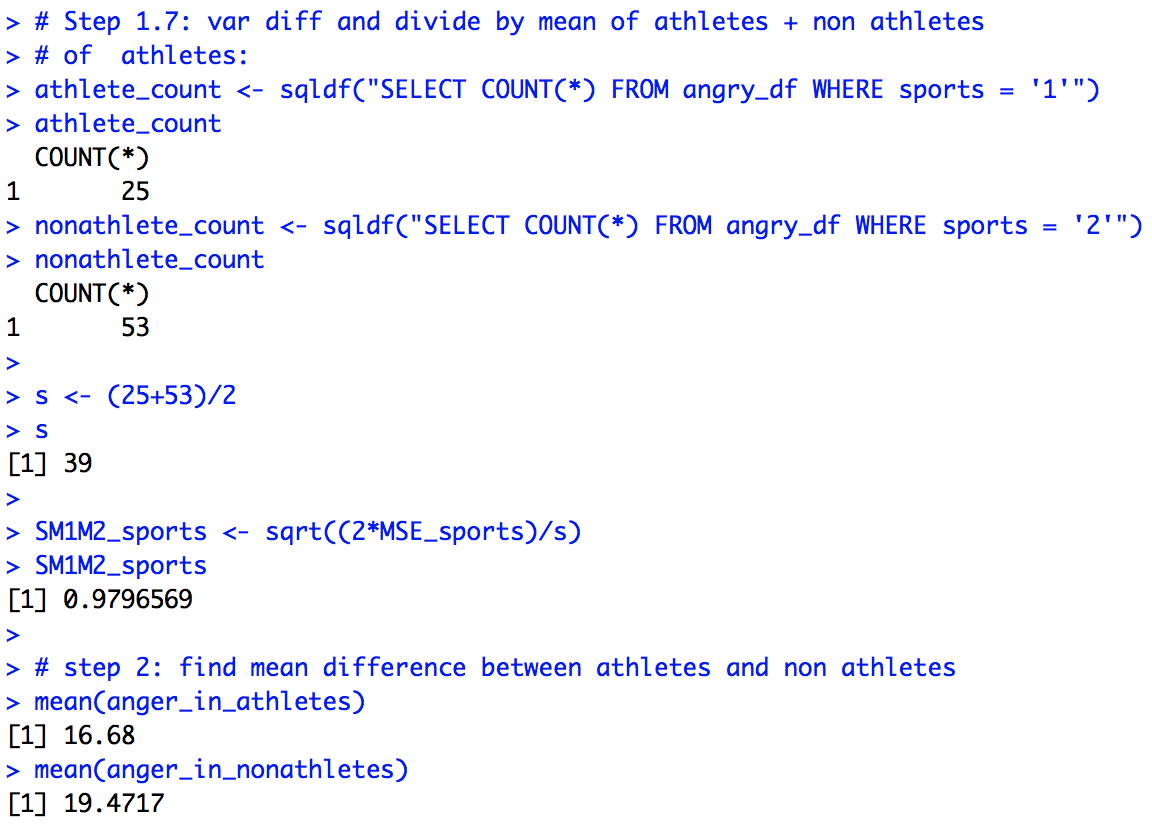


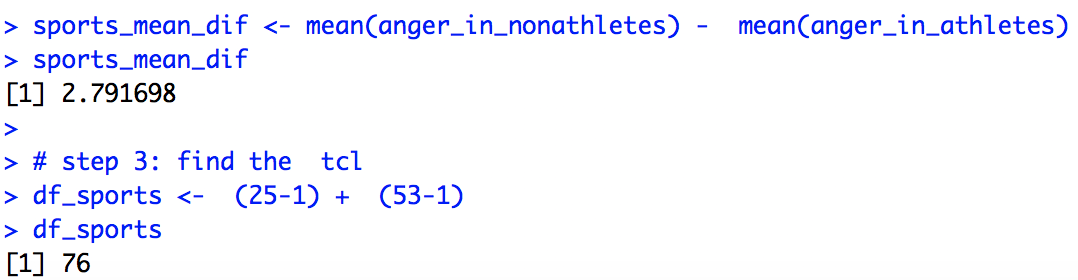
Now, we can go back to the formula above to compute the lower and upper limit using R, particularly with the t-test function.



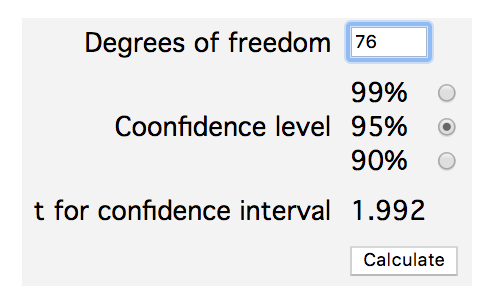
**Question 25: (AM#10) Calculate the 95% confidence interval for the difference between the mean Anger-In score for the athletes and non-athletes. What can you conclude? (**[**relevant section)**](http://onlinestatbook.com/2/estimation/difference_means.html)Using the methodology above, we solve as follows and we can conclude that athletes have a lower mean anger-in score as compared to non-athletes, in other words, we know that the difference between population means of athletes and non-athletes in terms of anger-in scores is not equal 0. We can also see that the confidence interval for the difference between their means is between -4.985 and -0.5977.



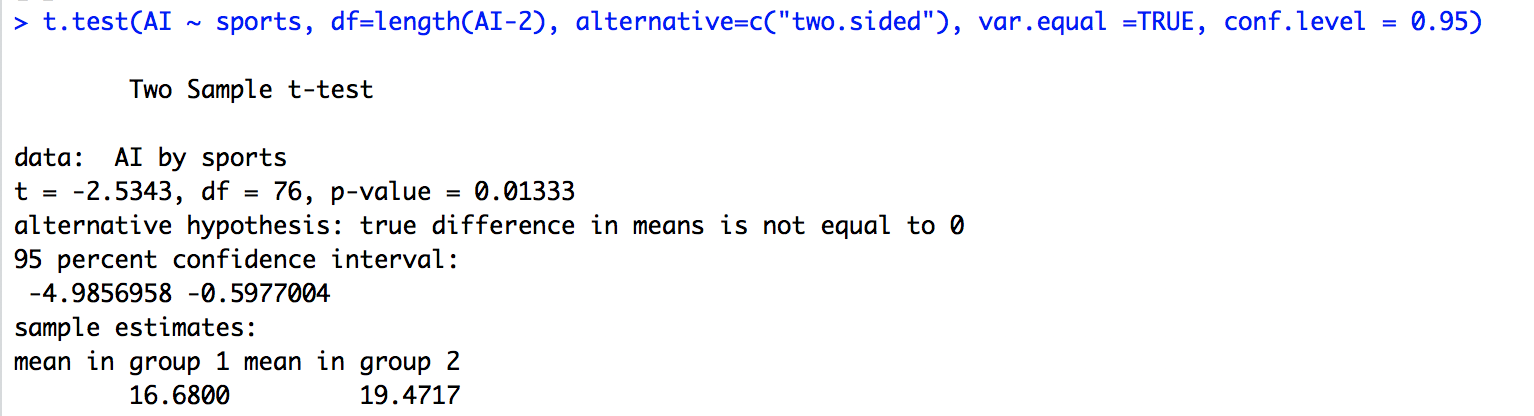




Finally, we find that the t value is 1.992.



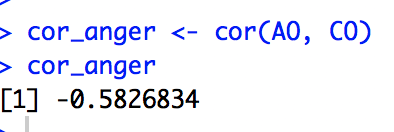
We can then plug these values into the formula to find the upper and lower limits of the confidence interval to get a lower limit of -4.985 and an upper limit of -0.5977.

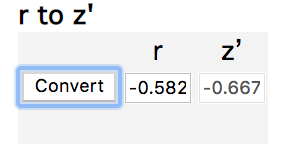


**Question 26: Find the 95% confidence interval on the population correlation between the Anger-Out and Control-Out scores. (**[**relevant section)**](http://onlinestatbook.com/2/estimation/correlation_ci.html)

To find the 95% confidence interval on the population correlation (ρ), we have three steps to get a confidence interval of (-0.713, -0.415)

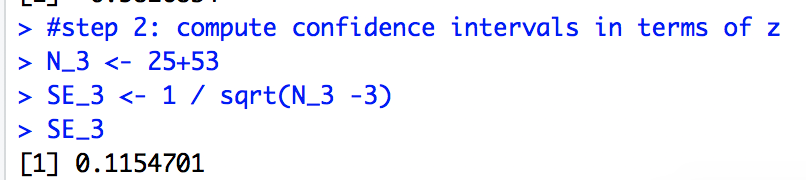
First, convert r to z’.



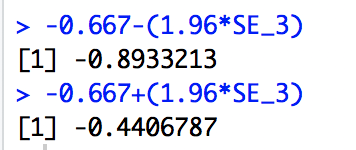
-

Then, compute a confidence interval in terms of z’. Using the formula for standard error below, we then get:

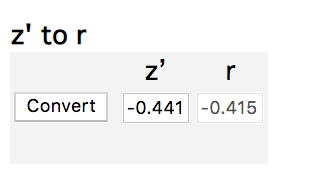
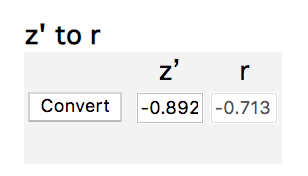
 , which gives us the following



We can then use this equation to find the right confidence intervals



Finally, convert the confidence interval back to r.



**Chapter 11: Logic of Hypothesis Testing**

**20. True/false: It is easier to reject the null hypothesis if the researcher uses a smaller alpha (α) level. (**[**relevant section**](http://onlinestatbook.com/2/logic_of_hypothesis_testing/significance.html)**&**[**relevant section**](http://onlinestatbook.com/2/logic_of_hypothesis_testing/errors.html)**)**

False—alpha is the probability of a Type I error given that the null hypothesis is true. Alpha assumes that we fail to reject the null hypothesis.

**21. True/false: You are more likely to make a Type I error when using a small sample than when using a large sample. (**[**relevant section**](http://onlinestatbook.com/2/logic_of_hypothesis_testing/errors.html)**)**

False, the probability of making a Type I error should not depend on sample size.

**22. True/false: You accept the alternative hypothesis when you reject the null hypothesis. (**[**relevant section**](http://onlinestatbook.com/2/logic_of_hypothesis_testing/significant.html)**)**

True, you can only reject the null hypothesis, but you can accept the alternative hypothesis.

**23. True/false: You do not accept the null hypothesis when you fail to reject it. (**[**relevant section**](http://onlinestatbook.com/2/logic_of_hypothesis_testing/nonsignificant.html)**)**

True, you do not necessarily accept the null hypothesis when you do not reject it.

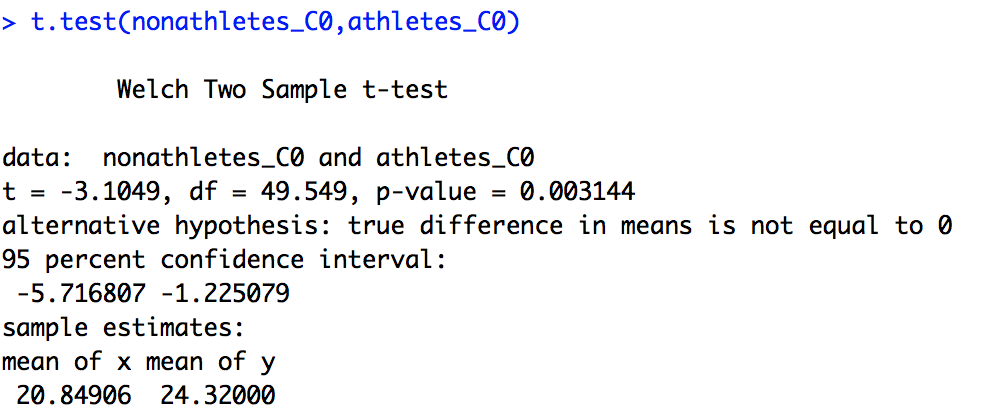
**24. True/false: A researcher risks making a Type I error any time the null hypothesis is rejected. (**[**relevant section**](http://onlinestatbook.com/2/logic_of_hypothesis_testing/errors.html)**)**

True, since when you reject the null hypothesis, you are assuming that H0 is true, and the type 1 error is the same as the alpha or the p-value. When a researcher rejects the null hypothesis when it is true, then she does face the risk of making a type 1 error.

**Chapter 12: Tests of Means**

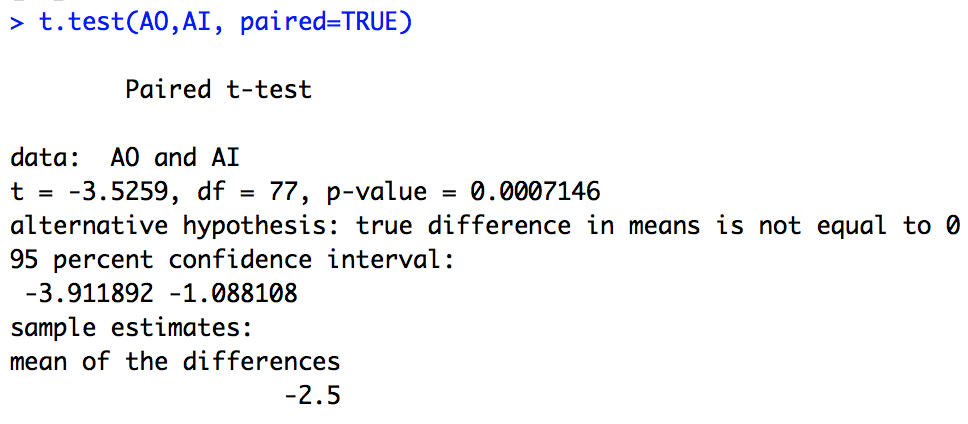
**The following questions use data from the**[**Angry Moods**](http://onlinestatbook.com/2/case_studies/angry_moods.html)**(AM) case study.**

**Question 21: (AM#17) Do athletes or non-athletes calm down more when angry? Conduct a t test to see if the difference between groups in Control-In scores is statistically significant.**



The difference is statistically significant, since the p value is below 0.05 and even below 0.01.

**Question 22: Do people in general have a higher Anger-Out or Anger-In score? Conduct a t test on the difference between means of these two scores. Are these two means independent or dependent? (**[**relevant section**](http://onlinestatbook.com/2/tests_of_means/correlated.html)**)**



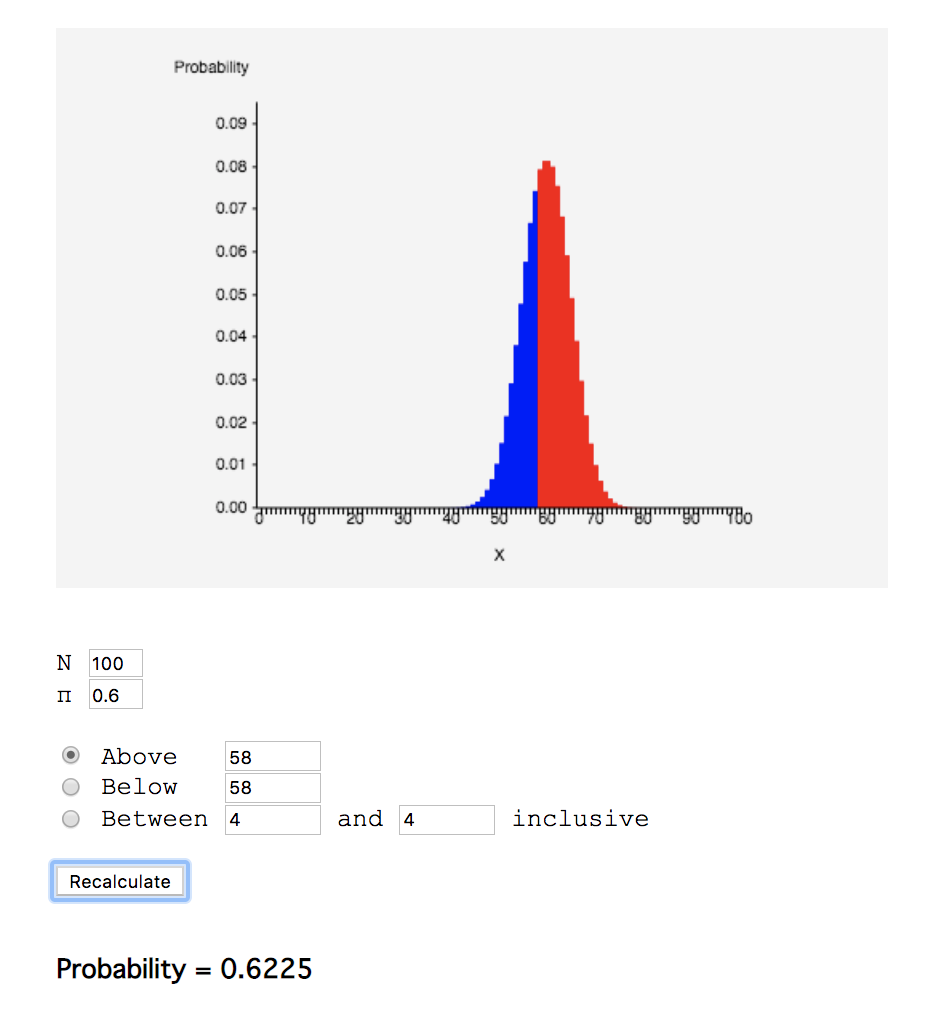
The Anger-In Score is generally higher. These two means are not independent as seen by the description of the alternative hypothesis.

**Chapter 13: Power**

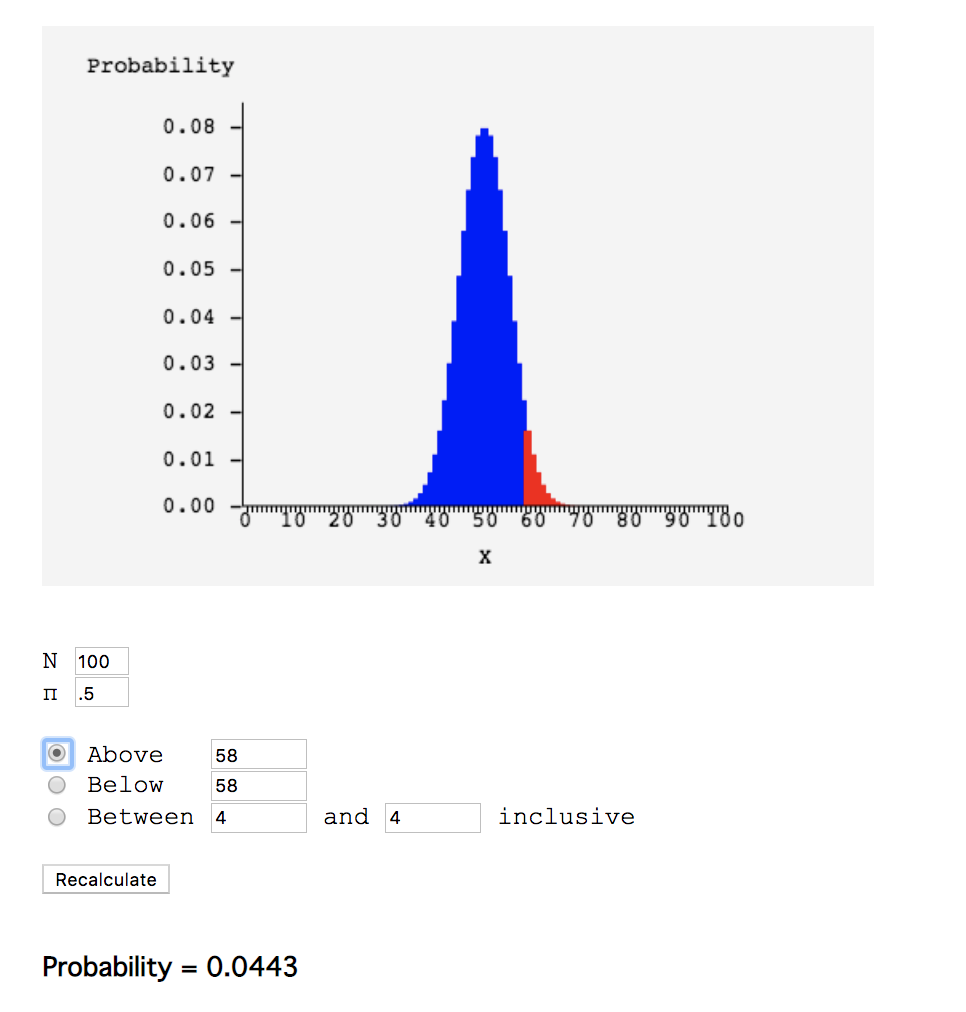
**5. Alan, while snooping around his grandmother's basement stumbled upon a shiny object protruding from under a stack of boxes. When he reached for the object a genie miraculously materialized and stated: "You have found my magic coin. If you flip this coin an infinite number of times you will notice that heads will show 60% of the time." Soon after the genie's declaration he vanished, never to be seen again. Alan, excited about his new magical discovery, approached his friend Ken and told him about what he had found. Ken was skeptical of his friend's story, however, he told Alan to flip the coin 100 times and to record how many flips resulted with heads.**

**(a) What is the probability that Alan will be able convince Ken that his coin has special powers by finding a p value below 0.05 (one tailed).   
Use the**[**Binomial Calculator**](http://onlinestatbook.com/2/calculators/binomial_dist.html)**(and some trial and error)**

Using trial and error on the binomial calculator, we can see that the greatest X value of P(X) that yields a probability of less than 0.05 is X =58. When we factor in the magic coin, with p = 0.6, we find that the probability of getting at least 58 heads (meaning that the coin has special powers), is 0.6225.



The above shows the final probability for the binomial distribution.

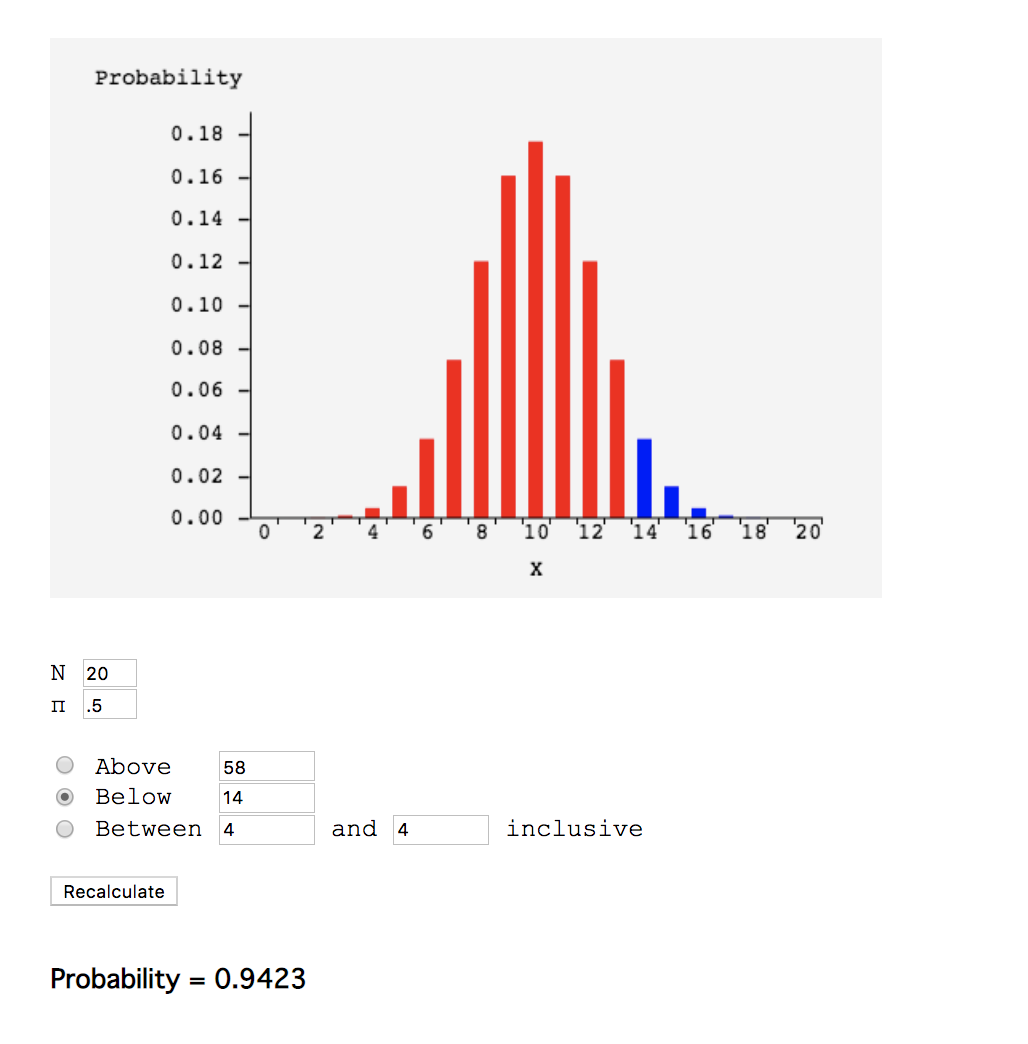


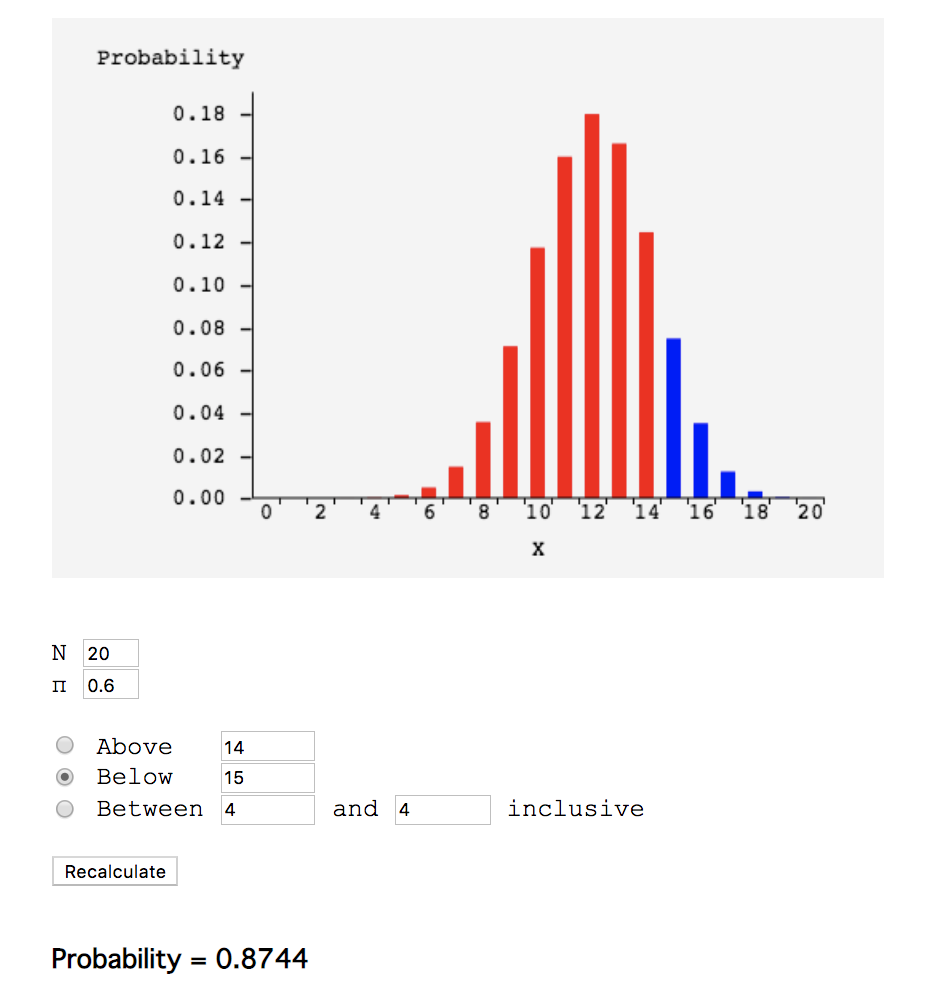
The above shows my trial and error method.

**(b) If Ken told Alan to flip the coin only 20 times, what is the probability that Alan will not be able to convince Ken (by failing to reject the null hypothesis at the 0.05 level)?**

If the coin was only flipped 20 times, then the probability that Alan would not be able to convince Ken would be 0.8744. Using trial and error, we see that the highest X we can find that has a probability less than 0.95 is X = 14.

Below X=14, we get a probability of 0.942, which is below 0.95, which is the value where we would reject the null hypothesis at the 0.05 level. Thus, plugging this X into the new distribution with p = 0.6, we get the probability of 0.8744. Note that the below graphs say X = 15, because the binomial probability is discrete and not continuous, thus we need numbers *below 15*, since it is 14 inclusive.

****

****