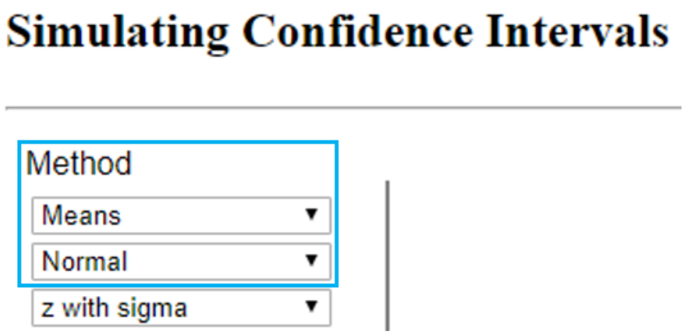
Lab 5 (100 pts.) - Interpretation of Confidence Intervals and Power Analysis

Objectives: A Better Understanding of Confidence Intervals and Power Curves.

This lab uses applets on the internet instead of your software package. Since the instructions will be included here, there will be no tutorials and only one solution.

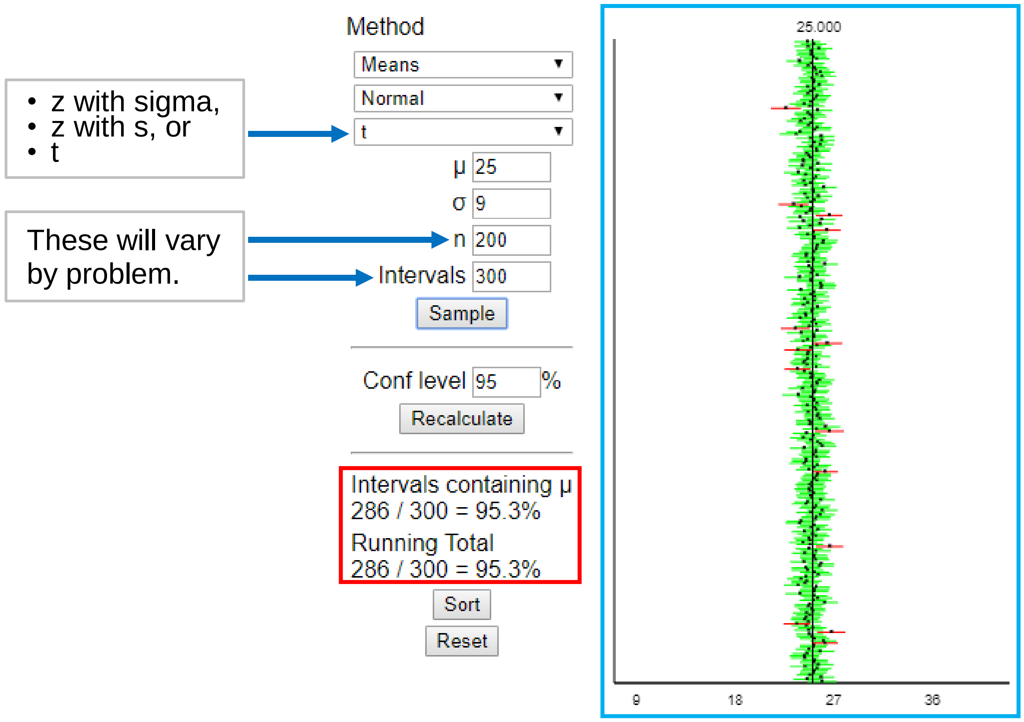
A. (10 points) Online Prelab

B. (45 points) Interpretation of a Confidence Interval. We are going to use one of the Rossman/Chance applets for this interpretation. The applet is located at: http://www.rossmanchance.com/applets/ConfSim.html. First, make sure the first drop-down box in the upper left is set to "Means" and the second is set to "Normal":



We will be performing simulations using three different distributions: a z-distribution (σ known), a z-distribution (σ not known, with s), and a t distribution (σ not known). All simulations will use a μ of 25 and a σ of 9 (If σ is not known, it is supposed to be s. This applet states the value as σ for both cases.). "n" refers to the sample size. Please use a confidence level ("Conf level") of 95%. You need to input theses values on the left panel of the applet (as shown in the screenshot below). The value of "Intervals" and "n" will be specified in each problem below. Once you have set all of the options, run the simulation by pressing Sample. To reset the simulation, press Reset.

You will be reporting two types of output which are shown below.

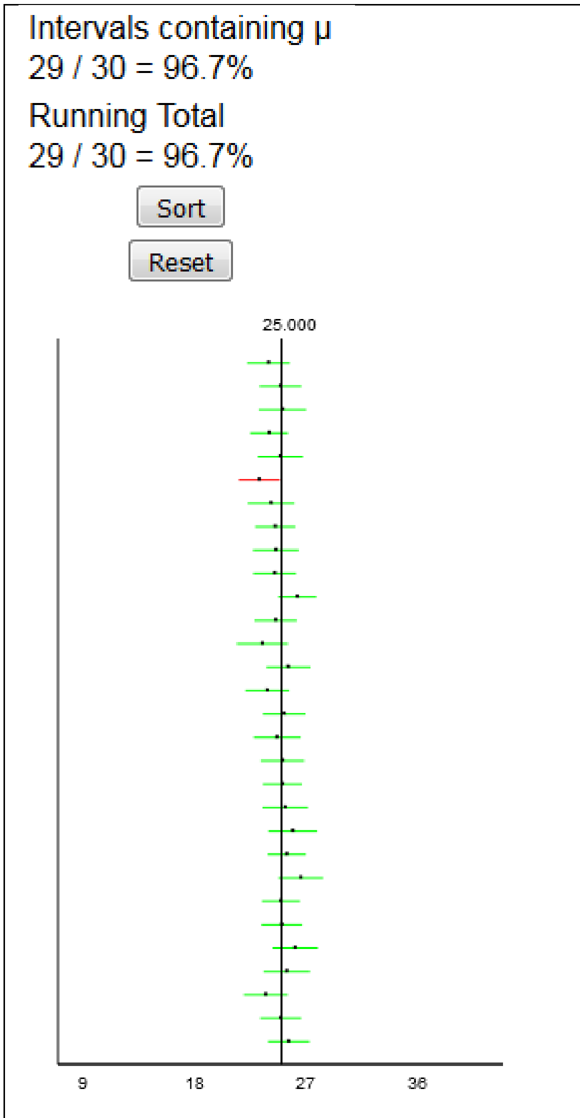


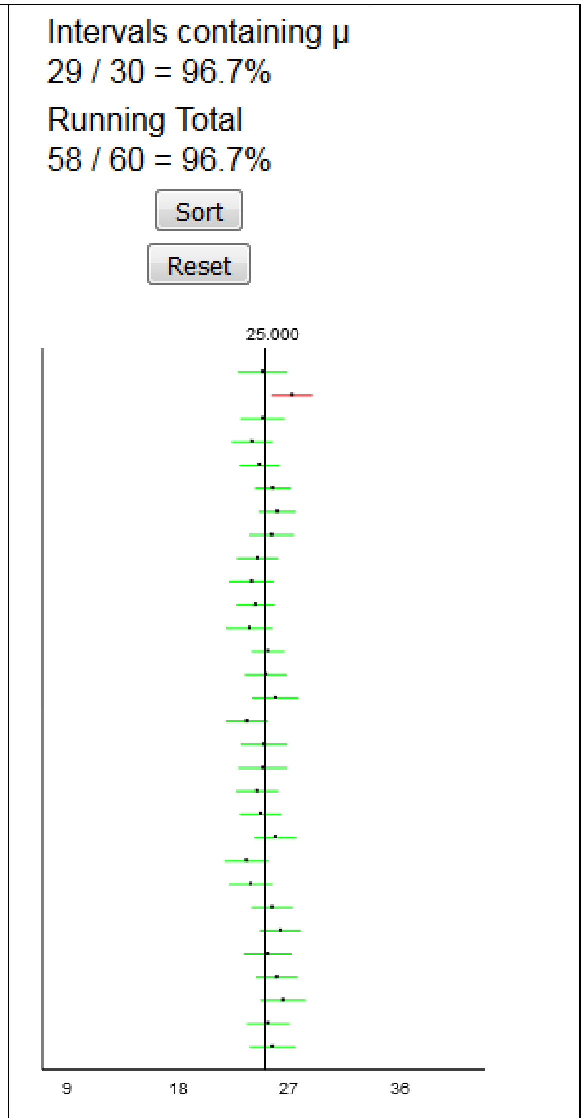
(1) The Intervals containing μ and the Running Total (lower left, in the red box).
(2) The figure of all the confidence intervals generated (right, in the blue box). If they contain the true mean, the line(s) will be in green. If they don't contain the true mean, the line(s) will be in red.

Note: only sample solutions are presented for this part because each person should have a different answer.

1. (10 pts.) Set the distribution to "z with sigma," "n" to 100, and the Intervals to 30. Press "Sample." Record both the lower left (red box), "Intervals containing μ" and "Running total", and the right (blue box), "the diagram showing all of the intervals", into your lab report. Repeat pressing "Sample" and recording the percentages and graphs until you have completed 300 total intervals. Since you will press sample 10 times, you will have 10 graphs and 10 reports of "Intervals containing μ" and the "Running Total." Please put this information into a table.

Solution:





Repeat this until there is total 10 graphs and the running total is 300.

2. (10 pts.) Were any of the percentages, either in the "Interval containing μ" or the "Running Total," in question 1) exactly 95%? If so were, how many? Does this imply that there is a problem with the calculations done with the applet? Please explain your answer.

Solution:

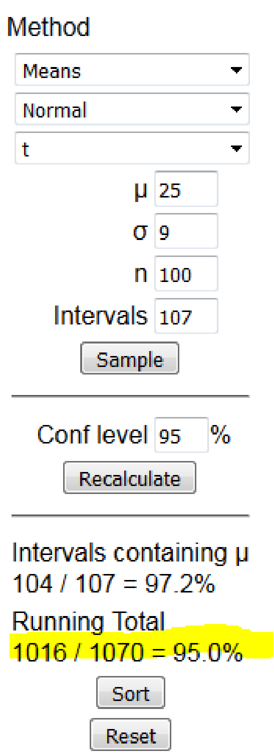
None of the percentages from "Intervals containing μ" can be exactly 95%.
There were no instances where 95% of the intervals contained the true mean. Though this is a slight possibility.

This is not a problem. The theory says that, when the number of intervals created is very large, the proportion of intervals containing the true mean will approach 0.95. On the one hand, when the number of intervals is too small, there is a substantial probability that the proportion will differ considerably from 0.95. On the other hand, this limiting result only says that, when the number intervals is large enough, the proportion will be in a very small neighborhood around 0.95. There is always a possibility that the probability is not 0.95, but the difference will get arbitrarily small.

3. (5 pts.) Reset the simulation and choose the t distribution method with Intervals equal to 107. Keep the sample size, n, at 100. Press Sample until the **Running Total** percentage is close to 95% (within 0.2%) and take a screen shot of the final cumulative total (red box). You do not need to report the individual totals or graphs each time. What do you think **mathematically large** is in this circumstance? That is, how many intervals were required so that your percentage of Running Total is close to 95%?

Solution:

Output required is the screen shot of what is highlighted in yellow.



I got it in 10 tries, for a running total of 1070 intervals. In this case, mathematically large is around 1000.

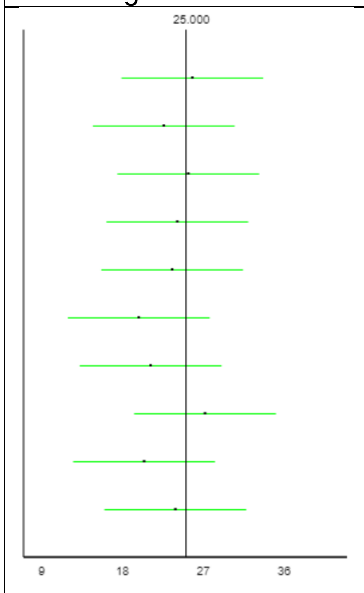
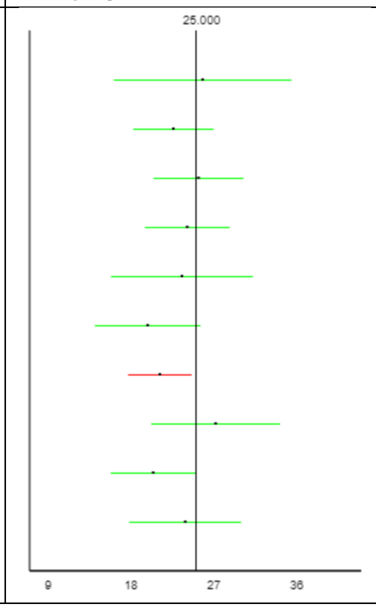
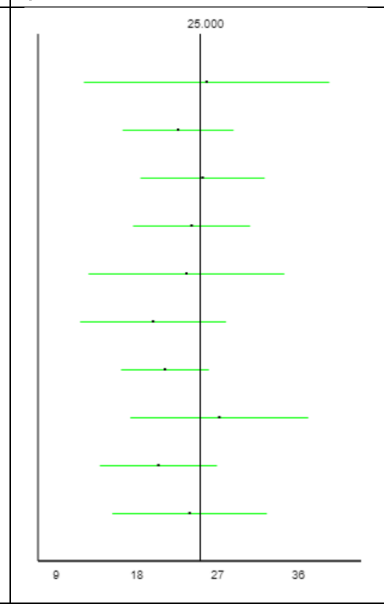
When I was testing this, I was getting running totals of 214 (2 tries) or as large as 1177 for the percentage to be within 0.2% of 95%.

4. (15 pts.) After pressing "Reset," change n to 5 and the intervals to 10 with the same mean, standard deviation, and confidence level. The only output required in this part

is the figure of all of the confidence intervals (blue box); we are not interested in the percentage of "Intervals containing μ."

a. (3 pts.) With the distribution set as "z with sigma," press "Sample" and copy the figure of all the confidence intervals to your report. Then, change the distribution to "z with s," and copy the figure after it has changed. Finally, change the distribution to "t" and copy the new figure. Be sure to clearly label which figure corresponds to which distribution. I recommend reducing the sizes of the figures so that all three of them fit on one line for ease of comparisons. Note that you only need to press "Sample" once for this part.

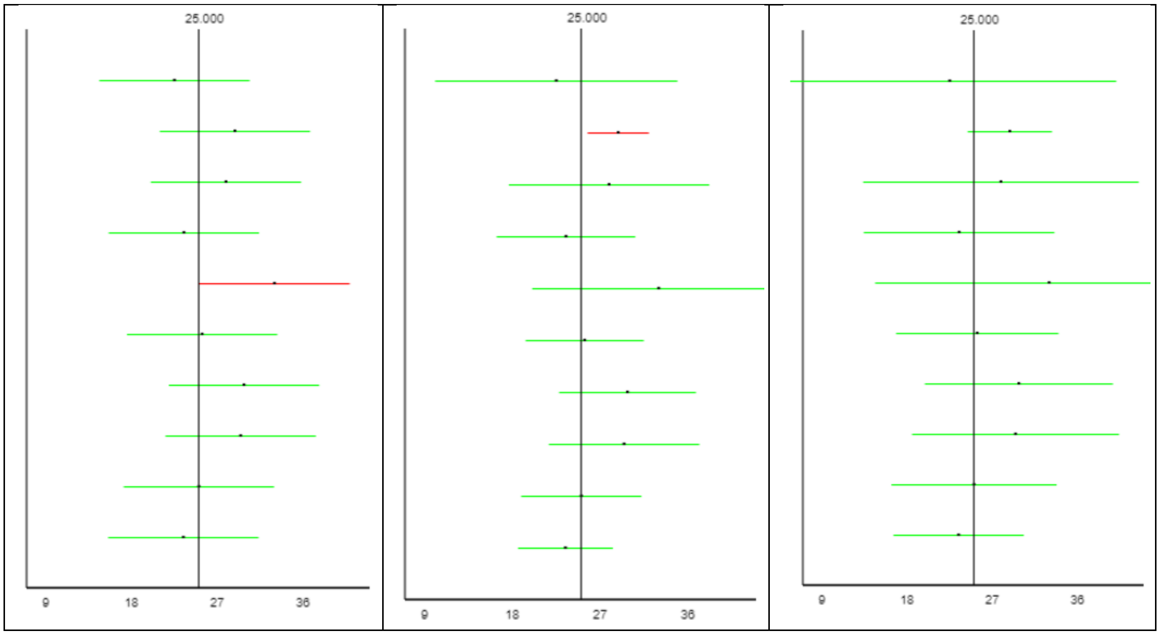
Solution:

z with sigma	z with s	t
		

b. (2 pts.) Repeat part a) once to obtain another set of the three figures (two sets in total: one from part a), and one from b)).

Solution:

z with sigma	z with s	t



c. (5 pts.) Comparing the "z with sigma" and "z with s" figures, describe the difference between the two figures using the population standard deviation and the two figures using the sample standard deviation. Please explain why they are different. In your response, be sure to indicate which option in the applet is causing the difference.

Solution:

The center of each interval is the same. For the distribution with sigma, all of the widths are the same. When the sample standard deviation is used, the widths vary. This is caused by the fact that s depends on the sample, but the population standard deviation is a constant.

d. (5 pts.) Comparing the "z with s" and the "t" figures, describe the difference between the two figures using the z distribution and the two figures using the t distribution. Please explain why they are different. In your response, be sure to indicate which option in the applet is causing the difference.

Solution:

The center of each interval is the same. The widths of the intervals are larger using the t distribution than the z distribution with s. This is caused by the fact that the critical value for z, $z_{\alpha/2}$, is less than the critical value for t, $t_{\alpha/2,n-1}$.

5. (5 pts.) In question 4, you generated Confidence Intervals based on 3 models (z with sigma, z with s, and t) for each of the two different data sets (samples), and used them to see how changing the model affects the Confidence Interval. In terms of study design, identify the study unit, treatment levels, and response in this exercise, and explain how the exercise is similar to a matched pair study design. In this case, why is this design preferable to using a new sample for each new set of Confidence Intervals (i.e., using 9 samples to make 9 graphs, instead of using 3 samples)?

Solution:

Each sample is a study unit, the different models are treatment levels, and the pattern of the CIs is the response. The exercise is similar to a matched pair because all of the treatments are applied to each study unit and compared to each other within the same study unit. Although random variability among samples does not introduce a confounding variable, it does add extra noise to the signal that we are interested in examining. Applying all the models to each data set (blocking) reduces the random noise that would result from using a new sample each time and makes it easier to see the effect that changing the model has on the CI (using a completely randomized design with 1 treatment per sample, we would probably need > 9 samples to get similar level of clarity in the result that we got in question 4 using only 3 samples). By reducing extraneous noise, blocking allows us to draw a valid conclusion with a smaller sample size.

C. (50 points) Interpretation of Power. R can also be used to create applets using a tool called Shiny. The applet that we are going to use is located at: http://shiny.stat.tamu.edu:3838/eykolo/power/. The inputs are on the left side. So that you can more easily see the results, please only shade the quantities that are requested. To make the grading easier, **please use the graph bounds from 400 to 650 ONLY** (use the last two input fields called "Lower Bound of Plot" and "Upper Bound of Plot").

The Deely Laboratory provides drinking-water testing and analysis services. In Indiana, the laboratory tests for roundup (glyposate). Roundup enters drinking water through runoff of the corn fields. The service knows that their analysis procedure is unbiased but not perfectly precise, so the laboratory analyzes each water sample ten times and reports the mean result. The measurements all follow a Normal distribution quite closely. The standard deviation of this distribution is a property of the analytic procedure and is known to be $\sigma = 75$ parts per billion (ppb).

The Deely Laboratory has been asked by the university to evaluate a claim that the drinking water in the Student Union has an average Roundup concentration below 500 ppb which is well below the Environmental Protection Agency's action level of 700 ppb. Therefore, the hypotheses that we are testing are

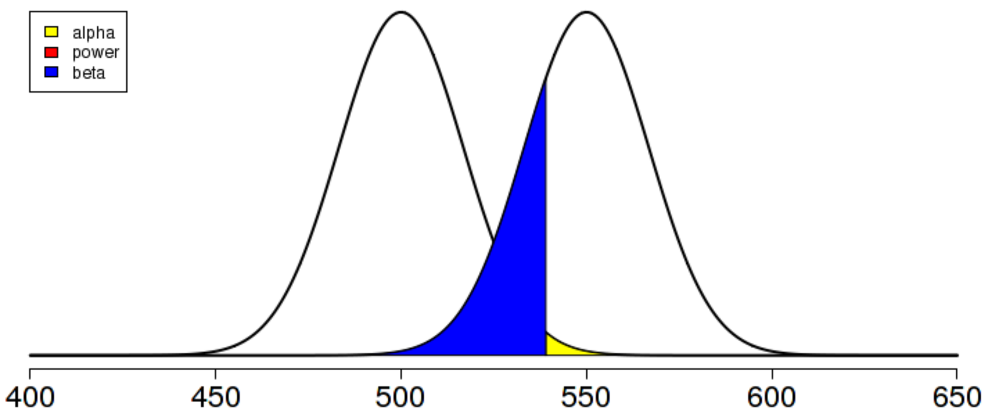
H_0: $\mu = 500$
H_a: $\mu > 500$

The lab chooses a significance level and plans to perform twenty analyses of one specimen (n = 20).

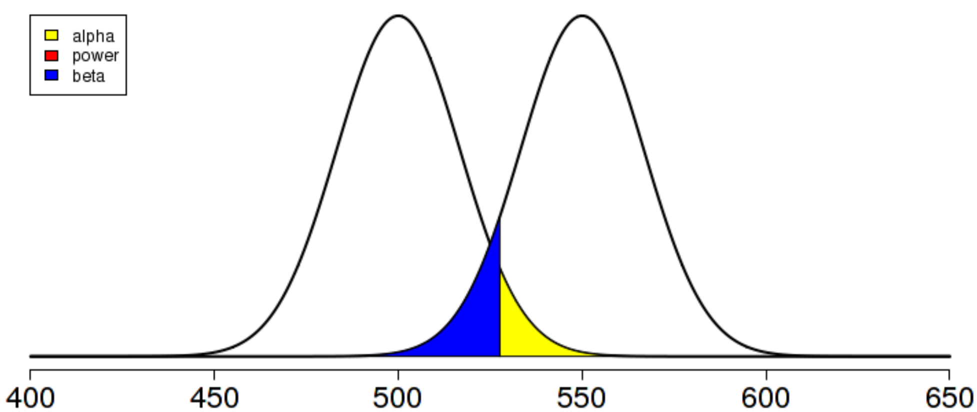
1. (20 pts., 5 pts. each) The answer to each part will be a plot with alpha and beta shaded.

   a. At the 1% level of significance, show the plot of this test against the specific alternative $\mu_a = 550$.

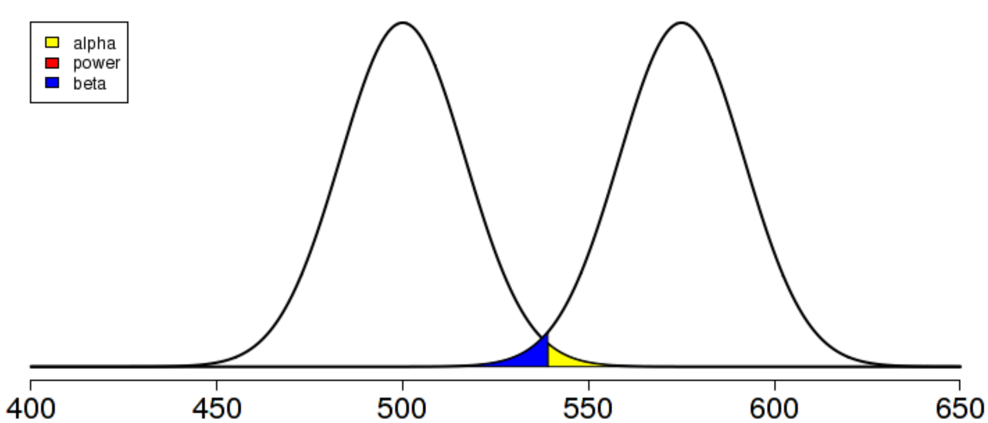
Solution:



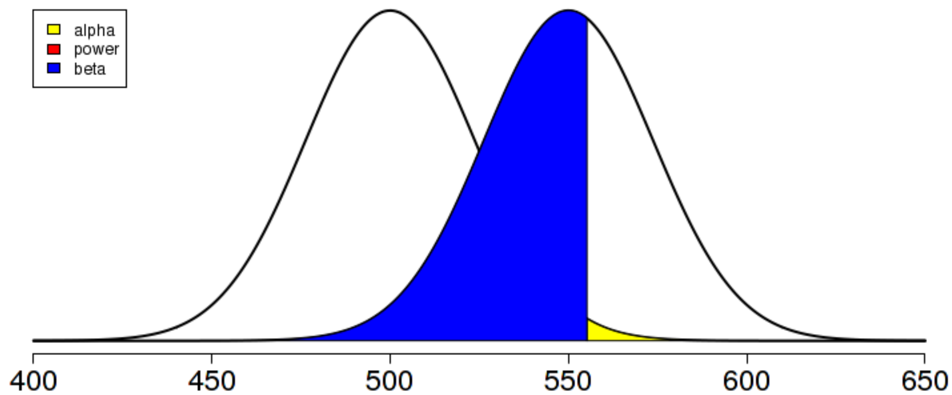
b. At the 5% level of significance, show the plot of this test against the specific alternative $\mu_a = 550$.



c. At the 1% level of significance, show the plot of this test against the specific alternative $\mu_a = 575$.



d. If the lab performs 10 analyses of one specimen (n = 10), show the plot of this test against the specific alternative $\mu_a = 550$? This is at a 1% significance level.



2. (10 pts.) Write a short paragraph explaining the consequences of changing the significance level (parts a vs. b), alternative μ_a (parts a vs. c), and sample size (parts a vs. d) on β. Please include in your discussion what causes the change in β, change of cutoff location, width of distribution, and/or distance between the distributions. There should be at least one sentence per pair.

Solution:

Between parts a) and b), everything stays the same except that the cutoff value moves to the left which decreases β.

Between a) and c), the distance between the distributions increases and everything else remains the same, causing a decrease in β .

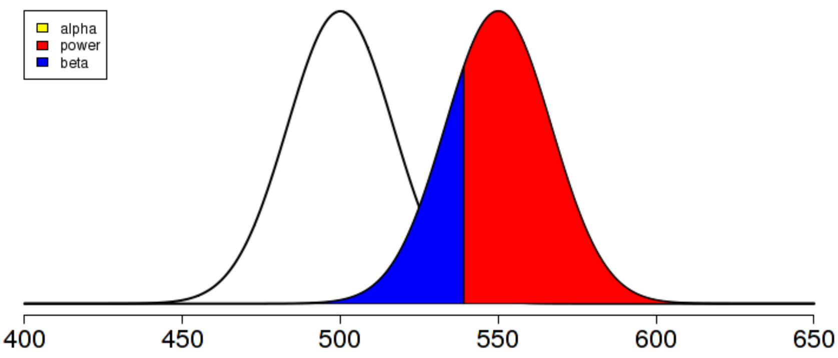
Between a) and d), the width of both distributions is increased, and β is increased.

Thus β will be decreased by (1) increasing α with all else equal, (2) Increasing $|\mu_0 - \mu_A|$ all else equal, and (3) increasing the sample size, all else equal.

3. (10 pts., 5 pts. each) How are β and power related?

a. Display the plot for any one of the situations that you obtained from question 1, but shade the power and beta and don't shade alpha.

Solution:



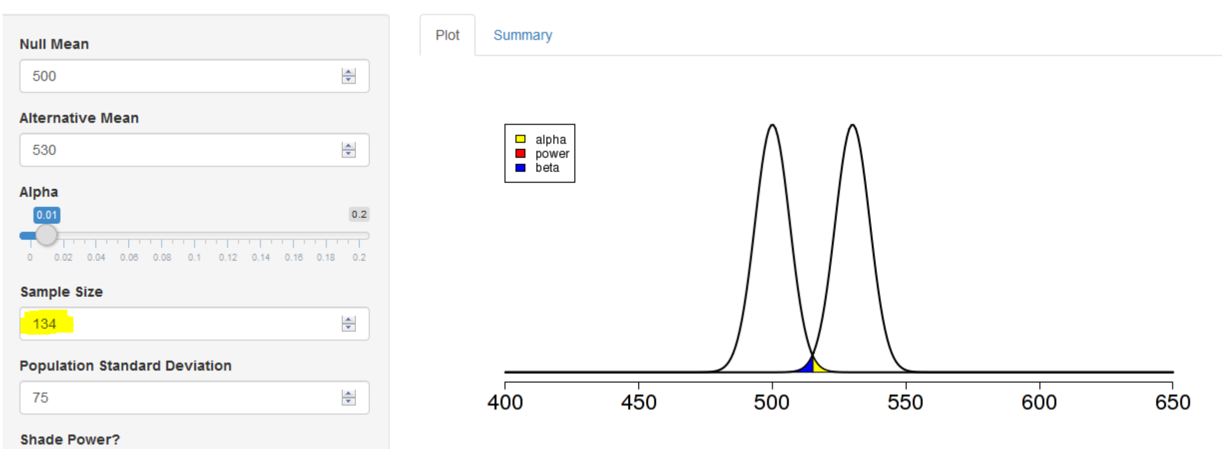
b. What is the formula relating power and β?

Solution:

We see that the blue and red areas combined constitute the whole area under the right curve. The area under a probability distribution is always 1. Hence, $power + \beta = 1$, or $power = 1 - \beta$

4. (5 pts.) What sample size would be required for β to be approximately 0.01 at the 1% level of significance against the specific alternative $\mu_a = 530$? To answer this question, manually change the sample size, with alpha and beta being shaded and the power un-shaded, until the areas corresponding to alpha and beta (blue and yellow) look approximately the same. Please provide the plot with your final sample size. Be sure to state what the sample size is.

Solution:



n = 134. (Any value around 130~140 are close enough thus will be consider as correct.)

5. (5 pts.) For each of the following probabilities; α, β, and power, use their respective definitions to answer which of the hypotheses, H_o or H_a, is assumed to be true. Hint: Which distribution is shaded with the probability?

Solution:

$\alpha = P(Reject\ |Null\ is\ True)$
$\beta = P(Fail\ to\ Reject\ |Alternate\ is\ True)$
$Power = 1 - \beta$

Hence α assumes H_0 is true while β and power assume that H_a is true.