Activity #20: One-sample t-test

In the last activity, we learned how to conduct a hypothesis test (a one-sample z-test, where we know σ:

1. State two competing hypotheses about a population parameter
   a. H_0: the null hypothesis (typically a statement of no effect)
   b. H_1: the alternative hypothesis (typically a statement of the effect you predict to see)
2. Consider potential decision errors and choose an appropriate level of significance (α)
3. Determine the statistic you will calculate from your sample to estimate the parameter of interest
4. Sketch the sampling distribution of that statistic (under the assumption that the null hypothesis is true)
   a. Calculate the statistic you observer from your sample
   b. Locate that statistic on your sampling distribution and compare it to a critical region on your sampling distribution
5. Make a decision to reject or retain the null hypothesis
   a. Calculate a p-value (the probability of observing data as or more extreme given a true null hypothesis)

Scenario: In the 1980s, many companies experimented with *flextime*, allowing employees to choose their work schedules within broad limits set by management. It was believed that flextime would reduce absenteeism.

Suppose a company is willing to try flextime for one year to see if it reduces absenteeism. Based on historical data, this company knows its employees have averaged 6.3 days absent. The company chose 100 employees at random to try flextime for one year. During that year, those 100 employees averaged 5.5 days absent with a standard deviation of 2.9 days.

1) State the null and alternative hypotheses. Is this study a one-tailed (confirmatory) or two-tailed (exploratory) study?

2) Express the consequences of both Type I and Type II errors. Which error has the most serious consequences? At what level should α be set?

3) According to the scenario, the distribution of "days absent" for the entire company has μ=6.3. We are not told the shape or the standard deviation of that distribution. We know even less about the distribution of days absent for flextime employees (we do not know the mean, shape, or standard deviation).

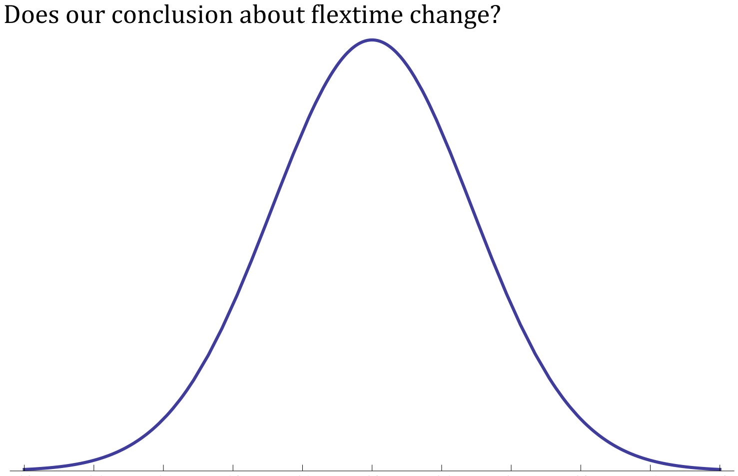
For our sample of 100 employees, we know the sample average is 5.5 with a sample standard deviation of 2.9. What distribution does this sample average of 5.5 come from? In other words, if we could repeatedly sample 100 employees, have them try flextime for a year, and calculate their average days absent, what would the distribution of those averages look like? What would be the mean and standard error of that distribution? What assumptions are you making? Sketch the sampling distribution on the next page.

4) Suppose we set α=0.05. Find the critical value(s) corresponding to this α-level and shade-in the appropriate critical region.

5) Draw an arrow to represent our observed sample mean (5.5) in your sampling distribution. Convert this observed sample mean to an observed t-statistic.

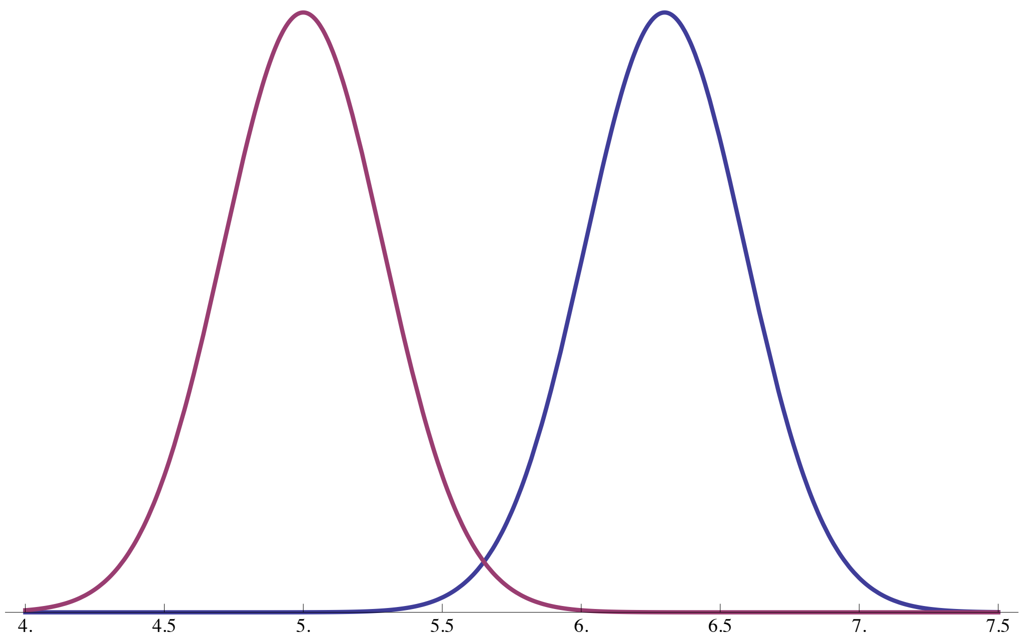
6) Compute the p-value from this study. What can we conclude from this p-value? Does this p-value represent the probability that the null hypothesis is true?

7) The sampling distribution of our sample averages (assuming a true null hypothesis) has been drawn below. Once again, draw an arrow to locate our observed mean of 5.5. Find the critical values for α=0.001, α=0.01, α=0.05, and α=0.10. What happens as we increase α? Does our conclusion about flextime change?

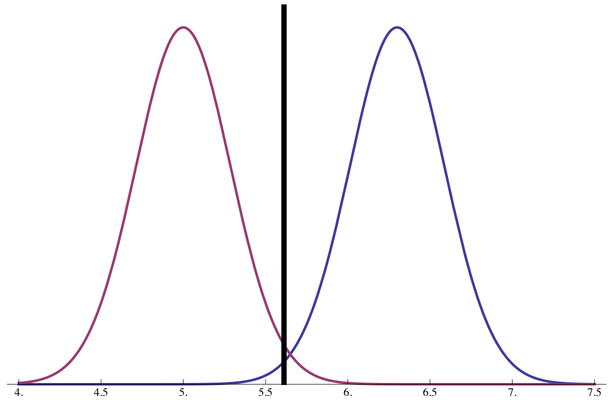


8) In any hypothesis test, what decision do we make about the null hypothesis if our p-value < α? What if our p-value > α?

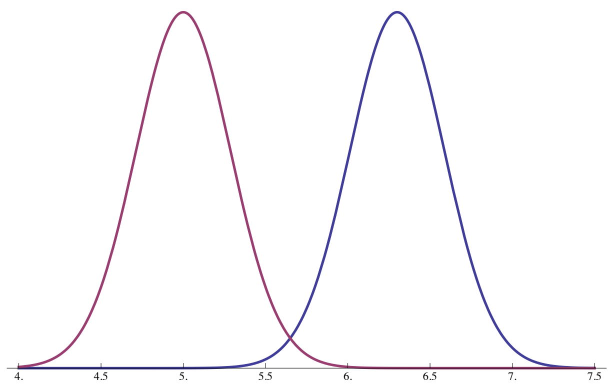
9) Suppose, in reality, flextime actually reduces absenteeism by 1.3 days (in other words, if we could have sampled our entire population, we would have found flextime employees at this company only miss 5 days each year). Given this reality (where the null hypothesis is no longer true), calculate the power of this study. Use the curves below to help:



10) Let's examine this concept of power in more depth. Below, I've drawn sampling distributions for both the null and alternative hypotheses. Using α=0.05, the critical value is also displayed. Shade-in α & β. Identify what power represents.



11) What happens to power as we choose smaller or larger values for α?



12) What happens to power if we increase our sample size? In the following curves, I increased n from 100 to 210.