Homework 6: Sections 5.1 - 6.5
KEY

STA209-04: Applied Statistics

Assigned: 03/08/2019
Due: 04/01/2019

Total Possible Points: 50

From the Book:

Questions: 5.31, 5.46, 6.37, 6.60, 6.75, 6.107, 6.124, 6.145, 6.173, 6.195, 6.221, 6.255, 6.257

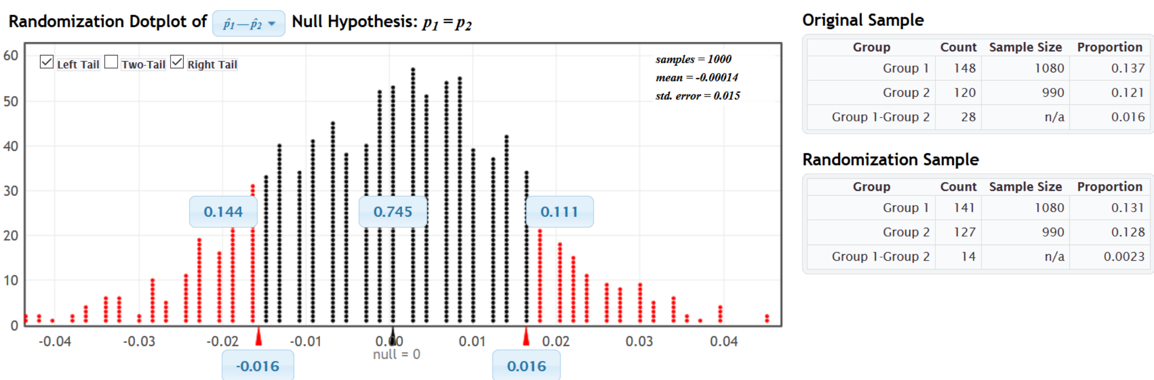
5.31 [**5 pts**] In a smoking cessation program, over 2000 smokers who were trying to quit were randomly assigned to either a group program or an individual program. After six months in the program, 148 of the 1080 in the group program were successfully abstaining from smoking, while 120 of the 990 in the individual program were successful. We wish to test to see if this data provide evidence of a difference in the proportion able to quit smoking in between smokers in a group program and smokers in an individual program.

a) State the null and alternative hypotheses, and give the notation and value of the sample statistic.

The null hypothesis is that the proportion of individuals able to quit smoking is the same in the group and individual programs. The alternative hypothesis is that the proportions are different. Symbolically, $H_0 : p_g - p_i = 0$; $H_A : p_g - p_i \neq 0$, where p_g represents the proportion of individuals in the group program that were able to quit smoking and p_i represents the proportion of individuals in the individual program that were able to quit smoking. Te value of the sample statistic, $\hat{p}_g - \hat{p}_i$, is 0.016.

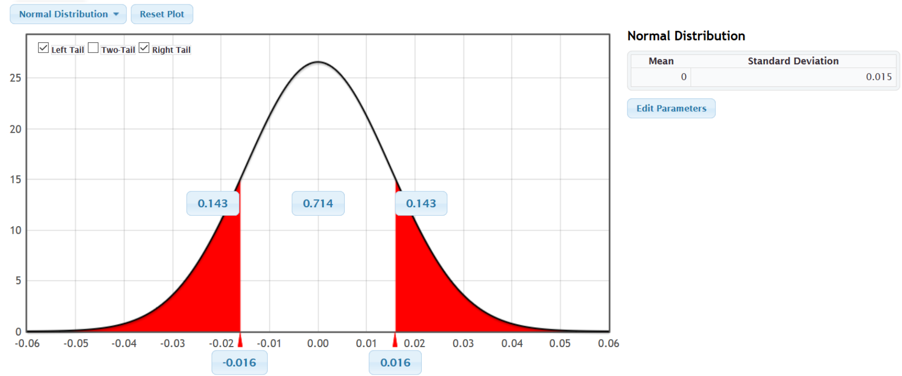
b) Use a randomization distribution and the observed sample statistic to find the p-value.

Based on the randomization distribution provided below, the p-value is $0.144 + 0.111 = 0.255$.



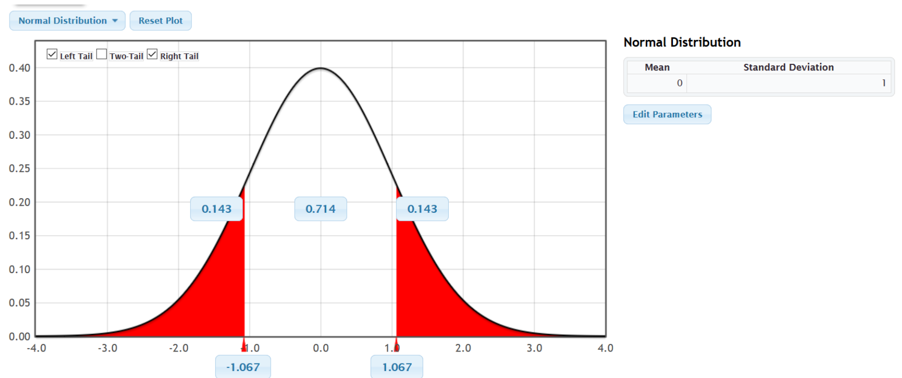
c) Give the mean and standard error of the normal distribution that most closely matches the randomization distribution, and then use this normal distribution with the observed sample statistic to find the p-value.

The normal mean and standard error that most closely matches the randomization distribution is $\mu = 0$ and $\sigma = 0.015$. This distribution is provided on the next page, and using this distribution, the p-value is 0.286.



d) Use the standard error found from the randomization distribution in part (b) to find the standardized test statistic, and then use that test statistic to find the p-value using a standard normal distribution.

Using the standard error from the randomization distribution, the standardized test statistic, z_{test}, is $\frac{(\hat{p}_g - \hat{p}_i) - 0}{SE} = 0.016/0.015 = 1.067$. Using this test statistic, the p-value is again 0.286.



e) Compare the p-values from parts (b), (c), and (d). Use any of these p-values to give the conclusion of the test.

All p-values are approximately the same. Using any of the three, we conclude that there is insufficient evidence to reject the null hypothesis. We may not conclude that there is a difference between programs in the proportion of individuals who quit smoking.

5.46 [4 pts] Hospital admissions for asthma in children younger than 15 years was studied in Scotland both before and after comprehensive smoke-free legislation was passed in March 2006. Monthly records were kept of the annualized percent change in asthma admissions. For the sample studied, before the legislation, admissions for asthma were increasing at a mean rate of 5.2% per year. The standard error for this estimate is 0.7% per year. After the legislation, admissions were decreasing at a mean rate of 18.2% per year, with a standard error for this mean of 1.79%. In both cases, the sample size is large enough to use a normal distribution.

a) Find and interpret a 95% confidence interval for the mean annual percent rate of change in childhood asthma hospital admissions in Scotland before the smoke-free legislation.

Using the normal approximation, the confidence interval is computed:

$$5.2 \pm 1.96 * 0.7 = (3.83, 6.57)$$

We are 95% confident that asthma admissions were increasing at a mean rate of at least 3.83% and at most 6.57%.

b) Find a 95% confidence interval for the same quantity after the legislation.

Using the normal approximation, the confidence interval is computed:

$$18.2 \pm 1.96 * 1.79 = (14.69, 21.71)$$

c) Is this an experiment or an observational study?

This is an observational study.

d) The evidence is quite compelling. Can we conclude cause and effect?

Since this is an observational study, we can not conclude cause an effect. There may be some confounding factor that explains the observed disparity.

6.37 [3 pts] A survey of 1000 adults in the US conducted in March 2011 asked "Do you favor or oppose 'sin taxes' on soda and junk food?" The proportion in favor of taxing these foods was 32%.

a) Find a 95% confidence interval for the proportion of US adults favoring taxes on soda and junk food.

Given the above information, we have $\hat{p} = 0.32$. Since $n\hat{p} = 1000 * 0.32 = 320 \geq 10$, we may use a normal approximation for constructing the 95% CI:

$$0.32 \pm 1.96\sqrt{\frac{0.32(1-0.32)}{1000}} = (0.29, 0.35)$$

b) What is the margin of error?

The margin of error is

$$1.96\sqrt{\frac{0.32(1-0.32)}{1000}} = 0.03$$

c) If we want a margin of error of only 1% (with 95% confidence), what sample size is needed?

To obtain a margin of error of 1% we would need a sample size of:

$$n = \hat{p}(1-\hat{p})\frac{z_{crit}^2}{MOE^2} = 0.32(1-0.32)\frac{1.96^2}{0.01^2} \approx 8360$$

6.60 [5 pts] Can babies reason probabilistically? A study investigates this by showing ten-to-twelve-month-old infants two jars of lollipop-shaped objects colored pink or black. Each infant first crawled or walked to whichever color they wanted, determining their "preferred" color. They were then given the choice between two jars that had the same number of preferred objects, but that differed in their *probability* of getting the preferred color; each jar had 12 in the preferred color and either 4 or 36 in the other color. Babies choosing randomly or based on the absolute number of their preferred color would choose equally between the two jars, while babies understand probability would more often choose the jar with the higher proportion of their preferred color. Of the 24 infants studied, 18 chose the jar with the higher proportion of their preferred color. Are infants more likely to choose the jar with the higher proportion of their preferred color?

a) State the null and alternative hypotheses.

The null hypothesis is that the proportion of babies who choose the jar with the higher proportion of their preferred color is 0.50. The alternative hypothesis is that babies will choose the jar with the higher proportion of their preferred color more often (i.e. the proportion is greater than 0.50).

b) Give the relevant sample statistic, using correct notation.

The relevant sample statistic is the sample proportion, $\hat{p}_{pref}$, where $\hat{p}_{pref}$ is the proportion who selected the jar with the higher proportion of the preferred color.

c) Which of the following should be used to calculate a p-value for this dataset? A randomization test, a test using the normal distribution, or either one? Why?

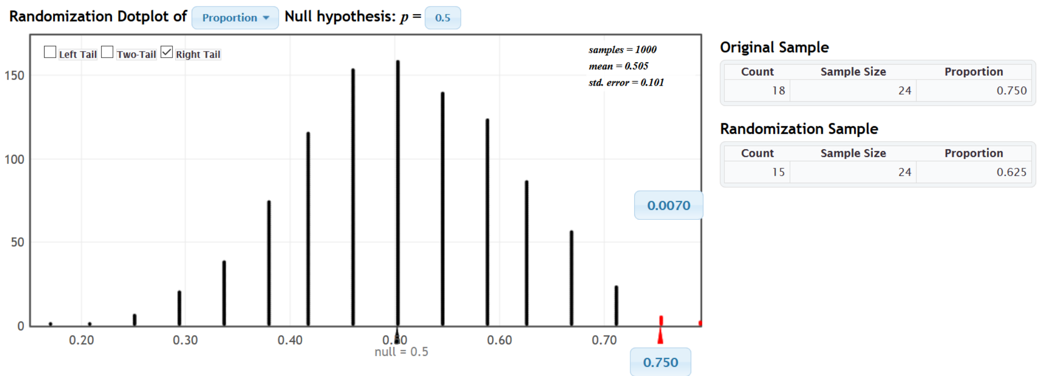
The normal approximation should only be used when $n\hat{p}$ and $n(1-\hat{p})$ are both greater than or equal to 10. Since $n(1-p_0) = 24(1-0.50) = 12 > 10$, we should use the normal distribution approximation.

d) Find a p-value using a method appropriate for this data situation.

Using the test for z-test for a single proportion we compute the following test statistic:

$$z_{test} = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = 2.45$$

Since this is a one-sided test, we obtain a right-tailed p-value of 0.0071.



e) Make a conclusion in context, using $\alpha = 0.05$.

We reject the null hypothesis. There is sufficient evidence to conclude that babies do not choose the two jar types with equal preference. Rather, babies tend to choose the jar with the greater proportion of their preferred color more often. This may suggest that babies do think probabilistically.

6.76 [5 pts] Compute the standard error for sample means from a population with mean $\mu = 100$ and standard deviation $\sigma = 25$ for sample sizes of $n = 30$, $n = 200$, and $n = 1000$. What effect does increasing the sample size have on the standard error? Using this information about the effect on the standard error, discuss the effect of increasing the sample size on the accuracy of using a sample mean to estimate a population mean.

To compute the standard error, we divide the standard deviation by the square root of the sample size. Doing so for each of the provided sample sizes, we obtain standard errors of 4.56, 1.77, and 0.79. Clearly, increasing the sample size leads to a decrease in the standard error. This immediately implies that larger samples correspond to more accurate estimation of the population mean through the sample mean.

6.107 [3 pts] Plastic microparticles are contaminating the world's shorelines, and much of this pollution appears to come from fibers from washing polyester clothes. The worst offender appears to be fleece, and a recent study found that the mean number of polyester fibers discharged into wastewater from washing fleece was 290 fibers per liter of wastewater, with a standard deviation of 87.6 and a sample size of 120.

a) Find and interpret a 99% confidence interval for the mean number of polyester microfibers per liter of wastewater when washing fleece.

The 99% confidence interval is:

$$290 \pm 2.58\frac{87.6}{120} = (269.40, 310.60)$$

We are 99% confident that the mean number of polyester microfibers per liter of wastewater when washing fleece is at least 269.40 and at most 310.60.

b) What is the margin of error?

The margin of error is $2.58\frac{87.6}{\sqrt{120}} = 20.60$

c) If we want a margin of error of only ± 5 with 99% confidence, what sample size is needed?

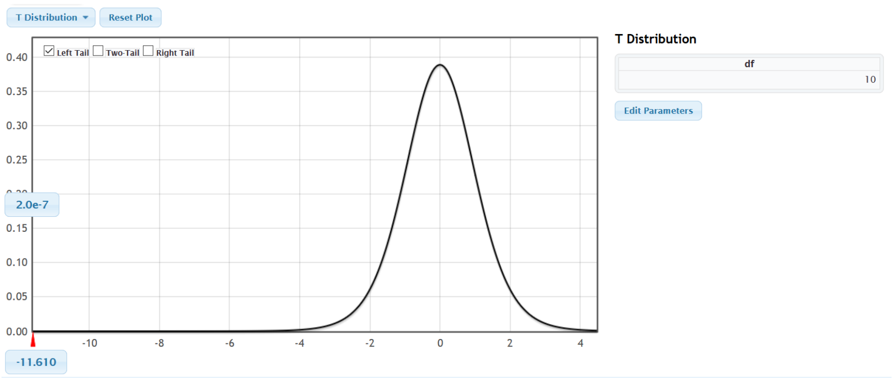
We would need a sample size of

$$n = \left(\frac{2.58 * 87.6}{5}\right)^2 \approx 2044$$

6.124 [2 pts] During the National Football League's 2014 AFC championship game, officials measured the air pressure on 11 of the game footballs being used by the New England Patriots. They found that the balls had an average air pressure of 11.1 psi, with a standard deviation of 0.40 psi.

a) Assuming this is a representative sample of all footballs used by the Patriots in the 2014 season, perform the appropriate test to determine if the average air pressure in footballs used by the Patriots was significantly less than the allowable limit of 12.5 psi. There is no extreme skewness or outliers in the data, so it is appropriate to use the t-distribution.

We compute the test statistic, $t_{test} = \frac{11.1-12.5}{0.40/sqrt11} = -11.61$, and then compute a p-value of 0.0000002 using the t-distribution with 10 degree of freedom shown below.



b) Is it fair to assume that this sample is representative of all footballs used by the Patriots during the 2014 season?

This is probably not a fair assumption. There are several games played during the season, each with varying conditions (weather probably being one relevant example).

6.145 [2 pts] Errors in medical prescription occur, and a study examined whether electronic prescribing may help reduce errors. Two groups of doctors used written prescriptions and had similar error rates before the study. One group switched to e-prescriptions while the other continued to use written prescriptions, and error rates were measured one year later. The results are given in Table 6.4. Find and interpret a 95% confidence interval for the difference in proportion of errors between the two groups. Is it plausible that there is no difference?

Table 6.4

	Error	**No Error**	**Total**
Electronic	254	3594	3848
Written	1478	2370	3848

Given the large "cell" counts (i.e each table entry count), we may use the normal approximation to construct the 95% confidence interval for the difference in proportions:

$$(0.07 - 0.38) \pm 1.96 * \sqrt{\frac{0.07(1-0.07)}{3848} + \frac{0.38(1-0.38)}{3848}} = (-0.33, -0.29)$$

We are 95% confident that the difference in error rate between electronic and written prescriptions is at least -0.33 and at most -0.29. Given this confidence interval, it is not plausible to conclude that there is no difference in error rate between written and electronic prescriptions.

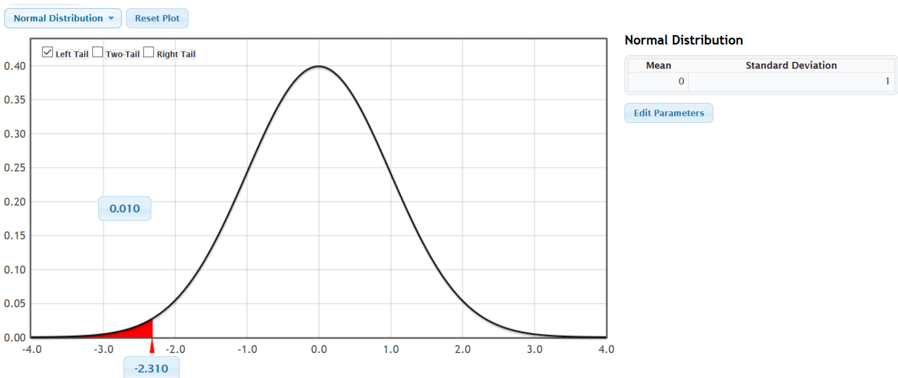
6.173 [6 pts] Are malaria parasites able to control mosquito behavior to their advantage? A study investigated this question by taking mosquitoes and giving them the opportunity to have their first "blood meal" from a mouse. The mosquitoes were randomized to either eat from a mouse infected with malaria or an uninfected mouse. At several time points after this, mosquitoes were put into a cage with a human and it was recorded whether or not each mosquito approached the human (presumably to bite, although mosquitoes were caught before biting). Once infected, the malaria parasites in the mosquitoes go through two stages: the Oocyst stage in which the mosquito has been infected but it is not yet infectious to others and then the Sporozite stage in which the mosquito is infectious to others. Malaria parasites would benefit if mosquitoes sought risky blood meals (such as biting a human) *less* often in the Oocyst stage (because mosquitoes are often killed while attempting a blood meal) and *more* often in the Sporozite stage after becoming infectious (because this is one of the primary ways in which malaria is transmitted). Does exposing mosquitoes to malaria actually impact their behavior in this way?

a) In the Oocyst stage (after eating from mouse but before becoming infectious), 20 out of 113 mosquitoes in the group exposed to malaria approached the human and 36 out of 117 mosquitoes in the group not exposed to malaria approached the human. Calculate the z-statistic.

$$z_{test} = \frac{(\frac{20}{113} - \frac{36}{117}) - 0}{\sqrt{\frac{\frac{20+36}{113+117}(1-\frac{20+36}{113+117})}{113} + \frac{\frac{20+36}{113+117}(1-\frac{20+36}{113+117})}{117}}} = -2.31$$

b) Calculate the p-value for testing whether this provides evidence that the proportion of mosquitoes in the Oocyst stage approaching the human is lower in the group exposed in malaria.

The p-value is 0.01.

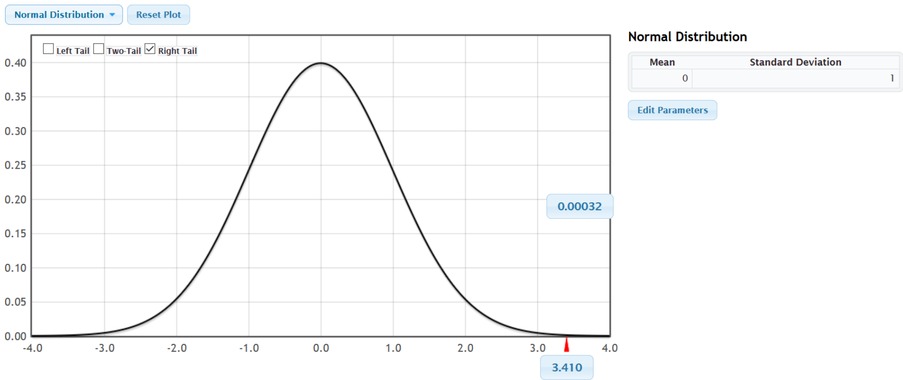


c) In the Sporozite stage (after becoming infectious), 37 out of 149 mosquitoes in the group exposed to malaria approached the human and 14 out of 144 mosquitoes in the group not exposed to malaria approached the human. Calculate the z-statistic.

$$z_{test} = \frac{(\frac{37}{149} - \frac{14}{144}) - 0}{\sqrt{\frac{\frac{37+14}{149+144}(1-\frac{37+14}{149+144})}{149} + \frac{\frac{37+14}{149+144}(1-\frac{37+14}{149+144})}{144}}} = 3.41$$

d) Calculate the p-value for testing whether this provides evidence that the proportion of mosquitoes in the Sporozite stage approaching the human is higher in the group exposed to malaria.

The p-value is 0.00032.



e) Based on your p-values, male conclusions about what you have learned about mosquito behavior, stage of infection, and exposure to malaria or not.

In the Oocyst stage, we may conclude that mosquitoes exposed to malaria are less likely to approach humans than those not exposed to malaria. The reverse is true in the Sporozite stage.

f) Can we conclude that being exposed to malaria (as opposed to not being exposed to malaria) *causes* these behavior changes in mosquitoes? Why or why not?

Given that this was a randomized experiment, we may conclude that these relationships are causal.

6.195 [3 pts] A study examines chocolate's effects on blood vessel function in healthy people. In the randomized, double-blind, placebo-controlled study, 11 people received 46 grams (1.6 ounces) of dark chocolate (which is naturally flavonoid-rich) every day for two weeks, while a control group of 10 people received a placebo consisting of dark chocolate with low flavonoid content. Participants had their vascular health measured (by means of flow-mediated dilation) before and after the two-week study. The increase over the two-week period was measured, with larger numbers indicating greater vascular health. For the group getting the good dark chocolate, the mean increase was 1.3 with a standard deviation of 2.32, while the control group had a mean change of -0.96 with a standard deviation of 1.58.

a) Explain what "randomized, double-blind, placebo-controlled study" means.

A randomized, double-blind, placebo-controlled study is one in which there are at least two treatment groups, one being a control group that serves as a comparator to the remaining treatment group. The control uses a placebo, hence "placebo-controlled". This study design is randomized and double-blinded if the treatment groups are randomly allocated to participants (randomized) and neither the participants nor participating study staff are aware of the treatment assignment (double-blinded).

b) Find and interpret a 95% confidence interval for the difference in means between the two groups. Be sure to clearly define the parameters you are estimating. You may assume that neither sample shows significant departures from normality.

Using the conservative approach for computing the degrees of freedom, the 95% confidence interval is:

$$(1.3 - (-0.96)) \pm 2.26\sqrt{\frac{2.32^2}{11} + \frac{1.58^2}{10}} = (0.32, 4.20)$$

We are then 95% confident that the difference in mean flow-mediated dilation between the dark chocolate and placebo groups is at least 0.32 and 4.20. The parameter we are estimating with this interval is $\mu_{dc} - \mu_p$, where μ_{dc} and μ_p refer to the mean flow-mediated dilation in the dark chocolate and placebo groups respectively.

c) Is it plausible that there is "no difference" between the two kinds of chocolate? Justify your answer using the confidence interval found in (b).

Since the interval we found does not contain 0, it is not plausible that there is "no difference" between the two kinds of chocolate.

6.221 [6 pts] Drinking tea appears to offer a strong boost to the immune system. IN a study introduced in Exercise 3.91 on page 239, we see that production of interferon gamma, a molecule that fights bacteria, viruses, and tumors, appears to be enhanced in tea drinkers. In the study, eleven healthy non-tea-drinking individuals were asked to drink five or six cups of tea a day, while ten healthy non-tea- and non-coffee-drinkers were asked to drink the same amount of coffee, which has caffeine but not the L-theanine that is in tea. The groups were randomly assigned. After two weeks, blood samples were exposed to an antigen and production of interferon gamma was measured. The results are shown in Table 6.19 and are available in ImmuneTea. The question of interest is whether the data provide evidence that production is enhanced in tea drinkers.

Table 6.19

Tea	Coffee
5	0
11	0
13	3
18	11
20	15
47	16
48	21
52	21
55	38
56	42
58	

a) Is this an experiment or an observational study?

This is a randomized experiment.

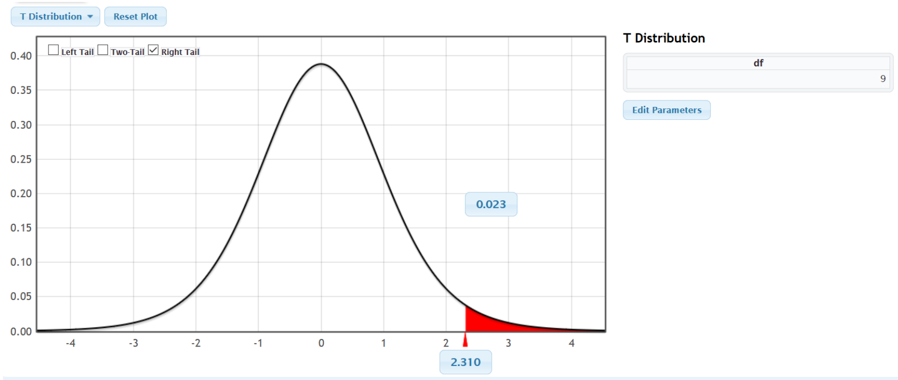
b) What are the null and alternative hypotheses?

The null hypothesis is that the difference in mean interferon gamma concentration between tea drinkers and coffee drinkers is 0. The alternative hypothesis is that this difference is positive.

c) Find a standardized test statistic and use the t-distribution to find the p-value and make a conclusion.

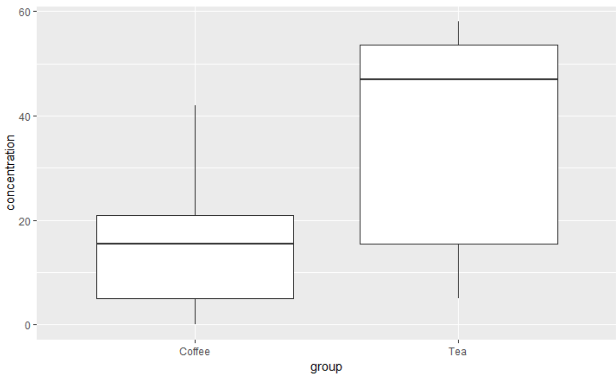
$$t_{test} = \frac{(34.82 - 16.70) - 0}{\sqrt{\frac{21.08^2}{11} + \frac{14.58^2}{10}}} = 2.31$$

Using this test statistic, and using the conservative approach, we obtain a p-value of 0.026. We reject the null hypothesis and conclude that there is a difference in interferon gamma production between the two groups, with tea drinkers having increased production.



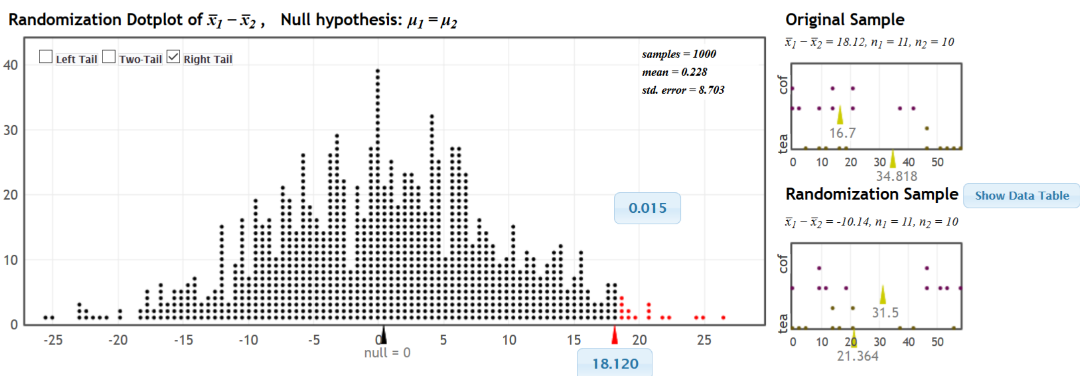
d) Always plot your data! Look at a graph of the data. Does it appear to satisfy a normality condition?

Plotting boxplots for each group, we obtain the following. These boxplots indicate our data is likely not normal. With tea drinkers in particular, the distribution is severely asymmetric. The coffee drinker data is also right skewed.



e) A randomization test might be a more appropriate test to use in this case. Construct a randomization distribution for this test and use it to find a p-value and make a conclusion.

Performing a randomization test, we obtain a p-value of 0.015. We reject the null hypothesis and conclude that there is a difference in interferon gamma production between the two groups, with tea drinkers having increased production.



f) What conclusion can we draw?

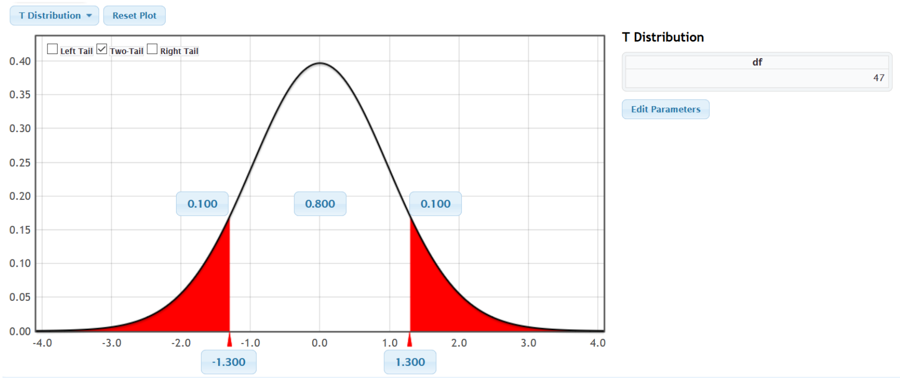
While both tests lead to the same conclusion, we should rely on the randomization test results in order to draw our conclusion. As such, may conclude that, on average, the interferon gamma production among tea drinkers is greater than among coffee drinkers.

6.255 [3 pts] As part of the same study described in Exercise 6.254, the researchers were also interested in whether babies preferred singing or speech. Forty-eight of the original fifty infants were exposed to both singing and speech by the same woman. Interest was again measured by the amount of time the baby looked at the woman while she made noise. In this case the mean time while speaking was 66.97 with standard deviation of 43.42, and the mean for singing was 56.58 with a standard deviation of 31.57 seconds. The mean of the differences was 10.39 more seconds for the speaking treatment with a standard deviation of 55.37 seconds. Perform the appropriate test to determine if this is sufficient evidence to conclude that babies have a preference (either way) between speaking and singing.

These data are paired. Each baby was exposed to both "treatments", singing and speech. We use the mean and standard deviation of the differences in attention time in order to compute our test statistic and perform a test:

$$t_{test} = \frac{10.39 - 0}{55.37/\sqrt{48}} = 1.30$$

We obtain a p-value of 0.2. Given this p-value we fail to reject the null hypothesis of no preference among babies between hearing speech and singing.



6.257 [3 pts] Table 6.25 gives a sample of grades on the first two quizzes in an introductory statistics course. We are interested in testing whether the mean grade on the second quiz is significantly higher than the mean grade on the first quiz.

Table 6.25

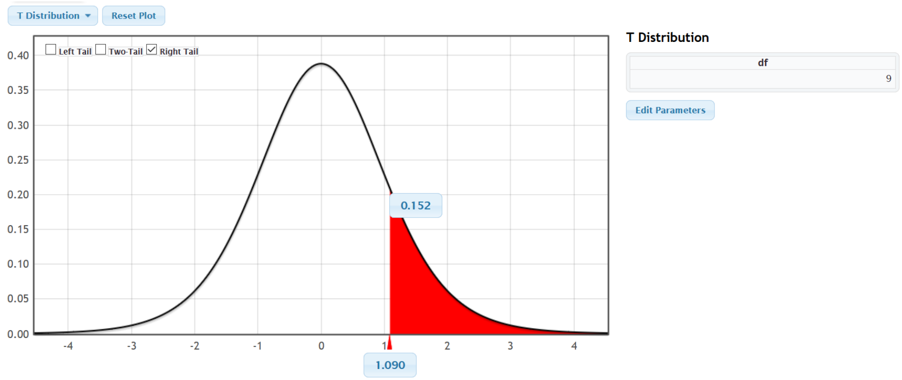
First Quiz	Second Quiz
72	78
95	96
56	72
87	89
80	80
98	95
74	86
85	87
77	82
62	75

a) Complete the test if we assume that the grades from the first quiz come from a random sample of 10 students in the course and the grades on the second quiz come from a different separate random sample of 10 students in the course. Clearly state the conclusion.

Treating the data as if from two independent samples,

$$t_{test} = \frac{(84 - 78.6) - 0}{\sqrt{\frac{13.38^2}{10} + \frac{8.06^2}{10}}} = 1.09.$$

We obtain a p-value of 0.152. We fail to reject the null hypothesis of no difference in average score between the two quizzes.

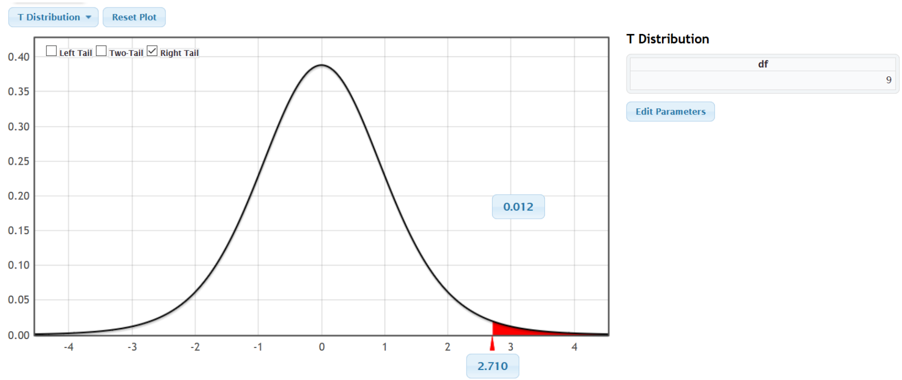


b) Now conduct the test if we assume that the grades recorded for the first quiz and the second quiz are from the same 10 students in the same order. (So the first student got a 72 on the first quiz and a 78 on the second quiz.)

Treating the data as paired (which they are),

$$t_{test} = \frac{5.4}{6.29/\sqrt{10}} = 2.71.$$

We obtain a p-value of 0.012. We reject the null hypothesis of no difference in average score between the two quizzes. The data suggest that, on average, students achieved higher scores on their second quiz.



c) Why are the results so different? Which is a better way to collect the data to answer the question of whether grades are higher on the second quiz?

The difference in results is attributed to the difference in variability. Treating the data as arising from two independent samples, the variability is substantially larger than the variability of the differences. The better way to collect data to answer the question of whether grades are higher on the second quiz is clearly the paired approach. The variability is substantially lower, resulting in lower standard errors and more powerful inference.