**Statistics 200: Second Lab Activity for Section 4.3**

**Note: This lab is round two for this section about performing hypothesis tests. Nothing new is introduced, but students need to reach proficiency in every step of the process. Students may continue to struggle with identifying the parameter of interest and writing the null/alternative hypotheses, but hopefully they are getting a handle on that by now. You need to continue reinforcing the important facts...**

* **the randomization distribution is generated assuming the null hypothesis is true**
* **p- values measure the evidence against the null and in favor of the alternative**
* **smaller p-values indicate more evidence against the null and in favor of the alternative**
* **Conclusion in context means they need to say something about the strength of the evidence and relate it back to the actual question of interest.**

**TAs and LAs need to know how to do all things in StatKey, including identify and use the original sample statistic.**

**Determining Statistical Significance - Learning objectives:**

* Recognize that smaller p-values give stronger evidence in support of the alternative hypothesis
* Demonstrate an understanding of the concept of statistical significance
* Make a formal decision in a hypothesis test by comparing a p-value to the significance level
* State the conclusion to a hypothesis test in context
* Make a less formal decision that reflects the strength of evidence in a p-value
* Conduct a hypothesis test for a variety of situations

For this lab we will use *NFL Scouting Combine* data for drafted running backs and wide receivers from years 2012-2014. The combine is a series of tests to evaluate college football players ahead of the *NFL Draft*. The dataset is available on Canvas. To perform the hypothesis tests below, you will need to upload the data to StatKey and select the appropriate columns. For guidance, refer to the StatKey guide.

The dataset contains the following variables:

year Year player participated in combine

position Running Back (RB) or Wide Receiver (WR)

height height in inches

weight weight in pounds

fortyyd forty yard dash time in seconds

threecone threecone time in seconds (run between 3 cones in an L shape)

vertical vertical jump height in inches

broad broad jump distance in inches

bench number of bench press repetitions with 225lbs

round round player was drafted (1-7)

**Activity 1: *Strength and Draft Round***

Is there a linear association between a player’s strength, measured using the *bench* variable, and the round he was drafted in?

1. State the hypotheses of interest.
2. What is the notation and value of the sample statistic?
3. Use StatKey to generate a randomization distribution for these hypotheses. Remember, you will have to upload the dataset to StatKey and select the correct variables. What is the p-value?
4. Complete the p-value interpretation below:

If there is no/zero linear association between strength and draft round in the population, the chance of seeing a sample correlation of -0.059 or more extreme is 0.564.

1. What is the formal conclusion at a significance level of 0.05? Because the p-value is greater than 0.05, we cannot reject the null hypothesis at a significance level of 0.05.
2. What is the conclusion in context? We do not have significant evidence that stronger running backs and wide receivers are drafted earlier.

**Activity 2*: How fast are NFL players?***

Prior to 2012, the average forty yard dash time for running backs and wide receivers was 4.53. On average, were NFL players between 2012-2014 significantly faster than in previous years?

1. State the hypotheses of interest.
2. What is the notation and value of the sample statistic?
3. Use StatKey to generate a randomization distribution for these hypotheses. What is the p-value?
4. Interpret the p-value. If the true mean forty yard dash time was 4.53, then the chance of seeing a sample mean of 4.511 or less is 0.034.
5. What is the formal conclusion and conclusion in context using a significance level of 0.05? Since the p-value is less than 0.05, we reject the null hypothesis at a significance level of 0.05. We have significant evidence that running backs and receivers in 2012-2014 were faster than in previous years.

**Activity 3: Are running backs shorter than wide receivers?**

Assuming that running backs are group 1 and wide receivers are group 2, determine if, on average, running backs are shorter than wide receivers.

1. State the hypotheses of interest.
2. What is the notation and value of the sample statistic?
3. Use StatKey to generate a randomization distribution for these hypotheses. What is the p-value? or the pvalue is less than 1/bootstrap sample size, which is 0.002 for 5000 or 0.0001 for 10000.
4. Interpret the p-value. If the running backs and receivers were really the same height on average, then the probability of seeing a difference in height of -2.55 inches is essentially zero (or less than 0.002/0.001 etc).
5. What is the formal conclusion and conclusion in context using a significance level of 0.05? At a significance level of 0.05 we reject the null hypothesis in favor of the alternative and conclude that there is very strong evidence that running backs are shorter on average than wide receivers.