This statistics worksheet examines the goals and shots in two prominent lacrosse leagues: the Premier Lacrosse League (PLL) and the National Lacrosse League (NLL). The PLL and NLL are highly regarded professional lacrosse leagues that feature top-tier athletes from around the world. The NLL is a box lacrosse league that is played in an indoor, smaller field while the PLL is a field lacrosse league that plays on an outdoor field with larger dimensions. This worksheet will provide an analysis of goal-scoring within these leagues. The data is collected from the PLL and NLL website for all games in the 2021-2022 season and is stored in the file **lacrosse\_pll\_nll\_2021-2022.csv**. The *League* variable codes whether each game was played in the NLL or PLL and the total goals scored (both teams) is in *goals*.

**Descriptive Statistics for Goals:**



|  |  |  |  |
| --- | --- | --- | --- |
| **League** | **N** | **Mean** | **StDev** |
| NLL | 126 | 22.262 | 4.847 |
| PLL | 36 | 23.889 | 4.496 |

1. Suppose you are interested in using these data to determine if there is statistically discernible evidence that, on average, goals scored in the PLL and NLL differ.
	1. Record the null and alternative hypotheses associated with the research question. Be sure to identify any parameters in context.

$H\_{0}:μ\_{PLL}=μ\_{NLL}$ where $μ\_{PLL}$ and $μ\_{NLL}$ are the mean goals per game rate in the PLL and NLL, respectively

$$H\_{a}:μ\_{PLL}\ne μ\_{NLL}$$

* 1. Construct a randomization distribution for doing this test. Draw a rough sketch of this distribution below, include a scale on the horizontal axis, and indicate how the p-value is found.



p-value=2(0.038)=0.076

* 1. Using the p-value, provide an appropriate conclusion for the research question. Be sure to provide full context of the application.

With a p-value of 0.076 we would not reject the null hypothesis at a 5% level but would reject it at a 10% level. At best, there is weak evidence to show a difference in mean goals between NLL and PLL games.

Upon further analysis of the leagues, you discover that the NLL and the PLL play their games for different amounts of time. The PLL plays games for 48 minutes, while the NLL’s games are 60 minutes long. In an effort to correct this mistake, we will “scale down” the NLL goals to a 48 minute rate. This variable is stored in *Goals\_per\_48*.

1. What is the difference in means (NLL – PLL) for *Goals\_per\_48* in this sample?

$$\overbar{x}\_{NLL}-\overbar{x}\_{PLL}=17.81-22.89=-6.09$$

1. Do the data provide evidence that, on average, goals scored per 48 minutes differ between the PLL and NLL? Include all the pieces of the appropriate hypothesis test. (Tip: Use a similar process to Question 1)

$H\_{0}:μ\_{PLL}=μ\_{NLL}$ where $μ\_{PLL}$ and $μ\_{NLL}$ are now the mean goals per 48 in the PLL and NLL, respectively

$$H\_{a}:μ\_{PLL}\ne μ\_{NLL}$$



None of the randomization samples have a difference anywhere near as large as 6.08 in either direction, so the p-value≈0.

We have very strong evidence that the mean goals per 48 minutes is lower in the NLL than in the PLL.

1. In many scenarios, it is useful to extend a conclusion of a hypothesis test by including a confidence interval with the results. Calculate and interpret a 95% confidence interval for the difference in average goals per 48 minutes for NLL and PLL lacrosse games. (Tip: Your interpretation should indicate the direction as well as the size of the difference.)



Using the middle 95% of a bootstrap distribution for the difference in mean goals per 48 minutes, we are 95% sure that the mean goals per 48 minutes for the NLL is between 7.68 and 4.47 less than the PLL.

Since this interval has only negative values, we have strong evidence that the mean goals per 48 minutes is lower for the NLL than for the PLL.