Activity #24: Independent Samples t-test examples

Scenario: There are two types of baseball fans: those who like the DH (designated hitter) rule and those who hate it. DH-supporters argue that games are more exciting with a designated hitter because there are no "easy outs" (the pitcher never has to bat). These people like the American League, which has the DH rule. DH-opponents argue that the DH rule takes strategy out of the game (sending in pinch hitters for the pitcher). These people support the National League, which doesn't have the DH rule.

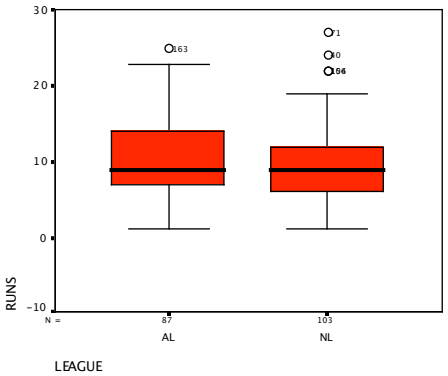
MLB Commissioner Bud Selig has hired you to investigate the impact the DH rule has on the game of baseball. Because fans like high-scoring games, he wants to know if the American League (DH) averages more runs per game than the National League (no-DH). You collect a sample of data from 190 baseball games in 1999 in order to complete this study.

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

2) Express the consequences of Type I and Type II errors for this study. Which error do you want most to avoid?

3) From your sample of 190 games, you calculate the following sample statistics. Sketch two boxplots on the same set of axes to compare these two distributions. Which league appears to have a higher number of runs per game? Does it appear as though the two "treatments" yielded the same variance?

League	n	$\overline{X}$	s	Min	Q25	Median	Q75	Max
AL	83	10.49	4.868	1	7	9	14	25
NL	107	9.65	4.785	1	6	9	12	27



4) The null hypothesis states that the two leagues do not differ in the average number of runs per game. If this hypothesis is true, what would the distribution of the difference between sample mean runs-per-game look like? Calculate the standard error of this sampling distribution and sketch a picture of it.

5) Calculate the value of our observed test statistic. Locate this observed test statistic on the distribution.

6) Using a significance level of 0.10, find the critical value of the test statistic from your table. Shade this area on your distribution.

7) What is your conclusion?

8) SPSS computed the p-value to be p = 0.12. Express what this represents.

9) Create a 90% confidence interval for the difference in average runs scored for the AL and NL. Express what it represents.

10) After hearing your report, Bud Selig decided to give you another assignment. This time, you'll check to see if the DH rule increases or decreases the length of a baseball game. Formulate the competing hypotheses for this study.

11) From your sample of 190 games, you calculate the following sample statistics. Which league appears to have a higher number of runs per game? Does it appear as though the two "treatments" yielded the same variance?

League	n	$\overline{X}$	s
AL	83	178.53	23.159
NL	107	170.98	25.469
	Difference	7.55	

12) Calculate a 95% confidence interval for the difference in average game lengths.

13) Make your conclusion (explaining how you made your decision).