Confidence Intervals

1. Recall the class survey. Thirteen female and twenty-four male students filled out the survey, reporting (among other variables) their GMAT scores and interest levels in the course. We will use this data to compare females and males.

   (a) What are the relevant populations?

   (b) For the 7 female respondents who reported their GMAT scores, the mean was 671 and the standard deviation was 27. For the 15 male respondents, the mean was 682 and the standard deviation was 38. Find a 95% confidence interval for the difference in population means.

   (c) For the 13 female respondents who reported their interest levels in the course (1–10), the mean was 6.2 and the standard deviation was 2.5. For the 24 male respondents, the mean was 6.5 and the standard deviation was 2.1. Find a 95% confidence interval for the difference in population means.

   (d) For the confidence intervals you constructed in parts (b) and (c) to be valid, what assumptions need to be satisfied? How could you check these assumptions?

Hypothesis Tests

2. Consider again the class survey data. We will use the data to evaluate whether or not there is a significant difference between the female and the male population means.

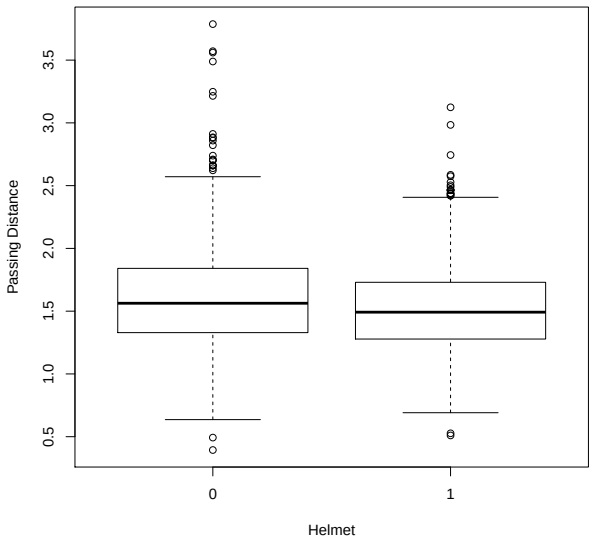
   (a) For the 7 female respondents who reported their GMAT scores, the mean was 671 and the standard deviation was 27. For the 15 male respondents, the mean was 682 and the standard deviation was 38. If the population means were equal what would be the chance of seeing a difference in sample means as large as observed?

   (b) For the 13 female respondents who reported their interest levels in the course (1–10), the mean was 6.2 and the standard deviation was 2.5. For the 24 male respondents, the mean was 6.5 and the standard deviation was 2.1. If the population means were equal what would be the chance of seeing a difference in sample means as large as observed?

   (c) What is the relationship between the confidence intervals in Question 1 and your answers to parts (a) and (b)?

Case Study: Bicycle Passing Distance

3. Here are boxplots of the passing distances (in meters) for a bike rider with and without a helmet. Is there evidence that the passing distance differs when the rider has a helmet?



Here are the sample statistics for the passing distance without a helmet: $n_1 = 1206$, $\bar{x}_1 = 1.61$, $s_1 = 0.405$. Here are the sample Here are the sample statistics for the passing distance with a helmet: $n_2 = 1149$, $\bar{x}_2 = 1.52$, $s_2 = 0.354$.

Formulate the problem as a hypothesis test, using significance level 5%.

(a) What are the populations?

(b) What are the null and alternative hypotheses?

(c) What are the samples?

(d) What is the test statistic?

(e) Approximately what is the p-value and the result of the test?

(f) Find a 95% confidence interval for the difference in passing difference with and without a helmet.