



Independent Samples T-test

1. Intro and Data Setup
 - a. If you've watched our video that explains the concept of an independent samples t-test, you may recall that an independent-samples t-test helps you figure out when you really do have one group of people in the data who are acting like apples, and a different group of people in the data who are acting like oranges. Or if the groups are acting like apples of different shades (and so really versions of the same thing).
 - b. In this example, we are going to take a look at average Overall Rating of the Festival by binary Gender, using two groups for gender –female and male. We want to determine if females and males rated the festival very differently, acting like apples and oranges, or not so differently, acting more like apples of different shades.
 - c. The Overall Rating question asked festival attendees, “How would you rate today's event overall?” Festival goers could choose an answer on a 5-point scale from Poor to Excellent. Here, Gender is the predictor variable – we want to know if Gender predicts differences in Overall Rating. Overall Rating is our outcome variable – it is the variable that we want to learn more about.
 - d. We will use the Overall by Gender tab of the EvalFest Database for Excel.xlsx file to analyze the data. Column A contains the Overall Festival Rating data for females, and Column B contains the Overall Festival Rating data for males.
2. Conducting the independent samples t-test
 - a. You will need to have the Data Analysis ToolPak loaded in Excel. Click on the Data tab, and if you don't see the Data Analysis icon on the top right, you will need to load this ToolPak.
 - i. To add it, go to Tools and click on Excel Add-ins...
 - ii. Check Analysis Toolpak and click OK.
 - iii. When you restart Excel, you should see the Data Analysis icon within the Data tab.
 - b. Click on Data Analysis.
 - c. Click on t-test: Two Sample Assuming Unequal Variances and click OK.
 - d. Click in the Variable 1 Range box and highlight all of the data in the OverallRatingFemale column.
 - e. Next, click in the Variable 2 Range box and highlight all of the data in the OverallRatingMale column. Assigning ratings of Females as Variable 1 and ratings of Males as Variable 2 tells Excel that these are the two groups of

- ratings you want to compare.
- f. Next, if you want the output in the same worksheet, click on the Output Range radio button and choose where in the current worksheet you want the output. If you don't do this, the default is to place the output in a different worksheet.
 - i. You only need to choose the cell in which you want the output to begin.
 - g. Click OK.
3. Interpretation of the Independent Samples T-test Output
- a. Excel will provide some output – these are your results. Excel provides the value of t, the degrees of freedom, and two p-values, among other things.
 - b. The t-test uses the average Festival rating for each group, and the variation in those ratings, to figure out if there are differences in the ratings of Male and Females (acting like apples and oranges), or if the Festival ratings of Males and Females are pretty similar (acting like apples of different shades).
 - c. If you've watched the conceptual video, you may recall that the p-value is a measure of the risk you take in making a decision based on your data. If the p-value is less than .05, that means there is less than 5% chance that the results are based on something other than the data. Thinking about this another way, there is a 1 minus the p-value chance that you DO have differences in ratings between Males and Females, and they are acting like apples and oranges.
 - d. In more statistical terms, you use the p-value to determine if the mean of the outcome variable is statistically significantly different between the groups in the predictor variable.
 - e. In the output, you're most interested in the p-value for a two-tail test. For this scenario, we can see that our p value is 5.46×10^{-8} , or .0000000546. That number is less than .05 and so that means that we have a statistically significant difference in the average ratings provided by Female and Male attendees.
 - f. Keep in mind that, just because a result is statistically significant, just because you found a p-value of less than .05, you still have to decide whether or not the result has *practical implications* for your program. Looking at the mean ratings for Males and Females, we see that Females had a higher mean rating than Males, at 4.44 vs. 4.33. This is a difference of 0.11 points on a 5-point scale, or a 2.2% difference in average rating. Is this difference meaningful for your program? Maybe, or maybe not – this is something you will have to decide.
 - g. So now what? How do you think about interpreting these results? Remember that these ratings were made on a five-point scale, where 1 was Poor and 5 was Excellent. Both of our average ratings were above 4.0, and so somewhere between Very Good and Excellent on the scale. If you have statistical significance and the results are meaningful for your program, you want to think about the kinds of ratings you have overall and the group differences.
 - h. **Slide 9:** So, to interpret these results, we might say something like: Ratings for the Festival were quite positive. There was a statistically significant difference in how males and females rated the event. Males rated the Festival between Very Good and Excellent, on average, with a rating of 4.33. Female ratings

were also in the Very Good to Excellent range, but were statistically significantly higher, and thus even closer to the top of the scale at 4.44.

- i. We used the two-tailed p-value to determine if there was a statistically significant difference in overall ratings between males and females. That is, we were looking for differences in either direction – females greater than males or vice versa. Once we found a statistically significant difference, we then went back to the mean for females and the mean for males to determine which was greater, and if the results had practical implications.
- j. If you want to conduct a one-tail, less than test (in this scenario, you believe that the mean overall rating for males is statistically significantly less than the mean overall rating for females), you will use the one-tail p-value.
- k. If you want to conduct a one-tail, greater than test (in this scenario, you believe that the mean overall rating for males is statistically significantly greater than the mean overall rating for females), your p-value will be 1 minus the one-tail p-value.
- l. Determining whether to use the two-tailed p-value, one-tailed p-value, or 1 minus the one-tailed p-value depends on the question you are trying to answer.

