Worked-Through Example

Independent Samples *t*-Test: Differences on Exercise (in hours) by Gender

***Introduction***

A two-tailed independent samples *t*-test was conducted to examine whether the mean of Exercise (hours) was significantly different between the Male and Female categories of Gender.

***Assumptions***

**Normality.** Shapiro-Wilk tests were conducted to determine whether Exercise(hours) could have been produced by a normal distribution for each category of Gender (Razali & Wah, 2011). The result of the Shapiro-Wilk test for Exercise(hours) in the Male category was not significant based on an alpha value of 0.05, *W* = 0.95, *p* = .134. This result suggests that a normal distribution cannot be ruled out as the underlying distribution for Exercise(hours)in the Male category. The result of the Shapiro-Wilk test Exercise(hours) in the Female category was not significant based on an alpha value of 0.05, *W* = 0.93, *p* = .126. This result suggests that a normal distribution cannot be ruled out as the underlying distribution for Exercise(hours) in the Female category. The Shapiro-Wilk test was not significant for either the Male or Female categories of Gender, indicating the normality assumption is met.

**Homogeneity of Variance.** Levene's test was conducted to assess whether the variance of Exercise(hours) was equal between the categories of Gender. The result of Levene's test for Exercise(hours) was not significant based on an alpha value of 0.05, *F*(1, 54) = 0.14, *p* = .709. This result suggests it is possible that the variance of Exercise(hours) is equal for each category of Gender, indicating the assumption of homogeneity of variance was met.

***Results***

The result of the two-tailed independent samples *t*-test was significant based on an alpha value of 0.05, *t*(54) = -8.29, *p* < .001, indicating the null hypothesis can be rejected. This finding suggests the mean of Exercise(hours) was significantly different between the Male and Female categories of Gender. The results are presented in Table 1. A bar plot of the means is presented in Figure 1.

**Table 1**

*Two-Tailed Independent Samples t-Test for Exercise(hours) by Gender*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|   | Male | Female |   |   |   |
| Variable | *M* | *SD* | *M* | *SD* | *t* | *p* | *d* |
| Exercise\_hours | 1.84 | 0.59 | 3.20 | 0.64 | -8.29 | < .001 | 2.22 |

*Note.* N = 56. Degrees of Freedom for the *t*-statistic = 54. *d* represents Cohen's *d.*

**Figure 1**

*The mean of Exercise(hours)by levels of Gender*



**References**

Intellectus Statistics [Online computer software]. (2020). Intellectus Statistics. https://analyze.intellectusstatistics.com/

Razali, N. M., & Wah, Y. B. (2011). Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. *Journal of Statistical Modeling and Analytics, 2*(1), 21-33.

**Glossaries**

**Independent Samples *t*-Test**

The independent samples *t*-test is used to determine if there is a significant difference between two groups (e.g., men vs. women) on a scale-level dependent variable. This test uses the difference between the average scores of the two groups to compute the *t* statistic, which is used with the *df* to compute the *p*-value (i.e., significance level). A significant result indicates the observed test statistic would be unlikely under the null hypothesis. The independent samples *t*-test carries the assumptions of independence of observations, normality, and equality (or homogeneity) of variance. ***Fun Fact!*** *William Sealy Gosset, who published a paper about the t distribution in 1908, worked for the Guinness Brewery in Dublin, Ireland.* **Cohen's *d*:** Effect size for the *t*-test; determines the strength of the differences between the matched scores. The larger the effect size, the greater the differences in the matched scores. **Degrees of Freedom (*df*):** Refers to the number of values used to compute a statistic. The *df* is determined by the number of observations in the sample and equal the number of observations - 1; used with *t* to compute the *p*-value. **Levene's Test:** Test to assess if the assumption of equality of variance is met; if significance is found, the groups differ in their spread of the dependent variable scores; this may differ from the output found from other statistical packages (such as SPSS), as Intellectus Statistics™ uses the median instead of the mean for calculations; the median tends to provide a more-robust choice that can account for non-normality. **Mean (*M*):** The average value of a scale-level variable. **Normality:** Refers to the distribution of the data. The assumption is that the data follows the bell-shaped curve. ***p*-value:** The probability of obtaining the observed results if the null hypothesis is true. A result is usually considered statistically significant if the *p*-value is ≤ .05. **Shapiro-Wilk Test:** A test to assess if the assumption of normality is met. If statistical significance is found in this test, the data is *not* normally distributed. **Standard Deviation (*SD*):** The spread of the data around the mean of a scale-level variable. ***t*-Test Statistic (*t*):** Used with the *df* to determine the *p* value.