10.7: Matched or Paired Samples

In most cases of economic or business data we have little or no control over the process of how the data are gathered. In this sense the data are not the result of a planned controlled experiment. In some cases, however, we can develop data that are part of a controlled experiment. This situation occurs frequently in quality control situations. Imagine that the production rates of two machines built to the same design, but at different manufacturing plants, are being tested for differences in some production metric such as speed of output or meeting some production specification such as strength of the product. The test is the same in format to what we have been testing, but here we can have matched pairs for which we can test if differences exist. Each observation has its matched pair against which differences are calculated. First, the differences in the metric to be tested between the two lists of observations must be calculated, and this is typically labeled with the letter "d." Then, the average of these matched differences, $\overline{X}_d$ is calculated as is its standard deviation, S_d. We expect that the standard deviation of the differences of the matched pairs will be smaller than unmatched pairs because presumably fewer differences should exist because of the correlation between the two groups.

Even though the match pair is covered as part of the two population comparison, I consider this a one sample comparison. Why? because you take the difference between the two values as the data you are analyzing, and apply the one sample hypothesis test.

When using a hypothesis test for matched or paired samples, the following characteristics may be present:

1. Simple random sampling is used.
2. Sample sizes are often small.
3. Two measurements (samples) are drawn from the same pair of individuals or objects.
4. Differences are calculated from the matched or paired samples.
5. The differences form the sample that is used for the hypothesis test.
6. Either the matched pairs have differences that come from a population that is normal or the number of differences is sufficiently large so that distribution of the sample mean of differences is approximately normal.

In a hypothesis test for matched or paired samples, subjects are matched in pairs and differences are calculated. The differences are the data. The population mean for the differences, μ_d, is then tested using a Student's-t test for a single population mean with $n-1$ degrees of freedom, where n is the number of differences, that is, the number of pairs not the number of observations.

$$\textbf{The null and alternative hypotheses for this test are:}$$

$$H_0 : \mu_d = 0$$

$$H_a : \mu_d \neq 0$$

$$\textbf{The test statistic is:}$$

$$t_c = \frac{\overline{x}_d - \mu_d}{\left(\frac{s_d}{\sqrt{n}}\right)}$$

At a 5% level of significance, from the sample data, there is not sufficient evidence to conclude that the strength development class helped to make the players stronger, on average.

10.7: Matched or Paired Samples is shared under a not declared license and was authored, remixed, and/or curated by LibreTexts.