Comparing Two Means

Independent Samples

- **Independent Samples:** implies data values obtained from one sample are unrelated to the values from the other sample
  - If we are comparing two populations using the independent sample design, we independently select samples from both populations.
  - If we are comparing two treatments using the independent sample design, we randomly assign the subjects to the treatments.
- **Examples**
  - Assignment of patients in medical trials.
  - Assignment of students to teaching methods.

Dependent Samples

- **Dependent Samples:** implies subjects are paired so that they are as much alike as possible before measurements are made.

- **Examples:**
- Do PF Fliers make one run faster?
- Does eating Wheaties make one stronger?
- Is one product display more effective in producing sales than another?

- **Pair based on the following:**
- • speed results
- • strength results
- • store locations

- In Dependent Sample Design, we randomly assign one member of the pair to treatment one and the other to treatment two.
  - The purpose of pairing is to explain subject to subject variability.
- If subject to subject variability is large relative to the expected treatment differences then we want to use a dependent sample design. Otherwise, use an independent sample design.
  - In some studies, we apply both treatments to the same subject.
    - For example: comparison of types of shoe heels

For dependent sample design

- **all analysis is based on differences**
- The differences are a sample from a distribution with mean, μ , and unknown variance, σ .
- Sample size, n, is the number of pairs.
- df = (n-1)
- See inference handout

Independent Samples

- Each subject is in only one group and there is no reason to believe that knowing about one subjects ‘performance’ in one of the groups would tell you anything about another subjects ‘performance’ in the other group
  - In this situation we are said to have independent samples or, as it is sometimes called, a between subjects design

Inference in the independent scenario

- To carry out procedures in the independent setting for two sample means, we need to calculate the following for each sample:
  - Sample mean
  - Standard deviation
  - Standard error

- We then look at the ratio of the difference between sample means and the standard error of the distribution of differences between sample means.
  - The larger this is the more significant the result will be
    - The less likely the result would occur by chance if there was no real difference between the samples.
- See Inference Handout