Comparing Two Means:

Comparative studies are more convincing than single single-sample investigations, so one-sample inference is not as common as comparative two-sample inference.

In a comparative study, we may want to compare two treatments, or we may want to compare two populations. In either case, the samples must be chosen randomly and independently in order to perform statistical inference.

Because matched pairs are NOT chosen independently, we will not use two-sample inference for a matched pairs design. For a matched pairs design, apply the one-sample t-procedures to the observed differences.

Otherwise, we may use two-sample inference to compare two treatments or two populations.

The null hypothesis is that there is no difference between the two parameters.

$H_0: \mu_1 = \mu_2$ or $H_0: \mu_1 - \mu_2 = 0$

The alternative hypothesis could be that

$H_A: \mu_1 \neq \mu_2$ (two-sided)

$H_A: \mu_1 > \mu_2$ or $H_A: \mu_1 - \mu_2 = 0$ (one-sided)

$H_A: \mu_1 < \mu_2$ or $H_A: \mu_1 - \mu_2 < 0$ (one-sided)

Before you begin, check your assumptions! For comparing two means, both samples must be an SRS and must be chosen independently. Also, both populations must be normally distributed. (Check the data for outliers or skewness.)

If these assumptions hold, then the difference in sample means is an unbiased estimator of the difference in population means, so $\bar{x}_1 - \bar{x}_2$ is equal to $\mu_1 - \mu_2$.

Also, the variance of $\bar{x}_1 - \bar{x}_2$ is the sum of the variances of $\bar{x}_1$ and $\bar{x}_2$, which is $\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$.

Furthermore, if both populations are normally distributed, then $\bar{x}_1 - \bar{x}_2$ is also normally distributed.

In order to standardize $\bar{x}_1 - \bar{x}_2$, subtract the mean and divide by the standard deviation

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$$

If we do not know σ_1 and σ_2, we will substitute the standard error s_1 and s_2 for the standard deviation. This gives the standardized t value:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$$

CAUTION! This statistic does NOT have a t distribution.

We will use the TI-83 to perform two-sample T tests.

2-SampTTest for a hypothesis
2-SampTInt for a confidence interval