10.1 and 11.1: Confidence Intervals and Hypothesis Tests for the Difference Between Two Means: Independent Samples

Often, we do not have enough information to hypothesize a value for the population mean. Instead of comparing the mean to a hypothesized value, we want to compare means for two different groups.

For example, we might want to compare the mean tumor shrinkage in two sets of cancer patients, one group receiving conventional treatment and the other group receiving a new treatment.

A marketing executive wishing to measure the success of a new marketing campaign could compare the mean sales for a sample of stores not using the marketing strategy to the mean sales for stores that used the marketing strategy.

Hypothesis testing for the difference between two means (critical-value approach):

When comparing the means of two samples, we assume each sample was taken from a different population. In other words, Sample 1 comes from a population with mean μ_1; Sample 2 comes from a population with mean μ_2.

To determine whether the two means are the same, we set up null and alternative hypotheses:

$$H_0 : \mu_1 = \mu_2$$
$$H_1 : \mu_1 \neq \mu_2$$

Note: This is equivalent to:

$$H_0 : \mu_1 - \mu_2 = 0$$
$$H_1 : \mu_1 - \mu \neq 0$$

If we do not have information about the standard deviations of the populations, we use the *Welch's approximate t*:

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

For the denominator, we use the sample standard deviations to approximate the standard error of the sampling distribution of $\bar{X}_1 - \bar{X}_2$:

$$\sigma_{\bar{X}_1 - \bar{X}_2} \approx s_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

As long as both populations are normally distributed, or both sample sizes are sufficiently large (at least 30), the Welch's approximate *t* approximately follows the Student's *t*-distribution, with the degrees of freedom corresponding to the <u>smaller</u> of the two sample sizes.
In order to use the procedure below, the following statements should be true:

- We know (or can reasonably assume) that the both populations follow the normal distribution, or have a large sample size ($n \geq 30$).
- Both samples are randomly obtained from their corresponding populations, or individuals are assigned randomly to one of the two groups.
- Observations within each sample are independent of one another.
- For each sample, the sample size is not over 5% of the population.
- The two samples are independent (data points from one sample are not paired with data points in the other sample).

<u>Hypothesis Testing for a the Difference of Two Independent Means:</u>

<u>Step 1</u>: Determine the significance level α .

<u>Step 2</u>: Determine the null and alternative hypotheses.

Two-Tailed Test (most common)	Left-Tailed Test (rare)	Right-Tailed Test (rare)
$H_0 : \mu_1 = \mu_2$ $H_1 : \mu_1 \neq \mu_2$	$H_0 : \mu_1 = \mu_2$ $H_1 : \mu_1 < \mu_2$	$H_0 : \mu_1 = \mu_2$ $H_1 : \mu_1 > \mu_2$
Rejection Region	Rejection Region	Rejection Region

Note: One-tailed tests assume that the scenario not listed ($\mu_1 > \mu_2$ for a left-tailed test or $\mu_1 < \mu_2$ for a right-tailed test) is not possible or is of zero interest.

<u>Step 3</u>: Use your α level and hypotheses, sketch the rejection region.

<u>Step 4</u>: Compute the test statistic $t = \dfrac{(\overline{x}_1 - \overline{x}_2) - (\mu_1 - \mu_2)}{s_{\overline{X}_1 - \overline{X}_2}} = \dfrac{(\overline{x}_1 - \overline{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$

<u>Step 5</u>: Use a *t*-distribution table to determine the <u>critical value for *t*</u> associated with your rejection region. For the degrees of freedom, use the smaller of the two sample sizes.

<u>Step 6</u>: Determine whether the value of *t* calculated from your sample is in the rejection region.
- If *t* is in the rejection region, reject the null hypothesis.
- If *t* is not in the rejection region, do not reject the null hypothesis.

<u>Step 7</u>: State your conclusion.

Constructing a confidence interval for the difference between means:

We can construct a confidence interval for the difference between means for two independent samples.

Upper bound: $(\bar{x}_1 - \bar{x}_2) + t_{\alpha/2} \cdot s_{\bar{X}_1 - \bar{X}_2} = (\bar{x}_1 - \bar{x}_2) + t_{\alpha/2} \cdot \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$

Lower bound: $(\bar{x}_1 - \bar{x}_2) - t_{\alpha/2} \cdot s_{\bar{X}_1 - \bar{X}_2} = (\bar{x}_1 - \bar{x}_2) - t_{\alpha/2} \cdot \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$

To calculate $t_{\alpha/2}$, we use the Student's *t*-table, choosing the degrees of freedom corresponding to the smaller sample.

Example 1: Suppose an inventor wants to test the effectiveness of a new produce container intended to extend the life of refrigerated fruit. The inventor buys three dozen apples, and randomly chooses 12 of them to test in the new containers. (Due to limited funding for creation of prototypes, she only had 12 containers.) The other 24 apples were placed in standard bags. The containers and the bags were placed in the refrigerator. She recorded the time it took for each apple to develop observable mold. For the apples in the new containers, the mean time to develop mold was 6.5 days, with a standard deviation of 0.9 days. For the apples in bags, the mean time to develop mold was 5.2 days, with a standard deviation of 1.2 days. Perform a hypothesis test to evaluate whether the new containers helped the apples last longer, using $\alpha = 0.05$. Construct the 95% confidence interval for the difference between the means. Assume that the time for mold to grow is normally distributed.

<u>Example 2:</u> The instructor for a keyboarding classes wishes to determine which of two sets of practice exercises is more effective for his students learning to type. One sample of 34 students used Exercise Set A and had a mean of 48.6 words per minute and a standard deviation of 6.2 words per minute on a typing test. A second sample of 42 students used Exercise Set B and had a mean of 51.9 words per minute and a standard deviation of 8.1 words per minute. Does this sample provide evidence that the two exercise sets resulted in different typing speeds? Use $\alpha = 0.05$. Construct the 95% confidence interval.