Today we will study

- An introduction to two-sample hypothesis testing, for the difference between two population parameters
- How to perform a two-sample z-test for the difference between two means $\mu_1$ and $\mu_2$ (large independent samples)

An Overview of Two-Sample Hypothesis Testing
Paraphrased from our text - section 8.1

**DEFINITION**
For a two-sample hypothesis test,

- the **null hypothesis** $H_0$ is a statistical hypothesis that usually states there is no difference between the parameters of two populations. The null hypothesis always contains the condition of equality $\leq, =, \geq$.
- the **alternative hypothesis** $H_a$ is a statistical hypothesis that must be true when $H_0$ is false. Contains the condition of inequality, such as $<$, $\neq$, $>$.

Two-Sample z-test for the Difference Between Means

Three conditions must be satisfied to perform this z-test.

- The two samples must be independent.
  Two samples are independent if the sample values selected from one population are not related to or somehow paired with the sample values selected from the other population.
- Each sample size must be at least 30 or, if not, each population must have a normal distribution with a known standard deviation.
- The samples must be randomly selected.

We have already learned to perform a hypothesis test for the population mean and the population proportion $p$ using a z-test. The process here is similar.

**Two-Sample z-Test for the Difference Between Means**

A **two-sample z-test** can be used to test the difference between two population means $\mu_1$ and $\mu_2$ when a large sample (at least 30) is randomly selected from each population and the samples are independent. The **test statistic** is $\bar{x}_1 - \bar{x}_2$, and the **standardized test statistic** is

$$ z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sigma_{\bar{x}_1 - \bar{x}_2}} $$

where

$$ \sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} $$

When the samples are large you may use $\sigma_1 \approx s_1$ and $\sigma_2 \approx s_2$.

If the samples are not large, you can still use a two-sample z-test, provided the populations are normally distributed and the population standard deviation is known.

**Remark**
If the null hypothesis states $\mu_1 = \mu_2$, $\mu_1 \leq \mu_2$, or $\mu_1 \geq \mu_2$, then $\mu_1 = \mu_2$ is assumed and the expression $\mu_1 - \mu_2$ above is equal to 0.
GUIDELINES
Using a Two-Sample z-Test for the Difference Between Means
(Large Independent Samples)

1. Write the null hypothesis \( H_0 \) and alternative hypothesis \( H_a \); then identify the claim.
2. Specify the level of significance \( \alpha \).
3. Sketch the sampling distribution, add the test statistic, critical value(s) and rejection region(s).
4. Determine the critical value(s), \( z_0 \).
5. Determine the rejection region(s).
6. Calculate the standardized test statistic \( z \).
7. Make a decision to reject \( H_0 \) or fail to reject \( H_0 \).
   (a) If \( z \) is in the rejection region - reject \( H_0 \).
   (b) If \( z \) is not in the rejection region - fail to reject \( H_0 \).
8. Interpret the decision in the context of the original claim.

<table>
<thead>
<tr>
<th>( \alpha )</th>
<th>Tail</th>
<th>( z_0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.10</td>
<td>Left</td>
<td>-1.28</td>
</tr>
<tr>
<td>Right</td>
<td>1.28</td>
<td></td>
</tr>
<tr>
<td>Two</td>
<td>±1.645</td>
<td></td>
</tr>
<tr>
<td>0.05</td>
<td>Left</td>
<td>-1.645</td>
</tr>
<tr>
<td>Right</td>
<td>1.645</td>
<td></td>
</tr>
<tr>
<td>Two</td>
<td>±1.96</td>
<td></td>
</tr>
<tr>
<td>0.01</td>
<td>Left</td>
<td>-2.33</td>
</tr>
<tr>
<td>Right</td>
<td>2.33</td>
<td></td>
</tr>
<tr>
<td>Two</td>
<td>±2.575</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Claim</th>
<th>( H_0 )</th>
<th>( H_a )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reject ( H_0 ).</td>
<td>There is enough evidence to reject the claim</td>
<td>There is enough evidence to support the claim</td>
</tr>
<tr>
<td>Fail to Reject ( H_0 ).</td>
<td>There is <strong>Not</strong> enough evidence to reject the claim</td>
<td>There is <strong>Not</strong> enough evidence to support the claim</td>
</tr>
</tbody>
</table>