D&P: 11.2

F. Comparing two population means:

1. Trivial case: Pairs of observations.
   a. Ex.: 
      i. before-and-after observations for subjects, and test whether on average treatment makes a change 
      ii. twins, and see if they react differently 
      iii. right eye–left eye comparisons 
   b. Idea: Members of pair will be more like each other than will two randomly-selected people 
   c. Reduce this to the known one variable problem 
   d. Do it by analyzing measurement differences. 
      i. Population mean of differences equals differences of population means 
      ii. Standard deviation of differences may be estimated from the data 
   e. Alternatively, one may analyze ratios or some other comparison 
      i. Mean of comparison may no longer have easy interpretation in terms of comparison of means. 
   f. Use regular t-test 
   g. Called paired t-test 
   h. Example: 15 before and after measurements of systolic blood pressure.

<table>
<thead>
<tr>
<th>Sample Mean</th>
<th>Sample Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before</td>
<td>176.933</td>
</tr>
<tr>
<td>After</td>
<td>158</td>
</tr>
<tr>
<td>Difference</td>
<td>-18.933</td>
</tr>
</tbody>
</table>

2. Harder case: Unpaired observations. 
   a. Setup: 
      i. 1st population has mean $\mu_1$, sd $\sigma_1$, $n_1$ obs. 
      ii. 2nd population has mean $\mu_2$, sd $\sigma_2$, $n_2$ obs. 
      iii. Want to learn about mean differences: $\mu_1 - \mu_2$ 
   b. Examples: 
      i. Do people exercising stay healthier than those who don’t? 
      ii. Does the market return more the day after an increase, or the day after a decrease? 
   c. Require them to be independent 
      i. Don’t treat paired observations in this way. 
   d. Solution: Estimate difference by $\bar{X}_1 - \bar{X}_2$: difference in sample averages. 
   e. Need sampling distribution 
   f. Rule from before: 
      i. If: 
         • Two variables have a normal distribution 
         • They are independent 
      ii. Then: 
         • means add (subtract) 
         • variances add 
         • sum (difference) has a normal distribution (new rule) 
   g. So: $\bar{X}_1 - \bar{X}_2$ is 

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

\[ Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \]

\[ 95\% \text{ CI for } \mu_1 - \mu_2 \text{ is } \bar{X}_1 - \bar{X}_2 \pm t^* \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \]