Lecture 4c: Practice Problems

John McGready
Johns Hopkins University
1. Eight counties were selected from State A
   - Each of these counties was matched with a county from State B, based on the following factors:
     - Mean income
     - Percentage of residents living below the poverty level
     - Violent crime rate
     - Infant Mortality Rate (IMR) in 2006
     - Information on the infant mortality rate in 2007 was collected on each set of eight counties
     - IMR is measured in deaths per 10,000 live births
     - A pre- and post-neonatal care program was implemented in State B at the beginning of 2007
1. This data is being used to compare the IMR rates in States A and B in 1997
   - This comparison will be used as part of the evaluation of the neonatal care program in State B, regarding its effectiveness on reducing infant mortality
### Practice: Paired t-test

1. The data is as follows:

<table>
<thead>
<tr>
<th>Pair</th>
<th>IMR: State A</th>
<th>IMR: State B</th>
<th>Difference (A - B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>80</td>
<td>76</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>130</td>
<td>112</td>
<td>18</td>
</tr>
<tr>
<td>3</td>
<td>88</td>
<td>97</td>
<td>-9</td>
</tr>
<tr>
<td>4</td>
<td>98</td>
<td>67</td>
<td>31</td>
</tr>
<tr>
<td>5</td>
<td>103</td>
<td>107</td>
<td>-4</td>
</tr>
<tr>
<td>6</td>
<td>121</td>
<td>116</td>
<td>5</td>
</tr>
<tr>
<td>7</td>
<td>83</td>
<td>94</td>
<td>-11</td>
</tr>
<tr>
<td>8</td>
<td>93</td>
<td>78</td>
<td>15</td>
</tr>
<tr>
<td>Mean</td>
<td>99.5</td>
<td>93.4</td>
<td>6.1</td>
</tr>
<tr>
<td>SD</td>
<td>17.9</td>
<td>18.1</td>
<td>14.5</td>
</tr>
</tbody>
</table>
1. What is the appropriate method for testing whether the mean IMR is the same for both states in 2007?
   - Estimate a 95% confidence for the true difference in mean IMR between the two state groupings
   - State your null and alternative hypotheses for the corresponding hypothesis test
   - Report a p-value for the hypothesis test
   - Do the results from the 95% confidence level and the p-value agree in terms of the null hypothesis (using $\alpha=.05$ for the hypothesis test)?
   - What would your results be for A - D if you had 32 county pairs, and the mean difference and standard deviation of the difference were the same?
2. What is the role of the $\alpha$-level in hypothesis testing?

3. What is the role of the p-value in hypothesis testing?

4. Which of the following quantities depend on data in a paired t-test situation?
   - The true mean difference
   - The p-value
   - The estimated standard error of the sample mean difference
   - The $\alpha$-level of the test
   - The sample mean difference
5. Seventy individuals were enrolled in a dietary counseling program intended to promote healthier eating. Each subject had his/her sodium levels measured on the day of enrollment, and after two weeks of counseling. The results of these measurements were as follows:

\[
\bar{x}_{\text{pre}} = 17.7 \text{ mEq/8hr} \\
\bar{x}_{\text{post}} = 16.5 \text{ mEq/8hr} \\
s_{\text{differences}} = 12.2 \text{ mEq/8hr}
\]

Was the change in average sodium excretion statistically significant at the .05 level? Justify your answer numerically.