Comparing the Means of Many Independent Samples (cont.)

V. Multiple Comparisons

In analysis of variance when we reject the null hypothesis then the question of interest is which means are different? It is possible to do pairwise t-test to find out the means that are different from one another. However, the overall level of significance of these tests is not fixed and depends on the number of performed comparisons. To remedy this problem, there are a few techniques to compare the means simultaneously and they are called *multiple comparison procedures*. Here we will introduce a few of these methods.

1. The Bonferroni Method: This method is simply based on the Bonferroni inequality, which states that the probability that at least one of several events will occur is less than or equal to the sum of the individual probabilities of these events. In hypothesis testing context, this means that if we do k tests each at level of significance of $\alpha/k$, then the joint level of significance of these tests cannot exceed $\alpha$. Similar argument applies to the confidence intervals. Table 11 gives critical values for the Bonferroni procedure.

   Example: Soybean data
2. **The Tukey Method:** The method is based on the *studentized range distribution* and provides all possible pair-wise comparisons. We will not go to the details of this method, but the procedure is available in computer packages such as MINITAB.

**Example:** Soybean data

Tukey's pairwise comparisons

Family error rate = 0.0500  
Individual error rate = 0.0200  
Critical value = 3.61

Intervals for (column level mean) - (row level mean)

<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th>Light</th>
</tr>
</thead>
<tbody>
<tr>
<td>Light</td>
<td>-108.08</td>
<td>-39.92</td>
</tr>
<tr>
<td>Stress</td>
<td>-17.08</td>
<td>56.92</td>
</tr>
<tr>
<td></td>
<td>51.08</td>
<td>125.08</td>
</tr>
</tbody>
</table>

3. **Fisher’s least significance difference (LSD)**

The $100(1-\alpha)\%$ confidence interval for the difference between two means is:
Example: Soybean data

<table>
<thead>
<tr>
<th>(stress)</th>
<th>(control)</th>
<th>(more light)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>218</td>
<td>235</td>
</tr>
</tbody>
</table>

VI. The Basic Assumptions of Analysis of Variance

In the analysis of variance we have the following assumptions:

- Observations are independent samples from their respective populations.
- The groups are independent from each other.
- Populations are normally distributed with equal standard deviations.

Normal Probability Plot