

DIET AND LONGEVITY STUDY

9. Multiple Comparisons

The diet restriction study is an example of a planned experiment. Indeed, before the experiment was carried out we planned to make several comparisons between the lifetimes of some mice groups fed different diets. In the previous section we estimated the individual differences by the 95% confidence intervals.

Suppose now that we wish to examine the differences between all possible pairs and we don't have any particular pairs to compare in mind before conducting the experiment. This setting calls for using multiple comparisons. Multiple comparison procedures have been developed as ways of constructing individual confidence intervals so that the simultaneous confidence level is controlled (at 95%, for example). The 95% simultaneous confidence level means that we can be 95% confident that all the intervals simultaneously contain the differences. That is, in 95% of all experiments, every confidence interval would include the true value of $\mu_i - \mu_j$, and only 5% of the time would at least one interval fail to cover the true value.

SPSS has several multiple comparison procedures that should be run after the experiment has been conducted. The most important are Tukey's HSD method, Bonferroni method, and the LSD (least significant difference) Fisher's method, and Duncan's method. The LSD method is the most liberal procedure (narrowest confidence intervals) and has good power to detect differences when the differences are present.

In the LSD method a set of simultaneous confidence intervals for *all possible* differences of the form $\mu_i - \mu_j$ is formed by computing

$$(\bar{x}_i - \bar{x}_j) \pm LSD_\alpha,$$

where

$$LSD_\alpha = t_{1-\alpha/2, N-k} * s_p * SQRT\left(\frac{1}{n_i} + \frac{1}{n_j}\right)$$

for each $i < j$, where s_p is the pooled standard deviation and RANGE is the studentized range statistic. The LSD test requires that, to achieve significance, any pairwise difference must exceed a critical value that is equal to LSD_α .

SPSS produces the following outputs for the LSD test:

Multiple Range Tests: LSD test with significance level .05						
The difference between two means is significant if $MEAN(J)-MEAN(I) \geq 4.7222 * RANGE * SQRT(1/N(I) + 1/N(J))$ with the following value(s) for RANGE: 2.78						
(*) Indicates significant differences which are shown in the lower triangle						
		Group				
		1	2	3	4	5 6
Mean	Group					
27.4020	1					
32.6912	2		*			
39.6857	3		* *			
42.2972	4		* * *			
42.8857	5		* * *			
45.1167	6		* * * *			

In the above table the means are ordered and displayed from smallest to largest in the rows, and the asterisks in the lower part of the matrix indicate which pairs of groups differ significantly at the 5% level.

Notice that SPSS doesn't produce confidence intervals for the differences between treatment means. Nevertheless, the confidence intervals can be calculated from the values provided by the SPSS output by the formula:

$$(\bar{x}_i - \bar{x}_j) \pm 4.7222 * 2.78 * SQRT\left(\frac{1}{n_i} + \frac{1}{n_j}\right)$$

or equivalently

$$(\bar{x}_i - \bar{x}_j) \pm 13.12772 * SQRT\left(\frac{1}{n_i} + \frac{1}{n_j}\right).$$

The confidence intervals are closely related to the multiple comparisons results for the same method. If one of the confidence intervals includes the value 0, then that pair of means will not be declared significantly different, and vice versa.

The LSD procedure fixes the probability of a false rejection for each single pair of means being compared. It does not control the overall probability of some false rejection among all pairs.

The Bonferroni method does not have the above disadvantage. Use of the procedure with $\alpha=0.05$, for example, guarantees that the probability of *any* false rejection among all comparisons made is no greater than 0.05. This is much stronger protection than controlling the probability of a false rejection at 0.05 for each separate comparison.

Multiple Range Tests: Modified LSD (Bonferroni) test with significance level .05

The difference between two means is significant if
 $MEAN(J)-MEAN(I) \geq 4.7222 * RANGE * \sqrt{1/N(I) + 1/N(J)}$
 with the following value(s) for RANGE: 4.18

(*) Indicates significant differences which are shown in the lower triangle

Mean	Group	Groups					
		1	2	3	4	5	6
27.4020	1						
32.6912	2		*				
39.6857	3		*	*			
42.2972	4		*	*	*		
42.8857	5		*	*	*	*	
45.1167	6		*	*	*	*	*

Variable LIFETIME
 By Variable TREATMT

Multiple Range Tests: Tukey-HSD test with significance level .050

The difference between two means is significant if
 $MEAN(J)-MEAN(I) \geq 4.7222 * RANGE * \sqrt{1/N(I) + 1/N(J)}$
 with the following value(s) for RANGE: 4.06

(*) Indicates significant differences which are shown in the lower triangle

Mean	Group	Groups					
		1	2	3	4	5	6
27.4020	1						
32.6912	2		*				
39.6857	3		*	*			
42.2972	4		*	*	*		
42.8857	5		*	*	*	*	
45.1167	6		*	*	*	*	*

Note that the LSD procedure has detected more significant differences between the treatment means than each of the remaining two multiple comparisons tests considered above.

The researchers in the study used the Duncan's multiple range to evaluate statistical significance of differences between means. Its output in SPSS is displayed below:

Multiple Range Tests: Duncan test with significance level .05

The difference between two means is significant if
 $MEAN(J) - MEAN(I) \geq 4.7222 * RANGE * \sqrt{1/N(I) + 1/N(J)}$
 with the following value(s) for RANGE:

Step	2	3	4	5	6
RANGE	2.79	2.93	3.02	3.09	3.15

(*) Indicates significant differences which are shown in the lower triangle

Mean	TREATMT	Groups					
		1	2	3	4	5	6
27.4020	1						
32.6912	2		*				
39.6857	3		*	*			
42.2972	4		*	*	*		
42.8857	5		*	*	*	*	
45.1167	6		*	*	*	*	*