Post Hoc Comparisons

Tukey's HSD procedure is appropriate for post-hoc pairwise comparisons between groups. The output lists all possible pairwise comparisons, including those that are redundant.

Dependent Variable: Outcome
Tukey HSD

(I) Factor	(J) Factor	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Level 1	Level 2	-4.00000	1.73205	.106	-8.8359	.8359
	Level 3	-5.00000*	1.73205	.043	-9.8359	-.1641
Level 2	Level 1	4.00000	1.73205	.106	-.8359	8.8359
	Level 3	-1.00000	1.73205	.835	-5.8359	3.8359
Level 3	Level 1	5.00000*	1.73205	.043	.1641	9.8359
	Level 2	1.00000	1.73205	.835	-3.8359	5.8359

* The mean difference is significant at the .05 level.

The "Sig." column provides the probability of the HSD statistic (which is not listed). The HSD statistic is a function of the "Mean Difference" and the "Std. Error". For the first comparison in the example:

$$HSD = \frac{M_1 - M_2}{SE_{DIFF}} = \frac{-4.000}{1.732} = 2.309$$

An HSD of 2.309 (with 2 df_{BETWEEN} and 9 df_{WITHIN} like in the ANOVA source table) has a two-tailed probability (p) of .106, which is not a statistically significant finding.

"Mean Difference (I-J)" is the difference between the means for the "I" and "J" groups. Even though half of the listed comparisons are redundant, the mean differences will have the opposite signs because of subtraction order. This will also change the signs of the associated confidence intervals.

These "Standard Errors" are for the difference between the two group means. The values are a function of the MS_{WITHIN} (from the ANOVA) and the sample sizes:

$$SE_{DIFF} = \sqrt{\left(\frac{MS_{WITHIN}}{n_{GROUP}}\right) + \left(\frac{MS_{WITHIN}}{n_{GROUP}}\right)}$$

$$SE_{DIFF} = \sqrt{\left(\frac{6.000}{4}\right) + \left(\frac{6.000}{4}\right)} = 1.732$$

In this case, because all groups are of the same size, the standard error for each comparison is the same.

This section provides confidence intervals around (centered on) the "Mean Differences." Calculation requires the appropriate critical value. Specifically, the HSD statistic (with 2 df_{BETWEEN} and 9 df_{WITHIN}) that has a probability of .05 equals 3.068. For the first comparison in the example:

$$CI_{DIFF} = M_{DIFF} \pm (HSD_{CRITICAL})(SE_{DIFF})$$
$$CI_{DIFF} = -4.000 \pm (2.792)(1.732)$$

Thus, the estimates that the true population mean difference is somewhere between -8.836 and 0.836 (knowing that the estimate could be incorrect).

Homogeneous Subsets

Tukey HSD

Factor	N	Subset for alpha = .05 1	2
Level 1	4	2.0000	
Level 2	4	6.0000	6.0000
Level 3	4		7.0000
Sig.		.106	.835

Means for groups in homogeneous subsets are displayed.
a. Uses Harmonic Mean Sample Size = 3.000.

"Homogeneous Subsets" provide groupings for the means. Means within the same subset are not significantly different from each other (note the "Sig." value at the bottom of the column for the subset). This offers a useful summary of the comparisons as analyzed above.