Print

Multiple Comparison Methods

Multiple comparisons can inflate the Type 1 error rate for the experiment if it's not controlled using post hoc techniques. This means that the probability of incorrectly rejecting the null hypothesis when evaluating differences in means will increase if adjustments are not made. For example, suppose your ANOVA results suggest that you reject the null hypothesis that the means are the same across groups. You then decide to conduct multiple pairwise comparisons in a post hoc analysis to learn which means differ. However, when you perform a single statistical test at the α level of 0.05, you have a 5% chance of incorrectly rejecting the null hypothesis, if the null hypothesis is in fact true.

If you make no adjustments to your analysis procedure and continue to use the α equals 0.05 criterion for each pairwise comparison, then your probability of making a Type 1 error on at least one of the pairwise tests goes up dramatically as the number of comparisons increases. The comparisonwise error rate, or the CER, is the probability of a Type 1 error on a single pairwise test.

The experimentwise error rate, or EER, is the probability of making at least one Type 1 error when you perform the entire set of comparisons. The EER considers the number of pairwise comparisons that you make, so it increases as the number of tests increases. Presuming the null hypothesis is true for each test, the chance that you falsely conclude that at least one difference exists is much higher when you consider all possible comparisons.

Comparisonwise Error Rate	Number of Comparisons	Experimentwise Error Rate
0.05	1	0.05
0.05	3	0.14
0.05	6	0.26
0.05	10	0.40

You can calculate the EER as 1 minus the complement, that is, 1 minus the probability of making no Type 1 errors. You calculate EER as $1 - (1 - \alpha)^{nc}$ where *nc* is the number of comparisons, assuming that the tests are independent.

If you're testing one hypothesis at a significance level of .05, meaning that the CER is .05, then the overall EER is .05 as well. If you're testing three hypotheses, each at a significance level of .05, then the overall EER is .14. This means that you have a 14% chance of rejecting at least one of your three null hypotheses just by chance, even if the null is true.

You can see how the EER increases even more if you conduct six and 10 pairwise comparisons. Comparing four groups, such as in the heating quality example in the last demonstration, requires you to make six pairwise comparisons. Without any adjustment, the chance of at least one false positive result is 26%. You need to use a method that controls the EER at a level like 0.05. Let's learn about two of these methods.

Close