A second alternative to the pairwise testing approach

Steve Simon

2022-07-04

In an earlier blog post, I mentioned Dunnett’s test, a comparison of multiple treatments to a control group as an alternative to the pairwise approach to testing that is commonly used in analysis of variance. Here’s a second interesting alternative.

### Comparison to an overall mean

You are in a setting where you are expecting to see a degree of uniformity among several groups. For example, you have a group of clinics who all treat the same type of patients. The time spent with each patient, the total cost of the medications prescribed, the length of the notes written in the medical record–these can differ from patient to patient. But because of how the patients are assigned to a particular clinic (more or less randomly), these should balance out on average.

If there is one clinic that has different average results, it has to be because that the physicians at that clinic have a different approach to treatment. If you can identify a clinic (it could be more than one clinic, of course) that differs from the norm, then investigate. Perhaps that clinic is doing something “wrong” that causes the discrepancy, or maybe that clinic is the only one doing things right. In any case understanding why you see unexpected deviations in a process that should be uniform across clinics provides you with an opportunity to learn something.

### The traditional hypothesis

You’ve seen this hypothesis for the standard ANOVA setting in your introductory Statistics class



This is a very common approach in a comparison of k independent groups. Test for any deviation from the null hypothesis using an F test. If this is statistically significant, then use a Tukey follow-up test to see which pairs of means differ from one another.

### The comparison to the overall mean hypothesis

Suppose you are in a setting like the one described above, where there is an expectation of similarity between each group mean and the overall mean. Then you would write your hypothesis as



This approach is not that common, but you can find it in most statistical software programs under the name “Analysis of Means.” It fits in well with a setting where you hope that all the groups produce consistent results. If any do not, then you want to identify the group or groups that deviate from the norm and study factors that may account for the deviation.

The analysis of means approach compares each group mean to the overall mean. It is easy to implement and it lends itself to a simple graphical display. You need a table of critical values, which depend on alpha (the overall Type I error rate), g (the number of groups), and n (the number of observations within each group). Some tables use the degrees of freedom for error in place of n.

You don’t need to compute the traditional F statistic for Analysis of Variance first, because the Analysis of Means approach controls the overall Type I error rate. This protects you from the accusation of p-hacking, even if the number of groups is very large.

It’s important to editorialize a bit here. Deviating from the norm could be a good thing or a bad thing or it could be an indifferent thing. Your goal is not to use statistics to hunt out different groups to reward or punish them. You are using statistics to help in understanding if deviations from the norm occur and then study those deviating groups to understand why they deviate.

### A simple example

Five genetically modified strains of eggplant were tested for presence of a natural insecticide, CryA1c, in the leaves of the plant. Four plants were assessed for each strain. Here are the descriptive statistics. The research question is whether any strain differs from the overall mean level of CryA1c.

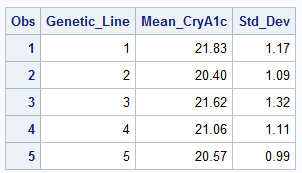


(Image taken from Wikipedia)

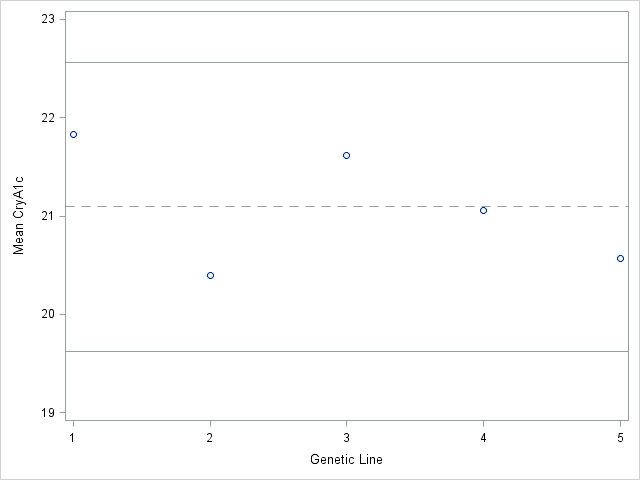
Desiree M. Hautea DM et al. Field Performance of Bt Eggplants (Solanum melongena L.) in the Philippines: Cry1Ac Expression and Control of the Eggplant Fruit and Shoot Borer (Leucinodes orbonalis Guenée). PLoS One. 2016; 11(6): e0157498.

<https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0157498>

Here are the summary statistics.



The raw data was not available, but you can calculate the analysis of means limits using the group means and standard deviations. You can plot the individual means versus the limits provided by analysis of means calculations.



Notice that all five means lie inside the limits. None of the five strains shows a statisticially significant difference from the overall mean.

You may prefer a one-sided test in this setting, such as testing whether any strain is deficient in the CryA1c levels. You can make a very easy modification to get one-sided tests.

### Caveats

The formulas get a bit more complex and the graph gets a bit messier if the sample size differs from one group to another.

More importantly, the only comparisons that you can make with the analysis of means approach is a comparison to the overall mean. If two groups are both statistically significantly higher than the overall mean, you cannnot make a comparison between those two groups without losing control over the overall Type I error rate.

You also need to specify the Analysis of Means hypothesis prior to looking at your data. Peeking at the data and then choosing your hypothesis is cheating.

The Analysis of Means, just like the Dunnett’s test described in an earlier blog post, provides a simple approach to testing a different type of hypothesis. Because it reduces the number of comparisons from all possible pairwise differences to a comparison of each group to the overall mean, you gain some precision and can summarize your results in a simple easily understood graph.