Using Least Squares Means and Multiple Comparison Tests in an ANCOVA Model

This demonstration shows how to request least squares means and do multiple comparison tests in ANCOVA.

Here is the plot produced in the previous demonstration: Analysis of Covariance for BP Change. The mean of **BaselineBP** is approximately 95. Because the slopes for the three treatments are not equal, we want to calculate least squares means at two other values of **BaselineBP**: 90 and 100.

Let's look at how we modify the PROC GLM code from the previous demonstration. As before, the CLASS variable is **Treatment**. The MODEL statement is simplified. The vertical bar notation indicates that the model includes the categorical predictor **Treatment**, the covariate **BaselineBP**, and the interaction of those two variables.

The three LSMEANS statements request least squares means. All three of these statements specify the categorical variable **Treatment**, followed by options. The first LSMEANS statement specifies only two options, which also appear in the other two LSMEANS statements. The PDIFF=ALL option requests all pairwise comparisons of the three treatments and *p*-values for the differences. The *p*-values help us determine whether the differences are significantly different from zero. (Note that the PDIFF option is set to =ALL by default, so specifying the =ALL is not required.) In all pairwise comparisons, we want to make adjustments in order to maintain the type 1 error rate. The ADJUST= option specifies the Tukey method of adjustment, which is recommended for looking at all pairwise comparisons. On the other hand, if we wanted to look at only the comparisons with the placebo, for example, the Tukey adjustment would be too much and we would want to select a different method. This first LSMEANS statement has no other options, so SAS will report the least squares means for the treatment groups at the mean value of the covariate.

The second and third LSMEANS statements contain an additional option -- the AT option -- to request the least squares means at values of the covariate other than the mean: 90 and 100, respectively.

Let's submit the code.

```
proc glm data=mydata.trials;
   class treatment;
   model bpchange = treatment|baselinebp;
   lsmeans treatment / pdiff=all adjust=tukey;
   lsmeans treatment / at baselinebp=90 pdiff=all adjust=tukey;
   lsmeans treatment / at baselinebp=100 pdiff=all adjust=tukey;
title 'Least Squares Means for ANCOVA Model';
run;
quit;
```

We've already seen the first part of the results (in the last demonstration): the ANOVA table, the R square, type 3 sums of squares, and the significant interaction. Let's focus on the results generated by the LSMEANS statements.

The first set of tables and graphs is generated by the first LSMEANS statement, and they report on the comparison of the least squares means for the three treatment groups at the mean value of the covariate. In the heading above the tables, notice the reference to the Tukey adjustment for multiple comparisons. Remember that the first table assigns numbers to the least squares means for the three treatment groups: *Approved Drug* is 1, *New Drug* is 2, and *Placebo* is 3. In the next table, the *p*-values indicate that all of the comparisons are significant.

Next is a graph of the group means, followed by a diffogram. The diffogram provides a graphical test of significance for mean comparisons at a specific value of the covariate (here, the mean value of **BaselineBP**). Each line segment represents a comparison of least squares means. The length of the lines indicates the confidence interval. The diagonal reference line is the zero line. Lines that do not cross the reference line represent groups that have least squares means that are significantly different from one another. Any line segments that cross the reference line are displayed in red. They indicate that the least squares means for these groups are not significantly different from each other at alpha = 0.05. In this diffogram, none of the lines crosses the reference line, which indicates that the least squares means for all treatment groups are significantly different from each other. As you can see, the diffogram provides a lot of information at a glance.

Next are the results from the second LSMEANS statement. The heading shows that the comparison is made at the **BaselineBP** value of 90. The *p*-values indicate that *Approved Drug* and *New Drug* are found to be statistically different from one another. However, based on the *p*-values, we see that one of the comparisons is highly nonsignificant: the comparison between treatment groups 1 (*Approved Drug*) and 3 (*Placebo*). The graph of the group means shows that the means for the approved drug and the placebo are almost identical. In the diffogram, the line for the *Approved Drug*-

Placebo comparison is red and crosses the reference line—again, indicating that there is no significant difference between the approved drug and the placebo at the **BaselineBP** value of 90. Is this result surprising? Let's take a quick look at the ANCOVA graph that we skipped over earlier. At **BaselineBP** = 90, the *Approved Drug* and *Placebo* lines intersect; they have the same value, so there is no significant difference. We only see differences between these two treatments at higher values of **BaselineBP**.

Finally, we look at the results from the third LSMEANS statement, for the **BaselineBP** value of 100. The *p*-values indicate that all pairwise comparisons are significant at 100. The diffogram shows this, as well.

Copyright © 2017 SAS Institute Inc., Cary, NC, USA. All rights reserved.

Close