Print

Practice 3.1 (Level 1): Generating a Two-Way Analysis of Variance

Task

In this practice, you generate a two-way analysis of variance.

A computer service center has four technicians who specialize in repairing three brands of computer disk drives for desktop computers. The service center wants to study the effects of the technician and brand of the disk drive on the service time. The data is stored in the data set **mydata.disks**.

Reminder: Make sure you've defined the **mydata** library.

1. Use PROC GLM to generate a two-way analysis of variance with **Time** as the dependent variable and **Technician** and **Brand** as the independent variables. Include the interaction between the independent variables in your model. Presuming a level of significance of 0.05, is the overall *F* test significant in your model? Is there a significant interaction?

```
proc glm data=mydata.disks;
   class Technician Brand;
   model Time=Technician|Brand;
run;
quit;
```

   Examine the results.

   The *p*-value for the overall *F* test is less than 0.0001, which is less than the significance level of 0.05. Therefore, the test is significant and you can conclude that at least one treatment mean is different from one other treatment mean. The *p*-value for the test for interaction is also less than 0.0001. Therefore, the interaction is significant.

2. Examine the interaction plot. Does this graph verify the conclusion reached in the test for interaction? Why or why not?

   The graph verifies that there is an interaction between **Technician** and **Brand**. This is evident because the lines on the graph are not parallel. In particular, notice that Bob repairs **Brand 3** in the shortest amount of time, but **Brand 3** takes Karen the longest amount of time to repair.

3. Is it appropriate to examine the tests for the main effects shown in the PROC GLM output?

   The tests for the main effects in the PROC GLM output can be misleading in the presence of an interaction. Therefore, it is not appropriate to examine these tests. Instead, you should test for simple effects using the SLICE= option.

4. Use the LSMEANS statement with the SLICE= option to determine whether there are differences between the technicians for each brand of disk drive. Also examine the differences between the brands of disk drive for each technician. What are your conclusions?

```
proc glm data=mydata.disks;
   class Technician Brand;
   model Time=Technician|Brand;
   lsmeans Technician*Brand / slice=Brand slice=Technician;
run;
quit;
```

Examine the results.

Based on the *p*-values shown in the Technician*Brand Effect Sliced by Brand for Time table, the service time for at least one technician is different from one other technician for each brand of disk drive.

In the Technician*Brand Effect Sliced by Technician for Time table, within each of the technicians, except for Justin, the average service time for at least one brand of disk drive is different from the service times for the other brands.

Hide Solution

Close