# CAKE-BAKING EXPERIMENT

9.   Checking Model Assumptions

The test procedures of GLM General Factorial Model described in the previous sections are valid only if it is assumed that all taste scores are independent of one another, the scores are normally distributed and that variances are equal for all treatment groups.

The most serious violation of the above assumptions is a lack of independence among observations. If the experiment is conducted according to the rules discussed in Section 3, there is no reason to suspect that the assumption might be violated.

There are nine treatment groups, each consisting of three observations. Under the circumstances, it is very difficult to detect non-normality or obtain strong evidence that the assumption of equal variances is violated. Even visual inspection of the data shows that the spread of observations in all the groups is similar. In particular, it is easy to check that the standard deviation takes on the three discrete values 1, 1.732, and 2. Taking into account a relatively small number of observations in each group, this does not provide any evidence that the assumption of equal variances is violated. The spread versus level plot displayed below also indicates strong support for the assumption.

Spread vs. Level Plot of RATING

Level (Mean)

Groups: TIME * TEMP
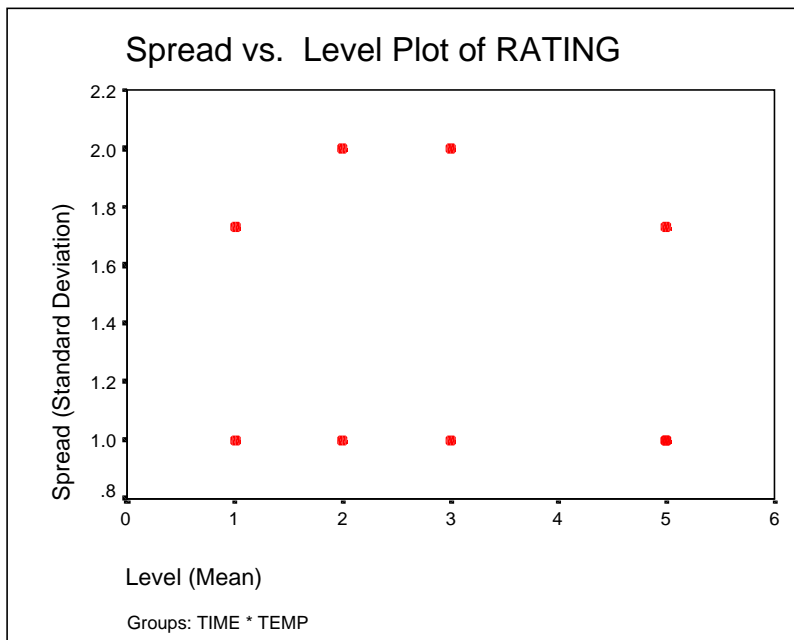
In general, analysis of variance is robust to departures from normality, although the data should be symmetric. It is hard to make strong claims about non-normality if the number of observations in each group is small.

The matrix scatterplot of residuals is shown below. The plots are also useful for checking assumptions about the data.

Dependent Variable: RATING
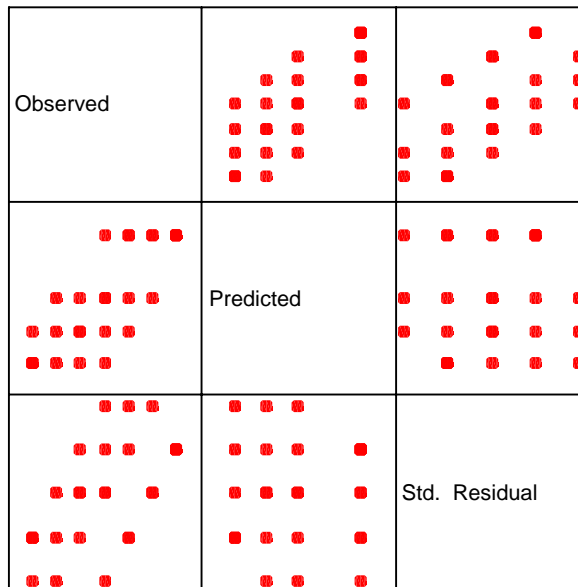
Model: Intercept + TIME + TEMP + TIME*TEMP

The plot of observed versus predicted values (middle-upper plot) is close to a 45° straight line passing through the origin, which indicates a good fit. The standardized residuals versus the predicted values (middle-lower plot) shows some non-random effects.