# ABSORBENCY OF PAPER TOWELS

10.   Checking the Assumptions

The test procedures of GLM General Factorial Model described in the previous sections are valid only if it is assumed that all absorbency measurements are independent of one another, the measurements 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 5, there is no reason to suspect that the assumption might be violated.

There are nine treatment groups, each consisting of just five observations. Under the circumstances, it is very difficult to detect non-normality or obtain strong evidence that the assumption of equal variances is violated. In Section 8 we found that the standard deviation for each brand decreases when taken across the three time levels from high of 0.6419 to low of 0.1942. Moreover, it seems that the inherent properties of the absorption process make the standard deviation smaller and smaller as immersion time increases. Thus, we can expect that any sample of 5 observations at the time level of 10 seconds will produce measurements with similar spread. However, this violation of the equality of variances assumption may make the F statistic unreliable.

The spread versus level plot displayed below also indicates violation of the assumption.
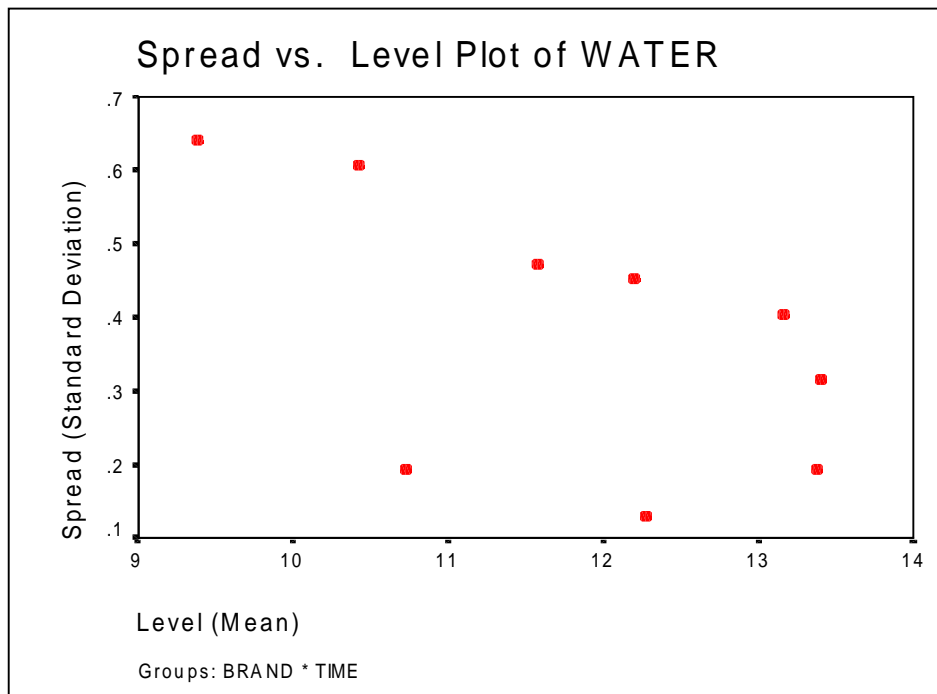

One solution is to transform the response variable. Observe that in our data set, variance appears to decrease as the means increase. This is why the logarithmic and square-root

transformation have turned out to be unsuccessful to make the spread of observations in the treatment groups more even. But also other transformations such as $1/\sqrt{X}$ were not able to make the spread of the data in the groups approximately equal.

Fortunately, a violation of this assumption has minimal impact if the groups are of approximately equal sizes. In our data set, all groups have the same size, and therefore we expect that unequal variances do not change the value of the F-statistic significantly.

Now we discuss the normality assumption. In general, analysis of variance is robust to departures from normality, although the data should be symmetric. The Q-Q plots in SPSS for each of the nine groups did not indicate that the assumption of normality might be violated for any of the groups. In general, it is hard to make strong claims about non-normality if the number of observations in each group is that small.

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

Dependent Variable: WATER

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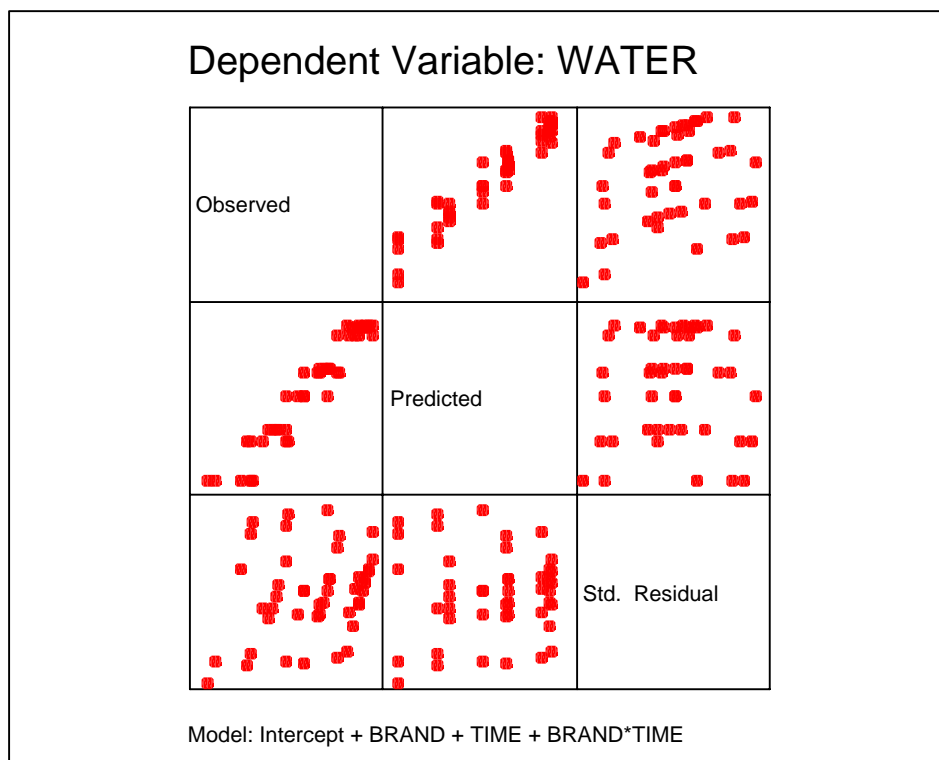
The plot of observed versus predicted values (middle upper plot) is close to a 45° straight line passing through the origin, which indicates a good fit. The standardized residuals versus the predicted values (middle lower plot) shows some non-random effects.