7. Checking the Regression Model Assumptions

In order to describe the relationship between breakdown time and voltage, we have used the simple regression model. However, the conclusions based on the model are valid only if the underlying assumptions are satisfied. Specifically we assume that:

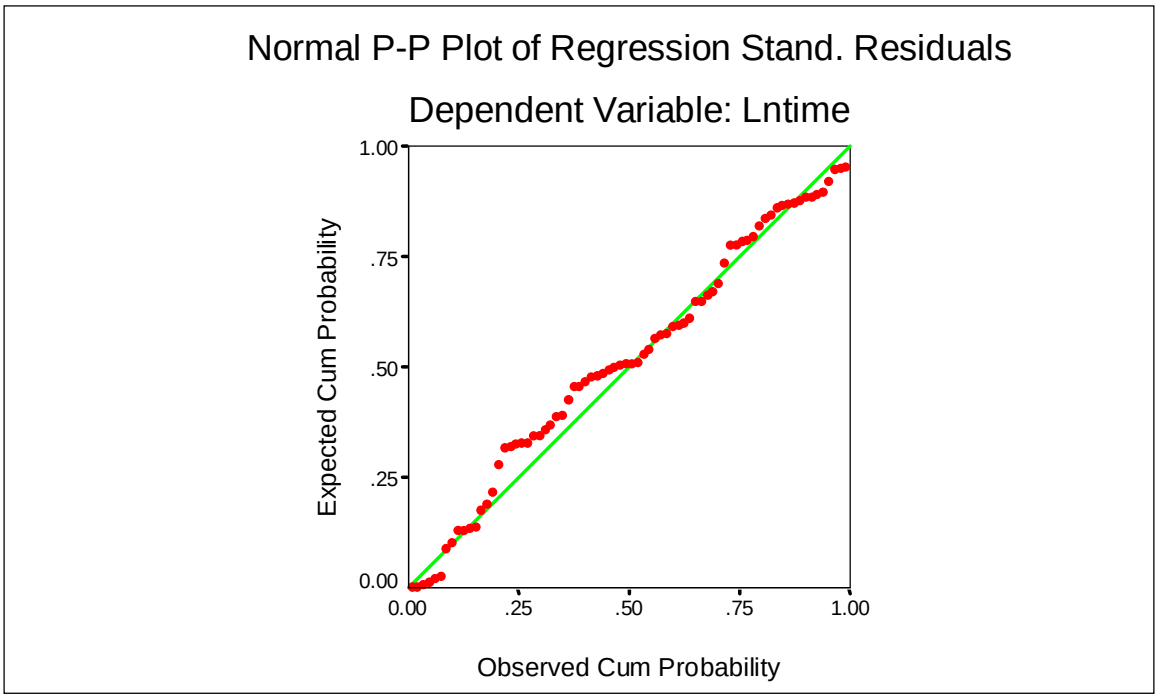
1. The plot of log-breakdown time means against voltage is a straight line.
2. The log-breakdown times at each voltage level follow a normal distribution.
3. The spread of the log-breakdown time around the straight line is the same at all voltage levels.
4. The log-breakdown times for one voltage level are independent of log-breakdown times in other groups.

7.1 Checking the Normality Assumption

Estimates of the coefficients and their standard errors are robust to nonnormal distributions. The consequences of violating this assumption are usually minor for the tests and confidence intervals. However, if prediction intervals are used, departures from normality become important. This is because the prediction intervals are based directly on the normality of the population distributions whereas tests and confidence intervals are based on the sampling distributions of the estimates which may be approximately normal even when the population distributions are not.

The normality assumption can be verified by looking at the plot residuals. Indeed, if the residuals at each voltage level follow a normal distribution, then the log-breakdown times follow a normal distribution and vice versa.

In order to assess whether the assumption is not violated with SPSS, the normal P-P plot of regression standardized residuals is obtained. The plot plots the cumulative proportions of standardized residuals against the cumulative proportions of the normal distribution. If the normality assumption is not violated, points will cluster around a straight line.

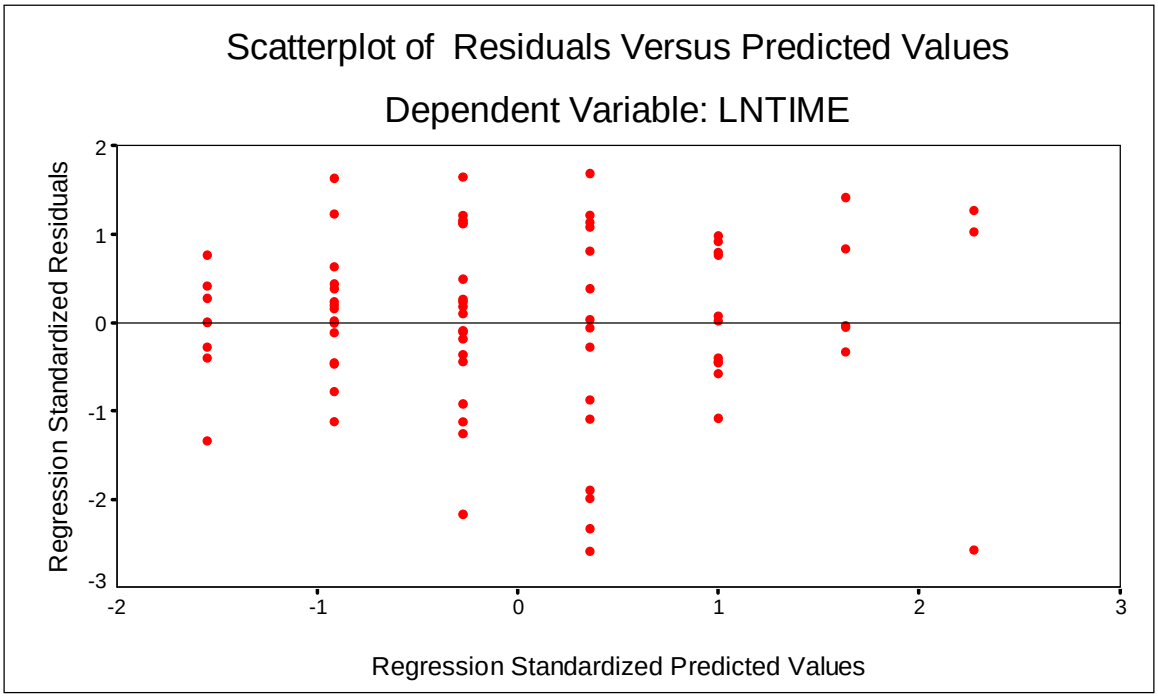


As you can see, the above plot supports the normality assumption. The pattern is close enough to a straight line.

7.2 Checking the Constant Variance Assumption

One method of checking whether the assumption of constant variance is not violated is to plot the residuals against the predicted values. We then look for a change in the spread or dispersion of the plotted points.

The scatterplot displayed below shows that the assumption of constant variance is not very likely to be violated. The spread of the plotted points is not significantly different at different voltage levels.



7.3 Checking the Independence Assumption

Of the all assumptions, independence is the most crucial. Lack of independence causes no bias in estimates of the coefficients, but standard errors are seriously violated. As a consequence the tests and confidence intervals van be effected. This assumption can be violated by the method in which the measurements were obtained. There is no sufficient information here to indicate that the assumption might be violated.

7.4 Summary

After a logarithmic transformation of the times to breakdown, a simple linear regression model fits the insulating fluid data well. No evidence (from the normal P-P plot or residual plot) indicates that either the normality or constant variance assumption is violated. Thus mean estimation and prediction can proceed from that model, with results back-transformed to the original scale.