8. One-Way Analysis of Variance Model

The simple linear regression model allows us to produce some estimates of the mean breakdown time at each voltage level. The model is based on the assumption that the log-breakdown time means lie on a straight line against voltage. If this is not the case, an analysis of variance can be used.

In this section we will apply the analysis of variance to our experiment. The separate means model (one-way ANOVA) applied to our experiment has the form

$$\mu\{Ln(Time)\,|\,Voltage\_i\} = \mu_i$$

for i=26, 28, …, 38, where μ_i denotes the mean of the group subjected to the voltage of i kV. This model is obviously more general than the simple regression model because the log-breakdown time means may or may not lie on the straight line -their values are not restricted.

The following display contains the output of the one-way analysis of variance applied to the log-transformed breakdown times.

Variable LNTIME
By Variable CODE

Analysis of Variance

Source	**D.F.**	**Sum of Squares**	**Mean Squares**	**F Ratio**	**F Prob.**
Between Groups	6	196.4774	32.7462	13.0043	.0000
Within Groups	69	173.7489	2.5181		
Total	75	370.2263			

Group	**Count**	**Mean**	**Standard Deviation**	**Standard Error**	**95% Conf Inter for Mean**		
26 kV	3	5.6240	3.3552	1.9371	-2.7109	TO	13.9589
28 kV	5	5.3295	1.1446	.5119	3.9084	TO	6.7507
30 kV	11	3.8220	1.1112	.3350	3.0755	TO	4.5685
32 kV	15	2.2285	2.1981	.5675	1.0113	TO	3.4458
34 kV	19	1.7864	1.5252	.3499	1.0513	TO	2.5215
36 kV	15	.9022	1.1099	.2866	.2876	TO	1.5169
38 kV	8	-.4243	.9917	.3506	-1.2534	TO	.4047
Total	76	2.1457	2.2218	.2549	1.6380	TO	2.6534

We remember that if the log-transformed data have a symmetric distribution, then taking the antilogarithm of the mean on the log scale gives an estimate of the median on the original scale. Hence, in order to obtain the estimates of the group medians on the original scale of measurement, you have to take the antilogarithm of the above group means.

The analysis of variance F-statistic is F=13.0043, with 6 and 69 degrees of freedom, giving a reported p-value of 0. The p-value indicates strong evidence against the null hypothesis of no difference among the average log-breakdown times for the seven groups. In other words, there is strong evidence of differences among the group medians on the original scale of measurement.

The way the experiment was conducted makes it possible to infer that the different voltage levels must be directly responsible for the observed differences in time to breakdown.

The within-group mean square is 2.5181, so the pooled estimate of a common standard deviation is the square root of the value, which is equal to 1.586852 minutes.

The estimated means in the model for each voltage level are simply the group averages. The standard errors displayed above are based on the individual standard deviations for each group, not the pooled standard errors. In order to obtain the pooled standard errors of the means, which is the one-way analysis of variance standard error, we divide the pooled estimate of standard deviation (1.586852) by the square root of the sample size.

The quality of estimated means in the insulating fluid problem from the separate-means model and the simple linear regression model is discussed in **Section 10**.