Report on the analysis of Lacquer Data

R. Nazim Khan
School of Mathematics and Statistics
The University of Western Australia

Telephone: (08) 6648 3369
email: nazim@maths.uwa.edu.au

3 October 2006
1 Executive Summary

Brass Products (henceforth referred as The Company) has received complaints from customers that its brass lamps and mailboxes corrode prematurely. The Company has identified that the problem lies with the poor quality lacquer currently used to coat the brass. Five possible alternatives are being considered by The Company. Data provided by The Company on the number of days until the first sign of rust on mailboxes coated with each of the lacquers was analysed using statistical techniques (ANOVA), with the aim to determine the best and most cost effective lacquer.

![Days to Rust against Lacquer](image.png)

Figure 1.1: Scatter plot of data.

A graph of the data (Figure 1.1) shows considerable overlap between the five lacquers, with perhaps lacquer 2 providing the highest and lacquer 1 lowest mean number of days before any visible signs of rust.

On the basis of the statistical analysis, we conclude that lacquers 2, 4 and 5 are of the same quality. We recommend lacquer 5, since it the the cheapest of the three.

2 Introduction

The Company has had numerous complaints from customers regarding the premature rusting of its brass lamps and mailboxes. It has come to the conclusion that the cause of the problem is the low quality of the lacquer used to coat the products. To solve this problem The Company is considering replacing the lacquer by one of five possible alternatives. In order to choose between them, an experiment was conducted in which each lacquer was used to coat 25 mailboxes selected at random, and these were then placed outside under similar weather exposure conditions. For each mailbox the number of days until the first visible signs of rust were recorded. The data is available as an Excel file lacquer.xls.
3 Data Analysis

3.1 Exploratory Data Analysis

The scatter diagram of the data is shown in Figure 1.1, and table 3.1 presents some summary statistics.

<table>
<thead>
<tr>
<th></th>
<th>Lacquer 2</th>
<th>Lacquer 4</th>
<th>Lacquer 5</th>
<th>Lacquer 1</th>
<th>Lacquer 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>185.6</td>
<td>182.6</td>
<td>178.8</td>
<td>162.4</td>
<td>155.8</td>
</tr>
<tr>
<td>Median</td>
<td>184</td>
<td>176</td>
<td>177</td>
<td>165</td>
<td>156</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>41.47</td>
<td>40.72</td>
<td>34.00</td>
<td>34.87</td>
<td>34.13</td>
</tr>
</tbody>
</table>

Table 3.1: Summary statistics for the four lacquers, listed in decreasing order of mean.

The scatterplot and table indicates that lacquers 2 and 4 have the largest mean, followed by lacquer 5, with lacquers 1 and 3 having the lowest means. The spread of the data for each lacquer is similar, and the distribution of the data appears to be fairly symmetric. There do not seem to be any extreme outliers in the data.

3.2 Statistical Analysis

An analysis of variance (ANOVA) was performed in Excel; Figure 3.1 shows the relevant output.

The hypotheses being tested here are

\[ H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5 \quad H_1 : \text{Not all the means are equal} \]

where \( \mu_i \), \( i = 1, 2, \ldots, 5 \) is the mean number of days before the first visible signs of rusting for lacquer \( i \). The p-value for the test is \( 0.0172 < 0.05 \), so the data provides sufficient evidence against the null hypothesis. Thus we conclude that the mean number of days before the first signs of rusting are not the same for the five lacquers.

To determine the best lacquer, we next find confidence intervals for the difference of the means. Note that the half width of the intervals is

\[
w = t_{0.025, 120} \sqrt{MSE \left( \frac{1}{25} + \frac{1}{25} \right)} = 1.9799 \sqrt{1382.981 \left( \frac{1}{25} + \frac{1}{25} \right)} = 20.83,
\]

Figure 3.1: ANOVA output from Excel.
and these half widths are the same for all the intervals since the sample sizes are equal.

<table>
<thead>
<tr>
<th>Difference of Means</th>
<th>Confidence interval</th>
<th>Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \mu_2 - \mu_3 )</td>
<td>(9.0, 50.7)</td>
<td>( \mu_2 &gt; \mu_3 )</td>
</tr>
<tr>
<td>( \mu_2 - \mu_1 )</td>
<td>(2.4, 44.1)</td>
<td>( \mu_2 &gt; \mu_1 )</td>
</tr>
<tr>
<td>( \mu_2 - \mu_5 )</td>
<td>(-14.0, 27.7)</td>
<td>( \mu_2 = \mu_5 )</td>
</tr>
<tr>
<td>( \mu_4 - \mu_3 )</td>
<td>(6.0, 47.6)</td>
<td>( \mu_4 &gt; \mu_3 )</td>
</tr>
<tr>
<td>( \mu_4 - \mu_1 )</td>
<td>(-0.6, 41.0)</td>
<td>( \mu_4 = \mu_1 )</td>
</tr>
<tr>
<td>( \mu_5 - \mu_3 )</td>
<td>(2.2, 43.8)</td>
<td>( \mu_5 &gt; \mu_3 )</td>
</tr>
<tr>
<td>( \mu_5 - \mu_1 )</td>
<td>(-4.4, 37.2)</td>
<td>( \mu_5 = \mu_1 )</td>
</tr>
<tr>
<td>( \mu_1 - \mu_3 )</td>
<td>(-14.2, 27.4)</td>
<td>( \mu_1 = \mu_3 )</td>
</tr>
</tbody>
</table>

Table 3.2: Confidence intervals for difference of means for the four designs.

The confidence intervals are listed in Table 3.2. We have listed them in order, first giving the difference between the largest and the smallest mean, then the next smallest mean and so on, until the means are not significantly different (at the 5% level of significance).

From these it can be seen that the means for lacquers 2, 5 and 4 are not different. Although lacquer 4 is not different from lacquer 1 on the basis of these confidence intervals, the evidence is not strong, as the confidence interval only just covers 0. On the other hand, it is clear on the basis of these confidence intervals that lacquer 5 is not different from lacquer 1 but is better than lacquer 3.

### 3.3 Model Diagnostics

We now perform an analysis of residuals to check model assumptions. Figure 3.2 shows the relevant plots.

The histogram of residuals is not unlike that expected for a normal distribution. Thus we conclude that the residuals are normally distributed.

The scatterplot of residuals against fitted values shows that the spreads of the five lacquers are not different. We conclude that the five lacquers have a common variance.

Finally the plot of residuals against fitted values shows no obvious pattern, and we conclude that the residuals are independent. Thus all the assumptions of the ANOVA model are satisfied.
4 Conclusion

The analysis of the data shows that lacquers 2, 4 and 5 provide longer term rust protection on average compared to the other lacquers and are statistically indistinguishable. On the basis of cost we would prefer Lacquer 5 from among these. Further, lacquers 5 and 1 are also statistically indistinguishable, and neither are lacquers 1 and 3, and we would prefer lacquer 3 on the basis of cost. However, lacquers 2 and 4 are statistically different from lacquer 1 and provide better protection on average than lacquer 1. Thus, the choice needs to be between lacquers 2 and 4, and we select lacquer 2 on the basis of cost.

5 Recommendation

We recommend that lacquer 2 be selected on the basis of cost. However, other aspects, such as availability and supply should also be considered.