Project Two

DAT 475

Southern New Hampshire University

Catherine Taft

5/29/23

**Procedure**

Upon evaluating the results from the previous investigation of the facility, we have decided to do a one-way ANOVA test to determine the degree to which the proportion of defects in each of the three models may differ. In this way we may be able to further narrow down our examination of the products and isolate deeper issues, as previously we did note problems with low quality components that were being purchased to save money, as well as inconsistent process documentation and training.

The dataset we were given consisted of a count of defects for each of three models, as well as percentages of defects. Given that the counts were somehow identical for all three models (a suspicious fact that warrants further investigation), that data is not useful for this, as there will obviously not be any difference between the means, and that is the very thing that ANOVA tests compare. Therefore, we decided to examine the proportions of defects. In the following test, we will look at the difference in means between defect proportions across all three models.

**Hypotheses**

Given that we are comparing means across three models of circuit board, the null hypothesis for this test will be: = µModel1 = µModel2 = µModel3. In other words, there is no statistically significant difference in the mean proportion of defects between any of the models.

The alternative hypothesis will be = µModel1 ≠ µModel2 ≠ µModel3, in other words these are not roughly equivalent and there is a statistically significant difference between the mean proportion of defects between at least two of these models.

Once running the test, if we find that there is indeed a statistically significant difference between at least two of the models, we can do a post-hoc test to determine which pairing of models has significant differences, as the initial ANOVA test doesn’t tell us that (Frost, 2021). At that point, we will be able to tell which models are more prone to defects than others, and we can then drill down on the root causes for any potential differences. The fishbone diagram we did previously can help us isolate potential problem areas to investigate that may pertain to manufacturing defects. Knowing which models to focus on will make our investigation more efficient.

**Hypothesis Test**

We performed the hypothesis test in SPSS using Defect Percentage as the dependent variable, and Model Number as the factor. We are using an α value of 0.05. I will add screenshots of the initial and post-hoc tests with brief explanations, and then provide analysis in the following section. I did need a tutorial to walk me through the process because I had never used SPSS before, and there was some manipulation of data required within the program before the analysis (O’Loughlin, 2016).

Here is the initial result from the one-way ANOVA. Confusingly, SPSS labels the p-value as “Sig”, but here we have a p-value of 0.023 for the between-groups result. Comparing that to our significance level of α = 0.05, we can see that the p-value is less than the significance level, which indicates that there is a statistically significant difference between at least two of these groups. Therefore, we will reject the null hypothesis and move on to a post-hoc test.

A picture containing text, screenshot, font, number

Description automatically generated

The Tukey test was chosen for this, and we see the following results. Please note that I accidentally labeled Model 3 as just “3”, so that is what “3” means here.

A screenshot of a spreadsheet

Description automatically generated with low confidence

We can remove the duplicates to reduce this to the following three pairings and their p-values:

Model 1 & Model 2: 0.036

Model 1 & Model 3: 0.041

Model 2 & Model 3: 0.997

**Results and Analysis**

As we saw, in the baseline ANOVA test the p-value was less than the significance level, which indicated to us that we can reject the null hypothesis, and that there is a statistically significant difference between at least two of the groups. Completing the post-hoc test gave us a pairwise comparison breakdown of the variables. Comparing the p-values of the resulting pairs to the significance level of 0.05, we find that Model 2 and Model 3 do not have a statistically significant difference, but the other two pairings do, and therefore we reject the null hypothesis for those two pairings specifically.

With that information we can then look at the Mean Difference column in the post-hoc chart for those two pairings. Model 1 and Model 2 have a mean difference of 10.65, and Model 1 and Model 3 have a mean difference of 10.4. The positive value of the mean difference, which is calculated as I-J where Model 1 is I, indicates that Model 1 has the higher mean proportion of defects in both cases. Given that these are the two pairings with statistically significant differences based on the p-values, we can determine that Model 1 is associated with a significantly higher number of defects than Model 2 or Model 3, and we should evaluate Model 1 and its associated processes based on the potential problems isolated previously in the fishbone chart. The fact that Model 2 and Model 3 do not have a statistically significant difference in mean defect proportion may also be useful as we compare similarities and differences that may exist between the production processes of these models.

Resources

Frost, J. (2021, September 17). *Using post hoc tests with ANOVA*. Statistics By Jim. https://statisticsbyjim.com/anova/post-hoc-tests-anova/

O'Loughlin, E. (2016, November 16). *How To Perform a One-Way ANOVA Test in SPSS.* YouTube. https://www.youtube.com/watch?v=OEOeXpxSjf8