BIA652D Multivariate Data Analysis

Manova Analysis of HATCO Dataset

GROUP4 Homework 7

2016

# **Summary of MANOVA Analysis**

A MANOVA analysis was undertaken for variables related to customer satisfaction level and usage level to see if these measures differ across the three groups of buying situations faced by the customer – new task, modified rebuy and straight rebuy. A second analysis was also undertaken to see if these measures differed across combination of buying situations and industry types. Buying situation has three groups as mentioned above and industry type has two groups – Industry A and Other Industries. The satisfaction and usage levels are metric and measured on a 10-point scale.

First an examination of each of the measures (satisfaction and usage levels) was undertaken separately and then across the groups of buying situations and industry types to check for validity of assumptions of normality and homoscedasticity. The boxplot, normal probability plot and histogram on page 8 show that usage level is normally distributed. The Shapiro Wilks test gives p-value of 0.3201 indicating that null hypothesis that usage level is normally distributed cannot be rejected at 0.05 significance level. For satisfaction level, the p-value given by this test is 0.0740 indicating that this variable is normally distributed too at 0.05 significance level. This is supported by the boxplot, normal probability plot and histogram on page 11. The boxplots on page 30 show constant variance of usage level across the three buying situations indicating assumption of homoscedasticity is met. The statistical tests for homogeneity of variance of usage level across the three buying situations, Levene, Brown and Forsythe and Bartlett, all have p-values greater than 0.05 (on page 47) also indicating assumption of homoscedasticity is met.

Satisfaction level appears to have minor violation of the assumption of homoscedasticity. Its variance across the third group of buying situation (straight rebuy) is slightly higher than the variance across the other two groups from the boxplots on page 31. Out of the three statistical tests for homogeneity of variance, two give p-value greater than 0.05 while the third one (Brown and Forsythe) has p-value less than 0.05 (0.0426). This also supports what is seen in the boxplots.

Next a set of multivariate tests is performed to see if the buying situation has an impact on the usage and satisfaction levels. All four tests Wilks' Lambda, Pillai's Trace, Hotelling-Lawley Trace and Roy's Greatest Root show that the collective set of usage level and satisfaction level have significant difference across the three groups of buying situations. The null hypothesis in this case is that no difference in collective set of usage level and satisfaction level exists across the three buying situations at 0.05 significance level. As all these tests have p-values less than 0.05, this hypothesis is rejected. Usage level and satisfaction level are examined individually too across the three groups of buying situations and individually too they have significant difference across the three groups. This means that the type of buying situation does have an impact on the usage level and satisfaction level both individually and collectively. It can be observed in the individual tests that the mean of usage level and satisfaction level are highest in buying situations that are straight rebuys and lowest in buying situations that are new tasks.

Although the above tests indicate that usage level and satisfaction level are not the same across the three buying situations it is not known for which two buying situations they are different. Therefore, a set of post-hoc tests is performed to find out which two buying situations result in a statistically significant difference in usage level and satisfaction level. When usage level and satisfaction level are examined individually, for usage level, the group difference between modified rebuy and straight rebuy is about -8.351 (minus sign indicates usage level is lower for modified rebuys), difference between new task and modified rebuy is -9.619 and difference between new task and straight rebuy is -17.97. All these differences are statistically significant at 0.05 level. These values also support what was observed earlier that usage level is lowest for new buys and highest for straight rebuys. This makes sense as usage level which is how much of the (customer) firm’s total product is purchased from HATCO would be low when customer is new to HATCO’s products and purchasing their product for the first time.

For satisfaction level, the group difference between modified rebuy and straight rebuy is -0.391 while the group difference between new buy and modified rebuy is -1.0737. The latter difference i.e. between new buy and modified rebuy is more than twice the difference between modified rebuy and straight rebuy. Both these differences are however statistically significant at 0.05 level. This is also conceptually logical as for new buys the customer would not have a lot of purchase experience to have satisfaction from. However, in case of rebuys, the customer has past purchase experience from HATCO and built a level of satisfaction or dissatisfaction based on that.

In the second analysis, the usage level and satisfaction level are again examined collectively and individually for differences but this time across different buying situations and industry types. All four tests Wilks' Lambda, Pillai's Trace, Hotelling-Lawley Trace and Roy's Greatest Root show that buying situation has a significant main effect on usage level and satisfaction level collectively however the industry type does not at 0.05 significance level. Moreover, the combined effect or interaction effect of industry type and buying situation is also not statistically significant at 0.05 level. When usage level and satisfaction level are tested individually however, industry type does have a significant main effect on usage level and so does buying situation. Their interaction effect is again not significant. For satisfaction level, individual tests show a non-significant main effect of industry type which is different from usage level where individual test showed significant main effect of industry type. Buying situation has a significant main effect on satisfaction level individually too. The interaction effect of buying situation and industry type is also non-significant on satisfaction level.

Looking at the plots on pages 6 and 8 which show the interaction effect of buying situation and industry type on usage level and satisfaction level respectively, the results of the statistical tests are more or less supported. In both plots, the difference between usage level and satisfaction level across the two different industry types is minimal. The lines are quite close. Moreover, this difference stays constant more or less constant across all three buying situations. In the usage level plot, we can see the difference in usage level between the two industry types slightly narrowing in the middle of the plot for the second group of buying situation i.e. modified rebuys. However, it does not appear to be a significant difference also supported by the statistical test results.

Thus, it can be concluded that buying situation of the customer does have a significant impact on usage level and satisfaction level both individually and collectively. Both are highest for straight rebuys and lowest for new buys. Industry type does not have a significant impact on these measures collectively or on satisfaction level individually. It only has a significant main effect on usage level individually. The usage level for other industries is slightly higher than for Industry A classified customers. Also, the interaction effect of buying situation and industry type on both these measures collectively and individually is non-significant. This means the effects of industry type and buying situation are quite independent of each other.