**Johannah Cushing**

**Statistics II, week 7: MANOVA and GLM**

**May 8, 2017**

**SUMMARY**

In this analysis each respondent is classified based on the levels (groups) of the treatment (independent variable). MANOVA is often employed to control the overall Type I error rate while assessing differences on each dependent variable collectively and individually.

HBAT aims to strengthen its customer relationship strategy and as such is looking at differences due to its distribution channel system. (x5) x5 is composed of two channels: direct through HBAT’s salesforce or indirect through a broker. Three purchase outcomes have been identified as focal issues in evaluating the impacts of the two distribution systems and will be treated as the analysis variables. These are: x19, satisfaction; x20; Likelihood of recommending HBAT; and x21, likelihood of future purchase. The purpose of the analysis is to identify if differences exist between these two systems across all or a subset of the purchase outcomes.

Table 5 “Descriptive Statistics of Purchase Outcome Measures” shows group profiles on each of the purchase outcomes (x19, x20, and x21 across the two groups- direct vs. indirect distribution systems). For all three purchase outcomes, the direct distribution channel has higher mean scores and seems to be a better selling channel. Also, x21 has higher mean scores than x20 which has higher mean scores than x19.

There is a concern for adequate sample sizes across the entire MANOVA analyses and thus 200 surveys are used instead of just the initial 100. With the larger dataset, 108 firms used the indirect broker system and 92 use the direct system. With these group sizes there will be adequate statistical power at the 80%.

With MANOVA the most important assumptions are the independence of observations, homoscedasticity across groups, and normality. Outliers can influence group means and should be considered.

A random sampling plan was employed to help ensure the independence of respondents. Levene’s test assesses the univariate homogeneity of variance across the two groups. For all three variables it is found to be nonsignificant (i.e. greater than .05). The Box’s M test for equality of the covariance matrices shows a nonsignificant value of .607, indicating that there is no significant difference between the two groups on the three dependent variables collectively. Therefore the assumption of homoscedasticity is met for each individual variable separately as well as the three variables collectively.

Bartlett’s test for sphericity is used to determine whether the dependent measures are significantly correlated. It examines the correlations among all dependent variables and assesses whether, collectively, there is significant intercorrelation. Here we find there is a significant degree of intercorrelation (significance=.000).

By looking at boxplots (distribution of x19, distribution of x20, distribution of x21) we can address outliers. There are only a few, if any extreme points across the groups. No observations had an extreme value on all three dependent measures, and no value was so extreme as to indicate that it should be excluded. Thus all 200 observations are retained.

In step 4 we assess whether the two groups have statistically significant differences for the three purchase outcomes variables, collectively as well as individually. We specify the maximum allowable Type I error rate, accepting that 5 of 100 (.05) times we might conclude that the type of distribution channel impacts the purchase outcome variables when in fact it does not. Then we can proceed to test for differences between the two groups, perform univariate tests on each purchase outcome, and then assess power levels.

Common multivariate tests include: Pillai’s criterion, Wilk’s lambda, Hotelling’s T^2, and Roy’s greatest characteristic root. Each test shows that the set of purchase outcomes have a highly significant difference (.000) between the two types of distribution channel. This is in alignment with what is seen in table 5 and the boxplots of the distributions of the individual variables. (This data is in table 7.7; not reproduced below as not able to do so in SAS).

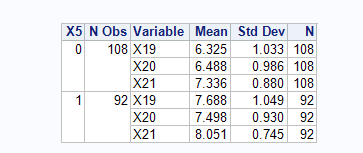
Table 7 also shows results for the univariate tests for each individual purchase outcome. All of the individual tests are highly significant (.000), indicating that each variable follows the pattern of higher purchase outcomes for those served by the direct distribution system than served indirectly through a broker (7.688, 7.498, and 8.051 vs. 6.325, 6.488, and 7.336 for x19, x20, and x21, respectively). The power for these tests was all 1.0, indicating that the sample sizes and effect size were sufficient to ensure that significant differences would be detected if they existed beyond the differences due to sampling error.

The Scheffe test and Tukey test are post-hocs tests to find out which pairs of means are significant. They suggest that the differences may not be significant. For further analysis we may wish to switch from x5 to x1.

**RESULTS**

**Table 5- Descriptive Statistics of Purchase Outcome Measures**

\*Here x5 0 is indirect through broker, and x5 1 is direct to customer



**Table 6- Multivariate and Univariate Measures for Testing Homoscedasticity of x5**

