A Comparison of Volume of People Entering and Exiting a Mall and Their Usage of the Automatic versus Manual Door

Alana Lesnansky

Course Number – Summer 2013, Instructor: Dr. Guang-Hwa Andy Chang

Executive Summary

Introduction

Automatic (handicapped) doors are becoming more and more prevalent in today's society. Some consumers prefer automatic doors to manual doors for the convenience they offer, and stores try to make the consumers comfortable. Business owners must balance the benefits versus the costs of the manual versus automatic doors and must take into consideration consumer preferences and even decide how many of which type of door, and whether to have different choices for the in versus the out door.

Methods

Two hundred and sixty one mall patrons were observed, and in order to collect this sample (considered a convenience sample), the researcher sat at a table in the Southern Park Mall near the entrance. Then, the researcher watched the doors that led to the atrium before the food court. The researcher tracked both whether the subjects were going into or out of the mall, and also took note of if they utilized the automatic (handicapped) door, or used the manual (traditional) door.

Results

Descriptive statistics were performed on both variables (out versus in and automatic versus manual). Then, confidence interval estimation for difference between two sample proportions, test hypothesis for difference between two sample proportions, and a Chi-Square test for independence were tested at the 95% confidence level to determine whether more people chose the automatic door or manual, and if they chose the automatic if they did so more often upon entering or exiting, and whether or not the two variables were independent. The results revealed that the population proportion of people who utilize the automatic door is less than .5. Also, the results showed that the variables are not independent (there is a correlation). Results also show that there are more people that use the manual door, and also, it seems as if more people who use the automatic door choose it on the way out rather than on the way in.

Conclusions

Just from the results in this study alone, if mall owners are deciding between doors, they should note that manual doors are more often used, but if they want automatic doors, they should be placed where people are exiting rather than entering.

Introduction

In our society, almost every public entrance is outfitted with an automatic (handicapped) door. Either the automatic doors are operated by a push button, or they have a sensor that lets the door open automatically as a person approaches. When there is a push button, generally there are also doors that offer a manual operation (pushing the door open). This juxtaposition of manual doors versus automatic doors is especially prevalent in public malls, where a large number of people are constantly going into and out of the mall. On the surface, which type of door or what combination of doors to have does not seem like a pressing issue, however, further investigations reveal several financial implications. When an automatic door is triggered, it usually stays open for a longer period of time than the person/people actually need to walk through the door. This is energy inefficient, letting either heat or air conditioning out, depending on the season. Accidental triggering of the automatic door also has this same effect, and in addition tends to shorten the life of the equipment (Yang, Lai, Sheu, and Chen, 2013). On the other hand, some consumers prefer automatic doors to manual doors for the convenience they offer, and stores try to make the consumers comfortable. One particular study found that, in general, doors seem more inviting when they swing open and pause before closing again (Ju and Takayama, 2009). Business owners must balance the benefits versus the costs of the doors and must take into consideration consumer preferences and even decide how many of which type of door, and whether to have different choices for the in versus the out door. When faced with a choice, will more people decide to use the automatic door on their way into or out of the mall? The research hypothesis is that more people use the automatic door on the way out, but more people use the manual doors in general, and the variables have a relationship. This is an observational study and the analyses that will be performed are confidence interval estimation for difference between two sample proportions, test hypothesis for difference between two sample proportions, and a Chi-Square test for independence. These will be performed on the statistical software package R-Commander.

Methods

This is an observational study. The population of interest is all the mall patrons in the Greater Youngstown area. The simple random sampling would have been the preferred method for data collection to answer the research question and hypothesis for this study. In order to collect this sample, the researcher sat at a table in the Southern Park Mall near the entrance. Then, the researcher watched the doors that led to the atrium before the food court. The researcher tracked both whether the subjects were going into or out of the mall, and also took note of if they utilized the automatic (handicapped) door, or used the manual (traditional) door. Both the “in” versus “out” variable and the “automatic” versus “manual” variable are qualitative. The time frame in which this experiment took place was 4:00 pm-4:50 pm. Even though simple random sampling would have resulted in a sample that more accurately depicted the population, due to time constraints, the best available form of sampling was convenience sampling. Due to its many limitations, by utilizing convenience sampling, it will be virtually impossible to draw conclusions about the whole population. A limitation of this data collection was the time frame in which the data was collected. It could be that time of day would change both the proportion of people entering and leaving the mall, and the proportion of people who utilize the automatic door over the manual. In addition, choosing all of the sample data from the same set of doors is problematic. The set of doors observed consisted of three manual doors, and only one automatic, so from the start of the experiment, when the doors became crowded, it follows from reason that more people would take the manual doors, purely because there is more space. This could be considered as a confounding variable. In addition, the sample of people who choose to enter and leave through the food court doors may not be representative of the population of all the doors at that mall as a whole. There was no non-response bias in this study, because every person who came into or out of the mall at that particular location had to choose one of the doors, no participants were able to take the option of "not responding". A sample size of 261 was chosen because it seemed reasonably large enough to capture as accurate an estimate as possible (in light of convenience sampling) of the population of interest.

Results

Descriptive Statistics

In order to understand the hypothesis and be able to analyze it fully, a more in depth look at both of the variables, both separately and together, first must be taken. A sample size of (n=261) was utilized in this study. Table 1 is a frequency chart for the variable of whether the mall patron went into or out of the mall and Figure 1 is a bar chart displaying the same data.

Table 1: Descriptive Statistics for Out or In Variable

	Frequency	**Relative Frequency**
Out	137	.5249
In	124	.4751

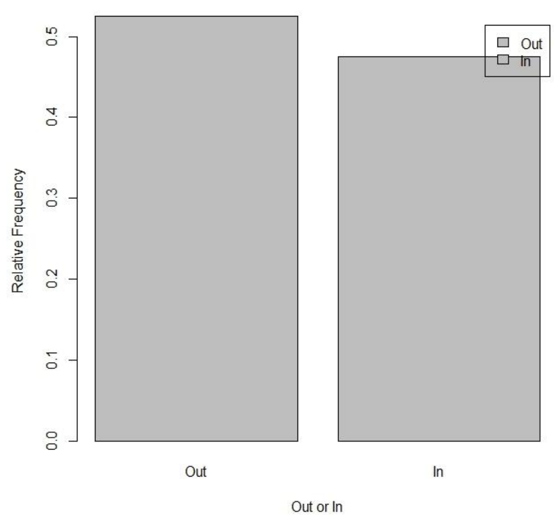

Figure 1: Bar Chart for Out or In Variable

The relative frequencies of the “in” and “out” category indicate that roughly the same volume of people entered and exited the mall during the given time period. The relative frequencies being about the same may result from the fact that during the time period observed, people who work in the mall could be getting off work to go home, and people who work elsewhere may be getting out of work, and coming to the mall to shop. Table 2 is a frequency chart for the variable of whether the patrons utilized the automatic or manual door and Figure 2 is a bar chart displaying the same data.

Table 2: Descriptive Statistics for Automatic or Manual Variable

	Frequency	Relative Frequency
Automatic	50	.1916
Manual	211	.8084

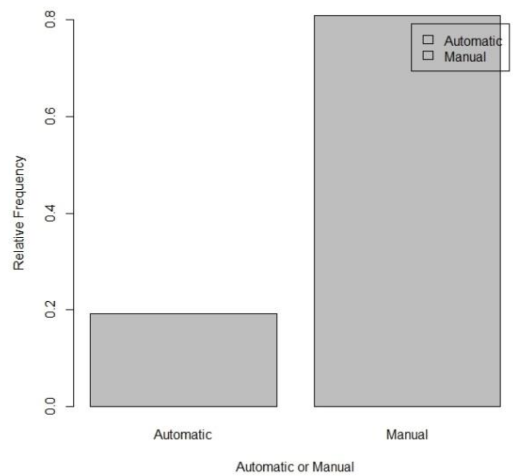

Figure 2: Bar Chart for Automatic or Manual Variable

The relative frequency of usage for the manual door is clearly larger than that of the automatic door. This could be possibly attributed to the fact that there are more manual doors to choose from than the automatic doors. Table 3 displays a frequency chart for the two variables together. Out of all people observed in this study, 19.16% of them used automatic door. Among those who were going out, 28.48% of them used automatics door. However, of those who were going in 8.87% of them used it automatics door.

Table 3: Descriptive Statistics for Out or In and Automatic or Manual

	Out	In	Total
Used Automatic Door	39	11	50
Used Manual Door	98	113	211
Total	137	124	261

It begins to become clear that more people chose the automatic door on the way out than on the way in. Figure 3 also depicts the two variables.

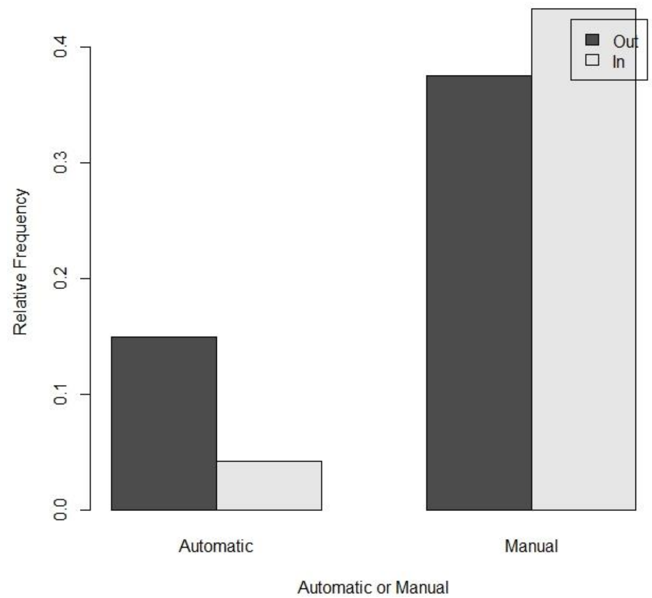

Figure 3: Bar Graph for Automatic or Manual and In or Out

The bar graph shows that more people chose the automatic door on the way out as opposed to on the way in. It also becomes evident that people chose the manual door about the same frequency, whether they were going in or out, as opposed to the automatic door (where there is a discrepancy).

Inferential Statistics

The inferential statistics for this particular set of data are confidence interval estimation for difference between two sample proportions (or percentages), test hypothesis for difference between two sample proportions, and a Chi-Square test for independence. The research hypothesis is that more people use the automatic door on the way out, but more people use the manual doors in general, and there is correlation between the two variables, “type of door used” and “customer is going in or going out”.

As reported earlier, out of all people observed in this study, 19.16% of them used automatic door. The z-test for population proportion was used for testing whether the proportion of people using automatic door is more than 50%. The assumption that the sample proportion was from a random sample was satisfied to the best of the researcher’s ability. The assumption that the sample size is large is satisfied (see Appendix). The sample size is less than 10% of the population, satisfying the last assumption. The p-value associated with this test is 1.03×10^{-23}. At the significance level of .05, there is sufficient evidence to support that less than 50% of people chose the automatic door. The 95% confidence interval estimate for percentage of people who used automatic door is 19.16% ± 4.77%.

In Table 4, the observed values are followed by the expected cell count in parenthesis. These numbers are needed in the Chi-Square Test for Independence.

Table 4: Observed and Expected Cell Counts for in or Out and Automatic or Manual Variables

	Out	**In**	**Total**
Used Automatic Door	39 (26.25)	11 (23.75)	50
Used Manual Door	98 (110.75)	113 (100.25)	211
Total	137	124	261

The assumption that the data are from a random sample is satisfied to the best of the researcher's ability. None of the expected cell counts are less than 5, so the conditions in Cochran's large sample size condition are satisfied. The computer software R-Commander produced the p-value of 5.886×10^{-5} from the chi-square test for independence. Therefore, at the .05 level of significance, there is sufficient evidence to support that, statistically, there is significant correlation between the two variables, "type of door used" and "customer is going in or going out". The evidence now suggests that whether or not people will choose the automatic door or manual door depends on whether or not they are exiting or entering. The people going out were more likely used the automatic door.

As reported, 28.48% of those who were going out used automatics door. However, of those who were going in 8.87% of them used it automatics door. For the confidence interval estimation for difference between two proportions, the assumption that the sample is randomly selected is satisfied to the best of the researcher's ability given time and project constraints. The assumption that both sample sizes are large is satisfied (see Appendix). The confidence interval estimate of the difference in proportions is (10.53%, 28.66%).This interval does not cover zero, which indicates that there is a statistical difference between the two proportions. This result matches the outcome from the chi-square test for independence.

Summary

From the descriptive statistics alone, a preliminary trend is being shown that more people do chose the automatic door when they are on the way out than when they are on the way in. After the inferential statistical analysis, the evidence supports that there is a statistically significant difference in the proportion of people who choose the automatic door upon entering versus exiting, that the proportion of people who choose the automatic door is less than .5, and also there is statistically significant correlation between the type of door used and whether the customer is going in or coming out of the mall.

More people may choose the automatic door on the way out because they are carrying bags and are tired from shopping and walking, so it seems like an easier and more convenient option. However, taken on a whole, it is clear more people use the manual doors. This could have been due to the fact that there were more manual doors than automatic doors in the area where the researchers gathered the data. It could also have been attributed to the fact that people simply prefer the manual doors. If a mall was looking at making a decision about the type of doors they would like to use for their mall, this research suggests using more manual doors than automatic, but it also suggests that if they were to use automatic doors, to place them on customers way out instead of in, though more research would clearly be necessary before deciding any sort of purchasing decision.

Some limitations of this study include the fact that it was a convenience sample, and it was looking at only one entrance during one time period of the day. This sample cannot necessarily be considered representative of the whole population of mall patrons in the Greater Youngstown Area. Further research could be conducted utilizing the simple random sampling method. In addition, further research could be done to verify the results of this study, and elaborate on the results by taking a larger sample in more places to determine consumer preferences and help businesses decide which door or set of doors is optimal for customer satisfaction.

References

Ju, W., & Takayama, L. (2009). Approachability: How People Interpret Automatic Door Movement as Gesture. *International Journal Of Design 3*(2): 77-86.

Yang, J., Lai, C., Sheu, H., & Chen, J. (2013). An Intelligent Automated Door Control System Based on a Smart Camera. *Sensors (14248220) 13*(5): 5923-5936. doi:10.3390/s130505923.

Appendix

Confidence Interval Estimation for Difference Between Two Sample Proportions

$\hat{p}_o$ = proportion of people using the automatic door in the way out (number of people choosing automatic door on the way out/number of people going out=39/137=.284672)

$\hat{p}_i$ = proportion of people that use the automatic door on the way in (number if people choosing automatic door on the way in/number of people going in=11/124=.088710)

Assumption for Large Sample Size procedure: $n_o\hat{p}_o \geq 5$, $n_o(1-\hat{p}_o) \geq 5$, $n_i\hat{p}_I \geq 5$, and $n_i(1-\hat{p}_i)\geq 5$

Based on the data observed, the assumptions are all satisfied.

Confidence Interval: $(\hat{p}_o-\hat{p}_i)\pm z_{\propto/2}\sqrt{\frac{\hat{p}_o(1-\hat{p}_o)}{n_o}+\frac{\hat{p}_i(1-\hat{p}_i)}{n_i}}$

Confidence Interval with Values: $(.284672-.088710)\pm 1.96\sqrt{\frac{.284672(1-.284672)}{137}+\frac{.088710(1-.088710)}{124}}$

.195962±.090634

(.105328, .286596)

R-Commander Output:

```
> prop.test(.Table, alternative='two.sided', conf.level=.95, correct=FALSE)

        2-sample test for equality of proportions without continuity
        correction

data:  .Table
X-squared = 16.1389, df = 1, p-value = 5.886e-05
alternative hypothesis: two.sided
95 percent confidence interval:
 0.1053294 0.2865943
sample estimates:
    prop 1     prop 2
0.28467153 0.08870968
```

Test Hypothesis for Difference Between Two Sample Proportions

$\hat{p}_a$ = proportion of people who utilized the automatic door as opposed to the manual door (number of people who choose automatic door/number of people in sample=50/261=.191571)

Null Hypothesis (H_0): the proportion of people who choose the automatic door is .5 (H_0: $\hat{p}_a$=.5)

Alternative Hypothesis (H_a): the proportion of people who chose the automatic door versus the manual door is not .5 (H_a: $\hat{p}_a$‹.5)

Assumption that Sample Size is Large:

n(hypothesized value)≥5 is 50(.5)=25

n(1-hypothesized value)≥ is 50(1-.5)=25

Test Statistic: $z=\frac{\hat{p}_a - hypothesized\ value}{\sqrt{\frac{(hypothesized\ value)(1-hypothesized\ value)}{n}}}$

Test Statistic with Values: $z=\frac{.191571-.5}{\sqrt{\frac{(.5)(1-.5)}{261}}}$ =-9.97

P-Value for a 2 Sided Test with z-score -9.97: $1.03\text{x}10^{-23}$

R-Commander Output:

Note: the P-Value calculated by the R-Commander Test for Single-Sample Proportion says "less than $2.2\text{x}10^{-16}$", the answer when calculated by hand is different but is more accurate

```
> prop.test(rbind(.Table), alternative='less', p=.5, conf.level=.95,
+   correct=FALSE)

        1-sample proportions test without continuity correction

data:  rbind(.Table), null probability 0.5
X-squared = 99.3142, df = 1, p-value < 2.2e-16
alternative hypothesis: true p is less than 0.5
95 percent confidence interval:
 0.0000000 0.2347222
sample estimates:
        p
0.1915709
```

Chi-Square Test for Independence

Null Hypothesis (H_0): the two variables are independent (no relationship)

Alternative Hypothesis (H_a): the two variables are not independent (relationship)

Test Statistic: $\sum_{all\ cells} \frac{(observed\ cell\ count - expected\ cell\ count)^2}{expected\ cell\ count}$

Degrees of Freedom: (rows-1)(colums-1)=(2-1)(2-1)=1

R-Commander Output:

```
> .Test <- chisq.test(.Table, correct=FALSE)

> .Test

        Pearson's Chi-squared test

data:  .Table
X-squared = 16.1389, df = 1, p-value = 5.886e-05
```