Body sizes and variations among different grocery stores

Katie Lancaster and Zach Wiese

Introduction

This study poses an interest to us because on occasional trips to the Wal-Mart in Ashland, WI, we both made the observation that, on average, a greater proportion of shoppers appeared overweight when compared to other grocery stores. We hypothesized that of the three grocery stores in Ashland, the Chequamegon Co-Op, County Market, and Wal-Mart, they would each have a different distribution of weight-class shoppers.

The results of the study are important in understanding demographics of our population, and may, perhaps, represent a store which supplies more "unhealthy" food as compared to another.

Methods

For this study we needed to collaborate subjectively in order create the three levels of body size. After we had determined what constituted as skinny, normal, and overweight we proceeded to gather samples for the study at each grocery store. At Wal-Mart, the first of the three stores, we had originally wanted to study shoppers going in the grocery section from the outside in the parking lot. However, we were unable to do this because there was no immediate parking next to the building, and therefore we had to enter the store and gather are our data from inside. For approximately thirty minutes we walked through the aisles collecting samples and for an additional twenty minutes we sat at a bench observing shoppers enter the store until we had our target of 35 individuals.

For both the County Market and the Chequamegon Food Co-op we were able to position ourselves directly in front of their entrances. We stayed at each location until we had gathered sufficient data (35 individuals) to finish our sampling. While our procedures were fairly biased, due to the fact that our categories were mostly subjective, we feel that we were able to maintain the same levels of criteria to determine which body size the individuals fell under.

Though we tried to gather randomized data, we realize that for this study we were unable to fully achieve this goal. Although the data itself is not random, we chose a busy time for the three stores in order to do our sampling. The time period in which we chose to complete the sampling was during relatively busy grocery shopping hours (around 5:00 pm) in order to observe a larger and possibly more diverse number of individuals. Also, if a group or family of individuals was observed, only one sample was taken from the group. This helped to ensure more random data because groups of people tended to be all of a similar body type.

Results

11-Step Hypothesis Test

1) A categorical response variable with three levels (body size) from three populations (grocery stores) was recorded; thus, a chi-square test is required.

2) H_O: The distribution of people of different body sizes is the same for all grocery stores.

H_A: The distribution of people of different body sizes is NOT the same for all grocery stores.

3) $\alpha = 0.05$

4) The design is observational, and somewhat random, given circumstances. The observational data is entered into R with,

```
> obs<-matrix(c(2,16,17,6,16,13,8,24,3),nrow=3,byrow=TRUE)
> rownames(obs) <- c("Wal-Mart","County Market","Co-Op")
> colnames(obs) <- c("skinny","normal","overweight")
> obs
              skinny normal overweight
Wal-Mart           2     16         17
County Market      6     16         13
Co-Op              8     24          3
```

In addition, the chi-square test is fit now so that the expected table can be easily observed.

```
> chi1 <- chisq.test(obs,correct=FALSE)
```

5) All cells in the expected table have values greater than five. Thus, the test statistic computed below should reasonably follow an χ^2-distribution.

```
> chi1$expected
                skinny   normal overweight
Wal-Mart      5.333333 18.66667         11
County Market 5.333333 18.66667         11
Co-Op         5.333333 18.66667         11
```

6) The table of observed frequencies is shown above.

7) The χ^2 test statistic is 15.24 with 4 df.

```
> chi1
        Pearson's Chi-squared test

data:  obs
X-squared = 15.2403, df = 4, p-value = 0.004228
```

8) The p-value is 0.004.

9) Reject H_O because p-value < α.

10) It appears that the distribution of body sizes does differ between the three grocery stores. An analysis of the residuals shows Wal-Mart as having more overweight customers than expected, County Market having more skinny and more overweight customers than expected, and the Co-Op having more skinny and normal-sized customers than expected.

```
> chi1$residuals
                 skinny     normal overweight
Wal-Mart      -1.4433757 -0.6172134  1.8090681
County Market  0.2886751 -0.6172134  0.6030227
Co-Op          1.1547005  1.2344268 -2.4120908
```

Discussion

Overall, we found that the distribution of body sizes does differ between the three grocery stores. Our main hypothesis was based on the observation of the abundance of overweight shoppers in Wal-Mart, and this proved true for our sample. The opposite-end observation we saw was a lack of overweight shoppers in the Chequamegon Co-Op, also shown in the residuals table above. The food offered at the Co-Op tends to be more local, organic products, and lacks an abundance of "cheap junk food" or low-nutrition product brands such as "Flavorite". On the other end of the spectrum is Wal-Mart, which carries GMOs, processed and far-from-local foods, and a large amount of products containing empty calories. The County Market could almost be looked at as the control or middle between the two stores, as it carries many generic brands of food, as well as some local products and a small natural-foods section. One could make the conclusion that healthier people shop at the Co-Op, and/or the Co-Op supplies healthier foods when compared to the Wal-Mart grocery.

However, all this data is admittedly somewhat bias because they weren't randomly allocated. If we were to do this project again, we would gather much more data at varied times during the day, on varied days of the week, and then take a randomized sample from each sample population.

Also, this data was very subjective. In order for an additional study to be considered more objective then this we would have to gather individuals' heights and weights and apply that to predetermined objective body-size groups. The optimal procedure for carrying out this observational experiment would be to find a way to randomly measure the body size of shoppers. Then again, this data would be very difficult to gather from people on a voluntary basis, as it is a sensitive subject for many people.

R Commands

```
library(NCStats)
obs<-matrix(c(2,16,17,6,16,13,8,24,3),nrow=3,byrow=TRUE)
obs
rownames(obs) <- c("Wal-Mart","County Market","Co-Op")
colnames(obs) <- c("skinny","normal","overweight")
obs
chi1 <- chisq.test(obs,correct=FALSE)
chi1
chi1$observed
chi1$expected
chi1$residuals
```