An Observational Study of the Usage of Reusable Bags at Safeway

By Leah Muncy and Miguel Ramirez

17 April 2017

Introduction: The purpose of this project was to either confirm or deny a relationship between the amount of groceries purchased and the usage of reusable shopping bags. We aimed to answer the following question: Do people have a greater tendency to bring their own reusable shopping bags for greater quantities of groceries purchased or for smaller quantities of groceries purchased?

To accomplish this, we looked at a simple random sampling of shoppers as they left Safeway in Sonora, CA on random days at random times of day, in order to potentially identify if the number of grocery bags in the shopper's cart is related to their usage of their own reusable bags. Shoppers were observed from one seating area near the checkout lanes, and were watched for 30 minute intervals. The store was visited a total of 5 times, with days and times chosen at random.

The shoppers were observed for the presence of either brown 10 cent shopping bags, or for their own reusable bags. The express self-checkout lanes were classified as the "small quantity of groceries purchased" group and the regular checkout lanes were the "large quantity of groceries purchased" group. Both shoppers with and without carts were considered, as persons purchasing only a small quantity of groceries were more likely to not use a cart.

Definitions and Assumptions: Our population was defined as the total number of shoppers in Safeway that made purchases and used a bag to carry those purchases within the selected intervals. Our sample are the shoppers observed within the five thirty-minute intervals. We assume that:
1. Shoppers using the 10 cent paper Safeway shopping bags were not brought beforehand by the shopper
2. Shoppers in the express self-checkout lanes are, in fact, buying a small quantity of groceries
3. Shoppers in the regular checkout lanes are buying a quantity of groceries greater than the amount recommended for the express lanes
4. Reusable bags are those that are plastic or paper bags not purchased from Safeway, or that the bags are plastic, nylon, etc
5. Some shoppers will have both reusable bags and purchased bags for larger purchases - these shoppers will be considered as having brought their own bags

Our initial hypothesis from our personal experiences shopping at Safeway is that people who make smaller purchases will purchase bags, while those who make larger purchases will have brought 1 or more of their own reusable bags.

Sampling Design and Methodology:
In order to gather this data, we used random.org to generate the five days in the month of March that would be chosen for observation, using numbers 1 through 31. The random number generator gave us the following days: March 4th, March 9th, March 12th, March 24th, March 28th. Next, using random.org, we randomly chose the hours at which the shoppers would be observed, labeling 7 to 21, as none of the members of the group could record data before 7am or after 8pm. The first time chosen was assigned to March 4th, and second to March 9th, etc. The following days and times are listed below:

March 4th - 8:00, March 9th - 17:00, March 12 - 14:00, March 24 - 11:00, March 28th - 14:00

Each group member was assigned 1 or 2 days to record data at the Safeway, observing both the self-checkout and regular lanes. If a shopper had at least 1 reusable bag, they were considered a "yes." If the shopper had no bags, or had no reusable bags, we marked "no." Because our sample size was 223 people, we used the two-sample z procedures. Because the shoppers were chosen randomly and our successes and failures, as a whole, were greater than 15, we found that it was safe to use the two-proportion z-tests of significance. Successes were those that had brought bags, and failures were those who did not bring bags.

Problems We Faced:
One of the main issues we discovered was that Safeway has both an "Express" lane and "Self-Checkout" lanes, making it unclear if we were going to draw a distinction between the two. We ended up classifying the "Express" lane as a regular lane, because it still required a cashier and bagger, and those in the Express lanes still often had more items than in the Self-Checkout lanes. Some shoppers also purchased reusable bags at the time of their purchase, which did confuse whether or not to categorize them as having brought their bags or not. We chose to categorize these people as having not brought bags, as they did not enter the store with them. Because of schedule restrictions, we also could not collect data during the early morning or times late at night. We would expect that most people are not doing large grocery shopping trips at those times, and therefore, there would have most likely been fewer people having brought their bags. Because we do not have data that represents these times, our data may be skewed to the right, representing perhaps more people on average having brought bags than the true average.

Results:
Using the two-proportion z hypothesis test and two-proportion z confidence interval, we first labeled our 2 proportions.

	TOTAL	# SUCCESSES	# FAILURES
$\hat{P}_1$ REGULAR LANE	n_1 138	79	59
$\hat{P}_2$ SELF-CHECK LANE	n_2 85	43	42

$$\hat{P}_1 = 0.57$$
$$\hat{P}_2 = 0.51$$

We then calculated a 95% confidence interval, in order to see if the value of 0 was within the interval. If 0 was in our interval, there would be significant evidence of no difference between the amount of groceries purchased and if shoppers will bring in reusable bags.

$$(p_1 - p_2) \pm z^* \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

Using a TI-84 Calculator, our 95% confidence interval for the difference between the proportions of those who bring bags to purchase a large quantity of groceries compared to those who purchase a small quantity of groceries is (-0.07, 0.20) or between -7% and 20%. Because the value of 0 lies within the interval, there may be significant evidence of no difference between the two proportions. Therefore, we calculated a two-proportion z-hypothesis test to test the significance of this finding using a TI-84 calculator.

$$z = \hat{P}_1 - \hat{P}_2 / [(\hat{P}(1-\hat{P})(1/n_1 + 1/n_2)]^{\frac{1}{2}}$$

$H_0: \hat{P}_1 = \hat{P}_2$

$H_a: \hat{P}_1 > \hat{P}_1$

As our hypothesis states, we believe that those who intend to purchase larger quantities of groceries, $\hat{P}_1$, will more frequently bring in their own reusable bags than those who intend on purchasing smaller quantities, $\hat{P}_2$.

The given values are from the TI-84 calculator using the 2-proportion z-test:

Z = 0.97

P = 0.17

At the $\alpha = 0.05$ level of significance, we keep H_0, and conclude that there is no significant evidence that a greater proportion of people will bring bags if they intend on making a larger purchase, rather than a smaller purchase. We therefore reject our initial hypothesis.

Discussion:

Our tests led us to conclude that there is not a significant difference in the amount of groceries purchased and whether or not a person will bring reusable bags to use for this purchase. Because we were only able to collect data during the hours of 7am to 9pm, our data may be skewed, favoring those who may make larger purchases, as larger purchases are typically not made during late nights or early mornings. Despite this, we simply do not have enough data to make any clear conclusion that those who make larger purchases are more likely to bring reusable bags than those who intend on making smaller purchases. We do have a P-value that is on the lower side, however, and thus, if more observations were to be performed, there is a possibility that there could ultimately be a significant difference between the two.

<u>Conclusion:</u>
Based on the data gathered, our results appear to reject our initial hypothesis that people who make smaller purchases will purchase bags, while those who make larger purchases will have brought 1 or more of their own reusable bags. Based on the data gathered during March 2017, during the days and hours observed at the Safeway in Sonora, California, there is no significant difference between the likelihood of a shopper to bring their bags depending upon intended quantity of groceries to purchase. Because we do not have a very large sample size, only 5 separate days were observed, and because not all hours were observed, there may be some right-skew in the data, as well as undercoverage. The margin of error of the data is also rather large, with proportions of each population initially very close, explaining our inability to confirm our initial hypothesis. While our two-proportion z-test did not yield a significant conclusion, with those intending on making larger purchases more likely to bring bags, if more data were to be collected, based on our previous observations, we believe that we would find that our hypothesis would be considered true. Logically, those who intend on needing lots of bags will prefer to supply their own and avoid paying the fee, and thus, it could be very likely that our hypothesis and conclusions could be confirmed with further testing.

Mar. 4, 2017: 17:00-19:00
__48 people all together__.

__23 brought bags__ : 15 in express lane, 8 in regular
__20 did not(bought 10 cent bags)__: 6 in express lane, 14 in regular
__5 people left without bags__: all five in express lane

REG 22:
8 yes
14 no
EXP 26:
15 yes
11 no

Mar. 12, 2017: 19:00-19:30
__29 people all together__.

__15 brought bags__: 4 in express lane, 11 in regular
__10 did not(bought 10 cent bags)__: 3 in express lane, 7 in regular
__4 people left without bags__: 3 in express lane, 1 in regular
REG 19:
11 yes
8 no
EXP 10:
4 yes
6 no

48 Mar. 28, 2017: 16:30
Self checkout 16:
9 no
7 yes

Regular 32:
20 yes
12 no

38 Mar. 9
Reg 28: 18 yes
10no

Self 10:
7 yes
3 no

60 Mar. 24
Reg37: 22 yes
15 no
Self 23:
10 yes
13 no

Thursday March 30
8:00-9:00 (somewhere between this time)
16 people total
9 brought bags: 6 in regular, 3 in express
3 didn't(bought): all regular
4 left without: all express