

# Project Part 3

Rachel Alexander

## **Question**

Do police officers treat black drivers with more suspicion than white drivers?

## **Data Description**

Interaction with police most commonly occurs at traffic stops. To uncover the prevalence of racial profiling in American policing, I have chosen police stop data from the state of North Carolina. The specific data set I am using for this project is the full population of all police searches conducted at stop light/signs in Charlotte, NC, in 2015. Each observation is the record of an individual traffic stop conducted by an officer of either the Charlotte-Mecklenburg Police Department or the UNC Charlotte University Police.

## **Data Relevance**

In this project, I will be comparing the hit rates of black and white and white drivers. The hit rate, defined here as the proportion of searches in which contraband is found, is a metric commonly used to test for racial bias in law enforcement (1). Given that one of the attributes recorded for each police stop is whether or not contraband was found, the hit rates of each racial group can easily be calculated and compared using my selected data set. If the hit rate of one group is higher than the other, we can conclude that members of that group are targeted only when officers are relatively certain contraband can be found, indicating that officers treat that group with less suspicion compared to other groups.

## **Generalization**

My police search data is a very small subset of the full population of searches recorded by Charlotte police. The observations I am examining, while forming the population of police searches at stop light/signs in Charlotte in 2015, are very specific to time and place. Police stop data has been collected in Charlotte and in cities across the U.S for over twenty years.

It would be difficult to make historical claims concerning traffic stop policing using evidence from any one single year. Additionally, drivers in Charlotte and across the United States are far more likely to be stopped for vehicle regulatory or equipment violations than for stop light/sign violations analyzed in this project. Approximately 45% of observations in my original raw data set occurred because of vehicle regulatory and equipment violations, while only 8% of observations occurred because of stop light/sign violations. Charlotte is also unique given its position as a major city and collegiate area. Attitudes towards diversity may not be reflective of the state North Carolina and the United States as a whole. Yet, it remains to be seen whether results from this data set can give us insight into policing in urban and/or college-town areas similar to where stops signs and traffic lights are most commonly found in Charlotte. Further research is required concerning the geographic locations of stop lights and stop signs represented by the searches in this data set.

## Test Selection

A **two sample z-test for proportions** will be conducted to determine if the hit rate, or proportion of searches which yield contraband, is higher for white drivers than black drivers.

Assumption 1: Populations are independent

In this test, the population of white drivers represented in the data is being compared to the population of black drivers represented in the data. Given that every observation in the overall data set is independent, we know that both populations of black and white drivers must be independent of each other as well. The race of one stopped driver does not affect the race of the next stopped driver, and thus the collection of black and white drivers in this data set are independent groups.

Assumption 2: Populations follow a binomial distribution

Mean and variance values of both populations are greater than 10, indicating both populations follow the binomial distribution which can be approximated by the normal distribution. The mean and variance of the white population are 27 and 49 respectively, and the mean and variance of the black population are 64 and 118.

## Test

```
#Setup: calculate population counts and corresponding hit rates
white_pop <- length(which(NC_stops$subject_race=="white"))
```

```
HR_white <- length(which(NC_stops$subject_race=="white" &
```

```

NC_stops$contraband_found==TRUE))/white_pop

black_pop <- length(which(NC_stops$subject_race=="black"))

HR_black <- length(which(NC_stops$subject_race=="black" &
                        NC_stops$contraband_found==TRUE))/black_pop

phat <- c(HR_white, HR_black)
n <- c(white_pop, black_pop)
x <- phat*n

prop.test(x, n, alternative="greater", correct=FALSE)

##
## 2-sample test for equality of proportions without continuity
## correction
##
## data:  x out of n
## X-squared = 0.0030684, df = 1, p-value = 0.4779
## alternative hypothesis: greater
## 95 percent confidence interval:
## -0.1038249  1.0000000
## sample estimates:
##      prop 1      prop 2
## 0.3552632 0.3516484

```

## Conclusions

Based on the results of the two-sample z test for proportions, there is not sufficient evidence to conclude that black drivers are treated with more suspicion than white drivers. In other words, the hit rate for white drivers is not greater than the hit rate for black drivers. Given that both groups had similar hit rates, it is reasonable that the difference between them would not be statistically significant. Thus, it follows the two-sample z test for proportions could not provide evidence of racial bias in policing practices at stop lights and stop signs in Charlotte, NC in 2015. Since these results are not generalizable outside this focused population, we cannot draw conclusions on the persistence of racial profiling in the state of North Carolina and the United States at large.

## References

1. <https://openpolicing.stanford.edu/>