1 Introduction

In this report our goal is to analyze the 2018 Boulder Police Data to answer two questions:

1. Is police bias occurring?

2. If it is occurring, where is it occurring most severely?

As we will see, the second answer is the more straightforward from a statistical perspective. Furthermore, given previous research and the testimony of policed groups, our priors favor it occurring, over it not occurring. We will therefore address the second first first. To do so we will use a rudimentary but tractable "lottery model" of police stops. The lottery model lacks rigor however, and is unable to address our first question. Furthermore, the 2018 Discretionary Stop Data is not sufficient to answer Question 1 in a general form. We will therefore use a more sophisticated "threshold model" to address the question not for all discretionary stops, but rather for speeding stops. We can also address this question for searches conducted during discretionary stops.

2 Models

Definitions:

- *Boulder Policed Population:* the population of all individuals whom the Boulder Police Department have access to, ie. who they could stop

2.1 Lottery Model

The lottery model of discretionary stops idealizes the stopping process, assuming that police perform stops entirely randomly. The process is modelled as if each individual of race/ethnicity r independently flips the same coin, and on a "heads" is subjected to a discretionary stop. From this it seeks to estimate the probability of a "heads" by race, denoted ρ_r, and more importantly, where ρ_r varies with r.

By definition, the discretionary stop data only includes information on individuals whose outcome was "heads", individuals who were stopped. We will therefore rely on the population estimates provided in the accompanying **?** to estimate the total number of "lottery participants" per demographic.

Letting s_r denote the number of stops of each race/ethnicity r, and n_r denote the total number of such individuals in the BPP. We then have the following:

$$\begin{pmatrix} \mathrm{P}\left[s_r = k\right] = \\ (n_r, k)\rho_r^k(1-\rho_r)^{n_r-k} \end{pmatrix} \tag{1}$$

We place uniform priors on each of the ρ_r to obtain the posterior distribution:

$$\begin{pmatrix} \mathrm{P}\left[\rho_r | s_r, n_r\right] \propto \\ (n_r, s_r)\rho_r^{s_r}(1-\rho_r)^{n_r-s_r} \end{pmatrix} \tag{2}$$

Eq. (2) could be rewritten as a join posterior over the multidimensional $[\rho_1, ..., \rho_N]$, however since each stop is assumed to be independent we can assume the same about the ρ_r.

We will use combine two metrics to determine where racial bias may be occurring most severely. First, we calculate the median black-white bias, that is, the posterior median of the distribution of $B_{\text{bw}} = \rho_b - \rho_w$. When B_{bw} is large and positive, it indicates a large racial disparity in stop occurrances. To determine the extent to which this represents a statistically signifcant bias, rather than due to simple noise, we will use the Bayes Factor (BF).

The BF is a (somewhat crude) method of comparing two models, M_1 and M_2, each modeling the same dataset D. For us, M_1 will denote model (2) while M_2 denotes a similar model, but where $\rho_b = \rho_w$. The BF is defined as the ratio:

$$\text{BF} = \frac{\mathrm{P}\left[M_1|D\right]}{\mathrm{P}\left[M_2|D\right]}$$

Applying Bayes Theorem to the top and bottom one can show:

$$\text{BF} = \frac{\text{P}\ [D|M_1]}{\text{P}\ [D|M_2]}$$

Where the P $[D|M]$ is the normalizing, "evidence" term from the posterior, P $[D|M] = \int \text{P}\ [D,\theta|M]d\theta = \int \text{P}\ [D|\theta, M]\text{P}\ [\theta|M]d\theta$.

When the BF is large, it indicates that M_1 is significantly more conistent with the data than model M_2, and vice versa when it is small. The BF can often exhibit extreme values, particularly when one model is heavily favored over another. For computational tractability we will thus work on the log scale.

2.2 Treshold Model

The lottery model is extremely simplified, and can only determine whether a discrepancy in stop rates exists, not if that discrepancy reflects bias on the part of the stopping officer. A slightly more

$\rho_r = \frac{1}{n_r} \sum_i \nVdash(\alpha_i > \tau_{r_i})$

Assume that the BPD have thresholds τ_r for each race/ethnicity pairing r, and that if individual i has $\alpha_i > \tau_{r_i}$ then the BPD will initiate a discretionary stop. The ultimate question of interest here, is whether the τ_r vary across r. However we will not be able to answer this question in general. Instead we will attempt to answer several, more limited questions, which are indicative that the τ_r do indeed vary.

First, we will asses $\rho_r = P[$

References