Title: A METHOD TO IDENTIFY CHANGES IN PEDESTRIAN AND CYCLIST COLLISION RATE

Presenting Author: Offer Grembek

Authors: O. Grembek 1;D. R. Ragland 1;

Affiliation
1. Traffic Safety Center, University of California, Berkeley, CA, USA,

Abstract:
The identification of high collision locations is well-studied in transportation safety. However, insights regarding high collision locations are not always downward scalable when collision counts are orders of magnitude smaller, as in the case of pedestrian and bicyclist collisions. The problem with such small numbers is that for an individual intersection, an annual increase of one or two collisions could be the only indication that an intersection's underlying risk has changed. It is therefore very difficult to determine what collision count would mandate further investigation. This is important because it may lead to inefficient use of resources. The objective here is not to predict high collision locations, but rather to identify locations that indicate a systematic change in the underlying pedestrian and bicyclist's collision rate.

In this study we use actual collision counts to create a longitudinal confidence region around the minimum collision rate of an intersection. This procedure takes into account both the mean and the variance of the collision counts to estimate the minimum underlying collision rate. However, we have shown that the variance for an individual intersection may still be too large to identify any statistical change in risk. Therefore, intersections are subsequently clustered and confidence intervals are constructed for each cluster.

We begin by explaining how we construct a confidence region for a single intersection. We first choose an initial collision rate and construct upper and lowers bounds assuming a Poisson distribution. We then follow the same procedure to construct the bounds around an average of two Poisson random variables and so on up to an average of 20 Poisson random variables. The result is a continuously decreasing confidence region around the initial collision rate. We then plot the actual collision counts for an intersection and check if they all fall within the boundaries. We developed an algorithm to modify the initial condition and determine the smallest collision rate that would contain all observations. When this happens we have obtained an estimate for the minimum collision rate as indicated by the collision data. Any future observations that fall within these boundaries are not statistically different from historical data, for a specific level of confidence. Since our bounds only represent the minimum collision rate we cannot determine if observations outside the bounds actually indicate a change in collision rate. However, we can alleviate some of this deficiency by systematically clustering intersections.

To cluster the intersections we obtain the confidence region around the minimum collision rate for all intersections of interest. We use K-means cluster analysis to cluster the intersection by similar collision rates and construct an aggregate confidence interval for each of the clusters. We use these clusters to evaluate whether future collision counts represent probabilistic variation or require further investigation by the responsible agency.

We tested this technique using annual collision counts from 1997 to 2007 for a 14 intersection section along San Pablo Avenue (a major arterial in the San Francisco Bay Area). The analysis showed that using four years of historical data can identify temporal changes in an intersection's collision rate using only collision counts. These results mandate further development of this procedure to evaluate the performance of alternative clustering approaches that capture the natural tendency for spatial clustering and to evaluate what we can gain from introducing additional information about the intersection's features.

The technique presented here serves two purposes: (i) identify when an intersection's collision rate is not statistically consistent with historical data; (ii) evaluate if modifications designed to affect the underlying collision rate resulted in a significant change in collision rate.