6. Safety Performance functions- tools for improved use of safety and exposure data

Lars Ekman

This is essentially the summary of my report "On the Treatment of Flow in Traffic Safety Analysis - a non parametric approach applied on vulnerable road users". The full report is available at our department.

The treatment of flow in the area of traffic safety has a long tradition. The influence flow has on the number of accidents is, however, often considered so obvious that it tends to be trivial. I can show that the relation between flow and accidents holds interesting information. Knowledge about such relations may be very useful for several purposes.

Traffic flow counts or other expressions of the traffic exposure are often only used to construct accident rates. Quite often the concept of accident rates is used synonymously with "risk". There is, however, a great risk '!' in mixing these two concepts. The word 'risk' is spontaneously interpreted as danger or hazard for someone involved in a certain activity. In common life, risk is used as a very broad concept including both the probabilities of an unwanted event, as well as the consequences of this event. The concept of accident rate is often used to avoid this "over-interpretation", but nevertheless also this concept needs to be treated with care.

Based on the way accident rates are used three main streams in the use of accident rates and exposure have been found. These three groups were derived from how different researchers used or interpreted their results. The three approaches are:

- The probability approach
- The effectiveness approach
- The standardization approach

The first approach is what is referred to when discussing accident rate in general. The other two approaches constitute a classification system based on the way in which the computed rates seemed to be interpreted.

In order to compare countermeasures, or conduct other types of traffic safety comparisons where the flow varies, it is vital to know the full shape of the relation between accidents and flow. Such relations are called Safety Performance Functions (SPF). If we focus on the situation for individual road users the equivalent Risk Performance Functions (RPF) may be used to increase the comparability. The traditional comparison of accident rates is equal to assuming the SPF is a straight line, and thus the RPF is a constant.

The main aim of this work is to develop a "transparent" system for estimating SPFs and RPFs. One step towards a transparent and, thus, interpretable treatment of accident and flow data is the development of a system for aggregating approaches in order to create aggregates with

equal, or at least manageable, precision. The aggregation is based on relevant flow and is made in such a way that all aggregates represent the same number of the relevant road users.

In order to generate a non-parametric function, without built-in presumptions of the overall relationship between flow and accidents, moving averages line is used. A series of computer programs is developed in order to describe the accuracy of the resulting functions. Two computer intensive methods: simulation and bootstrap, are used. With both these methods, "exact" confidence intervals are produced.

Bootstrap is based on random selection, with replacement, from the original data observations. With this method several new data sets are constructed. The procedure is completed all the way throughout the analyses, over and over again, each time producing a bootstrap replicate of the end result. In order to estimate the confidence intervals the procedure has to be repeated about 1000 times.

Confidence intervals are then computed by a process of interpolation. The confidence intervals may only be produced in the range where there are several observations in the original data set. A special routine to handle this "edge problem" has also been developed.

The stability of and the power of the computation of confidence intervals is tested with the use of synthetic data sets produced by a random process. The 80% confidence intervals seems to be slightly "conservative" i.e. covering the true value in more than 80% of the cases.

The method developed was applied to a data set consisting of accident records, conflict observations and traffic flow counts for different road user categories from 95 non-signalized intersections in the cities of Malmö and Lund.

The result on bicycle and pedestrian safety is illustrated by the following Risk Performance Functions:

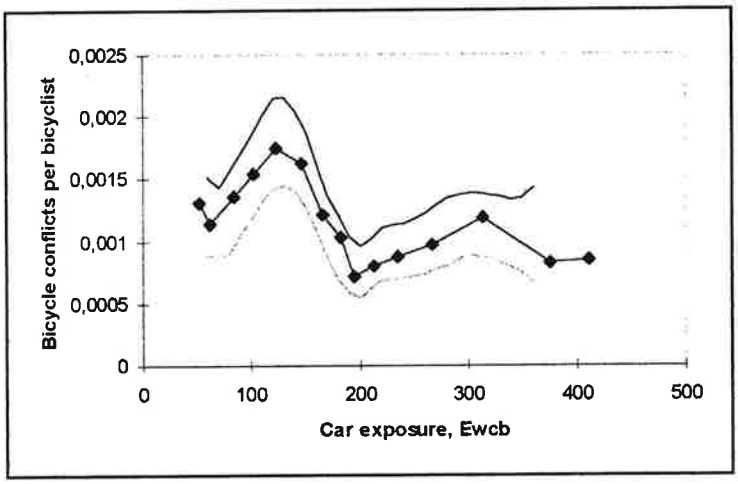

Figure 0.1:
Bicycle conflicts per bicyclist versus accumulated and weighted car flow (Ewcb). Moving average line (RPF) with an estimated 80% confidence interval

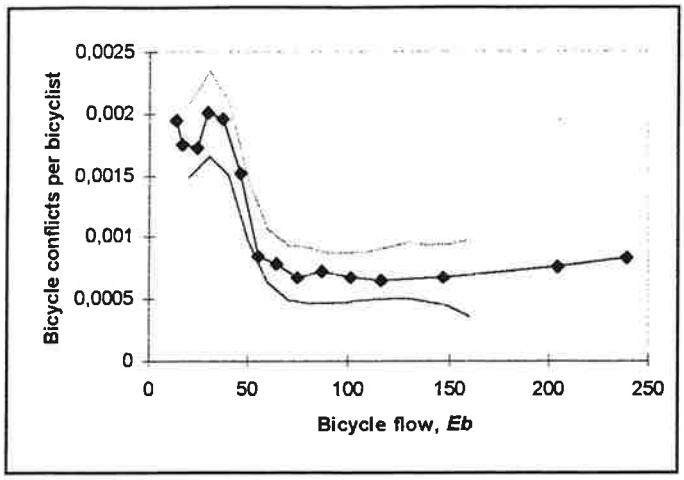

Figure 0.2:
Bicycle conflicts per bicyclist versus bicycle flow(Eb) Moving average line (RPF) with an estimated 80% confidence interval

The results regarding bicycle safety are:

- the conflict rate for bicyclists is twice as large at locations with low bicycle flow compared to locations with higher flow

- the conflict rate for bicyclists seems to be high at locations within a limited range of car flow

- bicycle flow seems more significant than car flow for the conflict rates for individual bicyclists

- the conflict rate is generally higher for bicyclists approaching intersections from the minor street compared to those coming from the main street

- the difference in conflict rate between low and high bicycle flow is larger than differences due to the design variables tested: the width of the road and major or minor street

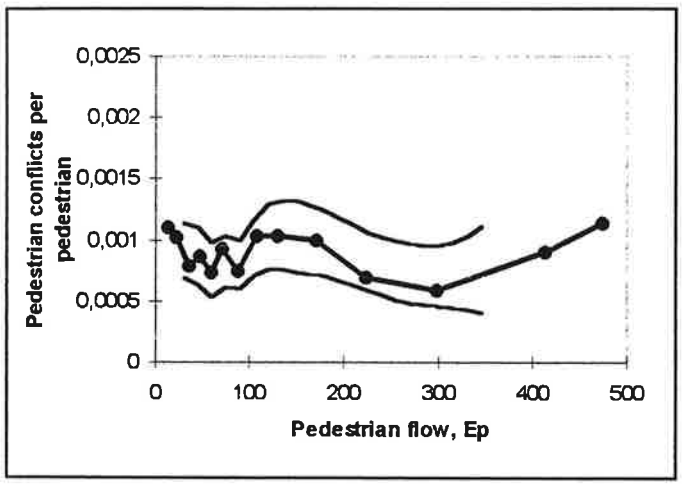

Figure 0.3:
Pedestrian conflicts per pedestrian versus pedestrian flow. Moving average line (RPF) with an estimated 80 % confidence interval

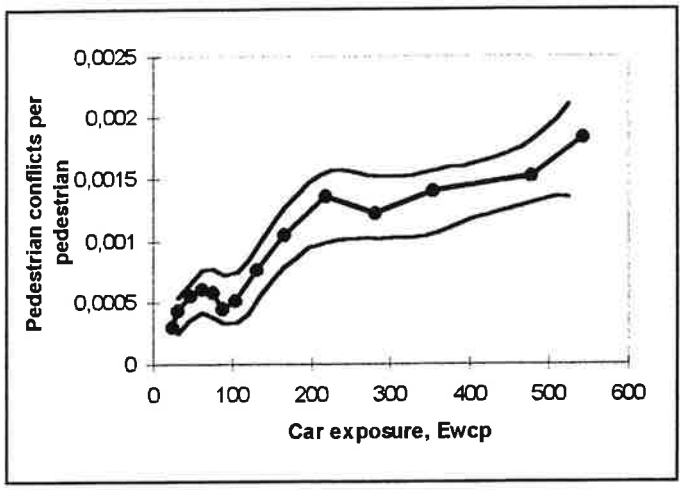

Figure 0.4:
Pedestrian conflicts per pedestrian versus car exposure (incoming cars per hour). Moving average line (RPF) with an estimated 80% confidence interval

The results regarding pedestrian safety are:

- the conflict rate for pedestrians is largely unaffected by pedestrian flow
- car flow seems to be of great importance for the conflict rate for pedestrians
- the increase in conflict rate by increasing car flow seems to be a "step function" rather than a linear function
- zebra marking seems to increase the conflict rate for locations with low pedestrian flow, irrespective of car flow
- refuge seems to give a decrease in conflict rate which is larger than the increase "caused" by the zebra marking
- it has not, however, in this study, been possible to show that the width of the street or whether the street is a main street or minor street, has any significant impact on conflict rate for pedestrians

The empirical application has confirmed that:

- the method of aggregating and averaging gives a function with good visual interpretability
- the bootstrap method gives an accurate and useful description of the stability of the estimated non-parametric function
- the relationship between conflicts and flow is complex
- knowledge about such relations could improve traffic safety evaluations that include comparisons between groups of locations with different flow
- analyses of the non-parametric functions may be the base for suggesting active flow manipulation as a traffic safety measure

- information about the effect of flow on safety may be used to generate hypotheses about the processes underlying traffic safety problems

Thus, the question is not anymore whether flow contains valuable information for traffic safety analyses, but rather, if we are going to make use of that information in improving traffic safety.