FINAL REPORT
A PROBABILISTIC APPROACH TO TRAFFIC PROBLEMS: PHASE III

October 31, 1971
The Florida State University
Tallahassee, Florida

Prepared under Contract No. DOT-FH-11-7661 with the U.S. Department of Transportation, Federal Highway Administration. The opinions, findings and conclusions expressed in this publication are those of the authors and not necessarily those of the Federal Highway Administration.

Ralph A. Bradley
Principal Investigator

Robert J. Serfling
Primary Co-Investigator

Ismail N. Shimi
Primary Co-Investigator

George S. Kimeldorf
Co-Investigator
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>General Activity</td>
<td>3</td>
</tr>
<tr>
<td>Description of Research</td>
<td>4</td>
</tr>
<tr>
<td>Abstracts of Technical Reports (appended)</td>
<td>A1</td>
</tr>
<tr>
<td>Technical Reports (appended)</td>
<td></td>
</tr>
<tr>
<td>Gordon Pledger and R. J. Serfling</td>
<td>M203</td>
</tr>
<tr>
<td>2. &quot;Improvements on Robust Estimation through Iteration,&quot;</td>
<td></td>
</tr>
<tr>
<td>I. N. Shimi and Douglas A. Wolfe</td>
<td>M205</td>
</tr>
<tr>
<td>3. &quot;Moment Formulas for ([X/m]),&quot;</td>
<td></td>
</tr>
<tr>
<td>I. R. Savage and R. J. Serfling</td>
<td>M206</td>
</tr>
<tr>
<td>R. J. Serfling and Constance L. Wood</td>
<td>M207</td>
</tr>
<tr>
<td>5. &quot;Comparison of Two Series of Traffic Flow Headways,&quot;</td>
<td></td>
</tr>
<tr>
<td>J. P. Lang and R. J. Serfling</td>
<td>M208</td>
</tr>
<tr>
<td>L. H. Crow and I. N. Shimi</td>
<td>M209</td>
</tr>
<tr>
<td>D. W. Fairweather and I. N. Shimi</td>
<td>M210</td>
</tr>
</tbody>
</table>
1. **Introduction**

This report summarizes the research and development study performed by the Department of Statistics, Florida State University, for the Federal Highway Administration, U. S. Department of Transportation, under Contract No. DOT-FH-11-7661. A follow-on to Contracts No. FH-11-6680 and No. FH-11-6890, the present contract began on July 15, 1970 and has a termination date of October 31, 1971.

Section 2 of this report catalogs general activity on the contract, while in Section 3 the research work is described in a technical way. Appended (A1) are abstracts of technical reports submitted under the contract, followed by the reports themselves (M203, M205, M206, M207, M208, M209 and M210).

Expanding the specification of work initiated under Contracts No. FH-11-6680 and No. FH-11-6890, the objectives of the project are to describe essential relationships of traffic flow models, to investigate heavy traffic situations, to design efficient statistical procedures for parameter estimation and tests of validity of models, and to solve specific problems in probabilistic traffic theory. The Contractor's research proposal (entitled "A Probabilistic Approach to Traffic Problems: Phase III," dated March 15, 1970) sets forth a variety of problem areas for potential investigation, among which the Contract designates as of high priority:

I. **PROBABILISTIC MODELING OF TRAFFIC FLOW**

A. A class of models relevant to traffic flow.

B. A model for flow produced by interacting Poisson flows.
C. Applications of alternating renewal models and mixtures of point processes.

E. Fundamental relationships for traffic flow models.

F. Counting models for flow grouped into categories.

II. HEAVY TRAFFIC THEORY

A. The waiting time of a vehicle in queue.

C. Models for formation and dissipation of congestion.

D. Detection of incidents on a freeway.

E. Control of access to heavy traffic systems.

III. STATISTICAL METHODOLOGY

A. General techniques for estimation of parameters and tests of hypotheses.

B. Parameters related to capacity and level-of-service.

C. Comparison of two series of gaps.

D. Counting time necessary for accurate estimation.

IV. MISCELLANEOUS TOPICS

C. A macroscopic study of traffic flow with stochastic inflow.

Research work on these objectives has led to seven technical reports issued under the present contract. These are:


5. "Comparison of Two Series of Traffic Flow Headways,"
   J. P. Lang and R. J. Serfling.

   I. N. Shimi.

   D. W. Fairweather and I. N. Shimi.

A number of articles based on research conducted under this series
of contracts have been published in professional journals. "The time
to drive through a no-passing zone," by Vincent Hodgson, has appeared
by R. J. Serfling, has appeared in Transportation Research, 1969. "The
variance function of the Erlang process," by R. J. Serfling, has appeared
relationships in a stationary point process," by R. J. Serfling (with
H. Cramér and M. R. Leadbetter), has appeared in Z. Wahrscheinlichkeitstheorie,
1971. "Point processes for traffic flow," by R. J. Serfling, has been
Further articles are in preparation.

2. General Activity

Activity commenced on July 15, 1970, whereas formal notification of
the awarding of the contract was received on December 1, 1970.

Some time was spent in business activity concerning the final report
for Contract No. FH-11-6890, the predecessor of the present contract,
and concerning the contractual details of the present contract. Much of
the latter activity was generated by the Contracts and Procurement Division
of the Federal Highway Administration.
In lieu of a site visit to Florida State University by members of the Traffic Systems Division, members of the Department of Statistics were invited to visit the Traffic Systems Division. The meeting took place on March 26, 1971, with Florida State University represented by Professors Serfling and Shimi and the Traffic Systems Division represented by Dr. John Eicher, Dr. Sidney Weiner and Mr. Harry Lum. The discussions dealt with the current research activity and with the need to communicate technical results in a qualitative way to a non-technical readership.

In addition to the technical reports submitted as completed, quarterly business reports were submitted as well as the appropriate INTERIM REPORT in accordance with the Contract.

The orientation of the Department of Statistics toward probabilistic traffic problems has been under continual development, and this has enhanced productivity under the Contract. In particular, collaborations have taken place between R. J. Serfling and Professor I. R. Savage, Dr. Gordon Pledger, Mr. Patrick Lang and Miss Constance Wood, and between I. N. Shimi and Drs. Douglas A. Wolfe, D. W. Fairweather and L. H. Crow.

In Summer, 1970, I. N. Shimi taught a course on Renewal Theory, in which emphasis was given to traffic problems.

The work of coordinating and supervising all of the above activities has been carried out by Professor Ralph A. Bradley.

3. Description of Research

In Section 1 the research objectives of the contract were cited and seven technical reports listed. While the content of these reports may
be found from the appended material, here we indicate relationships to the high-priority topic areas.

<table>
<thead>
<tr>
<th>Report</th>
<th>Relevant topics</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>II.A</td>
</tr>
<tr>
<td>2</td>
<td>III.A</td>
</tr>
<tr>
<td>3</td>
<td>I.D, III.A, III.D</td>
</tr>
<tr>
<td>4</td>
<td>III.A</td>
</tr>
<tr>
<td>5</td>
<td>III.C</td>
</tr>
<tr>
<td>6</td>
<td>III.A</td>
</tr>
<tr>
<td>7</td>
<td>I.F</td>
</tr>
</tbody>
</table>

Thus emphasis area I ("Probabilistic modeling of traffic flow") is involved in Reports 3 and 7, emphasis area II ("Heavy traffic theory") is involved in Report 1, and emphasis area III ("Statistical methodology") is involved in Reports 2, 3, 4, 5 and 6.

The thrusts of this research are as follows. Regarding the **modeling** of flow, progress has been made (i) in providing improvements to the Poisson model without sacrificing computational feasibility, (ii) in characterizing the time-varying lane decomposition of a given cluster of vehicles. Regarding **heavy traffic**, the effect of dependence among headways upon heavy traffic parameters has been characterized. Regarding **statistical methodology**, effort has concerned (i) estimation techniques not requiring Poisson assumptions, (ii) increased utility of count data as opposed to headway data, (iii) "before-and-after" comparisons of traffic flow properties, (iv) consideration of appropriate alternatives in testing Poisson goodness-of-fit.

On the following pages we give abstracts of the technical reports.
1. **The Waiting Time of a Vehicle in Queue.**

FSU Statistics Report M203, by Gordon Pledger and R. J. Serfling

**Abstract**

The distribution of the waiting time of the n-th customer in a single-server queue is approximated for n large and the queue in "heavy traffic." By the latter is meant that the excess of mean inter-arrival time over mean service time is so small that the stability of the queue is precarious. Generalizing Kingman's "heavy traffic" approximation, the present result applies under broader dependence restrictions and not only when the queue is in equilibrium but also during the transition period. The dependence structure of inter-arrival times and service times and the degree of progress toward equilibrium are reflected in the parameters of the approximation.

The motivating application is road traffic. Particular attention is devoted to the mean waiting time and to the probability of exceedance of a critical level. The influence of the headway correlations upon the mean waiting time is examined. Exceedance probabilities are compared under pre-equilibrium and equilibrium conditions.
2. Improvements on Robust Estimation Through Iteration, FSU Statistics Report M205, by I. N. Shimi and Douglas A. Wolfe

Abstract

In a large number of traffic flow problems it is sufficient to know, or estimate, a few of the moments of the distributions involved to get some useful results. In most cases one cannot determine the exact forms of the distributions involved but can guess either from previous experience or from information provided by the sample, that the distribution belongs to some family of distributions. Say heavy-tail or light-tail distributions, or one of a possible number of distributions. Estimators that are "best" for the moments of some distributions need not be suitable to estimate the same moments of some other type of distribution.

In the recent literature a number of articles are concerned with the problem of robust estimation. Although there have been many techniques proposed for obtaining robust estimators for the location of a distribution, few seem to take full advantage of the information provided by the sample. R. V. Hogg ("Some observations on robust estimation" JASA (1967), p. 1179 - 1186) proposed a class of robust estimators that are extremely appealing. He effectively uses the information in the sample concerning the "tails" of the underlying distribution, namely the sample kurtosis, to select an estimator for the location of the distribution from a suitable class of estimators. The robustness
properties of his class of estimators are a direct consequence of this "sample-peeking" selection procedure of the estimator and not necessarily because of any robustness properties of the actual statistic selected from the available class.

In this report, we propose iterative modifications of a slightly more general scheme than Hogg's. Such modifications maintain the excellent asymptotic properties of the point estimators selected by his procedure and greatly improve upon their small sample properties. Empirical investigations comparing the modified estimation procedure with Hogg's original proposal are presented. The results of these studies show that much can be gained by using the iterative procedure, particularly when there is some prior knowledge available about the weight of the "tails" of the underlying distribution.
Given a random variable \( X \) and a positive integer \( m \), let the random variable \( Y = \lfloor X/m \rfloor \) denote the greatest integer \( \leq X/m \).

This report gives formulas for the characteristic function of \( Y \) in terms of that of \( X \) and useful approximating formulas for the moments of \( Y \) in terms of those of \( X \).

This applies to point processes for traffic flow as follows. Given a model for a series of events, one may define a new model by considering every \( m \)-th event of the original series. For example, starting with a Poisson model, the resulting model is the so-called Erlang model of order \( m \). One may then derive properties of the counting distributions of the new model in terms of those of the original model, since the count in a given observation period in the new model is a random variable having the distribution of \( \lfloor N/m \rfloor \), where \( N \) is the corresponding count variable in the original model.

A special advantage of these results is that it provides an improved basis for using count data instead of headway data. In particular, regarding the Erlang model, the results are of use in fitting an Erlang model to count data, testing Poisson versus Erlang, and estimating the correlation of an Erlang model.

Applications to other "every \( m \)-th event" type models may also be effected.
4. Testing Poisson Versus Erlang, Using Count Data

FSU Statistics Report M207
by R. J. Serfling and Constance L. Wood

Abstract

A central role in modeling traffic flow is played by the Poisson model, which is accurate in the case of light density flow and is easy to manipulate mathematically. For moderately high density flow, the Erlang model is a generalization which has theoretical appeal but is somewhat difficult to manipulate. This report lays foundations supporting greater practical use of the Erlang model as an alternative to the Poisson.

A limitation of current procedures is that headway data is required. The present report develops a procedure utilizing count data, which is usually more convenient to acquire. The statistic receiving primary attention is the sample index of dispersion based on Erlang counts. Its asymptotic distribution theory is obtained and application to estimation of parameters and testing of hypotheses is treated.
5. Comparison of Two Series of Traffic Flow Headways
FSU Statistics Report M208
by J. P. Lang and R. J. Serfling

Abstract

A fundamental hypothesis in traffic flow studies is that two sequences of headways have identical properties. The two sequences may arise from two separate streams of traffic or from two portions of the same stream. For example, the hypothesis of stationarity of traffic flow may be tested by testing the hypothesis that two sequences of headways observed over different time periods have identical properties. Or, in investigating the effect of a new control installation, it may be relevant to compare a sequence of headways observed before the installation with a sequence observed afterwards.

It is of interest to test such a hypothesis without assuming that headways within a sequence are uncorrelated, and without assuming a knowledge of which general model (e.g., Poisson) has given rise to the headways studied. This report extends the applicability of the well-known Wilcoxon statistic, as a large-sample nonparametric test, to the problem just described. Further, the likelihood ratio test under Poisson assumptions is derived and its large-sample properties developed. Finally, the comparison of these competing procedures is discussed.
   by L. H. Crow and I. N. Shimi

   **Abstract**

   The present report concerns the maximum likelihood estimation of the parameters of a wide class of distributions relevant to gaps between cars. This class is characterized by a probability density function
   \[ f(x) = \lambda g'(x) \exp(-\lambda g(x)), \quad 0 < x < T, \]
   where \( g(x) \) is an increasing differentiable function which tends to infinity as \( x \) tends to infinity. This class includes for example the exponential, the Weibull and the extreme-value distributions. We are interested in estimating the unknown parameter \( \lambda \).

   What is novel about the results presented in this report is its special appropriateness to data collected in traffic studies. Two possible ways of taking the observations can be used. Considering a stretch of fixed length of a one-lane highway or of a particular lane of a multi-lane highway, overhead pictures are taken of the configuration of the cars on this stretch. This can be done either by the use of a helicopter or by a fixed camera lying directly over the highway. Another way of collecting the desired data might be by the use of a recording device measuring the gaps between cars that pass the device over fixed periods of time, each starting with the passing of a car.
It is shown that the usual techniques for estimating the parameter \( \lambda \) using observations collected in the ways considered in this report, and actually being utilized in data collecting for traffic studies, are not appropriate. Using these special sampling techniques, an estimator is given for the parameter \( \lambda \). This estimator is the maximum likelihood estimator for \( \lambda \) and it is shown to possess the usual "good" properties of estimators. For example, strong consistency, asymptotic normality and asymptotic efficiency are some of the properties of the given estimator. Also, some useful and basic formulas are derived.
7. A Counting Model for Flow Grouped into Categories,
FSU Statistics Report M210, by D. W. Fairweather and I. N. Shimi

Abstract

Consider a group of cars moving in the same direction on a multi-
lane highway where the maximum distance between any two in the group does not exceed a fixed value, one mile say. A car is said to be in the group if it is still within this fixed distance from each of the other cars in the group. The cars in the group are classified according to their lanes of travel. Suppose we start observing a group of cars with \( N_0 = (n_{01}, \ldots, n_{0k}) \) cars of different types, where \( n_{0i} \) is the number of cars of type \( i \) in the group at the start. As time passes a car of some type \( i \), say, might join the initial group. This can happen if the group overtakes a slow moving car or if the group is overtaken by a faster moving car or by a car merging within the group. When a car joins the group a perturbation occurs that might cause the cars of the different types to change their types, by changing lanes for example. This gives a new vector \( \mathbf{N}_1 = (N_{11}, N_{12}, \ldots, N_{1k}) \) representing the new numbers of cars in the different types. We call this the first generation group. Now this first generation group will be perturbed similarly by a new car that might join it, giving us a second generation vector \( \mathbf{N}_2 = (N_{21}, N_{22}, \ldots, N_{2k}) \). The process repeats itself. One might be interested in the properties of the \( n \)-th generation vector \( \mathbf{N}_n \).
In this report we find the joint distribution for the number of cars of the different types in the \( m \)-th generation, the expected number of cars of the different types in the \( m \)-th generation, the asymptotic expected values and limits of the random vectors \( \mathbf{N}_m \) as \( m \to \infty \). We assume that the type of car that joins the group is \( i \) with probability \( p_i \geq 0, \sum_{i=1}^k p_i = 1 \). We also assume that the type of car joining the group at any time depends only on the probability distribution \( \{p_i\} \) and not on the group it is joining.

In this report we have used some of the results in our paper "An Immigration and Fragmentation Stochastic Process" \textit{J. Math. Biosciences.} to appear.