Teknillinen korkeakoulu
Liikennetekniikka
Julkaisu 87

Statistical Analysis of VEHICLE TIME HEADWAYS

R. Tapio Luttinen

TL Consulting Engineers, Ltd.

Dissertation for the degree of Doctor of Technology to be presented with due permission for public examination and debate in the Auditorium at the Helsinki University of Technology Lahti Center (Neopoli, Lahti, Finland) on the 31st of May, 1996, at 12 o'clock noon.

Otaniemi 1996

Julkaisu—Publication:

Luttinen, R. Tapio, *Statistical Analysis of Vehicle Time Headways.* Teknillinen korkeakoulu, Liikennetekniikka, Julkaisu 87. Otaniemi, 1996.

Luttinen, R. Tapio, *Statistical Analysis of Vehicle Time Headways.* Helsinki University of Technology, Transportation Engineering, Publication 87. Otaniemi, 1996.

This book was prepared with LaTeX and reproduced from a Postscript file supplied by the author.

Second corrected printing

ISBN 951-22-3063-1
ISSN 0781-5816

Printed in Finland by Gummerus Printing, Saarijärvi, 1996

To Tarja, Mikko, Hanna, and Jaakko

For want of skilful strategy an army is lost;
victory is the fruit of long planning.

Proverbs 11:14 (The New English Bible)

Luttinen, R. Tapio, *Statistical Analysis of Vehicle Time Headways.* Teknillinen korkeakoulu, Liikennetekniikka, Julkaisu 87. Otaniemi, 1996. ISBN 951-22-3063-1. ISSN 0781-5816. UDC 656.021.

Keywords: traffic flow, headways, capacity, level of service, statistics, parameter estimation, goodness of fit, moving probability, Monte Carlo method

Abstract

The properties of vehicle time headways are fundamental in many traffic engineering applications, such as capacity and level of service studies on highways, unsignalized intersections, and roundabouts. The operation of modern vehicle-actuated traffic signals is based on the measurement of time headways in the arriving traffic flow. In addition, the vehicle generation in traffic flow simulation is usually based on some theoretical vehicle time headway model.

The statistical analysis of vehicle time headways has been inadequate in three important aspects: 1) There has been no standard procedure to collect headway data and to describe their statistical properties. 2) The goodness-of-fit tests have been either powerless or infeasible. 3) Test results from multi-sample data have not been combined properly.

A four-stage identification process is suggested to describe the headway data and to compare it with theoretical distributions. The process includes the estimation of the probability density function, the hazard function, the coefficient of variation, and the squared skewness and the kurtosis. The four-stage identification process effectively describes those properties of the distribution that are most helpful in selecting a theoretical headway model.

One of the major problems in the headway studies has been the method for goodness-of-fit tests. The two most commonly used tests are the chi-square test and the nonparametric Kolmogorov-Smirnov test. The chi-square test is not very powerful. The nonparametric Kolmogorov-Smirnov test should be applied only, when the parameters of the distribution are known. If the parameters are estimated from the data, as in typical headway studies, the nonparametric Kolmogorov-Smirnov test gives too conservative results. These problems are addressed by parametric goodness-of-fit tests based on Monte Carlo methods.

Another great problem has been the lack of theoretical foundation in dealing with multi-sample data. The headway data usually consist of several samples, and the null hypothesis is tested against each of them. Two methods are presented to strengthen the evidence of multi-sample tests: 1) The combined probability method gives a single significance probability based

on several independent tests. 2) The moving probability method is used to describe the variation of combined probabilities against traffic volume.

These methods were applied to time headway data from Finnish two-lane two-way roads. The independence of consecutive headways was tested using autocorrelation analysis, runs tests and goodness-of-fit tests for the geometric bunch size distribution. The results indicate that the renewal hypothesis should not be accepted in all traffic situations. This conclusion is partly supported by a further analysis of previous studies.

Five theoretical distributions were tested for goodness of fit: the negative exponential distribution, the shifted exponential distribution, the gamma distribution, the lognormal distribution and the semi-Poisson distribution. None of these passed the tests. In the parameter estimation the maximum likelihood method was preferred. For distributions having a location (threshold) parameter, a modified maximum likelihood method was shown to give good estimates.

The proposed procedures give a scientific foundation to identify and estimate statistical models for vehicle time headways, and to test the goodness of fit. It is shown that the statistical methods in the analysis of vehicle headways should be thoroughly revised following the guidelines presented here.

Contents

Glossary of Notation

A^2	Anderson-Darling statistic
A-D	Anderson-Darling
ANCOVA	Analysis of covariance
BAN	Best asymptotically normal
BLUE	Best linear unbiased estimator
C	Sample coefficient of variation
$C(T)$	Coefficient of variation of the distribution for T
corr(x,y)	Correlation coefficient between x and y
c.v.	Coefficient of variation
D	Deterministic distribution
D	Kolmogorov-Smirnov statistic
$E(T)$	Expectation for random variable T
EDF	Empirical distribution function
edf	Empirical density function
ehf	Empirical hazard function
eq.	Equation
exp(x)	Exponential function (e^x)
$F(t)$	Probability distribution function ($\mathbb{P}\{T \leq t\}$)
$F_n(t)$	Empirical probability distribution function for sample of size n
$f(t)$	Probability density function
$f^*(s)$	Laplace transform of $f(t)$
fig.	Figure
G	General distribution
HCM	Highway Capacity Manual
$h(t)$	Hazard function
i.i.d.	Independent and identically distributed
$K(x)$	Kernel function
K-S	Kolmogorov-Smirnov
M	Exponential distribution

M'	Modified exponential distribution
m	Number of replications, number of samples
m_k	The kth central sample moment $\sum_{j=1}^{n}(x_j - \bar{x})^k/n$
m'_k	The kth noncentral sample moment $\sum_{j=1}^{n} x_j^k/n$
MCS	Minimum chi-square
Md	Median
MLE	Maximum likelihood estimator
MME	Method of moments estimator
MMLE	Modified maximum likelihood estimator
MMME	Modified method of moments estimator
Mo	Mode
MVUE	Minimum variance unbiased estimators
$N(t)$	Number of events in $(0, t]$
$N(a,b]$	Number of events in $(a, b]$
$N(\mu,\sigma)$	Normal distribution with mean μ and standard deviation σ
n	Sample size
$o(\Delta)$	$\lim_{\Delta\to 0} o(\Delta)/\Delta = 0$
P	Significance probability (P-value of a test), combined probability
$P_k(\tilde{\lambda}_{k,j})$	The jth term of a k-point moving probability
p	Probability, proportion of follower headways
p_i	Significance probability (P-value of a test) for sample i
P(A-D)	Significance probability of the Anderson-Darling test
P(K-S)	Significance probability of the Kolmogorov-Smirnov test
PDF	Probability distribution function
pdf	Probability density function
Ph	Peak height
$\mathbb{P}\{A\}$	Probability of event A
Q	Kendall's rank correlation test statistic
$\mathrm{Re} = \mathrm{Re}^1$	Set of real numbers
Re_+	Set of positive real numbers
R^2	Coefficient of determination (multiple correlation coefficient)
r	Pearson's correlation coefficient
r_{xy}	Sample correlation coefficient between x and y
$r_{xy\cdot z}$	Sample correlation coefficient between x and y with z held constant
r.v.	Random variable
$rrmse$	Relative root mean square error
S	Weighted sign test statistic

s	Sample standard deviation, Laplace transform variable
S_r	Sum of r random variates
s.e.	Standard error
SL	Speed limit [km/h]
T	Time headway variable
t	Time headway (value)
$\bar{t}$	Sample mean of headways
t_i	The ith headway
$T_{(1)}$	The first order statistic of the distribution for T
$t_{(1)}$	The first order statistic of a sample; i.e., $\min\{t_1,\ldots,t_n\}$
$U(a,b)$	Uniform (rectangular) distribution between a and b
V	Exponential ordered scores test statistic
W^2	Cramér-von Mises statistic
X_+	Positive part of X; i.e., $(X+\|X\|)/2$
$x_j(\omega)$	Position of vehicle j at time ω
$\dot{x}_j(\omega)$	Differential of x_j with respect to ω (speed)
$\ddot{x}_j(\omega)$	Second differential of x_j with respect to ω (acceleration)
α	Shape parameter of the gamma distribution
$\alpha_3(T)$	Skewness of the distribution for T
α_3	Sample skewness
$\alpha_4(T)$	Kurtosis of the distribution for T
α_4	Sample kurtosis
β	Scale parameter of the gamma distribution
$\Gamma(x)$	Gamma function
$\gamma(a,x)$	Incomplete gamma function
δ	Threshold for exponential tail
θ	Parameter of a distribution
$\hat{\theta}$	Maximum likelihood estimator for θ
$\check{\theta}$	Modified maximum likelihood estimator for θ
$\breve{\theta}$	Minimum chi-square estimator for θ
$\tilde{\theta}$	Estimator for θ
λ	Flow rate, traffic intensity
$\tilde{\lambda}$	Traffic volume, estimated flow rate
μ	Scale parameter of the lognormal distribution
$\mu(T)$	Mean of the distribution for T
ν	Degrees of freedom
ρ	Server utilization factor
ρ_i	Autocorrelation coefficient at lag i
$\tilde{\rho}_i$	Sample autocorrelation coefficient at lag i

$\rho_{i \cdot x}$	Partial correlation coefficient for autocorrelation at lag i with x held constant
σ	Shape parameter of the lognormal distribution,
$\sigma(T)$	Standard deviation of the distribution for T
$\sigma^2(T)$	Variance of the distribution for T
τ	Location (threshold) parameter of a headway distribution, rank correlation coefficient (Kendall's τ)
υ	Speed difference
$\Phi(X)$	Probability distribution function of the standard normal distribution
$\phi(X)$	Probability density function of the standard normal distribution
ω	Time
$*$	Convolution operator
$\#\{\, x \mid A \,\}$	Number of x's fulfilling condition A

Preface

The work started with the aim of finding more effective methods for vehicle-actuated traffic signal control on two-lane roads. Part of the project was the development of a traffic signal control simulation program. In order to find a suitable headway variate generator, some field measurements were performed on two-lane roads with speed limits not greater than 70 km/h. (In Finland traffic signal control is not allowed on roads with higher speed limits.) The location and time of each measurement was selected so that the expected flow rate would not be much above 1,000 veh/h.

The analysis of the vehicle time headway data soon appeared to be a far more complex issue than was initially assumed. It became evident that the methods commonly used in the literature were not very efficient, and in many cases they were infeasible. Consequently, the focus of the study shifted from traffic signal control to the statistical analysis of vehicle headways. Some shortcomings in this report are due to this fact. When the field measurements were made, the extent and type of the analyses to be performed was not anticipated.

Many of the theoretical distributions and their basic properties have been described in an introductory paper (Luttinen 1990). The major results have been published earlier: the goodness-of-fit tests and parameter estimation for the gamma distribution (Luttinen 1991); the basic statistical properties of time headways (Luttinen 1992); and the methods of identification, estimation, and goodness-of-fit tests (Luttinen 1994). The comments of the reviewers of these papers have been very helpful.

The traffic analyzer measurements were performed by the Laboratory of Transportation Engineering at the Helsinki University of Technology. My special thanks to Mr. Kari Hintikka for all the care he has taken with the measurements. I gratefully acknowledge professors Arde Faghri, Sulevi Lyly, Hannu Niemi and Matti Pursula for their advice and support. Professor Faghri has given very valuable help to make the purpose and style of the text more clear. Professor Lyly has helped me in many practical matters.

Professor Niemi's detailed comments on statistical and mathematical aspects have been extremely helpful. Professor Pursula's expertise in traffic flow theory has saved me from many pitfalls.

Many of the analyses would have been impossible to carry out without the financial support from the Henry Ford Foundation in Finland.

Finally, my deepest thanks to my wife Tarja, to my children Mikko, Hanna, and Jaakko, and to my parents Raili and Matti for their continuous support during the past years. Without their patience and encouragement this report would never have seen the daylight.

Lahti, Finland
April 28, 1996

R. Tapio Luttinen

Chapter 1

Introduction

1.1 Definitions

Time headway, or headway for short, is the time interval between two vehicles passing a point as measured from the front bumper to the front bumper. A headway is accordingly the sum of the time used by a vehicle to pass the observation point (*occupancy time*) and the time interval (*gap*) to the arrival of the next vehicle.

In platoons the gap is usually more meaningful to the driver than the headway. Headways are, however, more important in the traffic flow theory. Because the flow rate[1] is reciprocal to the mean headway, the study of headways is closely connected to the study of highway capacity. In selecting a headway a driver establishes a flow rate for the vehicle (Dawson & Chimini 1968).

The headways are measured in *time* rather than in space. Time headways, rather than space headways or *spacings* between consecutive vehicles, have been selected as the subject because:

1. Time headways and flow rates are directly connected. There is a similar connection between spacings and densities, but flow rates are usually more meaningful and more easily measured by practicing traffic engineers.

2. In platoons the driver of a trailing vehicle adjusts the gap considering safety, one aspect of which is the reaction time. It can be assumed

[1]The *flow rate* or *traffic flow* means the expected number of vehicles per time unit passing the reference point. The *traffic volume* is the measured number of vehicles per time unit. In HCM (1994) the terms have been used in a different sense.

that the gap, and hence the headway of a trailing vehicle, is not as sensitive to vehicle speeds as the spacing.

3. An estimate of the spacing can be obtained as the product of speed and headway, assuming constant speed during the headway.

In the literature on point processes the *counting specification* is often taken as primary, because it allows a generalization to higher dimensional spaces (e.g., Daley & Vere-Jones 1988). This study, however, discusses the traffic flow as a process on the real line. For this purpose the *interval specification* is more informative.

1.2 Importance of headway studies

Many traffic operations (such as passing, merging, and crossing) depend on the availability of large headways in the traffic flow. Consequently, the distribution of headways has an effect on *platoon formation and delays*.

In the 1985 Highway Capacity Manual (HCM 1985) the *level of service* on two-lane rural highways is approximated by the proportion of headways less than five seconds—thus making a connection between the headway distribution and the level of service. In addition, the headways play a crucial role in the analysis of unsignalized intersections and roundabouts (Sullivan & Troutbeck 1994).

Although the average delay in fixed time traffic signals is not very sensitive to the detailed stochastic properties of the arrival model (Newell 1956, Newell 1965, Kingman 1962), in *vehicle actuated traffic signals* the control, the extension time in particular, is very sensitive to the arrival pattern (Darroch, Newell & Morris 1964). The characteristics of vehicle headways are accordingly quite relevant in the study of optimal traffic signal control.

The *mathematical analysis and simulation* of traffic operations are based on theoretical models, which must be evaluated against the properties of real world data. More advanced methods have increased the need for reliable knowledge of the statistical properties of vehicle headways.

1.3 History and problems

Traffic flow has been considered as a stochastic process at least since Adams (1936). He formulated the idea of arrivals as a random (i.e., Poisson) process and verified good agreement with theory and observations. Since then many more sophisticated models have been proposed.

The properties of headways have been extensively studied, especially in the 60's. Some of this earlier work is now drawn together and evaluated in the light of recent data and more powerful statistical methods.

No unified framework has been developed for traffic flow theory. Accordingly, the vehicle headway studies have been concentrated on the statistical analysis of headway data. There have been, however, some notable exceptions. Greenberg (1966) has found a connection between the microscopic traffic flow theory and the lognormal follower headway distribution.

An interesting new approach has been developed by Heidemann (1990, 1993). He has applied the theory of stochastic processes to analytically derive the headway distribution as a function of traffic density. Heidemann criticizes the statistical approach for lacking in detailed analytical reasons that would justify the various types of distributions. His approach may provide a link between the macroscopic traffic flow theory and the headway distributions, but so far it has not found many applications. The distribution function is rather complicated, and its statistical properties have not been widely explored. The Heidemann model is not discussed below, but it well deserves further studies.

The statistical analysis of vehicle headways has been inadequate in three important aspects:

1. There has been no standard procedures to collect headway data and to describe their statistical properties.

2. The goodness-of-fit tests have been either powerless or infeasible.

3. Test results from several samples have not been combined properly.

In the literature on vehicle headways the shape of the empirical (sample) distributions have been mostly described by cumulative distribution functions and histograms. In some studies measures such as the mode (May 1961, Summala & Vierimaa 1980, Rajalin & Hassel 1992), the median (May 1961) and the coefficient of variation (May 1965, Buckley 1968, Dunne, Rothery & Potts 1968, Breiman, Lawrence, Goodwin & Bailey 1977, Pursula & Sainio 1985, Griffiths & Hunt 1991) have been presented. The renewal hypothesis (mutually independent and identically distributed headways) has also drawn the attention of several researchers (e.g., Dunne et al. 1968, Breiman, Gafarian, Lichtenstein & Murthy 1969, Cowan 1975, Chrissikopoulos, Darzentas & McDowell 1982).

The work of Breiman et al. (1969) deserves special attention. They have made a serious attempt to standardize the analysis of headway data by presenting a method to determine trendless time periods in the traffic flow, and by introducing a set of three tests for the renewal hypothesis. The idea of trendless data has not, however, been unanimously accepted. In addition, the comparison of the headway data presented by various authors is difficult, because no standard set of analyses has been applied to identify the most important properties of the headway data.

The goodness of fit has been tested mainly by three methods: (1) graphical evaluation, (2) the chi-square test, and (3) the Kolmogorov-Smirnov test. The graphical evaluation has been based on the probability distribution function (e.g., Akçelik & Chung 1994) or the probability density function (e.g., Buckley 1968). These methods are highly subjective and should be used only in preliminary studies. The use of histograms to estimate the probability density function has sometimes even produced a distorted idea of the distribution.[2]

The chi-square test is theoretically justified, but it is based on the reduction of observations into counts, which decreases the power of the test considerably. The Kolmogorov-Smirnov test has been even more problematic. It has been used in its nonparametric form, which is based on the assumption that the parameters of the distribution are known. In the headway studies the parameters are, however, usually estimated from the data. Because the estimation makes the fit better, the critical values should be smaller. This correction has not been made in any of the studies reviewed. Consequently, the results have been too conservative. The problems associated with the goodness-of-fit tests have not received proper attention in previous studies.

Another great problem has been the lack of theoretical foundation in dealing with multi-sample data. The headway data usually consist of several samples, and the null hypothesis is tested against each sample. The result has often been presented as the number or proportion of tests which have led to the rejection of the null hypothesis (e.g., Buckley 1968, Breiman et al. 1969). In some cases even the average significance probability has been used (e.g., Branston 1976). In none of the studies reviewed have the individual tests been properly combined to give a scientifically valid evaluation of the overall acceptability of the hypothesis.

[2]The figure on page 459 of Griffiths & Hunt (1991) gives the impression of a bell shaped probability density function that is not supported by their equation 1.

1.4 Purpose of the study

The scope of this study is the statistical analysis of *vehicle headways on two-lane two-way roads*. The headways on both lanes are analyzed separately. The study has two purposes:

1. To introduce a *statistical methodology* for the analysis of the headway data and for the evaluation of the proposed models

2. To get information about the *statistical properties* of headways on two-lane two-way roads.

The main objective is to present proper statistical methods for the goodness-of-fit tests and for the combination of significance probabilities. The second objective is to present a set of methods to effectively describe the statistical properties of headway data. These tests and methods will be applied to some of the most commonly used headway distribution models. Combined with the discussion on the data collection and parameter estimation, the text presents a scientific foundation for the statistical analysis of vehicle headways.

Although the data are from two-lane two-way roads, the statistical methods presented can be applied to any headway data. Some methods have even more general interest.

1.5 Summary of contents

The presentation follows the *statistical modeling process* in figure 1.1. Chapter 2 reviews the microscopic traffic flow theory in the light of vehicle headways. There is also a short introduction to the main statistical methods used. The problem of combining the test results from several samples is solved by the method of combined probabilities (Fisher 1938). The moving probability method developed by Luttinen (1992) allows the analysis of test results over another variable, such as traffic volume.

Chapter 3 discusses the measurements and the preliminary analysis of the Finnish headway data. The methods introduced in earlier literature to obtain stationary headway data are evaluated. Because the combination of samples introduces a bias in the data, data collection based on trend analysis gives better results.

Chapter 4 introduces methods for model identification. In addition to the conventional probability distribution function and the probability density function, the shape of the headway distribution is analyzed using the

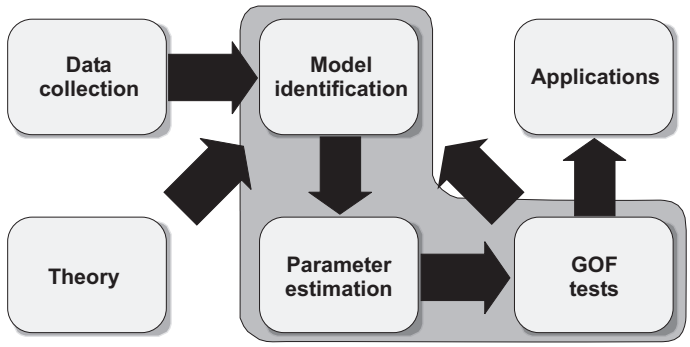

Figure 1.1: Statistical modeling process

hazard function, which has been extensively applied in life and reliability data analyses. The first four moments are shown to be very informative measures of location and dispersion. On the basis of these analyses a *four-stage identification process* (Luttinen 1994) is introduced. In addition, three methods are used to test, whether the renewal hypothesis can be accepted for the headway data.

Chapter 5 analyses the applicability of several statistical distributions as models for the headway distributions on two-lane two-way roads. Main emphasis is on the goodness-of-fit tests and the parameter[3] estimation. Parametric goodness-of-fit tests are compared with the nonparametric Kolmogorov-Smirnov tests.

It is assumed that the reader has a basic knowledge of the probability theory as well as statistical and numerical methods. The theory is described only to the extent that is essential to understand the text. More detailed descriptions can be found in the references given.

[3]The parameters are called location, scale and shape parameters following the terminology of D'Agostino & Stephens (1986), Johnson, Kotz & Balakrishnan (1994), and Stuart & Ord (1987). A parameter τ is called the *location parameter*, if the distribution has the form $f(t-\tau)$, and θ is called the *scale parameter*, if the form is $\theta f(\theta t)$ or $\frac{1}{\theta} f\left(\frac{t}{\theta}\right)$ (Stuart & Ord 1987). If higher moments (skewness and kurtosis) of a distribution are defined by a third parameter, it is called the *shape parameter*.

Chapter 2

Theoretical background

2.1 Microscopic traffic flow theory

The microscopic theory of traffic flow describes the movements of vehicles on road sections. The movement of vehicles on a lane can be displayed as a time-space digram (fig. 2.1). The distances between vehicles can be studied either at a fixed location (headways) or at a fixed time (spacings) or as a dynamic (car-following) model, which is the most comprehensive approach. A short review of the car-following models gives a broader view on the properties of vehicle headways.

Let the position of a vehicle j at a moment ω be $x_j(\omega)$. The time instant for the vehicle to pass the observation point (x) is:

$$\omega_j(x) = \{\, \omega \mid x_j(\omega) = x \,\}. \tag{2.1}$$

Thus, the time headway of vehicle j at point x is:

$$t_j(x) = \omega_j(x) - \omega_{j-1}(x). \tag{2.2}$$

Assuming constant speed for vehicle j during the headway, the headway of vehicle j at time ω is:

$$t_j(\omega) = \frac{x_{j-1}(\omega) - x_j(\omega)}{\dot{x}_j(\omega)}, \tag{2.3}$$

where the numerator is the spacing of vehicle j at time ω, and the dot denotes differentiation with respect to time. This equation gives the basic relationship between time headways, space headways, and car-following models.

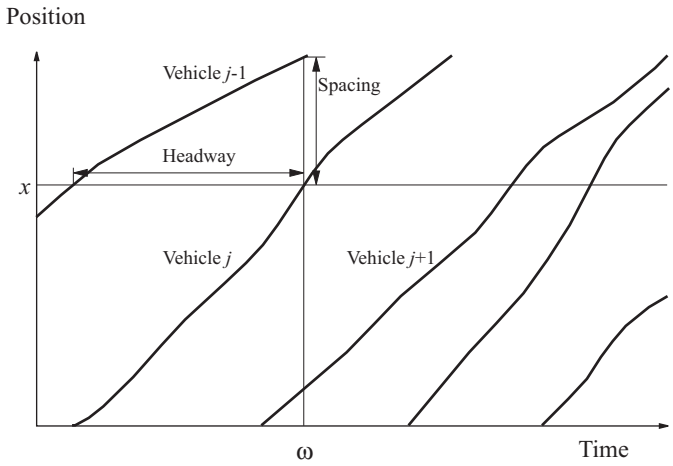

Figure 2.1: Time-space diagram of vehicle trajectories

Let us assume that the driver of vehicle j has reaction time R_j. The length of the vehicle ahead is L_{j-1}. The driver of vehicle j wants to keep a net safety spacing B_j in front of the vehicle. This spacing allows the driver to avoid too strong decelerations. The desired headway is then (Daou 1966):

$$t_j = R_j + \frac{L_{j-i} + B_j}{\dot{x}_{j-1}}. \tag{2.4}$$

Daou analyzed over 24,000 headways from the Holland Tunnel and obtained least squares estimates for reaction time $R_j = 1.488$ s and (gross) safety spacing $L_{j-1} + B_j = 35.0$ ft (10.7 m).

According to this theory, there is a correlation between desired headways and speeds as shown in figure 2.2. At higher speeds the follower headways are shorter.

In a car-following model the driver of a following vehicle can control the vehicle by accelerating and decelerating. The researchers in the General Motors (see May 1990) developed several models based on the assumption that the *response* (control) is a function of *stimuli* and *sensitivity*.

According to the *linear car-following model* the response is a linear function of the speed difference between the vehicles (Chandler, Herman & Montroll 1958):

$$\ddot{x}_j\left(\omega + R_j\right) = \alpha_0\left[\dot{x}_{j-1}\left(\omega\right) - \dot{x}_j\left(\omega\right)\right]. \tag{2.5}$$

Because the response does not depend on the headway or the spacing between vehicles, the model is unrealistic, and it cannot describe the desired

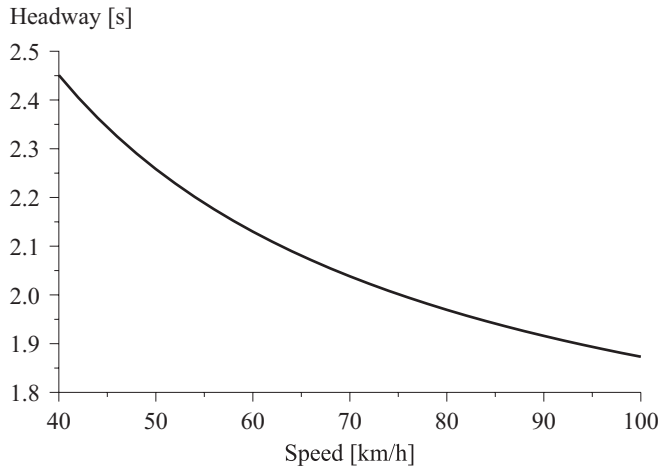

Figure 2.2: Speed and desired headway according to Daou (1966)

headway.

Gazis, Herman & Potts (1959) have proposed a *nonlinear car-following model*:

$$\ddot{x}_j(\omega + R_j) = \frac{\alpha_1}{x_{j-1}(\omega) - x_j(\omega)} \left[\dot{x}_{j-1}(\omega) - \dot{x}_j(\omega)\right], \tag{2.6}$$

where the sensitivity is a function of a parameter (α_1) and the spacing, and the stimulus is the speed difference between vehicles $j-1$ and j.

The model produces different equilibrium headways for different initial headways and relative speeds. Long initial headways and low initial relative speeds result in long equilibrium headways.

The headway of a vehicle approaching the vehicle ahead is shown in figure 2.3. The leading vehicle has constant speed of 80 km/h. The follower starts the car-following at initial headway 10 s and speed 100 km/h. Reaction time is 1.5 seconds. Sensitivity coefficients (α_1) are 3.0 m/s (upper curve), 2.6 m/s (middle curve), and 2.0 m/s (lower curve). The straight line shows the headway if the follower maintains constant speed (100 km/h).

Gazis, Herman & Rothery (1961) have further modified the nonlinear model:

$$\ddot{x}_j(\omega + R_j) = \alpha_2 \frac{\dot{x}_j^m(\omega + R_j)}{\left[x_{j-1}(\omega) - x_j(\omega)\right]^l} \left[\dot{x}_{j-1}(\omega) - \dot{x}_j(\omega)\right]. \tag{2.7}$$

Increasing l makes the model less sensitive to the initial headway, because the response is more strongly inversely proportional to the spacing. The

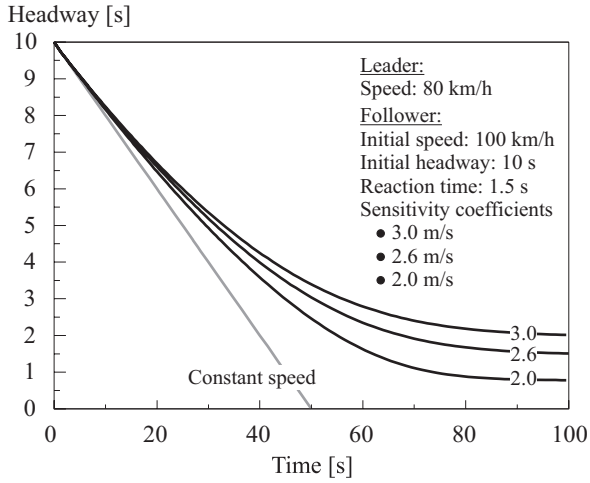

Figure 2.3: Headway as a function of time, when a vehicle is approaching the vehicle ahead according to the nonlinear car-following model (2.6)

response is rather weak until the spacing between vehicles is short. A larger value of m makes the steady state headway less dependent on the initial speed difference, because the sensitivity is proportional to the speed of the follower.

On the basis of the car-following models the variation in the follower headways can be explained by seven factors:

1. Differences in the sensitivity between followers
2. Delayed responses to speed variations of the leading car
3. Oscillation due to strong and delayed responses
4. Variation in the length of leading vehicles
5. Approaching (transition) phase in the beginning of the car-following process
6. Variation in the initial headway, when the car-following process starts
7. Variation in the initial speed difference, when the car-following process starts.

Factors 1–4 describe the typical car-following process, when a vehicle has finished the approaching phase and is following the leading vehicle. Factors 5–7 describe the situation, when a vehicle approaches the leading vehicle

and shifts from the free-moving state to the car-following state. Factors 1, 6, and 7 describe different types of drivers.

The car-following models have serious limitations in headway studies:

1. Even in a reasonably dense traffic some drivers will not attempt to follow the car preceding them (Herman & Potts 1961), because the desired speed of a driver can be lower than the speed of the vehicle ahead. The models can only be applied to vehicles that are following another vehicle.

2. At long spacings the speed difference does not influence the follower. The models do not, however, have a threshold for car-following and free driving.

3. Some short headways are caused by passing vehicles. The models do not, however, describe passing.

4. The models developed by the researchers in the General Motors do not have a connection between speed and spacing. The equilibrium headway is different depending on the initial speed of the follower (fig. 2.3). In addition, the models do not allow the possibility of differences in the "desired spacing" among followers.

5. According to Leutzbach (1988) "the perceptual threshold for positive speed differences (the spacing decreases) is smaller than that for negative speed differences (the spacing increases)". The models have the same sensitivity for negative and positive speed differences.

6. The car-following theory does not deal with the inherent fuzziness of the car-following process. The follower cannot observe and does not continuously react to variations smaller than a perception threshold in the spacing or the speed of the leading car (Leutzbach 1988). This "acceleration noise" causes random variation in the process.

7. A driver's reaction pattern is imprecise, but the reaction of the follower is deterministic in the models (Kikuchi & Chakroborty 1992).

These considerations have led to psycho-physical[1] (Leutzbach 1988) and fuzzy (Kikuchi & Chakroborty 1992) car-following models.

[1] *Psycho-physiologisch* in German.

2.2 Statistical methods

2.2.1 Sampling techniques

Headway (t_i) is the time between successive vehicles ($i-1$ and i) as they pass a point on a lane or roadway measured from front bumper to front bumper (HCM 1994). The mean headway ($\bar{t}$) and the traffic volume ($\tilde{\lambda}$) in a sample are connected in an obvious way:

$$\tilde{\lambda} = \frac{n}{\sum_{i=1}^{n} t_i} = \frac{1}{\bar{t}}, \tag{2.8}$$

where n is the number of observations in the sample. Because the expected traffic volume is, by definition, the flow rate, the expected mean headway at flow rate (λ) is:

$$E(\,\bar{t}\mid\lambda\,) = \frac{1}{\lambda}. \tag{2.9}$$

In order to get generally applicable results about the properties of headways, it is necessary to study *stationary* conditions.[2] Consequently, all nonrandom variation must be removed from the measurements as far as possible. Otherwise the parameters of the distribution change as a function of time.

In previous studies mainly two approaches have been used to overcome the problem of nonrandom variation: The samples have been collected either as fixed time slices or on the basis of trend analysis. Quite often the method of collecting and analyzing the data is not specified.

In some studies the measurements are investigated in fixed time slices of length short enough to exclude any significant trend (typically 30 s to 10 min). It is, however, possible to have a significant trend in a 5–10 minute time period. One minute period, on the other hand, is too short, at least under light flow conditions.

As the time period (S_n, sum of sample headways) gets shorter, sample mean ($\bar{t}$) longer or standard deviation (s) larger, the standard error in the

[2]The stationarity is considered to be equivalent to trendlessness. It is possible that the process is nonstationary, although the flow rate remains constant. This possibility is, however, ignored, as suggested by Cox & Lewis (1966).—Because traffic flow is a dynamic process in time and space, stationarity can be defined both in time and in space (McLean 1989). Only stationarity in time is considered here.

mean headway estimate increases:

$$\begin{aligned} \text{s.e.}[\tilde{\mu}(T)] &= \frac{s}{\sqrt{n}} \\ &= s\sqrt{\frac{\bar{t}}{S_n}}. \end{aligned} \tag{2.10}$$

Under light flow conditions the number of observations per period is small. The growth of the standard deviation as the mean headway increases[3] makes the mean estimate even more unreliable. As Botma (1986) has noted, the number of long headways in small samples is a highly fluctuating quantity.

Because the number of headways in a short observation period is usually too small for testing goodness of fit, it is necessary to group samples having nearly equal means. This may result in a sampling error as the sample means have an inappropriate distribution. The artificial homogenization of the data leads to the narrowing of the headway distributions, and it can happen that the tails of the underlying distributions are undersurveyed (Heidemann 1990). According to Greenberg (1966) the stratification of headways by flow, density or other criteria will introduce a truncation into the data.

Consider m samples, each sample having size n_i and mean and variance equal to:

$$\bar{t}_i = \frac{1}{n_i}\sum_{k=1}^{n_i} t_{ik} \tag{2.11}$$

$$s_i^2 = \frac{1}{n_i}\sum_{k=1}^{n_i}(t_{ik} - \bar{t}_i)^2, \tag{2.12}$$

where t_{ik} the kth headway in sample i. The combined sample mean is:

$$\bar{t} = \frac{1}{n}\sum_{i=1}^{m} n_i \bar{t}_i, \tag{2.13}$$

where

$$n = \sum_{i=1}^{m} n_i. \tag{2.13a}$$

[3]See section 4.3.3.

The variance of the combined sample is:

$$
\begin{aligned}
s_t^2 &= \frac{1}{n}\sum_{i=1}^{m}\sum_{k=1}^{n_i}(t_{ik}-\bar{t}\,)^2 \\
&= \frac{1}{n}\sum_{i=1}^{m}\sum_{k=1}^{n_i}(t_{ik}-\bar{t}_i+\bar{t}_i-\bar{t}\,)^2 \\
&= \frac{1}{n}\sum_{i=1}^{m}\left[n_i(\bar{t}_i-\bar{t}\,)^2+\sum_{k=1}^{n_i}(t_{ik}-\bar{t}_i)^2+2(\bar{t}_i-\bar{t}\,)\overbrace{\sum_{k=1}^{n_i}(t_{ik}-\bar{t}_i)}^{0}\right] \\
&= \frac{1}{n}\sum_{i=1}^{m}n_i\left[(\bar{t}_i-\bar{t}\,)^2+s_i^2\right] \\
&= \frac{1}{n}\sum_{i=1}^{m}n_i(\bar{t}_i-\bar{t}\,)^2+\frac{1}{n}\sum_{i=1}^{m}n_i s_i^2.
\end{aligned}
\tag{2.14}
$$

Consequently, the variance of the combined sample is the variance of sample means plus the mean of sample variances. If true mean of the combined sample is used, the mean of sample variances can be assumed rather constant. The range of sample means, however, should be defined so that the variance of sample means is not biased. Too narrow (wide) range results in underestimation (overestimation) of the variance of the headways.

By the central limit theorem, the distribution of sample means from a population with mean μ and variance σ^2 approaches normal distribution with mean μ and variance σ^2/n, when the sample size (n) is large (Stuart & Ord 1987). If the flow rate is changing during the time of measurement, the sample means between the selected lower (a) and upper (b) limits can be assumed to be uniformly distributed. The range can be selected so that the variance $[(b-a)^2/12]$ of the sample means is realistic, but the distribution of the sample means is still biased.

The other approach, adopted by Dunne et al. (1968), and by Breiman et al. (1969), is to analyze several long time periods showing no trend. If the number of observations per sample is large, the estimate of the mean is more reliable. It is not necessary to group samples. Greater variation and serial correlation is allowed without a strong impact on the mean estimate. Longer sampling periods enable the detection of lower frequencies in the spectral analysis, if it is performed. There is, however, a loss of data, because nonstationary time periods are discarded.

Despite the use of trend tests, there is a possibility of short deterministic patterns in the traffic flow. This additional component of variation may

cause a small increase in the variance of the headway data (Heidemann 1990).

2.2.2 Trend tests

In order to obtain trendless samples, it is necessary to subject the data to trend tests. Three tests are evaluated:

1. Weighted sign test
2. Kendall's rank correlation test
3. Exponential ordered scores test.

Cox & Stuart (1955) suggested the *weighted sign test* as a quick and simple but less efficient alternative to the rank correlation test. For a pair $\{ t_i, t_j \mid i < j \}$ of observations a score is defined:

$$h_{ij} = \begin{cases} 1, & \text{if } t_i > t_j \\ 0, & \text{if } t_i < t_j. \end{cases} \tag{2.15}$$

The test statistic is a weighted number of points of decrease:

$$S = \sum_{k=1}^{\frac{1}{2}n} (n - 2k + 1) h_{k,n-k+1}. \tag{2.16}$$

Under the null hypothesis (no trend) the observations are not correlated, and S is asymptotically normal (Brockwell & Davis 1987) with mean and variance equal to:

$$\mu(S) = \frac{1}{8} n^2 \tag{2.17}$$

$$\sigma^2(S) = \frac{1}{24} n(n^2 - 1). \tag{2.18}$$

The *Kendall's rank correlation test* (Kendall & Ord 1990, Stuart & Ord 1991) extends the pairwise comparison to all pairs $\{ t_i, t_j \mid i < j \}$:

$$Q = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} h_{ij}. \tag{2.19}$$

The rank correlation coefficient is:

$$\tau = 1 - \frac{4Q}{n(n-1)}. \tag{2.20}$$

Under the null hypothesis (no trend) τ tends to normality rapidly (Stuart & Ord 1991) with mean and variance given by:

$$\mu(\tau) = 0 \tag{2.21}$$

$$\sigma^2(\tau) = \frac{2(2n+5)}{9n(n-1)}. \tag{2.22}$$

The *exponential ordered scores test* of Cox & Lewis (1966) attaches the score

$$s_{r,n} = \frac{1}{n} + \cdots + \frac{1}{n-r+1}, \qquad r = 1,\ldots,n \tag{2.23}$$

to the rth order statistic ($t_{(r)}$, the rth longest headway) in a sample of size n. Let $s_n(i)$ be the score of headway t_i. The test statistic is:

$$V = \sum_{i=1}^{n} s_n(i) \left(i - \frac{n+1}{2} \right). \tag{2.24}$$

The test statistic is asymptotically normally distributed with mean zero and variance

$$\sigma^2(V) = \sum_{i=1}^{n} \left(i - \frac{n+1}{2} \right)^2 K_{2,n}, \tag{2.25}$$

where

$$K_{2,n} = 1 - \frac{1}{n-1} \left(\frac{1}{n} + \frac{1}{n-1} + \cdots + \frac{1}{2} \right). \tag{2.25a}$$

Empirical power curves[4] were evaluated for the three tests by a Monte Carlo method with shifted exponential pseudo-random variates generated by the following algorithm:

$$T_{j+1} = \tau - \left(\frac{1}{\lambda(\omega_j)} - \tau \right) \ln(U), \qquad j = 0,1,2,\ldots, \tag{2.26}$$

[4]See page 36.

where

$$\omega_j = \sum_{i=1}^{j} T_i \tag{2.26a}$$

$$\omega_0 = 0, \tag{2.26b}$$

$\lambda(\omega_j)$ is the current flow rate, τ is the location parameter of the shifted exponential distribution, and U is a pseudo-random number from the uniform distribution (0,1]. The algorithm allows negative random variates to be generated. Because negative vehicle headways are unrealizable, the power curves for negative trends (displayed on a gray background in figure 2.4) should be viewed for comparison only.

Linear trend was used. Consequently, the average headway at time instant ω_j is:

$$E(T_{j+1}) = \frac{1}{\lambda(\omega_j)} = \frac{1}{\lambda_0}(1 + \theta\omega_j), \tag{2.27}$$

where λ_0 is the initial traffic intensity and θ is the trend factor. Factorial design (Montgomery 1984) was used with λ_0 = 100 veh/h, 200 veh/h, 500 veh/h, and 1,000 veh/h, as well as θ = -500 %/h, -200 %/h, -100 %/h, -50 %/h, -10 %/h, 0 %/h, 10 %/h, 50 %/h 100 %/h, 200 %/h, and 500 %/h. The sample sizes were 50, 100 and 200, and the number of replicas was 100. The two-sided significance level was 0.05.

Figure 2.4 shows the empirical power curves of the tests based on Monte Carlo runs. A test is powerful if its power curve is equal to the significance level for trendless samples and near unity for samples with trend. The curves are at 0.05, when there is no trend ($\theta = 0$), that is the tests reject 5 % of the samples. As the absolute value of the trend increases, the tests can detect the trend more efficiently. The exponential ordered scores test is superior to the other tests, and the weighted sign test is the least powerful. Although the exponential ordered scores test is computationally more strenuous, it was selected as the prime method of testing trend.

According to Stuart (1956) the asymptotic relative efficiencies of the weighted sign test and the rank correlation test against the regression coefficient test are 0.83 and 0.98, respectively. The empirical power curves in figure 2.4 confirm the superiority of the rank correlation test against the weighted sign test.

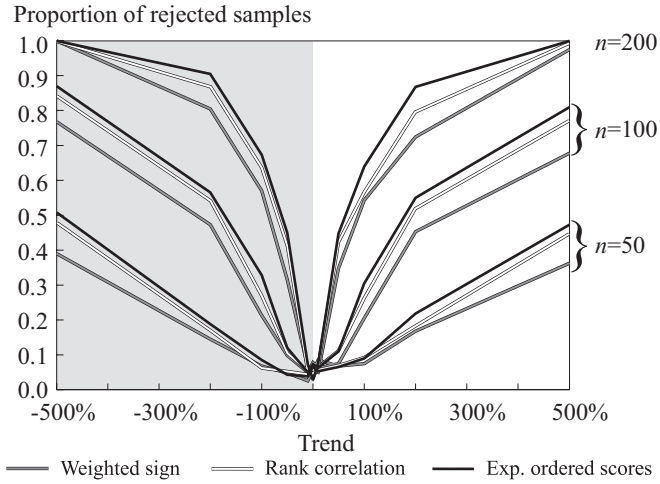

Figure 2.4: Empirical power curves for three trend tests

2.2.3 Parameter estimation

Method of moments

The parameters of a distribution can be estimated by equating the expressions for the first k population moments to the corresponding expressions for the k sample moments, and solving the k parameters from the set of equations. This method of moments has been widely used in headway studies because of its simplicity: Raw data is not required after the first k moments have been calculated, and the first moment (mean headway) is directly obtained from the traffic volume (see eq. 2.8). Also, the solution to the set of equations is usually simple.

The method of moments estimators are usually consistent (Dudewicz & Mishra 1988); i.e., the mean squared error of the estimator decreases to zero as the sample size increases. The method is, however, usually not efficient[5] (Stuart & Ord 1991), except for distributions closely resembling the normal (Greenwood & Durand 1960). This is to be expected, because the information in the data is reduced to the sample moments. Accordingly, the method of moments is usually not appropriate to estimate the parameters of skew distributions, like the headway distributions, but it can be used to give first approximations for more efficient methods.

[5]The estimator is (asymptotically) efficient if it has in large samples the smallest variance among all consistent and asymptotically normal estimators. (Stuart & Ord 1991)

Maximum likelihood method

The maximum likelihood method is the other traditional method to estimate the parameters ($\boldsymbol{\theta} = \theta_1, \ldots, \theta_k$) of a distribution. Let us assume that the probability density function is $f(\, t \,|\, \boldsymbol{\theta} \,)$. For a sample of n independent observations the parameter estimators ($\hat{\boldsymbol{\theta}}$) are obtained by maximizing the *likelihood function*:

$$L(\hat{\boldsymbol{\theta}}) = \prod_{i=1}^{n} f\Big(\, t_i \, \Big| \, \hat{\boldsymbol{\theta}} \, \Big) \, . \tag{2.28}$$

subject to parameter vector $\hat{\boldsymbol{\theta}}$. In contrast to the method of moments, the likelihood function contains all the information in the sample.

Usually the same estimators can be obtained more easily by finding the maximum of the *log-likelihood function*:

$$\ln L(\hat{\boldsymbol{\theta}}) = \sum_{i=1}^{n} \ln f\Big(\, t_i \, \Big| \, \hat{\boldsymbol{\theta}} \, \Big) \, . \tag{2.29}$$

A necessary condition for the maximum is that the partial derivatives equate to zero:

$$\left.\begin{aligned} \frac{\partial \ln L(\hat{\boldsymbol{\theta}})}{\partial \hat{\theta}_1} &= 0 \\ &\vdots \\ \frac{\partial \ln L(\hat{\boldsymbol{\theta}})}{\partial \hat{\theta}_k} &= 0. \end{aligned}\right\} \tag{2.30}$$

This set of equations is called the *likelihood equations*. A sufficient, though not a necessary, condition that a solution to these equations is a local maximum is:

$$\left.\begin{aligned} \frac{\partial^2 \ln L(\hat{\boldsymbol{\theta}})}{\partial \hat{\theta}_1^2} &< 0 \\ &\vdots \\ \frac{\partial^2 \ln L(\hat{\boldsymbol{\theta}})}{\partial \hat{\theta}_k^2} &< 0. \end{aligned}\right\} \tag{2.31}$$

If there are more than one maxima, the largest one of them is chosen. The resulting estimators ($\hat{\boldsymbol{\theta}} = \hat{\theta}_1, \ldots, \hat{\theta}_k$) are the *maximum likelihood estimators* (MLEs), provided that there is no terminal maximum at the extreme permissible values of the parameters. (Stuart & Ord 1991)

The maximum likelihood method has several desirable properties. Under very general conditions a MLE is consistent, and under certain regularity conditions it is efficient. Also, if there is an efficient estimator, the maximum likelihood method will produce it. Under some general conditions the MLE is asymptotically normally distributed with mean equal to the true parameter value, and among all estimators that are asymptotically normally distributed, the MLE possesses the minimum asymptotic variance; i.e. a MLE is a best asymptotically normal (BAN) estimator. The properties of the MLEs for large samples are, accordingly, excellent. (Stuart & Ord 1991)

Modified maximum likelihood method

According to Cohen & Whitten (1988) the first order statistic (the smallest observation) of a sample contains more information about the location[6] parameter ($\tau \equiv \theta_1$; threshold for the minimum headway) than any of the other sample observations, often more than all the other observations combined. This property can be used to modify the maximum likelihood method.

The modified maximum likelihood estimators (MMLEs, $\check{\boldsymbol{\theta}}$) are obtained by substituting the equation of the partial derivative with respect to the location parameter (τ) with the following equation:

$$F\big(\, t_{(1)} \,\big|\, \check{\boldsymbol{\theta}} \,\big) = F_n(t_{(1)}) = \frac{1}{n+1}, \tag{2.32}$$

where $t_{(1)}$ is the first order statistic of the sample. The equation can also be written in terms of the inverse probability distribution function:

$$F^{-1}\left(\frac{1}{n+1} \,\middle|\, \check{\boldsymbol{\theta}} \right) = t_{(1)} \tag{2.33}$$

In some situations the modified method gives better estimators than the maximum likelihood method. The modified method has been used to estimate the parameters of the gamma and the lognormal distributions. Cohen & Whitten (1980) and Cohen & Whitten (1982) give additional information about the properties of MMLEs.

Minimum chi-square method

For a sample classified into m mutually exclusive and exhaustive classes, the minimum chi-square (MCS) estimators are obtained by minimizing the

[6] See page 20, footnote 3.

chi-square statistic[7] (Cox & Hinkley 1974):

$$\chi^2 = \sum_{i=1}^{m} \frac{[n_i - np(\, i \mid \breve{\boldsymbol{\theta}}\,)]^2}{np(\, i \mid \breve{\boldsymbol{\theta}}\,)}, \tag{2.34}$$

where n_i is the number of observations in class i, n is the sample size, and $p(\, i \mid \breve{\boldsymbol{\theta}}\,)$ is the probability of an observation falling into class i conditional upon parameter vector ($\breve{\boldsymbol{\theta}}$).

Like the MLE, the MCS estimator is also asymptotically efficient and a best asymptotically normal (BAN) estimator (Stuart & Ord 1991). The MCS method, however, wastes some information because of the data classification.

The maximum likelihood method is considered computationally better and more effective in small samples than the MCS method. The classification of data is troublesome, especially in small samples. Also, the maximum likelihood method, unlike the MCS method, gives closed form solutions for some distributions. In the analysis of the headway data the maximum likelihood method is given preference. When it fails, the modified maximum likelihood method is used.[8]

2.2.4 Goodness-of-fit tests

Testing hypotheses

Goodness-of-fit tests are used to test the hypothesis that some particular distribution $F(\, t \,|\, \boldsymbol{\theta}\,)$ is the true unknown distribution $G(t)$, which has generated the sample:

$$H_0 \colon G(t) = F(\, t \,|\, \boldsymbol{\theta}\,)\,. \tag{2.35}$$

The *null hypothesis* (H_0) that the distributions are equal, is tested against the *alternative hypothesis* (H_A) that the distributions are not equal. The null hypothesis should be rejected if the distributions differ.

If $F(\, t \,|\, \boldsymbol{\theta}\,)$ is completely specified, the null hypothesis is called *simple*. If the parameters ($\boldsymbol{\theta}$) are partly or completely unspecified, H_0 is called *composite*—it is composed of various simple hypotheses. In the goodness-of-fit tests for headway distributions, H_0 is simple, while H_A is composite; i.e., the alternative headway distribution is unspecified. It is obvious that simple hypotheses are easier to test.

[7] See also page 37.

[8] For the shifted exponential distribution the modified method of moments estimators are used, because they are easier to calculate and give practically the same results.

The test fails if H_0 is rejected while it is true (*type I error*), or H_0 is accepted while it is false (*type II error*). The probability of type I error is called the *significance probability*[9], and the maximum probability of rejecting H_0 is called the *significance level* of the test. The probability of type II error is difficult to determine in headway studies, because H_A is composite.

The result of a test is either "H_0 is rejected" or "H_0 is not rejected". In the first case H_A is accepted, and the significance probability gives the probability of an erroneous judgment. The second case is interpreted that there is not enough evidence to show that H_0 is false. This does *not* necessarily mean that H_0 is true. Accordingly, it is possible for several distributions to pass the goodness-of-fit test, although only one or none of them is the true distribution $G(t)$. In the future larger data sets and more powerful tests may lead to the rejection of these hypotheses. This is typical of statistics: Much can be said about what is (probably[10]) *not* true, but very little about what *is* true.

The ability of a test to detect that H_0 is false, is called the *power* of the test. For a composite alternative hypothesis the power depends on the particular (simple) models in H_A. Consequently, the power is best described as a curve or function, which shows the probability of rejecting H_0 as a function of simple models in H_A. (Figure 2.4 on page 32 shows the power curves for three trend tests.) In an ideal test the probability of rejecting H_0 while it is true (type I error) is nul, and the probability of rejecting it while it is false is unity.

Because the true distribution $G(t)$ behind a headway sample is unknown, $G(t)$ cannot be used in the test, but $F(\,t\,|\,\boldsymbol{\theta}\,)$ must be tested against the properties of the sample:

$$H_0\text{: The random sample } \{t_1,\ldots,t_n\} \text{ is generated by } F(\,t\,|\,\boldsymbol{\theta}\,)\,. \qquad (2.36)$$

If the parameters of $F(t)$ are estimated from the sample, the null hypothesis becomes:

$$H_0\text{: The random sample } \{t_1,\ldots,t_n\} \text{ is generated by } F\Big(\,t\,\Big|\,\tilde{\boldsymbol{\theta}}\,\Big)\,. \qquad (2.37)$$

[9] Although the term "P-value" is more commonly used, the term "significance probability" is used here to emphasize the fact that the significance of a test is the probability of type I error.

[10] Popper (1983) states this even more strictly: "Probability hypotheses do not rule out anything observable... Probability estimates are not falsifiable. Neither, of course, are they verifiable." Probabilistic explanations are, however, important, because "often the best that can be established with some warrant is a statistical regularity" (Nagel 1982).

In both cases the problem is to find a powerful *test statistic* (B), which measures the goodness of fit. If the distribution $F_B(\cdot)$ of the test statistic under H_0 is known, the significance probability (p) is the probability that a random variate (B) from distribution $F_B(\cdot)$ is greater than or equal to b:

$$p = \mathbb{P}\{B \geq b\} = 1 - F_B(b). \tag{2.38}$$

If H_0 is true, the significance probabilities are uniformly (0,1) distributed (Fisher 1938). This is the consequence of the *probability integral transformation*: If X is a continuous random variate and $\mathbb{P}\{X \leq x\} = F(x)$, then $F(X)$ is distributed according to the uniform (0,1) distribution (Johnson et al. 1994).[11]

If the distribution of the test statistic is the same for all distributions and their parameters, the statistic is called *distribution-free*, and the hypothesis is called *nonparametric*[12]. A test using a statistic whose distribution is different depending on $F(\,t\,|\,\boldsymbol{\theta}\,)$ is called *parametric*. Such a test is usually more powerful, because it is based on the known properties of the statistical model.

Nonparametric tests

Chi-square test If a sample follows distribution $F(\,t\,|\,\boldsymbol{\theta}\,)$, it can be divided into m exclusive and exhaustive classes $A_1, \ldots, A_m$ so that the proportion of observations in a class corresponds to the probability of a random variate from $F(\,t\,|\,\boldsymbol{\theta}\,)$ falling into the same class. In headway studies an obvious classification is a set of consecutive headway ranges:

$$\{t_j \in A_i\} \Longleftrightarrow \{T_i < t_j \leq T_{i+1}\}\,. \tag{2.39}$$

[11] Consequently, the inverse probability distribution function can be used to generate a random variate X following distribution $F(x)$ from a uniformly $U(0,1)$ distributed random variate Y:

$$X = F^{-1}(Y).$$

[12] The terminology follows loosely that of Stuart & Ord (1991). They confine the adjectives "parametric" and "nonparametric" to statistical hypotheses and apply the adjective "distribution-free" to the test statistic, the distribution of the test statistic, and so on. Often the terms "nonparametric" and "distribution-free" are used interchangeably (Conover 1980).

In this case the probability $\mathbb{P}\{t_j \in A_i\}$ can be calculated as follows:

$$\begin{aligned} p(\, i \mid \boldsymbol{\theta}\,) &= \mathbb{P}\{t_j \in A_i\} \\ &= \mathbb{P}\{T_i < t_j \leq T_{i+1}\} \\ &= F(\, T_{i+1} \mid \boldsymbol{\theta}\,) - F(\, T_i \mid \boldsymbol{\theta}\,)\,. \end{aligned} \tag{2.40}$$

The goodness of fit can be measured by Pearson's chi-square statistic:

$$\chi^2 = \sum_{i=1}^{m} \frac{[n_i - np(\, i \mid \boldsymbol{\theta}\,)]^2}{np(\, i \mid \boldsymbol{\theta}\,)}, \tag{2.41}$$

where m is the number of classes, n_i is the number of observations in class i, and n is the sample size. If the null hypothesis is true, n_i is a binomial random variate with parameters n and $p(\, i \mid \boldsymbol{\theta}\,)$, and expectation $E(n_i) = np(\, i \mid \boldsymbol{\theta}\,)$. The statistic is then asymptotically (i.e., for large n) chi-square distributed with $\nu = m - 1$ degrees of freedom.

If some or all of the parameters ($\boldsymbol{\theta} = \theta_1, \ldots, \theta_k$) are estimated from the sample, the degrees of freedom becomes $\nu = m - 1 - r$, where r is the number of estimated parameters. When more parameters are estimated from the sample, the fit is bound to be better, and the test statistic must be smaller (Lilliefors 1967).

The significance probability p of the chi-square test is the probability that a random variate (χ^2) following the chi-square distribution is greater than or equal to x^2:

$$p = \mathbb{P}\{\chi^2 \geq x^2\} = 1 - F_{\chi^2}(\, x^2 \mid \nu\,), \tag{2.42}$$

where $F_{\chi^2}(\cdot)$ is the probability distribution function (PDF) of the chi-square distribution. The asymptotic distribution of χ^2 does not depend on the distribution $F(\, t \mid \boldsymbol{\theta}\,)$ tested or its parameters, but only on ν. Consequently, asymptotically the test statistic is distribution-free, and the test is nonparametric.

Because the test statistic is only asymptotically chi-square distributed, the sample size must be large enough. It has been found, that the approximation is rather good, when the sample size is about four times greater than the number of cells, even if some of the expected frequencies $[np(\, i \mid \boldsymbol{\theta}\,)]$ are quite small. If the cells are approximately equiprobable the sample size can be two times greater or even equal to the number of cells, depending on the significance level. (Conover 1980, Moore 1986)

Kolmogorov-Smirnov test For a sample $\{t_1, \ldots, t_n\}$ of n observations the *empirical distribution function* (EDF) gives the proportion of observations not greater than t:

$$F_n(t) = \left\{ \frac{j}{n} \;\middle|\; t_{(j)} \leq t < t_{(j+1)} \right\}, \qquad j = 1, \ldots, n. \tag{2.43}$$

If $\{t_1, \ldots, t_n\}$ is drawn from $F(\, t \,|\, \boldsymbol{\theta} \,)$, the EDF approaches $F(\, t \,|\, \boldsymbol{\theta} \,)$ as the sample size increases (Dudewicz & Mishra 1988):

$$\lim_{n \to \infty} F_n(t) = F(\, t \,|\, \boldsymbol{\theta} \,). \tag{2.44}$$

The null hypothesis is that the distributions are similar enough for $F_n(t)$ to be the EDF of a random sample generated by $F(\, t \,|\, \boldsymbol{\theta} \,)$:

$$H_0 \colon F_n(t) \sim F(\, t \,|\, \boldsymbol{\theta} \,). \tag{2.45}$$

It is possible to test the goodness of fit by measuring the discrepancy between the EDF and the proposed distribution. Such a statistic is called an EDF statistic (Stephens 1986*a*).

The Kolmogorov-Smirnov (K-S) statistic D is the largest absolute vertical difference between $F_n(t)$ and $F(\, t \,|\, \boldsymbol{\theta} \,)$:

$$D = \sup_{t \in \mathrm{Re}_+} |F_n(t) - F(\, t \,|\, \boldsymbol{\theta} \,)|. \tag{2.46}$$

Under H_0 the distribution of D is known, and it is independent of $F(\, t \,|\, \boldsymbol{\theta} \,)$. D can be calculated as follows:

$$D = \max \left\{ D^+, D^- \right\}, \tag{2.47}$$

where

$$D^+ = \max_{j \in \{1, \ldots, n\}} \left\{ \frac{j}{n} - F\big(t_{(j)} \,|\, \boldsymbol{\theta} \,\big) \right\} \tag{2.47a}$$

$$D^- = \max_{j \in \{1, \ldots, n\}} \left\{ F\big(t_{(j)} \,|\, \boldsymbol{\theta} \,\big) - \frac{j-1}{n} \right\}. \tag{2.47b}$$

The statistic is distribution-free, and its distribution is known. The significance probability of the test is

$$p = 1 - F_D(\, D \mid n \,), \tag{2.48}$$

where $F_D(\cdot)$ is the PDF of the K-S distribution.

The K-S test avoids the arbitrary discretizing of the distribution function, and the transformation of the observations to counts. This process, troublesome in small samples, is required by the chi-square test. The reduction of data makes the chi-square test less powerful than the K-S test (See Moore 1986).

Stuart & Ord (1991) report some comparisons between the chi-square and the K-S tests: For large samples with equal significance level (0.05) and equal power (0.5), the K-S test can detect a deviation about half as small as the chi-square test can. The ratio declines steadily in favor of the K-S test as the sample size increases. On the other hand, to detect an equal deviation, the K-S test requires the sample size to be of order $n^{4/5}$ compared to n for the chi-square test. As the sample size increases, the relative efficiency of the chi-square test tends to zero. The K-S test is a very much more sensitive test of fit for a continuous distribution. It is also asymptotically much more efficient.

The K-S test does not have a nonparametric modification for the case that the parameters are estimated from the sample. The alternative is a parametric test.

Parametric tests

Parametric Kolmogorov-Smirnov test If the parameters of a theoretical distribution must be estimated from the sample, the null hypothesis is:

$$H_0 \colon F_n(t) \sim F\Big(t \,\Big|\, \tilde{\boldsymbol{\theta}} \Big) . \tag{2.49}$$

The K-S test statistic is now:

$$D(\tilde{\boldsymbol{\theta}}) = \sup_{t \in \mathrm{Re}_+} \Big| F_n(t) - F\Big(t \,\Big|\, \tilde{\boldsymbol{\theta}} \Big) \Big| . \tag{2.50}$$

Because fitting the parameters makes it possible to adjust the tested distribution more efficiently than by using predetermined parameters, the critical values for a given significance probability must be smaller. The test statistic is no longer distribution free, but different critical values relate to different null hypotheses (Green & Hegazy 1976). The test depends on the distribution tested, the parameters estimated, the method of estimation, and the sample size (Stephens 1986*a*).

Although the distribution of $D(\tilde{\boldsymbol{\theta}})$ is unknown, the significance probability can be evaluated by a Monte Carlo method (see algorithm 2.1). For

Step 1 Estimate sample parameters $\tilde{\boldsymbol{\theta}}$.

Step 2 Calculate test statistic $D_n(\tilde{\boldsymbol{\theta}})$ for the sample.

Step 3 Set $i = 1$ and $k = 0$.

Step 4 Generate random replica (i) with distribution $F\Big(t \,\Big|\, \tilde{\boldsymbol{\theta}} \Big)$ and size n.

Step 5 Estimate the parameters $\tilde{\boldsymbol{\theta}}_i$ of replica i.

Step 6 Calculate test statistic $D_i(\tilde{\boldsymbol{\theta}}_i)$ for replica i.

Step 7 If $D_i(\tilde{\boldsymbol{\theta}}_i) \geq D(\tilde{\boldsymbol{\theta}})$ set $k = k + 1$.

Step 8 If $i < m$, set $i = i + 1$, and go to step 4. Else calculate

$$\tilde{p}_m = \frac{(k+1)}{(m+1)}$$

and stop.

Algorithm 2.1: A Monte Carlo method for testing goodness of fit

each sample, m random replicas are generated from $F\Big(t \,\Big|\, \tilde{\boldsymbol{\theta}} \Big)$. For each replica (i), the test statistic $[D_i(\tilde{\boldsymbol{\theta}}_i)]$ is calculated, and the number of replicas (k) having test statistic greater than or equal to $[D(\tilde{\boldsymbol{\theta}})]$ is counted. The estimator for the significance probability is:

$$\tilde{p}_m = \frac{k+1}{m+1}, \tag{2.51}$$

where

$$k = \#\big\{ i \;\big|\; D_i(\tilde{\boldsymbol{\theta}}_i) \geq D(\tilde{\boldsymbol{\theta}}) \big\}. \tag{2.51a}$$

Noreen (1989) has shown the validity of this test, and presented confidence intervals for the method. Assuming that all p_m are a priori equally likely (H_0 is assumed to be true), the probability that the true significance probability p is less than or equal to p_u is:

$$\mathbb{P}\{\, p \leq p_u \,|\, k, m \,\} = \frac{\int_0^{p_u} p^k (1-p)^{m-k} dp}{\int_0^1 p^k (1-p)^{m-k} dp}. \tag{2.52}$$

The upper confidence limit $p_{u,\alpha}$ at confidence level α can be obtained by solving the equation:

$$\mathbb{P}\{\, p \le p_{u,\alpha} \,|\, k, m \,\} = \alpha. \tag{2.53}$$

For $\tilde{p}_{9999} = 0.05$ the 95 % upper confidence limit is 0.059, for $\tilde{p}_{999} = 0.05$ the limit is 0.062, for $\tilde{p}_{499} = 0.05$ it is 0.067, and for $\tilde{p}_{99} = 0.05$ it is 0.089.

The results of the parametric K-S tests for the exponential distribution were compared with tables available in the statistical literature (Stephens 1986*a*, Durbin 1975). There were no significant differences.

Lilliefors (1967) compared the results of nonparametric and parametric K-S tests for the normal distribution: The 1 % critical values of the parametric test were approximately the same as the 20 % critical values of the nonparametric test. For the exponential distribution the percentages were 5 and 20, respectively (Lilliefors 1969). Accordingly, the nonparametric K-S test is much too conservative if the parameters are estimated from the sample.

Other parametric tests The K-S statistic measures the largest absolute difference between the EDF and the proposed model. Typically the largest differences are in the middle region of the distribution. Near the tails the absolute differences are smaller, but relative differences may be quite large. In this sense, the K-S statistic gives less weight to the tails. The statistic can be improved using a quadratic measure (Stephens 1986*a*):

$$Q = n \int_0^\infty [F_n(t) - F(t)]^2 \, \psi[F(t)] \, dF(t) \,, \tag{2.54}$$

where $\psi[F(t)]$ is a weight function. If $\psi[F(t)] \equiv 1$, the statistic is the Cramér-von Mises statistic (W^2), which can be calculated as follows (Green & Hegazy 1976):

$$W^2 = \sum_{j=1}^{n} \left(F\Big(t_{(j)} \,\Big|\, \tilde{\boldsymbol{\theta}} \Big) - \frac{j - \frac{1}{2}}{n} \right)^2 + \frac{1}{12n}. \tag{2.55}$$

Anderson & Darling (1952) suggested $\psi[F(t)] = \{F(t)\,[1-F(t)]\}^{-1}$ as the weight function. Because this is the reciprocal of the variance of $\sqrt{n}[F_n(t) - F(t)]$, in a certain sense, the statistic assigns equal weight to each point of the distribution $F(t)$. The Anderson-Darling (A-D) statistic (A^2) can be

calculated as (Green & Hegazy 1976):

$$A^2 = -n - 2\sum_{j=1}^{n}\left\{\frac{j-\frac{1}{2}}{n}\left[\ln F\Big(t_{(j)}\,\Big|\,\tilde{\boldsymbol{\theta}}\Big) + \ln\Big(1 - F\Big(t_{(n-j+1)}\,\Big|\,\tilde{\boldsymbol{\theta}}\Big)\Big)\right]\right\}. \tag{2.56}$$

The A-D test will fail if there are observations outside the range of the distribution. If even a single $F(t_j)$ is 0 or 1, the statistic grows to infinity, leading to the rejection of H_0. Consequently, the location parameter (threshold for minimum headway) should always be estimated smaller than the shortest headway in the sample.

Studies of these test have suggested that the parametric EDF tests are more powerful than the chi-square test. Among the parametric EDF tests, the K-S test is the least powerful, and the A-D test slightly better than the Cramér-von Mises. Stephens (1986*a*) suggests the A-D test as a omnibus test statistic for EDF tests with unknown parameters.

Based on these results, the goodness-of-fit tests for the headway data have been performed by the A-D method. The test procedure followed algorithm 2.1 with 9,999 replicas.[13] (In the tests for the semi-Poisson distribution the number of replicas was, however, 500.) The nonparametric K-S tests were calculated for comparison. Although these results could not be used to test H_0, they gave a measure of the closeness of fit comparable between different models. The comparison of the nonparametric K-S and parametric A-D tests also revealed the difference in the power of these tests.

2.2.5 Combination of probabilities

In the analysis of multi-sample data, it is desirable to find a single measure to describe the overall acceptability of the null hypothesis. The problem can be described by an example:

Example 2.1 The balance of a dice is tested. The null hypothesis is that the dice is balanced, and each outcome (1, 2, 3, 4, 5, and 6) has equal probability ($\frac{1}{6}$). The alternative hypothesis is that the dice favors large numbers. The dice is thrown five times, and the result is {5, 6, 6, 5, 6}. The probability for this or a better result is $\left(\frac{1}{6}\right)^5 \sum_{i=3}^{5}\binom{5}{i} = 0.002$.

It should be noted that the average probability (0.233) is not a proper measure of significance. Another interesting feature in the example is that

[13] The parameter estimation and the goodness-of-fit tests were performed by a computer program called Testfit. It is coded in C language, and it calls IMSL Fortran subroutines.

none of the individual test probabilities p_i is statistically significant, while the combined probability is highly significant.

In the example the probabilities were multiplied. An alternative method is to operate with the sum of the logarithmic probabilities ($\ln p_i$). The latter method will now be followed, as suggested by Fisher (1938) and Stephens (1986*b*).

In the study of headways, the overall null hypothesis (H_0) can be composite, such as:

$$H_0\text{: The headway data is generated by } F(\,t\,|\,\cdot\,)\,. \tag{2.57}$$

The hypothesis is composite, because the parameters of the distribution are unspecified. For each sample, the parameters are estimated, and the null hypothesis is simple:

$$H_{0i}\text{: The headway sample } i \text{ is generated by } F\Big(\,t\,\Big|\,\tilde{\boldsymbol{\theta}}\,\Big)\,. \tag{2.58}$$

H_0 is true if and only if H_{0i} is true for all samples.

If the null hypothesis (H_{0i}) is true for all samples, and the tests are independent, then the significance probabilities (p_i) are uniformly distributed between 0 and 1 [$U(0,1)$]. If k tests are made, the probabilities $p_1, \ldots, p_k$ are a random sample from $U(0,1)$. This can be tested by a method presented by Fisher (1938).

For the chi-square distribution with two degrees of freedom it holds:

$$p = F_{\chi^2}(\,x \mid \nu = 2\,) = e^{-\frac{1}{2}x}. \tag{2.59}$$

Taking the natural logarithm of the probability p and multiplying it by -2 gives the equivalent value of x for two degrees of freedom:

$$x = F^{-1}_{\chi^2}(\,p \mid \nu = 2\,) = -2\ln p. \tag{2.60}$$

If p is a uniform (0,1) random variate, then x follows the chi-square distribution with two degrees of freedom.

The sum of m variates following the chi-square distribution with two degrees of freedom is itself chi-square distributed with $2m$ degrees of freedom. If H_0 is true, and the samples are independent, then the significance probabilities p_i are $U(0,1)$, and the statistic

$$z = -2\sum_{i=1}^{m} \ln p_i \tag{2.61}$$

has chi-square distribution with $2m$ degrees of freedom (Stephens 1986*b*). The significance (P) of the combined test is now the probability that a random variate (Z) having chi-square distribution with $2m$ degrees of freedom, is greater than or equal to z:

$$P = \mathbb{P}\{Z \geq z\} = 1 - F_{\chi^2}(\, z \mid \nu = 2m\,). \tag{2.62}$$

This method can be used to obtain a single significance probability for a multi-sample data set.

2.2.6 Moving probability

Sometimes it is desirable to know, whether a null hypothesis should be rejected in some conditions but not rejected in other conditions. It is possible to divide the samples into classes according to some factor and to calculate the combined probability for each group. The classification of the data is not, however, always obvious. For example, there is no obvious way to classify samples by traffic volume.

Let us assume that a set of probabilities p_i is arranged in ascending order according to the traffic volume. The combined probability for the first k samples can be calculated as described above. It describes the combined significance probability corresponding to the traffic volume equal to the mean of the first k sample volumes. Next the second sample is selected as a starting point, and the combined probability for k consecutive samples and the corresponding mean volume is calculated. This procedure is repeated until the combined probability is calculated for the last k samples. This series of combined probabilities is called the *moving probability*. This method, suggested by Luttinen (1992), smoothes the variation of the significance probabilities and increases the power of the tests by combining the significance probabilities for several samples.

If the samples are arranged in an ascending order according to the traffic volume, the k-point moving probability is:

$$P_k(\tilde{\lambda}_{k,j}) = 1 - F_{\chi^2}(\, z_{k,j} \mid \nu = 2k\,), \tag{2.63}$$

where

$$z_{k,j} = -2 \sum_{i=j}^{j+k-1} \ln p_i \tag{2.63a}$$

$$\tilde{\lambda}_{k,j} = \frac{1}{k} \sum_{i=j}^{j+k-1} \tilde{\lambda}_i, \qquad j + k - 1 \leq m, \tag{2.63b}$$

$\tilde{\lambda}_i$ is the traffic volume of sample i, $\tilde{\lambda}_{k,j}$ is the traffic volume of the jth term, and m is the total number of samples. The traffic volume of term j is the mean of the sample volumes in the term.

A balance must be found between sensitivity and power. A large number of points (k) in each term makes the curve smoother and increases the power of each combined significance, but the curve is also less sensitive to deterministic variations. A small number of points makes the curve more sensitive and ragged, but the power of each combined significance decreases.

In a moving average the number of points should be selected as the lowest value producing a reasonably smooth curve. For a moving probability the situation is different: Assume that all tests give significance probabilities near 0.2. Even a small number of points produces a smooth curve, but the result is not realistic, because the combined probabilities are too large. (For a 5-point moving probability the values are near 0.1, while a 10-point moving probability is near 0.04.) Consequently, the number of points k in each term should be as large as possible, but such that the moving probability still covers a sufficiently wide range of traffic volumes, and that it is sensitive to local variations.

Chapter 3

Data collection

3.1 Measurements

The purpose of the field studies was to examine the time headways in freely flowing traffic on two-lane two-way roads. All the road sections selected were reasonably level, straight and with no nearby traffic signals.

Arrival characteristics were measured by a traffic analyzer. The data were collected by the staff of the Laboratory of Transportation Engineering at the Helsinki University of Technology. Using two inductive loops in both directions (fig. 3.1), the analyzer recorded for each passing vehicle its serial number, gross time headway (time interval from front bumper to front bumper in units of 1/100 s), net time headway (time interval from back bumper to front bumper in units of 1/100 s), speed (in units of 1 km/h) and length (in units of 1/10 m). The data were transmitted to a computer. The headways on both lanes were measured and analyzed separately.

The measurements were performed during the summer of 1988 in southern Finland on 9 locations having speed limits 50, 60, and 70 km/h. These roads are called "low speed roads". The locations and the measurement times were selected so that the expected flow rates were preferably below 1,000 veh/h. The weather was mostly dry, but two measurements failed because of rain or detector malfunction.

Also some data from year 1984 (Pursula & Sainio 1985) measured on 10 locations with speed limits 80 and 100 km/h ("high speed roads") have been used. These measurements were concentrated on higher volumes, because capacity was a major aspect of that study. Consequently, the opposite volumes tend also to be higher than on the low speed roads. The data have been previously analyzed by Pursula & Sainio (1985) and by Pursula

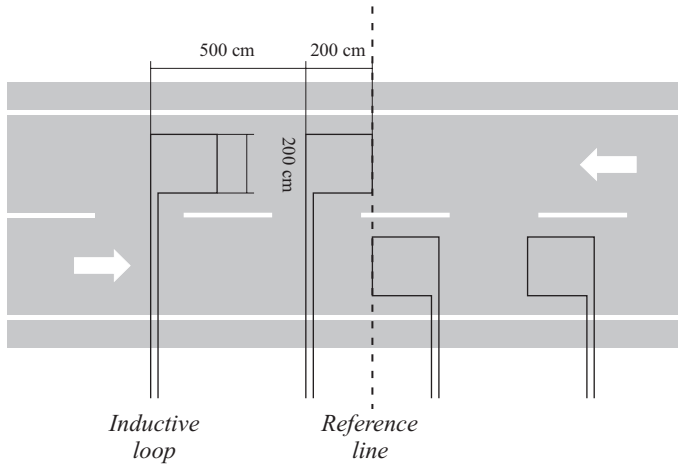

Figure 3.1: Inductive loops of the traffic analyzer

& Enberg (1991). In all of these measurements more than 73,000 headways were recorded.

In addition to the speed limits and the volumes, the road categories have other differences. The speed limit describes also the road conditions, most importantly the road geometry and the intersection density. High speed roads have usually a better geometry. Lower speed limits (50 and 60 km/h, in particular) are typically located near densely populated areas, where the intersection density is high.

3.2 Preliminary data analysis

The data were checked for consistency. The measured speeds (km/h) were corrected on the basis of radar measurements. The corrected speed (v_c) was calculated from the measured speed (v_m) using the relation:

$$v_c = 4.77 + 0.911 v_m. \tag{3.1}$$

Measuring periods having more than one percent passing vehicles, were discarded. Also, the first observation (time from the beginning of the measurement to the arrival of the first vehicle) and passing vehicles (crossing the detectors to the wrong direction) were removed from the data. Accordingly, the sampling was synchronous (see Luttinen 1990).

Detection and elimination of outliers was done during the trend analysis on the basis of graphical displays of headways and speeds. The outliers were

quite apparent.

3.3 Selection of trend free samples

In order to obtain samples with stationary data, trend analysis was performed by a computer program called Trendana[1] showing graphically each headway, the 15-point moving average, the cumulative vehicle count and the speed of each vehicle.

The data were analyzed sequentially using the exponential ordered scores trend test (section 2.2.2). The sample size was incremented by 50 vehicles until the test reported trend at 5 % level of significance. The sample was then decremented until the level of significance for trend was near 50 % or at least between 30 and 70 %, the sample size was greater than 100, and the sampling period was between 5 and 40 minutes. The 40 minute limit was set in order to keep the test sensitive to local trends, but under low flow conditions the sample size or the period length condition had to be relaxed sometimes.

If a satisfactory sample was not found, the first observations in the sample that seemed to cause the trend were removed, and the process was repeated. Also, periods with low speeds indicating disturbances were excluded, and the process was repeated starting after the excluded observations. Samples with more than 10 % heavy vehicles (longer than 6 m) were excluded from further analysis.

3.4 Description of selected samples

The preliminary analysis and the trend tests produced 65 samples consisting of 16,780 observations. The samples represent speed limits 50, 60, 70, 80 and 100 km/h (fig. 3.2). The per lane flow rates vary from 150 to 1,840 veh/h (figs. 3.3 and 3.4). The basic properties of the samples are listed in appendix A.

The observations from high speed roads were gathered in 1984. Because that study was more focused on the capacity and level of service issues, the observations concentrate on high volumes. On low speed roads the observations are more concentrated on low volumes. Consequently, speed limits and volumes are correlated. The high speed roads also have a higher overall standard and more opposing traffic. Hence, care must be taken

[1] The program is coded in C language, and it calls IMSL Fortran subroutines.

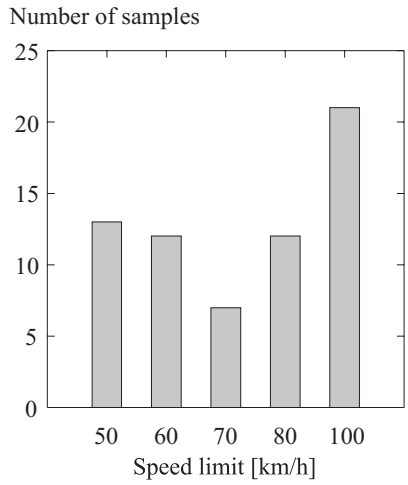

Figure 3.2: Number of samples by speed limit

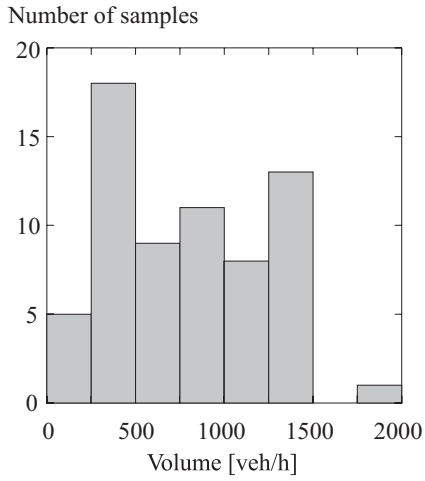

Figure 3.3: Number of samples by volume

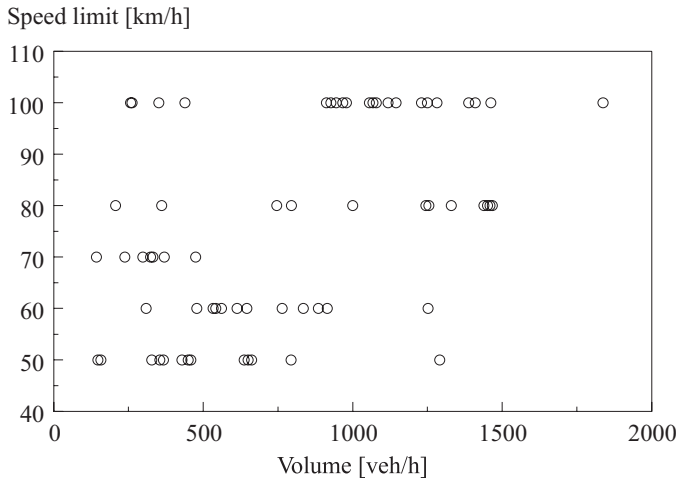

Figure 3.4: Sample volumes by speed limit

in conclusions about the influence of the speed limit and the volume on headways.

From low speed (50–70 km/h) roads there are no observations with volume range 1,000–1,250 veh/h. From high speed (80–100 km/h) roads the observations are missing for volume range about 500–700 veh/h. From roads having speed limit 70 km/h there are only low volume (< 500 veh/h) samples. In the whole data set, there is only one sample having volume greater than 1,500 veh/h.

Figure 3.5 shows the volume and the space mean speed (calculated as harmonic mean of spot speeds) of each sample. Because the samples are from different locations, no curve fitting has been done. The data, however, shows the decrease of mean speed with increasing volume. Except for two samples (12 and 33), the figure does not reveal any congestion or disturbances in the flow.

The low mean speed of sample 12 indicates congestion. Sample 33 has a low mean speed relative to flow rate (361 veh/h). Further analysis showed that the sample also has a very high variation in relative speeds of leading vehicles. This indicates some kind of a disturbance in the flow.

After the exclusion of samples 12 and 33, the final data set consists of 63 samples having 16,417 observations. Accordingly, less than 25 % of the total measurements were included in the analysis. These data describe traffic flow on two-lane two-way roads under stationary conditions with not more than 10 % heavy vehicles and very few passing vehicles.

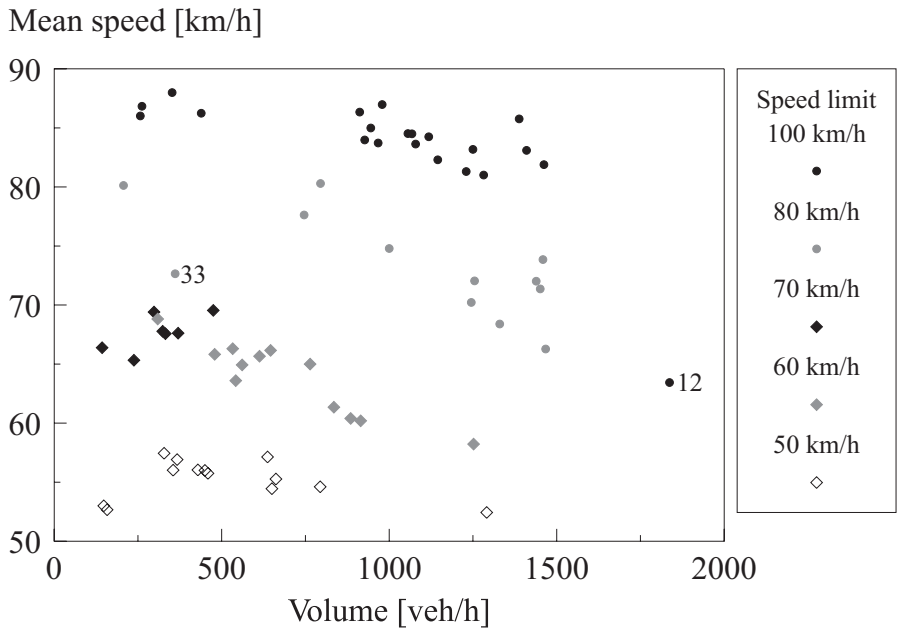

Figure 3.5: Volume-speed diagram

Chapter 4

Identification

4.1 The purpose of identification

A headway (T) describes the time interval between consecutive vehicles in traffic flow. For a traffic engineer it is interesting to examine the shape and structure of the headway distribution. The differences between follower and free flowing headways are of special interest. In capacity and level-of-service studies, the proportions of short and long headways deserve attention. A simulation designer is interested in the possible correlation between consecutive headways.

Identification is the process of finding, from a set of models, the best theoretical model to describe the data. It can be a stochastic model or, as here, a probability distribution. An effective description of the most important properties of the data is essential in order to find a theoretical model which has properties similar to the data.

This chapter discusses the shape of the headway distribution, measures of location and dispersion, and tests for the renewal hypothesis, which are concerned with the autocorrelation of consecutive headways. Some of the methods (e.g., the hazard function and the KS^2-chart) have not been used in headway studies—at least not widely. For some commonly used methods, more efficient procedures are presented. The concluding section presents a four stage identification process, which is suggested as a standard procedure in headway studies.

4.2 Shape of the distribution

4.2.1 Empirical distribution function

The probability distribution function (PDF) defines the probability that a headway random variate (T) is not greater than some given value t:

$$F(t) = \mathbb{P}\{T \leq t\}. \tag{4.1}$$

The PDF which has generated a sample of size n can be estimated by the empirical distribution function (EDF). It gives the proportion of observations less than or equal to t in the sample:

$$F_n(t) = \left\{ \frac{j}{n} \;\middle|\; t_{(j)} \leq t < t_{(j+1)} \right\}, \qquad j = 1, \ldots, n. \tag{4.2}$$

The EDF has a central role in the goodness-of-fit tests (section 2.2.4).

Figure 4.1 shows contour plots of empirical headway distribution functions. Separate plots for low speed (50–70 km/h) and high speed (80–100 km/h) roads were produced by smoothing the sample EDFs. Because the traffic volume of only three (low speed) samples is below 200 veh/h, the low volume contour lines are not very reliable.

The proportion of extremely short headways is small, but between one and three seconds the EDF rises steeply. On high speed roads the concentration on short headways is more substantial than on low speed roads. At 500 veh/h the low speed roads have about 30 % of headways shorter than two seconds. At the same volume the high speed roads have about 40 % of headways shorter than two seconds, and the proportion of short headways is rather constant at moderate to high flow rates. The 0.9 contours are almost identical in both road categories.

At high volumes the contour lines are very densely concentrated at short headways, but they curve toward longer headways at lower volumes, thus indicating a greater proportion of long headways. This is only natural, considering that traffic volume is the reciprocal of the mean headway. According to May (1990), the percentage of vehicles having headways less than or equal to the mean headway is rather constant, about 67 %. This approximation holds for the low speed data.

In HCM (1985) the percent time delay is used as a level of service measure. For practical purposes, the percent time delayed is approximated by the percentage of vehicles traveling in platoons ($t < 5$). The level of service is accordingly based directly on the empirical headway distribution. In figure 4.2 the proportion of vehicles having headways less than five seconds,

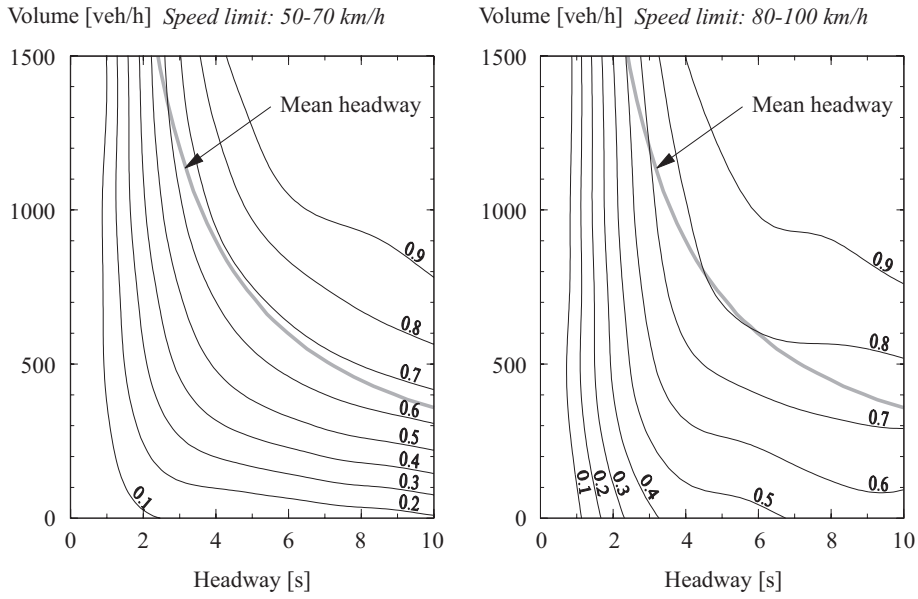

Figure 4.1: Contour plots of empirical headway distribution functions

is plotted against the traffic volume. Except for two low volume samples (45 and 46), the proportion is larger than in a totally random, or Poisson, process (negative exponential headways). The proportion is also larger on high speed roads than on low speed roads.

It should be noted that the data in figure 4.2 are from several sites. Thus, the variation is partly caused by local conditions. However, the figure shows that the data are consistent in respect of the level of service.

Figure 4.3 shows the proportion of headways less than or equal to one second. The proportion is 0.11 for high speed roads and 0.062 for low speed roads. There is no significant correlation with the traffic volume. For all but the lowest volumes, the proportion is lower than in a Poisson process.

Enberg & Pursula (1992) obtained similar results for Finnish high-class two-lane roads.[1] The data did not indicate a significant correlation between very short ($t < 1$ s) headways and traffic volume. There is, however, some evidence that on Finnish freeways the proportion of headways less than one second increases with increasing traffic volume (Salonen 1982).

[1] These roads have a very high standard of alignment and grade-separated intersections. The speed limit is 100 km/h, and only fast motor vehicles are allowed.

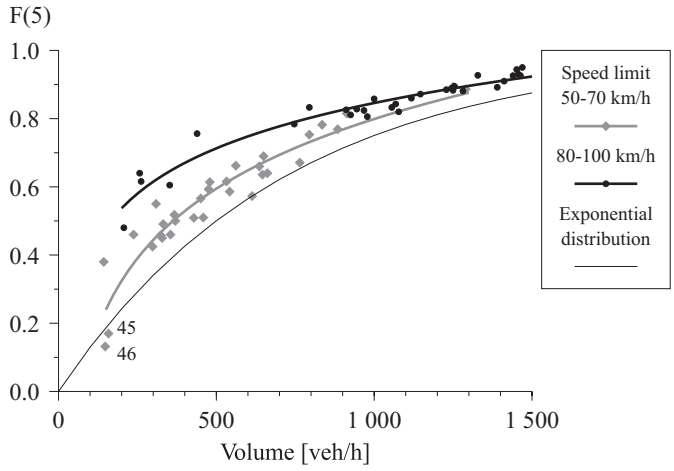

Figure 4.2: Proportion of headways less than five seconds

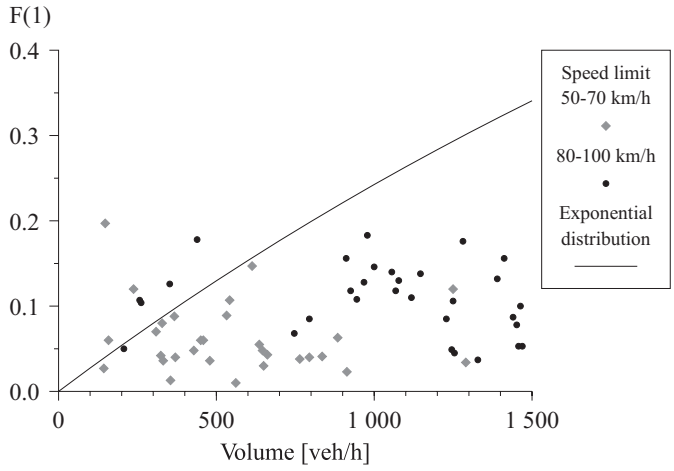

Figure 4.3: Proportion of headways less than or equal to one second

4.2.2 Empirical density function

The probability distribution function (PDF) can be defined in terms of the probability density function (pdf):

$$F(t) = \int_0^t f(x)\ dx. \tag{4.3}$$

Thus, the pdf is the derivative of the PDF, and:

$$\int_0^\infty f(x)\ dx = 1. \tag{4.4}$$

Both the PDF and the pdf have the same and complete information about the distribution, but the pdf is visually more informative. Because of this, the plotting of the pdf estimates (empirical density functions, edf's) is a standard procedure in headway studies.

The density function is usually estimated by the *histogram method*. Starting from an origin (t_0), the range of the data is partitioned into bins (class intervals) having the same width (h), which is also called the window width. The height of the histogram at point t is the number of observations belonging to the same bin as t divided by the number of total observations (n) and the bin width (h):

$$\begin{aligned} f_n(t) = \frac{1}{nh} \#\big\{ i \mid\ & [t_0 + jh \le t < t_0 + (j+1)h] \\ & \wedge [t_0 + jh \le t_i < t_0 + (j+1)h] \big\} \\ & i = 1, \ldots, n \text{ and } j = 0, 1, \ldots \end{aligned} \tag{4.5}$$

The total area under the histogram is unity. Once the origin and the bin width are selected, the method is very straightforward, and there are many programs to automate the process.

The histogram method has two major drawbacks: Firstly, the location of the origin (t_0) and the bin width (h) can have a considerable effect on the shape of the estimate, especially at low values of t. Secondly, the discontinuity of the histograms causes even more difficulties if the derivatives of the density estimate are needed, or if the density estimate is needed as an intermediate component for other methods. (Silverman 1986)

The histogram method calculates the frequency of observations in a window of width h. The window is first placed at the origin, and then shifted to the right by its own width. Each observation in the window is given equal weight. Better estimates are, however, obtained if the the window slides

smoothly along the x-axis, and the observations are weighted by an appropriate smoothing function (kernel). The observations in the middle of the window should have greater weight, while the observations near the sides of the window should be given less weight. This procedure is called the *kernel method*.

The kernel method produces a continuous estimate of the pdf, avoiding the pitfalls of the histogram method, but some noise in the tail cannot be avoided without losing essential details. The kernel density estimator with kernel $K(\cdot)$ at point t is defined as:

$$f_n(t) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{t - t_i}{h}\right). \tag{4.6}$$

Too small values of h produce noisy estimates. On the other hand, too large values smooth away essential details of the density.

For headway data, the estimate should reproduce the sharp peak of the pdf without too much noise in the tail. As the density estimates were further smoothed in order to produce the surface plots (figures 4.4 and 4.5), a small smoothing parameter ($h = 0.3$) was selected. This estimator was able to produce the sharp peak of the density, even though there was some noise in the tail.

The estimates are based on the Epanechnikov kernel (Silverman 1986):

$$K(x) = \begin{cases} \frac{3}{4\sqrt{5}} \left(1 - \frac{1}{5}x^2\right), & \text{if } |x| < \sqrt{5} \\ 0, & \text{otherwise.} \end{cases} \tag{4.7}$$

It is the most efficient kernel function, although the differences between the best kernels are very small (Silverman 1986).

Figures 4.4 and 4.5 show surface plots of headway density estimates. The surfaces were smoothed by a polynomial fit for each t in the grid. The method produces some wavelike nonsmoothness in the t direction of the surface, but the shape of the density functions is very clear. The method of distance weighted least squares was also tried. The surfaces were smoother, but the peak was lower, and it had saddle points in the regions of missing observations (see Luttinen 1994).

The density is unimodal, bell-shaped and skewed to the right. The proportion of headways less than one second is small. The mode is between one and two seconds, after which the frequencies diminish in an exponential way. The most striking difference between the figures is that under low volume conditions the frequency of short headways is much lower on low speed

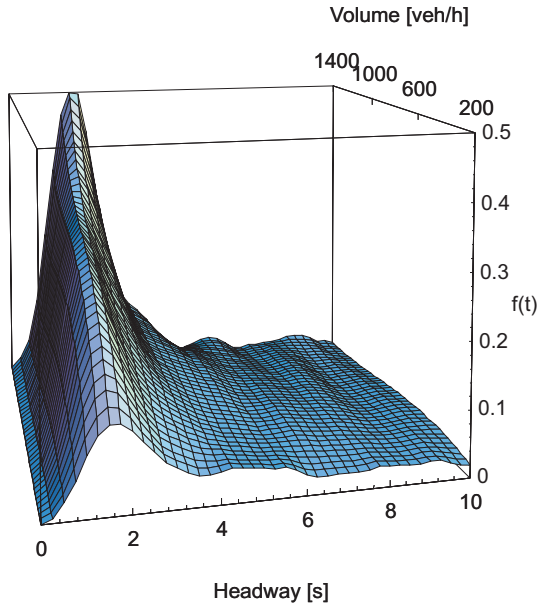

Figure 4.4: Empirical headway density function on low speed (50–70 km/h) roads

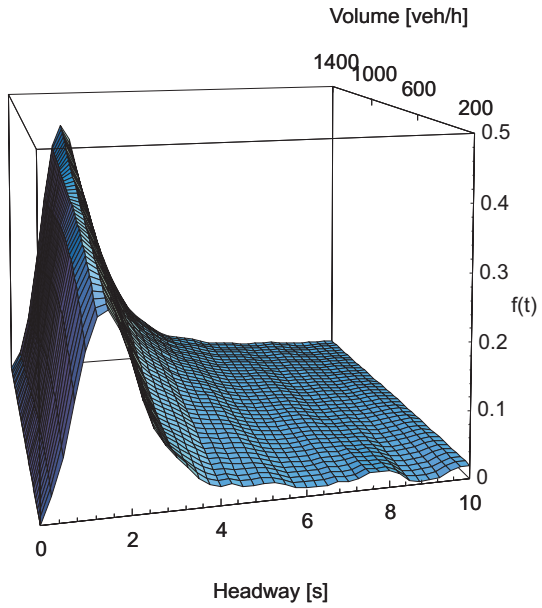

Figure 4.5: Empirical headway density function on high speed (80–100 km/h) roads

roads. As the volume increases, the frequency of short headways, of course, increases on both road categories.

The difference between the two road categories may be caused by greater variation in speed differences on high speed roads (fig. 5.30, page 133), different road conditions, and differences in the opposite flow level. The low speed (50–70 km/h) data were collected near densely populated areas, where intersection density was higher than on high speed (80–100 km/h) roads.

The edf's have some similarity with the pdf's of the gamma and the lognormal distributions. Also, several mixed distributions have similar densities. The densities of the negative exponential and the shifted exponential distributions are, however, different. These distributions have the greatest density at the shortest headways, whereas the edf increases smoothly from zero to the mode. Consequently, the exponential distributions are not appropriate models for short headways.

4.2.3 Empirical hazard function

The hazard function is widely used in reliability and life data analysis. It gives the instantaneous failure rate for a system that has not failed until t. The hazard function for a random variable T is defined as:

$$\begin{aligned} h(t) &= \lim_{\Delta t \to 0} \frac{\mathbb{P}\{\, t \le T < t + \Delta t \,|\, t \le T \,\}}{\Delta t} \\ &= \frac{f(t)}{1 - F(t)}. \end{aligned} \tag{4.8}$$

The pdf and the PDF can be expressed in terms of the hazard function as follows:

$$f(t) = h(t) \exp\left(- \int_0^t h(u)\, du \right) \tag{4.9}$$

$$F(t) = 1 - \exp\left(- \int_0^t h(u)\, du \right). \tag{4.10}$$

The hazard function also has the complete information of the distribution. If the hazard function has a simple shape, it is possible to form a hazard-based model using the equations above.

In terms of traffic flow, the hazard function $h(t)$ is the instantaneous termination rate of headways conditional upon lasting to time t. Besides of its theoretical significance, the hazard function shows properties of the distribution that are not so clearly seen in either the pdf or the PDF.

A crude estimate of the hazard function can be quickly calculated by classifying the data into intervals of length a, and calculating:

$$h_n(t) = \frac{\#\{\, t_i \mid t \le t_i < t + a \,\}}{\#\{\, t_j \mid t \le t_j \,\}}, \qquad i,j = 1,\ldots,n. \tag{4.11}$$

Better estimates can, however, be obtained by the kernel method. The smoothed empirical hazard function (ehf) of the headway data was calculated using a biweight kernel (IMSL 1989):

$$K(x) = \begin{cases} \dfrac{15}{16}\left(1 - x^2\right)^2, & \text{if } |x| < 1 \\ 0, & \text{otherwise.} \end{cases} \tag{4.12}$$

The smoothed kernel is:

$$K_s(t - t_{(i)}) = \frac{1}{\alpha d_{ik}} K\left(\frac{t - t_{(i)}}{\beta d_{ik}}\right), \tag{4.13}$$

where α, β, and k are smoothing parameters, d_{ik} is the distance to the kth nearest "failure" from $t_{(i)}$, and $t_{(i)}$ is the ith order statistic of the sample. The smoothing parameters were estimated by a modified maximum likelihood method.[2] Consequently, the ehf was calculated as follows:

$$h_n(t) = \sum_{i=1}^{n} \frac{1}{n - i + 1} K_s(t - t_{(i)}), \tag{4.14}$$

where n is the sample size.

Figures 4.6 and 4.7 show surface plots of the empirical hazard functions. The surfaces are smoothed by making a polynomial fit for each t in the grid. The method produced a little nonsmoothness in the t direction, but the shape is obvious. The method of distance weighted least squares was also tried. It produced smoother surfaces, but the peak had saddle points in the regions of no observations.

The figures have some common properties: The peak is between 1.5 and 2 seconds. It moves slightly to the right, as the volume increases. The peak is low at low volumes, and increases as the volume increases. After the peak, the ehf decreases, until it reaches a constant level. This level is higher at high volumes.

[2]IMSL subroutines DHAZEZ and DHAZST were used to estimate the parameters and to calculate the empirical hazard functions.

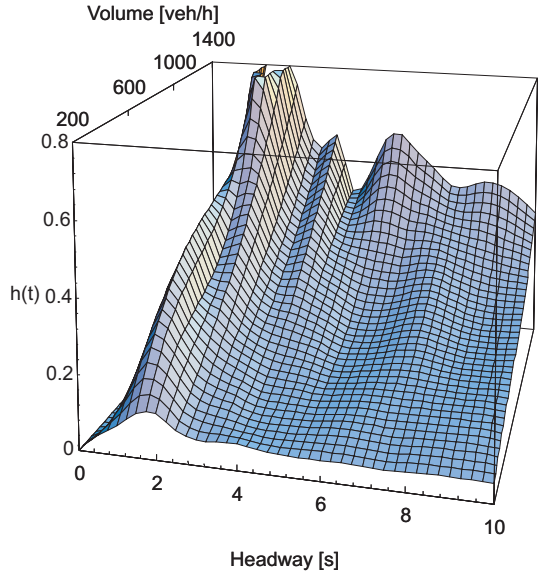

Figure 4.6: Empirical headway hazard functions for low speed (50–70 km/h) roads

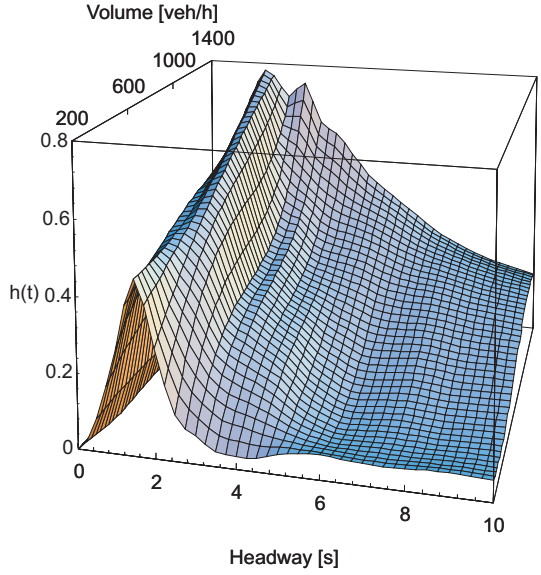

Figure 4.7: Empirical headway hazard functions for high speed (80–100 km/h) roads

At low volumes the peak is higher on high speed roads. When the volume on low speed roads is greater than about 1,000 veh/h, the peak rises very steeply. There are, however, only two low speed samples for volumes higher than 1,000 veh/h. The extremely high peak is caused by one small ($n = 150$) sample (no. 52), which has $h(2.0) = 1.8$. Thus, the high peak should be viewed with some skepticism. Although the ehf in figure 4.6 is nonsmooth at high volumes, the basic shape is obvious.

The hazard function of the *exponential distribution* is constant, and equal to the flow rate. The exponential tendency of low volume traffic is evident on low speed roads; the ehf becomes more level as the volume decreases.

If the tail of the distribution is exponential, the hazard function reaches a constant value at long headways. This value is proportional to the flow rate. Constant ehf values at large headways are accordingly an indication of an exponential tail in the headway distribution.

The hazard rate of the *gamma distribution* is monotone decreasing from infinity if the shape parameter (α) is less than unity, monotone increasing from zero if α is greater than unity, and in both cases approaches a constant value (scale parameter β) as t becomes large (Kalbfleisch & Prentice 1980). If the shape parameter is $\alpha = 1$, the distribution is exponential, and the hazard function remains constant [$h(t) = \beta$]. This kind of shape is seen at the lowest volumes on low speed roads, except for the low peak in the ehf. This suggests that at low volumes the headway distribution resembles a gamma distribution, and the shape parameter (α) approaches unity from above, as the flow rate decreases to zero. Because $h(0) = 0$ for $\alpha > 1$, the gamma distributions do not produce extremely short headways, and because $h(t)$ approaches β as t becomes large, the tail of the distributions approaches exponentiality.

The *lognormal* hazard function is zero at $t = 0$, increases to maximum, and then decreases, gradually approaching zero at long headways (Kalbfleisch & Prentice 1980). The basic shape is similar to the ehf, but the fall after the peak is more gradual.

The hazard function of the *semi-Poisson distribution* is similar to the headway hazard function estimate. The model in figure 5.49 (page 160) has gamma distributed short headways and exponential tail.

In conclusion, the ehf suggests a mixed distribution with an exponential tail as a theoretical headway distribution. The lognormal distribution has desirable properties at short headways only. The gamma distribution could be considered for low volumes.

4.3 Measures of location and dispersion

4.3.1 Mode and peak height

The mode (Mo) is the point where the density reaches its maximum. So, it is an approximation of the most frequent value in the distribution. Summala & Vierimaa (1980) describe the mode as an approximation of the headway that most drivers select when they are following the vehicle ahead. As such, the mode is a measure of typical platoon behavior and a crude estimate of highway capacity.

The peak height (Ph) is a measure of the frequency at mode. For unimodal distributions the mode and the peak height can be defined as:

$$Mo = \{t \in \mathrm{Re}_+ \mid f(t) = \max_{x \in \mathrm{Re}_+} f(x)\} \tag{4.15}$$

$$Ph = f(Mo). \tag{4.16}$$

The mode and the peak height of the headway data were estimated using kernel estimates for density. The smoothing parameter of the Epanechnikov kernel (4.7) was $h = 0.2$. A low smoothing parameter was used, because the density was estimated only at the range where the observations were most densely concentrated. Consequently, the peak height obtained was higher than in the density estimates. The golden section search (Gill, Murray & Wright 1981) with absolute precision of 0.01 was used to find the maximum of the edf.

The mode values are shown in figure 4.8. There is no significant correlation with flow rate. The mode is about 1.53 (with standard deviation of 0.29) on low speed roads and 1.41 (with standard deviation of 0.22) on high speed roads. The difference is, however, not statistically significant at 5 % risk level ($P = 0.08$), according to the t-test. Thus, the mode is a little below 1.5 (95 % confidence region is 1.46–1.48) at both high and low speed roads.

Summala & Vierimaa (1980) measured the modes of headway distributions on two-way two-lane roads with speed limits 50–100 km/h. The mode varied between 1.0 and 1.9, but it was rather constant at all traffic volumes. Rajalin & Hassel (1992) studied net time interval (gap) distributions on two-way two-lane roads with speed limits 80 km/h and 100 km/h. They reported a mode of 0.8 s at all traffic volumes.

The peak heights are shown in figure 4.9. The peak rises as the flow increases. On high speed roads (speed limit 80–100 km/h) the peak is higher than on low speed roads, especially at low volumes. At high volumes the speed limit looses its significance.

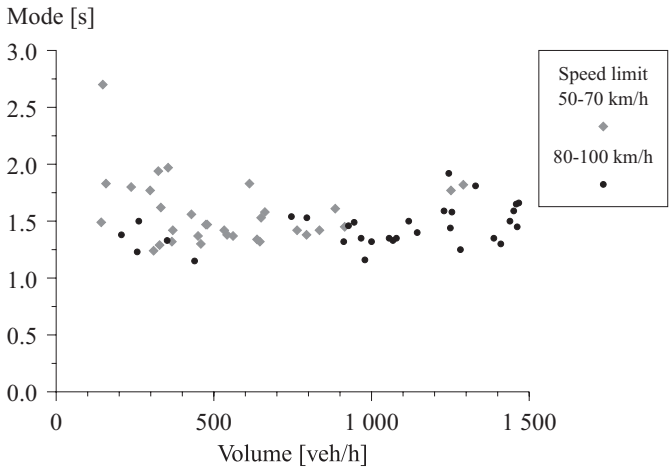

Figure 4.8: Mode of empirical headway distributions

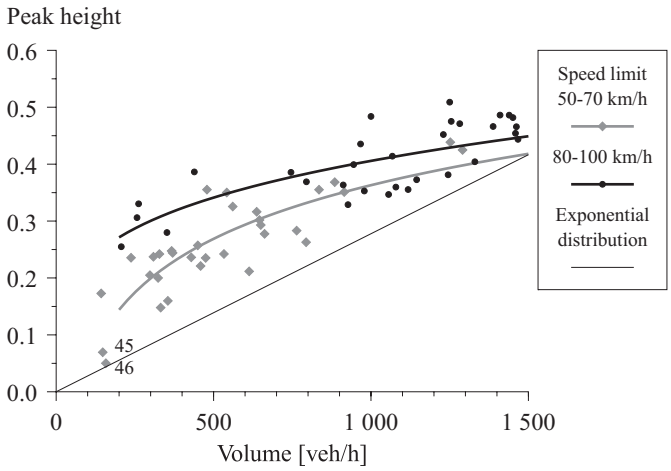

Figure 4.9: Peak height of empirical headway distributions

A regression analysis was performed by means of a logarithmic regression $[a + b\ln(\tilde{\lambda})]$ for low speed roads and a power regression $(a\tilde{\lambda}^b)$ for high speed roads. The fit is reasonable: $R^2 = 0.737$ on low speed roads and $R^2 = 0.620$ on high speed roads. The equality of residual means between high speed and low speed roads was tested using the t-test. The hypothesis should be rejected on risk level $P = 0.002$. The exclusion of the two very low (< 0.05) peak values (samples 45 and 46) makes the result statistically more significant ($P = 0.001$). Similar result ($P = 0.007$) was obtained using analysis of covariance (ANCOVA). Accordingly, separate models should be presented for lower and upper speed levels.

Enberg & Pursula (1991) reported only small differences in peak heights on high-class two-lane rural roads in Finland. The sampling period was 15 minutes. The mode was in the interval 1–2 s. When 60 minute samples were used, the increase in the peak height was more evident (Enberg & Pursula 1992).

4.3.2 Median

The median (Md) of a continuous random variate is the 50th percentile of the distribution:

$$\mathbb{P}\{T \leq Md\} = 0.5. \tag{4.17}$$

The sample median is a value such that half of the sample observations are to the left and half to the right of it:

$$\tilde{Md} = \begin{cases} t_{(k)}, & \text{if } n = 2k - 1 \\ \frac{1}{2}\left(t_{(k)} + t_{(k+1)}\right), & \text{if } n = 2k. \end{cases} \tag{4.18}$$

where $t_{(k)}$ is the kth order statistic of the sample, and n is the sample size. Sample medians of the headway data are shown in figure 4.10. The median on high speed roads is surprisingly level ($R^2 = 0.462$), but there are only few low volume samples. On low speed roads the volume and the median are more clearly correlated ($R^2 = 0.855$).

For symmetrical unimodal distributions the mean, the median and the mode coincide. It is a known property of nonsymmetric unimodal distributions that the mean, the median, and the mode occur in the same or the reverse order as in the dictionary, and that the median is nearer to the mean than to the mode (Stuart & Ord 1987). Because the headway pdf is positively skewed, these measures occur in the reverse dictionary order.

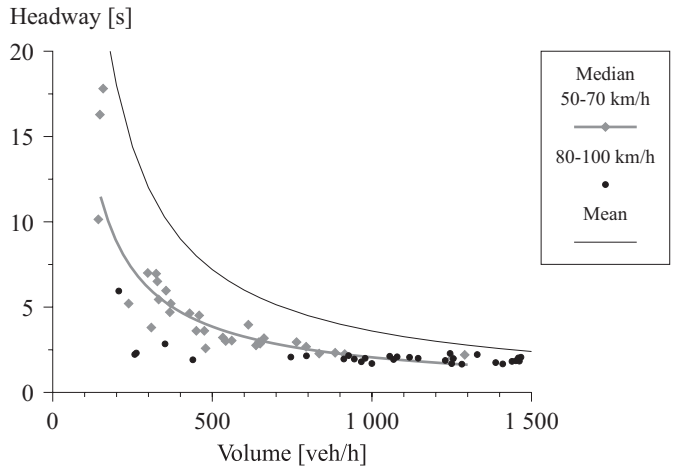

Figure 4.10: Median of empirical headway distributions

According to Stuart & Ord (1987) the following approximate relation holds for unimodal probability distributions of moderate asymmetry:

$$Md - Mo \approx 3[\,\mu(T) - Md\,]. \tag{4.19}$$

The headway distribution is, however, too asymmetric for this relation to hold, and the relation gives too high estimates of the median. The linear least squares estimates are (fig. 4.11):

$$\tilde{Md}_L = 0.529 + 0.465\bar{t} \tag{4.20}$$

$$\tilde{Md}_H = 1.601 + 0.107\bar{t}. \tag{4.21}$$

The accuracy of the estimate for high speed roads ($\tilde{Md}_H$) is, however, suspect, because there are only few low volume observations. The coefficient of determination is only $R^2 = 0.520$ for $\tilde{Md}_H$, as compared to $R^2 = 0.825$ for $\tilde{Md}_L$.

Separate models were estimated for low and high speed roads, although ANCOVA gave significance probability of only $P = 0.059$ for the homogeneity of slopes. The figure, however, shows that the slopes of the regression lines are quite different, but most of the points from the high speed roads are near the intersection of the lines. In order to stabilize the variance, a weighting factor of $1/\bar{t}^2$ was used. The R^2 for the weighted values was 0.966 for the low speed data and 0.990 for the high speed data. The adjusted significance probabilities were $P < 0.001$ and $P = 0.002$, respectively. The regression is accordingly statistically significant for both data sets.

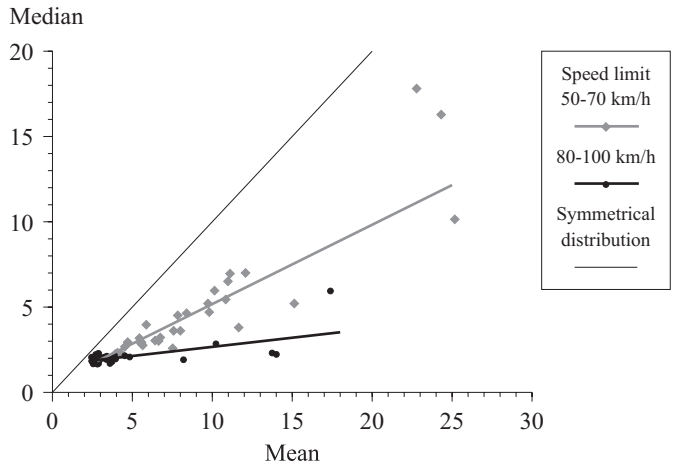

Figure 4.11: Mean and median of empirical headway distributions

4.3.3 Coefficient of variation

The coefficient of variation (c.v.) is the proportion of the standard deviation to the mean of a random variable (T):

$$C(T) = \frac{\sigma(T)}{\mu(T)}. \tag{4.22}$$

The sample c.v. is the proportion of the sample standard deviation to the sample mean:

$$C = \frac{s}{\bar{t}}. \tag{4.23}$$

The negative exponential distribution (headway distribution of a totally random arrival process) has $C(T)$ equal to unity. Because the negative exponential distribution is central in the theory of stochastic processes, and it is widely applied in traffic flow theory, $C(T) = 1$ provides a convenient point of comparison.

The CV-chart in figure 4.12 shows the sample coefficients of variation for the headway data. Some basic properties can be observed:

1. Under heavy traffic conditions, the proportion of freely moving vehicles is small. The variance of headways is accordingly small. In figure 4.12, the c.v. decreases below unity at high flow levels.

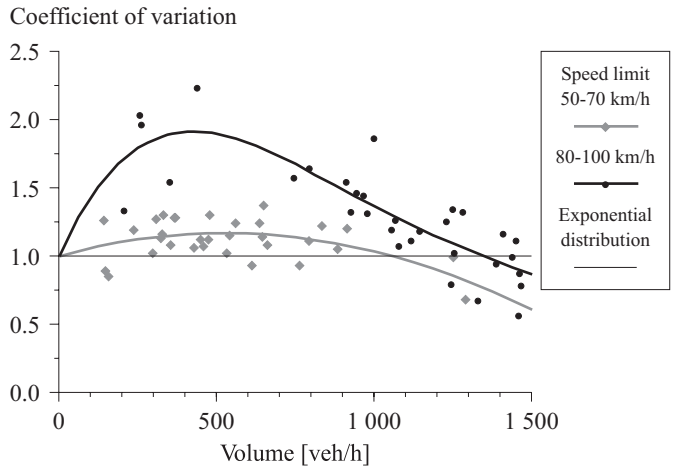

Figure 4.12: CV-chart (sample coefficient of variation) of the headway data

2. The Poisson tendency of low density traffic has a theoretical (Breiman 1963, Heidemann 1990) as well as an intuitive basis: Under low density conditions, the vehicles can move freely, and randomness of the process increases. Hence, the c.v. is expected to approach unity as the traffic flow approaches null. This phenomenon can be observed only at low speed roads, where more low volume samples are available.

3. Under medium traffic, there is a mixture of free-flowing vehicles and followers. This increases the variance above pure random process. Accordingly, the c.v. rises above unity. This is in contrast to the statement of May (1990) that the c.v. approaches unity under low flow conditions, but decreases continuously as the flow rate increases. According to Pursula & Sainio (1985) the result of May is valid for the first lane on a freeway.

4. The c.v. is larger on high speed roads than on low speed roads. In these data, the opposite flow rate was higher on high speed roads, thus reducing passing opportunities. Other explaining factors may be higher variation of speeds and greater willingness to pass on high speed roads. In addition, the intersection density was higher on the low speed roads. Consequently, there were more joining and departing vehicles, thus reducing the platooning.

To test whether it is appropriate to model the $C(T)$ on high speed and on low speed roads separately, a third degree polynomial curve was fit to the

c.v. data. The equality of the means of the residuals for high speed and low speed roads were tested using a t-test. Using residuals instead of actual c.v. values, increases the power of the test, because the effect of traffic volume can be circumvented. The equality of the means should be rejected at risk level $2.2 \cdot 10^{-6}$. Because there was a significant difference between the c.v. values on high speed and low speed roads, a separate model was construed for both.

Two polynomial curves were fitted on the data (figure 4.12): one for the high speed (80–100 km/h) roads and the other for the low speed (50–70 km/h) roads. The regression equations are:

$$\tilde{C}_L = -7.667\tilde{\lambda}^2 + 2.247\tilde{\lambda} + 1 \tag{4.24}$$

$$\tilde{C}_H = (1 - 1.8\tilde{\lambda})(1 + 20.173\tilde{\lambda} - 82.189\tilde{\lambda}^2 + 114.839\tilde{\lambda}^3). \tag{4.25}$$

The coefficient of determination for the mixed model is $R^2 = 0.685$. That is, about 69% of the variance in the c.v. can be explained by the traffic volume. A single model for all speed limits has $R^2 = 0.226$. Modeling high speed and low speed roads separately explains about 46% more of the variance.

There are some points worth consideration in the equations. The curves were forced to unity when the traffic volume was zero. This was based on the assumption of the Poisson tendency in low density traffic. The upper curve was forced to zero at volume 2,000 veh/h in order to prevent extensive curviness.

The curves have maxima at flow rates 530 veh/h (low speed roads) and 440 veh/h (high speed roads). This corresponds to mean headways 6.8 s and 8.2 s. Under lower flow rates, the long headways predominate and reduce the c.v. Traffic density is low, and there are ample opportunities for passing. Under higher flow rates, short headways predominate and reduce the c.v. Drivers are more content to follow in platoons (Pursula & Sainio 1985). Thus, the maximum is likely to be connected with the passing behavior. In Bureau of Public Roads (1950) it was shown that typically the demand for passing becomes greater than the opportunities at flow rate 500 veh/h (see also Wardrop 1952).

These observations gain at least partial support from other authors, as seen in figure 4.13, although the results are based on very different data. The data of Dunne et al. (1968) are from a two-lane rural road, but the data of Breiman et al. (1977), Buckley (1968) and May (1965) come from freeway lanes. The traffic volumes in all the samples of May are higher than in the samples of Dunne et al. (1968).

The traffic volumes in the samples of Dunne et al. (1968) are between 350

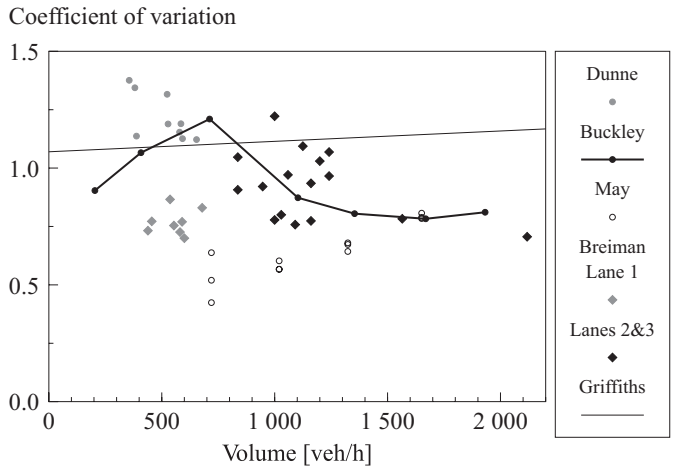

Figure 4.13: Coefficient of variation in other headway studies: Dunne et al. (1968), Buckley (1968), May (1965), Breiman et al. (1977), and Griffiths & Hunt (1991)

and 700 veh/h; i.e., the volume range that has the highest c.v. values in figure 4.12. In accordance with these results, all the samples of Dunne et al. have c.v. greater than unity. The c.v. is less than unity in the freeway samples of May (1965) and Breiman et al. (1977) (lanes one and two), and near unity in the samples of Buckley (1968) and Breiman et al. from freeway lane three. The data of Buckley represent a wide variety of volume levels, and the c.v. is similar to the results for the low speed roads in figure 4.12. The data of May were measured for gap availability studies near freeway ramps, and results are very different from the others. More general conclusions from the data of May should be drawn only with great care.

Griffiths & Hunt (1991) have studied headways on urban areas. They found a linear relation between the standard deviation and the mean headway, which can be expressed in terms of the sample c.v. and the traffic volume ($\tilde{\lambda}$, veh/h):

$$\tilde{C} = 1.07 + \frac{0.160\tilde{\lambda}}{3600}. \tag{4.26}$$

This line is shown in figure 4.13.

Finnish studies by Pursula & Sainio (1985) suggest that the c.v. on freeways, especially on the first lane, is lower than on two-lane highways with speed limit 80–100 km/h. The results of Breiman et al. (1977) also

show that lane one has lower c.v. than lane two, which has lower c.v. than lane three.

4.3.4 Skewness and kurtosis

The proportion of the first two moments was discussed above. The third and the fourth central moments, skewness (α_3) and kurtosis (α_4), give us more information about the shape of the distribution. Skewness is a measure of symmetry. For symmetric distributions $\alpha_3 = 0$. If the data are more concentrated on the low values, as in a headway distribution, skewness is positive. Kurtosis is a measure of how "heavy" the tails of a distribution are. Sample skewness and kurtosis are defined as follows:

$$\alpha_3 = \frac{\sum_{i=1}^{n}(t_i - \bar{t}\,)^3}{ns^3} \tag{4.27}$$

$$\alpha_4 = \frac{\sum_{i=1}^{n}(t_i - \bar{t}\,)^4}{ns^4}. \tag{4.28}$$

where $\bar{t}$ and s are sample mean and standard deviation, respectively.

The skewness of the headway data is displayed in figure 4.14. Analysis of covariance shows that the skewness of headway distributions on low speed and high speed roads differs significantly ($P < 0.001$). Consequently, the linear regression analysis was performed separately for the low speed and the high speed data. On low speed roads the linear model explains 36 % ($R^2 = 0.36$) of the variance, and the significance of the regression is $P < 0.001$. On high speed roads the respective values are $R^2 = 0.073$ and $P = 0.136$, indicating that the correlation is not significant, and the skewness is about 4.05, irrespective of the traffic volume.

On low speed roads, the skewness is near two at low volumes. This corresponds to the skewness of the negative exponential distribution. At higher volumes the skewness increases.

Following Ramberg, Tadikamalla, Dudewicz & Mykytka (1979), figure 4.15 shows the sample kurtosis against the squared sample skewness. (The figure is called here a KS^2-chart.) This relation is occasionally used as a guide in selecting theoretical distributions (Cochran & Cheng 1989). The relation in the KS^2-chart is almost linear, but a slightly better fit ($R^2 = 0.981$) was obtained by the power curve:

$$\alpha_4 = 2.297(\alpha_3^2)^{0.847}. \tag{4.29}$$

For comparison, figure 4.15 shows also corresponding points and curves of some theoretical distributions. The negative exponential distribution (along

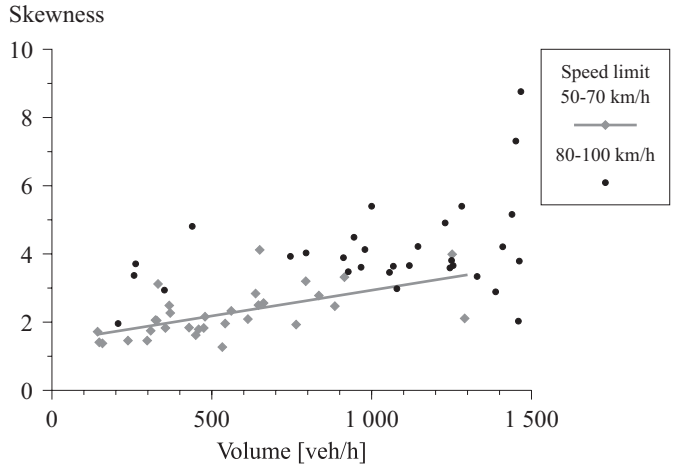

Figure 4.14: Sample skewness

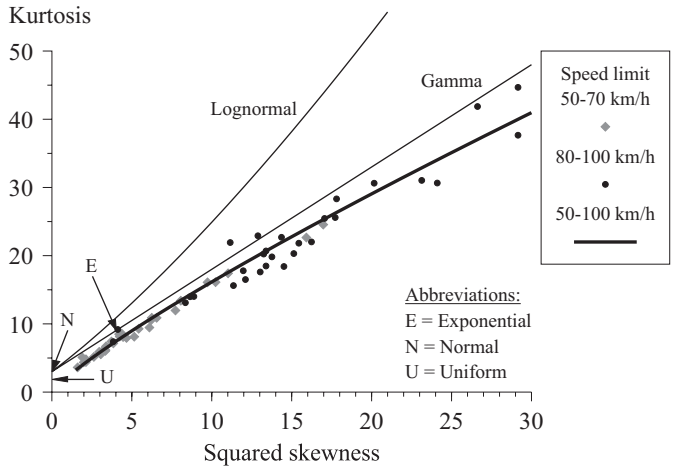

Figure 4.15: KS^2-chart (kurtosis and squared skewness) of the headway data and some theoretical distributions

with normal and uniform distributions) reduces to a single point. As the squared skewness grows the kurtosis grows more slowly than in the case of either the gamma or the log-normal distribution, but the gamma distribution is more close to the observed values. The strong correlation of the squared skewness and the kurtosis motivates a further search for the theoretical headway distribution.

4.4 Tests for the renewal hypothesis

4.4.1 Background

According to the renewal hypothesis, the headways are independent and identically distributed (i.i.d.):

$$H_0\colon\ \mathbb{P}\{\, T_i \le t \,|\, T_1,\ldots,T_{i-1} \,\} = \mathbb{P}\{T_i \le t\} = F(t)\,, \qquad i = 2,3,\ldots \tag{4.30}$$

and

$$\mathbb{P}\{T_1 \le t\} = F(t)\,. \tag{4.30a}$$

Such a stochastic process is called an ordinary renewal process. If the first headway (T_1) has a different probability distribution, the process is called a modified renewal process. In the headway data under study, the first headway in the sample is measured from an arbitrary arrival (synchronized sampling). Consequently, the first headway can be assumed to have the same distribution as the other headways.

The alternative hypothesis is that the headways are mutually correlated: If a short (long) headway has a higher probability after previous short (long) headways the data has positive autocorrelation. If a short (long) headway is more probable after long (short) headways the autocorrelation is negative. If the headways can be assumed i.i.d., the statistical analysis and modeling of vehicle headways is much less complicated.

At low flow rates, with only small differences in desired speeds and good passing opportunities, the vehicles can move freely, and correlation between consecutive headways is unlikely. At high flow rates, with large variation in desired speeds and restricted passing opportunities, platooning increases. Because the speeds of the followers are restricted, the platoons are likely to grow. A short headway is then more likely followed by another short headway. This makes the hypothesis of positive correlation plausible. On the other hand, negative autocorrelation would appear in such conditions

which favor platoons of size two (each leader has one follower). This situation could realize, when both the demand and the opportunities for passing are "reasonable". Otherwise the hypothesis of negative autocorrelation is not justifiable. Consequently, positive autocorrelation should be considered as an alternative to the renewal hypothesis.

If consecutive headways are positively correlated, short and long headways are not randomly distributed, but clustered. This can be tested by the runs test. In addition, the positive autocorrelation affects the distribution of platoon sizes. Consequently, three methods are presented to test the renewal hypothesis for vehicle headway: (1) autocorrelation analysis, (2) runs tests for randomness, and (3) tests for the distribution of the platoon length. This set of three tests was first proposed by Breiman et al. (1969). The power of these tests is now further enhanced by calculating the combined probabilities and the moving probabilities. The combined probabilities of some earlier test results are also presented.

4.4.2 Autocorrelation

The sample autocorrelation (or serial correlation) coefficient gives a measure of correlation between observations at different distances apart. For a sample of n observations the autocorrelation coefficient at lag k is:

$$\tilde{\rho}_k = \frac{\sum_{j=1}^{n-k}(t_j - \bar{t})(t_{j+k} - \bar{t})}{\sum_{j=1}^{n}(t_j - \bar{t})^2}, \tag{4.31}$$

where

$$\bar{t} = \frac{1}{n}\sum_{j=1}^{n} t_j. \tag{4.31a}$$

Note that $\tilde{\rho}_0 \equiv 1$. The sample autocorrelation coefficients are plotted in a correlogram (see figure 4.16).

In a sample of n observations the autocorrelation coefficients are asymptotically normally distributed with mean zero and standard deviation $1/n$ (Brockwell & Davis 1987). Thus, under null hypothesis, approximately 95 % of the coefficients are within bounds $\pm 1.96 n^{-1/2}$. The dashed lines in figure 4.16 show these bounds.

Autocorrelation coefficients give a rough guide, whether the observations are from a renewal process ($\rho_k = 0, k = 1,2,\dots$). The most important coefficient in this respect is ρ_1. Figure 4.17 displays $\tilde{\rho}_1$ for each headway

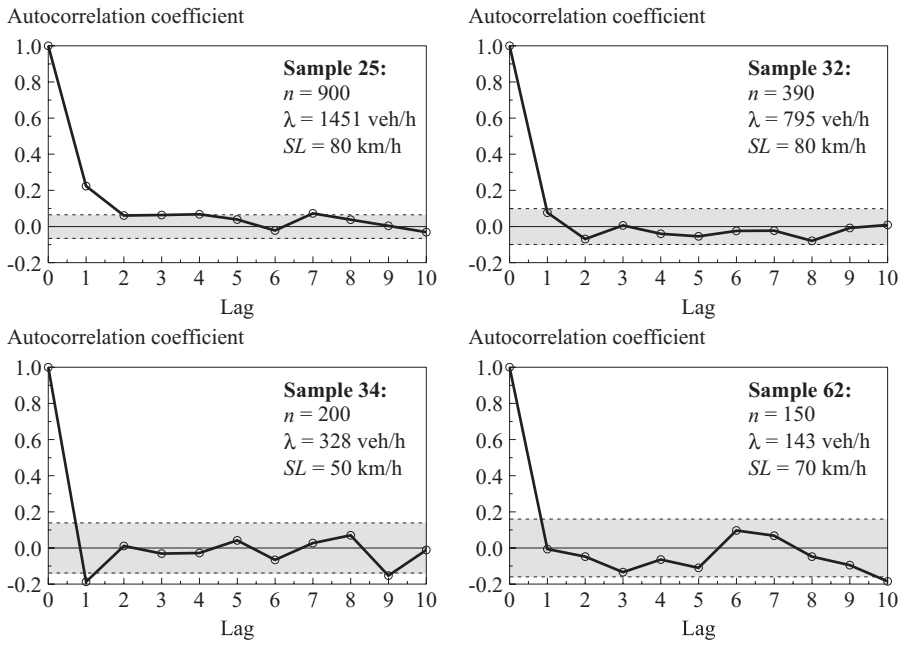

Figure 4.16: Correlograms with 0.95 confidence limits

sample. The coefficients have a large variance, but most values are positive. This indicates that there may be some small positive autocorrelation, but larger samples are needed to detect it.

Two methods have been used to make the evidence stronger: (1) The significance probabilities of $\tilde{\rho}_1$ for each sample were combined using the moving probabilities method. (2) Samples were grouped and partial correlation coefficients for autocorrelation were calculated. The methods and their results are described below.

The test for autocorrelation at lag 1 is:

$$H_0 \colon \rho_1 = 0 \tag{4.32}$$

against

$$H_A \colon \rho_1 > 0. \tag{4.33}$$

This is a one-sided test in contrast to the two sided tests normally used (Dunne et al. 1968, Breiman et al. 1969, Chrissikopoulos et al. 1982), because it is reasonable to assume that platooning causes positive autocorrelation.

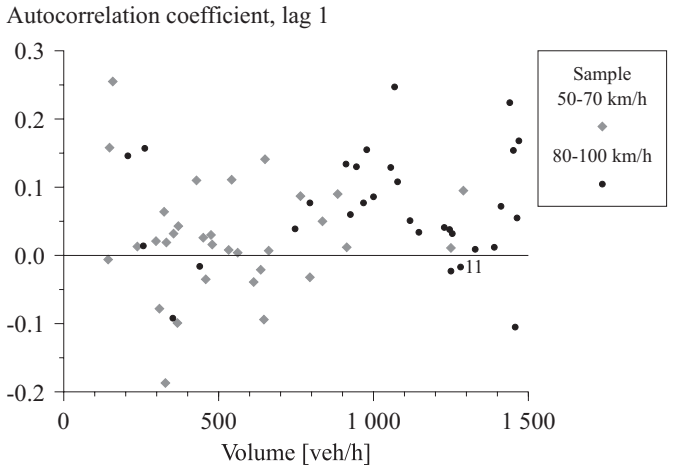

Figure 4.17: Autocorrelation coefficients at lag 1

The test for negative autocorrelation gave clearly nonsignificant values. The autocorrelation ($\tilde{\rho}_1$) is negative in some small samples only. For samples larger than 250 observations, only one negative autocorrelation was found (sample 11: $\rho_1 = -0.017$).

The variation of the significance probabilities over traffic volumes (fig. 4.18) is described by the *moving probability method* (see section 2.2.6). On high speed (80–100 km/h) roads there is significant autocorrelation between consecutive headways. There is, however, a spike in the moving probability curve, which may indicate better passing opportunities in some samples. (Similar, although lower spikes are in figures 4.19 and 4.20.) On low speed (50–70 km/h) roads there is no significant autocorrelation, at least under low volumes. The moving probability curve, however, goes down to significant values near 900 veh/h.

The *combined significance* of all high speed samples is about $3 \cdot 10^{-22}$, and for all low speed samples 0.04. Consequently, the renewal hypothesis should be rejected on high speed roads, at least when the traffic volume is above 500 veh/h. On low speed roads the possibility of autocorrelation should be considered at high volumes ($\tilde{\lambda} > 900$ veh/h).

On high speed roads the combined $\tilde{\rho}_1$ is 0.07. Since largest samples give more accurate estimates, the sample coefficients were weighted by the sample sizes. The weighted estimate is $\tilde{\rho}_1 = 0.09$, and the 95 % confidence interval is 0.06–0.12. On low speed roads the coefficient does not differ significantly from zero. Accordingly, the tests indicate a small, but statistically significant

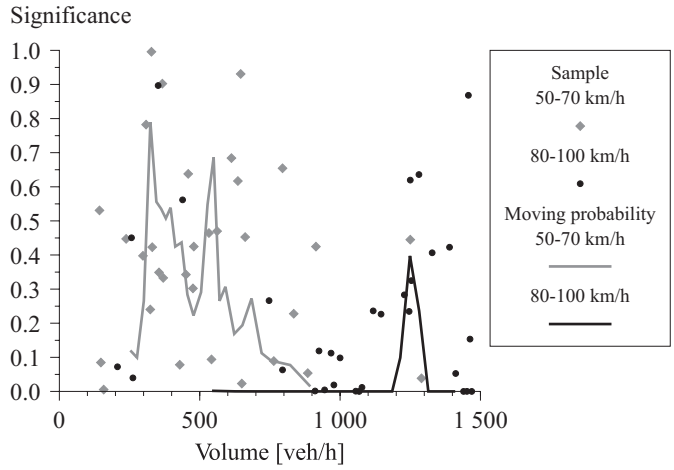

Figure 4.18: Significance of tests for renewal hypothesis against positive autocorrelation ($\rho_1 > 0$), and 9-point moving probabilities

autocorrelation at lag one on high speed roads.

The autocorrelation was also tested using the *partial correlation coefficients*. The samples of low speed and high speed roads were combined into three classes: (0, 500], (500, 1000], and (1000, 1500] veh/h. Because each sample has a different mean headway ($\bar{t}$), there is correlation between headways and their sample means $r_{t_i\bar{t}} = \mathrm{corr}(t_i, \bar{t})$. The correlation of headways and sample means increases the autocorrelation $[\tilde{\rho}_1 = \mathrm{corr}(t_i, t_{i-1})]$ in a combined sample. The effect of the mean headway can, however, be eliminated by the partial correlation coefficient (Fisher 1938):

$$\tilde{\rho}_{1\cdot\bar{t}} = \mathrm{corr}(\, t_i, t_{i-1} \mid \bar{t}\,) = \frac{\tilde{\rho}_1 - r_{t_i\bar{t}}\, r_{t_{i-1}\bar{t}}}{\left[(1 - r^2_{t_i\bar{t}})(1 - r^2_{t_{i-1}\bar{t}})\right]^{\frac{1}{2}}}. \qquad (4.34)$$

The partial correlation coefficient for autocorrelation at lag one gives the autocorrelation coefficient, when the mean headway is held constant. Because $r_{t_i\bar{t}}$ and $r_{t_{i-1}\bar{t}}$ can be assumed equal ($r_{t\bar{t}}$), the partial correlation coefficient can be calculated as:

$$\tilde{\rho}_{1\cdot\bar{t}} = \frac{\tilde{\rho}_1 - r^2_{t\bar{t}}}{1 - r^2_{t\bar{t}}}. \qquad (4.35)$$

The null hypothesis

$$H_0 \colon \rho_{1\cdot\bar{t}} = 0 \qquad (4.36)$$

can be tested using the following test statistic (Fisher 1938):

$$y = \frac{\tilde{\rho}_{1\cdot\bar{t}}}{\sqrt{1-\tilde{\rho}^2_{1\cdot\bar{t}}}}\sqrt{n-4}, \tag{4.37}$$

where n is the number of observations in the class. The statistic y follows t-distribution with $n-4$ degrees of freedom.[3] The test results are presented in table 4.1.

Table 4.1: Partial correlation coefficients ($\rho_{1\cdot\bar{t}}$)

Volume class	n	$\tilde{\rho}_{1\cdot\bar{t}}$	Significance
50–70 km/h			
0– 500 veh/h	2494	0.017	0.198
500–1000 veh/h	2832	0.025	0.092
1000–1500 veh/h	500	0.053	0.120
0–1500 veh/h	5826	0.018	0.081
80–100 km/h			
0– 500 veh/h	580	0.064	0.062
500–1000 veh/h	2361	0.093	$3.30 \cdot 10^{-6}$
1000–1500 veh/h	7650	0.092	$3.94 \cdot 10^{-16}$
0–1500 veh/h	10591	0.076	$3.24 \cdot 10^{-15}$

The partial correlation coefficients do not indicate any significant autocorrelation at lag one on low speed roads. On high speed roads the positive autocorrelation is extremely significant under volumes 500–1,500 veh/h. At lower volumes the autocorrelation is not significant, but the calculation is based on 580 observations only.

The results of the moving probability method and the partial correlation coefficients are similar. To sum up, the tests indicate that on high speed roads at volumes greater than 500 veh/h the autocorrelation coefficient at lag 1 is about 0.09. Under lower volumes and on low speed roads, there is no evidence of autocorrelation. It should be emphasized that the decisive factor may not be the speed limit, at least alone. Other road conditions, such as the intersection density and the opposite flow rate may have an impact on the autocorrelation, as well.

Some earlier results are discussed below. Combined probabilities and one-sided test results have been calculated by the present author on the

[3] The degrees of freedom are $k-3$, where k is the number of correlation pairs $\{t_i, t_{i-1}\}$, which is equal to $n-1$. This is also true for the rightmost rooted term in the equation for y.

basis of the original results. Dunne et al. (1968) studied the autocorrelation of trend-free samples from a two-lane road. The combined probability of their nine data sets (lag 1) was 0.702 (one-sided test), which is consistent with the renewal hypothesis. Two-sided test gives 0.284, which is also non-significant. Breiman et al. (1969) found in one of eight data sets (three-lane unidirectional section of John Lodge Expressway in Detroit) significant autocorrelation (lag 1) at 0.05 level. The hypothesis of independent intervals was not rejected. The combined significance is 0.23. Testing for positive autocorrelation only (one-sided test) the combined significance is 0.05, suggesting possible positive autocorrelation. In a later article Breiman et al. (1977) proposed that a small positive correlation between successive space headways may be present. Cowan (1975) studied 1,324 successive headways on a two-lane road in Australia. On the basis of the autocorrelation coefficient and a runs test, the renewal hypothesis was not rejected. Chrissikopoulos et al. (1982) studied six headway samples from U.K. trunk roads. The result of their unspecified tests was that the headways were independently distributed. However, the combined significance probability (lag 1) of their data is 0.012 using 2-sided test, and 0.0015, when testing against positive autocorrelation.

Previous studies have so far given support to the renewal hypothesis. Further analysis of this material has cast some doubt on the conclusions. Also the new data presented, shows that the possibility of a small positive autocorrelation between consecutive headways should be taken seriously. Although the autocorrelation is statistically significant, it is virtually insignificant in many applications.

4.4.3 Randomness

Randomness of the headway data was tested using the Wald & Wolfowitz (1940) runs test. It tests, whether long and short headways are randomly distributed, or if short headways are clustered. Testing runs above and below the median ($\tilde{M}d$) (Madansky 1988) is appropriate in this case (see also Breiman et al. 1969).

Let each observation t_i be associated with a variable x_i so that:

$$x_i = \begin{cases} 1, & \text{if } t_i > \tilde{M}d \\ 0, & \text{if } t_i < \tilde{M}d. \end{cases} \tag{4.38}$$

The observations equal to the median are ignored. A sequence $\{x_{j+1}, x_{j+2}, \ldots, x_{j+k}\}$ is called a run if $x_{j+1} = x_{j+2} = \cdots = x_{j+k}$ and $x_j \neq x_{j+1}$

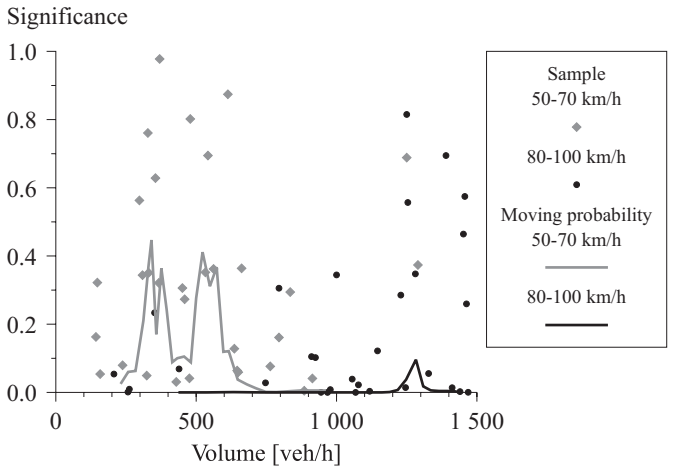

Figure 4.19: Significance probabilities of one-sided runs tests and 7-point moving probabilities

and $x_{j+k} \neq x_{j+k+1}$. The first and the last observation begin and end, respectively, a run.

The number of runs is asymptotically normally distributed with mean and variance equal to (Wald & Wolfowitz 1940):

$$\mu = \frac{2r(n-r)}{n} + 1 \tag{4.39}$$

$$\sigma^2 = \frac{2r(n-r)[2r(n-r)-n]}{n^2(n-1)}, \tag{4.40}$$

where n is the total number of observations and r is the number of observations below the median. Clustering reduces the number of runs. The number of runs is reduced also by a trend in the data. Hence, the trendlessness of the data is important.

The one-sided tests for the probability of less runs were used, in order to test if the platooning reduces the number of runs. Figure 4.19 shows the significance of the one-sided tests and the moving probabilities (section 2.2.6) for high speed and low speed roads. The combined probabilities are displayed in table 4.2. The combined probabilities for two-sided tests are presented for comparison.

On high speed roads the moving probabilities are significant almost throughout the volume range. On low speed roads the nonrandomness is significant for volumes greater than 700 veh/h. The combined significance

Table 4.2: Significance probabilities of the runs tests

Speed limit [km/h]	One-sided test	Two-sided test
50–70	0.001	0.074
80–100	0.000	0.000

on high speed roads is $1.6 \cdot 10^{-20}$ and on low speed roads 0.001. It is evident that the headways are not i.i.d., but clustered.

Breiman et al. (1969) found only one significant value (0.03) in eight runs tests. The combined probability (0.49) is also nonsignificant. This is contrary to the results presented above.

4.4.4 Platoon length

Vehicle (i) is defined to be a follower (0) if its headway is at most t_0, otherwise it is a leader (1). The status of a vehicle is accordingly defined as follows:

$$x_i = \begin{cases} 0, & \text{if } t_i \leq t_0 \\ 1, & \text{otherwise.} \end{cases} \tag{4.41}$$

The difference in speed is ignored.

The number of vehicles in a platoon is the number of consecutive gaps $t_i \leq t_0$ (followers) plus 1 (leader). The platoon has length k if $x_1 = 1$, $x_2 = 0$, $x_3 = 0, \ldots, x_k = 0$, and $x_{k+1} = 1$. The platoon may consist of a single vehicle ($k = 1$).

If the headways are i.i.d., the probability of vehicle i being a follower is:

$$p = \mathbb{P}\{t_i \leq t_0\} = F(t_0). \tag{4.42}$$

The platoon length is then geometrically distributed (Drew 1968), and the probability for length k is:

$$p_k = p^{k-1}(1-p), \qquad k = 1,2,\ldots \tag{4.43}$$

The renewal property (i.i.d. headways) can be tested using the null hypothesis:

$$H_0 \colon p_k = p^{k-1}(1-p) \tag{4.44}$$

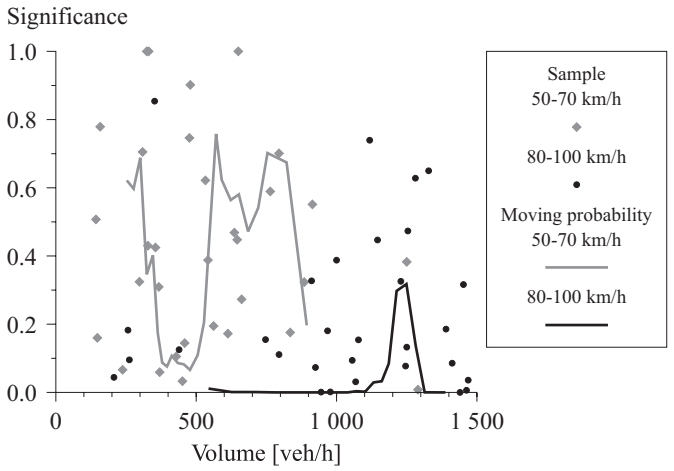

Figure 4.20: Chi-square tests for geometrically distributed platoon lengths and 9-point moving probabilities

against the alternative hypothesis:

$$H_A\colon p_k \neq p^{k-1}(1-p). \tag{4.45}$$

The tests are based on platoon criterion $t_0 = 5$ s, and the chi-square goodness-of-fit test is used. The chi-square test was performed with $k-2$ degrees of freedom. One degree of freedom was lost, because p was estimated from the sample. Platoon lengths 1, 2,..., 20 and > 20 were defined as distinct classes. They were combined so that the expectation for each class was not less than five, except for the last class the expectation was not less than unity. On high speed roads one sample was excluded from the analysis, because, after combining classes, there were no degrees of freedom left. The results of the tests are displayed in figure 4.20.

The significance probabilities of the low speed samples are rather evenly spread, and the moving probability stays above significant values. The significance probabilities of the high speed samples are more concentrated on the low values. The combined probability is 0.13 for low speed samples and $1.9 \cdot 10^{-9}$ for the high speed samples. Consequently, the hypothesis of geometric platoon length distribution should be rejected on high speed roads, at least for volumes above 500 veh/h.

In addition to the geometric distribution, several other distributions have been proposed as models for the platoon length distribution, such as the Borel-Tanner distribution (Tanner 1953, Tanner 1961) and the Miller distribution (Miller 1961) with one and two parameters. Miller (1961) found that

his one parameter model fitted the data about as well as the Borel-Tanner distribution. Galin (1980) found the Borel-Tanner and the two-parameter Miller models more accurate than the geometric distribution or the one-parameter Miller model. The headway data were from two-lane Israeli roads. The combined significance probability for the geometric distribution can be calculated as 0.0001. For comparison, the respective probability for the Borel-Tanner distribution is 0.739, for the one parameter Miller distribution 0.0007, and for the two parameter Miller distribution 0.889. These results confirm the rejection of the geometric distribution.

4.5 Discussion

After the measurements, trend tests and other preliminary analysis, 63 trendless samples consisting of 16,417 headways were analyzed. The samples were classified into two road categories according to the speed limit: the low speed roads (50–70 km/h) and the high speed roads (80–100 km/h).

The *renewal hypothesis* was tested using three methods: autocorrelation analysis, runs tests and tests for geometrically distributed platoon lengths. The significance probabilities of individual samples were combined, and the variation of significance over different flow levels was described by moving probabilities. In contrary to many previous papers, the headway data indicates that the possibility of a small positive autocorrelation should not be neglected.

A *four-stage identification process* (Luttinen 1994) is suggested to describe the properties of the headway distribution. This process combines the most powerful analyses presented above:

1. The *empirical probability density function* (edf) shows the basic shape of the headway distribution more clearly than the empirical distribution function. It should be estimated by a kernel method in order to obtain a continuous estimate, and to avoid the bias due to the histogram method.

2. The *empirical hazard function* (ehf) gives even more information about the statistical properties of headways than the edf. A graphical examination of both the edf and the ehf gives a clear indication about the shape of the headway distribution.

3. The moments of a distribution provide some numerical measures for the identification process. The sample moments may have large variation, but their relations provide proper measures for the identification.

The *coefficient of variation* is a scaled measure of the distribution spread. A CV-chart gives essential information about the headway data and proposed theoretical distributions. It is especially worth checking, whether the coefficient of variation is near, below, or above unity.

4. The *kurtosis* and the *squared skewness* of the headway samples have a close, linear or quasi-linear relation, which supports their application in the identification. The KS^2-chart shows very effectively the differences between theoretical distributions and the data.

The four-stage identification process reveals many shortcomings of the simple distributions. The exponential distribution does not pass any of the identification phases. The hazard function and the coefficient of variation reveal the problems of the gamma distribution most clearly. The deviation of the lognormal distribution from the headway data is most obvious, when the kurtosis and the squared skewness are measured. The empirical hazard function indicates that the tail of the headway distribution is exponential. Consequently, the four-stage identification process suggests mixed distributions with an exponential right tail. The process can be applied to further evaluate such models.

The identification process has demonstrated some general properties of headways. There is, however, wide variability among the samples and the road categories. High speed roads have a higher peak in the empirical density and hazard functions, a higher proportion of trailing vehicles, and a higher coefficient of variation, as well as more significant serial correlation between consecutive headways. These differences all relate to the higher platooning on high speed roads, which can be explained by many factors: The number of passing opportunities depends on the headway distribution of the opposite flow. If the passing rate is different from the catch-up rate, the flow is not in equilibrium but in a transition state (McLean 1989). The shape of the headway distribution changes, as the measurement point is moved along the road. When the passing demand exceeds the passing opportunities, it can be expected that, moving in the direction of the traffic stream on a lane, the speeds will decrease and the platooning increase (Botma 1986). On a non-uniform road, the equilibrium will not be achieved if the sub-section is shorter than the required transition length (McLean 1989). In addition, speed limit, passing sight distance, lane and shoulder width, climbing lanes, flow rate of the opposing traffic,[4] intersection density, and distance to major

[4]When the opposing volume is above 400–450 veh/h, the headway distribution does not change relevantly (Heidemann 1993).

traffic generators will have an effect on the headway distributions. Consequently, the traffic flow is not stationary in space, but each cross-section on each road has its own traffic characteristics. A headway model is necessarily a gross simplification of the reality.

The data from high speed roads have been analyzed previously by Pursula & Sainio (1985) and Pursula & Enberg (1991). Some comparisons of the results are presented below:

1. The density estimates were not calculated by Pursula & Sainio (1985), but frequency diagrams were displayed with bin width 0.5 s and the origin at zero. Because the frequencies were not corrected by the interval width, the values were lower by a factor of 0.5, as compared to the density estimates. The basic shape was, however, the same as presented here. As Pursula & Sainio (1985) observed, the mode of the headway distributions is between 1 and 2 seconds.

2. The peak heights are in the same range, when the correction due to the interval width is made. The comparison with the exponential distribution is different, because the correction is not made in figure 23 of Pursula & Sainio (1985). If the correction is made, the results are very similar.

3. The coefficients of variation have similar values. The c.v. is above unity at volumes less than about 1,500 veh/h. In the data of Pursula & Sainio (1985) the c.v. decreases steeply at volumes above 1,200 veh/h. In the present data, the decrease begins gradually and at lower volumes.

4. The platoon percentage is similar here as presented by Pursula & Sainio (1985) and Pursula & Enberg (1991). The curve in figure 4.2 is, however, in the upper part of the range in figure 24 of Pursula & Sainio (1985), especially at low volumes.

   A major reason for the differences may be in the data collection principle. Trend analysis was not used, but the headways were collected in 5 minute time slices. Moreover, the proportion of heavy vehicles and passing vehicles did not have any effect on the data collection. In the present analysis, the rejection of measure periods with more than 1 % passing vehicles may have favored samples with more vehicles in platoons and no “holes” in the flow. There may also be differences in the opposite traffic volumes.

Chapter 5

Theoretical headway models

5.1 Principles of evaluation

The evaluation of the headway distribution models is based on three considerations:

1. Reasonability. It is an advantage if the structure of the model is based on explicit theoretical reasoning about the characteristics of traffic flow. The parameters of such models can give additional information on the properties of traffic flow.

2. Applicability. In mathematical analysis the model should have a simple structure to avoid insurpassable problems, and the existence and simple form of the Laplace transform[1] is often an advantage. If simulation is considered as an alternative, the generation of pseudo-random variates should be fast and reliable. Parameter estimation should not be too complicated.

3. Validity. The model should give a good approximation of the real world phenomena; i.e., the empirical headway distributions. This is tested first by the identification process and finally by the goodness-of-fit tests.

[1] Laplace transforms, or some other transforms, are commonly used in probability theory, because they make many calculations easier. The Laplace transforms can be used to obtain the moments of a distribution (see footnote 5, page 129). When dealing with sums of random variates, the convolution property of the Laplace transform is most helpful:

$$f^*_{X+Y}(s) = f^*_X(s) f^*_Y(s).$$

(See Kleinrock 1975, Luttinen 1990).

Five theoretical distributions were selected for a more detailed analysis: the negative exponential distribution, the shifted exponential distribution, the gamma distribution, the lognormal distribution and the semi-Poisson distribution. The negative exponential distribution is the interarrival time distribution of a totally random arrival process, i.e. the Poisson process. The shifted exponential distribution avoids the problem of extremely short headways by setting a threshold for short headways. The gamma distribution is a generalization of the exponential distributions. A special case of the gamma distribution, namely the Erlangian distribution, has been widely used in the traffic flow theory. The lognormal distribution has a theoretical connection to the car-following models (Greenberg 1966). As an example of mixed models, the semi-Poisson distribution is studied. It has combined the follower headway distribution and the exponential tail most elegantly. Some other mixed models are discussed briefly.

The theoretical models are divided into two classes: the simple distributions and the mixed distributions. The *simple distributions* have the same model for all vehicles. The *mixed distributions* typically have two distinct headway distributions: one for free-moving and the other for trailing (following) vehicles.

5.2 Simple Distributions

5.2.1 Negative exponential distribution

Properties of the negative exponential distribution

The negative exponential distribution is the interarrival time distribution of the Poisson process. The Markov property[2] makes the distribution analytically simple to use. Consequently, the distribution is widely used in the theory of point processes. In traffic flow theory, the negative exponential distribution has been used since Adams (1936).

A process, which has the following four properties, is a Poisson process (Khintchine 1960):

1. The number of events at the beginning of the process is zero; i.e., $N(0) = 0$.

[2] Also called "lack of memory":

$$\mathbb{P}\{\, T > t + u \,|\, T > t \,\} = \mathbb{P}\{T > u\}\,.$$

2. The process has independent increments:

$$\mathbb{P}\{N(l_1,u_1] = n_1, \ldots, N(l_r,u_r] = n_r\} = \prod_{i=1}^{r} \mathbb{P}\{N(l_i,u_i] = n_i\}$$
$$\text{for all } l_1 < u_1 \leq l_2 < u_2 \leq \cdots \leq l_r < u_r, \quad r = 1,2,\ldots \tag{5.1}$$

3. The process has stationary increments:

$$\begin{aligned} \mathbb{P}\{N(a,b] = k\} &= \mathbb{P}\{N(a+c,b+c] = k\} \\ &= \mathbb{P}\{N(b-a) = k\} \qquad 0 \leq a < b,\ c \geq 0. \end{aligned} \tag{5.2}$$

4. At any time instant only one event can occur:

$$\mathbb{P}\{N(t,t+\Delta] \geq 2\} = o(\Delta) \qquad \text{as } \Delta \to 0. \tag{5.3}$$

The basic statistical properties of the negative exponential distribution are presented in table 5.1. Concerning the four-stage identification process, the following observations can be made:

1. The negative exponential density function (fig. 5.1) has mode at zero. The peak height is $f(0) = \lambda$, after which the density decreases exponentially. The distribution overestimates the frequency of extremely short intervals and underestimates the frequency of headways near the mode of the empirical headway distribution. Longer headways are again overestimated (fig. 5.2).

2. The hazard rate is equal to the flow rate [$h(t) = \lambda$].

3. The coefficient of variation is equal to unity.

4. The skewness and the kurtosis have constant values, independent of the scale parameter (λ).

The negative exponential distribution does not pass the four-stage identification process. In fact, it fails in all four stages.

Parameter estimation

The expectation of the distribution is the reciprocal of the scale parameter (Luttinen 1990). Accordingly, the *method of moments estimator* (MME) for the scale parameter (λ) is the reciprocal of the sample mean:

$$\tilde{\lambda} = \frac{1}{\bar{t}}. \tag{5.4}$$

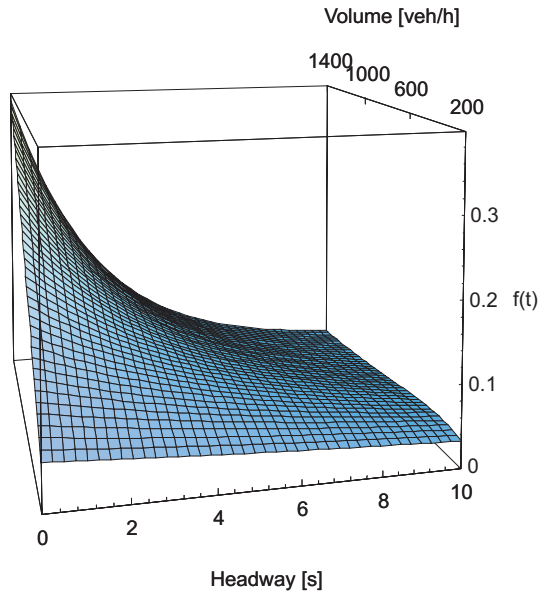

Figure 5.1: Negative exponential density function

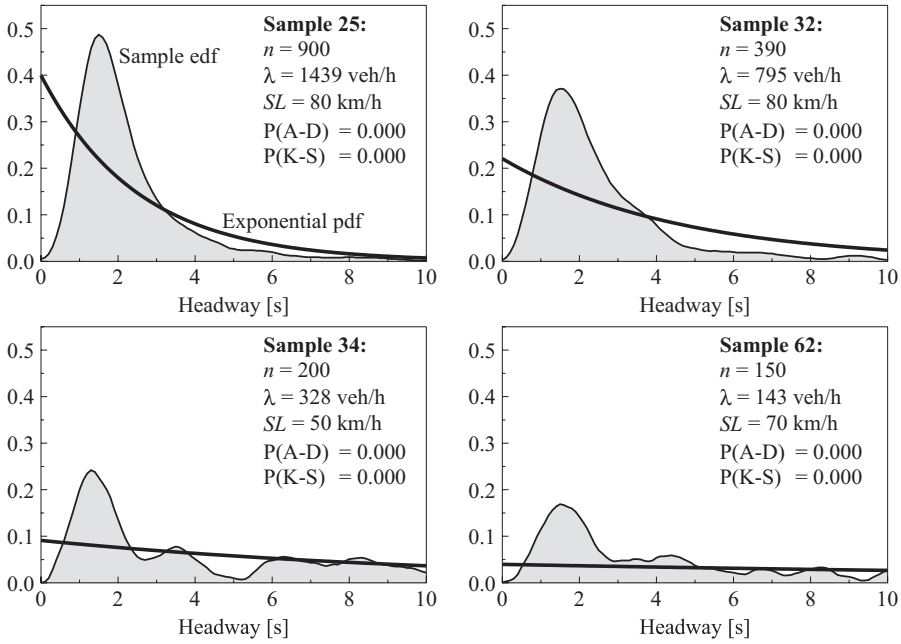

Figure 5.2: Negative exponential pdf's and sample edf's

Table 5.1: Properties of the negative exponential distribution

PDF	$F(t) = \begin{cases} 1 - e^{-\lambda t}, & \text{if } t \geq 0 \\ 0, & \text{otherwise} \end{cases}$
pdf	$f(t) = \begin{cases} \lambda e^{-\lambda t}, & \text{if } t \geq 0 \\ 0, & \text{otherwise} \end{cases}$
Hazard function	$h(t) = \begin{cases} \lambda, & \text{if } t \geq 0 \\ 0, & \text{otherwise} \end{cases}$
Laplace transform	$f^*(s) = \dfrac{\lambda}{s + \lambda}$
Mean	$\mu(T) = \dfrac{1}{\lambda}$
Median	$Md = \dfrac{1}{\lambda} \ln 2$
Mode	$Mo = 0$
Variance	$\sigma^2(T) = \dfrac{1}{\lambda^2}$
Coefficient of variation	$C(T) = 1$
Skewness	$\alpha_3(T) = 2$
Kurtosis	$\alpha_4(T) = 9$
$\lambda > 0$	

The likelihood function and the loglikelihood function are:

$$L(\lambda) = \lambda^n e^{-\lambda \sum_{j=1}^{n} t_j} \tag{5.5}$$

$$\ln L(\lambda) = n \ln \lambda - \lambda \sum_{j=1}^{n} t_j. \tag{5.6}$$

The *maximum likelihood estimator* (MLE) is obtained by differentiation

$$\frac{\partial \ln L(\hat{\lambda})}{\partial \hat{\lambda}} = \frac{n}{\hat{\lambda}} - \sum_{j=1}^{n} t_j = 0 \tag{5.7}$$

$$\hat{\lambda} = \frac{1}{\bar{t}}. \tag{5.7a}$$

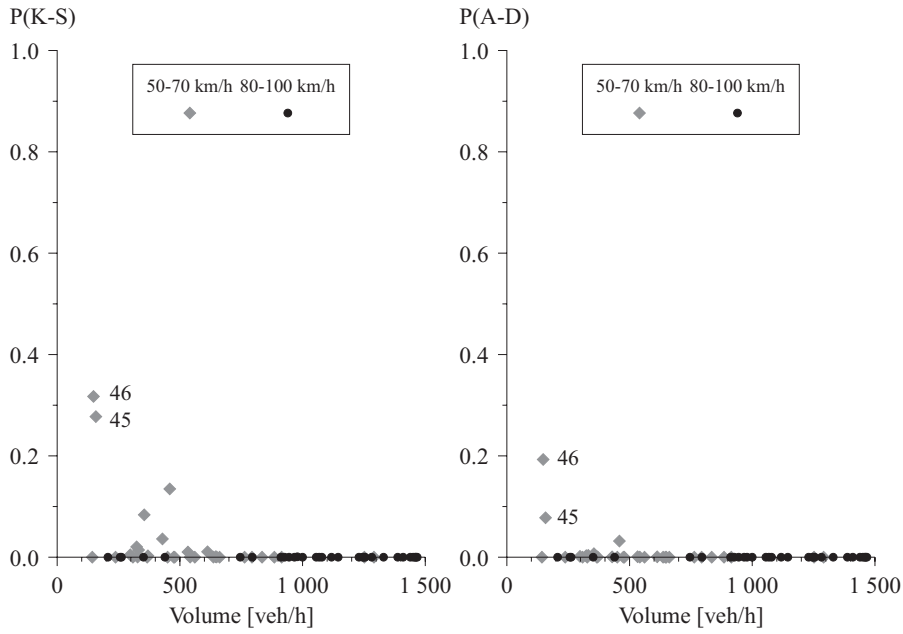

Figure 5.3: The nonparametric K-S and the parametric A-D goodness-of-fit tests for the negative exponential distribution

The MLE and the MME are identical for the negative exponential distribution. This estimator has been used in the following statistical analyses.

Goodness-of-fit tests

The goodness of fit was tested using the parametric Anderson-Darling (A-D) test. For each sample 9,999 replications were generated and analyzed. Figure 5.3 shows the test results. For comparison, the results of nonparametric Kolmogorov-Smirnov (K-S) tests are also presented.

The combined significance probability of the A-D tests is practically zero ($2.9 \cdot 10^{-150}$). Even for the three low volume (less than 160 veh/h) samples, the combined significance is 0.00016. Only two small samples (45 and 46) of no more than 100 observations yield nonsignificant results. Consequently, the negative exponential distribution should be rejected, at least for volumes greater than 140 veh/h.

Negative exponential distribution in headway studies

Reasonability The Poisson process describes situations, where vehicles without physical length can move freely. Several authors (e.g., Weiss & Herman 1962, Breiman 1963, Thedéen 1964) have shown that under rather weak assumptions (no driver interactions and i.i.d. speeds) the number of vehicles in an arbitrary interval will be asymptotically Poisson distributed as time tends to infinity. Thus, the headways will asymptotically follow the negative exponential distribution. This is called the *Poisson tendency* of low density traffic.

Some platooning nearly always occurs on two lane roads, because passing is restricted, and even in good passing conditions some drivers are content to follow behind. The Poisson process cannot deal with this phenomenon.

Applicability Because of the Markov property, the negative exponential distribution lends itself to demanding mathematical analyses. A further advantage is the simple form of the Laplace transform. Therefore, the negative exponential distribution is often preferred over other, more realistic but also more complex distributions. In many cases explicit solutions can only be found, when the arrival process is assumed Poisson.

Because the traffic volume is the only information needed, the parameter estimation is very straightforward. Also, the generation of exponential random variates is very effective. Hence, the applicability of the negative exponential distribution is excellent.

Validity The frequency of unrealistically short headways in the negative exponential distribution is too large. In fact, extremely short headways have the highest probability density. The model gets more distorted as the flow rate increases (fig. 5.2).

On the basis of the Poisson tendency theory, the exponential distribution could be used in the study of low density traffic conditions. Wattleworth (in Baerwald 1976) suggests flow rates of 500 veh/h or less. The coefficient of variation data and the goodness-of-fit tests above suggest even lower volumes (about 100 veh/h or less). There is, however, not enough low volume data to make accurate recommendations.

The negative exponential distribution has since Adams (1936) played a major role in the theoretical study of traffic flow, and especially in traffic signal control. The studies of Garwood (1940), Darroch et al. (1964), and many others rely on the exponential headway distribution. Also in some modern studies on adaptive traffic signal control, the queue length predictor

algorithm is based, at least partially, on Poisson arrivals (Baras, Levine & Lin 1979*b*, Betró, Schoen & Speranza 1987).

In conclusion, the negative exponential distribution can be considered as a model for vehicle headways under very low flow conditions and for applications that are not very sensitive to the shape of the headway distribution. In some cases the analytical solution of the problem may only be possible with a very simple model. In such situations the limitations of the model should be clearly indicated.

5.2.2 Shifted exponential distribution

Properties of the shifted exponential distribution

One possibility to avoid the extremely short headways predicted by the negative exponential distribution is, in queuing theory terminology, to add a phase with deterministic service time (τ) in series with the exponential server (Luttinen 1990). This is equivalent to shifting the negative exponential distribution τ seconds to the right. A headway is then considered as a random variate $T = \tau + X$, where τ is a nonnegative constant, and X is a random variate following the negative exponential distribution.

The density function is obtained from the negative exponential pdf by replacing t with $t - \tau$:

$$f(\, t \,|\, \tau, \theta \,) = \begin{cases} \theta e^{-\theta(t-\tau)}, & \text{if } t \geq \tau \\ 0, & \text{otherwise,} \end{cases} \tag{5.8}$$

where τ and θ are the location and scale parameters, respectively.

The basic statistical properties of the shifted exponential distribution are presented in table 5.2. Concerning the four-stage identification process, the following observations can be made:

1. The pdf (fig. 5.4) has mode at τ with peak height $f(\tau) = \theta$, after which the density decreases exponentially. Figure 5.5 shows the shifted exponential pdf estimates for four headway samples. The pdf underestimates the density near the sample mode, and overestimates the frequency of long headways.

2. The hazard rate is equal to the scale parameter; i.e., $h(t) = \theta$ for $t \geq \tau$.

3. In terms of flow rate [$\lambda = 1/\mu(T)$], the coefficient of variation is $C(T) = 1 - \lambda\tau$. Because the location parameter is always nonnegative ($\tau \geq 0$), the c.v. is less than or equal to unity.

Table 5.2: Properties of the shifted exponential distribution

PDF	$F(t) = \begin{cases} 1 - e^{-\theta(t-\tau)}, & \text{if } t \geq \tau \\ 0, & \text{otherwise} \end{cases}$
pdf	$f(t) = \begin{cases} \theta e^{-\theta(t-\tau)}, & \text{if } t \geq \tau \\ 0, & \text{otherwise} \end{cases}$
Hazard rate	$h(t) = \begin{cases} \theta, & \text{if } t \geq \tau \\ 0, & \text{otherwise} \end{cases}$
Laplace transform	$f^*(s) = e^{-s\tau} \dfrac{\theta}{s+\theta}$
Mean	$\mu(T) = \tau + \dfrac{1}{\theta}$
Median	$Md = \tau + \dfrac{\ln 2}{\theta}$
Mode	$Mo = \tau$
Variance	$\sigma^2(T) = \dfrac{1}{\theta^2}$
Coefficient of variation	$C(T) = \dfrac{1}{1+\theta\tau}$
Skewness	$\alpha_3(T) = 2$
Kurtosis	$\alpha_4(T) = 9$
$\theta > 0, \tau \geq 0$	

4. The skewness and the kurtosis have constant values, independent of the parameters.

The shifted exponential distribution does not pass the identification process. Like the negative exponential distribution, it fails in all the four stages.

Parameter estimation

Method of moments estimators The method of moments estimators (MMEs) are obtained by equating the expressions for the mean and the variance (table 5.2) to the sample mean ($\bar{t}$) and the sample variance (s^2):

$$\bar{t} = \tilde{\tau} + \frac{1}{\tilde{\theta}} \tag{5.9}$$

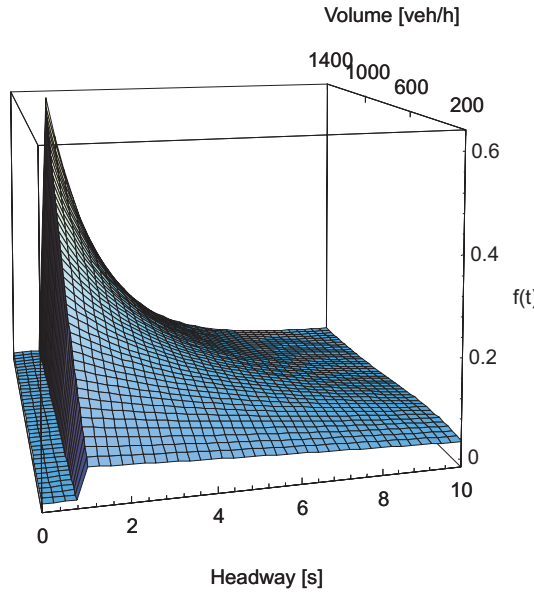

Figure 5.4: Shifted exponential density function

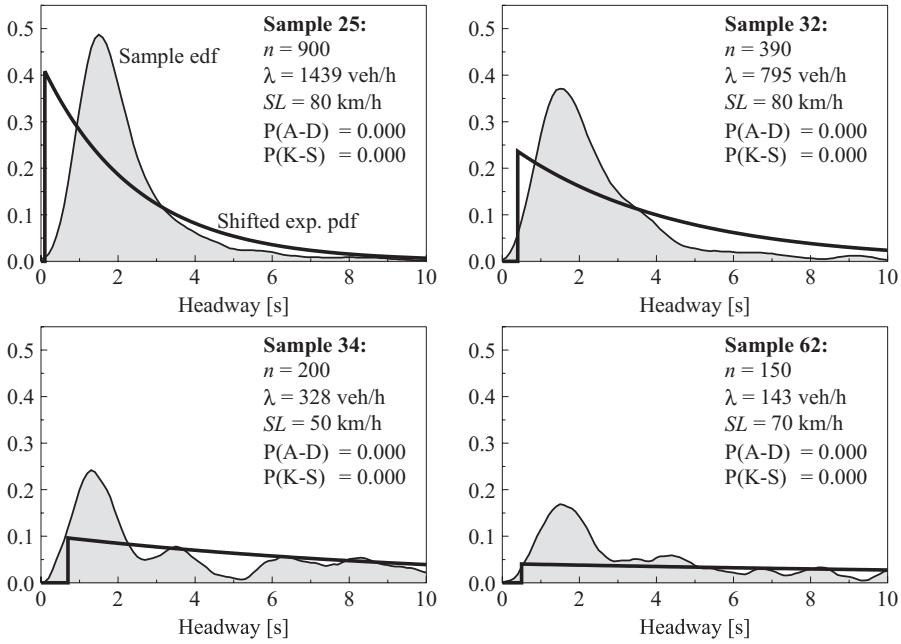

Figure 5.5: Shifted exponential pdf's and sample edf's

$$s^2 = \frac{1}{\tilde{\theta}^2}. \tag{5.10}$$

By solving this pair of equations the MMEs are obtained as:

$$\tilde{\tau} = \bar{t} - s \tag{5.11}$$

$$\tilde{\theta} = \frac{1}{s} = \frac{1}{\bar{t} - \tilde{\tau}}. \tag{5.12}$$

If the sample c.v. is greater than unity, the sample standard deviation (s) is greater than the sample mean ($\bar{t}$). The location parameter estimate is then infeasible ($\tilde{\tau} < 0$), and the best estimate, in the sense of the method of moments, is $\tau = 0$, which leads to the negative exponential distribution.

Modified method of moments estimators The first order statistic ($T_{(1)}$) in a sample of size n from a shifted exponential distribution $f(\, t \,|\, \tau, \theta \,)$ follows the shifted exponential distribution with the scale parameter equal to $n\theta$ (Cohen & Whitten 1988). Thus:

$$f\big(\, t_{(1)} \,|\, \tau, n\theta \,\big) = \begin{cases} n\theta e^{-n\theta(t_{(1)} - \tau)}, & \text{if } t_{(1)} \geq \tau \\ 0, & \text{otherwise.} \end{cases} \tag{5.13}$$

$T_{(1)}$ has expectation:

$$E(T_{(1)}) = \tau + \frac{1}{n\theta}. \tag{5.14}$$

The estimating equations of the modified method of moments estimators (MMMEs) are then:

$$E(T) = \bar{t} = \tilde{\tau} + \frac{1}{\tilde{\theta}} \tag{5.15}$$

$$E(T_{(1)}) = t_{(1)} = \tilde{\tau} + \frac{1}{n\tilde{\theta}}. \tag{5.16}$$

The MMMEs are now obtained as follows:

$$\tilde{\tau} = \frac{n t_{(1)} - \bar{t}}{n - 1} \tag{5.17}$$

$$\tilde{\theta} = \frac{1}{\bar{t} - \tilde{\tau}}. \tag{5.18}$$

According to Cohen & Whitten (1988) these are both the best linear unbiased estimators (BLUE) and the minimum variance unbiased estimators (MVUE), and therefore they are the preferred estimators.

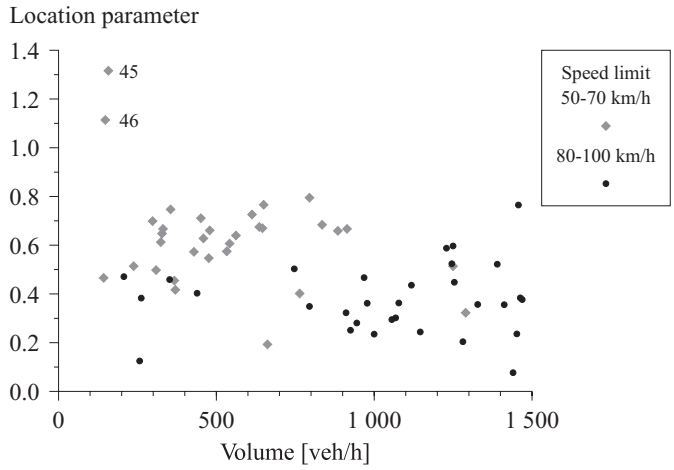

Figure 5.6: MMMEs for the location parameter of the shifted exponential distribution

Figure 5.6 shows the MMMEs for $\tilde{\tau}$. The estimates do not correlate with the volume. The mean of the estimates is 0.63 s on low speed roads and 0.38 s on high speed roads. The difference in the estimates is statistically significant ($P < 0.0005$). The two estimates (samples 45 and 46) greater than one second must be considered outliers due to small sample sizes ($n \leq 100$). If they are removed, the mean for the low speed roads is 0.59 s, which is still significantly larger than the mean for high speed roads.

Maximum likelihood estimators The likelihood function and its logarithm for the shifted exponential distribution are:

$$L(\tau, \theta) = \begin{cases} \theta^n e^{-\sum_{j=1}^{n} \theta(t_i - \tau)}, & \text{if } t_{(1)} \geq \tau \\ 0, & \text{otherwise} \end{cases} \tag{5.19}$$

$$\ln L(\tau, \theta) = \begin{cases} n \ln \theta^n - \sum_{j=1}^{n} \theta(t_i - \tau), & \text{if } t_{(1)} \geq \tau \\ -\infty, & \text{otherwise.} \end{cases} \tag{5.20}$$

For any fixed θ, the likelihood function is maximized by choosing τ as large as possible so that:

$$f(\, t_i \,|\, \tau, \theta \,) > 0 \qquad \text{for all } i = 1, \ldots, n. \tag{5.21}$$

This leads to:

$$\hat{\tau} = \min\{t_1, \ldots, t_n\} = t_{(1)}. \tag{5.22}$$

If the location parameter estimate $\hat{\tau}$ is known, the scale parameter estimate ($\hat{\theta}$) is obtained by solving (Dudewicz & Mishra 1988):

$$\frac{\partial L(\hat{\tau},\hat{\theta})}{\partial \hat{\theta}} = \frac{n}{\hat{\theta}} - \sum_{i=1}^{n}(t_i - \hat{\tau}) = 0, \tag{5.23}$$

which gives:

$$\hat{\theta} = \frac{1}{\bar{t} - \hat{\tau}}. \tag{5.24}$$

Because the probability $\mathbb{P}\{T \le t_{(1)}\} = F\Big(t_{(1)} \,\Big|\, \hat{\theta}, \hat{\tau} \Big)$ is null, the MLEs give bad goodness-of-fit test results.

Modified maximum likelihood estimators The modified maximum likelihood estimate (MMLE) for τ is calculated from the following equation:

$$F\big(t_{(1)} \,\big|\, \check{\tau}, \check{\theta} \big) = \frac{1}{n+1}. \tag{5.25}$$

The estimator for θ is the same as given by the maximum likelihood method, as well as by the other methods. The MMLEs are accordingly:

$$\check{\tau} = \frac{\ln\big(\frac{n}{n+1}\big)\bar{t} + t_{(1)}}{\ln\big(\frac{n}{n+1}\big) + 1} \tag{5.26}$$

$$\check{\theta} = \frac{1}{\bar{t} - \check{\tau}}. \tag{5.27}$$

For sample sizes near and above 100, the MMLEs are virtually equal to the MMMEs. The MMMEs were used, because they could be calculated more effectively.

Goodness-of-fit tests

The goodness of fit was tested using the parametric Anderson-Darling (A-D) test. For each sample 9,999 replications were generated and analyzed. Figure 5.7 shows the test results. For comparison, the results of nonparametric Kolmogorov-Smirnov (K-S) tests are also presented. The combined probability of the A-D tests is $5.3 \cdot 10^{-150}$. Even for the three low volume (less than 160 veh/h) samples, the combined significance is 0.002. The shifted exponential distribution does not score much better than the negative exponential distribution, and it should be rejected, at least for volumes greater than 140 veh/h.

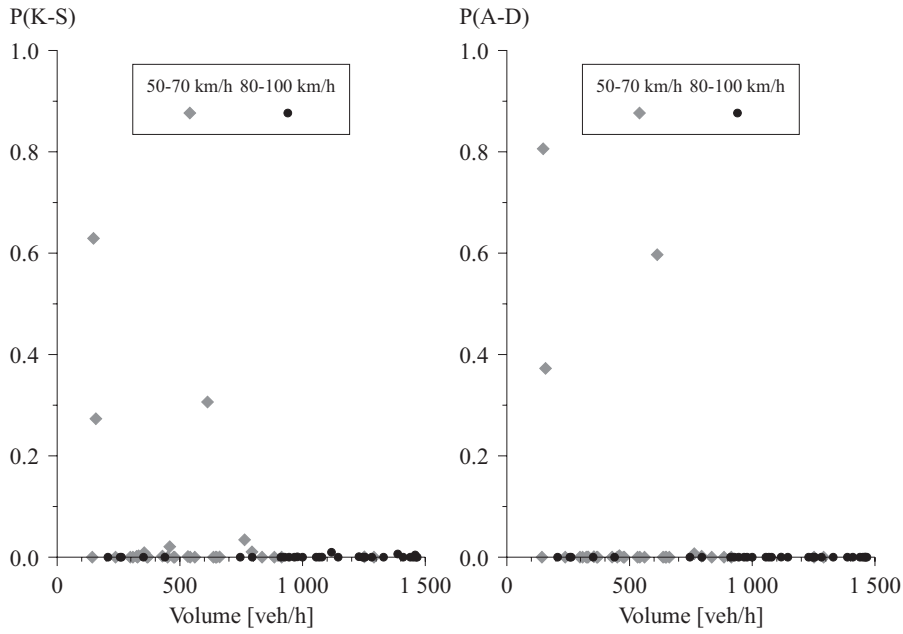

Figure 5.7: The nonparametric K-S and the parametric A-D goodness-of-fit tests for the shifted exponential distribution

Shifted exponential distribution in headway studies

Reasonability The location parameter (τ) prevents the occurrence of extremely short headways, which is a major problem of the negative exponential distribution. A fixed threshold for the minimum headway is a simplistic solution that is not in accordance with the car-following dynamics. The shifted exponential distribution is not a reasonable model for headways, but its properties are best under low flow rates.

Applicability If the exponential distribution has a positive location parameter, it loses the Markov property. The applicability, however, remains good. The Laplace transform is not complicated, the parameter estimation is straightforward, and pseudo-random variates can be generated efficiently. The shifted exponential distribution has been widely used in simulation studies (Lin 1985*b*, Lin 1985*a*, Shawaly, Ashworth & Laurence 1988).

Validity The goodness of fit is only a little better than for the negative exponential distribution. The shifted exponential distribution can be con-

sidered for very low flow conditions, or for applications that are sensitive to extremely short headways but not to other properties of the headway distribution.

In some applications, it may be better to choose a location parameter larger than $t_{(1)}$. Statistically such headway distributions are infeasible, but they may give better results, because the mode moves closer to the sample mode. The validity of the application should be judged by verifying the accuracy of the results.

5.2.3 Gamma distribution

Properties of the gamma distribution

The shifted exponential distribution avoids the problem of extremely short headways by adding a constant (τ) to an exponential random variate. In queuing theory terminology this means adding a server with constant service time in series with a server having negative exponential service time distribution (Luttinen 1990). Another approach is to put α exponential servers in series, thus reducing the probability of extremely short departure intervals (Luttinen 1990). In both cases, the service is in stages, so that only one server can be occupied at a time.

The resulting departure interval is a sum of α mutually i.i.d. exponential random variates:

$$S_\alpha = X_1 + X_2 + \cdots + X_\alpha, \qquad \alpha \in \{1,2,\ldots\}. \tag{5.28}$$

The new random variate (S_α) is said to follow the α-phase *Erlangian distribution.*

The probability distribution function of the Erlangian distribution is:

$$F(\, t \,|\, \beta, \alpha \,) = 1 - \sum_{j=0}^{\alpha-1} \frac{(\beta t)^j}{j!} e^{-\beta t}. \tag{5.29}$$

The probability density function (fig. 5.8) follows by differentiation:

$$f(\, t \,|\, \beta, \alpha \,) = \frac{(\beta t)^{\alpha-1}}{(\alpha-1)!} \beta e^{-\beta t}. \tag{5.30}$$

The negative exponential distribution is a special case ($\alpha = 1$) of the Erlangian distribution.

If the number of phases in the Erlangian distribution is allowed to be non-integer ($\alpha > 0$), the resulting generalization is called the *gamma distribution.*

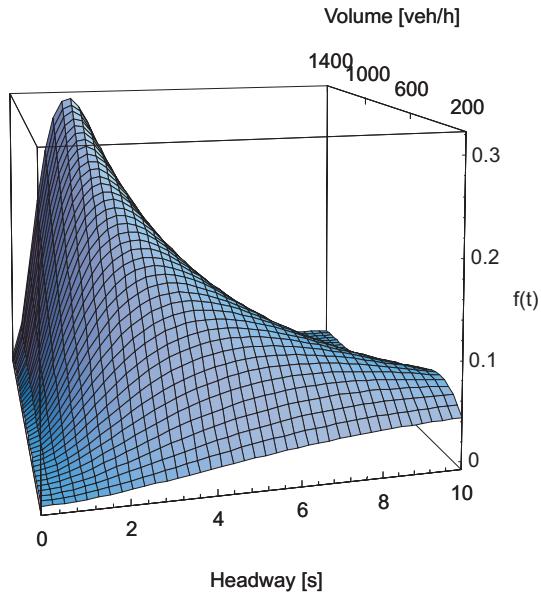

Figure 5.8: 3-phase Erlangian density function

The proportion of short headways can be further adjusted by introducing a location parameter (τ). The pdf is then:

$$f(\, t \,|\, \tau, \beta, \alpha \,) = \frac{\beta^{\alpha}(t-\tau)^{\alpha-1}}{\Gamma(\alpha)} e^{-\beta t}, \tag{5.31}$$

where

$$\Gamma(x) = \int_0^{\infty} y^{x-1} e^{y} dy \tag{5.31a}$$

is the gamma function. It is a generalization of the factorial function:

$$(x-1)! = \Gamma(x). \tag{5.32}$$

The parameters of the distribution are called the location parameter (τ), the scale parameter (β) and the shape parameter (α).

The basic statistical properties of the gamma distribution are presented in table 5.3. Concerning the four-stage identification process, the following observations can be made:

1. If the shape parameter (α) is equal to unity, pdf (fig. 5.9) has mode at τ with peak height $f(0) = \beta$ (i.e., the shifted exponential distribution). As α increases above unity, the mode approaches the mean. If α is

less than unity, the mode is at τ, and the peak height rises to infinity: $\lim_{t\to\tau+} f(t) = \infty$. Figure 5.10 shows gamma pdf estimates for four samples.

2. The hazard rate is $h(t) = \beta$ for $\alpha = 1$, monotone decreasing from infinity for $\alpha < 1$, and monotone increasing from zero for $\alpha > 1$ (see figures 5.11 and 5.12). If $\alpha \neq 1$, the hazard rate approaches β as t increases. For $t < \tau$ the hazard rate is $h(t) = 0$.

3. If $\tau = 0$, the coefficient of variation is equal to unity for $\alpha = 1$, less than unity for $\alpha > 1$ and greater than unity for $\alpha < 1$. As τ increases, the coefficient of variation decreases.

4. The skewness $[\alpha_3(T)]$ and the kurtosis $[\alpha_4(T)]$ are functions of the shape parameter (α) only. Both are positive and decrease, as the shape parameter increases. The kurtosis and the squared skewness have the following relation:

$$\alpha_4(T) = 3 + \frac{3}{2}\alpha_3^2(T). \tag{5.33}$$

For a corresponding value of squared skewness, the kurtosis is a little larger than the headway kurtosis (fig. 4.15, page 73).

The gamma distribution does not pass the four-stage identification process. The hazard function is most informative in this respect. Also, for $(\alpha > 1)$ the coefficient of variation decreases below unity and, as α increases, the distribution soon becomes too smooth and too symmetric.

Parameter estimation

Method of moments estimators The estimating equations for the MMEs are:

$$\bar{t} = \tilde{\tau} + \frac{\tilde{\alpha}}{\tilde{\beta}} \tag{5.34}$$

$$s^2 = \frac{\tilde{\alpha}}{\tilde{\beta}^2} \tag{5.35}$$

$$\alpha_3 = \frac{2}{\sqrt{\tilde{\alpha}}}. \tag{5.36}$$

Table 5.3: Properties of the gamma distribution

PDF	$F(t) = \begin{cases} \dfrac{\gamma[\alpha, \beta(t-\tau)]}{\Gamma(\alpha)}, & \text{if } t \geq \tau \\ 0, & \text{otherwise} \end{cases}$
pdf	$f(t) = \begin{cases} \dfrac{\beta^{\alpha}(t-\tau)^{\alpha-1}}{\Gamma(\alpha)} e^{-\beta(t-\tau)}, & \text{if } t \geq \tau \\ 0, & \text{otherwise} \end{cases}$
Hazard rate	$h(t) = \dfrac{f(t)}{1 - F(t)}$
Laplace transform	$f^*(s) = e^{-s\tau} \left(\dfrac{\beta}{s+\beta} \right)^{\alpha}$
Mean	$\mu(T) = \tau + \dfrac{\alpha}{\beta}$
Median	$Md = F^{-1}(0.5)$
Mode	$Mo = \begin{cases} \tau + \dfrac{\alpha - 1}{\beta}, & \text{if } \alpha > 1 \\ \tau, & \text{otherwise} \end{cases}$
Variance	$\sigma^2(T) = \dfrac{\alpha}{\beta^2}$
Coefficient of variation	$C(T) = \dfrac{\sqrt{\alpha}}{\alpha + \beta\tau}$
Skewness	$\alpha_3(T) = \dfrac{2}{\sqrt{\alpha}}$
Kurtosis	$\alpha_4(T) = 3 + \dfrac{6}{\alpha}$
$\alpha, \beta > 0, \tau \geq 0$	
$\gamma(\alpha, y) = \int_0^y x^{\alpha-1} e^{-x} dx$	

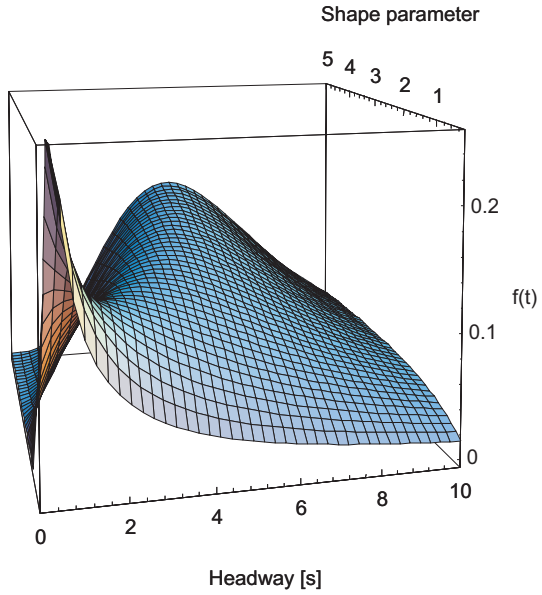

Figure 5.9: Gamma density functions ($\tau = 0$) for flow rate 600 veh/h

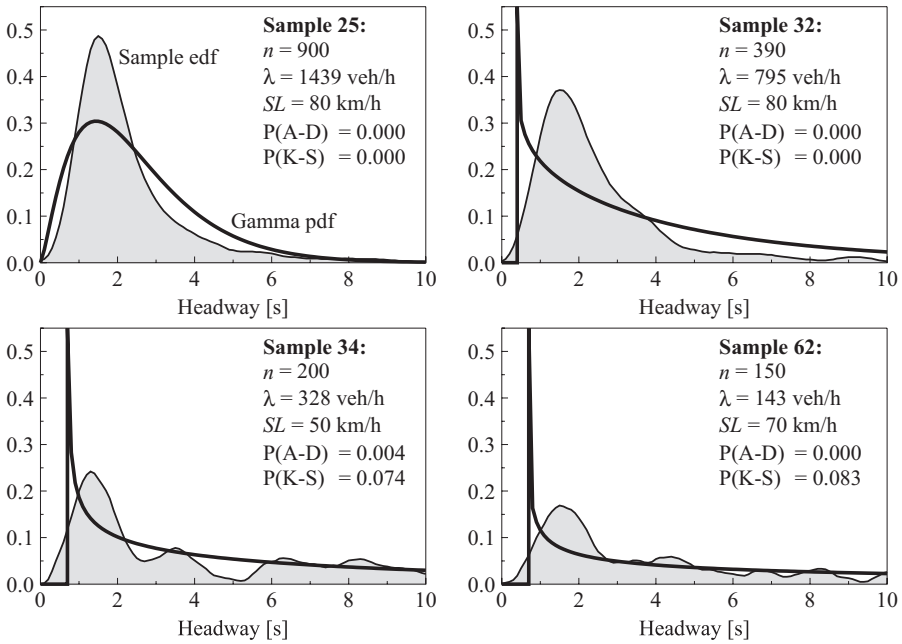

Figure 5.10: Gamma pdf's and sample edf's

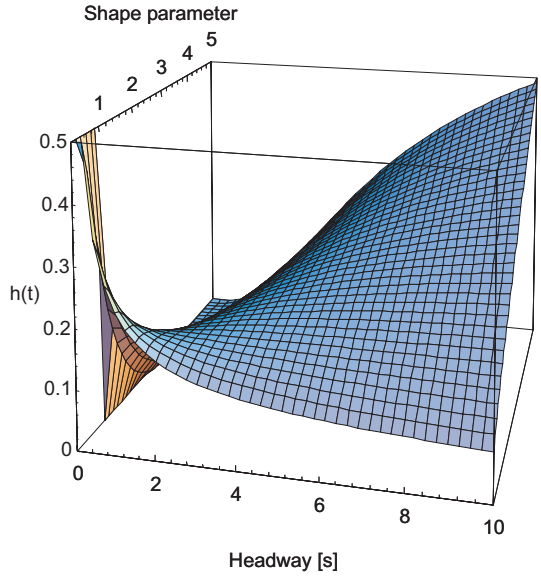

Figure 5.11: Gamma hazard functions ($\tau = 0$) for flow rate 600 veh/h

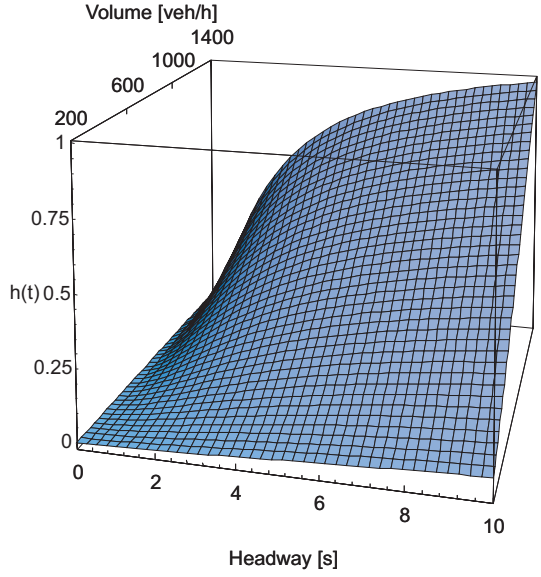

Figure 5.12: Gamma hazard functions ($\tau = 0,\ \alpha = 3$)

There is an explicit solution:

$$\tilde{\tau} = \bar{t} - \frac{2s}{\alpha_3} \tag{5.37}$$

$$\tilde{\beta} = \frac{2}{s\alpha_3} \tag{5.38}$$

$$\tilde{\alpha} = \frac{4}{\alpha_3^2}. \tag{5.39}$$

The estimate of the shape parameter (α) is directly proportional to the squared sample skewness (α_3) and independent of the location parameter. This may result in a serious fault (Bowman & Shenton 1988). The efficiency of the shape parameter estimate can be as low as 22 % (Stuart & Ord 1991). Furthermore, the location parameter estimate can be greater than the smallest sample observation $t_{(1)}$, and hence inadmissible (Cohen & Whitten 1988).

The method of moments is inefficient in parameter estimation, except for distributions closely resembling the normal (Greenwood & Durand 1960). Headway distributions are positively skew and do not resemble the normal distribution. Preliminary studies by the author have demonstrated the poor performance of the MMEs for the gamma distribution.

A MME can be used to make a crude estimate of the range of α. Figure 4.14 (page 73) shows that the sample skewness is mostly greater than two and in all cases greater than 1.25. This means that the shape parameter values are expected to be mostly below unity, and not much greater than three. This information is useful in the evaluation of other methods.

Cohen & Whitten (1988) have suggested *modified method of moments estimators* (MMMEs), because they are easy to calculate and applicable over the entire parameter space. The estimating equations are:

$$\bar{t} = \tilde{\tau} + \frac{\tilde{\alpha}}{\tilde{\beta}} \tag{5.40}$$

$$s^2 = \frac{\tilde{\alpha}}{\tilde{\beta}^2} \tag{5.41}$$

$$F\Big(t_{(1)} \,\Big|\, \tilde{\tau}, \tilde{\beta}, \tilde{\alpha} \Big) = \frac{1}{n+1}. \tag{5.42}$$

Maximum likelihood estimators

The likelihood and the log-likelihood functions of the gamma distribution are:

$$\begin{aligned} L(\tau,\beta,\alpha) &= \prod_{j=1}^{n} f(\, t_j \,|\, \tau,\beta,\alpha\,) \\ &= \left(\frac{\beta^{\alpha}}{\Gamma(\alpha)}\right)^{n} \prod_{j=1}^{n} (t_j-\tau)^{\alpha-1} e^{-\beta(t_j-\tau)} \end{aligned} \tag{5.43}$$

$$\ln L(\tau,\beta,\alpha) = n[\alpha \ln\beta - \ln\Gamma(\alpha)] + \sum_{j=1}^{n} (\alpha-1)\ln(t_j-\tau) - \beta(t_j-\tau). \tag{5.44}$$

The equations for the MLEs are obtained by equating to zero the partial derivatives and dividing by n:

$$\frac{1}{n(1-\hat{\alpha})} \frac{\partial \ln L(\hat{\tau},\hat{\beta},\hat{\alpha})}{\partial \hat{\tau}} = \frac{\hat{\beta}}{1-\hat{\alpha}} + \frac{1}{n}\sum_{j=1}^{n} \frac{1}{t_j-\hat{\tau}} = 0 \tag{5.45}$$

$$\frac{1}{n} \frac{\partial \ln L(\hat{\tau},\hat{\beta},\hat{\alpha})}{\partial \hat{\beta}} = \frac{\hat{\alpha}}{\hat{\beta}} - \frac{1}{n}\sum_{j=1}^{n} (t_j-\hat{\tau}) = 0 \tag{5.46}$$

$$\frac{1}{n} \frac{\partial \ln L(\hat{\tau},\hat{\beta},\hat{\alpha})}{\partial \hat{\alpha}} = \ln\hat{\beta} - \Psi(\hat{\alpha}) + \frac{1}{n}\sum_{j=1}^{n} \ln(t_j-\hat{\tau}) = 0, \tag{5.47}$$

where $\Psi(\cdot)$ is the digamma function:

$$\Psi(x) = \frac{d[\ln\Gamma(x)]}{dx}. \tag{5.47a}$$

The equations can be expressed in terms of modified arithmetic (A), geometric (G) and harmonic (H) means:

$$A = \frac{1}{n}\sum_{j=1}^{n} (t_j-\hat{\tau}) \tag{5.48}$$

$$G = \sqrt[n]{\prod_{j=1}^{n} (t_j-\hat{\tau})} \tag{5.49}$$

$$H = \frac{1}{n}\sum_{j=1}^{n} \frac{1}{t_j-\hat{\tau}} \tag{5.50}$$

as follows (Bowman & Shenton 1988):

$$\frac{\hat{\beta}}{\hat{\alpha}-1} = H \tag{5.51}$$

$$\frac{\hat{\alpha}}{\hat{\beta}} = A \tag{5.52}$$

$$\Psi(\hat{\alpha}) - \ln\hat{\beta} = \ln G. \tag{5.53}$$

If the shape parameter estimate $\hat{\alpha}$ is less than unity, the likelihood function (L) becomes infinite as the location parameter approaches the first order statistic of the sample:

$$\lim_{\hat{\tau}\to t_{(1)}} L(\hat{\tau},\hat{\beta},\hat{\alpha}) = \infty. \tag{5.54}$$

In this case $\hat{\tau} = t_{(1)}$, but $\hat{\beta}$ and $\hat{\alpha}$ fail to exist since (Cohen & Whitten 1982):

$$\lim_{\hat{\tau}\to t_{(1)}} \frac{\partial L(\hat{\tau},\hat{\beta},\hat{\alpha})}{\partial\hat{\alpha}} = -\infty. \tag{5.55}$$

If the shape parameter estimate $\hat{\alpha}$ is less than unity, the left hand side of equation (5.51) becomes negative, indicating that some observations should be less than $\hat{\tau}$, which is infeasible. As a conclusion, the maximum likelihood method cannot be applied for shape parameter values less than unity, which is mostly the case in the headway data.

Johnson et al. (1994) observe that the MLEs are rather unstable if $\hat{\alpha}$ is near to unity, even though it exceeds it. They recommend that the MLEs should be used only if it is expected that $\hat{\alpha}$ is at least 2.5 ($\alpha_3 < 1.27$).

Modified maximum likelihood estimators

As explained above, shape parameter estimates less than unity are to be expected in the headway data. The MLEs are then infeasible, but the maximum likelihood method can be modified (algorithm 5.1) by replacing the equation (5.51) by:

$$F\big(\, t_{(1)} \,\big|\, \check{\tau},\check{\beta},\check{\alpha} \,\big) = \frac{1}{n+1}. \tag{5.56}$$

Solving for the scale parameter estimator in equations (5.52) and (5.53), the following equations are obtained:

$$\ln(\check{\alpha}) - \Psi(\check{\alpha}) = \ln\left(\frac{A}{G}\right) \tag{5.57}$$

$$\check{\beta} = \frac{\check{\alpha}}{A}. \tag{5.58}$$

Step 1 Give two initial location parameter estimates. Set $\check{\tau}_0$ to some large value. Assign $\check{\tau}_1 = (1-\epsilon)t_{(1)}$, where ϵ is the stopping criterion (see step 2). Assign $j = 1$. Go to step 2.

Step 2 Determine shape and scale parameter estimates. The shape parameter estimate ($\check{\alpha}$) is computed using algorithm (5.59). The scale parameter estimate ($\check{\beta}_j$) is given by equation (5.58). If $|\check{\tau}_j - \check{\tau}_{j-1}|/\check{\tau}_j \leq \epsilon$, go to step 4, else go to step 3.

Step 3 Determine new location parameter estimate ($\check{\tau}_{j+1}$) from equation (5.61). If $\check{\tau}_{j+1} \geq t_{(1)}$, go to step 4, else increment j by 1 and go to step 2.

Step 4 Output parameter estimates $(\check{\tau}, \check{\beta}, \check{\alpha})$. Stop.

Algorithm 5.1: Calculation of the MMLEs for the gamma distribution (Luttinen 1991)

The shape parameter estimator $\check{\alpha}$ is solved using a rapidly converging algorithm (Bowman & Shenton 1988):

$$\check{\alpha}_j = \check{\alpha}_{j-1}[\ln(\check{\alpha}_{j-1}) - \Psi(\check{\alpha}_{j-1})]/\xi, \quad j = 1,2,\ldots, \tag{5.59}$$

where

$$\xi = \ln\left(\frac{A}{G}\right), \qquad 0 < \xi \tag{5.59a}$$

and

$$\check{\alpha}_0 = \frac{0.5000876 + 0.1648852\xi - 0.0544276\xi^2}{\xi}, \quad 0 \leq \xi < 0.5772 \tag{5.59b}$$

$$\check{\alpha}_0 = \frac{8.898919 + 9.059950\xi + 0.9775373\xi^2}{\xi(17.79728 + 11.968477\xi + \xi^2)}, \qquad 0.5772 \leq \xi. \tag{5.59c}$$

If $\check{\alpha} > 1$, then $0 < \ln(\check{\alpha}) - \Psi(\check{\alpha}) < 0.5772$ and $\xi \geq 0.5772$ leads to no solution.

The location parameter estimator can be solved by finding the $\check{\tau}$, which gives probability $1/(n+1)$ for a headway less than or equal to the shortest sample headway:

$$F^{-1}\left(\frac{1}{n+1} \,\middle|\, \check{\tau}, \check{\beta}, \check{\alpha}\right) = t_{(1)}, \tag{5.60}$$

where $F^{-1}(\cdot)$ is the inverse of the gamma PDF. Once an old estimator $\check{\tau}_j$ is known, the new estimator ($\check{\tau}_{j+1}$) can be obtained as follows:

$$\check{\tau}_{j+1} = \check{\tau}_j + t_{(1)} - F^{-1}\left(\frac{1}{n+1} \,\middle|\, \check{\tau}_j, \check{\beta}, \check{\alpha} \right). \tag{5.61}$$

Accuracy of parameter estimates

The accuracy of the parameter estimates was evaluated by a C-language computer program calling IMSL Fortran subroutines (IMSL 1989). For each combination of parameters, 1,000 replicas of 365 observations were generated. The accuracy demand for location parameter estimate was $\epsilon = 10^{-6}$. The subroutine RNGAM was used to generate the observations, and the parameters were estimated using the modified maximum likelihood method (algorithm 5.1). The accuracy of the estimates was measured by a loss function called relative root mean square error (*rrmse*) (Bai, Jakeman & Taylor 1990):

$$rrmse(\theta) = \left[\frac{1}{m} \sum_{j=1}^{m} \left(\frac{\check{\theta}_j - \theta}{\theta} \right)^2 \right]^{\frac{1}{2}}. \tag{5.62}$$

where m is the number of replications, θ is the true value of the parameter, and $\check{\theta}_j$ is the parameter estimate for replica j.

Table 5.4 presents the *rrmse*'s for different values of location (τ) and shape (α) parameters. The scale parameter (β) is set to unity to make the table comparable with the data presented by Bai et al. (1990).

Some accuracy is lost as compared to the maximum likelihood method, especially when the shape parameter (α) is large. The restriction of $\alpha < 1$ is, however, avoided. The accuracy is quite good for $\alpha \leq 1$. The loss of accuracy in the location parameter estimate ($\check{\tau}$) with large values of the shape parameter is unavoidable. Consequently, the accuracy of the MMLEs is considered appropriate for the shape parameter values near or below unity.

Bowman & Shenton (1988) consider the MMLEs particularly successful for $1/4 \leq \alpha < 1$, and quite acceptable for $1 \leq \alpha < 2$. (1/4 was the lowest shape parameter value that they tested.) In the goodness-of-fit tests below, the parameters were estimated by the modified maximum likelihood method. This method uses all the information in the samples, and it is not as sensitive to outliers as the moment methods (Bowman & Shenton 1988).

Figures 5.13, 5.14 and 5.15 show the parameter estimates for the gamma distribution. The location parameter estimates are mostly between 0.2 and

Table 5.4: Relative root mean square error (*rrmse*) of MMLEs for the gamma distribution. The *rrmse*'s of MLEs in (Bai et al. 1990) are shown in parentheses.

α	θ	$\tau = 0.2$	$\tau = 0.6$	$\tau = 1.0$
	α	0.063	0.061	0.062
$\alpha = 0.50$	β	0.098	0.094	0.094
	τ	0.000	0.000	0.000
	α	0.065	0.070	0.066
$\alpha = 0.75$	β	0.089	0.091	0.085
	τ	0.002	0.001	0.001
	α	0.073	0.073	0.075
$\alpha = 1.00$	β	0.086	0.085	0.088
	τ	0.015	0.005	0.003
	α	0.133	0.123 (0.11)	0.136
$\alpha = 2.00$	β	0.102	0.099 (0.10)	0.106
	τ	0.299	0.094 (0.05)	0.061
	α	0.156	0.230 (0.27)	0.312
$\alpha = 6.00$	β	0.124	0.144 (0.16)	0.170
	τ	2.011	0.944 (0.65)	0.751

0.8. The scale parameter rises exponentially as the volume increases. The shape parameter estimates are concentrated between 0.4 and 3.1. On high speed roads the shape parameter is larger at high volumes. On low speed roads the correlation is not significant.

Goodness-of-fit tests

For each sample the parameters were estimated using the modified maximum likelihood method. The goodness of fit was tested using Monte Carlo tests with the A-D statistic (section 2.2.4). For each sample 9,999 replications were generated and analyzed. Figure 5.16 shows the test results. For comparison, the results of nonparametric K-S tests are also presented.

As figure 5.16 shows, the goodness of fit is poor. The points are not evenly distributed between 0 and 1 at any flow level. The combined significance of the A-D tests is $4.1 \cdot 10^{-141}$. For volumes less than 160 veh/h the K-S tests give combined significance probability of 0.25. The corresponding probability of the A-D tests is 0.001.

In figure 5.17, the nonparametric K-S tests are compared with the para-

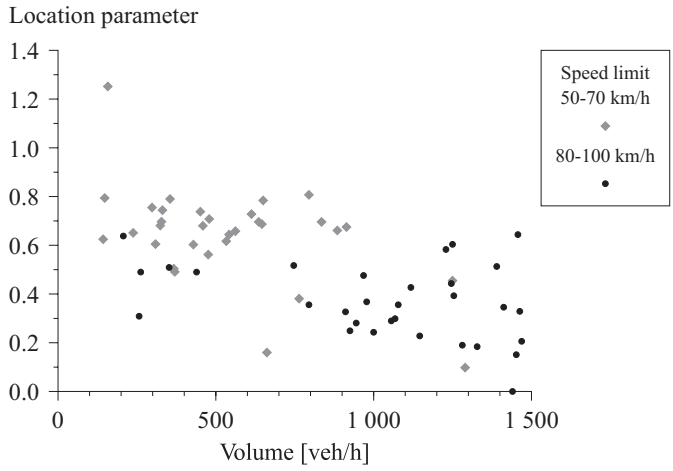

Figure 5.13: Location parameter estimates (MMLE) for the gamma distribution

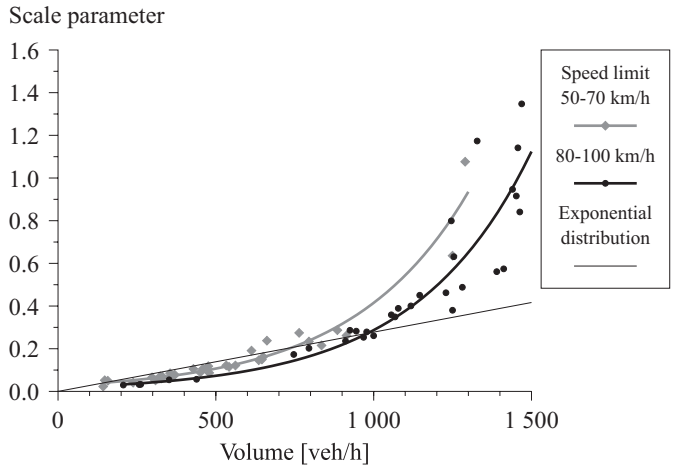

Figure 5.14: Scale parameter estimates (MMLE) for the gamma distribution

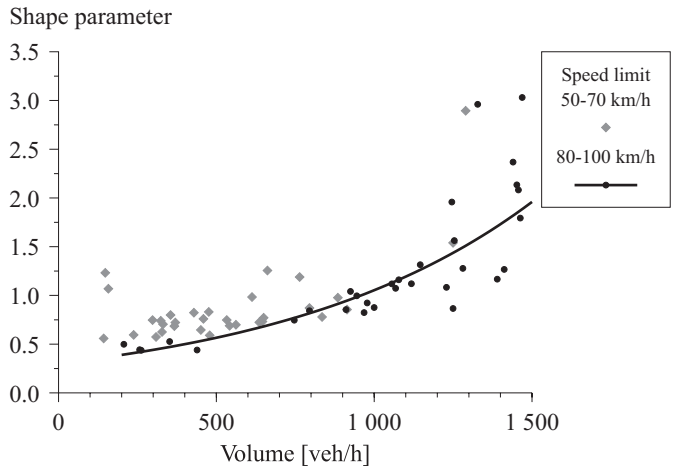

Figure 5.15: Shape parameter estimates (MMLE) for the gamma distribution

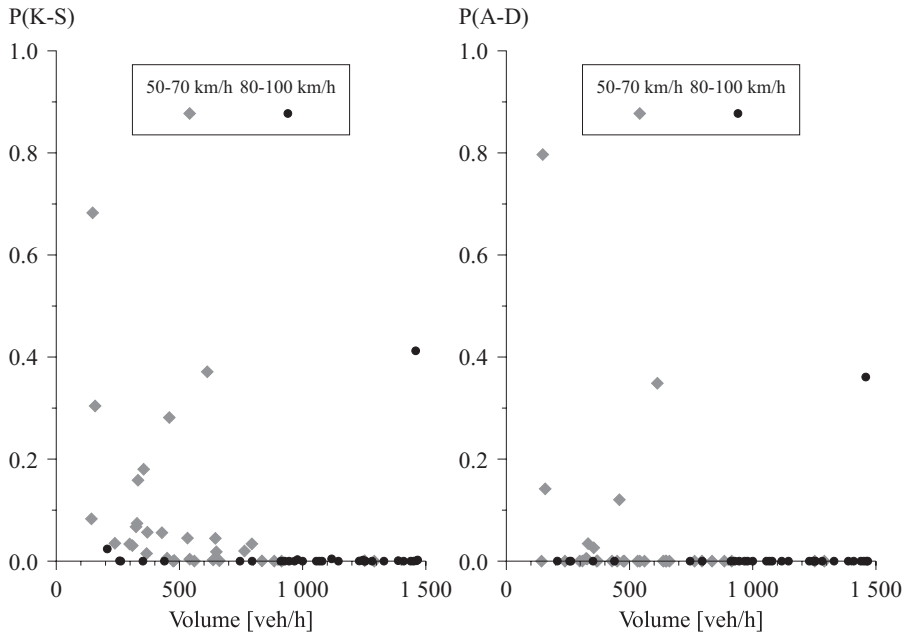

Figure 5.16: The nonparametric K-S and the parametric A-D goodness-of-fit tests for the gamma distribution

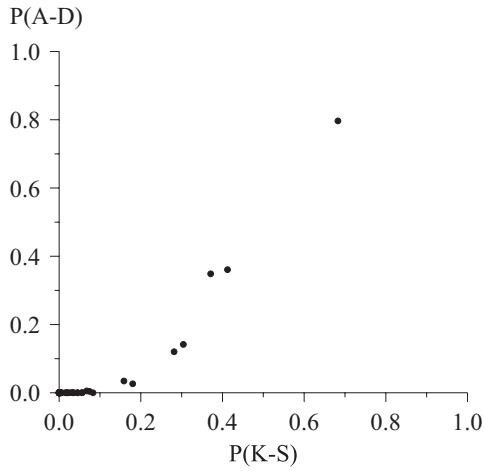

Figure 5.17: Comparison of the nonparametric K-S and the parametric A-D tests for the gamma distribution

metric A-D tests. The figure confirms the statement that the nonparametric K-S test gives too conservative results (section 2.2.4). The 5 % significance level of the parametric A-D test is about the same as the 20 % significance level of the nonparametric K-S test. A similar result was obtained by Lilliefors (1969) for the negative exponential distribution. Since the shape parameter estimates are concentrated near unity (mean 1.09, standard deviation 0.62), the similarity of the results is to be expected.

Gamma distribution in headway studies

Reasonability The gamma distribution with $\alpha > 1$ has a bell-like shape, which gives low probability to extremely short headways. The coefficient of variation, however, decreases below unity. Consequently, the the shape parameter α is estimated below unity in most samples. In these cases, the problem with extremely short headways is even more critical than with the negative exponential distribution. The location parameter can be used to ease the problem, but then the original motivation for the application of the gamma distribution is lost.

Figure 5.18 shows the skewness of the headway data and estimated gamma distributions. Although figure 4.15 indicated that the relation between the skewness and the kurtosis of the gamma distribution does not differ much from the headway data, figure 5.18 shows that a gamma distribution cannot model very skew ($\alpha_3 > 3$) headway distributions. In fact,

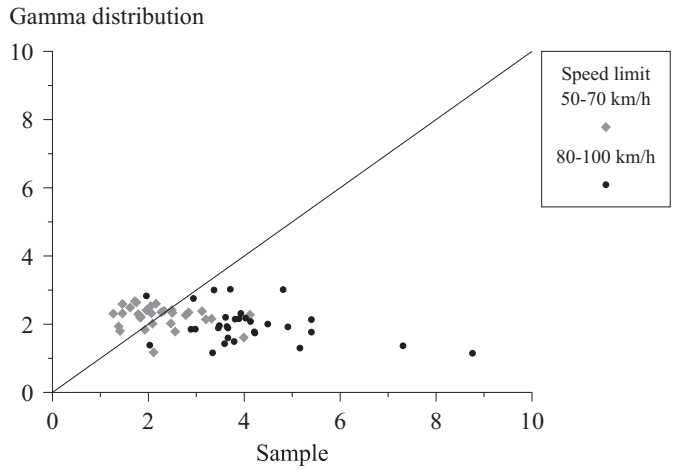

Figure 5.18: Skewness of the headway data and fitted gamma distributions

the skewness of the fitted gamma distributions does not correlate positively with the skewness of the empirical headway distributions. The MMEs would give different results, but the fact remains that the properties of the gamma distribution are very different from the properties of the headway distributions.

As $\alpha = 1$ the gamma distribution is equal to the negative exponential distribution ($\tau = 0$) or the shifted exponential distribution ($\tau > 0$). The discussion on these distributions is relevant here also. The gamma distribution can therefore be justified as a model for low traffic flow conditions.

When the flow rate is high, most vehicles are followers, and the headways are concentrated near the mode. The bell-like shape of the corresponding headway distribution is typical of the gamma distribution with large α. However, as the the flow rate increases, the headway distribution becomes more skew (fig. 4.14, page 73), whereas the skewness of the gamma distribution decreases for large α.

Applicability The gamma distribution is rather simple in mathematical analysis, especially when dealing with transforms, or if the analysis is restricted to the Erlangian distribution. Efficient subroutines can be found to generate gamma variates.

Parameter estimation of the gamma distribution presents some problems, but in many practical situations reasonable results can be obtained by letting $\hat{\tau} = t_{(1)} - \epsilon$ (where ϵ is some small value), and estimating α and β by the

method of moments.

Validity For shape parameter values greater than unity, the density function has the skewed bell-like shape typical of headway distributions. Goodness-of-fit tests do not, however, support the use of the gamma distribution in traffic flow studies, except perhaps at very low and very high flow rates. There is not, however, enough data to make accurate recommendations. At low flow conditions, the exponential distribution is usually more simple, and both theoretically and statistically justifiable. At higher volumes the gamma distribution can be given preference over the exponential distribution.

In the tradeoff between applicability and validity, the gamma distribution offers a little more credibility and a little more hard work than the negative exponential distribution. The pdf of the gamma distribution is appealing, when the shape parameter is greater than unity, but it fails to model both the strong accumulation of headways near the mode and the skewness of headway distributions. If the gamma distribution is used as a model for vehicle headways, the process studied should not be very sensitive to the shape of the headway distribution.

5.2.4 Lognormal distribution

Properties of the lognormal distribution

If a variate $Y = \ln(T - \tau)$ follows the normal distribution $N(\mu, \sigma^2)$, then T follows the lognormal distribution:

$$\begin{aligned} f(\, t \,|\, \tau, \mu, \sigma \,) &= \frac{1}{t-\tau}\phi\left(\frac{\ln(t-\tau)-\mu}{\sigma}\right) \\ &= \frac{1}{\sigma(t-\tau)\sqrt{2\pi}} \exp\left(-\frac{[\ln(t-\tau)-\mu]^2}{2\sigma^2}\right), \qquad t > \tau \end{aligned} \tag{5.63}$$

The lognormal relation holds if the change in a headway during a small time interval is a random proportion of the headway at the start of the interval, and the mean and the variance of the headway remain constant over time. Greenberg (1966) has shown the resulting connection between the lognormal headway distribution and car-following theory.

The most important properties of the lognormal distribution are presented in table 5.5. Concerning the four-stage identification process, the following observations can be made:

Table 5.5: Properties of the lognormal distribution

PDF	$F(t) = \begin{cases} \Phi\left(\frac{\ln(t-\tau)-\mu}{\sigma}\right), & \text{if } t \geq \tau \\ 0, & \text{otherwise} \end{cases}$
pdf	$f(t) = \begin{cases} \frac{1}{t-\tau}\phi\left(\frac{\ln(t-\tau)-\mu}{\sigma}\right), & \text{if } t \geq \tau \\ 0, & \text{otherwise} \end{cases}$
Hazard rate	$h(t) = \frac{f(t)}{1-F(t)}$
Laplace transform	—
Mean	$\mu(T) = \tau + e^{\mu+\frac{1}{2}\sigma^2}$
Median	$Md(T) = \tau + e^{\mu}$
Mode	$Mo(T) = \tau + e^{\mu-\sigma^2}$
Variance	$\sigma^2(T) = e^{2\mu+\sigma^2}\left(e^{\sigma^2}-1\right)$
Coefficient of variation	$C(T) = \frac{\sqrt{e^{s^2}-1}}{1+\tau e^{-\mu-\frac{1}{2}s^2}}$
Skewness	$\alpha_3(T) = (\omega+2)\sqrt{\omega-1}$
Kurtosis	$\alpha_4(T) = \omega^4 + 2\omega^3 + 3\omega^2 - 3$
$\mu, \sigma > 0, \tau \geq 0$	
$\omega = e^{\sigma^2}$	

1. The lognormal pdf (fig. 5.19) has a shape similar to the empirical headway density function, especially at high flow rates. As the coefficient of variation decreases, the mode moves towards the mean (fig. 5.20). A closer examination of the pdf (fig. 5.21), however, shows that the decrease of the density after the peak is too slow.

2. The lognormal hazard function (figs. 5.22 and 5.23) is zero at $t = 0$, increases to maximum, and decreases gradually, approaching zero as the headway becomes large (Kalbfleisch & Prentice 1980). The hazard function is similar to the empirical hazard function, but the fall after the peak is more gradual.

3. The coefficient of variation does not present any limitations for the lognormal distribution. It $\tau = 0$, the c.v. is:

$$C(T) = \sqrt{e^{s^2} - 1}. \tag{5.64}$$

4. For a corresponding value of the squared skewness, the kurtosis of the lognormal distribution is too large. See figure 4.15 on page 73.

The lognormal distribution differs most from the headway data, when the kurtosis and the squared skewness are measured. In addition, the pdf and the hazard function, in particular, fall too gradually after the peak. The lognormal distribution fails in three stages of the identification process. Stages 2 and 4 are most clear in this respect.

Parameter estimation

Maximum likelihood estimators The likelihood function of the lognormal distribution is:

$$L(\tau, \mu, \sigma) = \frac{\exp\Big(-\sum_{i=1}^{n} \frac{[\ln(t_i - \tau) - \mu]^2}{2\sigma^2}\Big)}{\Big(\sigma\sqrt{2\pi}\Big)^n \prod_{j=1}^{n}(t_j - \tau)}. \tag{5.65}$$

This function has global maximum at $\hat{\tau} = t_{(1)}, \hat{\mu} = -\infty, \hat{\sigma}^2 = \infty$ regardless of the sample (Crow & Shimizu 1988). Consequently, the maximum likelihood method is not feasible.

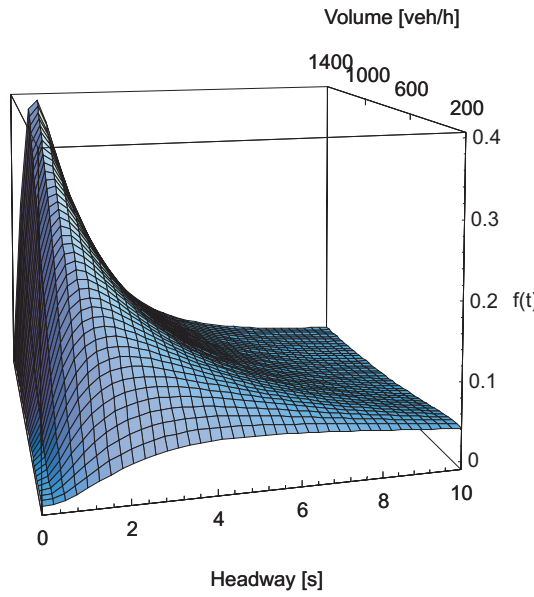

Figure 5.19: Lognormal density function $[C(T) = 1.2]$

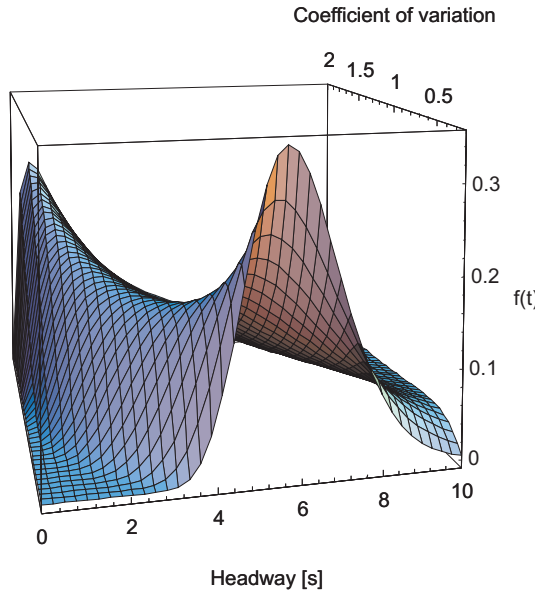

Figure 5.20: Lognormal density function for flow rate 600 veh/h

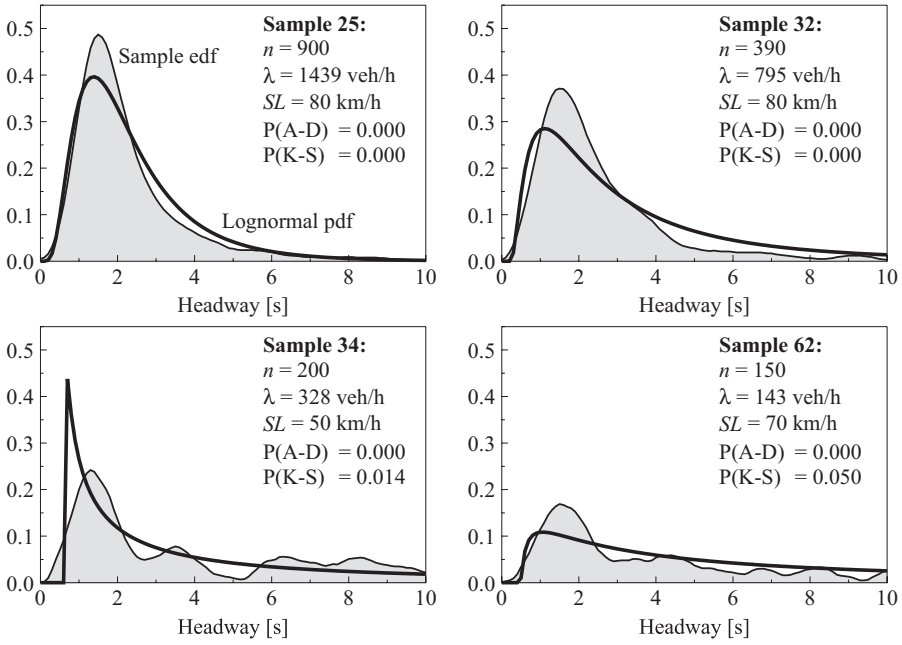

Figure 5.21: Lognormal pdf's and sample edf's

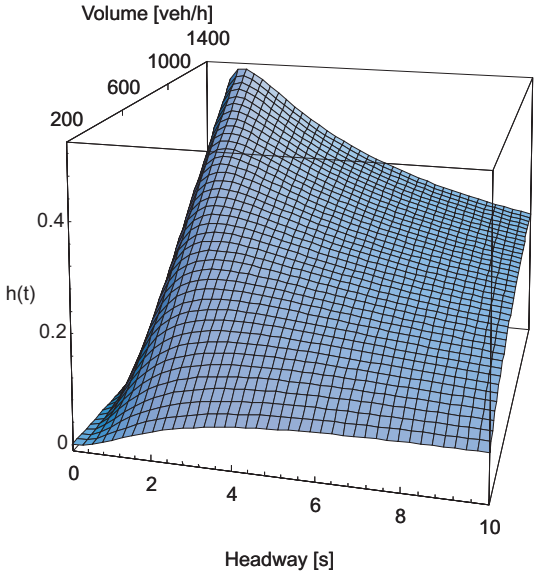

Figure 5.22: Lognormal hazard function $[C(T) = 1.2]$

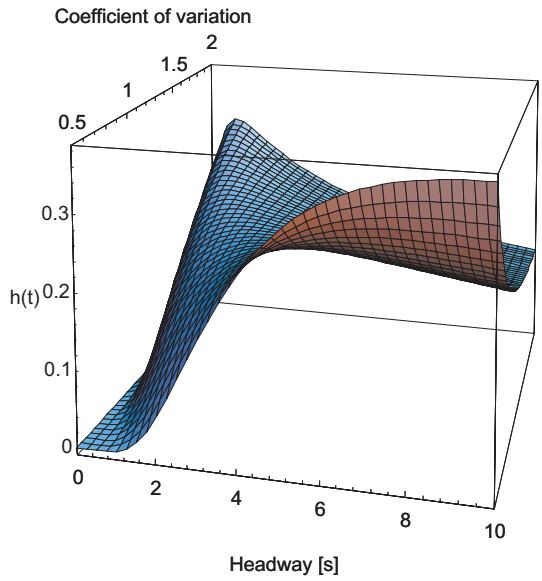

Figure 5.23: Lognormal hazard function for flow rate 600 veh/h

Local maximum likelihood estimators Local maximum likelihood estimating equations for the lognormal distribution are obtained by differentiating the loglikelihood function (Cohen & Whitten 1980):

$$\frac{\partial \ln L(\hat{\tau},\hat{\mu},\hat{\sigma})}{\partial \hat{\tau}} = \frac{1}{\hat{\sigma}^2}\sum_{j=1}^{n}\frac{\ln(t_j-\hat{\tau})-\hat{\mu}}{t_j-\hat{\tau}} + \sum_{j=1}^{n}\frac{1}{t_j-\hat{\tau}} = 0 \tag{5.66}$$

$$\frac{\partial \ln L(\hat{\tau},\hat{\mu},\hat{\sigma})}{\partial \hat{\mu}} = \frac{1}{\hat{\sigma}^2}\sum_{j=1}^{n}[\ln(t_j-\hat{\tau})-\hat{\mu}] = 0 \tag{5.67}$$

$$\frac{\partial \ln L(\hat{\tau},\hat{\mu},\hat{\sigma})}{\partial \hat{\sigma}} = \frac{n}{\hat{\sigma}} + \frac{1}{\hat{\sigma}^3}\sum_{j=1}^{n}[\ln(t_j-\hat{\tau})-\hat{\mu}]^2 = 0. \tag{5.68}$$

By eliminating $\hat{\mu}$ and $\hat{\sigma}$ from these equations, the equation for $\hat{\tau}$ is obtained:

$$\left(\sum_{j=1}^{n}\frac{1}{t_j-\hat{\tau}}\right)\left[\sum_{j=1}^{n}\ln(t_j-\hat{\tau}) - \sum_{j=1}^{n}\ln^2(t_j-\hat{\tau}) + \frac{1}{n}\left(\sum_{j=1}^{n}\ln(t_j-\hat{\tau})\right)^2\right] - n\sum_{j=1}^{n}\frac{\ln(t_j-\hat{\tau})}{t_j-\hat{\tau}} = 0. \tag{5.69}$$

The local maximum likelihood estimate (LMLE) for $\hat{\tau}$ can now be calculated iteratively. Only values $\hat{\tau} < t_{(1)}$ are accepted.

After $\hat{\tau}$ has been solved, the scale and shape[3] parameters can be estimated as follows (Cohen & Whitten 1980):

$$\hat{\mu} = \frac{1}{n}\sum_{j=1}^{n} \ln(t_j - \hat{\tau}) \tag{5.70}$$

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{j=1}^{n} \ln^2(t_j - \hat{\tau}) - \left(\frac{1}{n}\sum_{j=1}^{n} \ln(t_j - \hat{\tau})\right)^2. \tag{5.71}$$

Although the LMLEs gave the best estimates of all the methods examined, the method had convergence problems—and a series of Monte Carlo tests with 9,999 replications of 63 samples is very sensitive to convergence problems. As a result, the local maximum likelihood method was rejected.

Modified maximum likelihood estimators The local maximum likelihood method can be modified as explained on page 34. Equation (5.66) is replaced by:

$$F^{-1}\left(\frac{1}{n+1}\,\middle|\,\check{\tau}, \check{\mu}, \check{\sigma}\right) = t_{(1)}. \tag{5.72}$$

This can be expressed in terms of the inverse normal PDF:

$$\Phi^{-1}\left(\frac{1}{n+1}\right) = \frac{\ln(t_{(1)} - \check{\tau}) - \check{\mu}}{\check{\sigma}}. \tag{5.73}$$

Now the set of equations consists of equations (5.67) and (5.68), and the above equation. By eliminating $\check{\mu}$ and $\check{\sigma}$ from these equations, the estimating equation for $\check{\tau}$ is obtained (Cohen & Whitten 1980):

$$\ln(t_{(1)} - \check{\tau}) - \frac{1}{n}\sum_{j=1}^{n} \ln(t_j - \check{\tau}) - \Phi^{-1}\left(\frac{1}{n+1}\right)\left[\frac{1}{n}\sum_{j=1}^{n} \ln^2(t_j - \check{\tau}) - \left(\frac{1}{n}\sum_{j=1}^{n} \ln(t_j - \check{\tau})\right)^2\right]^{\frac{1}{2}} = 0. \tag{5.74}$$

This equation can be solved iteratively. The scale and shape parameter estimates follow from (5.70) and (5.71).

The moment estimators for the lognormal distribution have a large sampling error, and they are not uniquely determined. The modified moment

[3]For terminology see (Crow & Shimizu 1988).

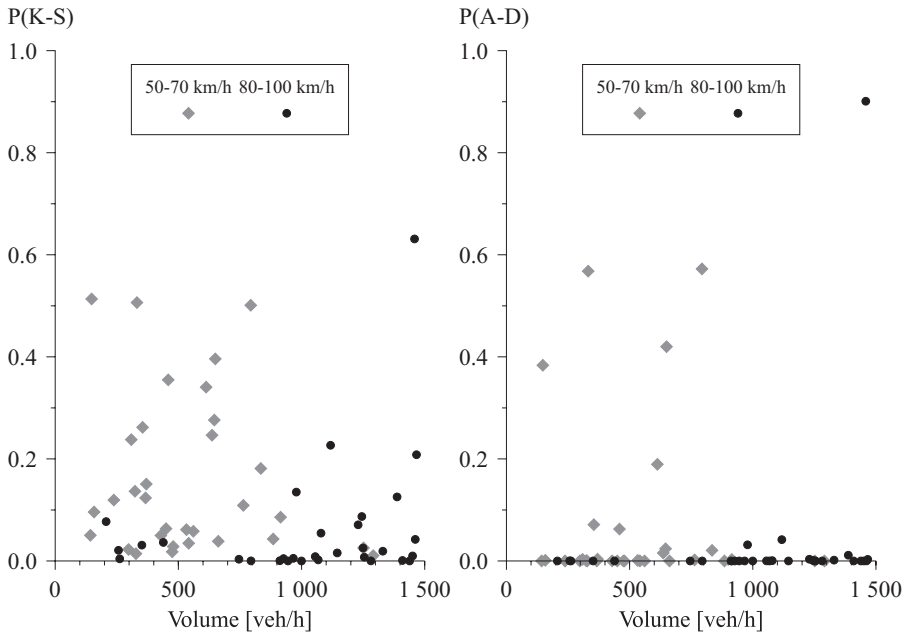

Figure 5.24: The nonparametric K-S and the parametric A-D goodness-of-fit tests for the lognormal distribution

estimation of Cohen & Whitten (1980) was also applied. The estimates were not as accurate as the modified maximum likelihood estimates. The local maximum likelihood method gave the most accurate estimates, but it was rejected because of convergence problems. In the goodness-of-fit tests below the parameters were estimated by the modified maximum likelihood method.

Goodness-of-fit tests

The goodness-of-fit was tested using the parametric Anderson-Darling (A-D) test. For each sample 9,999 replications were generated and analyzed. Figure 5.24 shows the test results. For comparison, the results of nonparametric Kolmogorov-Smirnov (K-S) tests are also presented.

The results are better than for the exponential or the gamma distribution, but the combined probabilities are still practically zero (A–D test: $1.80 \cdot 10^{-114}$, K–S test: $1.84 \cdot 10^{-36}$). The fit is better on low speed roads (A–D test: $1.78 \cdot 10^{-15}$), but the hypothesis of lognormal vehicle headways should be rejected.

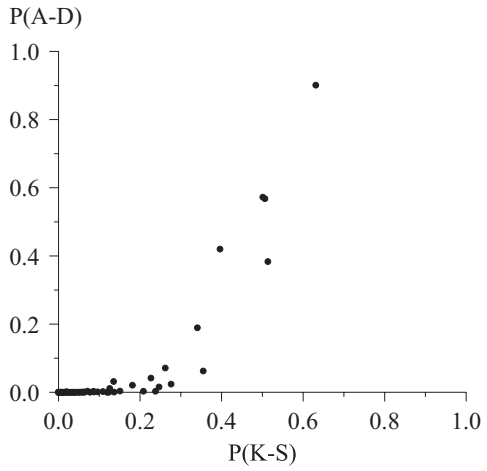

Figure 5.25: Comparison of the tests for the lognormal distribution

The comparison of the K-S and the A-D tests (fig. 5.25) shows that the 20 % significance level of the nonparametric K-S test is about equal to the 1–2 % significance level of the parametric A-D test. This, again, demonstrates the superiority of the parametric tests over the standard K-S test.

Lognormal distribution in headway studies

Reasonability Daou (1964) reported a good fit to lognormal distribution of spacings within platoons. He considered space headways less than 200 ft (61 m) with speed differences less than 5 ft/s (5.5 km/h). Two years later Daou (1966) presented a more detailed analysis of the data and also a theoretical justification for using lognormal distribution as a model of constrained headways.

Greenberg (1966) observed that "there may be some 'universal' law of traffic described (or at least approximated) by log-normal distribution". The lognormal relation holds if the change in the headway of a vehicle in any small interval of time is a random proportion of the headway at the start of the interval, and the mean headway of a vehicle and the variance of its headways remain constant over time. Greenberg also showed a connection between this model and the car-following models.

According to Baras, Dorsey & Levine (1979*a*), multiplicative, independent, and identically distributed errors by various drivers attempting to follow each other combine to give a lognormal density. Consequently, the

lognormal distribution appears to be an attractive model for follower headways.

Applicability The modified maximum likelihood method requires raw headway data and the application of numerical methods, but modified moment estimates are relatively easy to calculate, although with some loss of accuracy. Lognormal random variate generation is relatively efficient with currently available routines. The lack of Laplace transform, however, limits the applicability of the lognormal distribution.

With currently available routines to calculate the PDF of the normal distribution, even in pocket calculators, there is potential for a wider application of the lognormal distribution. In mathematical analysis the lognormal distribution is more complicated than, say, the gamma distribution.

Validity Tolle (1971) studied the three parameter lognormal distribution for headway data from a corridor on Interstate 71 in Ohio. The traffic volumes ranged from 800 to 1,900 veh/h. The location parameter was fixed ($\tau = 0.3$). The K-S tests gave nonsignificant results in all samples at 5 % level. Six out of eleven samples produced nonsignificant values at 5 % significance level using the chi-square test. The results were best under lowest flow (800–1,200 veh/h) conditions. The significance probabilities were not presented.

Branston (1976) studied headway data from M4 motorway in West London and from 50 sites on two-way roads around Bloomington, Indiana. The results of the chi-square tests[4] for the 3-parameter lognormal distribution were not very encouraging. Branston also used the lognormal distribution as a model for constrained headways in his lognormal generalized queuing model (see section 5.3.7).

As a statistical model for vehicle headways, the lognormal distribution has two major problems (Luttinen 1994):

1. The goodness of fit, although better than for the exponential or the gamma distributions, is not acceptable.

2. The mathematical analysis presents some problems, because the lognormal distribution does not have a Laplace transform. If it is neces-

[4]Branston used average probability as a combined measure for goodness of fit. This, however, gives an unduly optimistic view. As an example: The five samples from the slow lane on M4 gave average probability 0.105, but the combined probability is 0.0002. For the mixed model, the 11 samples from loop 10 (Indiana) gave the best average probability 0.327, but the combined probability is only 0.036.

sary to use numerical or Monte Carlo methods, one may as well use a distribution which fits the headway data better. Thus, the lognormal distribution is neither simple enough nor realistic enough.

Consequently, the lognormal distribution has a limited use as an overall headway distribution model, but it deserves proper attention in studies on follower headways.

5.3 Mixed distributions

5.3.1 Structure of the model

The main problem of the simple distributions is their inability to describe both the sharp peak and the long tail of the headway distribution. Even at low volume conditions, there is a considerable accumulation of headways near the mode. Furthermore, the empirical hazard functions (figs. 4.6 and 4.7, page 62) indicate that the tail is exponentially distributed. These properties of the headway distributions suggest two vehicle categories—free flowing and following. Because vehicles in different categories have different headway properties, the division should be included in the model, which accordingly becomes a mixture of two distributions.

According to Dawson & Chimini (1968) a vehicle is considered free, if:

1. The headway is of "adequate" duration.
2. The free vehicle is able to pass so that it does not have to modify its time-space trajectory, as it approaches the preceding vehicle.
3. A passing vehicle has sustained a positive speed difference after the passing maneuver so that the free vehicle is still able to operate as an independent unit.

The other vehicles are followers. The vehicles could also be classified into four or five categories:

1. Free flowing vehicles
2. Followers
3. Vehicles in a transition stage from free flowing to follower
4. Vehicles starting a passing maneuver.

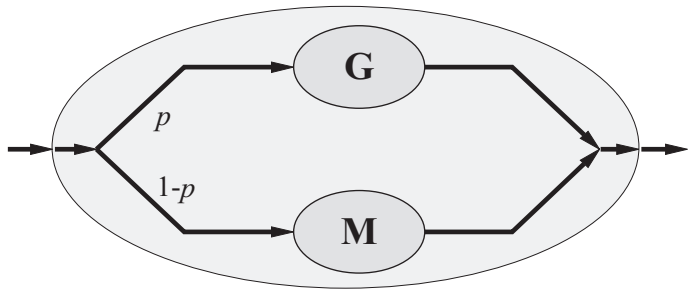

Figure 5.26: Queuing model for a mixed headway distribution

A fifth category would include the vehicles in the state of passing, but these vehicles are driving on the other lane.

The vehicles with moderate headways are a mixture of categories 1, 2, 3 and 4. Because it is difficult to make a distinction between vehicles in categories 2, 3 and 4, and in order to keep the model simple, the headway distribution is considered as a mixture of free headways and constrained (follower) headways only. A follower is then any vehicle in category 2, 3 or 4.

The most common approach with two vehicle categories is to consider a mixture of two probability distributions. A vehicle (X) is either a follower ($X = 1$) or a free flowing vehicle ($X = 2$). In terms of queuing theory, this model may be described as a system of two servers, only one of which may serve at a time (fig. 5.26). There is a continuous queue of vehicles to the system. Each vehicle is served by server 1 with probability p and by server 2 with probability $1 - p$. Server i has service time distribution $f_i(t)$. The headway distribution model is the distribution of departure intervals from the system. This system differs from an ordinary queuing system in two respects:

1. The system is always busy because of a continuous queue to the system.

2. Only one customer (vehicle) may be in the system at a time.

The model may be analyzed by the methods of the queuing theory, and especially by the *method of stages* (Cooper 1981, Luttinen 1990).

5.3.2 Properties of the model

The density functions for constrained headways $[f_1(t)]$ and free headways $[f_2(t)]$ are the conditional densities:

$$f_1(t) = f(\, t \,|\, X = 1\,) \tag{5.75}$$
$$f_2(t) = f(\, t \,|\, X = 2\,)\,. \tag{5.76}$$

If the proportion of constrained headways is p, the *probability density function* (pdf) of the headway distribution is:

$$f(t) = pf_1(t) + (1-p)f_2(t). \tag{5.77}$$

The *hazard function* is:

$$h(t) = \frac{f(t)}{1-F(t)} = \frac{pf_1(t) + (1-p)f_2(t)}{1-[pF_1(t) + (1-p)F_2(t)]}. \tag{5.78}$$

The distribution has *Laplace transform*:

$$f^*(s) = pf_1^*(s) + (1-p)f_2^*(s) \tag{5.79}$$

where $f_i^*(s)$ is the transform of $f_i(t)$, if the transform exists.

The *mean* is:

$$\mu(T) = p\mu(T_1) + (1-p)\mu(T_2). \tag{5.80}$$

where $\mu(T_i)$ is the expectation of headway variate T_i from distribution $f_i(t)$. Higher noncentral moments are obtained similarly:[5]

$$E(T^r) = pE(T_1^r) + (1-p)E(T_2^r). \tag{5.81}$$

Variance, skewness and kurtosis can be calculated using the relations be-

[5] The moments can also be obtained by using the derivatives of the Laplace transform:

$$E(T^r) = (-1)^r \frac{d^r f^*(s)}{ds^r}\bigg|_{s=0}.$$

tween central and noncentral moments:

$$\sigma^2(T) = E[T - \mu(T)]^2 \tag{5.82}$$
$$= E(T^2) - [\mu(T)]^2$$
$$\alpha_3(T) = \frac{E[T - \mu(T)]^3}{\sigma^3(T)} \tag{5.83}$$
$$= \frac{1}{\sigma^3(T)} \left\{ E(T^3) - 3\mu(T)E(T^2) + 2[\mu(T)]^3 \right\}$$
$$\alpha_4(T) = \frac{E[T - \mu(T)]^4}{\sigma^4(T)} \tag{5.84}$$
$$= \frac{1}{\sigma^4(T)} \left\{ E(T^4) - 4\mu(T)E(T^3) + 6[\mu(T)]^2 E(T^2) - 3[\mu(T)]^4 \right\}.$$

Some proposed mixed distribution models are presented below. Detailed analysis and goodness-of-fit tests for all these distributions are beyond the scope of this research. The semi-Poisson distribution is given considerable attention, but other mixed distributions are discussed more briefly, using only information that has been readily available. Because all the models are based on the assumption of exponentially distributed large headways, the exponential tail hypothesis is examined first.

5.3.3 The exponential tail hypothesis

Goodness-of-fit tests

The exponential tail hypothesis states that headways greater than a threshold (δ) follow the negative exponential distribution:

$$H_0 \colon 1 - F_n(t) \sim e^{-\hat{\theta}t}, \qquad t > \delta, \tag{5.85}$$

where $\hat{\theta}$ is the scale parameter estimate and n is the number of headways greater than δ in the sample. The test has two stages:

1. Determination of δ
2. Goodness-of-fit test for headways greater than δ.

The goodness-of-fit tests were conducted using threshold values 0–14.5 in increments of 0.5. Parametric Monte Carlo tests with the A-D statistic were used. The number of replications was 9,999. The significance probabilities were combined using the method described in section 2.2.5 (page 43).

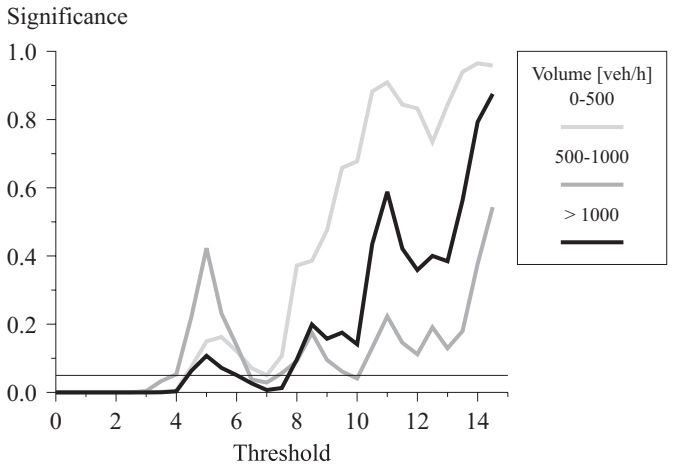

Figure 5.27: Goodness-of-fit tests for the exponential tail (50–70 km/h)

Figures 5.27 and 5.28 show the combined significance probabilities for different volume levels. Figure 5.29 shows the combined significance probabilities for different speed limits. The threshold value for the exponential tail hypothesis is eight seconds, after which the test results are nonsignificant. Because of such a large threshold value, a positively skew headway distribution should be considered for the follower headways.

Low speed roads (fig. 5.27 and 5.29) show nonsignificant values also at $4.5 \leq \delta \leq 6$ s. This phenomenon appears at all volume levels. It may indicate lower threshold values. The lower threshold ($\delta = 4.5$) is, however, not accepted, because there are significant deviations from the exponential distribution at ($4.5s < \delta < 8s$). The speed difference data (fig. 5.30, page 133) also support a higher threshold.

Miller (1961) tested for threshold values four, five and six seconds, and found $\delta = 6$ s about ideal. His data were from a straight section on a two-lane road between Stockholm and Uppsala in Sweden. He noted that there is a slight excess of observed intervals in the 6–8 s range as some drivers begin to slow down, when they approach this close to the vehicle ahead. He decided that some criterion based upon relative velocities is necessary for determining the platoons, when time intervals are less than eight seconds.

Wasielewski (1979) studied the exponential tail hypothesis of headway distributions from the center westbound lane of Route I-94, a six-lane divided urban expressway in Detroit. He found no significant deviations from the exponential distribution, when the threshold was 2.5 seconds for volumes

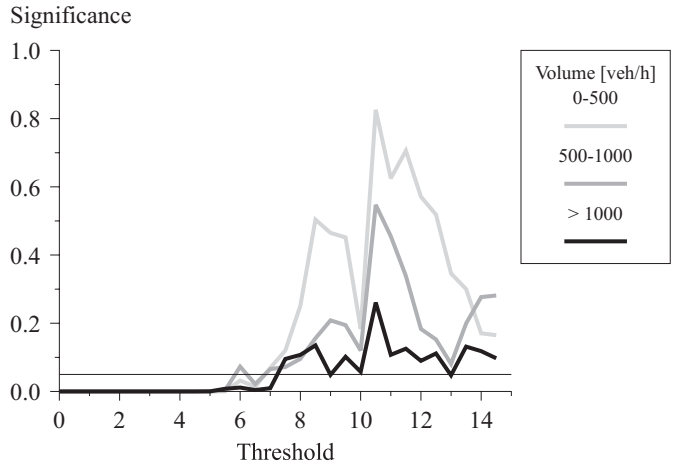

Figure 5.28: Goodness-of-fit tests for the exponential tail (80–100 km/h)

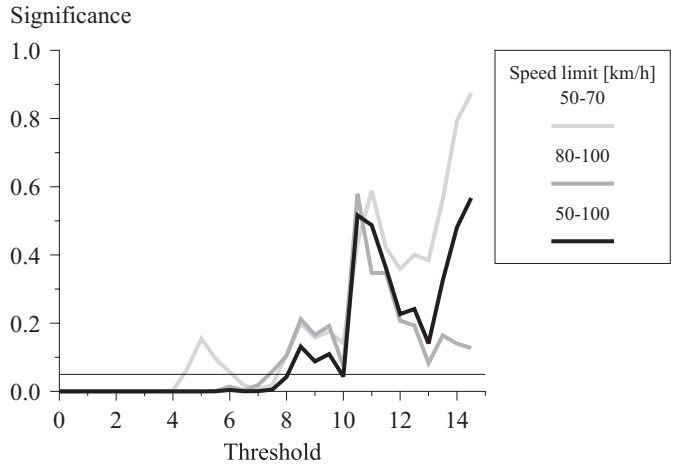

Figure 5.29: Goodness-of-fit tests for the exponential tail

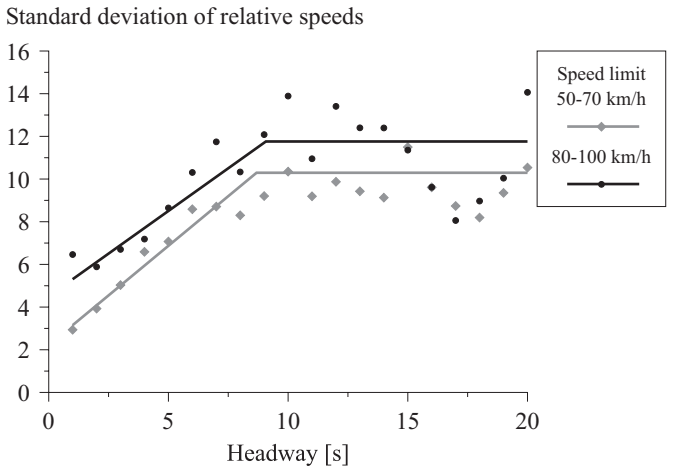

Figure 5.30: Standard deviation of speed differences

less than 1,450 veh/h and 3.5 seconds for volumes greater than 1,450 veh/h. These values are too low for two-lane roads.

It is possible that the goodness-of-fit tests give too small threshold estimates. The tests cannot detect small deviations from the exponential distribution for headways slightly greater than δ, if the tail is exponential. Consequently, an alternative approach was also examined.

Speed differences

Properties of speed differences Car-following models (section 2.1) suggest that the speed difference (relative speed) between the leader and the follower ($v_{i-1}-v_i$) is a decisive factor in the car-following process. A follower adjusts the speed of the vehicle to the speed of the leading vehicle. This speed adjustment decreases the variation of speeds among trailing vehicles. At some time distance the interaction of speeds vanishes.

Figure 5.30 shows the standard deviation (s_v) of speed differences of consecutive vehicles against the headway. Headways are combined in one second intervals $(t-1,t]$. At short headways (1–5 s) s_v is small and increases as the headway increases. At large headways ($t > 10$ s) s_v is rather constant. On high speed roads s_v is larger than on low speed roads.

A piecewise linear model was fitted to the data—rising slope for headways less than the threshold and a constant value for headways greater than the threshold. The s_v's were weighted by the number of observations in the interval. On both low and high speed roads the threshold value of about 9 s

was obtained. However, more extensive data sets and more accurate measuring equipment are needed for further analysis, because speed differences have larger measurement errors than absolute speeds.

This result is in agreement with the goodness-of-fit tests above. The original Bureau of Public Roads (1950) reports a similar analysis and virtually the same result. Chishaki & Tamura (1984) present a very similar figure of the variance of the speed differences. Branston (1979) reports threshold values 4.5 s and 3.75 s for the nearside and offside lanes on M4 motorway in West London.

Enberg & Pursula (1992) have shown that on high-class two-lane rural roads, the relation between the headways and the standard deviation of speed differences is similar to figure 5.30. For headways larger than about five seconds, there is no significant increase in the standard deviation. When the headway is less than one second, the standard deviation jumps to about 10 km/h. This phenomenon is also reported in Bureau of Public Roads (1950) and by Branston (1979). A similar, but not as strong, phenomenon can be seen on high speed roads in figure 5.30. It may be assumed, that at very short headways some vehicles are starting to pass the vehicle ahead. The increase in s_v is very small or nonexistent in figure 5.30, because there were only few passing vehicles.

Figure 5.30 shows that vehicles may be divided into three categories. Leading vehicles are not affected by vehicles ahead. The drivers select their own desired speed and headways can be assumed to follow the exponential distribution. Trailing vehicles follow the vehicle ahead and adjust their speeds accordingly. Because the mode of the headway distribution is about 1.5 seconds, this headway can be assumed to be the average desired headway of the followers. In addition to these two classes there are vehicles in a transition stage. These vehicles are neither following the vehicle ahead nor driving freely. The transition stage seems to begin, when the headway is about 8–9 seconds. Speed adjustment decreases the variation between speeds of consecutive vehicles. At headways less than one second, the speed variation increases again on high speed roads, because the vehicles starting to pass have greater speed than the vehicle ahead.

Figures 5.31 and 5.32 show some basic statistics of speed difference distributions. Mean is near zero at all headways. As the headway is between two and eight seconds, the mean is slightly positive, which is probably caused by passers that have returned to the right lane. For headways greater than eight seconds the variation in the mean increases, and the mean oscillates around zero. When the headways are greater than ten seconds, the mean speed difference is negative. In this case some proportion of the vehicles

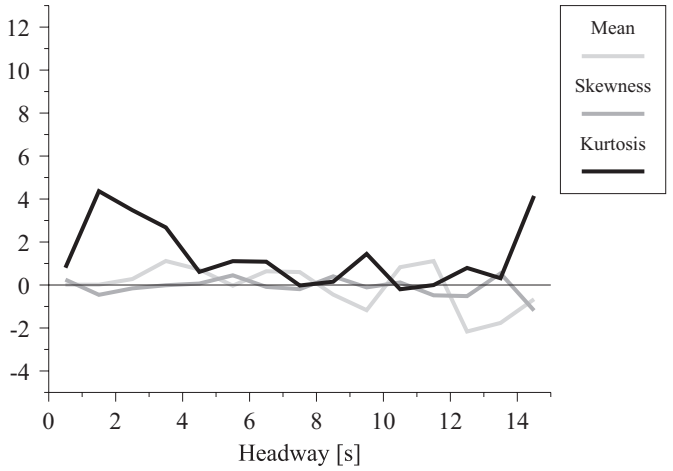

Figure 5.31: Mean, skewness and kurtosis of speed difference distributions (50–70 km/h)

ahead are followers having constrained speeds (v_{i-1}), while the speeds of the vehicles behind (v_i) are nonconstrained.

The skewness indicates that the speed difference distribution is practically symmetric. On high speed roads at headways less than one second, the skewness has a rather large negative value. Considering the slightly negative mean and the increased standard deviation, the negative skewness is most likely caused by vehicles beginning to pass, thus producing negative speed differences.

Summala & Vierimaa (1980) found that the speed difference distribution becomes skew, as the spacing grows over 50 m. At spacings longer than 200 m the distribution is again symmetric. This phenomenon is not found in figures 5.31 and 5.32.

The most significant feature in the figures is the excessively positive kurtosis at short headways. (The kurtosis is scaled so that the normal distribution has null kurtosis.) The speed difference distributions are more sharply peaked than the normal distribution (see also fig. 5.33). This is true especially for headways less than four seconds. The kurtosis is largest at headways between one and two seconds, and the peak is higher on high speed roads. Because at headways between one and two seconds the distribution is symmetric with zero mean, it is evident that at these headways most drivers are content to follow the leader at the same speed. The kurtosis approaches zero and the distribution approaches normality at headways

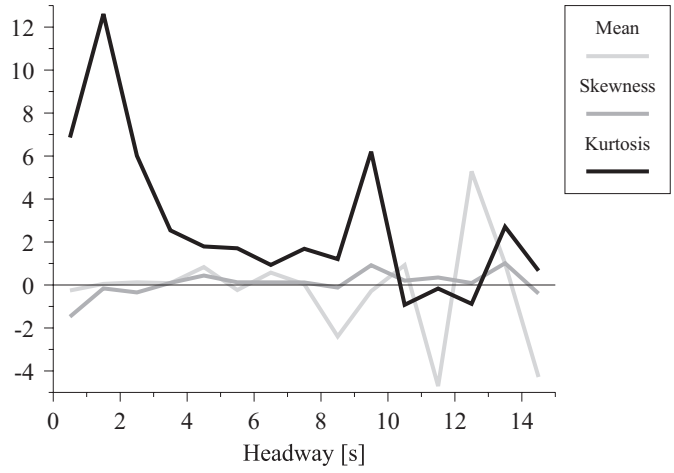

Figure 5.32: Mean, skewness and kurtosis of speed difference distributions (80–100 km/h)

larger than seven seconds on low speed roads, and at headways larger than ten seconds on high speed roads. The variation in the kurtosis increases at large headways due to fewer observations.

Proportion of free headways The distribution of speed differences at long headways can be assumed to be the distribution of speed differences for free flowing vehicles. The mean and standard deviation of speed differences at headways $t \geq 10$ s are $\bar{v} = -0.879$ and $s_v = 10.492$ on low speed roads, and $\bar{v} = -0.862$ and $s_v = 11.660$ on high speed roads. Assuming normality, the means do not differ significantly from zero, but the nonequality of the standard deviations is statistically significant (at level $P = 0.001$). On high speed roads the speed differences have larger variation.

Because not every vehicle having a headway shorter than eight seconds is a follower, the distribution of their speed differences is a mixture of free speed differences and constrained speed differences. If the speed difference distribution $f_2(\mu_1, \sigma_2)$ of free flowing vehicles is estimated, the speed difference distribution $f(\mu, \sigma)$ of all vehicles can be estimated as a mixed normal distribution (fig. 5.30):

$$f(\mu, \sigma) = p f_1(\mu_1, \sigma_1) + (1 - p) f_2(\mu_2, \sigma_2), \tag{5.86}$$

The proportion of trailing vehicles (p) and the parameters of the speed difference distribution (μ_1, σ_1) of trailing vehicles must be estimated from the data.

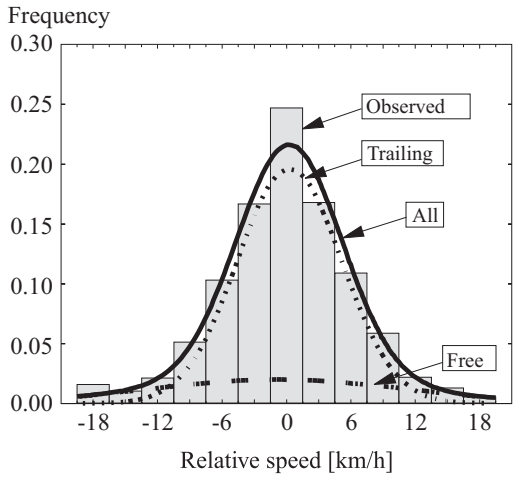

Figure 5.33: Mixed normal distribution for speed differences on high speed (80–100 km/h) roads

The minimum chi-square method (section 2.2.3) was used to estimate p, while maximum likelihood estimates were used for μ_1 and σ_1. The estimate for the proportion of free flowing vehicles $(1 - \tilde{p})$ among headways $t \leq 8$ s is 20 % on high speed roads and 15 % on low speed roads. The proportion of trailing vehicles is approximately the same as the proportion of headways $t \leq 3$ s on high speed roads and $t \leq 5$ s on low speed roads.

Parameter estimation for the exponential tail

The goodness-of-fit tests above (p. 130) indicated that headways larger than eight seconds belong to the exponential tail. Assuming the eight second threshold, the scale parameter estimate for the exponential tail is equal to the scale parameter estimate for the shifted exponential distribution with $\tau = 8$ s:

$$\hat{\theta} = \frac{1}{\bar{t}_8 - 8} \tag{5.87}$$

where $\bar{t}_8$ is the mean of headways greater than eight seconds.

The number of free vehicles can be assumed to increase with increasing flow rate. Consequently, the scale parameter (θ) of the exponential tail can be assumed to have a positive correlation with the traffic volume $(\tilde{\lambda})$.[6] A linear model $(\tilde{\theta} = a + b\tilde{\lambda})$ was tested using the linear regression analysis.

[6] See also section 4.2.3, page 60.

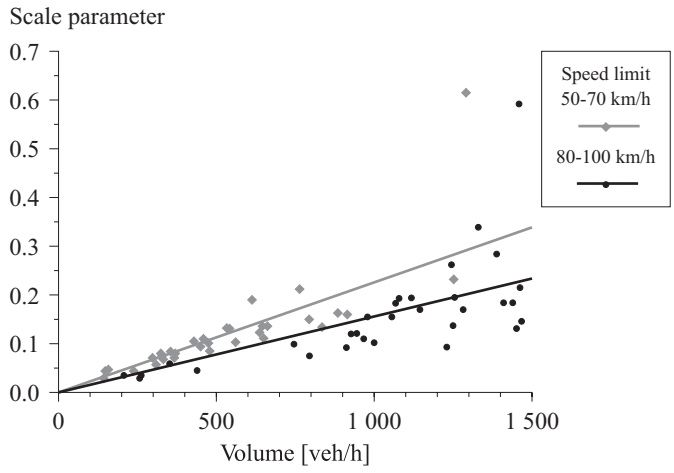

Figure 5.34: Scale parameter estimates for the exponential tail

Because the analysis of covariance (ANCOVA) indicated that there is a significant ($P = 0.006$) difference between the scale parameter estimates for low speed and high speed roads, separate models ($\tilde{\theta}_{50}$, $\tilde{\theta}_{80}$) were tested for both road categories. The constant term (a) was omitted, because it did not differ significantly from zero. (As the flow rate approaches zero, the scale parameter must also approach zero.) The following equations for the scale parameter estimates were obtained (fig. 5.34):

$$\tilde{\theta}_{50} = 0.813\tilde{\lambda} \tag{5.88}$$

$$\tilde{\theta}_{80} = 0.561\tilde{\lambda}, \tag{5.89}$$

where $\tilde{\lambda}$ is the traffic volume [veh/s].[7]

In order to stabilize the variance, which increases with growing traffic volume, $1/\tilde{\lambda}^2$ was used as a weighting factor. The coefficient of determination is $R^2 = 0.936$ for low speed roads and $R^2 = 0.823$ for high speed roads. The significance of the regression is $P < 0.0005$ in both cases.

On high speed roads the estimates are lower. This indicates a "longer tail" and accordingly higher coefficient of variation, skewness and kurtosis. The scale parameter estimates are also the estimates for the hazard rate at the tail of the headway distribution.

[7] Wasielewski (1979) obtained for a center expressway lane:

$$\tilde{\theta} = -0.058 + 1.401\tilde{\lambda}. \tag{5.90}$$

The scale parameter was estimated using the location parameter value of eight seconds. The estimate is, however, valid even if the true location parameter has a lower value. This can be shown easily. If the true location parameter is a, then:

$$\mathbb{P}\{T > t\} = e^{-\theta(t-a)}. \tag{5.91}$$

If only headways greater than b (and $a < b$) are examined, then:

$$\begin{aligned} \mathbb{P}\{\, T > t \,|\, T > b \,\} &= \frac{e^{-\theta(t-a)}}{e^{-\theta(b-a)}} \\ &= e^{-\theta(t-b)}. \end{aligned} \tag{5.92}$$

A larger location parameter does not cause any bias in the scale parameter estimate. The estimate is, however, based on fewer observations, thus reducing the accuracy.

5.3.4 Hyperexponential distribution

Properties of the hyperexponential distribution

Hyperexponential distribution was first applied to headway studies by Schuhl (1955), hence it is also known as Schuhl's (composite exponential) distribution. If both leader and follower headways are exponential, the headway distribution is:

$$f(\, t \,|\, p, \theta_1, \theta_2 \,) = p\theta_1 e^{-\theta_1 t} + (1-p)\theta_2 e^{-\theta_2 t}. \tag{5.93}$$

The headways are produced by a service system with two parallel servers, each having exponential service times (fig. 5.35). As noted above, there is a constant queue to the system.

The pdf of the hyperexponential distribution (fig. 5.36) starts from $f(0) = p\theta_1 + (1-p)\theta_2$ and decreases asymptotically to zero. The hazard function decreases from $h(0) = p\theta_1 + (1-p)\theta_2$ to $\lim_{t\to\infty} h(t) = \min\{\theta_1, \theta_2\}$ (Trivedi 1982). The coefficient of variation is $C(T) \geq 1$ (Trivedi 1982).

Schuhl (1955) used a location parameter (τ) for the follower headways. The resulting pdf was:

$$f(\, t \,|\, p, \tau, \theta_1, \theta_2 \,) = p\theta_1 e^{-\theta_1(t-\tau)} + (1-p)\theta_2 e^{-\theta_2 t}. \tag{5.94}$$

Free headways were not given a location parameter, because these vehicles were considered free to pass.

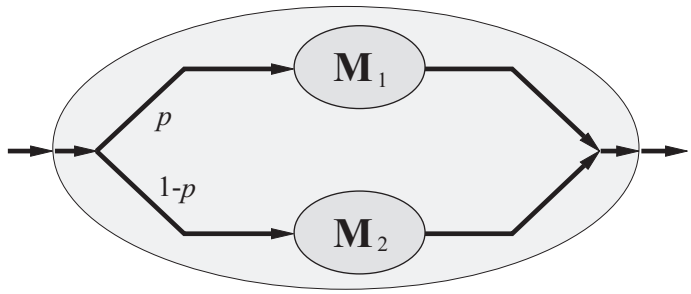

Figure 5.35: Queuing model for the hyperexponential distribution

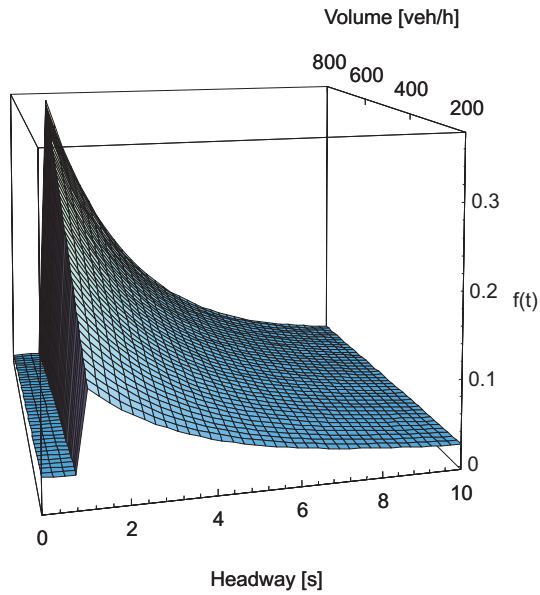

Figure 5.36: Hyperexponential density functions with parameters by Grecco & Sword (1968)

Hyperexponential distribution in headway studies

Reasonability As compared with the negative exponential and the shifted exponential distributions, the hyperexponential distribution has the advantage of being able to produce a larger coefficient of variation. It also has separate models for free and follower headways. The negative exponential distribution is, however, not an appropriate model for constrained headways. The problem of too short headways can be avoided by location parameters, but this introduces an artificial discontinuity in the density function.

Applicability Pseudo-random number generation is fast, although not as efficient as for the negative exponential distribution. The Laplace transform has a simple form:

$$f^*(\, s \mid p, \theta_1, \theta_2\,) = p\frac{\theta_1}{s+\theta_1} + (1-p)\frac{\theta_2}{s+\theta_2}. \tag{5.95}$$

Without location parameters the hyperexponential distribution has three parameters to estimate. Introducing one or two location parameters increases the total number of parameters to four or five, which makes the parameter estimation more difficult.

Grecco & Sword (1968) assumed a fixed value of $\tau = 1$ s for the location parameter, and estimated the three remaining parameters by minimizing the chi-square statistic. The regression equations for the parameters were:

$$\tilde{p} = \frac{0.115\tilde{\lambda}}{100} \tag{5.96}$$

$$\tilde{\tau} = 1 \tag{5.97}$$

$$\tilde{\theta_1} = 0.4 \tag{5.98}$$

$$\tilde{\theta_2} = \frac{1}{24 - 0.0122\tilde{\lambda}}, \tag{5.99}$$

where $\tilde{\lambda}$ is the traffic volume [veh/h]. These estimates are, however, valid only for $\tilde{\lambda} < 870$ veh/h, since $\tilde{p} > 1$ for $\tilde{\lambda} \geq 870$ veh/h. Figure 5.36 shows the pdf calculated with these parameter values.

Griffiths & Hunt (1991) applied the location parameter (τ) to both $f_1(t)$ and $f_2(t)$, and called the model *double displaced negative exponential distribution* (DDNED). The parameters were estimated by a "hybrid method", using the equations for the mean and the variance as constraints, and minimizing the chi-square statistic. The estimation was constrained so that $\tilde{p}$

remained in the range $0 < \tilde{p} \leq 0.5$, and $\tilde{\tau}$ was required to be greater [*sic*][8] than the smallest observed headway ($t_{(1)}$).

Validity Grecco & Sword (1968) collected headway data on a two-lane portion of the U.S. 52 bypass. The headway data were grouped in one second intervals up to twenty seconds. The data were collected in volume groups using the class length of 100 veh/h. The chi-square tests with 8–17 degrees of freedom gave acceptable fit for volumes less than 700 veh/h. The combined probability (calculated by the present author) of the six tests is 0.79.

Griffiths & Hunt (1991) examined the vehicle headways in urban areas. The model had the same location parameter τ for both $f_1(t)$ and $f_2(t)$. The mean of $\tilde{\tau}$ was 1.5 s, the minimum 0.58 s, and the maximum 2.66 s. The chi-square test with 5 % significance level gave acceptable fit for 78 of 86 data sets. Because $\tilde{\tau}$ was greater than the shortest observed headway in each sample, the model would have failed all A-D tests with null significance level.

Sainio (1984) tested the hyperexponential distribution[9] against headway data from high class Finnish highways. He presented only the chi-square test statistic values for the five samples. When the significance probabilities are calculated ($\nu = 10$), the model is rejected in three samples at 0.01 level. The combined significance probability is $3.9 \cdot 10^{-6}$.

The hyperexponential distribution does not pass the identification process. Although the coefficient of variation can be greater than unity, the shape of the pdf and the hazard function do not have the correct shape. The hyperexponential distribution has similar properties as the gamma distribution with shape parameter less than unity. The pdf has a similar shape, and the c.v. is larger than unity. The hyperexponential distribution, however, has a more reasonable hazard rate at $t = \tau$, and the kurtosis is slightly closer to the empirical kurtosis, when plotted against the squared skewness. Consequently, the identification process suggests that the hyperexponential distribution could perform slightly better than the gamma distribution as a headway model.

[8] If the location parameter (τ) is greater than the smallest headway ($t_{(1)}$), the probability $\mathbb{P}\{T \leq t_{(1)}\}$ is null. Consequently, the headway sample cannot be generated by the proposed model.

[9] Sainio (1984) called it the Kell distribution (see Kell 1962).

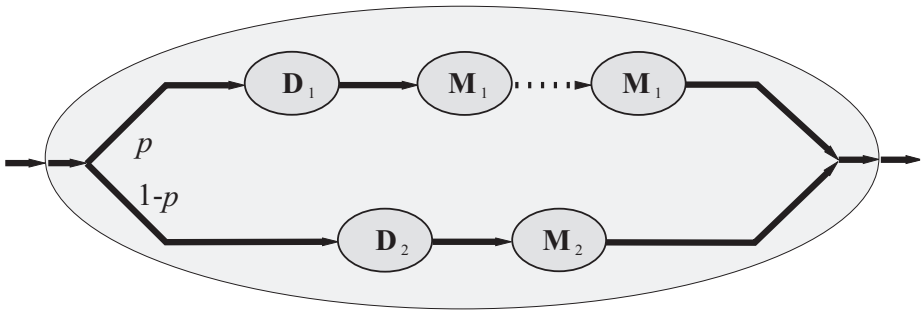

Figure 5.37: Queuing model for the hyperlang distribution

5.3.5 Hyperlang distribution

Properties of the hyperlang distribution

Dawson (in Drew 1967, Dawson & Chimini 1968, Dawson 1969) has suggested the Erlang-distribution as a model for constrained headways. The headways in the hyperlang (hyper-Erlang) model are produced by a service system with two alternative channels (fig. 5.37). Channel 1 is selected with probability p, and it has one server with Erlang-distributed service times, or alternatively r servers in series each having the same exponential service time distribution. The service time distribution in channel 2 is exponential. A location (threshold) parameter is shown as a server (D) with a constant service time.

The hyperexponential distribution is naturally a special case of the hyperlang distribution ($r = 1$). The generalization to *exponential-gamma mixture distribution* is straightforward.

The pdf of the hyperlang distribution is:[10]

$$f(\,t\,|\,p,\tau_1,\theta_1,\tau_2,\theta_2,r\,) = p\frac{\theta_1^r(t-\tau_1)^{r-1}}{(r-1)!}e^{-\theta_1(t-\tau_1)} + (1-p)\theta_2 e^{-\theta_2(t-\tau_2)}. \tag{5.100}$$

[10]In place of θ_i Dawson (1969) used parameter γ_i, so that:

$$\theta_i = \frac{1}{\gamma_i - \tau_i},$$

where γ_i is the average headway in channel i.

The distribution has Laplace transform (Luttinen 1990):

$$f^*(\, s \mid p, \tau_1, \theta_1, \tau_2, \theta_2, r \,) = p \left(\frac{\theta_1}{s + \theta_1} \right)^r e^{-s\tau_1} + (1-p) \frac{\theta_2}{s + \theta_2} e^{-s\tau_2}. \quad (5.101)$$

The mean and the variance are (Luttinen 1990):

$$\mu(T) = p \left(\frac{1}{\theta_1} + \tau_1 \right) + (1-p) \left(\frac{r}{\theta_2} + \tau_2 \right) \quad (5.102)$$

$$\sigma^2(T) = p \left[\left(\frac{1}{\theta_1} + \tau_1 \right)^2 + \frac{1}{\theta_1} \right] + (1-p) \left[\left(\frac{r}{\theta_2} + \tau_2 \right)^2 + \frac{r}{\theta_2} \right] - \left[p \left(\frac{1}{\theta_1} + \tau_1 \right) + (1-p) \left(\frac{r}{\theta_2} + \tau_2 \right) \right]^2 . \quad (5.103)$$

The hazard function of the hyperlang distribution is similar to the empirical hazard function, except that at zero headway $h(\, 0 \mid p, \theta_1, \theta_2, r \,) = \theta_2$, assuming $\tau_1 = \tau_2 = 0$ (Trivedi 1982).

Hyperlang distribution in headway studies

Reasonability The hyperlang distribution has exponential tail, and the shape of the pdf is similar to the empirical headway distributions. Some additional flexibility could be obtained using gamma distribution for constrained headways.

The distribution avoids extremely short headways, but the location parameter (τ_2) of the free headway distribution makes the pdf and the hazard function discontinuous. The model assumes that the frequency of free headways is highest, when the headway is only slightly greater than τ_2, and there are no free headways, when t is a little smaller than τ_2. The problem is more outstanding at low flow conditions, when the proportion of free headways is large.

Applicability The large number of parameters (six) causes problems in parameter estimation. Dawson & Chimini (1968) and Summala & Vierimaa (1980) have estimated the parameters using least squares fit. For the headway data in HCM (1965), Dawson & Chimini (1968) estimated $r = 2$ and $\tau_2 = 0.75$ for all samples. The estimated density and hazard functions are displayed in figures 5.38 and 5.39.

The generation of pseudorandom numbers is reasonably effective. The Laplace transform exists, but the mathematical analysis is more tedious than the analysis of the exponential and the gamma distributions.

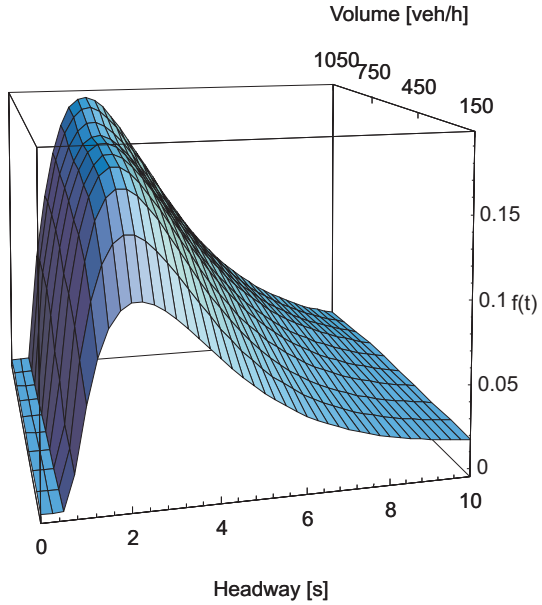

Figure 5.38: Hyperlang density functions [parameters by Dawson & Chimini (1968)]

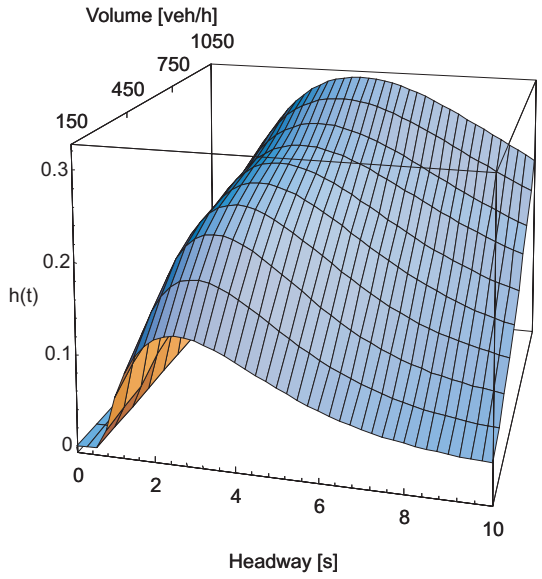

Figure 5.39: Hyperlang hazard functions [parameters by Dawson & Chimini (1968)]

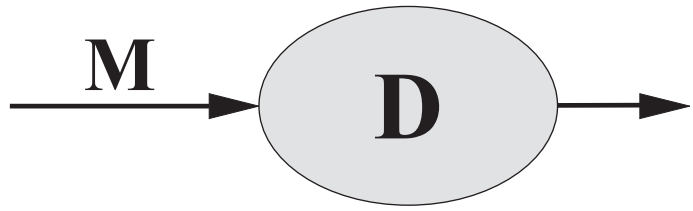

Figure 5.40: M/D/1 queuing model

Validity Dawson & Chimini (1968) and Dawson (1969) did not report any goodness-of-fit tests. Because the hyperlang distribution is a generalization of negative exponential, shifted exponential, Erlang and hyperexponential distributions, it is likely to fit the headway data better. But for a six parameter distribution, the goodness-of-fit requirements are extremely high.

5.3.6 M/D/1 queuing model

Tanner (1953, 1961) has modeled the arrival process as a M/D/1 queuing process (fig. 5.40). The service system has one server with constant service time (τ). There is not, however, a continuous queue to the service point, but the arrival process is Poisson.

The vehicles in platoons have constant headways, and the free headways follow the shifted exponential distribution. The probability for a follower headway ($t = \tau$) is equal to the utilization factor[11] (ρ) of the server:

$$\mathbb{P}\{T = \tau\} = \lambda\tau = \rho \tag{5.104}$$

For steady state conditions, the utilization factor must be less than unity. The platoon length follows the Borel distribution, and the probability for platoon length k is (Tanner 1961):

$$p_k = \frac{(\lambda\tau k)^{k-1}}{n!} e^{-\lambda\tau k}. \tag{5.105}$$

Because the platoon length distribution is not geometric, the headways are not i.i.d.

[11] The utilization factor (ρ) is the expected number of arrivals per mean service time in the limit (Gross & Harris 1985), or equivalently the fraction of time that the server is busy (Kleinrock 1975).

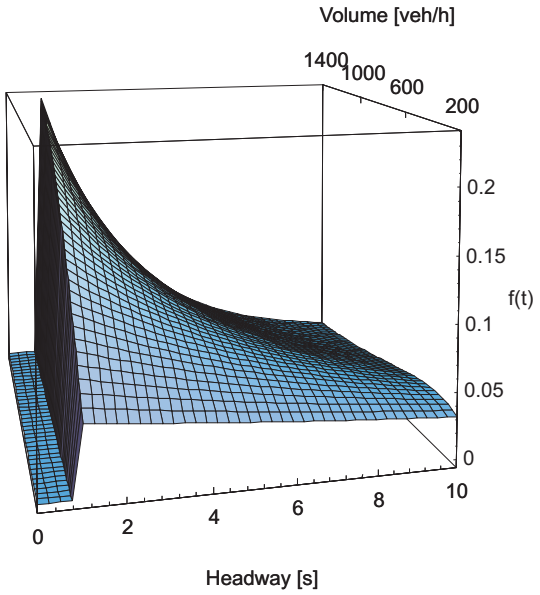

Figure 5.41: M/D/1 density function ($\tau = 1$ s)

The PDF of departure intervals (headways) is discontinuous (Luttinen 1990):

$$F(\,t\,|\,\lambda,\tau\,) = \begin{cases} 0, & \text{if } t < \tau \\ 1-(1-\lambda\tau)e^{-\lambda(t-\tau)}, & \text{otherwise.} \end{cases} \tag{5.106}$$

The pdf is obtained by differentiation (Luttinen 1990):

$$f(\,t\,|\,\lambda,\tau\,) = \begin{cases} 0, & \text{if } t < \tau \\ (1-\lambda\tau)\lambda e^{-\lambda(t-\tau)}, & \text{otherwise.} \end{cases} \tag{5.107}$$

The pdf is equal to the pdf of the shifted exponential distribution multiplied by the server idleness factor $(1-\rho)$. The mean, the variance, and the coefficient of variation are (Luttinen 1990):

$$\mu(T) = \frac{1}{\lambda} \tag{5.108}$$

$$\sigma^2(T) = \frac{1}{\lambda^2} - \tau^2 \tag{5.109}$$

$$C(T) = \sqrt{1-\lambda^2\tau^2} \leq 1. \tag{5.110}$$

Cowan (1975, 1978), used the headway distribution assuming a renewal process. The Cowan M3 model is accordingly a generalization of the M/D/1

model:

$$F(\,t\,|\,\theta,\lambda\,) = \begin{cases} 0, & \text{if } t < \tau \\ 1-(1-\theta)e^{-\lambda(t-\tau)}, & \text{otherwise.} \end{cases} \tag{5.111}$$

The parameters of the M/D/1 headway model can be easily estimated by the method of moments. The scale parameter estimate ($\tilde{\lambda}$) is equal to the traffic volume:

$$\tilde{\lambda} = \frac{1}{\bar{t}}, \tag{5.112}$$

and it does not depend on τ. The trailing headway (τ) can be estimated using the sample variance:

$$\tilde{\tau} = \sqrt{\frac{1}{\tilde{\lambda}^2} - s^2} \tag{5.113}$$

M/D/1 queuing model in headway studies

Reasonability In some sense the M/D/1 model follows logically from the shifted exponential distribution: If the vehicles move freely at headways greater than τ, and headways shorter than τ are not possible, there must be some platooning with headways equal to τ. The model separates leading and following vehicles, but the hypothesis of constant headways for following vehicles is unrealistic.

Applicability The M/D/1 model is rather simple to analyze mathematically. The parameter estimation is straightforward, and the generation of pseudorandom variates does not present any problems.

Validity Sullivan & Troutbeck (1994) applied the Cowan's M3 model to urban traffic flow. They concluded that the model is appropriate to evaluate the performance of roundabouts and unsignalized intersections, where accuracy is only required for headways greater than the critical gap. They did not present any goodness-of-fit tests.

Akçelik & Chung (1994) described the calibration of the M3 model, which they called the "bunched exponential distribution". On the basis of visual evaluation they considered that the model provides good estimates of arrival headways, and they strongly recommended its use in stead of the negative exponential and the shifted exponential distributions.

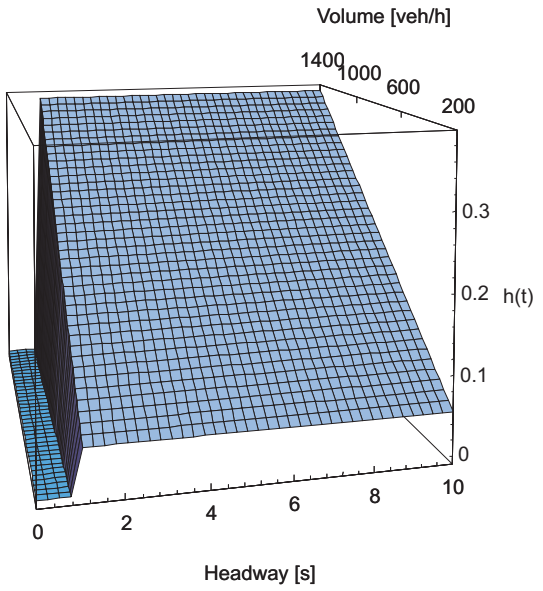

Figure 5.42: M/D/1 hazard function ($\tau = 1$ s)

The discontinuity of the pdf and even the PDF is a major problem in the model. The goodness-of-fit becomes worse as the proportion of trailing vehicles (ρ) increases. The unfeasibility of the model is, however, most obvious in the hazard function plot of figure 5.42. In addition, the coefficient of variation is too low for most flow conditions.

5.3.7 Generalized queuing model

A queuing model with Poisson arrivals and a general service time distribution (fig. 5.43) is called the M/G/1 queuing system (fig. 5.43). It produces follower headways (T_1) as long as the server is busy. With probability $1-p$, the server experiences some idleness during the interdeparture period. The length of an interdeparture time with idleness is the sum of the service time (T_1) and the idle time (T_2) (Gross & Harris 1985). The density of this sum is the convolution ($f_1 * f_2$) of the service time density and the interarrival time density, where interarrival time density (f_2) is exponential.

The pdf of the interdeparture time distribution is:

$$\begin{aligned} f(t) &= pf_1(t) + (1-p)\int_0^t f_1(u)\theta e^{-\theta(t-u)}du \\ &= pf_1(t) + (1-p)\theta e^{-\theta t}\int_0^t f_1(u)e^{\theta u}du. \end{aligned} \tag{5.114}$$

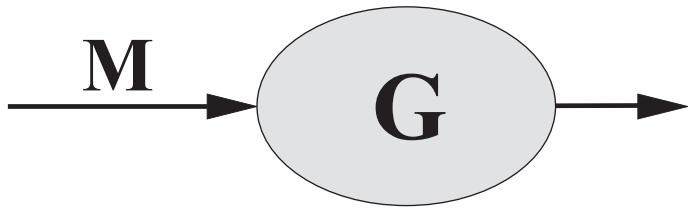

Figure 5.43: M/G/1 queuing system

The PDF is (Gross & Harris 1985):

$$F(t) = pF_1(t) + (1-p)\int_0^t F_1(t-u)\theta e^{-\theta u}du, \tag{5.115}$$

or equivalently:

$$\begin{aligned} F(t) &= pF_1(t) + (1-p)\int_0^t f_1(u)\left(1 - e^{-\theta(t-u)}\right)du \\ &= F_1(t) - (1-p)e^{-\theta t}\int_0^t f_1(u)e^{\theta u}du. \end{aligned} \tag{5.116}$$

The Laplace transform is obtained using the convolution property[12] of the transform:

$$\begin{aligned} f^*(s) &= pf_1^*(s) + (1-p)f_1^*(s)f_2^*(s) \\ &= pf_1^*(s) + (1-p)f_1^*(s)\frac{\theta}{s+\theta}. \end{aligned} \tag{5.117}$$

The proportion of interdeparture times with no idleness (p) is equal to the utilization factor (ρ) of the server (Gross & Harris 1985):

$$\rho = \frac{\mu(T_1)}{\mu(T_2)}. \tag{5.118}$$

Cowan (1975) used gamma distributed service times. Branston (1976) applied both gamma and lognormal service time distributions. The models were modified so that p was not equal to the utilization factor, but it was estimated from the data. Cowan called this model M4, while Branston called it the *generalized queuing model*. Figure 5.44 shows the pdf of the model with lognormal follower headways. The parameters follow the approximations given by Branston. Figure 5.44 shows the corresponding hazard function.

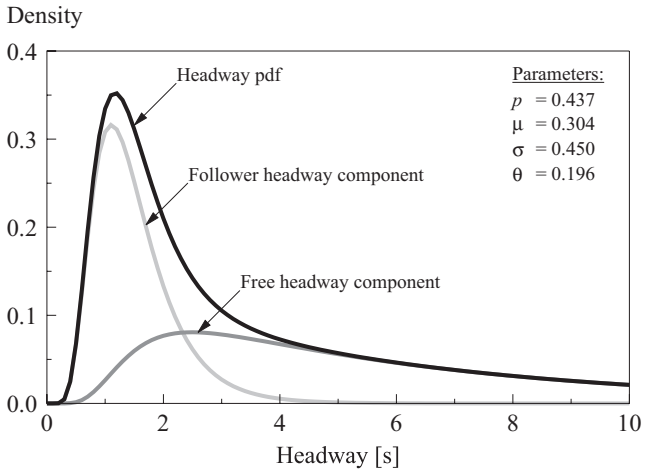

Figure 5.44: Density function of the lognormal generalized queuing model

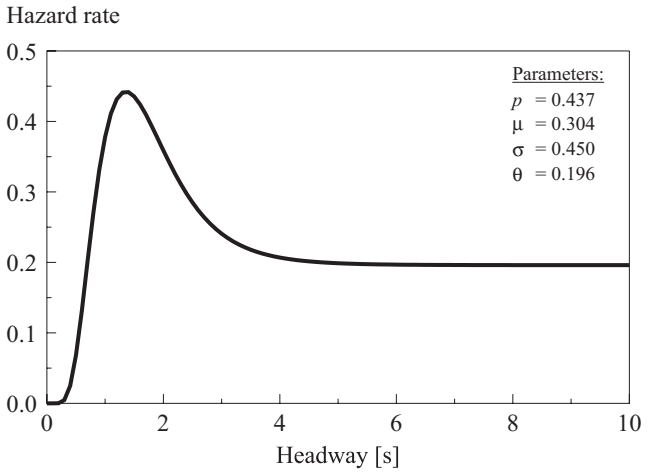

Figure 5.45: Hazard function of the lognormal generalized queuing model

For comparison with the semi-Poisson distribution, some properties of the generalized queuing model with gamma distributed service times are presented. The pdf is:[13]

$$\begin{aligned} f(\,t\,|\,p,\beta,\alpha,\theta\,) &= p\frac{\beta^\alpha t^{\alpha-1}}{\Gamma(\alpha)}e^{-\beta t} + (1-p)\frac{\beta^\alpha}{\Gamma(\alpha)}\theta e^{-\theta t}\int_0^t u^{\alpha-1}e^{-u(\beta-\theta)}du \\ &= p\frac{\beta^\alpha t^{\alpha-1}}{\Gamma(\alpha)}e^{-\beta t} \\ &\quad + (1-p)\left(\frac{\beta}{\beta-\theta}\right)^\alpha \frac{\gamma[\alpha,t(\beta-\theta)]}{\Gamma(\alpha)}\theta e^{-\theta t}, \end{aligned} \tag{5.119}$$

where $\gamma(\cdot)$ is the incomplete gamma function (see table 5.3, page 104). The pdf may also be expressed as:

$$\begin{aligned} f(\,t\,|\,p,\beta,\alpha,\theta\,) &= pf_1(\,t\,|\,\beta,\alpha\,) \\ &\quad + (1-p)\left(\frac{\beta}{\beta-\theta}\right)^\alpha F_1(\,t\,|\,\beta-\theta,\alpha\,)f_2(\,t\,|\,\theta\,). \end{aligned} \tag{5.120}$$

Correspondingly, the PDF can be expressed as:

$$\begin{aligned} F(\,t\,|\,p,\beta,\alpha,\theta\,) &= \frac{\gamma(\alpha,\beta t)}{\Gamma(\alpha)} - (1-p)\left(\frac{\beta}{\beta-\theta}\right)^\alpha \frac{\gamma[\alpha,t(\beta-\theta)]}{\Gamma(\alpha)}e^{-\theta t} \\ &= F_1(\,t\,|\,\beta,\alpha\,) - (1-p)\left(\frac{\beta}{\beta-\theta}\right)^\alpha F_1(\,t\,|\,\beta-\theta,\alpha\,)[1-F_2(\,t\,|\,\theta\,)]. \end{aligned} \tag{5.121}$$

The Laplace transform is:

$$f^*(s) = \left(\frac{\beta}{s+\beta}\right)^\alpha \left(p + (1-p)\frac{\theta}{s+\theta}\right). \tag{5.122}$$

The moments as well as the measures of variation, skewness and kurtosis

[12] Convolution property of the Laplace transform:

$$f^*_{X+Y} = f^*_X f^*_Y.$$

See Kleinrock (1975).

[13] See relation 3.381.1 in Gradshteyn & Ryzhik (1980).

are now easily obtained:

$$\mu(T) = \frac{\alpha}{\beta} + \frac{1-p}{\theta} \tag{5.123}$$

$$\sigma^2(T) = \frac{\alpha}{\beta^2} + \frac{1-p^2}{\theta^2} \tag{5.124}$$

$$C(T) = \frac{\sqrt{\alpha\theta^2 + \beta^2(1-p^2)}}{\alpha\theta + \beta(1-p)} \tag{5.125}$$

$$\alpha_3(T) = \frac{2[\beta^3(1-p^3) + \alpha\theta^3]}{[\beta^2(1-p^2) + \alpha\theta^2]^{3/2}} \tag{5.126}$$

$$\alpha_4(T) = \frac{3\left\{\beta^4\left[4 - (1+p^2)^2\right] + \alpha\theta^2[2\beta^2(1-p^2) + \theta^2(\alpha+2)]\right\}}{[\beta^2(1-p^2) + \alpha\theta^2]^2}. \tag{5.127}$$

Three special cases are worth mentioning:

1. $p = 1$. The model is equal to the gamma distribution [$\mu(T) = \alpha/\beta$].

2. $p = 0$. The headway distribution is the sum of a gamma and an exponential random variate [$\mu(T) = \alpha/\beta + 1/\theta$]. As the mean follower headway (α/β) approaches null, the distribution of all headways approaches exponentiality [$\mu(T) = 1/\theta$].

3. $p = \rho < 1$. The departure rate is equal to the arrival rate [$1/\mu(T) = \theta$]. This is the M/G/1 model.

Generalized queuing model in headway studies

Reasonability According to Cowan (1975) the generalized queuing model describes a situation that an ∞-lane road merges into one lane according to the first-come–first-merge rule such that each driver leaves a randomly chosen tracking headway. The headways at the entry point can be described by the generalized queuing model.

Wasielewski (1979) criticizes the generalized queuing model for lacking in a physical basis. A free headway is a sum of a constrained headway (T_1) and an exponential variate (T_2). But the rationale for selecting exponential distribution for free headways lies in the fact that these vehicles are not influenced by the vehicle ahead.

Wasielewski prefers the semi-Poisson distribution, but as will be shown, the free headway in the semi-Poisson model is also a sum of two random variates. In fact, the two distributions are extremely similar: The generalized

queuing model with gamma distributed follower headways was plotted using the same parameter values as for the semi-Poisson distribution in figure 5.48. The two pdf's were virtually indistinguishable.

Although the generalized queuing model gives separate models for free and follower headways, it is not based on the principles of the traffic flow dynamics. The model is, however, very flexible, or actually a family of models.

Applicability One of the benefits of this model is that it enables the use of queuing theory tools for the M/G/1 systems. The Laplace transform exists, if it exists for the follower headway distribution. In addition, the random number generation is straightforward.

The mathematical tractability of the generalized queuing model depends on the service time distribution. The gamma distribution does not present unsurpassable problems. The lognormal distribution is theoretically interesting as a model for follower headways, but it leads to more complicated formulas. Moreover, it does not have the Laplace transform.

Parameter estimation for the generalized queuing model is problematic. Branston (1976) estimated the parameters by minimizing the chi-square statistic. Cowan (1975) tried the maximum likelihood method, but could not solve the resulting set of four likelihood equations. The method of moments yielded easier equations, but the estimates were infeasible. Finally, Cowan estimated the parameters by minimizing the K-S statistic.

Validity The probability density function (fig. 5.44) and the hazard function (fig. 5.45) have the correct basic shape. The model is flexible enough in terms of the coefficient of variation. Figure 5.46 shows the relation of the kurtosis and the squared skewness for the generalized queuing model with gamma distributed follower headways. The parameters were generated randomly within the limits displayed in the figure. The model can produce distributions such that the kurtosis is only slightly greater than the kurtosis of empirical headway distributions.

Cowan (1975) fitted the generalized queuing model (he called it M4) to a data set of 1,324 successive headways from Mona Valley Road in Sydney. The headways were observed over approximately two hours on the westbound lane of a two lane road. He used gamma distributed service time headways. The minimized K-S statistic was 0.015, which is not significant at 1 % level.

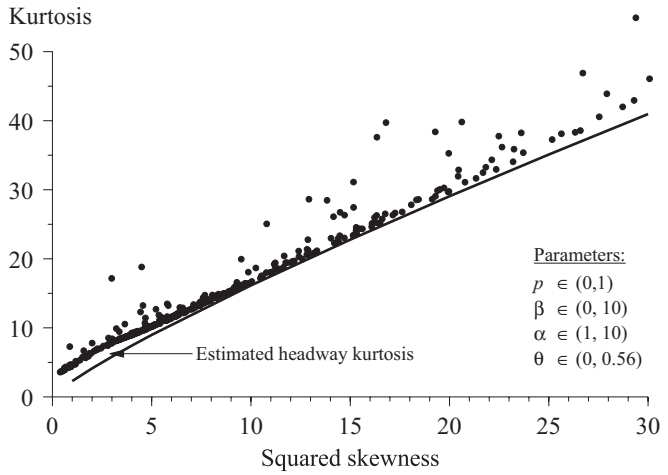

Figure 5.46: KS^2-chart of the generalized queuing model with gamma distributed follower headways (simulated data)

Branston (1976) compared the generalized queuing model and the semi-Poisson distribution. The data were from the fast lane of M4 motorway in England. He reported average probability of the 16 chi-square tests to be 0.196 for the semi-Poisson distribution with gamma distributed zone of emptiness. For the generalized queuing model with gamma and lognormal service time distributions Branston reported average probabilities of 0.193 and 0.359, respectively. The combined probabilities are, however, 0.0004, 0.0002, and 0.0985, respectively. For the slow lane, the generalized queuing model did not give substantially better fit.

5.3.8 Semi-Poisson distribution

Properties of the semi-Poisson distribution

The semi-Poisson model of Buckley(1962, 1968) is based on the conjecture that in front of each vehicle there is a zone of emptiness, which describes the fluctuations in the car following (Wasielewski 1979), and has probability density $f_1(t)$. The exponential free headway distribution is modified so that it includes only headways greater than a random variate (Z) sampled from $f_1(t)$. Figure 5.47 shows the semi-Poisson model as a queuing system, where M' stands for a modified exponential service time distribution.

The free headway PDF is:

$$F_2(t) = \mathbb{P}\{\, Y \leq t \,|\, Y > Z \,\}\,, \tag{5.128}$$

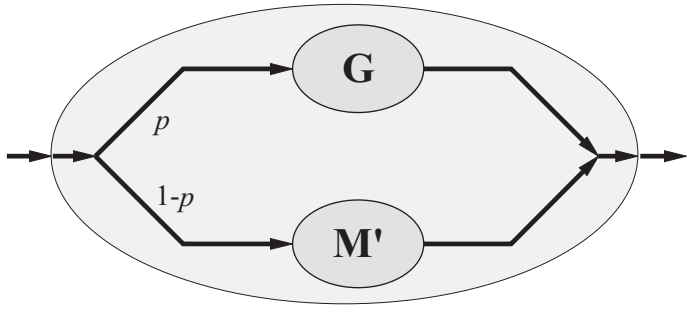

Figure 5.47: Queuing model for the semi-Poisson distribution

where Y is a random variate from the negative exponential distribution. Using the definition of the conditional probability, the PDF can be expressed as:

$$\begin{aligned} F_2(t) &= \frac{\mathbb{P}\{Y \leq t, Z < Y\}}{\mathbb{P}\{Z < Y\}} \\ &= \frac{F_{YZ}(t,Y)}{F_1(Y)}, \end{aligned} \tag{5.129}$$

where $F_1(Y)$ is the probability that an exponential random variate (Y) is greater than a random variate from the follower headway distribution. $F_{YZ}(t,Y)$ is the probability that t is greater than or equal to an exponential random variate Y, which, in turn, is greater than a random variate Z from the follower distribution.

For a fixed value $Y = y$, the probability $\mathbb{P}\{Z < Y\}$ is the product $\mathbb{P}\{Y = y\}\,\mathbb{P}\{Z < y\}$. If Y is exponentially distributed with pdf $g(\cdot)$, the probability $F_{YZ}(t,Y) = \mathbb{P}\{Y \leq t, Z < Y\}$ can be expressed by integrating from null to t:

$$\begin{aligned} F_{YZ}(t,Y) &= \int_0^t g(y) \int_0^y f_1(z)\, dz\, dy \\ &= \int_0^t g(y) F_1(y)\, dy. \end{aligned} \tag{5.130}$$

The pdf is obtained by differentiation:

$$\begin{aligned} f_{YZ}(t,Y) &= \frac{d}{dt} F_{YZ}(t,Y) \\ &= g(t) F_1(t). \end{aligned} \tag{5.131}$$

Consequently, the pdf for free headways is:

$$\begin{aligned} f_2(t) &= \frac{f_{YZ}(t,Y)}{F_1(Y)} \\ &= \frac{F_1(t)}{F_1(Y)} g(t). \end{aligned} \tag{5.132}$$

$F_1(Y)$ is the probability that an exponential random variate (Y) is greater than a random variate (Z) from the follower distribution. For a fixed value $Y = y$, the probability $\mathbb{P}\{Y > Z\}$ is equal to the product $\mathbb{P}\{Y = y\}\,\mathbb{P}\{Z < y\}$. If Y is exponentially distributed with scale parameter θ, the probability $F_1(Y)$ can be expressed by integrating from null to infinity:

$$F_1(Y) = \int_0^\infty \theta e^{-\theta y} \int_0^y f_1(z)\, dz\, dy. \tag{5.133}$$

Using the relation 29.2.6 of Abramowitz & Stegun (1972), this probability is the same as the Laplace transform of the follower distribution (Buckley 1968):

$$F_1(Y) = f_1^*(\theta). \tag{5.134}$$

This leads to the equation for the free headway pdf:

$$f_2(t) = \frac{F_1(t)}{f_1^*(\theta)} g(t). \tag{5.135}$$

The Laplace transform can be derived using the same property 29.2.6 of Abramowitz & Stegun (1972):

$$\begin{aligned} f_2^*(t) &= \frac{\theta}{f_1^*(\theta)} \int_0^\infty e^{-(s+\theta)t} \int_0^t f_1(x)\, dx \\ &= \frac{\theta}{s+\theta} \cdot \frac{f_1^*(s+\theta)}{f_1^*(\theta)}. \end{aligned} \tag{5.136}$$

As Buckley (1968) has observed, the first term is the Laplace transform of the exponential distribution $g(t) = \theta e^{-\theta t}$. Accordingly, the Laplace transform of the free headway distribution is the convolution of two distributions, namely $g(t)$ and, say, $c(t)$. Because $c(t)$ has the Laplace transform:

$$c^*(s) = \frac{f_1^*(s+\theta)}{f_1^*(\theta)}, \tag{5.137}$$

it can be derived by the inverse transform (Buckley 1968):

$$c(t) = \frac{f_1(t)}{f_1^*(\theta)} e^{-\theta t}. \tag{5.138}$$

A free headway is thus a sum of two random variates, which follow the distributions $g(t)$ and $c(t)$.

The pdf for the semi-Poisson distribution is (Buckley 1968):

$$f(t) = pf_1(t) + (1-p)\frac{F_1(t)}{f_1^*(\theta)}\theta e^{-\theta t}, \qquad t \geq 0,\ \theta > 0,\ 0 \leq p \leq 1. \tag{5.139}$$

The PDF is obtained by integration:

$$F(t) = pF_1(t) + \frac{1-p}{f_1^*(\theta)} \int_0^t F_1(z)\theta e^{-\theta z}\, dz, \qquad t \geq 0,\ \theta > 0,\ 0 \leq p \leq 1. \tag{5.140}$$

The distribution has the Laplace transform (Buckley 1968):

$$f^*(s) = pf_1^*(t) + (1-p)\frac{\theta f_1^*(s+\theta)}{(s+\theta)f_1^*(\theta)}. \tag{5.141}$$

Follower headway distribution

Buckley (1962) suggested the normal distribution as a model for followers' headways. He found that the error due to the non-zero probability of negative headways was very small. Later Buckley (1968) applied gamma and truncated normal distributions. He observed that gamma distribution has the general shape which is assumed to be correct for follower headways. On the other hand, he assumed that the truncated normal distribution could describe the error in the tracking task, when a driver attempts to maintain a desired spacing. With gamma distributed follower headways Buckley could not obtain solutions for the parameters at low and low-medium flows. When the parameters for both models were obtained, the model with gamma distributed follower headways gave similar or better goodness of fit than the truncated normal distribution.

Ashton (1971) saw no strong theoretical justification for choosing the normal distribution. He chose gamma distribution because of its connection with the exponential distribution. Such a model should be theoretically more tractable. In the data analysis Ashton, however, used exponential zone of emptiness.

Branston (1976) applied the model with the normal and the gamma distributions for the follower headways. He estimated the parameters by combining the method of moments for p and θ with the minimum chi-square method for β and α. He did not have similar difficulties as Buckley (1968). The goodness-of-fit was not acceptable, when the the normal distribution was used for the follower headways. For gamma distributed follower headways he reported an acceptable goodness of fit.

Wasielewski (1974, 1979) calculated the distribution for the follower headways directly from the empirical headway distribution without introducing a parametric form. This nonparametric approach is not followed here, because in theoretical work and simulation a parametric model is preferable.

To sum up, normal, truncated normal, exponential, and gamma distributions have been suggested as models for the follower headways. The exponential distribution can describe the traffic flow with no vehicle interaction, but it is unsuitable as a follower headway model. The studies of Buckley and Branston suggest that the gamma distribution gives better results than the normal and the truncated normal distributions. Lognormal distribution would be theoretically an interesting alternative, but its usefulness is limited, because it does not have a Laplace transform. Consequently, the *gamma distribution* was selected as the follower headway model.

The pdf (fig. 5.48) of the semi-Poisson distribution with gamma distributed follower headways is:

$$f(\,t\,|\,p,\beta,\alpha,\theta\,) = p\frac{(\beta t)^{\alpha-1}}{\Gamma(\alpha)}\beta e^{-\beta t} + (1-p)\frac{\gamma(\alpha,\beta t)}{\Gamma(\alpha)}\left(1+\frac{\theta}{\beta}\right)^{\alpha}\theta e^{-\theta t},$$
$$t>0;\ \beta,\alpha,\theta\geq 0;\ 0\leq p\leq 1, \quad (5.142)$$

where $\gamma(\cdot)$ is the incomplete gamma function (see table 5.3, page 104). The PDF is:

$$F(\,t\,|\,p,\beta,\alpha,\theta\,) = p\frac{\gamma(\alpha,\beta t)}{\Gamma(\alpha)} + (1-p)\frac{\theta}{\Gamma(\alpha)}\left(1+\frac{\theta}{\beta}\right)^{\alpha}\int_0^t \gamma(\alpha,\beta u)e^{-\theta u}du,$$
$$t>0;\ \beta,\alpha,\theta\geq 0;\ 0\leq p\leq 1. \quad (5.143)$$

Figure 5.49 shows the estimated semi-Poisson hazard function for sample 4.

The Laplace transform of the semi-Poisson distribution is:

$$f^*(s) = p\left(\frac{\beta}{\beta+s}\right)^{\alpha} + (1-p)\frac{\theta}{\theta+s}\left(\frac{\beta+\theta}{\beta+\theta+s}\right)^{\alpha}. \quad (5.144)$$

It is obvious that a free headway is a sum of two random variates: a variate from the exponential distribution and a variate from the gamma distribution

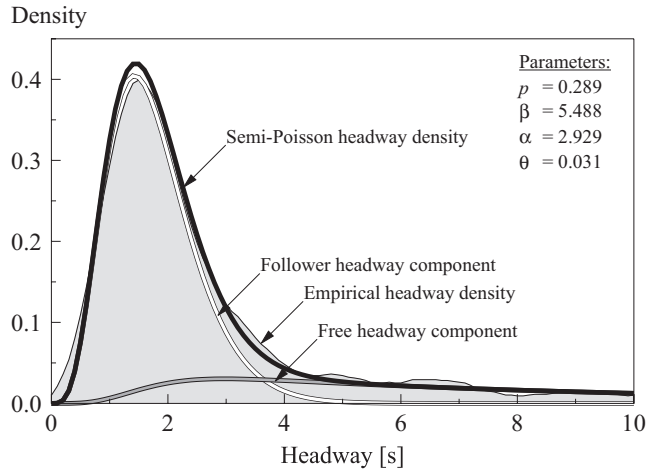

Figure 5.48: Semi-Poisson density function for sample 4

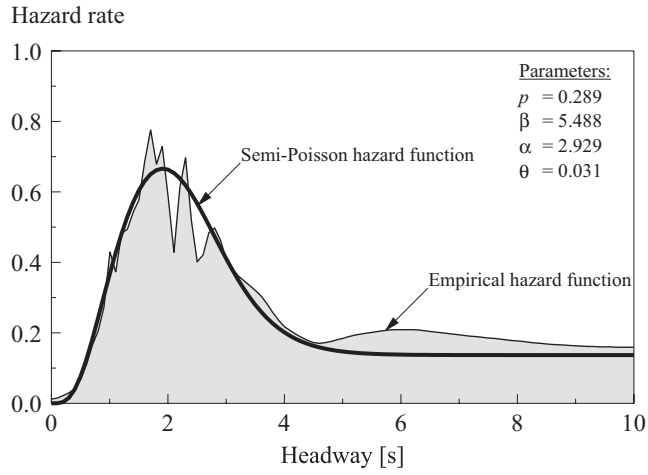

Figure 5.49: Semi-Poisson hazard function for sample 4

with scale parameter equal to the sum $(\beta + \theta)$ of the scale parameters from the follower gamma distribution and the exponential distribution. This result can be used in the generation of semi-Poisson pseudo-random numbers.

The moments of the distribution can be obtained using the Laplace transform:

$$E(T^r) = (-1)^r \frac{d^r f^*(s)}{ds^r}\bigg|_{s=0}. \tag{5.145}$$

The first four moments are:

$$E(T) = \mu(T) = p\frac{\alpha}{\beta} + (1-p)\left(\frac{1}{\theta} + \frac{\alpha}{\beta+\theta}\right) \tag{5.146}$$

$$E(T^2) = p\frac{\alpha(\alpha+1)}{\beta^2} + (1-p)\left(\frac{2}{\theta^2} + \frac{2\alpha}{\theta(\beta+\theta)} + \frac{\alpha(\alpha+1)}{(\beta+\theta)^2}\right) \tag{5.147}$$

$$\begin{aligned} E(T^3) = {} & p\frac{\alpha(\alpha+1)(\alpha+2)}{\beta^3} \\ & + (1-p)\left(\frac{6}{\theta^3} + \frac{6\alpha}{\theta^2(\beta+\theta)} + \frac{3\alpha(\alpha+1)}{\theta(\beta+\theta)^2} + \frac{\alpha(\alpha+1)(\alpha+2)}{(\beta+\theta)^3}\right) \end{aligned} \tag{5.148}$$

$$\begin{aligned} E(T^4) = {} & p\frac{\alpha(\alpha+1)(\alpha+2)(\alpha+3)}{\beta^4} \\ & + (1-p)\bigg(\frac{24}{\theta^4} + \frac{24\alpha}{\theta^3(\beta+\theta)} + \frac{12\alpha(\alpha+1)}{\theta^2(\beta+\theta)^2} \\ & + \frac{4\alpha(\alpha+1)(\alpha+2)}{\theta(\beta+\theta)^3} + \frac{\alpha(\alpha+1)(\alpha+2)(\alpha+3)}{(\beta+\theta)^4}\bigg). \end{aligned} \tag{5.149}$$

The moments can also be expressed recursively:

$$E(T^r) = p\frac{A_r}{\beta^r} + (1-p)B_r, \qquad r = 0,1,2,\ldots, \tag{5.150}$$

where

$$A_0 = B_0 = 1 \tag{5.150a}$$

$$A_r = (\alpha + r - 1)A_{r-1} \tag{5.150b}$$

$$B_r = \frac{r}{\theta}B_{r-1} + \frac{A_r}{(\beta+\theta)^r}. \tag{5.150c}$$

The central moments can be expressed in terms of the noncentral moments. The rth central moment about the mean is (Stuart & Ord 1987):

$$E\left\{[T-\mu(T)]^r\right\} = \sum_{j=0}^{r}\binom{r}{j}E(T^{r-j})[-\mu(T)]^j. \tag{5.151}$$

Thus, the measures of dispersion can be calculated using the following equations:

$$\sigma^2(T) = E\left\{[T - \mu(T)]^2\right\} \tag{5.152}$$

$$\alpha_3(T) = \frac{E\left\{[T - \mu(T)]^3\right\}}{\sigma^3(T)} \tag{5.153}$$

$$\alpha_4(T) = \frac{E\left\{[T - \mu(T)]^4\right\}}{\sigma^4(T)}. \tag{5.154}$$

Concerning the four-stage identification process, the following observations can be made:

1. The pdf of the semi-Poisson distribution (figs. 5.48 and 5.50) is very similar to the empirical density function of the headway data.

2. The hazard rate (fig. 5.49) has the same properties as the empirical hazard function. It starts from zero, rises to the maximum, falls, and reaches a constant level.

3. In terms of the coefficient of variation the semi-Poisson distribution is very flexible.

4. Figure 5.51 shows the kurtosis of the semi-Poisson distribution against the squared skewness. The parameters were generated randomly within the limits displayed in the figure. The model can produce samples such that the kurtosis corresponding to the squared skewness is only slightly greater than the kurtosis of headway samples. When the distribution approaches symmetry ($\alpha_3 \to 0$), the kurtosis approaches the values corresponding to the gamma distribution (see fig. 4.15 on page 73). This is to be expected considering that $\alpha_4 \to 3$ (see table 5.3 on page 104), and the distribution approaches normality as p approaches unity, and α tends to infinity (Johnson et al. 1994). The figure is very similar to the corresponding figure (5.46) of the generalized queuing model.

The identification process indicates that the properties of the semi-Poisson distribution are very similar with the headway data. The KS^2-chart displays most clearly the slight differences between the semi-Poisson distribution and the headway data.

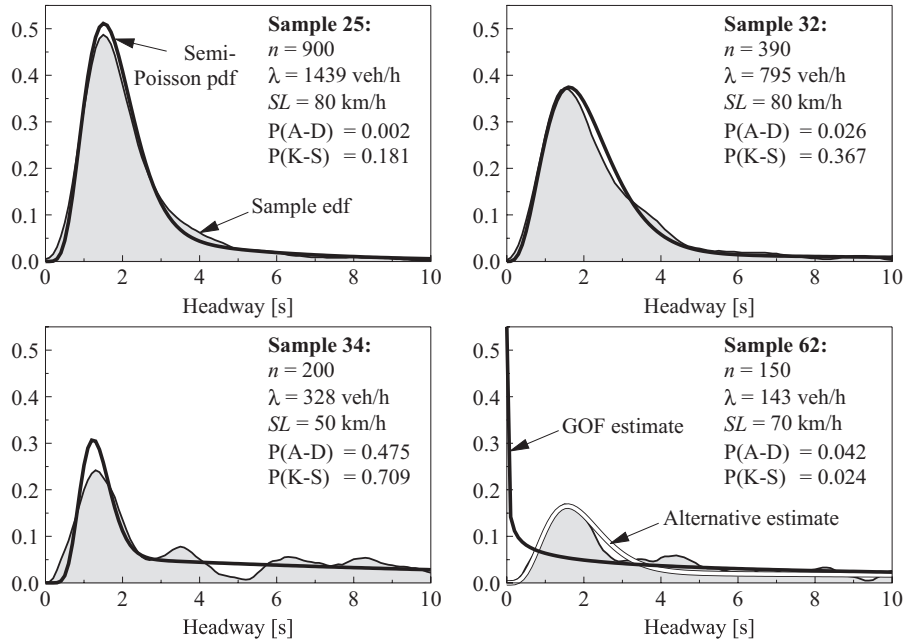

Figure 5.50: Semi-Poisson pdf and sample edf. In sample 62 the "GOF estimate" shows the estimated pdf that was used in the goodness-of-fit tests. The "alternative estimate" shows a better estimate, which was found using another starting point to find the maximum of the loglikelihood function.

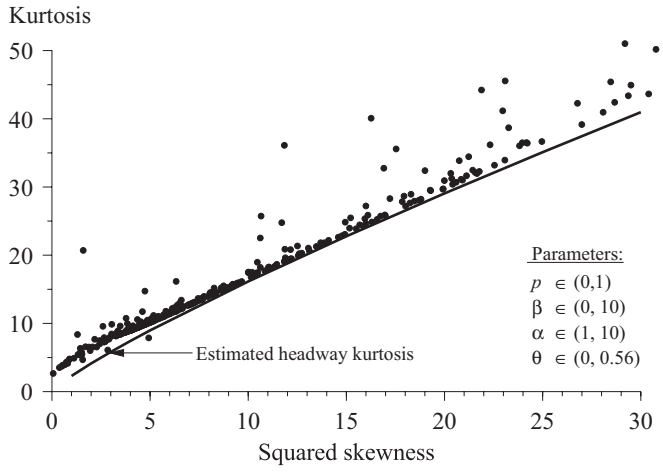

Figure 5.51: KS^2-chart of the semi-Poisson distribution with gamma distributed follower headways (simulated data)

Parameter estimation

Method of moments estimators According to Greenwood & Durand (1960) the method of moments is inefficient, except for distributions closely resembling the normal. In this case the gamma distribution for follower headways is expected to have a large shape parameter, which gives the distribution the bell-like shape typical of the normal distribution. On the other hand, the method of moments gives good parameter estimates for the negative exponential distribution also. Consequently, there are grounds for using the method of moments to estimate the parameters of the gamma distribution.

The MME's are obtained by equating the first four moments to the sample moments and solving simultaneously the parameter values. Buckley (1968) used this method to estimate the parameters. Because the simultaneous equations were not explicitly solvable, he used numerical methods. He found no physically sensible solutions for flows less than 21 vehicles per minute (1,260 veh/h). The complicated form of the equations anticipate severe difficulties in finding feasible parameter estimates.

Modified method of moments estimators To make the estimators tractable, and to avoid the use of higher moments the method of moments can be modified: Practically all headways in the exponential tail ($t_i > \delta$) belong to the free headway distribution. The PDF of the follower headways is then:

$$F_1(t) = 1 - \epsilon(t), \qquad \text{when } t > \delta, \tag{5.155}$$

where $\epsilon(t)$ is small. Now, p and θ can be estimated using the properties of the tail of the headway distribution, assuming that there are enough observations in the tail .

Let us assume a sample of size n with m observations greater than δ. The scale parameter estimate $\tilde{\theta}$ is the inverse of the modified sample mean of the tail (Wasielewski 1974, Wasielewski 1979):

$$\tilde{\theta} = \left(\frac{1}{m} \sum_{j=n-m+1}^{n} t_{(j)} - \delta \right)^{-1}, \tag{5.156}$$

where $t_{(j)}$ is the jth order statistic of the sample.

The probability that a headway belongs to the tail of the distribution is:

$$\begin{aligned}\mathbb{P}\{T > \delta\} &= \int_{\delta}^{\infty} f(t)\, dt \\ &= p \int_{\delta}^{\infty} f_1(t)\, dt + \frac{(1-p)}{f_1^*(\theta)} \int_{\delta}^{\infty} F_1(t)\theta e^{-\theta t} dt \\ &= p[1 - F_1(\delta)] + \frac{(1-p)}{f_1^*(\theta)} \int_{\delta}^{\infty} F_1(t)\theta e^{-\theta t} dt.\end{aligned} \tag{5.157}$$

If $t > \delta$, the follower headway PDF is near unity ($F_1(t) \approx 1$), and the equation simplifies to

$$\begin{aligned}\mathbb{P}\{T > \delta\} &= \frac{1-p}{f_1^*(\theta)} e^{-\theta\delta} \\ &= (1-p)\left(1 + \frac{\theta}{\beta}\right)^{\alpha} e^{-\theta\delta}\end{aligned} \tag{5.158}$$

The proportion of headways in the tail can be used as an estimate for the probability $\mathbb{P}\{T > \delta\}$. This leads to the following estimator for p:

$$\tilde{p} = 1 - \frac{m}{n}\left(1 + \frac{\tilde{\theta}}{\tilde{\beta}}\right)^{-\tilde{\alpha}} e^{\tilde{\theta}\delta}. \tag{5.159}$$

The system of equations can be completed by equating the first two moments to the corresponding sample moments:

$$m_1' = \bar{t} = \tilde{p}\frac{\tilde{\alpha}}{\tilde{\beta}} + (1-\tilde{p})\left(\frac{1}{\tilde{\theta}} + \frac{\tilde{\alpha}}{\tilde{\beta}+\tilde{\theta}}\right) \tag{5.160}$$

$$m_2' = \tilde{p}\frac{\tilde{\alpha}(\tilde{\alpha}+1)}{\tilde{\beta}^2} + (1-\tilde{p})\left(\frac{2}{\tilde{\theta}^2} + \frac{2\tilde{\alpha}}{\tilde{\theta}+(\tilde{\beta}+\tilde{\theta})} + \frac{\tilde{\alpha}(\tilde{\alpha}+1)}{(\tilde{\beta}+\tilde{\theta})^2}\right). \tag{5.161}$$

Branston (1976) estimated θ and p from these two equations. He then estimated β and α by minimizing the chi-square statistic.

The equation for the first moment can be solved to obtain the estimator for the shape parameter:

$$\tilde{\alpha} = \frac{\tilde{\beta}(\tilde{\beta}+\tilde{\theta})[m_1'\tilde{\theta} - (1-\tilde{p})]}{\tilde{\theta}(\tilde{\beta}+\tilde{p}\tilde{\theta})}. \tag{5.162}$$

The he scale parameter estimator is the solution to the polynomial equation, which is obtained by substituting for $\tilde{\alpha}$ in the equation for the second

moment:

$$\begin{aligned}
&[(m'^2_1 - m_2)\tilde{\theta}^2 + 1 - \tilde{p}^2]\tilde{\beta}^4 \\
&\quad + \{(1-2\tilde{p})(m'^2_1 - m_2)\tilde{\theta}^2 + [1 - 2\tilde{p}(1-\tilde{p})]m'_1\tilde{\theta} + 5\tilde{p}(1-\tilde{p})\}\tilde{\theta}\tilde{\beta}^3 \\
&\quad + \{[3m'^2_1 - (2+\tilde{p})m'_2]\tilde{\theta}^2 + (4\tilde{p}-1)m'_1\tilde{\theta} + 2 - \tilde{p}(1-\tilde{p})\}\tilde{p}\tilde{\theta}^2\tilde{\beta}^2 \\
&\quad + [(m'^2_1 - m'_2\tilde{p})\tilde{\theta}^2 + (4\tilde{p}-1)m'_1\tilde{\theta} - \tilde{p}(1-\tilde{p})]\tilde{p}\tilde{\theta}^3\tilde{\beta} \\
&\quad + [m'_1\tilde{\theta} - (1-\tilde{p})]\tilde{p}^2\tilde{\theta}^4 = 0. \quad (5.163)
\end{aligned}$$

This equation has two unknown parameter estimates ($\tilde{p}$ and $\tilde{\beta}$). Because this is a fourth degree polynomial, its roots can be solved, assuming either $\tilde{p}$ or $\tilde{\beta}$ known. There is an analytic solution to the equation with hundreds of terms (obtained by the Maple computer algebra system), but it was not used, because there are well tested, fast and accurate numerical methods for finding the roots of a polynomial.

Assuming $\tilde{p}$ known, $\tilde{\alpha}$ and $\tilde{\beta}$ can be solved using equations (5.162) and (5.163). The estimate is accepted, if the absolute error in the estimate of p is smaller than the criterion of precision:

$$\left| 1 - \frac{m}{n}\left(1 + \frac{\tilde{\theta}}{\tilde{\beta}}\right)^{-\tilde{\alpha}} e^{\tilde{\theta}\delta} - \tilde{p} \right| < \epsilon. \quad (5.164)$$

The problem with this method is that $\tilde{\beta}$ is not a real number for all values of $\tilde{p}$. Unfortunately, this often occurs with such values of $\tilde{p}$ that would result in smallest estimation errors. Consequently, the modified method of moments was rejected.

Maximum likelihood estimators The likelihood function of the semi-Poisson distribution with gamma distributed follower headways is:

$$\begin{aligned}
L(p,\beta,\alpha,\theta) &= \prod_{j=1}^{n} f(\, t_j \,|\, p,\beta,\alpha,\theta\,) \\
&= \left(\frac{1}{\Gamma(\alpha)}\right)^n \prod_{j=1}^{n} \Big[p(\beta t_j)^{\alpha-1}\beta e^{-\beta t_j} \\
&\qquad + (1-p)\gamma(\alpha,\beta t_j)\left(1+\frac{\theta}{\beta}\right)^{\alpha}\theta e^{-\theta t_j}\Big]. \quad (5.165)
\end{aligned}$$

The log-likelihood function is:

$$\ln L(p,\beta,\alpha,\theta) = -n\ln\Gamma(\alpha) + \sum_{j=1}^{n} \ln\Bigg[p(\beta t_j)^{\alpha-1}\beta e^{-\beta t_j} + (1-p)\gamma(\alpha,\beta t_j)\left(1+\frac{\theta}{\beta}\right)^{\alpha}\theta e^{-\theta t_j}\Bigg]. \quad (5.166)$$

The maximum likelihood estimators are obtained by equating to zero the gradient of the log-likelihood function and solving the resulting set of equations. This must be done by numerical methods.

A modified method was also tested, in which the exponential scale parameter was estimated from the tail using equation (5.156). This method decreased the dimensionality of the problem, but there was no significant increase in speed and the results were less accurate. Accordingly, the maximum likelihood method was used without any modifications.

In some cases, the results were sensitive to the starting point (see fig. 5.50).[14] Because the calculations were very time consuming, only one starting point was tried. The parameter estimation of sample 59 failed. In five samples (26, 38, 45, 46, and 62) $\hat{p}$ was either unity or null.

Figures 5.52 and 5.53 show the parameter estimates for the headway data, excluding the six samples listed above. On low speed roads there is significant correlation between $\hat{p}$ and volume ($r = 0.734$, $P < 0.0005$), and between $\hat{\theta}$ and volume ($r = 0.920$, $P < 0.0005$). That is, the proportion of follower headways increases with increasing traffic volume. Also, the proportion of short free headways increases. On high speed roads the correlation between $\hat{\theta}$ and the volume is also strong ($r = 0.810$, $P < 0.0005$). For $\hat{p}$ the correlation is not as strong, but it is still statistically significant ($r = 0.511$, $P = 0.004$).

The parameters of the gamma distribution do not correlate with traffic volume, but they have a strong mutual correlation (figure 5.54). The ratio $\hat{\beta}/\hat{\alpha}$ is relatively constant, being 0.60 on low speed roads and 0.59 on high speed roads. The difference is not statistically significant ($P = 0.62$), but the follower flow rate is about 2,140 veh/h on both low speed and high speed roads. The estimate does not change with the overall traffic flow rate. The follower flow rate is near the capacity estimate of 2,200 person cars per hour per lane on multilane highways (HCM 1994).

[14] Branston (1976) also encountered instability in the parameter estimates.

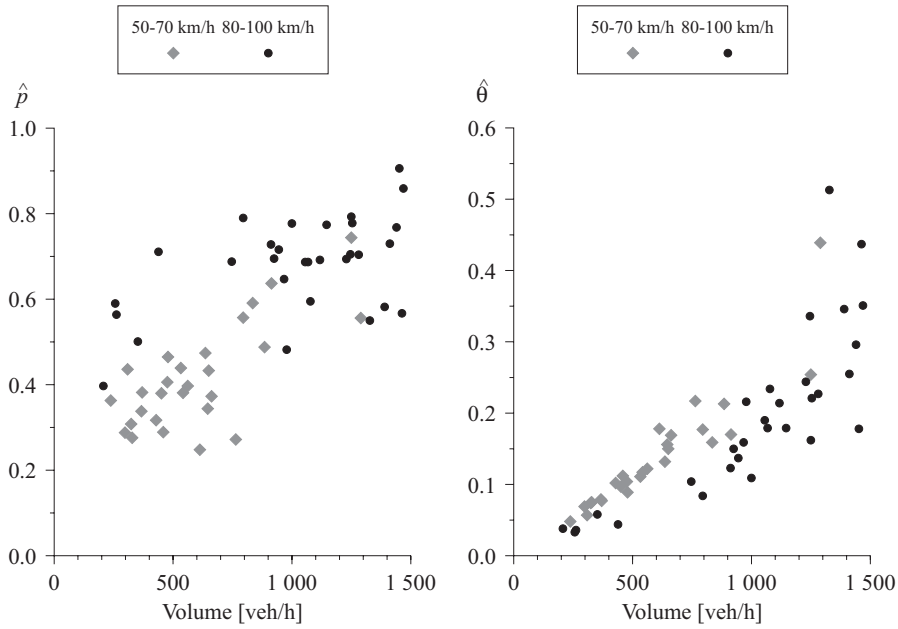

Figure 5.52: Maximum likelihood estimates for p and θ

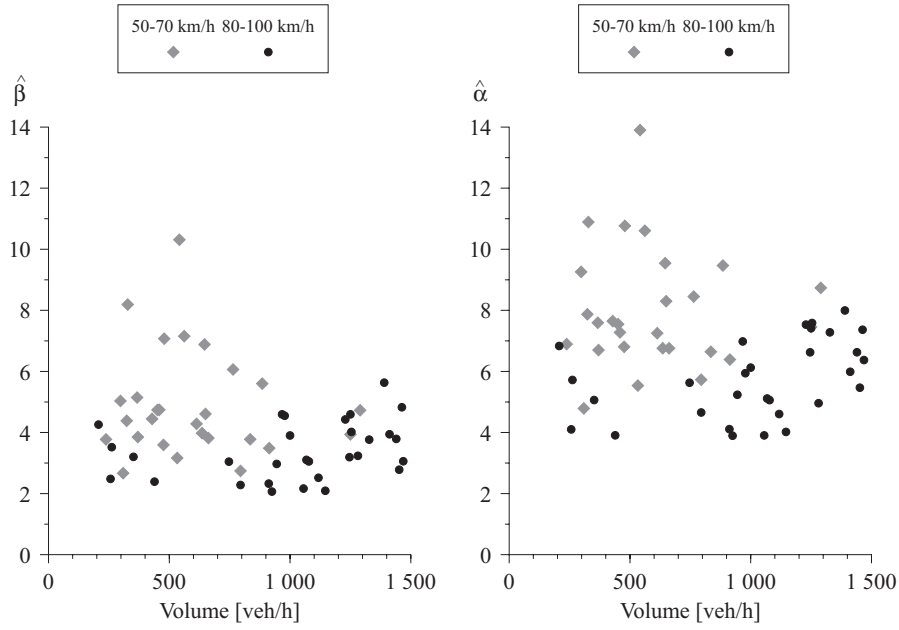

Figure 5.53: Maximum likelihood estimates for β and α

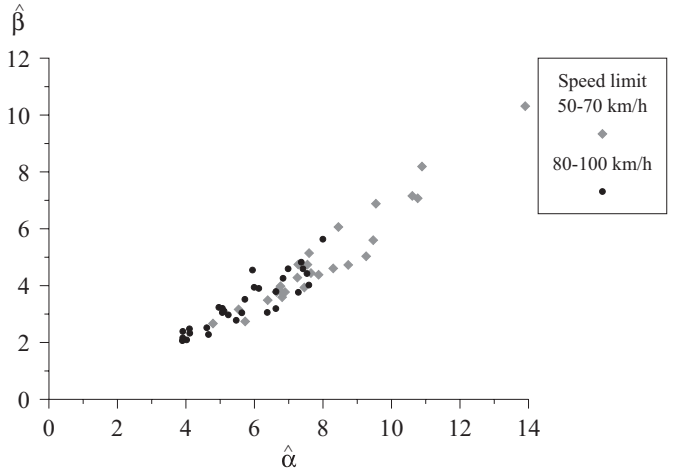

Figure 5.54: $\hat{\alpha}$ against $\hat{\beta}$

Goodness-of-fit tests

The parameters were estimated using the maximum likelihood method. Sample 59 was excluded, because its parameter estimation failed. The goodness of fit was tested using both the nonparametric K-S test and the parametric A-D test. The parametric test was conducted using the Monte Carlo method described in section 2.2.4.

For each of the 62 samples, 500 replications were generated and analyzed. The number of replications was low, because the parameter estimation was very time consuming.[15] If the parameter estimation of a replication failed, a new replication was generated. If more than 25 (5 %) failures were recorded, the goodness-of-fit test was discarded. Four samples (38, 56, 61, and 64) failed the test procedure.

Figure 5.55 shows the goodness-of-fit test results for the semi-Poisson distribution with gamma distributed zone of emptiness. There is an interesting difference in the figures: The significance probabilities of the K-S tests are mostly between 0.2 and 0.8, while the A-D tests produce significance probabilities which are mostly either below 0.3 or above 0.6. The comparison of the results (fig. 5.56) shows that the 5 % significance level of the parametric A-D test is about the same as the 40 % significance level of the nonparametric K-S test. Consequently, the semi-Poisson distribution

[15]It took three weeks to perform the goodness-of-fit tests on three i386/i486 personal computers.

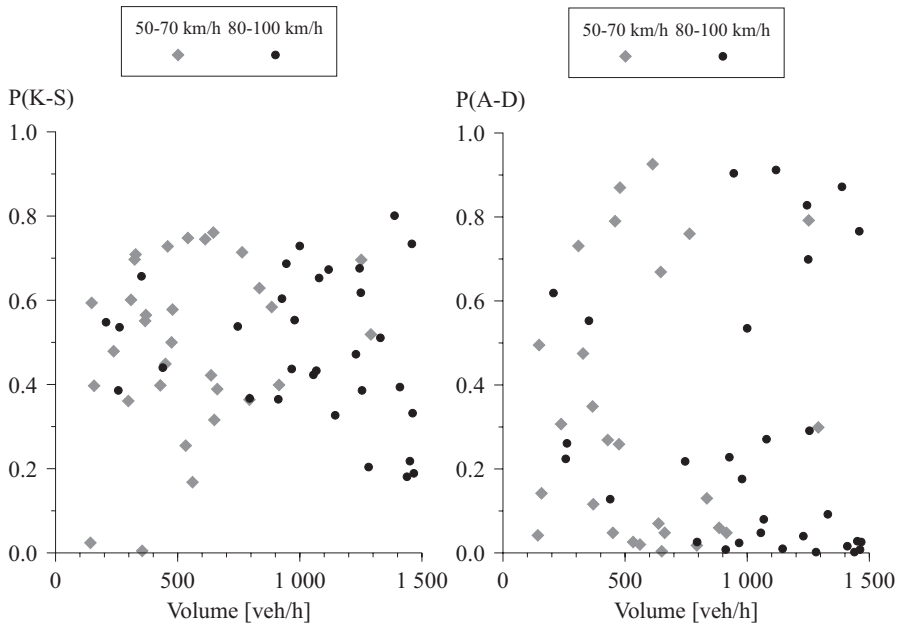

Figure 5.55: Goodness-of-fit tests for the semi-Poisson distribution

fits the data extremely well, but it appears to be caused more by the flexibility of the model than by its inherently correct shape. As the flexibility of the model increases, the bias in the nonparametric EDF tests also increases.

The combined probability of the nonparametric K-S tests is 0.85. The parametric A-D tests, however, give combined probability of only $4.3 \cdot 10^{-11}$. The fit is best, when the data have large sample variance. When the variance is larger than 100, the significance probability of the sample is greater than 0.1. Only sample 62 is an exception, but it is probably caused by the unsuccessful parameter estimation. On the other hand, all the eight samples of at least 500 observations have significance probability less than 0.1. These samples are from high speed roads. Accordingly, the semi-Poisson model fits best small samples with large variances.[16] In such cases the "noise" in the data hides small bias in the model.

Semi-Poisson distribution in headway studies

Reasonability The modification of the free headway distribution $f_2(t)$ in the semi-Poisson model can be criticized: The condition that a free head-

[16]The data do not contain large (> 500) samples with large variances (> 100).

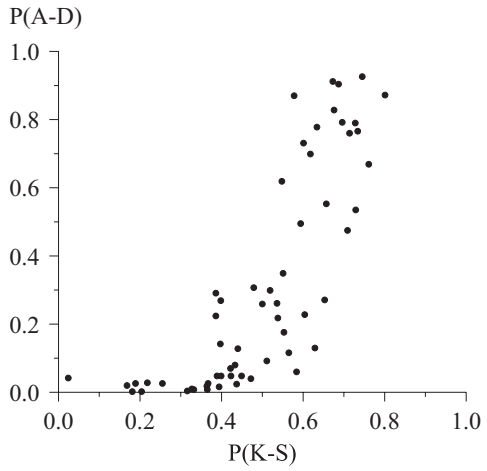

Figure 5.56: Comparison of the nonparametric K-S and the parametric A-D tests for the semi-Poisson distribution

way should be greater than a random variate from the constrained headway distribution $f_1(t)$ seems arbitrary, because the shape of the constrained headway distribution is not necessarily the way that the frequency of free headways decreases near the origin. The process of becoming a follower is not equal to adjusting the headway while being a follower.

Branston (1976) describes the difference between the semi-Poisson model and the M/G/1 queuing model as follows:

> In the semi-Poisson model each free headway is obtained by comparing an exponential headway with a following headway, while in the queuing model each free headway is obtained by adding an exponential gap to a following headway.

This description overstates the difference between these distributions. The nonfollower headway variate in the semi-Poisson distribution is also a sum of an exponential random variate and another random variate (see equation 5.138). In fact, the M/G/1 model and the semi-Poisson distribution do not differ significantly when realistic parameter values are used (Botma & Fi 1991).

Applicability Parameter estimation is time consuming, but semi-Poisson pseudorandom numbers can be generated efficiently (Buckley 1962). The pdf is rather complicated, and the PDF must be solved numerically. Now that numerical integration is widely available in personal computers and

even in pocket calculators, there are better potentials for the application of the semi-Poisson distribution. The Laplace transform exists, if it exists for the follower headway distribution.

Validity The chi-square tests of Buckley (1968) gave all "good" results. Branston (1976) reported an average probability of 0.196 for 16 chi-square tests (M4 fast lane). The combined probability, however, is only 0.0004. For normal zone of emptiness the results were even worse.

The semi-Poisson distribution yields a better fit than any of the simple distributions. If the K-S tests give as good results as shown above, but the hypothesis is still rejected, the distribution is likely to be so flexible that an extremely good fit is required. So good a fit may be impossible for headway data, where some disturbances cannot always be avoided.

Chapter 6

Conclusions and recommendations

The main purpose of the study was to present methods in order to improve the statistical analysis of vehicle headways. A more firm scientific foundation has been laid for further studies. Also, some new information about the properties of vehicle headways on two-lane two-way roads has been obtained.

The *combined probabilities method* of Fisher (1938) provided a scientifically sound method to test a hypothesis on a multi-sample data set. The method was used extensively throughout the research. It was also demonstrated that in many earlier headway studies this method would have been of great help—either further confirming the results or guarding from false conclusions. The combined probabilities method should be routinely used in the tests for multi-sample data.

The combined probabilities method forms the basis of the *moving probability method* (Luttinen 1992), which appeared to be very helpful in demonstrating the variation of test results over traffic volumes. This new method provided results not directly available with the traditional methods. It can be assumed that the method will have much wider applications in statistics. However, the experience with the method is still limited, and further research is needed.

The first step in the empirical research (fig. 1.1, page 20) is the *data collection*. A major problem in the collection of headway data is to obtain stationary data. It was shown that the combination of small samples with nearly equal means produces a biased data set. This bias can be avoided, if the data collection is based on trend analysis. Tests against trend should become a standard procedure in headway studies.

In model identification it is important to find a few methods that show the most important properties of the data. For the identification of headway models a *four-stage identification process* was proposed. It includes the estimation of:

1. Probability density function
2. Hazard function
3. Coefficient of variation (CV-chart)
4. Kurtosis and squared skewness (KS^2-chart).

The probability density function displays the shape of the distribution better than the probability distribution function. The hazard function is mostly used in life and reliability data analysis, but it also very informative in headway studies. The kernel method was shown to produce smooth estimates for density and hazard functions. It should be preferred over the more traditional histogram method.

The coefficient of variation is a scaled measure of the variation in headways. It is a useful measure of dispersion in the headway distribution. The correlation between the kurtosis and the squared skewness was very high in the headway data. A KS^2-chart displays clearly differences between the headway data and statistical models.

The four-stage identification process allows an effective statistical description of the headway data. The process also shows many advantages and disadvantages of the theoretical statistical distributions when compared to the headway data. Moreover, the comparison of data from different sources would be easier, if the data were submitted to standard analyses. The four-stage identification process is suggested as a framework for future studies.

The *independence* of consecutive headways was tested using the autocorrelation analysis, runs tests, and goodness-of-fit tests for the geometric platoon length distribution. The power of the tests was enhanced by combining the significance probabilities of each sample into a single significance measure. The variation of significance over traffic volumes was described and the evidence made stronger by the moving probability methods. The results indicate that the renewal hypothesis should not be accepted in all traffic situations, although the possible autocorrelation is weak.

Concerning the *parameter estimation*, two properties of the headway distributions are worth consideration:

1. The headway distributions are strongly skewed to the right.

2. The headways of trailing vehicles have a practical minimum.

Because of the skewness of the headway distribution, the method of moments does not give as good estimates as the *maximum likelihood method*. Because of the practical minimum headway, a location parameter (τ) is often included in the headway models. In some cases no solutions can be found for the maximum likelihood equations of such models. The *modified maximum likelihood method* was shown to produce good parameter estimates for distributions with a location parameter.

The main problems of the *goodness-of-fit tests* in headway studies are:

1. The chi-square test is based on grouping of the data. This makes the test less powerful than the tests based on the empirical distribution function (EDF).

2. The nonparametric goodness-of-fit tests based on the EDF can be applied only, when the parameters of the distribution are known. If the parameters of the headway distributions must be estimated from the data, the nonparametric tests give too conservative results.

3. When the parameters are estimated from the data, parametric EDF tests should be used. The distribution of the test statistic is, however, not known.

These problems were conquered using parametric EDF test and estimating the distribution of the test statistic by Monte Carlo methods. The superiority of these tests over the nonparametric Kolmogorov-Smirnov tests was demonstrated. The power of the tests was further enhanced by combining the significance probabilities of all samples. It is suggested that the goodness of fit should be measured by both the nonparametric Kolmogorov-Smirnov test and by some parametric EDF test. The parametric Anderson-Darling test is recommended. The conclusions should be based on the parametric test, but the nonparametric test gives an overall goodness-of-fit measure against which different models can be compared. This helps to find out if the problem in the model is lack of fit, or flexibility of the model, which presents higher requirements for the fit.

Three *principles of evaluation* were suggested to evaluate the headway distribution models:

1. Reasonability. It is an advantage if the structure of the model is based on explicit theoretical reasoning about the characteristics of traffic flow.

2. APPLICABILITY. The model should have a simple structure to avoid insurpassable problems in mathematical analysis. If simulation is considered, the generation of pseudo-random variates should be fast and reliable. Parameter estimation should not be too complicated.

3. VALIDITY. The model should give a good approximation of the real world phenomena. This can be tested first by the four-stage identification process and finally by the goodness-of-fit tests.

Five *statistical distributions* were evaluated as models for the headway distributions on two-lane roads. The results of the goodness-of-fit tests[1] are presented in table 6.1.

Table 6.1: Summary of the goodness-of-fit tests

Distribution	P(K-S)	P(A-D)
Negative exponential	$2.1 \cdot 10^{-162}$	$2.9 \cdot 10^{-150}$
Shifted exponential	$3.8 \cdot 10^{-157}$	$5.3 \cdot 10^{-150}$
Gamma	$8.7 \cdot 10^{-106}$	$4.1 \cdot 10^{-141}$
Lognormal	$1.8 \cdot 10^{-36}$	$1.8 \cdot 10^{-114}$
Semi-Poisson	0.85	$4.3 \cdot 10^{-11}$

The *exponential distribution* can be considered as a model for vehicle headways under extremely low flow conditions, and for applications that are not very sensitive to the shape of the headway distribution. In some cases the analytical solution of the problem may only be possible with a very simple model. In such situations the limitations of the model should be clearly indicated.

The *shifted exponential distribution* does not produce extremely short headways, which is a major problem in the negative exponential distribution. The goodness of fit is, however, not significantly better. The shifted exponential distribution can be used in similar situations as the negative exponential distribution, especially when it is important to set a minimum headway.

The *gamma distribution* can be used, if the distribution should have a simple form and low probability for short headways, and the results are not too sensitive to other properties of the distribution. Under low-to-moderate

[1] The precision of the Kolmogorov-Smirnov tests was four decimals. If a test result was 0.0000, the combined probability was calculated using 0.000025 in stead.

traffic volumes the use of the gamma distribution can be given preference over the exponential distribution.

The *lognormal distribution* falls between the simple-nonrealistic and complex-realistic models. The goodness-of-fit results are better than for the other simple distributions, but mathematical analysis presents some problems. There are, however, some theoretical reasons to consider the lognormal distribution as a model for the follower headway distribution (Daou 1966, Greenberg 1966).

The *semi-Poisson distribution* can be recommended for demanding applications with adequate computational facilities. It produced the best test results, but even it failed the parametric test.

The data in this study are from Finnish two-lane two-way roads. Such factors as speed limit, passing sight distance, lane and shoulder width, climbing lanes, proportion of heavy vehicles, flow rate of opposite traffic, traffic signals, distance to major traffic generators, and intersection density have an effect on the headway distribution. Our knowledge of these relationships is, however, very limited. The headway distribution model should be very adaptive, but not too flexible. It should have all the right parameters, but none too many. In this respect all the work toward a unified traffic flow theory is most welcome.

The proposed procedures give a scientific foundation to identify and estimate statistical models for vehicle headways, and to test the goodness of fit. It has been shown that the statistical methods in the analysis of vehicle headways should be thoroughly revised following the guidelines presented in this text.

References

Abramowitz, M. & Stegun, I. A. (1972), *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Dover Publications, Inc., New York. 9th printing.

Adams, W. F. (1936), 'Road traffic considered as a random series', *Journal of the Institution of Civil Engineers* **4**(1), 121–130.

Akçelik, R. & Chung, E. (1994), 'Calibration of the bunched exponential distribution of arrival headways', *Road & Transport Research* **3**(1), 42–59.

Anderson, T. & Darling, D. (1952), 'Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes', *The Annals of Mathematical Statistics* **23**, 193–212.

Ashton, W. D. (1971), 'Distribution for gaps in road traffic', *J. Inst. Maths Applics* **7**, 37–46.

Baerwald, J. E., ed. (1976), *Transportation and Traffic Engineering Handbook*, Prentice-Hall, Inc., Englewood Cliffs, N.J.

Bai, J., Jakeman, A. J. & Taylor, J. A. (1990), 'Percentile estimation of the three-parameter gamma and lognormal distributions: Methods of moments versus maximum likelihood', *Mathematics and Computers in Simulation* **32**(1&2), 167–172.

Baras, J. S., Dorsey, A. J. & Levine, W. S. (1979*a*), 'Estimation of traffic platoon structure from headway statistics', *IEEE Transactions on Automatic Control* **24**(4), 553–559.

Baras, J. S., Levine, W. S. & Lin, T. L. (1979*b*), 'Discrete-time point processes in urban traffic queue estimation', *IEEE Transactions on Automatic Control.* **24**(1), 12–27.

Betró, B., Schoen, F. & Speranza, M. G. (1987), A stochastic environment for the adaptive control of single intersections, *in* N. H. Gartner & N. H. M. Wilson, eds, 'Transportation and Traffic Theory', Proceedings of the Tenth International Symposium on Transportation and Traffic Theory, held July 8–10, 1987, at the Massachusetts Institute of Technology, Cambridge, Massachusetts, Elsevier, New York.

Botma, H. (1986), Traffic operations on busy two-lane rural roads in the Netherlands, *in* 'Traffic Flow Theory, Characteristics, and Highway Capacity', number 1091 *in* 'Transportation Research Record', Transportation Research Board, Washington, D.C., pp. 126–131.

Botma, H. & Fi, I. (1991), Traffic operation on 2-lane roads in Hungary and the Netherlands, *in* Brannolte (1991), pp. 41–53.

Bowman, K. O. & Shenton, L. R. (1988), *Properties of Estimators for the Gamma Distribution*, Marcel Dekker, Inc., New York.

Brannolte, U., ed. (1991), *Highway Capacity and Level of Service*, Balkema, Rotterdam.

Branston, D. (1976), 'Models of single lane time headway distributions', *Transportation Science* **10**(2), 125–148.

Branston, D. (1979), 'A method of estimating the free speed distribution for a road', *Transportation Science* **13**(2), 130–145.

Breiman, L. (1963), 'The Poisson tendency in traffic distribution', *The Annals of Mathematical Statistics* **32**, 308–311.

Breiman, L., Gafarian, A. V., Lichtenstein, R. & Murthy, V. K. (1969), An experimental analysis of single-lane time headways in freely flowing traffic, *in* Leutzbach & Baron (1969), pp. 22–29.

Breiman, L., Lawrence, R., Goodwin, D. & Bailey, B. (1977), 'The statistical properties of freeway traffic', *Transportation Research* **11A**(4), 221–228.

Brockwell, P. J. & Davis, R. A. (1987), *Time Series: Theory and Methods*, Springer Verlag, New York.

Buckley, D. J. (1962), Road traffic headway distributions, *in* 'Proceedings', Vol. 1, part 1, Australian Road Research Board, Victoria, pp. 153–187.

Buckley, D. J. (1968), 'A semi-Poisson model of traffic flow', *Transportation Science* **2**(2), 107–133.

Bureau of Public Roads (1950), *Highway Capacity Manual*, U.S. Department of Commerce, Washington, D.C.

Chandler, R. E., Herman, R. & Montroll, E. W. (1958), 'Traffic dynamics: Studies in car following', *Operations Research* **6**(2), 165–184.

Chishaki, T. & Tamura, Y. (1984), Headway distribution model based on the distinction between leaders and followers, *in* J. Volmuller & R. Hamerslag, eds, 'Proceedings of the Ninth International Symposium on Transportation and Traffic Theory', VNU Science Press, Utrecht, pp. 43–63.

Chrissikopoulos, V., Darzentas, J. & McDowell, M. (1982), 'Aspects of headway distributions and platooning on major roads', *Traffic Engineering and Control* **23**(5), 268–271.

Cochran, J. K. & Cheng, C.-S. (1989), 'Automating the procedure for analyzing univariate statistics in computer simulations contexts', *Transactions of the Society for Computer Simulation* **6**(3), 173–188.

Cohen, A. C. & Whitten, B. J. (1980), 'Estimation in the three-parameter lognormal distribution', *Journal of the American Statistical Association* **75**(370), 399–404.

Cohen, A. C. & Whitten, B. J. (1982), 'Modified moment and maximum likelihood estimators for parameters of the three-parameter gamma distribution', *Communications in Statistics: Simulation and Computation* **11**(2), 197–216.

Cohen, A. C. & Whitten, B. J. (1988), *Parameter Estimation in Reliability and Life Span Models*, Marcel Dekker, Inc., New York.

Conover, W. (1980), *Practical Nonparametric Statistics*, second edn, John Wiley & Sons, New York.

Cooper, R. B. (1981), *Introduction to Queueing Theory*, second edn, Edward Arnold (Publishers) Limited, London.

Cowan, R. (1978), 'An improved model for signalized intersections with vehicle-actuated control', *Journal of Applied Probability* **15**, 384–396.

Cowan, R. J. (1975), 'Useful headway models', *Transportation Research* **9**(6), 371–375.

Cox, D. R. & Hinkley, D. (1974), *Theoretical Statistics*, Chapman and Hall, London.

Cox, D. R. & Lewis, P. A. W. (1966), *The Statistical Analysis of Series of Events*, Methuen & Co Ltd., London.

Cox, D. R. & Stuart, A. (1955), 'Some quick sign tests for trend in location and dispersion', *Biometrika* **42**, 80–95.

Crow, E. L. & Shimizu, K., eds (1988), *Lognormal Distributions: Theory and Applications*, Marcel Dekker, Inc., New York.

D'Agostino, R. B. & Stephens, M. A., eds (1986), *Goodness-of-Fit Techniques*, Marcel Dekker, Inc., New York.

Daley, D. & Vere-Jones, D., eds (1988), *An Introduction to the Theory of Point Processes*, Springer Verlag, New York.

Daou, A. (1964), 'The distribution of headways in a platoon', *Operations Research* **12**(2), 360–361.

Daou, A. (1966), 'On flow within platoons', *Australian Road Research* **2**(7), 4–13.

Darroch, J. N., Newell, G. F. & Morris, R. W. J. (1964), 'Queues for a vehicle-actuated traffic light', *Operations Research* **12**(6), 882–895.

Dawson, R. F. (1969), The hyperlang probability distribution - a generalized traffic headway model, *in* Leutzbach & Baron (1969), pp. 30–36.

Dawson, R. F. & Chimini, L. A. (1968), The hyperlang probability distribution—A generalized traffic headway model, *in* 'Characteristics of Traffic Flow', number 230 *in* 'Highway Research Record', Highway Research Board, Washington, D.C., pp. 1–14.

Drew, D. R. (1967), Gap acceptance characteristics for ramp-freeway surveillance and control, *in* 'Freeway Traffic Characteristics and Control', number 157 *in* 'Highway Research Record', Highway Research Board, Washington, D.C., pp. 108–143. With discussion.

Drew, D. R. (1968), *Traffic Flow Theory and Control*, McGraw-Hill, New York.

Dudewicz, E. J. & Mishra, S. N. (1988), *Modern Mathematical Statistics*, John Wiley & Sons, New York.

Dunne, M. C., Rothery, R. W. & Potts, R. B. (1968), 'A discrete Markov model of vehicular traffic', *Transportation Science* **2**(3), 233–251.

Durbin, J. (1975), 'Kolmogorov-Smirnov tests when parameters are estimated with applications to tests of exponentiality and tests on spacings', *Biometrika* **62**(1), 5–22.

Enberg, Å. & Pursula, M. (1991), Traffic flow and level of service on high-class two-lane rural roads in Finland, *in* Brannolte (1991), pp. 119–126.

Enberg, Å. & Pursula, M. (1992), *Moottoriliikennetien liikennevirran ominaisuudet* [Basic characteristics of traffic flow on high-class two-lane rural roads in Finland], number 12/1992 *in* 'Tielaitoksen selvityksiä', Tielaitos, Helsinki.

Fisher, R. A. (1938), *Statistical Methods for Research Workers*, seventh edn, Oliver and Boyd, Edinburgh.

Galin, D. (1980), 'A comparison of some bunch-size models for two-lane rural roads - an Israeli experience', *Transportation Research* **14A**(1), 51–56.

Garwood, F. (1940), 'An application of the theory of probability to the operation of vehicular-controlled traffic signals', *Journal of the Royal Statistical Society* **7**(1), 65–77.

Gazis, D. G., Herman, R. & Potts, Renfrey, B. (1959), 'Car-following theory of steady-state traffic flow', *Operations Research* **7**(4), 499–505.

Gazis, D. G., Herman, R. & Rothery, R. W. (1961), 'Nonlinear follow-the-leader models of traffic flow', *Operations Research* **9**(4), 545–567.

Gill, P. E., Murray, W. & Wright, M. H. (1981), *Practical Optimization*, Academic Press, London.

Gradshteyn, I. S. & Ryzhik, I. M. (1980), *Table of Integrals, Series, and Products*, Academic Press, New York. Corrected and enlarged edition prepared by Alan Jeffrey.

Grecco, W. L. & Sword, E. C. (1968), 'Prediction of parameters for Schuhl's headway distribution', *Traffic Engineering* **38**, 36–38.

Green, J. R. & Hegazy, Y. A. S. (1976), 'Powerful modified-EDF goodness-of-fit tests', *Journal of the American Statistical Association* **71**(353), 204–209.

Greenberg, I. (1966), 'The log-normal distribution of headways', *Australian Road Research* **2**(7), 14–18.

Greenwood, J. A. & Durand, D. (1960), 'Aids for fitting the gamma distribution by maximum likelihood', *Technometrics* **2**(1), 55–65.

Griffiths, J. D. & Hunt, J. G. (1991), 'Vehicle headways in urban areas', *Traffic Engineering and Control* **32**(10), 458–462.

Gross, D. & Harris, C. M. (1985), *Fundamentals of Queueing Theory*, second edn, John Wiley & Sons, New York.

HCM (1965), *Highway Capacity Manual*, number 87 *in* 'Special Report', Highway Research Board,National Research Council, Washington, D.C.

HCM (1985), *Highway Capacity Manual*, number 209 *in* 'Special Report', Transportation Research Board, National Research Council, Washington, D.C.

HCM (1994), *Highway Capacity Manual*, number 209 *in* 'Special Report', third edn, Transportation Research Board, National Research Council, Washington, D.C. Updated October 1994.

Heidemann, D. (1990), A theoretical model to calculate time-headway distributions as a function of traffic density, *in* M. Koshi, ed., 'Transportation and Traffic Theory', Proceedings of the Eleventh International Symposium on Transportation and Traffic Theory, held July 18–20, 1990, in Yokohama, Japan, Elsevier, New York, pp. 1–17.

Heidemann, D. (1993), A theoretical model for distributions of speeds and time-headways on two-lane roads, *in* C. F. Daganzo, ed., 'Transportation and Traffic Theory', Proceedings of the 12^{th} International Symposium on Transportation and Traffic Theory, Berkeley, California, USA, 21–23 July, 1993, Elsevier, Amsterdam, pp. 523–537.

Herman, R. & Potts, R. B. (1961), Single-lane traffic theory and experiment, *in* R. Herman, ed., 'Theory of Traffic Flow', Proceedings of the Symposium on the Theory of Traffic Flow, Elsevier, Amsterdam, pp. 120–146.

IMSL (1989), *IMSL STAT/LIBRARY. Version 1.1*, IMSL, Inc., Houston, Texas.

Johnson, N. L., Kotz, S. & Balakrishnan, N. (1994), *Continuous Univariate Distributions*, Vol. 1, second edn, John Wiley & Sons, New York.

Kalbfleisch, J. D. & Prentice, R. L. (1980), *The Statistical Analysis of Failure Time Data*, John Wiley & Sons, New York.

Kell, J. H. (1962), Analyzing vehicular delay at intersections through simulation, *in* 'Theory of Traffic Flow', number 356 *in* 'Highway Research Board Bulletin', Highway Research Board, Washington, D.C., pp. 28–39.

Kendall, M. & Ord, J. K. (1990), *Time-Series*, third edn, Edward Arnold, London.

Khintchine, A. Y. (1960), *Mathematical Methods in the Theory of Queueing*, Charles Griffin Ltd., London. Translated by D. M. Andrews and M. H. Quenouille.

Kikuchi, S. & Chakroborty, P. (1992), Car-following model based on fuzzy inference system, *in* 'Highway Capacity and Traffic Flow', number 1365 *in* 'Transportation Research Record', Transportation Research Board, Washington, D.C., pp. 82–91.

Kingman, J. F. C. (1962), 'On queues in heavy traffic', *Journal of the Royal Statistical Society* **24**(2), 383–392.

Kleinrock, L. (1975), *Queueing Systems. Volume I: Theory*, John Wiley & Sons, New York.

Leutzbach, W. (1988), *Introduction to the Theory of Traffic Flow*, Springer Verlag, Berlin.

Leutzbach, W. & Baron, P., eds (1969), *Beiträge zur Theorie des Verkehrsflusses*, number 86 *in* 'Strassenbau und Strassenverkehrstechnik', Bundesminister für Verkehr, Bundesminister für Verkehr, Bonn.

Lilliefors, H. W. (1967), 'On the Kolmogorov-Smirnov test for normality with mean and variance unknown', *American Statistical Association Journal* **62**, 399–402.

Lilliefors, H. W. (1969), 'On the Kolmogorov-Smirnov test for the exponential distribution with mean unknown', *American Statistical Association Journal* **64**, 387–389.

Lin, F. (1985*a*), 'Relationships between queueing flows and presence detectors', *ITE Journal* pp. 42–46.

Lin, F.-B. (1985*b*), Evaluation of queue dissipation simulation models for analysis of presence-mode full-actuated signal control, *in* 'Highway Capacity, Traffic Characteristics, and Flow Theory', number 1005 *in* 'Transportation Research Record', Transportation Research Board, Washington, D.C., pp. 46–54.

Luttinen, R. T. (1990), Johdatus aikavälijakaumien teoriaan [Introduction to the theory of headway distributions], Julkaisu 71, Teknillinen korkeakoulu, Liikennetekniikka, Otaniemi.

Luttinen, R. T. (1991), Testing goodness of fit for 3-parameter gamma distribution, *in* 'Proceedings of the Fourth IMSL User Group Europe Conference', IMSL and CRPE-CNET/CNRS, Paris, pp. B10/1–13.

Luttinen, R. T. (1992), Statistical properties of vehicle time headways, *in* 'Highway Capacity and Traffic Flow', number 1365 *in* 'Transportation Research Record', Transportation Research Board, Washington, D.C., pp. 92–98.

Luttinen, R. T. (1994), Identification and estimation of headway distributions, *in* R. Akçelik, ed., 'Second International Symposium on Highway Capacity', Australian Road Research Board, Transportation Research Board, Sydney, pp. 427–446.

Madansky, A. (1988), *Prescriptions for Working Statisticians*, Springer Verlag, New York.

May, A. (1961), 'Traffic characteristics and phenomena on high density controlled access facilities', *Traffic Engineering* **31**(6), 11–19.

May, A. D. (1965), Gap availability studies, *in* 'Traffic Flow Characteristics 1963 and 1964', number 72 *in* 'Highway Research Record', Highway Research Board, Washington, D.C., pp. 101–136.

May, A. D. (1990), *Traffic Flow Fundamentals*, Prentice-Hall, Inc., Englewood Cliffs, N.J.

McLean, J. R. (1989), *Two-Lane Highway Traffic Operations: Theory and Practice*, Gordon and Breach Science Publishers, New York.

Miller, A. J. (1961), 'A queueing model for road traffic flow', *Journal of the Royal Statistical Society* **23B**(1), 64–75.

Montgomery, D. C. (1984), *Design and Analysis of Experiments*, second edn, John Wiley & Sons, New York.

Moore, D. S. (1986), *Tests of Chi-Square Type*, In D'Agostino & Stephens (1986), chapter 3, pp. 63–95.

Nagel, E. (1982), *The Structure of Science: Problems in the Logic of Scientific Explanation*, Routledge & Kegan Paul, London. First published as a paperback 1979.

Newell, G. F. (1956), 'Statistical analysis of the flow of highway traffic through a signalized intersection', *Quarterly of Applied Mathematics* **13**(4), 353–369.

Newell, G. F. (1965), 'Approximation methods for queues with application to the fixed-cycle traffic light', *SIAM Review* **7**(2), 223–240.

Noreen, E. W. (1989), *Computer Intensive Methods for Testing Hypotheses. An Introduction.*, John Wiley & Sons, New York.

Popper, K. R. (1983), *The Logic of Scientific Discovery*, Hutchinson, London. 11th impression (revised).

Pursula, M. & Enberg, Å. (1991), Characteristics and level of service estimation of traffic flow on two-lane rural roads in Finland, *in* 'Freeway Operations, Highway Capacity, and Traffic Flow 1991', number 1320 *in* 'Transportation Research Record', Transportation Research Board, Washington, D.C., pp. 135–143.

Pursula, M. & Sainio, H. (1985), *Kaksikaistaisten teiden liikennevirran perusominaisuudet* [Basic characteristics of traffic flow on two-lane rural roads in Finland], Helsinki University of Technology, Finnish Roads and Waterways Administration, Helsinki.

Rajalin, S. & Hassel, S.-O. (1992), Ajoneuvojen aikavälit ja lähellä perässä ajamisen syyt [Vehicle intervals and the reasons for tailgating], Liikenneturvan tutkimuksia 108, The Central Organization for Traffic Safety in Finland, Helsinki.

Ramberg, J. S., Tadikamalla, P. R., Dudewicz, E. J. & Mykytka, E. F. (1979), 'A probability distribution and its uses in fitting data', *Technometrics* **21**(2), 201–209.

Sainio, H. (1984), Liittymisrampin toiminta moottoriliikennetiellä [The function of the entrance ramp of an undivided two-lane road for motor traffic only], Master's thesis, Helsinki University of Technology, Espoo.

Salonen, M. (1982), Nopeusrajoituksen ja sään vaikutus liikennevirtaan - Kirjallisuustutkimus ja analyysi Jorvaksentien liikennevirrasta [The effects of speed limits and weather to the traffic flow - literature study and analysis of the traffic flow on the Western motorway of Helsinki], Master's thesis, Helsinki University of Technology, Espoo.

Schuhl, A. (1955), The probability theory applied to distribution of vehicles on two-lane highways, *in* 'Poisson and Traffic', Eno Foundation, Saugatuck, pp. 59–75.

Shawaly, E. A. A., Ashworth, R. & Laurence, C. J. D. (1988), 'A comparison of observed, estimated and simulated queue lengths and delays at oversaturated signalized junctions', *Traffic Engineering and Control* **29**(12), 637–643.

Silverman, B. (1986), *Density Estimation for Statistics and Data Analysis*, Chapman and Hall Ltd., London.

Stephens, M. A. (1986*a*), *Tests Based on EDF Statistics*, In D'Agostino & Stephens (1986), chapter 4, pp. 97–193.

Stephens, M. A. (1986*b*), *Tests for the Uniform Distribution*, In D'Agostino & Stephens (1986), chapter 8, pp. 331–366.

Stuart, A. (1956), 'The efficiences of tests of randomness against normal regression', *Journal of the American Statistical Association* **51**, 285–287.

Stuart, A. & Ord, J. K. (1987), *Kendall's Advanced Theory of Statistics. Volume 1: Distribution Theory*, fifth edn, Charles Griffin & Company Ltd., London.

Stuart, A. & Ord, J. K. (1991), *Kendall's Advanced Theory of Statistics. Volume 2: Classical Inference and Relationship*, fifth edn, Edward Arnold, London.

Sullivan, D. P. & Troutbeck, R. J. (1994), 'The use of Cowan's M3 headway distribution for modelling urban traffic flow', *Traffic Engineering and Control* **35**(7/8), 445–450.

Summala, H. & Vierimaa, J. (1980), *Ajoneuvojen nopeudet, nopeuserot ja aikavälit 10 mittauspisteessä syksyllä 1978: Analyysejä autonkuljettajien käyttäytymisestä* [Vehicle speeds, speed differences and headways at 10 locations in autumn 1978: Analyses of driver behavior], Finnish National Board of Public Roads and Waterways, Helsinki.

Tanner, J. C. (1953), 'A problem of interference between two queues', *Biometrika* **40**, 58–69.

Tanner, J. C. (1961), 'Delays on a two-lane road', *Journal of the Royal Statistical Society* **23B**(1), 38–63.

Thedéen, T. (1964), 'A note on the Poisson tendency in traffic distribution', *The Annals of Mathematical Statistics* **35**(4), 1823–1824.

Tolle, J. E. (1971), 'The lognormal headway distribution model', *Traffic Engineering and Control* **12**(5), 22–24.

Trivedi, K. S. (1982), *Probability and Statistics with Reliability, Queueing, and Computer Science Applications*, Prentice-Hall, Inc., Englewood Cliffs, NJ.

Wald, A. & Wolfowitz, J. (1940), 'On a test whether two samples are from the same population', *Annals of Mathematical Statistics* **2**, 147–162.

Wardrop, J. G. (1952), Some theoretical aspects of road traffic research, *in* 'Proceedings of the Institution of Civil Engineers', Vol. 1, Institution of Civil Engineers, London, pp. 325–378. With discussion.

Wasielewski, P. (1974), 'An integral equation for the semi-poisson headway distribution model', *Transportation Science* **8**, 237–247.

Wasielewski, P. (1979), 'Car-following headways on freeways interpreted by the semi-Poisson headway distribution model', *Transportation Science* **13**(1), 36–55.

Weiss, G. & Herman, R. (1962), 'Statistical properties of low-density traffic', *Quarterly of Applied Mathematics* **20**, 121–130.

Appendix A

Data description

Table A.1: Headway samples from 80–100 km/h roads

No.	SL	n	Volume	s^2	C	α_3	α_4
1	100	500	912	36.91	1.54	3.89	20.31
2	100	250	967	28.90	1.44	3.61	17.61
3	100	400	1068	18.04	1.26	3.64	20.24
4	100	400	945	30.75	1.46	4.49	30.63
5	100	391	927	26.34	1.32	3.48	16.50
6	100	125	262	725.61	1.96	3.71	19.83
7	100	190	352	247.15	1.54	2.94	13.98
8	100	90	439	332.74	2.23	4.81	31.04
9	100	75	257	805.51	2.03	3.37	15.62
10	100	180	979	23.07	1.31	4.13	25.46
11	100	420	1282	13.68	1.32	5.40	44.68
12	100	153	1837	1.33	0.59	2.95	16.34
13	100	500	1410	8.74	1.16	4.21	25.59
14	100	350	1462	4.64	0.87	3.79	22.70
15	100	250	1388	5.94	0.94	2.89	13.11
16	100	180	1250	15.01	1.34	3.81	18.42
17	100	600	1056	16.46	1.19	3.46	17.79
18	100	500	1145	13.82	1.18	4.22	28.33
19	100	200	1230	13.36	1.25	4.91	30.65
20	100	440	1079	12.84	1.07	2.98	14.02
21	100	200	1118	12.80	1.11	3.66	20.70
22	80	100	207	534.03	1.33	1.96	7.40
23	80	200	1255	8.63	1.02	3.66	18.49
24	80	500	1451	7.62	1.11	7.31	77.03
25	80	900	1439	6.13	0.99	5.16	41.87
26	80	114	1459	1.91	0.56	2.03	9.18
27	80	700	1330	3.29	0.67	3.34	21.93
28	80	370	1245	5.19	0.79	3.59	22.92
29	80	600	1467	3.71	0.78	8.76	132.67
30	80	226	1000	44.65	1.86	5.40	37.66
31	80	250	746	57.48	1.57	3.93	21.83
32	80	390	795	55.47	1.64	4.03	22.02
33	80	210	361	175.94	1.33	2.20	8.58

Table A.2: Headway samples from 50–70 km/h roads

No.	SL	n	Volume	s^2	C	α_3	α_4
34	50	200	328	160.87	1.16	2.05	8.50
35	50	170	367	156.78	1.28	2.49	10.54
36	50	200	637	48.95	1.24	2.84	13.40
37	50	251	450	80.59	1.12	1.62	5.16
38	50	150	355	120.85	1.08	1.83	6.17
39	50	167	429	78.92	1.06	1.84	6.71
40	50	100	459	70.88	1.07	1.79	6.15
41	50	200	650	57.29	1.37	4.12	24.56
42	50	150	794	25.49	1.11	3.20	16.12
43	50	350	1291	3.55	0.68	2.11	8.08
44	50	300	662	34.24	1.08	2.56	10.89
45	50	100	158	369.72	0.85	1.38	5.09
46	50	76	148	465.14	0.89	1.41	4.37
47	60	150	613	29.76	0.93	2.09	8.58
48	60	240	764	19.31	0.93	1.93	7.42
49	60	250	646	40.13	1.14	2.50	10.90
50	60	112	533	47.72	1.02	1.27	3.62
51	60	320	885	18.33	1.05	2.47	9.47
52	60	150	1252	8.12	0.99	3.99	22.66
53	60	260	915	22.34	1.20	3.32	17.42
54	60	220	835	27.51	1.22	2.78	11.97
55	60	100	309	218.68	1.27	1.75	5.55
56	60	140	542	58.50	1.15	1.96	7.15
57	60	140	479	95.84	1.30	2.16	7.94
58	60	290	561	62.95	1.24	2.33	9.28
59	70	110	332	200.75	1.30	3.12	16.09
60	70	100	370	154.69	1.28	2.27	8.12
61	70	120	324	156.80	1.13	2.06	8.14
62	70	150	143	1001.61	1.26	1.72	5.90
63	70	100	238	324.00	1.19	1.46	4.41
64	70	160	298	153.24	1.02	1.46	4.81
65	70	300	475	72.28	1.12	1.83	6.05