Preface

This thesis is submitted in partial fulfillment of the requirements for the Ph.D. degree at the Technical university of Denmark (DTU). The project is under the Industrial Ph.D. Fellowship Program administered by the Danish Academy of Technical Sciences (ATV) and financially supported by the Ministry of Science, Technology and Innovation. The main participants are Carl Bro and the Department of Informatics and Mathematical Modelling (IMM) at DTU. The Centre for Traffic and Transport (CTT), also at DTU, has been third party in the project. The project has been carried out in the period 1 June 1999 - 31 May 2002 with a four months stay at University College London (UCL) in the fall 2000.

This thesis is concerned with models and methods for improving hot spot safety work on the Danish road network.

Aknowledgement

I had no prior knowledge of the area of traffic safety, and the past three years have been very educational. In addition, my statistical knowledge have improved considerably in the process. I wish to express my gratitude to the number of people who helped and supported me in my studies.

First of all my main supervisor Poul Thyregod at IMM, DTU for his help and invaluable guidance throughout the last three years. All meetings with him have been constructive and inspirational. I am grateful to Pierre E. Aagaard, formerly at Carl Bro, for constructive criticism and encouragement. Even after leaving Carl Bro he played an active part in the project. Also, thanks to my supervisors Michael Aakjer Nielsen, at Carl Bro, and N. O. Jørgensen and Jan Grubb Laursen both at CTT. I wish to thank Stig Hemdorff and the Danish Road Directorate for assisting with the data collection and analysis. Special thanks to Conni Diget Fuglsang, at Carl Bro, for her time and effort in proof-reading the thesis.

Half-way through the project I visited the Centre for Transport Studies (CTS) at UCL, for four months. I wish to thank my supervisor at CTS, Ben Heydecker, for useful feedback on the models and to everyone at the department for making my stay there so pleasant.

I would like to express my gratitude to ATV and Carl Bro for making this project possible. Also, thanks to all my colleagues at Carl Bro and at the Section for Statistics at IMM for creating a comfortable and welcoming environment. In particular my officemate at IMM, Camilla Madsen, for our many pleasant hours. Last but not least, I am grateful to family and friends, especially my husband Nicolai, for their love and support. The last three years have been enjoyable and interesting.

Kgs. Lyngby, May 2002

Dorte Vistisen

Resume

Selvom der årligt anvendes mange millioner kroner på at gøre det danske vejnet mere sikkert at færdes i, er ressourcerne til trafiksikkerhed begrænsede. Det er derfor vigtigt, at disse ressourcer anvendes så effektivt som overhovedet muligt. Nærværende afhandling omhandler den del af trafiksikkerhedsarbejdet, der benævnes *sortpletarbejde*. Sortpletarbejde er arbejdet med at forbedre trafiksikkerheden gennem ændringer af geometriske og miljømæssige karakteristika på det eksisterende vejnet.

De anvendte modeller og metoder i sortpletarbejde i Danmark idag er udviklet for 20-30 år siden, hvor datamængden var mere begrænset og edb og statistiske metoder ikke så udviklede. Målet med denne afhandling er at bidrage til at forbedre det aktuelle danske tekniske niveau i sortpletarbejdet.

Basis for det systematiske sortpletarbejde er de opstillede modeller til beskrivelse af variationen i antallet af trafikuheld på vejnettet. I afhandlingen opstilles hierarkiske modeller disaggregerede over tid. Det er vist at de foreslåede uheldsmodeller beskriver variation i uheldstal bedre end de modeller, der idag benyttes i Danmark. Parametre for de opstillede modeller er estimeret for stats- og amtsveje udfra data fra Vejsektorens Informations System, VIS. Specifikke uheldsmodeller for det kommunale vejnet er ikke estimeret, da vejdata for kommuneveje ikke foreligger på systematiseret form.

Der er i projektet udviklet metoder til udpegning af særligt uheldsbelastede kryds og vejstrækninger, de såkaldte *sorte pletter*. Metoderne er baseret på de opstillede uheldsmodeller og er vist at være mere effektive end de anvendte eksisterende metoder i Danmark. Derudover gives der retningslinier for hvorledes prioriteringen af sorte pletter og sikkerhedsforanstaltninger kan forbedres.

En ny model til estimation af effekten af et trafiksikkerhedstiltag er foreslået. Modellen tager højde for den såkaldte *regressionseffekt* og giver bedre effektestimater end modellen benyttet på det danske vejnet idag. Den foreslåede model er ligeledes vist at være bedre end de hidtige modeller beskrevet i den internationale litteratur.

Foruden afhandlingen er der tillige publiceret følgende:

- Modeller og metoder til sortpletarbejde i DK. *Dansk Vejtidsskrift*, maj 2001.

- Metoder til detektering og vurdering af trafiksikkerhedsproblemer i vejnettet.

Trafikdage på Aalborg Universitet, august 2001.

- A consistent method for estimating the effect of hot spot safety work. *Traffic Engineering + Control*, marts 2002.

Abstract

Despite the fact that millions DKK each year are spent on improving road safety in Denmark, funds for traffic safety are limited. It is therefore vital to spend the resources as effectively as possible. This thesis is concerned with the area of traffic safety denoted *hot spot safety work*, which is the task of improving road safety through alterations of the geometrical and environmental characteristics of the existing road network.

The presently applied models and methods in hot spot safety work on the Danish road network were developed about two decades ago, when data was more limited and software and statistical methods less developed. The purpose of this thesis is to contribute to improving *State of the art* in Denmark.

Basis for the systematic hot spot safety work are the models describing the variation in accident counts on the road network. In the thesis hierarchical models disaggregated on time are derived. The proposed models are shown to describe variation in accident counts better than the models currently at use in Denmark. The parameters of the models are estimated for the national and regional road network using data from the Road Sector Information system, VIS. No specific accident models are estimated for the local road network as road data on local roads are not collected in a systematic manner.

Methods are developed for targeting intersections and road sections in the road network with an unusual high number of accidents, the so-called *hot spots*. The methods are based on the proposed accident models and they are shown to outperform the methods used in Denmark today. Also, guidelines on how to improve the prioritizing of hot spots and safety improving measures are provided.

A new model for estimating the effect of treating hot spots is proposed. The model takes into account the so-called *regression to the mean effect* and results in better estimates of the effect of treatment than the model currently at use on the Danish road network. The proposed method is also shown to outperform the methods as yet suggested in the international literature.

In addition to this thesis, the following papers are published:

- Modeller og metoder til sortpletarbejde i DK (in Danish). *Dansk Vejtidsskrift*, May 2001.

- Metoder til detektering og vurdering af trafiksikkerhedsproblemer i vejnettet (in Danish). *Trafikdage på Aalborg Universitet*, August 2001.
- A consistent method for estimating the effect of hot spot safety work. *Traffic Engineering + Control*, March 2002.

Contents

Chapter 1

Introduction

This first chapter provides an introduction to and overview of the topics in the thesis. This thesis is concerned with the area of traffic safety denoted *hot spot safety work*, which is the task of improving road safety through alterations of the geometrical and environmental characteristics of existing roads. The concept of *road safety* is used to describe how safe, or rather the opposite, how dangerous or risky, it is to pass a particular section of the road network. Road safety may be divided into:

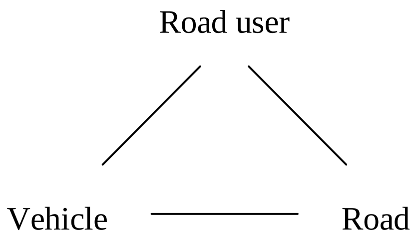
Perception of safety which is a feeling of security.

Observed safety which is reflected in the prevalence of accidents and their harm.

In this project, only the observed measure of safety is considered. The reason being that perception of safety is subjective and unpredictable. At the same site, the perception of safety may differ significantly from one person to another. A road user's feeling of security at a site may even change over time without changes to the road geometry or traffic flow at this site. Furthermore, the perception of safety can only be described and not measured, and it is an area of which little knowledge exists. This is not to say that road users' perception of safety is unimportant from a social point of view[1], but it should not be used as guidance in road safety work. Accident frequency on the other hand is factual and measurable. It is an objective measure, which is universal to all sites, and it is a highly researched area.

Traffic safety work aims at reducing both the risk and consequences of road traffic accidents. An accident may be seen as an interaction of three parties; the road user, the vehicle and the road.

[1] As pointed out in Adams (1988); A low number of reported accidents does not necessarily mean that the site is safe, it may be so terrifyingly dangerous that few people try to cross it.



From several studies of road accidents, it has become the general belief that a traffic accident is a consequence of a failure in this system. In-depth studies of accidents will often reveal the factors that have contributed to the accident occurrence. An accident factor may be assigned to the road user, the road or the vehicle. Studies in the UK have shown the *road user* factors to be predominant, followed by road environment factors (see T10 (2000)). However, this does not mean that efforts put into accident prevention work should be allocated accordingly. Often it will be easier to change the physical features on roads than changing the behavior of the road users.

Traffic safety work related to the physical characteristics of the roads may be divided into two phases. The first phase, the so-called *road safety audit* (see Vejdirektoratet (1999a)), is a systematic procedure applied to new road constructions. Knowledge of traffic safety is incorporated into the road planning and design phase while still on paper. The purpose of the road safety audit is to avoid creating new high risk areas by assuring that only relatively safe roads are built. The road safety audit scheme in Denmark has been developed by the national Road Directorate. The second phase is the so-called *hot spot*[2] *safety work*, in which road safety is improved through alterations of the geometrical and environmental characteristics of the problematic sites in the existing road network. This thesis is focused on hot spot safety work.

Hot spot safety work is the task of targeting and treating intersections and road sections with an unusual high number of accidents, the so-called hot spots. This work may be divided into three phases:

1. Targeting hot spots on the road network.

2. Prioritizing the hot spots to treat with safety improving measures.

3. Before and after studies of the effect of treatment.

The foundation for the hot spot safety work is mapping safety and setting up models describing the variation in accident counts between different sites in the road network. The assumption behind hot spot safety work is that most road accidents

[2] Earlier, the term *black spot* was used. However, due to the politically incorrectness of using the word *black* negatively to indicate a poor level of safety, the term *hot spot* is now applied instead.

are somehow related to the geographical-, geometrical-, environmental- and traffic characteristics of the location of the accident. In addition, on a location with a history of site-related accidents, one may anticipate new accidents to happen as long as the characteristics of this site remain unaltered.

1.1 Background

Since 1988, the number of traffic accidents has dropped despite an increase in the traffic flow. However, traffic accidents are still one of the primary reasons for the loss of living years in Denmark. Around 500 are killed and 10,000 are injured on the road network each year. Numbers, which, compared to the number of inhabitants, are considerably larger than in Norway and Sweden (see Vejdirektoratet (2002b)). In addition, only about 20% of all injury accidents are reported by the police (see Vejdirektoratet (2001b)), thus leading to an underestimation of the actual cost to society of road accidents. The Commission on Traffic Safety in Denmark has set up the *Action Plan for Traffic Safety in 2000* with an overall goal of reducing the number of road injuries by 40% in the period 1998-2012 (see Færdselssikkerhedskommissionen (2000)). Hot spot safety work is identified as one of the cost effective means for reaching this goal.

Despite the fact that millions DKK each year are spent on improving road safety in Denmark, funds for traffic safety are limited. It is therefore vital to spend the resources as effectively as possible. The presently applied models and methods for monitoring and analyzing safety on the Danish road network have been developed about two decades ago, when data were more limited and software and statistical methods less developed.

The general purpose of this Ph.D. study is to increase the level of knowledge in Denmark in the area of statistical models and methods in hot spot safety work. The aim being to provide improved models for describing the variation in accident counts, to propose better methods for targeting and prioritizing hot spots for treatment and a better model for estimating the effect of hot spot treatment work.

1.2 Outline of the thesis

The outline of the thesis is as follows. Chapter 2 proposes models for describing the variation in accident counts at intersections and on road sections respectively. A method for estimating and predicting the safety level at a site is provided. The models are illustrated by an example.

A method for targeting *hot spots* on the road network is proposed in chapter 3. The method is based on the models of chapter 2 and is illustrated by an example. Guidelines for prioritizing between hot spots and remedial safety measures are supplied at the end of the chapter.

Chapter 4 discusses the problems connected with estimating the effect of hot spot treatment work, in particular the problem of the so-called *regression to the mean effect*. A new model for estimating the effect of treatment is proposed. The effect model is based on the site safety estimates of chapter 2 and is illustrated by an example.

In chapter 5, the *state of the art* in Danish hot spot safety work is described and compared to the models and methods proposed in this thesis. Comparisons are made through simulation studies, which are described in greater detail in appendix E.

Finally, the parameters of the models proposed in chapter 2 are estimated for the national and regional road network in chapter 6.

The remaining appendices of the thesis provide the methodological and mathematical details of chapters 2 to 6.

Chapter 2

Modelling variation

Traffic accidents are the unintentional results of human behavior. They may be considered random events in the sense that the time and location of the next accident cannot be predicted. A study of accidents on the road network is not a scientific discipline in which one can perform experiments and repeat accident occurrences. Instead, one may view traffic accidents as events in a stochastic process with statistical modelling as our key source of knowledge.

The purpose of this chapter is to set up a combined cross-section and time dependent model for describing the variation in reported accident counts. Such a model may improve the ability to predict future accident levels at different sites.

For a given time period and place in the road network, the number of reported accidents at a site, x, may be considered a realization of a random variable X, which varies around its mean, λ. This mean is partly dependent on observable quantities, expressed through a parameter μ, and on a non-observable dispersion effect s:

$$\lambda = \mu s$$

The mean, λ, is the expected number of accidents at the site in a given time period. It will be suggested that a suitable model for describing the variation in reported accidents is a so-called hierarchical generalized linear model disaggregated on time, in which accidents are conditionally Poisson distributed with mean λ:

$$X|s \in Poiss\left(\lambda\right)$$

The dispersion effect, s, is described by a random variable, S, which is modelled by a gamma distribution with mean 1. This model is an extension to the generalized linear models (GLIM), presently used by the Road Directorate (see chapter 5), in that a dispersion effect, s, is now included. The argument for including a dispersion effect is that even though a model may include the main determinants of the expected number of accidents at a site, there are still features other than those included in the model, which distinguish one site from another. In other words, the explanatory variables in the model do not provide a complete explanation of the between-site variation. The dispersion effect, S, is random between sites but remains constant within the site from one time period to another. The hierarchical generalized linear model is designed to account for similarities between traffic sites of the same type (through μ) and still express the significance of each site (through s).

2.1 Variation in accident counts

The concept of *road safety* is primarily linked to accidents and their harm[1]. The usual index of safety of a site is the number of accidents expected to occur at this site within a given time period (and road length on road sections). Below, this safety index of a site is merely denoted *site safety*.

Accident counts differ between sites and within a site over time. Factors influencing the reported number of accidents at a site may be divided into six broad categories (see OECD (1997)):

- Autonomous factors determined outside the national social system, such as weather, state of technology, oil prices, population size.
- General socio-economic conditions, such as industrial development, unemployment, taxation, inflation, public education.
- Transportation sector, such as infrastructure, service level of public transportation, travel demand, modal choice, vehicle park.
- Data collection, such as underreporting etc.
- Accident countermeasures.

This thesis is primarily focused on factors related to traffic and road geometry. Basis for the systematic hot spot safety work is the models describing the between-site and within-site variation in road accident counts. Data used for deriving such models are the accident counts and the characteristics of the sites. Site characteristics included in accident models are denoted *traits*. Traffic flow as well as information on the road geometry are examples of characteristics traditionally included in road accident models.

The variation in accident counts may conceptually be separated into explained and unexplained variations:

[1] See chapter 1 for a discussion.

The explained variation in accident counts is the part of variation that may be ascribed to differences in the traits. All variation not ascribed to the traits is termed unexplained.

The unexplained variation in accident counts is the phenomenon that the reported number of accidents varies between sites and over time without differences in the traits.

Accident models assume accidents at a site to vary randomly around the expected number of accidents at this site, i.e. around the site safety level at the site. The models aim at describing any differences in safety levels between sites and within sites over time by differences in the traits. The remaining unexplained variation is modelled as random and is traditionally described by the Poisson distribution.

In practice, it is impossible to include all site-specific conditions as traits in the models. Either because they are unobserved or due to economical and practical considerations, which are limiting a systematic collection of such data. It is thus important to notice that the unexplained variation between and within sites covers variation related to non-observable quantities or observable quantities not included as traits, as well as random variation within sites. Figure 2.1 illustrates the difference in reported number of accidents between two sites i and i' in two different time periods t and t'.

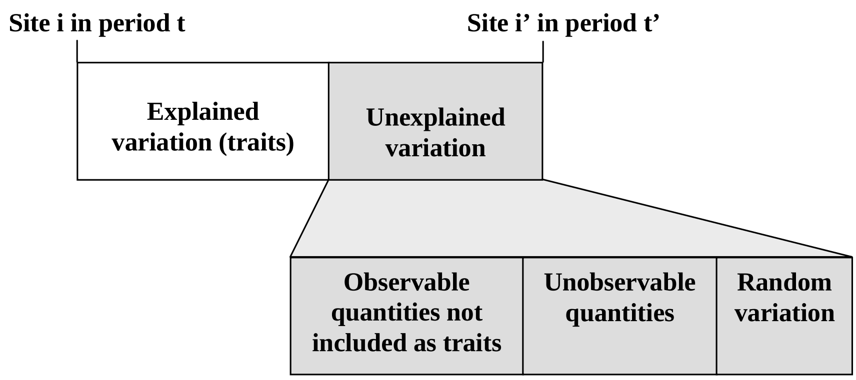


Figure 2.1: Difference in accident counts at two different sites in two different time periods.

By modelling road safety one seeks a better understanding of the factors affecting both the explained and unexplained part of the variation in accident counts, and consequently an increasing ability to separate their effects. The reason for emphasizing this distinction is the fact that while random variation is not influenced by external measures, the non-random variation may be affected by remedial treatment work, thus influencing and hopefully improving the level of safety.

2.2 State of the art review

Road accidents seem to occur at random in time and space. In addition, the probability of an accident occurring at a site in a short period of time (e.g. a second) is constant within this period. These physical conditions match the properties of the Poisson probability distribution. It is thus natural to attempt to model the variation in accident counts by the Poisson distribution. In a group of sites, let X denote the reported number of accidents at a site and let μ denote the common mean for the group:

$$X \in Poiss\,(\mu)$$

However, it soon became evident that the homogeneous Poisson distribution was inadequate because, accident data are more dispersed, i.e. the observed variance in the group of sites exceeds the mean[2]. This is due to the fact that the site safety levels differ between sites in the group. In Ashton (1966) an attempt was given to account for the differences between sites by allowing the Poisson mean to vary randomly between sites in the group. The between-site variation was taken to be gamma distributed with shape parameter α and common mean, μ, for the group:

$$\begin{aligned} X|\lambda &\in Poiss\,(\lambda) \\ \lambda &\in gamma\left(\alpha, \frac{\mu}{\alpha}\right) \end{aligned}$$

The Poisson mean, λ, is thus site-specific, and the number of accidents at a site is negative binomially distributed. The Poisson-gamma model was later supported by a study in Satterthwaite (1976) of the variation in road accident counts in Britain. The study showed the variance to be several times larger than the mean, indicating that the homogeneous Poisson distribution is not applicable for road accidents. Abbess et al. (1981) tested the gamma distribution in the model, set up by Ashton, to be a suitable prior for the mean accident rate, λ.

In the above studies, no site-specific conditions are included in the models. As a consequence, the variation in accident counts between sites is modelled random and given as a mixture of between-site and within-site random variation. Maycock and Hall (1984) allowed for systematic differences between sites in a group by relating the mean in the Poisson distribution to a number of traits, $z_1, ..., z_J$, such as traffic flow and various geometric variables. They suggested a generalized linear model using a pure Poisson error structure:

$$\begin{aligned} X &\in Poiss\,(\mu) \\ \mu &= f_{\beta}\,(z_1, ..., z_J) \end{aligned}$$

[2] In the Poisson distribution the variance equals the mean.

This model belongs to the class of generalized linear models (GLIM) with $\boldsymbol{\beta}$ in $f_{\boldsymbol{\beta}}(\cdot)$ as a set of parameters. Systematic variation between sites is included through the traits, but the remaining differences in accident counts are modelled as random variation within the site. However, Nicholson (1985) found the pure Poisson distribution to be inadequate even when differences in the traits were accounted for. In other words, not all site-specific conditions are included as traits in the model by Maycock and Hall (see figure 2.1). Hauer and Persaud (1987) extended the generalized linear model in Maycock and Hall (1984) by assuming a Poisson-gamma error structure.

$$\begin{aligned} X|\lambda &\in Poiss(\lambda) \\ \lambda &\in gamma\left(\alpha, \frac{\mu}{\alpha}\right) \\ \mu &= f_{\boldsymbol{\beta}}(z_1, ..., z_J) \end{aligned}$$

In this model, the Poisson mean was allowed to vary between sites in the group beyond what may be explained by differences in the traits. The Poisson-gamma GLM set up in Hauer and Persaud (1987) has been supported by Maher and Summersgill (1996), and is now widely accepted (see e.g. Kulmala (1995) and Hauer (1997)).

A few extensions to this model have been pursued. As an example, in Tunaru (1999), the parameter in the gamma distribution as well as the regression variables are modelled as random variables themselves, thus adding several levels to the model. However, most modelling of traffic accidents use a generalized linear model with negative binomial error structure, and no specific modelling of dispersion effects (see e.g. Abdel-Aty and Radwan (2000)).

In the models above, only the total reported number of accidents in the observation period is used, and traits are thus modelled as constant within this period. In practice, traits (in particular traffic flows) often change over time. In order to account for such changes, one may wish to divide the observation period into sub-periods. However, because accident counts, in different sub-periods at the same site, depend on the same site-specific conditions not reflected in the traits, they are not independent (see Maher and Summersgill (1996)). This poses difficulties in estimating the models, as accident counts no longer are independently negative binomially distributed. Consequently, models disaggregated on time have not been developed. An exception is Persaud (1994) where each year's accident count is used as separate records. However, Persaud made the unrealistic assumption of independence between yearly accident counts at the same site.

The models proposed in this chapter are disaggregated on sub-periods of one year, but assume yearly accident counts at the same site to be dependent. The random variation is described by a hierarchical Poisson-gamma distribution, but the Poisson mean is separated into a fixed and a dispersion part (a parametrization also used in Hauer (2001)).

2.3 The accident models

The models proposed below for describing variation in accident counts belong to a class of models called hierarchical generalized linear models (HGLMs), introduced by Lee and Nelder (1996) (see appendix A for a general description).

The road network consists of two main categories of sites; *intersections* and *road sections*. Their main difference being that road sections have a spatial dimension in form of a length. In modelling variation, one distinguishes between intersections and road sections but the outline of the accident models is similar. Both model types are based on data disaggregated over time, i.e. on data corresponding to successive sub-periods. The totality of the sub-periods at a site forms the observation period for this site. In this study, sub-periods of one calendar year is used. This way yearly changes in traffic as well as in other traits may be accounted for. In addition, general trends in the accident count may also be included in the model. In other words, models disaggregated on time allow the accident generating process to be non-stationary.

Intersections and road sections may further be divided into groups of sites of the same type, e.g. into junctions and roundabouts and into motorways and motortrafficways[3]. In the remainder of the thesis, these groups are referred to as *site-groups*. Each site-group has its own set of traits, and the variation in the reported number of accidents is modelled within the site-group.

Let H be a site-group of I sites $\{i\}_{i=1,...,I}$ and let x_{it} denote the number of reported accidents at site i in a given year t in the observation period $]0; T_i]$. The observation period and corresponding years $t \in \{1, 2, ..., T_i\}$ at site i are illustrated in figure 2.2.As

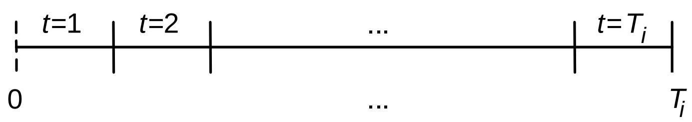


Figure 2.2: Observation period for site i.

an example, a site with yearly accident counts for the period 1994-98 (both years included) has the observation period $]1993; 1998]$ and years $t \in \{1994, 1995, ..., 1998\}$. The length of the observation period is 5 years. The intersection and road section models are developed below.

[3]Motortrafficways are high standard single or dual carriageway roads with similar user constraints to those on motorways.

2.3.1 Intersections

First, a model for describing the variation in accident counts at intersections is developed. For a given intersection i and year t in a site-group H in the road network, the number of reported accidents, x_{it}, may be considered a realization of a random variable X_{it}, which varies randomly around its mean, λ_{it}. This mean is partly dependent on the traits of site i in year t, through μ_{it}, and on a non-observable dispersion effect s_i:

$$\lambda_{it} = \mu_{it} s_i \tag{2.1}$$

Using the terminology of Lee and Nelder (1996), μ_{it} and s_i are denoted *the fixed effect part* respectively the *dispersion part* of the mean, λ_{it}, at site i in year t.

The variation in accident counts is modelled by the Poisson distribution (see Nicholson and Wong (1993)). The conditional distribution of X_{it} given the dispersion effect s_i is thus modelled by the Poisson distribution with mean λ_{it}:

$$X_{it}|s_i \in Poiss\left(\lambda_{it}\right)$$

The fixed effect part, μ_{it}, in (2.1) represents the explained variation in accident counts (see figure 2.1). Assume the existence of a group of intersections with similar traits[4] as site i, i.e. a reference population (using the terminology of Hauer (1997)). The element μ_{it} is then the mean accident frequency at sites with similar traits as site i and is denoted the *reference safety* at site i in year t. In accordance, the expected number of accidents at site i, λ_{it}, is denoted the *site safety* at site i in year t.

The dispersion effect, s_i, in (2.1) represents the site-specific conditions not included as traits, i.e. the non-random part of the unexplained variation (see figure 2.1). It expresses the deviation of the expected accident frequency at site i, the site safety level λ_{it}, from the expected accident frequency at sites with similar traits, i.e. from the reference safety level μ_{it}. The dispersion effect is modelled by a random variable, S_i, which is gamma distributed with mean 1. For intersection i in site-group H, the accident model in year t is:

$$\begin{aligned} X_{it}|s_i &\in Poiss\left(\lambda_{it}\right) \\ S_i &\in gamma\left(\alpha, \frac{1}{\alpha}\right) \\ \lambda_{it} &= \mu_{it} s_i \end{aligned} \tag{2.2}$$

The dispersion effect, s_i, is modelled as random between sites but remains constant within a site from one year to another. It models interdependence between accident counts at the same site in different years over the observation period. In other words, accident counts are independent from site to site but not within the same site. The

[4] Similar traits do in fact mean similar *values* of the traits.

mean of S_i is assumed to be 1 in order to normalize the site safety level to the corresponding reference safety level, $E(\lambda_{it}) = \mu_{it}$.

The parameter, α, in the gamma distribution for s_i is denoted the dispersion parameter and is specific for the site-group. The reciprocal of the dispersion parameter is the variance of the dispersion effect:

$$V(S_i) = \frac{1}{\alpha} \tag{2.3}$$

The parameter, α, is thus a measure of dispersion of the individual site safety levels from the corresponding reference safety levels within the site-group.

The reference safety indices of the sites in the site-group are modelled by a number of traits, specific for this site-group. Assume J different traits, $z_1, ..., z_J$, are included. The reference safety at site i in year t is calculated as:

$$\mu_{it} = f_{\boldsymbol{\beta}}(z_{it1}, ..., z_{itJ}) \tag{2.4}$$

where z_{itj} represents the value of trait j at site i in year t. The function $f_{\boldsymbol{\beta}}(\cdot)$ relates the traits to the reference safety level through a set of parameters $\boldsymbol{\beta} = \{\beta_1, ..., \beta_j\}$ (see appendix A for a description of the form of (2.4)). The set of parameters $\boldsymbol{\beta}$ is specific for each site-group.

According to Lee and Nelder (1996), for given accident counts and parameters α and $\boldsymbol{\beta}$, the dispersion effect at site i is estimated as:

$$s_i = \frac{\alpha + x_{i\cdot}}{\alpha + \mu_{i\cdot}} \tag{2.5}$$

where $x_{i\cdot}$ is the total reported number of accidents at site i within the observation period, and $\mu_{i\cdot}$ is the corresponding expected number in the reference population.

The model described by (2.2) and (2.4) is called the Poisson-gamma hierarchical generalized linear model and was developed by Lee and Nelder (1996). A detailed description of the model along with methods for estimating the dispersion effects and parameters are developed in appendix A.

The marginal distribution of accidents at intersection i in year t, X_{it}, is the negative binomial distribution (see e.g. Lee (1994)):

$$X_{it} \in NB\left(\alpha, \frac{\alpha}{\alpha + \mu_{it}}\right)$$

with mean $E(X_{it}) = \mu_{it}$ and variance $V(X_{it}) = \mu_{it} + \mu_{it}^2/\alpha$. However, accident counts are not independently negative binomially distributed. The variance in the marginal distribution may be expressed as:

$$V(X_{it}) = E(V(X_{it}|s_i)) + V(E(X_{it}|s_i)) = E(\lambda_{it}) + V(\lambda_{it})$$

Because the mean of the dispersion effect is 1, the mean of the site safety level, λ_{it}, equals the mean of X_{it}, i.e. $E(\lambda_{it}) = \mu_{it}$. As a consequence, including a dispersion effect in the model causes the variance of X_{it} to exceed its mean. The variance of the site safety level, $V(\lambda_{it}) = \mu_{it}^2/\alpha$, is thus an expression of overdispersion in a Poisson distribution with mean μ_{it}.

2.3.2 Road sections

The structure of the accident model for road sections differs from the intersection model only in that the length of the road section now is included in the distribution of the dispersion effect. In addition, the expected number of accidents on a road section is modelled proportional with the length of the road.

Hence, for a given road section i of length L_i in a site-group H in the road network, the number of reported accidents, x_{it}, in year t may be considered a realization of a random variable X_{it}, which varies randomly around its mean, $\lambda_{it} L_i$. Again, the mean is divided into a fixed and a random part:

$$\lambda_{it} = \mu_{it} s_i \tag{2.6}$$

and the site and reference safety levels, λ_{it} and μ_{it}, now represent the expected accident number per unit length in year t. In this study a unit length of one kilometer is used.

As for intersections, the conditional distribution of X_{it} given the dispersion effect s_i is modelled by the Poisson distribution. The dispersion effect is modelled by a random variable, S_i, which is gamma distributed with shape parameter αL_i and mean 1 (see appendix A). For road section i of length L_i in site-group H, the accident model in year t is:

$$\begin{aligned} X_{it}|s_i &\in Poiss\left(\lambda_{it} L_i\right) \\ S_i &\in gamma\left(\alpha L_i, \frac{1}{\alpha L_i}\right) \\ \lambda_{it} &= \mu_{it} s_i \end{aligned} \tag{2.7}$$

Analogous to the intersection model, the reference safety indices of the road sections in the site-group are modelled by a number of traits, $z_1, ..., z_J$, specific for this site-group. The reference safety level at road section i in year t is calculated as:

$$\mu_{it} = f_{\boldsymbol{\beta}}\left(z_{it1}, ..., z_{itJ}\right) \tag{2.8}$$

with z_{itj} representing the value of trait j at site i in year t. The shape parameter in the gamma distribution now varies between the road sections in the site-group, by a known factor L_i. The Poisson-gamma hierarchical generalized linear model in (2.7) and (2.8) is thus a weighted model with dispersion parameter α and weight L_i. The above structure of the model was also suggested in Hauer (2001) for describing the variation in accident counts on road sections on a more aggregated level. Including the length, L_i, of road section i in the distribution of S_i, has the effect that the variance of the dispersion effect decreases with the length of the road:

$$V(S_i) = \frac{1}{\alpha L_i} \tag{2.9}$$

The interpretation is that the longer the road, the lesser is the deviation of site safety from the reference safety level. Consequently, the influence of a road section on the estimates of the model parameters is proportional to its length.

For accident counts $x_{i\cdot}$ and given parameters α and $\boldsymbol{\beta}$, the dispersion effect at road section i of length L_i is estimated as:

$$s_i = \frac{\alpha L_i + x_{i\cdot}}{\alpha L_i + \mu_{i\cdot} L_i} = \frac{\alpha + x_{i\cdot}/L_i}{\alpha + \mu_{i\cdot}} \tag{2.10}$$

where $x_{i\cdot}/L_i$ is the total reported number of accidents per kilometer at site i in the observation period, and $\mu_{i\cdot}$ is the corresponding expected number at sites with similar traits as road section i.

The marginal distribution of accidents at road section i in year t, X_{it}, is the negative binomial distribution:

$$X_{it} \in NB\left(\alpha L_i, \frac{\alpha L_i}{\alpha L_i + \mu_{it} L_i}\right)$$

with mean $E(X_{it}) = \mu_{it} L_i$ and variance $V(X_{it}) = \mu_{it} L_i + \mu_{it}^2 L_i/\alpha$. Here $\mu_{it}^2 L_i/\alpha$ is an expression of overdispersion in a Poisson distribution with mean $\mu_{it} L_i$. Again accident counts are not independently negative binomial distributed.

2.3.3 Example

The intersection model above is illustrated by an example based on accident and site data from 2,944 state and regional junctions for the period 1994-98. Within this period, a total of 6,351 accidents have been reported. For each junction, yearly data on the reported number of accidents and traffic flow are available[5]. In addition road geometry information such as the frontage, channelisation and yield relations is given. The annual reported number of accidents at a site in the dataset ranges from 0 to

[5] For more than 2/3 of the junction arms, the average annual daily traffic has been recorded for each of the years 1994-98. The AADT in the remaining years is calculated from the AADT available, using the Danish national traffic growth index for average yearly increases in traffic (see Vejdirektoratet (2002b)).

18 accidents with an average of 0.45. The average annual daily traffic ranges from 250 to 60,206 vehicles per day on major arms and from 250 to 26,647 on minor arms. The road geometry traits of the first 4 junctions in the dataset are listed in table 2.1.

Site	**No. arms**	**Frontage**	**Yield relations**		**Channelisation**	
			Major	**Minor**	**Major**	**Minor**
1	3	None	None	Other	Yes	None
2	3	Industry	None	Other	None	None
3	4	Industry	Signal	Signal	Yes	Yes
4	3	None	None	Other	Yes	None
⋮	⋮	⋮	⋮	⋮	⋮	⋮

Table 2.1: Road geometry traits of the first 4 junctions in the dataset.

In table 2.1, abbreviations are used to indicate the road geometry traits. A detailed description of the abbreviations is listed in table 2.2. Number of arms, frontage as well as yield relations and channelisation are categorical variables, and may assume the levels listed in table 2.2.

Variable name	**Abbrev.**	**Variable**	
No. arms			
3	3	-	
4	4	no.arms4	
5	5	no.arms5	
Frontage			
Na, none/scarce ribbon development or road side development with no frontage	None	-	
Industry or shops	Industry	front1	
Urban/low buildings or residences and flats	Urban	front2	
Yield relations		**Major**	**Minor**
NA or none	None	-	-
Signal controlled	Signal	yield.ma1	yield.mi1
Other	Other	yield.ma2	yield.mi2
Channelisation		**Major**	**Minor**
NA or none	None	-	-
Yes	Yes	chan.ma	chan.mi

Table 2.2: Levels of the categorical variables.

The first level of a variable is used as reference[6]. The reference level is not given a variable name. As an example, the presence of 3 arms in a junction is not indicated,

[6]Included in the intercept.

while number of arms equal to 4 or 5 is indicated in table 2.2. Individual site observation periods range from 3 to 5 years. Annually measured traffic flow and reported accident number for the first 4 junctions are given in table 2.3.

Site		**1994**	**1995**	**1996**	**1997**	**1998**
1	Major AADT	25,428	29,496	27,726	28,558	28,178
	Minor AADT	2,580	2,657	2,684	2,392	2,265
	Accident count	1	2	2	2	1
2	Major AADT			3,518	3,588	3,854
	Minor AADT			692	716	733
	Accident count			0	1	0
3	Major AADT	23,756	26,364	26,891	27,698	27,824
	Minor AADT	10,341	10,502	11,063	11,419	11,620
	Accident count	10	6	10	9	6
4	Major AADT		21,226	19,952	20,153	20,355
	Minor AADT		1,232	1,245	1,420	2,826
	Accident count		0	1	1	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮

Table 2.3: Accident and traffic data for the first 4 junctions.

A variable, γ, indicating the average annual trend in accident development is included in the model. The first year of the study, 1994, is used as base value, e.g. the years 1994,...,1998 are entered with values 0,...,4. The annual change in safety due to trends in time is thus $\gamma - 1$, with a negative value indicating an average annual decrease in the expected accident count. The accumulated change $\triangle t$ years after the base year is:

$$\gamma^{\triangle t} - 1$$

For categorical variables, let $I(z_j)$ indicate whether or not trait j is present at the site, then the model structure of the fixed effect part of the mean at a junction in a given year is (see appendix D for the general structure of the junction model):

$$\begin{aligned} \mu &= a \cdot \gamma^{\triangle t} \cdot AADT_{ma}^{b_1} \cdot AADT_{mi}^{b_2} \cdot \exp\left(\beta_1(1) \cdot I(\text{no.arms4})\right. \\ &\quad + \beta_1(2) \cdot I(\text{no.arms5}) + \beta_2(1) \cdot I(\text{front1}) + \beta_2(2) \cdot I(\text{front2}) \\ &\quad + \beta_3(1) \cdot I(\text{yield.ma1}) + \beta_3(2) \cdot I(\text{yield.ma2}) + \beta_4(1) \cdot I(\text{yield.mi1}) \\ &\quad \left. + \beta_4(2) \cdot I(\text{yield.mi2}) + \beta_4(2) \cdot I(\text{chan.ma}) + \beta_6(1) \cdot I(\text{chan.mi})\right) \end{aligned}$$

Here $AADT_{ma}$ and $AADT_{mi}$ represent the traffic flow on major and minor arms respectively.

2.4 Estimating the site safety index

For a given year τ in the observation period of site i, one may estimate the site safety level, $\lambda_{i\tau}$, at the site for this period. For a year τ beyond the observation period, the site safety level may be predicted, provided the value of the traits is known for this future year. The observation period, $]0; T_i]$, and beyond for site i is illustrated in figure 2.3.

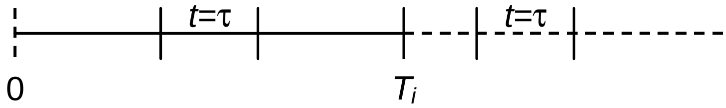


Figure 2.3: Observation period and beyond for site i.

In general, given estimates of α and $\boldsymbol{\beta}$, the site safety level at site i in year τ is estimated (or predicted) as (see (2.1) and (2.6)):

$$\widehat{\lambda}_{i\tau} = \widehat{\mu}_{i\tau}\widehat{s}_i$$

The reference safety level, $\mu_{i\tau}$, is estimated (or predicted) from the traits in year τ using (2.4) or (2.8):

$$\widehat{\mu}_{i\tau} = f_{\widehat{\beta}}\left(z_{i\tau 1}, ..., z_{i\tau J}\right) \tag{2.11}$$

while the dispersion effect, s_i, is estimated from the traits and accident counts of all years in the observation period $]0; T_i]$ (see (2.5) and (2.10)). Consequently, the predicted site safety level of a year beyond the observation period is still based on the accident and site data of the observation period. For simplicity, in the remainder of this chapter τ is assumed to be a year within the observation period.

From the estimates in (2.5) and (2.10), the expected accident frequency at a junction i, respectively per kilometer at a road section i, in year τ is estimated as:

$$\widehat{\lambda}_{i\tau} = \widehat{\mu}_{i\tau}\widehat{s}_i = \begin{cases} \widehat{w}_i\widehat{\mu}_{i\tau} + (1-\widehat{w}_i)\frac{x_{i\cdot}}{\widehat{\mu}_{i\cdot}}\widehat{\mu}_{i\tau} \\ \widehat{w}_i\widehat{\mu}_{i\tau} + (1-\widehat{w}_i)\frac{x_{i\cdot}/L_i}{\widehat{\mu}_{i\cdot}}\widehat{\mu}_{i\tau} \end{cases} \tag{2.12}$$

The site safety estimates in (2.12) are, in an empirical Bayesian framework, the empirical Bayes estimators of site safety (see appendix A). The estimate $\widehat{\lambda}_{i\tau}$ is based on the total reported number of accidents, $x_{i\cdot} = \sum_{t=1}^{T_i} x_{it}$, as well as on the corresponding estimated reference safety level, $\widehat{\mu}_{i\cdot} = \sum_{t=1}^{T_i} \widehat{\mu}_{it}$. The reference safety for the observation period, $\mu_{i\cdot}$,is estimated from the traits of each year in $\{1, ..., T_i\}$:

$$\widehat{\mu}_{i\cdot} = \sum_{t=1}^{T_i} \widehat{\mu}_{it} = \sum_{t=1}^{T_i} f_{\widehat{\beta}}\left(z_{it1}, ..., z_{itJ}\right) \tag{2.13}$$

In other words, the estimated site safety for the year τ is based on information for the whole observation period. As a consequence, changes in traffic flow as well as in other traits from one year to another are taken into account. The element $\widehat{w}_i$ in (2.12) is a weight between 0 and 1 and will be discussed below.

Assume site i has an observation period of T_i years. The site safety levels for the whole observation period $]0; T_i]$ are estimated as:

$$\widehat{\lambda}_{i\cdot} = \sum_{\tau=1}^{T_i} \widehat{\lambda}_{i\tau} = \widehat{\mu}_{i\cdot} \widehat{s}_i = \begin{cases} \widehat{w}_i \widehat{\mu}_{i\cdot} + (1 - \widehat{w}_i)\, x_{i\cdot} \\ \widehat{w}_i \widehat{\mu}_{i\cdot} + (1 - \widehat{w}_i)\, x_{i\cdot}/L_i \end{cases} \tag{2.14}$$

The expected number of accidents at site i in the observation period, $\widehat{\lambda}_{i\cdot}$, is calculated as a weighted average of the total expected number in the reference population, $\widehat{\mu}_{i\cdot}$, and the total reported number at the site, $x_{i\cdot}$. On road sections, $\widehat{\lambda}_{i\cdot}$ and $\widehat{\mu}_{i\cdot}$ represent the total expected accident numbers per kilometer, and hence the reported number per kilometer $x_{i\cdot}/L_i$ comes out in the expression.

It is well established that accident counts are uncertain estimates of site safety because of the random variation in accident counts. This fact is in accordance with the so-called *Stein result* (see appendix A). The Stein result states that for a group of sites, the empirical Bayes estimator of a random variable at a site i, such as the site safety $\widehat{\lambda}_{i\cdot}$, based on data for the whole site-group, through $\widehat{\mu}_{i\cdot}$, has a smaller mean squared error than the maximum likelihood estimator, $x_{i\cdot}$, based only on data for the site. The estimates in (2.12) and (2.14) are thus better estimators of safety than the corresponding accident counts.

The weight, $\widehat{w}_i$, in (2.12) and (2.14) depends on the estimated reference safety at site i and on the estimated variation of the dispersion effects within the site-group. For both intersections and road sections, the weight is estimated as:

$$\widehat{w}_i = \frac{\widehat{\alpha}}{\widehat{\alpha} + \widehat{\mu}_{i\cdot}}$$

In the estimates of the site safety levels in (2.14), the total reported accident number is regressed towards the mean accident frequency in the reference population. Thus the weight, $\widehat{w}_i$, may be interpreted as an estimated *regression parameter* representing the so-called *regression to the mean effect* (see appendix B). Figure 2.4 illustrates the regression towards the mean for 5 intersections with identical traits.The picture in figure 2.4 also applies to road sections. In that case, the total reported number of accidents, $x_{i\cdot}$, at a road section i is divided by its length, L_i.

The regression parameter, w_i is increasing with the dispersion parameter, α. The value of α is an indication of the variation in the dispersion effect, s_i, at site i (see (2.3) and (2.9)). Using the fact that $\lambda_{i\tau} = \mu_{i\tau} s_i$, then for $\mu_{i\tau}$ given, the variation over

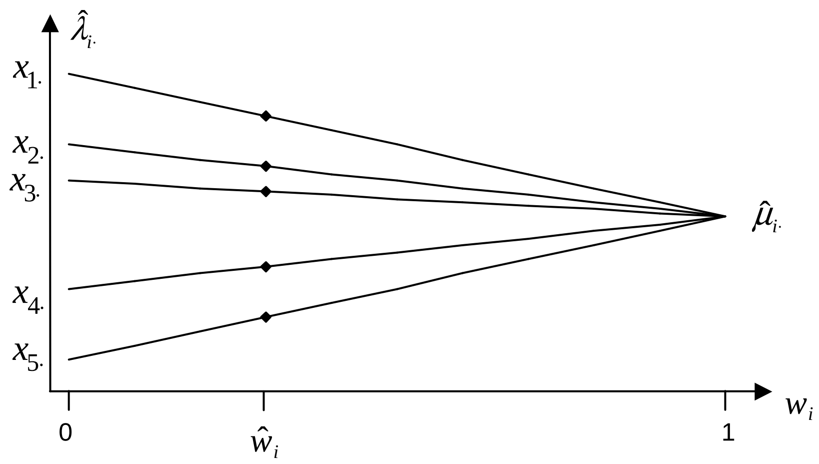


Figure 2.4: The importance of the regression parameter w on the estimated site safety.

sites in the site safety index, $\lambda_{i\tau}$, may be expressed as:

$$Var\left(\lambda_{i\tau}\right) = Var\left(\mu_{i\tau} s_i\right) = \mu_{i\tau}^2 Var\left(s_i\right) = \begin{cases} \mu_{i\tau}^2/\alpha \\ \mu_{i\tau}^2/\alpha L_i \end{cases}$$

Hence, the variance of $\lambda_{i\tau}$ is proportional to the variance of the dispersion effect, s_i. Thus α is an indication of the magnitude by which the site safety varies around the corresponding references safety. The interpretation is, that for a small value of α, i.e. a large variation in the dispersion effect, the reported number of accidents (per kilometer for road sections) at a site is a better estimator of safety at the site than the corresponding reference safety level, and vice versa. The reference safety for the observation period, $\mu_{i\cdot} = \sum_{t=1}^{T_i} \mu_{it}$, at site i is increasing in T_i. Thus for a stable α over time, the regression parameter, w_i, is decreasing in the length of the observation period, T_i, and the site safety index for the observation period, $\lambda_{i\cdot}$, approaches the reported accident count, $x_{i\cdot}$ ($x_{i\cdot}/L_i$ for road sections). The interpretation is that the confidence in accident counts for expressing safety at a site increases over time.

The estimates in (2.12) and (2.14) combines the two sources of information (reference and site) in a way that regards are given to both the length of the observation period at the site (trough the reference safety level) and to the variation of the dispersion effect (through the dispersion parameter).

2.4.1 Example continued

Assume the parameters of the model in the example above have been estimated (see appendix D). The reference safety level for the total number of accidents at a junction in a given year may now be estimated as:

$$\begin{aligned}\widehat{\mu} &= 0.000127 \cdot 0.97^{\triangle t} \cdot AADT_{ma}^{0.43} \cdot AADT_{mi}^{0.44} \cdot \exp\left(0.54 \cdot I\left(\text{no.arms4}\right)\right. \\ &\quad -0.45 \cdot I\left(\text{no.arms5}\right) - 0.30 \cdot I\left(\text{front1}\right) - 0.24 \cdot I\left(\text{front2}\right) \\ &\quad -1.95 \cdot I\left(\text{yield.ma1}\right) - 1.10 \cdot I\left(\text{yield.ma2}\right) + 2.92 \cdot I\left(\text{yield.mi1}\right) \\ &\quad \left.+0.81 \cdot I\left(\text{yield.mi2}\right) + 0.14 \cdot I\left(\text{chan.ma}\right) + 0.33 \cdot I\left(\text{chan.mi}\right)\right)\end{aligned}$$

The time trend variable, γ, is estimated to be 0.97, indicating an average annual decrease in the expected number of accidents of approximately 3%. Other things being equal, this results in an estimated decrease of 14% in the expected number of accidents over a five year period. As an example, because the year 1998 is 4 years after the base value 1994 ($\triangle t = 4$), the reference safety level at junction 3 in 1998 is estimated as:

$$\begin{aligned}\widehat{\mu}_{3,1998} &= 0.000127 \cdot 0.97^{4} \cdot 27824^{0.43} \cdot 11620^{0.44} \cdot \exp\left(0.54-\right. \\ &\quad \left.0.30 - 1.95 + 2.92 + 0.14 + 0.33\right) = 3.11\end{aligned}$$

Consequently, one would expect 3.11 accidents per year at junctions with the same traits as junction 3 in 1998. The reported number of accidents at the site that year was 6. Table 2.4 lists the estimated reference safety levels for the 4 sites in table 2.1.

Site	$\widehat{\mu}_{i,1994}$	$\widehat{\mu}_{i,1995}$	$\widehat{\mu}_{i,1996}$	$\widehat{\mu}_{i,1997}$	$\widehat{\mu}_{i,1998}$
1	0.82	0.85	0.81	0.76	0.71
2			0.12	0.12	0.12
3	3.03	3.09	3.10	3.09	3.02
4		0.53	0.50	0.52	0.68

Table 2.4: Estimated reference safety levels.

The total reported number of accidents at site 3 in the observation period is 41, while the corresponding reference safety level is 15.33. The dispersion parameter for junctions, α, is estimated to be 1.83 (see appendix D). Hence, using (2.5), the dispersion effect, s_3, at junction 3 is estimated as:

$$\widehat{s}_3 = \frac{1.83 + 41}{1.83 + 15.33} = 2.50$$

Hence, the site safety level at site 3 is estimated to be 2.50 times its reference safety level. The estimated dispersion effects and site safety indices etc. for the 4 junctions in table 2.1 are given in table 2.5.

Site	T_i	$\widehat{s}_i$	$\widehat{\lambda}_{i,1994}$	$\widehat{\lambda}_{i,1995}$	$\widehat{\lambda}_{i,1996}$	$\widehat{\lambda}_{i,1997}$	$\widehat{\lambda}_{i,1998}$	$\widehat{w}_i$	$x_{i\cdot}$	$\widehat{\mu}_{i\cdot}$	$\widehat{\lambda}_{i\cdot}$
1	5	1.70	1.38	1.45	1.38	1.29	1.21	0.32	8	3.95	6.72
2	3	1.30			0.15	0.15	0.15	0.84	1	0.35	0.46
3	5	2.50	7.56	7.72	7.73	7.70	7.54	0.11	41	15.33	38.26
4	4	0.94		0.50	0.47	0.49	0.64	0.45	2	2.23	2.10

Table 2.5: Estimated dispersion effects and site safety levels etc.

The estimated regression parameter, $\widehat{w}$, indicates the relative weight of the reference safety level over the corresponding accident count in the site safety estimate. From table 2.5 it can be seen that at site 2 very little weight is put on the accident count. This is due to the fact that the observation period at site 2 is relatively short together with a relatively low estimated reference safety level.

The estimated site safety indices in relation to the reference safety indices are illustrated in figure 2.5.The time indices are omitted, as the values of the site safety

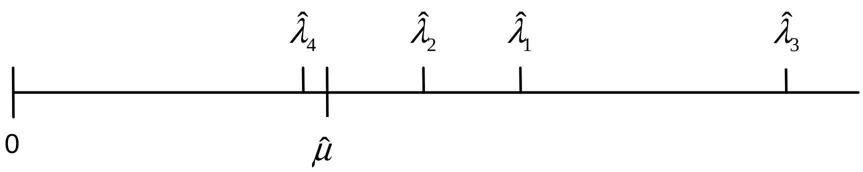


Figure 2.5: The estimated site safety indices relative to the reference safety at the sites.

levels relative to the reference safety levels are unchanged over the observation periods and determined by the estimated dispersion effects, $\widehat{s}$. The estimated site safety index at site 3 is relatively farthest away from its estimated reference safety level.

Chapter 3

Targeting hot spots

The purpose of this chapter is to set up methods for targeting accident-prone locations in the road network, the so-called hot spots. The foundation for the methods is the models developed in chapter 2.

The definition of safety used in this chapter is equal to that of chapter 2. Hence, *site safety* and *reference safety* are defined as the expected number of accidents at a site in a given year (and kilometer for road sections) and the corresponding expected number at sites with similar traits respectively.

A hot spot is defined as a site with an unusually low level of safety. In order to detect whether or not a site is a hot spot, it is compared to a reference site with similar traits. Under the models in chapter 2, this corresponds to comparing the site safety level to the corresponding reference safety level. Methods for targeting hot intersections and hot spots on road sections are developed. The proposed methods use the estimated dispersion effect to indicate whether or not a site should be targeted as a hot spot. The corresponding uncertainty in the estimate is determined from the estimated conditional distribution of the dispersion effect.

A review of the development in the methods used for targeting and ranking accident hot spots is given below.

3.1 State of the art review

At first, sites were ranked according to their reported number of accidents, x. Sites with a number exceeding a chosen threshold value were targeted as hot spots (see e.g. Jorgensen (1966)). This method is very sensitive to random variation in accident counts, and later, the expected number of accidents, λ, estimated from a model, was used instead. However, it is well established that, in general, there are considerable differences between the expected number of accidents at different types of intersections and road sections. As an example, motorways are generally safer than rural roads. The method, based on a threshold value for the expected accident number at sites, will result in the same types of sites being targeted each time. This may be inexpedient, as the most effective solution will end up altering the site into a different road type. Such alterations are often too expensive and/or impossible. In addition, one may argue that the fact that some types of roads are safer than others is already known in the road planning phase. Hence, one might expect that the most suitable type of roads has already been selected, and only minor alterations of the site are feasible.

Instead McGuigan (1981) suggested ranking sites according to their potential for accident reduction (PAR). At a site, this measure is calculated as the difference between the reported number of accidents at the site and the expected number at sites with similar traits, μ (μL for road sections):

$$PAR = x - \mu$$

McGuigan estimated μ from a standard regression model[1] including traffic flow as the only trait. The accident count, x, is the estimated level of site safety at the site.

With the developments made in accident models, the estimates of site safety were improved. As an example, Persaud et al. (1999) suggested using an empirical Bayes estimate, λ, instead of the accident count in PAR. Persaud et al. used a Poisson-gamma generalized linear model with traits such as traffic flow and various geometric variables (see Hauer and Persaud (1987) for details of the model).

The task of targeting hot spots, may be viewed as a ranking and selection problem (see e.g. Dudewicz and Koo (1987)), and parallel with the PAR-method, Gupta and Hsu (1980) introduced the so-called probability of correct selection (PCS). In a group of sites, H, a subset $\widetilde{H} \subseteq H$ is targeted as hot spots, if the probability of hereby selecting the *worst* site is above a chosen threshold value:

$$PCS\left(\widetilde{H}\right) = \Pr\left(worst\ site \in \widetilde{H}\right)$$

By the *worst* site is meant the site with the largest expected number of accidents, λ. Gupta assumed accidents to be normal distributed. Later Hauer and Persaud (1984) derived the probability of correct selection for a Poisson-gamma model (see Ashton (1966) for a description of the model). However, the PCS in Hauer and Persaud (1984) was used as a measure of the overall efficiency of the targeting method and not directly used for targeting hot spots. Schlüter et al. (1997) derived the PCS for an individual site as the posterior probability of being "worst" in a Poisson-gamma model with no traits. Heydecker and Wu (2001) later extended the PCS measure in Schlüter et al. (1997) to include traits, and defined PCS as the probability of the Poisson rate exceeding a chosen threshold value.

[1] Assuming normal distributed errors.

A few alternative methods for targeting hot spots have been proposed. For a given treatment measure, Heydecker and Wu (1993) suggested ranking sites according to the posterior probability that accidents occurring at the site involve the feature the measure is aimed at. Heydecker and Wu assumed a Poisson-beta model with no traits. In Persaud and Kazakov (1994) sites are targeted, if the estimated economical benefit of treating the site exceeds a threshold value based on the allocated budget.

3.2 Methods for targeting hot spots

The methods proposed in this section for targeting intersections and road sections are based on the following general definition of a hot spot:

Definition 1 *A hot spot is a place in the road network with a site safety level below a critical level of safety.*

The term *critical* in definition 1 is not to be taken literally. It does not necessarily mean that levels of site safety above the critical level are non-critical, but it is simply a level used for comparison. The critical level of safety at a site may be calculated as the expected level at sites with similar traits multiplied by a constant c. Here c reflects the amount by which, the decision-maker believes, the site safety level at a site must differ from the safety level at sites with similar traits in order to be critical. The value of c is thus politically determined.

3.2.1 Intersections

First, a method for targeting intersections is derived. In accordance with chapter 2, let λ_{it} and μ_{it} denote the site safety level and reference safety level respectively at an intersection i in year t. Intersection i is then a hot spot in year t if $\lambda_{it} > c\mu_{it}$. Thus, c times the reference safety level, μ_{it}, is used as the critical level. Using (2.1) of chapter 2, one gets the following result:

$$\lambda_{it} > c\mu_{it} \Leftrightarrow s_i > c$$

Hence, $s_i > c$ indicates that intersection i is a hot spot. Because s_i applies to all years t in the observation period, the value of the dispersion effect, s_i, is an expression of the general level of hotness at intersection i in the observation period. A large s_i indicates a high level of hotness. Consequently, definition 1 of a hot spot may be restated for intersections:

Definition 2 *A hot spot is an intersection in the road network with a dispersion effect exceeding c.*

For a group of intersections, plotting the dispersion effects for each intersection gives a quick overview of the sites being hot spots, as depicted in figure 3.1. Here an s_i in the shaded area above c corresponds to a hot spot i.The value of c expresses the

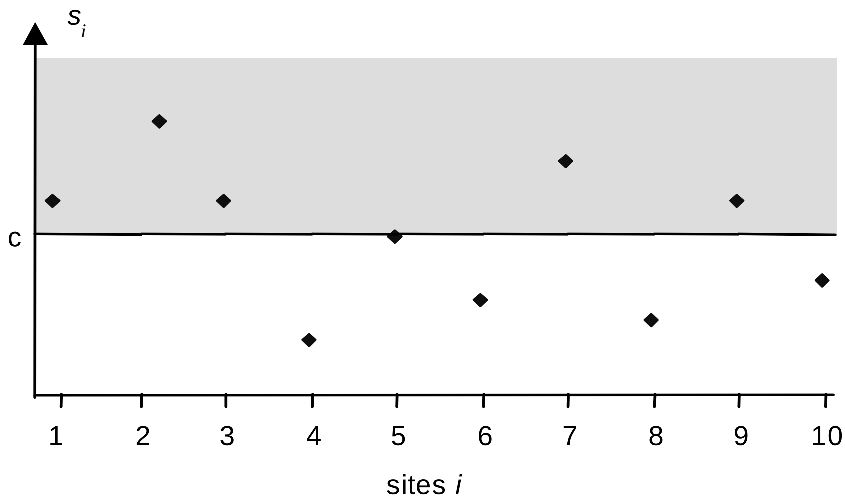


Figure 3.1: Values of the dispersion effect for 10 different sites.

proportion of intersections one wishes to identify as hot spots. The larger the value of c, the smaller is the proportion of intersections targeted.

In reality, the dispersion effect, s_i, is an unknown unobservable model quantity that can only be estimated. Intersections are targeted from the estimated dispersion effects, which are based on the accident counts and traits of the sites. It is important to distinguish between true hot spots in the road network and intersections targeted as hot spots. The latter only includes a subset of the first. Below, the term *accident hot spot* defines a site targeted as a hot spot.

A dispersion effect that measures the deviation of site safety from the reference safety level may be interpreted as a fixed parameter or as a random variable. Under the Poisson-gamma hierarchical generalized linear model for intersections, proposed in chapter 2, dispersion effects are random variables, and for estimated model parameters α and β, the dispersion effect at an intersection i is estimated as (see (2.5)):

$$\widehat{s}_i = \frac{\widehat{\lambda}_{it}}{\widehat{\mu}_{it}} = \frac{\widehat{\alpha} + x_{i\cdot}}{\widehat{\alpha} + \widehat{\mu}_{i\cdot}} \tag{3.1}$$

where $x_{i\cdot}$ is the total number of reported accidents at intersection i in the observation period and $\widehat{\mu}_{i\cdot}$ is the corresponding reference safety estimated from the traits. The estimated site safety level, $\widehat{\lambda}_{it}$, is based on accident and site data for all intersections in the site-group through the estimated dispersion parameter $\widehat{\alpha}$. However, a method for targeting hot spots based solely upon the average observed deviation of accident count from the estimated reference safety level, implicitly assumes that dispersion effects are fixed. In that case, the dispersion effect at an intersection i is estimated as[2]:

$$\widetilde{s}_i = \frac{x_{i\cdot}}{\widehat{\mu}_{i\cdot}} \tag{3.2}$$

Here $\widetilde{s}_i$ is the maximum likelihood estimator of the dispersion effect, and $x_{i\cdot}$ is the estimated site safety level at intersection i based only on data for this site. For $\widehat{\mu}_{i\cdot} \neq 0$, the estimated dispersion effect in (3.1) may be expressed as:

$$\widehat{s}_i = \widehat{w}_i \cdot 1 + (1 - \widehat{w}_i) \frac{x_{i\cdot}}{\widehat{\mu}_{i\cdot}}$$

where $\widehat{w}_i$ is a weight between zero and unity. It is seen that the estimated dispersion effect is a weighted sum of unity and the average observed deviation of accident count from the estimated reference safety level, i.e. the ratio $x_{i\cdot}/\widehat{\mu}_{i\cdot}$. This ratio is equal to the maximum likelihood estimate in (3.2). Because $x_{i\cdot} > \widehat{\mu}_{i\cdot}$ implies $\widehat{s}_i > 1$, (3.1) and (3.2) will, for $c = 1$, target the same set of intersections as hot spots. However, this will not be the case for $c \neq 1$.

The magnitude of the level of hotness, $\widehat{s}_i$, depends upon both the amount of data at site i (through $x_{i\cdot}$ and $\widehat{\mu}_{i\cdot}$) and the variation in accident data within the site-group (through $\widehat{\alpha}$). This is reflected in the weight $\widehat{w}_i$:

$$\widehat{w}_i = \frac{\widehat{\alpha}}{\widehat{\alpha} + \widehat{\mu}_{i\cdot}}$$

As an example, consider two intersections i and j in the same site-group. Let $x_{i\cdot}$ and $x_{j\cdot}$ denote the reported number of accidents at the sites, and let $\widehat{\mu}_{i\cdot}$ and $\widehat{\mu}_{j\cdot}$ denote the corresponding estimated reference safety levels. Assume $x_{j\cdot}/\widehat{\mu}_{j\cdot} = x_{i\cdot}/\widehat{\mu}_{i\cdot} > 1$, and define n as the rate between the length of the observation periods at site j and i respectively, $n = T_j/T_i$. If the observation period at site j includes more years than at site i ($n > 1$), the estimated level of hotness at site j is higher than at site i:

$$\widehat{s}_j = \frac{\widehat{\alpha} + x_{j\cdot}}{\widehat{\alpha} + \widehat{\mu}_{j\cdot}} = \frac{\widehat{\alpha} + n x_{i\cdot}}{\widehat{\alpha} + n\widehat{\mu}_{i\cdot}} > \frac{\widehat{\alpha} + x_{i\cdot}}{\widehat{\alpha} + \widehat{\mu}_{i\cdot}} = \widehat{s}_i$$

In general, the level of hotness is increasing with n, when the accident count exceeds the reference safety level and vice versa. This is illustrated in figure 3.2. Also, the estimated level of hotness, $\widehat{s}_i$, approaches the maximum likelihood estimate, $\widetilde{s}_i$, as n approaches infinity.In targeting methods based on average evaluations, such as $\widetilde{s}_i$ in

[2]The Danish Road Directorate and regional authorities use $\widetilde{s}_i$ as the estimated level of hotness at a site i (see chapter 5).

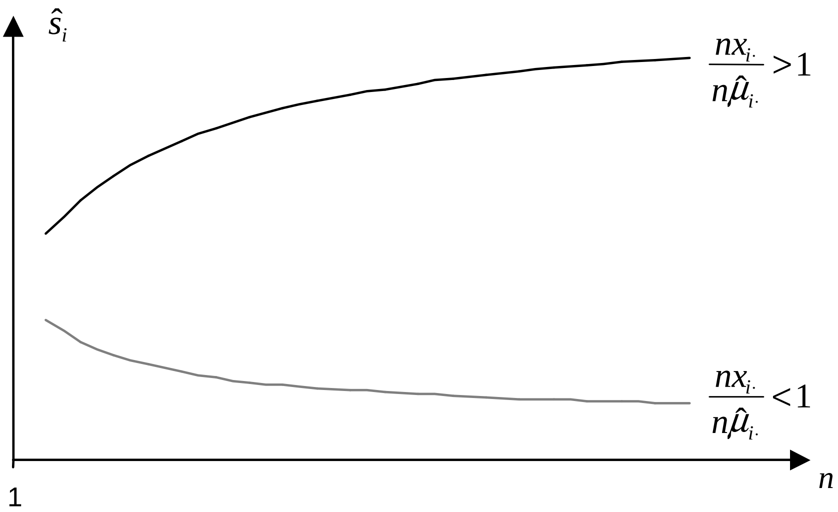


Figure 3.2: Change in level of hotness with increasing n.

(3.2), no considerations are given to the amount of data available at the site or to the variation in accident data within the site-group.

If both the accident count and reference safety level at a site i are relatively small compared to the estimated dispersion parameter, $\widehat{\alpha}$, then the estimated level of hotness, $\widehat{s}_i$, is close to unity even when the accident count deviates markedly from the reference safety level. Consequently, dependent on the chosen critical level of dispersion, c, intersection i must have a relatively large number of reported accidents in order to be targeted. Furthermore, the estimated dispersion effect, $\widehat{s}_i$, is a dimensional quantity, as it measures the ratio between the estimated site and reference safety level based on the same traits. Hence, estimated dispersion effects at different sites in different site-groups are directly comparable. The estimated dispersion effect is a function of the estimated model parameters, $\widehat{\alpha}$ and $\widehat{\boldsymbol{\beta}}$, the traits and the reported number of accidents at the site. Once the model parameters α and $\boldsymbol{\beta}$ are estimated, $\widehat{s}_i$ is easily calculated from (3.1).

In the estimated dispersion effect, the weight of the accident count, $x_{i\cdot}$, at site i is reduced by a factor $\widehat{w}_i$, which is dependent on the length of the observation period at the site. In other words; the ratio $x_{i\cdot}/\widehat{\mu}_{i\cdot}$ is regressed[3] towards the common mean of 1. According to the so-called *Stein result* and results of simulation (see appendix A and E), the estimate in (3.1) is a better[4] estimate of dispersion than the ratio $\widetilde{s}_i = x_{i\cdot}/\widehat{\mu}_{i\cdot}$ in (3.2). This is a result of the fact that the estimate $\widetilde{s}_i$ uses the crude accident count

[3]Also known as *Bayesian shrinkage*.

[4]In the sense that that is has a smaller mean squared error of estimation.

$x_{i\cdot}$ as measure of site safety at site i in the observation period. In fact, the estimate $\widehat{s}_i$ is the best linear unbiased estimator of s_i in terms of the *mean squared error* of estimation (see appendix A). However, the estimate of the dispersion effect, $\widehat{s}_i$, is still subject to uncertainty because of the random variation in accident data.

Uncertainty in the estimated dispersion effect

In order to make sure that intersections targeted as hot spots do in fact have a level of hotness exceeding the critical level, c, one needs to take the uncertainty in $\widehat{s}_i$ into account. Let $c_i^* = c + k\xi_i$ be the adjusted critical level for $\widehat{s}_i$, with ξ_i representing the uncertainty in the estimate, $\widehat{s}_i$. The parameter $k \in \mathbb{R}$ expresses the demanded level of certainty for targeting an intersection as a hot spot. A large positive k indicates that only intersections with a relatively high evidence of being a hot spot are targeted, while a large negative k indicates that all sites are under suspicion, and only intersections with a high evidence of not being a hot spot are omitted. The adjusted critical level is now site-dependent through the uncertainty of the estimated dispersion effect. Consequently, definition 2 may be restated as a rule for targeting intersections as hot spots:

Rule 1 *Intersection i is an accident hot spot, if the estimated dispersion effect, $\widehat{s}_i$, exceeds c_i^*.*

The degree of uncertainty, ξ_i, is an indication of how well $\widehat{s}_i$ estimates the true dispersion effect s_i. The level of uncertainty differs from intersection to intersection, i.e. some $\widehat{s}_i$ are estimated with more precision than others. For estimated parameters α and $\boldsymbol{\beta}$, this uncertainty is expressed in the estimated variance of $\widehat{s}_i$:

$$\widehat{V}\left(S_i|\mathbf{x}_i\right) = \frac{\widehat{\alpha} + x_{i\cdot}}{\left(\widehat{\alpha} + \widehat{\mu}_{i\cdot}\right)^2}$$

However, in the case of targeting hot spots, one is only interested in whether or not the targeted intersection is in fact a hot spot. Under definition 2, this corresponds to whether or not the dispersion effect is greater than c. The obvious measure is the estimate (see appendix A):

$$\widehat{P}_i \equiv \widehat{P}_i\left(S > c \mid x_{i\cdot}\right) = \int_c^\infty \frac{\left(\widehat{\alpha} + \widehat{\mu}_{i\cdot}\right)^{\widehat{\alpha}+x_{i\cdot}}}{\Gamma\left(\widehat{\alpha} + x_{i\cdot}\right)} s^{\widehat{\alpha}+x_{i\cdot}-1} e^{-(\widehat{\alpha}+\widehat{\mu}_{i\cdot})s} ds \tag{3.3}$$

The value of $\widehat{P}_i$ in (3.3) indicates the strength of evidence of intersection i being a hot spot. In the following, $\widehat{P}_i$ is denoted the evidence of hotness (EOH) of $\widehat{s}_i$, and the interval $]c;\infty]$ is the estimated empirical Bayes confidence interval for $\widehat{s}_i$ of level $\widehat{P}_i$. The EOH of $\widehat{s}_i$ takes into consideration both the position and the uncertainty in the estimate. Hence, intersections may be targeted directly from the evidence of hotness. Assume that for a given c, one demands an evidence of hotness of at least d, then rule 1 may be restated as:

Rule 2 *Intersection i is an accident hot spot, if the evidence of hotness is at least d.*

Analogous to k in c_i^* of rule 1, d is a threshold value expressing the demanded level of certainty of a targeted intersection for being a hot spot. However, in the targeting phase one is still at the risk of making two kinds of wrong decisions:

1. Targeting an intersection which in reality **is not** a hot spot (a cool spot).
2. Not targeting an intersection which in reality **is** a hot spot.

The value of d expresses the concern of the decision-maker for targeting cool spots relative to the concern for not targeting a hot spot (see Aagaard (1997) for a discussion). Analogous to diagnostic procedures (see e.g. Fletcher et al. (1996)), the *sensitivity* of the targeting method is described by the probability that a given hot spot is in fact targeted (equal to 1 minus the probability of making wrong decision 2). The *specificity* of the method is described by the probability that a given *cool* spot is not targeted (equal to 1 minus the probability of making wrong decision 1). The targeting method is good, if both the sensitivity and the specificity are high. However, there is a trade off between the sensitivity and specificity of a targeting method, which may be influenced by the value of d. A small value of d increases the number of intersections targeted as hot spots. This increases the probability of a given hot spot being targeted, hence increasing the sensitivity. Unfortunately, it also increases the probability of a given cool spot being targeted, hereby decreasing the specificity of the targeting method. Methods for choosing the value of d will not be discussed further[5] in this thesis.

Two different accident hot spots may have the same estimated level of hotness but with different evidence of hotness. As an example, consider two sites i and j from the same site-group with $(x_{i\cdot}, \widehat{\mu}_{i\cdot}) = (2,1)$ and $\left(x_{j\cdot}, \widehat{\mu}_{j\cdot}\right) = (5,3)$ respectively. Assuming an $\widehat{\alpha} = 1$ one has $\widehat{s}_i = \widehat{s}_j = 1.5$, but for $c = 1$ one has $\widehat{P}_i = 0.68$ while $\widehat{P}_j = 0.79$. Intersection j has a higher evidence of being a hot spot than intersection i, despite the fact, that the ratio, $x_{i\cdot}/\widehat{\mu}_{i\cdot}$, between accident count and corresponding reference safety at site i is larger here than at intersection j. Using (3.1) one gets the following general expression:

$$\widehat{s}_i = k \Leftrightarrow \widehat{\mu}_{i\cdot} = \frac{1}{k} x_{i\cdot} + \left(\frac{1}{k} - 1\right) \widehat{\alpha} \tag{3.4}$$

[5] However, the *Reciever Operating Characteristics* (ROC) curves is a technique used in diagnostic procedures for selecting an optimal value of d. The ROC curves are based on the sensitivity and specificity of the method for different values of d (see Swets and Pickett (1992)).

where k is a constant. Expression (3.4) shows, that an increase in the total reported number of accidents, $x_{i\cdot}$, by n units, requires an increase in $\widehat{\mu}_{i\cdot}$ of n/k units to maintain the same level of hotness, $\widehat{s}_i$. It may be shown, that if the relation between $\widehat{\mu}_{i\cdot}$ and $x_{i\cdot}$ in (3.4) is maintained, $\widehat{P}_i$ is increasing with n for $k > c$ and decreasing for $k < c$. As an example, let $c = 1$ and $\widehat{\alpha} = 2$ then figure 3.3 illustrates the change in evidence of hotness for increasing values of n. The shade of the contour area indicates the value of EOH, i.e. the darker the colour the higher is the evidence of hotness. The value of n is increasing in the direction of the arrows.

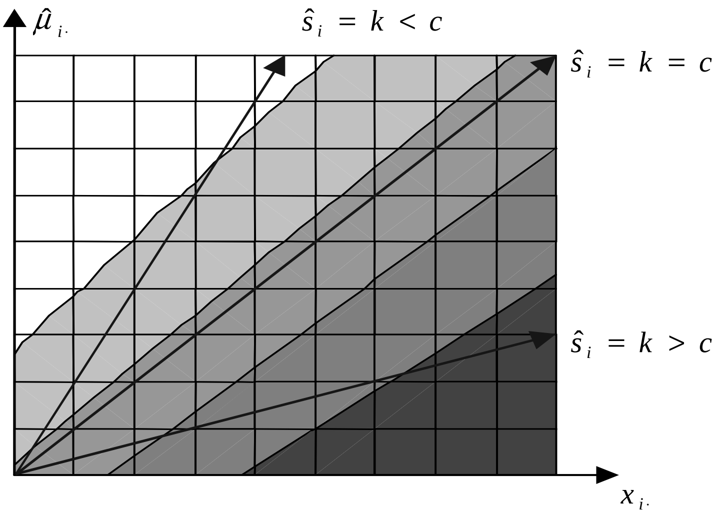


Figure 3.3: Change in evidence of hotness (EOH) with n, for different fixed values of the estimated dispersion effect, $\widehat{s}_i = (\widehat{\alpha} + nx_{i\cdot}) / (\widehat{\alpha} + n\widehat{\mu}_{i\cdot})$. The EOH is increasing in the shade of the contour area, and the value of n is increasing in the direction of the arrows.

The interpretation is that the confidence in the estimated dispersion effect, for expressing whether or not intersection i is an accident hot spot, increases with more accident data. Furthermore, like the estimated variance of the dispersion effect, the evidence of hotness depends on the variation in site safety within the site-group through $\widehat{\alpha}$. As an example, consider two sites i and j from two different site-groups, with estimated dispersion parameters $\widehat{\alpha}_i = 1$ and $\widehat{\alpha}_j = 10$ respectively. Assume the two sites have the same total number of reported accidents and estimated reference safety level, $(x_{i\cdot}, \widehat{\mu}_{i\cdot}) = (x_{j\cdot}, \widehat{\mu}_{j\cdot}) = (2, 1)$. For $c = 1$ this results in $\widehat{P}_i = 0.68$ while $\widehat{P}_j = 0.58$. In general, for fixed values of $x_{i\cdot}$ and $\widehat{\mu}_{i\cdot}$ the evidence of hotness is decreasing with $\widehat{\alpha}$ for $c > 1$, and increasing for $c < 1$. This is illustrated in figure 3.4.A large $\widehat{\alpha}$ indicates that the reference safety is a good measure of site safety,

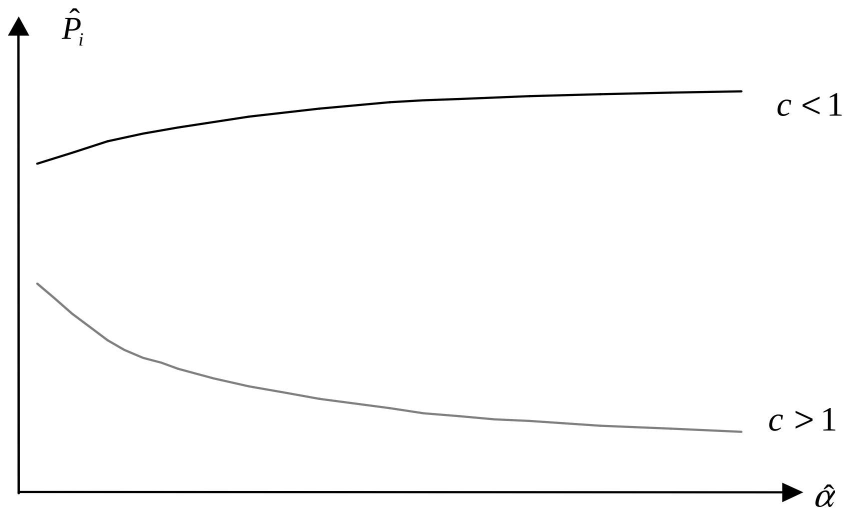


Figure 3.4: Change in evidence of hotness for fixed values of $x_{i\cdot}$ and $\widehat{\mu}_{i\cdot}$.

i.e. the confidence in accident data at the site for expressing the site safety level is relatively low and vice versa. In addition, like the level of hotness, the evidence of hotness is a dimensional quantity, which is directly comparable at different sites in different site-groups.

Summarizing the properties above, the advantages of a method for targeting hot spots defined by rule 2 are:

- Both the amount of data at the intersection and the variation between intersections in the site-group are considered in the method.
- Only intersections with a relatively[6] large number of reported accidents may be targeted.
- The estimated dispersion effect, $\widehat{s}_i$, is a better estimator of dispersion than the crude ratio $\widetilde{s}_i = x_{i\cdot}/\widehat{\mu}_{i\cdot}$. Consequently, the method based on the evidence of hotness is less sensitive towards random variation than a method based on $\widetilde{s}_i$.
- Intersections from different site-groups are directly comparable.

The advantages listed above are closely connected with that of the model describing variation in accidents on intersections (see chapter 2).

Simulations show that for intersections, targeting hot spots from the estimated dispersion effect, $\widehat{s}$, leads to the same sensitivity as when the evidence of hotness

[6] Dependent on the chosen values of c and d.

(EOH) is used (see appendix E). In fact, both methods will target the same group of sites. However, this is not the case for road sections as described below.

Example continued

Consider the junctions in the example of chapter 2. Assume all junctions with an estimated site safety level, $\widehat{\lambda}$, exceeding the corresponding reference safety level, $\widehat{\mu}$, are considered potential hot spots ($c = 1$). In practice, however, one is able to treat only a subset of the junctions. Hence, an evidence of hotness, $\widehat{P}$, of at least 90% is required in order to target a site as a hot spot. As an example, from the estimated dispersion parameter, $\widehat{\alpha} = 1.83$, accident count and estimated reference safety levels, the evidence of hotness at junction 1 is calculated as:

$$\widehat{P}_1 \equiv \widehat{P}_1 \left(S > 1 \mid x_{1\cdot} \right) = \int_1^{\infty} \frac{(1.83 + 3.95)^{1.83+8}}{\Gamma\left(1.83 + 8\right)} s^{1.83+8-1} e^{-(1.83+3.95)s} ds = 0.92$$

The calculated evidence of hotness for the 4 junctions in table 2.1 are listed in table 3.1 along with the estimated dispersion effect etc.

Site	T_i	$x_{i\cdot}$	$\widehat{\mu}_{i\cdot}$	$\overline{x}_i$	$x_{i\cdot}/\widehat{\mu}_{i\cdot}$	$\widehat{s}_i$	$\widehat{P}_i$
1	5	8	3.95	1.60	2.02	1.70	**0.92**
2	3	1	0.35	0.33	2.82	1.30	0.58
3	5	41	15.33	8.20	2.67	2.50	**1.00**
4	4	2	2.23	0.50	0.90	0.94	0.39

Table 3.1: Evidence of hotness etc.

From table 3.1, it can be seen that junctions 1 and 3 are targeted as hot spots for treatment. In this case, they are also the sites with the highest annual average reported number of accidents. The 4 sites are ranked similarly using the estimated dispersion effect, $\widehat{s}_i$, or the evidence of hotness, $\widehat{P}_i$. It can be seen that even though junction 2 has the highest ratio $x_{i\cdot}/\widehat{\mu}_{i\cdot}$, it is not the site with the highest evidence of hotness $\widehat{P}_i$. The reason for this is the relatively short period of observation at junction 2 (3 years), i.e. the data basis is relatively low compared to e.g. site 1 with an observation period of 5 years.

The purpose of targeting hot spots is to improve safety at the sites through remedial treatment. For each accident hot spot, the potential safety measures are to be found. Thus one needs a sufficient amount of accident occurrences to detect a common risk factor, which again will indicate the relevant measure for this site. Hence, in practice, only sites with a reported number of accidents exceeding a fixed $x_{\min}$ are considered in the targeting phase.

3.2.2 Road sections

As described in appendix A, the site safety on a road section may vary over the length of the road. As for intersections, site safety on a road section is given as the product of its reference safety and its dispersion effect (see chapter 2). The reference safety is modelled as constant over the length of the road because the traits are constant[7]. Consequently, the dispersion effect varies with position as well. The variation in site safety and dispersion effect on a road section i of length L_i in year t is illustrated in figure 3.5.

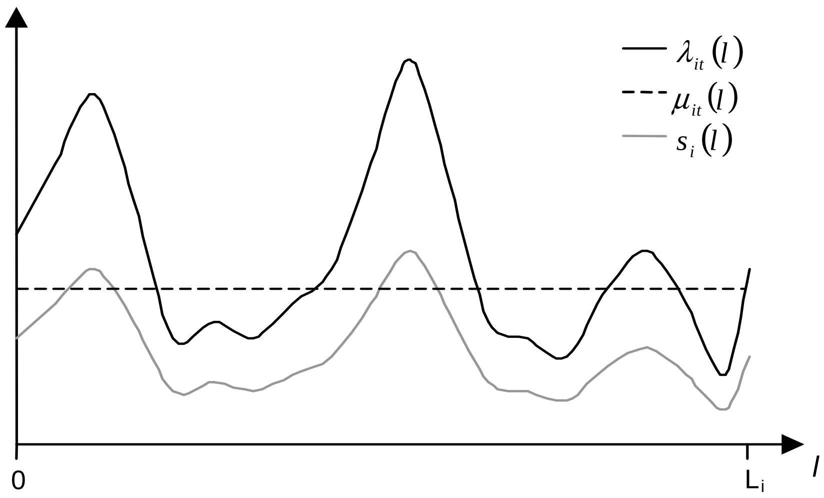


Figure 3.5: The variation in site safety and in the dispersion effect over road section i in year t.

In accordance with definition 1, a *hot section* is defined as a road section for which any part of the road has a site safety level below the critical level of safety. Again, the critical level may be expressed as the reference safety level multiplied by a constant, c. Because the reference safety level is constant over the length of the road section, this corresponds to comparing the dispersion effect to the constant c. For a chosen level, c, definition 1 may be restated for road sections:

Definition 3 *A hot section is a road section in the road network where at least one sub-section of the road has a dispersion effect exceeding c.*

A sub-section on a road section with a dispersion effect exceeding c is referred to as a hot spot. There may be several hot spots on one hot section, each with a potential for safety improvement.

[7] The change in site characteristics defines the division of the road network into road sections.

The change in the value of the dispersion effect on a road section i, s_i, is illustrated in figure 3.6.It is seen, that on sub-sections A_1 and A_2 the dispersion effect exceeds

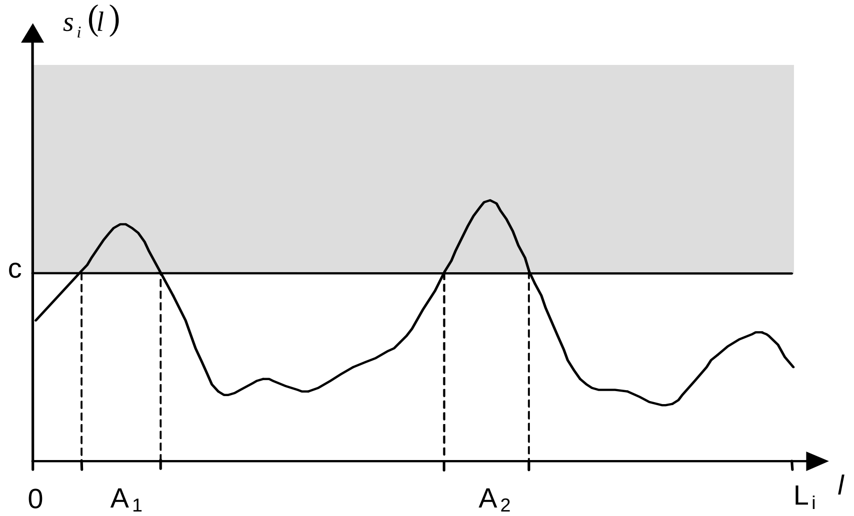


Figure 3.6: Variation in the dispersion effect, s_i, over road section i.

c. Consequently, A_1 and A_2 are hot spots and site i is a hot section. In figure 3.6, let D_{A_j} denote the area beneath the graph of s_i on sub-section A_j:

$$D_{A_j} = \int_{A_j} s_i(l)\, dl$$

If l_{A_j} is the length of sub-section A_j, then D_{A_j}/l_{A_j} is an expression of the level of hotness on sub-section A_j, and $D_{A_j}/l_{A_j} > c$ is a sufficient condition of sub-section A_j being a hot spot. Define $s_{iA_j} = D_{A_j}/l_{A_j}$, then s_{iA_j} is the average dispersion effect on sub-section A_j, i.e. the level of hotness, and is calculated as:

$$s_{iA_j} = \frac{\int_{A_j} s_i(l)\, dl}{l_{A_j}}$$

The parameter s_{iA_j} is an unknown unobservable quantity that can only be estimated. In analogy with intersections, an *accident hot spot* is defined as the part of a road section targeted as a hot spot. Consequently, *accident hot sections* are road sections with one or more accident hot spots.

For estimated model parameters α and β, the dispersion effect of a sub-section of length l on road section i is estimated as (see appendix A):

$$\widehat{s}_{il} = \frac{\widehat{\alpha} l + x_{i\cdot l}}{\widehat{\alpha} l + \widehat{\mu}_{i\cdot} l} = \widehat{w}_i \cdot 1 + (1 - \widehat{w}_i)\, \frac{x_{i\cdot l}}{\widehat{\mu}_{i\cdot} l}, \tag{3.5}$$

where $x_{i \cdot l}$ denotes the total number of reported accidents on road section i within the sub-section of length l, and $\widehat{\mu}_{i.}l$ denotes the corresponding expected number on road sections with similar traits and length. Again, $\widehat{w}_i$ in (3.5) is a weight between zero and unity:

$$\widehat{w}_i = \frac{\widehat{\alpha}}{\widehat{\alpha} + \widehat{\mu}_{i.}}$$

The estimated dispersion effect, $\widehat{s}_{il}$, is the estimated level of hotness of a sub-section of length l on road section i. As for intersections, the estimate $\widehat{s}_{il}$ is subject to uncertainty because of the random variation in accident data. Instead, one may use the evidence of hotness of $\widehat{s}_{il}$ for targeting hot spots:

$$\widehat{P}_{il} = \widehat{P}_{il}\left(S > c \mid x_{i \cdot l}\right) = \int_c^{\infty} \frac{\left(\widehat{\alpha} l + \widehat{\mu}_{i.} l\right)^{\widehat{\alpha} l + x_{i \cdot l}}}{\Gamma\left(\widehat{\alpha} l + x_{i \cdot l}\right)} s^{\widehat{\alpha} l + x_{i \cdot l} - 1} e^{-\left(\widehat{\alpha} l + \widehat{\mu}_{i.} l\right) s} ds \tag{3.6}$$

The value of $\widehat{P}_{il}$ expresses the strength of evidence of the sub-section of length l being a hot spot. Assume, that for a given c, one demands an evidence of hotness of at least d in order to target a sub-section as a hot spot. Then definition 3 may be restated as a rule for targeting hot sections:

Rule 3 *An accident hot section is a road section in the road network where at least one sub-section of the road has an evidence of hotness of at least d.*

Assume $\widehat{\alpha}$, $\widehat{\mu}_{i.}$ and c are known, then for given length, l, of a sub-section, the evidence of hotness (EOH) increases for increasing accident count, $x_{i \cdot l}$, and for given accident count, the EOH decreases for increasing length, l.

On a road section i, two sub-sections of length l_1 and l_2 respectively may have the same estimated level of hotness, $\widehat{s}_{il_1} = \widehat{s}_{il_2}$, but still have different evidence of hotness, $\widehat{P}_{il_1} \neq \widehat{P}_{il_2}$. An example is when the first sub-section is twice the length, $l_1 = 2l_2$, and has twice the number of reported accidents, $x_{i \cdot l_1} = 2x_{i \cdot l_2}$ as the latter. In this case, the sub-section of length l_1 has a higher evidence of hotness when the ratio $x_{i \cdot l_1}/\widehat{\mu}_{i.}l_1$ is above the critical level c and vice versa. In general, for a sub-section of length l on road section i, with a reported accident rate per kilometer of $x_{i \cdot l}/l$, one has:

$$\begin{aligned} x_{i \cdot l}/l &> \widehat{\mu}_{i.}c \Rightarrow \widehat{P}_{i,nl}\left(S > c \mid n x_{i \cdot l}\right) > \widehat{P}_{il}\left(S > c \mid x_{i \cdot l}\right), \forall n > 1 \\ x_{i \cdot l}/l &< \widehat{\mu}_{i.}c \Rightarrow \widehat{P}_{i,nl}\left(S > c \mid n x_{i \cdot l}\right) < \widehat{P}_{il}\left(S > c \mid x_{i \cdot l}\right), \forall n > 1 \end{aligned}$$

The reason for this fact is that the confidence in $x_{i \cdot l}/l$ for expressing the true accident rate per kilometer at a sub-section on road section i increases with longer sub-sections. Assume $\widehat{\mu}_{i.} = c = 1$ and $\widehat{\alpha} = 2$, then figure 3.7 illustrates the change in evidence of hotness for increasing values of n. The shade of the contour area indicates the value of EOH, i.e. the darker the colour the higher is the evidence of hotness. The value of n is increasing in the direction of the arrows.

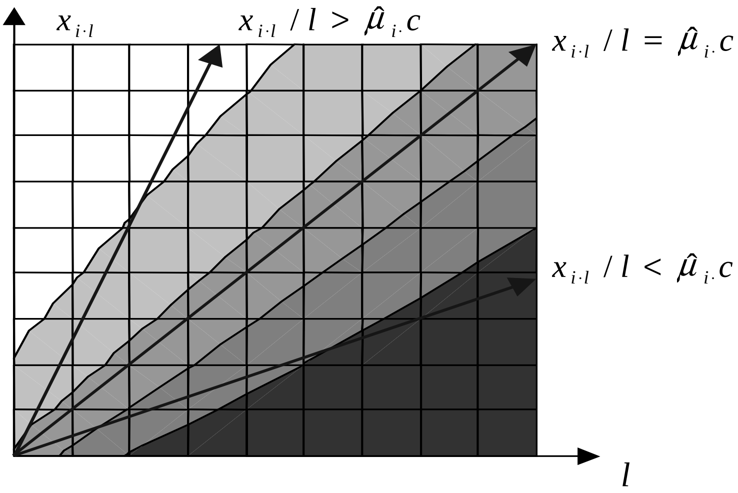


Figure 3.7: Change in evidence of hotness (EOH) with n for different fixed values of $nx_{i\cdot l}/nl$. The EOH is increasing in the shade of the contour area, and the value of n is increasing in the direction of the arrows.

For a road section, the evidence of hotness should be calculated for all sub-sections with a relatively high number of reported accidents. A method for targeting hot spots on road sections is given below.

Targeting hot spots on road sections

The outline of the proposed method for targeting hot spots on road sections is as follows:

(i) For a site-group of road sections, choose a fixed minimum number of reported accidents and a minimum evidence of hotness.

(ii) For each road section in the site-group, the evidence of hotness is calculated for all sub-sections with a reported accident number exceeding the fixed minimum. Subsequently, sub-sections with an EOH exceeding the fixed minimum are selected.

(iii) The sub-sections selected in (ii) are ranked according to their EOH, and hot spots are targeted successively from the top such that no overlap between sub-sections is present.

The method is described in further detail below.

Ad. (i) A fixed minimum, $x_{\min}$, of reported number of accidents is chosen, and a minimum level of evidence of hotness, d, is selected in $[0;1]$. The fixed minimum $x_{\min}$ is to ensure a sufficient number of accident to indicate a trend in the accidents, i.e. to indicate safety measures for treatment. The fixed minimum d is to ensure a sufficient potential for accident reduction.

Ad. (ii) All possible sub-sections j with at least $x_{\min}$ reported number of accidents in the observation period are marked, and the corresponding evidences of hotness are calculated. Assume, road section i has m marked sub-sections, where the evidence of hotness exceeds d, $j = 1, ..., m$. Let $l_1, l_2, ..., l_m$ and $x_{i \cdot l_1}, x_{i \cdot l_2}, ..., x_{i \cdot l_m}$ denote the corresponding lengths and accident counts of the sub-sections on road section i:

Sub-section	Evidence of hotness
1	$\widehat{P}_{il_1} = \widehat{P}_{il_1} \left(S > c \mid x_{i \cdot l_1}\right)$
2	$\widehat{P}_{il_2} = \widehat{P}_{il_2} \left(S > c \mid x_{i \cdot l_2}\right)$
$\vdots$	$\vdots$
m	$\widehat{P}_{il_m} = \widehat{P}_{il_m} \left(S > c \mid x_{i \cdot l_m}\right)$

Ad. (iii) The sub-sections on road section i selected in (ii) are then ranked according to their evidence of hotness in descending order, with $[j]$ being the sub-section on road section i with the j^{th} highest evidence of hotness:

Sub-section	Evidence of hotness
$[1]$	$\widehat{P}_{il_{[1]}} = \widehat{P}_{il_{[1]}} \left(S > c \mid x_{i \cdot l_{[1]}}\right)$
$[2]$	$\widehat{P}_{il_{[2]}} = \widehat{P}_{il_{[2]}} \left(S > c \mid x_{i \cdot l_{[2]}}\right)$
$\vdots$	$\vdots$
$[m]$	$\widehat{P}_{il_{[m]}} = \widehat{P}_{il_{[m]}} \left(S > c \mid x_{i \cdot l_{[m]}}\right)$

sub-section $[1]$ is targeted as a hot spot. The following sub-sections are targeted successively according to the list above, such that sub-section $[j]$ is targeted if no overlap is present with any of the prior targeted sub-sections of higher evidence of hotness than sub-section $[j]$.

The advantage of the method above is that all potential hot spots, i.e. all sub-sections with $x_{\min}$ or more reported number of accidents are evaluated, and the sub-sections with the highest evidence of hotness are targeted as hot spots.

As an example, assume one has chosen $x_{\min} = 3$ and $d = 0.5$. Thus, all sub-sections with 3 or more reported accidents within the observation period are marked. On a road section i of length $L_i = 7.5$ kilometers with a total of 7 accidents on 7 different locations in the observation period, 15 sub-sections are marked[8]. Figure 3.8 illustrates the 15 marked sub-sections on road section i.Assume $\widehat{\mu}_{i\cdot} = c = 1$ and

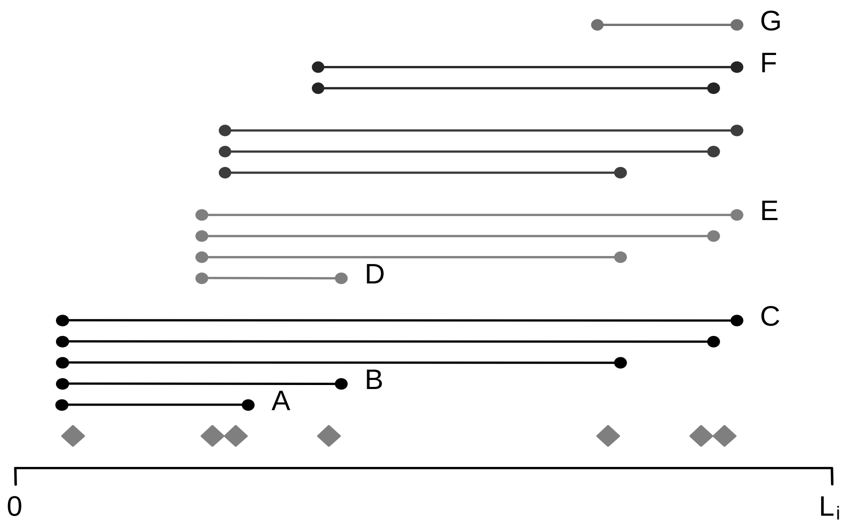


Figure 3.8: Sub-sections on road section i for $x_{\min} = 3$.

$\widehat{\alpha} = 2$ as above, then for each of the 15 marked sub-sections, the evidence of hotness is calculated. Sub-sections with an evidence of hotness exceeding 0.5 are selected. Table 3.2 lists the evidence of hotness for the 7 selected sub-sections corresponding to sub-sections A-G in figure 3.8.

Sub-section	**Accident count.**	**length (km)**	**EOH**
A	3	1.68	0.67
B	4	2.24	0.70
C	7	6.51	0.52
D	3	1.40	**0.72**
E	6	5.04	0.56
F	4	3.64	0.51
G	3	1.54	**0.69**

Table 3.2: Evidence of hotness for the 7 selected sub-sections on road section i.

The selected sub-sections are ranked according to their evidence of hotness in descending order. Sub-section D has the highest evidence of hotness among the 7 selected sub-sections and is targeted as a hot spot. This rules out all remaining

[8]In general, on a road section i with $x_{i\cdot}$ accidents in the observation period, $\sum_{n=1}^{x_{i\cdot}-x_{\min}+1} n$ sub-sections will be marked for further investigation.

sub-sections except for sub-section G. Thus, sub-sections D and G are accident hot spots.

A road between two intersections in the road network may consist of several road sections. There is no reason to believe, that the site safety level will suddenly shift at the border of two successive road sections. Instead, one may assume the site safety level to be a continuous function of position on a road between two intersections. However, the reference safety level is estimated from the traits, which may change from one road section to another. Consequently, the reference safety level may shift downwards or upwards when entering the subsequent road section. As a result, the dispersion effect will shift in level accordingly. Figure 3.9 illustrates such changes over two successive road sections.The proposed method for targeting hot spots on

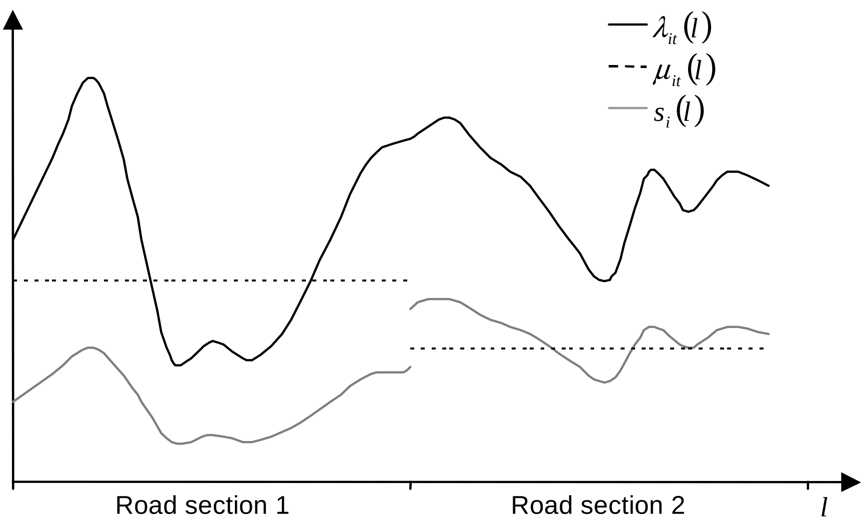


Figure 3.9: The variation in site and reference safety, and in the dispersion effect over two successive road sections in year t.

road sections may be extended to roads consisting of several road sections. In that case, the evidence of hotness of a sub-section may be calculated from accident count and reference safety levels of several different road sections.

As an example, consider a sub-section overlapping two successive road sections as illustrated in figure 3.10.The sub-section consists of two parts, A and B, on road section 1 and 2 respectively, and a total of 4 accidents have been reported on the sub-section within the observation period. Let l_A and l_B denote the lengths of parts A and B, and let $\widehat{\mu}_{1\cdot}$ and $\widehat{\mu}_{2\cdot}$ denote the corresponding estimated reference safety per kilometer within the observation periods of road sections 1 and 2. Let $l = l_A + l_B$ be

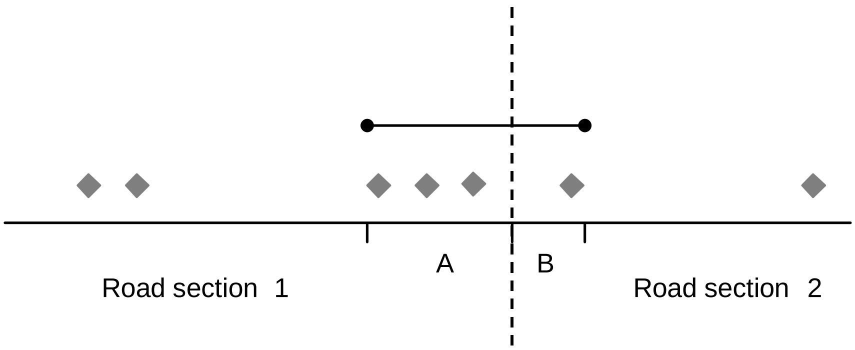


Figure 3.10: A sub-section overlapping road sections 1 and 2 respectively.

the total length of the sub-section, then the evidence of hotness is calculated as:

$$\widehat{P}_{(1,2),l} = \widehat{P}_{(1,2),l}\left(S > c \mid 4\right) = \int_c^\infty \frac{\left(\widehat{\alpha}l + \widehat{\mu}_1.l_A + \widehat{\mu}_2.l_B\right)^{\widehat{\alpha}l+4}}{\Gamma\left(\widehat{\alpha}l + 4\right)} s^{\widehat{\alpha}l+4-1} e^{-(\widehat{\alpha}l+\widehat{\mu}_1.l_A+\widehat{\mu}_2.l_B)s} ds$$

It is assumed that road sections 1 and 2 belong to the same site-group. The evidence of hotness for a sub-section enclosing the border of two road sections is thus calculated as weighted average of the evidence of hotness for the parts in the first and second section respectively.

Contrary to intersections, simulation studies show that for road sections, targeting hot spots from the evidence of hotness leads to a few percent higher sensitivity compared to the method using the estimated dispersion effect (see appendix E).

3.3 Extending the definition of accident hot spots

The rules for targeting hot spots derived above are based on the total number of reported accidents at a site. However, road accidents are connected with a number of features, which may be considered in the targeting phase:

- The severity of the accident, e.g. fatal, injury or property damage only accident.
- The accident category, e.g. rear-end collisions.
- The accident contributory factors present in the accident, e.g. ice on the road.

An accident contributory factor may cause an accident to occur, but will not necessarily do so. As an example, a driver collides with the crash barrier on an icy bend. The obvious conclusion is that the icy curve is the cause of the accident. The driver would probably not have crashed had the curve not been icy, but presumably hundreds of vehicles have passed the curve without crashing. Hence, the ice on the road is not the sole cause of the accident. Accident contributory factors may be defined as (see Transportforskningsudvalget (1968)):

Definition 4 *An accident contributory factor is a necessary but not sufficient condition for certain types of accident occurrences.*

An accident contributory factor obviously affects the probability of accident occurrence and may be assigned to the driver, the site or the vehicle. There are a variety of accident factors and several factors may be present in a single accident. Determining the accident contributory factors will often involve a reconstruction of the accident course, which is a rather complicated and costly task. As a result, information on accident factors is scarce. Accident severity and category are, however, stated in the police report and are thus easily accessible. As an example, injury accidents are considered to have a higher cost to society than property damage only accidents, and they usually have a higher coverage in accident statistics as well. Consequently, it may be efficient to target hot spots from the reported number of injury accidents instead of from the total accident count.

Heydecker and Wu (1991) & (1993) suggest a method for indicating treatments for accident hot spots already targeted, based on the probability of observing a certain feature in an accident. One may use the same line of reasoning for targeting hot spots. However, instead of investigating the probability of observing a certain feature, one may as well investigate the expected number of accidents with the certain feature at a site.

Accidents involving a certain accident feature are subject to random variation as described in chapter 2. The variation in accidents involving a certain accident feature may be described by a model similar to that of chapter 2, by assigning an additional subscript, k, to the variables indicating the feature (i.e. the severity of the accident, the accident situation or the accident factor). Sites are classified into groups of similar road type and accident feature. These groups are sub-sets of the previous site-groups and are denoted *site–feature-groups*[9]. The distributional assumptions of chapter 2 are assumed, and a set of fixed effect parameters, $\boldsymbol{\beta}_k$, and dispersion parameter, α_k, of the site-feature-group are estimated within this group. The estimation is based on accidents in the site-feature-group involving feature k. Hot spots on intersections may be targeted from the corresponding evidence of hotness of an intersection i:

$$\widehat{P}_{ik} \equiv \widehat{P}_{ik}\left(S > c \mid x_{i\cdot k}\right) = \int_c^{\infty} \frac{\left(\widehat{\alpha}_k + \widehat{\mu}_{i\cdot k}\right)^{\widehat{\alpha}_k + x_{i\cdot k}}}{\Gamma\left(\widehat{\alpha}_k + x_{i\cdot k}\right)} s^{\widehat{\alpha}_k + x_{i\cdot k} - 1} e^{-(\widehat{\alpha}_k + \widehat{\mu}_{i\cdot k})s} ds$$

For road sections, the length of the sub-section is included in the evidence of hotness in analogy with the section above.

[9] A site may appear in several site-feature-groups within its site-group.

In targeting hot spots, several features may be taken into account. As an example, assume one wishes to target hot spots based on the total number of accidents, but also taking into account the number of injury accidents. The subscript k then denotes the severity of the accident, and models are set up for each of the levels of severity, injury and property damage only accidents. The corresponding evidences of hotness (EOHs), $\widehat{P}_{i,injury}$ and $\widehat{P}_{i,pdo}$, are used to derive a common measure for comparison, calculated as:

$$\widehat{P}_i \equiv \theta \widehat{P}_{i,injury} + (1-\theta)\, \widehat{P}_{i,pdo}$$

Here, the value of θ reflects the relative importance of injury to property damage only accidents. As an example, one may equal θ to the ratio of the cost to society of the two accident types. Let d denote the critical level for $\widehat{P}_i$, such that site i is targeted as a hot spot, if $\widehat{P}_i$ exceeds d. The parameter set $(\theta, 1-\theta, d)$ defines a hyperplane, and accident hot spots are sites with a $\left(\widehat{P}_{i,injury}, \widehat{P}_{i,pdo}\right)$ above the hyperplane as illustrated in figure 3.11.

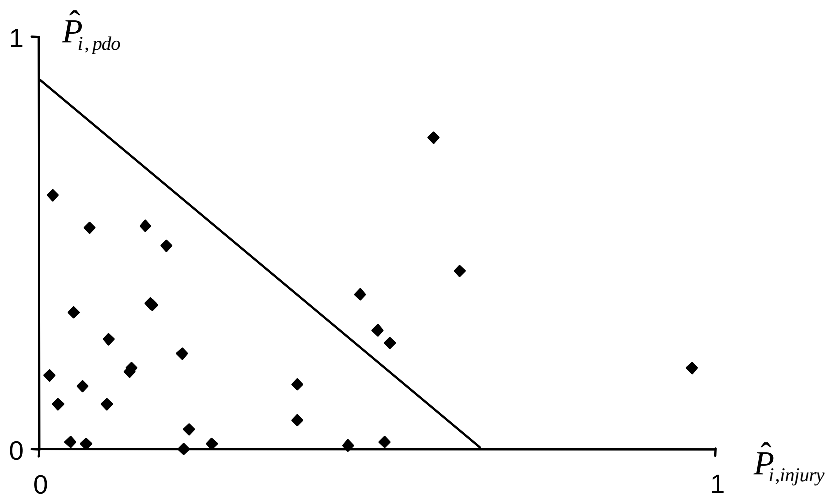


Figure 3.11: Targeting hot spots from the evidences of hotness.

A method for targeting hot spots based on accidents involving certain features may seem more sophisticated than one based on the total number of accidents. However, depending on the number of features included, the number of accidents involved in the targeting method decreases. This may increase the uncertainty in the corresponding estimated dispersion effect and with that the uncertainty in the targeting method.

3.4 Prioritizing

High risk sites are targeted with the aim of improving safety on the road network through remedial treatment of the sites. Safety measures are implemented at accident hot spots in order to improve safety, and any achieved positive effects on safety are denoted the benefits of the implemented measures. Implementing safety measures is costly, but in theory, all measures generating a positive net-benefit should be implemented. However, the restricted funding for hot spot safety work does put a limit to the number of measures that may be implemented, regardless of the fact that in their implementation a surplus is generated. In order to utilize the limited funds as effectively as possible, it is necessary to prioritize between sites and safety measures.

The general objective of remedial work is to improve safety, but different constraints may apply. For instance, the objective may be to reach a certain tolerable level of safety with as few means as possible, or it may be to improve safety as much as possible within a given budget. Additional objectives and constraints may be of relevance, but the general task of prioritizing is to maximize the benefit-cost ratio of a portfolio of safety measures under given constraints.

Let $\mathbf{Y}$ represent a portfolio of safety measures and let $C(\mathbf{Y})$ and $B(\mathbf{Y})$ denote the corresponding overall cost and benefit of $\mathbf{Y}$. Then, the general aim of prioritizing may be expressed as:

$$\max_{\mathbf{Y}} \frac{B(\mathbf{Y})}{C(\mathbf{Y})}$$

Today, the major hot spots have already been eliminated due to previous road safety remedial work, hence the safety tolerability criterion is usually not the decisive criteria. Instead the objective is to improve safety as much as possible with the funds allocated, but without a given target level of safety[10]. This approach is known as the *as low as reasonably practicable* (ALARP) principle (see Melchers (2001)). Here *reasonably practicable* refers to *within the budget*.

Ideally, the cost and benefit of implementing a portfolio of safety measures should be calculated for the entire lifetime of the portfolio, because in maintaining a safety measure, the number of accidents at a site may be reduced for a period of several years. However, it is difficult to estimate the lifetime of all safety measures and often only costs and benefits for the first year are used in the prioritizing process[11] (see e.g. Bernhardt and Virkler (2002)).

The benefit of treatment may be expressed in objective measures such as saved number of accidents, saved accident costs etc. Common to all measures of benefit

[10] Even the Danish vision "Each accident is one too many" or the Swedish "0-vision" (see Færdselssikkerhedskommissionen (2000)) is constrained by a budget rather than by the tolerability criteria of zero traffic accidents.

[11] This is the case in hot spot safety work in Denmark today (see chapter 5 for details).

used in hot spot safety work is that they are based on the current safety levels, $\boldsymbol{\lambda}$, at the sites considered for treatment:

$$B(\mathbf{Y}) = f_{\mathbf{Y}}(\boldsymbol{\lambda}, \boldsymbol{\varepsilon}, \mathbf{p}, ...)$$

In addition to the site safety levels, the reduction rate, ε, in accidents gained from implementing a safety measure is included in $f_{\mathbf{Y}}(\cdot)$. The saved accident costs from implementing a safety scheme is a common measure of benefit, as accidents of different severity may be combined to produce one objective measure. In that case, a price p of each considered type of accident severity is included in $f_{\mathbf{Y}}(\cdot)$.

Pricing accidents of different severity is not a straight forward matter. Fatal accidents are considered to be more costly than non-fatal injury accidents which again are considered more costly than property damage only accidents. The cost of a property damage only accident is relatively easy to obtain, as it is calculated as the sum of the value of material damage. The problem of pricing injury accidents is that no market exists for e.g. a human life, thus no market prices are stated, and instead the pricing of injury accidents is done indirectly (see Dasgupta and Pearce (1972)). There are basically three methods for estimating the costs of injury and death to society (see T10 (2000)):

Implicit values where accidents are priced according to the cost of the methods implemented in trying to avoid them, e.g. priced according to the average cost of the given medical treatment in trying to avoid a person dying, divided by the probability of the treatment being successful (in saving the patient's life). In practice, implicit values of life differ significantly between private roads and public transport because of differences in the sizes of the funds allocated to prevent accidents from happening.

Human capital where the major part of the cost of an injury is the discounted present value of the victim's future output or income lost due to the injury. The additional cost contributors are involving medical treatment, police, property damage and administration costs.

Willingness to pay where the price of accidents are calculated from people's trade-off between road safety and other commodities, e.g. deciding between different modes of transportation with different safety levels. The approach is similar to the method used for valuation of travel time. Evidence on willingness to pay is most commonly obtained by questionnaires, and often the values are much higher than those estimated from the human capital approach (see Jones-Lee (1990)).

The willingness to pay method has been adopted by e.g. Britain, Sweden and Finland, while Denmark still uses the human capital approach for estimating costs of different accident types to society. Values of different accident types are given in Vejdirektoratet (1999c). The advantage of the willingness to pay (WTP) approach is that it reflects the public's concern for safety. Also, because WTP values tend to be higher than implicit values or human capital values, estimated benefits of remedial work are increased, which may increase the priority given to road safety. In practice, the actual cost of injury accident is dependent on the number of people involved in the accident, e.g. on the number of car occupants. However, because this number is independent of the site of the accident, the average cost of an accident type is used.

In addition to the problem of pricing accidents, the obtained accident reduction from implementing a safety measure is uncertain. This uncertainty may be divided into:

(i) The uncertainty concerning the reduction rate in accidents due to the treatment portfolio.

(ii) The uncertainty concerning the extent to which the accident types, the measure is aimed, are in fact present at the site.

(iii) The uncertainty concerning whether or not a site is in fact a hot spot.

The uncertainties in (ii) and (iii) are site-related, while the uncertainty in (i) is linked to the safety measure. It is assumed that sites with a high certainty of being a hot spot have relatively higher potential for accident reduction than sites with a low certainty (see Persaud et al. (1999)). Also, a safety measure aimed at a particular type of accident is assumed to have a relatively higher effect at sites where such accidents are predominant.

In general, before and after studies of the effect of treatment are few. Hence, knowledge of the reduction rate, ε, in accidents from implementing a certain safety measure is often not available, and instead an imputed number is used. The uncertainty concerning the reduction rate due to the treatment portfolio, described in (i), is thus difficult to account for. Additional before and after studies of the effect of treatment are needed to increase knowledge in this area.

The uncertainty described in (ii) about the extent to which the accident types, the measure is aimed, are present at the site, is due to the fact that site safety levels are unknown, unobservable quantities that can only be estimated. As described in the previous chapter, the Poisson-gamma hierarchical generalized linear models (Pg-HGLMs) give better[12] estimates of safety than the crude accident counts (see chapter 2). In addition, the Pg-HGLM estimates of site safety take into account the level of hotness at the sites through the estimated dispersion effects, $\widehat{\mathbf{s}}$. If sites are targeted from their evidence of hotness (EOH) of the dispersion effect, as proposed in this

[12] In terms of the mean squared error of estimation.

chapter, the sites under consideration in the prioritizing phase all have a relatively high certainty of being hot spots. The site safety estimates will thus account for the uncertainties described in (ii) and (iii).

Summing up, targeting hot spots from the method proposed in this chapter and subsequently prioritizing between accident hot spots and measures using the estimated safety levels in (2.12) of chapter 2 has the following advantages:

1. The uncertainty in whether or not a site is in fact an accident hot spot is taken into consideration, because only sites targeted as hot spots from rule 2 or 3 are considered.

2. The potential for accident reduction is considered in the prioritizing phase through the estimated dispersion effects in the site safety estimates.

3.4.1 Example continued

Consider the two accident hot spots, junctions 1 and 3, targeted above. At junction 1 the major part of accidents was reported as collisions between vehicles on major arms and approaching vehicles from the minor arm. The junction has no yield relations on the major arms, while the minor arm has a *give way marking.* A potential safety scheme, Y_1, at this site is to convert the give way marking on the minor arm into a stop sign, and to implement give way markings on the major arms. At junction 3, a large number of accidents were caused by red light violations. A potential safety scheme, Y_3, at this site is to implement a controller to briefly extend the red light for cross traffic.

Due to the limited funds, a budget constraint is the decisive criteria in the prioritizing phase, and for simplicity, the first year benefit (FYB) is used as measure of performance of potential safety schemes. In the FYB-value, the benefit of implementing a safety scheme Y in a junction i is calculated as the expected saved accident cost, AC_{iY}, the first year due to the scheme:

$$AC_{iY} = \varepsilon_Y \cdot \lambda_i \cdot p$$

where ε_Y is the reduction rate in the number of accidents at junction i due to safety scheme Y, and λ_i is the expected number of accidents at junction i (the site safety level). The element p is the average cost to society of one road accident, which in year 2000 was estimated to be 889,000 DKK (see Vejdirektoratet (2002b)). Let C_{iY} denote the cost of implementing scheme Y in junction i. The first year benefit of scheme Y in junction i is then calculated as:

$$FYB_{iY} = \frac{AC_{iY}}{C_{iY}}$$

Assume that prior studies of the effects of scheme Y_1 and Y_3 are available, and that the effect of treatment (ε_Y) on average was 0.75 and 0.20 respectively. The estimated site safety levels in year 1998 of table 2.5 are used as the expected number of accidents at junctions 1 and 3, as no major changes to either traffic flow or road geometry are anticipated in the following years. The benefit of implementing scheme Y_1 in junction 1 is estimated as:

$$AC_{1Y_1} = 0.75 \cdot 1.21 \cdot 889000 = 808276$$

The saved accident costs of implementing scheme Y_1 in junction 1 are thus $808,276$ DKK. The corresponding implementation costs are 1,092,266 DKK, resulting in a first year benefit of:

$$FYB_{1Y_1} = \frac{808,276}{1,092,266} = 74\%$$

The expected saved accident costs and first year benefit of scheme Y_3 in junction 3 are calculated in a similar way. Table 3.3 lists the expected saved accident cost, cost of implementation and first year benefit of safety schemes Y_1 and Y_3.

Scheme	$\widehat{\lambda}_{i,1998}$	ε_Y	AC_{iY}	C_{iY}	FYB_{iY}
1	1.21	0.75	808,276	1,092,266	74%
3	7.54	0.20	1,341,152	1,230,414	109%

Table 3.3: The benefit, cost and first year benefit of the potential safety schemes at junctions 1 and 3.

Table 3.3 shows that safety scheme Y_1 has a higher reduction rate than scheme Y_3. However, because the expected number of accidents at junction 3 is much larger than at intersection 1, implementing scheme Y_3 results in a relatively larger first year benefit. Consider a budget constraint of 1.5 million DKK, then the solution would be to implement safety scheme Y_3 rather than scheme Y_1.

Chapter 4

Before and after studies

The aim of this chapter is to set up a method for estimating the effect of remedial work on road safety. The proposed model compensates for the so-called *regression to the mean effect* in both the period before and after treatment. Let λ_{with} and $\lambda_{without}$ denote the site safety levels with and without treatment at a site in a given time period. Under the models in chapter 2, the effect, ε, of treating the site is calculated as the proportional change in site safety at the site:

$$\varepsilon = 1 - \frac{\lambda_{with}}{\lambda_{without}}$$

The purpose of monitoring the effect of remedial safety work is to understand the effects of treatments, and use the experience as a future guidance for prioritizing between safety measures. Furthermore, estimated effects of different treatments may also be used in the road planning phase. One may imagine a simplified cause-effect diagram of the accident process as depicted in figure 4.1.The diagram may be in-

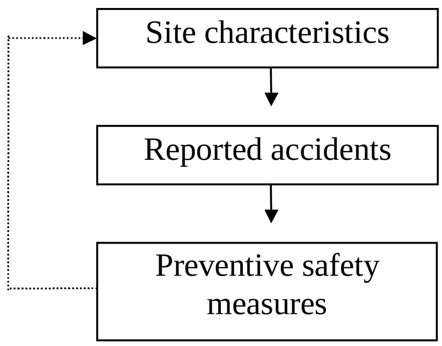


Figure 4.1: Cause-effect diagram.

terpreted as a learning process in which the site characteristics have an impact on the number and types of accidents occurring at the site, which again influence the choice of treatment measures in remedial work. The safety measures may change the site characteristics, often leading to a change in the reported number of accidents. Knowledge of the relations between the elements in figure 4.1 may be used to point out areas with potential accident contributory factors, to be used in both the planning (traffic safety audits[1]) and treatment phase (accident hot spot safety work).

4.1 Disturbing elements

Treating a site may lead to changes in traffic volume, road geometry and other site characteristics thus affecting the accident frequency at the site. However, the literature on before and after studies has pointed out a number of other observable and non-observable elements affecting the change in reported accident count before and after the implementation of preventive safety measures (e.g. Elvik et al. (1997)):

- Overlap in effects
- Negative effects
- Road user behavioural adjustment
- Regression to the mean effect
- Migration effect
- General trends
- Change in the level of reporting

Similar problems are connected with the estimation of the traits in accident models developed in chapter 2 (see appendix D). A detailed description of each of the disturbing elements connected with before and after studies is given below.

4.1.1 Overlap in effects

Different safety measures may be aimed at the same type of accidents thus creating an effect overlap if implemented at the same site. In a before and after study of the effect on the total number of accidents at a site, only the total effect of the measures will show. However, if the purpose is to estimate the individual effects of measures, one needs to account for effect overlap.

[1] See chapter 1 for a description of traffic safety audits.

4.1.2 Negative effects

Accident treatment measures may have negative as well as positive effects on safety. As an example, implementing signal lights in junctions often reduces the number of accidents of crossing vehicles, while increasing the number of rear-end collisions in the junction (see Lahrmann and Leleur (1994)). In a before and after study of the effect on the total number of accidents at a site only the *net-effect* will show.

4.1.3 Road user behavioural adjustment

Road user behavioural adjustment is the problem of road users adjusting their behaviour to a safety measure in such a way that the actual effect of the measure deviates from what was expected. Behavioural adjustment leading to an unchanged level of safety at a site after treatment, is known as risk homeostasis (see Wilde (1986) and Assum et al. (1999)). This phenomenon is strongly connected to the road user's subjective perception of safety. As an example, improving road friction will reduce the stopping distances of cars, which is anticipated to have a substantial positive effect on safety. However, this increased feeling of safety has lead some car users to decrease their distance to the car in front of them, and hereby reducing the actual effect of the measure below the expected (Evans (1994)).

4.1.4 Regression to the mean effect

The regression to the mean (RTM) effect is a statistical phenomenon, which is known in various areas (see e.g. Schall and Smith (2000) for a description of RTM exemplified in baseball performance). In the area of traffic safety, the regression to the mean effect may be explained as follows. Sites with accident counts above (or below) the expected at sites with similar traits one year, will in the following year have accident counts which on average are closer to the expected number of accidents at sites with similar traits. In other words, the accident counts have regressed towards the mean (see Davis (1986) for a general description of RTM).

For illustrative purposes, consider a group of sites with similar traits in a given year. The dispersion of the accident counts at individual sites in the group around the reported overall group mean, $\overline{x}$, is illustrated in figure 4.2. Each square in figure 4.2 corresponds to a site.Assume, one selects a sub-set, M, of sites with accident counts equal to m. If the site characteristics of the whole group remain unchanged in the following year, one would expect the mean accident count in the group to be constant as well, i.e. the overall group mean is still $\overline{x}$.

What about accident counts at sites in sub-set M? If accident counts were perfect estimates of site safety, the accident counts in different years at the same site would be perfectly correlated. Consequently, with unchanged site characteristics one would

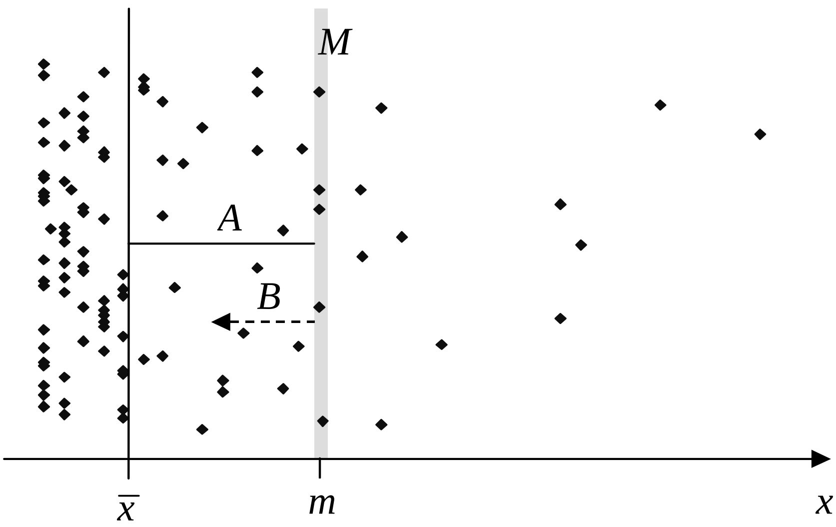


Figure 4.2: Accident counts in a group, M, of sites with similar traits and overall reported mean $\overline{x}$.

expect sites in M with m accident the first year to have m accidents in the following year as well. In that case, no regression to the mean effect is present. However, perfect correlation is not the case in practice because of random variation in accident counts. If, on the other hand, there is no correlation between accident counts at the same site, any deviation between the reported number of accidents in sub-set M and the overall group mean, $\overline{x}$, is purely random. Thus, sites in M will in the following year on average have accident counts $\overline{x}$. In that case, one may say that accidents have regressed all the way back to the group mean. In figure 4.2, this corresponds to the distance A.

In reality, there is some, but imperfect, correlation. The correlation between yearly accident counts indicates site-specific conditions influencing the safety level at the site, which are not described by the traits (see figure 2.1 of chapter 2). This again indicates that the expected number of accidents at a site (the site safety level) is somewhere in between the overall group mean and its accident counts. In other words, accident counts at sites in M will in the following year on average be closer to the overall group mean, $\overline{x}$, than m was. In figure 4.2 this is illustrated by the distance B and the phenomenon is termed the *regression to the mean* effect. The magnitude of the regression to the mean effect at a site in M is thus a function of the correlation between its yearly accident counts (see appendix B for mathematical details). According to figure 4.2, the relative magnitude of the regression to the mean

(RTM) effect at a site in M may be calculated as:

$$RTM = \frac{B}{A}$$

The more extreme m is as compared to $\overline{x}$, the more pronounced is the regression to the mean effect calculated in absolute values, i.e. calculated in number of accidents B on figure 4.2. It is important to stress, however, that while A and B may be small quantities at some sites, the regression to the mean effect, RTM, is present at all sites and in every year. Let x_{before} denote the accident counts at a site the first year and let $E\left(X_{after}|x_{before}\right)$ denote the expected number of accidents the following year given accident counts x_{before}. The regression to the mean effect at sites in sub-set M for different values of m is illustrated in figure 4.3.The figure illustrates that accident

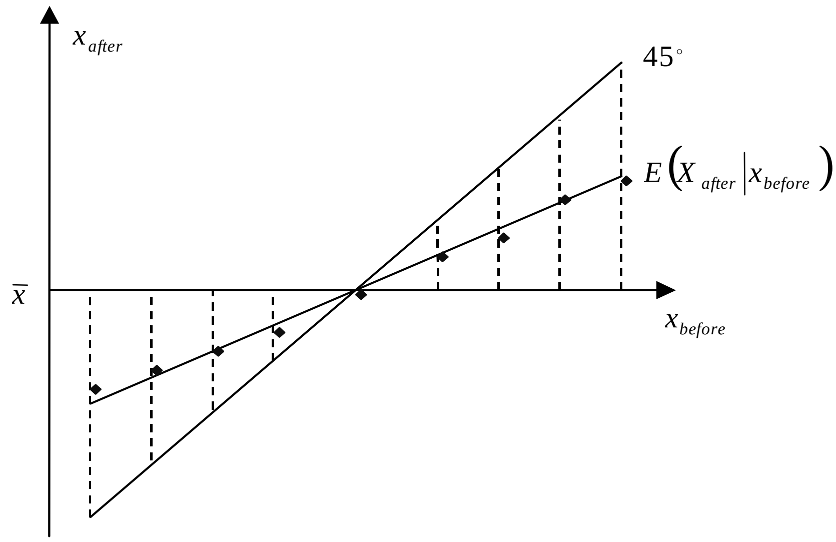


Figure 4.3: Expected accident counts given reported number of accidents in the year before.

counts in the period after, x_{after}, on average are closer to the overall mean $\overline{x}$ than x_{before} was. Hence, if accident counts the first year, x_{before}, was lower than $\overline{x}$, one would expect accident counts the following year, x_{after}, to be above x_{before} and vice versa.

In practice, one does not have groups of sites with similar traits to study. However, under the models in chapter 2, μ_{it} denotes the expected number of accidents at sites with similar traits as site i in year t (and per kilometer for road sections). Hence, μ_{it} is the mean of the reference population of site i in year t (corresponding to the group mean). If accident counts at the site in year t, x_{it}, is above μ_{it} (above $\mu_{it}L_i$ for road sections) one would expect accident counts, x_{t+1}, in year $t+1$ to be closer to the corresponding reference safety level, $\mu_{i,t+1}$ $(\mu_{i,t+1}L_i)$, than x_{it} was to μ_{it} $(\mu_{it}L_i)$. According to (2.12) of chapter 2, the predicted number of accidents in year $t+1$ given accident counts in year t is calculated as a weighted average of x_t and the reference safety level, $\mu_{i,t+1}$ $(\mu_{i,t+1}L_i)$, in year $t+1$. The weight in (2.12) is in fact the estimated regression to the mean effect (see appendix B). Hence, the regression to the mean effect is removed as accurately as possible from the site safety predictions in (2.12).

In hot spot safety work, sites are often targeted as hot spots because of an abnormally high number of reported accidents in the period under investigation. According to the regression to the mean effect, one may anticipate a decrease in accident counts at accident hot spots in the following period, even with no treatment. The regression to the mean effect is in this case also known as *bias-by-selection*. Thus, accident counts before treatment are likely to exaggerate the site safety level at accident hot spots before treatment. Not accounting for this may lead to an over-estimated effect of treatment (Hauer (1997)). On the other hand, sites not targeted because of an abnormally low number of accidents are expected to have an increase in accident counts the following period. In this case the RTM effect may be denoted a *bias-by-not-selection*. In addition, as noted above, the phenomenon resulting in the regression to the mean effect is present at all sites in all periods. Hence, accident counts in the period after treatment may also to some extent exaggerate the safety level at a site. Consequently, crude accident counts before and after treatment are *both* uncertain estimates of site safety. Summing up, before and after studies of the effect treatment should account for the regression to the mean effect in both periods.

4.1.5 Migration effect

The migration effect is the phenomenon that the accident frequency apparently rises at sites that are untreated but adjacent to treated sites. Boyle and Wright (1984) proposed a hypothesis that migration effect was due to a behavioural mechanism based on the idea of road user behavioural adjustment. However, the existence of a migration effect has not been verified (see Elvik (1997)), and a study by Maher (1990) indicates that the apparent migration effects may to a large extent be explained by a regression to the mean effect caused by a bias-by-not-selection. In other words, sites are not targeted as hot spots because of unusually low accident counts. Unusual in the sense that the reported number of accidents is below the expected at sites with similar traits. At these sites, one would anticipate an increase in accident counts in the following period due to the RTM effect. In addition, incorrect coding of the location of accidents in the network as well as changes in traffic flow due to treatment of adjacent sites may be contributing factors.

However, if such a migration effect exists, which may not be attributed to the regression to the mean effect or changes in site characteristics, it will violate the assumption in the models of chapter 2, that the dispersion effects at untreated sites are constant over time.

4.1.6 General trends

A study of time series of accident counts often reveals a trend in the accident development over a period of time[2]. As an example, since 1980 the yearly reported number of accidents on the Danish road network has been decreasing (see Vejdirektoratet (2002b)). A before and after study needs to account for the effect of such trends that are otherwise attributed to the treatment. However, part of the general decrease in the aggregated number of reported accidents since 1980 is in fact due to successful hot spot safety work. In principle, this should not be included in the general trend adjustment in a before and after study or one may overcompensate. There are many factors influencing general increases or decreases in accident counts. For instance, road users are changing their choice of modes and attitude in traffic, e.g. less people are walking and cycling and less drivers are drink driving. In addition more drivers and passengers are wearing seat belts. Also, the construction of safer cars and roads affects the general level of safety. In practice, however, it is virtually impossible to distinguish and correct for the individual effects of such factors.

4.1.7 Change in the level of reporting

Incomplete reporting of accidents leads to a general underestimation of the road safety problems. A change in the level of reporting at a site will also change its estimated level of safety. Thus, a decrease after treatment in the proportion reported will decrease the reported number of accidents. If not accounted for, the effect of treatment is overestimated[3]. However, the proportion of accidents reported at a site is rarely known, which in practice makes it impossible to observe and to calculate any changes in its level.

4.2 State of the art review

The development in before and after studies of the effect of remedial treatment is very much linked to developments made in modelling the variation in traffic accidents. Initially, the effect, ε, of treating a site was simply calculated as the proportional change in accident counts, x, before and after treatment:

$$\varepsilon = 1 - \frac{x_{after}}{x_{before}}$$

[2] Trends may even differ between site-groups.

[3] See chapter 6 for a description of the level of reporting in Denmark.

Hereby, all changes were implicitly assumed attributable to the treatment. This method is known as a *naive before and after study* (see Hauer (1997)). A chi-square test of whether or not the change in accident counts was purely random was sometimes performed along with this method.

However, in the naive before and after study, one cannot say what part of the change is due to the treatment and what part is due to changes in other factors, such as change in traffic flow etc. As an attempt to account for changes not attributable to the treatment, the before count was later multiplied by a constant, C, reflecting a number of factors such as the general development in accident counts, changes in traffic flow at the site and the regression to the mean effect (see Hauer (1992)):

$$\varepsilon = 1 - \frac{x_{after}}{x_{before} \cdot C}$$

The basis for the methods used today, however, is developed by Persaud (1986) and Hauer and Persaud (1987). They compared the reported number of accidents after treatment with the expected number, $\lambda_{without}$, had the treatment not been implemented. The latter was estimated as the posterior mean in a Poisson-gamma model, thus individually adjusting for the regression to the mean effect in the period before treatment:

$$\varepsilon = 1 - \frac{x_{after}}{\lambda_{without}}$$

Further work in this area has primarily been improvements to the estimate of the expected number of accidents without treatment, $\lambda_{without}$, in the expression above. Hauer (1997) included traits such as traffic flow, and the estimated number of accidents had the treatment not been implemented was estimated as the posterior mean in a Poisson-gamma generalized linear model. Below it will be argued that the method in Hauer (1997) is inconsistent and still suffers from the regression to the mean effect.

A few other attempts have been made to estimate the effect of treatment. In Hauer (1983) accidents are modelled by the Poisson distribution and the proportional change in site safety, $\theta = 1 - \varepsilon$, is regarded as a random variable. The uncertainty in θ is described by a gamma distribution. If λ denotes the Poisson mean before treatment, the Poisson mean after treatment is $\theta\lambda$, and the value of θ was estimated as the maximum likelihood estimator in the posterior distribution of θ. Site characteristics were not included in the model. Later Kulmala (1995) applied the method to a Poisson-gamma generalized linear model.

A great effort has been put into developing an estimate of site safety without treatment, which takes into account the regression to the mean effect. However, the crude accident count in the period after treatment is still used as the estimate of site safety with treatment. Hence, two different methods are used for estimating the site safety level at a site. This is done by the same authors who claim that the crude accident counts are unreliable estimates of safety because of the random variation in accident counts. Below, an alternative effect model is proposed, which compensates for the regression to the mean effect in both the period before and after treatment.

4.3 Effect of treatment

The task in before and after studies is to distinguish between the apparent changes in site safety caused by random fluctuations of accidents around the true underlying safety level at a site, and changes attributable to the safety treatment (Nicholson (1985)). In practice, accident data are scarce and it is impossible to distinguish between these effects with absolute certainty. However, the aim of a before and after study is to eliminate the effect of random variation as accurately as possible. The method proposed in this chapter will be based on the following general definition of the safety effect of treating a site:

Definition 5 *The safety effect of treating a site is the change in site safety at the site.*

The foundation for the before and after study is hence a comparison between the site safety level at the site in the period after treatment and the site safety level in the same period had the site not been treated. These levels of site safety will be referred to as:

(a) The site safety level with treatment

(b) The site safety level without treatment

The indications a and b will be used as indices in the measures below. The site safety level with treatment is expressed in the accident period following the treatment, while indication for the site safety level without treatment is expressed in the accident period preceding treatment. However, in order to measure the effect of treatment, safety with and without treatment needs to be based on the same traits except for changes caused by the treatment. In addition, one should compensate for the disturbing elements described above. Definition 5 implies that any remaining effect after all possible correction is attributed to the safety treatment.

4.3.1 Handling of disturbing elements

The definition of site safety at a site is the same as in chapter 2, i.e. it is a measure underlying in the number of reported accidents and not a subjective perception of safety (feeling of security). One may argue that the measure of treatment effect should take into account the change in severity of the accidents, as a treatment reducing the severity of accidents at a site, without decreasing its total number, still has had a positive effect on safety. The ability of a method to show such effects is linked to the level of detail by which accidents are classified into groups based on severity.

In definition 5, the safety effect of treatment is linked to the site in question and not to the individual safety measures implemented. Only the total net-effect achieved from treating a specific site is of interest. No attempt is made to identify the individual contributions of the different safety measures implemented. The whole issue of effect overlap and negative effects of measures are thus left aside in this thesis.

Also, no explicit consideration will be given to the effect of road user behavioural adjustment. It is merely included in the safety effect. This derives from the assumption that the safety effect of a treatment is an objective measure reflected in the reported accident number, which should not be influenced by subjective speculations as to what potential effect the treatment might have had, had the road users behaved differently.

Under the model in chapter 2, the effect of accident migration, which may not be explained purely in probalistic terms, will be expressed in the value of the dispersion effects, $\mathbf{s}$, at adjacent sites. Hence, its existence may be monitored by testing whether or not the dispersion effects at adjacent sites are unaltered by the treatment of the neighboring site. For now, it is assumed that accident migration between sites may be explained by the regression to the mean effect, by changes in traffic flow or by incorrect coding of accident locations in accordance with Maher (1990). Furthermore, because the safety effect of treatment is linked to a site rather than to a safety measure, no explicit consideration will be given to the adjacent sites of the site treated.

The regression to the mean effect (RTM) is a highly debated issue, and many attempts have been made to estimate its value (see e.g. Kulmala (1995)). Because RTM is essentially a consequence of the imperfect correlation between accident counts at a site, RTM should be considered site-specific. Thus, the magnitude of the regression to the mean effect at a particular site may not apply to other sites. Consequently, it serves no purpose first to estimate its value, and then eliminate it from the estimated effect, rather than to remove it immediately. As noted above, the phenomenon resulting in the regression to the mean effect is present at all sites in all years. In this chapter, an attempt will be given to remove it as accurately as possible from the estimated effect by removing it from the period before as well as from the period after treatment.

The effect of trends in time is attempted modelled by including it as a trait in the accident models (see chapter 6). Consequently, the time-trend will be accounted for in the effect of treatment. Changes in the levels of reporting, not captured by the time-trend, is not accounted for in the effect. This is due to the lack of information on such data.

4.3.2 A method for estimating the effect of treatment

Let T be the year of implementation of the treatment measure(s) at a site i and let $]0; T-1] = \{1, ..., T-1\}$ and $]T; U] = \{T+1, ..., U\}$ be the periods before and after treatment respectively as illustrated in figure 4.4. The year of implementation, T, is excluded from the study because the exact time of implementation often is not available (see Vejdirektoratet (1999d)), and because sub-periods of one year is used.

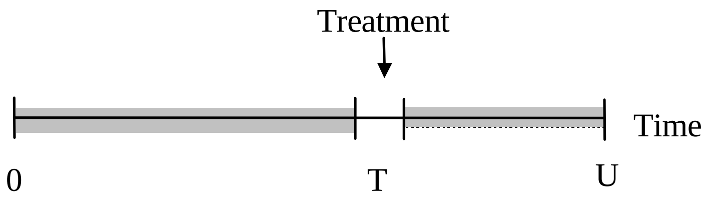


Figure 4.4: Periods before and after treatment of site i.

In order for the site safety levels with and without treatment at a site to be based on the same traits (in particular on the same traffic flow), except for those changed by treatment, they are estimated for the same year. In this study, the year, $T+1$, subsequent to the year of implementation is used. At a site i, let $\lambda_{i,T+1,a}$ and $\lambda_{i,T+1,b}$ denote the site safety levels with and without treatment in year $T+1$. On road sections, $\lambda_{i,T+1,a}$ and $\lambda_{i,T+1,b}$ are the site safety levels per kilometer. The safety effect, ε_i, of treating site i may be defined as the change in site safety:

$$\varepsilon_i = 1 - \frac{\lambda_{i,T+1,a}}{\lambda_{i,T+1,b}} \tag{4.1}$$

Hence, the safety effect, ε_i, is based on the traits in year $T+1$, e.g. on the traffic flow in year $T+1$. Because the levels of site safety with and without treatment are calculated for the same year, any difference in the lengths of the observation period before respectively after treatment needs not be accounted for. From (2.1) and (2.6) of chapter 2, the safety effect in (4.1) may be expressed as:

$$\varepsilon_i = 1 - \frac{\mu_{i,T+1,a}}{\mu_{i,T+1,b}} \cdot \frac{s_{i,a}}{s_{i,b}} \tag{4.2}$$

The total effect of treatment may conceptually be divided into explained and unexplained effect:

Explained effect measured by changes in the traits through changes in the reference safety levels.

Unexplained effect measured by changes in the dispersion effect.

The site safety levels, $\lambda_{i,T+1,a}$ and $\lambda_{i,T+1,b}$, are unobservable quantities that have to be estimated, respectively predicted. The site safety level without treatment, $\lambda_{i,T+1,b}$, is based on the accident period before treatment, $]0;T-1]$, but its value is predicted for the year $T+1$. The site safety level with treatment, $\lambda_{i,T+1,a}$, is based on the accident period after treatment, $]T;U]$, and its value is estimated for the same year $T+1$. Under the models in chapter 2, the safety effect, ε_i, in (4.1) is estimated as:

$$\widehat{\varepsilon}_i = 1 - \frac{\widehat{\lambda}_{i,T+1,a}}{\widehat{\lambda}_{i,T+1,b}} \tag{4.3}$$

The estimates $\widehat{\lambda}_{i,T+1,b}$ and $\widehat{\lambda}_{i,T+1,a}$ are calculated in a similar way using (2.12) of chapter 2. Consequently, both estimates have the properties described in chapter 2.

It is assumed that for each year t in the study, the value of the traits are known. Assume, that site i prior to treatment belongs to a site-group H_b. Let $x_{i\cdot,b} = \sum_{t=1}^{T-1} x_{it}$ and $\widehat{\mu}_{i\cdot,b} = \sum_{t=1}^{T-1} \widehat{\mu}_{it}$ denote the total reported number of accidents and corresponding estimated reference safety level for the period, $]0;T-1]$, before treatment. The reference safety level, μ_{it}, in year $t \in \{1,...,T-1\}$ is estimated from the estimated fixed effect parameters and traits of year t using (2.11) of chapter 2. Under the models of chapter 2, the site safety level without treatment, at an intersection i and on a road section of length L_i respectively, may be predicted for the year $T+1$:

$$\widehat{\lambda}_{i,T+1,b} = \widehat{\mu}_{i,T+1,b}\widehat{s}_{i,b} = \begin{cases} \widehat{w}_{i,b}\widehat{\mu}_{i,T+1,b} + (1-\widehat{w}_{i,b})\,\frac{x_{i\cdot,b}}{\widehat{\mu}_{i\cdot,b}}\widehat{\mu}_{i,T+1,b} \\ \widehat{w}_{i,b}\widehat{\mu}_{i,T+1,b} + (1-\widehat{w}_{i,b})\,\frac{x_{i\cdot,b}/L_i}{\widehat{\mu}_{i\cdot,b}}\widehat{\mu}_{i,T+1,b} \end{cases} \tag{4.4}$$

The predicted reference safety level, $\widehat{\mu}_{i,T+1,b}$, in year $T+1$ is calculated from the traits of this period had the site not been treated. For both intersections and road sections, the weight, $\widehat{w}_{i,b}$, is estimated as:

$$\widehat{w}_{i,b} = \frac{\widehat{\alpha}_b}{\widehat{\alpha}_b + \widehat{\mu}_{i\cdot,b}}$$

where $\widehat{\alpha}_b$ is the estimated dispersion parameter in the distribution of the dispersion effect before treatment. The estimated site safety levels without treatment in (4.4) are in accordance with the theory of Hauer (1997). The estimates are the empirical Bayes estimators of site safety given accident data of the period before treatment. In (4.4), the reported accident number is regressed[4] towards the expected number of accidents at sites with similar traits as site i, using regression parameter $\widehat{w}_{i,b}$. Consequently, the regression to the mean effect of this period is removed as accurately as possible (see Davis (1986)). According to the so-called *Stein Result* and results of simulation (see

[4]Also known as *Bayesian shrinkage.*

chapter 2), the estimates in (4.4) are better[5] estimates of safety without treatment, than the crude accident count, $x_{i\cdot,b}$, before treatment.

Remedial treatment work often changes the traits of a site (e.g. replacing yield signs by signal controls at intersections). Sometimes the traits are changed to such an extent that the site changes site-group (e.g. converting intersections to roundabouts). Assume, that site i after treatment belongs to a site-group H_a. Let $x_{i\cdot,a} = \sum_{t=T+1}^{U} x_{it}$ and $\widehat{\mu}_{i\cdot,a} = \sum_{t=T+1}^{U} \widehat{\mu}_{it}$ denote the total reported number of accidents and corresponding estimated reference safety level for the period, $]T;U]$, after treatment. Analogue to the above, the reference safety level, μ_{it}, in year $t \in \{T+1,...,U\}$ is estimated from the estimated fixed effect parameters and traits of year t using (2.11). Under the models of chapter 2, the site safety level with treatment, at an intersection i and on a road section of length L_i respectively, may be estimated for the year $T+1$:

$$\widehat{\lambda}_{i,T+1,a} = \widehat{\mu}_{i,T+1,a}\widehat{s}_{i,a} = \begin{cases} \widehat{w}_{i,a}\widehat{\mu}_{i,T+1,a} + (1-\widehat{w}_{i,a})\,\frac{x_{i\cdot,a}}{\widehat{\mu}_{i\cdot,a}}\widehat{\mu}_{i,T+1,a} \\ \widehat{w}_{i,a}\widehat{\mu}_{i,T+1,a} + (1-\widehat{w}_{i,a})\,\frac{x_{i\cdot,a}/L_i}{\widehat{\mu}_{i\cdot,a}}\widehat{\mu}_{i,T+1,a} \end{cases} \tag{4.5}$$

The estimated reference safety level, $\widehat{\mu}_{i,T+1,a}$, in year $T+1$ is calculated from the traits at site i with treatment. For both intersections and road sections, the weight, $\widehat{w}_{i,a}$, is estimated as:

$$\widehat{w}_{i,a} = \frac{\widehat{\alpha}_a}{\widehat{\alpha}_a + \widehat{\mu}_{i\cdot,a}}$$

In the case, where site i does not change site-group after treatment, the estimated dispersion parameter and fixed effect parameters are the same in both periods. As before, the estimates in (4.5) are the empirical Bayes estimators of site safety with treatment and regression parameter $\widehat{w}_{i,a}$. Hence, the regression to the mean effect of this period is removed as accurately as possible. Again, the estimates in (4.5) are better estimates of safety with treatment, than the crude accident count, $x_{i\cdot,a}$, after treatment.

Summing up, the estimated effect in (4.3) uses the site in question as reference for the change in site safety. However, due to the scarcity of accident data, knowledge of safety levels with and without treatment at sites with similar traits are included in the estimate.

In the literature, the crude accident count after treatment has been used as estimate of safety with treatment (e.g. Hauer (1997)). This is with the argument that there is no selection bias in the period after treatment, and therefore no regression to the mean effect to account for. As mentioned above, however, the phenomenon resulting in the regression to the mean effect is present at all sites in all years, and consequently also in the period after treatment. Because accident data are scarce,

[5] In terms of the mean squared error of estimation.

the observation period after treatment needs to be very long in order for the accident count of this period to equal to the level of safety at the site after treatment. This is rarely the case in practice. An estimated level of site safety with treatment, based solely on the crude accident count at the site after treatment, will not take into account the regression to the mean effect in the period after treatment.

Both $\widehat{\lambda}_{i,T+1,a}$ and $\widehat{\lambda}_{i,T+1,b}$ are better estimates of site safety with and without treatment than the corresponding crude accident counts. Consequently, one would expect the estimate in (4.3) to be a better estimator of the effect of treatment than the method suggested by Hauer (1997). A conclusion supported by simulation studies, which show that the mean squared error of effect estimation in the method suggested by Hauer (1997) is several times that of the proposed estimate in (4.3) (see appendix E for details).

Furthermore it seems inconsistent first to consider the estimate in (4.4) a proper measure of the site safety level without treatment, and then not use (4.5) as the estimated site safety with treatment. Consider an example, in which a given site has a total observation period $]0; W]$, as depicted in figure 4.5.The site has been treated

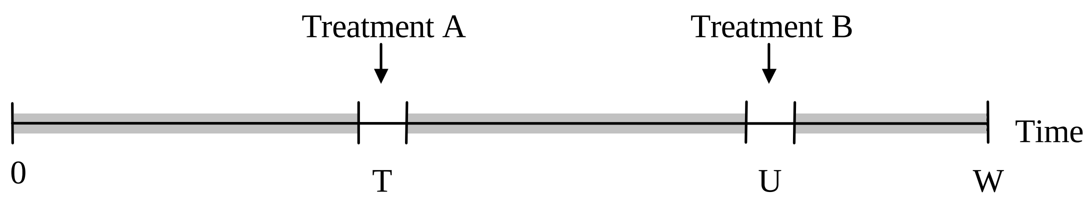


Figure 4.5: Observation period for a site treated twice within a period.

twice within the observation period (in year T and U). Again, the years of treatment are excluded from the study. A before and after study of the effect of treatment A involves the following measures:

1) Predicting the site safety level in year $T+1$ without treatment A, based on the accident period $]0; T-1]$.

2) Estimating the site safety level in year $T+1$ with treatment A, based on the accident period $]T; U-1]$.

and a before and after study of the effect of treatment B involves the measures:

3) Predicting the site safety level in year $U+1$ without treatment B, based on the accident period $]T; U-1]$.

4) Estimating the site safety level in year $U+1$ with treatment B, based on the accident period $]U; W]$.

The measures in 2) and 3) both estimate the site safety level, i.e. the expected number of accidents, at the site after treatment A but before treatment B, and should consequently be estimated using the same method. This is exactly the case when the proposed method in (4.3) is applied. The method suggested in Hauer (1997), however, uses two different methods for estimating the site safety level after treatment A but before treatment B. In the before and after study of treatment B, the site safety level is estimated using formula (4.4). This is in line with the method proposed in (4.3). In the before and after study of treatment A, however, only the crude accident count of period $]T; U-1]$ is used. The method in Hauer (1997) seems inconsistent because it uses two different methods for estimating the same level of site safety.

Summing up, the method proposed in (4.3) for estimating the effect of treating a site in the road network has the following advantages:

- The method is consistent, in the sense that the estimated level of site safety with treatment and the predicted level of site safety without treatment are calculated in a similar way.

- Because the estimates of site safety with and without treatment are calculated for the same year, they are based on the exact same traits except for those changed by the treatment. Consequently, no additional adjustments for changes in traits or for differences in the length of the period before and after treatment are needed.

- The regression to the mean effect is removed as accurately as possible from the accident count in both the period before and after treatment. Thus the regression to the mean effect is removed as accurately as possible from the estimated effect of treatment.

- According to the so-called *Stein result* and results of simulation, the mean squared error of estimation using the proposed method is less than in the method suggested by Hauer (1997).

An improvement of the method used in before and after studies will contribute to hot spot safety work in general, as the improved estimates of the effect of treatment may improve the foundation for prioritizing of hot spots and safety measures as well as for the road safety audit.

4.3.3 Example continued

Ideally, only relevant data should be used in before and after studies. Hauer (1997) defines relevant accident data as types of accidents, which may be affected by the treatment measure in question. In practise, this classification of data into relevant and irrelevant accidents is difficult, because all aspects of the accident process are not yet fully understood. One still lacks information on the factors that causes or prevents accidents, and there is a grey area between the two groups of accident data where the treatment may or may not have an effect. Even though the choice of relevant data for the study is of great importance[6], all available accident data for a site are usually taken into account in the before and after study. Hence, the before and after study illustrated in this example is based on the total reported accident counts.

Assume, that the accident hot spots, junctions 1 and 3, from the example in chapter 3 are treated in year 1999 with the potential safety schemes described above. Hence, the treatment of junction 1 involved converting the give way marking on the minor road into a stop sign and implementing give way markings on the major arms. At junction 3, a controller to briefly extend the red light for cross traffic was implemented. Table 4.1 lists the road geometry traits at junctions 1 and 3 after treatment.

Site	**No. arms**	**Frontage**	**Yield relations**		**Channelisation**	
			Major	**Minor**	**Major**	**Minor**
1	3	None	*Other*	Other	Yes	None
3	4	Industry	Signal	Signal	Yes	Yes

Table 4.1: Road geometry data after treatment.

Comparing table 4.1 to table 2.1 of chapter 2, one observes that the only change in site characteristics described by the traits is at junction 1, where yield relations on major roads are changed from level *None* to level *Other*. Neither of the junctions change site-group after treatment. Corresponding traffic flow and accident counts in the period after treatment are listed in table 4.2

Site		**2000**	**2001**	**2002**
1	Major AADT	29,712	30,258	29,121
	Minor AADT	2,834	2,952	3,235
	Accident count	1	1	0
3	Major AADT	28,191	28,376	28,877
	Minor AADT	11,990	12,011	12,236
	Accident count	6	4	5

Table 4.2: Traffic flow and accident counts after treatment.

The reference safety levels with and without treatment are estimated as well as

[6]See Hauer (1997) for examples of dependency between the choice of relevant data and the estimated effect of treatment.

predicted for the year 2000 based on the measured traffic flows for this year. As an example, the reference safety with treatment at junction 1 in year 2000 is estimated as (see appendix D):

$$\widehat{\mu}_{1,2000,a} = 0.000127 \cdot 0.97^6 \cdot 29712^{0.43} \cdot 2834^{0.44} \cdot \exp\left(-1.10 + 0.14\right) = 0.25$$

and the corresponding reference safety level without treatment in year 2000 is predicted as:

$$\widehat{\mu}_{1,2000,b} = 0.000127 \cdot 0.97^6 \cdot 29712^{0.43} \cdot 2834^{0.44} \cdot \exp\left(0.14\right) = 0.76$$

The estimated dispersion effect, $\widehat{s}_{1,a}$, at junction 1 after treatment is based on the accident counts and corresponding reference safety levels in the period after treatment (see (2.5) of chapter 2):

$$\widehat{s}_{1,a} = \frac{1.83 + 2}{1.83 + 0.75} = 1.48$$

The total expected number of accidents at sites with similar traits as junction 1 in the 3-year period after treatment is 0.75. Because junction 1 remains in the same site-group after treatment, the estimated dispersion effect, $\widehat{\alpha}$, is still 1.83. The reference safety levels and dispersion effect at junction 3 is calculated in a similar way. Table 4.3 lists the average annual number of accidents, estimated dispersion effects and safety indices with and without treatment in year 2000.

Site	$\overline{x}_{i,b}$	$\overline{x}_{i,a}$	$\widehat{s}_{i,b}$	$\widehat{s}_{i,a}$	$\widehat{\mu}_{i,2000,b}$	$\widehat{\mu}_{i,2000,a}$	$\widehat{\lambda}_{i,2000,b}$	$\widehat{\lambda}_{i,2000,a}$
1	1.60	0.67	1.70	1.48	0.76	0.25	1.29	0.37
3	8.20	5.00	2.50	1.63	2.90	2.90	7.24	4.72

Table 4.3: The accident number, estimated dispersion effects and safety indeces with and without treatment.

The estimated and predicted site safety levels, $\widehat{\lambda}_{i,2000,a}$ and $\widehat{\lambda}_{i,2000,b}$, with and without treatment in year 2000 are calculated from (4.5) and (4.4) respectively. The effect of treatment, $\widehat{\varepsilon}_i$, is estimated from (4.3):

$$\begin{aligned} \widehat{\varepsilon}_1 &= 1 - \frac{0.37}{1.29} = 0.71 \\ \widehat{\varepsilon}_3 &= 1 - \frac{4.72}{7.24} = 0.35 \end{aligned}$$

The effect at junction 1 is mainly gained from the proportional reduction in the estimated reference safety level, as the proportional reduction in the estimated dispersion effect is relatively small. The effect at junction 3, however, is only from the proportional reduction in the estimated dispersion effect after treatment, because the improvement of the existing signal control is not reflected in the traits. The corresponding effect estimates using the method in Hauer (1997) are $\widehat{\varepsilon}_1 = 0.48$ and $\widehat{\varepsilon}_3 = 0.31$ respectively, and are both below the estimates using the proposed method above. This is due to the fact, that in the estimated site safety levels with treatment, the accident count is regressed downwards towards the reference mean.

Chapter 5

State of the art in Denmark

The purpose of this chapter is to give a coherent and structured description of the state of the art in hot spot safety work in Denmark, and to compare it to the models and methods proposed in this thesis. Simulation studies based on accident and road data for 3- and 4-arm signal controlled junctions on Danish state and regional roads are used in the comparison (see appendix E).

Hot spot safety work on the Danish road network is managed by three sub-administrations:

- The national Road Directorate
- The regional authorities (the 14 counties)
- The local authorities (the 275 municipalities)

The Road Directorate is the primary developer of the theoretical foundation of hot spot safety work in Denmark today. In addition, they offer technical assistance to the entire road sector. The models and methods developed by the Road Directorate are used by most counties and partly by larger municipalities. In general, the level of theoretical sophistication in hot spot safety work of the different sub-administrations is decreasing in the list above.

Each sub-administration administrates and operates part of the total road network. The division of the road network and the corresponding traffic volume and people killed or injured in year 2000 are listed in table 5.1 (see Vejsektoren (2002b)).

The Road Directorate is in charge of the national road network. In length, state roads only add up to about 3% of the total road network, but carry more than a quarter of the total traffic volume (kilometers driven). Traffic volume wise, the national road network is relatively safe as only about 9% of the accidents is on state roads. The reason is the fact that a large part of the state roads is motorways and motortrafficways which in general are relatively safe. The regional road network is

	National	Regional	Local
Road network	3%	14%	83%
Traffic volume	27%	32%	41%
Killed in traffic	14%	45%	41%
Road accidents	9%	27%	64%

Table 5.1: The road network divided into sub-administrations in year 2000.

distributed over the 14 counties, each administrating the part within its own county limits. The vast majority of the road network is classified as local roads and is managed by the 275 municipalities. Only about 41% of the traffic volume is, however, carried on local roads, but almost two thirds of all road accidents happen on the local road network. In addition, a large amount of people killed in traffic are killed on the local road network.

For each sub-administration, the elements of its hot spot safety work are described and discussed below. A detailed mathematical presentation of the models and methods is given in appendix C. It should be noted that some of the regional and local authorities do not consistently make use of the methods described below.

5.1 The national and regional authorities

The accident models of the national Road Directorate (RD), the so-called *ap-models*, are developed in cooperation with the regional authorities, and they apply to both state and regional roads. Except for a few minor counties, both sub-administrations use the same models and methods in hot spot safety work.

The data foundation for the ap-models below, is the Road Sector Information System, VIS, a nationwide road data bank owned jointly by RD and the regional authorities. VIS contains road and accident data for the national and regional road network, based on the police reports. In addition, it contains accident data for local roads but without the corresponding site characteristics (see chapter 6 for a description of VIS). In the future, VIS is supposed to include road data for the local road network as well (see Vejdirektoratet (1998b)).

5.1.1 Modelling variation

The Road Directorate has classified the state and regional road network into road sections, roundabouts and intersections. Road sections and intersections are further classified into a number of so-called *ap-groups*, defined by geographical-, geometrical- and environmental characteristics (site characteristics), as illustrated in figure 5.1 (see Vejdirektoratet (2001b) for a definition of the current ap-groups).For each ap-group,

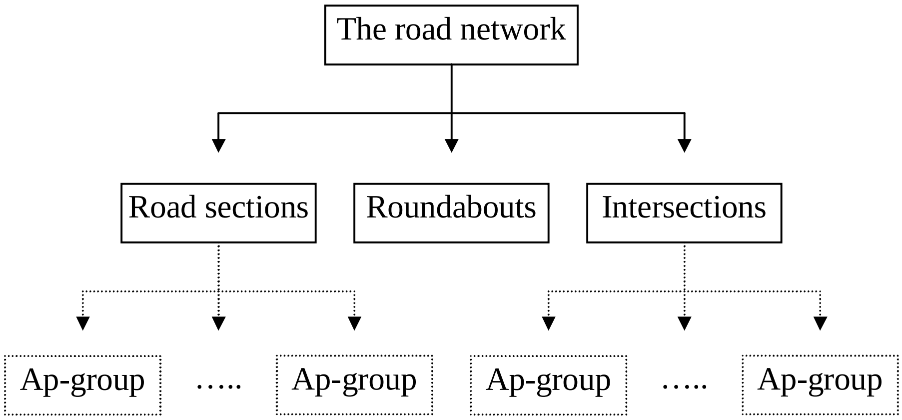


Figure 5.1: The classification of the national and regional road network into ap-groups.

a model describing variation in accident counts within the ap-group is used, that has been developed by the Road Directorate (except for roundabouts, where the model was developed by Aagaard (1995)).

The accident models are fairly simple (few parameters) because site characteristics, other than traffic flow, are not directly included as traits in the models[1]. The models are the foundation for the national and regional hot spot safety work. The variation in accident counts is separated into explained variation between the expected number of accidents at different sites within an ap-group, and random variation at the sites around the expected numbers.

The models are based on accident counts for state and regional roads. The observation period at a site ranges between 3 and 5 years, and the site characteristics (e.g. traffic flow) of a site are modelled constant within its observation period. Let μ_i denote the expected number of accidents at a site i in a year t in the observation period (and kilometer for road sections), and let x_{it} denote the corresponding reported number of accidents. Any deviation in the number of reported accidents from the expected number is assumed random and is described by the Poisson distribution. Thus, x_{it} is a realization of a random variable X_{it}. Let L_i denote the length of a road section i, hence the distribution of X_{it} at an intersection or roundabout i, respectively on a road section i, in year t is modelled by the Poisson distribution:

$$X_{it} \in \begin{cases} Poiss\,(\mu_i) \\ Poiss\,(\mu_i L_i) \end{cases} \tag{5.1}$$

[1] An attempt has been made to include site characteristics, besides traffic flow, as traits in a model for urban road sections (see Vejdirektoratet (1998a)). However, such a model is presently not in use.

The accident generating process is assumed stationary within the observation period. In other words, the yearly expected number of accidents (the site safety level) at a site is modelled constant over the observation period. Currently, traffic flow is the only trait in the models. For road sections, the structure of the site safety level per kilometer is:

$$\mu = a \cdot AADT^b \tag{5.2}$$

where $AADT$ is the average annual daily traffic (number of vehicles) at the site, and a and b are parameters estimated for each ap-group. Differences in the expected number of accidents at road sections in an ap-group are described solely by differences in traffic flow. Models are estimated for the total number of accidents, injury accidents and fatal accidents (see Vejdirektoratet (2001b)). The site safety level for roundabouts is similar in structure to (5.2) and models are estimated for the total number of accidents (see Aagaard (1995)). For intersections, the structure of the site safety level is:

$$\mu = a \cdot AADT_{ma}^{b_1} \cdot AADT_{mi}^{b_2} \tag{5.3}$$

where $AADT_{ma}$ and $AADT_{mi}$ are the average annual daily incoming traffic on major and minor roads respectively (see appendix D for a definition of major and minor arms). The parameters a, b_1 and b_2 are estimated for each ap-group. Again models are estimated for the total number of accidents, injury accidents and fatal accidents. Analogous to (5.2), differences in the expected number of accidents at intersections in an ap-group are described solely by differences in traffic flow. A detailed mathematical description of the models is given in appendix C.

Discussion

The models defined in (5.1)-(5.3) belong to the class of generalized linear models (GLIM) with a Poisson error distribution (see appendix A for a description of GLIM).

Within an ap-group, all variation in accident counts not ascribed to differences in traffic flow are thus modelled random and described by the Poisson distribution. However, it seems reasonable to assume that not all site-specific conditions, besides traffic flow, are accounted for in the ap-group definitions (see chapter 2 for further discussion). In other words, two sites in the same ap-group and with the same traffic flows are still likely to have different site safety levels[2]. The models developed by the Road Directorate do not allow for such site-specific differences in an ap-group.

Simulation studies comparing the models above with the Poisson-gamma hierarchical generalized linear models (Pg-HGLMs) proposed in chapter 2 show that the Pg-HGLM estimates of safety are better in terms of the mean squared error of estimation (see appendix E). The reason for this result is the inclusion of the dispersion

[2] Nicholson (1985) found the Poisson distribution to be inadequate even when differences in traits were accounted for.

effect, s, in Pg-HGLM, which accounts for site-specific conditions not described by the traits.

Furthermore, the models used by the Road Directorate and regional authorities are aggregated over time. This implies constant values of the traits at a site over its observation period (e.g. constant traffic flow). A study of the measured traffic flows on the national and regional road network show that at certain sites, the average annual daily traffic varies considerable over the observation period. Hence, in order for the model to be representative of the current traffic situation at a site, short observation periods may be required, leading to a high uncertainty in the estimated model parameters. The Poisson-gamma hierarchical generalized linear models are disaggregated on sub-periods of one year. Consequently, changes in traffic flow and other traits from one year to another are taken into account.

5.1.2 Targeting hot spots

The Road Directorate and regional authorities usually target hot spots once a year on basis of the total number of accidents or injury accidents only. An accident hot spot is identified by comparing a site to the other sites in the respective ap-group. The following definition of an accident hot spot is used by the national and regional authorities (see Vejdirektoratet (2001b)):

Definition 6 *An accident hot spot is a site in the state or regional road network with a reported number of accidents, which is both beyond a fixed minimum and significantly larger than its expected number of accidents.*

The fixed minimum is a threshold value supposedly assuring that the sites targeted have sufficient data to indicate a trend in the accidents, i.e. to indicate safety measures for treatment. The expected number of accidents in definition 6 refers to the expressions in (5.2) and (5.3) respectively. The requirement that the total number of reported accidents significantly exceeds the corresponding expected number, assures that one do not target the same type of sites each time, as comparisons are made within the ap-group. The chosen values of the fixed minimum and the level of significance indicate the proportion of sites, the decision-maker wishes to identify as accident hot spots. The level of significance reflects the evidence of hotness at a site. The mathematics behind definition 6 are described in appendix C.

The Road Directorate has developed a computer system, *VISplet*, for targeting hot spots on state and regional roads (see Vejdirektoratet (1996)). The system is used in combination with the Road Sector Information system, VIS, and is based on definition 6 and the accident models above.

Discussion

In the targeting method above, sites are classified into accident hot spots and accident cool spots. The method does not, however, indicate the amount by which the safety level at an accident hot spot exceeds the safety level at sites with similar traits. This amount reflects the level of hotness at a site (see chapter 3) and indicates the potential for accident reduction. The level of hotness may be used for ranking and prioritizing accident hot spots for treatment. Under the models developed by the Road Directorate, the level of hotness at an accident hot spot is estimated as the ratio between the crude accident count, x, and the corresponding safety level, μ (μL for road sections), at sites with similar traits. However, results of simulation show this estimate of the level of hotness to be inferior[3] to the proposed estimate in chapter 3 (see appendix E).

Simulation studies also show that the method for targeting hot spots proposed in chapter 3 gives a marginally higher sensitivity than the method used at present by the Road Directorate and regional authorities (see appendix E). The specificity is approximately the same in the two methods. The higher sensitivity in the proposed method is due to the fact that in a group of sites (e.g. in an ap-group), the magnitude by which site safety levels vary from their corresponding group means (reference safety levels) is considered[4].

5.1.3 Prioritizing

Due to the limited sources of funding for hot spot safety work, not all accident hot spots may be treated. Consequently, one has to prioritize between sites and safety measures. The task of prioritizing is to select a cost-efficient portfolio of safety measures within the given budget. In practice, the prioritizing is performed between groups of safety measures (safety measure schemes). The performance of a safety scheme is reflected in its ability to improve safety at a site. The Road Directorate and regional authorities use the *first year benefit (FYB)* and *marginal benefit (MB)* as measures of performance of potential safety schemes. The FYB and MB values are measures common in prioritizing of hot spot safety measures (see e.g. Bernhardt and Virkler (2002)). The relevance of a scheme is based on the accident category and factors involved. The selection of an optimal portfolio of safety schemes is a prioritizing process in two steps (see Vejdirektoratet (1992)):

A) Comparing the individual performance of potential safety schemes *within* each accident hot spot.

[3] In the sense that it has a larger mean squared error of estimation.

[4] Through the estimated dispersion parameter, $\widehat{\alpha}$.

B) Comparing the individual performance of potential safety schemes *between* accident hot spots.

Part A in the prioritizing process is based on the first year benefit (FYB) of the potential safety schemes at the site in question. The first year benefit is calculated as the ratio of the expected saved accident costs the first year due to the scheme (AC) and its cost of implementation (C):

$$FYB = \frac{AC}{C}$$

The expected saved accident costs the first year are calculated as the product of the expected accident reduction, and the cost to society of one accident. The costs of injury and property damage only accidents are repeatedly estimated by the Road Directorate (see Vejdirektoratet (1999c) & (2002b)). The expected accident reduction is based on the crude accident counts and on the expected reduction rate of the safety scheme. The expected reduction rate in accidents due to a safety scheme is either based on experience from previous studies or roughly estimated as either 0%, 33% or 50%, dependent on the accident categories and factors involved at the site. At each accident hot spot, the potential safety schemes are ranked according to their individual FYB-values, and the scheme with the highest first year benefit is denoted the *primary* scheme at this site.

In the second part of the prioritizing process, the potential safety schemes at different accident hot spots are compared using the marginal benefit (MB) of the excess investment needed for implementing a more costly scheme (lower FYB), but with a higher saved accident cost. The argument is then that a non-primary safety scheme at a site, but with a higher marginal benefit than the first year benefit of a primary scheme at another site is more cost-efficient. A detailed mathematical description of the prioritizing method is given in appendix C.

The potential safety schemes at a site are assumed to be true alternatives, and two different schemes may not be implemented at the same site. Hence, effect overlap of schemes is not an issue. The optimal portfolio of safety schemes is found by successively selecting the most dominant safety scheme until the budget or another constraint is reached.

Discussion

The inherent uncertainty with the targeting of hot spots imply that some accident hot spots are *hotter* than others in the sense that they have a higher statistical certainty of being a hot spot. The level of hotness at a site indicates its relative potential for accident reduction, and is thus important in the prioritizing phase. However, in the prioritizing method above, no considerations are given to the level of hotness at accident hot spots. Instead, the method is based on the crude accident count and on measures not dependent on the site.

Sites are targeted as hot spots because of an abnormally high number of reported accidents in a period. According to the regression to the mean effect discussed in chapter 4, one may anticipate a decrease in accident counts in the following period. This indicates that the accident counts at sites targeted as hot spots exaggerate the underlying site safety level. A prioritizing method based only on the crude accident count is thus likely to overestimate the first year benefit and marginal benefit measures. Instead, the prioritizing method should be based on the model estimates of site safety. However, the predicted safety levels in (5.2) and (5.3) are poor estimates of site safety at accident hot spots. A site is targeted as a hot spot if its accident count, x, is significantly higher than the safety level at similar sites, μ (μL on road sections). In other words, accident hot spots are sites, where the models defined in (5.1)-(5.3) do not apply.

As mentioned above, the Poisson-gamma hierarchical generalized models give better[5] estimates of safety than the models currently used by the Road Directorate and regional authorities. A prioritizing method based on the Pg-HGLM estimates of site safety is thus better than a method based on the crude accident count or on the safety levels in (5.2) and (5.3). In addition, the Pg-HGLM site safety estimates take into account the level of hotness at a site through the estimated dispersion effect, $\widehat{s}$.

5.1.4 Before and after studies

Analogous to chapter 4, the foundation of before and after studies on state and regional roads is a comparison between the site safety level at the site in the period after treatment and the site safety level in the same period had the site not been treated. Besides changes inflicted by the treatment, it is assumed that no other changes have been made to the road geometry in these periods. Again, the year of implementation of the safety scheme is excluded from the study (see Vejdirektoratet (1999d)). The safety effect of treating a site is generally defined as:

Definition 7 *The safety effect of treating a site is the change in site safety at the site.*

The site safety level after treatment is estimated as the crude accident count of the period after treatment. The adjusted accident count of the period before treatment is used as the predicted site safety level had the safety scheme not been implemented.

Let $x_{i\cdot,b}$ and $x_{i\cdot,a}$ denote the reported number of accidents at site i in the period before and after treatment. The effect of treatment at site i is estimated as:

$$\widehat{\varepsilon}_i = 1 - \frac{x_{i\cdot,a}}{x_{i\cdot,b} \cdot C} \tag{5.4}$$

[5]In terms of the mean squared error of estimation.

where C in (5.4) is a constant consisting of a number of correction factors:

Change in traffic flow calculated as the ratio of the total traffic flow at the site in the period after and before treatment respectively.

The regression to the mean effect a fixed estimated percentage between 20-30% (see Vejdirektoratet (2001b)).

General change in safety from the period before to the period after treatment in a control group[6]. The general change is estimated as the ratio of the accident count in the control group in the period after and before treatment respectively.

The estimated effect of treatment is furthermore tested for significance in a chi-square distribution. Examples of the method are given in Vejdirektoratet (1994) and (1999d). The mathematics behind the method are described in appendix C.

Discussion

In the estimated effect of treatment in (5.4), the accident count of the period before treatment is adjusted by a number of factors in order to estimate the level of site safety without treatment. However, the adjustment factors are connected with a number of problems.

The regression to the mean (RTM) effect is modelled constant within an ap-group. However, the RTM effect is not a constant figure but varies with traits such as traffic flow (see appendix B). The regression to the mean effect is thus site-specific. Applying a constant RTM effect to a group of sites results in poor estimates of the effect of treatment, independent on its value. In addition, the phenomenon resulting in the regression to the mean effect is also present in the period after treatment, and the accident count in this period should be adjusted accordingly (see chapter 4).

The correction factor for changes in traffic flow does not reflect its connection with site safety expressed in (5.2) and (5.3). In the effect estimate, proportionality between site safety and traffic flow is implied, while this is only the case in (5.2) for $p = 1$ and not the case at all in (5.3).

Hot spots are targeted on basis of a relatively high reported accident number. The change in accident count in the control group may thus not apply to the accident hot spots (see the description of bias-by-not-selection in chapter 4). Furthermore, no considerations are given to changes in traffic flow in the control group from one period to another. Consequently, the estimated general change in safety is not adjusted for changes contributively to changes in traffic flow (see Hauer (1997)).

[6] A control group is a sub-set of the sites in the ap-group where no safety schemes are implemented in the period in question.

The proposed effect model in chapter 4 is based on traits for the same year (except for those changed by treatment). In addition, traits describing trend in accident count are included in the Poisson-gamma hierarchical generalized models (see chapter 6). Simulation studies show that the method proposed in chapter 4 is better[7] at estimating the true effect of treatment, than the method currently used by the Road Directorate and regional authorities (see appendix E). The reason is that the proposed method in chapter 4 estimates safety at sites better, both with and without treatment.

5.2 The local authorities

At present, only one third of the municipalities uses a structured hot spot safety work (see Vejdirektoratet (1999b)). A study by the national Road Directorate in corporation with Odense University shows that local authorities are especially burdened with accident costs because of the high number of accidents on local roads. As a consequence, focus on hot spot safety work on local roads has increased, and the Road Directorate has compiled a manual to safety work for municipalities (see Vejdirektoratet (1998b)). The methods described in this section are analogous to the methods given in this handbook.

5.2.1 Modelling variation

The registration of accidents on local roads is less detailed than on state and regional roads, and road data are not collected in a systematic manner. This results in difficulty in developing models for describing the variation in traffic accidents on local roads. Consequently, most local authorities do not make use of statistically based methods in accident hot spot safety work.

5.2.2 Targeting hot spots

Due to the lack of models describing the variation in accident counts on local roads, the municipalities use a so-called *accident frequency/accident rate*-method, based on the following definition of accident hot spots:

Definition 8 *An accident hot spot is a site on the local road network with both a high accident frequency (number of reported accidents) and a high accident rate (number of accidents per passing vehicle).*

[7] In terms of the mean squared error of estimation.

On road sections, definition 8 applies to a fixed road length (usually 400 meters). The structure in the accident frequency/accident rate method for targeting accident hot spots is:

(i) The local road network is classified according to road sections, intersections and roundabouts.

(ii) Within each road class, sites are listed with decreasing accident frequency, and the n sites with the highest accident frequencies are selected.

(iii) The n sites selected in (ii) are re-listed with decreasing accident rate, and the $m < n$ sites with the highest accident rates are targeted as hot spots.

Occasionally, hot spots are targeted from their accident frequency only, i.e. phase (iii) is ignored. In addition, some municipalities only consider a subset of sites in the targeting process, e.g. sites with very low traffic flows are omitted. The values of n and m are politically determined. The Road Directorate has developed a PC-based system for targeting hot spots on local roads using the method above. The system is combined with the road management system for municipalities, VEJMAN (see Vejdirektoratet (2001b) for a description of VEJMAN). A minor part of the municipalities has very few reported accidents a year and instead they use the populations' perception of safety as the basis for hot spot safety work.

5.2.3 Prioritizing

A cost-efficient portfolio of safety schemes is selected from step A of the prioritizing method used by the Road Directorate and regional authorities. For each accident hot spot, the potential safety schemes are identified and the corresponding first year benefits (FYBs) are calculated. The safety schemes are ranked according to their FYB-values and the optimal portfolio of schemes is found by successively selecting the safety scheme with the highest FYB until the budget or another constraint is reached.

5.2.4 Before and after studies

Before and after studies of preventive safety schemes on local roads are rare, and there is no systematic collection of the estimated effects. If performed, they are based on the reported accident counts before and after treatment of the site in question, and on a control group. Corrections for the regression to the mean effect and changes in traffic flow are often not included (see Vejdirektoratet (1992)).

5.2.5 Discussion

The local authorities use the crude accident counts as estimates of safety. Consequently, the random variation in accidents is not accounted for. This fact is reflected in all phases of the hot spot safety work on local roads. In addition, the prioritizing method and before and after studies suffer from the same deficiencies as the methods used by the Road Directorate and regional authorities. In addition, the use of perception of safety as basis for traffic safety work is highly disputable (see chapter 1 for a discussion). A more systematic collection of site characteristics as well as better location of accidents is needed, in order to be able to estimate models for the variation in accident counts on local roads. However, according to the *Stein result* (see appendix A), even a Poisson-gamma hierarchical generalized linear model with no traits[8] gives better estimates of site safety compared to the crude accident counts.

[8] Corresponding to all sites in a site-group having the same reference safety level. The reference safety level is then the mean accident count in the site-group.

Chapter 6

Analysis of accident counts in Denmark

The objective of this chapter is to estimate the parameters of the hierarchical generalized linear models developed in chapter 2. Relationships between safety, traffic flow, time and road geometry are derived, in order to have the ability to estimate and predict safety at sites for the hot spot safety work. Data used in the analysis are based on accident and site data for the national and regional road network in Denmark.

6.1 Data sources

The parameters of the accident models are estimated from the accident and site data available. Hence, the quality of the accident models is linked to the quality of the data sources. Both accident and site data suffer from inaccuracies and misclassifications as well as misrecording of information. As an example, the recorded traffic flow may be inaccurate and the accident severity and road geometry characteristics misclassified. In addition, accident data suffer from underreporting, which is further described below.

Each traffic sub-administration, under national, regional or local authorities, is responsible for collecting site characteristics for their part of the road network. The national Road Directorate usually gathers information for both state and regional roads and collects it in the Road Sector Information System, VIS. Site characteristics for local roads are collected by the municipalities.

The official accident statistics are collected on a national level by *Statistics Denmark* from the police reports. However, there are other sources of accident data. The three main sources of accident data in Denmark are:

1. Police reports

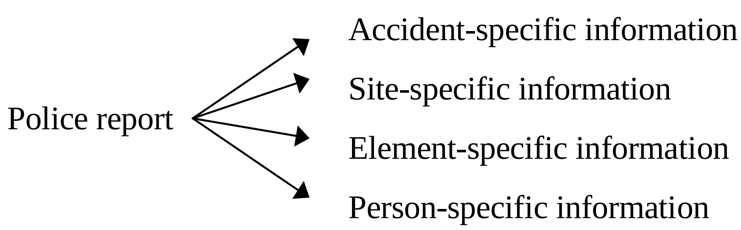
2. Hospital admission statistics

3. Insurance companies

None of these data sources completely covers the traffic accident occurrences in Denmark. The advantage of the police reports is that they in principle are recorded consistently on a national level. They contain detailed information on both the locality of the accident, accident severity and on the parties involved. The disadvantage is the fact that reporting here is incomplete, especially for property damage only accidents. This problem of underreporting is discussed below. Some hospital admission statistics contain extensive data on those traffic accidents involving personal injury, which are severe enough to require hospitalization. Unfortunately the reporting system in hospitals is inconsistent and the information is not related to the location of accidents. However, in a project by the Funen region, a system, which in the future may be used for determining the location of accidents in the hospital admission statistics, has been developed (see Lauritsen (2001)). Insurance companies have a large amount of data on traffic accidents involving insured cars. This data source has a higher level of reporting of property damage only accidents, compared to the police reports, but also lacks the location of accidents. Because the locality of accidents is important for deriving relationships between safety and the site characteristics, hospitals and insurance companies have little practical value as sources for accident data. Hence, the accident data used in this study are based on the official accident statistics, i.e. on the police reports. Provided that the system by Lauritsen for determining the location of accidents in the hospital admission statistics is applied nationally, these data may in the future be used together with the police reports.

6.1.1 The police report

The police shall report all injury accidents, and all property damage only accidents with damage exceeding 10,000 DKK. The following is recorded:



The accident-specific information includes information on the time of the accident, the accident class (e.g. rear-end and turning accidents) and on the severity of the accident (fatal, serious injury or slight injury). The site-specific information supplies the location of the accident and with that the road- and weather conditions at the time of the accident. The elements involved in the accident, refer to the vehicle(s) and possible parts of the road geometry such as crash barriers etc., and the police report furthermore states the condition of these elements. Finally, the report has information on the road users involved, including age, sex and possible intoxication of the driver and any passengers. The only part of the police report used in this study is information on the locality and severity of the accident.

Underreporting

Not all accidents on the road network are observed and not all observed accidents that are reportable are in fact reported by the police. Incomplete reporting is a problem in all of the motorized countries. Unfortunately, the level of reporting in Denmark appears to be lower than in the countries to which Denmark is usually compared[1] (see Elvik and Mysen (1999)). The level of coverage varies with accident severity and location (police jurisdiction). In addition, the level of reporting varies for the different road user classes involved in the accidents, as listed in table 6.1 (see Elvik et al. (1997)).

Road user class	**Level of reporting**
Cars	48.8%
Pedestrians	36.4%
Mopeds/motorcycles	28.8%
Pedal cycles	8.5%

Table 6.1: Level of reporting in Denmark for different road users involved in accidents.

Table 6.1 shows, that the level of reporting for accidents involving pedal cycles is particularly low, which is primarily due to the fact that such accidents often have a low level of personal injury and material damage. In general, the level of reporting is increasing with the severity of the accident, i.e. lowest for property damage only accidents and increasing to almost 100% for fatal accidents. A survey by Odense University hospital (see Ulykkes Analyse Gruppen (1997)) has shown that in 1996 only about 45% of hospitalized casualties of road accidents were included in the police reports. In addition, the level of reporting varies between jurisdictions, with a tendency of accidents in rural areas not being reported to the same extent as in urban areas.

Incomplete reporting affects the reported number of accidents, and with that the estimated level of safety at a given site. The main concerns of underreporting are:

[1]Sweden, Norway, Great Britain, Germany and Australia.

- The national total number of traffic accidents is underestimated, resulting in an underestimation of the true economical costs of traffic accidents to society.

- The variation in the level of reporting between police jurisdictions leading to an inconsistent determination of the relative level of safety at sites in different jurisdictions.

The underestimated number of accidents occurring in Denmark may lead to a low political prioritizing of traffic safety. Fortunately, accidents not reported are mainly of low personal injury and property damage. In other words, they are in a group of accidents with a relatively low cost to society.

For hot spot safety work, the main problem of underreporting is that the level of reporting may vary between police jurisdictions. It has the immediate effect that jurisdictions with a low level of reporting may unjustly appear to have a relatively high level of safety. In the Poisson-gamma hierarchical generalized linear models, differences in accident coverage are expressed in the dispersion effect. Hence, the dispersion effect at sites in jurisdictions with a low level of reporting may be unjustly low. As a consequence, hot spot safety work involving several jurisdictions with different levels of reporting will result in part of the hot spots not being targeted. At present, little information exists on the variation in the level of reporting between police jurisdictions in Denmark.

6.2 Data

Data used in this study are from the Road Sector Information System, VIS, based on the police reports. Data include information on site characteristics, such as traffic flow and road geometry, and on accidents for road sections and intersections on the national and regional road network. Site data for local roads are not collected in a systematic manner and consequently not included in this study.

6.2.1 Accident data

In the present study, traffic accidents are defined as accidents involving at least one motorized vehicle. They are classified into three main groups:

- Injury accidents

- Property damage only accidents

- Extra accidents

Injury accidents are accidents involving personal injury requiring medical attention. For each injury accident, the number of killed[2], seriously injured and slightly injured persons are recorded (see Vejdirektoratet (2001b) for further definition of severity). Property damage only accidents are those involving no personal injury but with material damage exceeding 10,000 DKK. Extra accidents are accidents involving little property damage, no injuries and no gross breach of the law. No official police report is recorded for extra accidents, and information on this class of accidents is thus sporadic. Extra accidents are therefore not included in this study.

The development in the reported number of injury- and property damage only accidents for the 10 year period, 1989-98, is illustrated in figure 6.1. It can be seen that the numbers have dropped by more than 10% from 1989 to 1998.

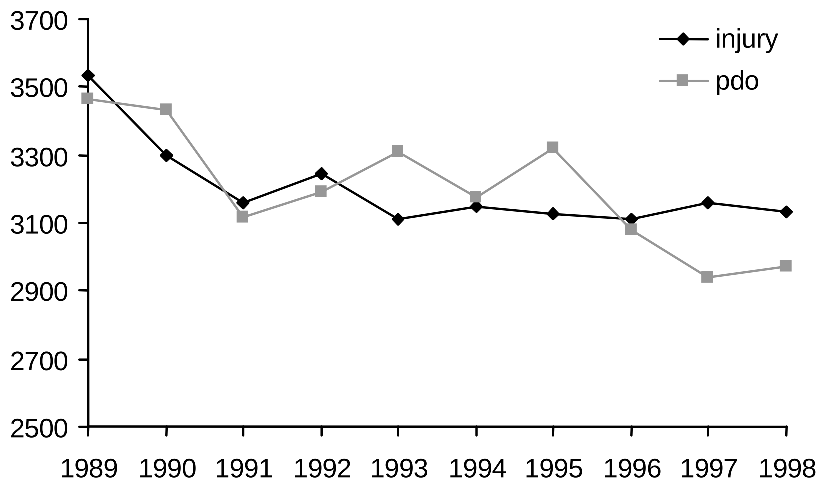


Figure 6.1: The development in the number of reported accidents on state and regional roads over the period 1989-98.

Accidents are classified into 10 main accident categories. The distribution of the number of accidents between these categories is very stable over the 10 year period (see appendix D). The most frequent accident categories in Denmark are single-vehicle accidents and accidents between vehicles going in the same direction such as rear-end collisions. Accidents occurring within 20 meters of the center of an intersection are classified as intersection accidents.

[2]People dying from the accident within 30 days are labelled *killed*. People dying from the accident after 30 days are categorized as *seriously injured*.

6.2.2 Site data

The road network is classified in intersections and road sections. This is due to the fact, that the typical conflicts arising between road users in intersections are of a different character than those arising on road sections. For intersections, the database contains data on traffic flow as well as information on the channelisation and yields relations of each of the arms. On road sections, information on the total traffic flow in each direction as well as data on a number of geometry variables and speed limit at the site is available.

The measure of traffic flow at a site is the average annual daily traffic flow of motorized vehicles (AADT). In practice, the traffic flow varies in size and composition[3] over time, and Grimmer et al. (1986) suggest dividing the traffic flow into flow movements. However, the VIS database lacks such detailed information.

6.2.3 Limitations of data

VIS only lists data on the present road geometry, and no history of geometry data prior to the last alteration of the sites is saved. This lack of historical site data makes it impossible to include years prior to the last alteration year of a site. However, since 1994 yearly counts of the average annual daily traffic are saved for the majority of the roads.

The overall quality of the data base depends on the number of errors found in accident and road data. Studies by Jarret et al. (1997) and Austin (1995a) & (1995b) show that recorded site data in Britain contain large errors. Examples of errors are inaccuracies in numerical variables, such as traffic flow, erroneously coded variables representing categories and erroneously location of accidents. Errors affect the relationship between accident counts and site characteristics, hence affecting the quality of hot spot safety work. In addition, misreporting of accident severity leads to a wrong classification of accidents. A study in Britain (see Adams (1988)) shows that only one in four casualties recorded as seriously injured is in fact correctly classified according to the definitions. The quality of models describing variation in different accident severity groups is thus affected. The types of errors found in British recorded accident and site date are likely to apply to data in VIS as well.

Even though, accident records also contain some data on site characteristics, the road network database in VIS is used as source for site characteristics. The reason is, that the road network database is likely to be more accurate than the accident records, because a more detailed investigation has been undertaken to obtain this information.

[3] E.g. the rate of heavy vehicles to lighter vehicles varies.

6.2.4 Data selection

Selection of data from VIS is based on the standard procedures used by the Road Directorate (RD). The selected data are subsequently compared to the RD grouping of data for the period 1989-93 (see Vejdirektoratet (1996)). Inconsistencies are found mainly to be attributable to the fact, that some sites have been altered after 1993. Sites altered after 1993 will not have any records on road geometry and speed limit for the period 1989-93. The Road Directorate traditionally classifies sites into a number of so-called ap-groups depending on their site characteristics (see chapter 5). Within each ap-group, an accident model is estimated with the annual daily traffic as the only trait. In this study however, fewer groups are used, and instead the site characteristics are included as traits in the models.

The outcome of the estimation of the model parameters is highly dependent on the selection procedure applied. The selection procedure used in this study is described in detail in appendix D.

6.3 Statistical modelling

The models describing the between-site and within-site variation in road accident counts are the basis for hot spot work. The purpose of such accident models is to estimate and predict safety at sites. The estimated safety levels may be used to target and prioritize between hot spots (see chapter 3).

The foundation for accident modelling is the accident counts and corresponding site-specific conditions of the sites (site characteristics). Hence, in modelling variation, one wishes to derive relationships between safety and site characteristics. One should keep in mind that accident models are in principle all wrong, but some may prove useful in describing safety (see Box (1976)). Hence, with given data, the task is to select the models which best describes the present as well as the future safety levels at sites. In theory, alternative models may describe safety equally well, as different site characteristics may reflect the same phenomenon at a site. In that case, the principle of parsimony is used, i.e. the simplest of the best models is selected.

6.3.1 Delimiting the problem

In the Road Sector Information System, VIS, accident and site data are available for junctions, roundabouts and slip roads as well as motorways, motortrafficways and other roads[4]. The period 1989-98 is considered in the study. In delimiting the problem one should select the sites as well as the time period to include in the study.

[4] Road sections not defined as motorways or motortrafficways.

In order for the models to be able to predict future safety levels, the precision of the estimated parameters is important. Because precision increases with the amount of data, it is important to incorporate as much data as possible into the models. However, the ability of the models to describe the present safety level at sites depends on how well data represent the present safety situation. It seems fair to assume that safety in e.g. year 1998 is best represented by data for this year, while the first year of the period 1989-98 is probably not at all representative for the safety situation in 1998. As a consequence, the data period 1989-98 should be limited to include only data representative for the present safety situation.

Both observable and non-observable site-specific conditions at a site may change considerably over a time period. While changes in characteristics included as traits may be accommodated in the models, changes in other site characteristics are not. As a consequence, the overall study period was limited to the years 1994-98[5]. Also, since 1994 yearly traffic counts have been saved for most of the sites. Within the 1994-98 period, however, some sites have been treated or altered, which leads to changes in the site characteristics, traits as well as others. The observation periods for these sites are reduced accordingly. As an example, for a site treated in 1996, only the years 1997 and 1998 are included in the study. Hence, observation periods for the sites in the study varied between 1 and 5 years, with more than 90% of the sites having a 5 year observation period.

The main use of the accident models is to target and prioritize between hot spots, and subsequently to estimate effects of treatment. Consequently, only site-groups of potential interest for hot spot safety work should be included in the analysis. As pointed out in chapter 3, one needs a sufficient amount of accident occurrences to detect a common risk factor pointing at relevant safety measures for this site. As a result, models are not estimated for slip roads (usually ramps), as none of these have more than 2 accident occurrences within their observation periods. At present, site data for roundabouts are not available in a manner immediately suitable for modelling. The Road Directorate does not estimate models for this site-group, and so far only Aagaard (1995) has undertaken this task. In addition roundabouts have a relatively low accident count with little personal injury. They are thus not likely to be targeted as hot spots. As a consequence, roundabouts are omitted from this study and only 3-, 4- and 5- arm junctions are included in the model for intersections.

Road sections are classified into the site-groups; *motorways*, *motortrafficways* and *other roads*. Only normal road paths are included, and accident models are estimated for each of these groups. The road sections excluded from these three site-groups are assembled in a fourth group, *remaining roads*, and an accident model is estimated for this site-group as well.

Table 6.2 lists the number of junctions and road sections together with the total

[5] A time span of 5 years is traditionally used by the Road Directorate.

Site-groups	**No. sites**	**Km**	**Accidents**	**Injury**	**Fatal**	**Rate**
Junctions	2,944	-	6,351	2,779	140	0.45
Motorways	5,970	875	2,414	1,073	119	0.56
Motortrafficways	2,714	309	426	205	34	0.28
Other roads	73,452	9,977	19,713	10,619	1,004	0.40
Remaining roads	7,375	753	1,006	414	37	0.28
Total	92,455	11,914	29,910	15,090	1,334	

Table 6.2: Site-groups included in the study for the period 1994-98.

number of accidents, injury accidents and fatal accidents for each site-group. For road sections, the total length in kilometer is also listed. The *rate* in table 6.2 is the average annual number of accidents (injury and property damage only accidents) at a site in the group. For road sections, the rate is the average annual number per kilometer.

6.3.2 Response and explanatory variables

The measure of site safety at a site is the number of accidents expected to occur at this site within a year (and kilometer for road sections). Consequently, the number of reported accidents are used as the response variable for the models. Accident models are estimated for injury accidents, which on average have a relatively high cost to society. Injury accidents have a relatively high coverage but are rare events. Including property damage only accidents, will increase variation in accident data and hereby increase the potential for the models to explain part of this variation. Hence, models are estimated for the total number of reported accidents (injury and property damage only accidents) as well.

For the site-groups and study period above, the task is to select the traits best describing the present and future safety levels at the sites. Sites are classified into different site-groups because each group has its own set of potential traits. Thus, for different site-groups, different site characteristics were attempted included in the model. As an example, all motorways have medians. Including a trait indicating the presence of a median is thus not of interest to this site-group, while such a trait is included in the motortrafficway model. The group of remaining road sections is sites excluded from the other groups because of suspicion of misclassifications of the number of lanes, width of lanes or pavement width. As a consequence, such variables were not considered for this group of sites.

The site characteristics may be classified into *numerical variables* such as traffic flow and width of the road, and into *categorical variables*. The categorical variables are particular by the fact that they only assume a limited number of values (levels).

Among the large number of site characteristics available, only a subset is included as traits in the models. From the perspective of estimating the effect of remedial work, it may seem of value to include all possible site characteristics as traits in the model. However, the statistical explanatory power of a trait is also an indication of its effect in remedial work. In general, only site characteristics with statistical significant coefficients were included.

Some of the site characteristics showed a high level of mutual covariation, e.g. for motortrafficways the pavement width of the road was almost proportional to the width of the lanes. In this case, the pavement width was omitted. In general, mutual covariation of variables indicates that there may be an alternative set of traits with similar explanatory power as the one chosen. In that case, the principle of parsimony is applied.

Only main effects are included in the models, i.e. no variables describing interactions between two or more of the traits are used. This is partly due to the lack of sufficient data and from the perspective that the model should be kept simple and thus easy to interpret. Instead, any interaction between traits is considered site-specific and is thus modelled by the dispersion effect.

Previous studies have found the traffic flow to be the single most important determinant in accident models (see OECD (1997)). This may be explained by the fact that traffic is what creates the conflicts leading to accidents, i.e. no traffic no accidents. The common measure of traffic flow is the average annual daily traffic (AADT). Intersections and road sections with the latest traffic counts prior to 1989 are omitted, as such counts are considered unrepresentative for the present traffic situation at these sites.

At intersections, most accidents occur between crossing traffic flows, and it is therefore appropriate to divide the traffic into major and minor flows in order to capture the potential conflicts. Hence, the arms forming the junctions are classified into major and minor roads using the definition of the Road Directorate (see Vejdirektoratet (2001a)). The definition is based on the yield relations, in the sense, that arms with the right of way are major roads. However, when all or none of the arms have the right of way (e.g. signal controlled junctions), the two roads with the largest AADT are assumed to be the major arms. For most intersections, the arms categorized as major as well as minor roads have the same channelisation and yield relations within each category. When this is not the case, the arm with the largest average annual daily traffic defines the yield relations and channelisation of its category.

6.3.3 Model structure

The hierarchical generalized linear models set up in chapter 2 are the foundation for the data analysis, i.e. it is assumed that accidents, x, at a site are Poisson distributed with a gamma distributed dispersion effect. For intersections, the following model structure applies:

$$
\begin{aligned}
p\left(x|s\right) &= \frac{\lambda^x}{x!}e^{-\lambda} \\
f\left(s\right) &= \frac{\alpha^\alpha}{\Gamma\left(\alpha\right)}s^{\alpha-1}e^{-\alpha s}
\end{aligned}
$$

where $\lambda = \mu s$ is the expected number of accidents at the site and μ is the reference safety described by traits for traffic flow and road geometry etc. Using the terminology of chapter 2, the reference safety is denoted *the fixed effect part* of the mean. For road sections the following model structure is used:

$$
\begin{aligned}
p\left(x|s\right) &= \frac{\left(\lambda L\right)^x}{x!}e^{-\lambda L} \\
f\left(s\right) &= \frac{\left(\alpha L\right)^{\alpha L}}{\Gamma\left(\alpha L\right)}s^{\alpha L-1}e^{-\alpha L s}
\end{aligned}
$$

The expected number of accidents on a road section is traditionally assumed proportional to the length of the road. However, a study by Mountain et al. (1998) has shown non-linear relationship between site safety and the road length. This is mainly due to the fact that the deviation of site safety from the corresponding reference safety level is decreasing in the length of the road. To accommodate this, the length of the road is included as a weight in the gamma distribution for the dispersion effect[6]. The expected number of accidents is hence still considered proportional to the length of the road section in the road section models.

The models are set up for time periods of one year, i.e. a site with an observation period 1994-98 is represented by 5 records in the study. As a consequence, yearly changes at the site in e.g. traffic flow may be accounted for. This allows for better estimates of safety at a site in a particular year within the observation period.

In addition to local changes in site safety, the total number of reported accidents change from one year to another as depicted in figure 6.1 above. The figure implies general yearly changes in the expected number of accidents not accounted for by changes in the other traits. This trend may be due to changes in the state of technology of vehicles and accident coverage etc. Describing such yearly changes in the models is important for the ability of the model for predicting future safety levels, as it ensures that models are not rapidly outdated (Mountain et al. (1998)). Furthermore, in before and after studies it allows the general changes in safety levels to be separated from the treatment effects. One possibility is to include a categorical variable with levels corresponding to each year in the study period, thus allowing for each year to be treated separately. However, this considerably increases the number

[6] An approach also recently pursued in Hauer (2001).

of variables in the model. Furthermore, the level of years is limited to the study period, i.e. the trend variable cannot be projected beyond 1998, which makes future predictions of safety impossible.

Instead, within a site-group, a constant annual change in safety is assumed, and a variable, γ, indicating this time trend is included in the fixed effect part of the mean. The time trend variable represents yearly changes in safety not accounted for by the other traits. The first year of the study is used as base value, e.g. the years 1994,...,1998 are entered with values 0,...,4. Let $\gamma - 1$ be the annual relative change in safety due to trends in time, with a negative value indicating an average annual decrease in the expected accident count. The accumulated change $\triangle t$ years after the base year is:

$$\gamma^{\triangle t} - 1$$

The structure of the fixed effect part of the mean is based on the structure of the models developed by the Road Directorate, but with additional traits. For intersections, the general structure of the fixed effect part of the mean is:

$$\mu = a \cdot \gamma^{\triangle t} \cdot AADT_{ma}^{b_1} \cdot AADT_{mi}^{b_2} \cdot \exp\left(\sum_{k,l} \beta_k(l) Z_k(l)\right)$$

where a is a constant (the so-called *intercept*) and γ is the annual proportional change in safety due to trends in time. The regression variables $AADT_{ma}$ and $AADT_{mi}$ are the average annual daily incoming traffic on major and minor arms in the intersection. The part, $\exp\left(\sum_{k,l} \beta_k(l) Z_k(l)\right)$, is the contribution from the categorical variables with $Z_k(l)$ indicating whether or not variable k on level l is present at the site:

$$Z_k(l) = \begin{cases} 1, & \text{if variable } k \text{ is on level l} \\ 0, & \text{otherwise} \end{cases}$$

For road sections, the general structure of the fixed effect part of the mean is:

$$\mu = a \cdot \gamma^{\triangle t} \cdot AADT^b \cdot \prod_j Y_j^{\beta_j} \cdot \exp\left(\sum_{k,l} \beta_k(l) Z_k(l)\right)$$

where $AADT$ is the average annual daily traffic on the road section. The part $\prod_j Y_j^{\beta_j}$ is the contribution from numerical traits other than γ and $AADT$, with Y_j as the value of variable j. As pointed out by Jarret et al. (1997) and Austin (1995a) & (1995b), recorded site data may contain errors. Inaccuracies occurring at random are included in the random variation of accident counts. Systematic errors in data for a group of sites may, on the other hand, be considered part of the site-specific conditions and are thus modelled by the dispersion effect, s.

The models are first estimated for the total reported number of accidents. The selected traits for these models are retained for the models using injury accidents only. This is mainly in order for the models to be comparable. However, due to the lesser amount of injury accidents to the total accident count, some of these traits may now fail to show statistical significance. Despite this fact and because insignificant variables still contribute to the explained variation they are kept in the model. The fitting of the models is described in appendix D.

6.4 Output analysis

The selected traits for the models are listed in table 6.3. Estimated coefficients and a description of the selection of traits are given in appendix D.

Intersections	**Road sections**			
Junctions	Motorways	Motortraffic-ways	Other roads	Remaining
Year	Year	Year	Year	Year
No. of arms	Length	Length	Length	Length
Frontage	AADT	AADT	AADT	AADT
Major	Width rest	Width of lanes	Width	Speed
AADT	Speed	Speed	Width of lanes	Frontage
Yield relations	No. of lanes	No. of lanes	Speed	Edge
Channelisation	Barrier middle	Median	No. of lanes	Bicycle path
Minor			Barrier	Barrier middle
AADT			Barrier middle	
Yield relations			Median	
Channelisation			Frontage	

Table 6.3: The selected traits for each site-group.

The estimated average annual proportional change, γ, in safety attributable to time trends is found to be rather small. However, the accumulated change over the observation period is relatively high (reflected in figure 6.1). Assuming γ is constant, this implies that accident models which do not allow for trends in time will become rapidly outdated for prediction purposes. As an example, motorways have an estimated average annual decrease in safety due to trends in time of 2%. Other circumstances being equal, this results in an estimated decrease of 10% over a five year period. The time trend variable, γ, was not found to be statistically significant within all the site-groups. However, for site-groups with a relatively large[7]

[7] At least 5% over the study period 1994-98.

accumulated proportional change in safety attributable to time trends, γ was still included in the fixed effect part of the mean.

One must be careful not to over-interpret the values of the estimated coefficients for the traits in the model. Some traits vary simultaneously, and the coefficient of one trait does not directly reflect its effect on safety. Instead, it is an indication of the effect of this variable when the other variables in the model are accounted for. However, one should keep in mind that models describe the data included in the study. Varying one trait while other traits are kept constant is a fictitious example, which may not be meaningful in this data set.

In addition, a trait may vary with one or more site characteristics not included in the model. In that case, the effect due to the omitted variables may be ascribed to the included trait and hereby inflate or deflate its coefficient. As an example, the speed limit on road sections is normally correlated with the lane width on the road. Higher speed limits are usually allowed on roads with wide lanes, because they in general are relatively safe. Omitting the geometry variable describing the lane width from the model will affect the coefficient for the speed limit variable, and may give the misleading interpretation that safety increases with speed. As Kulmala (1995) points out, conclusions about effects of traits on safety should only be made on the basis of controlled before and after studies.

These facts make it difficult to compare the estimated coefficients in this study to findings of other studies with different traits. However, the coefficients for the most important determinant in accident models, the traffic flow, are of the same magnitude as the corresponding parameters in the models in Vejdirektoratet (2001b).

In general, not all relevant site characteristics are included as traits in the model. The dispersion part of the mean, s, will however account for at least part of the effect attributable to site characteristics not included in the model (see chapter 2 for a discussion).

Any changes in the categorizing of accident severity have not been included either. As an example, since 1997 concussions have been labelled slightly injuries as opposed to prior 1997 when they where considered serious injuries (see Vejdirektoratet (2001a)). However, the time-trend variable will account for some of the general effects of this change in categorization.

Chapter 7

Conclusions and suggestions

This chapter summarizes the work presented in the thesis and outlines the conclusions. At the end of the chapter some suggestions for future work are provided.

The general purpose of the study was to improve *state of the art* in hot spot safety work in Denmark. This aim has been achieved by proposing improved models and methods for each of the phases in hot spot safety work. The superiority of the models and methods has been documented through simulation studies.

7.1 Summary and conclusions

The basis for hot spot safety work is the models describing variation in accident counts. This variation may conceptually be separated into explained and unexplained variation. The explained variation is the part of variation that may be ascribed to differences in the traits. All variation not ascribed to the traits is termed unexplained and is modelled random. The thesis points out the fact that not all site-specific conditions are included as traits. Hence, the unexplained part of variation covers variation related to non-observable quantities or observable quantities not included as traits, as well as random variation within sites as illustrated in figure 2.1.

In the thesis, a class of models called *Poisson-gamma hierarchical generalized linear models* (Pg-HGLMs) has been proposed for describing the variation in road accident counts. The accident models are disaggregated on time periods of one year, which assures that yearly changes in traffic as well as in other traits may be accounted for. General trends in accident counts may also be included in the models. The Pg-HGLMs allow extra error components in the mean, the so-called dispersion effects. The dispersion effects represent the site-specific conditions not included as traits in the model, i.e. the non-random part of the unexplained variation described above. In addition, a dispersion effect models interdependence between yearly accident counts at the same site. The proposed Poisson-gamma hierarchical generalized linear models have been found to provide better estimates of site safety than the models currently at use in Denmark.

Lee and Nelder (2001) suggest a technique for estimating the parameters of models corresponding to the intersection model, which, after a few modifications, may also be used to estimate the parameters of the road section models. From these estimation procedures, the parameters of the proposed accident models have been estimated for the national and regional road network in Denmark using data from the Road Sector Information System, VIS. No specific accident models are estimated for the local road network, as site data on local roads are not collected in a systematic manner.

The dispersion effect in the proposed accident models expresses how the expected accident frequency at a site deviates from the expected accident frequency at sites with similar traits. It is thus an indication of the *level of hotness* at a site, and may be used for targeting the so-called *hot spots* in the road network. Because the dispersion effect is an unknown unobservable model quantity that can only be estimated, an additional measure denoted the *evidence of hotness* is proposed, which takes the uncertainty in the estimate into consideration. On road sections, targeting hot spots from the evidence of hotness results in a marginally higher sensitivity than a method using the level of hotness. On intersections, the two measures result in the same sensitivity. In general, targeting hot spots from the methods proposed in the thesis were shown to give a marginally higher sensitivity than the method used on the Danish road network today.

Funds for hot spot safety work are limited, and one needs to prioritize between hot spots and safety improving measures. The proposed accident models provide estimates of site safety, which incorporate the estimated dispersion effect at a site. Hence, prioritizing methods based on these site safety estimates will take into account the level of hotness at a site as well as random variation in accident counts.

A new model for estimating the effect of hot spot treatment work is proposed. The model is based on the site safety estimates provided by the accident models, and takes into account the so-called *regression to the mean effect* as well as changes in traffic flow and other traits. The proposed method is found to give better estimates of the effect of treatment than the method currently used in Denmark. In addition, it outperforms the methods as yet suggested in the international literature. The improved estimates of treatment effect will improve the foundation for prioritizing of hot spots and safety measures as well as for the road safety audit.

The proposed models and methods in this thesis are believed to contribute to the foundation for improvement of hot spot safety work in Denmark in general.

7.2 Suggestions for future work

In the analysis of accident counts in Denmark, differences in the level of reporting between police jurisdictions are not accounted for, as such information is not available. However, differences in coverage affect the model estimates considerably when site and accident data of different jurisdictions are pooled. Studies of differences in the level of reporting between police jurisdictions are thus of great importance for future hot spot safety work.

Also, in the accident models, traits such as traffic flow are assumed measured without error. However, the annual daily traffic at a site in the national or regional road network is estimated from relatively few days of traffic measurements, and may thus contain a considerable amount of errors. In addition, the Danish national traffic growth index for average yearly increases in traffic has been used to project traffic flows for years with no measurements. The national growth index is based on developments in the total traffic flow and is thus a rather crude estimate of the growth in traffic at individual sites. Studies of how errors in the measured traits affect the site safety estimates may be of interest. However, the need for a more differentiated traffic growth index has been recognized by the Road Directorate (see Vejdirektoratet (2002a)).

A systematic collection of road data and improved location of accident occurrences are needed for the local road network. A structured site and accident database for local roads, corresponding to the Road Sector Information System, VIS, for state and regional roads, will allow the proposed models and methods of this thesis to be applied to the local road network as well.

A minor part of the local authorities have very few reported accidents a year, and instead they use the populations' perception of safety as the basis for hot spot safety work. However, the perception of safety is subjective and unpredictable and should not be used as guidance in road safety work. Instead, studies of how the term *accident frequency* may be extended to include other objective measures such as *near-misses* etc. are needed.

Appendix A

The framework of the models

The objective of this chapter is to briefly introduce and describe developments made in the theory of generalized linear models and to derive methods for estimating the components of the Poisson-gamma hierarchical generalized linear models set up in chapter 2. In addition, the mathematical framework behind the road section model is described.

A.1 Generalized linear models

The value of an observation x may be described as a result of two contributors; a systematic part described through a set of explanatory variables, and an uncontrolled random part. The systematic component corresponds to the expectation of x, $\mu = E(x)$. The random part may be described by a probability distribution with mean value μ.

An early attempt to model this has been through general linear models (GLM) where the mean is given as a linear model of the explanatory variables. Let $\mathbf{Z}$ and $\boldsymbol{\beta}$ be a set of explanatory variables and the set of corresponding fixed effect parameters respectively, such that $\mathbf{Z}\boldsymbol{\beta}$ describes the systematic part of the observations $\mathbf{x}$. The response variables are assumed to be independent, and the random part to be normal distributed:

$$\begin{aligned} \boldsymbol{\mu} &= \mathbf{Z}\boldsymbol{\beta} \\ \mathbf{x} &\in N\left(\boldsymbol{\mu}, \sigma^2 \mathbf{I}\right) \end{aligned} \tag{A.1}$$

The first general extension of this model was the generalized linear models (GLIM), which allowed for a non-normal error distribution (see Nelder and Wedderburn (1972) or McCullagh and Nelder (1989)):

$$\begin{aligned} \boldsymbol{\eta} &= g(\boldsymbol{\mu}) = \mathbf{Z}\boldsymbol{\beta} \\ \mathbf{x} &\in f(\boldsymbol{\mu}) \end{aligned} \tag{A.2}$$

The additivity of the systematic effects occurs on a monotonic transformation (the link function) of the mean rather than on the mean itself. This enables the simple use of likelihood-based procedures for the estimation of parameters as in the GLM model. Hence, the link function $g(\cdot)$ is thus used to linearize the systematic part, and the linear predictor, η, describes the linear dependency of the explanatory variables. The random part is described through the distribution of the response $\mathbf{x}$, where $f(\cdot)$ is an arbitrary distribution belonging to the exponential family.

Later, the general linear models were extended to hierarchical models or general linear mixed models (GLMMIX). The name refers to the fact that the model for $\boldsymbol{\mu}$ contains both fixed effect parameters $\boldsymbol{\beta}$ and random effects $\mathbf{s}$. Let $\mathbf{Z}_1\boldsymbol{\beta}$ describe the fixed effect part of $\boldsymbol{\mu}$, while $\mathbf{s}$ are random effects describing the variation of the mean around $\mathbf{Z}_1\boldsymbol{\beta}$. Both $\mathbf{x}$ and $\mathbf{s}$ are assumed normal distributed:

$$\begin{aligned} \boldsymbol{\mu} &= \mathbf{Z}_1\boldsymbol{\beta} + \mathbf{Z}_2\mathbf{s} \\ \mathbf{s} &\in N\left(\mathbf{0}, \sigma^2\mathbf{I}\right) \\ \mathbf{x} &\in N\left(\boldsymbol{\mu}, \sigma^2\mathbf{I}\right) \end{aligned}$$

Recently this hierarchical structure has been developed for the generalized linear models, producing a class of generalized linear mixed models (GLMM), in which random effects, $\mathbf{s}$, with assumed normal distribution are allowed in the linear predictor:

$$\begin{aligned} \boldsymbol{\eta} &= g(\boldsymbol{\mu}) = \mathbf{Z}_1\boldsymbol{\beta} + \mathbf{Z}_2\mathbf{s} \\ \mathbf{s} &\in N\left(\mathbf{0}, \sigma^2\mathbf{I}\right) \\ \mathbf{x} &\in f(\boldsymbol{\mu}) \end{aligned}$$

The random effects, $\mathbf{s}$, express the overdispersion. Breslow and Clayton (1993) have provided a unifying approach on how to estimate the parameters and random effects for this type of model.

The most recent extension is that the random effects, $\mathbf{s}$, in the linear predictor of GLMM are not restricted to follow a normal distribution. It may come from an arbitrary distribution conjugate to that of the response $\mathbf{x}$. Solutions for estimating the fixed effect parameters and random effects in such models are given in Lee and Nelder (1996) & (2001). This class of models is denoted *hierarchical generalized linear models* (HGLM).

The following sections describe the hierarchical generalized linear model for modelling variation in traffic accident counts in intersections and on road sections.

A.2 Modelling variation in intersections

Let $\mathbf{x} = \{x_{it}\}_{i=1,\ldots,I,t=1,\ldots,T_i}$ be a set of conditionally independent observations such that the conditional distribution of x_{it} given a dispersion effect, s_i, is the Poisson distribution with mean $\lambda_{it} = \mu_{it} s_i$. The observation period of site i is $]0; T_i]$. Assume the dispersion effects, $\mathbf{s} = \{s_i\}_{i=1,...,I}$, are independent and identically gamma distributed with shape parameter α and mean 1:

$$\begin{aligned} p(x_{it}|s_i) &= p(X_{it} = x_{it}|s_i) = \frac{(\lambda_{it})^{x_{it}}}{x_{it}!} e^{-\lambda_{it}},\ \lambda_{it} \in \mathbb{R}_+, x_{it} \in \mathbb{N}_0 \\ f(s_i) &= f(S_i = s_i) = \frac{\alpha^\alpha}{\Gamma(\alpha)} s_i^{\alpha-1} e^{-\alpha s_i},\ \alpha \in \mathbb{R}_+, s_i \in \mathbb{R}_+ \end{aligned} \tag{A.3}$$

The random effect, s, describes the overdispersion in the Poisson error distribution. Because $E(S_i) = 1$, a unique set of parameters for the gamma distribution exists. Let the conditional mean of x_{it} given s_i be modelled by a linear model:

$$\eta'_{it} = g(\lambda_{it}) = g(\mu_{it} s_i) = \eta_{it} + v_i \tag{A.4}$$

where η'_{it} is called the linear predictor for λ_{it} with link function $g(\cdot)$. In a Poisson-gamma distribution, the canonical link function for the conditional mean, $g(\cdot)$, is the logarithm, $g(\cdot) = \ln(\cdot)$. The linear predictor, η_{it}, is the fixed effect part of the mean corresponding to the mean in the GLIM model (A.2):

$$\eta_{it} = g(\mu_{it}) = \mathbf{Z}_{it}\boldsymbol{\beta} = \sum_{j=1}^{J} z_{itj}\beta_j \Leftrightarrow \mu_{it} = g^{-1}(\mathbf{Z}_{it}\boldsymbol{\beta}) \tag{A.5}$$

while $v_i = \ln(s_i)$ is the random term describing the deviation from η_{it}. Below, s_i is denoted the *dispersion part* of the mean. Thus (A.4) models both the fixed effects for η_{it} and the overdispersion described by v_i. The model (A.3)-(A.5) is called a *Poisson-gamma hierarchical generalized linear model* (Pg-HGLM) and belongs to the class of conjugate HGLMs. The dispersion effect, s_i, is both a parameter in the conditional distribution of x_{it}, and at the same time a random variable with its own distribution.

Because accident counts are conditionally independent, the joint distribution of $x_{i1}, ..., x_{iT_i}|s_i$ is calculated as the product of the individual Poisson distributions. The marginal distribution of all accident observations $\mathbf{x} = (\mathbf{x}_1, ..., \mathbf{x}_I)$, $\mathbf{x}_i = (x_{i1}, ..., x_{iT_i})$, is:

$$\begin{aligned} P(\mathbf{x}) &= \prod_{i=1}^{I} P(\mathbf{X}_i = \mathbf{x}_i) = \prod_{i=1}^{I} \int_0^\infty p(\mathbf{X}_i = \mathbf{x}_i|s_i) f(S_i = s_i)\, \partial s_i \\ &= \prod_{i=1}^{I} \int_0^\infty \left[\prod_{t=1}^{T_i} \frac{(\mu_{it}s_i)^{x_{it}}}{x_{it}!} e^{-\mu_{it}s_i} \right] \frac{\alpha^\alpha}{\Gamma(\alpha)} s_i^{\alpha-1} e^{-\alpha s_i} \partial s_i \\ &= \prod_{i=1}^{I} \left[\left(\prod_{t=1}^{T_i} \frac{\mu_{it}^{x_{it}}}{x_{it}!} \right) \frac{\alpha^\alpha}{\Gamma(\alpha)} \int_0^\infty s_i^{\alpha+x_{i\cdot}-1} e^{-(\alpha+\mu_{i\cdot})s_i} \partial s_i \right] \\ &= \left[\prod_{i=1}^{I} \prod_{t=1}^{T_i} \frac{\mu_{it}^{x_{it}}}{x_{it}!} \right] \prod_{i=1}^{I} \left[\frac{\alpha^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha + x_{i\cdot})}{(\alpha + \mu_{i\cdot})^{\alpha + x_{i\cdot}}} \right] \end{aligned}$$

where $x_{i\cdot} = \sum_{t=1}^{T_i} x_{it}$ and $\mu_{i\cdot} = \sum_{t=1}^{T_i} \mu_{it}$ are the total reported number of accidents at site i and the corresponding reference safety level respectively.

A.2.1 Estimation of fixed and random effects

The fixed effect parameters of models are usually estimated in the marginal likelihood function of the observed response variables $\mathbf{x}$:

$$\begin{aligned} l(\boldsymbol{\beta}, \alpha; \mathbf{x}) &= \ln P(\mathbf{x}) \propto \ln \left[\prod_{i=1}^{I} \prod_{t=1}^{T_i} \mu_{it}^{x_{it}} \right] \prod_{i=1}^{I} \left[\frac{1}{(\alpha + \mu_{i\cdot})^{\alpha + x_{i\cdot}}} \right] \\ &= \sum_{i=1}^{I} \sum_{t=1}^{T_i} x_{it} \ln \mu_{it} - \sum_{i=1}^{I} (\alpha + x_{i\cdot}) \ln (\alpha + \mu_{i\cdot}) \end{aligned}$$

The marginal score function for the determination of the fixed effect parameter β_j becomes:

$$\begin{aligned} \frac{\partial l}{\partial \beta_j} &= \frac{\partial}{\partial \beta_j} \left[\sum_{i=1}^{I} \sum_{t=1}^{T_i} x_{it} \ln \mu_{it} - \sum_{i=1}^{I} (\alpha + x_{i\cdot}) \ln (\alpha + \mu_{i\cdot}) \right] \qquad \text{(A.6)} \\ &= \sum_{i=1}^{I} \sum_{t=1}^{T_i} x_{it} \frac{1}{\mu_{it}} \frac{\partial \mu_{it}}{\partial \beta_j} - \sum_{i=1}^{I} (\alpha + x_{i\cdot}) \frac{1}{\alpha + \mu_{i\cdot}} \frac{\partial (\alpha + \mu_{i\cdot})}{\partial \beta_j} \\ &= \sum_{i=1}^{I} \sum_{t=1}^{T_i} x_{it} z_{itj} - \sum_{i=1}^{I} \frac{\alpha + x_{i\cdot}}{\alpha + \mu_{i\cdot}} \left(\sum_{t=1}^{T_i} \mu_{it} z_{itj} \right) \\ &= \sum_{i=1}^{I} \sum_{t=1}^{T_i} \left(x_{it} - \frac{\alpha + x_{i\cdot}}{\alpha + \mu_{i\cdot}} \mu_{it} \right) z_{itj} \end{aligned}$$

In (A.6), the fact that because $g^{-1}(\cdot) = \exp(\cdot)$, then $\partial \mu_{it} / \partial \beta_j = \mu_{it} z_{itj}$ is used. Furthermore $\partial \mu_{i\cdot} / \partial \beta_j = \sum_{t=1}^{T_i} \partial \mu_{it} / \partial \beta_j$ is applied in the equation. Solving $\partial l / \partial \beta_j = 0, \forall j = 1, ..., J$ results in the maximum likelihood estimators (MLEs) for the fixed effects parameters $\boldsymbol{\beta}$.

Estimation of the dispersion effects has received increasing attention in recent years. Because $\mathbf{s}$ is not a part of the marginal distribution of $\mathbf{x}$, Lee and Nelder (1996) suggest using the so-called *h-likelihood* for estimating both the fixed effect parameters, $\boldsymbol{\beta}$, and random effects $\mathbf{s}$. For given α, the h-likelihood is the joint distribution of the observed $\mathbf{x}$ and unobserved variables $\mathbf{s}$:

$$\begin{aligned} h(\boldsymbol{\beta}, \mathbf{s}; \alpha, \mathbf{x}) &\equiv l_0(\boldsymbol{\beta}; \mathbf{x}|\mathbf{s}) + l_1(\alpha; \mathbf{s}) = \ln [p(\mathbf{x}|\mathbf{s}, \boldsymbol{\beta}) f(\mathbf{s}|\alpha)] \qquad \text{(A.7)} \\ &= \sum_{i=1}^{I} \sum_{t=1}^{T_i} \ln p(x_{it}|s_i) + \sum_{i=1}^{I} \ln f(s_i|\alpha) \end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=1}^{I}\sum_{t=1}^{T_i} \ln\left(\frac{(\mu_{it}s_i)^{x_{it}}}{x_{it}!}e^{-\mu_{it}s_i}\right) + \sum_{i=1}^{I}\ln\left(\frac{\alpha^\alpha}{\Gamma(\alpha)}s_i^{\alpha-1}e^{-\alpha s_i}\right) \\
&\propto \sum_{i=1}^{I}\sum_{t=1}^{T_i}\left(x_{it}\ln(\mu_{it}s_i) - \mu_{it}s_i\right) + \sum_{i=1}^{I}\left(\alpha\ln s_i - \alpha s_i\right) \\
&= \sum_{i=1}^{I}\sum_{t=1}^{T_i}\left(x_{it}\ln\mu_{it}\right) + \sum_{i=1}^{I}\left((\alpha + x_{i\cdot})\ln s_i - (\alpha + \mu_{i\cdot})\, s_i\right)
\end{aligned}
$$

The last line in expression (A.7) is the kernel of the h-likelihood function for $\boldsymbol{\beta}$ and $\mathbf{s}$, where $\boldsymbol{\beta}$ enters via (A.5). The estimates of $\boldsymbol{\beta}$ and $\mathbf{s}$ are called the maximum h-likelihood estimates (MHLEs) and are obtained by equating the components of the score function, $D_{(\boldsymbol{\beta},\mathbf{s})}h(\boldsymbol{\beta},\mathbf{s};\alpha,\mathbf{x}) = (\partial h/\partial\boldsymbol{\beta}, \partial h/\partial\mathbf{s})$, to zero. Assuming the fixed effect parameters, $\boldsymbol{\beta}$, are given, i.e. the reference safety levels, $\boldsymbol{\mu}$, are known, the h-score function for the determination of the dispersion effect, s_i, becomes:

$$
\begin{aligned}
\frac{\partial h}{\partial s_i} &= \frac{\partial}{\partial s_i}\left[\sum_{i=1}^{I}\left((\alpha + x_{i\cdot})\ln s_i - (\mu_{i\cdot} + \alpha)\, s_i\right)\right] \\
&= (\alpha + x_{i\cdot})\frac{1}{s_i} - (\alpha + \mu_{i\cdot})
\end{aligned}
$$

Equating the h-score function, $\partial h/\partial s_i$, to zero results in the following maximum h-likelihood estimate for s_i:

$$
\partial h/\partial s_i = (\alpha + x_{i\cdot})\frac{1}{s_i} - (\alpha + \mu_{i\cdot}) = 0 \Leftrightarrow \widehat{s}_i = \frac{\alpha + x_{i\cdot}}{\alpha + \mu_{i\cdot}} \tag{A.8}
$$

The MHLE for s_i may be interpreted as an estimate of $E(S_i|\mathbf{x}_i)$, where $E(S_i|\mathbf{x}_i)$ is the empirical Bayes posterior estimator of s_i when α is estimated from data as described below.

The corresponding h-score function for the determination of the fixed effect parameter β_j becomes:

$$
\frac{\partial h}{\partial \beta_j} = \frac{\partial}{\partial \beta_j}\left[\sum_{i=1}^{I}\sum_{t=1}^{T_i}\left(x_{it}\ln\mu_{it}\right) + \sum_{i=1}^{I}\left((\alpha + x_{i\cdot})\ln s_i - (\alpha + \mu_{i\cdot})\, s_i\right)\right] \tag{A.9}
$$

which is equal to the marginal score function (A.6) for β_j. Hence, the MHLE for $\boldsymbol{\beta}$ is equal to the MLE. Using (A.6) together with the maximum h-likelihood estimate for s_i in (A.8), the h-likelihood equation for the fixed effect parameter β_j becomes:

$$
\frac{\partial h}{\partial \beta_j} = \sum_{i=1}^{I}\sum_{t=1}^{T_i}\left(x_{it} - \widehat{s}_i\mu_{it}\right) z_{itj} = 0 \tag{A.10}
$$

which is the Poisson estimating equation with offset $\widehat{v}_i = \log(\widehat{s}_i)$. The MHLEs for $\mathbf{s}$ and $\boldsymbol{\beta}$ are similar to the estimators suggested by Tsutakawa (1988) and Van Duijn (1993) for a multiplicative Poisson-gamma model. The MHLEs for s_i and β_j are interdependent, because in (A.8) $\mu_{i\cdot} = \sum_{t=1}^{T_i} g^{-1}\left(\sum_{j=1}^{J} \beta_j z_{itj}\right)$.

The maximum likelihood estimator for the dispersion component α may be biased due to the estimation of the fixed effect parameters[1]. Instead Lee and Nelder (1996) suggest using an adjusted profile h-likelihood in analogy with the restricted maximum likelihood (REML) estimation (see Schall (1991)):

$$\begin{aligned} h_A &= h + \frac{1}{2}\ln\left[\det\left(2\pi H^{-1}\right)\right] \\ &= l_1(\alpha; \mathbf{s}) - \frac{1}{2}\ln\left[\det(H)\right] \end{aligned}$$

with H being the Hessian matrix corresponding to the estimation of $\boldsymbol{\beta}$ and $\mathbf{s}$:

$$H = \begin{bmatrix} \frac{\partial^2 h}{\partial \beta^2} & \frac{\partial^2 h}{\partial \beta \partial \mathbf{s}} \\ \frac{\partial^2 h}{\partial \mathbf{s} \partial \beta} & \frac{\partial^2 h}{\partial \mathbf{s}^2} \end{bmatrix} = \begin{bmatrix} Z^T W Z & Z^T W Y \\ Y^T W Z & Y^T W Y + U \end{bmatrix}$$

Here Z is the matrix of explanatory variables and $W = (\partial\boldsymbol{\lambda}/\partial\boldsymbol{\eta}')^2 (V(\boldsymbol{\lambda}))^{-1} = \text{diag}(\boldsymbol{\lambda})$ is the GLM weight function with variance function $V(\boldsymbol{\lambda}) = \boldsymbol{\lambda}$. The matrix Y is a $\left(\sum_{i=1}^{I} T_i\right) \times I$ group indicator matrix $Y = (\partial\boldsymbol{\eta}'/\partial\mathbf{v})$ which in the Pg-HGLM has elements:

$$[Y]_{(i,t),k} = \begin{cases} 1, \ k = i \\ 0, otherwise \end{cases}$$

In the Hessian, U is an $I \times I$ diagonal matrix that for conjugate hierarchical generalized linear models has elements:

$$[U]_{(i,i)} = -\left(\frac{\partial^2 l_1(\alpha; \mathbf{s})}{\partial v_i^2}\right)$$

Hence, in a Poisson-gamma hierarchical generalized linear model $[U]_{(i,i)} = \alpha s_i$. Because $\partial(\ln(\det H))/\partial\alpha = \text{trace}(H^{-1}(\partial H/\partial\alpha)) = \text{trace}(K(\partial U/\partial\alpha))$, the score function for α becomes:

$$\frac{\partial h_A}{\partial \alpha} = \frac{\partial l_1(\alpha; \mathbf{s})}{\partial \alpha} - \frac{1}{2}\text{trace}\left[K\left(\frac{\partial U}{\partial \alpha}\right)\right] \tag{A.11}$$

[1]Simulation studies show that using the maximum likelihood estimate of α will result in the estimate $\widehat{\alpha}$ converging to infinity and the estimated dispersion effects, $\widehat{\mathbf{s}}$, converging to their means $E(\mathbf{s}) = \mathbf{1}$. This is due to the fact that, as the estimated dispersion effects, $\widehat{\mathbf{s}}$, are regressed towards unity, the estimated variance of the dispersion effects decreases, i.e. the estimated dispersion component, $\widehat{\alpha}$, increases because $V(s) = 1/\alpha$. The regression of $\mathbf{s}$ towards unity is increasing with $\widehat{\alpha}$ (see A.8). Thus in each iteration the estimated value of α is increased and $\widehat{\mathbf{s}}$ is further regressed towards unity.

In (A.11) the $I \times I$ matrix K is given as the bottom right hand corner of H^{-1} corresponding to the dispersion effects (see Lee and Nelder (1996)) and $[\partial U/\partial\alpha]_{(i,i)} = s_i$. The maximum adjusted profile h-likelihood estimator (MAPHLE) for α is found by equating the score function (A.11) for α to 0. The estimate of α reflects the homogeneity of site safety within the site-group.

The size of the Hessian matrix increases with the number of observations, i.e. with the number of sites and observation periods. Hence, the adjusted profile h-likelihood, h_A, for α may become considerably hard to calculate for large datasets. Instead Lee and Nelder (2001) suggest estimating α from a second generalized linear model with responses $\mathbf{d}^*$, prior weights $\mathbf{1} - \mathbf{q}^a$ and scale parameter 2:

$$\boldsymbol{\eta} = g(\boldsymbol{\delta}) = \log(\boldsymbol{\delta}) \tag{A.12}$$

$$\mathbf{d}^* \in gamma\left(\frac{1}{2}, 2\boldsymbol{\delta}\right)$$

with $\boldsymbol{\delta} = \mathbf{1}/\boldsymbol{\alpha}$ and linear predictor $\boldsymbol{\eta}$. The response, d_i^*, for intersection i is the standardized adjusted deviance for estimating s_i and is calculated as:

$$d_i^* = \frac{d_i^a}{1 - q_i^a}$$

The deviance for site i in the estimation of the dispersion effects $\mathbf{s}$ is calculated as:

$$d_i = 2\left(s_i - \log(s_i) - 1\right)$$

From the Taylor approximation, $d_i \simeq V(s_i) = \delta$, hence d_i is chosen as a suitable statistic for measuring δ. The variance of the deviances differs from site to site:

$$V(d_i) \simeq \delta^2 (1 - p_i)^2$$

where p_i is the i^{th} diagonal element in the hat matrix[2] for estimating $\mathbf{s}$. For p_i small, the variance is large. This may be seen as an indication of a relatively large amount of observations behind the estimate $\widehat{s}_i$, i.e. a relatively large number of accidents. In order to compare deviances at different intersections, the deviances are scaled to ensure equal variances. Because the same dispersion parameter α apply to all intersections in a site-group, it is sufficient to divide the deviance, d_i, by $1 - p_i$. The weight $1 - p_i$ is further generalized to $1 - q_i$ for discrete data:

$$1 - q_i = \frac{\delta x_{i\cdot}}{\delta x_{i\cdot} + 1}$$

[2]The $I \times I$ hat matrix for estimating $\mathbf{s}$ is the bottom right hand corner of $V\left(V^T M V\right)^{-1} V^T M$, where $V = \begin{bmatrix} Z & Y \\ 0 & E \end{bmatrix}$ and $M = \begin{bmatrix} W & 0 \\ 0 & U \end{bmatrix}$ with E as the identity matrix and with Z, Y, W and U defined as above.

which again is close to unity for a relatively large amount of accidents. Because a $x_{i\cdot} = 0$ in $1 - q_i$ results in the absence of the i^{th} datum, the weights are adjusted to ensure that all data are used in the estimation procedure:

$$1 - q_i^a = \frac{\delta x_{i\cdot}}{\delta x_{i\cdot} + 1} + I\left(x_{i\cdot} = 0\right)\epsilon_i$$

The adjustment component, ϵ_i, at site i is calculated as:

$$\epsilon_i = \frac{\lambda_{i\cdot}}{\lambda_{i\cdot} + 1}$$

In accordance with this, the deviances for the estimation of the dispersion effects $\mathbf{s}$ are also adjusted

$$d_i^a = 2\left(s_i - \log\left(s_i\right) - 1\right) + I\left(x_{i\cdot} = 0\right)\epsilon_i\delta$$

The solution to (A.12) may be calculated as:

$$\widehat{\delta} = \frac{\sum_{i=1}^{I} d_i^a}{\sum_{i=1}^{I}\left(1 - q_i^a\right)} \qquad \text{(A.13)}$$

Because the estimates of $\mathbf{s}$, $\boldsymbol{\beta}$ and α are interdependent, they will have to be derived iteratively. The components of the P-g HGLM describing variation in road accident counts in intersections are estimated from (A.8), (A.10) and (A.13). For α given, the task of estimating $\boldsymbol{\beta}$ and $\mathbf{s}$ is the problem of iteratively evaluating (A.8) and solving (A.10) till convergence is reached. Given estimates of $\boldsymbol{\beta}$ and $\mathbf{s}$, the dispersion parameter α is reestimated from equation (A.11) or (A.13), and the process is repeated.

An initial value of α may be found by equating the sample variance of the reported accident counts, $\mathbf{x}$, with the variance in the negative binomial distribution, using the sample mean, $\overline{x}$, as the estimated reference mean for all sites. This is the so-called method of moments (see Maycock and Maher (1988)):

$$\frac{1}{n}\sum_{i=1}^{I}\sum_{t=1}^{T_i}\left(x_{it} - \overline{x}\right)^2 = \overline{x} + \frac{\overline{x}^2}{\alpha} \Leftrightarrow \alpha = \frac{\overline{x}^2}{\frac{1}{n}\sum_{i=1}^{I}\sum_{t=1}^{T_i}\left(x_{it} - \overline{x}\right)^2 - \overline{x}}$$

where n is the total number of records in the sample. In order to perform the first evaluation of $\mathbf{s}$, one needs an initial value of $\boldsymbol{\beta}$ in addition to α. The initial value of $\boldsymbol{\beta}$ may be given as the solution to the generalized linear model (GLIM) in (A.2) with a Poisson error distribution. The algorithm for estimating α, $\boldsymbol{\beta}$ and $\mathbf{s}$ is illustrated in figure A.1.

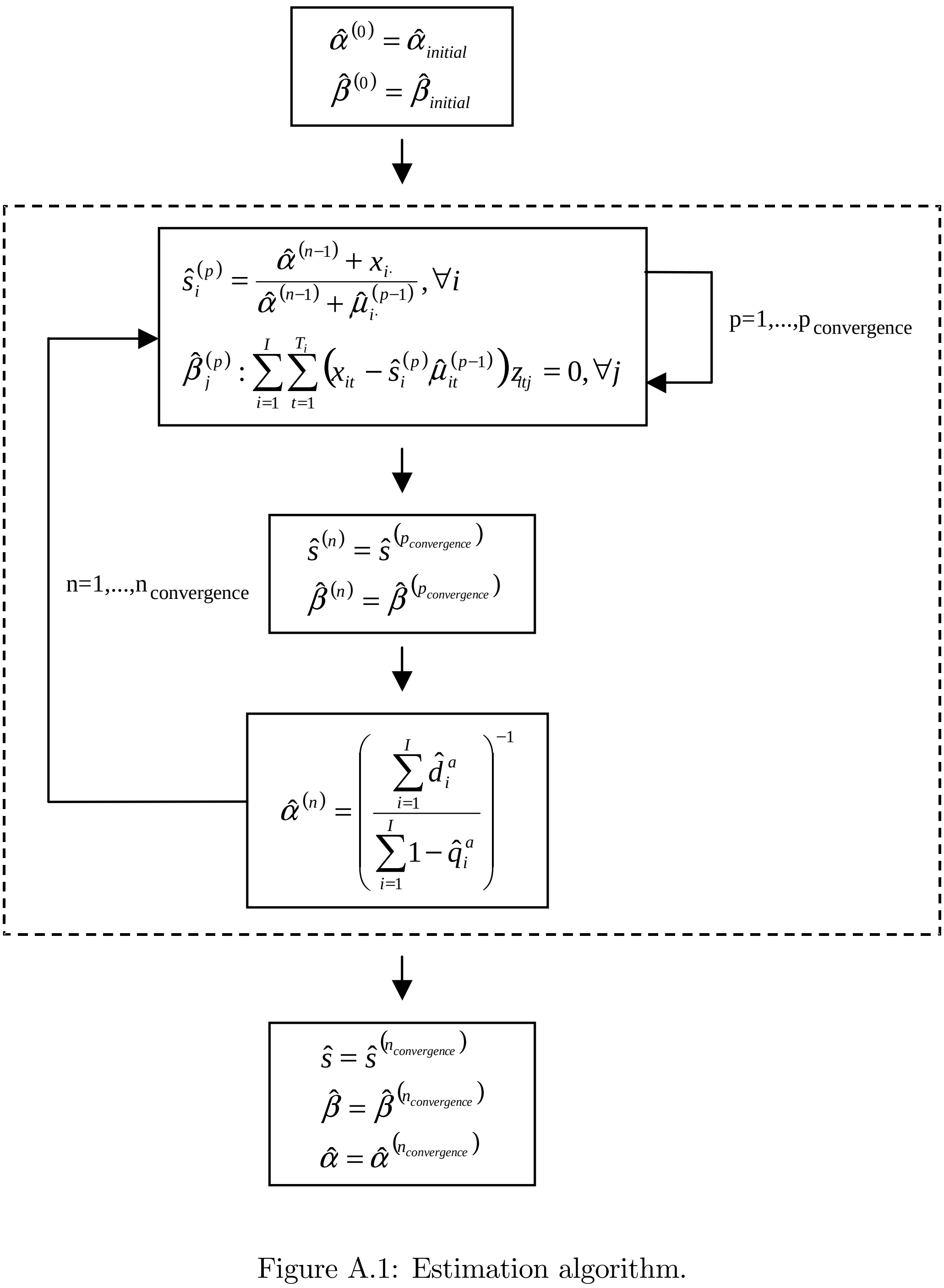


Figure A.1: Estimation algorithm.

A.2.2 In an empirical Bayesian framework

As above, let the variation in accident counts at a site i in a site-group H be described by the model defined by (A.3)-(A.5). The h-likelihood in (A.7) can be viewed as a Bayesian posterior under an improper uniform prior for $\boldsymbol{\beta}$ and α. Accordingly, in a Bayesian framework the distribution of S_i is called the prior distribution. At each site i, the prior distribution is updated with the given accident information (the observed deviations from the reference safety) using Bayes' theorem:

$$f(s_i|\mathbf{x}_i) \propto p(\mathbf{x}_i|s_i)\, f(s_i) \tag{A.14}$$

The conditional distribution of S_i is a gamma distribution with updated parameters $(\alpha + x_{i\cdot}, 1/(\alpha + \mu_{i\cdot}))$, which in a Bayesian framework is denoted the posterior distribution of S_i (see Lee (1994)). The gamma distribution is thus conjugate[3] to the Poisson distribution. For intersections, the cumulative posterior density function is:

$$P(S_i \leq c|\mathbf{x}_i) = \int_0^c \frac{(\alpha + \mu_{i\cdot})^{\alpha + x_{i\cdot}}}{\Gamma(\alpha + x_{i\cdot})} s^{\alpha + x_{i\cdot} - 1} e^{-(\alpha + \mu_{i\cdot})s} ds \tag{A.15}$$

For known α and β, the interval $]0; c]$ is the Bayes credible interval for S_i of level $P(S_i \leq c|\mathbf{x}_i)$ (see Bernardo and Smith (1994)). When α is estimated from data, Carlin and Louis (1996) call $]0; c]$ the empirical Bayes confidence interval for S_i. The cumulative density in (A.15) takes into account both the posterior position and dispersion of S_i[4]. The mean and variance of the posterior distribution are:

$$\begin{aligned} E(S_i|\mathbf{x}_i) &= \frac{\alpha + x_{i\cdot}}{\alpha + \mu_{i\cdot}} = w_i + (1 - w_i)\frac{x_{i\cdot}}{\mu_{i\cdot}}, \; w_i = \frac{\alpha}{\alpha + \mu_{i\cdot}} \\ V(S_i|\mathbf{x}_i) &= \frac{\alpha + x_{i\cdot}}{(\alpha + \mu_{i\cdot})^2} \end{aligned} \tag{A.16}$$

Because the gamma distribution is conjugate to the Poisson distribution, the mean in the posterior distribution is linear in the observations, $\mathbf{x}$. For known α and β the posterior mean, $E(S_i|\mathbf{x}_i)$, is the Bayes estimate of the dispersion effect, S_i. It is well known that the *mean squared error*[5]:

$$E\left[(\delta - S_i)^2\right]$$

[3] The prior and posterior are in the same paramteric family of distributions.

[4] The posterior uncertainty about α is ignored. However, different methods have been proposed for "correcting" the empirical Bayes confidence interval of this (see Carlin and Louis (1996)).

[5] The mean squared error of estimation is the sum of the squared bias, $S_i - E(\delta)$, and variance, $V(\delta)$, of the estimator δ. In selecting an estimator, there is a trade-off between the two measures of error.

is minimized for $\delta = E(S_i|\mathbf{x}_i)$. The estimate, $E(S_i|\mathbf{x}_i)$, is an unbiased predictor of the random effect, S_i, in the sense that the average value of the estimator is equal to the average value of the quantity being estimated[6]:

$$E\left[E(S_i|\mathbf{x}_i) - S_i\right] = E(S_i) - E(S_i) = 0$$

For normally distributed observations with random effects, s, the linear estimators, which are unbiased in the above sense and minimizing the *mean squared error*, are called the best linear unbiased estimators (BLUP) for s (see Robinson (1991)). For estimated parameters α and β, $\widehat{s}_i = (\widehat{\alpha} + x_{i\cdot}) / (\widehat{\alpha} + \widehat{\mu}_{i\cdot})$ is the parametric empirical Bayes estimator for s_i. Lee and Nelder (1996) call $\widehat{s}_i$ the BLUP in the Poisson-gamma hierarchical generalized linear model[7].

For normally distributed observations, $\mathbf{x} = (\mathbf{x}_1, ..., \mathbf{x}_I)$, with random effects, $\mathbf{s} = (s_1, ..., s_I)$, the empirical Bayes estimator for s_i based on data for all the sites, $\mathbf{x}$, is a better estimator than the maximum likelihood estimator based only on data for site i, $\mathbf{x}_i$, in terms of the mean squared error. This fact has become known as the *Stein effect* or *Stein result* (see e.g. Robert (1994)). Efron and Morris (1973) have shown that this result holds for any distributional assumption. Furthermore, the greatest gain in mean squared error from the Stein result seems to occur precisely in hierarchical Bayes or empirical Bayes settings (see Morris (1983)).

A.3 Modelling variation on road sections

Below, the mathematical framework of Poisson processes on intervals is described. The theory is applied to accidents on road sections in order to develop the road section model of chapter 2. A method for estimating the components of the road section model is given at the end of the chapter.

A.3.1 The mathematical framework

In a one-dimensional space, e.g. a line, one may define a counting process, $N(L)$, which counts the number of point occurrences along the interval $[0; L]$ (see Andersen et al. (1993)). A point process, whether in time or space, may be perceived as a realization of a stochastic process. The homogeneous Poisson process is an example of a simple point process in which points occur randomly. It has the following properties:

[6] It has become a convention that estimators of fixed effects are called *estimators* and are unbiased in the sense that the mean value of the estimator equals the value of the quantity being estimated $E\left(\widehat{\delta}\right) = \delta$, while "estimators" of random effects are called *predictors* and are unbiased in the sense that $E\left(\widehat{\delta}\right) = E(\delta)$ (see Robinson (1991)).

[7] Harville (1991) suggests denoting $\widehat{s}_i$ the empirical BLUP.

1. For a given interval, $[l_1; l_2] : N(l_2) - N(l_1) \in P(\lambda(l_2 - l_1))$.

2. The number of occurrences in non-overlapping intervals are mutually independent.

The intensity, λ, in the homogeneous Poisson process is thus constant over the interval $[l_1; l_2]$. A generalization is the non-homogeneous Poisson process (see Solomon (1987)) in which the intensity varies with position, $\lambda(l)$. The properties of the non-homogeneous Poisson process are:

1. For a given interval, $[l_1; l_2] : N(l_2) - N(l_1) \in P\left(\int_{l_1}^{l_2} \lambda(l)\, dl\right)$.

2. The number of occurrences in non-overlapping intervals are mutually independent.

Below, $\Delta_{[l_1;l_2]}$ denotes the Poisson mean for the interval $[l_1; l_2]$, $\Delta_{[l_1;l_2]} = \int_{l_1}^{l_2} \lambda(l)\, dl$. The variation in intensity for the non-homogeneous Poisson process is illustrated in figure A.2.

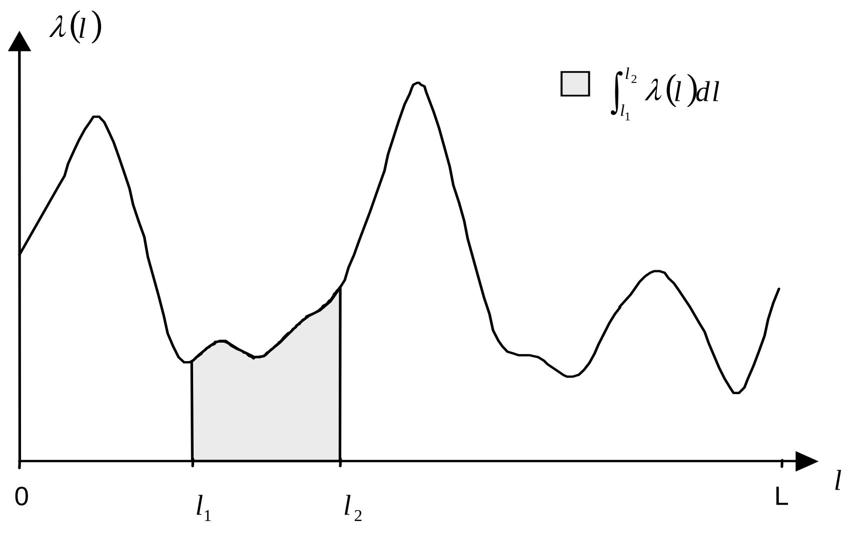


Figure A.2: The variation in Poisson intensity on an interval.

A further extension is the doubly stochastic Poisson process[8] in which $\lambda(\cdot)$ is the realization of an unobserved random intensity function $\Lambda(\cdot)$ (see Cox and Isham

[8] Also known as frailty models in newer literature (see Andersen et al. (1993)).

(1992)). The intensity function $\Lambda(\cdot)$ is denoted the intensity process. Suppose $[0; L[$ is divided into disjoint adjacent sub-intervals $A_1, A_2, ..., A_n$ of the same length L_A:

$$[0; L[= [A_1 \mid A_2 \mid ... \mid A_n]$$

Let $N(A_j)$ denote the number of points in interval A_j, then the following conditional distribution appears:

$$N(A_j) \mid \lambda(\cdot) \in P\left(\Delta_{A_j}\right)$$

where the Poisson mean, Δ_{A_j}, is calculated as:

$$\Delta_{A_j} = \int_{A_j} \lambda(l)\, dl$$

Define $\lambda_j = \Delta_{A_j}/L_A$, then the conditional distribution may be expressed as:

$$N(A_j) \mid \lambda(\cdot) \in P(\lambda_j L_A)$$

where λ_j is the Poisson intensity of interval A_j, and the number of occurrences in non-overlapping sub-intervals is conditionally independent. The Poisson means for the sub-intervals, $[A_1 \mid A_2 \mid ... \mid A_n]$, are illustrated in figure A.3.

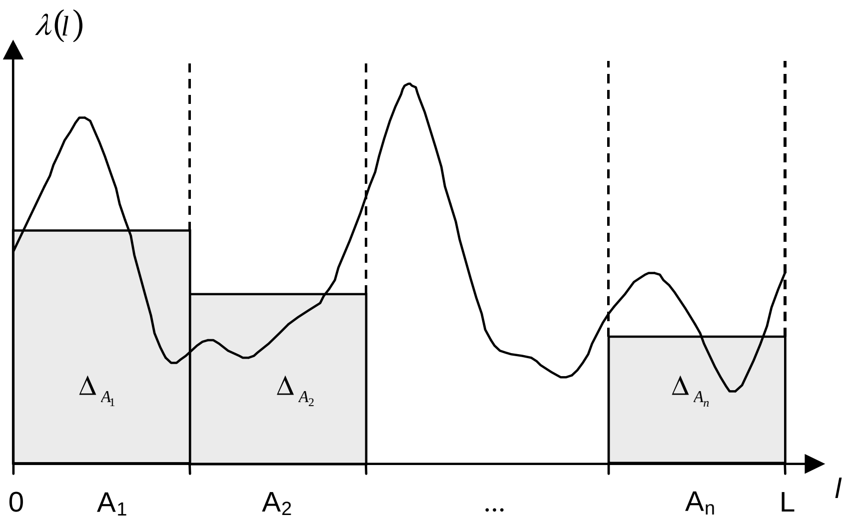


Figure A.3: The Poisson means for the sub-intervals $A_1, A_2, ..., A_n$.

Suppose one has a group of independent intervals of the same length, each corresponding to its own realization, $\lambda(\cdot)$, of the intensity process $\Lambda(\cdot)$. Assume $\Lambda(\cdot)$ to be second-order stationary, then the mean intensity at position l, $E(\Lambda(l))$, is constant, $\overline{\lambda}$, and the mean number of occurrences in a sub-interval of length L_A within the group is calculated as:

$$E(N(l+L_A) - N(l)) = E\left(\int_l^{l+L_A} \Lambda(l)\, dl\right) = \overline{\lambda} L_A$$

where $\overline{\lambda}$ is constant within the group of intervals, i.e. the mean function, $E(\Lambda(\cdot))$, only depends on the length of the interval and not on its position l. Hence, for sub-interval A_j of a given interval within the group of intervals, the corresponding conditional Poisson mean, Δ_{A_j}, is calculated as:

$$\Delta_{A_j} = \lambda_j L_A = s_{A_j} \overline{\lambda} L_A \tag{A.17}$$

where s_A may be interpreted as a dispersion effect indicating how Δ_{A_j} deviates from the group mean $E\left(\Delta_{A_j}\right) = \overline{\lambda} L_A$:

$$s_{A_j} = \frac{\Delta_{A_j}}{E\left(\Delta_{A_j}\right)} = \frac{\lambda_j}{\overline{\lambda}}$$

Because $\Lambda(\cdot)$ is stationary and the sub-intervals are of the same length, L_A, the Poisson intensities of two disjoint sub-intervals in $[A_1 \mid A_2 \mid ... \mid A_n]$ are drawn from the same distribution, hence the corresponding dispersion effects $s_{A_1}, s_{A_2}, ..., s_{A_n}$ are drawn from a common distribution. As intervals within the group are assumed independent, and $\overline{\lambda}$ applies to the whole group, the result may be extended to include all disjoint sub-intervals of length L_A within the group of intervals.

Assume the intensity process $\Lambda(\cdot)$ is such that the autocovariance between intensities of two endpoints in a sub-interval is vanishing, then the intensity of two non-overlapping sub-intervals may be considered to be independent, and hence also their dispersion effects. Thus $s_{A_1}, s_{A_2}, ..., s_{A_n}$ for sub-intervals $A_1, A_2, ..., A_n$ are values of independent and identically distributed random variables with mean 1[9].

Consider the case where $A_1, A_2, ..., A_n$ are basis sub-intervals of length $L_A = 1$, i.e. $A_j = [j-1; j[$ and $n = L$. Then $\Delta_{A_j} = \lambda_j = s_{A_j}\overline{\lambda}$ and the dispersion effects of two non-overlapping sub-intervals are still assumed to be independent. Let the dispersion effect, s_{A_j}, of basis sub-interval, A_j, be modelled by a random variable, S_{A_j}, with its own distribution described by a gamma density function with shape parameter α and mean 1:

$$S_{A_j} \in gamma\left(\alpha, \frac{1}{\alpha}\right), \forall j = 1, ..., L$$

[9] $\forall j : E\left(s_{A_j}\right) = E\left(\lambda_j/\overline{\lambda}\right) = E\left(\Delta_{A_j}/L_A\right)/\overline{\lambda} = E\left(\Delta_{A_j}\right)/\left(\overline{\lambda}L_A\right) = E\left(\Delta_{A_j}\right)/E\left(\Delta_{A_j}\right) = 1$

The dispersion effect, $s_{[0;L[}$, of the whole interval $[0;L[$ may be calculated as the average of the dispersion effects of the basis sub-intervals:

$$\begin{aligned} s_{[0;L[} &= \frac{\Delta_{[0;L[}}{E\left(\Delta_{[0;L[}\right)} = \frac{\int_0^L \lambda(l)\,dl}{\overline{\lambda} L} \\ &\quad \frac{1}{L}\left(\frac{\int_{A_1}\lambda(l)\,dl}{\overline{\lambda}} + \frac{\int_{A_2}\lambda(l)\,dl}{\overline{\lambda}} + \ldots + \frac{\int_{A_L}\lambda(l)\,dl}{\overline{\lambda}}\right) \\ &= \frac{1}{L}\left(\frac{\lambda_1}{\overline{\lambda}} + \frac{\lambda_2}{\overline{\lambda}} + \ldots + \frac{\lambda_L}{\overline{\lambda}}\right) \\ &= \frac{1}{L}\left(s_{A_1} + s_{A_2} + \ldots + s_{A_L}\right) \end{aligned}$$

Hence, the distribution of the dispersion effect, $s_{[0;L[}$, is described by a gamma density function with shape parameter αL and mean 1:

$$S_{[0;L[} \in gamma\left(\alpha L, \frac{1}{\alpha L}\right)$$

Consider two overlapping sub-intervals, $[l_1;l_2[$ and $[l_1';l_2'[$, in $[0;L[$ as illustrated in figure A.4.In analogy with the above, the dispersion effects of the two sub-intervals,

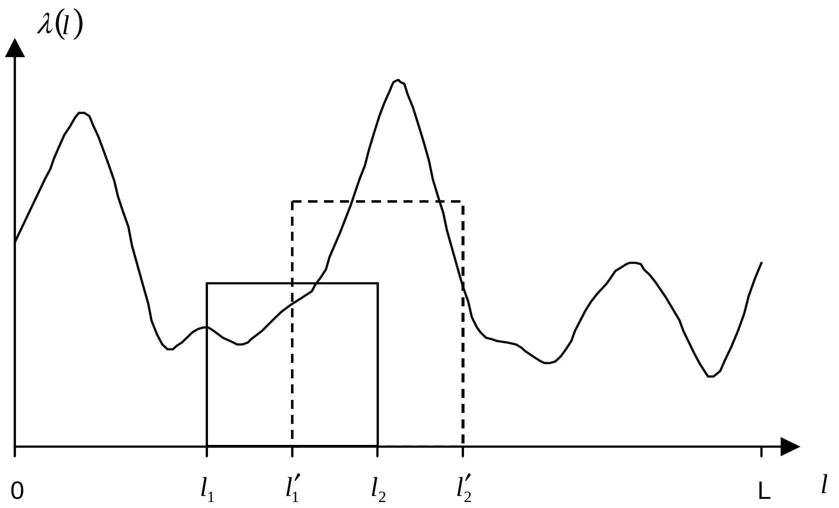


Figure A.4: Two overlapping sub-intervals $[l_1;l_2[$ and $[l_1';l_2'[$ in $[0;L[$.

$s_{[l_1;l_2[}$ and $s_{[l_1';l_2'[}$, may be calculated as the average of the dispersion effects of the corresponding basis sub-intervals:

$$
\begin{aligned}
s_{[l_1;l_2[} &= \frac{1}{(l_2-l_1)}\sum_{i=l_1}^{l_2} s_{A_i} = \frac{1}{(l_2-l_1)}\sum_{i=l_1}^{l_1'-1} s_{A_i} + \frac{1}{(l_2-l_1)}\sum_{i=l_1'}^{l_2} s_{A_i} \\
s_{[l_1';l_2'[} &= \frac{1}{(l_2'-l_1')}\sum_{j=l_1'}^{l_2'} s_{A_j} = \frac{1}{(l_2'-l_1')}\sum_{j=l_1'}^{l_2} s_{A_j} + \frac{1}{(l_2'-l_1')}\sum_{j=l_2+1}^{l_2'} s_{A_j}
\end{aligned}
$$

and the distribution of the dispersion effects is described by gamma density functions:

$$
\begin{aligned}
S_{[l_1;l_2[} &\in gamma\left(\alpha\left(l_2-l_1\right), \frac{1}{\alpha\left(l_2-l_1\right)}\right) \\
S_{[l_1';l_2'[} &\in gamma\left(\alpha\left(l_2'-l_1'\right), \frac{1}{\alpha\left(l_2'-l_1'\right)}\right)
\end{aligned}
$$

However, because of the overlap between the two sub-intervals, the dispersion effects are not independent. The covariance of $S_{[l_1;l_2[}$ and $S_{[l_1';l_2'[}$ may be calculated as:

$$
\begin{aligned}
\operatorname{cov}\left(S_{[l_1;l_2[}, S_{[l_1';l_2'[}\right) &= E\left(\left(\sum_{i=l_1}^{l_2}\left(S_{A_i}-1\right)\right)\left(\sum_{j=l_1'}^{l_2'}\left(S_{A_j}-1\right)\right)\right) \\
&= E\left(\sum_{i=l_1}^{l_2}\sum_{j=l_1'}^{l_2'}\left(S_{A_i}-1\right)\left(S_{A_j}-1\right)\right) \\
&= E\left(\sum_{i=l_1'}^{l_2}\left(S_{A_i}-1\right)^2\right) = \frac{1}{\alpha}\left(l_2-l_1'\right)
\end{aligned}
$$

Below, the mathematical framework derived above is applied to accidents on road sections.

A.3.2 Accidents on road sections

Accidents on a road section may be interpreted as a realization of a stochastic Poisson point process. Consider a road section i of length L_i, i.e. the interval $[0;L_i[$. On this section a counting process, $x_{it}(l)$, which counts the number of accidents in year t for the sub-section $[0;l[$ is defined.

Accidents tend to accumulate in certain sections of the road indicating that the intensity varies over the road section. Let $\lambda_{it}(l)$ denote the Poisson intensity at position l, and let $\Delta_{[0;L_i[}$ denote the corresponding Poisson mean for the interval

$[0; L_i[$. Furthermore, as the intensity is unknown one may model the variation in accident counts on road sections by a frailty model with intensity, $\lambda_{it}(\cdot)$. The intensity is a realization of an unobserved random intensity function $\Lambda_{it}(\cdot)$. Let x_{it} denote the total number of accidents on road section i in year t, $x_{it} \equiv x_{it}(L)$, hence, the following conditional distribution appears:

$$x_{it}|\lambda_{it}(\cdot) \in P\left(\Delta_{[0;L_i[}\right)$$

where the Poisson mean is calculated as:

$$\Delta_{[0;L_i[} = \int_0^{L_i} \lambda_{it}(l)\, dl$$

Define $\lambda_{it} = \Delta_{[0;L_i[}/L_i$, then the conditional distribution may be expressed as:

$$x_{it} \mid \lambda_{it}(\cdot) \in P(\lambda_{it} L_i)$$

where λ_{it} is the expected number of accidents per kilometer at site i in year t The intensity, λ_{it}, is denoted the *site safety* per kilometer at site i in year t.

Suppose one has a group of independent road sections with the same traits, each corresponding to its own realization, $\lambda_{it}(\cdot)$, of the intensity process $\Lambda_{it}(\cdot)$. Assume $\Lambda_{it}(\cdot)$ to be second-order stationary, then the mean intensity at position l, $E(\Lambda_{it}(l))$, is constant, μ_{it}, and the mean number of accidents on a road section with the same traits as section i in year t is:

$$E\left(\Delta_{[0;L_i[}\right) = \mu_{it} L_i$$

The mean intensity, μ_{it}, is the expected number of accidents per kilometer at sites with the same traits as i in year t. Hence, the intensity, μ_{it}, is denoted the *reference safety level* per kilometer at site i in year t. The assumption, that the mean intensity, μ_{it}, is constant over an interval, is in line with the assumption that traits are constant over the entire length of the road section. The Poisson mean for road section i may be expressed as:

$$\Delta_{[0;L_i[} = \lambda_{it} L_i = s_i \mu_{it} L_i$$

where s_i may be interpreted as a dispersion effect indicating how site safety on road section i in year t deviates from the corresponding reference safety:

$$s_i = \frac{\Delta_{[0;L_i[}}{\mu_{it} L_i} = \frac{1}{L_i}\left(\frac{\Delta_{[0;1[}}{\mu_{it}} + \frac{\Delta_{[1;2[}}{\mu_{it}} + ... + \frac{\Delta_{[L_i-1;L_i[}}{\mu_{it}}\right)$$

The dispersion effect for road section i is given as the mean of the L_i number of the corresponding unit length[10] dispersion effects. The dispersion effect, s_i, may be

[10] In the road section model, a unit length corresponds to one kilometer.

modelled by a random variable, S_i, with its own distribution described by a gamma density function with shape parameter αL_i and mean 1:

$$S_i \in gamma\left(\alpha L_i, \frac{1}{\alpha L_i}\right)$$

A.3.3 Estimation of fixed and random effects

Let $\mathbf{x} = \{x_{it}\}_{i=1,...,I,t=1,...,T_i}$ be a set of conditionally independent observations such that the conditional distribution of x_{it} given a dispersion effect, s_i, is the Poisson distribution with mean $\lambda_{it} L_i = \mu_{it} L_i s_i$. Assume that the dispersion effects, $\mathbf{s} = \{s_i\}_{i=1,...,I}$, are independent and identically gamma distributed with shape parameter αL_i and mean 1:

$$\begin{aligned} p\left(x_{it}|s_i\right) &= p\left(X_{it} = x_{it}|s_i\right) = \frac{\left(\lambda_{it} L_i\right)^{x_{it}}}{x_{it}!} e^{-\lambda_{it} L_i},\ \lambda_{it} \in \mathbb{R}_+, x_{it} \in \mathbb{N}_0 \quad \text{(A.18)} \\ f\left(s_i\right) &= f\left(S_i = s_i\right) = \frac{\left(\alpha L_i\right)^{\alpha L_i}}{\Gamma\left(\alpha L_i\right)} s_i^{\alpha L_i - 1} e^{-\alpha L_i s_i},\ \alpha \in \mathbb{R}_+, s_i \in \mathbb{R}_+ \end{aligned}$$

Analogue to the intersection model (A.3), the conditional mean per kilometer, λ_{it}, of x_{it} given s_i is modelled by a linear model:

$$\eta'_{it} = g\left(\lambda_{it}\right) = g\left(\mu_{it} s_i\right) = \eta_{it} + v_i \quad \text{(A.19)}$$

again, $g(\cdot) = \ln(\cdot)$ is the canonical link function for the conditional mean per kilometer and the fixed effect part of the mean is:

$$\eta_{it} = g\left(\mu_{it}\right) = \mathbf{Z}_{it}\boldsymbol{\beta} = \sum_{j=1}^{J} z_{itj}\beta_j \Leftrightarrow \mu_{it} = g^{-1}\left(\mathbf{Z}_{it}\boldsymbol{\beta}\right) \quad \text{(A.20)}$$

The model (A.18)-(A.20) is called a Poisson-gamma hierarchical generalized linear model. The marginal distribution of all accident observations $\mathbf{x} = (\mathbf{x}_1, ..., \mathbf{x}_I)$, $\mathbf{x}_i = (x_{i1}, ..., x_{iT_i})$, is:

$$\begin{aligned} P\left(\mathbf{x}\right) &= \prod_{i=1}^{I} P\left(\mathbf{X}_i = \mathbf{x}_i\right) = \prod_{i=1}^{I} \int_0^\infty p\left(\mathbf{X}_i = \mathbf{x}_i|s_i\right) f\left(S_i = s_i\right) \partial s_i \\ &= \prod_{i=1}^{I} \int_0^\infty \left[\prod_{t=1}^{T_i} \frac{\left(\mu_{it} L_i s_i\right)^{x_{it}}}{x_{it}!} e^{-\mu_{it} L_i s_i}\right] \frac{\left(\alpha L_i\right)^{\alpha L_i}}{\Gamma\left(\alpha L_i\right)} s_i^{\alpha L_i - 1} e^{-\alpha L_i s_i} \partial s_i \\ &= \prod_{i=1}^{I} \left[\left(\prod_{t=1}^{T_i} \frac{\left(\mu_{it} L_i\right)^{x_{it}}}{x_{it}!}\right) \frac{\left(\alpha L_i\right)^{\alpha L_i}}{\Gamma\left(\alpha L_i\right)} \int_0^\infty s_i^{\alpha L_i + x_{i\cdot} - 1} e^{-\left(\alpha L_i + \mu_{i\cdot} L_i\right) s_i} \partial s_i\right] \\ &= \left[\prod_{i=1}^{I} \prod_{t=1}^{T_i} \frac{\left(\mu_{it} L_i\right)^{x_{it}}}{x_{it}!}\right] \prod_{i=1}^{I} \left[\frac{\left(\alpha L_i\right)^{\alpha L_i}}{\Gamma\left(\alpha L_i\right)} \frac{\Gamma\left(\alpha L_i + x_{i\cdot}\right)}{\left(\alpha L_i + \mu_{i\cdot} L_i\right)^{\alpha L_i + x_{i\cdot}}}\right] \end{aligned}$$

where $x_{i\cdot} = \sum_{t=1}^{T_i} x_{it}$ and $\mu_{i\cdot} = \sum_{t=1}^{T_i} \mu_{it}$ is the total reported number of accidents at site i and corresponding reference safety respectively. The marginal likelihood function of the observed response variables $\mathbf{x}$ is:

$$\begin{aligned} l(\boldsymbol{\beta}, \alpha; \mathbf{x}) &= \ln P(\mathbf{x}) \propto \ln \left[\prod_{i=1}^{I} \prod_{t=1}^{T_i} (\mu_{it} L_i)^{x_{it}} \right] \prod_{i=1}^{I} \left[\frac{1}{(\alpha L_i + \mu_{i\cdot} L_i)^{\alpha L_i + x_{i\cdot}}} \right] \\ &\propto \sum_{i=1}^{I} \sum_{t=1}^{T_i} x_{it} \ln \mu_{it} - \sum_{i=1}^{I} (\alpha L_i + x_{i\cdot}) \ln (\alpha L_i + \mu_{i\cdot} L_i) \end{aligned}$$

Hence, the marginal score function for the determination of the fixed effect parameter β_j becomes:

$$\begin{aligned} \frac{\partial l}{\partial \beta_j} &= \frac{\partial}{\partial \beta_j} \left[\sum_{i=1}^{I} \sum_{t=1}^{T_i} x_{it} \ln \mu_{it} - (\alpha L_i + x_{i\cdot}) \ln (\alpha L_i + \mu_{i\cdot} L_i) \right] \qquad \text{(A.21)} \\ &= \sum_{i=1}^{I} \sum_{t=1}^{T_i} x_{it} \frac{1}{\mu_{it}} \frac{\partial \mu_{it}}{\partial \beta_j} - \sum_{i=1}^{I} (\alpha L_i + x_{i\cdot}) \frac{1}{\alpha L_i + \mu_{i\cdot} L_i} \frac{\partial (\alpha L_i + \mu_{i\cdot} L_i)}{\partial \beta_j} \\ &= \sum_{i=1}^{I} \sum_{t=1}^{T_i} x_{it} z_{itj} - \sum_{i=1}^{I} \frac{\alpha L_i + x_{i\cdot}}{\alpha L_i + \mu_{i\cdot} L_i} \left(L_i \sum_{t=1}^{T_i} \mu_{it} z_{itj} \right) \\ &= \sum_{i=1}^{I} \sum_{t=1}^{T_i} \left(x_{it} - \frac{\alpha L_i + x_{i\cdot}}{\alpha L_i + \mu_{i\cdot} L_i} \mu_{it} L_i \right) z_{itj} \end{aligned}$$

For given α, the h-likelihood for road sections is given as the joint distribution of the observed $\mathbf{x}$ and unobserved variables $\mathbf{s}$:

$$\begin{aligned} h(\boldsymbol{\beta}, \alpha; \mathbf{x}, \mathbf{s}) &\equiv l_0(\boldsymbol{\beta}; \mathbf{x}|\mathbf{s}) + l_1(\alpha; \mathbf{s}) = \ln [p(\mathbf{x}|\mathbf{s}, \boldsymbol{\beta}) f(\mathbf{s}|\alpha)] \qquad \text{(A.22)} \\ &= \sum_{i=1}^{I} \sum_{t=1}^{T_i} \ln p(x_{it}|s_i) + \sum_{i=1}^{I} \ln f(s_i|\alpha) \\ &= \sum_{i=1}^{I} \sum_{t=1}^{T_i} \ln \left(\frac{(\mu_{it} L_i s_i)^{x_{it}}}{x_{it}!} e^{-\mu_{it} L_i s_i} \right) + \sum_{i=1}^{I} \ln \left(\frac{(\alpha L_i)^{\alpha L_i}}{\Gamma(\alpha L_i)} s_i^{\alpha L_i - 1} e^{-\alpha L_i s_i} \right) \\ &\propto \sum_{i=1}^{I} \sum_{t=1}^{T_i} (x_{it} \ln (\mu_{it} s_i) - \mu_{it} L_i s_i) + \sum_{i=1}^{I} (\alpha L_i \ln s_i - \alpha L_i s_i) \\ &= \sum_{i=1}^{I} \sum_{t=1}^{T_i} (x_{it} \ln \mu_{it}) + \sum_{i=1}^{I} ((\alpha L_i + x_{i\cdot}) \ln s_i - (\alpha L_i + \mu_{i\cdot} L_i) s_i) \end{aligned}$$

The last line in the expression is the kernel of the h-likelihood function for $\boldsymbol{\beta}$ and $\mathbf{s}$, where $\boldsymbol{\beta}$ enters via (A.20). The estimates of $\boldsymbol{\beta}$ and $\mathbf{s}$ are the maximum h-likelihood estimates (MHLEs), and are obtained by equating the components of the score function,

$D_{(\boldsymbol{\beta},\mathbf{s})}h\left(\boldsymbol{\beta},\alpha;\mathbf{x},\mathbf{s}\right)=\left(\partial h/\partial\boldsymbol{\beta},\partial h/\partial\mathbf{s}\right)$, to zero. Assuming the fixed effect parameters, $\boldsymbol{\beta}$, are given, i.e. the reference safety levels, $\boldsymbol{\mu}$, are known, the h-score function for the determination of the dispersion effect s_i becomes:

$$\begin{aligned}\frac{\partial h}{\partial s_i} &= \frac{\partial}{\partial s_i}\left[\sum_{i=1}^{I}\left(\left(\alpha L_i+x_{i\cdot}\right)\ln s_i-\left(\alpha L_i+\mu_{i\cdot}L_i\right)s_i\right)\right]\\ &= \left(\alpha L_i+x_{i\cdot}\right)\frac{1}{s_i}-\left(\alpha L_i+\mu_{i\cdot}L_i\right)\end{aligned}$$

Equating $\partial h/\partial s_i$ to zero gives the following maximum h-likelihood estimate for s_i:

$$\partial h/\partial s_i=\left(\alpha L_i+x_{i\cdot}\right)\frac{1}{s_i}-\left(\alpha L_i+\mu_{i\cdot}L_i\right)=0\Leftrightarrow s_i=\frac{\alpha L_i+x_{i\cdot}}{\alpha L_i+\mu_{i\cdot}L_i} \tag{A.23}$$

Again, the MHLE for s_i may be interpreted as the empirical Bayes posterior estimate of s_i when α is estimated from data. The corresponding h-score function for the determination of the fixed effect parameter β_j becomes:

$$\frac{\partial h}{\partial \beta_j}\propto\frac{\partial}{\partial \beta_j}\left[\sum_{i=1}^{I}\sum_{t=1}^{T_i}\left(x_{it}\ln\mu_{it}\right)+\sum_{i=1}^{I}\left(\left(\alpha L_i+x_{i\cdot}\right)\ln s_i-\left(\alpha L_i+\mu_{i\cdot}L_i\right)s_i\right)\right] \tag{A.24}$$

which is equal to the marginal score function for β_j in (A.21), thus the MHLE for $\boldsymbol{\beta}$ are equal to the MLE. Using (A.21) along with the maximum h-likelihood estimate for s_i in (A.23), the h-likelihood equation for the fixed effect parameter β_j becomes:

$$\frac{\partial h}{\partial \beta_j}=\sum_{i=1}^{I}\sum_{t=1}^{T_i}\left(x_{it}-\widehat{s}_i\mu_{it}L_i\right)z_{itj}=0 \tag{A.25}$$

The MHLEs for s_i and β_j are interdependent, because $\mu_{i\cdot}=\sum_{t=1}^{T_i}g^{-1}\left(\sum_{j=1}^{J}\beta_j z_{itj}\right)$ in (A.23). The technique used to estimate fixed and random effects in the model for road sections is similar to the procedure in Lee and Nelder (2001) used for intersections. However, in the estimation of the dispersion parameter, α, the second generalized linear model is slightly modified. The model in (A.18) may be seen as weighted model with dispersion parameter α and known weights $\mathbf{L}=\left(L_1,...L_I\right)$. Hence, in the estimation of α the weighted deviances d_iL_i are used, i.e. one now has responses $\mathbf{d}^*\mathbf{L}$, prior weights $\mathbf{1}-\mathbf{q}^a$ and scale parameter 2:

$$\boldsymbol{\eta}=g\left(\boldsymbol{\delta}\right)=\log\left(\boldsymbol{\delta}\right) \tag{A.26}$$

$$\mathbf{d}^*\mathbf{L}\sim gamma\left(\frac{1}{2},2\boldsymbol{\delta}\right)$$

with $\mathbf{L} = (L_1, ..., L_I)$ and L_i as the length of road section i. Again $\boldsymbol{\delta} = \mathbf{1}/\boldsymbol{\alpha}$ and the linear predictor is $\boldsymbol{\eta}$. The elements $\mathbf{d}^*$ and $\mathbf{1} - \mathbf{q}^a$ are calculated as above using $\delta_i = \delta / L_i$. Hence, the response for site i is calculated as:

$$d_i^* L_i = \frac{d_i^a L_i}{1 - q_i^a} \tag{A.27}$$

In expression (A.27), the adjusted deviance at site i, d_i^a, is calculated as:

$$d_i^a = 2\left(s_i - \log\left(s_i\right) - 1\right) + I\left(x_{i\cdot} = 0\right)\epsilon_i \frac{\delta}{L_i}$$

with adjustment component, ϵ_i:

$$\epsilon_i = \frac{\lambda_{i\cdot} L_i}{\lambda_{i\cdot} L_i + 1}$$

Again, $d_i \simeq V\left(s_i\right) = 1/\alpha L_i \equiv \delta_i$ from the Taylor approximation, thus $d_i L_i$ is chosen as a suitable statistic for measuring $\delta = \delta_i L_i$ at road section i. Analogue to intersections, the variation of the deviances differs from site to site:

$$V\left(d_i\right) \simeq \delta_i^2 \left(1 - p_i\right)^2 = \frac{\delta^2}{L_i^2}\left(1 - p_i\right)^2 \Rightarrow V\left(d_i L_i\right) \simeq \delta^2 \left(1 - p_i\right)^2$$

with p_i as the i^{th} diagonal element in the hat matrix for estimating $\mathbf{s}$. Because δ applies to the whole site-group, it is sufficient to scale the responses $d_i L_i$ by $1 - p_i$ in order to ensure equal variances. Again, the weight $1 - p_i$ is further generalized to $1 - q_i$, and the adjusted weights may be expressed as:

$$1 - q_i^a = \frac{\delta_i x_{i\cdot}}{\delta_i x_{i\cdot} + 1} + I\left(x_{i\cdot} = 0\right)\epsilon_i = \frac{\delta x_{i\cdot}}{\delta x_{i\cdot} + L_i} + I\left(x_{i\cdot} = 0\right)\epsilon_i$$

Analogue to intersections, the solution to (A.26) may be calculated as:

$$\widehat{\delta} = \frac{\sum_{i=1}^{I} d_i^a L_i}{\sum_{i=1}^{I}\left(1 - q_i^a\right)}$$

Appendix B

The regression to the mean effect

The regression to the mean (RTM) effect is known in various areas. More than a century ago, Galton showed that the height of the offspring of short or tall parents on average was closer to the mean height of the offspring generation, than the parents had been to the mean height of their generation. Schall and Smith (2000) showed baseball players with a high batting average one season on average had a lower batting average the next season. In short, the RTM effect is the observation that what happened *after* on average is not what happened *before*.

The purpose of this chapter is to investigate the concept of the regression to the mean effect used by e.g. Abbess et al. (1981) and Kulmala (1995), and to estimate the RTM effect expressed in the hierarchical generalized linear model proposed in chapter 2. The latter is the regression to the mean effect defined in chapter 4.

B.1 An intuitive definition of a RTM

The definition of the regression to the mean effect used by e.g. Abbess et al. (1981) and Kulmala (1995) is an intuitive measure of the proportion of accident counts attributable to random variation, i.e. a residual for the individual observation. In a given year, let λ denote the expected number of accidents at a site (the site safety level), and let x denote the corresponding reported accident counts. The regression to the mean effect is in this terminology defined as:

$$RTM^* = \frac{x - \lambda}{x} \tag{B.1}$$

The RTM effect in (B.1) is hence dependent on the individual accident count and may even assume negative values. Figure B.1 illustrates the change in regression to the mean effect for different values of x and λ, when the accident count exceeds the site safety level. The shade of the contour area indicates the value of RTM^*, i.e. the

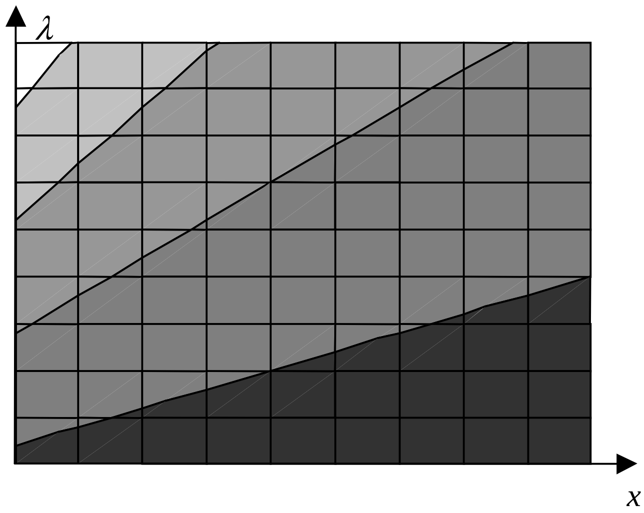


Figure B.1: Regression to the mean (RTM) effect for different values of x and λ. The RTM effect is increasing in the shade of the contour area.

darker the colour the larger is the regression to the mean effect.Figure B.1 indicates that the regression to the mean effect is not a constant number within a group of sites. Even sites with the same level of site safety, λ, may have different RTM effects. However, because the regression to the mean effect in (B.1) is a measure of random variation in accident counts, the average RTM in a group of sites with the same site safety level is close to zero.

B.2 RTM in the HGLM model

In the literature since Galton, the regression to the mean effect has been used as a statistical (average) measure of a regression relation between two variables. This RTM effect is expressed in the Poisson-gamma hierarchical linear models.

Let x_t and x_{t+1} denote the reported accident counts at an intersection in year t and $t+1$ respectively. The reported accident counts are realizations of random variables X_t and X_{t+1} with means μ_t and μ_{t+1} respectively. Assume for simplicity, that the mean and variance are unchanged, i.e. that $\mu_t = \mu_{t+1} = \mu$ and $V(X_t) = V(X_{t+1})$. Then the relative magnitude of the regression to the mean effect is defined as:

$$RTM \equiv \frac{B}{A} = \frac{x_t - E(X_{t+1}|x_t)}{x_t - \mu}$$

The quantities A and B correspond to A and B in figure B.2.Because accident counts

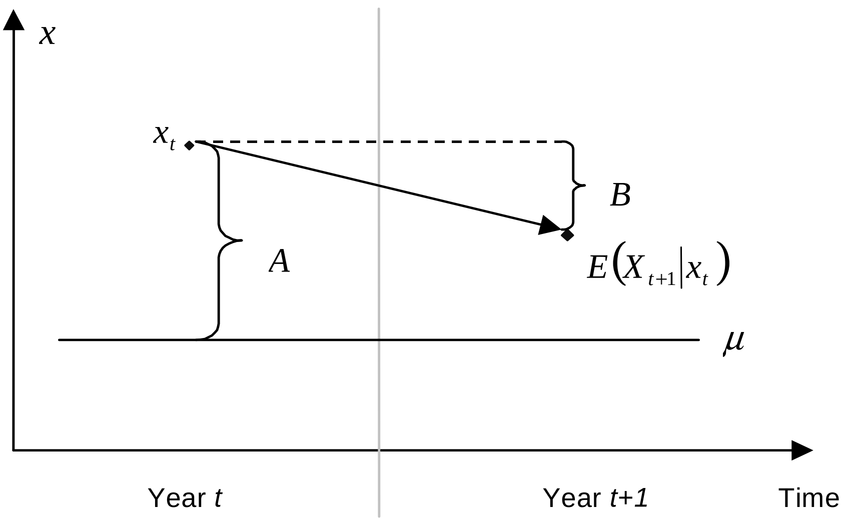


Figure B.2: Regression to the mean effect with constant mean μ.

X_t and X_{t+1} are assumed to have equal variances, the regression of X_{t+1} upon X_t is (see Davis (1986)):

$$E\left(X_{t+1}|x_t\right) - \mu = \rho\left(x_t - \mu\right)$$

with ρ denoting the correlation between X_t and X_{t+1}. The regression to the mean effect is thus a function of the correlation between accident counts:

$$RTM = 1 - \rho$$

Hence, the RTM effect does not depend on the accident counts x_t. If no correlation is present, $\rho = 0$, then one has:

$$E\left(X_{t+1}|x_t\right) - \mu = 0 \Leftrightarrow E\left(X_{t+1}|x_t\right) = \mu$$

indicating that accident counts in year $t+1$ regress all the way back to the mean, i.e. $RTM = 100\%$. If on the other hand, accident counts were perfectly correlated, $\rho = 1$, one has:

$$E\left(X_{t+1}|x_t\right) - \mu = x_t - \mu \Leftrightarrow E\left(X_{t+1}|x_t\right) = x_t$$

indicating no regression to the mean, i.e. $RTM = 0\%$. In practice, however, imperfect correlation exists, $\rho \in \left]0;1\right[$, which means that:

$$\left|E\left(X_{t+1}|x_t\right) - \mu\right| < \left|x_t - \mu\right|$$

In other words, accident counts in year $t+1$ are on average closer to the mean than their counterpart accident counts in year t.

Assuming constant traits as above, the accident counts X_t and X_{t+1} are, under the model in chapter 2, both conditionally Poisson distributed with mean $\lambda = \mu s$:

$$\begin{aligned} X_t|s &\in Poiss(\mu s) \\ X_{t+1}|s &\in Poiss(\mu s) \end{aligned}$$

Because the dispersion effect, S, is gamma distributed with mean 1:

$$S \in gamma\left(\alpha, \frac{1}{\alpha}\right)$$

the marginal distribution of X_t and X_{t+1} respectively, is the negative binomial distribution with parameters $(\alpha, \alpha/(\alpha+\mu))$. The mean and variance in the marginal distribution are:

$$\begin{aligned} E(X_t) &= E(X_{t+1}) = \mu \\ V(X_t) &= V(X_{t+1}) = \mu + \frac{\mu^2}{\alpha} \end{aligned}$$

The correlation between two random variables X_t and X_{t+1} is in general defined as:

$$\rho \equiv \frac{\mathrm{cov}(X_t, X_{t+1})}{\sqrt{V(X_t)}\sqrt{V(X_{t+1})}}$$

Using the fact that X_t and X_{t+1} are conditionally independent with equal variances, the covariance of accident counts X_t and X_{t+1} is found as:

$$\begin{aligned} \mathrm{cov}(X_t, X_{t+1}) &= E[\mathrm{cov}(X_t|s, X_{t+1}|s)] + \mathrm{cov}[E(X_t|s), E(X_{t+1}|s)] \\ &= 0 + \mathrm{cov}(\mu s, \mu s) = V(\mu s) = \mu^2 V(s) = \frac{\mu^2}{\alpha} \end{aligned}$$

Assuming constant traits, then under the models in chapter 2, the correlation of X_t and X_{t+1} is thus calculated as:

$$\rho = \frac{\mu^2}{\alpha} \Big/ \left(\mu + \frac{\mu^2}{\alpha}\right) = \frac{\mu}{\alpha+\mu}$$

The similar result is derived for road sections. The regression to the mean effect is:

$$RTM = 1 - \rho = \frac{\alpha}{\alpha+\mu} \tag{B.2}$$

Because RTM in (B.2) is dependent on the traits at the site through μ, the regression to the mean effect is site-specific. However, it is not dependent on the individual accident counts, and sites with similar traits and from the same site-group will have the same RTM effect. The weight in the predicted site safety levels in (2.12) and (2.14) of chapter 2 is the estimated regression to the mean effect as defined in (B.2). Consequently, the regression to the mean effect is removed as accurately as possible from the predictions in (2.12) and (2.14). The definition of RTM in (B.2) is in line with Hauer (1997).

Appendix C

The framework of state of the art

The objective of this appendix is to give a detailed mathematical presentation of the models and methods used by the national Road Directorate and regional authorities in hot spot safety work. The models are derived in a mathematical context similar to that of the Poisson-gamma hierarchical generalized linear model developed in chapter 2.

C.1 Modelling variation

Assume sites are divided into groups of similar type, the so called *ap-groups*. Let H be such a group of sites $\{i\}_{i=1,\ldots,I}$, and let x_{it} denote the reported number of accidents at site $i \in H$ in a given year t in the observation period $]0;T_i]$. The distribution of x_{it} on intersections and road sections of length L_i is the Poisson distribution with mean μ_i respectively $\mu_i L_i$:

$$p\left(x_{it}|\mu_i\right) = p\left(X_{it} = x_{it}|\mu_i\right) = \begin{cases} \frac{(\mu_i)^{x_{it}}}{x_{it}!} e^{-\mu_i}, \ \mu_i \in \mathbb{R}_+, \ x_{it} \in \mathbb{N}_0 \\ \frac{(\mu_i L_i)^{x_{it}}}{x_{it}!} e^{-\mu_i L_i}, \ \mu_i \in \mathbb{R}_+, \ x_{it} \in \mathbb{N}_0 \end{cases} \tag{C.1}$$

The distribution of x_{it} describes the variation around the mean μ_i $(\mu_i L_i)$. The Poisson mean is described by a log-linear model of the traits, z_{ij}, and fixed effect parameters $\beta_j, j = 1, .., J$:

$$\eta_i = g\left(\mu_i\right) = \mathbf{Z}_i \boldsymbol{\beta} = \sum_{j=1}^{J} z_{ij}\beta_j \Leftrightarrow \mu_i = g^{-1}\left(\mathbf{Z}_i \boldsymbol{\beta}\right) \tag{C.2}$$

where η_i is called the linear predictor for μ_i with link function $g\left(\cdot\right)$, and $g\left(\cdot\right)$ is the logarithm. The traits at site i, $\mathbf{Z}_i$, are modelled constant over the observation period with $z_{i1} = 1, \forall i$. The fixed effect parameters, $\boldsymbol{\beta}$, are associated with all data in the ap-group. The models described by (C.1) and (C.2) is called a generalized linear model

(GLIM) with a Poisson error distribution (see appendix A for a general description of GLIM).

C.1.1 Road sections

The site safety level on a road section is defined as the expected number of accidents per year and kilometer. The only trait in the model is the traffic flow (see Vejdirektoratet (2001b)):

$$\mu = exp\left(\beta_1 + z_2\beta_2\right) = a \cdot AADT^b \tag{C.3}$$

where $AADT$ is the average annual daily traffic (number of vehicles), and a and b are parameters. Models are estimated for the total number of accidents, injury accidents and fatal accidents. Estimated model parameters may be found in Vejdirektoratet (2001b).

C.1.2 Intersections

The site safety level at an intersection is defined as the expected number of accidents per year. The only traits in the model are the traffic flows on the major and minor arms respectively (see Vejdirektoratet (2001b)):

$$\mu = exp\left(\beta_1 + z_2\beta_2 + z_3\beta_3\right) = a \cdot AADT_{ma}^{b_1} \cdot AADT_{mi}^{b_2} \tag{C.4}$$

where $AADT_{ma}$ and $AADT_{mi}$ are the average annual daily incoming traffic on major and minor roads respectively, and a, b_1 and b_2 are parameters. At signal controlled intersections, the major arms are the two roads with the largest annual daily traffic, and in intersections with no signal control they are the two roads with the right of way. Models are estimated for the total number of accidents, injury accidents and fatal accidents. Estimated model parameters may be found in Vejdirektoratet (2001b).

C.1.3 Roundabouts

The site safety level in a roundabout is defined as the expected number of accidents per year. The structure of the Poisson mean is similar to (C.3), i.e. the only trait in the model is the traffic flow (see Aagaard (1995)):

$$\mu = exp\left(\beta_1 + z_2\beta_2\right) = a \cdot AADT^b \tag{C.5}$$

The daily annual traffic is calculated as the total of all arms divided by four. The parameters in (C.5) are not estimated on a regular basis by the Road Directorate.

C.2 Targeting hot spots

Assuming the models above, the total number of accidents, $x_{i\cdot}$, at a site i in the observation period $]0;T_i]$ is Poisson distributed with mean $\lambda_i T_i = \mu_i T_i$ ($\mu_i T_i L_i$ for road sections). Accident hot spots are identified by determining whether the so-called *p-value*, $p\left(X \geq x_{i\cdot} | \mu_i T_i\right)$, in a test of the hypothesis:

$$H_0 : \lambda_i = \mu_i \text{ vs } \lambda_i > \mu_i$$

is below a given value (the level of significance). Let the level of significance be α and let $x_{\min}$ denote the minimum threshold value, then under the models defined by (C.1) and (C.2), definition 6 of chapter 5 may be restated as a rule for targeting hot spots:

Rule 4 *Site i on the state or regional road network is an accident hot spot, if $x_{i\cdot} > x_{\min}$ and $p\left(X \geq x_{i\cdot} | \mu_i T_i\right) < \alpha$.*

Currently the Road Directorate and regional authorities use $x_{\min} = 4$ (for a 5 year period) and $\alpha = 5\%$ or less. For intersections and roundabouts, $p\left(X \geq x_{i\cdot} | \mu_i T_i\right)$ is easily calculated, and a site may be targeted directly from rule 4. The probability constraint in rule 4 corresponds to the accident counts, $x_{i\cdot}$, exceeding the $1 - \alpha$ quantile in the Poisson distribution with mean $\mu_i T_i$ ($\mu_i T_i L_i$ for road sections).

Road sections differ in length, and accidents tend to accumulate in certain sections of the road, indicating that the level of safety varies over the road section (see appendix A for details). If such an accumulation of accidents is large enough within a sub-section of the road, the national and regional authorities define the sub-section as an accident hot spot. In order to target hot spots on road sections, a so-called *slide method*, combining the two requirements in rule 4 is used. The road length, l, corresponding to a probability of α for observing $x_{\min}$ or more accidents at sites with similar traits (traffic flow) in the *ap-group*, is found as the solution to the equation:

$$p\left(X \geq x_{\min} | \mu_i T_i\right) = 1 - \sum_{x=0}^{x_{\min}-1} \frac{\left(\mu_i \cdot T_i \cdot l\right)^x}{x!} \cdot e^{-(\mu_i \cdot T_i \cdot l)} = \alpha$$

where μ_i is the expected number of accidents per kilometer at road section i. A slide of length l is then sledded over the road section. If at any point, the number of reported accidents within the slide exceeds $x_{\min}$, the corresponding sub-section of road section i is targeted as a hot spot.

C.3 Prioritizing

Assume one has a set of potential safety schemes $\mathbf{Y} = \{Y_1, Y_2, Y_3, ...\}$ and a set of accident hot spots $\mathbf{i} = \{i_1, i_2, i_3, ...\}$ such that all potential safety schemes at an accident hot spot in **i** are contained in **Y**. The selection of an optimal portfolio of safety schemes is a prioritizing process in two steps (see Vejdirektoratet (1992)). The first part is based on the first year benefit (FYB) of the potential safety scheme at the site in question. For a potential scheme, Y, at an accident hot spot i, the FYB is calculated as the ratio of the expected saved accident cost, AC_{iY}, the first year at site i due to scheme Y and its cost of implementation, C_{iY}:

$$FYB_{iY} = \frac{AC_{iY}}{C_{iY}}$$

For two safety schemes, Y_1 and Y_2, scheme Y_1 dominates Y_2 at site i if $FYB_{iY_1} > FYB_{iY_2}$. For each accident hot spot, the potential safety schemes are ranked according to their FYB-value, and the scheme with the highest FYB is denoted the primary scheme at this site.

In the second part of the prioritizing process, the potential safety schemes at different sites are compared. For each scheme Y at a site i, the marginal benefit (MB) of the excess investment needed for implementing a more costly scheme (lower FYB), but with a higher saved accident cost, is calculated. As an example, consider two potential safety schemes Y_1 and Y_2 at site i, with Y_1 as the primary scheme, but with $AC_{iY_1} < AC_{iY_2}$. The marginal benefit of implementing scheme Y_2 instead of scheme Y_1 is calculated as:

$$MB_{i,Y_1Y_2} = \frac{AC_{iY_2} - AC_{iY_1}}{C_{iY_2} - C_{iY_1}}$$

The argument is then, that a non-primary safety scheme, Y_2, at site i with a higher MB than the FYB of a primary measure at another site, is more cost-efficient. As an example, assume one has two accident hot spots i and i' with the following first year benefits of three schemes Y_1, Y_2 and Y_3:

$$FYB_{iY_1} > FYB_{iY_2} > FYB_{i'Y_3}$$

where Y_1 and Y_3 are the primary measures at site i and i' respectively. The measure Y_2 is a non-primary potential safety scheme at site i. If the marginal benefit of Y_2 at site i exceeds the first year benefit of scheme Y_3 at site i', i.e. $MB_{i,Y_1Y_2} > FYB_{i'Y_3}$, then Y_2 *dominates* Y_3, and it is altogether more efficient to implement the non-primary scheme Y_2 at site i instead of implementing the primary schemes Y_1 and Y_3 at site i and i' respectively.

Let ε_Y be the reduction rate in accidents due to the implementation of scheme Y, and let $x_{i\cdot}$ be the total reported number of accidents within the observation period. Then, the expected saved accident costs the first year at site i due to scheme Y, AC_{iY}, are generally estimated as:

$$AC_{iY} = \frac{\varepsilon_Y \cdot x_{i\cdot} \cdot p}{T_i}$$

where T_i is the length of the observation period at site i and p is the cost to society of one accident[1]. Previous studies of the reduction rate, ε_Y, in accidents due to safety scheme Y may be available (e.g. from Elvik et al. (1997)). However, if this is not the case, the reduction rate in each of the reported accidents at the site due to the implementing measure Y is roughly imputed:

$$\widehat{\varepsilon}_Y = \begin{cases} 50\%, \text{ on accidents certain to be affected by measure } Y \\ 33\%, \text{ on accidents which might be affected by measure } Y \\ 0\%, \text{ on accidents not affected by measure } Y \end{cases}$$

The average estimated reduction rate in accident counts at site i is given as a weighted average of the individual relative effects.

If the potential safety scheme results in the site changing to another ap-group, the reduction rate in accident counts is calculated as the difference between the estimated site safety levels before and after treatment:

$$\widehat{\varepsilon}_Y = 1 - \frac{\widehat{\mu}_{i,after}}{\widehat{\mu}_{i,before}}$$

The optimal portfolio of safety schemes is found by successively selecting the most dominant preventive safety scheme in $\mathbf{l}$ till the budget or another constraint is reached.

C.4 Before and after studies

Analogous to chapter 4, let T be the year of implementation of the scheme of safety measure(s) at site i in an ap-group H. The periods before and after treatment are $]0; T-1]$ respectively $]T; U]$ (see figure 4.4). The year of implementation, T, is excluded from the study. Let $x_{i\cdot,b}$ and $x_{i\cdot,a}$ be the reported number of accidents in the periods before and after treatment respectively. Besides changes inflicted by the treatment, it is assumed that no other changes have been made to the road geometry in these periods. The effect of treating a site i is estimated as:

$$\widehat{\varepsilon}_i = 1 - \frac{x_{i\cdot,a}}{x_{i\cdot,b} \cdot \frac{U-T}{T-1} \cdot C_{i,traffic} \cdot (1 - C_{RTM}) \cdot C_{general}} \tag{C.6}$$

where $x_{i\cdot,b} = \sum_{t=1}^{T-1} x_{it}$ and $x_{i\cdot,a} = \sum_{t=T+1}^{U} x_{it}$. In (C.6), the crude accident count of the period after treatment, $x_{i\cdot,a}$, is used as the estimated site safety level at the

[1] Accidents may be classified into different groups of severity, e.g. into injury and property damage only accidents.

site with treatment. The adjusted accident count of the period before treatment is used as the corresponding predicted site safety level had the safety scheme not been implemented. The accident count, $x_{i\cdot,b}$, of the period before treatment is adjusted for a number of factors.

First of all, if the periods before and after treatment are of different lengths, i.e. if $(U-T) \neq (T-1)$, then the total number of reported accidents in the period before is adjusted by:

$$\frac{U-T}{T-1}$$

Secondly, the correction factor, $C_{i,traffic}$, is the estimated change in site safety due to a general change in traffic flow at site i, and is calculated as:

$$C_{i,traffic} = \frac{T-1}{U-T} \frac{\sum_{t=T+1}^{U} AADT_{it}}{\sum_{t=1}^{T-1} AADT_{it}}$$

where $AADT_{it}$ is the total annual average daily traffic in year t regardless of the fact whether site i is an intersection or a road section. The correction factor, C_{RTM}, is the change in site safety due to the regression to the mean effect[2]. The definition of the regression to the mean effect used in (C.6) is in line with the RTM effect in Kulmala (1995) (see appendix B). Hence, this definition is different from the definition of RTM used in chapter 4.

The correction factor, $C_{general}$, is used to describe the general change in site safety from the period before to the period after treatment. It is estimated as the change in accident counts at sites in ap-group H, where no safety schemes are implemented (the control group). Let $J \subseteq H$ denote the control group, then $C_{general}$ is calculated as:

$$C_{general} = \frac{T-1}{U-T} \frac{\sum_{j \in J} \sum_{t=T+1}^{U} x_{jt}}{\sum_{j \in J} \sum_{t=1}^{T-1} x_{jt}}$$

The factor $C_{general}$ must also be adjusted for differences in length of before and after periods in the control group. This factor is used to describe the trend in accident counts over time.

An estimated positive effect of treatment, $\widehat{\varepsilon}_i > 0$, indicates a positive effect of the scheme on the site safety level at site i. The percentage change in site safety is tested in a chi-square distribution with statistic (see Vejdirektoratet (1999d)):

$$\chi^2 = \frac{\left(x_{i\cdot,b} \cdot \frac{U-T}{T-1} \cdot C_{i,traffic} \cdot (1 - C_{RTM}) \cdot C_{general} - x_{i\cdot,after}\right)^2}{x_{i\cdot,b} \cdot \frac{U-T}{T-1} \cdot C_{i,traffic} \cdot (1 - C_{RTM}) \cdot C_{general}} \tag{C.7}$$

[2]Estimated to be 25% for intersections with no signal control (see Vejdirektoratet (1999d)).

A positive effect of the treatment is present, if the accident count of the period after treatment is significantly[3] lower than the adjusted accident count of the period before treatment.

[3] The level of significance is usually chosen to be $1\%, 2.5\%$ or 5%.

Appendix D

The framework of the data analysis

The purpose of this chapter is to go through the estimation of the parameters of the models proposed in chapter 2. In the process, the selection of data from VIS is described, as well as the statistical modelling procedures. At the end of the chapter, the resulting estimated parameters of the models are listed.

D.1 Data

Recorded traffic accidents are classified into 10 main accident categories listed in table D.1. The most common types of accidents on state and regional roads are single-vehicle accidents, followed by accidents of vehicles going in the same direction with no turning and accidents of intersecting vehicles with no turning.

Main accident categories	
0	Single-vehicle accidents
1	Accidents of vehicles going in the same direction with no turning
2	Accidents of vehicles going in the opposite direction with no turning
3	Accidents of vehicles going in the same direction with turning
4	Accidents of vehicles going in the opposite direction with turning
5	Accidents of intersecting vehicles with turning
6	Accidents of intersecting vehicles with no turning
7	Accidents involving parked vehicles
8	Accidents involving pedestrians
9	Accidents involving animals or objects on the road

Table D.1: Main accident categories.

Figure D.1 illustrates the percentage distribution of the main accident categories over the period 1989-98.The relative distribution of the number of accidents between

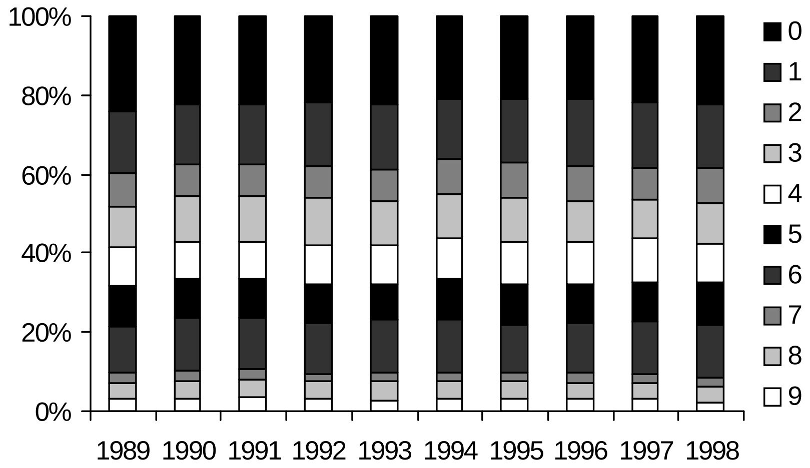


Figure D.1: Total number of accidents by accident category for the period 1989-98.

these categories seem to be very stable over the 10 year period.

D.2 Data selection

Selection of data from VIS is based on the standard procedures used by the Road Directorate. For each site, the latest alteration date as well as the latest date of hot spot treatment is recorded. Hot spot treatment does not necessarily change the recorded site characteristics. The site characteristics may be classified into *numerical variables* such as traffic flow and width of the road, and into *categorical variables*. The categorical variables are particular by the fact that they only assume a limited number of values (levels). In general, records with missing numerical variables are omitted, while missing categorical variables are included, but indicated by a *not available (NA)* level. There is a large overlap of records for the missing variables, e.g. the majority of records with no information on the width of the lanes also lacks information on the width of the whole road (pavement width). Below, the procedures used for selecting intersection and road section data are described.

D.2.1 Intersections

The base dataset consists of 74,126 intersections. The details of the data selection procedure are as follows:

- Only intersections categorized as 3-, 4- or >4-arm junctions are selected. Hereby, railway crossings and local and regional boundaries are omitted.

- For each arm, the frontage of the road is selected from the corresponding road section in the database. The frontage of the junction is selected as the frontage with the highest priority of the arms (see below).

- Junctions with measured traffic flow (AADT) on less than 3 of the arms are omitted.

- Of the remaining sites, junctions with the latest measured traffic flows prior to 1989 are omitted, because these measures are considered outdated and thus not representative of the current site.

- The categorization of the junctions is revised according to the number of arms in the junction with measured traffic flow. This results in three categories; 3-, 4- and 5-arm junctions.

- For more than 2/3 of the junction arms, the average annual daily traffic has been recorded for each of the years 1994-98. The AADT in the remaining years is calculated from the AADT available, using the Danish national traffic growth index for average yearly increases in traffic (see Vejdirektoratet (2002b)).

- The arms of a junction are sorted by yield relations (asc.), road ID (asc.) and traffic flow (desc.). This is in line with the sorting by the Road Directorate. Subsequently, the 2 upper arms are categorized as major arms and the remaining as minor arms.

- The upper arm in the major and the minor arm category respectively determines the channelisation, yield relations and road ID of its class. The traffic flows are divided by 2 and summed within each arm category.

- Junctions with major or minor road ID not available are removed (about 8% of the remaining junctions).

- 4- or 5-arm junctions with a total AADT of less than 500 on minor arms and 3-arm junctions with less than 250 are omitted (about 68% of the remaining junctions).

- Only the period after the year of the last alteration/treatment is included. Hence, junctions altered in 1998 are completely omitted from the set (about 5% of the remaining).

The selection procedure above results in a dataset of 2,944 junctions of 3-, 4- or 5-arms for the period 1989-98. However, only data for the period 1994-98 are used. This results in a dataset of 14,182 records with individual observation periods ranging from 1 to 5 years as listed in table D.2.

	Junctions	
Years	Number	%
1	55	2
2	46	2
3	53	2
4	74	3
5	2,716	92
Total	2,944	100

Table D.2: Observation periods for the intersection dataset.

It can be seen that more than 90% of the junctions in the study have an observation period of 5 years.

D.2.2 Road sections

The base dataset consists of 92,633 road sections. For part of the road sections, data for both sides of the road are supplied (mostly road sections with a median). However, because accident location does not indicate the side of the road, general site characteristics applying to the entire road section are derived. In general, values of numerical variables are added, and the most influential of the categorical variables is selected. The selection procedure is described below:

- For each section, the speed limit is taken to be the local speed limit. If no local speed limit is present, the general speed limit for the road class is applied.
- The frontage, crash barrier and crash barrier in the median of the road section is selected as the one with highest priority (see below).
- The width of the road sides, the width of the lanes and the number of lanes are summed.
- Road sections with no information on the width of the road or lanes are removed.
- Sections with the latest measured traffic prior to 1989 or where AADT is less than 500 are removed (less than 2%).

- For almost all of the road sections, the average annual daily traffic is measured for each of the years 1994-98. The AADT for the remaining years are calculated from the AADT available, using the Danish national traffic growth index for average yearly increases in traffic (see Vejdirektoratet (2002b)).
- Only the period after the year of the last alteration/treatment is included. Thus road sections altered in 1998 are completely omitted from the set (about 1% of the remaining).

The dataset now contains records for 89,545 road sections for the time period 1989-98. However, only the period 1994-98 is used, which results in a dataset of 437,683 records. The road sections were further classified into site-groups based on the stated road class:

Motorways defined as normal road paths (e.g. ramps are excluded), and with a stated general speed limit of 110 km/h. Because motorways all have medians, only sites where the total width of the road exceeds the total width of the lanes, and where the total number of lanes is between 4 and 8 are selected. In addition, sites with an average lane width less than 2.5 meters are omitted.

Motortrafficways defined as normal road paths, and with a stated general speed limit of 80 km/h. Only sites where the total width of the road is at least that of the total width of the lanes, and where the total number of lanes is 2 or more are selected. In addition, sites with an average lane width less than 2.5 meters are omitted.

Other roads defined as normal road paths and with a stated general speed limit of 50 or 80 km/h. Only sites where the total width of the road is at least that of the total width of the lanes, and where the total number of lanes is 2 or more are selected. In addition, sites with an average lane width less than 2.5 meters are omitted.

The data selection above is applied to ensure that the site-groups are relatively homogenous, and to exclude records with misclassifications. The remainder of the selected road sections was gathered in a fourth group, *remaining roads*, and parameters of a model were estimated for this site-group as well. In the group of remaining roads, information on the number of lanes and the width of the road and lanes was not included, due to the risk of these being misclassified. Table D.3 lists the observation periods for sites included in the study. It appears that more than 90% of the road sections in the study have an observation period of 5 years.

Accidents are assigned to intersections and road sections using the road identification number and the kilometrage. Only accidents reported after the last year of alteration/treatment of the site are included.

	Motorways		**Motortrafficways**		**Other roads**		**Remaining**	
Years	Number	%	Number	%	Number	%	Number	%
1	0	0	0	0	620	1	104	1
2	0	0	41	2	686	1	61	1
3	0	0	37	1	1,152	2	137	2
4	239	4	20	1	1,444	2	267	4
5	5,731	96	2,616	96	69,550	94	6,806	92
Total	5,970	100	2,714	100	73,452	100	7,375	100

Table D.3: Observation periods for the road section dataset.

D.3 Statistical modelling

Above, sites are classified into different site-groups, each with its own set of potential traits. Consequently, different explanatory variables were attempted included in the model for different site-groups as listed in table D.4. The variables; *Year*, *No. of arms*, *AADT*, *length*, *width* and *width of lanes* are numerical variables. The remainder is categorical variables and may only assume a limited number of values (levels).

Intersections	**Road sections**			
Junctions	Motorways	Motortraffic-ways	Other roads	Remaining
Year	Year	Year	Year	Year
No. of arms	Length	Length	Length	Length
Frontage	AADT	AADT	AADT	AADT
Major	Width	Width	Width	Speed
AADT	Width of lanes	Width of lanes	Width of lanes	Barrier
Yield relations	Speed	Speed	Speed	Barrier middle
Channelisation	No. of lanes	No. of lanes	No. of lanes	Median
Minor	Barrier	Barrier	Barrier	Frontage
AADT	Barrier middle	Barrier middle	Barrier middle	Edge
Yield relations		Median	Median	Bicycle path
Channelisation			Frontage	
			Edge	
			Bicycle path	

Table D.4: The potential traits for each site-group.

One level of the categorical variable is used as reference[1]. For many records, information on one or more of the categorical variables was missing. As a result,

[1] So-called *treatment coding* of contrasts.

these variables were given a level indicating that no information was available on this site characteristic (NA). In general, the level *NA* is used as reference. For categorical variables with no NA-level, the most common, or the most meaningful level is used as reference. As an example, the speed limit equal to 110 is used as reference for motorways.

Sites with an unusually small or large value of one or more of its traits may have a considerable effect on the estimation of the corresponding coefficient (high leverage). As a consequence, these sites are called *outliers* and are sometimes omitted from the accident models. In this study however, these records are included, because deviation from the reference safety will be explained by the dispersion effect. In addition, such sites are likely to be accident hot spots.

The generalized linear model with Poisson errors (see appendix A) is a standard option in the statistical software S-PLUS. Furthermore, simulation studies have shown that the estimated parameters of this model will typically only differ by a few percent from the estimated parameters of the fixed effect part of the mean in the Poisson-gamma hierarchical generalized linear model (in line with findings in Maycock and Hall (1984) and Kulmala (1995)). Hence, the same set of traits is likely to be selected for both model types. This was also the case in this study. As a consequence, the traits for the fixed effect part of the mean were found by fitting an ordinary generalized linear model (GLIM) with Poisson errors to data.

Different site characteristics may express the same condition at sites, hence alternative models exist for describing variation in accident counts. In order to discriminate between model alternatives, a measure of the deviation between model and data is needed. Because Poisson errors are assumed, the deviance is used as measure of goodness-of-fit (see McCullagh and Nelder (1989)). Given accident data, $\mathbf{x}$, the deviance of a model with *fitted* values $\widehat{\boldsymbol{\mu}}$ and Poisson errors is determined as:

$$\begin{aligned} SD\left(\mathbf{x};\widehat{\boldsymbol{\mu}}\right) &= 2l\left(\mathbf{x},\mathbf{x}\right) - 2l\left(\mathbf{x};\widehat{\boldsymbol{\mu}}\right) = 2\sum_{i=1}^{I}\sum_{t=1}^{T_i}\left(x_{it}\log\left(\frac{x_{it}}{\widehat{\mu}_{it}}\right) - \left(x_{it}-\widehat{\mu}_{it}\right)\right) \\ &\equiv \sum_{i=1}^{I}\sum_{t=1}^{T_i} SD\left(x_{it};\widehat{\mu}_{it}\right) \end{aligned}$$

where $\widehat{\mu}_{it}$ is the estimated reference safety at site i in year t. In the case of road section i of length L_i, the estimated safety in year t is $\widehat{\mu}_{it}L_i$ instead of $\widehat{\mu}_{it}$ in $SD(\cdot)$. The deviance is a measure of the deviation between accident data and the fitted values under the model[2].

Each time an explanatory variable is omitted from the model, the deviance increases. However, this does not mean that all available site characteristics should be

[2]The goodness-of-fit of the model is tested by the χ^2 distribution. However, because of the large number of records with zero accidents, a modified deviance (assuming $\sigma = 1$) is used to ensure

included in the model as traits. Including a large amount of explanatory variables may decrease the ability of the model to predict safety levels outside the dataset used in the estimation (see Akaike (1983) for a discussion). Selecting traits for the hierarchical generalized model in chapter 2 is in practice the choice between including a particular site characteristic in the fixed effect part of the mean, μ, or in the dispersion part of the mean, s. Informally placed, the aim is to select enough traits for the model to be able to predict safety levels for before and after studies, and to select so few traits that abnormalities of accident hot spots are reflected in the dispersion effect. In order to keep the model simple and easy to interpret, traits describing possible interactions between different site characteristics were not considered, i.e. only main effects are included in the model. Because data do not result in an orthogonal design, correlation of effects indicates that different traits to a certain degree describe the same conditions at the site.

Statistical significance is used as guidance for selecting traits for the model. In general, only statistically significant traits are included. A trait is statistically significant in the sense that its estimated coefficient differs significantly from zero. The significance of a variable is measured by its so-called deviance residual. The deviance residual of a variable z_k is measured by the decrease in deviance from including variable z_k with the other $z_1, ..., z_{k-1}$ variables in a model:

$$SD_{z_k} \equiv SD\left(\mathbf{x};\widehat{\boldsymbol{\mu}}_{z_1,\ldots,z_{k-1}}\right) - SD\left(\mathbf{x};\widehat{\boldsymbol{\mu}}_{z_1,\ldots,z_{k-1},z_k}\right) \geq 0 \tag{D.1}$$

The measure SD_{z_k} is approximately χ^2_{df} distributed, where *df* are the degrees of freedom associated with variable z_k. If SD_{z_k} is less than the 95% quantile of the χ^2_{df} distribution, the *null hyphotesis*, that the coefficient of z_k is zero, cannot be rejected on a 5% level. In this case, variable z_k is termed statistically insignificant. It is important to note that one does not know, if the coefficient of variable z_k is in fact zero. One may only conclude that in the given dataset, the coefficient of variable z_k was not shown to be statistically different from zero. This implies, that one runs the risk of committing the error of omitting a variable, which in reality is statistically significant and vice versa. Also known as type II and type I errors. The level of the test (5% in this study) may thus be interpreted as the probability of type II error. This value reflects the decision-maker's concern for committing error

approximate normality (see Jørgensen (1997)):

$$SD\left(\mathbf{x};\widehat{\boldsymbol{\mu}}\right)^* = \sum_{i=1}^{I}\sum_{t=1}^{T_i}\begin{cases}\left(\sqrt{SD\left(x_{it};\widehat{\mu}_{it}\right)} + \frac{1}{\sqrt{SD\left(x_{it};\widehat{\mu}_{it}\right)}}\log\left(\frac{\sqrt{x_{it}}\log\left(\frac{x_{it}}{\widehat{\mu}_{it}}\right)}{\sqrt{SD\left(x_{it};\widehat{\mu}_{it}\right)}}\right)\right)^2, x_{it} > 0 \\ \left(\Phi^{-1}\left(\frac{1}{2}\exp\left(-\widehat{\mu}_{it}\right)\right)\right)^2, x_{it} = 0\end{cases}$$

where Φ is the cumulative distribution function of the standard normal variable. For road sections, $\widehat{\boldsymbol{\mu}}\mathbf{L}$ is used instead of $\widehat{\boldsymbol{\mu}}$.

II relative to the concern of committing error I (see Aagaard (1997) for a discussion). Also, variables failing to show statistical significance still contribute to the explained part variation as reflected in the fixed effect part of the mean, μ (see figure 2.1 of chapter 2). Consequently, despite its failure to demonstrate significance in some of the site-groups, the time trend variable, γ, is included in all of the models for the purpose of predicting future safety levels.

Of the remaining set of available site characteristics, however, only statistically significant variables are included as traits. The procedure applied is a so-called *backwards elimination*. For each of the site characteristics, the deviance residual in (D.1) is calculated[3]. If one or more variables are statistically insignificant, the site characteristic of least significance is removed and the procedure is repeated until all remaining variables are found to be significant.

If possible, levels of categorical variables failing to demonstrate significance were combined into fewer levels, in order to possibly demonstrate statistical significance. This was applied to categorical variables with a high correlation between some of its levels, or where the estimated coefficients were almost identical. Also, if one or more of the levels of a categorical variable had very few observations, they were grouped before fitting the model. As an example, the number of lanes on motorways is grouped into 4, 5, 6 or above 6 lanes from the beginning. The first three groups were the most common number of lanes for this site-group, while the latter group had very few sites. Both statistical as well as expert knowledge[4] have been considered when combining levels of a categorical variable. As an example, the variable describing the frontage of the road was classified into 3 groups, based on the type of road users appearing on roads with different frontage (no vulnerable road users, vulnerable road users and vulnerable road users together with children).

Upon selecting the traits for the fixed effect part of the mean, the components of the Poisson-gamma hierarchical generalized linear models proposed in chapter 2 were estimated. The components were estimated in S-PLUS using the estimation procedures in appendix A. The estimation procedures are not standard options in S-PLUS but had to be programmed. The estimated model parameters for each site-group are given below.

D.3.1 Intersections

In a 3-, 4- or 5-arm junction, the number of accidents, x, in a given year is modelled by the hierarchical generalized linear model set up in chapter 2 (site and time indices

[3] Also known as a type III test.

[4] The Road Directorate and other traffic safety experts.

are omitted):

$$\begin{aligned} p\left(x|s\right) &= \frac{\lambda^{x}}{x!}e^{-\lambda} \\ f\left(s\right) &= \frac{\alpha^{\alpha}}{\Gamma\left(\alpha\right)}s^{\alpha-1}e^{-\alpha s} \\ \lambda &= \mu s \end{aligned}$$

The structure of the fixed effect part of the mean is:

$$\mu = a \cdot \gamma^{\triangle t} \cdot AADT_{ma}^{b_1} \cdot AADT_{mi}^{b_2} \cdot \exp\left(\sum_{k,l} \beta_k\left(l\right) Z_k\left(l\right)\right)$$

and the dispersion effect, s, is estimated as (see chapter 2):

$$\widehat{s} = \frac{\alpha + x.}{\alpha + \mu.}$$

where, $x.$ is the total reported number of accidents within the observation period and $\mu.$ is the total corresponding reference safety. The site safety level, λ, is estimated as $\lambda = \mu s$. The value of the dispersion parameter, α, is an indication of the variation in the dispersion effect, s, at a site, but it is also an overall indication of how well the model fits data. A large α indicates that the model fits data well and vice versa.

The average annual decrease, $\gamma - 1$, in the expected number of accidents due to trends in time, is estimated to be 3% for the total number of accidents and 0% for injury accidents. The part, $\exp\left(\sum_{k,l} \beta_k\left(l\right) Z_k\left(l\right)\right)$, is the contribution from the categorical variables, where $Z_k\left(l\right)$ indicates whether or not variable k on level l is present at the site:

$$\begin{aligned} \sum_{k,l} \beta_k\left(l\right) Z_k\left(l\right) = {} & \beta_1\left(1\right) \cdot I\left(\text{no.arms4}\right) + \beta_1\left(2\right) \cdot I\left(\text{no.arms5}\right) + \beta_2\left(1\right) \cdot I\left(\text{front1}\right) \\ & + \beta_2\left(2\right) \cdot I\left(\text{front2}\right) + \beta_3\left(1\right) \cdot I\left(\text{yield.ma1}\right) \\ & + \beta_3\left(2\right) \cdot I\left(\text{yield.ma2}\right) + \beta_4\left(1\right) \cdot I\left(\text{yield.mi1}\right) \\ & + \beta_4\left(2\right) \cdot I\left(\text{yield.mi2}\right) + \beta_5\left(1\right) \cdot I\left(\text{chan.ma}\right) \\ & + \beta_6\left(1\right) \cdot I\left(\text{chan.mi}\right) \end{aligned}$$

The estimated coefficients for the traits are listed in table D.5.

For the categorical variable *frontage*, the levels are listed in increasing priority. Thus, for a junction with both shops and flats, one should select the level *urban/low buildings or residences and flats* for this site.

Variable name	**Variable**	**Total**	**Injury**
Dispersion parameter	α	1.83	1.67
Intercept		0.000127	0.000060
Time trend	γ	0.97	1.00
AADT on major arms	$AADT_{ma}$	0.43	0.45
AADT on minor arms	$AADT_{mi}$	0.44	0.44
No. arms			
3	-		
4	no.arms4	0.54	0.55
5	no.arms5	−0.45	−0.77
Frontage			
NA, none/scarce ribbon development or road side development with no frontage	-		
Industry or shops	front1	−0.30	−0.35
Urban/low buildings or residences and flats	front2	−0.24	−0.28
Yield relations on major arms			
NA or none	-		
Signal controlled	yield.ma1	−1.95	−1.16
Other	yield.ma2	−1.10	−1.38
Yield relations on minor arms			
NA or none	-		
Signal controlled	yield.mi1	2.92	2.07
Other	yield.mi2	0.81	0.93
Channelisation on major arms			
NA or none	-		
Yes	chan.ma	0.14	0.12
Channelisation on minor arms			
NA or none	-		
Yes	chan.mi	0.33	0.36

Table D.5: Estimated coefficients for 3-, 4- and 5-arm junctions.

The estimated coefficients in table D.5 indicate that signal controlled junctions on average have 2.6 as many accidents[5] (injury or property damage only) than junctions with no yield relations but with otherwise similar traits. In addition, junctions with channelisation on major and/or minor arms have on average a higher number of accidents than junctions with no channelisation but with otherwise similar traits. The sign of the coefficients for yield relations is in line with findings in Kulmala (1995), while the opposite is the case for channelisation coefficients. However, before and after studies in Kulmala (1995) did show that left and right turn lanes had a negative effect on safety in 3-arm junctions. Vejdirektoratet (1995) also found positive coefficients for channelisation in junctions in urban areas. In general, the sign and magnitude of the coefficients in table D.5 are approximately the same for the total number of accidents and for injury accidents. Also, the estimated dispersion parameter, $\widehat{\alpha}$, almost has the same value for injury accidents as for the total number of accidents. The standard error of the coefficients is on average about 30% larger for injury accidents.

Example

Consider a 4-arm signal controlled junction in an industrial area with channelisation on the major arms. The observation period for this site is three years, i.e. 1996-98. In 1996, the average annual daily traffic on major and minor arms is 12,976 and 4,830 respectively. The reference safety level for this year is estimated as:

$$\widehat{\mu} = 0.000127 \cdot 0.97^2 \cdot 12976_{ma}^{0.43} \cdot 4830_{mi}^{0.44} \cdot \exp(0.54 - 0.30 - 1.95 + 2.92 + 0.14) = 1.54$$

Hence, one has an expected number of accidents at sites with similar traits as the one in question of 1.54 in 1996[6]. The reference safety levels for the years 1997 and 1998 are estimated in a similar way. Assume a total of 8 accidents have been reported in the period 1996-98. The corresponding reference safety level for the observation period is found to be 4.25. In order to estimate the site safety level in 1996, the dispersion effect is estimated:

$$\widehat{s} = \frac{1.83 + 8}{1.83 + 4.25} = 1.62$$

The site safety at the site in 1996 is $\widehat{\lambda} = \widehat{\mu}\widehat{s} = 1.54 \cdot 1.62 = 2.50$, which is 62% higher than expected at similar sites.

[5] $\exp(-1.95 + 2.92) = 2.6$

[6] The corresponding reference safety for injury accidents is 0.61.

D.3.2 Road sections

For a road section of a certain road class, the number of reported accidents, x, in a given year is modelled by the hierarchical generalized linear model (site and time indices are omitted):

$$\begin{aligned} p(x|s) &= \frac{(\lambda L)^x}{x!} e^{-\lambda L} \\ f(s) &= \frac{(\alpha L)^{\alpha L}}{\Gamma(\alpha L)} s^{\alpha L - 1} e^{-\alpha L s} \\ \lambda &= \mu s \end{aligned}$$

where λ is the site safety level per kilometer and μ is the corresponding reference safety level. The variable L is the length of the road section in kilometers.. The dispersion effect, s, is modelled by a gamma distribution with shape and scale parameter both equal to αL. The general expression for the reference safety per kilometer, μ, on road sections is:

$$\mu = a \cdot \gamma^{\triangle t} \cdot AADT^b \cdot \prod_j Y_j^{\beta_j} \cdot \exp\left(\sum_{k,l} \beta_k(l) Z_k(l)\right)$$

where a is a constant (the intercept) and $\gamma - 1$ is the annual change in the expected number of accidents due to trends in time, and $AADT$ is the average annual daily traffic on the road section. The part $\prod_j Y_j^{\beta_j}$ is the contribution of the numerical variables other than γ and $AADT$, with Y_j representing the value of variable j. The part, $\exp\left(\sum_{k,l} \beta_k(l) Z_k(l)\right)$, is the contribution from the categorical variables as described above. The traits included in the model are different for each site-group.

The dispersion effect, s, is estimated as (see chapter 2):

$$s = \frac{\alpha L + x./L}{\alpha L + \mu.}$$

Here $x.$ is the total reported number of accidents within the observation period and $\mu.$ is the total corresponding reference safety per kilometer for the period. The site safety level, λ, is calculated as $\lambda = \mu s$.

Motorways

On motorways, the structure of the fixed effect part of the mean is:

$$\mu = a \cdot \gamma^{\triangle t} \cdot AADT^{b_1} \cdot WR^{b_2} \cdot \exp\left(\sum_{k,l} \beta_k(l) Z_k(l)\right)$$

The average annual decrease, $\gamma - 1$, in the expected number of accidents due to trends in time, is estimated to be 2% for the total number of accidents and 1% for injury accidents. The numerical variable WR is the remaining width in meters of the road, when the width of the lanes is deducted. Hence, WR expresses the width of the median. The contribution from the categorical variables is:

$$\begin{aligned}\sum_{k,l} \beta_k(l)\, Z_k(l) &= \beta_1(1) \cdot I(\text{speed}<110) + \beta_2(1) \cdot I(\text{no.lanes5,6}) \\ &\quad + \beta_2(2) \cdot I(\text{no.lanes}>6) + \beta_3(1) \cdot I(\text{bar.mid1}) \\ &\quad + \beta_3(2) \cdot I(\text{bar.mid2})\end{aligned}$$

The estimated coefficients for the variables are listed in table D.6.

Variable name	**Variable**	**Total**	**Injury**
Dispersion parameter	α	4.60	1.53
Intercept		0.000026	0.000048
Time trend	γ	0.98	0.99
AADT	$AADT$	1.02	0.89
Width rest	WR	−0.22	−0.12
Speed			
Below 110	speed<110	0.51	0.38
110	-		
Number of lanes			
4	-		
5 or 6	no.lanes5,6	0.19	0.12
Above 6	no.lanes>6	0.39	0.31
Crash barrier median			
NA or none	-		
Cable barrier	bar.mid1	−0.07	−0.19
Other	bar.mid2	0.06	−0.43

Table D.6: Estimated coefficients for motorways.

Table D.6 indicates that motorways with a relatively large WR (typically a wide median) on average have a lower total number of accidents than motorways with a relatively small WR but with otherwise similar traits. This is in line with findings in Abdel-Aty and Radwan (2000). In general, the sign and magnitude of the coefficients in table D.6 are approximately the same for the total number of accidents and for injury accidents. The exception is crash barrier in the median other than cable. Motorways with such a crash barrier in the median do on average have a higher total number of accidents (about 6%), but a lower number of injury accidents (about 35%), than motorways with no crash barrier in the median but with otherwise similar traits. In other words, accidents here seem to be of less severity. Table D.6 shows that the estimated dispersion parameter, $\widehat{\alpha}$, is about three times smaller for injury accidents than for the total number of accidents. Consequently, the variation in the dispersion effect, s, is relatively larger in a model estimated for injury accidents. Also, the standard error of the coefficients is on average about 50% larger for injury accidents.

Example Consider a 4-lane motorway section of length 0.35 kilometer. The total width of the road is 17 meters with a total of 15 meters for the road lanes. The speed limit is 110 km/h, and the crash barrier in the median is other than cable. The observation period is 1994-98. In 1998, the average annual daily traffic was 21,656, and the reference safety per kilometer for the total number of accidents in this year is estimated as:

$$\widehat{\mu} = 0.000026 \cdot 0.98^4 \cdot 21656^{1.02} \cdot 2^{-0.22} \cdot \exp(-0.07) = 0.59$$

Hence, in 1998 one has an expected number of accidents at sites with similar traits as the one in question of 0.59 per kilometer[7]. The reference safety levels for the remaining period 1994-97 are estimated in a similar way. Assume no accidents have been reported. The corresponding reference safety for the whole observation period is estimated to be 1.98. In order to estimate the site safety per kilometer in 1998, the dispersion effect is estimated:

$$\widehat{s} = \frac{4.60 \cdot 0.35 + 0}{4.60 \cdot 0.35 + 1.98 \cdot 0.35} = 0.70$$

The site safety per kilometer at the site in 1998 is $\widehat{\lambda} = \widehat{\mu}\widehat{s} = 0.59 \cdot 0.70 = 0.41$, which is 30% lower than expected at similar sites. The road section has a length of 0.35 kilometer which results in $0.41 \cdot 0.35 = 0.14$ expected number of accidents at the site in year 1998.

Motortrafficways

On motortrafficways, the structure of the fixed effect part of the mean is:

$$\mu = a \cdot \gamma^{\triangle t} \cdot AADT^{b_1} \cdot WL^{b_2} \cdot \exp\left(\sum_{k,l} \beta_k(l) Z_k(l)\right)$$

The average annual increase, $\gamma - 1$, in the expected number of accidents due to trends in time, is estimated to be 1% for the total number of accidents and 11% for

[7]The corresponding reference safety for injury accidents is 0.25.

injury accidents. The regression variable WL is the width of the road lanes. The contribution from the categorical variables is:

$$\sum_{k,l} \beta_k(l) Z_k(l) = \beta_1(1) \cdot I(\text{speed90}) + \beta_2(1) \cdot I(\text{no.lanes3}) + \beta_2(2) \cdot I(\text{no.lanes4}) + \beta_2(3) \cdot I(\text{no.lanes>4}) + \beta_3(1) \cdot I(\text{median})$$

The estimated coefficients for the variables are listed in table D.7.

Variable name	**Variable**	**Total**	**Injury**
Dispersion parameter	α	0.98	0.52
Intercept		0.000006	0.000023
Time trend	γ	1.01	1.11
AADT	$AADT$	0.49	0.42
Width lanes	WL	3.16	2.31
Speed			
70 or 80	-		
90	speed90	-0.16	0.08
Number of lanes			
2	-		
3	no.lanes3	-0.76	-0.48
4	no.lanes4	-1.46	-1.15
Above 4	no.lanes>4	-1.05	-0.86
Median			
NA or none	-		
Yes	median	-0.84	-0.35

Table D.7: Estimated coefficients for motortrafficways.

Table D.7 indicates that the number of accidents on motortrafficways with a median on average have a lower total number (about 57%) of accidents than motorways with no median but with otherwise similar traits. In general, the sign and magnitude of the coefficients in table D.7 are approximately the same for the total number of accidents and for injury accidents. The exception is a speed limit of 90. Motortrafficways with a speed limit of 90 km/h do on average have a lower total number of accidents (about 15%), but a higher number of injury accidents (about 8%), than motortrafficways with a lower speed limit, but with otherwise similar traits. In other words, accidents here seem to have a higher degree of severity. Table D.7 shows that the estimated dispersion parameter, $\widehat{\alpha}$, for injury accidents is only about half the value of the corresponding parameter for the total number of accidents. Consequently, the variation in the dispersion effect, s, is relatively larger in a model estimated for injury accidents. In addition, the standard error of the coefficients are on average about 50% larger for injury accidents.

Other roads

On other roads, the structure of the fixed effect part of the mean is:

$$\mu = a \cdot \gamma^{\triangle t} \cdot AADT^{b} \cdot W^{b_1} \cdot WL^{b_2} \cdot \exp\left(\sum_{k,l} \beta_k(l)\, Z_k(l)\right)$$

The estimated average annual change in the expected number of accidents due to trends in time is less than 1% (in fact less than 1% over a five year period) and has consequently no influence on the reference safety level. The regression variable W is the width of the entire road and WL is the width of the road lanes. The contribution from the categorical variables is:

$$\begin{aligned}\sum_{k,l} \beta_k(l)\, Z_k(l) = {} & \beta_1(1) \cdot I(\text{speed}{<}50) + \beta_1(2) \cdot I(\text{speed60}) + \beta_1(3) \cdot I(\text{speed50}) \\ & + \beta_1(4) \cdot I(\text{speed70}) + \beta_2(1) \cdot I(\text{no.lanes3}) \\ & + \beta_2(2) \cdot I(\text{no.lanes4}) + \beta_2(3) \cdot I(\text{no.lanes5}) \\ & + \beta_2(4) \cdot I(\text{no.lanes6}) + \beta_2(5) \cdot I(\text{no.lanes}{>}6) \\ & + \beta_3(1) \cdot I(\text{front1}) + \beta_3(2) \cdot I(\text{front2}) + \beta_4(1) \cdot I(\text{median}) \\ & + \beta_5(1) \cdot I(\text{barrier}) + \beta_6(1) \cdot I(\text{bar.mid}) + \beta_7(1) \cdot I(\text{edge1}) \\ & + \beta_7(2) \cdot I(\text{edge2})\end{aligned}$$

The estimated coefficients for the variables are listed in table D.8.

Table D.8 indicates that the number of accidents on other roads with marginal strips in both sides of the road on average have a lower total number of accidents (about 6%) than other roads with no marginal strips but with otherwise similar traits. The larger part of the coefficients for the total number of accidents and injury accidents in table D.8 are of approximately similar sign and magnitude. One of the exceptions is the presence of marginal strips in one side of the road. Other roads with marginal strips in one side of the road do on average have a higher total number of accidents (about 7%), but a lower number of injury accidents (about 4%), than other roads with no marginal strips but with otherwise similar traits. In other words, accidents here seem to have a relatively low degree of severity. Table D.8 shows that the estimated dispersion parameter, $\widehat{\alpha}$, for injury accidents is less than half the value of the corresponding parameter for the total number of accidents. Consequently, the variation in the dispersion effect, s, is relatively larger in a model estimated for injury accidents. In addition, the standard error of the coefficients is on average about 50% larger for injury accidents.

Variable name	**Variable**	**Total**	**Injury**
Dispersion parameter	α	1.54	0.68
Intercept		0.000672	0.000663
Time trend	γ	1.00	1.00
AADT	$AADT$	0.49	0.51
Width	W	0.01	-0.06
Width lanes	WL	1.06	0.81
Speed			
Below 50	speed<50	0.71	0.34
50	speed50	0.43	0.40
60	speed60	0.24	0.23
70	speed70	0.07	0.33
80	-		
Number of lanes			
2	-		
3	no.lanes3	0.36	0.41
4	no.lanes4	-0.21	-0.16
5	no.lanes5	0.08	-0.15
6	no.lanes6	0.11	0.21
Above 6	no.lanes>6	-2.27	-2.30
Frontage			
NA, none/scarce ribbon development or road side development with no frontage	-		
Industry or shops	front1	0.54	0.46
Urban/low buildings or residences and flats	front2	0.31	0.19
Median			
NA or none	-		
Yes	median	-0.32	-0.46
Crash barrier			
NA or none	-		
Yes	barrier	-0.19	-0.31
Crash barrier median			
NA or none	-		
Yes	bar.mid	-0.41	0.21
Edge			
NA or none	-		
Marginal strips in one side	edge1	0.07	-0.04
Marginal strips in both sides	edge2	-0.06	-0.05

Table D.8: Estimated coefficients for other roads.

Remaining roads

On the remaining roads, the structure of the fixed effect part of the mean is:

$$\mu = a \cdot \gamma^{\triangle t} \cdot AADT^{b} \cdot \exp\left(\sum_{k,l} \beta_k(l)\, Z_k(l)\right)$$

Similar to other roads, the estimated average annual change in the expected number of accidents due to trends in time is less than 1% over a five year period. The contribution from the categorical variables is:

$$\begin{aligned}\sum_{k,l} \beta_k(l)\, Z_k(l) &= \beta_1(1) \cdot I(\text{speed}<70) + \beta_1(2) \cdot I(\text{speed70}) + \beta_1(3) \cdot I(\text{speed}>80) \\ &\quad + \beta_2(1) \cdot I(\text{front1}) + \beta_2(2) \cdot I(\text{front2}) + \beta_3(1) \cdot I(\text{edge1}) \\ &\quad + \beta_3(2) \cdot I(\text{edge2}) + \beta_4(1) \cdot I(\text{cycle}) + \beta_5(1) \cdot I(\text{median})\end{aligned}$$

The estimated coefficients for the variables are listed in table D.9.

Table D.9 indicates that the number of accidents on remaining roads with a bicycle path, on average have a higher total number of accidents (about 51%) than remaining roads with no bicycle path, but with otherwise similar traits. The opposite is the case for remaining roads with medians, which is in line with the estimated parameters of the models for motortrafficways and other roads. In general, the sign and magnitude of the coefficients in table D.9 are approximately the same for the total number of accidents and for injury accidents. The exception is remaining roads with marginal strips in both sides of the road. Such roads do on average have a lower total number of accidents (about 14%), but a higher number of injury accidents (about 42%) than remaining roads with no marginal strips, but with otherwise similar traits. In other words, accidents here seem to have a higher degree of severity. Table D.9 shows that the estimated dispersion parameter, $\widehat{\alpha}$, for injury accidents is less than half the value of the corresponding parameter for the total number of accidents. Consequently, the variation in the dispersion effect, s, is relatively larger in a model estimated for injury accidents. In addition, the standard error of the coefficients is on average about 60% larger for injury accidents.

Variable name	**Variable**	**Total**	**Injury**
Dispersion parameter	α	0.79	0.35
Intercept		0.000391	0.000024
Time trend	γ	1.00	1.00
AADT	$AADT$	0.84	1.08
Speed			
Below 70	speed<70	0.48	0.31
70	speed70	0.99	0.41
80	-		
Above 80	speed>80	-0.81	-0.94
Frontage			
NA, none/scarce ribbon development or road side development with no frontage	-		
Industry or shops	front1	0.90	0.39
Urban/low buildings or residences and flats	front2	0.21	0.40
Median			
NA or none	-		
Yes	median	-0.27	-0.82
Edge			
NA or none	-		
Marginal strips in one side	edge1	-0.51	-0.81
Marginal strips in both sides	edge2	-0.15	0.35
Bicycle path			
NA or none	-		
Yes	cycle	0.41	0.34

Table D.9: Estimated coefficients for the remaining roads.

Appendix E

Simulation program

The purpose of this chapter is to asses the performance of the model and methods proposed in chapters 2-4, and to compare them to the model and methods developed by the Danish Road Directorate (RD), described in chapter 5.

The basis of the simulation studies are two datasets of 3- and 4-arm signal controlled junctions from the VIS database for the period 1994-98 (see chapter 6 for a description of VIS). Sites for which no changes have been made after 1993 were chosen, i.e. sites with observation periods of at least 3 years. The study is based on 168 and 458 signal controlled 3- and 4-arm junctions respectively, and simulations are performed for each of the two datasets individually. From the dataset, samples of 100 junctions were drawn randomly and for each sample, the simulation was repeated 100 times. The overall structure of the simulation study is illustrated in figure E.1.

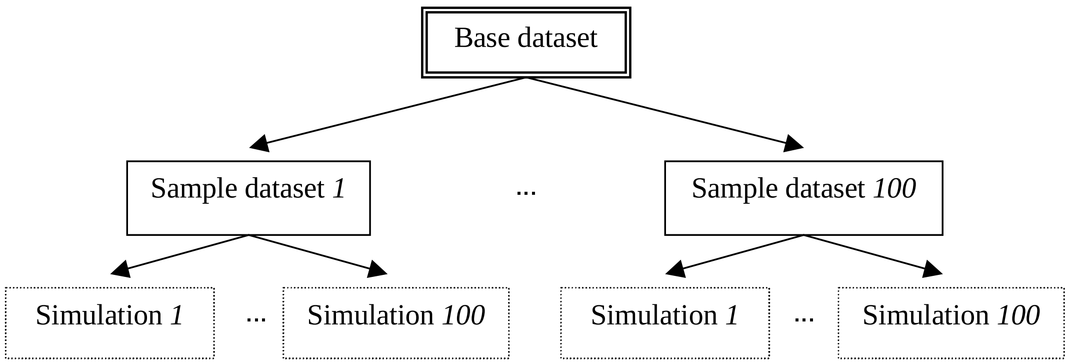


Figure E.1: The overall structure of the simulation study.

Because the observation periods vary from 3 to 5 years, the individual sample sizes also vary. The traits chosen for the simulation correspond to the traits selected in the analysis of accident counts in Denmark (see chapter 6), but with different levels of the categorical variables[1].

The simulation studies are performed in the Statistical software S-PLUS 2000 Professional Release 2.

E.1 The structure of the study

The outline of the simulation study is given in steps (i)-(vii), and each step is described in further details below.

(i) A sample set of sites is drawn and the *true* parameters are chosen.

(ii) Accidents in the period before treatment are generated.

(iii) The parameters of the Poisson-gamma hierarchical generalized linear model proposed in chapter 2 are estimated.

(vi) Hot spots are targeted from the method proposed in chapter 3.

(v) The accident hot spots are treated, and accidents for the period after treatment are generated.

(vi) The safety effect of the treatment is estimated using the method proposed in chapter 4.

(vii) Steps (iii)-(vi) are repeated for the model and methods developed by the Danish Road Directorate (RD), described in chapter 5.

The structure of the program is depicted in figure E.2. The details of the steps (i)-(vii) are:

Ad. (i) A sample set of 100 sites is drawn randomly from the base dataset of all 3-arm signal controlled junctions. The sample dataset contains actual traits, $\mathbf{Z}$, and reported accidents for the respective observation periods ranging from 1996-98 to 1994-98. The dispersion parameter, α, is determined using the so-called method of moments (see appendix A) and the fixed effect parameters, $\boldsymbol{\beta}$, are chosen as the solution to a generalized linear model (GLIM) with a Poisson error distribution. For each site and year in the observation period, the reference safety level, $\mu = g^{-1}(\mathbf{Z}\boldsymbol{\beta})$, is calculated from (A.5). Furthermore, given a chosen preferred proportion of sites identified as hot spots, σ, (in this case 10%) and the value of α, the critical level of dispersion, c, is calculated as the $1-\sigma$ quantile in the gamma distribution with shape and scale parameter

[1] The data analysis was not completed at the time the simulation studies were started.

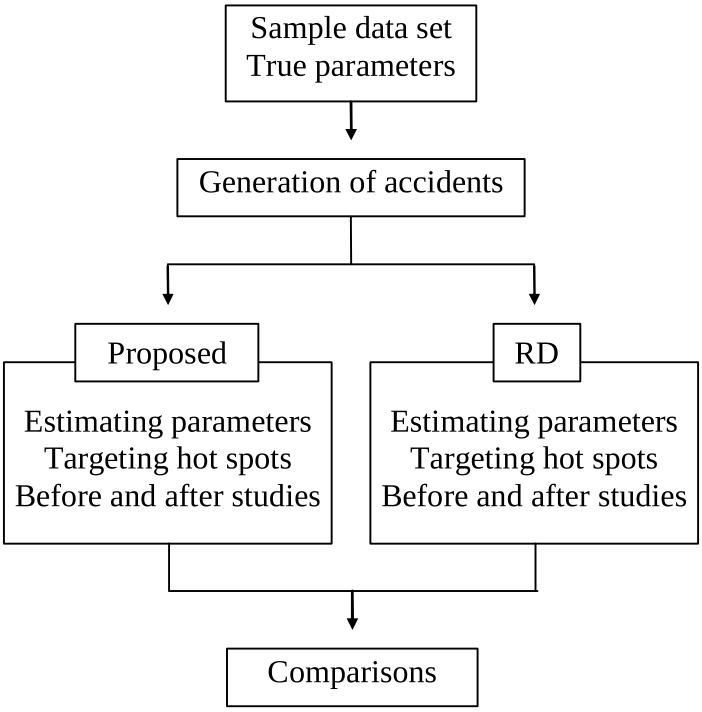


Figure E.2: The structure of each simulation.

both equal to α. The effect of treating a site, ε, is chosen to be 0.3[2]. The dispersion effects, $\mathbf{s}$, are drawn from the gamma distribution with shape and scale parameters both equal to α, and the hot spots are sites with $s > c$. Let n denote the number of hot spots. The site safety level, $\lambda = \mu s$, for each site and year is calculated. Below, the values of $\alpha, \boldsymbol{\beta}, \boldsymbol{\mu}, \mathbf{s}$ and $\boldsymbol{\lambda}$ are denoted the true model values.

Ad. (ii) Accidents in the period before treatment, $\mathbf{x}$, are drawn at random from the Poisson distribution with parameters $\boldsymbol{\lambda}$.

Ad. (iii) The parameters of the Poisson-gamma hierarchical generalized linear model proposed in chapter 2 are estimated from the traits, $\mathbf{Z}$, and the generated accidents, $\mathbf{x}$, using the estimation algorithm set up in appendix A.

Ad. (iv) From the estimated dispersion parameter, $\widehat{\alpha}$, and reference safety levels, $\widehat{\boldsymbol{\mu}}$, the evidence of hotness (EOH) for each site is calculated from (3.3), and the n sites with the highest EOH are targeted as hot spots (see chapter 3).

[2] A preliminary study showed the chosen value of ε had no effect on the relative performance results.

Ad. (v) All sites targeted in (iv) are treated, and for simplicity it is assumed that the treatments do not alter the values of the traits, $\mathbf{Z}$, but replicate themselves in the period after treatment. Consequently, the traits at a site in the first year after treatment are equal to the traits in the first year of the observation period before treatment[3] etc. This causes the reference safety levels, $\boldsymbol{\mu}$, in the period after treatment to stay unchanged, and changes in safety due to the treatment are only reflected in the dispersion effects, $\mathbf{s}$. The true dispersion effects of the accident hot spots are adjusted accordingly, i.e. if a site i is targeted in (iv), then s_i is multiplied by $1 - \varepsilon$, and accidents for this period are generated as in (ii) but with the updated dispersion effects.

Ad (vi) The effect of treatment is calculated using the method proposed in chapter 4. Because the reference safety levels are unchanged from the period before to the period after treatment, the effect may be estimated directly from the percentage change in the estimated dispersion effects, i.e. $\widehat{\varepsilon} = 1 - \widehat{s}_{after}/\widehat{s}_{before}$. In addition, an alternative estimate of the effect of treatment is calculated using the rate between reported accidents after treatment and the corresponding estimated site safety without treatment in this period (in line with Hauer (1997)). One of the side effects of replicating the traits of the period before treatment in the period after treatment is that the length of the period after treatment varies as in the period before treatment.

Ad (vii) The steps (iii)-(vi) are repeated for the model and methods developed by the Danish Road Directorate (RD), described in chapter 5. However, in order for the proposed and RD models to be based on the same set of traits, the RD model is also disaggregated on sub-periods of one year. From the sub-set of sites, where the probability of an accident count above the reported number is less than 10%, RD targets the n sites with the highest accident counts as hot spots (see appendix C). The threshold value of 10% is a relaxation of the 5% level of significance used by the Road Directorate. The purpose of this relaxation is to be able to target the same number of sites as above. Again, all accident hot spots are treated and their dispersion effects adjusted accordingly. A set of accidents in the period after treatment is generated in such a way, that if the proposed and the RD targeting method have targeted the same site, accident counts at this site in the period after treatment is the same in both cases. The effect of treatment is estimated using (5.4) of chapter 5 with an estimated regression to the mean effect of 25%[4].

[3] This is not possible in reality because at least the trait representing trend in time will change. However, for the purposes of this simulation it is of no importance.

[4] A preliminary study showed the chosen value of RTM in the range 20-30% had no effect on the relative performance results.

For each sample of sites, the simulation is repeated 100 times in order to check the sensitivity of the models and methods towards random variation in accidents. A total of 100 samples of sites are drawn to check robustness towards different samples.

E.2 Results

The overall results of the simulation study described above are listed in table E.1. The results are the average values of all samples and simulations.

	3-arm			**4-arm**		
	Proposed	RD	Hauer	Proposed	RD	Hauer
$Mse\left(\frac{1}{\widehat{\alpha}}\right)$	0.01			0.02		
$Mse\left(\widehat{\beta}\right)$	0.95	0.94		0.91	0.91	
$Mse\left(\widehat{\lambda}\right)$	0.10	0.14		0.24	0.89	
$Mse\left(\widehat{s}\right)$	0.11	0.33		0.21	0.51	
$Mean\left(s\right)$	1.40	1.39		1.83	1.80	
$Sensitivity$	0.37	0.36		0.49	0.44	
$Specificity$	0.93	0.93		0.94	0.94	
$\widehat{\varepsilon}$	0.21	0.55	0.23	0.32	0.41	0.25
$Mse\left(\widehat{\varepsilon}\right)$	0.03	0.13	0.27	0.04	0.11	0.28

Table E.1: Overall results of the simulation study.

Because the variance of the dispersion effects, $\widehat{s}$, is $1/\alpha$ at intersections, the estimation error of $1/\alpha$ is of more interest than of α itself. In table E.1, $\widehat{\lambda}$ denotes the estimated site safety level. However, in the models currently at use on state and regional roads, the estimated site safety level is denoted $\widehat{\mu}$. The $Mean\left(s\right)$ in table E.1 is the average value of the true dispersion effects of sites targeted as hot spots. Hence, the value of $Mean\left(s\right)$ reflects the average level of hotness at accident hot spots.

The models and methods proposed in chapters 2-4 outperform the model and methods currently used by the Road Directorate and the regional authorities in almost all phases of hot spot safety work. The results of the simulation study are described in detail below.

E.2.1 Model results

In general, the values of the estimated fixed effect parameters, $\widehat{\beta}$, change very little between the RD and the proposed model. This is in line with earlier studies for intersections (see e.g. Maycock and Hall (1984) and Kulmala (1995)). The RD model is a few percent better at estimating $\widehat{\beta}$, but the mean squared error of the estimated site safety level is considerably larger using the RD model. This is illustrated in figure E.3 for 3-arm signal controlled junctions.The RD model is particularly poor

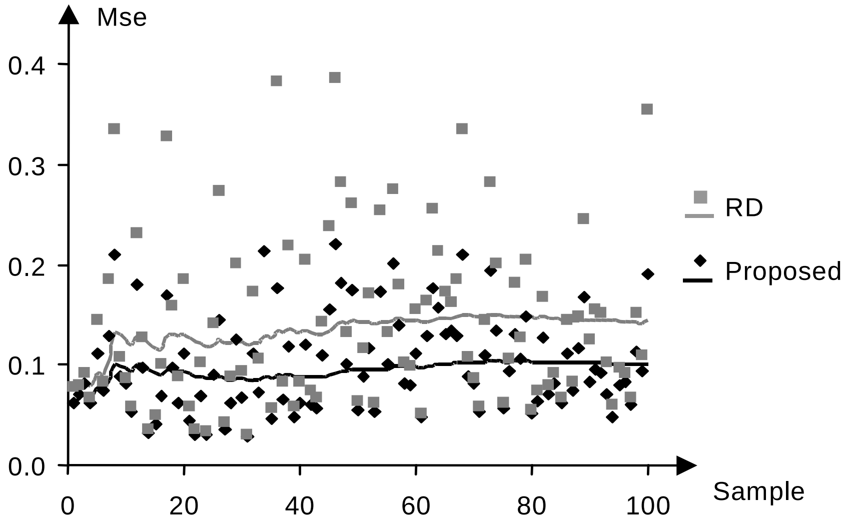


Figure E.3: Mean squared error of the estimated site safety for 3-arm signal controlled junctions. The lines indicate the averages of the previous samples.

in estimating site safety for 4-arm signal controlled junctions, where the variation in accident counts is larger. The inclusion of the dispersion effect, s, in the Poisson-gamma hierarchical generalized linear model is the reason why this proposed model outperforms the RD model. The dispersion effect accounts for site-specific conditions not described by the traits (see figure 2.1 of chapter 2). Figure E.4 illustrates the mean squared error of the estimated dispersion effect for 3-arm signal controlled junctions. The RD estimate of dispersion is the ratio between accident counts and the corresponding estimated level of safety, $x/\widehat{\mu}$.The error of the RD estimate of dispersion is not only larger than the proposed estimate, it also varies more between repeated simulations of the same sample.

E.2.2 Targeting results

None of the targeting methods are very robust towards random variation in accidents, and in a few cases the sensitivity varies between 0% and 90% in the simulations of the same sample. The sensitivity also varies considerably between the samples. Figure E.5 illustrates the sensitivity of the proposed and RD targeting methods for 4-arm signal controlled junctions. It appears that the sensitivity of the proposed method in

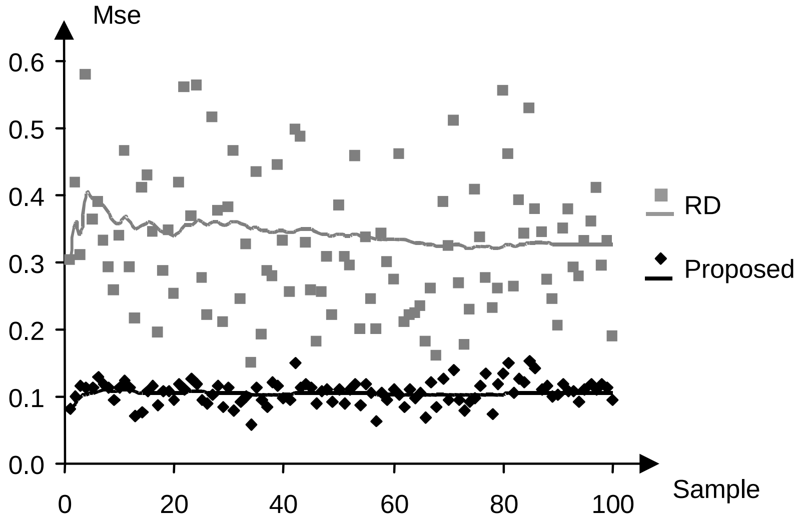


Figure E.4: Mean squared error of the estimated dispersion effect for 3-arm signal controlled junctions. The lines indicate the averages of the previous samples.

chapter 3 is higher than in the RD method.The specificity of the methods is almost the same and varies very little between samples. In addition, the true level of hotness, s, at accident hot spots is marginally higher for hot spots targeted from the proposed method. In other words, the sites targeted under the proposed method do on average have a higher potential for accident reduction.

An additional simulation study was performed to investigate the difference between targeting hot spots from the level of hotness, $\widehat{s}$, and from the evidence of hotness, $\widehat{P}$. For intersections, the methods targeted the same group of sites and consequently resulted in the same sensitivity and specificity. On road sections, however, the methods did not target the exact same group of sites (about 80% overlap). The method using the evidence of hotness, $\widehat{P}$, resulted in a few percent higher sensitivity than the method using the level of hotness, $\widehat{s}$. In addition, the true level of hotness, s, at sites targeted from the evidence of hotness was on average a few percent higher. The specificity was the same for both methods.

E.2.3 Before and after studies results

The effect estimate of treatment used by the Road Directorate and regional authorities gravely overestimates the effect of treatment, as illustrated in figure E.6 for 3-arm signal controlled junctions.The mean squared error of estimation is larger and varies more for the RD estimated effect of treatment, as illustrated in figure E.7.

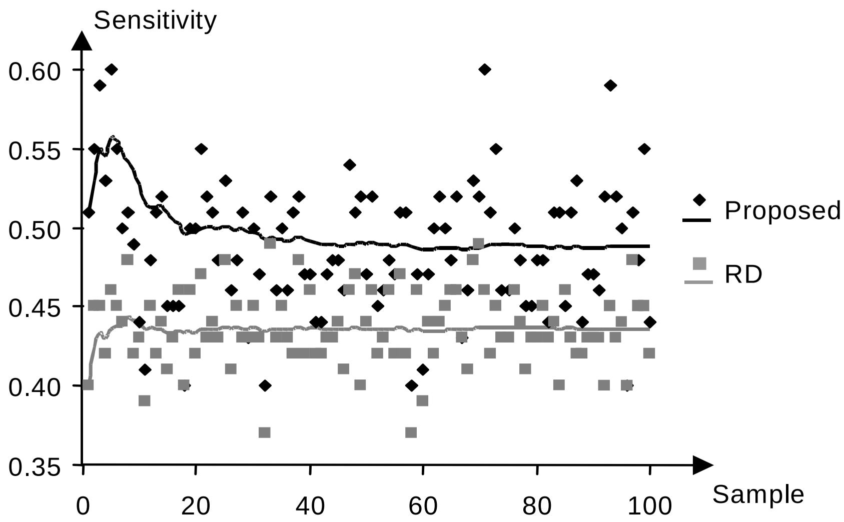


Figure E.5: Sensitivity of the targeting methods for 4-arm signal controlled junctions. The lines indicate the averages of the previous samples.

For 3-arm signal controlled junctions, the method proposed by Hauer (1997)[5] does on average give estimates of safety, which are closer to the true effect of treatment than the method proposed in chapter 4. However, for both 3-arm and 4-arm signal controlled junctions, the Hauer method has the highest mean squared error of estimation of all the methods considered in the study. The reason for this result is that the method proposed in Hauer (1997) results in larger variation in the effect estimates[6].

[5]See chapter 4 for a description of the method used by Hauer.

[6]This is in line with the empirical Bayes trade-off between bias and variance of the estimator (see appendix A).

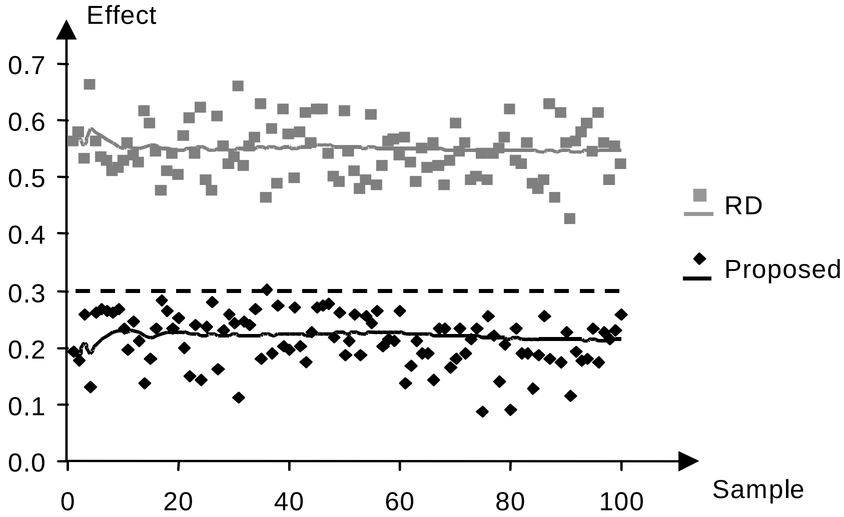


Figure E.6: Estimated effect of treatment of the targeting methods for 3-arm signal controlled junctions. The lines indicate the averages of the previous samples.

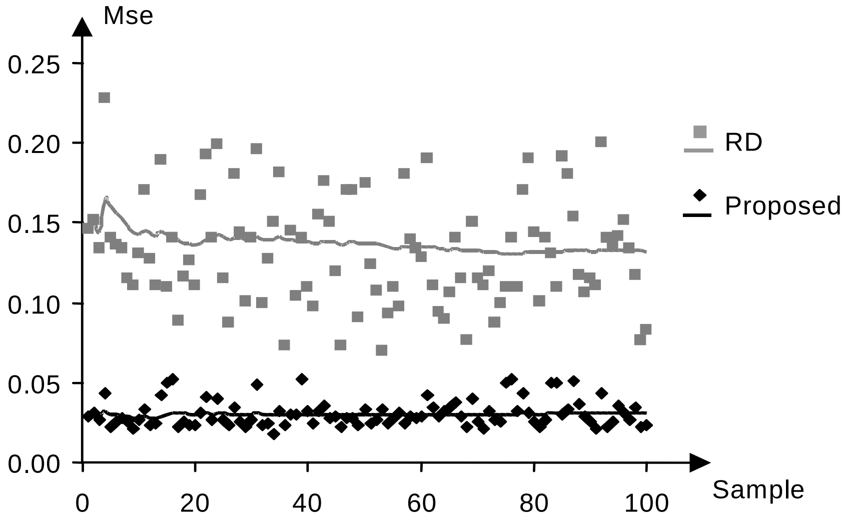


Figure E.7: Mean squared error of estimation of the effect of treatment for 3-arm signal controlled junctions. The lines indicate the averages of the previous samples.

Bibliography

[1] Aagaard, P.E. (1995) *Metoder til valg af reguleringsform for vejkryds.* IVTB-rapport 77, IFP, DTU, Denmark.

[2] Aagaard, P.E. (1997) Management of uncertainty in the safety evaluation of individual risk. *Trafikdage på Aalborg Universitet 1997*, Carl Bro as, Denmark.

[3] Abbess, C., D. Jarrett and C.C. Wright (1981) Accidents at black spots: estimating the effectiveness of remedial treatment, with special reference to the 'regression-to-mean' effect. *Traffic Engineering + Control*, **22**(10), 535-542.

[4] Abdel-Aty, M.A. and A.E. Radwan (2000) Modeling traffic accident occurrence and involvement. *Accident Analysis and Prevention*, **32**(5), 633-642.

[5] Adams, J.G.U. (1988) Evaluating the effectiveness of road safety measures. *Traffic Engineering + Control*, **29**(6), 344-352.

[6] Akaike, H. (1983) Information measures and model selection. *Bulletin of the International Statistical Institute. Proceedings of the 44th Session*, **1**, 277-290.

[7] Andersen, P.K., Ø. Borgan, R.D. Gill and N. Keiding (1993) *Statistical models based on counting processes.* Springer-Verlag, New York Inc.

[8] Ashton, W.D. (1966) *The theory of road traffic flow.* King's College, London, England.

[9] Assum, T., T. Bjørnskau, S. Fosser and F. Sagberg (1999) Risk compensation - the case of road lighting. *Accident Analysis and Prevention*, **31**(5), 545-553.

[10] Austin, K. (1995a) The identification of mistakes in road accident records: part1, locational variables. *Accident Analysis and Prevention*, **27**(2), 261-276.

[11] Austin, K. (1995b) The identification of mistakes in road accident records: part2, casualty variables. *Accident Analysis and Prevention*, **27**(2), 277-282.

[12] Bernardo, J. M. and A. F. M. Smith (1994) *Bayesian theory*. John Wiley & Sons, Chichester, England.

[13] Bernhardt, K.L.S. and M.R. Virkler (2002) Improving the identification, analysis and correction of high-crash locations. *ITE Journal*, January, 38-42.

[14] Box, G.E.P. (1976) Science and statistics. *Journal of the American Statistical Association*, **71**(355), 791-799.

[15] Boyle, A.J. and C.C. Wright (1984) Accident 'migration' after remedial treatment of accident blackspots. *Traffic Engineering + Control*, **25**(5), 260-267.

[16] Breslow, N.R. and D.G. Clayton (1993) Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association*, **88**(421), 9-25.

[17] Carlin, B.P. and T.A. Louis (1996) *Bayes and empirical Bayes methods for data analysis*. Chapman & Hall, Great Britain.

[18] Cox, D.R. and V. Isham (1992) *Point processes. Monographs on applied probability and statistics 12*. Chapman & Hall. England.

[19] Dasgupta, A.K. and D.W. Pearce (1972) *Cost-benefit analysis, theory and practice*. Macmillian, England.

[20] Davis, C.E. (1986) Regression to the mean. *Encyclopedia of statistical sciences*, **7**, 706-708. John Wiley and Sons, Inc.

[21] Dudewicz, E.J. and J.O. Koo (1982) *The complete categorized guide to statistical selection and ranking procedures*. American Sciences Press, Columbus OH.

[22] Efron, B. and C. Morris (1973) Stein's estimation rule and its competitors - an empirical Bayes approach. *Journal of the American Statistical Association*, **68**(341), 117-130.

[23] Elvik, R. (1997) Evaluations of road accident blackspot treatment: a case of the iron law of evaluation studies? *Accident Analysis and Prevention*, **29**(2), 191-199.

[24] Elvik, R., A.B. Mysen and T. Vaa (1997) *Trafikksikkerhetshåndbok*. Transportøkonomisk institut (TØI), Oslo, Norway.

[25] Elvik, R. and A.B. Mysen (1999) Incomplete accident reporting, meta-analysis of studies made in 13 countries. *Journal of the Transportation Research Board*. Transportation Research Record 1665, 133-139. Transportation Research Board, Washington D.C.

[26] Evans, L. (1994) *The 5th Westminster lecture on transport safety.* St Thomas' Hospital, Westminster, London.

[27] Fletcher, R.H., S.W. Fletcher and E.H. Wagner (1996) *Clinical epidemiology: the essentials.* Lippincott Williams & Wilkins.

[28] Færdselssikkerhedskommissionen (2000) *Hver ulykke er én for meget. Trafiksikkerhed starter med dig. Mod nye mål 2000-2012.* Trafikministeriet for Færdselssikkerhedskommissionen, Denmark.

[29] Grimmer, M., R.D. Hall and D. Pickering (1986) *Accidents at rural T-junctions.* Research report 65, Transport and Road Research Laboratory, Berkshire.

[30] Gupta, S.S. and J.C. Hsu (1980) Subset Selection Procedures With Application to Motor Vehicle Fatality Data in a Two-Way Layout. Technometrics, **22**(4), 543-546.

[31] Harville, D.A. (1991) Comment on "That BLUP is a good thing: the estimation of random effects". *Statistical Science*, **6**(1), 35-39.

[32] Hauer, E. (1983) An application of the likelihood/Bayes approach to the estimation of safety countermeasure effectiveness. *Accident Analysis and Prevention*, **15**(4), 287-298.

[33] Hauer, E. and B.N. Persaud (1984) Problem of identifying hazardous locations using accident data. *Journal of the Transportation Research Board.* Transportation Research Record 975, 36-43. Transportation Research Board, Washington D.C.

[34] Hauer, E. and B.N. Persaud (1987) How to estimate the safety of rail-highway grade crossings and the safety effect of warning devices. *Journal of the Transportation Research Board.* Transportation Research Record 1114, 131-140. Transportation Research Board, Washington D.C.

[35] Hauer, E. (1992) A note on three estimators of safety effect. *Traffic Engineering + Control*, **33**(6), 388-393.

[36] Hauer, E. (1997) *Observational before-after studies in road safety.* Pergamon, Oxford.

[37] Hauer, E. (2001) Overdispersion in modelling accidents on road sections and in empirical Bayes estimation. *Accident Analysis and Prevention*, **33**(6), 799-808.

[38] Heydecker, B.G. and J. Wu (1991) *Using the information in road accident records.* Transport Studies Group, University College London, England.

[39] Heydecker, B.G. and J. Wu (1993) A knowledge-based system for road accident remedial work. *Computing Systems in Engineering.* **4**(2-3), 337-348.

[40] Heydecker, B.G. and J. Wu (2001) Identification of sites for road accident remedial work by Bayesian statistical methods: an example of uncertain inference. *Advances in Engineering Software.* **32**(10-11), 859-869.

[41] Jarret, D.F., K. Lupton and C.C. Wright (1997) *The consistency of road accident variables in Great Britain: 1995.* Middlesex University.

[42] Jones-Lee, M.W. (1990) The value of transport safety. *Oxford review on economic policy*, **6**, 39-60.

[43] Jorgensen, R.E. (1966) *Evaluation of criteria for safety improvements on the highway.* Westat Research Analysts Inc., Gaithersburg, Maryland, USA.

[44] Jørgensen, B. (1997) *The theory of dispersion models.* Chapman & Hall Ltd, London.

[45] Kulmala, R. (1995) *Safety at rural three- and four-arm junctions.* Technical Research Centre of Finland (VTT), Espoo, Finland.

[46] Lahrmann, H. and S. Leleur (1994) *Vejtrafik, trafikteknik og trafikplanlægning.* Polyteknisk forlag, Denmark.

[47] Lauritsen, J.M. (2001) Registrering af trafikskader på sygehuse baseret på geografisk information (elektronisk kort). *Trafikdage på Aalborg Universitet 2001*, Fyns Amt, Denmark.

[48] Lee, P.M. (1994) *Bayesian statistics: an introduction.* University of York, England.

[49] Lee, Y. and J.A. Nelder (1996) Hierarchical generalized linear models. *Journal of the Royal Statistical Society B*, **58**(4), 619-678.

[50] Lee, Y. and J.A. Nelder (2001) Hierarchical generalized linear models: a synthesis of generalized linear models, random-effect models and structured dispersions. *Biometrika*, **88**(4), 987–1006.

[51] Maher, M.J. (1990) A bivariate negative binomial model to explain traffic accident migration. *Accident Analysis and Prevention*, **22**(5), 487-498.

[52] Maher, M.J. and I. Summersgill (1996) A comprehensive methodology for the fitting of predictive accident models. *Accident Analysis and Prevention*, **28**(3), 281-296.

[53] Maycock, G. and R.D. Hall (1984) *Accidents at 4-arm roundabouts*. Laboratory report 1120, Transport and Road Research Laboratory, Berkshire, England.

[54] Maycock, G. and M.J. Maher (1988) *Generalized linear models in the analysis of road accidents*. Working paper, Transport and Road Research Laboratory, Berkshire, England.

[55] McCullagh, P. and J.A. Nelder (1989) *Generalized linear models*. Chapman & Hall, Great Britain.

[56] McGuigan, D.R.D. (1981) The use of relationships between road accidents and traffic flow in "black-spot" identification. *Traffic Engineering + Control*, **22**(8-9), 448-453.

[57] Melchers, R.E. (2001) On the ALARP approach to risk management. *Reliability Engineering and System Safety*, **71** (2), 201-208.

[58] Morris, C.N. (1983) Parametric empirical Bayes inference: theory and applications. *Journal of the American Statistical Association*, **78**(381), 47-54.

[59] Mountain, L., M. Maher and B. Fawaz (1998) The influence of trend on estimates of accidents at junctions. *Accident Analysis and Prevention*, **30**(5), 641-649.

[60] Nelder, J.A. and R.W.M. Wedderburn (1972) Generalized Linear Models. *Journal of the Royal Statistical Society A*, **135**, 370 -384.

[61] Nicholson, A. (1985) The variability of accident counts. *Accident Analysis and Prevention*, **17**(1), 47-56.

[62] Nicholson, A. and Y.D. Wong (1993) Are accidents Poisson distributed? A statistical test. *Accident Analysis and Prevention*, **25**(1), 91-97.

[63] OECD (1997) *Road transport research. Road safety principles and models: review of descriptive, predictive, risk and accident consequence models*. OECD, Paris, France.

[64] Persaud, B.N. (1986) Relating the effect of safety measures to expected number of accidents. *Accident Analysis and Prevention*, **18**(1), 63-70.

[65] Persaud, B.N. (1994) Accident prediction models for rural roads. *Canadian Journal of Civil Engineering*, **21**(4), 547-554.

[66] Persaud, B.N. and A. Kazakov (1994) A procedure for allocating a safety improvement budget among treatment sites. *Accident Analysis and Prevention*, **26**(1), 121-126.

[67] Persaud, B.N., C. Lyon and T. Nguyen (1999) Empirical Bayes procedure for ranking sites for safety investigation by potential safety improvements. *Journal of the Transportation Research Board*. Transportation Research Record 1665, 7-12. Transportation Research Board, Washington D.C.

[68] Robert, C.P. (1994) *The Bayesian choice, a decision-theoretic motivation*. Springer-Verlag, New York.

[69] Robinson, G.K. (1991) That BLUP is a good thing: the estimation of random effects. *Statistical Science*, **6**(1), 15-51.

[70] Satterthwaite, S. P. (1976) *Day to day variations in road accident frequency in Britain*. Traffic Studies Group, University College London, England.

[71] Schall, R. (1991) Estimation in generalized linear models with random effects. *Biometrika*, **78**(4), 719-728.

[72] Schall, T. and G. Smith (2000) Do baseball players regress toward the mean? *The American Statistician*. **54**(4), 231-235.

[73] Schlüter, P.J., J.J. Deely and A.J. Nicholson (1997) Ranking and selecting motor vehicle accident sites by using a hierarchical Bayesian model. *The Statistician*, **46**(3), 293-316.

[74] Solomon, F. (1987) *Probability and stochastic processes*. Prentice-Hall, New Jersey.

[75] Swets, J.A. and R.M. Pickett (1992) *Evaluation of diagnostic systems: Methods from signal detection theory*. Academic Press, New York.

[76] T10 (2000) *Notes for the unit T10 of the Intercollegiates MSc course in transport*. University College London, England.

[77] Transportforskningsudvalget (1968) *Fynsundersøgelsen om trafikuheld*. Transportforskningsudvalget, Denmark.

[78] Tsutakawa, R.K. (1988) Mixed models for analyzing geographic variability in mortality rates. *Journal of the American Statistical Association*. **83**(401), 37-42.

[79] Tunaru, R. (1999) Hierarchical Bayesian models for road accident data. *Traffic Engineering + Control*, **40**(6), 318-324.

[80] Ulykkes Analyse Gruppen (1997) *Ulykker 1997, tilskadekomne registreret på skadestuen, Odense Universitetshospital*. Odense, Denmark.

[81] Van Duijn, M.A.J. (1993) *Mixed models for repeated count data.* Ph.D. dissertation. Leiden University, Leiden.

[82] Vejdirektoratet (1992) *4-benede ikke-signalregulerede kryds i Fyns og Viborg amter, undersøgelse af 37 kryds.* Trafiksikkerhedsafdelingen - SSV, Vejdirektoratet, Denmark.

[83] Vejdirektoratet (1994) *Cyklisters sikkerhed i byer, trafiksikkerhed og miljø.* Rapport 10, Vejdirektoratet, Denmark.

[84] Vejdirektoratet (1995) *Uheldsmodel for bygader, Del 1: Model for 3- og 4-benede kryds.* Notat 22, Vejdirektoratet, Denmark.

[85] Vejdirektoratet (1996) *Introduktion til VISplet: AP-parametre baseret på 1989-93.* Notat 29, Vejdirektoratet, Denmark.

[86] Vejdirektoratet (1998a) *Uheldsmodel for bygader, Del 2: Model for strækninger.* Notat 59, Vejdirektoratet, Denmark.

[87] Vejdirektoratet (1998b) *Håndbog i lokale trafiksikkerhedsplaner.* Vejdirektoratet, Denmark.

[88] Vejdirektoratet (1999a) *Håndbog i trafiksikkerhedsrevision.* Rapport 189, Vejdirektoratet, Denmark.

[89] Vejdirektoratet (1999b) *Trafiksikkerhedsarbejdet i kommunerne.* Rapport 189, Vejdirektoratet, Denmark.

[90] Vejdirektoratet (1999c) *Trafikøkonomiske enhedspriser.* Rapport 186, Vejdirektoratet, Denmark.

[91] Vejdirektoratet (1999d) *Avanceret signalstyring i signalreguleret kryds.* Notat 61, Vejdirektoratet, Denmark.

[92] Vejdirektoratet (2001a) *Korrektionsfaktorer 1997-2000 som følge af ændring af definition på hjernerystelse i 1997.* Vejdirektoratet, Denmark.

[93] Vejdirektoratet (2001b) *Håndbog i trafiksikkerhedsberegninger. Brug af uheldsmodeller og andre vurderinger.* Rapport 220, Vejdirektoratet, Denmark.

[94] Vejdirektoratet (2002a) *Indeks af udviklingen i biltrafikken.* Rapport 241, Vejdirektoratet, Denmark.

[95] Vejdirektoratet (2002b) *www.vd.dk.*

[96] Vejsektoren (2002) *www.vejsektoren.dk.*

[97] Wilde, G.J.S. (1986) Beyond the concept of risk homeostasis: suggestions for research and application towards the prevention of accidents and lifestyle-related disease. *Accident Analysis and Prevention*, **18**(5), 377-401.