Final Report 2002-23

Research

Development and Testing of a Vehicle/Pedestrian Collision Model for Neighborhood Traffic Control

Minnesota Local Road Research Board

Technical Report Documentation Page

1. Report No.	2.	3. Recipients Accession No.
MN/RC – 2002-23		

4. Title and Subtitle	5. Report Date
DEVELOPMENT AND TESTING OF A VEHICLE/PEDESTRIAN COLLISION MODEL FOR NEIGHBORHOOD TRAFFIC CONTROL	February 2002
	6.
7. Author(s)	8. Performing Organization Report No.
Gary A. Davis, Kate Sanderson, Sujay Davuluri	
9. Performing Organization Name and Address	10. Project/Task/Work Unit No.
Department of Civil Engineering University of Minnesota 500 Pillsbury Drive S.E. Minneapolis, MN 55455	11. Contract (C) or Grant (G) No. C) 74708 wo) 93
12. Sponsoring Organization Name and Address	13. Type of Report and Period Covered
Minnesota Department of Transportation 395 John Ireland Boulevard Mail Stop 330 St. Paul, Minnesota 55155	Final Report 1998-2001
	14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract (Limit: 200 words)

This report presents an approach to assess the effect of vehicle traffic volumes and speeds on pedestrian safety. It shows that the probability of standardized pedestrian conflict resulting in a collision can be computed given data on the distribution of vehicle speeds and headways on a residential street.

Researchers applied this method to data collected on a sample of 25 residential streets in the Twin Cities and found that collision rates varied between four and 64 collisions per 1,000 pedestrian conflicts, depending primarily on the street¹s traffic volume. Using a model that relates the impact speed of a vehicle to the severity of pedestrian injury, they computed the probabilities of a severe collision. Sensitive to both traffic volume and traffic speed, the severe collision rate varied between one and 25 collisions between 1,000 conflicts.

Using the same data, researchers also computed the crash reduction factor, used to assess the potential safety effect of a 25 miles per hour speed limit on the sample of residential streets. The estimated crash reductions ranged between .2 and 45 percent, depending primarily on the degree to which the vehicle speeds currently exceeded 25 miles per hour. Researchers also showed how this computation assists with the reconstruction of actual vehicle/pedestrian collisions.

17. Document Analysis/Descriptors		18. Availability Statement
Bayesian networks	Crash reduction factor	No restrictions. Document available from:
Markov Chain Monte Carlo	Traffic calming	National Technical Information Services,
Pedestrian crashes	Neighborhood traffic control	Springfield, Virginia 22161

19. Security Class (this report)	20. Security Class (this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	92	

DEVELOPMENT AND TESTING OF A VEHICLE/PEDESTRIAN COLLISION MODEL FOR NEIGHBORHOOD TRAFFIC CONTROL

Final Report

Prepared by

Gary A. Davis
Associate Professor and Principal Investigator

Kate Sanderson
Research Assistant

Sujay Davuluri
Research Assistant

Department of Civil Engineering
University of Minnesota
500 Pillsbury Drive SE
Minneapolis, MN 55455

February 2002

Published by

Minnesota Department of Transportation
Office of Research Services
395 John Ireland Boulevard
St. Paul, MN 55155

TABLE OF CONTENTS

List Of Figures

LIST OF TABLES

EXECUTIVE SUMMARY

The standard model for programming traffic safety improvements uses estimated crash rates to identify potential high-hazard sites, and estimated crash reduction factors to select the most effective safety projects. When one needs to consider neighborhood traffic control measures on local or residential streets, however, the standard model is not applicable because of (a) difficulties in obtaining routine, site-specific measures of pedestrian exposure, and (b) the sparsity of pedestrian collisions, especially on local streets. Similar problems are encountered when attempting to program roadside safety improvements, and the recommended solution is to simulate the effect of a proposed improvement by coupling a deterministic model of a run-off-road event with probability distributions on the model's variables.

This report illustrates how a similar approach can be used to assess the effect of vehicle traffic volumes and speeds on pedestrian safety. In particular, it is shown how analogue of the estimated crash rate, the probability a standardized pedestrian conflict results in a collision, can be computer given data on the distribution of vehicle speeds and headways on a residential street. This method was applied to data collected on a sample of 25 residential streets in the Twin Cities, and it was found that the collision rates varied between 4 and 64 collisions per 1000 pedestrian conflicts depending primarily on the street's traffic volume. Next, by developing a model relating the impact speed of a vehicle to the severity of injury suffered by the pedestrian, it WAS shown how to compute the probabilities of a severe collision. The severe collision rate varied between 1 and 25 collisions per 1000 conflicts and was sensitive to both traffic volume and traffic speed.

It was also shown how an analogue of the crash reduction factor, called the "Probability of Necessity", can be computed using the same data. This was used to assess the potential safety effect of 25 mph speed limit on the sample residential streets. The estimated crash reductions ranged between 0.2% and 46% depending primarily on the degree to which vehicle speeds currently exceeded 25 mph. Finally it was shown how computation of "Probabilities of Necessity" could also be incorporated in the reconstruction of actual vehicle/pedestrian collisions to assess the degree to which the vehicle might have been speeding and if speeding could be considered a causal factor in the collision. This method was applied to a sample of 8 collisions

occurring in 60 km/h (37 mph) in Adelaide and Australia; and it was found that speeding could be considered a casual factor in about half of these.

The probability calculations were carried out using Monte Carlo simulation. For the collision probabilities, standard Monte Carlo methods are adequate and implementation using stand-alone programs or spread sheets should be relatively straightforward. It is recommended that this method could become a standard analysis tool for traffic engineers. For computing probabilities of necessity however, more sophisticated Markov Chain Monte Carlo methods are needed and these can be difficult to use correctly absent a solid understanding of probability and statistics. It is recommended that these methods be limited to special studies possibly conducted by qualified consultants or specially trained personnel.

CHAPTER ONE

SETTING THE STAGE

In this report, "Neighborhood Traffic Controls" (NTC) will offer to engineering modifications applies to local and residential streets in order to reduce traffic volumes or speeds. Such modifications might include regulatory actions, such as creation of reduced speed zones or installation of stop signs and physical changes to the roadways such as construction of speed humps, traffic circles, or cul-de-sacs or more innovative actions such as the use of photo enforcement. These modifications include those traditionally known as "traffic calming". Although requests for NTC usually originate from local resident (Wallwork 1993), the use of NTC can become highly politicized with organizations such as the National Motorists Association opposing measures which "… inconvenience and hinder the legitimate travel of responsible motorist", to the extent of providing advices on how to organize opposition to NTC measures (NMA 2001). As with most traffic engineering actions, NTC must be physically realized on particular streets and traffic conditions on some streets may make them more promising candidates for NTC than others. Pedestrian safety, especially the safety of child pedestrian, is often mentioned as a reason for applying NTC but, at present, there is no commonly accepted procedure for rationally assessing the risk of pedestrians generated by vehicle traffic volumes and speeds. In the remainder of this chapter, we will review the problem of relating vehicle traffic speed and volume to pedestrian safety on local of residential streets, while the rest of this report will describe a procedure for making just such assessments.

A basic distinction central to highway and traffic engineering practice is that of the four functional classes of roads: freeways, arterial, collectors and local streets (Minnesota Department of Transportation (Mn/DOT) 1996a, p.2-5 (3)). Further, a road is said to provide high mobility when it can support a high volume of traffic moving at high speeds, while it supports high access when vehicles can move directly from the road to most, if not all, adjoining land uses. It is generally recognized that these two features conflict with each other with no road design providing simultaneously high mobility and access. As one moves from left to right in the least functional classes, one roughly goes from facilities high mobility but low access, to facilities providing high mobility but low access and to facilities providing low mobility but high access. The primary function of local streets then is to support access to adjacent land and to provide connections with streets higher in the functional hierarchy. Local streets generally have low speeds and volumes and the use of the street for through movement may be deliberated discouraged (Mn/DOT 1996a, p.2-5 (4). Constraints on vehicle mobility can become even more prominent on local residential streets in urban and suburban areas, where in addition to serving motor vehicles the street might provide the main (or only) corridor for pedestrians and cyclists, and may also serve as an area where neighbors socialize or children play (Homburger et at. 1989). For such streets, it has been recognized that " … the overriding consideration is to foster a safe and pleasant environment. The convenience of the motorists is a second consideration." (AASHTO 1994, p.428).

One can conclude that as vehicle speeds and volumes increase on a local or residential street, the ability of the street to serve these basic functions will be impaired. Figure 1.1 illustrates trends in Minnesota's population, vehicle registration and vehicle-miles of travel over the past 35 years (Mn/DPS 2000). Here we see that while the population has increased by about 40% the number of registered vehicles more than doubled, while the total vehicle-miles of travel more than tripled. Figure 1.1 shows not only a trend of increasing population but also that each person is increasingly more likely to have vehicle and each of these vehicles is likely to travel greater distances. If this trend continues, we can expect increasing congestion and delay on urban arterials resulting in more drivers being tempted to divert onto residential or local streets in pursuit of personally optimal routes. This in turn would lead us to expect increasing pressure to apply to serve its primary functions.

Regulation vehicle-speeds has long been and continues to be a contentious topic (TRB 1998). In Minnesota, the legislated speed limit in urban districts, which include most urban local and residential streets, is 30 mph (50 km/h) but in the early 1990's consideration was given to legislating a 25 mph speed limit on residential streets. This was dropped, at least in part because of concerns about the resulting enforcement problems but also because of uncertainty concerning what, if any, safety benefits might result. The Commissioner of Transportation has the responsibility for determining variances from the legislated speed limits, and consideration of a variance begins with a spot speed study to estimate the street's 85th percentile speed – the speed that is exceeded by only 15% of the vehicles using that street. The recommended speed limit is then determined by rounding the 85-percentile speed to the nearest speed evenly divisible by 5. A posted speed limit 5 mph (10 km/h) below the 85th percentile speed may be justified if "… there is a high accident record involving accidents of a type that would be reduced by enforcement of a lower speed limit." (Mn/DOT 1996b, p.6-40).

The use of the 85th percentile speed in determining speed limit is recommended by the traffic engineering profession (ITE 1993), is supported by motorists' lobbying groups (e.g. NMA 2001) and is used by state and local governments across the United States (Parker 1985). This practice avoids the enforcement dilemma of having either to cite unpopularly large numbers of drivers for speeding or to tolerate widespread violation of speed limits. It is justified by the belief that a majority of drivers on a road are capable of choosing a speed that is "reasonable and prudent" for that road. Although it is not clear what exactly is meant by "reasonable and prudent", the essence of this notion appears to be that most drivers are able to reasonable balance the benefits obtained by the shorter travel times provided by higher speeds against the expected costs associated with higher speeds including the costs of traffic crashes. An important assumption of this view is that all costs and benefits related to a speed choice are economists call "internal" – that is paid to or by the driver alone. In this case, it can be argued that the speed selected by the driver is an optimal one. However, if some of the costs are "external" – that is paid by other road users, then it can be argued that the speed selected by the driver, although optimal for him or her, will be higher than the socially optimal speed which takes into account the welfare of other road users. When we consider crashes between vehicles and pedestrians, physical injuries are born primarily by the pedestrian while other monetary costs of the crashes are often distributed over the customers of the driver's insurance company. Thus on residential streets where an important source of costs associated with higher speeds is that resulting from collision with pedestrians, it would be predicted that vehicle speed will be "over-consumed", with the burden for avoiding collision falling primarily on pedestrian.

Some indirect support for this view comes from studies of driver behavior in the presence of pedestrians. Field studies from different countries have found that in vehicle/pedestrian interactions, a substantial fraction of the drivers will not slow down or stop for pedestrians unless the pedestrian is actually in the vehicle path and that this tendency is more prevalent among drivers traveling at higher speeds (Katz et al. 1975; Thompson et al. 1985). This has been observed even where drivers are required by law to stop for pedestrians (Britt et al. 1995), and despite the fact that in surveys most drivers claim to yield the right-of-way to pedestrians (Varhelyi 1998). Varhelyi also reported a significant tendency for some drivers to increase their speed as a pedestrian approaches a zebra crossing, apparently to intimidate the pedestrian from asserting right-a-way. A recent crackdown on speeding and failure to yield to pedestrians on an arterial in St. Paul, Minnesota, even caught a state legislator driving at 42 mph in a 30 mph zone (Burson 1999).

When significant externalities exist, closing the gap between individual choices and socially optimal choices generally requires some sort of regulatory intervention. Although economic theory indicates that it should in principle be possible to apply an optimal toll or tax to the drivers to cause their individual decisions to match the social optimum, the practical difficulties arising in implementing such schemes are often prohibitive. For the interventions characteristic of traffic engineering, a more tractable alternative to optimal taxation is to optimize a specific measure of effectiveness subject to various constraints. For example, at the intersection of two streets the individual decision that summarizes delay is to enter the intersection as soon as one arrives. This requires though that drivers on other approaches do not do the same thing at the same time because if they do a situation that maximizes each person's delay by closing the intersection with wrecked vehicles, can result. In designing a regulatory scheme for the intersection, the objective could be to minimize total approach delay during peak-period operation subject to constraints on the maximum acceptable delay on minor approaches and on the overall cost of the project. For another example, using the 85th percentile speed for the speed limit on a tangent section leading to the inside lane of a horizontal curve may lead to speeds at which the curve cannot be successfully negotiated. A driver who fails to negotiate the curve and collides with a vehicle traveling the other way imposes costs not only on himself but also on the agencies, which must respond to the crash, on other customers of his insurance company and on the occupants of the other vehicle. A regulatory intervention then may seek to restrict approach speeds to the maximum, which keeps either the frequency of crashes or the probability of a crash below some acceptable level. Note that in both examples, selecting the engineering intervention requires knowledge of how the intervention would be expected to affect the performance measure.

On local or residential streets where a primary source of external costs can be of those resulting from collisions between vehicles and pedestrians, a legitimate regularly objective can be to manage vehicle traffic volume and/or speed so that the risk of a crash to a pedestrian is acceptably low. Determining an appropriate volume or speed thus requires being able to relate to levels of these variables to pedestrian crash risk, which in turn requires that we be specific about what we mean by crash risk. The first thing to note is that crash risk is not the same as crash frequency. To take an extreme example, Davis and Wessels (1999) reported that in Washington State between 1990 and 1995: out of a total of 11,160 vehicle/pedestrian collisions approximately 3% occurred on freeways while 65% occurred on city streets. If risk is equated with crash frequency one would be tempted to conclude that freeways are safer for pedestrians than are urban streets, which most would agree, is false. A given pedestrian is more likely to be struck trying to cross a busy freeway than when crossing a city street and the source of the increased risk is the very much higher traffic volume and speed on the freeway. The reason for this supposed paradox is that the opportunities for vehicle/pedestrian crashes are much lower on freeways because pedestrians are generally prevented from being there. Similarly, a residential street may have no reported pedestrian crashes not because the traffic volume and speed conditions are benign but because pedestrians, in self-defense, never enter the street. Although exclusion of pedestrians is appropriate on freeways, such a situation would generally be viewed as inconsistent with the stated goal "… to foster a safe and pleasant environment" where "The convenience of the motorist is a secondary consideration."

In more standard traffic safety programming, a key measure of crash risk is a location's estimated crash rate:

$$r = \frac{x}{e} \tag{1.1}$$

Where:

r = Estimated crash rate

x = Total reported crashes occurring over some time

e = Total exposure to crashes during that same period of time

The exposure is in turn usually estimated at intersections as total entering vehicles or as total vehicle-miles of travel for road sections. By comparing estimated crash rates for a number of similar locations, those with atypically high estimated crash rates can be identified and designated as potential high-hazard sites. Crash patterns at these high-hazard sites are then studied in more detail together with results from other pertinent engineering studies to identify plausible causes from these crashes. Countermeasures appropriate to these causes can then be identified and the crash reduction expected to occur without the countermeasure by an appropriate crash reduction factor (RF). That is:

$$Reduction = \theta x_p \tag{1.2}$$

Where

x_p = Predicted crash count

θ = Crash reduction factor

The predicted crash count is in turn obtained by applying the estimated crash rate to a predicted level of exposure while the RF is usually an estimate computed from a previously conducted before and after study via the equation.

$$\theta = \frac{r_B - r_A}{r_B} \tag{1.3}$$

Here r_B and r_A respectively denotes crash rates estimated before and after application of the countermeasure (Mn/DOT 1996b, pp. (11-10)–(11-14)). The value of the countermeasure can then be assessed by using the predicted crash reduction in a cost-benefit analysis (Mn/DOT 1996a, p.2-4 (3)).

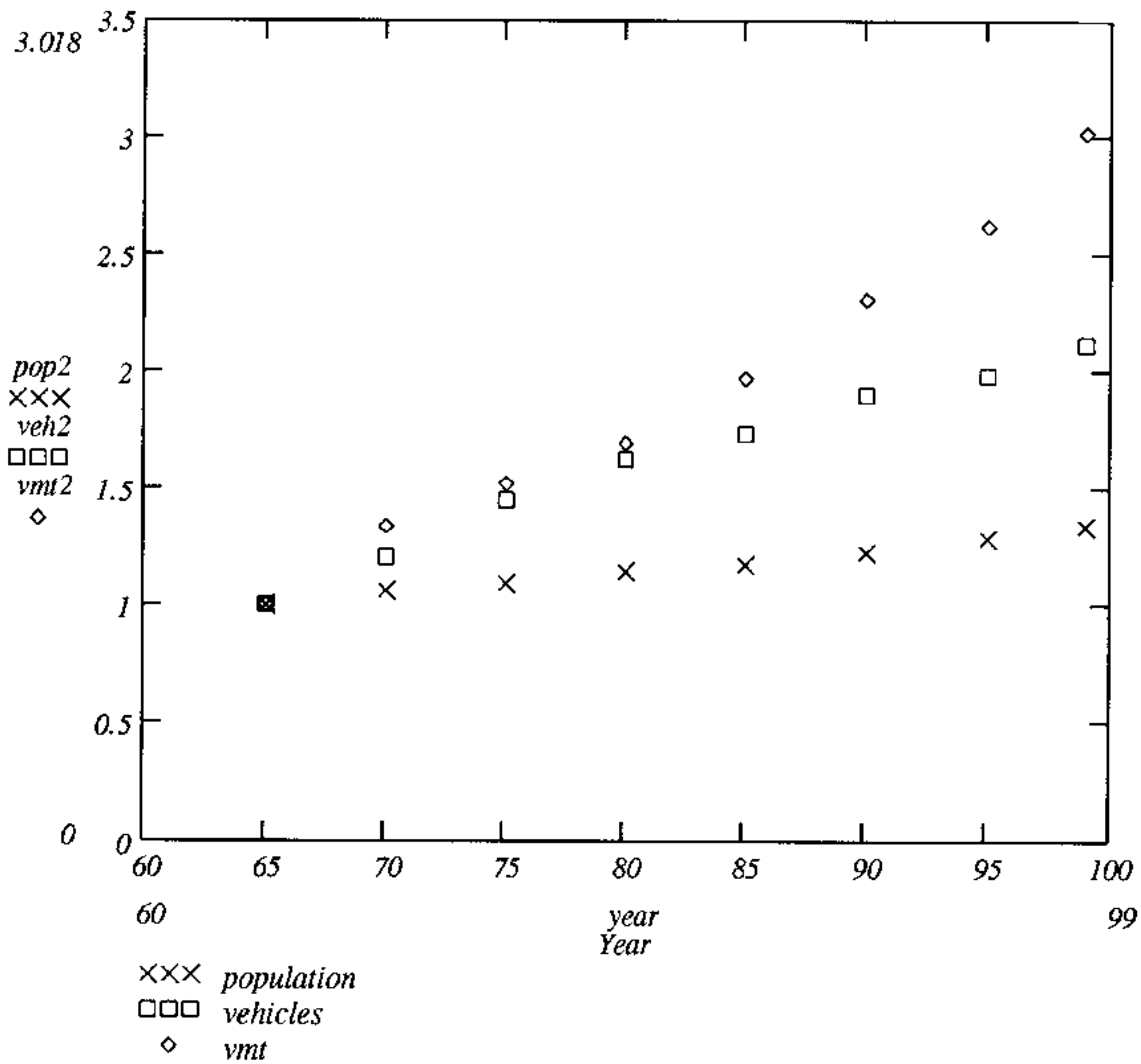
From the above, it is clear that the standard model for programming safety improvements relies heavily on estimated crash rates. However, when we attempt to use this model to program traffic management interventions on local and residential street two difficulties arise. First, although ideally the denominator of the estimate crash rate should be a complete count of entities using the facility, in practice even complete counting of vehicle traffic is not possible for most roads and intersections. Instead, portable traffic counts are used to obtain sample counts to traffic volumes on a facility. These sample counts are used to estimate the facility's average daily traffic (ADT) that in turn is used to compute exposure estimates. For pedestrian activity there is, at present, no cost-effective method for obtaining reliable automatic counts so routine development of pedestrian exposure measures is not possible. Second, using estimates of crash rate to identify potential high-hazard locations works best when crashes tend to be unevenly distributed across locations but for pedestrian crashes on local streets this is not the case. For instance, Hoque and Andreassen (1986) found that the 192 crashes occurring on local streets in Melbourne, Australia during 1981 took place on 189 separate road sections. This means that in one year, the overwhelming majority of local streets had zero or one pedestrian crashes making it difficult to use crash counts to clearly identify high-hazard locations.

Similar sorts of difficulties arise when one attempts to estimate benefits due to roadside safety improvements where the frequency of collisions with any specific object, such as a tree or a bridge abutment, tends to be too low to support reliable estimate of crash rates. An alternative solution is to use a deterministic model to describe whether or not a vehicle strikes a roadside object after encroaching onto the roadside, depending on the shape and location of the object as well as on variables such as the vehicle's speed, point of encroachment, encroachment angle and deceleration rate. By placing probability distributions on these variables, it is then possible, at least in principle, to compute the probability that an encroaching vehicle strikes the object. This method has been implemented in the computer program ROADSIDE (AASHTO 1996) and an enhanced version is current being developed and tested as the Roadside Analysis Program (RSAP) (Mak et al, 1998).

We will use a similar approach, in place of unavailable crash counts, to develop estimates of pedestrian crash rate and crash reduction effects on local and residential streets with structural knowledge of how a crash occur. Chapter 2 will describe development of a model relating the speed of the striking vehicle to the severity of injury suffered by a stuck pedestrian. Chapter 3 will then describe how to estimate pedestrian collision probabilities on local or residential streets and apply the method to a sample of 25 streets from the Twins Cities. Chapter 4 will then describe a similar method for estimating the effect of speeding in individually actually occurred pedestrian/vehicle crashes. Chapter 5 will present our conclusions and recommendations.

Figure 1.1

Proportional Change in Minnesota's Population, Total Registered Vehicles and Estimated Vehicle-Miles of Travel from 1965-1999. (1965 is the Base Year)



CHAPTER TWO

MODELING THE RELATION BETWEEN IMPACT SPEED AND INJURY

The objective of this chapter is to develop a method for predicting the probability distribution across categories of pedestrian injury severity given the vehicle's impact speed. Vilenius et at. (1994) show, at least to the first approximation, that the deceleration experienced by a pedestrian's head when contacting a vehicle is roughly proportional to the speed at which the head is traveling at the time of contact. Since Wood (1988) has shown that this speed is in turn roughly proportional to the speed of the impacting vehicle when the pedestrian is originally stuck, we can hypothesize that the "damage" suffered by the pedestrian in a crash is on average roughly proportional to the impact speed. More formally,

$$Damages = bv + e \tag{2.1}$$

Where

v = Vehicle's speed at impact

b = Average increase in damage for each unit increase in impact speed

e = Random damage due to individual differences

Now suppose that the severity of the injury suffered by the pedestrian can be classed into one of the n ordered categories, defined by n-1 "damage thresholds", denoted by $a_1, a_2, \ldots a_{n-1}$. Injury severity is then determined by the rule:

Severity = 1 if damage $\leq a_1$

i if $a_{i-1} <$ damage $\leq a_i$

n if damage $> a_{n-1}$

For a given impact speed the probability distribution across injury severity categories is then:

$$p\left[severity = 1 \middle| v\right] = p\left[bv + e \leq a_1\right] = p\left[e \leq a_1 - bv\right] = \int_{-\infty}^{a_1 - bv} f(e)\,de$$

$$p\left[severity = i\,|v\right] \quad = \quad \int_{a_{i-1}-bv}^{a_i-bv} f(e)\,de, \qquad i = 2,\ldots,n-1$$

$$p\left[severity = n|v\right] \quad = \quad \int_{a_{n-1}-bv}^{\infty} f(e)\,de \qquad (2.2)$$

Where f (e) denotes the probability density function of the random damage. The model specification is completed by identifying values for the parameters b, a_1, …, a_{n-1}, and a functional form for the probability distribution of the random damage. For example, if the random damages are assumed to be normal we obtain an ordered form of the prohibit model, while if the random damage has a logistic distribution we obtain an ordered logit model.

Since neither the values for the model parameters nor the functional form for the random damage distribution are likely to be know as a *priori,* these must be identified from the data. In principle, measurements of impact speed and injury severity in a random sample from the population of all pedestrian/vehicle crashes should be sufficient to carry out this identification but in practice two technical difficulties arise. First, except in certain special cases (e.g. as described in Randles et at. (2001)), direct measurements of the vehicle's speed at impact will not be available so these values must be estimated from a reconstruction of the crash. Several data sets containing impact speed estimates and injury severity classifications have been published over the years (Tharp and Tsongas 1977, Pasanen 1992, MacLaughlin et al. 1993) but before using these data one must seriously consider whether or not they were collected according to an outcome-based sampling scheme. For instance, in the late 1970's the National Highway Traffic Safety Administration sponsored crash investigation teams to collect information on vehicle/pedestrian crashes in several U.S, cities but it was simply not possible to investigate all crashes occurring during the study period. Efforts were made to investigate all crashes in which a pedestrian was fatally injured while crashes producing less severe injuries were samples so that the probability a given crash was included in the sample depended on the severity of the injury it caused (Lawrason et at. 1980). Outcome-based sampling allows one to use limited resources in order to collect acceptable samples on relatively rare events, such as fatal crashes or rare diseases, and is often used in epidemiology and econometrics. An outcome-based sample of vehicle/pedestrian crashes can be used to directly estimate the conditional probability of the vehicle's impact speed given the pedestrian's injury severity, P [v|severity] but for prediction of the effects of speed policies we need the prospective conditional probability p [severity|v]. Because the probability that a crash is included in the sample depends on the resulting injury severity, naively fitting a prospective model using an outcome-based sample will almost always produce inconsistent estimates of the model's intercept parameters leading in turn to biased predictions (Manski and Lerman 1977, Prentice and Pyke 1979). So unless it can be ascertained that a given impact speed/injury severity data set was generated using a representative sampling plan, these data cannot, by themselves, be used to fit a prospective model. Fortunately, if one can supplement the outcome-based sample with data that allow consistent estimate of either the marginal distribution of injury severities in the population or the marginal distribution of impact speed, methods do exist which five consistent estimates of the prospective model's parameters (Hsieh et al. 1985).

The availability and quality of supplemental data then determines the usefulness of an outcome-based sample. Ashton (1982) and Ashton et al. (1977) describe crash investigation studies conducted in Great Britain, in which estimates of the impact speed of vehicles involved in pedestrian crashes were made using accident reconstruction methods. The pedestrian injuries were classified as slight, serious or fatal by consultant police and hospital records. A serious injury was one resulting in either hospitalization or in one (or more) of a set of particular injury types even if not resulting in hospitalization. A fatal injury produced death within 30 days and a slight injury as then any injury not judged as serious or fatal (Ashton, 1982 p. 613). These data are reproduced here in Table 2.1 and were used by Pasanen (1992, Pasanen and Salmibaara 1993) to fit a prospective model relating impact speed to fatality risk. It must be noted though that Pasanen does not appear to have accounted for the possibility of outcome-based bias in his fitted model, while Ashton himself pointed out that the marginal distributions of injury severity in these data tended to show greater proportions of more severe injuries when compared to nationwide tabulations indicating that outcome-based sampling has probably operating. Nationwide tabulations for 1976 were also given (Ashton 1982) and these are reproduced in Table 2.2. If one makes the plausible assumption that the marginal distribution of injury severity for the sample areas is similar to the nationwide distribution, these data can be used to fit the prospective model.

The second technical difficulty stems from the fact that it is generating not possible to make a precise estimate of the vehicle's impact speed in an accident reconstruction, at best, one can only determine a plausible range of speeds and the widely of this range depends on the circumstances of the individual crash (Niedere 1991). If, for model fitting, one picks a representative value from this range. This leads to an error-in-variables regression problem and fails to account for the resulting co-variable measurement error. This can also produce inconsistent parameter estimates (Carroll et al. 1995). This difficulty is compounded somewhat since in Table 2.1 the speed data have been grouped into 10 km/h bins and failure to account for this grouping can lead one to overstate the precision of a parameter estimate (Armstrong 1990). In this study, we will (partially) account for measurement error issues by assuming that a crash's unobserved actual impact speed is uniformly distributed within the speed bun to which that crash was assigned. The probability of an injury severity in category i, given that the impact speed is in bin k, is then:

$$P\left[severity = i \,\middle|\, v \in bin\ k\right] = \frac{\int_{v_{1,k}}^{v_{2,k}} P\left[severity = i \mid v\right] dv}{v_{2,k} - v_{1,k}} \qquad (2.3)$$

Where $v_{1,k}$ and $v_{2,k}$ are the lower and upper limits defining speed bin k.

The data given in Table 2.1 and 2.3 are sufficient to fit threshold models with three injury severity categories: the models being determined by the slope parameter b, and two thresholds, a_1 and a_2. Speeds in the 71 + km/h bin were taken to be uniformly distributed between 70 and 100 km/hr. The conditional maximum likelihood method described by Heieh et at. (1985) was implemented using MathCAD 6+ (Mathsoft 1995) by numerically solving the resulting likelihood equations. The covariance matrix for the parameter estimates was then estimated using the expressions given in the Appendix A of Hsieh et al. (1985). Both logit and probit model were fit. For the logit models closed form expressions for the integrals appearing in equation (2.2) and (2.3) are available leading to expressions of the form:

$$P\left[severity = i \mid v \in bin\ k\right] = \frac{\int_{v_{v,k}}^{v_{2,k}} \int_{a_i - bv}^{a_{i+1} - bv} f(e)\, de\, dv}{v_{2,k} - v_{1,k}}$$

$$= \frac{\log\left(\frac{1 + \exp\left(a_i - bv_{2,k}\right)}{1 + \exp\left(a_i - bv_{1,k}\right)}\right) - \log\left(\frac{1 + \exp\left(a_{i+1} - bv_{2,k}\right)}{1 + \exp\left(a_{i+1} - bv_{1,k}\right)}\right)}{b\left(v_{2,k} - v_{1,k}\right)} \qquad (2.4)$$

While for the profit models the integrals were evaluated numerically. Goodness of fit was assessed using the deviance statistic:

$$D = -2\left(\sum_{i=1}^{3}\sum_{k=1}^{8} m_{i,k} \log\left(\frac{m_{i,k}}{m_{i,+}}\right) - \sum_{i=1}^{3}\sum_{k=1}^{8} m_{i,k} \log\left(\frac{p_{i,k}\,\pi_k}{q_i}\right)\right) \qquad (2.5)$$

Where:

$m_{i,k}$ = Number of investigated crashes with severity i in speed bin k

$m_{i,+}$ = Total number of investigate crashes with severity i

$p_{i,k}$ = $P\left[severity = i \mid v \in bin\ k\right]$ evaluated at the maximum likelihood estimates (MLE) for b, a_1 and a_2

π_k = MLE of the marginal probability of an impact speed falling in bin k

q_i = MLE of the marginal probability of injury severity i

Large values for the deviance statistic indicate a lack if fit between the data in Table 2.1 and what would be predicted by our threshold models. Table 2.3 displays the deviance measures for each pedestrian age group and for both the logit and probit models together with significance levels from an approximate chi-squared test of the null hypothesis that the proposed models fits the data. Both the logit and probit models provide acceptable fits with the deviance values for the logit models being slightly lower. Since the resulting prospective models for both logit and probit turned out to be very similar, in what follows, we will focus on the logit models.

Table 2.4 gives the parameter estimates and corresponding standard error for the fitted logit models while Figures 2.1, 2.2 and 2.3 show the distribution of injury severities as functions of impact speed for each pedestrian age group. To see how the values in the figures were arrived at, consider a child pedestrian who is struck by an automobile traveling at 50 km/hr, using the parameter estimates from Table 2.4, the probabilities of slight, serious and fatal injuries would be:

$$\begin{aligned}
P\left[slight \mid v = 50\,km/h\right] &= \frac{\exp(4.68-(0.12)50)}{1+\exp(4.68-(0.12)50)} = 0.21 \\
P\left[serious \mid v = 50\,km/h\right] &= \frac{\exp(8.85-(0.12)50)}{1+\exp(8.85-(0.12)50)} - \frac{\exp(4.68-(0.12)50)}{1+\exp(4.68-(0.12)50)} = 0.74 \\
P\left[fatal \mid v = 50\,km/h\right] &= 1-\frac{\exp(8.85-(0.12)50)}{1+\exp(8.85-(0.12)50)} = 0.05
\end{aligned} \qquad (2.6)$$

Inspection of Table 2.4 and Figures 2.1-2.3 shows that while the “Children” and “Adult” groups have essentially similar injury severity distributions, that for the “Elderly” group is noticeable different, with older pedestrians being much more likely to suffer severe injuries at lower impact speed. This accords with the increased “frailty” of older persons identified by Hauer (1988).

It should also be noted that the injury severity models for the ‘Children” and “Adult” presented in Figures 2.1 and 2.2 are noticeable different from models presented elsewhere in the literature. In Figures 2.1 and 2.2, the impact speed at which the chance of fatal injury is 50% occurs between 70 and 75 km/h, as opposed to 53 km/h (Pasanen and Salmivaara 1993) or 45 km/h (Anderson et al. 1997). On the other hand, the model for “Elderly” pedestrians, presented in Figure 2.3, is similar to the previously published curves. The same data were used here as were used by Pasanen (1992) in developing his model, the main differences being that in our application the data were not aggregated across pedestrian age groups and estimation methods which explicitly accounted for outcome-based biases were employed. The resulting differences between Pasanen and Salmivaara’s model and our Figures 2.1 and 2.2 must then be due to a combination of aggregation and outcome-based biases. The pedigree of the data used in Anderson et al. (1997) is less clear cut but given the similarity of their curve to that of Pasanen and Salmivaara one should at least suspect that similar factors were present. Thus, although the previously published curves are consistent with what we obtained for the aged 60+ pedestrian group, they do not appear representative of what one can expect for younger pedestrians.

Finally, as noted in Chapter 1, there has been discussion in Minnesota as to whether or not 25 mph (40 km/h) is more appropriate than 30 mph (50 km/h) as a speed limit for residential streets, but that was not possible to identify the predicted safety benefits resulting from this change. Knoblauch (1977) indicates that on local residential streets the most frequent type of pedestrian crash involves a child running in front of a moving vehicle. If one assumes that such streets should function as mixed-use facilities rather than solely as avenues for vehicle travel, then a residual baseline frequency of such crashes is to be expected. A rational speed limit policy might then seek to keep the chances of severe injuries resulting to the pedestrian to an acceptably low level. Looking at Figure 2.1, we can see that for impact speeds below 25 mph (40 km/h) it is most likely that the pedestrian injuries will be slight; but that for impact speeds above 25 mph (40 km/h) serious injuries becomes most frequent and above about 47 (75 km/h) fatal injuries are most likely. Since dart-out type crashes often give the driver little time to react to the emergence of the pedestrian, one can argue that the initial and impact speeds should be similar, and so Figure 2.1 suggests that 25 mph (40 km/h) provides a plausible limit for vehicle speeds on local residential streets.

Table 2.1

Tabulations of Investigated Pedestrian Crashes Reported in Ashton (1982)

Children (Ages 0-14 years)

Injury Severity	Estimated Impact Speed (km/h)							
	0-10	11-20	21-30	31-40	41-50	51-60	61-70	71+
Slight	9	18	25	12	5	2	0	0
Serious	0	2	18	25	18	9	0	0
Fatal	0	0	1	0	5	3	2	1

Adults (Ages 15-59 years)

Injury Severity	Estimated Impact Speed (km/h)							
	0-10	11-20	21-30	31-40	41-50	51-60	61-70	71+
Slight	2	5	11	10	2	1	0	0
Serious	0	1	18	14	14	7	1	0
Fatal	0	0	3	1	3	11	8	9

Elderly (Ages 60+ years)

Injury Severity	Estimated Impact Speed (km/h)							
	0-10	11-20	21-30	31-40	41-50	51-60	61-70	71+
Slight	6	1	3	0	0	0	0	0
Serious	0	2	13	16	6	1	0	0
Fatal	0	0	3	2	15	10	3	1

Table 2.2

Pedestrian Casualties by Age and Injury Severity for Great Britain in 1976 (from Ashton 1982)

Injury Severity	Age Group		
	0-14 years	15-59 years	60+ years
Slight	21,072	17,873	7,272
Serious	7,461	6,276	6,276
Fatal	405	720	1,208

Table 2.3

Goodness of Fit for Logit and Probit Models

(D = Computed Deviance Statistic; df = Degree of Freedom; p = Significance level)

Age Group	Logit Models			Probit Model		
	D	Df	P	D	Df	P
0-14	11.37	10	0.33	11.55	10	0.32
15-59	9.62	10	0.47	11.97	10	0.29
60+	11.48	10	0.32	12.61	10	0.25

Table 2.4

Parameter Estimates for Logit Models

(Estimated Standard Errors in Parentheses)

Age Group	Model Parameter		
	B	a_1	a_2
0-14	0.120 (0.019)	4.678 (0.543)	8.846 (0.809)
15-59	0.127 (0.018)	4.971 (0.531)	8.866 (0.822)
60+	0.204 (0.035)	5.290 (0.811)	9.728 (1.433)

Figure 2.1

Injury Severity Versus Impact Speed for Pedestrians Aged 0-14 Years

Key:

P [slight]: _______

P [serious]: +++++++

P [fatal]: - - - - - - -

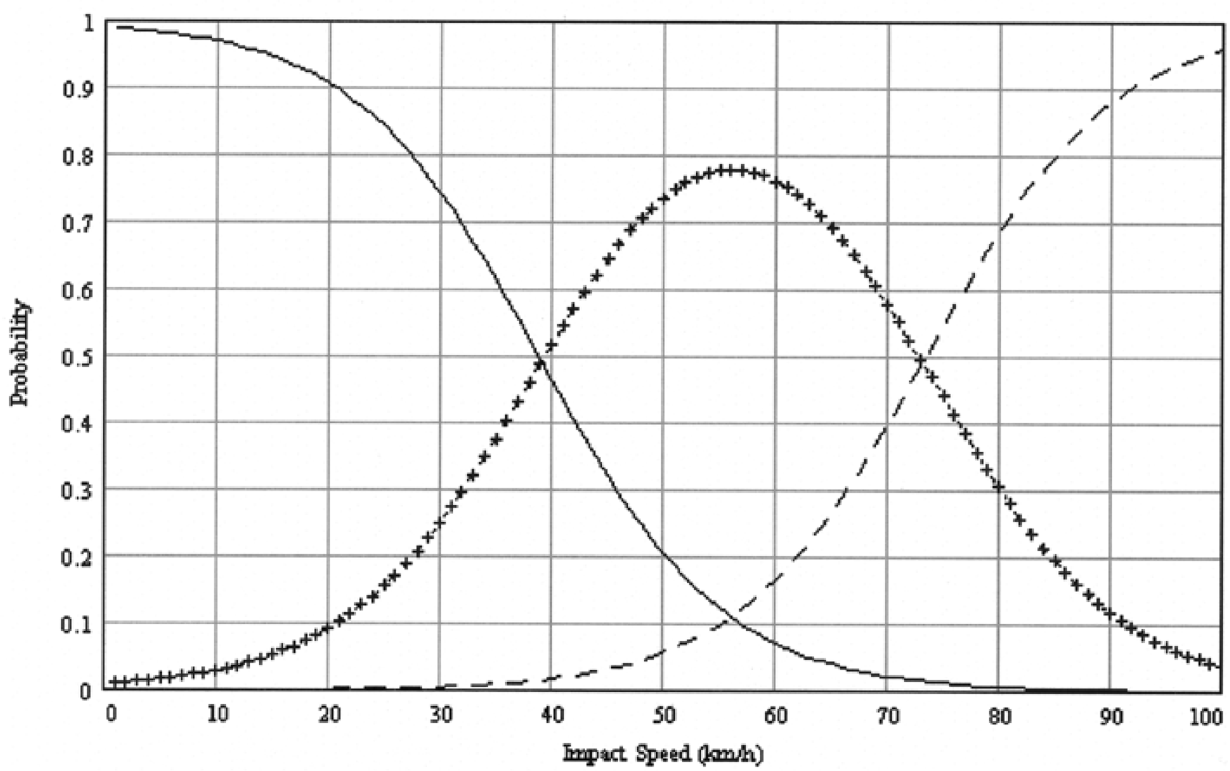


Figure 2.2

Injury Severity Versus Impact Speed for Pedestrians Aged 15-59 Years

Key:

P [slight]: _______

P [serious]: +++++++

P [fatal]: - - - - - - -

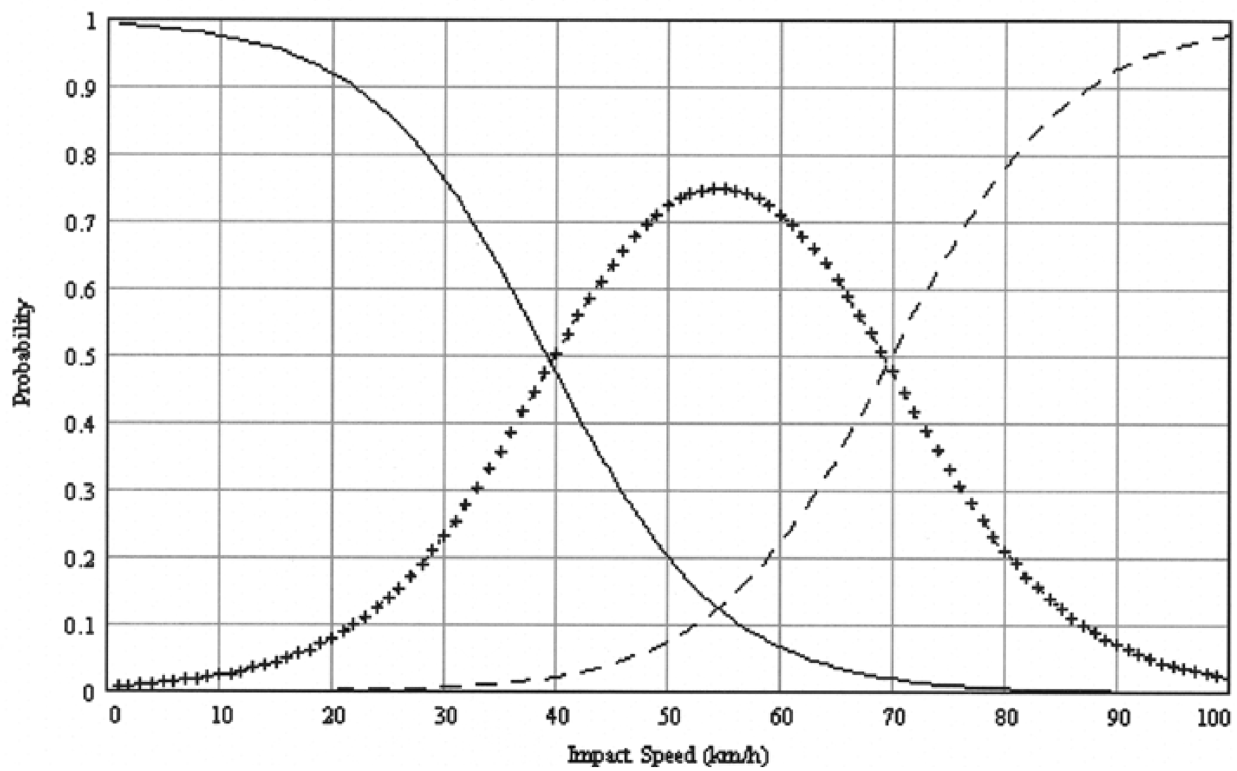


Figure 2.3

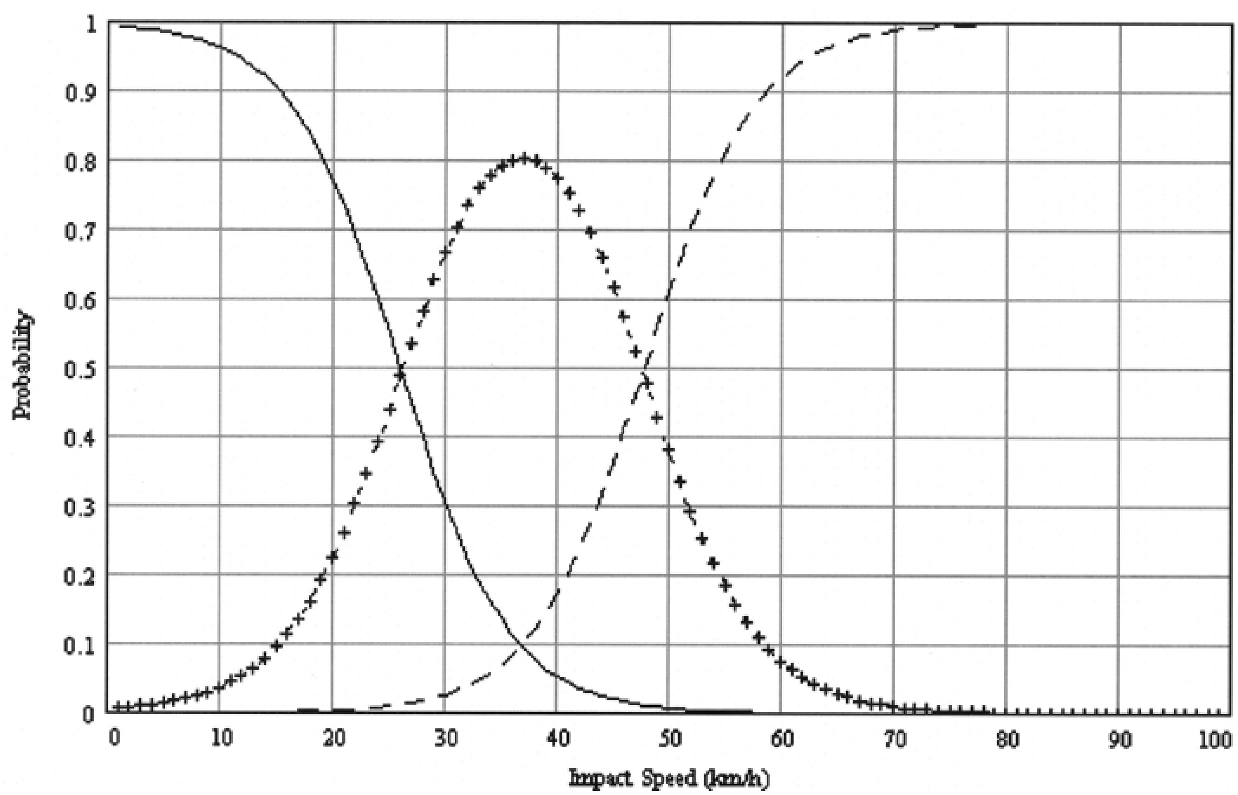
Injury Severity Versus Impact Speed for Pedestrians Aged 60+ Years

Key:

P [slight]: _______

P [serious]: +++++++

P [fatal]: - - - - - - -



CHAPTER THREE

DEVELOPMENT AND APPLICATION OF PEDESTRIAN CRASH MODEL

As noted in Chapter 1, pedestrian crashes on local and residential streets tend to be spatially diffuse with the large majority of such streets likely to show, at most, one crash in any reasonable time period. It was also noted that crash risk is not synonymous with crash count so that a low crash count does not necessarily guarantee that traffic conditions are such as to provide a “safe and pleasant environment” for pedestrians. But the sparseness of pedestrian crashes makes it difficult to compute reliable estimates of crash rates which in turn make it difficult to use estimated rates to measure crash risk and to prioritize sites as to their suitability for NTC. To get out of this impasse we can begin by taking note of Hauer’s (1997) distinction between safety of a site and a method for estimating safety. In particular, an estimating equation such as (1.1) does not *define* what is meant by crash rate. Rather a site’s crash rate is better seen as an underlying property of the site which the equation (1.1) attempts to measure. This underlying property may in turn depend on other features of the site such as traffic volumes and speeds, road geometry or pavement condition. Once we distinguish between crash rate and a method for estimating it we can ask whether or not it is possible to find some other method for estimating the crash rate, but this requires getting a clearer view of what the crash rate is. Imagine for the moment that for a residential street we knew what over the course of a year, pedestrians entered the street a total of n times and that there were x collisions between pedestrians and vehicles, we can ask: “Under what conditions would the ratio x/n give us a reasonable estimate of crash risk?” Probably the simplest mechanism for justifying this estimate arises by assuming that each entering pedestrian has the same probability, p, of being hit and that each pedestrian’s chance of being hit is independent of what happens to other pedestrians. It follows, then, that the crash count x can model as the outcome of a binomial random variable, and then x/n estimates the underlying crash probability p. Given two streets with significantly different values for their estimated crash probabilities, we would be inclined to conclude that the street with the higher crash probability is, other than equal, more dangerous for pedestrians. The question is then how to obtain an estimate of this probability when reliable values for x or n are unavailable.

In the simple binomial model the crash probability, *p*, is taken as an irreducible characteristic of the site in the sense that while it may be possible to identify other measurable variables which allow us to predict how these probabilities vary across sites, no attempt is made to explain (or derive) these probabilities from more fundamental mechanisms. As noted in Chapter 1, although engineers seeking to prioritize roadside improvement projects have encountered difficulties similar to ours and have developed a solution which has been implemented in computer programs such as ROADSIDE and RSAP. In essence this solution uses a deterministic model to describe whether or not a vehicle after encroaching onto the roadside will strike a roadside object as a function of the shape and location of the object as well as of variables such as the vehicle's speed, point of encroaching, encroachment angle and deceleration rate. By placing probability distributions on these variables, it is then possible, at least in principle, to compute the probability that an encroaching vehicle strikes the object.

We will use a similar approach to determine pedestrian collision probabilities but before moving to the details of our procedure, it may be helpful to take a closer look at its underlying rationale. Most would agree that crash risk to pedestrians increase as vehicle traffic volumes and speeds increase. Consider the risk to a pedestrian trying to cross a 12-foot lane of an urban freeway when the traffic volume is 2000 vehicle/hour and the average speed is 60 mph in contract to the risk to the same pedestrian trying to cross a 12-lane of a residential street when the traffic volume is 10 vehicle/hour and the average speed of 20 mph. Lying between these extremes is a variety of intermediate cases and the problem at hand is to somehow rank these intermediates cases as to pedestrian risk. One way we might try to solve this problem would be to develop a standardized test procedure. This could involve first specifying the properties of a standard "test pedestrian" then having a specified number of test pedestrians attempt to enter the street of interest and finally recording the number of collisions. The ratio of collisions to entries would then give an estimate of that street's crash probability and the number of entries could be chosen beforehand so that a given level of precision for this estimate could be guaranteed with a given confidence. Like any test procedure this one does not encompass all possible scenarios but achieves its validity by being (1) standardized so that different streets could be evaluated on the same criteria and (2) representative of the type of event of practical importance. The obstacles to implementing such a test however are obvious even if we could replace the human test pedestrians with robots.

Given that such a standardized test is not feasible, one alternative is to notice that entries and collisions do in fact occur even if we cannot stage them, so that we simply need to record the number of pedestrian entries and collisions produced by such "natural" tests. This, in essence, is what is done when safety programming is based on estimated crash rates. Some care needs to be taken so to guarantee that nature's tests occur under approximately similar conditions and to make appropriate corrections when they have not. As noted earlier though this approach is foreclosed for pedestrian crashes on local and residential streets because (1) unlike vehicle volumes, measures of pedestrian street entries cannot at present be collected automatically and (2) pedestrian crashes do not tend to show spatial clustering. If we were completely ignorant concerning the underlying mechanisms governing traffic crashes we would be at a stand. But if such knowledge were available, another alternative would be to simulate the outcome of a standardized test on a computer. This is the approach used in assessing roadside improvements and the one we will develop.

The first task is to define a standardized test. Pedestrian crashes tend to fall into several different types depending primarily on what the pedestrian is doing prior to the crash. These include when a pedestrian is walking along a road when a pedestrian is working on a stopped vehicle or when a pedestrian is crossing at an intersection (Zeeger 1993). Thus, no single physical model is likely to give an accurate representation of all crash types. The most common crash type (around 30% of pedestrian crashes) is a midblock dart-out in which the pedestrian runs into the street at some midblock location (Knoblauch et al. 1984, Baltes 1999) and these most commonly involve a child running onto a two-lane local street (Knoblauch 1977). Since, as noted in Chapter 1, local and residential street often function a multi-use facilities. A non-zero baseline of dart-outs by children is to be expected and the probability that a dart-out action by a child pedestrian results in a crash should then be indicative of the safety on a local or residential street.

Figure 3.1 illustrates a dart-out crash in which a car is traveling at an initial speed of v_1 when the pedestrian begins moving toward the street at a speed v_2. At the time the driver notices the pedestrian the vehicle is a distance of x_1 from the point of collision while the pedestrian is at distance x_2. If it is possible, without changing speed, to pass the collision point before it is reached by the pedestrian. This is what driver does. Otherwise after a perception/reaction interval t_p, the driver brakes to a stop at a constant deceleration *a*. It is assumed that if the pedestrian reaches the collision point before the vehicle the pedestrian stops at this point, while if the vehicle reaches the collision point first the pedestrian stops before running into the vehicle. This analytic approach to characterizing vehicle/pedestrian crashes appear to have first been treated by Mayne (1965) who also considered a more complicated case where the pedestrian continues to moving and could be struck by cars traveling in the far lane. Howarth et al. (1974) subsequently used this approach to develop a rough measure of children's exposure to crash events. More recently, similar models have been used by Pasanen and Salmivaara (1993) to consider the likelihood of fatal pedestrian crashes as a function of vehicle speed and by Vaughn (1997) to identify the point at which a pedestrian could first seen by a driver.

In figure 3.1, $x_N = (v_1 v_2)/v_2$ is the maximum distance the vehicle can be from the collision point and still avoid the collision by arriving before the pedestrian while $x_S = v_1 t_p + \left(v_1^2 / 2a \right)$ is the vehicle's stopping distance. Clearly, if the initial distance x_1 satisfies $x_N \langle\ x_1 \langle\ x_S$, a collision will occur and in that case the impact speed *vi* of the vehicle will be give by:

$$
\begin{aligned}
vi &= v_1 \quad if \quad x_N \langle\ x_1 \langle\ v_1 t_p \\
vi &= \sqrt{v_1^2 - 2a(v_1 - x_1 t_p)} \quad if \quad v_1 t_p \leq x_1 \langle\ x_s \\
vi &= o \quad otherwise
\end{aligned}
\qquad (3.1)
$$

If values for the variables x_1, v_1, t_p, a , x_2, and v_2, were known then whether or not a collision occurred and the resulting impact speed should be determined. If the age of the pedestrian were also known then a probability distribution over pedestrian injury severities could be obtained using the appropriate model from Chapter 2. In practice, though all of these variables are likely to differ from encounter to encounter and although one could select plausible nominal values, these would always be to some extend arbitrary, as would the conclusions reached from using them. On the other hand, if we knew a probability distribution $F(x_1, v_1, t_p, a, x_2, v_2)$ over the values of the collision variables then the probability of a collision could in principle be computed via multivariate integration:

$$
p \;=\; Prob\left[x_N \langle\ x_1 \langle\ x_S\right] \;=\; \int I\left(x_N \langle\ x_1 \langle\ x_S\right) dF\left(x_1, v_1, t_p, a, x_2, v_2\right) \qquad (3.2)
$$

Where I (.) denotes the indicator function

I (x) = 1 if statement x is true

I (x) = 0 if statement x is false

Similarly the probability that the collision speed is less than some value v′, given that a collision occurs can be computed as:

$$F(v') = Prob\left[vi \le v' \middle| x_N \langle x_1 \langle x_S\right] = \frac{Prob\left[vi \le v' \wedge x_N \langle x_1 \langle x_S\right]}{Prob\left[x_N \langle x_1 \langle x_S\right]} \tag{3.3}$$

Where

$$Prob\left[vi \le \wedge x_N \langle x_1 \langle x_S\right] = \int I(vi \le v')\, I\left(x_N \langle x_1 \langle x_S\right) dF\left(x_1, v_1, t_p, a, a_2, v_2\right) \tag{3.4}$$

Finally the probability an injury of severity category I results given that a collision occurs could be computed as:

$$Prob\left[severity = i\right] = \int Prob\left[severity = i \mid vi\right] df\,(vi) \tag{3.5}$$

Although the deterministic model of a pedestrian vehicle crash together with a probability distribution on the model's input variables should allow us to compute the desire probability evaluating the integrals appearing in equation (3.2–3.5) tends to be numerically difficult and to date all attempts at implementing this approach have been forced to allow only some of the input variables to be random with the others being fixed at plausible, but arbitrary, nominal values (Davis and Corkle 1997; Davis 1998).

Rather than attempting to find closed form expressions for above integrals or to evaluate them numerically using multidimensional quadrature we will use Monte Carlo (MC) methods. In a standard MC implementation a pseudo-ransom sample of values from the input variables' distribution is generated on a computer and then relevant integrals are evaluated by simply averaging over the appropriate functions of these random outcomes. Monte Carlo methods have in fact been employed in the RSAP model for evaluating roadside improvement (Mak et al. 1998). Tradition MC methods are limited though in the degree to which probabilities can be conditioned on crash-specific information. Although this is not a problem when computing the crash and injury severity probabilities in equations (3.2–3.5) it will become more important when we seek to compute analogues of crash reduction factors. During the last 15 years or so a more flexible set of techniques called Markov Chain Monte Carlo (MCMC) methods have been used by statisticians to solve increasingly more difficult estimation problems (Carlin and Louis 1996). In an MCMC method, a pseudo-random sequence of outcomes from a type of stochastic process called an ergondic Markov chain is simulated on a computer. With the Markov Chain being constructed so that its long run or stationary, distribution is the same as that for one desires probability statements. Since one of the properties of an ergotic Markov Chain is that averages taken over a sequential realization converge to averages computed from it stationary distribution, probability estimates can again be computed by averaging over the simulated outcomes. MCMC methods have recently been used to estimate crash rates while accounting for uncertainty in the exposure measures (Davis 2000) and to identify intersections where older drivers appear to have increased crash risk (Davis and Yang 2001). MCMC methods are even more attractive because a computer program called BUGS can carry out the necessary computations for a wide variety of estimation problems including the more standard MC problem (Gilks et at. 1995)

Figure 3.2 shows a representation of the pedestrian/vehicle collision model as a directed acyclic graph (DAG). In Figure 3.2, circles represent variables, while the arrows represent the fact that the values of some variables cause the values of others. For example, the arrow leading from v_1 to *vi* indicates that the vehicle=s initial speed is a causal determinant of the value for *vi*. Variables for which no arrows enter are called exogenous variables, since their values do not depend on other model variables, while the variables that are dependent are called endogenous variables. To specify a model for BUGS, it is necessary to specify a functional relationship for each of the endogenous variables, and a probability distribution over the exogenous variables. The functional relationships have already been defined above, so all that remains is to determine plausible probability distributions for x_1, v_1, t_p, a, x_2, and v_2.

In an earlier study, investigating the relationships between landscape design variables and vehicle speed (Giese 1996), spot speed data, as well as data on a number of physical measures, were collected on 50 residential street segments located in the Twin Cities metropolitan region. Since it was desired, capitalize on this earlier work if possible, these segments were also chosen as the initial sample for this study. However, locating the segments on a street map revealed that often two or more segments were contiguous, and so could not be treated as independent entities. Eliminating contiguous segments produced 25 segments for this study=s final sample. The locations of these segments, together with the average building setback determined in the previous study, are presented in Table 3.1. During the summer of 2000, a TimeMark Delta III portable traffic counter was placed on each of these segments for at least 24 hours on a typical weekday. From the raw pulse data collected by the counter, TimeMark=s TMWin software extracted measurements of individual vehicle speed, individual vehicle headways, hourly traffic volume by direction, and total traffic volume. The resulting 24-hour, two-directional traffic volumes are also presented in Table 3.1. From these data, it was possible to identify, for each segment, the hour and direction showing the highest traffic volume, and for this peak hour, the measurements of individual vehicle speeds and headways were extracted. Summary measures for the peak-hour, peak-direction traffic at each site are shown in Table 3.2.

In earlier work, (Davis 1998) it was assumed that vehicle speeds on residential streets followed a normal probability distribution, while vehicle-headways followed an exponential distribution. The first task then was to test the plausibility of these assumptions. Figure 3.3 shows histograms for the speed and headway distributions from site 17, while Figure 3.4 shows a normal and an exponential probability plot for the speed and headway data, respectively. It can be seen that while the speeds appear to be normally distributed, the headways are not exponential. The skewed shape of the headway distribution suggests the possibility that the headways are log-normally distributed. Table 3.3 shows the results of statistical tests of the hypothesis that the speeds and the log headways are normally distributed, for the peak-hour, peak-direction at each site. The Ryan-Joiner test implemented in Minitab (Minitab 1995) was used, with a significance cutoff of 0.05. This test in essence checks for the existence of a linear relationship in a normal probability plot, with a linear relationship being consistent with the hypothesis that the data are normally distributed. The statistic R in Table 3.3 is a correlation coefficient which measures the deviation from linearity, R=1.0 indicating a perfectly linear relationship, while R=0 indicates no relationship. The column P in Table 3.3 gives the corresponding significance levels, and for $P<0.05$ we would reject the hypothesis that the data are normally distributed. For the speeds, at sites 4 and 18 the computed value for R falls below the critical level, but since at the 0.05 level, out 25 tests, the probability of two or more rejections by chance alone equals 0.36, we can plausibly conclude that overall the speeds tend to be normally distributed. For the log-headways, we would reject the normal hypothesis for five of the 25 sites, and the probability of five or more rejections by chance alone is equal to 0.007. This second probability is low enough that we must conclude that at least some sites the lognormal model does not fit. Thus while in most cases a lognormal model is plausible for vehicle headways, one should check the degree to which one's conclusions are sensitive to that assumption.

Using the distributions for vehicle speeds and headways, the distribution for the initial distance x_1 can be determined, as follows. The assumption that the test pedestrian is heedless, and so darts out without looking at traffic, can be formalized by if the time point at which the dart-out begins is uniformly distributed between successive vehicle arrivals. That is, if we let t_1 denote the time between when the dart-out begins and when the vehicle would arrive at the collision point when traveling at speed v_1, then t_1 is uniformly distributed between 0 and the duration of the inter-vehicle gap. The initial distance is then given by $x_1 = v_1 t_1$.

Using portable traffic counters, data relevant to vehicle speed and headway distributions can be collected automatically at a site, but data concerning the model's other exogenous variables, t_p, a, v_2 and x_2, are more difficult to come by. However, researchers at the Texas Transportation Institute have recently completed a study of driver behavior in emergency braking situations, and have published data relevant to drivers' reaction times and deceleration rates (Fambro et al. 1997). In particular, in 21 tests where the subject was driving his or her own car and had to make an emergency stop to avoid an obstacle appearing suddenly in the roadway (studies 3 and 4 in Fambro et al 1997), the average time between appearance of the obstacle and the driver's application of the brakes was 1.07 seconds, with a standard deviation of 0.248 seconds. Although a detailed analysis of the reaction time distribution was not provided, the data did show evidence of positive skew, consistent with a lognormal distribution. Since prior work has suggested that reaction times tend to be approximately log-normally distributed (Olson 1996), we used a lognormal distribution for the reaction times, with mean and standard deviation equal to the values listed above. In tests where an average deceleration was recorded (study 2), the average (of these averages) was about -0.63g, with a standard deviation of about 0.08g (g = 32.2 feet/sec^2 being the acceleration due to gravity). Although the reported results were not sufficient to determine a distribution for the braking rates, there was again some evidence of positive skew. Taking the braking drag factor f = -a/g as our primary variable, we modeled f as being a lognormal outcome, with mean equal to 0.63 and standard deviation equal to 0.08.

The variables v_2 and x_2 refer the pedestrian's movement speed and the pedestrian's distance for the point of collision when noticed by the driver. As with driver reaction time and braking rate, site-specific data on these tend to be difficult to collect without special experiments. However, measurements of the average setback (distance from the street curb to building front) were available from the previous study, and were listed in Table 3.1. Absent any more detailed information on when drivers tend to notice and react to pedestrians, and taking the point of collision as being 1.5 meters (4.9 feet) from the curb, a reasonable solution is to treat the pedestrian's initial distance as being uniformly distributed between the point of collision and the building front. Finally, pedestrian movement speeds depend both on the pedestrian's age and whether or not he or she is running or walking. As noted earlier, Knoblauch (1977) has reported that on dart-out crashes most frequently involve child pedestrians. Times to run 50 yards by for fourth, fifth, and sixth grade children, collected in their physical education classes, were obtained from a Twin Cities elementary school. A running fifth grade boy was selected as our test pedestrian, with an average speed of 5.4 meters/sec (17.7 feet/sec) and a standard deviation of .45 meters/sec (1.5 feet/sec). Treating these running speeds as normal random variables then completed the specification of our model.

Using the WINBUGS implementation of the BUGS software, the probability that a pedestrian running into the street and hit by a vehicle, under the scenario described above, was estimated. In each case, a 4,000 iteration burnin was followed by 20,000 iterations used to compute the estimates. In addition, the probability of a collision with a severe (serious or fatal) injury was also estimated, and these values were in turn used to compute the conditional probability of a resulting severe injury, given that a collision occurs. These results are presented in Table 3.4, under the "Parametric Model" columns. As an example of how they interpret these results, at site #11 a typical heedless fifth grade boy running into the street would be hit by a car with probability 0.037, and with probability 0.013 he would be hit and severely injured. Considering only the population of occurring collisions, about 36% of these would result in severe injuries. As noted earlier, for some sites the assumption that the headways are log-normally distributed is questionable, and so the sensitivity of the results to deviations from log normality should be investigated. To do this, the various collision probabilities were re-computed, this time using the empirical speed and headway distributions generated by the actual samples. That is, at each iteration of the Gibbs sampler, instead of sampling from a normal or lognormal distribution, we randomly picked a speed/headway combination from the site's sample. In this way, the speed and headway distributions are identical to those collected at the site, and we also dispense with the assumption that speeds and headways are independent. The results of these computations are also displayed in Table 3.4, in the "Nonparametric Models" columns.

The first thing to note is the magnitude of the collision probabilities, which range from 0.004 to 0.064 for the parametric models, and from 0.004 to 0.077 for the nonparametric models. That is, we would expect that for every thousand times a heedless test child runs into the street, between 4 and 64–77 collisions with a car would result, depending on the site. These are similar to results presented by Howarth et al (1974), who found estimated collision rates of between 23 and 60 per thousand when neither the driver nor the child pedestrian takes evasive action. The second thing to note is that while the results developed using the parametric and nonparametric models are similar, they are not identical. This tendency is shown clearly in Figure 3.5, which is a scatter plot of the parametric collision probabilities versus their nonparametric counterparts. Since the primary practical use of the estimated collision probabilities will be to rank sites with respect to hazard to pedestrians, it is useful see if the two computational models yield similar identifications. The parametric model picked sites 27b, 55, 27a, 17, and 2 as the five with highest collision probabilities, while the nonparametric model picked sites 27b, 55, 46, 4, and 27a. Looking now at the 10 sites with highest collision probability, the parametric method added sites 4, 11, 12, 18, and 50 to the list, while the nonparametric method added sites 11, 2, 17, 12, 50. So, the two methods have three sites in common in their top five, and nine sites in common in their top ten. For the purposes of ranking sites the two different models for speed headway distributions tend to produce similar results, especially if the sample sizes were small.

Finally, before moving to the problem of estimating crash reduction effects, it may be of interest to consider the relative contributions of traffic volumes and average traffic speed to pedestrian risk. Figure 3.6 shows the collision probabilities, computed for both methods, plotted against the peak-hour peak-direction traffic volume for each site, and against average speed. Overall, we can see that while collision probability tends to increase as traffic volume increases, it is relatively independent of average vehicle speed. Figure 3.7 shows similar plots, but this time for the conditional probability of a severe injury given that a collision occur, and in Figure 3.7 we see an opposite pattern, with severe injury probability being relatively uninfluenced by traffic volume, but definitely tending to increase as average traffic speed increases. Thus, at least for these sites, traffic volume apparently influences the chance that our test pedestrian is hit, but has little influence on the severity of any resulting injuries. Average vehicle speeds on the other hand strongly affect the likelihood of severe injuries.

In Chapter 1 it was noted that two probabilistic quantities are commonly used in standard traffic safety programming: the estimated crash rate, which is used to identify potential high-hazard sites and to predict future crash experience in the absence of substantive change at a site, and the crash reduction factor (RF), which is used to predict the effect of a safety countermeasure. Like crash rate, reduction factors are usually defined operationally via a formula for estimating them, such as (1.3). With regard to crash rate, we found it important to distinguish between a feature or property of a site, and a method for measuring or estimating that property. This led us to identify the probability of a crash as the underlying property of interest, which the standard crash rate formula estimates. This then leads to the question: What is the underlying property that the RF formula estimates?

As noted elsewhere (Davis 2000b), reduction factors are used to predict the effect that a countermeasure is expected cause when implemented, so we need to be precise about what we mean by a "causal effect". Since we are interested in the "causal effect" of speed management, we will couch the discussion in those terms, although the approach is more generally applicable. Imagine then a street, and an event consisting of a pedestrian entering the traveled way. Define the response variable:

h1 = 1 if the pedestrian is hit by a vehicle
= 0 if the pedestrian is not hit by a vehicle

Now image a similar event except that speed controls are present on the street. Define response variable:

h2 = 1 if the pedestrian would have been hit if speed controls had been present
= 0 if the pedestrian would not have been hit if speed controls had been present

Clearly then, speed controls reduce the crash count at the site (by one) if h1=1 and h2=0. Hudea Pear (2000) has presented an extensive analysis of causal ideas in empirical research and three aspects of Pearl's work are relevant here. First Pearl calls conditional probabilities of the form $P[h2=0 \mid h1=1]$, the probability that a struck pedestrian would not have been hit if the speed controls have been in place –"Probabilities of Necessity". That is, if the probability that the absence of speed controls is a necessary condition for the occurrence of the crash. Second pearl shows that formulas of the form:

$$\vartheta = \frac{r_B - r_A}{r_B} \qquad (3.6)$$

can be interpreted when data are collected under appropriately controlled conditions as estimating Probability of Necessity.

Third, Pearl shows how structural knowledge of situation such as that used to define the pedestrian crash test in the previous section can also be used to compute Probabilities of Necessity. To compute an alternative estimate of the reduction factor due to speed management this would involve:

(1.) Using Bayes Theorem to compute posterior distributions for the exogenous variables given that a crash has occurred.

(2.) Setting the value of the vehicle's initial speed to what it would have been under speed management.

(3.) Compute the crash probability using the results of (1.) and (2.).

Balk and Pearl (1994 have also pointed out that these computations can be carried out more or less automatically for graphical models by augmenting the model to include counterfactual variables.

As noted in Chapter 1 several years back there was informal consideration in Minnesota of 25 mph speed limit for residential streets but there as uncertainty concerning the safety effects of the new limit. To illustrate how Pearl's methods could be applied in practice, we will BUGS to compute for each of our sample sites the probability that a pedestrian crash would be prevented under strict adherence to a 25 mph speed limit. The variable h1 is determined by the exogenous variables via:

h1 = 1 if $(v_1 x_2)/v_2 \langle x_1 \langle v_1 t_p + v_1^2/(2a)$

= 0 otherwise

The variable v_1^* is determined by the speed limit and the v_1

$v_1^* = v_1$ if $v_1 \leq 25mph$

= 25mph if $v_1 \rangle 25mph$

h2 = 1 if $v_1^* x_2 / v_2 \; \langle \; x_1^* \; \langle \; v_1^* t_p + (v_1^*)^2 / 2a$

= 0 otherwise

Where $x_1^* = v_1^* t_1$

Table 3.5 shows the results of these computations for each of the 25 sites along with the peak-hour peak-direction traffic characteristics and collision probabilities that were reported in previous tables. As noted above, P [h2=0|h1=1] is the crash reduction factor for a 25 mph speed limit. This is the probability that, other things being equal, a crash occurring to a test pedestrian under the street's current operating conditions would be avoided under a 25 mph maximum speed. Looking at this column the crash reductions vary from a low of 0.002 for site 58 to a high of 0.463 for site 22. Comparing the reduction probabilities to the average speed, we see that sites having an average speed below 25 mph tend to have low reduction probabilities and sites having higher average speeds tend to have higher reduction probabilities. This is as would be expected since a 25 mph speed limit should have little effect on a street where the majority of vehicles are already traveling at or below this limit.

The column P [h1=1] is simply the probability that a test pedestrian is hit and the entries in this column are identical to the collision probabilities appearing in Table 3.4. The last column, P [h2=0 and h1=1] gives the probability that a test pedestrian would be hit under the current conditions and that the same pedestrian would not be hit when 25 mph is the maximum speed. Since, in principle, the product of this last probability and the number of pedestrian entries would give a predicted number of crashes prevented we can call this last probability the crash potential of the speed policy. For example at site 11 we would expect 37 collisions per 1000 test pedestrian entries under the existing conditions, and other things equal, we would expect 25 mph speed limit to prevent 4 of these 37 collisions. The site with the greatest reduction potential is number 27b where the speed limit we would currently expect 64 collisions per 1000 entries, of which 9 would be prevented by the lower speeds. Increasingly, while P [h2=0|h1=1] depends primarily on average vehicle speed, while P [h1=1] depends primarily on traffic volume, P [h2=0 & h1=1] depends nontrivially on both of these traffic characteristics.

To summarize, estimated crash rates are often used to prioritize roadway sites as candidates for safety improvements. In roadside safety design where traffic crashes tend to be spatially sparse, an alternative method based on a physical model of a run-off-road crash, is used to help identify safety improvement projects. A similar problem arises when we attempt to prioritize local and residential streets with regard to traffic management activities because (a) measurements of pedestrian exposure are difficult to obtain in a cost-effective manner and (b) pedestrian crashed also tend to be spatially sparse. We have shown how the probability values related to a hypothetical pedestrian crash in which a selected type of pedestrian sets the stage for a midblock dart-out crash can be computed using (1.) vehicle speed and headway data collected automatically at the site of interest, (2) driver reaction time and braking rate data obtained from the literature, (3) building setback data measured from aerial photographs for the site of interest, and (4) pedestrian running speed data. The chief technical difficulty involves performing the numerical integrations that the method requires, but this is take care of by exploiting recent development of Monte Carlo computational method.

Table 3.1

Locations, Average Setbacks and Daily Traffic Volume for Sample Sites

Site ID	Street	From	To	City	Setback (ft)	Daily Traffic
11	Arden	Bridge	Country Club	Edina	53.3	1042
12	Drexel	SunnySide	Bridge	Edina	51.1	499
13	Bruce	SunnySide	Bridge	Edina	52.8	360
17	Irving	Lincoln	Franklin	Minneapolis	41.8	1143
18	Fremont	Franklin	22nd	Minneapolis	47.8	951
2	Humboldt	22nd	24th	Minneapolis	39.0	565
20	Princeton	Savannah	Ticonderoga	Eagan	42.5	301
21	Savannah	Gibraltor	“bend”	Eagan	57.4	530
22	Pennsylvania	States	“bend”	Eagan	41.0	176
24	35th Ave	53rd	54th	Minneapolis	41.0	126
27a	Highland EB	Cleveland	Cretin	St. Paul	46.7	755
27b	Highland WB	Cleveland	Cretin	St. Paul	46.7	911
3	Edina Blvd	Bridge	Country Club	Edina	66.8	973
32	Hill Crest	Davern	Fairview	St. Paul	47.7	209
34	16th Ave	69th	70th	Richfield	49.4	179
38	Boardman	34th	38th	Minneapolis	45.2	245
4	Wooddale	Bridge	Country Club	Edina	62.6	1512
40	Chicago	84th	86th	Bloomington	46.9	188
45	10th Ave	86th	88th	Bloomington	51.8	154
46	Parkwood Rd	Schaefer	Blake	Edina	74.7	503
49	17th Ave	67th	68th	Richfield	48.6	148
50	31st Ave	55th	56th	Minneapolis	46.6	388
55	Bryant	24th	25th	Minneapolis	37.6	1328
58	Bayard	Wheeler	Fairview	St. Paul	50.0	296
61	11th Ave	88th	90th	Bloomington	45.0	174

Table 3.2

Peak-Hour, Peak-Direction Characteristics of Sample Sites

Site ID	Peak Hour Peak-Direction				Sample Size	Speed (mph)		Headway (Log-sec)	
	Date	Hour	Direction	Volume		Mean	S.D.	Mean	S.D.
11	9/10	5 PM	NB	61	57	27.8	3.8	3.5	1.3
12	9/10	5 PM	NB	67	64	25.4	4.4	3.4	1.2
13	9/12	6 PM	NB	21	17	22.4	4.0	4.3	1.3
17	8/11	7 AM	NB	75	46	20.7	3.3	3.0	1.1
18	8/10	5 PM	NB	70	57	24.8	4.5	3.4	1.0
2	8/10	5 PM	SB	41	33	25.8	4.3	3.6	1.3
20	8/1	5 PM	SB	29	27	26.7	4.6	4.1	1.3
21	8/1	5 PM	EB	41	33	22.6	4.9	3.5	1.3
22	7/31	5 PM	NB	10	9	42.0	11.5	5.4	1.2
24	8/15	6 PM	SB	8	5	27.5	2.6	4.3	1.5
27a	7/18	5 PM	EB	78	78	28.3	3.8	3.3	1.2
27b	7/18	5 PM	WB	125	120	28.5	4.0	2.8	1.1
3	9/10	10 AM	NB	55	53	26.8	4.0	3.8	1.0
32	7/28	9 AM	WB	11	10	23.4	5.0	5.2	1.0
34	7/25	4 PM	SB	18	17	23.1	6.2	4.7	1.5
38	8/12	5 PM	WB	17	15	28.2	6.0	4.5	1.5
4	9/10	12 PM	SB	82	68	28.2	3.9	3.0	1.1
40	7/31	4 PM	SB	12	11	30.3	4.8	4.9	1.3
45	7/30	12 PM	NB	17	15	25.7	6.7	4.7	1.5
46	9/12	5 PM	EB	80	76	29.3	3.9	3.3	1.3
49	7/24	6 PM	NB	10	4	22.2	3.3	5.0	0.5
50	8/12	4 PM	NB	26	25	29.0	3.4	4.0	1.5
55	8/11	1 PM	NB	86	81	20.9	4.7	2.8	1.2
58	7/28	9 AM	EB	24	18	20.2	2.8	3.8	1.4
61	7/29	5 PM	NB	10	9	21.0	5.8	5.4	1.5

Table 3.3

Normality Tests for Speed and Log-Headway Distributions

Site ID	Sample Size	Speed Distribution		Log-Headway Distribution	
		Correlation (R)	Significance (P)	Correlation (R)	Significance (P)
11	57	0.984	0.099	0.971	0.015
12	64	0.986	>0.1	0.977	0.031
13	17	0.970	>0.10	0.980	>0.10
17	46	0.987	>0.10	0.990	>0.10
18	57	0.978	0.028	0.997	>0.10
2	33	0.973	0.054	0.990	>0.10
20	27	0.985	>0.10	0.974	>0.10
21	33	0.987	>0.10	0.957	0.023
22	9	0.936	>0.10	0.969	>0.10
24	5	0.936	>0.10	0.924	>0.10
27a	78	0.993	>0.10	0.987	0.099
27b	120	0.997	>0.10	0.993	>0.10
3	53	0.982	0.093	0.984	>0.10
32	10	0.984	>0.10	0.954	>0.10
34	17	0.954	0.086	0.976	>0.10
38	15	0.965	>0.10	0.939	0.055
4	68	0.972	<0.01	0.086	>0.10
40	11	0.968	>0.10	0.968	>0.10
45	15	0.950	0.078	0.923	0.027
46	76	0.991	>0.10	0.979	0.021
49	4	0.981	>0.10	0.987	>0.10
50	25	0.981	>0.10	0.983	>0.10
55	81	0.988	0.094	0.987	0.097
58	18	0.983	>0.10	0.989	>0.10
61	9	0.937	>0.10	0.969	>0.10

Table 3.4

Probability of Collision, Probability of Severe Collision and Probability of Severe Injury Given a Collision Computed Using Parametric and Nonparametric Models for Speeds and Headways.

Site ID	Parametric Model			NonParametric Model		
	P [collision]	P [severe]	P [severe|collision]	P [collision]	P [severe]	P [severe|collision]
11	.037	.013	.360	.037	.015	.401
12	.034	.010	.308	.034	.012	.360
13	.013	.003	.225	.013	.003	.235
17	.044	.009	.210	.043	.009	.207
18	.032	.010	.302	.030	.0008	.278
2	.041	.013	.328	.045	.014	.306
20	.027	.009	.330	.031	.009	.300
21	.026	.007	.253	.024	.006	.256
22	.010	.006	.593	.009	.006	.628
24	.027	.009	.341	.023	.007	.282
27a	.046	.017	.375	.047	.017	.357
27b	.064	.025	.393	.069	.027	.392
3	.018	.006	.318	.018	.006	.364
32	.004	.001	.287	.005	.001	.275
34	.013	.004	.274	.011	.003	.279
38	.022	.008	.371	.034	.014	.406
4	.039	.014	.361	.038	.015	.395
40	.011	.004	.398	.011	.004	.382
45	.013	.004	.314	.015	.004	.241
46	.030	.012	.404	.034	.014	.414
49	.004	.001	.267	.003	.001	.247
50	.032	.013	.391	.037	.016	.429
55	.059	.014	.237	.069	.016	.233
58	.025	.005	.195	.025	.004	.162
61	.007	.001	.209	.004	.001	.264

Table 3.5

Peak-Hour Peak-Direction Traffic Conditions Versus Collision-Related Probabilities

Site ID	Setback (ft)	Peak-Hour Peak-Direction			Collision-Related Probabilities		
		Volume	Average	S.D.	P [h2=0|h1=1]	P [h1=1]	P [h2=0 and h1=1]
11	53.3	61	27.8	3.8	.116	.037	.004
12	51.1	67	25.4	4.4	.080	.034	.003
13	52.8	21	22.4	4.0	.029	.013	.000
17	41.8	75	20.7	3.3	.006	.044	.000
18	47.8	70	24.8	4.5	.070	.032	.000
2	39.0	41	25.8	4.3	.085	.041	.003
20	42.5	29	26.7	4.6	.110	.027	.003
21	57.4	41	22.6	4.9	.040	.026	.001
22	41.0	10	42.0	11.5	.463	.010	.005
24	41.0	8	27.5	2.6	.106	.027	.003
27a	46.7	78	28.3	3.8	.113	.046	.005
27b	46.7	125	28.5	4.0	.140	.064	.009
3	66.8	55	26.8	4.0	.110	.018	.002
32	47.7	11	23.4	5.0	.058	.004	.000
34	49.4	18	23.1	6.2	.069	.013	.001
38	45.2	17	28.2	6.0	.167	.022	.004
4	62.6	82	28.2	3.9	.136	.039	.005
40	46.9	12	30.3	4.8	.204	.011	.002
45	51.8	17	25.7	6.7	.125	.013	.002
46	74.7	80	29.3	3.9	.156	.030	.005
49	48.6	10	22.2	3.3	.017	.004	.000
50	46.6	26	29.0	3.4	.151	.032	.005
55	37.6	86	20.9	4.7	.021	.059	.001
58	50.0	24	20.2	2.8	.002	.025	.000
61	45.0	10	21.0	5.8	.041	.007	.000

Figure 3.1

Hypothetical Crash between a Vehicle and a Pedestrian

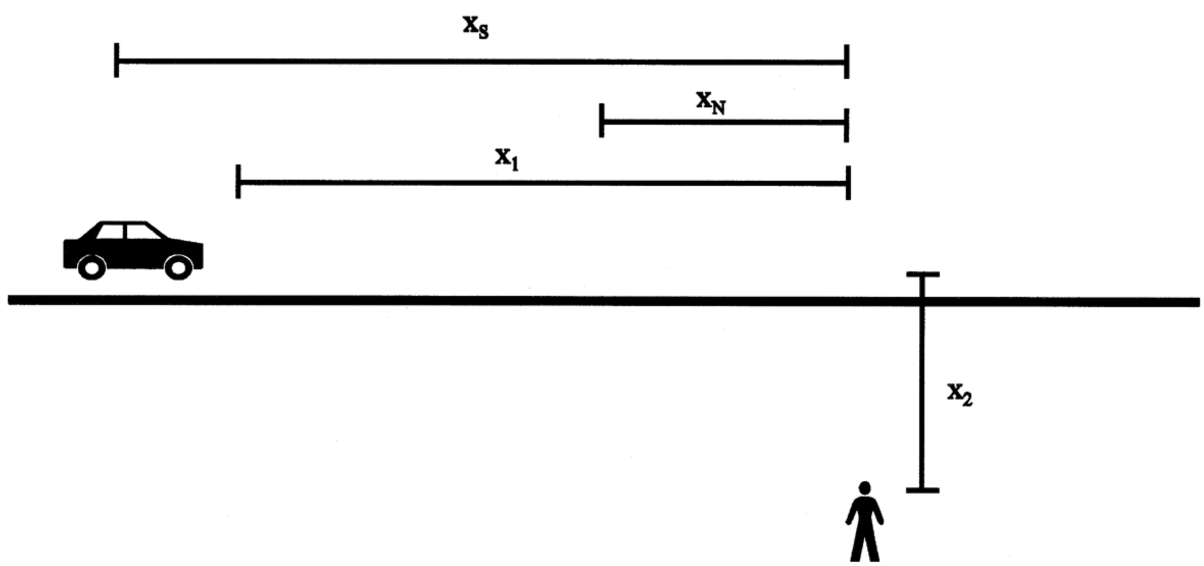


Figure 3.2

Directed Acyclic Graph (DAG) Model Showing Inter-relationships or Collision Model Variables.

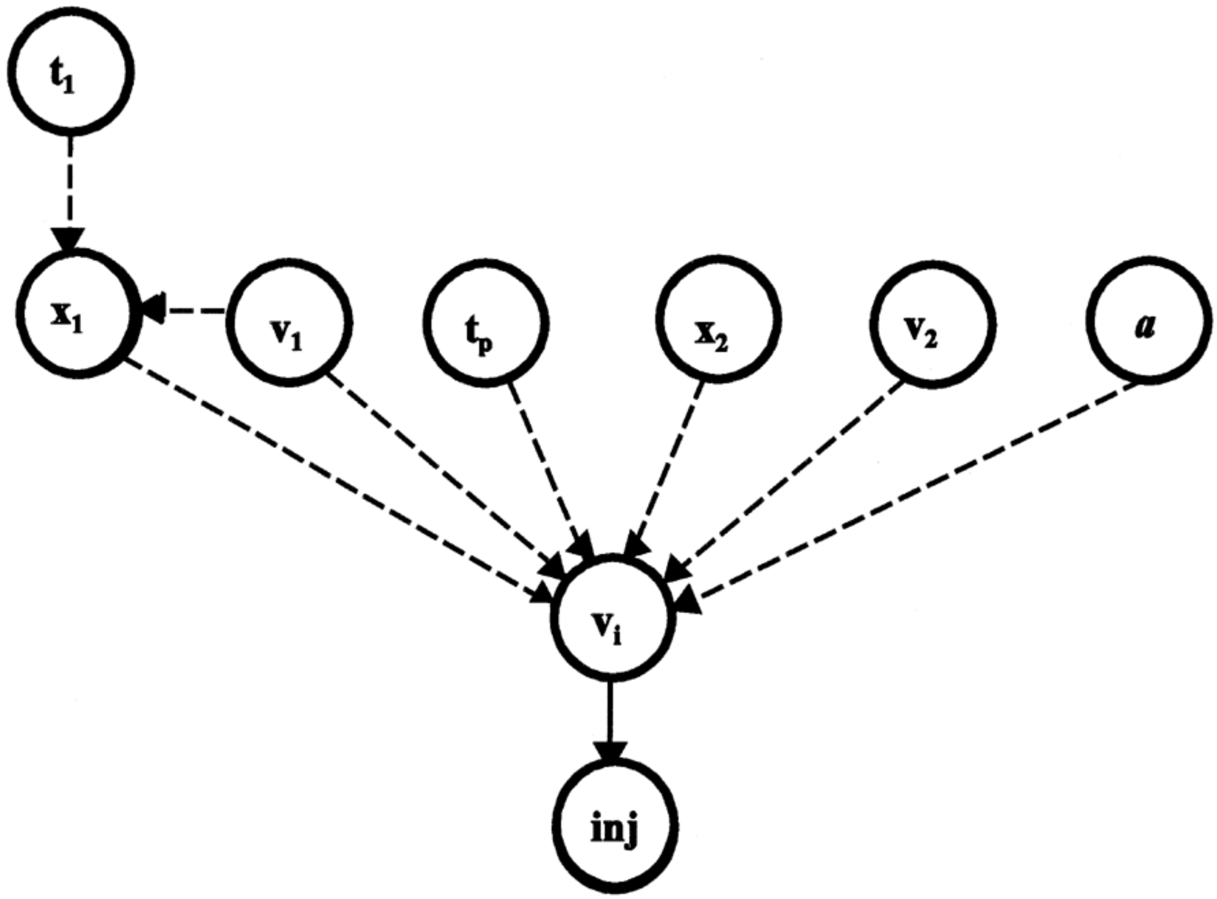
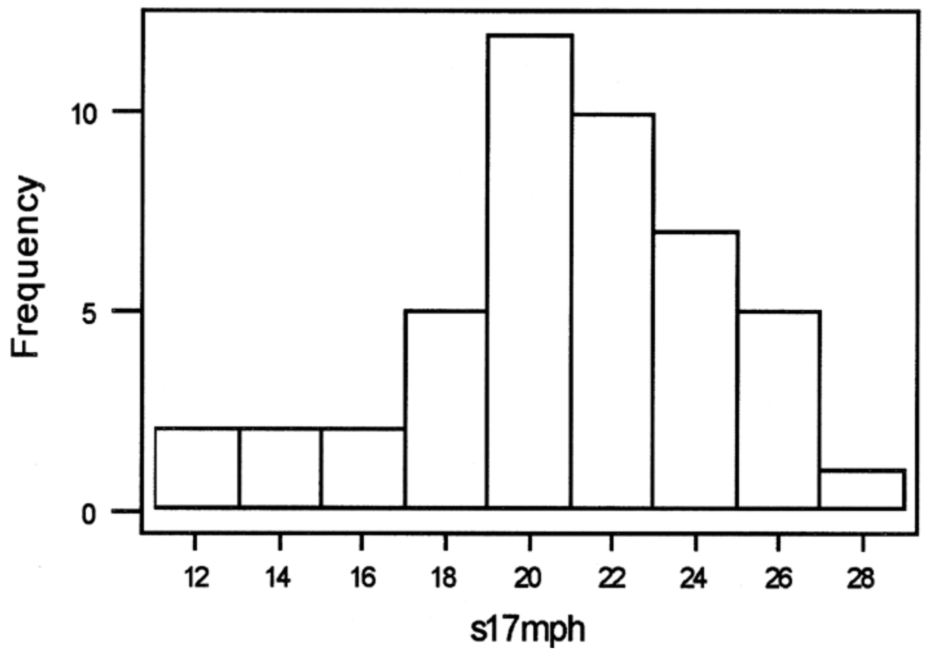


Figure 3.3

Histograms Showing the Distribution of Individual Vehicle Speeds and Headways at Site 17.



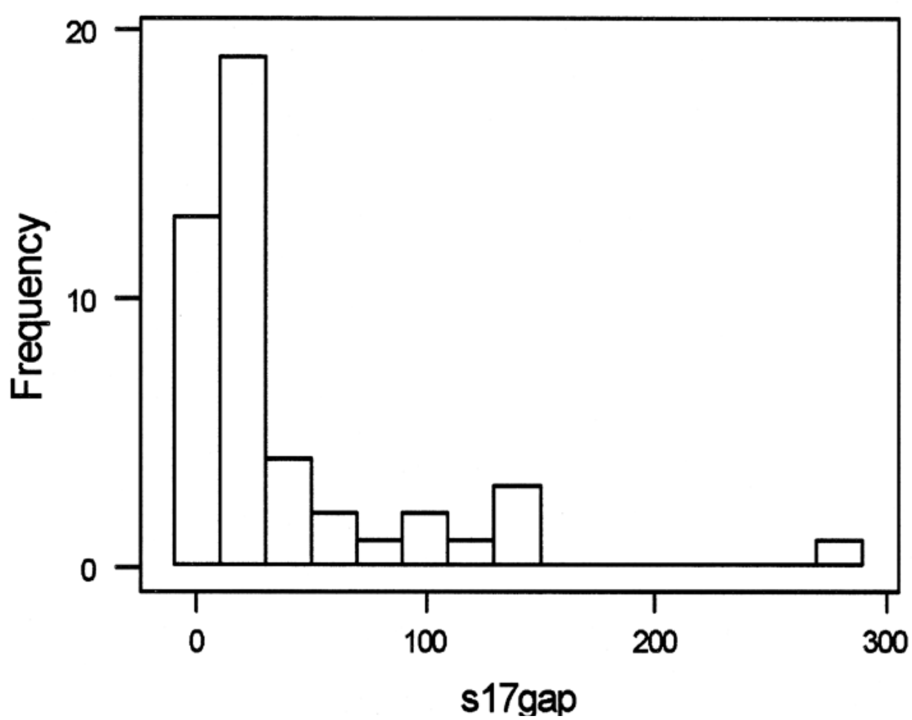
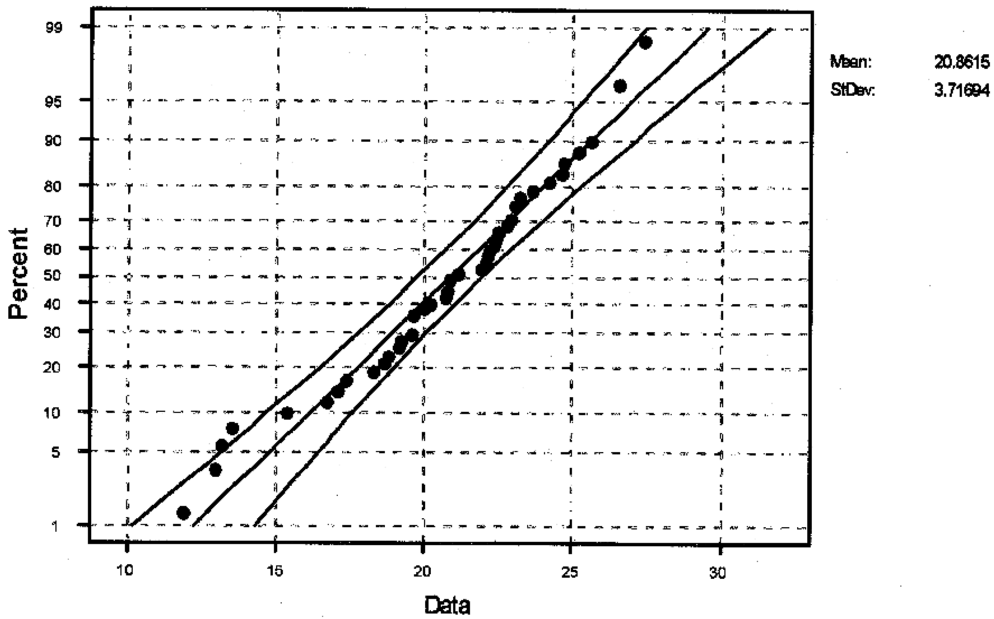


Figure 3.4

Probability Plots for Individual Speeds and Headways at Site 17

Normal Probability Plot for s17mph



Exponential Probability Plot for s17gap

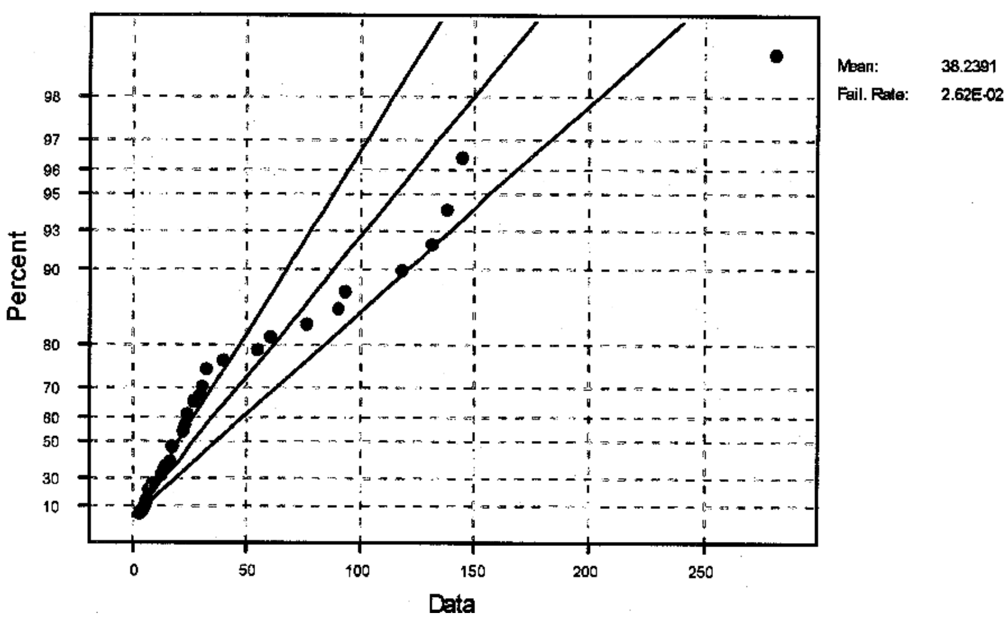


Figure 3.5

Parametric (phit1) Versus Nonparametric (phit2) Estimates of Collision Probabilities at Sample Sites.

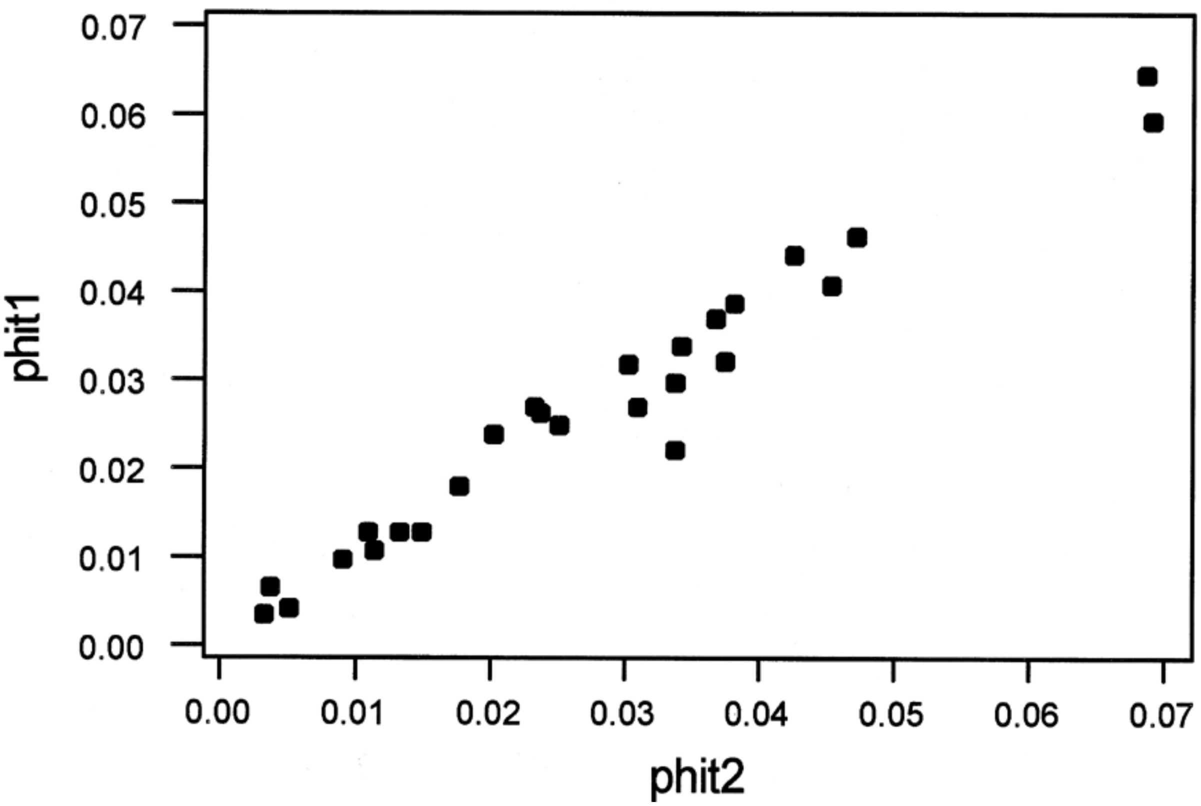
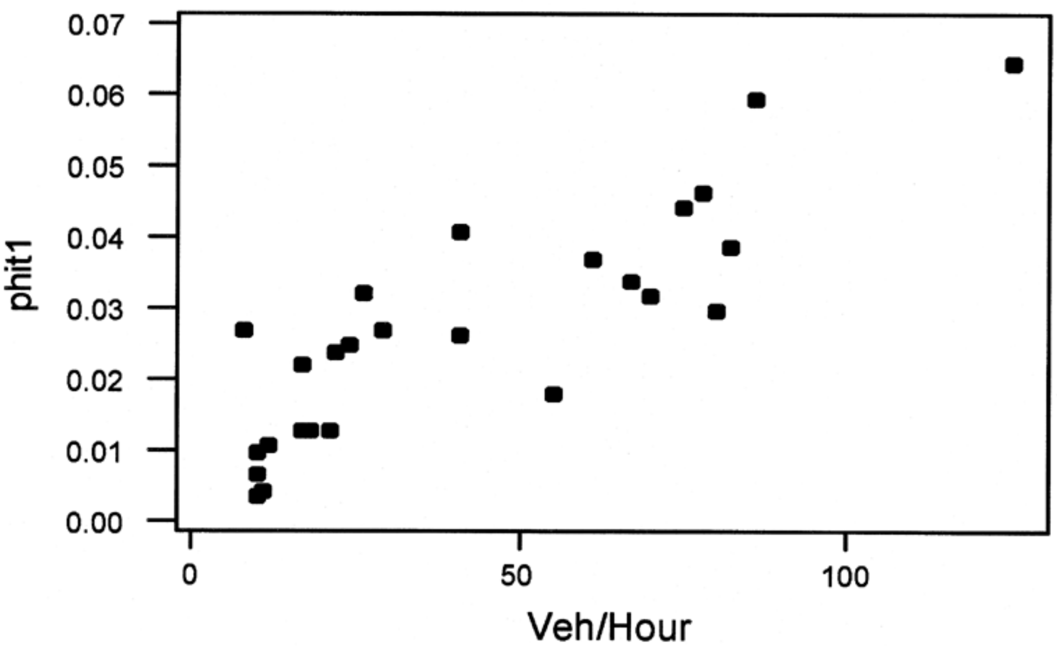


Figure 3.6

Parametric Collision Probabilities (phit1) Versus Traffic Volume and Average Traffic Speed at Sample Sites



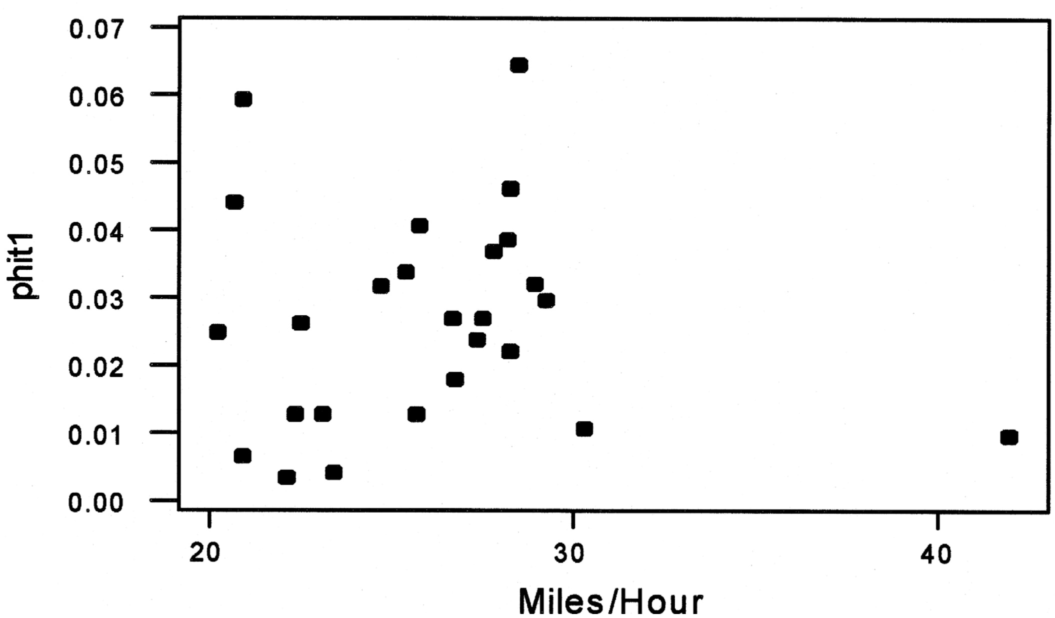
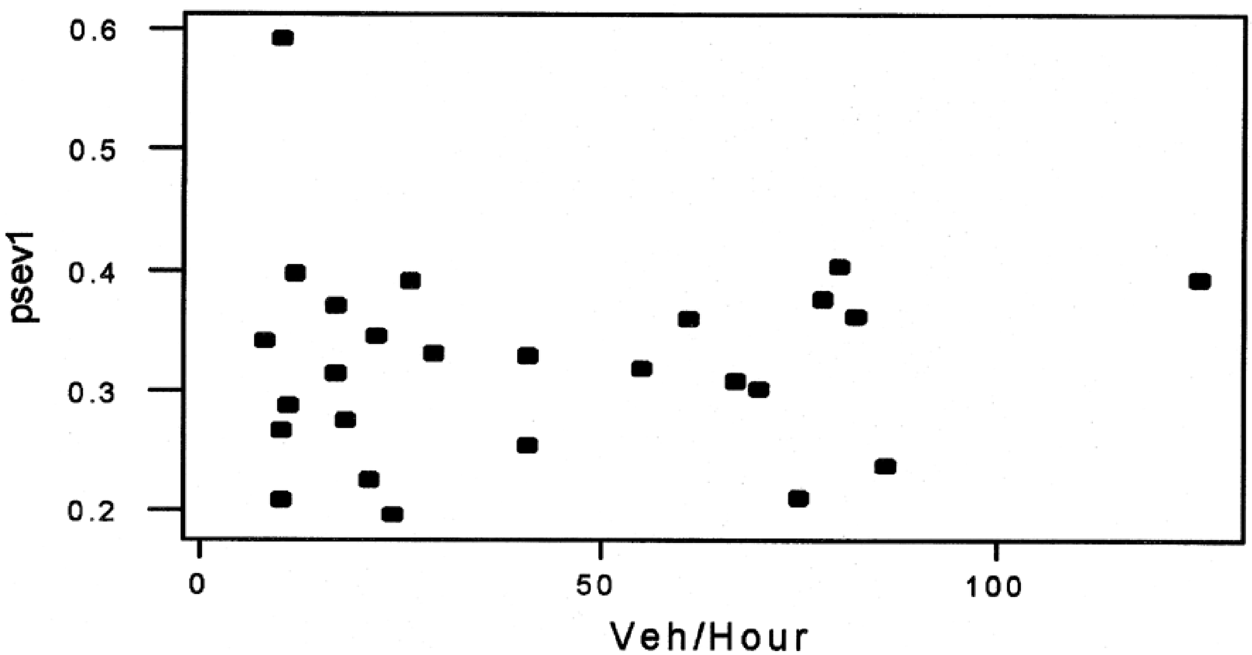
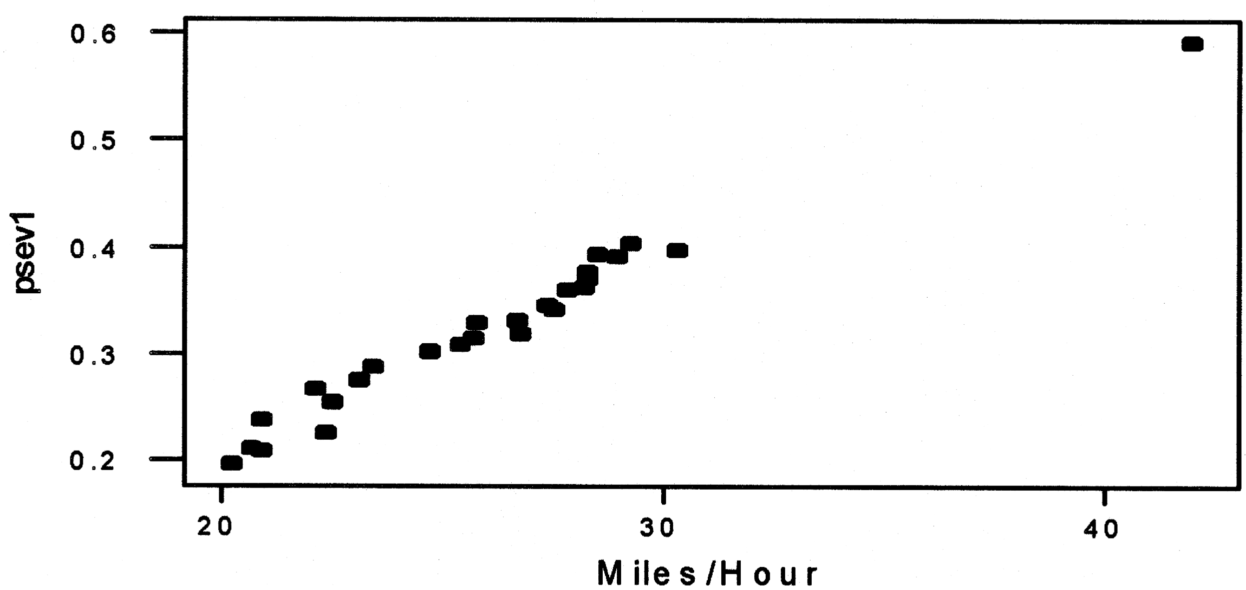


Figure 3.7

Probabilities of Severe Injury Given a Crash Occurs (psev1) Versus Traffic Volume and Average Speed at Sample Sites





CHAPTER FOUR

ESTIMATING THE CASUAL EFFECT OF SPEEDING IN ACTUAL VEHICLE/PEDESTRIAN COLLISIONS

As noted earlier we are in the midst of a vigorous international debate concerning the relationship (if any) between speed, speed limits and the probability of being involved in a crash. In Chapter 3 we saw that the commonly used measure of a countermeasure's safety effect, the crash reduction factor, can be interpreted as a probability of necessity, estimated from data collected with good experimental controls. We also saw that the lack of reliable measures of pedestrian exposure together with the sparsity of pedestrian crashes making it difficult to develop empirical estimates of the effect of speeding in pedestrian crashes. The solution proposed in Chapter 3 was to use a structural model of how such crashes occur in order to simulate the effect of a standardized by hypothetical test. This was used to compute estimates of collision probabilities and the Probabilities of Necessity corresponding to a 25 mph speed limit. In this Chapter, we will show how to use a similar structural approach to compute Probabilities of Necessity for actually occurring rather than hypothetical pedestrian crashes.

A formal statement of the problem at hand can be given using Rubin's Causal Model (Rubin 1974, Holland 1986). Let j = 1, …, N index a set of vehicle/pedestrian conflicts which include actually occurring collisions as well as all instances where a driver had to brake in order to avoid hitting a pedestrian. For each conflict, define the potential response variables:

$h1_j$	=1 if pedestrian j was actually struck
	=0 if pedestrian j was not struck
$h2_j$	=1 if pedestrian j would have been struck under strict adherence to the posted speed limit
	=0 if pedestrian j would not have been struck under strict adherence

Assuming enforcement of the speed limit causes obedience to the limit the change in crash frequency that would result if speed limits had been enforced would then be:

$$\Delta = \sum_{(j=1)}^{N}\left(h1_j - h2_j\right) = \sum_{(j:h1_j=1)}\left(h1_j = h2_j\right) + \sum_{(j:h1_j=0)}\left(h1_j - h2_j\right) \qquad (4.1)$$

With the first sum on the right hand side of equation (4.1) giving crashes prevented by enforcement, while the second sum give crashes caused by enforcement. If drivers who were normally traveling at or below the speed limit would not have increased their speeds under a condition of strict enforcement, the mechanic of vehicle braking indicates that the monotonic condition $h1_j \geq h2_j$ is plausible and the total crashes caused by enforcement will be zero. The change in crash frequency then becomes:

$$\Delta = \sum_{(j:h1_j=1)}(h1_j - h2_j) = n - \sum_{(j:h1_j=1)} h2_j \qquad (4.2)$$

Where n is the number of observed crashes.

The $h2_j$ reflect counterfactual outcomes and so cannot be directly observed and the problem becomes one of producing a defensible estimate of their sum. As noted earlier, the most common solution employed in traffic engineering is based on the assumption that under speed limit enforcement each crash has a common probability $\vartheta = P[h2_j = 0 \mid h1_j = 1]$ of being prevented. The $h2_j$ can then be modeled as Bernoulli outcomes and the expected reduction would equal to $n\vartheta$. The quantity ϑ is an example of what traffic engineers call a "Crash Reduction Factor" and what Pearl (2000) calls a "Probability of Necessity". The availability of an estimate ϑ from a previous study which (a.) satisfies the conditions of exogeneity and montonicity (Pearl 2000, p292) and for which (b.) the road in question could be considered as exchangeable with the units of the study (Draper et al. 1993) would then justify the standard practice of using $\Delta_1 = n\vartheta$ to estimate the crash reduction. Unfortunately, the majority of available estimates of crash reduction factors have been based on observational studies for which the validity of condition (a.) is questionable (Hauer 1980, 1997; Davis 200b), and even more troublesome is the fact that studies of speeding and pedestrian crashes on minor or local streets satisfying condition (b.) are simply unavailable (TRB 1998, p58).

A still actively debated topic in philosophy concerns the relation (if any) between general casual claims, such as "Drinking poison kills", and particular casual claims, such as "Socrates was killed by drinking hemlock". Although resolution of this debate should not be expected any time soon, Pearl (2000) has argued that different types of casual claims can be usefully distinguished according to the degree to which they presuppose situation-specific information. Then generic crash reduction factor: $\vartheta = P[h2_j = 0 \mid h1_j = 1]$ which assumes very little about either the details concerning crash j or about the mechanisms underlying this type of crash is thus closer to a general rather than a specific causal claim. It may be however that crash investigations have produced data relevant to the conditions of this crash and by coupling this with prior knowledge about crash mechanisms it may be possible to compute individual probabilities of necessity, P [$h2_j = 0 \mid h1_j = 1, y_j$] where y_j denotes the data available for crash j. This leads to an alternative estimate of the expected crash reduction:

$$\Delta_2 = \sum_{(j:h1_j=1)} E\left[h1_j - h2_j \,\middle|\, h1_j = 1, y_j\right] = \sum_{(j:h1_j=1)} P\left[h2_j = 0 \,\middle|\, h1_j = 1, y_j\right] \quad (4.3)$$

Questions concerning individual crash causation frequently arise in legal contexts where a jury must assess "Cause in Fact", (Robertson 1997) and so it is not surprising that over the past 60 years the discipline of traffic accident reconstruction has developed primarily to assist the legal system in answering such questions (Hicks 1989). The definition of "Cause" is generally accepted in reconstruction work has been given by Baker (1975) who first defines a casual factor as "… any circumstances contributing to a result without which the result could not have occurred", (p.274), and then defines the cause of a crash as the complete combination of factors which, if reproduced, would result in another identical crash (p.284). Although determination of the complete cause is often impossible, the more limited objective of identifying salient casual factors can sometimes be accomplished. Accident reconstruction has also, on occasion, been employed in support of traffic safety research (e.g. Treat et al. 1979) and, in fact, researchers at the University of Adelaide's Road Accident Research Unit have used standard deterministic reconstruction methods to assess the degree to which speeding was a casual factor in fatal vehicle/pedestrian crashes (Anderson et al. 1997).

To illustrated how deterministic accident reconstruction might be used to answer questions concerning individual causation. Consider vehicle/pedestrian crashes that occur as follows:

The driver of a vehicle traveling at a speed, v, notices an impending collision with a pedestrian when the front of the vehicle is a distance, x, from the potential point of impact. After a perception/reaction time of t_p the driver locks the brakes, the vehicle decelerates with braking coefficient of friction equal to f and after a transition time of t_s, (Neptune et al. 1995) the tires begin making skid marks. The vehicle comes to a stop leaving skid mark of length *sl*. Before stopping, the vehicle strikes the pedestrian at a speed of *vi* and the measured skid mark from the point of impact to the vehicle's stopping point is *s2*. Variants of the kinematical formula:

$$s = vt + at^2)/2$$

Giving the distance, *s*, traveled during time interval, *t*, by an object with initial velocity, *v*, and undergoing a constant acceleration of *a* can then be used to determine whether or not speeding was a factor for this crash. For example: suppose that tire skid marks left at the scene indicated that the vehicle began skidding 20 meters (65.6 ft) before hitting the pedestrian and after impact continues to skid for another 10 meter (32.8 ft) before coming to a stop. Using $f = 0.7$ (so that a = – 6.9 meters/second^2 (–22.6 ft/sec^2)) gives the speed of the vehicle at the start of the skid mark as 20.3 meters/second (66.6 fps) and using $t_s = 0.2$ seconds gives the speed at the start of braking as 21.7 meters/second (71.2 fps). Then, using $t_p = 1.5$ seconds, one deduces that the vehicle was 56.7 meters (186 ft) from the collision point when the driver notices the pedestrian. If the vehicle has been traveling instead at the posted speed of 60 km/h (37 mph), the driver would have needed only 45.4 meters (149 ft) to stop and so, other things equal, would not have hit the pedestrian. At this point though it must be pointed out that there is no privileged status attached to the values $f = 0.7$, $t_s = 0.2$ seconds and $t_p = 1.5$ seconds, and that if one uses the not unreasonable values $f = 0.5$, $t_s = 0.2$ seconds and $t_p = 1.0$ seconds, one finds that the vehicle still hits the pedestrian when the initial speed is 60 km/h (37 mph). So, whether or not speeding was a casual factor in this crash remains uncertain.

It is often that the variables appearing in a casual assessment are underdetermined by variable measurements and constrains, and the problem of how to proceed when one cannot defend a particular nominal value for an unmeasured causal variable has received consideration in the accident reconstruction literature. The standard recommendation is to perform a sensitivity analysis in which estimates are recomputed as selected input variables vary over selected ranges (Fricke 1990, Niederer 1991). When the basic conclusions of the reconstruction (e.g. that the driver was speeding) are insensitive to this variation they can be regarded as well supported but if the basic conclusion differ for different but plausible combinations of input values, as happened above, this approach is inconclusive. It has been recognized though that when uncertainty is described using probability measures (Lindley 1987) more discriminating approaches are available. Brach (1994) illustrates how the method of statistical differentials can be used to compute the variance of an estimate if one has a differentiable expression for the desired estimate and if one can specify means and variances for distributions characterizing the uncertainty in the expression's arguments. Kost and Werner (1994) and Wood and O'Riordain (1994) illustrate how Monte Carlo simulation can be used for more complicated models where differentiable, closed form solutions are not available and where different measurements combine to provide information on a target variable. This approach also produces approximate probability distributions characterizing the uncertainty of estimates rather than just variances. Particularly interesting is the method used by Wood and O'Riordain where simulated outcomes inconsistent with measurements were rejected, in essence, producing posterior conditional distributions for the quantities of interest.

The deterministic assessment of whether or not obeying a special speed limit would have prevented a pedestrian crash consisted of three steps:

(1.) Estimating the vehicle's initial speed and location using the measured skid marks and nominal values for f, t_s and t_p

(2.) Setting the vehicle's initial speed to the counterfactual value

(3.) Using the same values of f, t_s and t_p along with the counterfactual speed to predict if the vehicle would then have stopped before hitting the pedestrian.

Step (1.) falls under what is generally regarded as accident reconstruction proper but step (2.) and (3.) are examples of what has been referred to variously as avoidance analysis (Limpert 1989), sequence of events analysis (Hicks 1989) or the hypothetical crash outcome method (Kloeden et al. 1997). Uncertainty assessment in accident reconstruction has for the most part been limited to step (1.) but the process outlined in steps (1-3) above can be seen as a special case of a method for counterfactual analysis that has been formalized by Pearl and his associates (Balke and Pearl 1994, Pearl 2000). In fact, Pearl refers to the operations in steps (1-3) as respectively abduction, action and prediction. To apply Pearl's method to the pedestrian crash problem an analyst would have to supplement the deterministic model with a prior probability distribution for the variables x, v, t_p, t_s and f along with a directed acyclic graph (DAG) model for the dependency structure among the model variables forming a Bayesian network. Bayes theorem could then be used to compute a posterior distribution conditioned on the skid mark length. The initial speed would then be set to its counterfactual value and the probability that the vehicle stops before hitting the pedestrian could then be computed using the posterior distribution for t_p, t_s, x and f. Balke and Pearl (1994) have further pointed out that technical difficulties arising from the need to describe or store the posterior distribution can be circumvented by applying methods for updating Bayesian networks to a DAG model that has been augmented to include nodes representing counterfactual outcomes.

An alternative estimate of the crash reduction due to prevention of speeding can be summarized by computing the corresponding "Probability of Necessity" for each set of crashes. This is an instance for assessing the "Cause of an Effect" which is generally not possible absent of some prior knowledge of causal structure (Dawid 2000, Pearl 2000). For traffic crashes, such prior knowledge is often available and forms the basis of the discipline of accident reconstruction.

In what follows, we will use an extended version of the collision model described above which allows for data on pedestrian injury and throw distance as well as skidding distances. As before, the driver of a vehicle traveling a speed of v notices an impending collision with a pedestrian when the front of the vehicle is at distance x from the potential point of impact. After a perception/reaction time of t_p the driver locks the brakes and after a transaction time to t_s the tires begin making skid marks. The vehicle comes to a stop leaving skid mark of length *sl*. Before stopping, the vehicle strikes the pedestrian at a speed of *vi*, the pedestrian is thrown into the air and comes to rest at a distance of d from the point of impact. The pedestrian is injured and the severity of injury can be classed as in Chapter 2 and 3 as slight, serious or fatal. In addition, if the pedestrian was struck after the vehicle began skidding it may be possible to measure a distance of *s2* running from the point of impact to the end of the skid mark. Figure 4.1 illustrates the collision scenario with x_s denoting the distance traveled during the braking transient. The basic inference problem is to characterize the posterior uncertainty in causal variables such as *v, vi* and x given some subset of the measurements *d, sl* and the pedestrian's injury severity. Figure 4.2 displays a DAG summarizing the conditional dependence structure of the collision model. To complete the model it is necessary to specify deterministic or stochastic relations for the arrows appearing in Figure 4.2 and prior distributions for the background variables x, v, t_p, t_s and f.

As noted above, it is assumed that the vehicle's braking occurs in three stages, the first consisting of the driver's perception/reaction time, denoted t_p during which the vehicle continues traveling at its original speed. Given the initial distance between the point of perception and the point of impact, x, it can be determined if the impact occurs during this first stage, in which case the impact speed is equal to the initial speed. The second stage consists of a transient time, t_s, during which deceleration occurs but no skid mark is left, while the third stage consists of the locked-wheel stop which leaves a visible skid mark. A constant deceleration characterized by a friction coefficient, f, is assumed to apply to both the transient and skid mark phases and if the impact occurs during one of the later phases the impact speed is computed by decelerating the vehicle at a constant rate over the distance running from the end of the perception/reaction phase to the point of impact. More explicitly:

$$vi = \begin{cases} v & \text{if } x < vt_p \\ 0 & \text{if } x > vt_p + v^2/(2fg) \\ \sqrt{v^2 - 2fg(x - vt_p)} & \text{otherwise} \end{cases} \tag{4.4}$$

Where g denotes the gravitational acceleration. The theoretical length of the skid mark is then computed as the difference between the total braking distance and the distance traversed during the transition phase:

$$\textit{Theoretical skid mark} = v^2/(2fg) - \left(vt_s - fgt_s^2/2 \right) \tag{4.5}$$

Most commonly in reconstruction practice the measured skid mark is treated as a deterministic function of the speed at the beginning of the skid mark, that is, the possibility of measurement error is discounted. However, Garrot and Fuenther (1982) described results from a series of controlled braking tests reported that the standard deviation of measured skid marks tended to increase as the initial speed increased and that the average coefficient of variation for the difference between the measured and theoretical skid lengths was approximately equal to 0.11. In the reconstructions described later the measured skid mark, denoted by sl, was assumed to be a lognormal random variable with an underlying normal mean equal to the natural log of the theoretical length and underlying normal variance equal to 0.01 giving a coefficient of variation of the measurement error in the skid marks of about 0.10. This yields approximate consistency with the Garrot and Guenther findings. For those collisions where impact occurs after the vehicle starts to skid, a theoretical value for the length of skid beginning at the point of impact is computed by applying the braking distance formula to the impact speed, *vi*. The measured second skid, *s2*, is then also assumed to be lognormal with underlying normal mean equal to the natural log of the theoretical value and underlying variance equal to 0.01.

Next, in accident reconstruction "throw distance" is usually taken to mean the distance measured parallel to the vehicle's path between a pedestrian's final resting point and the point of impact with the vehicle. Although in principle a deterministic simulation model of a pedestrian's trajectory can be developed by careful application of the principles of mechanics (Van Wiji et al. 1983), in practice both the computations and especially the data needed for such simulations make them ill suited for most accident reconstructions. A considerable effort has been expended on trying to develop tractable physical models with acceptable accuracy, which relates throw distance to the speed of the vehicle at impact. Probably the most sophisticated treatment of this problem has been given by Wood (1988, 1991) while Eubank and Hill (1998) have assembled a non-critical compendium of these efforts. Rather than attempting to resolve this issue, the tactic used here will be to employ statistical modeling to relate impact speed and throw distance. Measured throw distances and impact speeds from 55 crash tests between cars and adult pedestrian dummies reported in several earlier studies and tabulated in Eubanks and Hill (1998, pp.704, 708, 710-711) were plotted and are shown in Figure 4.3. The increasing scatter in measured throw distance as impact speed increases suggested the possibility of a lognormal relationship and ordinary least-squares was used to fit a linear model relating the natural log of the measured throw to the natural log of the measured speed. The residuals from this fit passed a test for being normally distributed and the final fitted model took the form of:

$$\begin{aligned} log\,(d) &= -\underset{(0.30)}{(3.43)} + \underset{(0.09)}{(1.61)}\, log\,(vi) + e \end{aligned} \tag{4.6}$$

Where:

d = Throw distance in meters

vi = Vehicle speed at impact in km/h

e = Normal random variable with mean equal to 0 and estimated variance equal to 0.06

Estimated standard errors for the regression parameters are shown in parenthesis below the corresponding parameter values.

Modeling the relationship between impact speed, *vi*, and degree of injury severity was treated in detail in Chapter 2 and the ordered logit model developed there was used here. Weighted exogenous sample maximum likelihood estimates of the model parameters were used. For the adult (ages 15-59) pedestrian group when the impact speed is in km/h these were a_1 = 4.07 (0.73), a_2 = 7.21 (1.01) and b = 0.095 (0.02) where the approximate standard errors are given in parentheses after each estimate.

Finally, the Bayesian network requires prior distributions for the background variables: *x, v*, t_p, t_s and *f*. Although prior distributions for the background arbitrariness of using fixed nominal values, they bring with them the problem how to select these distributions in some reasonable manner. Because in most cases the background variables will be under-identified by the available data, objectivity arguments based on variants of Jefferys' prior or on stable estimation are not applicable and it is to be expected that the conclusions will depend on the chosen priors. Like the problem of general versus specific causation, the appropriate representation and use of prior information lead to foundational issues that are still being actively debated (Freedman 1997). Absent a resolution to this debate, one can, at least, use priors that on their face are consistent with accident reconstruction practice and then follow Berger's (1985, p112) recommendation of reporting models, priors, data and Bayesian conclusions. In deterministic sensitivity analysis, it is often possible to identify defensible prior ranges for background variables (Niederer 1991). Wood and O'Riordain argue that in the absence of more specific information uniform distributions restricted to these ranges offer a plausible extension of the deterministic sensitivity methods (1994, p137). Following Wood and O'Riordian's suggestion, the reconstructions described in this paper used uniform prior distributions. Specifically, the range for *f* was (0.45, 1.0) and was taken from Fricke (1990, p62-14) where the lower bound corresponds to a dry, travel polished asphalt pavement and the upper bound to a dry, new concrete pavement. The range for the perception/reaction time, t_p, was (0.5 seconds, 2.5 seconds) which brackets the values obtained by Fambro et al. (1997) in surprise braking tests and the midpoint of which (1.5 seconds) equals a commonly chosen nominal value (Stewart-Morris 1995). For the braking transient time, Neptune et al. (1995) reported values ranging between 0.1 and 0.35 seconds for a well-tuned braking system while Reed and Keskin (1989) reported values in the range of 0.4–0.5 seconds, so the chosen range was 0.1–0.5 seconds. The strategy for the initial distance and initial speeds was to use wide enough ranges that no reasonable possibility was excluded a prior. The range for v was [5 meter/second, 50 meters/second] (16.4 fps, 164.1 fps) and that for x was [0 meters, 200 meters] (0 ft, 656.2 ft). The upper bound for x is approximately the recommended design stopping sight distance for a speed of 100 km/h (62 mph) (AASHTO 1964). Since this uses a 2.5 second reaction time and braking friction of 0.29, initial distances outside this range should almost always result in stopping before collision. Finally, uncertainty in the parameters for the throw distance and injury severity models were incorporated by treating the actual parameter values as normal random variables with means and standard deviations equal to the estimated presented earlier. In Figure 4.2 the nodes labeled, p_1 and p_2 represent these parameters.

As indicated earlier, accident reconstruction proper involves using data collected from an individual crash to estimate unmeasured quantities of interest such as initial and impact speeds. Uncertainly analysis for reconstruction proper has received some attention in the literature but the connection between this problem and that of updating a Bayesian network does not appear to have been recognized. As a check on the reconstruction model described above and to illustrate the use of Bayesian network methods in reconstruction proper, a random sample of crash tests involving vehicles and pedestrian dummies detailed in Eubanks and Hill (1998) was drawn. Since a number of these tests involved side-wipe impacts, post impact braking or were otherwise inconsistent with the scenario described as above. These were rejected when drawn and the sampling continued until 15 crash tests consistent with the scenario were obtained. It should be noted that these crash tests were not part of the sample used to develop the throw distance relationship given in Equation 4.6. Skid mark lengths and throw distances were obtained from the data tabulated for each crash test and for each tests a measurement of the vehicle's speed at impact was also available. The pedestrian collision model was specified for the Gibbs sampling program BUGS with the two skid measurements and the throw distance providing the data. For each test, a 5,000 iteration burnin was followed by a 50,000 iteration Gibbs sample with the results from every 10th iteration being saved for analysis. Figure 4.4 displays the measured impact speeds from each of the sample tests along with posterior medians for the impact speeds and posterior 95% credible intervals. Note that in 10 of 15 crashes the measured impact speed was lower than the posterior median while the posterior 95% credible intervals captured the measured values in 14 out of 15 crashes. Although not exhaustive, these results are consistent with what one would expect if the collision model were reasonably well calibrated.

To estimate the crash reduction due to speed limit adherence one requires values for the individual Probability of Necessity appearing in equation (4.3) where y_j denotes measurements of some subset of *sl, s2, d* and the injury severity. These can be computed by (i.) updating the distributions for x, t_p, t_s and *f* using the data y_p (ii.) setting the initial speed equal to the speed limit and (iii.) computing the probability that *vi = 0* (since *h2 = 0* and only if *vi = 0*) using the updated distributions for x, t_p, t_s and *f* together with the condition that *v* is set to the speed limit. For DAG models that have casual interpretations, Balke and Pearl (1994) describe how by appropriately augmenting the DAG with additional nodes representing counterfactual variables Bayesian network methods can be used to compute counterfactual probabilities. Figure 4.2 also shows the augmented DAG needed to evaluate Equation 4.3 for the pedestrian collision model where *v*, vi** denote the counterfactual variables.

Detailed investigation of vehicle/pedestrian crashes from the Twin Cities were unavailable so the application of this approach will be illustrated using data collected by Kloeden et al. (1997) where accident reconstruction methods were used to estimate the speed of crash-involved vehicles as part of a study seeking to relate speed to crash risk. The crash study were carefully selected to exclude possible confounding due to alcohol use, medical conditions, illegal maneuvers and weather and all the investigated crashes occurred on roads with a 60 km/h (37 mph) speed limit. A sample of speeds from vehicles not involved in crashes showed an average speed of about 60 km/h (37 mph) with an 85% percentile speed of about 80 km/h (50 mph) so it could be argued that the speed limit on these roads should be raised to 80 km/h (50 mph). The question then is how many vehicle/pedestrian crashes would have been prevented if all vehicles had obeyed the 60 km/h (37 mph) speed limit.

The sample of investigated crashes included eleven collisions between vehicles and pedestrian and eight satisfied the conditions of the collision model described above, being frontal impacts by a single vehicle with evidence of pre-impact braking. Information on each of the investigated crashes was published in Volume 2 of Kloeden et al. including scale drawings showing the rest positions of pedestrians and vehicles, the length and location of skid marks and an approximate point of impact. From this information, measurements of *sl, s2* and *d* were obtained along the degree of injury suffered by the pedestrian, and Table 4.1 displays these data. As a comparison, deterministic estimates of the initial speed, the impact speed and the impact speed that would have occurred if the initial speed had been 60 km/h (37 mph) were computed using the midpoint values t_p = 1.5 seconds, t_s = 0.3 seconds, f = 0.725 and assuming *sl* and *s2* were measured without error. Using BUGS, posterior distributions were estimated with the collision model augmented to include the counterfactual variable *v** and *vi**, and with *v** set equal to 60 km/h (37 mph). Again, a 5,000 iteration burnin was followed by 50,000 iterations with every 10th iteration being saved for analysis. Three separate Gibbs sampling chains were generated from different initial values and random number seeds, convergence was checked using the Gelman and Rubin (1992) test and sample size was checked using the Raftery and Lewis (1992) test, as implemented in CODA (Best et al. 1995). Table 4.2 displays results from these simulations along with the corresponding deterministic estimates.

Looking first at the deterministic estimates in Table 4.2, it can be seen that a standard deterministic analysis would conclude that in three of the cases: 025, 027 and 154, the vehicle was clearly speeding and that had it been traveling at 60 km/h (37 mph), other things being equal, then the crash would have been prevented. For case 218, the vehicle was clearly traveling at below the speed limit, while for the remaining four cases the vehicle was traveling near the speed limit. For these last five cased speeding would not be regarded as a causal factor. Looking next at the results from the Gibbs sampler, it appears that the vehicles in cased 025, 027 and 254 were probably speeding and these crashes would probably have been prevented had the vehicle been traveling at 60 km/h (37 mph). The vehicle in case 218 was probably not speeding and the crash would probably not have been prevented by speed limit enforcement. For the remaining four cases the effect of speeding is uncertain. Of particular interest is the rightmost column of Table 4.2, which gives the individual probabilities of necessity for each of the crashes required by the estimator given in Equation 4.3. Summing the entries in this column gives an estimated crash reduction due to speed limit adherence of 3.8 compared to the deterministic finding that three crashes would have been prevented. Overall, the results from the Bayesian network approach were consistent with the results form a more standard deterministic analysis that used midpoints of the priors for t_s, t_p and f as nominal values. This is expected since both approaches use the same braking model. The results from the Bayesian network approach are arguably more robust however since they are not tied to specific nominal values for un-identified variables. If one admits the existence of crashes that were either not investigated or could not be reconstructed, the estimate of 3.8 crashes gives a lower bound on the crash reduction due to speed limit adherence. Although not addressed here, it would, in principle, be possible to compare the benefits derived from this estimated crash reduction to the overall increase in motorist travel time produced by speed limit adherence.

Table 4.1

Data for the Reconstruction of Eight Vehicle/Pedestrian Collisions

Case	*s1* (meters)	*S2* (meters)	*d* (meters)	Pedestrian Injury
Cn015	15.0	5.0	9.0	Fatal
Cn025	22.2	5.9	7.8	Serious
Cn027	20.0	7.8	11.5	Serious
Cn057	15.6	5.0	---	Serious
Cn074	13.4	1.1	4.2	Serious
Cn121	11.2	2.1	5.0	Serious
Cn154	21.0	10.0	16.3	Serious
Cn218	11.1	4.6	3.6	Serious

Table 4.2

Posterior Estimates for Eight Vehicle/Pedestrian Collisions

(All speeds are in km/h)

	Deterministic Estimates			Gibbs Sampler Estimates							
				Initial Speed (*v*)			Impact Speed (*vi*)			Probabilities	
Case	*V*	*vi*	*vi**	Mean	2.5%	97.5%	Mean	2.5%	97.5%	P [*v*>60]	P [*vi**=0]
cn015	60	30	30	64	49	79	32	26	39	0.71	0.45
cn025	72	33	0	73	55	91	33	26	40	0.90	0.76
cn027	68	38	0	71	54	87	39	31	46	0.87	0.65
cn057	61	30	26	63	47	78	31	24	38	0.64	0.43
cn074	57	14	14	62	50	73	16	13	18	0.63	0.55
cn121	57	21	21	59	45	73	24	19	28	0.42	0.29
cn154	70	43	0	72	55	89	45	36	59	0.91	0.63
cn218	53	29	29	50	39	65	28	22	34	0.09	0.03

Figure 4.1

Graphic Representation of a Vehicle/Pedestrian Collision

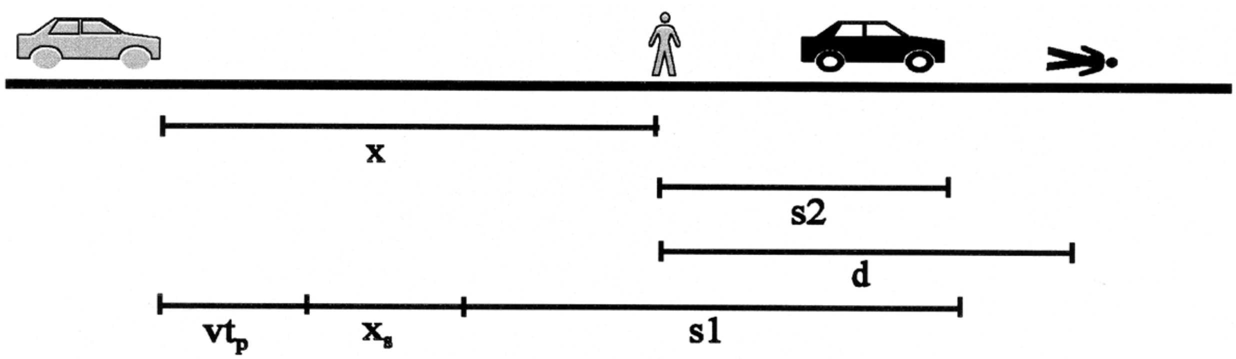


Figure 4.2

Directed Acyclic Graph (DAG) of Inter-relationships of Variables in the Reconstruction Model

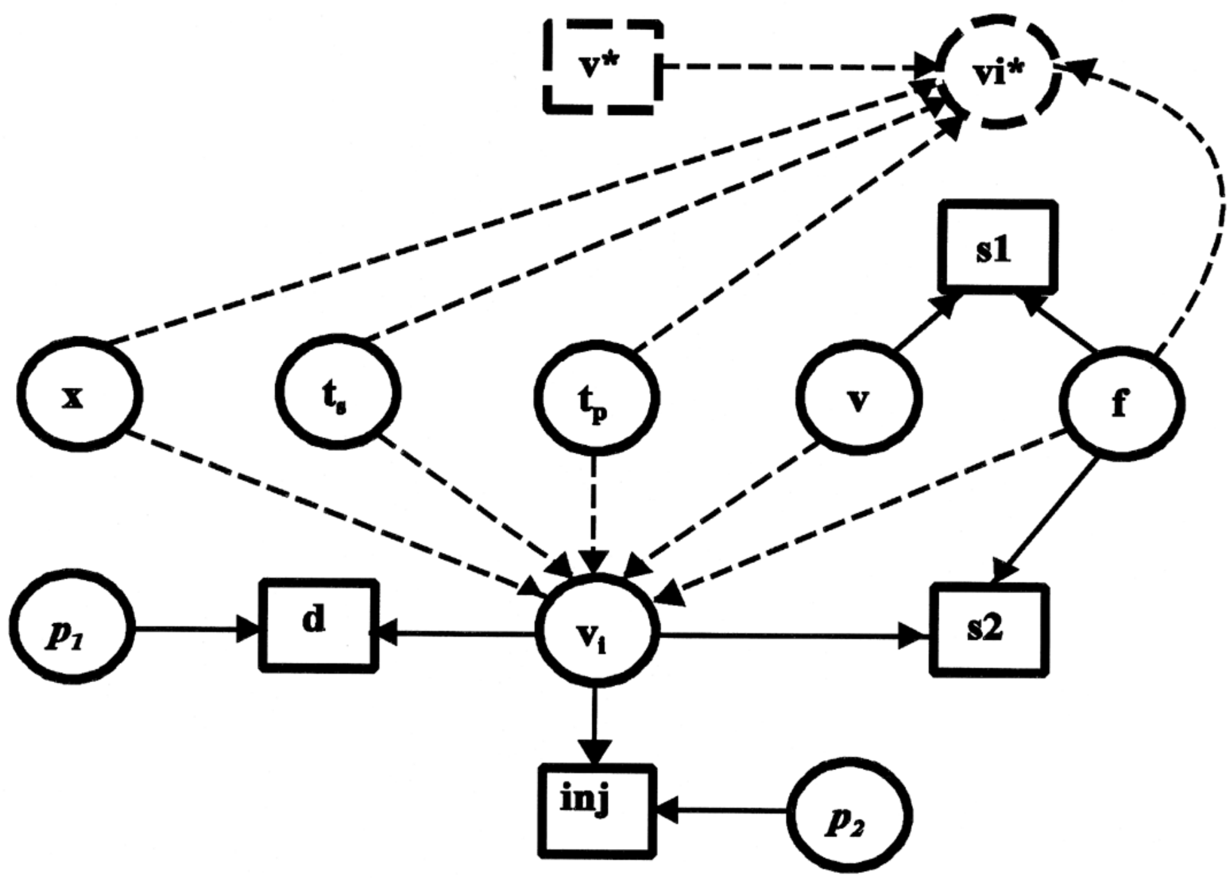


Figure 4.3

Fitting and Measured Throw Distances Versus Impact Speed for 55 Pedestrian Dummy Collisions.

(Throw is in meters. Speed is in Kilomer/Hour)

Measured Throw (*x*)

Mean Throw (___)

90% Confidence Region (- - -)

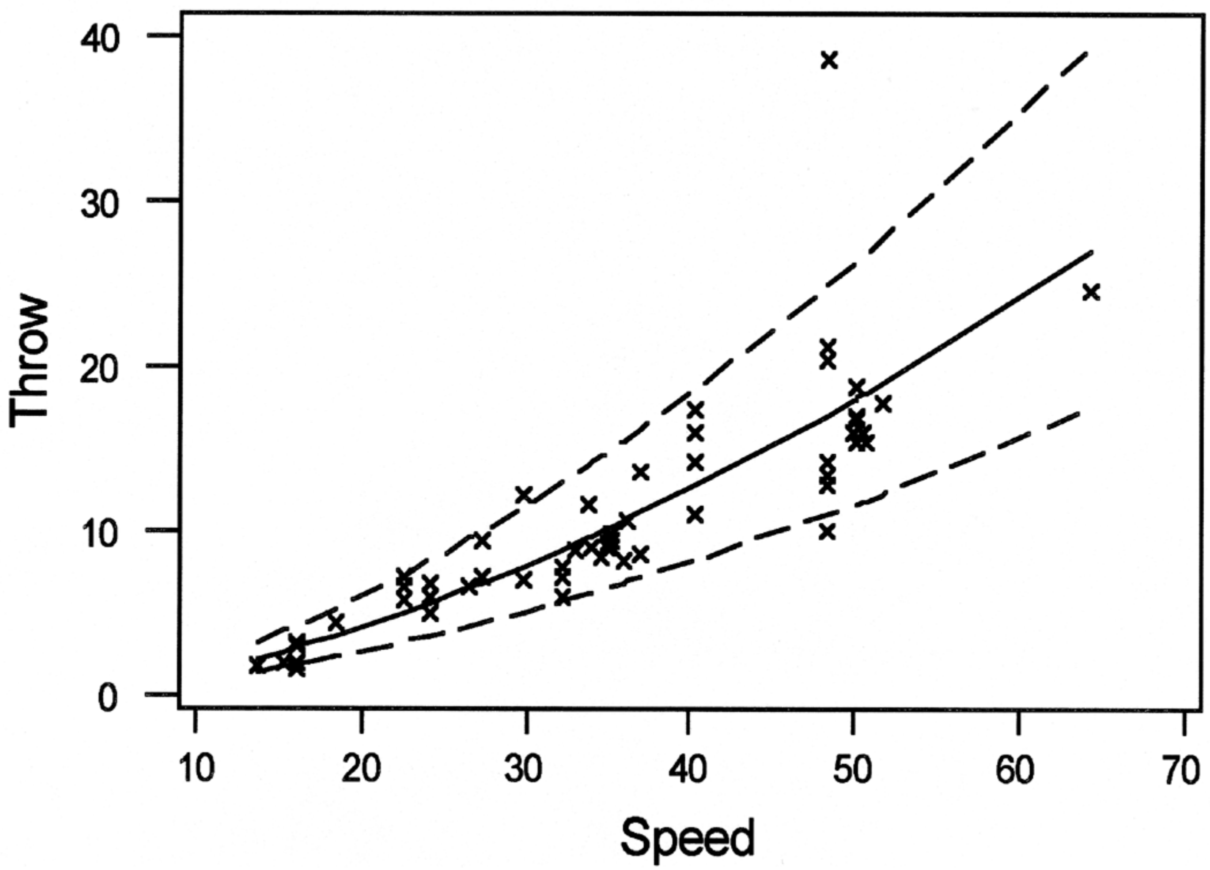
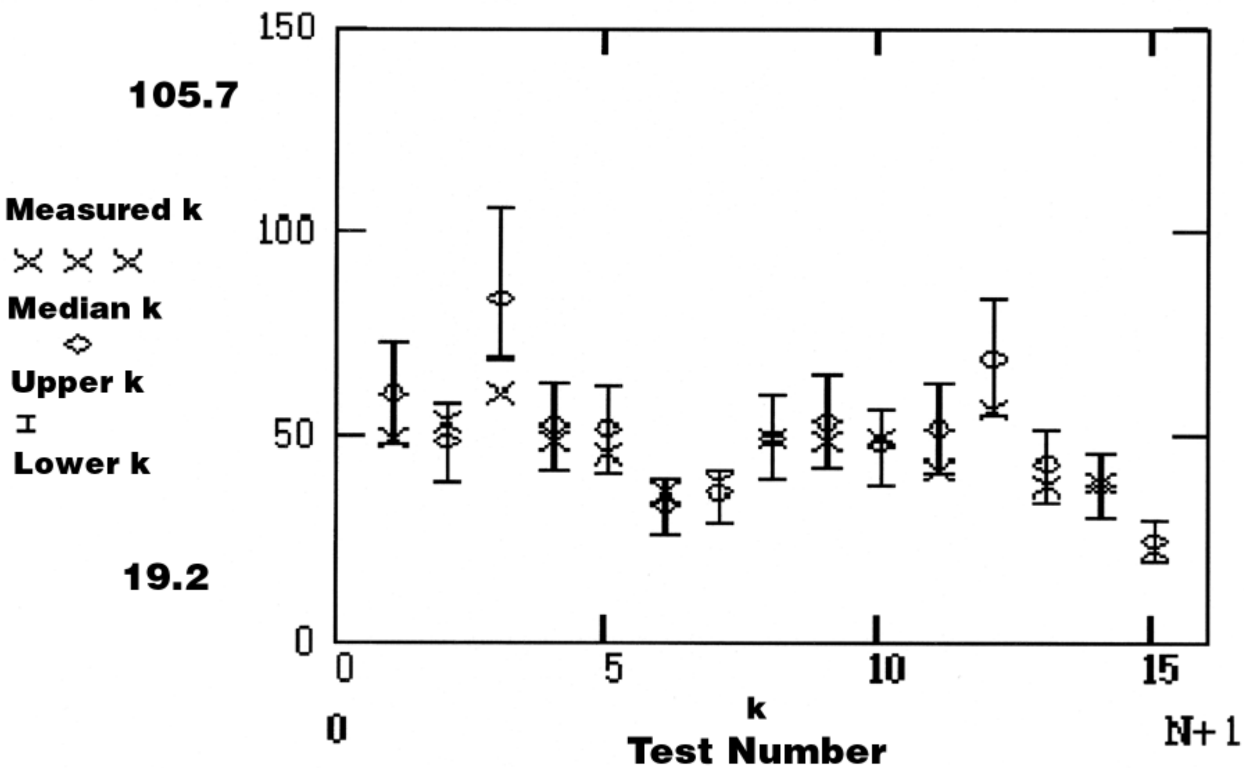


Figure 4.4

Measured Initial Speeds, Posterior Median Speeds and 95% Credible Intervals for 15 Pedestrian Dummy Crash Tests

(Speeds are in Kilometers/Hour)



CHAPTER FIVE

SUMMARY AND CONCLUSION

Neighborhood traffic control or "Traffic Calming" refers to engineering interventions expressly designed to lower vehicle speeds or volumes on residential streets. Consideration of a street for NTC is usually initiated by residents and the decision of whether or not to implement NTC measures often has a strong political component with both vigorous local support and opposition being a not uncommon occurrence. Traffic engineers can find themselves caught between opposing sides and their position is made more difficult by the fact that the profession lacks clear out warrants identifying those situations where NTC might have an overall beneficial effect. This is at least in part due to the difficulties inherent in quantifying the relationships between traffic conditions and safety on residential streets which arise primarily because the sparsity of crashes on residential streets, coupled with the problems inherent in estimating the exposure of non-vehicle road users making it difficult to estimate crash rates.

The primary objective of this research project has been to show how readily measured traffic characteristics, in particular the distributions of vehicle speeds and headways, can be coupled with a structure model of a standardized pedestrian/vehicle conflict in order to simulate how these traffic characteristics influence the outcome of the conflicts. In essence, one uses the structural model to define those combinations of initial conditions that lead to a crash and then computes the probability attached to this set. This is not a new idea but the approach has had limited practical application because of the computational difficulties arising when one attempts to compute the probabilities. We have hopefully shown, however, that these problems can be overcome by employing Monte Carlo simulation methods. In particular, we have shown how to compute the probability that a standardized vehicle/pedestrian conflict results in a collision and the probability that such a collision results in a severe injury to the pedestrian. These computational procedures were applied to data collected on 25 residential streets in the Twin Cities showing a range of peak-hour directional traffic volumes (between 10 and 125 vehicles/hour) and a range of average peak-hour speeds (between 20 and 42 mph), but with the majority of sites having average speeds lower than the nominal 30 mph speed limit. At least as far as our sample of streets is concerned the probabilistic measures were sensitive to traffic conditions. The collision probability estimates tended to increase as traffic volume increased but were fairly insensitive to changes in the average speed. On the other hand, the proportion of collisions resulting in severe injuries tended to be insensitive to traffic volume but very sensitive to average speed. It is suggested that the marginal probability of a crash with a severe injury might be used as a composite measure combining the effects of traffic volume and traffic speed. We have also shown how these probabilities could change as the result of a hypothetical 25 mph speed limit. In particular, the so called "Probability of Necessity", the conditional probability that, other things equal, a crash which occurs without under existing conditions would have been prevented with the speed limit, turns out to be (roughly) equivalent to the more commonly used crash reduction factor. As might be expected, those streets with average speeds above 25 mph showed the greatest reduction effect and this reduction was relatively insensitive to traffic volume. However, another composite measure, the probability of a preventable crash appears to be sensitive to both traffic volume and traffic speed.

The methods described in Chapter 3 can be used to "put a number" on the degree to which traffic conditions on a residential street pose a risk to child pedestrians. The methods do not, however, provide guidance as to which level of risk is acceptable or when risk is so great that some form of intervention is called for. This moves us beyond technical engineering into the area of social decision-making. As is well recognized, a genuinely zero-risk environment is neither feasible nor particularly desirable but we have, at present, little guidance as to where a line should be drawn. In the standard model for traffic safety programming, a road's risk is assessed relative to what appears to be typical on similar facilities. That is, crash rates for a large number of roads are estimated using crash and traffic count data and those roads with atypically high estimated rates are designated as potential high-hazard sites. A similar approach could be applied to residential streets using estimation methods described in Chapter 3. Automatic traffic recorders would be used to collect speed and headway data for a large, representative sample of residential streets and then standardized crash probabilities would be computed as described above. These results would provide a baseline against which the traffic conditions on some problematic streets could be evaluated.

This brings us to issues of implementation. In carrying out these computations we used the Markov Chain Monte Carlo software BUGS. However, for the collision probability computations reported in Table 3.4 this is something of an over-kill. Standard Monte Carlo simulation is sufficient to compute these probabilities and stand-alone computer code would not be that difficult to develop. In fact, the computations could probably be carried out using a reasonable sophisticated spreadsheet program as long as it supported random number generation. One would simply generate values for v_1, t_1, f, t_p, x_2 and v_2 from the appropriate distributions and then test whether or not a collision occurs (and compute the impact speed) using Equation 3.1. The degree of injury suffered by the pedestrian could then be simulated by sampling from a multinomial random variable and then computing probabilities using Equation 2.6. For Probabilities of Necessity that appear in Table 3.5 and in Chapter 4, however, MCMC methods are required because the desired probabilities are conditional on the occurrence of a collision. Unlike standard Monte Carlo methods MCMC methods are more difficult to use and it is usually recommended that potential users have a good working knowledge of probability and statistics. Thus, computation of collision probabilities and injury severity distributions could be done routinely by most traffic engineers in order to rank sites with regard to pedestrian risk. Computation of potential casual effects however should probably be relegated to special studies, carried out either by qualified consultants or by personnel who have received special training.

In support of the above work, this project also developed a quantitative model expressing the probability of a slight, serious or fatal injury to a pedestrian as a function of the speed of an impacting vehicle. Interestingly this modeling work revealed that for child pedestrians, the impact speed at which slight or severe injures are equal is about 25 mph suggesting that 25 mph is a reasonable speed limit for residential streets. In addition, we have shown how a modified version of the structural model can be used to compute Probabilities of Necessity for individual, actually occurring, vehicle/pedestrian crashes. This approach requires information from the sort of detailed investigation usually done to support a reconstruction of the crash. When applies to data from eight crashes from Adelaide, Australia that occurred on streets with a 60 km/hr (37 mph) speed limit, the method revealed that speeding was probably a factor in at least three of the crashes and the expected crash reduction due to adherence to the 60 km/h (37 mph) speed limit was 3.8 out of eight crashes.

Finally, it is worth emphasizing that the probabilities appearing in Chapter 3 have a standard frequent interpretation as limiting relative frequencies from an infinite number of repetitions of a random experiment. The probabilities reported in Table 4.2 do not, but rather, reflect degrees of belief accorded to propositions such as “The initial speed of the vehicle in case

015 was between 48.5 and 78.1 km/h" or "If the vehicle in case 154 had been traveling at 60 km/h, the driver would have stopped before hitting the pedestrian." Because these probabilities are specific to individual accidents for which the available information is limited, objective claims based on convergence of belief are not available and the product of this analysis should be seen as an expert opinion. This accords with the view of accident reconstruction expressed by Baker and Frick (1990): "Opinions or conclusions are the products of accident reconstruction. To the extend that reports of direct observations are available and can be depended on as facts, reconstruction is unnecessary." (p. 50–4). Thus, the methods presented in Chapter 4 have less in common with traditional statistical analysis than they do with recent applications of Bayesian methods to forensic problems (e.g. Aitken at al. 1997, Dawid and Evett 1997, Kadane and Terrin 1997) in which probability calculations are used to bind what can reasonably be concluded from imperfect evidence.

REFERENCES

AASHTO (1994) *A Policy on Geometric Design of Highways and Streets,* American Association of State Highway and Transportation Officials, New York.

ASHTO (1996) *Roadside Design Guide,* American Association of State Highway and Transportation Official, New York.

Aitkin, C., Bring, J., Leonard, T., and Papasouliostis, O. (1997) Estimation of Quantities of Drugs Handled and the Burden of Proof, *Journal of the Royal Statistical Society A,* 160, 333-350.

Anderson, R., McLean, A., Farmer, M., Lee, B., and Brooks, C., (1997) Vehicle Travel Speeds and the Incidence of Fatal Pedestrian Crashes. *Accident Analysis and Prevention,* 29, 667-674.

Armstrong, B. (1990) The Effects of Measurement Errors on Relative Risk Regressions, *American Journal of Epidemiology,* 132, 1176-1184.

Ashton, S. (1980) A Preliminary Assessment of the Potential for Pedestrian Injury Reduction through Vehicle Design, *Proceedings of 24th Stapp Car Crash Conference,* SAE, Inc., Warrendale, PA, 609-635.

Ashton, S., Peddar, J., and Mackay, G. (1977) Pedestrian Injuries and the Car Exterior, *SAE Technical Paper 770092*, SAE, Inc., Warrendale, PA.

Baker, J., (1975) *Traffic Accident Investigation Manual*, Traffic Institute, Northwestern University, Evanston, IL.

Balke, A., and Pearl, J., (1994) Probabilistic Evaluation of Counterfactual Queries, *Proceedings of 12th National Conference on Artificial Intelligence,* AAAI Press, Menlo Park, NY.

Baltes, M. (1998) Descriptive Analysis of Crashes Involving Pedestrians in Florida, 1990-1994, *Transportation Research Record,* 1636, 138-145.

Berge, J. (1985) *Statistical Decision Theory and Bayesian Analysis,* Springer-Verlarg, New York.

Best, N., Cowles, M. and Vines, K. (1995) CODA: *Convergence Diagnosis and Output Analysis Software for Gibbs Sampling*, MRC Biostatistics Unit, Cambridge, UK.

Brach, R. (1994) Uncertainty in Accident Reconstruction Calculations, *Accident Reconstruction: Technology and Animation IV,* SAE Inc., Warrendale, PA, 147-153.

Britt, J., Bergman, A., and Moffat, J., (1995) Law Enforcement, Pedestrian Safety and Driver Compliance with Crosswalk Laws: Evaluation of a Four-Year Campaign in Seattle, *Transportation Research Record*, 1485, 160-167.

Burson, P. (1999) Pedestrians Take Safety Message to the Street, *St. Paul Pioneer Press*, September 14, 1999, p. 2B.

Carroll, D., Ruppert, D., and Stefanski, L. (1995) *Measurement Error in Nonlinear Models*, Chapman and Hall, London, UK.

Davis, D., and Wessels, R. (1999) Automated Pedestrian Collision Typology in Washington State, 1990 to 1995, paper 991236 presented at the *78th Annual Meeting of the Transportation Research Board,* January 10-14, 1999, Washington, DC.

Davis, G. (1998) Method for Estimating the Effect of Traffic Volume and Speed on Pedestrian Safety for Residential Streets, *Transportation Research Record*, 1636, 110-115.

Davis, G. (2000a) Estimating Traffic Accident Rates while Accounting for Traffic-Volume Measurement Error: A Gibbs Sampling Approach, *Transportation Research Record*, 1717, 94-101.

Davis, G. (2000b) Accident Reduction Factors and Causal Inference in Traffic Safety Studies: *A Review, Accident Analysis and Prevention*, 32, 95-109.

Davis, G., and Corkle, P. (1997) Probabilistic Rating of Safety on Local Streets, in R. Benekohal (Ed.) *Traffic Congestion and Traffic Safety in the 21st Century*, American Society of Civil Engineers, New York, 465-471.

Davis, G., and Yang, S. (2001) Bayesian Identification of High-Risk Intersections for Older Drivers via Gibbs Sampling, *Transportation Research Record*, 1746, 84-89.

Dawid, P. (2000) Causal Inference without Counterfactuals, *Journal of American Statistical Association*, 95, 407-423.

Dawid, A., and Evett, I. (1997) Using a Graphical Method to Assist the Evaluation of Complicated Patterns of Evidence, *Journal of Forensic Science*, 42, 226-231.

Draper, D., Hodges, J., Mallows, C., and Pregibon, D. (1993) Exchangeability and Data Analysis, *Journal of the Royal Statistical Society A*, 156, 9-37.

Eubanks, J., and Hill, P. (1998) *Pedestrian Accident Reconstruction and Litigation*, Lawyers and Judges Publishing Co., Tuscon, AZ.

Fambro, D., Fitzpatrick, K., and Koppa, R. (1997) *Determination of Stopping Sight Distances, NCHPR Report 400*, Transportation Research Board, Washington, DC.

Freedman, D. (1997) Some Issues in the Foundations of Statistics, (with discussion) in *Topics in the Foundations of Statistics*, ed. B. Van Fraasen, Kluwer Academic Publishers, Norwell, MA, 19-83.

Fricke, L. (1990) *Traffic Accident Reconstruction*, Traffic Institute, Northwestern University, Evanston, IL.

Garrott, W., and Guenther, D. (1982) Determination of Precrash Parameters from Skid Mark Analysis, *Transportation Research Record*, 893, 38-46.

Gelman, A., and Rubin, D. (1992) Inference from Iterative Simulation Using Multiple Sequences, *Statistical Science*, 7, 457-472.

Giese, J. (1996) *The Relationship between Residential Street Design and Pedestrian Safety*, *Master's Thesis*, Dept. of Landscape Architecture, University of Minnesota, Minneapolis, MN.

Gilks, W., Thomas, A., and Spiegelhalter, D. (1994) A Language and Program for Complex Bayesian Modeling*, The Statistician*, 43, 169-197.

Hauer, E. (1980) Bias-by-Selection: Overestimation of the Effectiveness of Safety Countermeasures Caused by the Process of Selection for Treatment, *Accident Analysis and Prevention*, 12, 113-117.

Hauer, E. (1988) The Safety of Older Persons at Intersections, in *Transportation in an Aging Society, Volume 2*, Special Report 218, Transportation Research Board, Washington, DC, 194-252.

Hauer, E. (1997) *Observational Before-After Studies in Road Safety*, Elsevier Science, Oxford, UK.

Hicks, J. (1989) Traffic Accident Reconstruction, in K. Carper (Ed.), *Forensic Engineering*, Elsevier, New York, 101-129.

Holland, P. (1986) Statistics and Causal Inference, *Journal of American Statistical Association*, 81, 945-960.

Homburger, W., Deakin, E., Bosselmann, P., Smith, D., and Burkers, B. (1989) *Residential Street Design and Traffic Control*, Institute of Transportation Engineers, Washington, DC.

Hoque, M., and Andreassen, D. (1986) Pedestrian Accidents: an Examination by Road Class with Special Reference to Accident 'Cluster'," *Traffic Engineering and Control*, 27, No. 7/8, 391-397.

Howarth, C., Routledge, D., and Repetto-Wright, R. (1974) An Analysis of Road Accidents Involving Child Pedestrians, *Ergonomics*, 17, 319-330.

Hsieh, D., Manski, C., and McFadden, D. (1985) "Estimation of Response Probabilities from Augmented Retrospective Observations," *Journal of American Statistical Association*, 80, 651-662.

ITE (1993) *Speed Zone Guidelines: A Proposed Recommended Practice*, Institute of Transportation Engineers, Washington, DC.

Kadane, J., and Terrin, N. (1997) Missing Data in the Forensic Context, *Journal of the Royal Statistical Society A,* 160, 351-357.

Katz, A., Zaidel, D., and Elgrishi, A. (1975) An Experimental Study of Driver and Pedestrian Interaction During the Crossing Conflict, *Human Factors*, 17, 514-527.

Kloeden, C., McLean, A., Moore, V., and Ponte, G. (1997) *Traveling Speed and the Risk of Crash Involvement*, NHMRC Road Accident Research Unit, University of Adelaide, Adelaide, Australia.

Knoblauch, R. (1977) Causative Factors and Countermeasures for Rural and Suburban Pedestrian Accidents: Accident Data Collection and Analysis, *National Highway Traffic Safety Administration Report DOT-HS-355-3-728*, Dept. of Transportation, Washington, DC.

Knoblauch, R., Tobey, H., and Shunman, E. (1984) Pedestrian Characteristics and Exposure Measures, *Transportation Research Record*, 959, 35-41.

Koppa, R., Fambro, D., and Zimmer, R., (1996) Measuring Driver Performance in Braking Maneuvers, *Transportation Research Record*, 1550, 8-15.

Kost, G., and Werner, S. (1994) Use of Monte Carlo Simulation Techniques in Accident Reconstruction, *SAE Technical Paper 940719*, SAE, Inc. Warrendale, PA.

Lawrason, G., Swiercinsky, R., and Mason, R. (1980) *Pedestrian Injury Causation Study, National Highway Traffic Safety Administration Report DOT-HS-7-01580*, Dept. of Transportation, Washington, DC.

Limpert, R. (1989) *Motor Vehicle Accident Reconstruction and Cause Analysis, Third Edition*, Michie, Charlottesville, VA.

Lindley, D. (1987) The Probability Approach to the Treatment of Uncertainty in Artificial Intelligence and Expert Systems, *Statistical Science*, 2, 17-24.

MacLaughlin, T., Wiechel, J., and Guenther, D. (1993) Head Impact Reconstruction-HIC Validation and Pedestrian Injury Risk, *SAE Technical Paper 930895*, SAE Inc. Warrendale, PA.

Mak, K., Sicking, D., and Zimmerman, K. (1998) Roadside Safety Analysis Program: A Cost-Effectiveness Procedure, *Transportation Research Record*, 1647, 67-74.

Manski, C., and Lerman, S. (1977) The Estimation of Choice Probabilities from Choice Based Sample, *Econometrica*, 45, 1977-1988.

Mathsoft (1995) *Mathcad User's Guide*, Mathsoft, Inc. Cambridge, MA.

Mayne, A. (1965) The Problem of the Careless Pedestrian, in J. Almond (Ed.), *Proceedings of the Second International Symposium on the Theory of Road Traffic Flow*, OECD, Paris, 279-285.

Minitab (1995) *MINITAB User's Guide*, Minitab, Inc., State College, PA.

Mn/DOT (1996a) *Road Design Manual*, Minnesota Dept. of Transportation, St. Paul, MN.

Mn/DOT (1996b) *Traffic Engineering Manual*, Minnesota Dept. of Transportation, St. Paul, MN.

Mn/DPS (2000) *1999 Crash Facts*, Minnesota Dept. of Public Safety, St. Paul, MN.

Neptune, J., Flynn, J., Chavez, P., and Underwood, H. (1995) Speed from Skids: A Modern Approach, in *Accident Reconstruction: Technology and Animation V*, SAE, Inc. Warrendale, PA, 189-204.

Niederer, P. (1991) The Accuracy and Reliability of Accident Reconstruction, in *Automotive Engineering and Litigation*, vol. 4, eds. G. Peters and G. Peters, Wiley and Sons, New York, 257-303.

NMA (2001) "NMA's Position on Speed Limits," and "NMA's Position on Traffic Calming," *National Motorists Association*, http:\\www.motorists.com.

Olson, P. (1996) *Forensic Aspects of Driver Perception and Response*, Lawyers and Judges Publishing Co., Tuscon, AZ.

Parker, M (1985) *Synthesis of Speed Zoning Practices, Report FHWA/RD-85/096*, Federal Highway Administration, Washington, DC.

Pasanen, E. (1992) *Driving Speeds and Pedestrian Safety: A Mathematical Model*, National Technical Information Service Document PB93-141018, Dept. of Commerce, Washington, DC.

Pasanen, E., and Salmivaara, H. (1993) Driving Speed and Pedestrian Safety in the City of Helsinki, *Traffic Engineering and Control*, June.

Pearl, J. (2000) *Causality: Models, Reasoning, and Inference*, Cambridge University Press, Cambridge, UK.

Prentice, R., and Pyke, R. (1979) Logistic Disease Incidence Models and Case-Control Studies, *Biometrika*, 66, 403-411.

Raftery, A., and Lewis, S. (1992) How Many Iterations of the Gibbs Sampler J. Bernardo, J. Berger, A. Dawid, and A. Smith 9Eds.), *Bayesian Statistics 4*, Oxford University Press, Oxford, UK, 763-774.

Randles, B., Fugger, T., Eubanks, J., and Pasanen, E. (2001) Investigation and Analysis of Real-Life Pedestrian Collisions, *SAE Technical Paper 01-0171*, SAE, Inc., Warrendale, PA.

Reed, W., and Keskin, A. (1989) Vehicular Deceleration and its Relationship to Friction, *SAE Technical Paper 890736*, SAE, Inc., Warrendale, PA.

Robertson, D. (1997) The Common Sense of Cause in Fact, *Texas Law Review*, 75, 1765-1800.

Rubin, D. (1974) Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies, *Journal of Education Psychology*, 66, 688-701.

Stewart-Morris, M. (1995) Real Time, Accurate Recall, and Other Myths, in T. Bohan and A. Damask (Eds.), *Forensic Accident Investigation: Motor Vehicles*, Michie Butterworth, Charlottesville, VA, 413-438.

Tharp, K., and Tsongas, N. (1977) Injury Severity Factors-Traffic Pedestrian Collisions, *SAE Technical Paper 770093*, SAE, Inc. Warrendale, PA.

Thompson, S., Fraser, E., and Howard, C. (1985) Driver Behavior in the Presence of Child and Adult Pedestrians, *Ergonomics*, 28, 1469-1474.

Treat, J., Tumbas, N., McDonald, S., Shinar, D., Hume, R., Mayer, R., Stansifer, R., and Castellan, N. (1979) *Tri-Level Study of the Causes of Traffic Accidents, Report HS-034-3-535*, National Highway Traffic Safety Administration, Washington, DC.

TRB (1998) *Managing Speed: Review of Current Practice for Setting and Enforcing Speed Limits, Special Report 254*, Transportation Research Board, Washington, DC.

Van Wijk, J., Wismans, J., Maltha, J., and Wittebrood, L. (1983) MADYMO Pedestrian Simulations, *SAE Technical Paper 830060*, SAE, Inc. Warrendale, PA.

Varhelyi, A. (1998) Driver Behavior at a Zebra Crossing: *A Case Study, Accident Analysis and Prevention*, 30, 731-743.

Vaughn, R. (1997) Avoiding the Emerging Pedestrian: A Mathematical Model, *SAE Technical Paper 970962*, SAE, Inc., Warrendale, PA.

Vilenius, A., Ryan, G., Kloeden, C., McLean, A., and Dolinis, J. (1994) A Method of Estimating Linear and Angular Accelerations in Head Impacts to Pedestrians, *Accident Analysis and Prevention*, 26, 563-570.

Wallwork, M. (1993) Traffic Calming, in *The Traffic Safety Toolbox*, Institute of Transportation Engineers, Washington, DC, 235-246.

Wood, D. (1988) Impact and Movement of Pedestrians in Frontal Collisions with Vehicles, *Proceedings of the Institution of Mechanical Engineers*, 202, D2, 101-110.

Wood, D. (1991) Application of a Pedestrian Impact Model to the Determination of Impact Speed, *SAE Technical Paper 910814*, SAE, Inc. Warrendale, PA.

Wood, D., and O'Riordain, S. (1994) Monte Carlo Simulation Methods Applied to Accident Reconstruction and Avoidance Analysis, in *Accident Reconstruction: Technology and Animation IV*, SAE, Inc. Warrendale, PA, 129-136.

Zegeer, C. (1993) Designing for Pedestrians, in *The Traffic Safety Toolbox*, Institute of Transportation Engineers, Washington, DC, 187-202.

APPENDIX

EXAMPLES OF CODE AND OUTPUT FOR BUGS MODELS

Example: WINBUGS Code File

Used to Compute Probability of Necessity for Site 27b in Chapter 3

```
model pedhit5
{
	a <- f*9.81
	x1pass <- (x2*v1)/v2
	x1stop <- v1*tp + pow(v1,2)/(2*a)
	hit.bar  <- step(x1-x1pass)*step(x1stop-x1)
	phit <- .99999*hit.bar
	hit ~ dbern(phit)
	start <- (sb+instreet)

	v1.tau <- 1/(pow(v1.sigma,2))
	v1 ~ dnorm(v1.bar,v1.tau)

	x1 <- v1*t1

	t1.bar ~ dunif(0,1)
	t1 <- t1.bar*gap

	logap.tau <- 1/(pow(logap.sigma,2))
	gap ~ dlnorm(logap.mu,logap.tau)

	x2 ~ dunif(instreet,start);
	v2tau <- 1/(pow(v2sig,2))
	v2 ~ dnorm(v2bar,v2tau)

	tp.sigma2 <- log((pow(tp.sd,2)/pow(tp.bar,2))+1)
	tp.mu <- log(tp.bar)-0.5*tp.sigma2
	tp.tau <- 1/tp.sigma2
	tp ~ dlnorm(tp.mu,tp.tau)

	f.sigma2 <- log((pow(f.sd,2)/pow(f.bar,2))+1)
	f.mu <- log(f.bar)-0.5*f.sigma2
	f.tau <- 1/f.sigma2
	f ~ dlnorm(f.mu,f.tau)

	v1.star <- min(v1,11.2)
	x1.star <- v1.star*t1
	x1pass.star <- (x2*v1.star)/v2
```

```
        x1stop.star <- v1.star*tp + pow(v1.star,2)/(2*a)
        hit.star <- (step(x1.star-x1pass.star)*step(x1stop.star-x1.star))
        nohit.star <- 1-hit.star
}

Data list(v1.bar=12.731
v1.sigma=1.808,
logap.mu=2.778,
logap.sigma=1.111,
sb=14.2,
instreet=1.5,
v2bar= 5.4,
v2sig=.45,
tp.bar= 1.07,
tp.sd= 0.248,
f.bar=0.63,
f.sd=0.08,
hit=1)

Inits  list(
gap=40.0,
t1.bar=.5,
x2= 7.5,
v2= 4.6 ,
f=.63,
tp=1.0,
v1=15.0 )
```

Example: Statistical Output from WINBUGS

The mean of Nohit.star estimates the Probability of Necessity at Site 27b

Node	Mean	Sd	2.5%	Median	97.5%	Start	Sample
Nohit.Star	0.140	0.3478	0.0	0.0	1.0	4002	20000

Example: Classic BUGS Code and Output
Used in Reconstruction of CN015 in Chapter 4.

```
Welcome to BUGS on 6 th Jun 2000  at 13:18:0
BUGS : Copyright (c) 1992 .. 1995 MRC Biostatistics Unit.
All rights reserved.
Version 0.600 for 32 Bit PC.
For general release: please see documentation for disclaimer.
The support of the Economic and Social Research Council (UK)-
is gratefully acknowledged.
Bugs>compile("cntrfac3.bug")
model cntrfac3;

const
   M=1;

var
a, x.brake,x.tran,
x.prt,xx,x.init,nohit,fullhit,skidmark, v2.imp, vi,vi.kph, v.kph, speeding,
a1, a2, b, a1.bar,a2.bar,b.bar,a1.sig2,a2.sig2,b.sig2,a1.tau,a2.tau,b.tau,
pt1, pt2, psev[3],injury, ahat, bhat,ahat.bar,bhat.bar,ahat.sig2,bhat.sig2,
ahat.tau,bhat.tau,throw.bar,throw.sig2, throw.tau,throw,
skid1.bar, logskid1.bar, lvi, logskid2.bar,skid.tau,skid1,skid2,
tp.lower,tp.upper,tp, ts, ts.lower, ts.upper,
x.lower,x.upper, x, v.lower,v.upper,v, f.lower,f.upper,f,
v.star, xbrake.star, xprt.star, xstop.star, nohit.star;

data  in "cntrfac.dat";

inits in "cntrfac.in";
{

# counterfactual world
        xbrake.star <- pow(v.star,2)/(2*a);
        xprt.star <- v.star*tp;
        xstop.star <- xbrake.star + xprt.star;
        nohit.star <- step(x.init-xstop.star);

# data models
        logit(pt1)<- b*vi.kph-a1;
        logit(pt2) <- b*vi.kph-a2;
        psev[1]<- (1-pt1);
        psev[2]<- (pt1-pt2);
        psev[3]<- pt2;
        injury ~ dcat(psev[]);
```

```
	throw.bar <- ahat+bhat*log(vi);
	throw.tau <- 1/(throw.sig2 + .01);
	throw ~ dnorm(throw.bar,throw.tau);

	skid1.bar <- max(0.1,x.brake-x.tran);
	lvi <- max(0.1,log(vi));
	logskid1.bar <- log(skid1.bar);
	logskid2.bar <- 2*lvi-log(2*a);
#				skid2.bar <- pow(vi,2)/(2*a);
	skid1 ~ dnorm(logskid1.bar,skid.tau);
	skid2 ~ dnorm(logskid2.bar,skid.tau);

# braking model
	a <- f*(9.807);
	x.brake <- pow(v,2)/(2*a);
	x.tran <- v*ts-(a*pow(ts,2))/2;
	x.prt <- v*tp;
	xx <- max(x,-x.prt);
	nohit <- step(x-x.brake);
	fullhit <- step(-x);
	skidmark <- step(x.brake-x.tran);
	v2.imp <- (1-nohit)*((fullhit*pow(v,2)) +
		(1-fullhit)*(pow(v,2)-2*a*(x)));
	x.init <- xx + x.prt;
	vi				<- sqrt(v2.imp);
	vi.kph <- (3.6)*vi;
	v.kph <- (3.6)*v;
	speeding <- step(v-v.star);

# prior distributions
	tp ~ dunif(tp.lower,tp.upper);
	ts ~ dunif(ts.lower,ts.upper);
	x ~ dunif(x.lower, x.upper);
	v ~ dunif(v.lower, v.upper);
	f ~ dunif(f.lower, f.upper);
	a1.tau <- (1/a1.sig2);
	a2.tau <-(1/a2.sig2);
	b.tau <- (1/b.sig2);
	ahat.tau <- (1/ahat.sig2);
	bhat.tau <- (1/bhat.sig2);
	a1~dnorm(a1.bar,a1.tau)I(,a2);
	a2~dnorm(a2.bar,a2.tau)I(a1,);
	b~dnorm(b.bar,b.tau);
	ahat~dnorm(ahat.bar,ahat.tau);
	bhat~dnorm(bhat.bar,bhat.tau);
```

```
}

Parsing model declarations.
Loading data value file(s).
Loading initial value file(s).
Parsing model specification.
Checking model graph for directed cycles.
Possible directed cycle or undirected link in model
Generating code.
Generating sampling distributions.
Checking model specification.
Choosing update methods.
Metropolis method choosen for node ts
Metropolis method choosen for node x
Metropolis method choosen for node v
Metropolis method choosen for node f
compilation took  00:00:00
Bugs>
Bugs>inits()
list(tp=.5,
x=0,
v=10,
f=.45,
ts=.1,
a1=2.6,
a2=5.2,
b=.05,
ahat=-1.8,
bhat=1.4,
seed=123456789)

Bugs>
Bugs>data()
list(skid1=2.71,
skid2=1.61,
injury=3,
throw=2.20,
v.lower=10,
v.upper=25;
f.lower=0.45,
f.upper=1.0,
x.lower=0,
x.upper=30,
ts.lower=0.1,
ts.upper=0.5,
```

```
tp.lower=0.5,
tp.upper=2.5,
a1.bar=4.07214,
a1.sig2=0.5266,
a2.bar=7.20865,
a2.sig2=1.02124,
b.bar=0.0948,
b.sig2=.000519,
ahat.bar=-1.3662,
ahat.sig2=.0399,
bhat.bar=1.6078,
bhat.sig2=.00781,
throw.sig2 = 0.0616,
skid.tau= 100,
v.star=16.667)


Bugs>
Bugs>update(5000)      time for    5000   updates was  00:00:18
Bugs>
Bugs>monitor(tp,10)
Bugs>
Bugs>monitor(ts,10)
Bugs>
Bugs>monitor(f,10)
Bugs>
Bugs>monitor(v,10)
Bugs>
Bugs>monitor(x,10)
Bugs>
Bugs>monitor(xx,10)
Bugs>
Bugs>monitor(v.kph,10)
Bugs>
Bugs>monitor(vi.kph,10)
Bugs>
Bugs>monitor(x.init,10)
Bugs>
Bugs>monitor(nohit.star,10)
Bugs>
Bugs>monitor(speeding,10)
Bugs>
Bugs>update(50000)      time for    50000  updates was  00:03:00
Bugs>
Bugs>diag(tp)
            mean      sd      mean      sd      Z      sample
```

```
          1.46      7.23E-2   1.49      1.60E-3  -4.83E-1     5000
Bugs>
Bugs>diag(ts)
            mean       sd       mean       sd        Z       sample
          3.02E-1   7.56E-3   2.96E-1   2.82E-4   2.89E-1     5000
Bugs>
Bugs>diag(f)
            mean       sd       mean       sd        Z       sample
          7.97E-1   1.97E-2   8.10E-1   1.08E-3  -4.53E-1     5000
Bugs>
Bugs>diag(v)
            mean       sd       mean       sd        Z       sample
          1.79E+1   6.81      1.78E+1   3.06E-1   2.24E-1     5000
Bugs>
Bugs>diag(x)
            mean       sd       mean       sd        Z       sample
          1.55E+1   8.59      1.50E+1   2.93E-1   9.62E-1     5000
Bugs>
Bugs>diag(xx)
            mean       sd       mean       sd        Z       sample
          1.55E+1   8.59      1.50E+1   2.93E-1   9.62E-1     5000
Bugs>
Bugs>stats(tp)
            mean        sd       2.5% :  97.5%  CI    median      sample
          1.494E+0   5.759E-1   5.438E-1   2.442E+0   1.500E+0     5000
Bugs>
Bugs>stats(ts)
            mean        sd       2.5% :  97.5%  CI    median      sample
          2.969E-1   1.154E-1   1.097E-1   4.890E-1   2.931E-1     5000
Bugs>
Bugs>stats(f)
            mean        sd       2.5% :  97.5%  CI    median      sample
          8.095E-1   1.338E-1   5.138E-1   9.922E-1   8.332E-1     5000
Bugs>
Bugs>stats(v)
            mean        sd       2.5% :  97.5%  CI    median      sample
          1.779E+1   2.050E+0   1.358E+1   2.153E+1   1.786E+1     5000
Bugs>
Bugs>stats(x)
            mean        sd       2.5% :  97.5%  CI    median      sample
          1.500E+1   2.767E+0   1.009E+1   2.058E+1   1.480E+1     5000
Bugs>
Bugs>stats(xx)
            mean        sd       2.5% :  97.5%  CI    median      sample
          1.500E+1   2.767E+0   1.009E+1   2.058E+1   1.480E+1     5000
Bugs>
```

```
Bugs>stats(v.kph)
        mean       sd      2.5% :  97.5%  CI    median     sample
     6.407E+1   7.381E+0   4.892E+1   7.751E+1   6.429E+1     5000
Bugs>
Bugs>stats(vi.kph)
        mean       sd      2.5% :  97.5%  CI    median     sample
     3.240E+1   3.116E+0   2.579E+1   3.782E+1   3.264E+1     5000
Bugs>
Bugs>stats(x.init)
        mean       sd      2.5% :  97.5%  CI    median     sample
     4.160E+1   1.159E+1   2.219E+1   6.415E+1   4.096E+1     5000
Bugs>
Bugs>stats(nohit.star)
        mean       sd      2.5% :  97.5%  CI    median     sample
     4.500E-1   4.974E-1   0.000E+0   1.000E+0   0.000E+0    5000
Bugs>
    Bugs>stats(speeding)
        mean       sd      2.5% :  97.5%  CI    median     sample
     7.188E-1   4.495E-1   0.000E+0   1.000E+0   1.000E+0    5000
Bugs>
Bugs>q()
```