Technical Report Documentation Page

1. Report No.
FHWA-RD-90-006

2. Government Accession No.
PB90208489

3. Recipient's Catalog No.

4. Title and Subtitle
Empirical Bayes Approach to the Estimation of "Unsafety:" the Multivariate Method.

5. Report Date
March 1990

6. Performing Organization Code

7. Author(s)
Ezra Hauer

8. Performing Organization Report No.

9. Performing Organization Name and Address
Safety Studies Group, Department of Civil Engineering, University of Toronto, Toronto, M5S 1A4, Canada

10. Work Unit No. (TRAIS)
3A3b0032

11. Contract or Grant No.
DTFH61-86-C-00075

12. Sponsoring Agency Name and Address
-Office of Safety and Traffic Operations R&D, Federal Highway Administration, 6300 Georgetown Pike, McLean, Virginia 22101-2296
-Transport Canada, Ottawa, K1A 0N5

13. Type of Report and Period Covered
Final Report 1986-1990

14. Sponsoring Agency Code

15. Supplementary Notes
FHWA Contract Manager (COTR) John Viner (HSR-30)

16. Abstract
There are two kinds of clues to the unsafety of an entity: its traits (such as traffic, geometry, age or gender) and its historical accident record. The essence of the Empirical Bayes (EB) approach to the estimation of unsafety is that it uses both clues. How this is accomplished is described.

To estimate the unsafety of an entity using the EB approach, information is needed about the mean and the variance of the unsafety of similar entities which form its reference population. The Method of Sample Moments has been used for this purpose in the past. It suffers from three shortcomings. First, to yield usable estimates a very large reference population is required. Second, the choice of reference population is to some extent arbitrary. Third, entities in the chosen reference population usually cannot match the traits of the entity the unsafety of which is estimated.

To alleviate these shortcomings the Multivariate Method for estimating the mean and variance of unsafety in reference populations is offered. Its logical foundations are described and its soundness is demonstrated. The use of the Multivariate Method makes the EB approach to unsafety estimation applicable to a wider range of circumstances, it makes the decision about what entities to include in the reference population less arbitrary and it yields better estimates of unsafety. The applications of the EB and Multivariate Methods to tasks of identifying deviant entities and estimating the effect of interventions on unsafety are discussed and illustrated by numerical examples.

17. Key Words
Statistical methods, safety, estimation, Empirical Bayes, blackspots, treatment effectiveness.

18. Distribution Statement
No restrictions. This document is available to the public through the National Technical Information Service, Springfield, Virginia 22161.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	70	

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

GLOSSARY OF TERMS AND NOTATION

Compound Poisson distribution	A probability distribution in which the parameter (m) of the Poisson probability law is itself a random variable.
Crossbucks	Sign used as a warning device at rail-highway grade crossings.
EB	See Empirical Bayes
Empirical Bayes	The Empirical Bayes (EB) approach postulates that parameters of a probability distribution come from a certain probability distribution the parameters of which can be estimated on the basis of data.
Estimation of effect	The task of estimating what effect some intervention had on the unsafety of an entity.
Flashers	Flashing light and bells used as a warning device at rail-highway grade crossings.
Identification	The task of indentifying entities the unsafety of which may require attention.
Method of Sample Moments	A method to estimate E{m} and VAR{m} from the sample moments $\bar{x}$ and s^2.
Multivariate Method	A method to estimate E{m} and VAR{m} from the residuals of a multivariate statistical model.
Reference population	A set of entities the observed traits of which are similar to the traits of the entity about the unsafety of which we inquire.

Residual	The difference between the accident count on some entity and the estimate of E{m} calculated from the fitted model.
Poisson probability law	$P(X=x) = e^{-m}m^{x}/x!$
Unsafety of an entity	the number of accidents (perhaps classified by kind and severity) expected to occur on that entity per unit of time in a certain period.

^	The caret above a letter is used to indicate that this is an estimate of the quantity which the letter represents.
E{m}	When each of a set of entities has its own m, E{m} is their expected value.
$E_I\{m\}$	is the expected value of m in reference population I.
$E\{m\|x\}$	is the expected value of m in a population in which all entities has the same accident count x.
F_1, F_2	are traffic flows in vehicles/hour.
k	is a dimensionless parameter in the relationship $VAR\{m\}=[E\{m\}]^2/k$.
m	is the unsafety of an entity i.e., the number of accidents perhaps of given severity expected to occur on the entity per unit of time during a certain period.
$\hat{m}\|x$	is the estimate of m when x is known.
M	is the largest count of accidents.
n	number of entities in a reference population.
n_x	number of entities which recorded x accidents.

s^2 is the sample variance (see equation 4).

VAR{m} When each of a set of entities has its own m, VAR{m} is their variance.

$VAR_I\{m\}$ is the variance of m in reference population I.

VAR{m|x} is the variance of m in a population in which all entities have the same accident count x.

VAR{x} is the variance of the count of accidents in a reference population (around E{m}).

x is the count of accidents.

$\bar{x}$ is the sample mean (see equation 3).

α is defined in equation 1 as E{m}/[E{m}+VAR{m}].

$\hat{\sigma}$ is the sample standard deviation, the estimate of standard deviation σ.

$\hat{\sigma}/\sqrt{\nu}$ is the standard error, the estimate of the standard deviation of the sample mean with ν as the sample size.

1. INTRODUCTION

One should think of "unsafety" as a property of some entity (person, intersection) which becomes manifest in the occurrence of accidents. We are interested in the "unsafety" of an entity either because we want to know whether to intervene in order to make it safer or to assess how its "unsafety" has been affected by some intervention. In both cases we need to estimate what the "unsafety" of the entity was or is.

The promise of the Empirical Bayes (EB) method for the estimation of "unsafety" has been explored in earlier work.[1,2,3,4] It is founded on the premise that the entity about the "unsafety" of which we inquire, not only has a record of accidents but, in addition, belongs to a population of similar entities - a "reference population" and that what we know about the unsafety of entities which form the reference population also tells us something about the unsafety of the entity under scrutiny.

To establish a vocabulary, the "unsafety" of an entity will mean the number of accidents (perhaps classified by kind and severity) expected to occur on that entity per unit of time in a certain period. This quantity will be designated by "m" and the task is to estimate m accurately. Each entity which belongs to a certain reference population has its own m. Thus, the m's in the reference populations form a distribution with a mean E{m} and variance VAR{m}.

With this vocabulary in place one can be more specific about the essence of the EB method. To estimate the unsafety of a specific entity, the EB method uses not only the accident record of the entity but also the E{m} and VAR{m} of the reference population to which this entity is said to belong. Thus, to use the EB method, estimates of E{m} and VAR{m} are required.

Several linked difficulties arise; some are conceptual and some are practical in nature. The principal conceptual difficulty is that of deciding what is a suitable reference population for a certain entity. If this decision is merely a matter of judgement, different analysts may choose different reference populations, use different values for E{m} and VAR{m} and therefore arrive at discrepant estimates of unsafety for the same entity. The practical difficulty is that if one wishes to diminish the element of arbitrariness by choosing members for the reference population so that they are very similar to the entity the safety is to be estimated, the number of entities in the reference population will be too small to allow estimation of E{m} and VAR{m} by the methods used till now. These and similar questions were raised by Wright et al., in more detail by Elvik and most recently by Mountain and Fawaz.[5,6,7]

It seems that the sting can be taken out of both difficulties, conceptual and practical, by a suitable interpretation of the results of multivariate statistical analysis of data. This has been done in earlier work performed under this contract but without emphasis on the logical foundations of the approach and its empirical validity.[8,9] This report is an attempt to fill the gap.

2. ESTIMATING "UNSAFETY" BY THE EMPIRICAL BAYES METHOD

The unsafety of an entity has been defined to mean the number of accidents by kind and severity, expected to occur on the entity per unit of time during a specific period. In this definition the term "entity" can refer to a Mr. Smith, to the intersection of Main and High Streets, to all signalized intersection in Toronto, to automobiles of a certain vintage and make in Ontario etc. The word "expected" in this sentence refers to a long-term average which would materialize if it were possible to freeze a person's age and to perpetuate the traffic flow, driver demography, vehicle characteristics and other relevant conditions of the period. Because unsafety is defined to be an expected value, it can not be directly observed or measured; it has to be estimated.

Thus, e.g., we might be interested in an estimate of the unsafety in 1984 of Mr. Smith, a 22 year old Ontario driver who in that year had no accidents; or we might wish to estimate the unsafety of a specific rail-highway grade crossing between 1970 and 1980, which served an average of 800 vehicles/day, 2 trains/-day, had one track, was equipped with crossbucks and recorded in that period two accidents. Estimates of this kind are required both to assess the need for road safety management and to assess the results of attempts at managing road safety.

a. The Two Clues to Unsafety

Two kinds of clues contain information about the unsafety of an entity. Clues of the first kind are contained in traits such as gender, age, traffic volume or road geometry. Clues of the second kind are derived from the history of accident occurrence of that entity.

If some knowledge of the unsafety of entities with similar traits exists, one can make informative statements about the unsafety of a specific entity with the same traits. For example, if it is only known that Smith was licensed to drive in Ontario in 1984 but that person's age or gender is not known, what can be said is that in Ontario there were 0.0606 accidents per driver in that year. On this basis it can be asserted that 0.0606 is the estimate of the expected number of accidents for Smith in that year. The long-hand reasoning which supports this assertion is as follows:

> Consider the population of people licensed to drive in Ontario in 1984 to be the reference population for Smith because nothing is known to distinguish Smith from others in this population. Were one to make such assertions about any other member of this population, one should make the same claim as for Smith, since all members of this population have the same traits. Because it is desirable to make that claim which promises to be associated with the least error, it is best to use the mean. In this case, the mean is 0.0606 accidents/year. This is why it is proper to assert that 0.0606 is the estimate of unsafety for Smith.

If it is known, in addition, that Smith was 22, this added trait defines a new reference population. It is a subset of the population of all Ontario drivers and contains only those who in 1984 were 22 years old. For this new reference population the average number of accidents is 0.0840 and this now the new estimate for a 22 year old Smith. If it be also known that Smith is male, yet another reference population is defined. The traits which define this latest reference population are: being a driver in Ontario in 1984, being 22, and being male. For this population there was an average of 0.1152 accidents per driver per year. If nothing more is known about Mr. Smith, 0.1152 must be the estimate of his unsafety in 1984. Some may find it disconcerting that the estimate Smith's unsafety is a function of what is known about this driver. However, it is merely a manifestation of the property of all estimators that they depend on data.

In short, what is known about the unsafety of entities with similar traits is an important and legitimate clue (of the first kind) to the unsafety of a specific entity and therefore should be used for unsafety estimation. In this, the concept of reference population is seen to play a pivotal role. What is said about Smith depends on what is known about his traits, for it is these traits which define the reference population. To say something about the unsafety of Smith we have to know what was the unsafety of entities in the reference population. This, at once, links estimation to the reality of data and brings in its wake a string of difficulties.

It is desirable that scientific measurements be free of the arbitrary. Statistical estimation is a form of scientific measurement and therefore it is not pleasing to see its results to depend on judgement. However, contemplation of what is known about the traits of an entity and what is known about accidents in a corresponding reference population quickly leads to the realization that arbitrariness is not avoidable, either in practice or in principle, as is shown below.

Most would agree that to estimate Mr. Smith's unsafety, it is important to know whether he was single or married in 1984. Married men tend to have fewer accidents. However, marital status is not in the data base which Ontario maintains. Therefore, neither is this trait known for Mr. Smith nor is it possible to ascertain what is the mean number of accidents for a corresponding reference population. Thus, what has earlier been offered as a legitimate estimate of Smith's unsafety (0.1152 accidents per year), is in fact a reflection of what happens to be in the data base. What is in the data base, in turn, depends on what is measurable, what is easy to measure and on additional spurious events and considerations which had little to do with the estimation task at hand. Had different information been in the data base, one would have made different assertions about Mr. Smith's unsafety. Thus, <u>the first difficulty is that the estimate depends on what data happen to be available</u>.

It has been recognized above that marital status is a trait which is related to a person's unsafety. However, there must be other traits which influence unsafety and remain unrecognized. Thus, not only is the estimate of unsafety prescribed by what happens to be in the data base, it is also circumscribed by the extent of the acuity of the estimator's intuition and understanding. Accordingly, <u>the second difficulty is that the estimate of unsafety depends on what is thought to be relevant.</u>

Suppose now that Mr. Smith reveals that in 1984 he was married. Now he is no more a legitimate member of the reference population of male, 22 year old Ontario drivers whose marital status is unknown and who had, on the average 0.1152 accidents in 1984. Since the marital status of drivers in Ontario is not recorded, it is impossible to know the unsafety of the reference population of such married drivers. But it is now known that Smith was married. What should we estimate his unsafety to be? Thus, the third difficulty is that for any specific entity it is always possible to think of it as having some relevant trait which sets it apart from all available reference populations. Is then some available reference population still a legitimate source of information about the unsafety of this entity?

Thus, the use of "clues of the first kind" for unsafety estimation is tied to a dilemma. On one hand, it is both humanly reasonable and scientifically proper to adopt an estimation procedure which makes use of all relevant information. Conversely, an estimation procedure which makes no use of that information about the unsafety of an entity which is contained in the knowledge of its traits seems inefficient and contrary to common sense. On the other hand, what information about the unsafety of an entity can be extracted from the knowledge of its traits is entirely due to what is known about many other entities with the same or similar traits. Thus, the use of clues of the first kind is tied to the use of a reference population by links of logical necessity and the use of a reference population can not be divorced from the use of judgement. It is as if the desideratum of objectivity in estimation requires the sacrifice of common reason while the exercise of good reason by attempting to make the most of available information, leads one to relinquish estimation which is demonstrably clear of the judgemental.

The second kind of clue which bears on unsafety estimation is the historical accident record of the entity. It has been said earlier that Mr. Smith has had no accidents in 1984. Surely this tells something about his unsafety in 1984. Since only 1 in about 10 drivers of this kind have an accident in a certain year, it may tell little. Still, in principle, what information there is in the accident record of an entity, needs to be brought to bear on the estimation task.

If the count of accidents is considered to be the only basis for the statistical estimation of unsafety, a person with 0 accidents is estimated to have unsafety = 0. However, it makes no good sense to claim that because Mr. Smith has had no accidents in 1984, his unsafety in 1984 is estimated to be 0 accidents per year. Only people not driving have a zero chance of being in an accident as drivers. It is possible to decline making an estimate of unsafety in these circumstances, insisting that the count of accidents has to be substantial before a statistical estimate can be of any interest. Such a position has the unfortunate consequence that many entities about the unsafety of which we inquire (drivers, grade crossings, road sections) never have an accident count which can be deemed substantial and, as a consequence, one would have to concede that nothing can be said about their unsafety.

If the option of remaining silent about Smith's unsafety is unpalatable, it might be possible to make use of Mr. Smith's earlier accident record, hoping that he has had at least one accident in the past, and thus to avoid the embarrassment of having to declare that his unsafety in 1984 is estimated to be 0 accidents per year. Unfortunately, it is not clear how an accident in, say, 1982 is to affect the estimate of Smith's unsafety in 1984. Surely, Smith's unsafety changes with the passage of time; he matures, he gains experience, he was a student in 1982 and is gainfully employed in 1984. Consequently it is necessary to make some assumptions about <u>how</u> his unsafety changes with time, before the 1982 accident can be brought to bear on the task of estimating his unsafety in 1984. The assumption made most frequently and tacitly is that the unsafety of the entity does not change in time. While this assumption has the merit of being convenient, it lacks the virtue of being sensible.

Classical statistical estimation which relies only on clues of the second kind, the occurrence of accidents in our case, is usually thought to be entirely "objective" and free of the arbitrary. Indeed it would be so if one could claim that the unsafety of Smith in 1984 is the same as in any other year. While such a claim may be approximately true of a coin-tossing machine, it is an obvious untruth for any entity in the unsafety of which there is practical interest. However, if it is admitted that the unsafety of all entities is changing as conditions (such as age, exposure, traffic flow, driving culture, vehicle population, road characteristics) change, it is necessary to face the problem of using, say, Smith's 1982 accident record to estimate his unsafety in 1984. No matter how this is done, judgement and arbitrariness enter into classical estimation in the same manner as they have been seen earlier to infest estimation based on traits. This is shown below.

First, assume that Smith's exposure (amount of driving) in 1982 has something to do with his accident record in that year. However, data about Smith's exposure are not kept. It might be possible to find data about how much an average driver drove in 1982 and 1984 and also use data about the number of vehicles per kilometre of road in those years and apply these to the task of estimation. However, just as before, the problem is that the estimate depends on what data happens to be available.

Second, exposure has been recognized above as related to Smith's unsafety and its change from 1982 to 1984. However, there must be other conditions of Smith's life which are related to his unsafety and which have changed from 1982 to 1984. It is not that these have not been measured, it is that what they are is not known. Thus, for "conditions" just as for "traits", the second difficulty is that the estimate of unsafety depends on what is thought to be relevant.

Suppose now that it is possible to account for the change in a set of common "conditions." However, Smith tells that he graduated from school in 1983 and was a salesman in 1984. How this affects a person's unsafety is not known. Thus, just as before, the third difficulty is that for any specific entity it is always possible to think of some trait (here, change in conditions) which sets it apart from all available reference populations.

In summary, the unsafety of each entity changes in time. Therefore when estimation is based solely on the accident history of an entity and classical statistical methods of estimation are used, their use is seen to depend on the arbitrary and the judgmental in precisely the same manner as the method of estimation which is based on traits and reference populations. To estimate Smith's unsafety only on the basis of his accident record, as if we did not know that he has been driving in Ontario, was male and was 22 years old, is to disregard important information. To consider Smith's traits but not to take note of his actual accident record also amounts to a loss of information. It ought to be obvious that it is best to make use of both kinds of clues: those which derive from traits such as gender, age, traffic volume or road geometry and those which derive from the count of accidents.

At the same time, it seems proper to recognize that if the traits of entities are to affect estimation, or if the unsafety of an entity changes in time, the process of estimation will be unlike that of measuring distance, pressure, or rate of reaction; it will depend on what data is available, on what is recognized as relevant and on other manifestations of judgment. One should not pretend otherwise.

b. Mixing the Two Clues

The art of estimating unsafety using both kinds of clues is described in a handful of fairly recent articles.[1,10,11,12] A brief summary of the main results will be provided here.

Denote the unsafety of an entity by m. Consider a reference population to which this entity belongs. Denote by E{m} and VAR{m} the mean and the variance of the m's of entities in this reference population. The entity has recorded x accidents in the period of time for which we wish to estimate its unsafety. The task is to estimate its m.

It can be shown that if the count of accidents (x) obeys the Poisson probability law and when the distribution of the m's in the reference population can be described by a Gamma probability density function, a good estimator of the m for a specific entity is:[11]

$$\alpha E\{m\}+(1-\alpha)x, \quad \text{with } \alpha = E\{m\}/[E\{m\}+VAR\{m\}] \qquad (1)$$

Numerical example 1 serves to illustrate the use and meaning of this simple equation.

Numerical Example 1.

Consider a highway-rail grade crossing which is in an urban area, has a single track, is used by 2 trains per day, 550 cars per day and is signed by crossbucks. In the 5 years 1981 through 1985 it has recorded two accidents. What is the estimate $\hat{m}$ of its unsafety m during that period of time? (^ above a letter to signifies that this is an estimate.)

Later in section 3 it will be shown that in a reference population of such crossings, $\hat{E}\{m\}=0.0239$ accidents/year and $\hat{VAR}\{m\}=0.0011$ [acc./year]^2. For 5 years, $\hat{E}\{m\}= 5\times0.0239=0.1195$ accidents and $\hat{VAR}\{m\}= 5^2\times0.0011=0.0275$ [accidents]^2. Therefore, the estimate of α is 0.1195/(0.1195+0.0275)=0.81. By equation 1, $\hat{m}=0.81\times0.1195+0.19\times2=0.10+0.38=0.48$ accidents in 5 years.

The essence of the EB method is now in plain view. The two clues to m (0.1195 describing an average entity of the reference population and 2 being an "outcome" on the entity of interest) are combined and the weight attached to each clue depends only on the ratio VAR{m}/E{m}. Note that if in the reference population all m's are the same, VAR{m}=0 and therefore α=1 with the effect that the count of accidents does not influence the estimate. While this may seem curious at first, one soon recognizes that if all entities have the same m the estimate of E{m} is also the estimate of the unsafety for the entity under scrutiny, as the estimator indicates. Similarly, if the m's in the reference population are very diverse such that VAR{m}>>E{m}, α will be very small and what is known about the reference population exerts little influence on estimation. This again is as should be. Thus, the estimator in equation 1 which has been deduced by logic from specified premises, mixes the two kinds of clues about unsafety and does so in a manner which is in accord with common sense.

A more unified conceptual framework can be attained if instead of regarding the two kinds of clues as belonging to two separate and different domains, the accident count is taken to be just another trait. Thus, a reference population can be subdivided such that within each subpopulation all entities have the same accident count x. The subpopulation is defined by all the traits common to the entities of the original population and, in addition, by the "trait" of having recorded x accidents. Each entity belonging to this new reference population has unique m and their mean (E{m|x}) is estimated by equation 1.

The m's in the new reference population have also a variance (VAR{m|x}). It can be estimated by

$$\alpha(1-\alpha)E\{m\} + (1-\alpha)^2 x \quad . \tag{2}$$

Since ~~the~~ m of the entity of interest is one from a distribution in which the m's have a variance given by equation 2, this is also the variance of $\hat{m}|x$. This closes the loop. The starting point was a reference population which was defined by traits only. Entities in it were said to have m's with a mean E{m} and variance VAR{m}. Then a new trait has been added, that of having x accidents. In this new reference population the mean of the m's (denoted by $E\{m|x\}$) is estimated by equation 1 and this is used to estimate the m of the entity of interest; the variance of the m's in this new reference population (denoted by $VAR\{m|x\}$) is estimated by equation 2 and this is used to estimate the variance of the estimate of that m .

Numerical Example 2.

> Consider again the highway-rail grade crossing from Numerical example 1. For the 5 years 1981 through 1985, there were x=2 accidents. Calculations showed that $\hat{E}\{m\}=0.1195$ accidents and the estimate of α is 0.81.
>
> By equation 2, $\hat{VAR}\{\hat{m}|2\}=0.81\times0.19\times0.1195+0.19^2\times2=0.018+0.072=0.09$ [accidents]^2. Thus, from numerical example 1 $\hat{m}=0.48$ accidents in 5 years and the standard deviation of this is $\sqrt{0.09}=0.30$ accidents in 5 years.

The benefit of using both clues to unsafety (instead of relying on the accident count only) is now obvious. Had only the accident count been used, the estimate of m would be two accidents in 5 years and the standard deviation of that estimate is $\sqrt{2}=1.4$. Using the EB method (and the added information about the reference population) the estimate of m is 0.48 accidents in 5 years and the standard deviation of this estimate is 0.30.

The superior efficiency of the EB estimation can be shown by a graph. In figure 1, E{m} is set to be unity. At $\alpha=0$ ignorance about the distribution of m's in the reference population is so complete that it is disregarded. At this point, the variance of the estimate of m (shown on the vertical axis) is estimated by the count of accidents, x, as is customary. (This leads to conflict with common sense when x=0). At the other end, where $\alpha=1$, the variance of m's in the reference population is 0. It follows that as soon as an entity is known to be one of the reference population with mean E{m}, its m is known and there is no uncertainty about it. Between these extremes, the variance of the estimate assumes intermediate values. Except for the anomalous case when x=0 and $\alpha<0.5$, the variance of the estimate obtained by the joint use of both clues is always below that when estimation is by accident counts only.

To summarize, both clues to the unsafety of an entity should be used in estimation. One of the ways by which this can be done is given by equations 1 and 2. In these equations, the clues of the first kind are represented by the first two central moments of the m's in the reference population - E{m} and VAR{m}. Thus, estimates of E{m} and VAR{m} which pertain to the m's of the reference population are needed. Two alternative methods for the estimation of E{m} and VAR{m} are discussed next.

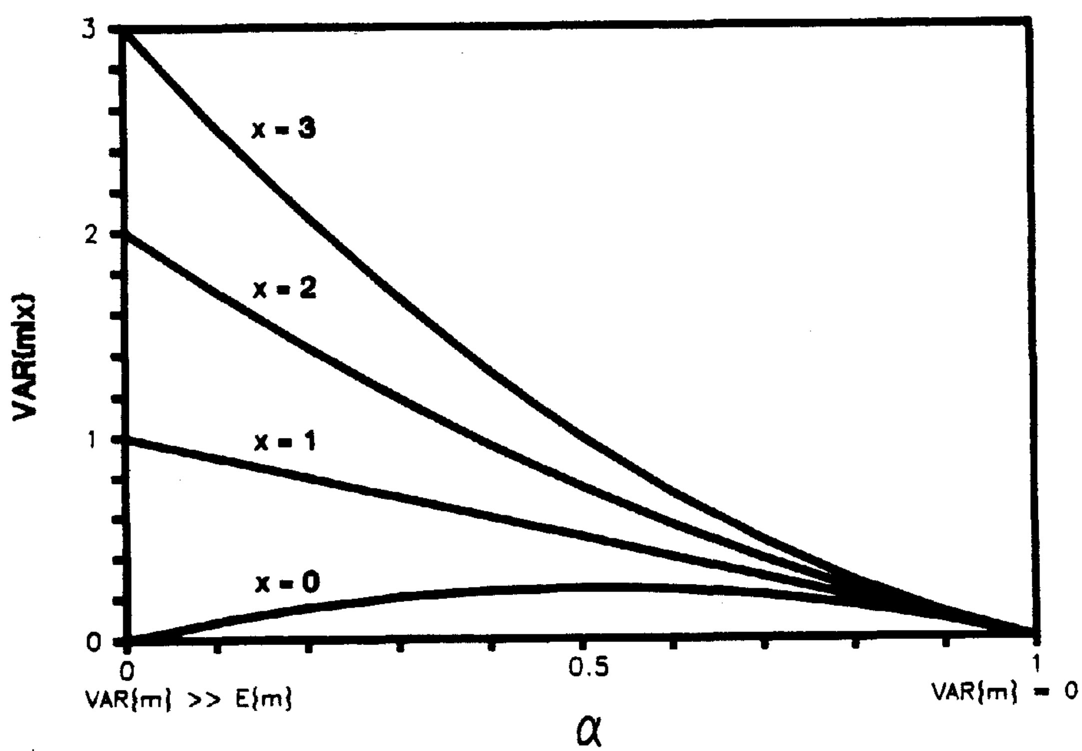

Figure 1. The efficiency of EB estimation.

~~3.~~ TWO METHODS FOR ESTIMATING E{m} AND VAR{m}.

Entities which share a specific set of chosen traits were said to form a reference population. Each entity of the reference population has its own m. The set of m's in the reference population has a mean E{m} and a variance VAR{m}. Estimates of these two moments of m in the reference population are needed to estimate the unsafety of any entity of this reference population (see equations 1 and 2). Two alternative methods to estimate E{m} and VAR{m} are discussed below.

a. The Method of Sample Moments.

Consider a reference population of n entities of which n_x have recorded x = 0,1,2,...,M accidents during a specified period of time. (Note that $\Sigma_0^M n_x=n$). With this notation, the sample mean ($\bar{x}$) and sample variance (s^2) are calculated by:

$$\bar{x} = \Sigma_0^M x n_x/n \qquad (3)$$

$$s^2 = \Sigma_0^M (x-\bar{x})^2 n_x/n \qquad (4)$$

As n increases, $\bar{x}$ approaches E{m}. In addition, it can be shown that if the count of accidents for each entity is Poisson distributed, as n increases, s^2 approaches VAR{m}+E{m}.[13] Thus, $\bar{x}$ can serve as an estimator of E{m} and $s^2-\bar{x}$ as an estimator of VAR{m}. The larger the reference population, the more accurate these estimates are. Because the estimators are based on $\bar{x}$ and s^2, this approach to estimation will be called the Method of Sample Moments.

Numerical Example 3.

In the previous numerical examples discussion revolved around the unsafety of a highway-rail grade crossing which is in an urban area, has a single track, is used by 2 trains per day, 550 vehicles per day and is equipped by crossbucks. Estimates of E{m} and of VAR{m} were used and the promise made to explain later where they come from. One can now show how such estimates are obtained by the Method of Sample Moments.

Consider a reference population of rail-highway grade crossings the defining traits of which are: {urban area, 1 track, 1 to 2 trains per day, 0 to 1000 vehicles per day, signed by crossbucks}. Data kept by the Federal Railroad Administration show that in 1980 there were about 10,000 crossings with such traits. In this reference population, 9770 crossings had 0 accidents, 160 had 1 accident, 8 had 2 accidents and 1 had 3 accidents. For this data, $\bar{x}=0.0181$ [acc.], $s^2=0.0199$ [acc.]^2 and therefore the estimates are $\hat{E}\{m\}=0.0181$ [acc./year] and $V\hat{A}R\{m\}=0.0018$ [acc./year]^2.

The attraction of the Method of Sample Moments is that its validity rests on a single assumption: that for each entity, if m remained constant, the occurrence of accidents would be well described by the Poisson probability law.

However, two practical difficulties are inherent in the Method of Sample Moments. First, it is only seldom that a sufficiently large data set can be found to allow for adequately accurate estimation of E{m} and VAR{m}. This limits the circumstances in which the EB method of unsafety estimation can be used in practice. Second, even with very large data sets, one can not find adequate reference populations for entities described by several traits. Thus, e.g., even with more than 200,000 grade crossings in the USA, use had to be made of the categories "1 to 2 trains per day" and "0 to 1000 vehicles per day" while the corresponding traits of the entity are "2 trains per day" and "550 vehicles per day." (It is for this reason that the estimates given here were not used in the previous numerical examples).

The Multivariate Method discussed below intends to obviate these difficulties.

b. The Multivariate Method.

In practice it is often necessary to estimate the unsafety of entities for which a sizeable reference population does not or can not exists. To illustrate if it is known about Mr. Smith (who has been introduced to the reader in section 2) that he has driven in 1984 about 13500 kilometres and has been convicted three times for speeding and once for impaired driving, conceptually there still is a reference population. Its defining traits are: {Ontario driver in 1984, male, 22 years old, 13500 km/year (8388 miles/year), 3 speeding convictions, 1 impaired driving conviction}. However, even though there are more than 5 million drivers in Ontario, the number who match all these traits is too small to allow reliable estimation of E{m} and VAR{m} by the Method of Sample Moments. In the limit, when traits are many or are continuous in nature (as are age or average traffic flow), a real reference population does not exist. In this case we can only ask what would be the mean and the variance of the m if a large reference population did exist. The concept of a reference population is still clearly defined but it is now an imaginary one. Since no data about it can exist, the Method of Sample Moments can not be used. It appears, however, that even in this case it is possible to estimate E{m} and VAR{m}. The idea for doing so is the same as that on which much of multivariate statistical analysis rests.

It is common to estimate the unsafety of entities with specified traits by multivariate statistical methods. Thus, e.g., there are "models" to estimate the unsafety of a grade crossing as a function of the train and car traffic, its location (urban or rural), its geometric characteristics (number of tracks, angle) and the type of warning devices used (crossbucks, flashers or gates). Similarly, there are multivariate models which estimate the unsafety of a driver as a function of age, gender and number of convictions for offenses associated with the use of the road. Multivariate models exist also for intersections and roads. The purpose of this section is to elucidate the relationship between such multivariate models, imaginary reference populations and the two moments E{m} and VAR{m}.

Imagine a population of entities with the same traits. In this imaginary reference population, as in a real one, m would still vary from entity to entity. To see why, consider a group of Ontario drivers who, just like Mr. Smith, are male, 22 years old, have driven 13500 km (8388 miles/year) and got 3 speeding and 1 impaired driving convictions 1984. One should still expect them to be different from each other and from Mr. Smith in many important respects - education, personality, where they drive, what cars they use, and so on. Therefore drivers in this imaginary reference population should also be expected to have m's which are not all the same. For this reason it is still meaningful to speak of the mean and the variance of m's in such a reference population, albeit an imaginary one.

When a multivariate model is being fitted to accident data, it is to estimate the E{m} as a function of independent variables. This is based on the belief that E{m} depends on the independent variables in some systematic way which can be captured by an equation or a set of equations - a model. These <u>independent variables</u> are precisely what has so far been called <u>traits</u>. Thus, each combination of traits (independent variables) defines an imaginary reference population.

When a multivariate model is used to estimate the m of some specific entity, the reasoning mimics that used earlier in section 2. To illustrate, if the aim is to estimate the m of Mr. Smith, the argument is that the estimate of E{m} obtained from the multivariate model pertains to that imaginary reference population which matches exactly Smith's traits. Because Mr. Smith is indistinguishable from others in that imaginary reference population, the best estimate for Smith's m is what E{m} for the imaginary reference population is estimated to be by the multivariate model.

Seeing the multivariate model in this light, as providing an estimate of E{m} for a continuum of imaginary reference populations, logically entails also a specific view of the residuals. A residual, in our case, is the difference between an accident count on some specific entity which served as "datum" for model fitting and the estimate of E{m} calculated from the fitted model equation. This residual consists of two parts. One is the difference between the m of the specific entity which happened to serve as datum and the count of accidents on it. This part is usually taken to be the difference between a Poisson-distributed count (x) and its mean (m). The other part is the difference between the m of the specific entity and which served as a datum and the estimate of E{m} for the imaginary reference population to which the entity belongs. In other words, x is a realization of a Poisson random variable the mean of which is m while m is a random variable from a reference population in which m's have a distribution with E{m} and VAR{m} as central moments. In short, the residuals are viewed as coming from a family of compound Poisson distributions.

For compound Poisson distributions (see, e.g., reference 13)

$$VAR\{x\} = VAR\{m\} + E\{m\} \quad . \tag{5}$$

For the reference population to which the entity the accident count of which was used as a datum belongs, E{m} can be estimated using the model equation and VAR{x} can be estimated by the squared residual. Therefore, based on the relationship in equation 5, the difference [squared residual - estimate of E{m}] can be used to estimate the VAR{m} for the imaginary reference population to which this datum point belongs.

Thus, for each datum point there is an estimate of VAR{m} which is associated with a set of traits (independent variables) defining an imaginary reference population. Just as E{m} is related to the independent variables by some regularities which can be captured by a model equation, so might VAR{m} be. Therefore, the estimates of VAR{m} for every datum point can also be subjected to a multivariate analysis the aim of which is to express VAR{m} as a function of the independent variables.

In previous work it has been found that when estimates of VAR{m} are plotted against the estimates of E{m}, the two are systematically associated.[8] Their relationship could be adequately represented by the simple quadratic function $VAR\{m\} = [E\{m\}]^2/k$. When, in view of the data at hand, such a quadratic relationship is justified, this is a particularly simple way to link VAR{m} with the independent variables since these are used to estimate E{m}. Only one added parameter, 1/k, needs to be estimated.

Before such estimation is attempted it is advisable to examine the data in the following manner. For each datum point calculate $\hat{E}\{m\}$ and $V\hat{A}R\{m\}$. Arrange the datum points in the ascending order by $\hat{E}\{m\}$. Form groups and for each group calculate the average of the $\hat{E}\{m\}$ and $V\hat{A}R\{m\}$.[1] Plot these averages one against

---

[1]As shown in 3c, a correction term C has to be subtracted from the average of the $V\hat{A}R\{m\}$. It is given by $C=\Delta^2/12$ where Δ is the difference between the smallest and largest $\hat{E}\{m\}$ in the group.

the other to see whether the simple quadratic relationship fits the data. If yes, use $[\hat{E}\{m\}]^2$ as the independent variable in a simple linear regression with zero intercept to obtain an initial estimate of 1/k. A better estimate of k can than be obtained by the method of maximum likelihood. The theoretical considerations and the empirical support for the choice of this simple quadratic relationship are given in the appendix of reference 14.

Whether the Multivariate Method for estimating E{m} and VAR{m} is sound remains to be examined. However, at this point it is possible to illustrate how it might be used.

Numerical Example 4.

Consider again the highway rail grade crossing of the preceding numerical examples. Using a multivariate model developed earlier E{m} is estimated by $\hat{E}\{m\}=$ $a\times(\text{cars/day})^{b}\times(\text{trains/day})^{c}\times(\text{trains/day})^{d\times\ln(\text{trains/day})}$ =0.0239 accidents/year. In this, a=-5.345, b=0.405, c=1.039 and d=-0.115, which are parameter estimates specific to the traits: {urban, single track, cross-bucks}. In the same multivariate model for the set of traits {urban, single track, urban} the estimate of k was 0.52. Thus, VAR{m} is estimated to be $V\hat{A}R\{m\}=$ $0.0239^2/0.52=0.0011$ $[\text{acc./year}]^2$. These are the estimates of E{m} and VAR{m} used in numerical example 1.

The advantages of the Multivariate Method are two and they are linked. First, it provides estimates of E{m} and VAR{m} which can exactly match many traits. Traits which are continuous in nature cease to be a problem. Second, one does not need to have a large reference population for any particular combination of traits. However, because Multivariate Method uses familiar statistical techniques in unfamiliar ways, its correctness has to be examined and established.

c. Comparing the two Methods

The conceptual foundations of the Multivariate Method are straightforward and the method should work. However, there might difficulties of pragmatic nature. First, when fitting a model to data, there are seldom good reasons for choosing this or that form of model equation. However, the choice of model form determines what the residuals are. Second, the estimation of model parameters involves a minimization of the sum of squared residuals (or something close to it). Thus, there might be a systematic underestimation of the VAR{m}. Third, it could be that VAR{m} does not depend on the independent variables in some regular manner and in particular that $VAR\{m\} \neq [E\{m\}]^2/k$. For all these reasons it is necessary to demonstrate that the multivariate method and the method of sample moments give similar results.

The estimates of E{m} and VAR{m} obtained by the Method of Sample Moments can not be directly compared to estimates produced by Multivariate Method. The imaginary reference population of which the Multivariate Method speaks is always a subset of the reference population used by the Method of Sample Moments. Since the two reference populations do not coincide, the estimates of E{m} and VAR{m} will differ. This can be illustrated by comparing the results of the numerical examples introduced earlier.

The reference population in numerical example 3 uses the categories 1 to 2 trains/day and 0 to 1000 vehicles/day. For that reference population, E{m} was estimated to be 0.0181 acc./year. However, in numerical example 4, the specific values of 2 trains/day and 550 cars per day were used. This makes for an imaginary reference population which is more homogeneous in its traits for which $\hat{E}\{m\}=0.0239$ acc./year. Similarly, for the reference population in numerical example 3 $\hat{VAR}\{m\}=0.0018$ $[acc./year]^2$ whereas for the imaginary reference population of numerical example 4 we estimate 0.0011 $[acc./year]^2$. This is in agreement with what one should expect in general. The more homogeneous is a group of entities in terms of its traits, the lesser should be the variance of its m's.

Thus, in order to compare estimates obtained by the Method of Sample Moments to estimates obtained by the Multivariate Method, it is necessary to establish what is the relationship between the estimates of E{m} and VAR{m} obtained by the two methods.

Consider two entities, i and j which differ in the magnitude of an independent variable as is shown in figure 2. These entities have m_i and m_j. Because of the difference in the magnitude of the independent variable, i and j belong separate imaginary reference populations, I and J. The m's in reference population I have moments $E_I\{m\}$, $VAR_I\{m\}$ and the m's in reference population J have moments $E_J\{m\}$ and $VAR_J\{m\}$.

If the two imaginary reference populations I and J are merged to form one joint reference population N and if I and J are thought to have the same number of entities,

$$E_N\{m\} = (E_I\{m\} + E_J\{m\})/2 \qquad (6)$$

The aim is to find the $VAR_N\{m\}$ of the joint reference population N. The contribution of the entities of reference population I to $VAR_N\{m\}$ is the $VAR_I\{m\}$ (which was calculated with $E_I\{m\}$ as mean) to which $(E_N\{m\}-E_I\{m\})^2$ must be added. Similarly, entities of reference population J contributed, $VAR_J\{m\}$ to which $(E_N\{m\}-E_J\{m\})^2$ is added. If I and J are thought to have the same number of entities, the variance of m's in N will be

$$VAR_N\{m\} = [VAR_I\{m\}+VAR_J\{m\}+(E_N\{m\}-E_I\{m\})^2+(E_N\{m\}-E_J\{m\})^2]/2 \qquad (7)$$

Equations 6 and 7 are easily generalized to any number of reference populations which are being joined. If N is composed of reference populations I, J, K,..., $VAR_N\{m\}$ is the sum of the average of $VAR_I\{m\}$, $VAR_J\{m\}$, $VAR_K\{m\}$,... and the average of $(E_N\{m\}-E_I\{m\})^2$, $(E_N\{m\}-E_J\{m\})^2$, $(E_N\{m\}-E_K\{m\})^2$, ... In the specific case when N is made up of many reference populations and their E{m}s are distributed uniformly over an interval of size Δ, the average of the terms $(E_N\{m\}-E_I\{m\})^2$, $(E_N\{m\}-E_J\{m\})^2$, $(E_N\{m\}-E_K\{m\})^2$,... is $\Delta^2/12$.

Equations 6 and 7 facilitate the comparisons of estimates by the two methods. Estimates for the left hand side of equations 6 and 7 will be obtained by the Method of Sample Moments (as, e.g., in numerical example 1). Estimates for the right hand side of equations 6 and 7 will be obtained (for the same data) by the Multivariate Method (as, e.g., in numerical example 2) If the interpretation in section 3b is correct, and if the potential problems mentioned in the first paragraph of 3c are not serious, there should be an approximate numerical equality between the two sides.

This kind of comparison requires data sets which allow the estimation of E(m) and VAR(m) by both methods for many reference populations. Furthermore, successful multivariate models for the same data sets have to be available. In section 4 the data sets and the multivariate models used will be described.

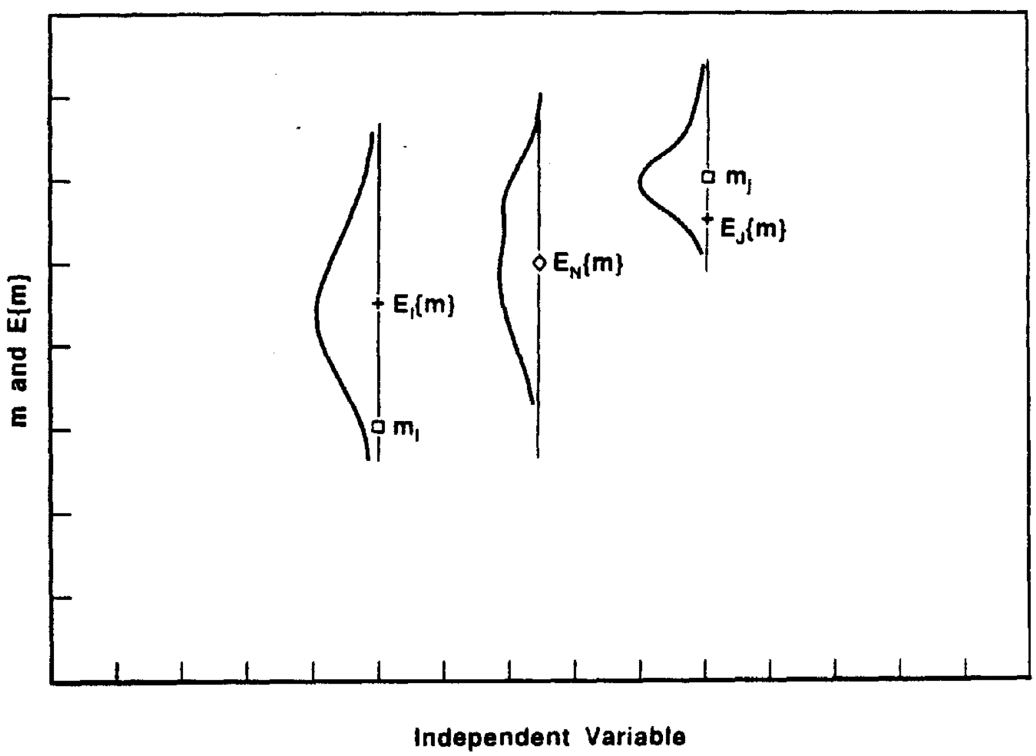

Figure 2. The relationship between m and E(m) in reference populations I, J and their union N.

4. DATA, MODELS AND HOW THEY ARE USED

Having created the machinery by which the performance of the Multivariate Method can be judged, the next step is to assemble the data which will allow the empirical investigation to be carried out and to specify the multivariate models which are needed for that purpose. Two data sets were used. One pertains to rail-highway grade crossings in the U.S. the other to drivers in Ontario. For each data set we have specified numerous multivariate models and estimated their parameters.

a. Rail-highway grade crossings in the U.S..

The first data set contains information about some 200,000 rail-highway grade crossings in the United States during the period 1980 through 1984. For each crossing we know its accident history, train traffic, vehicular traffic, number of tracks, setting, warning device used and other traits.

To estimate E{m} use has been made of the model equation

$$\hat{E}\{m\}=a\times(\text{cars/day})^{b}\times(\text{trains/day})^{c}\times(\text{trains/day})^{d\times\ln(\text{trains/day})} \quad (8)$$

Eight sets of parameters a,b,c and d have been estimated, one for each combination of the three traits : setting (urban or rural), number of tracks (single or multiple) and warning device (crossbuck or flashers).[8] To illustrate how VAR(m) is estimated consider figure 3 which shows what is revealed by the data for the 41,822 crossings which are in an urban setting and are equipped with crossbucks.

To produce figure 3, for each crossing the E{m} of its imaginary reference population has been estimated. The dimension of Ê{m} is accidents in 5 years. When a crossing has a single track one set of parameters values is used to estimate E{m}, when it has a multiple track, another. Using the estimates Ê{m} we calculated for each crossing the squared residual (accidents in 5 years - $\hat{E}\{m\})^2$. From this squared residual Ê{m} is subtracted to give an estimate of VAR(m) in accordance with equation 5. All crossings are then ranked by their Ê{m} and groups of about 1000 crossings with similar Ê{m}'s are created. For each group the

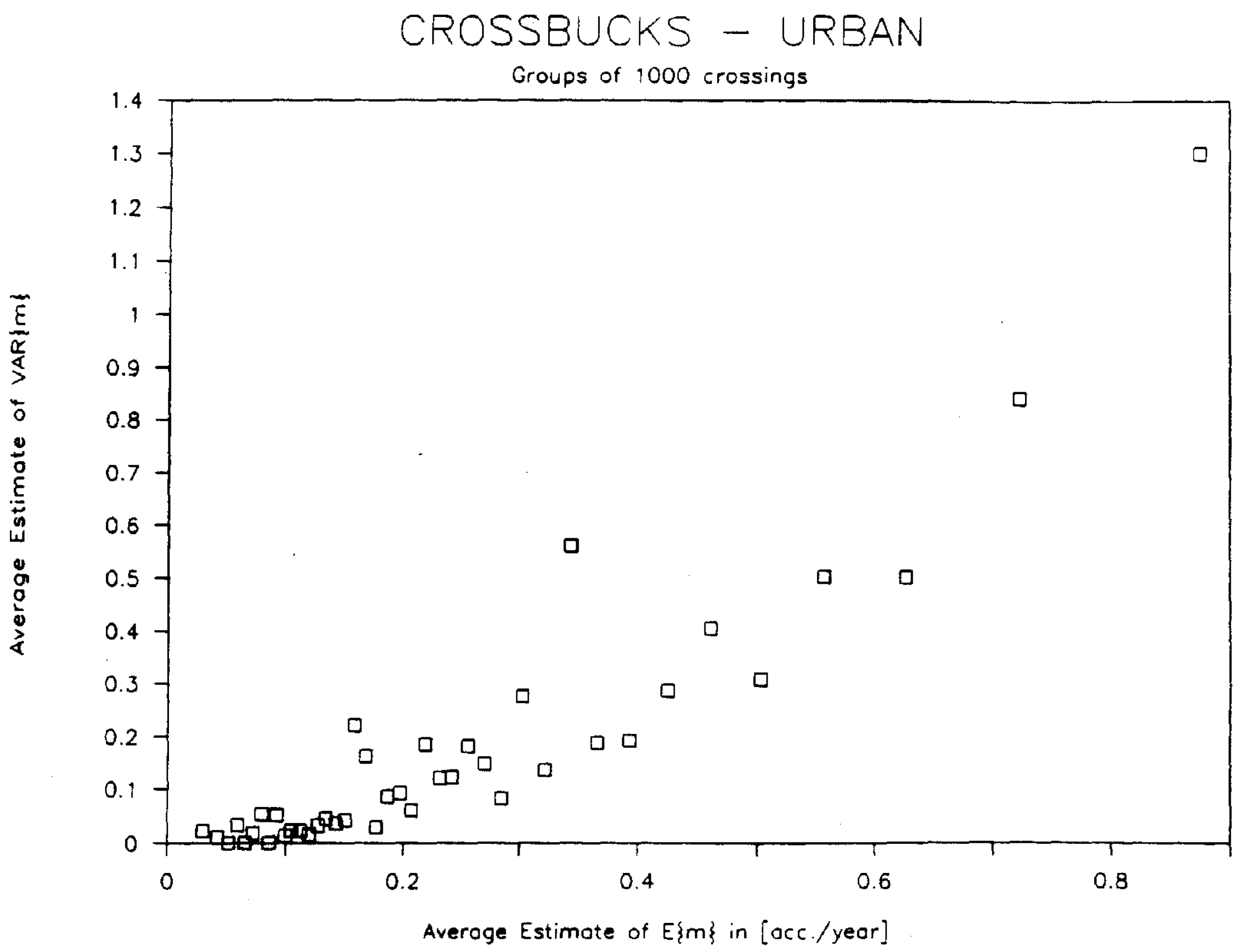

Figure 3. The relationship between E{m} and VAR(m).

average $\hat{E}\{m\}$ and the average $V\hat{A}R\{m\}$ are calculated. From the average of the $V\hat{A}R\{m\}$ the correction C (see footnote in section 3b) is subtracted. The results are shown in figure 3. As in most other cases, a relationship of the form $VAR\{m\} = (E\{m\})^2/k$ is indicated and k is estimated to be 0.61.

Similar relationship between the average estimates of E{m} and VAR{m} have been found when all eight cases have been examined separately.

b. Ontario Drivers

The second data set contains information about a random sample of some 170,000 Ontario drivers for the period 1981 to 1984. These are drivers who had at least one conviction for a traffic law violation in either 1981 or 1982. For each driver information about age, gender, kind of license, record of accidents, record of convictions is available. For further details see reference 15.

To estimate the E{m} of a driver (for 1983-1984 based on 1981-1982 data) dozens of linear multivariate models have been formulated using various combinations of age, gender, 14 conviction types and 2 accident types as independent variables. Estimates of the parameters for one such model are given in table 1 .

To illustrate the estimation of E{m} with this model, consider a female driver, 24 years old, who during 1981 and 1982 had two speeding convictions and one conviction for failing to yield the right of way. For her the expected number of accidents in the two years 1983 and 1984 is estimated to be: 0.181 (intercept or "base driver") - 0.059 (for being female) - 0.065 (for being 24) + 2*0.029 (for two speeding convictions) + 0.027 (for failing to yield) = 0.142.

Figure 4 shows what is revealed by the data for these 170,000 Ontario drivers. To produce figure 4 for each of the 170,000 drivers the $\hat{E}\{m\}$ has been calculated using the parameters in table 1. Next the squared residual has been calculated for each driver using his or her accident record. From the squared residual $\hat{E}\{m\}$ has been subtracted to get $V\hat{A}R\{m\}$. After ranking

Table 1. Estimates of regression coefficients.[15]

Regression Run A1: Age, sex as dummies, no accidents

Description of variable	Estimated Coefficient	Standard Error
Intercept (Male, Age < 21)	0.18135	0.002632
Dummy variable for age 21-25	-0.06523	0.001591
Dummy variable for age 26-30	-0.03552	0.002618
Dummy variable for age 31-35	-0.05356	0.002681
Dummy variable for age 36-40	-0.05783	0.002804
Dummy variable for age 41-50	-0.04957	0.003239
Dummy variable for age 51-60	-0.05717	0.003234
Dummy variable for age > 60	-0.05588	0.004423
Dummy variable for female	-0.05899	0.006264
mm1: Fail to yield, imp. turns, PXO, amber violations, etc	0.02736	0.001646
mm2: Disobey red lights; rail crossing violations.	0.045535	0.002791
mm3: Fail report accid.; Careless driving; Dang. driving;Crim. neg. caus. death	0.048755	0.004567
mm4: Unsafe move;imp. o'taking; Disobey signs	0.059365	0.003723
mm5: Fail to remain; breath test; alcohol; impairment.	0.068515	0.015328
mm6: Seat belt; F.T.C.; parking; divided h'way offences.	0.03421	0.001961
nn1: Minor neglect: no drivers license; permits; insurance, address change	0.031975	0.009674
nn2: Neglect· D/Lic. suspended or not produced; plates, insurance.	0.03312	0.003617
vv1: Minor vehicle neglect: lamps windows obstructed, etc.	0.051595	0.006479
vv2: Vehicle neglect: Unsafe veh., brakes, tires	0.08548	0.009013
vv3: Truck weight offences.	0.187235	0.025898
vv4: Truck dimension offences.	0.11554	0.011842
ee1: Environmental offences: noise, fumes.	0.091265	0.008340
mm7: Speeding offences	0.02902	0.000840

all drivers in the order of their $\hat{E}\{m\}$, groups of 1000 drivers were created and the average $\hat{E}\{m\}$ and $V\hat{A}R\{m\}$ calculated for each. From the average $V\hat{A}R\{m\}$ the correction C has been subtracted. The corrected values were used in figure 4.

Since at E{m}=0, VAR{m} must be 0, the squares clearly indicate a non-linear relationship. As in the other cases, the quadratic relations $VAR\{m\}=[E\{m\}]^2/k$ can be used to fit the squares in figure 4. The best fitting curve is not shown in order to preserve the impression that even with very large data sets the detailed nature of the relationship between E{m} and VAR{m} can be difficult to establish.

With these large data sets ready and the corresponding multivariate models available, it is possible to compare estimates of E{m} and VAR{m} obtained by the Multivariate Method and the Method of Sample Moments.

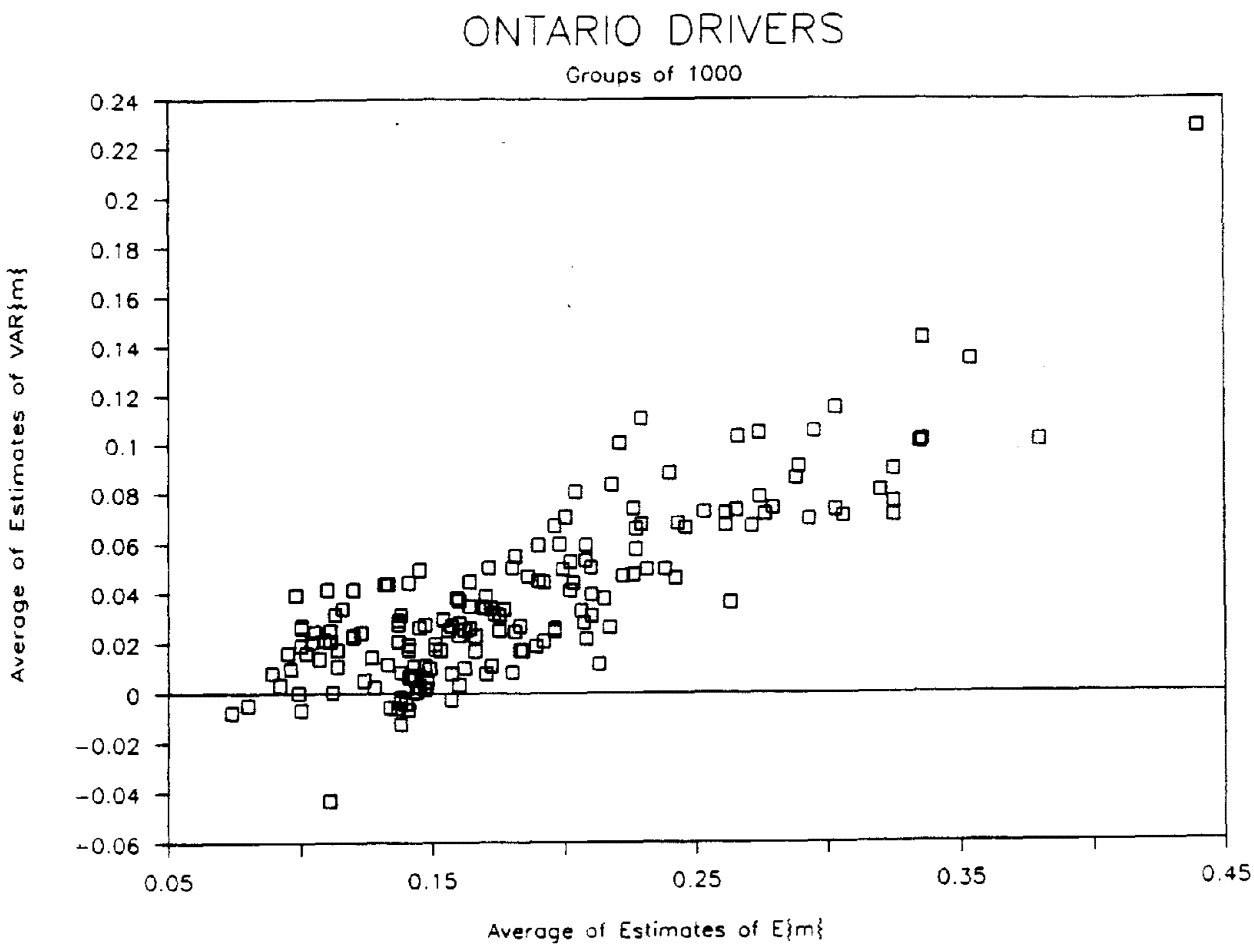

Figure 4. The relationship between E{m} and VAR{m}.

5. IS THE MULTIVARIATE METHOD SOUND?

To use the clues to unsafety which are contained in the information about the traits of an entity, estimates of E{m} and VAR{m} for its reference population are required. When the reference population is large, estimates of E{m} and VAR{m} can be obtained by the Method of Sample Moments (section 3a). However, often the requisite reference population does not exist or is too small to yield useful estimates. In those cases the Multivariate Method (section 3b) has to be used. While the conceptual foundations of the Multivariate Method appear to be sound, an empirical proof of the pudding is in order. This will be done by juxtaposing estimates obtained by the Method of Sample Moments to those obtained by Multivariate Method in several large reference populations. Should both methods yield similar results, this will be construed as evidence in support of the soundness of the Multivariate Method.

a. Rail-Highway Grade Crossings.

Consider a group of 10,000 grade crossings which have been selected at random from the data base. The crossings in this group have diverse traits. Some are in an urban area, equipped with flashers and serve heavy traffic; others are in a rural area, protected by crossbucks and lightly used. Over a period of 5 years these crossings recorded 1750 accidents, an average of 0.1750 accidents per crossing. Thus, by the Method of Sample Moments, $\bar{x}=\hat{E}\{m\}=0.1750$.

Alternatively, for every one of these 10,000 grade crossings m can be estimated using model equation 8 with parameters a,b,c and d which fit the traits of a crossing. The average of such estimates of m for all 10,000 crossings is 0.1752 accidents per crossing. Thus, by the Multivariate Method, $\hat{E}\{m\}=0.1752$. The conclusions is that for this group of 10,000 crossings the estimates of E{m} by the two methods correspond closely.

A similar comparison for 10 different groups of 10,000 grade crossings each is given in table 2.

Table 2.

Comparison of $\hat{E}\{m\}$ obtained by the 2 methods in 10 groups of 10,000 crossings. [Accidents in 5 years per crossing]

Group Number	Sample Mean	Average of Model Estimates
1	0.1750	0.1752
2	0.1750	0.1758
3	0.1706	0.1771
4	0.1801	0.1751
5	0.1780	0.1773
6	0.1819	0.1772
7	0.1764	0.1765
8	0.1737	0.1763
9	0.1765	0.1727
10	0.1664	0.1726
Mean	0.1754	0.1756
$\hat{\sigma}$	0.0042	0.0016
$\hat{\sigma}/\sqrt{10}$	0.0013	0.0005

The satisfactory correspondence between the two columns in table 2 indicates merely that the multivariate model has been correctly fitted to the data and that it performs satisfactorily in randomly selected subsets of it. The question is whether the estimates of VAR{m} obtained by both methods also match.

Consider again the group of 10,000 crossings examined earlier. Estimating by the Method of Sample Moments $\hat{VAR}\{m\}=0.1624$ [accidents in 5 years]^2. The Multivariate Method requires that estimate $\hat{VAR}_i\{m\}$ and $(E\{m\}-E_i\{m\})^2$ for i=1,2,...,10,000 be obtained, added, and their averages calculated (see section 3c). In this case, $\hat{VAR}\{m\}=0.1119+0.0385=0.1504$ [accidents in 5 years]^2. Thus, the two methods also give similar estimates of VAR{m}. A comparison for 10 groups of 10,000 grade crossings is given in table 3.

Table 3.

Comparison of the $\hat{VAR}(m)$ obtained by 2 methods in 10 groups of 10,000 crossings. [Accidents in 5 years per crossing]^2

Group No.	Method of Moments $\hat{VAR}(m)$	Multivariate Method $\hat{VAR}(m)$	$\hat{VAR}_i(m)$
1	0.1624	0.1504	= 0.1119+0.0385
2	0.1344	0.1492	= 0.1112+0.0380
3	0.1523	0.1520	= 0.1133+0.0387
4	0.1514	0.1496	= 0.1111+0.0385
5	0.1379	0.1535	= 0.1139+0.0396
6	0.1651	0.1576	= 0.1168+0.0408
7	0.1599	0.1574	= 0.1163+0.0411
8	0.1366	0.1474	= 0.1106+0.0368
9	0.1466	0.1494	= 0.1108+0.0386
10	0.1419	0.1453	= 0.1081+0.0372
Mean	0.1488	0.1512	
$\hat{\sigma}$	0.0106	0.0038	
$\hat{\sigma}/\sqrt{10}$	0.0034	0.0012	

Several observations follow. First, in the population of U.S. rail-highway grade crossings and in its randomly selected subsets, both methods produce practically identical estimates of VAR(m). Viewing the difference of the means (0.1488-0.1512=-0.0024) in the light of $\hat{\sigma}/\sqrt{10}$, no difference between the estimates can be detected. Therefore, since the Method of Sample Moments gives sound estimates, so does the Multivariate Method. Second, comparing the σ's one is led to conclude that the Multivariate Method yields estimates which have the lesser standard error. Third, when the population is diverse (as this one is) there is a substantial difference between the VAR(m) which applies to the entire group of 10,000 crossings and the $VAR_i(m)$ which applies to the imaginary reference populations for each crossing from which the group is constituted. The former (≈0.15) is always larger than the latter (≈0.11).

The last observation requires elaboration. When the unsafety of entity i is to be estimated, its $\hat{VAR}_i\{m\}$ is needed. If the Method of Sample Moments is used to provide an estimate of $\hat{VAR}_i\{m\}$ and the reference population is diverse in its traits, a correction "C" needs to be applied. In this case, C≈0.04.

Next several subpopulations of the U.S. grade crossings are examined. However, the mode of comparison is slightly altered. In tables 2 and 3 crossings were chosen at random to form groups of 10,000. As a result, each group consisted of crossings which are diverse in their traits and therefore in their m's. In contrast, for the examination below, groups of crossings which are as homogeneous in their m's as possible will be used.

The first such set of crossings are all in a rural setting and equipped with flashers. In figure 5 estimates of E{m} obtained by the two methods are compared, figure 6 is the comparison of the two estimates of VAR{m}. The slope of the regression line in figure 5 is 1.009 with a standard error of 0.020. This merely signifies that the model has been properly fitted. The slope of the regression line in figure 6 is 1.028 with a standard error of 0.025. The good correspondence of the two estimates of VAR{m} supports the soundness of the Multivariate Method.

Figure 7 shows how the two estimates of VAR{m} match for crossings in a rural setting but equipped with crossbucks. Here the slope of the regression line is 0.972 with a standard error of 0.0256. In figure 8 a similar comparison is shown for crossings in an urban area which are equipped by flashers. Here the regression line for the two estimates has a slope of 1.003 with a standard error of 0.021. In figure 9 estimates of VAR{m} for crossings in an urban area and equipped with crossbucks are compared. The slope of the regression line is 1.002 and the standard error is 0.014.

It can be concluded that for this data the two methods produce similar estimates of E{m} and VAR{m}.

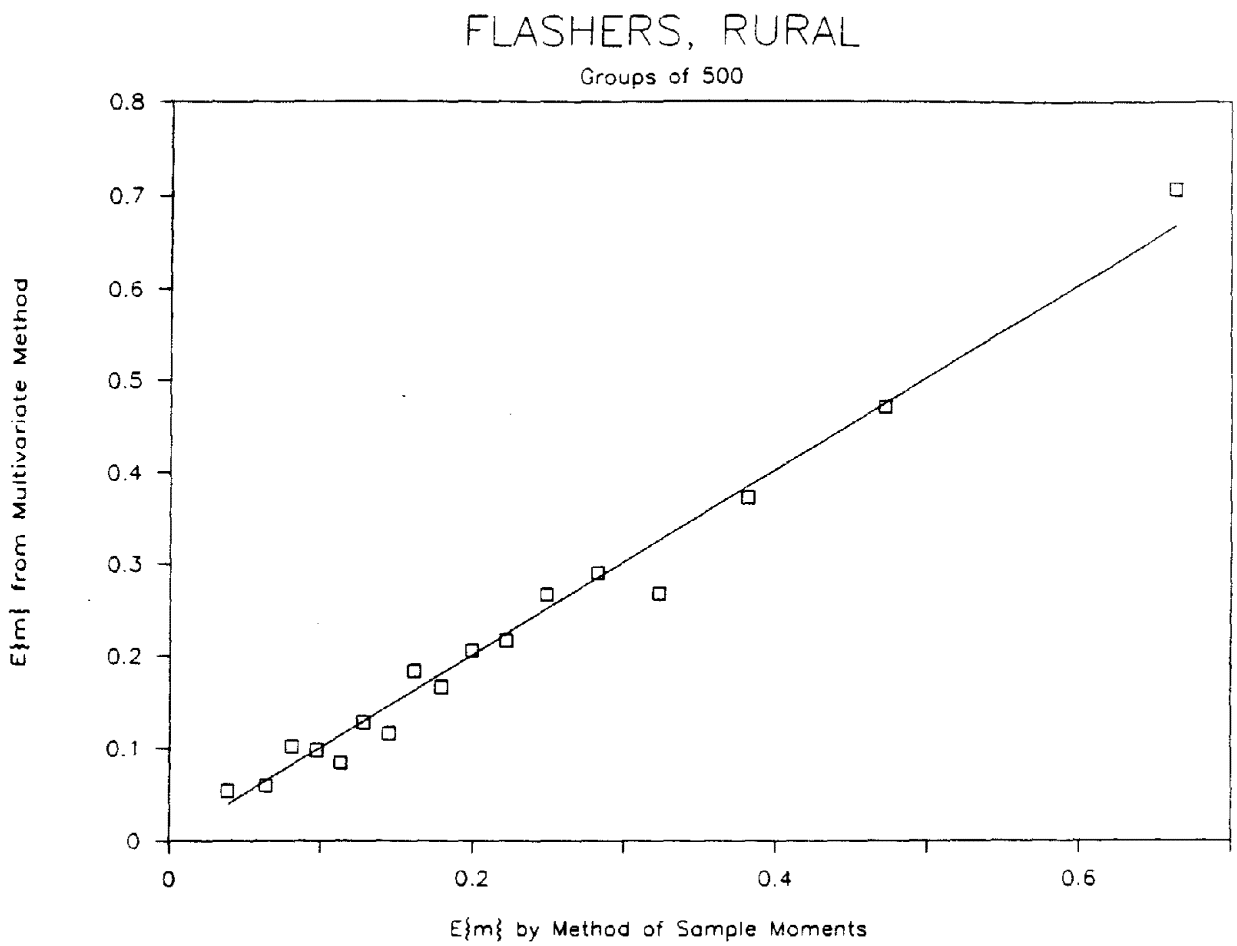

Figure 5. Comparison of E{m} for rural flashers estimated by two **methods.**

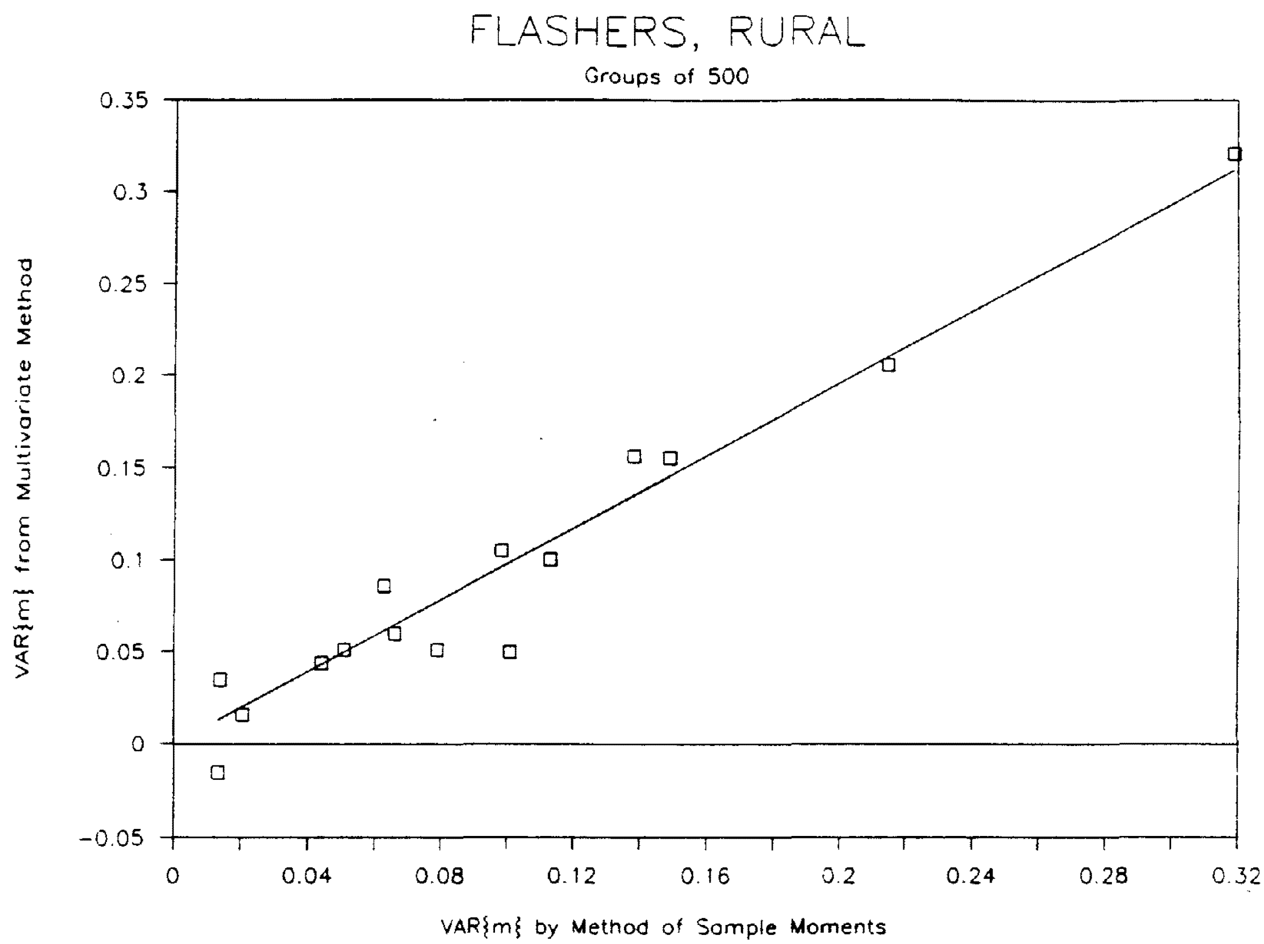

Figure 6. Comparison of VAR{m} for rural flashers estimated by two methods.

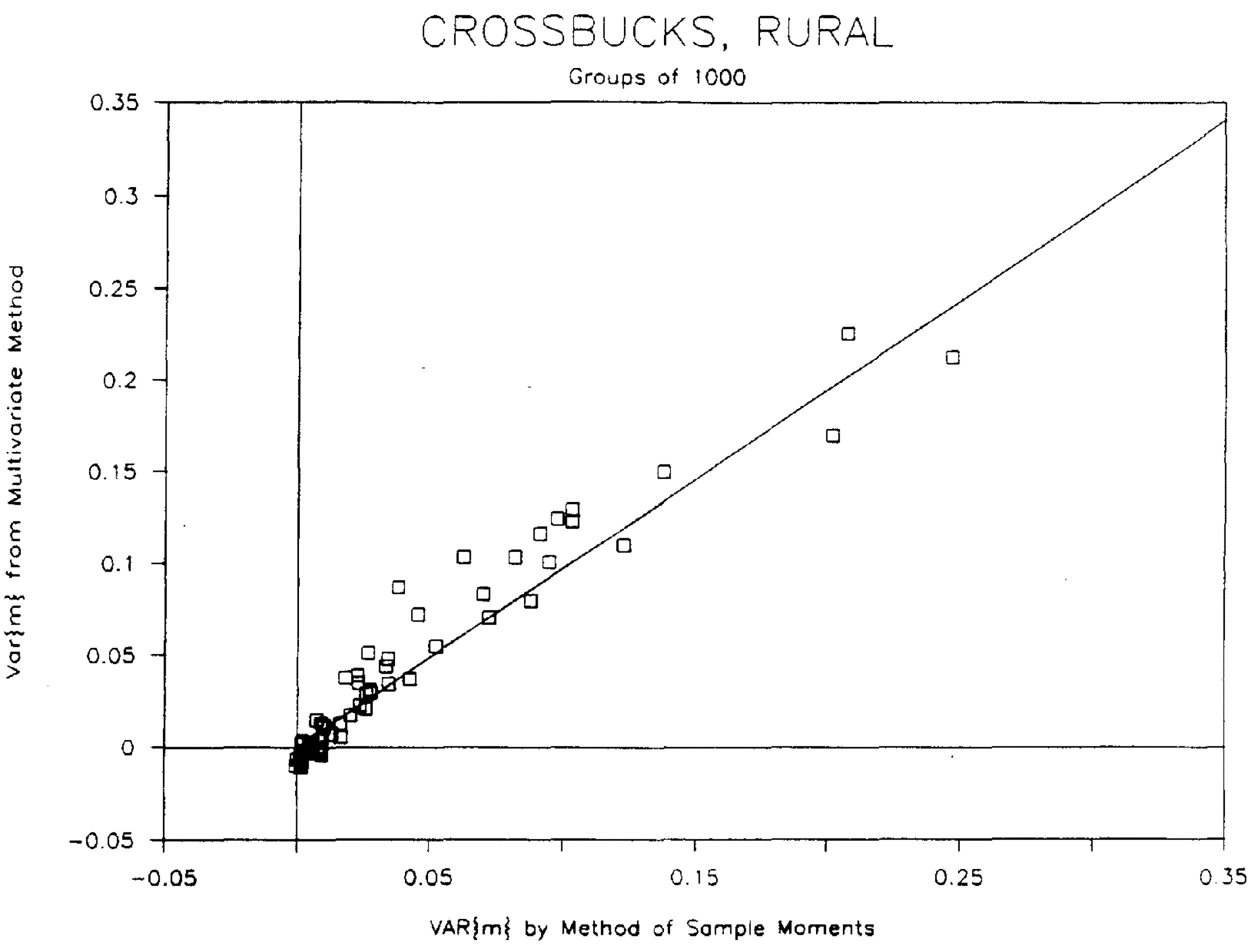

Figure 7. Comparison of VAR{m} for rural crossbucks estimated by two methods.

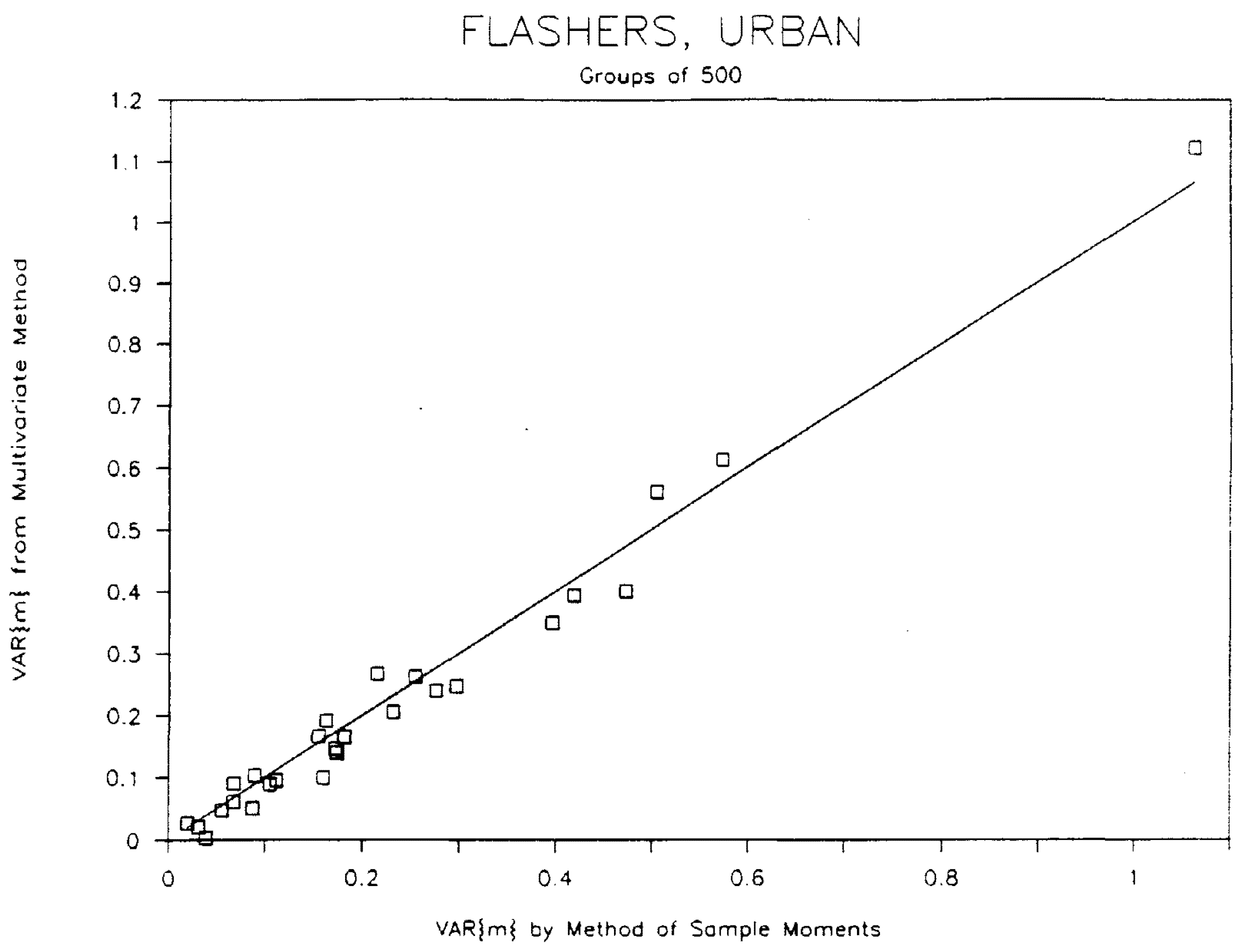

Figure 8. Comparisons of VAR(m) for urban flashers estimated by two methods.

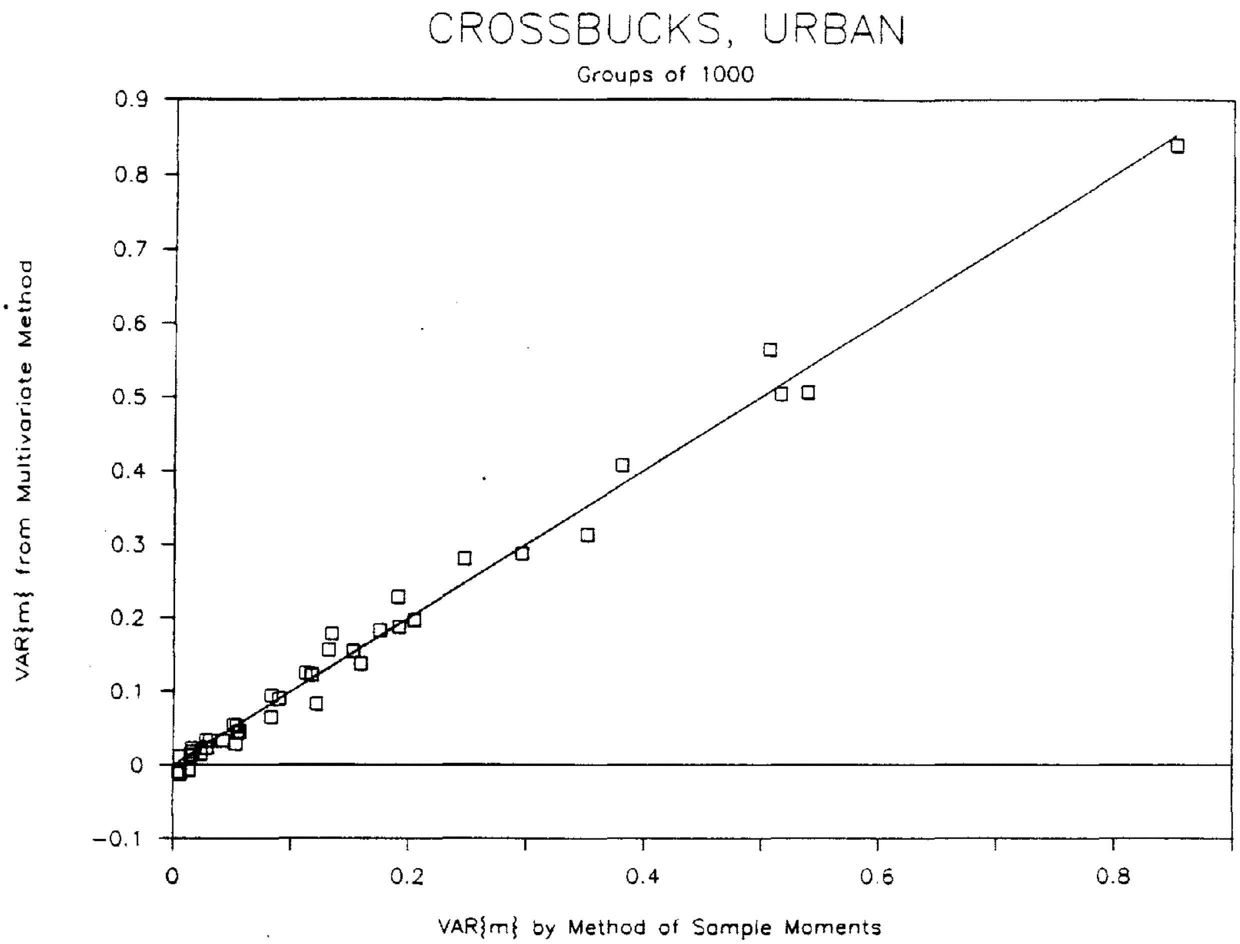

Figure 9. Comparison of VAR(m) for urban crossbucks estimated by two methods.

b. Ontario Drivers

Consider a group of 10,000 drivers selected at random from our data base. These drivers have diverse traits. Some may be male, 18 year old, with many convictions and accidents in 1981 through 1982; others may be female, 40 years old with no convictions and accidents in 1981 through 1982. In the 2 years 1983 through 1984, drivers of this group recorded 1648 accidents, an average of 0.1648 accidents per driver. This is the estimate of E{m} by the Method of Sample Moments. Alternatively, the m of every one of these 10,000 drivers can be estimated using the multivariate model. The average of these 10,000 estimates of m is 0.1691 accidents per driver. This is the estimate of E{m} by the Multivariate Method. Thus, the two methods give similar estimates of E{m}. A similar comparison for 16 groups of 10,000 drivers is given in table 4. The results show only that the multivariate model has been correctly fitted to the data and performs well in random subsets of it.

To see whether the estimates of VAR{m} obtained by the Multivariate Method match those obtained by the Method of Sample Moments, consider again the same group of 10,000 drivers. Estimating by the Method of Sample Moments, $\hat{VAR}\{m\}=0.0118$ [accidents in 2 years]^2. The Multivariate Method requires that estimates of $\hat{VAR}_i\{m\}$ and $(E\{m\}-Ei\{m\})^2$ be obtained for drivers i=1,2,..., 10,000 and then averaged. In this case, $\hat{VAR}\{m\}= 0.0132+0.0036 =0.0168$ [accidents in 2 years]^2. Similar comparisons for 16 different groups of 10,000 drivers are given in table 5.

Several observations follow. First, in the population of Ontario drivers and in its randomly selected subsets, both methods produce similar estimates of VAR{m}. Viewing the difference of the means (0.0174-0.0168=0.0006) in the light of $\hat{\sigma}/\sqrt{16}$, no difference between the estimates can be detected. It follows that the Multivariate Method performs well in this data set. Second, comparing the $\hat{\sigma}$'s one is led to conclude that the Multivariate Method yields estimates which have the lesser standard error. Third, when the groups are diverse there is a substantial difference between the VAR{m} of the group the $VAR_i\{m\}$ which applies to the imaginary reference populations for each driver. The former (≈0.017) is always larger than the latter (≈0.013).

Table 4.

Comparison of $\hat{E}\{m\}$ obtained by the 2 methods
in 16 groups of 10,000 drivers.
[Accidents in 2 years per driver]

Group Number	Sample Mean	Average of Model Estimates
1	0.1648	0.1691
2	0.1711	0.1712
3	0.1723	0.1711
4	0.1750	0.1713
5	0.1631	0.1682
6	0.1704	0.1700
7	0.1720	0.1691
8	0.1720	0.1694
9	0.1726	0.1694
10	0.1723	0.1687
11	0.1629	0.1697
12	0.1699	0.1700
13	0.1711	0.1691
14	0.1604	0.1692
15	0.1773	0.1698
16	0.1649	0.1681
Mean	0.1695	0.1696
$\hat{\sigma}$	0.0048	0.0010
$\hat{\sigma}/\sqrt{16}$	0.0012	0.0002

Table 5.

Comparison of the $\hat{VAR}\{m\}$ obtained by 2 methods in 16 groups of 10,000 drivers. [Accidents in 2 years per driver]^2

Group No.	Method of Moments $\hat{VAR}\{m\}$	Multivariate Method $\hat{VAR}\{m\}$	= $\hat{VAR}_i\{m\}$	+correction
1	0.0118	0.0168	0.0132	0.0036
2	0.0147	0.0171	0.0135	0.0036
3	0.0147	0.0170	0.0135	0.0035
4	0.0178	0.0173	0.0136	0.0037
5	0.0166	0.0168	0.0131	0.0037
6	0.0166	0.0168	0.0133	0.0035
7	0.0236	0.0168	0.0133	0.0036
8	0.0170	0.0166	0.0132	0.0034
9	0.0214	0.0168	0.0133	0.0036
10	0.0145	0.0166	0.0132	0.0035
11	0.0187	0.0167	0.0133	0.0035
12	0.0213	0.0170	0.0134	0.0036
13	0.0193	0.0167	0.0132	0.0035
14	0.0172	0.0168	0.0132	0.0035
15	0.0188	0.0168	0.0133	0.0035
16	0.0144	0.0165	0.0131	0.0035
Mean	0.0174	0.0168	0.0133	0.0035
$\hat{\sigma}$	0.0031	0.0002	0.0001	0.0001
$\hat{\sigma}/\sqrt{16}$	0.0008	0.0000	0.0000	0.0000

The groups in tables 4 and 5 were formed by selecting drivers at random. As a result, drivers of diverse traits belong to the same group. For the examination to follow, drivers which are as homogeneous in their m's as possible placed in the same group. For groups of 1000 Ontario drivers who within each group are estimated to have similar m's, figure 10 compares estimates of E{m} by the 2 methods, figure 11 is the comparison of the 2 estimates of VAR{m}.

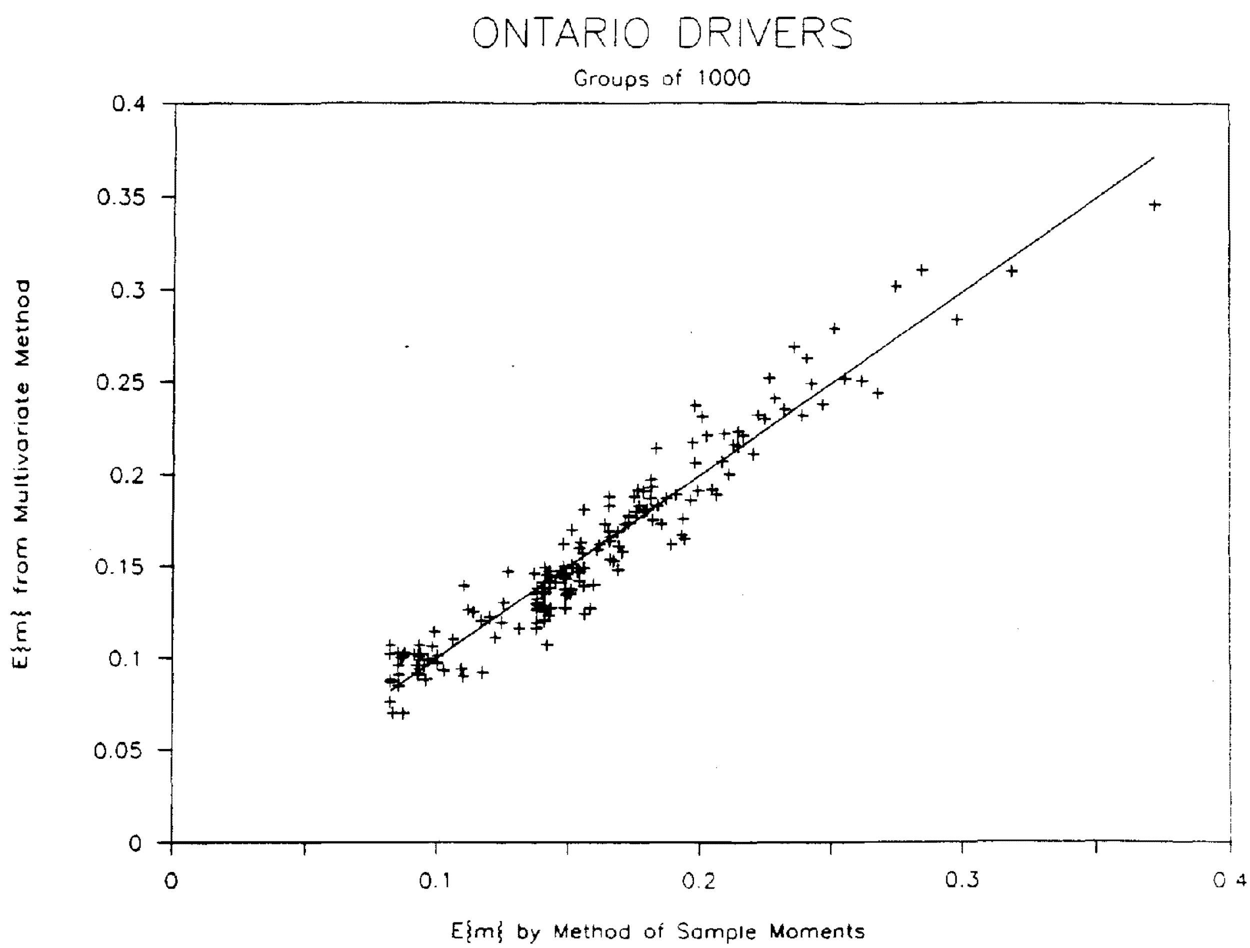

Figure 10. Comparison of E{m} for Ontario drivers estimated by two methods.

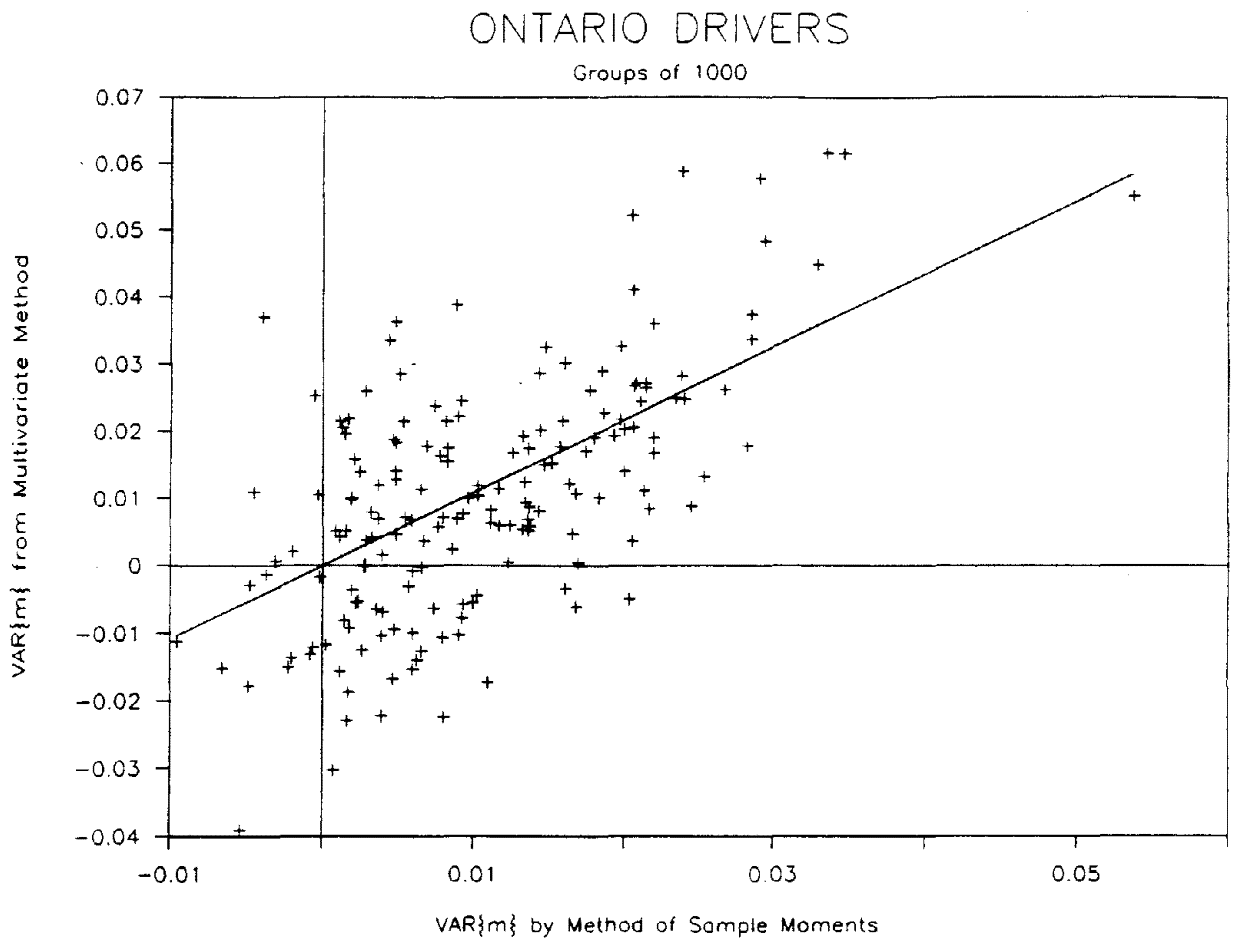

Figure 11. Comparison of VAR{m} for Ontario drivers estimated by two methods.

It is ~~quite~~ obvious that in this mode of comparison too, the two methods produce very similar estimates of E{m}. The slope of the regression line in figure 10 is 0.999 with a standard error of 0.006. This merely signifies that the model has been properly fitted.

The slope of the regression line in figure 11 is 1.085 with a standard error of 0.075. Thus, on the average, there is also a good correspondence between the two estimates of VAR{m}. However, while the visual impact of figures 6 through 9 for the grade crossings was that of a good correspondence between each two estimates of VAR{m}, figure 11 gives a different impression. This is partly an illusion because the scale in figure 11 differs from that in figures 6 through 9 by a factor of 5 to 10. Were one to rescale the vertical axis in figure 11 to range from -0.4 to 0.7, instead of -.04 to 0.07, it would give the same impression of good fit to a straight line as figures 6 through 9.

Still it is true to say that the two estimates of VAR{m} are better correlated in the population of, say, Rural Flashers (figure 6, R=0.99) than in the population of Ontario Drivers (figure 11, R=0.63). Part of the reason is that the multivariate model for grade crossings is based on better information about "traits" than that for Ontario Drivers. Thus, e.g., train and car traffic volumes are used to estimate the unsafety of a grade crossing while for Ontario drivers no information about their exposure exists and, as a proxy, the sparse and noisy information about convictions had to be used.

Figure 10 and 11 are based on multivariate model A1 (see section 4b, table 1) which applies to the entire population of Ontario drivers. In this model, age and gender are accounted for by additive parameters while no distinction is made between drivers licensed to drive trucks or cars only. In figures 12 and 13 the correspondence of the two estimates of VAR{m} is shown for two additional multivariate models. One set of parameters has been estimated for the subpopulation of Ontario drivers who are male, 28 to 60 years old and licensed to drive cars only, another set of parameters for females with the same traits. Only accidents and convictions in 1981 and 1982 were used as independent variables for accidents in 1983 and 1984 as the dependent variable.

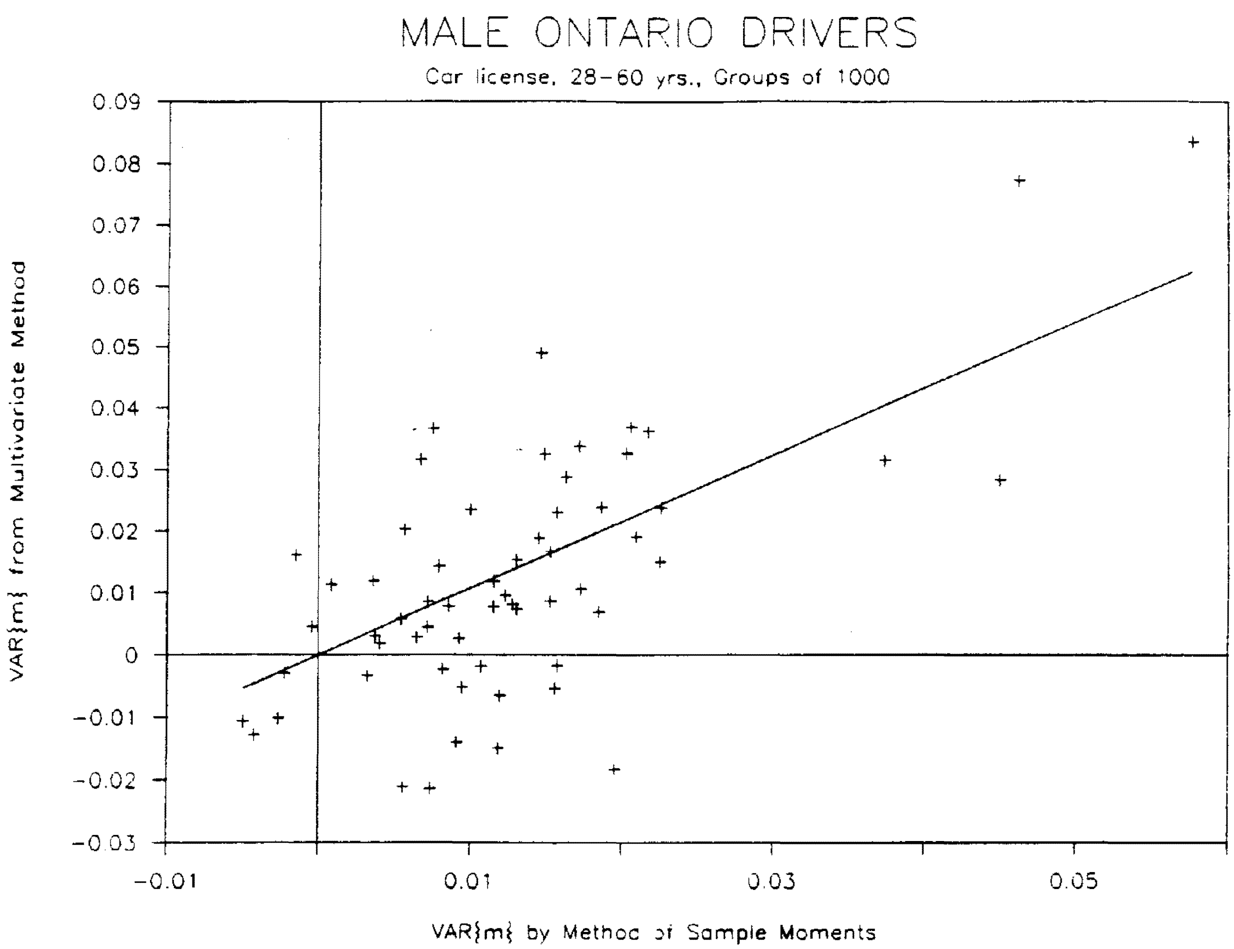

Figure 12. Comparison of VAR(m) for male Ontario drivers estimated by two methods.

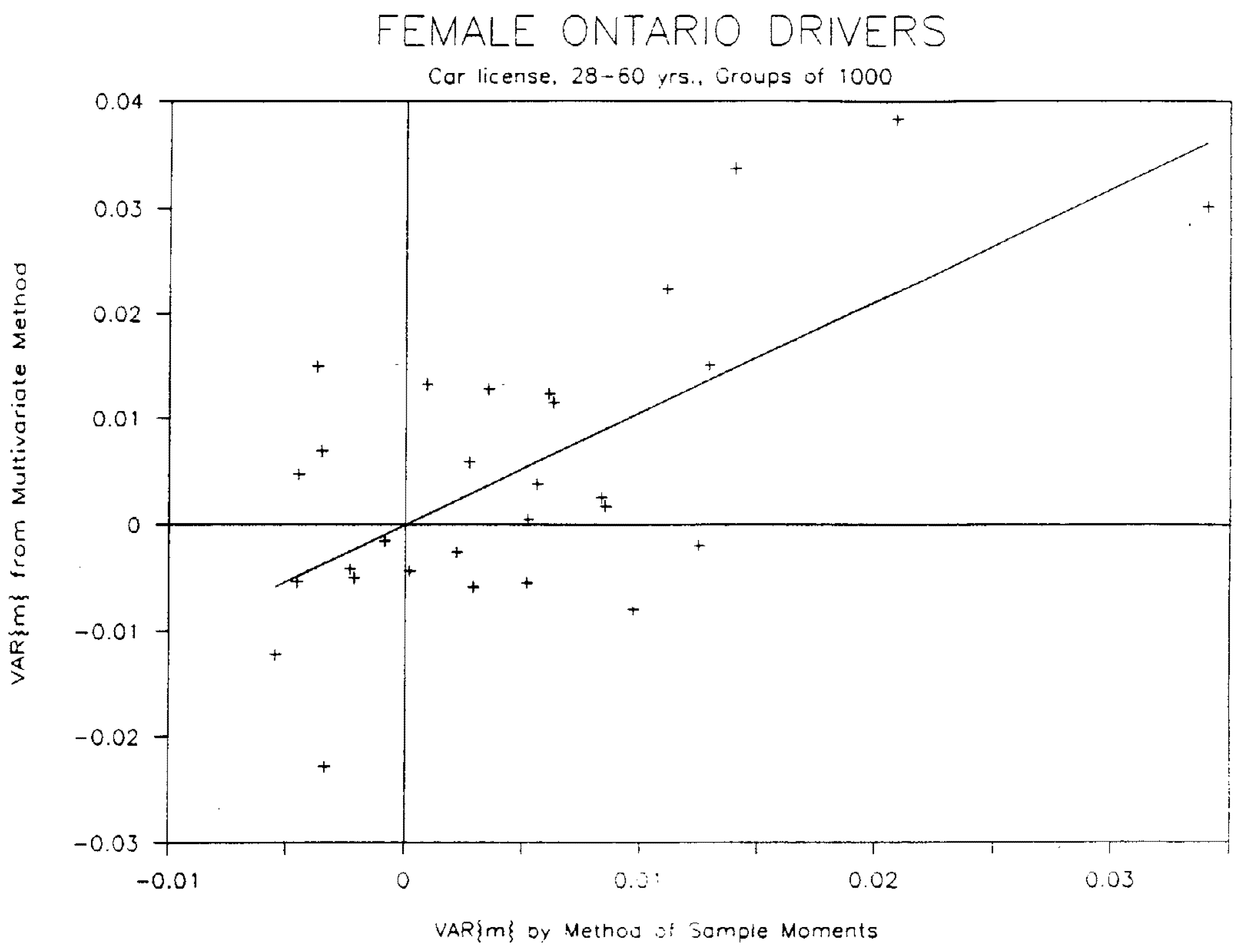

Figure 13. Comparison of VAR{m} for female Ontario drivers estimated two methods.

The slope of the regression line in figure 12 is 1.08 with a standard error of 0.11; in figure 13 it is 1.06 with a standard error of 0.19. In both cases R≈0.67. Thus, as far as one can tell, once again the two methods yield corresponding estimates of VAR{m}.

c. The Multivariate Method is Sound

The conceptual foundations of the Multivariate Method have been laid in section 3b. The idea is simple. A multivariate statistical model links the unsafety of entities to their traits. After the parameters are estimated, the mean of the m's in the imaginary reference population to which an entity belongs (E{m}) can be calculated. Consider an entity which serves as one datum for the estimation of model parameters. The squared residual of this datum point is an estimate of VAR{x}, the variance of its accident count around E{m}. Because for compound Poisson distributions $VAR\{x\}=VAR\{m\}+E\{m\}$, $\hat{VAR}\{m\}=\hat{VAR}\{x\}-\hat{E}\{m\}$.

The main question explored in this section was whether estimates of VAR{m} obtained by the Multivariate Method match the estimates obtained by the venerable Method of Sample Moments. Since both estimates are subject to the inaccuracies which go with estimation, the correspondence one may expect is less than perfect. With this in mind and in view of the empirical findings in sections 5a and 5b, it can be concluded without hesitation that for the data and models examined, the Multivariate Method is a sound way to estimate E{m} and VAR{m}. The question of the accuracy of each estimate has not been explored. However, there is an indication in tables 2, 3, 4 and 5 that the estimates produced by the Multivariate Method have a lesser variance. Thus, in the cases examined, the Multivariate Method seems to yield better estimates of E{m} and VAR{m} than the Method of Sample Moments.

6. APPLICATIONS

In what way then are the preceding results useful?

The interest in the estimation of unsafety stems from the need to perform two practical tasks. First, if one cannot say how unsafe something is, one will not know whether some intervention or remedial treatment should be considered. This is the "identification" task. Second, if one cannot say how much safer something has been made by the remedial treatment, one will continuously waste resources doing what is of little merit and waste life by failing to do what is more useful. This is the "estimating the effect of interventions" task. How the results of the preceding sections serve these twin tasks will be illustrated in the two subsections below. A sequence of illustrative examples will be used and their ramifications discussed.

a. Identification

To decide which entities need attention one usually asks: "what is the unsafety of this entity" and "by how much does it deviate from what is normal or from what could be attained".

By equation 1, the estimator of the unsafety m of an entity is $\alpha E\{m\}+(1-\alpha)x$ where x is the count of accidents and with $\alpha=E\{m\}/[E\{m\}+VAR\{m\}]$. Since $E\{m\}$ measures what is "normal" in the reference population, the difference between m and $E\{m\}$ is expressed by:

$$(1-\alpha)(x-E\{m\}) \qquad (9)$$

Thus, to say how unsafe an entity is and how it deviates from what is normal, one needs to have an estimate of α which, in turn, requires estimates of $E\{m\}$ and $VAR\{m\}$. It is at this point that the results of this inquiry are of use. Numerical example 5 is used to illustrate (numerical examples 1 to 4 are in sections 2 and 3).

Numerical Example 5.

A certain signalized intersection in Metropolitan Toronto has recorded in the last 3 years (on weekdays from 7:00 to 9:00 a.m.) 5 accidents of the type shown in figure 14. The average flows are F_1=450 vehicles/hour and F_2=120 vehicles per hour. What is the estimate of the expected number of such accidents for this intersection and how does it compare to what is "normal"?

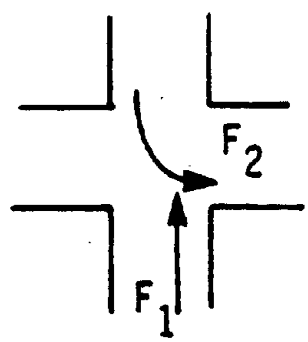

Figure 14. Accidents between vehicles of flows F_1 and F_2.

These are simple questions, and it is disconcerting to realize that a traffic engineer would find it difficult to give a satisfactory answer. Most would estimate m to be 5/3. But $\hat{m}$=5/3 is an inefficient estimate; it disregards all "clues of the first kind" such as: this is a four-legged signalized intersection in Toronto, the accidents are between a left-turning and a straight through vehicle, the corresponding flows are 450 and 120 vehicles per hour. To capture the information content of these clues, one has to have an estimate of E{m} for a reference population with these traits. However, in all Metropolitan Toronto there are but a few signalized intersections with very similar flows F_1 and F_2. Thus, no useful real reference population exists.

Because there is no useful real reference population, the traffic engineer would find it even more difficult to answer the second part of the question, namely, to say what is "normal" under these conditions. This must be a severe handicap for engineering practice. If it is not known what "normal" is, how can one judge what is "deviant"? The Multivariate Method helps to alleviate these difficulties as is shown below.

Numerical Example 5 Continued

A multivariate model has been fitted to data from 145 signalized intersections in Metropolitan Toronto.[9] For accidents of this type the model equation is $\hat{E}\{m\}=0.0283\times10^{-6}F_1\times F_2^{0.5163}$ accidents/hour which at this intersection makes 150.8×10^{-6} acc./hour. With 261 weekdays in a year, in 3 years there are 2×261×3=1566 hours in the morning peak. Thus, in 3 years, an average intersection with similar traits should record $150.8\times1566\times10^{-6}=0.236$ morning-peak accidents.

This gives an answer to the question what is normal. Therefore the count of five morning-peak accidents in three years can be seen in a new perspective. Thus, even though no real reference population exists it proved possible to say how many accidents of this type should be expected at an average intersection with these specific flows. To do that, a multivariate statistical model had to be fitted to data.

However, the original question of what is the expected number of such accidents at this intersection is still without answer. The essence of the Multivariate Method now comes into play. It provides a way to estimate VAR{m} for each datum point used in a multivariate statistical analysis and asserts that in similar circumstances VAR{m} can be represented by:

$$VAR\{m\}=[E\{m\}]^2/k \qquad (10)$$

This makes $\alpha=(1+E\{m\}/k)^{-1}$.

Numerical Example 5 Continued

When the multivariate model has been fitted to the Toronto intersections, it has been found that for this type of accident, $\hat{\kappa}$=1.39. Thus, $\hat{VAR}(m)=0.236^2/1.39=0.040$ [accidents in 3 years]^2. As a consequence, the estimate of α is $(1+0.236/1.39)^{-1}=0.855$. Accordingly the estimate of the expected number of such accidents at this intersection is $\hat{m}=0.855\times0.236+0.145\times5=0.202+0.725=0.927$ accidents in 3 years.

This completes the answer to the questions originally asked. The estimate of the expected number of morning-peak accidents in the 3-year period at this intersection was 0.927 while an average intersection with these flows should be expected to have 0.236 such accidents in the same period of time. This answer can be given because it has been shown in this report that:

a. Estimates of E{m} and VAR{m} obtained by the Multivariate Method are sound and

b. Such estimates of VAR{m} are usually related to E{m} as in equation 10.

It is useful to carry the numerical example a step further.

Numerical Example 5 Continued

By equation 2, the variance of the m's in the reference population of entities which recorded x accidents is estimated by $\alpha(1-\alpha)E\{m\}+(1-\alpha)^2x$. In the present case, $\hat{VAR}(m|5)=0.855\times0.145\times0.236+0.145^2\times5=0.029+0.105=0.134$ [accidents in 3 years]^2. Thus, m is estimated to be 0.927 accidents in 3 years with a standard error of $\sqrt{0.134}=0.366$. On the other hand, an average intersection of this kind is estimated to have 0.236 accidents in 3 years with a standard error of $\sqrt{0.04}=0.2$. In summary, the difference between the m of our intersection and what is "normal" is estimated to be 0.927-0.236=0.691 accidents in 3 years. If one could assume that the standard errors are independent (which they are not) the standard error of 0.691 would be $(0.134+0.040)^{1/2}=0.42$.

This is the kind of information needed to decide whether the m of the intersection under scrutiny deviates from what is normal and therefore, whether it should be given detailed attention. A more satisfactory way to deal with this question is suggested in reference 9 and is illustrated below.

Numerical Example 5 Continued

It has been established earlier that intersections of this kind have $\hat{E}\{m\}=0.236$ and $V\hat{A}R\{m\}=0.04$. If the m's are Gamma distributed, it can be found that 5% of such intersections have m's larger than 0.64 accidents in 3 years. These will be called "deviant".

The intersection under scrutiny is estimated to have $\hat{E}\{m|5\}=0.927$ and $V\hat{A}R\{m|5\}=0.134$. Using the Gamma probability distribution again, the probability that this intersection has an $m>0.64$ is 0.77. Thus, the probability that it is a deviant intersection is 77%, the probability that it is not deviant (a "false positive") is 23%.

In this subsection the following simple question has been posed: "What is the unsafety of a specific entity (intersection) and by how much does it deviate from what is 'normal'?" To provide an answer requires estimates of $E\{m\}$ and $VAR\{m\}$ for a reference population. However, in the numerical example, as in many other cases of practical interest, a sufficiently large reference population could not be found. The Multivariate Method has been used to provide an answer. For such an answer to inspire confidence, it had to be demonstrated that the Multivariate Method is valid. This was the main purpose of this report.

b. Estimation of Effect

The second circumstance in which estimates of the unsafety of an entity are needed is when it has been subjected to some intervention or remedial treatment and the effect on unsafety has to be determined. Here the principal task is to estimate what would have been the unsafety of this entity during the 'after' period had the treatment not been implemented. This task can be separated into two steps. First, to estimate what the unsafety was during the "before" period and next to estimate how it would have changed from the "before" period to the "after" period had treatment not been implemented.

Numerical Example 6.

> At the signalized intersection considered in numerical example 5, the left turn phase during the morning peak has been changed from "permissive" (solid green) to "protected" (green arrow). During 1 year after the change no accidents of the type shown in figure 14 were recorded while F_1 has changed from 450 to 500 vehicles per hour and F_2 has changed from 120 to 160 vehicles per hour. What is the estimated effect of the change in signal phasing on unsafety?

To answer, the first step is to estimate what the m was during the "before" period. But this is precisely what already has been done in numerical example 3.

> Thus, the m of this intersection during the "before" period is estimated to be 0.927 accidents in 3 years with a standard error of 0.366. (Or, equivalently, 0.309 accidents per year with a standard error of 0.122).

Before proceeding it is worth pointing out the essential unity of the "identification" and the "estimation of intervention effect" tasks. Both hinge on precisely the same act of unsafety estimation. At the peril of belaboring a point, a comment about the usefulness of the results of the preceding sections is in order. The estimation of unsafety requires that estimates of E{m} and VAR{m} for a reference population be available. Inasmuch as no adequate real reference population exists for the intersection under scrutiny, an alternative way to obtain estimates of E{m} and VAR{m} is from a multivariate statistical analysis. In the preceding sections of this report it has been shown that estimates so obtained are sound because, on the average, they are virtually identical to what has been be obtained by the Method of Sample Moments in many cases in which a real and large reference population did exist.

It follows that had the Multivariate Method not been developed and its soundness not established, the Empirical Bayes method could be applied to "identification" and to "estimation of effect" only in those cases in which real and large reference populations exist. This would restrict its scope. Since the Multivariate Method has been shown sound, the scope of application of the EB approach in road safety has been broadened to encompass those circumstances in which real and large reference populations are not available but a multivariate statistical model either already exists or can be fitted.

An additional figure of merit deserves mention. A real and large reference population exists only when the traits which characterize an entity are few. Thus, e.g., one could perhaps find a sufficiently large reference population for the accidents in figure 14 if one included in it all signalized intersection of Metropolitan Toronto irrespective of the flows F_1 and F_2. However, now one could rightly follow Elvik and ask whether estimates of E{m} and VAR{m} for such a reference population apply at an intersection with F_1=450 v/h and F_2=120 v/h.[6] The use of the Multivariate Method takes the sting out of this question.

The remainder of the numerical example is given for completeness.

Numerical Example 6 Continued

The estimate of m before the signal phasing was changed is now known to be 0.309 accidents per year with a standard error of 0.122. The next step is to estimate what would the unsafety be, had the signal phasing remained unchanged but the flows F_1 and F_2 changed from 450 v/h and 120 v/h to 500 v/h and 160 v/h respectively. If $\hat{E}\{m\}=0.0283\times10^{-6}F_1\times F_2^{0.5163}$ accidents/hour captures the manner in which the expected number of accidents depends on the two flows, the change in flows can be thought to increase m by a factor $(500/450)\times(160/120)^{0.5163}=1.288$.

Thus, had signal phasing remained unchanged, during 1 year of the "after" period one should expect 0.309×1.288=0.398 a.m. peak accidents/year, the standard error of which is ±0.122×1.288=±0.157. It is to this that the "0 after accidents" has to be compared. Of course, one would not attempt to reach any conclusions on the basis of this single intersection. However, if results of many such comparisons are pooled, defensible conclusions will emerge.

The original applications of the EB method in road safety had to do with the need to cleanse "regression to the mean" out of estimates of intervention effectiveness. However, these early efforts were restricted to cases in which large reference populations could be found. Even when a large reference population could be found, the collection of data was often laborious. Thus, e.g., in the work about conversion from two-way to four-way stop control in San Francisco it was necessary to unearth old accident counts in the basement of the City Hall.[3] These counts for more than 1000 two-way stop controlled intersections and for several years had to be manually coded and only then could the estimation of E{m} and VAR{m}begin. Even then, once the process was complete, $\hat{E}\{m\}$ and $\hat{VAR}\{m\}$ could not be tailored to the specific traffic volumes of the 49 intersections which were converted from two-way to four-way stop control in San Francisco.

If faced with a similar task now it would not be necessary to be confined to those instances of conversion from two-way to four-way stop control for which a large reference population could be identified. In addition, it would be possible to find estimates of E{m} and VAR{m} which are more closely tailored to the traits of the intersections at which such conversions have been implemented. Still, it should be noted, it would be necessary to specify and estimate the parameters of a multivariate statistical model for two-way stop controlled intersections or have the results of such an activity available.

c. Extensions

The Multivariate Method extends the scope of the Empirical Bayes method for the estimation of unsafety to those cases for which a large reference population does not exist. In addition the estimates of unsafety obtained by the Multivariate Method are more closely tailored to the traits of the entity the unsafety of which is of interest. However, the implementation of the Multivariate Method within an EB framework presupposes the existence of a multivariate statistical model for E{m} and VAR{m}.

To properly specify a multivariate statistical model and to estimate its parameters requires data, expertise and work. The hope is that once such a model is built, it applies to a broad segment of reality. Thus, e.g., once such a model for U.S. grade crossings has been built, the results can be applied to the tasks of "identification" and "estimation of intervention effect" for any set of grade crossings in the U.S. In this case, the EB framework can be used without any additional data collection or analysis.

The possibility of using preexisting multivariate models as the source of $\hat{E}\{m\}$ and $\hat{VAR}\{m\}$ is rich in promise and heavy in responsibility. The promise is in enabling the practitioner to engage in safety analyses without having to make every case into an extensive data collection effort and research project. The responsibility is in the need to build multivariate models with broad applicability. Thus, e.g., one hopes that the model which estimates E{m} and k for a grade crossing in Florida is also good for a crossing in Michigan (perhaps up to a multiplicative correction factor). This hope rests on the belief that if there is some regularity in the relationship between accident occurrence and the measured traits of grade crossings, the regularity is common to both States. If this be so and the multivariate model has captured the regularity, it will have broad applicability. Whether the same regularity holds in different States is an empirical question which requires empirical exploration.

Over the years many multivariate models have been built and the results published. Thus, e.g., references 8, 16, 17 and 18 estimate the unsafety of rail-highway grade crossings as a function of vehicle and train flows, surrounding area, warning device used and geometry. References 9, 19, 20 and 21 give multivariate models for the unsafety of intersections. A review of seven multivariate models of highway accidents on road sections may be found in reference 22, published in 1983. Other studies have been published since then.[23,24] Many multivariate models to estimate the unsafety of drivers have been built (see e.g., references 15, 25, 26).

While many multitude of multivariate models have been built in the past, they were all concerned with the estimation E(m) only. To use their results to estimate unsafety by EB methods, estimates of VAR(m) or of k are also needed. However, since the need to have estimates of VAR(m) or k was not evident in the past and the technique for their estimation were not developed, only few estimates of k now exist.

For Ontario drivers estimates of k are between 1.9 and 2.4.[15] For grade crossings the estimates of k were found to be between 0.49 to 0.74.[8] For signalized intersections estimates of k around 2 were obtained.[9] Using the data of reference 24 estimates of values of k for rural two lane roads are between 2 and 3.

It should be remembered that k is a function of the traits which are used in a multivariate model. Therefore multivariate models which differ in the traits they use as independent variables, will have different values for k. Still, k is not just constant without meaning; it is a measure of how wide is the distribution of m's in a reference population defined by specific traits. Because k represents and measures something real one can rightly expect that as more estimates of k will become available, it will become possible to judge what the k of a family of models is. The hope is that models which apply to the same kinds of entities (say, drivers or signalized intersections) and use the same independent variables (say, past accidents and violations or traffic flow and geometry) will have similar k's. If this will prove to be the case, the practitioner will be in a position to estimate k for some reference populations from a catalogue of accumulated research experience.

7. SUMMARY

If one can not estimate how unsafe something is, how much safer it has been made and how much safer it could be made, one can not claim to be managing road safety. Earlier work has shown that the Empirical Bayes frame of thought improves the ability to estimate how unsafe an entity is and how much safer an intervention has made it. However, to apply the EB method was at times difficult. Application required data for large reference populations which are often not easy to find and in many cases do not exist. In addition, questions were raised about what makes reference population I better suited for estimating the safety of a specific entity than reference population J. Without a satisfactory answer, the EB method was open to arbitrariness. This report is an attempt to make the EB method applicable to a larger set of circumstances, and at the same time take the sting out of the question of how to choose a suitable reference population.

The EB method is founded on the claim that clues to the unsafety of an entity are contained not only in its accident history but also in its traits (age, gender, traffic volume, geometry etc.). Therefore, when the unsafety of an entity is estimated, both clues should be used. To make use of those clues to unsafety which are derived from traits, an estimate of the mean of the m's and their variance in a population of entities with identical traits (the reference population) is needed.

The concept of reference population, while central in importance, is fraught with difficulties. What is eventually select as a reference population for some entity will depend on what data happens to be available, on what traits are though to be important and on other circumstances or judgements which are to some extent arbitrary. However, it is shown that precisely the same kind of arbitrariness plagues unsafety estimation which uses only the historical record of accidents as a clue. Therefore, it is concluded that the process of unsafety estimation is not akin to the measurement of distance or pressure; it will always depend on what data is available, what is thought to be relevant and on other manifestations of judgement. It follows that clues to unsafety which are based on traits are just as legitimate as the clues which are based on accident history.

How to estimate unsafety using both kinds of clues is discussed in section 2. In essence, if x is the count of accidents on the entity which belongs to a reference population with E{m} and VAR{m}, the m of this entity is estimated by $\alpha E\{m\}+(1-\alpha)x$ where $\alpha=E\{m\}/[E\{m\}+VAR\{m\}]$. The variance of the estimate of m is then $\alpha(1-\alpha)E\{m\}+(1-\alpha)^2x$.

Thus, to use the EB method for unsafety estimation one has to have estimates of E{m} and VAR{m}. In earlier work the venerable Method of Sample Moments (section 3a) has been used for this purpose. This meant that the EB method could be implemented only when a sufficiently large reference population could be found. It also meant that the estimates of E{m} and VAR{m} were for reference populations which could not closely match the traits of the entity the unsafety of which is of interest. Furthermore, since the traits of the entity and the reference population did not match, one could often find several alternative reference populations, each with its own E{m} and VAR{m}. Which of these should be chosen was not clear.

The thinking behind the Multivariate Method is explained in section 3b. It is offered as an alternative to the Method of Sample Moments and comes to alleviate some of its shortcomings. An imaginary reference population is defined to be one which matches all measured traits of the entity under scrutiny. It is shown how estimates of E{m} and VAR{m} for an imaginary reference population can be obtained from a multivariate statistical analysis.

The main idea is this. The accident record (dependent variable) and traits (independent variables) of many entities are the data from which a multivariate statistical model is built. First a functional form is chosen, then parameters are estimated. With this done, for each entity which served as datum the model equation is used to estimate its E{m}. The E{m} is the mean of the m's in the imaginary reference population to which this "datum" entity belongs. The squared residual for this entity is an estimate of VAR{x}, the variance of its accident count around E{m}. Because for compound Poisson distributions VAR{x}=VAR{m}+E{m}, the estimate of VAR{m} from one "datum" entity is $\hat{VAR}\{m\}=\hat{VAR}\{x\}-\hat{E}\{m\}$. These estimates of VAR{m} for each datum entity can than be subjected to a further multivariate statistical analysis intended to represent the relationship between VAR{m} and the measured traits. The need for such further multivariate analysis can often be avoided. It appears that in many cases the points [$\hat{VAR}\{m\}$, $\hat{E}\{m\}$] fit the simple relationship $VAR\{m\}=[E\{m\}]^2/k$. Thus, only an estimate of the single parameter k needs to be found.

To check the soundness of the Multivariate Method, the correspondence between the estimates obtained by the Method of Sample Moments and those obtained from the same data by the Multivariate Method has been examined. The juxtaposition of the estimates by the two methods requires that data sets be available which are both very large (so that the Method of Sample Moments can be used) and to which multivariate models have been fitted (so the estimate be the Multivariate Method could be produced). Two such data sets have been used. One is for 200,000 rail-highway grade crossings in the U.S., the other for some 170,000 Ontario drivers. These data sets, the multivariate models fitted to them and the manner in which they are used in this report are described in section 4.

In section 5 it is shown that the estimates of E{m} and VAR{m} obtained by the Multivariate Method match very closely those which are obtained by the Method of Sample Moments. This correspondence holds irrespective of whether groups of entities are formed at random or whether the groups are made homogeneous in terms of their m. The correspondence holds for each data set as a whole and also for many specific subsets. Furthermore, there is an indication that the Multivariate Method produces estimates with a lesser variance.

The Multivariate Method gives estimates of E{m} and VAR{m} for an imaginary reference population the defining traits of which are identical to the measured traits of the entity about the unsafety of which we inquire. In contrast, estimates by the Method of Sample Moments reflect the E{m} and VAR{m} of a reference population which is diverse in its traits. In consequence, the estimate of VAR{m} by the Method of Sample Moments can be substantially larger than the estimate obtained by the Multivariate Method. In short, the Multivariate Method yields sound estimates of E{m} and VAR{m} and these are more accurate and more appropriate than the estimates produced by the Method of Sample Moments.

Based on these results the use of the EB approach to the tasks of "identification" and "estimation of effect" is illustrated in section 6 using numerical examples. Thus, given the accident history and vehicular flows of a signalized intersection the question is asked what is its expected number of accidents and how it corresponds to what is normal at such intersections. The practitioner would find it difficult to answer both questions. The question about the expected number of accidents would be answered inefficiently, using only the accident history for this purpose and disregarding all else that is known about signalized intersections with such flows. The question about what is normal for such sites would most likely get no response because not many intersections in the practitioner's jurisdiction will have similar flows and traits. Thus, the source of the problem usually is that no real large reference population exists. It is shown how to answer both questions within the Empirical Bayes frame of reference. For this purpose a multivariate model for signalized intersections is used.

To illustrate how the EB approach applies to the "estimation of effect" task, the same story line is continued. Some aspect of signal phasing is said to have been changed, and record of accidents following this change is given. The task now is to estimate the effect of the change in signal phasing on unsafety. It becomes evident, that the tasks of "identification" and "estimation of effect" both hinge on the same act of unsafety estimation. Since to estimate unsafety estimates of E{m} and VAR{m} are needed and the Multivariate Method is the only practical way to obtain them in this circumstance, a multivariate statistical model for signalized intersections has been used. This illustrates how the obstacle of not having a real large reference population has been overcome.

In summary, the Multivariate Method for the estimation of E{m} and VAR{m} is sound. It extends the applicability of the Empirical Bayes method of unsafety estimation to circumstances in which large reference populations do not exist while data is sufficient to build a multivariate statistical model or when such a model has already been built. Estimates of E{m} and VAR{m} obtained by the Multivariate Method exactly match the traits of the entity of interest. In this manner, the Multivariate Method obviates the principal difficulties of the Method of Sample Moments. Once a multivariate statistical model has been built and is shown to be of wide applicability, the Empirical Bayes approach to unsafety estimation can be implemented without the need to collect and analyze data for the purpose of estimating E{m} and VAR{m}.

REFERENCES

(1) C. Abbess, D. Jarrett, and C.C. Wright, C.C., Accidents at blackspots: estimating the effectiveness of remedial treatment, with special reference to the "regression-to-the-mean" effect. Traffic Engineering and Control, Vol. 22, No. 10, pp.535-542, 1981.

(2) E. Hauer, On the estimation of the expected number of accidents. Accid. Anal. & Prev., Vol. 18, pp. 1-12. 1986.

(3) Hauer, E. and Lovell, J., New directions for learning about the safety effect of measures, Transportation Research Record 1068, pp. 96-102, 1986.

(4) Hauer, E., Lovell, J. and Persaud, B.N., New directions for learning about safety effectiveness, Report No. FHWA/RD-86/015, Federal Highway Administration, 1986.

(5) C.C. Wright, C.R. Abbess, and D.F. Jarrett, Estimating the regression to mean effect associated with road accident Black Spot treatment. Towards a more realistic approach. Accid. Anal. & Prev., Vol. 20, No. 3, pp.199-214, 1988.

(6) R. Elvik, R., Some difficulties in defining populations of "entities" for estimating the expected number of accidents. Accid. Anal.& Prev., Vol. 20, No. 4, pp.261-276, 1988.

(7) L. Mountain, and B. Fawaz, The area-wide effects of engineering measures on road accident occurrence, Traffic Engineering and Control, Vol. 30, No. 7/8, pp. 355-360, July/August 1989.

(8) E. Hauer and B.N. Persaud, How to estimate the safety of rail-highway grade crossings and the safety effect of warning devices. Transportation Research Record 1114, pp. 131-140, 1987.

(9) E. Hauer, J.C.N Ng and J. Lovell, Estimation of safety at signalized intersections. Transportation Research Record 1185, 1989.

(10) E. Hauer and B. Persaud, Theoretical Foundations of a Rational Procedure for Blackspot Identification and its Application, Proceedings of the 11th Annual Summer Meeting, PTRC, July 1983, pp. 231-244.

(11) E. Hauer, On the Estimation of the Expected Number of Accidents", Accident Analysis and Prevention, Vol.18, No.1, 1986, pp. 1-12.

(12) C.N. Morris, O.J. Pendelton, M.K. Bishop and C.L. Scaff, Empirical Bayes methodology in traffic accident analyses. Paper presented at the American Statistical Association Winter Conference, 1989.

(13) H. Robbins, An empirical Bayes estimation problem, Proc. Natl. Acad. Sci., USA, Vol 77, No. 12, pp.6988-6989, 1980.

(14) E. Hauer, The multivariate method for "unsafety" estimation, Department of Civil Engineering, University of Toronto, Report No. 89-09, ISBN-0-7727-7556-7, Sept. 1989.

(15) E. Hauer, A. Smiley A. and Persaud B., Accidents, Convictions and Demerit Points: An Ontario Driver Study -"Summary Report", Ministry of Transportation Ontario, Report 491-01, 1988.

(16) J. Coleman, and G.R. Stewart, Investigation of accident data for railroad highway grade crossings, Transportation Research Record 611, National Research Council, Washington, pp. 60-67, 1976.

(17) P. Mengert, Rail-highway crossing hazard prediction results, Report FRA-RRS-80-02, PB 80 170749, Transportation Systems Center, U.S. DOT, 1980.

3267
11

(18) G. van Belle, D. Meeter, and W. Farr, W., Influencing factors for railroad-highway grade crossing accidents in Florida, Accid. Anal. & Prev. Vol. 7, pp. 103-112, 1975.

(19) M.Y.K. Lau and A.D. May, Injury accident prediction models at signalized intersections, Transportation Research Record 1172, 1989.

(20) R.D. Hall, Accidents at four-arm single carriageway urban traffic signals, Transport and Road Research Laboratory, Contractor Report 65, 1986.

(21) D. Pickering, R.D. Hall, and M. Grimmer, M, Accidents at rural T-junctions, TRRL Report 65, Transport and Road Research Laboratory, 1986.

(22) F. Luanda, R.W. Smith, M. Padron, P. Resto, J. Guttierez and L. Fernandez, L., Multivariate statistical analysis of highway accident and highway conditions, Report DOT-RSPA-DMA-50/84/9, U.S. Department of Transportation, 1983.

(23) R. Hamerslag, J.P. Roos, M. Kwakernaak, Analysis of accidents in traffic situations by means of multiproportional weighed Poisson model, Transportation Research Record 847, pp.29-36, 1987.

(24) C.V. Zegeer, J. Hummer, D. Reinfurt, L. Herf, and W. Hunter, Safety effects of cross-section design for two-lane roads. Final Report. FHWA/RD-87/008. Federal Highway Administration, 1987.

(25) R.S. Coppin, and R.C. Peck, The California driver Record study, Parts 1-9, Sacramento, Department of Motor Vehicles, 1965.

(26) R.C. Peck, and J. Kuan, A statistical model of individual accident risk using driver record, territory and other biographical factors. Accid. Anal. and Prev. Vol. 15 No. 5, 371-393, 1983.