On the Use of Accident or Conviction Counts To Trigger Action

E. HAUER, K. QUAYE, AND Z. LIU

The probabilistic properties of the process of identifying entities, such as drivers or intersections, for some form of remedial action when they experience N accidents within D units of time are explored. This mechanism for triggering action is referred to as an N-D trigger. On the basis of the probability distribution of the "time-to-trigger," it is concluded that in road safety the problem of false positives is severe, and therefore entities identified on the basis of accident or conviction counts should be subjected to further safety diagnosis. Moreover, the longer the N-D trigger is applied to a population, the less useful it becomes. The performance of the trigger depends on the choice of N and D, and guidance is offered on how best to choose them.

When at least five correctable accidents occur at an intersection within 12 months, the accident experience warrant for a traffic signal is satisfied (*1*, 4C-6). The same kind of warrant pertains to multiway stop signs (*1*, 2B-4). Similarly, it is common practice to flag a road section or intersection for selective enforcement, engineering study, or remedial action if the count of accidents occurring on it per unit of time or exposure exceeds a certain number. Also, a driver's license is suspended when the driver has accumulated and exceeded a certain number of demerit points, with points being erased from the driver's record a fixed time after conviction. All these situations have the following common conceptual structure:

> There exists a group of "entities," be they intersections, road sections, or drivers. On each entity occur events such as accidents or convictions. The process of event occurrence is characterized by a degree of randomness. If on an entity N or more events occur within a time (or distance or exposure) "window" that is of duration (or length or size) D, some action is triggered. The action may be the performance of an engineering study, increased enforcement, implementation of remedial measures, revocation of driving privileges, or the like. This method of triggering action will be referred to as the N-D trigger.

Because of the random nature of event occurrence, the N-D trigger has certain probabilistic properties that are of practical interest. For example, on occasion a flurry of accidents may occur at safer-than-average intersections. Whereas the accident warrant is thereby met and thus the installation of a higher-level traffic control device may be considered, in reality this is a false alarm. How often is the accident warrant met in this spurious manner? How does the frequency of false alarms depend on the choice of N and D? The same questions arise when remedial work is applied to falsely identified blackspots or when merely unlucky drivers lose their driving privileges.

The probabilistic properties of the N-D trigger have been explored earlier (*2*). Unfortunately, an error in one of the key equations makes all those results numerically incorrect. The purpose of this paper is to correct this error and to present results of practical interest to a broader audience.

NOTATION AND ANALYSIS

Consider an entity on which "events" occur in accord with the Poisson probability law with a rate m per unit of time (or distance). When or where these events occur could be noted on a time or distance axis. Imagine a "window" of size D sliding along this axis. The time or distance from the origin until N events show in a window of size D for the first time is a random variable T. Thus T may denote the time when at an intersection a signal is warranted for the first time, it may designate the end of a road section that seems to be a blackspot, or T may stand for the instant when a driver has accumulated sufficient demerit points to warrant the suspension of the driver's license. The random variable T is the "time to trigger"; specific values of T will be denoted by t. The probability distribution $F_T(t)$ or the probability density function $F'_T(t) \equiv f_T(t)$ is needed to answer questions of practical interest.

It has been shown elsewhere (*2*) (a summary is given later) that in general,

$$F_T(t) = 1 - Ce^{-m\int_0^t p(\tau)d\tau} \tag{1}$$

where C is a constant of integration and $p(\tau)$ is the probability that there are $N - 1$ events in a window of size D when its right edge is on τ [the error mentioned earlier was to assume that $p(\tau)$ was independent of τ]. To find particular solutions and to determine what the constant of integration C is, initial conditions need to be specified. Two specific initial conditions are examined below.

Consider first an entity that has been in existence for some time. At time $\tau = 0$ when examination begins, fewer than N events are in the window extending from $-D$ to 0. For this, the "surviving entity case,"

$$F_T(t) = 1 - Ce^{-m\int_0^t p(\tau)d\tau} = 1 - e^{-m\bar{p}(t)t} \tag{2}$$

where $\bar{p}(t)$ is the mean value of $p(\tau)$ in $[0,t]$.

Safety Studies Group, Department of Civil Engineering, University of Toronto, Toronto, Ontario M5S 1A4, Canada.

A good approximation to $F_T(t)$ is obtained when $\bar{p}(t)$ is replaced by the constant

$$\bar{p} = \frac{\dfrac{(mD)^{N-1}}{(N-1)!}}{\sum_{0}^{N-1} \dfrac{(\mathrm{mD})^i}{i!}} \times k \quad (3)$$

where

$$k = (0.0009 \times 0.5941^{-N} + 0.8754) + mD(0.0917 \times 1.3256^{-N} - 0.1273) + (mD)^2(0.0339 \times 1.4009^{-N} + 0.0073)$$

This approximation was developed by Liu (*3*) for values of N from 3 to 8 and various values of mD subject to the condition that $0 \leq \bar{p} \leq 1$. Typical values of $\bar{p}$ are shown in Figure 1.

Consider next an intersection or road that has been newly opened to traffic or a driver just licensed. We will call this the "new entity" case. Here,

$$F_T(t) = \begin{cases} 1 - \sum_{0}^{N-1} \dfrac{e^{-mt}(mt)^i}{i!} & \text{when } t \leq D \\ 1 - [1 - F_T(D)]\, e^{-\bar{p}(t)m(t-D)} & \text{when } t > D \end{cases} \quad (4)$$

where $\bar{p}(t)$ is the mean value of $p(\tau)$ in $[D,t]$. A good approximation is obtained when $\bar{p}(t)$ is replaced by $\bar{p}$. The nature of $F(t)$ is best illustrated by a numerical example.

Numerical Example 1: Consider an intersection, presently equipped with stop signs on the minor approaches, that has two correctable accidents per year as a long-term average ($m = 2$). Conversion to multiway stop control is warranted when five correctable accidents occur in a 12-month period (1, 2B-6). Thus $N = 5$ accidents and $D = 1$ year. The $F(t)$ curves corresponding to Equations 2 and 4 are shown in Figure 2.

Thus, if such an intersection was newly opened (or equipped), the probability that multiway stops will be warranted within 1 year is 5 percent (Point A); fully 73 percent of two-way stop-controlled intersections that in reality have the acceptable average of two correctable accidents per year will meet the accident warrant in the first 10 years of operation (Point B). If the intersection was in existence at time 0, the corresponding probabilities are 13 and 75 percent as shown by Points C and D.

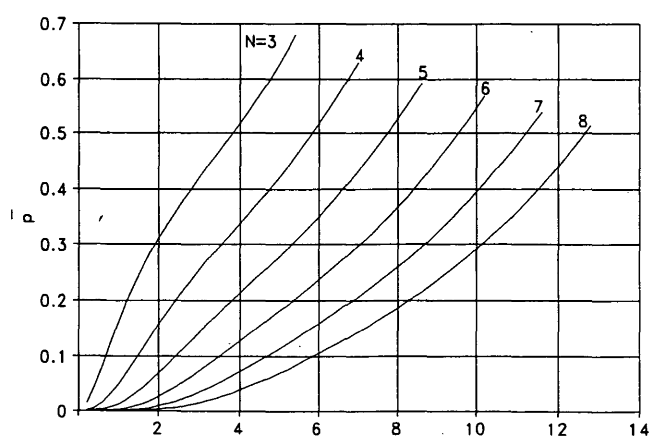

FIGURE 1 Values of $\bar{p}$ as a function of *mD* and *N*.

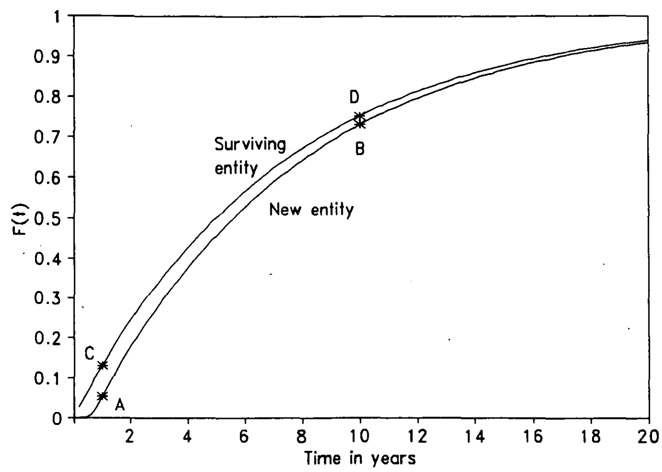

FIGURE 2 The probability distribution of the "time-to-trigger" when *mD* = 2 and *N* = 5.

A MICROCOSM

Having obtained the probability distribution of T, the functioning of the *N-D* trigger can be explored. It is planted into a simplified microcosm of professional practices that surround its use, and in this manner the workings of the trigger are illustrated.

Application to a Population

In Numerical Example 1 the m of the intersection was given. In reality, the *N-D* trigger is applied to a population of entities in which each entity has its own distinctive but unknown m. The purpose of the next example is to examine how an *N-D* trigger would perform if applied to such a population of entities.

Numerical Example 2: Consider a population of 1,000 intersections, all serving similar traffic flows. Assume that 900 have $m = 1$ accident per year, 90 have $m = 2$ accidents per year, and 10 have $m = 3$ accidents per year. It is not known which intersections have what m. Let $N = 5$ accidents and $D = 1$ year, as the MUTCD (*1*) recommends. At time = 0 no intersection had five or more accidents in the previous year. Table 1 indicates what is expected to happen by the end of the first year.

The entry in Column 5 is the product of the number of intersections in the population (Column 1) and the probability of an intersection to record five or more accidents before the end of the first year (Column 4). The hope was that the warrant will help to identify mainly the 10 "deviant" intersections that have an unusually high m. It turns out that in this particular population, the warrant may be expected to

TABLE 1 Performance of the 5–1 Trigger in the First Year

1	2	3	4	5
Number of Intersections	m (accidents/ year	$\bar{p}$ eqn. 3	$F_T(1)$ eqn. 2	Number expected to meet warrant in first year
900	1.0	0.0122	0.0121	11.0
90	2.0	0.0697	0.1301	11.7
10	3.0	0.1433	0.3497	3.5
1000				26.2

identify during the first year nearly one-third of these. Thus, the number of "correct positives" in the first year is expected to be 3.5. This leaves the remaining 6.5 deviants unidentified in the first year—these are the "false negatives." Of the 26.2 intersections at which the warrant is expected to be met, 11 are in fact safer than average and do not deserve attention or treatment—these may be called the "false positives."

The main merit of this illustration is in showing how false positives and false negatives arise when an *N-D* trigger is applied to some real population. By definition, "deviant" is what constitutes a small minority, whereas "normal" is always the large majority. Even though only 1.2 percent of the 900 normal intersections will spuriously meet the warrant in the first year, because they are many, the number caught by the 5-1 trigger (MUTCD warrant) is still substantial. The upshot is that of the 26.2 intersections at which the warrant is expected to be met, we really wanted to catch only 3.5.

Whether a clever choice of N and D can improve the performance of the warrant will be examined later. At this point we merely note that the problem of false positives and false negatives is inherent in the *N-D* trigger. Furthermore, because normal entities must be many, the number of false positives is bound to be large.

Depletion

The discussion in Numerical Example 2 pertains to the use of the *N-D* trigger in a population of entities for 1 year. In practice, however, demerit point systems or blackspot identification procedures are being applied continuously. It is therefore important to examine the performance of the trigger over time.

Numerical Example 3: In Numerical Example 2, of the 900 better-than-average intersections with $m = 1$ accidents per year, $900 - 11.0 = 889.0$ are expected to survive 1 year. Of these, $889.0 \times 0.0122 = 10.8$ are expected to meet the warrant during the second year, and so on. Table 2 gives the number of intersections identified for inspection by the 5-1 trigger year after year.

The number of safe intersections (those with $m = 1$) expected to meet the warrant in consecutive years is 11.0, 10.8, 10.7, 10.6, 10.4, and so on. This group is a steady source of false positives for a long time. Of the 10 deviant intersections with $m = 3$, the number expected to meet the warrant in consecutive years is 3.5, 2.3, 1.5, 1.0, 0.6, and so on. Thus, in the course of a few years most of the truly deviant intersections would have been identified, and continued application of the *N-D* warrant becomes useless. Whereas in the first year 13 percent of the identified sites are expected to have $m = 3$, in Year 5 this declines to only 3 percent.

These are the main features of the process that arises when an *N-D* warrant is applied to a population of entities the m's of which change slowly. First, since normal entities are by definition many, this group supplies false positives at a nearly constant rate for a long time. Second, because deviant entities are by definition few and the *N-D* warrant is relatively good at catching them, their supply in the population is depleted in short order. Therefore, continued application of the warrant is bound to be unproductive.

Two practical conclusions can now be articulated. First, because an entity identified by the *N-D* warrant has a good chance of being a false positive, it is essential that it be subjected to a detailed diagnostic examination before any costly action is taken. To the extent that defensible procedures for such diagnostic examination are not yet part of the professional repertoire, they need to be developed.

Second, continued application of an *N-D* trigger to a population of entities that remains essentially unchanged is of dubious merit. It gives rise to a process by which some entities are always caught in the net, but ever fewer of these are of real interest. Thus while action (more traffic control, treatment of additional blackspots, revocation of licenses) continues apace, the entities that are its subject approximate an almost random selection from the general population. It seems that if the aim is to identify entities with unusually high m, the *N-D* trigger in the form of accident warrants or demerit points should be applied only to new or changed entities and always for a limited period of time.

Regression to the Mean

When remedial action is based on the *N-D* trigger, there is the danger of illusory success due to "regression to the mean."

TABLE 2 Performance of the 5–1 Trigger in the Course of 5 Years

m accidents/year	Expected number of entities identified				
	year 1	year 2	year 3	year 4	year 5
1	11.0	10.8	10.7	10.6	10.4
2	11.7	10.2	8.9	7.7	6.7
3	3.5	2.3	1.5	1.0	0.6
	26.2	23.3	21.1	19.3	17.8

This should by now be well known [even though reports ignoring it are still being published (*4,5*)]. The earlier numerical examples provide the setting for a particularly clear illustration of how this bias comes about.

Numerical Example 4: Imagine that traffic control has indeed been upgraded at the 26 intersections where the 5-1 warrant was met (see Table 1). At some later time, the local authority conducts a naive before-and-after study to assess the effect of this upgrading on safety. Assuming that the change in traffic control did not affect the m's and thus did not change the safety of these intersections, what are the conclusions of such a study likely to be?

Since $N = 5$, each converted intersection has had five correctable accidents in the year before conversion. If the m's of these sites did not change, reading from Table 1, we should expect to find after conversion $(1 \times 11.0 + 2 \times 11.7 + 3 \times 3.5)/26.2 = 1.71$ correctable accidents per intersection in a year. Thus, whereas there was no change in the m's, a simple before-and-after study is expected to indicate an impressive reduction from 5 to 1.71 accidents per intersection. This "improvement" is usually illusory because, by assumption, the long-term average number of accidents has not changed.

Thus the illusion of success due to "regression to the mean" in naive before-after studies is inherent in the automatic application of accident warrants that are of the *N-D* type. The hope is that accident warrants are not applied automatically. However, one application that is decidedly automatic occurs when demerit points are given to drivers who are convicted for violations of the highway traffic act. Typically, a set of actions (warning letter, interview, mandatory retraining, license suspension) is automatically triggered at some preset count of demerit points. This setting is explored next.

How Long Before You Lose Your License?

Most "good drivers" think it grossly unfair that one's driving privileges should be removed because of a run of bad luck. For a system of demerit points to be acceptable, the large majority of drivers whose licenses are revoked should have a genuinely high rate of convictions for the violation of traffic laws. Unfortunately, as will be shown, to devise such a system on the basis of convictions is nearly impossible. Just as in the case of intersection accidents, most drivers of a population who reach a predetermined count of convictions are false positives.

Rather than treading over old ground, by calculating the number of false positives for a typical population, the issue will be illumined from a different angle. The chosen tool of inquiry will be the "mean-time-to-trigger."

Consider the $F_T(t)$ for the surviving entity case (Equation 2). Integrating $tF'_T(t)dt$ from zero to infinity, it can be shown that the mean-time-to-trigger is

$$E\{T\} = \frac{1}{\bar{p}m} \qquad (5)$$

This result is sufficient for a numerical example.

Numerical Example 5: In Ontario a driver's license is suspended when a person accumulates 15 demerit points. The number of demerit points depends on the type of offense and varies from 2 to 7. Demerit points are stricken from a person's record after 2 years ($D = 2$). With some 5 million drivers and 700,000 pointable convictions per year, the average Ontario driver receives about 0.14 pointable convictions per year. If the average number of points per conviction is 4, it takes about 4 pointable convictions in a 2-year window to reach suspension level.

Consider a better-than-average driver with $m = 0.1$ convictions per year. For this person, $\bar{p} = 0.000943$ (as calculated with $m = 0.1$, $N = 4$, and $D = 2$). For such drivers $E\{T\}$ is about 10,600 years. During a 50-year driving career, 1 of 212 such better-than-average drivers will have the license suspended.

Consider now another driver with a long-term conviction rate 5 times the population average ($m = 0.14 \times 5 = 0.7$ convictions per year). For this person, $E\{T\} = 15$ years. Thus, drivers of this kind will drive, on the average, 15 years before their license is suspended for the first time.

This kind of procedure may not be thought very satisfactory. First, there are millions of "better-than-average" drivers in Ontario. One in 212 such drivers will have the license suspended at some time. This seems unfair. Second, drivers who truly have an unusually large conviction rate may drive for a very long time before detection. The question arises whether it is possible to choose values of N and D to improve the performance of the trigger.

This question is best examined in light of the graphs in Figure 3, which is based on Equation 4. Point A pertains to the better-than-average drivers whose $m = 0.1$ convictions per year. With $N = 4$ convictions and $D = 2$ years, their $E\{T\} \simeq 10,600$ years. Point B describes the drivers whose long-term conviction rate is 5 times the population average ($m = 0.7$ convictions per year). With the same N and D, for these drivers $E\{T\} \simeq 15$ years.

If N is increased from 4 to 5, A moves to A′ and B moves to B′. This indeed decreases the probability of suspending the license of a better-than-average driver by a factor of 20. Unfortunately, this move further undermines our ability to detect bad drivers, tripling their $E\{T\}$ to 46 years.

The other option is to change D. If D is made 1 year instead of 2 while N remains 4, Point A moves to A″ and B to B″. For the better-than-average driver $E\{T\}$ is now 152,000 years, a shift in the desired direction. However the $E\{T\}$ for the

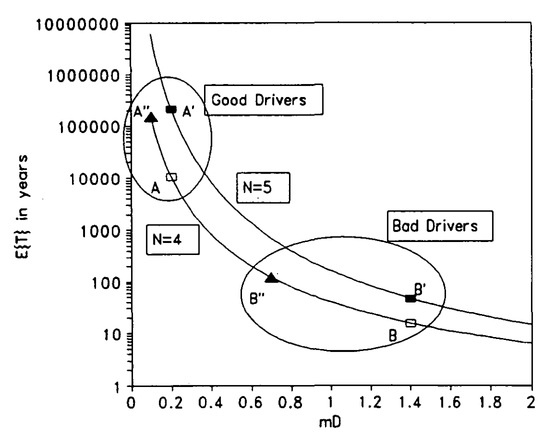

FIGURE 3 How $E\{T\}$ depends on N and D.

deviant drivers increases from 15 to 122 years. This is not what we wished to accomplish.

Thus, if one wishes to suspend the license of fewer better-than-average drivers, it will take even longer to catch those who have a truly deviant conviction rate and vice versa. Even though the performance of the *N-D* trigger in the demerit point context is rather dismal, there may be situations in which its performance is acceptable. Therefore it is still of interest to examine which choices of *N-D* are better than others.

CHOOSING THE BEST *N-D* PAIR

The main aim of the trigger is to identify entities having an unusually high m (m_{high}) and to do so, on the average, within a short time $E\{T|m_{\text{high}}\}$. There are many *N-D* pairs that all give the same $E\{T|m_{\text{high}}\}$. We propose to choose the *N-D* pair that makes $E\{T|m_{\text{low}}\}$ as long as possible for those entities having a relatively low m. This guidance for choosing an optimal *N-D* pair is applicable to the "new entity case" only. Finding an optimal *N-D* for the surviving entity case on the basis of the preceding proposition is meaningless since its structure suggests that one could always increase $E\{T|m_{\text{low}}\}$ by increasing D.

For a given m, N, and D, it can be shown that for the new entity case,

$$E\{T\} = [1 - F_T(D)]\left(D + \frac{1}{\bar{p}m}\right) + \frac{N}{m}\left(1 - \sum_{i=0}^{N}\frac{(mD)^i e^{-mD}}{i!}\right) \quad (6)$$

The way in which N and D influence $E\{T\}$ is shown in Figure 4. The figure shows $E\{T|m = 5\}$ for values of N equal to 10 and 15 and various values of D.

A number of observations can be made about $E\{T|m\}$ from this figure:

- $E\{T|m\}$ is generally high for small values of D;
- For any value of N, $E\{T|m\}$ has a minimum value, $E_{\min}\{T|m\}$;
- For large values of D, $E\{T|m\}$ approaches an asymptotic value of N/m; and
- $E_{\min}\{T|m\}$ increases as N increases.

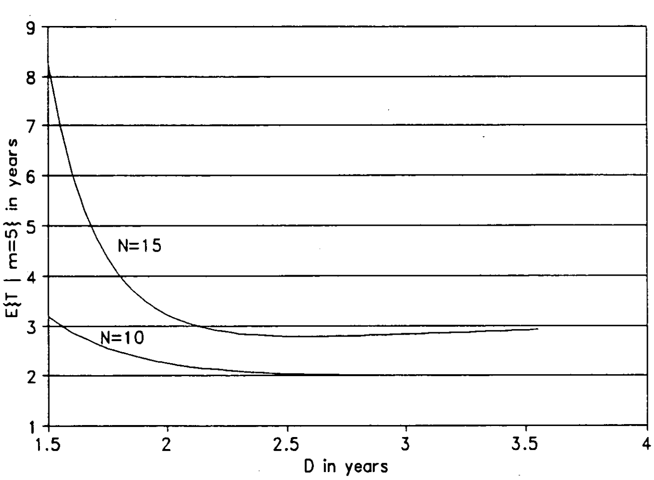

FIGURE 4 $E\{T|m = 5\}$ **as a function of** N **and** D **for the new entity case.**

As stated earlier, different *N-D* pairs can be found that yield the same values of a prespecified $E\{T|m_{\text{high}}\}$. However, from the foregoing, it is observed that the feasible values of N are those that meet the condition that the specified $E\{T|m_{\text{high}}\}$ is greater than or equal to $E_{\min}\{T|m_{\text{high}}\}$. This defines an upper bound for N. How an optimal *N-D* can be found will be illustrated with the following numerical example.

Numerical Example 6: In Numerical Example 2, of 1,000 similar intersections, 10 had a long-term average $m = 3$ accidents per year and 900 had $m = 1$ accident per year. Let these be the m_{high} and m_{low}, respectively. The aim is to identify the dangerous intersections within, say, 2 years. Several combinations of N and D that all have $E\{T|m = 3\} = 2$ years are given in Table 3.

Thus, the best *N-D* pair in this case is $N = 4$ and $D = 0.83$ years, which is quite close to the MUTCD warrant ($N = 5$, $D = 1$). Whereas intersections with $m = 3$ will be detected, on the average, in 2 years, intersections with $m = 1$ will survive, on the average, 30 years. Stated differently, one of 30 intersections with $m = 1$ accident per year will meet the warrant per year.

If the aim was to identify intersections with, say, $m_{\text{high}} = 5$ accidents per year, the optimal trigger would be $N = 9$, $D = 1.9$, which is quite different from that recommended in the MUTCD. In this case $E\{T|m = 1\} = 2{,}000$ years. A computer program for determining the best *N-D* pair is available on request. Some results are given in Tables 4 and 5.

DETECTION OF JUMPS IN *m*

So far we have imagined mainly the circumstance in which the m of each entity remains nearly constant over time and the task is to identify entities having an unusually large m. In this case, the longer the history of event occurrence for each such entity, the better one can detect the sought-after entities. One must therefore ask what sense it makes to use only the information contained in a window of size D and to disregard the portion of the event history to the left of the window.

There are, however, circumstances in which the use of a relatively short D might have appeal. One such circumstance is when a large and perhaps sudden increase in m might occur at some unknown time and the aim is to recognize it as soon as possible. Another is when D denotes distance, not time, and the window is sliding along some stretch of road. Here

TABLE 3 Combinations of N **and** D **Yielding** $E\{T|m = 3\} = 2$ **Years**

N	D	$E\{T\|m=1\}$
2	0.08	15.22
3	0.37	24.81
4	***0.83***	***29.85***
5	1.53	27.38
6	3.60	11.43

TABLE 4 Values of N, D, and $E\{T|m_{low}\}$ When $E\{T|m_{high}\}$ = 2 Years

m_{high}	m_{low}	N	D	$E\{T\|mlow\}$
3	1.5	4	0.83	9.31e+00
3	0.75	5	1.52	7.60e+01
3	0.3	5	1.52	3.47e+03
4	2.0	5	0.80	1.23e+01
4	1.0	6	1.23	1.88e+02
4	0.4	7	1.79	3.28e+04
5	2.5	6	0.80	1.58e+01
5	1.25	8	1.47	4.79e+02
5	0.5	9	1.89	3.76e+05
6	3.0	8	1.29	2.05e+01
6	1.5	11	1.82	1.84e+03
6	0.6	11	1.82	7.28e+06

the aim is to recognize when on a short stretch of road m is inordinately high.

It is unlikely that a trigger that makes no use of data outside the window of length or duration D can be an efficient device even in these circumstances. However, to give constructive advice on this matter is not simple. What can be shown with the results obtained so far is that in road safety the detection of jumps in m is a difficult task. In examining this question consider again the surviving entity case for which $E\{T\} = 1/(\bar{p}m)$ (see Equation 5).

In Numerical Example 5 a better-than-average Ontario driver was said to have 0.10 pointable convictions per year. To suspend the license of such a driver relatively rarely, $N = 4$ and $D = 2$ were chosen to make $E\{T|m = 0.1\} = 10{,}600$ years. Imagine that for some reason the m of this driver increases sevenfold to 0.7 convictions per year and the aim is to detect this jump not long after it occurs. For our purpose the instant of the jump in m can be taken as the time origin of a "surviving entity case." As was shown in Numerical Example 5, it will be 15 years on the average before the jump is detected. So, even with a relatively short D, a big jump in m still takes a long time to detect.

Similarly, imagine that the aim is to detect intersections where the m has jumped from 2 to 4. Using the best trigger ($N = 5$, $D = 0.8$ years, Table 4), it will still be 2 years on the average before the jump is recognized. Furthermore, in 1 out of every 12 intersections where no jump in m occurred (i.e., m remained two accidents per year), the trigger will be met.

In summary, when the m's are assumed not to change in time, the entire available historical record is relevant. To use the N-D trigger in this case makes little sense. The use of the N-D trigger might be justified when it is suspected that m can change suddenly and the aim is to detect such a change soon. Unfortunately, in the circumstances characteristic of road safety, the aim seems to be unattainable. The rate of event occurrence is too small to detect a sudden increase in m within few years of its occurrence without incurring the penalty of many false positives.

TABLE 5 Values of N, D, and $E\{T|m_{low}\}$ When $E\{T|m_{high}\}$ = 3 Years

m_{high}	m_{low}	N	D	$E\{T\|m_{low}\}$
3	1.5	5	0.96	2.09e+01
3	0.75	7	2.06	4.38e+02
3	0.3	8	2.80	1.57e+05
4	2	9	1.56	3.07e+01
4	1	11	2.73	2.76e+03
4	0.4	11	2.73	1.09e+07
5	2.5	29	2.71	1.78e+04
5	1.25	29	2.71	3.23e+12
5	0.5	29	2.71	1.47e+23
6	3	41	2.738	1.56e+08
6	1.5	41	2.738	5.65e+18
6	0.6	41	2.738	9.94e+33

SUMMARY

Procedures of the N-D trigger kind seem to be used in accident warrants and demerit point systems. The aim is usually to identify entities that are unsafe. In this paper the statistical properties of the N-D trigger are established and their repercussions for practical use are illustrated.

When the trigger is applied to a population of entities, inevitably there are false positives and false negatives. Since, by definition, normal entities are many and deviants are few, the trigger is often pulled unnecessarily. This leads to the conclusion that entities identified by the N-D trigger should be subsequently subjected to a sound diagnostic procedure capable of separating the wheat from the chaff.

The normal use of the N-D trigger is to scan the population of entities continuously or periodically. It is shown that if the unsafety of entities in the population changes slowly, repeated use of the trigger will deplete the supply of deviants while the majority of normal entities will provide a steady supply of false positives. After a few applications, the usefulness of the trigger may be exhausted.

A simple way to describe the performance of a trigger is by stating its mean-time-to-trigger as a function of the m of entities. Thus, for example, if a typical demerit point system is applied to a population of drivers, it takes 15 years on the average to remove the license of a driver whose conviction rate is 5 times the normal, and still 1 in 212 better-than-average drivers will lose the driver's license during the driving career.

If an N-D trigger is to be used, one may want to know how to choose the N-D pair sensibly. The choice can be made by first specifying the desired mean-time-to-trigger for entities considered deviant and then making the mean-time-to-trigger for a better-than-average entity as long as possible.

Whether to use an N-D trigger at all is an open question. It seems contrary to plain sense to disregard events that occurred outside the window of size D even if the intent is to detect sudden and large jumps in the rate of event occurrence. In any case, one can show that for the events of interest

(accidents at intersections or road sections, convictions for drivers) the rate of event occurrence is too small to have a realistic hope that sudden jumps can be detected soon after occurrence. It appears that the *N-D* trigger is not a particularly good device for identifying entities meriting attention.

DERIVATION OF $F_T(t)$

The equations in this section are based on the notation and description of the *N-D* trigger as given earlier. The probability $F'_T(t)\Delta t$ that N events show in a window of size D, for the first time, as its right edge moves from t to $t + \Delta t$ is the probability of the conjunction of the following three occurrences:

- The probability that N events did not materialize in a window of size D from when its right edge was at the origin until its right edge reached t [this is $1 - F'_T(t)$];
- The probability, $p(t)$, that when the right edge was on t, there were exactly $N - 1$ events in the window, given that the number of events in the window can be 0, 1, 2, . . ., $N - 1$; and
- The probability that an event was added to the window as the right edge moved from t to $t + \Delta t$ while none was deleted at the left edge (this is approximately $m\Delta t$).

The product of these three constituent probabilities is

$$F'_T(t)\,\Delta t = [1 - F'_T(t)] \times p(t) \times m \times \Delta t \qquad (7)$$

When Δt approaches 0 this relationship turns into the simple separable differential equation $F'_T(t)/[1 - F'_T(t)] = p(t)m$, for which the general solution is

$$F_T(t) = 1 - Ce^{-m\int_0^t p(\tau)d\tau} = 1 - Ce^{-m\bar{p}t} \qquad (8)$$

where C is a constant of integration and $\bar{p}(t)$ is the mean value of $p(\tau)$ in $[0,t]$. Elsewhere (*2*) it was assumed that $\bar{p}(t)$ was independent of t and was equal to the conditional probability

$$\bar{p}(t) = \frac{\dfrac{(mD)^{N-1}}{(N-1)!}}{\sum_0^{N-1} \dfrac{(mD)^i}{i!}} \qquad (9)$$

This was an erroneous assumption since $\bar{p}(t)$ is in fact dependent on t.

Liu (*3*) developed correction factors for the expression for $\bar{p}(t)$ given in Equation 9. This was done by a simulation of the *N-D* trigger using various values of m, N, and D. Through an elaborate function fitting process, Liu found that a good approximation for $\bar{p}(t)$ is obtained when the expression in Equation 9 is multiplied by a correction factor k, where

$$\begin{aligned} k = {} & (0.0009 \times 0.5941^{-N} + 0.8754) \\ & + mD(0.0917 \times 1.3256^{-N} - 0.1273) \\ & + (mD)^2(0.0339 \times 1.4009^{-N} + 0.0073) \end{aligned} \qquad (10)$$

REFERENCES

1. *Manual on Uniform Traffic Control Devices for Streets and Highways.* Federal Highway Administration, U.S. Department of Transportation, 1988.
2. E. Hauer and K. Quaye. On the Use of Accident and Conviction Counts To Trigger Action. In *Transportation and Traffic Theory* (M. Koshi, ed.), Elsevier, 1990, pp. 153–172.
3. Z. Liu. *N-D Trigger Simulation.* Master's thesis. Department of Civil Engineering, University of Toronto, Toronto, Ontario, Canada, 1992.
4. R. P. Bhesania. Impact of Mast-Mounted Signal Heads on Accident Reduction. *ITE Journal,* Vol. 61, No. 10, 1991, pp. 25–30.
5. N. Lalani. Comprehensive Safety Program Produces Dramatic Results. *ITE Journal,* Vol. 61, No. 10, 1991, pp. 31–34.

DISCUSSION

PATRICK BUTLER
National Organization for Women, 1000 16th Street, N.W., Suite 700, Washington, D.C. 20036–5705

By means of hypothetical accident and conviction rates, the paper evaluates the use of random events to single out individual intersections or drivers from apparently homogeneous populations for remedial treatment. For the intersection calculations, the necessity for equal exposure to risk is explicitly recognized by specifying that the intersections are "all serving similar traffic flows." The question of the known wide range in kilometers of exposure among drivers is ignored, however, in the model for the distribution of traffic convictions among drivers. By focusing on identifying "good" and "bad" drivers through annual conviction rates, the paper supports the erroneous insurance notion that driving risk can be accurately assigned to individual cars on an annual basis. In fact, however, the risk of traffic convictions and accidents is zero when a car is parked and increases kilometer by kilometer as a car is driven. The significant, predictable effect on insurance costs for drivers, especially for low-kilometer drivers, of ignoring individual exposure will be examined through consideration of the paper's Numerical Example 5.

With reference to a conviction rate of 0.14 per year averaged over all Ontario drivers, Example 5 specifies for the sake of its calculations that drivers with a conviction rate of 0.1 per year are "good" and that drivers with a conviction rate of 0.7 per year are "bad" and should be subject to having their licenses suspended. The Poisson calculations of Example 5 show that the current method specifying four convictions in 2 years (and other number-year criteria as well) for license suspension is very inefficient because too few 0.7-rate drivers and too many 0.1-rate drivers would meet this criterion. Nonetheless, the paper encourages continued efforts "to detect bad drivers," who, it appears to suggest, are those with annual conviction rates of 0.7 (or more). Table 6 tests the capability of such annual rate criteria for measuring driving risk.

Table 6 gives the annual kilometer exposures and conviction rates per kilometer chosen for four hypothetical drivers to produce the annual conviction rates used in Numerical Example 5. The kilometer-traveled value assigned each driver is within the range of what many drivers and cars typically travel in a year.

TABLE 6 One Year's Experience for Four Hypothetical Drivers

Driver	Km traveled during year (1)	Class average convictions per 10^6 km (2)	Convictions per year (1) x (2)	Paper's driver assessment based on convictions-per-year rate
1	10,000	10	0.1	"good"
2	70,000	10	0.7	"bad"
3	100,000	7	0.7	"bad"
4	10,000	70	0.7	"bad"

The table shows that a wide range of driving-exposure and conviction-rate combinations can produce a given low or high annual conviction rate. Although the class conviction rates per kilometer assumed for Drivers 1, 2, and 3 are close to the Ontario average (8.4 per 1 million km in 1990), the 0.7 annual rate calculated for Drivers 2 and 3 makes them "bad" drivers and, according to the paper, deserving of license suspension if they could be identified. The idea that individual "good" and "bad" drivers can be defined on the basis of annual rates of convictions or accidents is demonstrably not valid.

It may seem that those qualifying as "bad" drivers because of high annual kilometers driven are disadvantaged by being more likely to receive convictions and license suspensions than the low-kilometer "good" drivers. This is not true in the insurance pricing context, however. Since premiums are charged at annual class rates that are little affected by kilometers driven, the more driving that is done, the less is paid per kilometer for on-the-road insurance protection. With the same coverage and in the same class, Driver 2 would pay one-seventh the per-kilometer rate for insurance that Driver 1 does. Viewed another way, Driver 1 would pay 7 years of premiums for the 70,000 km of exposure to risk that Driver 2 pays one premium for.

Although reliable records of kilometers driven are not routinely kept for individual drivers, the kilometers driven by individual cars are registered on their odometers. Legislation requiring conversion of automobile insurance class rates from dollars per car year to cents per car kilometer as an anti–price discrimination measure is under consideration in several states because low-kilometer drivers are forced under the present system to pay much more per kilometer for identical insurance protection than high-kilometer drivers with cars in the same class. As now, premium would continue to be paid in advance to keep insurance in force, with the odometer limit to the prepaid protection displayed on the car's insurance card. Illegal odometer tampering, detected at the insurance company's annual audit or during accident investigation, would automatically void the policy.

A major block to such reform, however, is insurers' specious proof that classification of cars based on claim and conviction records provides cost justification for providing discounts for cars with "good" drivers and surcharges for cars with "bad" drivers. The criteria for these classifications are all variants of the Ontario number-year criterion for license suspension, such as years since last claim for "good" driver discounts and "bad" driver surcharges tied to claims and convictions within the last 3 years. The surcharge classes consistently experience more claims per car than the discount classes, both in fact and as modeled through Poisson calculations (*1*).

Although accidents (and traffic convictions) are random events, cars driven more kilometers than the average for their price class are more exposed to chance of accident and will be overrepresented in the class minority that has accidents and convictions. It is not "bad" drivers but overrepresentation of higher-kilometer cars that produces the higher accident averages that insurers invoke to justify the surcharging. Unlucky lower-kilometer insureds, already overcharged at "good" driver class rates, are more heavily overcharged by "bad" driver surcharges.

REFERENCE

1. P. Butler and T. Butler. Driver Record: A Political Red Herring that Reveals the Basic Flaw in Automobile Insurance Pricing. *Journal of Insurance Regulation,* Vol. 8, No. 2, Dec. 1989, pp. 200–234.

AUTHORS' CLOSURE

The object of our Numerical Example 5 is to quantify the implication of using an *N-D* trigger to identify "deviant" drivers. We base our analysis on an *N-D* trigger similar to the one used in Ontario (i.e., one with $N = 4$ pointable convictions and $D = 2$ years). In the hypothetical population used in Numerical Example 5, we therefore used the conviction experience of drivers in Ontario as a basis for defining a good driver as someone with an expected value of 0.1 convictions per year. To further explore the performance of the *N-D* trigger, we define a bad driver for this population as one whose expected number of convictions per year is 7 times that of our good driver (i.e., 0.7 convictions per year).

The discussant seems to be uncomfortable with the notion that drivers can thus be classified as good or bad on the basis of their accident or conviction record. This is because driving exposure (henceforth referred to as exposure) as measured by the number of kilometers driven per year by various drivers may be vastly different.

For a population of drivers with vastly different exposure, defining safety as the expected number of accidents or convictions alone can be problematic and misleading. Nonetheless, a nonhomogeneous group of this nature could be divided into sets of homogeneous groups in which drivers have fairly similar exposure. Within each homogeneous subpopulation, it is expected that the long-run average of the number of accidents or convictions per year will vary from driver to driver. Thus conditional on exposure, one is still in a position to classify some drivers as good and others as bad on the basis of their accident or conviction records.

For instance, in the table provided by the discussant, Drivers 1 and 4 have identical exposures of 1000 km/year. If the number of convictions in the table is representative of the long-run conviction experience of these two types of drivers, it would be appropriate to say that Driver 4 is worse than Driver 1. Let us suppose that an analyst desires to use the driving record of a population of drivers similar to these two to identify the bad drivers in the group for some form of remedial action. If an approach such as the *N-D* trigger is used, our study provides the analytical tool for examining the implication of using any chosen *N-D* pair.

Essentially, our goal is to provide the quantitative framework for examining the effect of identifying entities (drivers or intersections) with an *N-D* trigger. In the case of a population with different exposures, the entities should be subdivided into homogeneous groups before exploring different choices of *N* and *D*.

Since the setting of insurance premiums and the fairness of premiums are beyond the scope of our study, we reserve our comments on those issues. In closing, though, it is worthwhile to note that the discussant appears to favor the use of exposure-based accident or conviction rates (e.g., accidents per vehicle-kilometer) as a basis for defining good or bad drivers. However, such a measure has its drawbacks, especially in cases where the relationship between exposure and the expected number of accidents or convictions is nonlinear.

Publication of this paper sponsored by Task Force on Statistical Methods in Transportation.