APPLICATION OF STATISTICAL SAMPLING METHODS TO TRAFFIC PERFORMANCE AT URBAN INTERSECTIONS

By Dr. Bruce D. Greenshields

Associate Professor of Civil Engineering, George Washington University[1]

SYNOPSIS

Recently, (1942) the frequency of time spacings between successive vehicles obtained by O. K. Normann of the Public Roads Administration, was found to coincide very closely with the theoretical spacing given by the Poisson law. The differences in the spacings are insignificant except for those less than 1 sec. Due to the lengths of vehicles and the spacing required by drivers to avoid rear end collisions, it is obvious that there should be relatively few spacings in moving traffic under one second compared to the theoretical series which refers to dimensionless points.

In certain calculations, such as delay at signalized intersections, it is not the distribution of spacings that is used, but the number of vehicles appearing in a given time interval. For relatively large time intervals, say about 15 seconds, the distribution holds very closely to the random series.

Often very little calculation is necessary to arrive at results, for there are published tables giving both the individual and the cumulated terms of the Poisson series. With traffic behavior conforming to the Poisson series, it is expected that the techniques of sampling and control found valuable in industry will find increasing use in traffic control and regulation.

In a study of the geometry of traffic movements at urban intersections conducted by the Yale University Bureau of Highway Traffic, during the past two years, new data were found to confirm the assumption that the time and space distributions of vehicles along a roadway conform very closely to the Poisson random series.

It is the purpose of this paper, after first showing that road traffic follows a random series closely conforming to the Poisson series (*1*)[2], to examine the nature of such a series, and then to demonstrate how the statistical principles involved may be applied in the solution of some typical traffic problems.

The theory that traffic follows a random distribution, as depicted by the Poisson law, is not new. Mr. John P. Kinzer made such an assumption in calculating the probability of a car picked at random going a mile without interference on a two-lane road with a given volume of traffic (*2*).

In England, Mr. William F. Adams found that in general the flow of traffic conforms so well to the random series distribution that such a distribution may be described as normal. According to Mr. Adams, departure from this normal does not become pronounced until a volume of over 1,000 vehicles per hr is reached (*3*). Control devices tend to distort the distribution, but only in the immediate vicinity of the device.

In a study of rural traffic conducted by the writer in Ohio in 1934, it was found that traffic may be considered to be purely random on a two-lane roadway with volumes not exceeding 400 vehicles per hr for the two lanes. For volumes greater than this, there was a break in the distribution curve that was judged to be caused by the mutual interference of vehicles traveling at less than about 200-ft spacing which corresponded to about a 4-sec headway at the prevailing speed (*4*).

More recently, the frequency of time spacings between successive vehicles obtained by O. K. Normann (*5*), as shown in Table 1 is found to coincide very closely with the theoretical spacings given by the Poisson law. The differences in the spacings are insignificant except for those less than 1 sec. Due to the physical lengths of vehicles and the spacing required by drivers to avoid rear end collisions, it is to be expected that there would be relatively fewer spacings in moving traffic

[1] Formerly, 1944–46, with the Bureau of Highway Traffic, Yale University.

[2] Italicized figures in parentheses refer to list of references at the end of the paper.

under 1 sec than indicated by the theoretical series which refers to dimensionless points.

In the present study of urban traffic, the data as plotted in Figures 1, 2, and 3, show characteristic variations from a random series in that, there are scarcely any vehicles less than two frames (1.36 sec) apart, that below

TABLE 1

COMPARISON OF SPACINGS BETWEEN VEHICLES CALCULATED BY THE POISSON LAW, WITH SPACINGS FOUND BY NORMANN (5)

Vehicles Per Hour	Spacing in Seconds					
	1		5		10	
	Observed	Calculated	Observed	Calculated	Observed	Calculated
100	2	3	13	12	27	24
200	4	5	24	24	44	43
400	7	10	43	43	68	67
800	14	20	67	67	90	89
1,600	26	36	90	89		

Values shown are percentages of spacings less than given number of seconds.

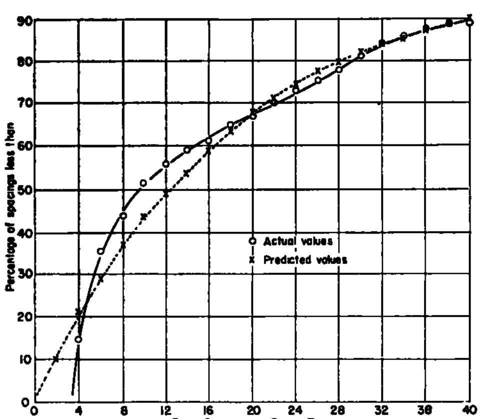

Figure 1. Comparative Time Spacings Between Vehicles—Traffic Volume 300 Vehicles per Hour—213 Spacings Recorded—Eastbound Traffic at Orange and George Streets, New Haven, Conn.

four frames there are fewer spacings than predicted by the random curve, that between 5 and 25 frames there are more spacings, and that above 25 frames the distributions are practically identical. (One frame equals $\frac{1}{88}$ min.) The heavier the volume the more spacings between 5 and 25 frames as compared to the Poisson distribution.

In certain calculations, such as delay at signalized intersections, it is not the distribution of spacings that is used, but the number

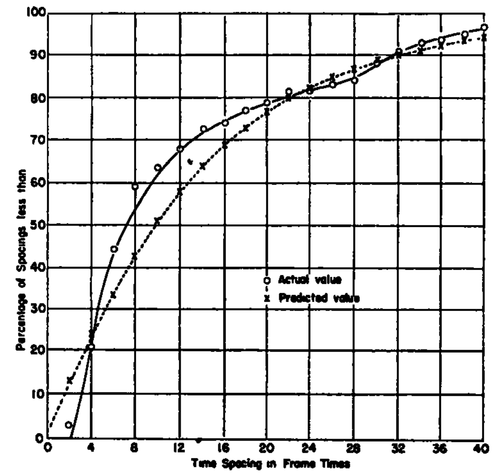

Figure 2. Cumulative Time Spacings Between Vehicles—Traffic Volume 378 Vehicles per Hour—174 Spaces Recorded—Orange and George Streets, New Haven, Conn.

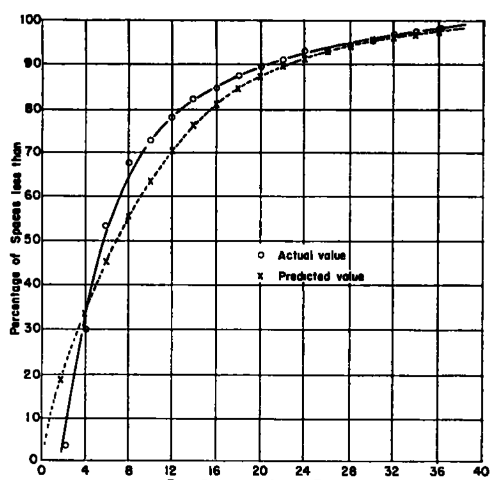

Figure 3. Cumulative Time Spacings Between Vehicles—Traffic Volume 540 Vehicles per Hour—348 Vehicle Spacings Recorded—Southbound Traffic at William and Fulton Streets, New York City.

of vehicles arriving in a given time interval. For relatively large time intervals, say about 15 sec, as shown by Figure 4, the distribution holds very losely to the random series.

In brief, it may be said that the Poisson series, when used to describe the irregular spacings between vehicles, gives results that are sufficiently accurate for practical purposes.

Poissons Exponential Expansion—The concept that traffic moving along a road is scattered at random is easily gained from a distant view of the road. In general, each vehicle is quite small compared to its distance from any other vehicle, and each vehicle moves independently. Different sections of the road of equal length may contain different numbers of vehicles and different numbers of vehicles may pass a given point in equal intervals of time. These stipulations with the added condition that the Poisson series applies to a discreet variable, dealing with integral numbers, such as the number of vehicles which may not be fractionalized, constitute the type of distribution to which the Poisson series applies.

For the purpose of stating the Poisson law in terms of traffic, let a road have P vehicles scattered along it at random so that any vehicle is completely independent of any other, and equal segments of the road are equally likely to contain the same number of vehicles.

The total number of vehicles, divided by the number of segments $= \frac{P}{z} = m$ the average number of vehicles in a segment.

According to the general term of the Poisson series, the probability of x vehicles appearing in a segment z is equal to

$$\frac{e^{-m} m^x}{x!}$$

wherein $x!$ equals factorial x and m equals the average number in a segment, and e equals the base of natural logarithms.

From the above, the probability (P_o) of zero vehicles appearing in a segment z is equal to

$$P_o = \frac{e^{-m} m^o}{o!} = e^{-m}$$

The probability (P_1) of one vehicle appearing is equal to

$$P_1 = \frac{e^{-m} m}{1!}$$

Likewise the probability of two vehicles appearing is equal to

$$P_2 = \frac{e^{-m} m^2}{2!}$$

The sum of all the probabilities is equal to

$$e^{-m}\left(\frac{m^o}{o!} + \frac{m^1}{1!} + \frac{m^2}{2!} + \cdots \frac{m^n}{n!} + \cdots\right)$$

This is equal to 1
for

$$\left(\frac{m^0}{o!} + \frac{m^1}{1!} + \frac{m^2}{2!} + \cdots \frac{m^n}{n!} + \cdots\right) = e^m$$

and therefore

$$e^{-m} e^{m} = e^o = 1 \ (6)$$

Mathematically speaking, this means that if an event is certain to happen, the probability of its happening is one or 100 percent. If an event has a 50–50 chance of happening, the probability of its happening is 0.5 or 50 percent.

ILLUSTRATIVE EXAMPLES

To illustrate the use of the theory just outlined, let a road 5 mi in length have 60 vehicles scattered along it at random and let it be required to find the probability of there being four vehicles on any half mile section. Using the notation above:

$$P = 60$$

$$z = 10$$

$$\frac{P}{z} = m = 6$$

The probability, P_4, of there being four vehicles is:

$$P_4 = \frac{e^{-m} m^x}{x!} = \frac{e^{-6} \cdot 6^4}{4!} = \frac{e^{-6} \cdot 6^4}{4 \cdot 3 \cdot 2 \cdot 1}$$

$$= 0.00248 \cdot 54 = 0.1339$$

Thus, on the average in one out of about 7.5 sections there will be four vehicles.

These calculations need not be made, however, for there are published tables that give the probability for any value of x corresponding to a given value of m (7).

The probability of there being four or more vehicles in a half mile section is found from the summation

$$\sum_{x=0}^{\infty} \frac{e^{-m} m^x}{x!} = \sum_{4}^{\infty} \frac{e^{-4} \cdot 6^x}{x!} = 0.8488$$

This value, also, need not be calculated but may be found from published tables (*1*) (*7*). Entering the Molina Table II with a value of $m = 6$ and $x = 4$, there is found the value, 0.848796. Thus out of 100 samplings of half mile sections about 85 (84.879) would have four or more vehicles.

The Probability of Time Gaps in a Traffic Stream—As a further example, given a volume of 360 vehicles per hr, let it be required to determine the probability P_o of there being no vehicles in a 1-sec time segment.

The average number of vehicles per sec is equal to $\frac{360}{3600} = \frac{1}{10} = m$. Therefore the probability of there being no vehicles in a one sec interval is equal to

$$P_o = \frac{e^{-m} m^x}{x!} = \frac{e^{-0.1} \cdot 0.1^0}{0!} = e^{-0.1} = 0.905$$

This value may also be found from the tables with m equal to 0.1 and $x = 0$.

In like manner, the probabilities of 2-, 3-, and 4-sec intervals with values of m equalling 0.2, 0.3, and 0.4 are found to be

0.8187 for 2 sec

0.7408 for 3 sec and

0.6703 for 4 sec

The Stop Sign—Another useful feature that may be used in connection with the Poisson series is that it is immaterial as to where a space or time segment is chosen. If the segments are chosen to be those immediately following the passage of a vehicle, there is obtained the probability (percentage of time) that an empty segment of given length (or longer) will appear.

The calculation of the time lost at a stop sign offers a useful application of this concept. For example given a volume of 900 vehicles per hr on the main street, let it be required to determine the percentage of the time a cross street vehicle will be able to cross without delay. The fact that the distribution is random in both directions means that the flow might just as well be considered to be one directional at 900 vehicles per hr. The Yale investigation revealed that a vehicle will cross without delay if there is no vehicle closer than 6 sec to the intersection. The average number of vehicles appearing on the main street in a 6-sec segment of time is equal to $900 \cdot \frac{6}{3600} = 1.5 = m$. From the Poisson table, it is found that a 6-sec gap will have no vehicles 22.3 percent of the time. This means that 22 3 percent of the time a vehicle crosses, it will suffer no delay and that for 77.7 percent of the time it will be blocked by one or more vehicles. Since there are 900 spaces in the moving stream of traffic per hr and 22.3 percent of these are 6 sec or longer, there are 201 opportunities per hr for crossing or one every 17.9 sec. The average wait is then $\frac{17.9}{2} = 8.95$ sec.

An entrance ramp to a main thoroughfare presents the same type of problem as the stop sign and may be solved in an identical manner.

The Theory of Random Distribution Applied to Signalized Intersections—For heavy volumes of traffic, stop signs are usually replaced by traffic signals. The application of the Poisson law to the analyses of traffic performance at signalized intersections will now be considered.

Before the time loss per vehicle at a signalized intersection for a given volume of traffic may be determined, it is necessary first to ascertain how often 0, 1, 2, 3, 4, 5 . . . vehicles will be retarded at the red signal. Traffic approaching an intersection may be thought of as a series of points scattered at random along a ribbon of time that moves through the intersection, while the blocking period of the traffic signal may be visualized as sweeping up all vehicles in a segment of the ribbon equal in length to the blocking interval. The longer the block, the more vehicles are swept up. Inasmuch as each vehicle in turn must wait for the one ahead to clear, the blocking period is equal to the red period of the signal for only the first vehicle and is longer for the second vehicle than for the first, and longer for the third than for the second, etc.

To determine the exact probabilities of retarding one or more vehicles, let

r = stop period of signal,

a = interval added by first vehicle (= 3.8 + 1.7 = 5.5 sec)[3],
b = interval added by second vehicle (= 3.1 sec),
c = interval added by third vehicle (= 2.7 sec),
d = interval added by fourth vehicle (= 2.4 sec),
e = interval added by fifth vehicle (= 2.2 sec),
f = interval added by sixth vehicle (= 1.7 sec), etc. (1.7 sec is the average minimum trailing interval between vehicles traveling within one mi per hr of the same speed.)

The first vehicle faces a stop period of r, but once retarded, the blocking period for the second vehicle becomes $r + a$ and the chance of its being retarded is increased. To determine the probability of retarding just one vehicle it is necessary to use the blocking period $r + a$ for in order to retard just one vehicle, two conditions must be met: (1) Only one vehicle must arrive during the red r period; and (2) No vehicle must arrive during the additional blocking period a.

In order for just two vehicles to arrive, the following circumstances must occur (using the blocking period $r + a + b$): (1) two in r, none in a, and none in b; or (2) one in r, one in a, and none in b.

In order to retard just three vehicles, any one of the following combinations of circumstances must occur:

1. 3 in r, 0 in a, 0 in b, 0 in c,
2. 2 in r, 1 in a, 0 in b, 0 in c,
3. 2 in r, 0 in a, 1 in b, 0 in c,
4. 1 in r, 2 in a, 0 in b, 0 in c,
5. 1 in r, 1 in a, 1 in b, 0 in c.

In all five cases, no vehicle may be retarded in period c the new period created by the third vehicle.

Following the same procedure, it would be found that there are 14 ways to retard the fourth vehicle and correspondingly more for the fifth and so on. This indicates that this approach to the problem is not very practical. For the sake of clarity however a numerical example is included.

Let it be required to find how many times in 100 cycles just two vehicles will be stopped if the volume of traffic is 200 vehicles per hr for the single lane considered, and the stop period of the signal is 20 sec.

In this case
r = 20 sec and as given above
a = 5.5 sec and
b = 3.1 sec

$$m \text{ for } r = \frac{20 \cdot 200}{3600} = 1.1$$

$$m \text{ for } a = \frac{5.5 \cdot 200}{3600} = 0.3$$

$$m \text{ for } b = \frac{3.1 \cdot 200}{3600} = 0.17$$

Then from the formula or Poisson table, it is found that:

Probability of 2 vehicles appearing in $r = 0.20$
Probability of 0 vehicles appearing in $a = 0.73$
Probability of 1 vehicle appearing in $r = 0.37$
Probability of 1 vehicle appearing in $a = 0.22$
Probability of 0 vehicles appearing in $b = 0.84$

Substituting these values in the criteria given above, that there must be two vehicles in r and none in a or one vehicle in r and one in a and none in b, the total probability becomes

$$\begin{aligned} 0.20 \times 0.74 \times 0.84 &= 0.1243 \\ 0.37 \times 0.22 \times 0.84 &= 0.0684 \\ \hline &\ 0.1927 \end{aligned}$$

Thus in approximately 19 out of 100 cycles, exactly two vehicles will be retarded.

Theoretical Method of Calculating Delay[4]—It has just been shown that it is theoretically possible to calculate the probabilities of retarding various numbers of vehicles. This is one of the two steps requisite to determining the delay per light cycle at a red signal. The next step is to compute the average delay for

[3] These are the constants found in the study by the Yale Bureau of Highway Traffic.

[4] The mathematical approach outlined in this section was suggested by Dr. Oystein Ore, Sterling Professor of Mathematics, Yale University. For the specific mathematic treatment in this section, the writer is indebted to Aadne Ore, F. E. Loomis Fellow in Physics, Yale University.

the first, second, third, etc., vehicles in line. These two steps may be combined into one operation.

The delay of the first vehicle in line is the difference between the time it would have entered the intersection had there been no red signal and the time it actually enters the intersection. Let:

T_1 = Potential blocking period for the first vehicle (i.e. the red interval)

x = Time after beginning of red interval that first vehicle arrives

t_1 = Time after beginning of the red interval that first vehicle enters intersection

$\phi = \dfrac{3600}{\text{vph}}$ or average spacing in seconds between vehicles

$u = \dfrac{1}{\phi}$ = average number of vehicles per sec

L_1 = Average time loss of first vehicle.

If the first vehicle arrives during the red signal or before T_1 (that is if x is less than T_1), it suffers a delay equal to $(t_1 - x)$. The amount of delay varies with x. For any particular value of x there is a measurable probability of a vehicle's arrival. If the delay a vehicle suffers when it does arrive at a particular x be multiplied by the probability of its arrival, the average delay is obtained.

To arrive at a time x after the signal change, the vehicle must not arrive during x but during an additional time dx as dx approaches zero.

The probability of the vehicle's not arriving during x is

$$\frac{e^{-xu}\cdot(ux)^0}{0!}\,e^{-ux}$$

wherein xu is the average number m of vehicles arriving in x seconds. The probability of the vehicle's arriving in dx is

$$\frac{e^{-udx}\cdot(udx)^1}{1!}.$$

If dx approaches zero then

$$e^{-udx}\text{ approaches } 1 \qquad \text{and}$$

$$\frac{e^{-udx}\cdot(udx)_1}{1!} = udx$$

The probability of one vehicle arriving at x is then the probability of zero vehicles arriving during x multiplied by the probability of one vehicle arriving in dx, as dx approaches zero in the limit, or expressed mathematically it is

$$e^{-ux}\cdot u\,dx$$

The loss of time suffered by the first vehicle arriving at x is

$$t_1 - x = L_1$$

as long as x is less than T_1. That is, as long as the first vehicle arrives before the end of the blocking period. If x is greater than T_1 then of course there is no delay to the first vehicle since it arrives after the end of the red period.

Combining terms and summing up (integrating) the various probabilities as x varies from zero to T_1:

$$L_1 = \int_0^{T_1} (t_1 - x)e^{-xu}u\,dx$$

$$= t_1 - t_1e^{-T_1u} + \frac{e^{-T_1u}}{u} + T_1e^{-T_1u}$$

Combining terms and substituting ϕ for $\dfrac{1}{u}$

$$L_1 = t_1 - \phi + [\phi - (t_1 - T_1)]e^{-uT_1}$$

For the second vehicle the average time loss is computed in a similar manner. Using the symbols as above:

T_2 = Potential blocking period for second vehicle, which is $T_1 + a$ or $r + a$

t_2 = Time in seconds after light change to red that second vehicle enters intersection

L_2 = Average loss of time for second vehicle.

The second vehicle is retarded if it arrives any time during T_2, that is, x must be less than T_2. But there must always be a first vehicle and the first vehicle must be in T_1. Accordingly there are two possibilities of retarding the second vehicle:

1. The first arrives in T_1 and the second arrives in T_1 (both during the red period)
2. The first arrives in T_1 and the second arrives in $T_2 - T_1$, that is one must arrive during T_1, none in $x. - T_1$ and one at x or in dx.

These probabilities and their corresponding delays must be added. The mathematical expression for the probability of (1) above is the product of

$e^{-xu}xu$ (one in x) and $u\,dx$ (one in dx)

or $e^{-xu}xu^2\,dx$

and the corresponding average delay is:

$$\int_0^{T_1} (t_2 - x)e^{-xu}\,xu^2\,dx$$

The probability of (2) above occuring is the product of

$\frac{e^{-T_1u}(T_1u)}{1!}$ (one in T_1),

$e^{-(x-T_1)u}$ (none in $x - T_1$), and $u\,dx$ (one in dx)

and the corresponding delay is:

$$\int_{T_1}^{T_2} (t_2 - x)e^{-T_1u}\,T_1ue^{-(x-T_1)u}\,u\,dx$$

The average loss of time to the second vehicle is the sum of the above or

$$L_2 = \int_0^{T_1} (t_2 - x)e^{-xu}(xu)^2\,dx + \int_{T_1}^{T_2} (t_2 - x)e^{-T_1u}\,T_1e^{-(x-T_1)u}\,u^2\,dx$$

$$= t_2 - 2\phi + T_1[1 - u(t_2 - T_2)]e^{-T_2u} + [2\phi - (t_2 - T_1)]e^{-T_1u}$$

In order to compute the total delay caused by the signal, it would be necessary to calculate the average delays for the third, fourth, fifth, etc., cars and then add them all. As the number of vehicles increases, however, the mathematics becomes too involved to be practical. There is only one way in which one vehicle may be retarded, but there are two ways to retard the second, five ways to retard the third, fourteen ways to retard the fourth, etc. At the present time, no simplification has been found which would make this theoretically correct method workable enough for ordinary computation. The foregoing calculations are shown with the thought that they may provide a clue to a subsequent solution. Until a theoretically correct solution is derived, however, a practical approach is suggested.

A More Practical Method of Estimating the Average Number of Vehicles Retarded—As an illustrative sample of a practical method of calculating the average number of vehicles retarded, let it be required to find the average number of vehicles retarded for a traffic volume of 228 vehicles per hr on a single lane road with the signal set for 30-sec "go" and 20-sec "stop". The average number of

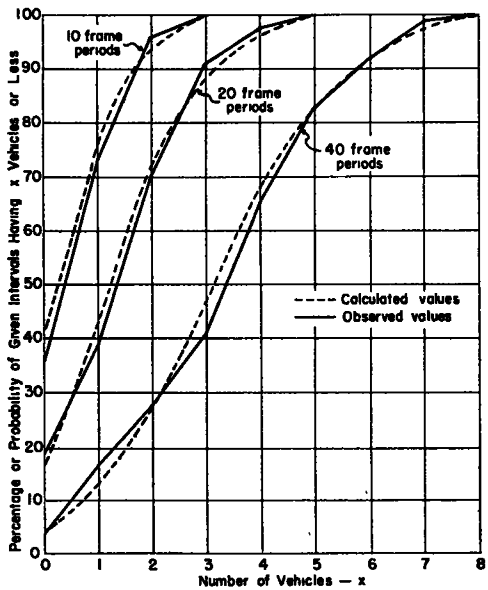

Figure 4. Probability that Given Time Intervals will Contain x Vehicles or Less Theoretical values are compared to observed data of 345 vehicles at Orange and George Streets, New Haven, Conn., with traffic volume of 490 vehicles per hr.

vehicles arriving during the 20-sec red period $= \frac{20 \times 228}{3600}$ vehicles $= 1.27$. This could result from approximately one for each of 3 cycles and 2 for the fourth cycle. As explained previously, however, these 1.27 vehicles tend to increase the effective length of the red signal. Reference to Figure 5 shows that 1.27 vehicles increase the blocking period by about 6.4 sec. The blocking period may now be considered to be 26.4 sec (20 + 6.4). A 26.4 second blocking period, will retard about $\frac{26.4 \times 228}{3600}$ vehicles or 1.67 vehicles.

The increase of the blocking period due to 1.67 vehicles is 7.7 sec and the blocking period may now be taken to be 27.7 sec. During the 27.7 sec blocking period 1.75 vehicles will be retarded to increase the blocking period to

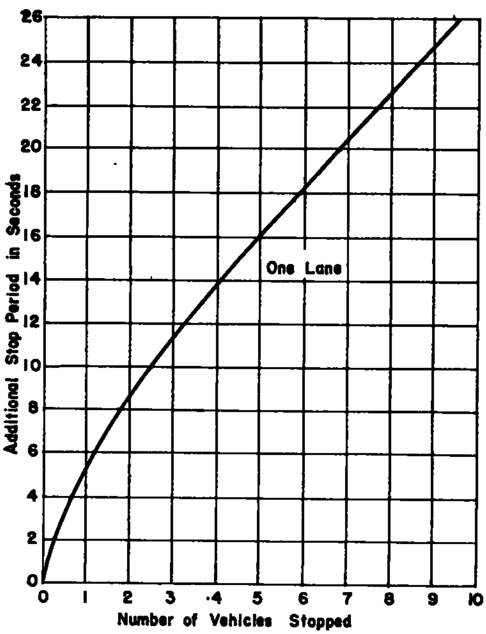

Figure 5. Additional Blocking Periods Created When Various Numbers of Vehicles are Retarded.

TABLE 2
AVERAGE NUMBER OF VEHICLES STOPPED WITH 228 VEHICLES PER HR PER LANE ON A THREE LANE ROAD, AND 20-SEC RED PERIOD

	Length of Blocking Period	Average Number of Vehicles Retarded
	sec	
1st approximation	20	1.27
2nd approximation	26 4	1.67
3rd approximation	27.7	1.75
4th approximation	27.95	1.77
5th approximation	28	1.77

27.95 sec. By further successive approximation, the number of vehicles retarded can be obtained with any degree of accuracy desired.

This information may be arranged in tabular form as in Table 2. For this particular example, it is sufficiently accurate to use an average of 1.77 vehicles per red signal. As shown in the table the 20-sec red signal corresponds to a 28-sec blocking period when the volume is 228 vehicles per hr.

Practical Method of Determining Delay—Having determined the average number of vehicles retarded and the effective blocking period, it is possible to calculate the delay to traffic at a signalized intersection if it is assumed that each vehicle on the average, arrives at the midpoint of the blocking period. It is easily comprehended that if traffic is approaching an intersection at random intervals, the average arrival time of the first vehicle is at the midpoint of the signal. But since the second vehicle always arrives after the first, it is not so easy to understand the assumption that the blocking period for the second vehicle starts at the same time as the first and that the second vehicle on the average arrives at the midpoint of its blocking period. This difficulty disappears if it is kept in mind that to arrive at random, vehicles must be regarded as having no physical size, so that they could all arrive at the same time.

The time loss for any vehicle is the difference between its arrival time and its leaving time. The first vehicle will arrive, on the average, at $\frac{r}{2}$ and leave at $r + 3.8$, wherein r equals the red period of the signal in seconds and 3.8 equals the time in seconds to enter the intersection after the signal changes to green. To this loss must be added 2 sec for accelerating to normal speed. Thus the loss amounts to:

$$r + 3.8 - \frac{r}{2} + 2$$

The stop period for the second vehicle will be $r + 3.8 + 1.7$, wherein 1.7 sec is the time in seconds that the second vehicle may enter after the first. The second vehicle enters (after being stopped), on the average, 3.1 sec after the first. The acceleration for the second vehicle is assumed to be 1 sec. The time loss for the second vehicle thus amounts to:

$$r + 38 + 3.1 - \frac{(r + 3.8 + 1.7)}{2} + 1$$

$$= r + 6.9 - \frac{(r + 5.5)}{2} + 1$$

In like manner the loss for the third vehicle is:

$$r + 9.6 - \frac{(r + 3.8 + 3.1 + 1.7)}{2} + 1$$

No acceleration loss is added for the fourth vehicle since it has reached normal speed by the time it enters the intersection. The loss for the fourth vehicle equals:

$$r + 14.1 - \frac{(r + 3.8 + 3.1 + 2.7 + 2.4 + 1.7)}{2}$$

For the fifth the loss equals:

$$r + 16.2 - \frac{(r + 3.8 + 3.1 + 2.7 + 2.4 + 2.1 + 1.7)}{2}$$

For the sixth the loss equals:

$$r + 18.3 - \frac{(r + 3.8 + 3.1 + 2.7 + 2.4 + 2.1 + 2.1 + 1.7)}{2}$$

And so on for others.

Comparing Results of Theoretical Calculations of Delay with Actually Observed Values—In order to check the theoretical results with actual ones, comparisons were made with data collected in a study conducted at the intersection of Constitution Ave and 17th St in Washington, D. C., by E. H. Holmes (*8*). For different signal settings the time to travel from 300 ft back of center line of the intersection to 300 ft past it, a distance totaling 600 ft, was measured. The traffic volume was fairly constant. The number of vehicles making left turns and the number of trucks were light. Consequently only the change in signal settings affected the change in traveling time.

For the multiple lane traffic at this intersection, the assumption was made that while stopping, traffic will distribute itself evenly over the road width. For instance, if ten vehicles are retarded on a two lane road, it was assumed that the traffic will behave so that five are retarded on each lane. Twelve cars retarded on a three lane road is then equivalent to four cars retarded on each lane of the road.

Therefore, to calculate the delay at an intersection with more than one lane per direction, the hourly volume of traffic in each direction is reduced to hourly volume of traffic per lane.

On Constitution Avenue, there were an average of 683 vehicles per hr using three lanes or about 228 vehicles per lane per hr. On 17th Street an average of 335 vehicles per hr were traveling on two lanes or about 168 vehicles per lane. In Table 3 the theoretical delay was calculated by the approximate method just outlined.

Unfortunately there is no way in which the delay as computed could be checked directly with Mr. Holmes' data since the latter gives only the total time required to traverse the 600 ft of intersection area. Information is lacking as to what the average traveling speed would be without the intersection delay. In order to get a basis for comparison, the computed delay was subtracted from the average total time recorded to give the "residual" running time. This residual running time was then translated into average residual running speed. The reliability of the calculated delays may be judged by the agreement of the residual speeds (or times). The range is from 17.2 to 20 mph. It will be noted that changing the green period on Constitution Avenue with the red period held constant at 15 sec changed the overall time to travel 600 ft by only 0.7 sec. This small variation is inconsistent with the other differences in the table and suggests that the observed value of 17.2 mph is questionable. If this value is omitted the range is from 18.4 to 20 mph.

The practicing traffic engineer is interested in knowing how a change in signal setting will affect delay. The comparisons given in Table 3

TABLE 3

COMPUTED INTERSECTION DELAY COMPARED WITH OBSERVED FIELD DATA—CONSTITUTION AVENUE AND M STREET, WASHINGTON, D. C.

Red Period on Constitution Ave.	Total Cycle Length	Observed Average Overall Time to Travel 600 Ft.	Calculated Delay	Residual Running Time	Residual Average Running Speed
sec.	*sec.*	*sec.*	*sec.*	*sec*	*m p.h.*
15	45	30.2	6.4	23.8	17 2
15	30	30.9	9.6	21.3	19.2
40	60	28.3	7.3	21.0	19.5
30	50	30.9	8.7	22.2	18.4
20	40	31 9	10.9	21 0	19 5
30	100	28.7	8.3	20.4	20.0
30	60	34.3	13.8	20 5	19.8

indicate that at least under the conditions observed, the relative values of delay may be determined with sufficient accuracy for signal timing.

Failure of Signals to Clear Traffic—The above calculations do not take into account the fact that the green signal may fail to clear some of the vehicles retarded by the red signal. It is important to determine the percentage of time a failure occurs for then the delay is not proportional to the number of vehicles retarded, since the last few cars have to wait through an additional red period before they can cross the intersection.

For an example, to illustrate how the chances of signal failures may be found, let it be required to find the probability of the cycle failure for 360 vehicles per hr on a single lane road with a 20 sec green and 20 sec red signal cycle. Since it requires 20 sec or more after the light changes to green for seven vehicles to enter the intersection, it will be assumed that the cycle will fail whenever seven or more vehicles appear in 40 sec.

The average number of vehicles appearing in 40 sec $= \dfrac{40 \times 360}{3600} = 4 = m$. With this value of m, the probability of seven or more vehicles appearing in 40 sec (found from table) equals 11.06 per cent. Therefore, the traffic signal will fail to clear the waiting traffic 11.06 percent of the time. Part of the time however, these failures may cause succeeding failures.

Any signal failure will affect the chances of a succeeding failure since there will be vehicles left over from the first cycle. In the present example, the second signal will fail if:

1. Seven vehicles arrive during the first and six or more during the second cycle
2. Eight vehicles arrive during the first and five or more during the second
3. Nine vehicles arrive during the first and four or more during the second
4. Ten vehicles arrive during the first and three or more during the second
5. Eleven vehicles arrive during the first and two or more during the second
6. Twelve vehicles arrive during the first and two or more during the second.

If the probabilities of the arrivals of the vehicles, as found in the Poisson tables, are multiplied together and added to give the total probability, the result is as follows:

1. $0.059540 \times 0.214870 = 0.012793$
2. $0.029770 \times 0.371163 = 0.011049$
3. $0.013231 \times 0.566530 = 0.007495$
4. $0.005292 \times 0.761897 = 0.004032$
5. $0.001925 \times 0.908422 = 0.001747$
6. $0.000642 \times 0.981684 = 0.000630$

$$0.037746$$

This means that two signals will fail in succession 3.77 percent of the time. In order to have three successive failures, there would need to be:

Thirteen vehicles in the first two cycles and six or more in the third,
Fourteen vehicles in the first two cycles and five or more in the third,
Fifteen vehicles in the first two cycles and four or more in the third, etc.

with the added condition that there be seven or more in the first cycle. While it is possible to compute the probabilities for these, it is not very practical. Therefore a much less tedious method that gives results that agree closely with the more exact procedure will now be described.

In the example just given the two cycles would fail in succession if 13 or more vehicles appeared during the two cycles, provided that seven or more appeared in the first cycle.

The average number appearing in two cycles (80 sec) equals $\dfrac{80 \times 360}{3600} = 8 = m$.

The probability of 13 or more appearing in the two cycles is 0.0638 as found in the Poisson tables (4 places is considered sufficient).

The average flow for the two failing cycles is not eight, the average flow on the roadway, but "13 or more vehicles". If it were known just how many vehicles "13 or more" amounts to it would be possible with this value of m to determine the probability of seven or more vehicles appearing in the first cycle. The next step is to find the mean value of "13 or more". Finding the arithmetical average requires extensive multiplication, but the mean value can be found very quickly. From the Poisson table it is found that the probability of:

13 or more vehicles appearing equals 0.0638
14 or more vehicles appearing equals 0.0342
15 or more vehicles appearing equals 0.0173
16 or more vehicles appearing equals 0.0082
17 or more vehicles appearing equals 0.0037
18 or more vehicles appearing equals 0.0016

19 or more vehicles appearing equals 0.0006 etc.

The mean of 0.0638 (the probability of 13 or more vehicles appearing) is 0.0319. According to the Poisson table above the number of vehicles corresponding to 0.0319 falls between 14 and 15. The values from the table above are plotted on semi-log paper in Figure 6.

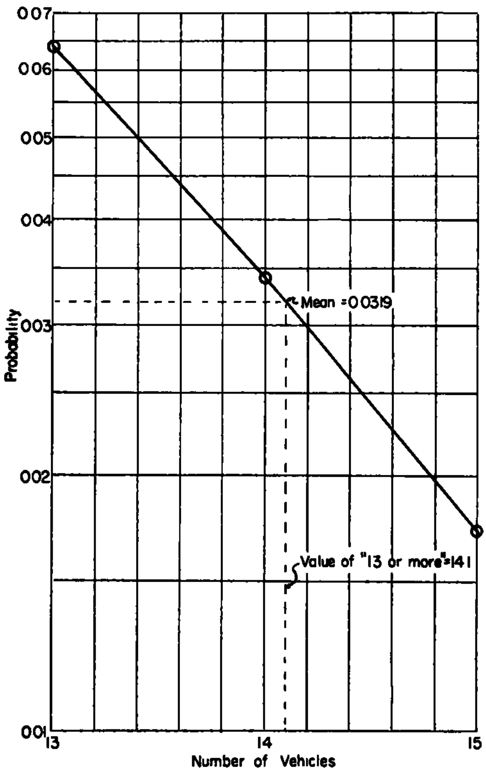

Figure 6. Volume = 360 Vehicles per Hour = 4 Vehicles per Cycle. Length of Cycle = 20 Sec Red; 20 Sec Green.

Note that the points fall on a nearly straight line. This fact makes it possible to interpolate between 14 and 15. The number of vehicles shown on the abscissa corresponding to 0.0319 is equal to approximately 14.1 which is the mean of "13 or more" for the two cycles or approximately 7.05 for one cycle. With this new m the probability of seven or more vehicles appearing in the first cycle is equal to 0.5704.

The probability of the two cycles failing is equal to the probability of there being 13 or more in the two cycles multiplied by the probability of there being seven or more in the first cycle or $0.0638 \times 0.5704 = 0.0360$. This may be compared with the correct value of 0.0377.

The probability of three cycles failing in succession would be equal to the probability of 19 or more vehicles appearing in three cycles times the probability of 13 or more in two cycles $\left(\text{with } m \text{ equal to } \frac{19}{3}\right)$, times the probability of seven or more in the first cycle.

In order to get a measure of the range of the variations of the approximate values from the true ones, let the average flow be increased from four to six vehicles for a 40 sec cycle, and assume that seven or more vehicles will cause a cycle to fail. Proceeding as above, the probability of two cycles failing in succession is found as follows:

Probability of 7 in 1st × probability of 6 or more in second $= 0.1377 \times 0.5543 = 0.07633$

Probability of 8 in 1st × probability of 5 or more in second $= 0.1032 \times 0.7149 = 0.07377$

Probability of 9 in 1st × probability of 4 or more in second $= 0.0688 \times 0.8488 = 0.05840$

Probability of 10 in 1st × probability of 3 or more in second $= 0.0413 \times 0.9380 = 0.03874$

Probability of 11 in 1st × probability of 2 or more in second $= 0.0225 \times 0.9826 = 0.02211$

Probability of 12 in 1st × probability of 1 or more in second $= 0.0113 \times 0.9975 = 0.01127$

Probability of 13 in 1st × probability of 0 or more in second $= 0.0052 \times 1.0000 = 0.00520$

Probability of 14 in 1st × probability of 0 or more in second $= 0.0022 \times 1.0000 = 0.00220$

Probability of 15 in 1st × probability of 0 or more in second $= 0.0009 \times 1.0000 = 0.00090$

Probability of 16 in 1st × probability of 0 or more in second $= 0.0003 \times 1.0000 = 0.00030$

Probability of 17 in 1st × probability of 0 or more in second $= 0.0001 \times 1.0000 = 0.00010$

Total probability = 0.2893

In the shorter method, using the Poisson table for an average flow of 12 vehicles for two cycles, it is found that the probability of "13 or more" vehicles appearing is equal to 0.4240. The mean value or one half of this is 0.2120. An examination of the table shows that the number of vehicles corresponding to this prob-

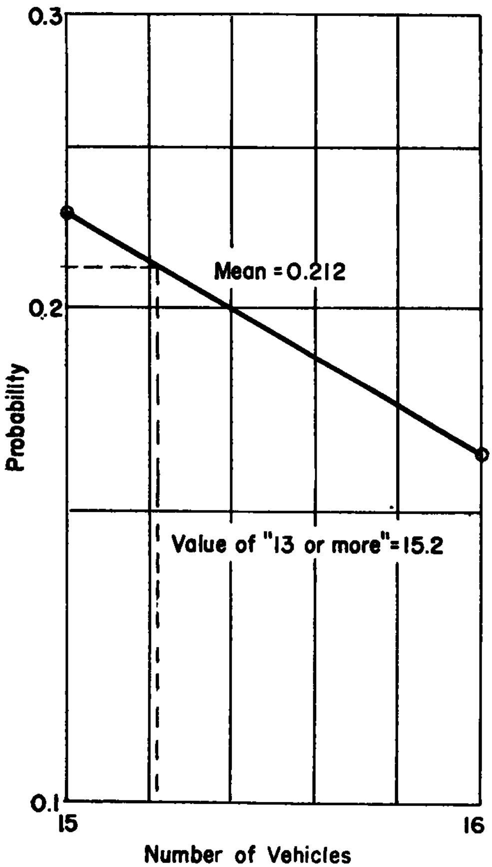

Figure 7. Volume = 450 Vehicles per Hour = 6 Vehicles per Cycle.

ability falls between 15 and 16. The mean value is found from the graphical interpolation to be 15.2 (Fig. 7). The mean flow for one cycle is therefore equal to 7.6 vehicles. With this value of m, the Poisson table gives 0.6354 as the probability of seven or more appearing in the first cycle. The probability of the failure of two successive cycles thus equals $0.6354 \times 0.4240 = 0.2694$. This may be compared with 0.2893, the probability obtained by the more exact method.

With it possible to calculate the number of vehicles retarded for different signal cycles and the amount of time they are delayed, it follows that the optimum cycle for any given volume of traffic may be determined.

The examples given show that although the theoretical approach outlined in this paper does not always lend itself readily to the solution of traffic problems, it may well serve as a spring-board to less exact but more practical procedures.

The results of such calculations could be combined into tables giving the optimum signal timing for any given volume of traffic. Before this is done, however. more field observations should be made to obtain more accurate constants of traffic performance.

No attempt has been made in this paper to include all the problems to which the methods outlined might be applied, or to carry all the suggested calculations to conclusion. Rather the paper has been presented with the thought that it might serve as an introduction to the use of the random series in analyzing traffic data.

Mr. E. C. Molina, in the introduction to the tables for "Poisson's Exponential Binomial Limit", says, "They are basic to techniques for inspection and sampling which have been developed in Bell Telephone Laboratories by its quality control engineers". It seems logical that the sampling methods found to work so well in industry should be applied to secure a like quality control in traffic. As new field studies yield new data to supply needed constants of performance, it is expected that many applications will be found for the Poisson series.

REFERENCES

1. Bruce D. Greenshields, Donald Shapiro, and Elroy L. Ericksen, "Traffic Performance at Urban Intersections" Bureau of Highway Traffic, Yale University, New Haven, Connecticut, (1947).
2. "Digest of Application of the Theory of Probability to Problems of Highway Traffic", *Proceedings*, Institute of Traffic Engineers, 118, (1934).
3. William F. Adams, "Road Traffic Considered as a Random Series", *Journal*, Institution of Civil Engineers, 121, November 1936.
4. Bruce D. Greenshields, "Initial Traffic Interference" presented (not published).

at Annual Meeting Highway Research Board, Washington, D. C., 1936. Mr. W. F. Adams, in a letter to the writer, July 28, 1937, commented on this paper and pointed out the applicability of the Poisson series in such a study.

5. O. K. Normann, "Results of Highway Capacity Studies", *Public Roads* Figure 18, Page 72, June 1942.
6. "Tables for Statisticians and Biometricians", edited by Carl Pearson, Cambridge University Press, page 76, (1930).
7. Edward Charles Molina, "Poisson's Exponential Binomial Limit", Table I, D. Van Nostrand Company, Inc., (1942). The tables by Molina give not only the individual values of the general term of Poisson's Exponential Expansion, but also the cumulated terms up to $m = 100$. Reference (1) contains tables for values of m to 15.
8. E. H. Holmes, "Effect of Control Methods on Traffic Flow", *Public Roads*, Vol. 14, No. 12, p. 233.

SOME CRITERIA FOR SCHEDULING MECHANICAL TRAFFIC COUNTS

By Boris B. Petroff[1]

Highway Economist, Public Roads Administration

SYNOPSIS

Traffic surveys will be needed to appraise the postwar use of rural roads. Most of the traffic counting work, particularly on low volume roads, is expected to be done with machines which do not record the traffic by the hour. Some criteria are needed for efficient scheduling of such counts. This study is only the beginning in a search for necessary information.

Practical considerations, indicated by experience, determine the minimum duration of a count to be 24 hr, or if longer, then in multiples of 24 hr. Coverage and control station counts are discussed. Coverage counts are defined as single observations which through the application of factors can be expanded to the annual average daily traffic. Control counts are defined as a system of observations from which expansion factors can be derived for application to coverage counts.

The prewar data from automatic traffic recorders at permanent locations were utilized. Ten stations with lowest volumes in the northern States and similarly ten stations in the southern States were selected. The application of the method of coefficients of variation which offer a measure of relative dispersion enabled the comparison of stations with different traffic volumes.

The study was divided into two parts. Part one deals with coverage counts and part two with control counts. Only weekdays were studied.

In part one it was sought to evaluate the length of a count in relation to the average weekday of each month. Counts of 24-, 48-, and 72-hr duration were studied. It was found that the range in terms of coefficients of variation is from 8.38 percent for the 72-hr counts in October in the southern States, to 32.73 percent for the 24-hr counts in March in the northern States. It was also found that October is one of the months when smallest coefficients of variation were observed and January is one of the months showing the largest. It was observed that the rate of increase in accuracy was higher when counts were prolonged from 24 to 48 hr than when the counts were prolonged from 48 to 72 hr.

In part two, various combinations of seasonally spaced weekday counts were compared with the annual average weekday. The number of all possible combinations is too great for use in this study. A preliminary analysis established certain limits to the study. In accordance with the results of the preliminary

[1] Data compiled and analyzed by: Alice W. Bradford, Elizabeth B. Hardesty, and Mary E. Kipp, Statisticians.