Statistical Analysis of Output Ratios in Traffic Simulation

A. V. GAFARIAN AND A. HALATI

Simulation models are increasingly becoming the most convenient tool for traffic studies. Users of such models need valid statistical methods to draw correct inferences. Presented in this paper is one such method applicable to several important traffic parameters. The motivation for this research arose from a study sponsored by the FHWA, U. S. Department of Transportation, to develop statistical guidelines for simulation experiments with traffic models. NETSIM, widely used for simulating vehicular traffic flow on urban streets, was used in the study. The output of the NETSIM model includes estimates of average speed, average delay per vehicle, and average travel time per vehicle mile. Because NETSIM uses the ratio of sample means to estimate these parameters, a situation exists that involves the ratios of observations that are in fact autocorrelated and cross correlated. In this paper, the efficacy of the ratio of sample means (used in NETSIM) as an estimator of the ratio of steady state means is discussed. Monte Carlo experiments have demonstrated that the user of the NETSIM model, in estimating these parameters from the model output, must apply statistical techniques based on ratio estimators. A technique that provides a measure of the accuracy of the estimate with a confidence interval is developed and demonstrated. The efficacy of the method is assessed through Monte Carlo experiments. The method is easy to use and can be applied just as readily to field data. It can be extended to the comparison of model outputs to field observations for simulation validation studies.

NETSIM is a widely accepted simulation tool for simulating traffic behavior on urban networks *(1, 2)*. The basic input requirements of the model are the network geometry, signalization information, and traffic counts, which consist of both input flow rates and turning movements. The standard output of the NETSIM model includes estimates of important traffic parameters such as

- Total vehicle minutes of travel time,
- Number of vehicles discharged,
- Total vehicle miles of travel distance,
- Average travel time per vehicle
- Average travel time per vehicle mile,
- Average speed, and
- Average delay time per vehicle.

The estimates of the traffic parameters are provided both on a link-by-link basis (links represent a one-way direction of flow on a street typically between two successive stop bars) and on a network basis.

A. V. Gafarian, Department of Industrial and Systems Engineering, University of Southern California, University Park, Los Angeles, Calif. 90089. A. Halati, California State Polytechnic University at Pomona, Calif. 91768.

It will be demonstrated that (a) each of the last four measures of effectiveness (MOEs) is a parameter that is the ratio of means of two random variables X and Y (X and Y are used generically here), that is, the MOE itself is μ_X/μ_Y, and (b) the natural estimate NETSIM provides is the ratio of the sample means of the X and Y random variables. This will be done in some detail for two parameters: average speed and average delay time per vehicle. The extension to two other parameters, average travel time per vehicle and average travel time per vehicle mile, will be obvious.

The discussion begins by noting that NETSIM is a stochastic microscopic traffic simulation model with a basic sampling interval of 1 second. Thus, the status of each individual vehicle is sampled at the rate of once every second and all required statistics are updated at the end of every second.

Example 1: Average Speed on a Link. In the simplest case, after the initial warm-up period, NETSIM produces for each 1-sec Δt time period the following observations on the totality of vehicles exiting the link during the time period under consideration: (a) number of vehicle miles in the link (equal to link length times the number exiting) and (b) vehicle minutes in the link (amount of time spent by all vehicles traversing the link). Running the model after warm-up for some integral multiple T of $\Delta t = 1$ provides the following as an estimate of average speed:

$$\sum_{j=1}^{T} (\text{vehicle miles in } j\text{th } \Delta t \text{ time period}) \div$$
$$\sum_{j=1}^{T} (\text{vehicle minutes in } j\text{th } \Delta t \text{ time period})$$

To understand why there is a ratio of two means $\overline{X}$ and $\overline{Y}$ that are estimates of μ_X and μ_Y, more work needs to be done. First, observe that as things stand now these observations, in both the numerator and the denominator, are not identically distributed. For example, because travel distance on a link in NETSIM is proportional to the number of vehicles discharged, during the red interval of the downstream signal the travel distance on the link will be accumulated at a low rate; during the early portion of the green interval, the travel distance will be accumulated at a large rate while the queue is dissipating.

Thus, there are observations on random variables that do not even have the same mean let alone the same distribution. Therefore, dividing the numerator and the denominator in the above expression by T does not give an estimate of the mean of any well-defined random variable.

It should be noted at the outset that it is important to deal with identically distributed observations because the problem of making valid statistical statements becomes tractable. To achieve identically distributed observations in both the numerator and the denominator of the ratio just given (so that dividing the numerator and the denominator by the number of observations gives estimates of the numerator and denominator means) is easy. All the observations for each $\Delta t = 1$ during one cycle of the link's downstream signal are summed. Thus, if the cycle length is 60 sec, then 60 sets of vehicle miles are added to produce one observation of vehicle miles. Likewise, the same is done with vehicle minutes in the denominator. A little reflection shows that these sums, from cycle to cycle, are certainly identically distributed after the warm-up. The remaining MOEs can be treated similarly.

Hereafter, it will be assumed that the collection interval will equal the downstream cycle length of each link. If there are two cycle lengths present in the network, for example, 60 and 90 sec, then running the model for (180)k seconds would provide 3k cycles' worth of observations for 60-sec links and 2k cycles' worth of observations for 90-sec cycle links.

To continue, take the above ratio, group the data as described, and end up with the following ratio for the estimate of average speed:

$$(X_1 + X_2 + \dots + X_n)/(Y_1 + Y_2 + \dots + Y_n)$$

where

X_i = accrued vehicle miles of vehicles departing during ith cycle, $i = 1,2, \dots, n$; and
Y_i = accrued vehicle minutes in the link of vehicles departing during ith cycle, $i = 1,2, \dots, n$.

Because the X_i's and Y_i's are identically distributed, this may be written

$$[(X_1 + X_2 + \dots + X_n)/n]/[(Y_1 + Y_2 + \dots + Y_n)/n] = \overline{X} / \overline{Y}$$

As $n \to \infty$, the numerator and denominator converge to μ_X and μ_Y, respectively, both with probability 1, where μ_X is the average vehicle miles per cycle and μ_Y is the average vehicle minutes per cycle. Thus, the problem of estimating link average speed is the same as estimating μ_X/μ_Y, the ratio of two means.

Example 2: Average Delay on a Link. Here again the following would be an estimate of average delay:

$$\sum_{j=1}^{T} \text{(accrued delay of vehicles departing the link during } j\text{th } \Delta t \text{ time period)} \div \sum_{j=1}^{T} \text{(number of vehicles departing the link during } j\text{th } \Delta t \text{ time period)}$$

As in Example 1, the same arguments could be used to get X_i's and Y_i's each identically distributed where

X_i = accrued delay of vehicles departing during the ith cycle, $i = 1,2, \dots, n$; and
Y_i = N_i = number of vehicles departing during the ith cycle, $i = 1,2, \dots, n$, (note that Y_i in this case is an integer-valued random variable).

and to produce an estimate of average delay $\sum_{i=1}^{n} X_i/\sum_{i=1}^{n} N_i = \overline{X} / \overline{N}$, which converges with probability 1 to μ_X/μ_N as $n \to \infty$. So again the ratio of two means is estimated. In this case, μ_X is the average delay per cycle and μ_N is the average number of vehicles discharged per cycle.

The principal objective of this paper is to develop a statistically valid method for using $\overline{X} / \overline{Y}$ as a point estimate for μ_X/μ_Y and to provide, with a confidence interval, a measure of its accuracy. What X and Y are depends on the particular MOE being estimated.

STATISTICAL PROPERTIES OF THE OBSERVATIONS

In this section, some important statistical properties of the observations will be described.

Observations Tend To Be Normal

This property follows from the fact they are sums of random variables, obtained by adding up all the individual observations for each Δt. Thus, the Central Limit Theorem, which holds for fairly unrestrictive conditions (even when the variables being added are not identically distributed or independent), comes into play and it can be stated that asymptotic normality is obtained. This includes integer-valued observations, such as the number of vehicles discharged during a cycle length.

In this connection, it should be mentioned that the method developed in the paper is based on the t-statistic and that this statistic is robust with respect to normality; that is, inferences using it are not seriously invalidated by the violation of the normality assumption. This will be demonstrated in the Monte Carlo experiment presented later in the paper.

Observations Are Not Independent

The observations of travel time and travel distance, for example, are each autocorrelated. Figure 1 shows estimates of autocorrelation for travel time on a link of a simple star network consisting of essentially an isolated four-legged intersection with pretimed signal control. The simulation run consisted of 130 cycles of a common signal cycle length of 80 sec (i.e., 10,400 sec). Estimates of autocorrelation

$$r_{XX}(k) = c_{XX}(k)/c_{XX}(0) \quad k = 0, 1, \dots, n/10$$

were obtained by using

$$c_{XX}(k) = (1/n)\sum_{i=1}^{n-k} (X_i - \overline{X})(X_{i+k} - \overline{X}) \quad k = 0, 1, \dots, n/10$$

where

$r_{XX}(k)$ = sample autocorrelation of the X series for lag k,
$c_{XX}(k)$ = sample autocovariance of the X series for lag k,
n = number of cycles that made up the simulation run, in this case 130, and

X_i = observed value of total travel time during cycle i.

The maximum lag was restricted to $n/10$ to obtain accurate estimates of the autocorrelations.

If it is assumed that travel time observations are independently, identically, and normally distributed random variables, the standard deviation of the autocorrelation estimates are approximately equal to $\sqrt{1/n}$, (*3*, pp.34, 35). In this case, the standard deviation would be approximately .0877. Because the estimate of the first lag autocorrelation is .321, almost four times the standard deviation, it can be concluded that the first lag correlation is not zero. Moreover, there is strong indication that there is autocorrelation up to lag 10. Thus, it is reasonable to assume that successive travel times are autocorrelated.

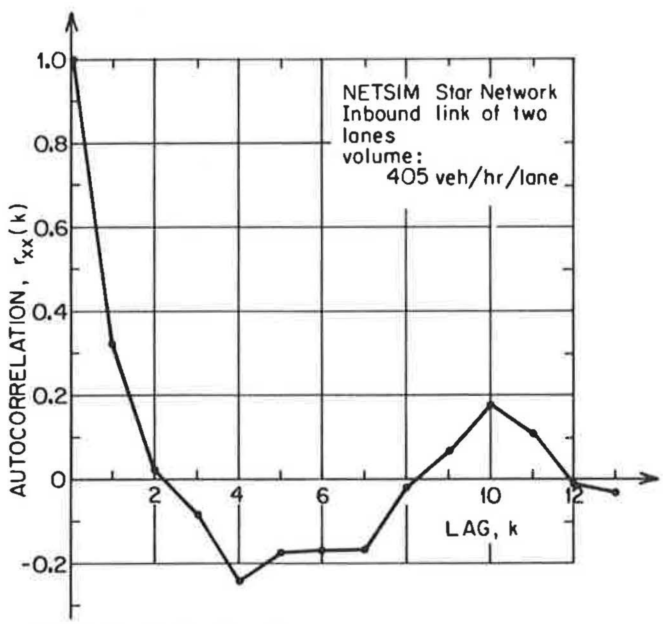

FIGURE 1 Estimates of autocorrelation for travel time on a link of a simple star network.

Observations of Random Variables of the Numerator and Denominator Are Cross Correlated

Figure 2, for example, shows the cross-correlation estimates between the total travel time per cycle and the total travel distance per cycle. Estimates of cross correlation

$$r_{XY}(k) = c_{XY}(k)/c_{XX}(0)c_{YY}(0) \quad k = 0, \pm1, \ldots, \pm(n/10)$$

were obtained using

$$c_{XY}(k) = (1/n) \sum_{i=1}^{n-k} (X_i - \bar{X})(Y_{i+k} - \bar{Y}) \quad k = 0, \quad 1, \ldots, n/10$$

$$= (1/n) \sum_{i=1}^{n+k} (X_{i-k} - \bar{X})(Y_i - \bar{Y}) \quad k = -1, -2, \ldots, n/10$$

where

$r_{XY}(k)$ = sample cross correlation of the X and Y series for lag k,

$c_{XY}(k)$ = sample cross covariance of the X (travel time) and Y (travel distance) series for lag k,

n = number of cycles that made up the simulation run, in this case 130,

X_i = travel time during cycle i, and

Y_i = travel distance during cycle i.

Again the maximum lag was restricted to $\pm n/10$ to obtain accurate estimates of the cross correlations.

Note that the cross covariance has both positive lags (where Y leads X) and negative lags (where X leads Y); and that, in general, $c_{XY}(k) \neq c_{XY}(-k)$. This is not the case for the autocovariance, where $c_{XX}(k) = c_{XX}(-k)$ [or $c_{YY}(k) = c_{YY}(-k)$]. The large cross-correlation estimate of lag 0 is expected because a large observation for travel distance indicates that a large number of vehicles have traversed the link and thus a large value of travel time has been incurred. In addition, significant cross correlation at larger lags is also observed.

PROBLEM DEFINITION

There are two common methods for performing simulation experiments, and the problem will be defined for each of these methods.

Method 1: A Single Long Run

The first method consists of running the simulation model for a long duration and using the observations generated in this single, continuous, long run to estimate the parameters of interest and to obtain a measure of the accuracy of the estimate.

In the case of the NETSIM model, as it pertains to the traffic parameters that are estimated as the ratio of two random variables (which in this case happens to be means), it was demonstrated in the preceding section that successive observations obtained on the random variables (at the end of each cycle) are autocorrelated and cross correlated. In the presence of these correlations, estimating the parameters from a single contin-

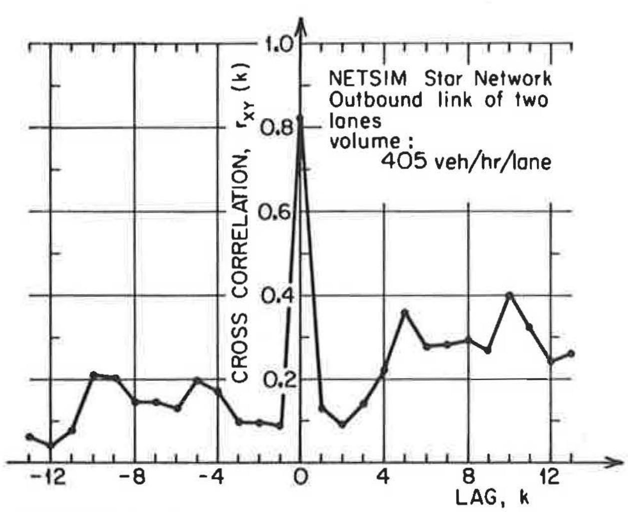

FIGURE 2 Cross-correlation estimates between total travel time per cycle and total travel distance per cycle.

uous run of the NETSIM model may be defined as the following statistical problem.

Let (X,Y) be a bivariate random variable and suppose a sequence of identically distributed observations $[(X_1,Y_1), (X_2,Y_2), \ldots, (X_n,Y_n)]$ has been obtained. The X_i's and Y_i's correspond to the numerator and denominator observations, respectively. Furthermore, assume the following:

$$E[X] = \mu_X \qquad \mathrm{Var}[X] = \sigma_X^2$$

$$E[Y] = \mu_Y \qquad \mathrm{Var}[Y] = \sigma_Y^2$$

$$E[(X_i - \mu_X)(X_{i+k} - \mu_X)]/\sigma_X^2 \neq 0 \qquad \text{for all } k$$

(observations are not independent);

$$E[(Y_i - \mu_Y)(Y_{i+k} - \mu_Y)]/\sigma_Y^2 \neq 0 \qquad \text{for all } k$$

(observations are not independent);

$$E[(X_i - \mu_X)(Y_{i+k} - \mu_Y)]/\sigma_X\sigma_Y \neq 0 \qquad k \geq 0$$

(observations of random variables of the numerator and denominator are cross correlated);

$$E[(X_{i-k} - \mu_X)(Y_i - \mu_Y)]/\sigma_X\sigma_Y \neq 0 \qquad k < 0$$

(observations of random variables of the numerator and denominator are cross correlated).

The problem then becomes that of using this information to estimate $R = \mu_X/\mu_Y$ and of assessing the accuracy of the estimate by constructing a confidence interval.

The two examples of this generic problem discussed in the introduction were average link speed and mean delay:

1. Average link speed. Here the point estimate is $\overline{X} / \overline{Y}$, where X_i is accrued vehicle miles of vehicles departing during the ith cycle and Y_i is accrued vehicle minutes in the link of departing vehicles during the ith cycle, $i = 1,2, \ldots, n$.
2. Mean delay. Here the point estimate is $\overline{X} / \overline{N}$, where X_i is accrued delay of vehicles departing the link during the ith cycle and N_i is the number of vehicles departing during the ith cycle, $i = 1,2, \ldots, n$.

The problem of using the observations from a single run and developing a confidence interval for μ_X/μ_Y (average speed in the first example) or μ_X/μ_N (mean delay in the second example) is extremely complex and involves estimating autocorrelations of the two numerator and denominator variables and the cross correlations of the numerator variables with the denominator variables [see Halati (*4*, pp.65–69)]. This requires an extremely long run to get reliable estimates of all the needed correlations, as well as to reduce the inherent bias present in the estimate $\overline{X} / \overline{Y}$ [Halati (*4*, p.63)].

In addition, the use of the method is predicated on collecting observations that are identically distributed. It was noted that identically distributed observations may be obtained by summing the statistics, collected at the end of each 1-sec sampling interval, over the period of one cycle length. This is obviously applicable if the network consists solely of pretimed controls. When actuated controls are present, there is no immediate alternative for producing identically distributed observations.

The method of independent replications, which will be discussed next, does not have this disadvantage. Each replication will result in a single observation that is the sum of the statistics over the duration of each run. The notion of the cycle will not be needed.

Method 2: Several Independent Replications

The second method for conducting simulation experiments is to perform independent replications. In this method, repeated runs of the model are performed in such a way that the output of the model in each run is independent of the others by using a different random generator seed in each run.

In this method, the great difficulty of getting reliable estimates of all the autocorrelations and cross correlation among successive observations is circumvented. However, the method has the disadvantage of requiring a warm-up time for each replication during which no data may be collected.

In the case of the NETSIM model, and again as it pertains to the analysis of those traffic parameters that are the ratio of two random variables, the problem of estimating these parameters and assessing the accuracy of the estimate may now be defined statistically in the following way.

Suppose n independent replications of the NETSIM model are performed and also assume that each run has a prescribed duration of k cycles. In this case the observations, (X_i, Y_i), $i = 1,2, \ldots, n$, would be the cumulative values of the observed statistics at the end of each run and

- X_i, $i = 1,2, \ldots, n$, would be a sequence of IID observations,
- Y_i, $i = 1,2, \ldots, n$, would be a sequence of IID observations,
- $E[X_i] = k\mu_X$ and $E[Y_i] = k\mu_Y$ because now X_i and Y_i are the cumulative values obtained by adding the statistics over k cycles, and
- The only cross correlation present is between X_i and Y_i.

The problem in this form is that of using the cumulative statistics $\{(X_i,Y_i), i = 1,2, \ldots, n\}$ to estimate $R = \mu_X/\mu_Y$ and assess the accuracy of the estimate.

Two points should be noted here. The first point is that $\sum_{i=1}^n X_i/\sum_{i=1}^n Y_i$ is still an estimator of R even though the X_i's and Y_i's are cumulative values when independent replications are performed. This is because, with probability 1,

$$\lim_{n\to\infty} \left(\sum_{i=1}^{n} X_i / \sum_{i=1}^{n} Y_i \right) = k\mu_X/k\mu_Y = R$$

The second point is that the normality assumption discussed in the section on statistical properties of the observations becomes better. This is because each X_i and Y_i is now the sum of a larger number of observations.

Thus, in the first example above

X_i = accrued travel distance of vehicles departing the link during the ith replication, $i = 1,2, \ldots, n$, and

Y_i = accrued vehicle minutes in the link of vehicles departing the link during the ith replication, $i = 1,2, \ldots, n$

and for an estimate of average speed for the link one would take

$$\sum_{i=1}^{n} X_i / \sum_{i=1}^{n} Y_i = \bar{X} / \bar{Y}$$

In the second example,

X_i = total accrued delay of the ith replication, $i = 1,2, \ldots, n$, and

Y_i = N_i = total number of vehicles departing the link during the ith replication, $i = 1,2, \ldots, n$, and for an estimate of mean delay one would take

$$\sum_{i=1}^{n} X_i / \sum_{i=1}^{n} N_i = \bar{X} / \bar{N}$$

To reiterate the important points with respect to independent replications, there are no autocorrelations among the X_i's or the Y_i's and there is only a cross correlation between X_i and Y_i. It is the problem in this form that will be studied in this paper.

PERFORMANCE DEGRADATION WHEN OBSERVATIONS FROM INDEPENDENT REPLICATIONS ARE NOT TREATED AS A RATIO OF RANDOM VARIABLES

Before proceeding with the development of the procedure for computing a confidence interval for μ_X/μ_Y by using the point estimator $\sum_{i=1}^{n} X_i/\sum_{i=1}^{n} Y_i = \bar{X} / \bar{Y}$, consider an important question of the degradation that occurs when the problem is not treated as one in ratio estimation. The reason for doing this is that one might easily be tempted to develop a confidence interval for, say, mean speed from independent replications by a method that goes as follows. Because each replication gives an independent estimate of mean speed $Z_i = X_i/Y_i$, $i = 1,2, \ldots, n$, one would estimate mean speed as

$$\bar{Z} = (1/n) \sum_{i=1}^{n} Z_i$$

and assess the accuracy of this estimate with a $(1 - \alpha) \times 100$ percent classical confidence interval

$$\left\{ \left[\bar{Z} - t_{1-(\alpha/2),(n-1)} \right] S/(n)^{1/2}, \left[\bar{Z} + t_{1-(\alpha/2),(n-1)} \right] S/(n)^{1/2} \right\}$$

where $t_{1-(\alpha/2),(n-1)}$ is the upper $\alpha/2$ point of the t-statistic with $n - 1$ degrees of freedom and

$$S = [1/(n-1)] \sum_{i=1}^{n} (Z_i - \bar{Z})^2$$

In effect, the observations from the model and field were treated as just described in the validation studies performed on NETSIM (2, pp.147–248). The thing that is wrong with this procedure is that one is really estimating (or doing hypothesis testing) on $E[X/Y]$ and not μ_X/μ_Y. It is well known that in general $E[X/Y] \neq \mu_X/\mu_Y$. To demonstrate that this approach to the analysis of ratio estimators of the NETSIM model may produce results that are greatly in error, a Monte Carlo study was conducted. The study consisted of generating bivariate normal random variables (X_i, Y_i) for selected sample sizes n with the variance-covariance matrix

$$\begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$$

X_i and Y_i correspond to the numerator and the denominator observations respectively of the ith replication and n corresponds to the number of replications. μ_X was chosen as 100 and μ_Y as 5. Thus, the known and true value of the ratio of means was 20.

For each sample size n and using the above procedure, a 95 percent confidence interval was constructed. To assess the goodness of the confidence interval, the experiment was repeated 500 times for each sample size and the following four measures of effectiveness on the behavior of the constructed confidence intervals were computed:

1. Coverage probability. This measure of effectiveness is the fraction of the confidence interval produced in the 500 repetitions of the experiment that covered the true value of the ratio of the means, which was 20. Closeness of this value to .95 is obviously a desired property of the method.
2. Coefficient of variation of coverage probability. This statistic is the ratio of the standard deviation of the estimate of coverage probability to the estimated coverage probability. It is a measure of how good the estimate of coverage probability is—the smaller the value the better the estimate. Thus, .010 for $n = 5$ means that the standard deviation of the estimate is only 1 percent of the estimate.
3. Average confidence interval length. In each repetition of the experiment, the length of the constructed confidence interval was recorded. This statistic represents the average of the recorded confidence interval lengths over the 500 repetitions. Obviously, the smaller the length the better.
4. Coefficient of variation of expected confidence interval length. This is the ratio of the standard deviation of the estimate of the average confidence interval length to the average confidence interval length. Again, it is a measure of how good the estimate of average confidence interval length is. For $n = 5$, .023 means that the standard deviation of the estimate is about 2.3 percent of the estimate.

The study was conducted for replication sizes of 5,6,7,8,9,10,20,50,100, and 200 observations per replication. The results are given in Table 1.

To obtain a basis for comparison of the results, the study was repeated identically by using the proposed method (to be developed in the next section) of analysis. The results of that study are given in Table 2.

It is noted that the coverage probabilities are substantially reduced by using the incorrect method of analysis described at the beginning of this section. It is seen here that as the number

TABLE 1 RESULTS OBTAINED FROM MONTE CARLO EXPERIMENTS USING INCORRECT METHOD

No. of Replications	Coverage Probability	Coefficient of Variation of Coverage Probability	Avg Confidence Interval Length	Coefficient of Variation of Avg Confidence Interval Length
5	.950	.010	10.48	.023
6	.936	.009	9.04	.022
7	.938	.011	7.80	.021
8	.946	.010	7.48	.022
9	.934	.011	6.82	.020
10	.936	.011	6.48	.020
20	.904	.013	4.34	.014
50	.754	.019	2.70	.010
100	.574	.022	1.92	.009
200	.206	.018	1.32	.006

of replications increases smaller confidence intervals result, and they begin to miss μ_X/μ_Y in increasing numbers {because they are really covering $E[X/Y] \neq (\mu_X/\mu_Y)$}. However, the proposed method of analysis, which is based on a ratio estimation technique, produced coverage probabilities close to the desired 95 percent percent for all sample sizes. At smaller replication sizes, when both methods appear to have coverage probabilities close to the desired .95, the proposed method consistently resulted in more precise (lower coefficient of variation) and smaller average confidence interval lengths.

PROPOSED METHOD

How a confidence interval may be developed for μ_X/μ_Y based on observations $[(X_i, Y_i), i = 1,2, \ldots, n]$ obtained from each independent replication of the model will now be considered. Keep in mind as an example that X_i is total travel distance in vehicle-miles and Y_i is total travel time in vehicle-minutes (or X_i could be total delay and $Y_i = N_i$ total number of vehicles discharged). The method is a small sample extension of the Fieller method *(5)*.

Suppose a confidence interval is wanted for $R = \mu_X/\mu_Y$. Then the estimator would be

$$\hat{R} = \sum_{i=1}^{n} X_i / \sum_{i=1}^{n} Y_i = (1/n) \sum_{i=1}^{n} X_i / (1/n) \sum_{i=1}^{n} Y_i = \bar{X} / \bar{Y}$$

Next, a new variable is considered:

$$Z_i = X_i - RY_i$$

and then

$$\bar{Z} = \bar{X} - R\,\bar{Y}$$

If it is assumed that X_i and Y_i are normally distributed (this assumption has been discussed in the section on statistical properties of the observations), then Z_i and $\bar{Z}$ will be normally distributed. Because $E\lfloor Z_i \rfloor = E\lfloor \bar{Z} \rfloor = 0$,

$$\bar{Z} / \left\{ (1/n) \bullet [1/(n-1)] \sum_{i=1}^{n} (Z_i - \bar{Z})^2 \right\}^{1/2}$$

TABLE 2 RESULTS OBTAINED FROM MONTE CARLO EXPERIMENTS USING PROPOSED METHOD

No. of Replications	Coverage Probability	Coefficient of Variation of Coverage Probability	Avg Confidence Interval Length	Coefficient of Variation of Avg Confidence Interval Length
5	.940	.011	10.48	.023
6	.928	.012	8.54	.019
7	.938	.011	7.18	.016
8	.944	.010	6.68	.016
9	.936	.011	6.06	.014
10	.942	.010	5.70	.013
20	.942	.010	3.70	.009
50	.960	.009	2.24	.005
100	.968	.008	1.56	.004
200	.938	.011	1.08	.003

has a Student t-distribution with $n - 1$ degrees of freedom. But

$$[1/(n-1)] \sum_{i=1}^{n} (Z_i - \overline{Z})^2 = [1/(n-1)] \sum_{i=1}^{n} (X_i - RY_i - \overline{X} + R\,\overline{Y})^2$$

$$= [1/(n-1)] \sum_{i=1}^{n} [(X_i - \overline{X}) - R(Y_i - \overline{Y})]^2$$

$$= [1/(n-1)] \sum_{i=1}^{n} [(X_i - \overline{X})^2 + R^2(Y_i - \overline{Y})^2 - 2R(X_i - \overline{X})(Y_i - \overline{Y})]$$

$$= (S_X^2 + R^2 S_Y^2 - 2RS_{XY})$$

where

$$S_X^2 = [1/(n-1)] \sum_{i=1}^{n} (X_i - \overline{X})^2 = \text{sample variance of the } X_i\text{'s}$$

$$S_Y^2 = [1/(n-1)] \sum_{i=1}^{n} (Y_i - \overline{Y})^2 = \text{sample variance of the } Y_i\text{'s}$$

$$S_{XY} = [1/(n-1)] \sum_{i=1}^{n} (X_i - \overline{X})(Y_i - \overline{Y}) = \text{sample covariance of the } (X_i, Y_i)\text{'s}$$

Hence at the $(1 - \alpha)$ level,

$$Pr\left(\left\{(\overline{X} - R\,\overline{Y}) / \left[(1/n)(S_X^2 + R^2 S_Y^2 - 2RS_{XY})\right]^{1/2}\right\} \le t_{1-(\alpha/2),n-1}\right) = 1 - \alpha$$

Note that when both sides of the argument in the above probability statement are squared, the result is a quadratic inequality in the unknown $R = \mu_X/\mu_Y$, the known estimates $\overline{X}$, $\overline{Y}$, S_X^2, S_Y^2, S_{XY}, and $t_{1-(\alpha/2),(n-1)}$. The roots of this quadratic inequality are

$$\left[[\,\overline{X}\,\overline{Y} - g(\alpha)S_{XY}] \pm \left([\,\overline{X}\,\overline{Y} - g(\alpha)S_{XY}]^2 - \{[\,\overline{Y}^2 - g(\alpha)S_Y^2][\,\overline{X}^2 - g(\alpha)S_X^2]\}\right)^{1/2}\right] / [\,\overline{Y}^2 - g(\alpha)S_Y^2]$$

where

$$g(\alpha) = t_{1-\alpha/2;\, n-1}^2/n$$

The $(1 - \alpha) \times 100$ percent confidence interval is then (r_1, r_2), where r_1 is the smaller root and r_2 is the larger root. It should be noted that $\overline{X} / \overline{Y}$ is not the midpoint of the confidence interval.

The efficacy of the proposed method in any application depends on how well the assumptions are met; namely, system in steady state, normality of the numerator and the denominator observations, and independent replications.

Steady state in NETSIM is achieved by a warm-up procedure that appears to work. Independent replications are achieved by starting each run with a different random number generator seed. The only assumption that is approximately met in applications is that of normality, and it is claimed that the method is not sensitive to this requirement.

To demonstrate the method's robustness to this assumption, it was applied to the M/M/1 queuing system to estimate mean delay (excluding service time) per customer. This is a single-server system with exponential interarrival and service times. This system was selected because it is known that delay in queue is extremely nonnormal.

The arrival rate was taken to be $\lambda = 36$ customers per hour and the service rate $\mu = 40$ customers per hour for a traffic intensity $\rho = \lambda/\mu = 36/40 = 0.90$. Each replication was started in steady state. This was accomplished by having an initial number of customers in the system obtained from sampling the steady state probability mass function given by $p_n = (1 - \rho)\,\rho^n$, $n = 0,1,2, \ldots$.

Each replication consisted of 2 hours of simulated time. The numerator and denominator recordings were, respectively, X_i = accumulated delay of departing customers, and $Y_i = N_i$ = number of departing customers, $i = 1,2, \ldots, n$. Thus, the result is a ratio estimation situation in which $\sum_{i=1}^{n} X_i / \sum_{i=1}^{n} N_i$ converges as $n \to \infty$ to μ_X/μ_N, where μ_X is the expected total delay per unit time and μ_N is the expected number of departures per unit time.

The number of replications was selected to be 5,6,7,8,9,10,20, and 40. For each replication size, the experiment was repeated 500 times and for each repetition a 95 percent confidence interval was constructed by the proposed method. Because for this system the true value of the steady state mean delay is known to be $\lambda/[\mu(\mu - \lambda)]$ (in this case 13.5 min), an estimate of coverage probability is the fraction of the 500 repetitions, which cover the true value. Also, estimates of the average confidence interval length and of the coefficients of variation of both coverage probability and average confidence interval length were computed and the results are given in Table 3.

The experiment was then repeated with an arrival rate of λ = 32 customers per hour and a service rate of $\mu = 40$ customers per hour for a traffic intensity of $\rho = \lambda/\mu = 32/40 = 0.80$. These results are given in Table 4.

Looking at the data in these two tables, some degradation of coverage probability due to the extreme nonnormality of the data can be seen, but the results are not bad. The worst case is an 81.4 percent coverage when 95 percent was expected, for n = 5 and $\rho = 0.8$. However, it improves rapidly and for 40 replications it is up 93.6 percent. The situation is considerably better for $\rho = 0.9$. It should be noted that the problem of nonnormality can be ameliorated by making the replications longer.

SUMMARY AND CONCLUSIONS

In this paper a number of important traffic parameters were identified as being ratio of means of two random variables. These parameters may be estimated by the ratio of sample means. In particular, it was noted that NETSIM uses this type of estimator. It was demonstrated that, in general, the numerator and denominator random variables that comprise the sample means are autocorrelated and cross correlated. Therefore,

TABLE 3 M/M/1 QUEUING SYSTEM, AVERAGE DELAY PER CUSTOMER (ρ = 0.9)

No. of Replications	Coverage Probability	Coefficient of Variation of Coverage Probability	Avg Confidence Interval Length	Coefficient of Variation of Avg Confidence Interval Length
5	.882	.016	29.86	.030
6	.876	.017	26.90	.028
7	.860	.018	23.18	.027
8	.888	.016	21.94	.023
9	.864	.018	18.68	.023
10	.874	.017	18.36	.023
20	.920	.013	12.96	.017
40	.950	.010	8.66	.011

TABLE 4 M/M/1 QUEUING SYSTEM, AVERAGE DELAY PER CUSTOMER (ρ = 0.8)

No. of Replications	Coverage Probability	Coefficient of Variation of Coverage Probability	Avg Confidence Interval Length	Coefficient of Variation of Avg Confidence Interval Length
5	.814	.021	12.06	.036
6	.874	.017	10.02	.030
7	.856	.018	9.80	.032
8	.870	.017	7.72	.028
9	.892	.016	8.04	.025
10	.872	.017	7.34	.025
20	.924	.013	5.08	.025
40	.936	.012	2.74	.014

obtaining estimates from a single continuous set of observations and assessing the accuracy of the estimate by a confidence interval is an extremely complicated statistical problem.

Because the method of independent replications simplifies the problem considerably, a method based on it was developed and its efficacy was demonstrated through Monte Carlo experiments. The method may be applied to the estimation of parameters and assessment of the accuracy of the estimates from field data, to the analysis of traffic simulation outputs, and to the comparison of field data with simulated data for validation studies.

REFERENCES

1. E. Lieberman, R. D. Worrall, D. Wicks, and J. Woo. *NETSIM Model*, Vol. 4: User's Guide. Report FHWA-RD-77-44. FHWA, U. S. Department of Transportation, Oct. 1977.
2. *Network Flow Simulation for Urban Traffic Control System.* General Applied Science Laboratories, Inc., Final Report FH-11-7462-2. Peat, Marwick, Mitchell and Co., Washington, D. C., 1972.
3. G. E. P. Box and G. M. Jenkins. *Time Series Analysis—Forecasting and Control.* Holden-Day, San Francisco, 1970.
4. A. Halati. *Estimation of the Ratio of the Steady-State Means of a Bivariate Stochastic Process.* Ph. D. thesis. University of Southern California, Los Angeles. Jan. 1985.
5. E. C. Fieller. "Some Problems in Interval Estimation." In *Journal of the Royal Statistical Society*, Series B, No. 2, 1954, pp. 175–185.

Publication of this paper sponsored by Committee on Traffic Flow Theory and Characteristics.