

DEPARTMENT OF TRAFFIC AND OPERATIONS

Traffic Inputs for Simulation on a Digital Computer

D. L. GERLOUGH,
*Institute of Transportation and Traffic Engineering,
University of California, Los Angeles*

The relative merits of observed inputs and generated inputs are discussed. It is pointed out that where generated inputs can be used, they have a decided advantage from the standpoint of computer time.

A method of generating Poisson inputs is described. Several sets of inputs have been generated by this method and tested statistically. The results of these are given.

Time spacings of vehicles entering a system are approximately exponential. Methods of adjusting the exponential to obtain greater realism are discussed. Methods of computer generation are described.

• IN SIMULATING the flow of traffic by means of a digital computer, one problem which arises is the generation of traffic inputs to the simulation. If empirical data are available, a straightforward method is to feed these data to the computer at appropriate times as simulation inputs. In most computers, however, the operation time of input units is many times greater than the computational time of the computer, so that whenever feasible, time should be saved by the generation of inputs within the computer itself. Inasmuch as most traffic events are random in nature, this generation consists of producing random inputs in accordance with probability distributions which have been found adequate to describe traffic phenomena.

It is the purpose of this paper to discuss some methods of accomplishing this artificial generation of traffic phenomena and to compare the results with observed traffic situations. The procedure will be first to describe methods for the generation of any random events and then to relate these techniques to problems of vehicular traffic flow.

COMPUTER GENERATION OF RANDOM PHENOMENA

The generation of random phenomena by a digital computer requires the cumulative probability distribution of the ap-

propriate form. The technique may be illustrated with the generalized probability distribution shown in Figure 1. Using a random fraction which has been previously generated, the distribution is entered along the probability axis and the corresponding value obtained from the variable axis. Thus, the generation of random phenomena may be divided into two parts: generation of random fractions (having a uniform distribu-

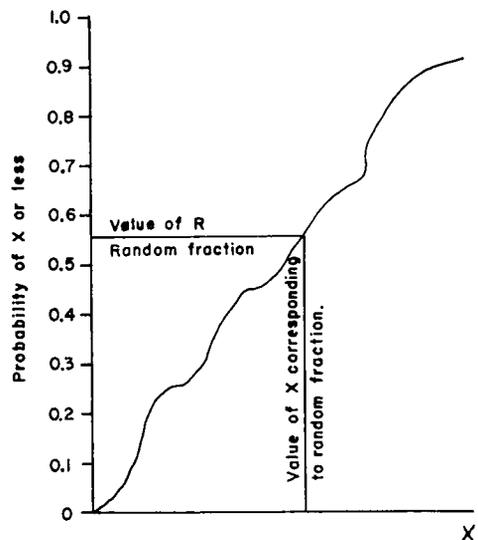


Figure 1. Generalized cumulative probability distribution.

tion), and conversion to random deviates having the desired distribution.

GENERATION OF RANDOM NUMBERS

Several investigators have studied the generation of pseudo-random numbers by computers and have found that the simplest and most reliable method is as follows (1). An assumed starting number is multiplied by an appropriate multiplier, and the low order or less significant half of the product is taken as the random number. The second random number is formed using the first as a starting number and the same multiplier as before, etc. This technique is usually stated as

$$R_m = \rho R_{m-1} \text{ Mod } b^n \quad (1)$$

R_m = m th random number;

ρ = multiplier;

n = number of digits in a normal word on the particular computer used;

b = number base of computer;

Mod b^n = instruction to use only the low order or less significant half of the full ($2n$ -digit) product (the remainder after dividing the product by b^n the maximum integral number of times); and

R_0 = any odd number selected as a starting number.

The multiplier ρ may be selected by taking a base which is prime relative to the number base of the computer and raising it to the highest power which can be held by one word of the computer. Table 1 gives a few appropriate multipliers.

The numbers resulting from this generation technique form a series of

TABLE 1

MULTIPLIERS FOR RANDOM NUMBER GENERATION

Number Base and Capacity of Computer, b^n	Random Number Multiplier, ρ
10 ¹⁰	7 ¹¹
10 ⁸	7 ⁹
2 ²⁵	5 ¹⁸
2 ³⁶	5 ¹⁸

pseudo-random numbers; that is, the numbers are generated in a non-random manner, but behave as though they were random. Tests by several investigators indicate no evidence that the numbers are non-random. Random numbers so generated may be interpreted either as random integers with the point at the extreme right, or as random fractions with the point at the extreme left. If the random numbers are interpreted as fractions, the result is a rectangular or uniform probability distribution.

CONVERSION TO DESIRED DISTRIBUTION

Conversion of random fractions to random deviates of the desired form through the use of the cumulative probability distribution can be accomplished in a variety of ways. If the cumulative probability distribution is a continuous function which can be represented by an equation, the operation of inserting the random fraction and obtaining the random deviate is purely a matter of calculation. (In certain specialized cases it is faster, and may even be more accurate, to obtain the random deviates in an indirect trial-and-error, Monte Carlo, method. Such a technique is used in the routine of Appendix C for generation of exponential arrivals.)

If the cumulative probability distribution is not represented by an equation but by tabular data, the generation of the random deviates may be performed by a form of table-lookup operation. In Figure 2 tabular values of a cumulative distribution are indicated by points. The distribution for discrete values of the random deviate, x , is shown as a broken line. The random fraction is compared with the ordinates of the various points until the first point satisfying the following condition is found:

$$R \leq P_i \quad (2)$$

in which

R = random fraction; and
 P_i = the probability (ordinate) value of the i th point.

The value of x for the i th point is taken as the desired deviate.

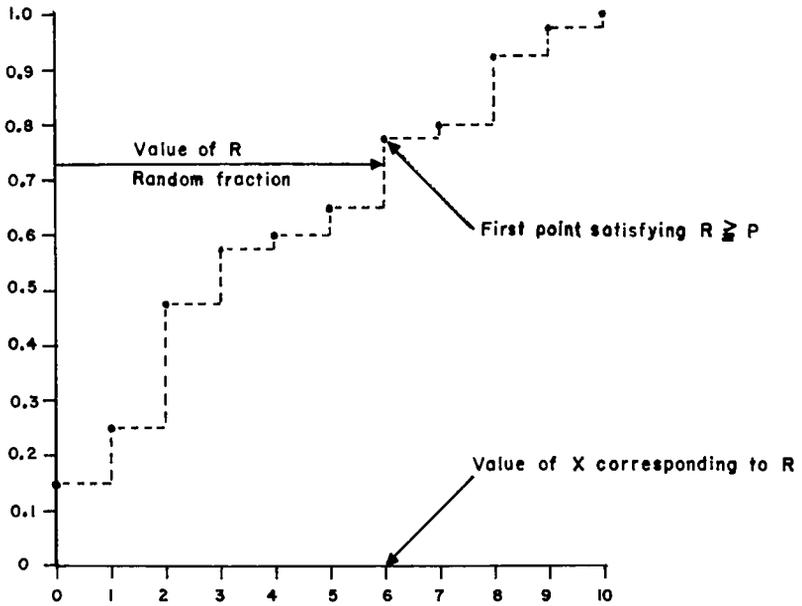


Figure 2. Cumulative probability distribution. Points represent tabular values. Broken line represents distribution for discrete values of X .

If the random deviate x is a continuous distribution, the tabular values may be interpolated along straight-line segments between the points, obtaining a unique value of x for each value of random fraction.

POISSON DISTRIBUTION

The usefulness of the Poisson distribution for describing various phenomena, including traffic, has been discussed elsewhere (2, 3, 4). The Poisson distribution is expressed as

$$p(x) = \frac{m^x e^{-m}}{x!} \tag{3}$$

in which

x = number of vehicles arriving during an interval of time of specified length T . Thus x takes on the values 0, 1, 2, . . . ;

m = average number of vehicles arriving during intervals of length T ;

$p(x)$ = probability of occurrence of x ;
and

e = natural base of logarithms.

The cumulative Poisson distribution is expressed:

$$P(x) = \sum_{i=0}^x \frac{m^i e^{-m}}{i!} \tag{4}$$

where $P(x)$ = probability of x or fewer during interval T .

The generation of random deviates (arrivals, for instance) obeying the Poisson distribution must be carried out by a trial-and-error process. First a random fraction, R , is generated. The cumulative Poisson distribution is then formed term by term, using Eq. 4. At each step the cumulation is compared to R . When the first value of $P(x)$ satisfying the relationship

$$P(x) \geq R \tag{5}$$

is found, the corresponding value of x is taken as the random variate (number of arrivals).

The flow diagram of a computer program for accomplishing this generation is given in Appendix A. Such a program has been coded for the IBM 650, and numerous tests have been run. The results of some of these tests are discussed

later in the paper. A description of this program is being prepared for release through the U.S. Bureau of Public Roads.

EXPONENTIAL DISTRIBUTION

Many phenomena characterized by sequences of arrivals, as in traffic situations, may be treated by means of the exponential distribution

$$P(g \geq t) = e^{-t/k} \tag{6}$$

in which

g = gap between successive arrivals, in time units;

t = time, usually in seconds;

k = average time spacing between arrivals (may be thought of as the abscissa of the center of gravity of the area under the exponential curve);

$1/k$ = volume (number of arrivals per unit time, the unit of

time being the same as that used for t ; and

$P(g \geq t)$ = probability that $g \geq t$.

Eq. 6 expresses the probability that the spacing between arrivals is equal to or greater than the specified time; a plot for $k=1$ is shown in Figure 3. The cumulative exponential probability distribution is the complement of this relationship; namely,

$$P(g < t) = 1 - e^{-t/k} \tag{7}$$

which is shown graphically in Figure 4. By considerations related to the Monte Carlo method, it is possible for the present purposes to simplify this expression to

$$P = 1 - e^{-t/k} \tag{8}$$

where t is taken as the time spacing between arrivals.

Solving Eq. 8 for t gives

$$t = -k \ln(1 - P) \tag{9}$$

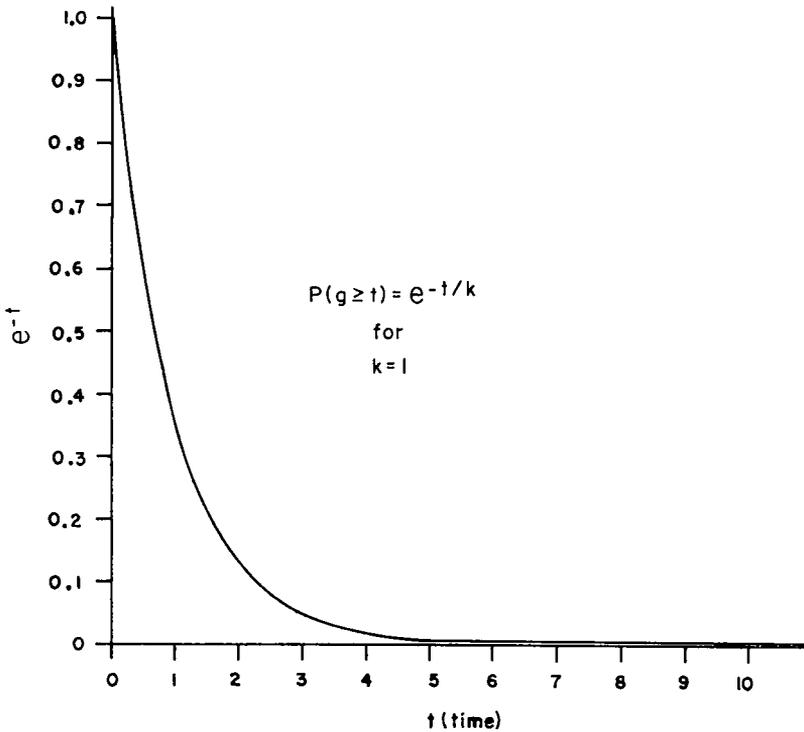


Figure 3.

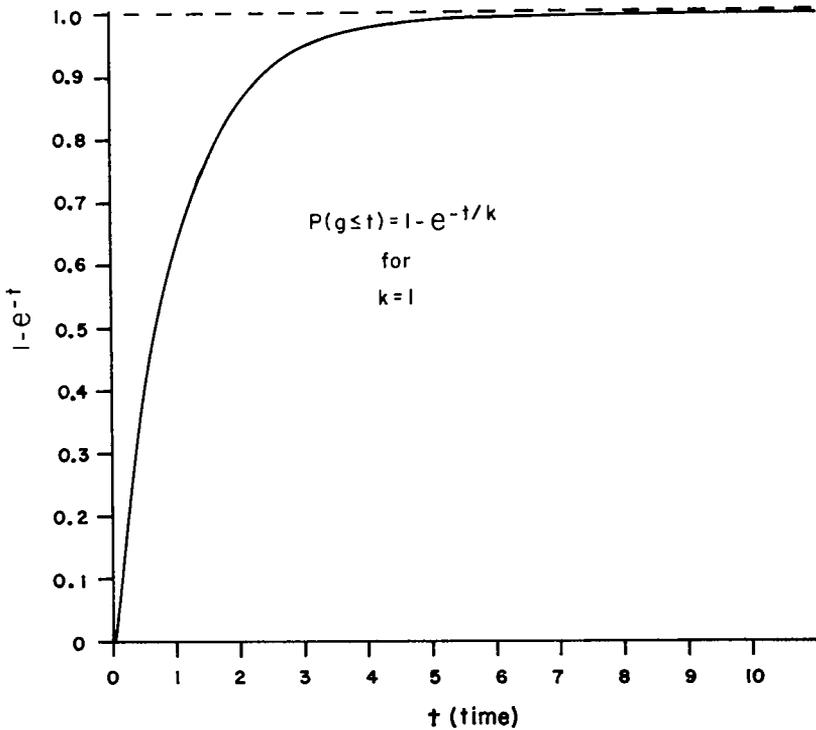


Figure 4.

By substituting the random fraction, R , for $(1-P)$ it is possible to solve for t . Flow diagrams for computer programs employing two different methods of solution are given in Appendices C and D. At least one of these will be released through the U.S. Bureau of Public Roads in a code usable on the IBM 650.

SHIFTED EXPONENTIAL DISTRIBUTION

Where all vehicles are free-flowing (that is, can pass at will) the exponential distribution appears to describe time-spacings adequately. However, when vehicles are flowing in platoons or are constrained so that they cannot pass at will, a modified distribution must be used.

From observations it is known that there is a certain minimum headway, τ , which can be maintained by vehicles. This may be stated as follows: "The probability of a gap between successive vehicles of less than τ is zero." This

phenomenon may be represented by an exponential curve shifted to the right by an amount τ , or

$$P_2 = 1 - e^{-(t-\tau)/k_2} \quad (10a)$$

which is plotted in Figure 5 with $k_2=1$. In Eq. 10, k_2 is the average time-spacing measured from the point where the curve intersects the t axis. The average spacing between vehicles, T_2 , is the average time-spacing measured from origin. Thus,

$$T_2 = k_2 + \tau \quad (11a)$$

or

$$k_2 = T_2 - \tau \quad (11b)$$

Thus,

$$P_2 = 1 - e^{-(t-\tau)/(T_2-\tau)} \quad (12)$$

Solving Eq. 10 for t gives

$$t = k_2 [-\ln(1-P)] + \tau \quad (10b)$$

Again, substituting the random fraction, R , for $(1-P)$ permits solution for t once values are available (or assumed) for k_2 and τ .

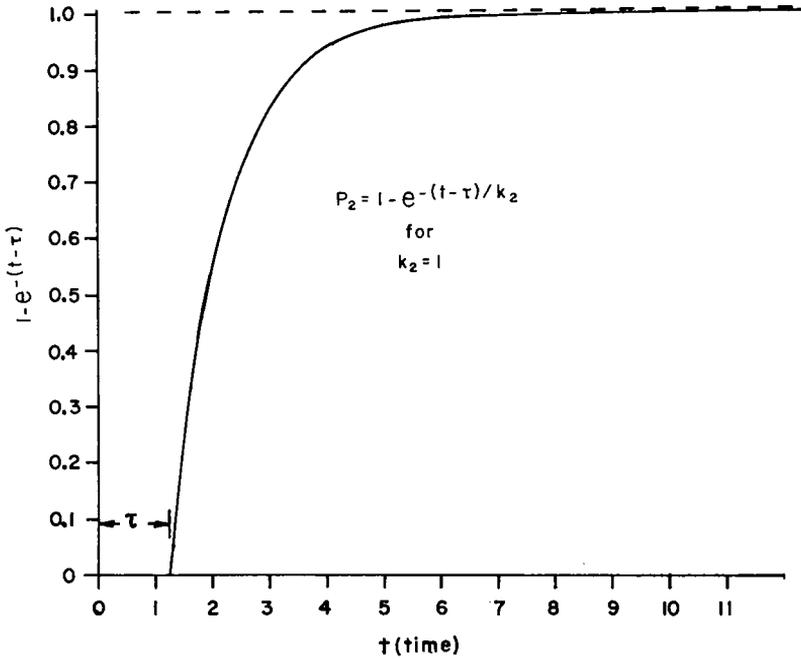


Figure 5.

COMPOSITE EXPONENTIAL DISTRIBUTION

Schuhl (5) has pointed out that a traffic stream can be composed of a combination of free-flowing vehicles and constrained vehicles. If a is the fraction of total volume made up of constrained vehicles, and $(1-a)$ is the fraction of total volume made up of free-flowing vehicles,

$$P_{1,2} = (1-a) (1 - e^{-t/k_1}) + a [1 - e^{-(t-\tau)/k_2}] \tag{13}$$

in which

- $P_{1,2}$ = probability of arrival from a composite distribution of constrained and unconstrained;
- k_1 = average time spacing of free-flowing (unconstrained vehicles);
- $k_2 = T_2 - \tau$; and
- T_2 = average time spacing of constrained vehicles.

Schuhl (5) gives a comparison between field observations and a curve of Eq. 13; this comparison is reproduced in Figure 6, which shows very satisfactory agreement.

A routine has been written for generating composite exponential distributions on an IBM 650. Several runs have been made; the results of one typical run are shown as plotted points in Figure 6.

RESULTS OF TESTS

Programs for the generation of Poisson inputs, exponential inputs, and composite exponential inputs have been run on the IBM 650, and the results compared with computed results and with field observations. Typical comparisons are presented in the following paragraphs.

Poisson Inputs

The program for the generation of Poisson inputs has been tested at several values of the parameter, m , ranging from 0.31 to 9.99. The results were tested with the chi-square test and substantially all were found acceptable at the 0.05 level. Tables 2 and 3 present typical comparisons of generated Poisson inputs, observed Poisson inputs, and ex-

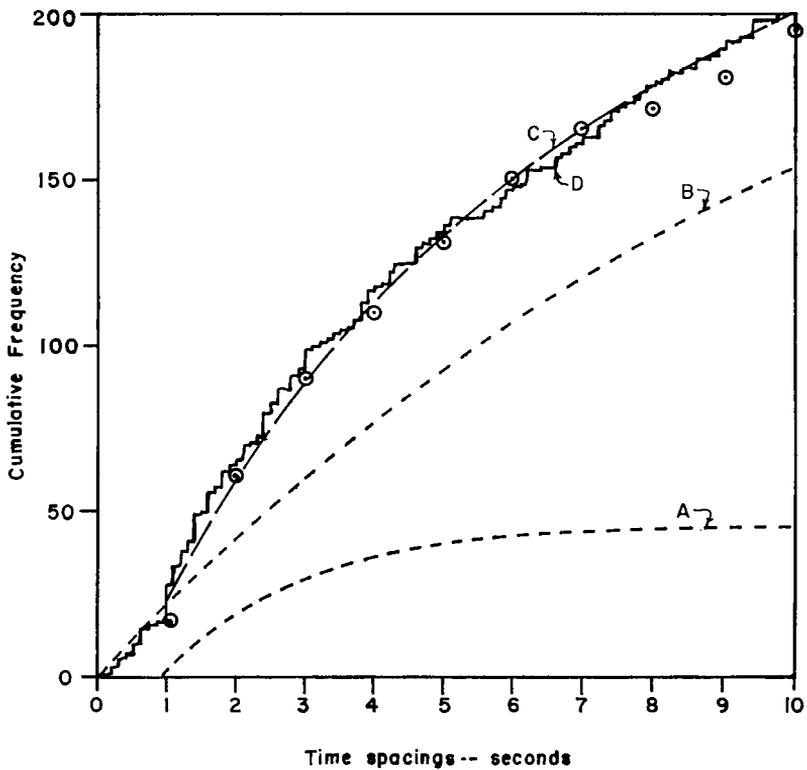


Figure 6. Composite exponential arrivals: Curve A computed for constrained vehicles; Curve B, for unconstrained vehicles; Curve C, for composite flow; Curve D, from field data (taken from Schuhl) Points in circles are results of computer generation.

TABLE 2
EXAMPLE OF POISSON DISTRIBUTION OF VEHICLE ARRIVALS

Number of Vehicles Arriving During Interval	Frequency		
	Expected	Observed In Field ¹	From Computer Generation
0	2.98	4	3
1	14.02	10	14
2	32.95	33	30
3	51.62	53	41
4	60.65	54	61
5	57.02	55	69
6	44.66	44	46
7	29.99	34	31
8	17.62	22	22
8	16.40	19	11
	328.00	328	328

Total intervals observed = 328
 $m = 4.75$
 χ^2 observed = 7.75
 χ^2 generated = 3.77
 χ^2 0.05 = 14.067

¹ Field data from Greenshields and Weida (4).

TABLE 3
EXAMPLE OF POISSON DISTRIBUTION OF SINGLE-CAR ACCIDENTS; $m = 1.00$; TOTAL ACCIDENTS = 44

Number of Single-Car Accidents	No. of Roads		
	Expected	Observed, 1950	Generated
0	16.2	18	19
1	16.2	14	18
2	8.1	7	6
≥ 3	3.5	5	1
Total	44.0	44	44

χ^2 observed = 0.5
 χ^2 generated = 3.4
 χ^2 0.05 = 3.84

pected values from the Poisson distribution.

Exponential Inputs

Several runs were made with the computer program for generation of exponential arrivals. Table 4 presents a typical comparison between generated results and expected results.

Composite Exponential Inputs

Several runs were made with a program for generating composite exponential arrivals. A comparison between field data, expected results, and generated results is made in Figure 6.

CONCLUSIONS

The foregoing examples demonstrate that many traffic situations can be represented by distributions which may be explicitly stated. It is clear that generation of traffic inputs by the computer appears to be acceptable in such situations. It should be further pointed out that traffic situations which can be represented by frequency distributions in tabular form lend themselves to computer generation. Thus, for many traffic

situations, generation of traffic inputs by the computer is thought to be entirely satisfactory for simulation and is to be preferred inasmuch as acceptable results are produced with a substantial saving in time.

REFERENCES

1. TAUSSKY, O., AND TODD, J., "Generation and Testing of Pseudo-Random Numbers." In "Symposium on Monte Carlo Methods," held at Univ. of Florida, Mar. 16 and 17, 1954. Ed. by H. A. Meyer. John Wiley and Sons (1956).
2. GERLOUGH, D. L., "Use of Poisson Distribution in Highway Traffic." In "Poisson and Traffic," pp. 1-58. The Eno Foundation for Highway Traffic Control (1955).
3. GREENSHIELDS, B. D., SHAPIRO, D., AND ERICKSEN, E. L., "Traffic Performance at Urban Street Intersections." Tech. Report No. 1, Yale Bureau of Highway Traffic (1947).
4. GREENSHIELDS, B. D., AND WEIDA, F. M., "Statistics with Application to Highway Traffic Analyses." The Eno Foundation for Highway Traffic Control, 238 pp. (1952).
5. SCHUHL, A., "The Probability Theory Applied to Distributions of Vehicles on Two-Lane Highways." In "Poisson and Traffic," pp. 59-75. The Eno Foundation for Highway Traffic Control (1955).
6. HASTINGS, C., "Approximations for Digital Computers." Princeton Univ. Press (1955).
7. VON NEUMANN, J., "Various Techniques Used in Connection with Random Digits." In "Monte Carlo Method," Nat. Bur. of Standards, Applied Math. Series, 12 (June 11, 1951).

TABLE 4
COMPARISON OF GENERATED AND EXPECTED
VALUES OF CUMULATIVE EXPONENTIAL
ARRIVALS (200 TRIALS)

Time Spacing (sec)	Cumulative Number of Gaps Less Than Specified Time Spacing	
	Expected	Generated
0	0	0
0.2	36	57
0.4	66	76
0.6	90	99
0.8	110	115
1.0	126	132
1.2	140	145
1.4	151	153
1.6	160	161
1.8	167	168
2.0	173	174
2.2	178	178

APPENDIX A

PROGRAM FOR GENERATION OF POISSON INPUTS

The Poisson distribution is expressed

$$P(x) = \frac{m^x}{x!} e^{-m} \quad x=0, 1, 2, \dots \quad (14)$$

where x is the number of occurrences during an interval, and m , the Poisson parameter, is the average occurrence per interval. If it is desired that the interval be different from the interval of averaging,

$$m = \beta t \quad (15)$$

where β is the average number of occurrences per unit time and t is the

number of time units in the interval.

The cumulative Poisson distribution is expressed

$$P(\leq x) = \sum_{i=0}^x \frac{m^i}{i!} e^{-m} \quad \text{where } i=0, 1, 2, \dots \quad (16a)$$

or

$$P(\leq x) = e^{-m} + \sum_{i=1}^x \frac{m^i}{i!} e^{-m} \quad \text{where } i=1, 2, 3, \dots \quad (16b)$$

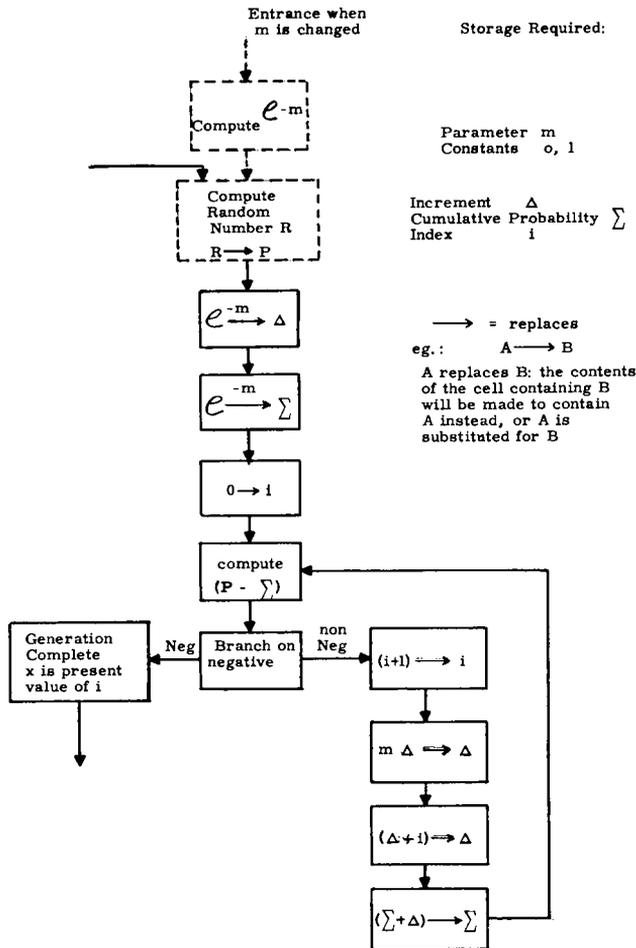


Figure 7. Flow diagram for trial-and-error generation of Poisson events.

where $P(\leq x)$ is the probability that x or fewer occur.

On a high-speed computer Poisson-distributed inputs may be generated by:

1. Generating or reading a random

number, P , which is interpreted as being between 0 and 1.

2. Performing a trial-and-error solution of Eq. 16b in order to obtain the value of x corresponding to P . This process is shown in Figure 7.

APPENDIX B

COMPUTATION OF e^{-m}

For fractional values of m , e^{-m} can be computed by a McLaurin series. However, fewer terms, and therefore greater computer speed, can be obtained by using one of the approximations of Hastings (6). The series used in the program for Poisson generation is as follows:

$$e^{-m} = \frac{1}{[1 + a_1 m + a_2 m^2 + a_3 m^3 + a_4 m^4 + a_5 m^5]^4} \quad (17)$$

in which

$$\begin{aligned} a_1 &= 0.2500, 1903, 6 \\ a_2 &= 0.0311, 9805, 6 \\ a_3 &= 0.0026, 7325, 5 \\ a_4 &= 0.0001, 2799, 3 \\ a_5 &= 0.0000, 1487, 6 \end{aligned}$$

For ease of computation, the quantity inside the brackets is arranged as follows:

$$\left[\left((a_5 m + a_4) m + a_3 \right) m + a_2 \right] m + a_1 \Big] m + 1.$$

APPENDIX C

GENERATION OF $t = -\ln(1-P)$ BY METHOD OF VON NEUMANN

Von Neumann (7) has demonstrated the following method of generating times in accordance with the exponential distribution:

1. Draw a series of random numbers, R , continuing as long as

$$R_1 > R_2 > \dots > R_n \leq R_{n+1}$$

2. If n is odd, R_1 is taken as the fractional part of t .

3. If n is even, the series is discarded, the integral part of t is incremented by 1, and a new series is drawn.

4. When a fractional part of t is accepted, the value of t is taken as the sum of the fractional part and the integral part.

A suitable program (including initialization) may be described by the flow diagram of Figure 8.

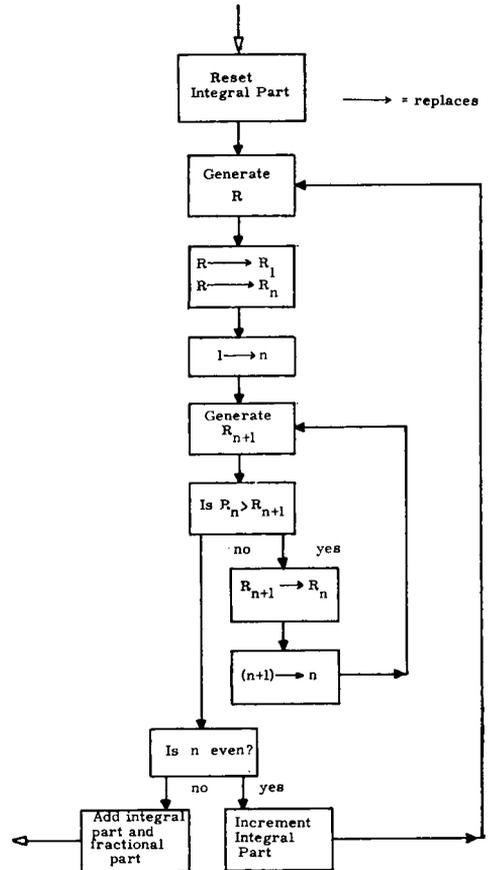


Figure 8. Flow diagram for obtaining $-\ln R$ by Von Neumann method.

APPENDIX D

GENERATION OF $t = -\ln(1-P)$ BY DIRECT METHOD

Hastings' approximation used:

$$\log_{10}x = C_1 \left(\frac{x-1}{x+1}\right) + C_3 \left(\frac{x-1}{x+1}\right)^3 + C_5 \left(\frac{x-1}{x+1}\right)^5 \quad (18)$$

for $\frac{1}{\sqrt{10}} \leq x \leq \sqrt{10}$

where

$$\begin{aligned} C_1 &= 0.8690286 \\ C_3 &= 0.2773839 \\ C_5 &= 0.2543275 \end{aligned}$$

Conversion relationship:

$$\ln W = 2.302585093 \log_{10} W = C_c \log_{10} W \quad (19)$$

In this program the portion of the computation called the mantissa is treated as x in the foregoing (see Figure 9). The desired result is $-\ln R$, where R is a random fraction.

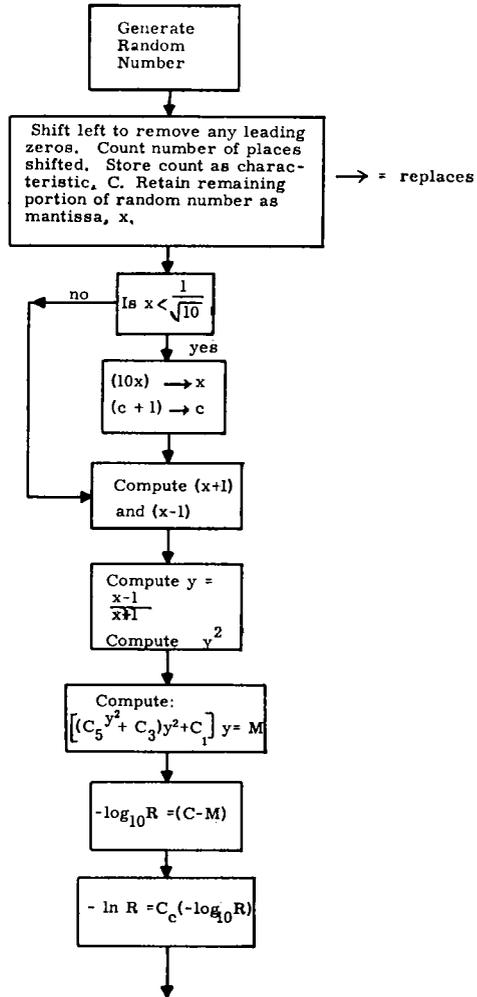


Figure 9. Flow diagram for obtaining $-\ln R$ by direct method.

APPENDIX E

GENERATION OF COMPOSITE EXPONENTIAL ARRIVALS

Any given arrival can belong to only one distribution. The fraction a of the arrivals will belong to the shifted exponential, while $(1 - a)$ will belong to the unshifted exponential.

For the unshifted exponential:

$$t = T_1 (-\ln R) \tag{20}$$

where T_1 is the average time-spacing of the vehicles in this group.

For the shifted exponential.

$$t = (T_2 - \tau) (-\ln R) + \tau \tag{21}$$

For this problem, several random generations are used as follows (Fig. 10):

- R_0 = random number to determine whether arrival belongs to shifted or unshifted exponential;
- R_1 = random number for unshifted exponential; and
- R_2 = random number for shifted exponential.

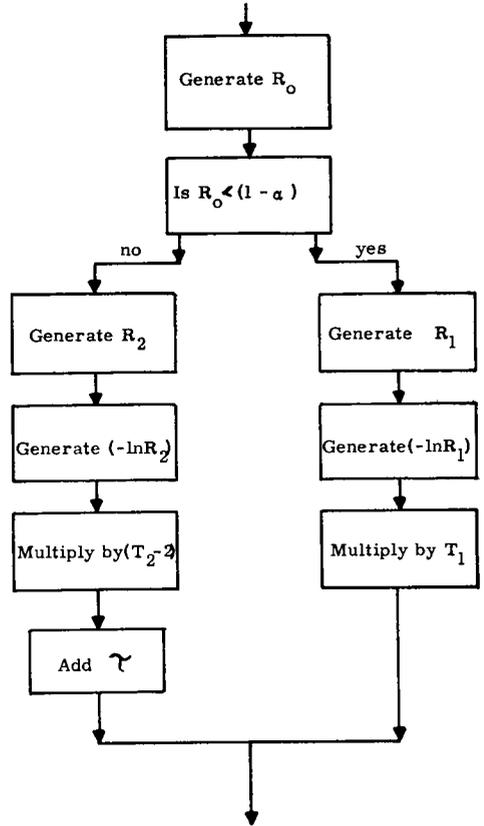


Figure 10. Flow diagram for generating composite exponential arrivals.