Lognormal Distribution for High Traffic Flows

MINJIE MEI AND A. GRAHAM R. BULLEN

The headway probability distributions for very high traffic flows are studied theoretically and empirically. Theoretical analyses show that the lognormal mechanism is applicable to individual headways for drivers in a car-following state. Thus the headway distribution of a traffic stream at high flows should follow the shifted lognormal distribution. Tests on a freeway data set with lane flow rates of 2,500 to 2,900 vehicles per day gave an excellent fit for the shifted lognormal distribution with a 0.3- or 0.4-sec shift. Previously the normal model had been the preferred headway distribution for high traffic flows, but in the present study it did not fit the data very well.

The time headway between vehicles in a traffic stream is an important and well-studied flow characteristic. Many models have been proposed for the probability distribution of headways and tested against highway traffic data. These distributions are widely used in traffic analysis methodologies and in traffic simulations.

Traffic displays distinct characteristics at different flow levels. There have been many studies of the distributions at low and medium traffic flow rates and many distributions have been calibrated. Very few studies, however, have been devoted to headway distributions at high flows for which all of the vehicles are in a car-following state. In this situation it has been most frequently assumed that the headway distribution follows a normal distribution (*1*), but the normal model does not fit the observed traffic data well and lacks a sound explanation of the traffic phenomenona.

The focus of this study is on the mechanism of time headways and their distribution at high flows. The lognormal model is proposed and is tested with field data, and it is compared with other models that have been previously presented.

LOGNORMAL MECHANISM FOR TIME HEADWAYS

The lognormal distribution has been found by many researchers to be the best simple model for headway distributions. Greenberg (*2*) and Daou (*3*) tried to find a theoretical basis for its validity as a headway model. Greenberg (*2*) and Tolle (*4*) found that the model gave good fits to the headway data that they collected. The lognormal model, however, has not been completely established for headway distributions for two reasons. First, no sound theoretical basis has been shown to relate the model to traffic behavior. Second, some statistical analyses with observed data have not been sufficiently rigorous.

The lognormal distribution arises as the result of a multiplicative mechanism acting on a number of factors. For this study, a unique case is of special interest, that is, the law of proportionate effect. This law deals with a variable whose value varies in a step-by-step sequence, such as in a time frame. Suppose that a variable is initially X_0, and after the jth time step it is X_j, reaching its final value X_n after n time steps. At the jth time step the change in the variable is a random proportion of the momentary value X_{j-1} already attained, thus

$$X_j - X_{j-1} = \varepsilon_j * X_{j-1} \quad (1)$$

where the set $\{\varepsilon_j\}$ is mutually independent and also independent of the set $\{X_j\}$. The law of proportionate effect then is,

> A variable subject to a process of change is said to obey the law of proportionate effect if the change in the variable at any step of the process is a random proportion of the previous value of the variable. (*5*)

The importance of the law is its link with the central limit theorem. Expression 1 may be rewritten as

$$\frac{X_j - X_{j-1}}{X_{j-1}} = \varepsilon_j \quad (2)$$

so

$$\sum_{j=1}^{n} \frac{X_j - X_{j-1}}{X_{j-1}} = \sum_{j=1}^{n} \varepsilon_j \quad (3)$$

Supposing the effect at each step to be small, then

$$\sum_{j=1}^{n} \frac{X_j - X_{j-1}}{X_{j-1}} \sim \int_{X_0}^{X_n} \frac{dX}{X} = \ln X_n - \ln X_0$$

giving

$$\ln X_n = \ln X_0 + \varepsilon_1 + \varepsilon_2 + \ldots + \varepsilon_n \quad (4)$$

The central limit theorem may be stated as, "Under very general conditions, as the number of variables in the sum becomes large, the distribution of the sum of random variables will approach the normal distribution" (*6*). The theorem is valid only when each individual variable has a small effect on

M. Mei, New York Department of Transportation, Region 8, Four Burnett Boulevard, Poughkeepsie, N.Y. 12602. A. G. R. Bullen, Department of Civil Engineering, University of Pittsburgh, Pittsburgh, Pa. 15261.

the sum. In Expression 4 the initial value X_0 will be close to 1.0 so ln X_0 will be of the order of the epsilons. The preceding process, therefore, meets all the requirements of the central limit theorem, so ln X_n is normally distributed. The random variable X_n is then lognormally distributed. Therefore, a variable obeying the law of proportionate effect is lognormally distributed provided that the change in each step is small.

Lognormal Distribution of Individual Time Headway

Assume that a driver is in a car-following situation on one lane of a freeway, and the driver either does not attempt to overtake the leading vehicle or does not have the chance to do so.

The headway of the vehicle, denoted by H, is always changing with time because drivers cannot maintain an absolute constant spacing and must adjust their speeds to maintain a safe distance and to keep up with traffic. Therefore H is a random variable changing with time. By specifying a point of time to be the starting point, and the headway at that point as H_0, the headway value after a small interval of time will change a random portion of its original value to become H_1

$$H_1 - H_0 = H_0 * \varepsilon_1 \tag{5}$$

After j time intervals, the change of headway value in the jth interval can be expressed as

$$H_j - H_{j-1} = H_{j-1} * \varepsilon_j \tag{6}$$

where

H_j = time headway at end of jth interval,
H_{j-1} = time headway at end of $j - 1$ interval,
ε_j = random proportion of change of H_{j-1} in jth interval.

With no knowledge about the distribution of ε_j the headway H will be lognormally distributed as long as the ε_j's are small and the initial headway is close to 1.0. The individual headway in a car-following situation, therefore, is lognormally distributed.

Shifted Lognormal Headway Distribution

Headway is the time interval between the arrivals of the corresponding point in a pair of moving vehicles, such as from front bumper to front bumper. The minimum value for a headway is the physical occupancy time of the leading vehicle plus a safety buffer. Then the domain of definition of headways is (d,∞). The random variable following the shifted lognormal distribution will be H', which is the headway, H, subtracted by the minimum spacing, d. The shift, d, can be determined from field data.

The mean, m, and standard deviation, σ, of headways from the observed data are given by

$$m = \frac{1}{n}\sum_{i=1}^{n} H_i \tag{7}$$

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(H_i - m)^2} \tag{8}$$

TABLE 1 Observed Headway Data and Their Parameters

Data Set No.	1	2	3	4	5	6	7	8
Lane No.	2	2	2	2	1	1	1	1
Average Volume(vpm)	46	49	39	42	35	36	35	34
No. of Observations	230	243	155	209	173	178	174	170
Mean (seconds)	1.27	1.24	1.54	1.43	1.70	1.67	1.71	1.75
Standard Deviation	0.60	0.52	0.59	0.63	0.88	0.65	0.85	0.94

Interval (Seconds)	Measured Distribution				Measured Distribution			
0.0 - 0.25	0	0	0	0	0	0	0	0
0.25 - 0.50	1	0	0	0	0	0	1	0
0.50 - 0.75	33	21	6	17	5	2	7	8
0.75 - 1.00	56	72	22	41	21	21	16	20
1.00 - 1.25	43	58	32	39	35	29	27	30
1.25 - 1.50	46	40	26	39	33	35	32	31
1.50 - 1.75	16	21	21	23	21	25	29	20
1.75 - 2.00	10	9	17	17	19	20	23	15
2.00 - 2.25	6	8	11	12	8	17	13	12
2.25 - 2.50	9	5	8	11	4	10	6	7
2.50 - 2.75	2	2	5	2	5	6	4	7
2.75 - 3.00	4	4	4	2	5	6	3	1
3.00 - 3.25	2	0	0	3	7	2	0	4
3.25 - 3.50	0	2	2	1	2	2	2	3
3.50 - 3.75	0	1	1	1	2	1	3	2
3.75 - 4.00	0	0	0	1	0	1	4	2
4.00 - 4.25	2	0	0	0	1	1	1	3
4.25 - 4.50	0	0	0	0	2	0	1	3
4.50 - 4.75	0	0	0	0	2	0	1	1
>4.75	0	0	0	0	1	0	1	1

Denoting the mean and standard deviation of H' as m' and σ' respectively, their values are given by

$$m' = m = d \tag{9}$$

$$\sigma' = \sigma \tag{10}$$

Lognormal Distribution and Time Headway Distribution of a Traffic Stream

To analyze the time headway distribution of a traffic stream, some knowledge about the drivers who make up the traffic stream is needed. Individual drivers will have their own unique desired headways and driving habits. Thus each driver will have a time headway distribution with a unique mean and standard deviation that may vary under different traffic conditions. The traffic stream is made up of individual time headways of different drivers with these different distributions. In a car-following condition, the measured time headway values from a traffic stream are the momentary values of each of the individual lognormal distributions.

The measured time headway distribution of a traffic stream is therefore the combination of the individual time headway distributions. The combined distribution may no longer be lognormal, and its characteristic is rather complicated. Consider the extreme case when the traffic volume is very high and all drivers have to drive at the car following headway. The mean of each individual time headway may still be different but they will converge toward a single value because of the close spacing. If, at very high volumes, the differences among the individual time headway distributions are very small and approach a single distribution, then the measured time headway distribution should converge to the lognormal or to the shifted lognormal distribution.

TESTING LOGNORMAL DISTRIBUTION WITH OBSERVED DATA

Data Collection and Reduction

Time headway data was taken on the two southbound lanes of the four-lane freeway I-279 at the Milroy overpass near

TABLE 2 Results of χ^2 Test of Shifted Lognormal and Normal Distribution with Observed Data

Data Set	#1	#2	#3	#4	#5	#6	#7	#8	
Test Level				Critical Values					
10%	12.0	13.5	12.0	12.0	14.6	12.0	12.0	12.0	
5%	14.2	15.5	14.2	14.2	16.6	14.2	14.2	14.2	
1%	18.2	20.0	18.2	18.2	21.5	18.2	18.2	18.2	
				Quantitative Test Results					
									Accept at 10%
S=0.0	17.67	19.35	1.90	4.61	18.70	1.97	10.72	10.72	12.00
S=0.1	15.81	15.93	2.08	3.52	16.82	2.74	9.87	8.86	12.00
S=0.2	13.21	11.22	2.54	3.71	14.20	2.70	8.61	7.59	12.00
S=0.3	11.68	6.74	4.30	5.31	12.37	3.52	9.54	6.52	12.00
S=0.4	11.87	3.42	4.15	8.65	10.42	5.12	6.84	6.58	12.00
S=0.5	20.69	1.98	6.37	21.94	10.10	8.59	6.21	5.63	12.00
S=0.6		11.26	10.88		10.05	14.55	6.91	7.23	12.00
S=0.7		>20	19.98		12.554	-	10.69	12.33	12.00
S=0.8					25.559		>25		12.00
Normal	50.27	84.22	26.05	27.26	77.375	33.00	48.12	69.72	12.00
				Acceptable or Not					
S=0.0	-	-	Y	Y	-	Y	Y	Y	
S=0.1	-	-	Y	Y	-	Y	Y	Y	
S=0.2	-	Y	Y	Y	Y	Y	Y	Y	
S=0.3	Y	Y	Y	Y	Y	Y	Y	Y	
S=0.4	Y	Y	Y	Y	Y	Y	Y	Y	
S=0.5	-	Y	Y	-	Y	Y	Y	Y	
S=0.6	-	Y	Y	-	Y	-	Y	Y	
S=0.7	-	-	-	-	Y	-	Y	-	
S=0.8	-	-	-	-	-	-	-	-	
Normal	-	-	-	-	-	-	-	-	

Y ... Acceptable at All Test Levels
\- ... Not Acceptable At All Test Levels
S ... Shift (Seconds)

downtown Pittsburgh during morning rush hour. There is an approximately 1 percent downgrade in the section. The average volume was 2,400 vehicles per hour per lane with 5-min flow rates reaching 2,900 vehicles per hour per lane. The average speed was about 75 km/hr, and the traffic flow was smooth and free of shock waves. There were few trucks and little lane changing in the traffic stream.

About 10,000 headways were measured, of which 1,375 were used for model testing in 8 data sets. The right lane is denoted as Lane 1, and the left lane as Lane 2. Data were grouped according to volumes. Data from the literature were used to test the model over other ranges of traffic flows.

Maximum likelihood methods were used to estimate the mean and the standard deviation of the model. The eight data sets and their parameters are given in Table 1.

Testing Models with Observed Data

In the collected data in Table 1, there are only 2 headways of less than 0.5 sec out of 1,375 measured headways. The average occupancy time is about 0.25 sec. Therefore, the shift should be in the neighborhood of 0.2 to 0.5 sec. Tests were carried out with varying shifts from 0.0 to 0.7 sec.

The most common methods for testing goodness-of-fit are the χ^2 test and Kolmogorov-Smirnov test.

Tolle did his analysis on the lognormal model mainly with Kolmogorov-Smirnov test because he found that

> The χ^2 test is not a very forgiving analysis and may be thrown off by only a few "bad" points. In reality, obtainment of actual "good" χ^2 fits from data which are influenced by so many unpredictable variables is not fully expected. (*4*,p.83)

The significance levels used were 1, 5, and 10 percent. The Kolmogorov-Smirnov tests gave acceptable results at all test levels for the eight data sets regardless of the shift. χ^2 tests with all the data sets were acceptable at all the test levels for shifts of 0.3 or 0.4 sec, but for no other shifts. The results are given in Table 2.

The normal distribution was also tested with the same data. None of the tests with the eight data sets gave acceptable results at the test levels.

Another model that has been used for headway distributions is the Pearson Type 3 distribution. With several parameters to determine the location, scale, and shape of the distribution, the model gives acceptable results for the data with χ^2 tests at the test levels. The calibration is more difficult, however, and the model also lacks any explanation of the traffic phenomenon.

CONCLUSION

Theoretical analyses have shown that the lognormal mechanism is applicable to individual headways with traffic in a car-following state. Thus the headway distribution of a traffic stream should converge to the shifted lognormal distribution as flow increases. Statistical tests on a high flow data set gave an excellent fit for the shifted lognormal distribution with a 0.3- or 0.4-sec shift. The lognormal model was superior to the normal model. Although the Pearson Type 3 model also gives good fits in most cases, it lacks a sound conceptual basis. The shifted lognormal model is an excellent and simple model for headway distributions at high flows.

ACKNOWLEDGMENT

This work is based partially on an M.S. thesis submitted by the author in partial fulfillment of the requirements for the degree of Master of Science in Civil Engineering at the University of Pittsburgh.

REFERENCES

1. A. D. May. *Traffic Flow Fundamentals.* Prentice-Hall, Englewood Cliffs, N.J., 1990, pp. 19–22.
2. I. Greenberg. The Log-Normal Distribution of Headways. *Australian Road Research*, Vol. 2, No. 7, March 1966, pp. 14–18.
3. A. Daou. On Flow Within Platoons. *Australian Road Research*, Vol. 2, No. 7, March 1966, pp. 4–13.
4. J. E. Tolle. *The Lognormal Headway Distribution Model.* Ph.D. dissertation. Ohio State University, Columbus, 1969.
5. J. Aitchison and A. C. Brown. *The Lognormal Distribution.* Cambridge University Press, Cambridge, 1963, pp. 22–23.
6. J. R. Benjamin. *Probability, Statistics and Decision for Civil Engineers.* McGraw-Hill, New York, N.Y., 1970.

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