Route Costing: A Review. U.S. Urban Mass Transportation Administration, May 1981.
7. R. J. Gephart. Development of a Bus Operating Cost Model Based on Disaggregate Data. In *Transportation Research Record 1011*, TRB, National Research Council, Washington, D.C., 1984, pp. 16–23.
8. B. Williams and W. Wells. Economics of Commuter Express Bus Operations. In *Transportation Research Record 915*, TRB, National Research Council, Washington, D.C., 1983, pp. 13–18.
9. R. B. Cervero, M. Wachs, R. Berlin, and R. J. Gephart. *Efficiency and Equity Implications of Alternative Transit Fare Policies.* Report CA-11-0019. U.S. Urban Mass Transportation Administration, July 1980.
10. D. L. Dornan. *Analysis of Commuter Rail Costs and Cost Allocation Methods.* Report TSC-1758-13. Transportation Systems Center, Peat, Marwick, Mitchell and Co., July 1983.
11. J. H. Walder. *Commuter Van Service on Staten Island.* Paper for Kennedy School of Government, Harvard University, Cambridge, Mass., April 1983.
12. R. Carey and B. Campbell. *Reducing the MBTA Deficit.* Paper for Program in City and Regional Planning, Harvard University, Cambridge, Mass., May 1983.
13. W. Cherwony and S. Mundle. Peak-Base Cost Allocation Models. In *Transportation Research Record 663*, TRB, National Research Council, Washington, D.C., 1978, pp. 52–56.
14. J. M. Reilly. Transit Costs During Peak and Off-Peak Hours. In *Transportation Research Record 625*, TRB, National Research Council, Washington, D.C., 1977, pp. 24–25.
15. J. A. Dawson. Segmentation of the Transit Market, *Transportation Quarterly*, Vol. 37, No. 1, Jan. 1983, p. 82.
16. D. H. Pickrell. *The Causes of Rising Transit Operating Deficits.* Report MA-11-0037. U.S. Urban Mass Transportation Administration, July 1983.

The views expressed here are strictly those of the author and in no way represent an official position of the U.S. Department of Transportation. Helpful comments were offered by Jose A. Gomez-Ibanez, John R. Meyer, and several TRB referees; but the author is solely responsible for any errors.

Publication of this paper sponsored by Committee on Public Transportation Planning and Development.

On-Time Performance and the Exponential Probability Distribution

WAYNE K. TALLEY AND A. JEFF BECKER

In spite of the seemingly strong support for research in on-time performance of bus service, previous research has largely been informal with little statistical basis. In this paper, it is concluded that the distribution of late and early time intervals between actual and scheduled time arrivals for buses at bus stops on a particular route conforms to the exponential probability distribution. The probability equation of the distribution can be used to compute the probability or percentage of buses arriving at a given bus stop that will be more than x minutes early or more than y minutes late. These probabilities may be interpreted as failure rates. The probability equation allows flexibility in interpreting results and setting standards for on-time performance.

The significant amounts of government funding being provided to public transit firms has given rise to concern regarding public return from such funding. This has led to an interest in studying the performance of public transit firms. Because bus service is the most common to be provided by public transit firms, particular attention has been given to studying its performance.

One area in the performance evaluation of bus service that has received a great deal of attention is on-time performance. On-time performance of bus service has been defined by John Bates (*1*) as "a motorbus passing or leaving a predetermined point along its routing within a time envelope that is no more than x minutes earlier and no more than y minutes later than a published schedule time." Recently, a survey (*1*) was conducted to determine basic practices and attitudes concerning on-time performance of bus service. The general conclusions of the survey are:

1. There is wide variation in the definition of on-time performance; however, a definition of no more than 1 min early and no more than 5 min late is the most commonly used.
2. Determination of on-time performance appears to be a largely informal practice with little statistical basis.

W. K. Talley, Old Dominion University, Norfolk, Va. 23508. A. J. Becker, Tidewater Transportation District Commission, Norfolk, Va. 23501.

3. On-time performance is considered to be an important bus performance characteristic.
4. There is strong support for research in on-time performance of bus service.

The purpose of this paper is to provide a statistical basis for analyzing on-time performance. Specifically, the paper proposes that the exponential probability distribution be used to compute the probabilities that buses (or percentage of buses) on a particular route and arriving at a particular bus stop will be more than *x* minutes early and more than *y* minutes late. This approach to evaluating on-time performance of bus service differs from the traditional approach in two major ways: (*a*) the former focuses on those bus arrivals that lie outside of the on-time performance interval as previously discussed, while the latter focuses on those arrivals that lie inside of the on-time performance interval; and (*b*) instead of the traditional approach in providing a formal definition of on-time performance, which a bus stop on a particular route must adhere to, the proposed approach provides probabilities, which buses at a bus stop on a particular route will be more than *x* minutes early and more than *y* minutes late; this information, in turn, may be used by transit management for obtaining different definitions of on-time performance depending on the bus stop (and thus route) in question or for evaluating the performance of a route with respect to the stated formal definition of on-time performance.

The methodology based on the exponential probability distribution for evaluating the on-time performance of bus service is presented followed by a section on the application of the methodology to bus routes of the Tidewater Transportation District Commission (TTDC), a public transit firm.

METHODOLOGY

The exponential probability distribution is a continuous distribution with respect to the variable x, where x is the interval (e.g., time) between events. The arithmetic mean (μ) and variance (σ^2) for the exponential probability distribution are computed as follows:

$$\mu = 1/\alpha \quad (1)$$

$$\sigma^2 = 1/\alpha^2 \quad (2)$$

The exponential probability distribution restricts the values of x from being negative. A plot of an exponential probability distribution with a mean of 2 ($\mu = 2$) is shown in Figure 1.

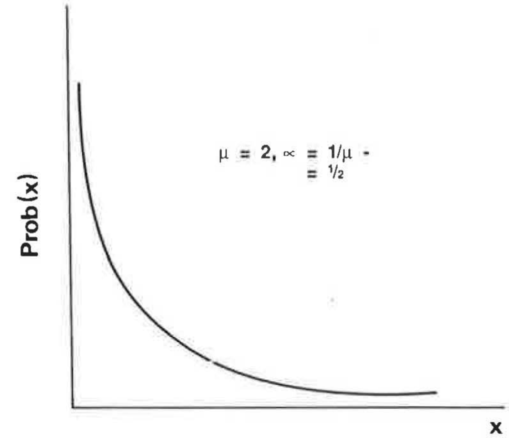

FIGURE 1 Exponential probability distribution.

A sample of y was obtained in order to endorse the proposal to use exponential probability distribution for computing the probability that buses on a particular route will be more than y minutes late at a bus stop.

The y values for motorbuses arriving at a bus stop on a particular route of the TTDC for a given day were assumed to be distributed as an exponential probability distribution. Let y represent the time interval in minutes for which a bus is late at the bus stop. This time interval is computed by taking the scheduled time of arrival for a bus at the bus stop from the actual time of arrival. Because the exponential probability distribution restricts the values of the variable from being negative, the sample was restricted to consist of only positive values for the variable y (which is concerned only with lateness). This sample in turn, was used in a Kolmogorov-Smirnov Goodness-of-Fit Test to make inference to the null hypothesis that this sample was taken from an exponential probability distribution. This test is described by Sidney Siegel (*2*).

Based on a 5 percent level of significance, the null hypothesis could not be rejected. This test was also performed with several other samples with respect to lateness and had the same result. Hence, the sample evidence supports the assumption that late time intervals for motorbuses arriving at a bus stop on a given route are distributed as the exponential probability distribution.

In order to provide support for whether the exponential probability distribution can be used for computing the probability that buses on a given route will be more than x minutes early at a particular bus stop, the same analysis was conducted for lateness. Based on a 5 percent level of significance, the null hypothesis could not be rejected. It was concluded that the sample evidence supports the assumption that the distribution of early time intervals for buses arriving at a bus stop on a particular route is determined by exponential probability.

The samples of time intervals for buses have been divided into two sorts: a sample for lateness and a sample for earliness. On-time buses are included in both samples. This follows from the definitions of x and y variables, and from the elimination of negative values for the variables in both samples. x is computed by taking the actual time of arrival from the scheduled time of arrival, and y is computed by taking the scheduled time of arrival from the actual time of arrival, therefore, a motorbus that is on time will be in both samples because $x = y = 0$.

Instead of including the on-time buses in both samples, it may be that by measuring on-time performance more accurately (in terms of seconds) no on-time buses for the route will be found. In that case, the problem of what to do with the on-time buses is solved. Alternatively, it may be assumed that the on-time performance of the initially classified on-time buses was not measured correctly and, furthermore, that these buses are equally likely to be late or early. Hence, one-half of these buses may be placed on the early bus sample and the other one-half in the late bus sample for the given route. A final alternative is to eliminate those buses that are actually found to be on time (measured in seconds) from either sample for the route.

Based on the preceding discussion, exponential probability distribution can be relied on for computing the probability that daily buses on a particular route will be more than x minutes early or more than y minutes late at a given bus stop. The following discussion concerns late arrivals; however, the same conclusions also pertain to early arrivals.

The probability that buses will be more than b minutes late can be computed from the following formula

$$\text{Prob}\ (y > b) = e^{-ab} \qquad (3)$$

Where, e is the base of natural logarithms, and a equals 1 divided by the arithmetic mean of the values of y in the sample.

Suppose in a sample of daily late arrivals (also including on-time arrivals) that the arithmetic mean of the arrivals was 1.63 min late. If b is specified to be 5 min late (the generally accepted standard at present), then the probability that buses on this route will be more than 5 min late will be 0.0498 or

$$\text{Prob}\ (y > 5) = e^{(-1/1.63)5} = 0.0498 \qquad (4)$$

For the purpose of evaluating the on-time performance of bus service, the exponential probability distribution provides a great deal of flexibility for transit management. For example, there are three unknowns in Equation 3: (*a*) the Prob ($y > b$), (*b*) the parameter a, and (*c*) the parameter b. Given any two of these unknowns, Equation 3 can be used to determine the third. In Equation 4 values for a and b are specified and solved for Prob ($y > b$). Values for Prob ($y > b$) and a can also be specified and solved for b. The final possibility is to specify Prob ($y > b$) and b and solve for a or the reciprocal of the arithmetic mean of the late arrival observations.

TABLE 1 ON-TIME PERFORMANCE FOR LATENESS

Route	No. of Observations > or = 0 Min Late	Average Late Adherence (min)	Probability Bus More Than 5 Min Late (%)	5 Percent Buses May Be This Late (min)	Buses 1 Min Early to 5 Min Late (%)
1	71	2.93	18.2	8.8	78
2	27	1.97	7.9	5.9	90
3	49	2.12	9.5	6.3	87
4	69	2.55	14.1	7.6	84
5	19	1.84	6.6	5.5	96
6	29	7.84	52.8	23.4	77
8	17	4.88	35.9	14.6	78
9	15	2.67	15.4	8.0	88
10	18	1.67	5.0	5.0	100
11	30	2.20	10.3	6.6	85
12	17	3.30	22.0	9.9	74
13	24	3.96	28.3	11.8	68
15	17	0.86	0.3	2.6	92
16	14	2.86	17.4	8.6	88
17	16	2.75	16.2	8.2	89
18	30	4.73	34.7	14.1	69
19	5	1.40	2.8	4.2	100
20	39	3.72	6.1	11.1	52
20X	5	1.60	4.4	4.8	86
22	10	0.60	0.0	1.8	100
23	33	3.22	21.2	9.6	89
26	41	2.02	8.4	6.0	100
35	2	1.50	3.6	4.5	100
36	6	8.00	53.5	23.9	70
37	2	3.50	24.0	10.5	33
39	29	0.90	0.4	2.7	100
40	39	0.79	0.2	2.4	94
41	8	2.00	8.2	6.0	78
44	10	0.60	0.0	1.8	82
45	64	5.47	40.1	16.4	63
46	10	3.60	24.9	10.8	82
47	22	2.27	11.1	6.8	86
49	17	2.25	10.8	6.7	94
50	11	0.73	0.1	2.2	91
71	11	7.18	49.8	21.5	26
72	11	0.57	0.0	1.7	38
73	15	8.14	54.1	24.3	41
74	22	3.37	22.7	10.1	51
75	15	5.67	41.4	17.0	44
76	17	8.24	54.5	24.6	38
80	16	2.19	10.2	6.5	89

APPLICATION OF METHODOLOGY TO THE TTDC

The TTDC is a public transit firm chartered in the Commonwealth of Virginia to plan, operate, and regulate public transportation services. Five cities (Chesapeake, Norfolk, Portsmouth, Suffolk, and Virginia Beach) are members of the TTDC and receive public passenger transportation from the TTDC. The TTDC provides a variety of public passenger transportation services, including bus, dial-a-ride, elderly and handicapped, ferry, and vanpool. With respect to bus service, the TTDC is a medium-sized system with 120 peak-hour buses on 41 routes

Time intervals between scheduled and actual time of arrival for buses arriving at a bus stop on each of the 41 routes were obtained for a typical day. The arithmetic means of the late and early arrivals (where on-time arrivals are included in both types of arrivals) for the 41 routes were computed (see Table 1 and Table 2, respectively).

As stated previously, the most commonly used definition of on-time performance by U.S. bus firms is no more than 1 min early and no more than 5 min late. The rationale for selecting this definition is not clear, but appears to be based on a reasonable waiting time for buses delayed due to travel conditions. Early departure is regarded as unnecessary and undesirable.

The failure rate with respect to the preceding definition of on-time performance may be obtained with the aid of the exponential probability distribution. The failure rates (5 min late) for TTDC's 41 routes are given in the fourth column of Table 1; failure rates (1 min early) for TTDC's 41 routes are found in the fourth column of Table 2. Note that these columns are not additive. With respect to the TTDC's lateness samples, the probability that buses will be more than 5 min late ranges from 0 to 54.5 percent for the routes. With respect to the TTDC's earliness samples, the probability that buses will be more than 1 min early ranges from zero to 81.9 percent for the routes. Given the wide range in the failure rates in being late

TABLE 2 ON-TIME PERFORMANCE FOR EARLINESS

Route	No. of Observations < or = 0 Min Early	Average Early Adherence (min)	Probability Bus More Than 5 Min Early (%)	5 Percent Buses May Be This Early (min)
1	27	0.30	3.6	0.9
2	17	0.29	3.2	0.9
3	24	0.38	7.2	1.1
4	36	0.26	2.1	0.8
5	8	0.75	26.4	2.2
6	10	0.30	3.6	0.9
8	2	1.00	36.8	3.0
9	6	0.33	4.8	1.0
10	7	0.43	9.8	1.3
11	12	0.25	1.8	0.7
12	7	0.29	3.2	0.9
13	6	0.17	0.3	0.5
15	19	0.66	22.0	2.0
16	3	0.67	22.5	2.0
17	5	0.60	18.9	1.8
18	15	2.00	60.7	6.0
19	2	0.50	13.5	1.5
20	31	1.86	58.4	5.6
20X	3	2.00	60.7	6.0
22	4	0.00	0.0	0.0
23	6	0.83	30.0	2.5
26	23	0.13	0.0	0.4
35	0	0.00	0.0	0.0
36	7	1.29	46.1	3.9
37	1	3.00	71.7	9.0
39	17	0.00	0.0	0.0
40	33	0.51	14.1	1.5
41	5	0.20	0.7	0.6
44	7	1.14	41.6	3.4
45	22	0.85	30.8	2.5
46	2	0.50	13.5	1.5
47	14	0.50	13.5	1.5
49	10	0.50	13.5	1.5
50	9	0.00	0.0	0.0
71	14	3.71	76.4	11.1
72	15	2.07	61.7	6.2
73	4	3.60	75.7	10.8
74	25	2.74	69.4	8.2
75	7	1.57	52.9	4.7
76	3	5.00	81.9	15.0
80	10	0.30	3.6	0.9

and early, the question arises as to whether the restrictions of on-time performance as defined (5 min late and 1 min early) are appropriate for every route. A higher value for *b* (late or early) may be acceptable for a more congestion-prone route.

Instead of focusing on the probability of failure rates, management may want to examine the magnitude of the failure to adhere to a schedule. Suppose transit management initially considers setting the limits for the definition of on-time performance based on a failure rate of 5 percent for being late and early. For the TTDC (see Column 5, Table 1), lateness limits will range from 1.7 min to 24.6 min for the routes; earliness limits (see Column 5, Table 2) will range from 0 to 15 min. If transit management chooses a higher failure rate of 10 percent, the ranges of the limits will become smaller.

Practically speaking, transit management may seek to keep the commonly used definition of on-time performance: no more than 5 min late and no more than 1 min early (see Column 6, Table 1). In that case, management may use the fourth columns in Tables 1 and 2 to determine whether a route is adhering to the definition of on-time performance with respect to certain probability limits. Suppose management states that a given route is adhering to the definition of on-time performance if the failure rate for lateness and earliness does not exceed 10 percent. In Table 1, 25 of the routes exceed this failure rate with respect to lateness; in Table 2, 24 of the routes exceed this failure rate with respect to earliness.

By increasing the acceptable failure rate, the number of routes with failure rates exceeding the acceptable rate will decrease. Hence, transit management might begin by stating a relatively high acceptable failure rate and concentrate on those few routes that exceed this rate by attempting to correct the problems that have caused the routes' relatively high failure rate. Over time the acceptable failure rate could be lowered in order to select those routes that are to be investigated for the purpose of improving their on-time performance. Furthermore, because probabilities are available for lateness and earliness, transit management can detect a route that may have an early on-time performance problem but not a late problem, and vice versa.

CONCLUSIONS

Although there appears to be strong support for research in on-time performance of bus service, previous research has largely been an informal practice with little statistical basis. In this paper, it is concluded that distribution of late and early schedule adherence intervals for buses arriving at a bus stop on a particular route conforms to the exponential probability. If the definition of on-time performance states that a bus at a bus stop will be no more than x minutes early and y minutes late, the probability equation of the exponential probability distribution can be used for computing the failure rate with respect to these limits. Specifically, the probability equation can be used to compute the probability or percentage of buses arriving at a given bus stop that will be more than x minutes early and more than y minutes late. Once transit management specifies the acceptable failure rates for being late and early, the probabilities given here can be compared with rates used for detecting bus routes experiencing problems that cause relatively high failure rates.

A major problem that arises when applying the methodology described is the categorization of samples (early or late) of buses that are initially classified as on time. This problem may be solved by measuring the on-time performance of buses more accurately (in terms of seconds), so that no buses are classified as on time for the route. The appropriate method for handling this problem will be deferred to future research. However, whichever method is selected, researchers should be aware that the proposed methodology serves as a stastistically rigorous screening device for detecting bus routes with possible on-time performance problems rather than as a forecasting model of schedule adherence.

REFERENCES

1. J. W. Bates. Definition of Practices for Bus Transit On-Time Performance: Preliminary Study. *Transportation Research Circular*, No. 300, Feb. 1986, pp. 1–5.
2. S. Siegel. *Nonparametric Statistics for the Behavioral Sciences.* McGraw-Hill Book Company, New York, N.Y., 1956.

Publication of this paper sponsored by Committee on Bus Transit Systems.