METHODS OF ESTIMATING HIGHWAY TRAFFIC VOLUME

By W Arthur Shelton
U S Bureau of Public Roads

SYNOPSIS

This paper compares the accuracy of results of two types of traffic sampling, 8 hour counts and 1 hour observations, applied to the known traffic of the Holland Tunnel and the George Washington Bridge of the Hudson River Samples of 70 separate and independent hours of Holland Tunnel Traffic yielded estimates of which the error was only 63 per cent of that for samples of 128 hours taken in 16 watches of 8 hours each Smaller samples of separate hours gave somewhat greater errors Similar results were secured for the George Washington Bridge traffic, samples of 140 separate hours gave estimates of which the error was only 40 per cent of that for samples of 128 hours in 16 clusters of 8 hours each

Samples consisting of from 28 to 140 separate and independent hours were found to give more precise estimates than samples of 16 watches of 8 hours each The index of precision ranged from 101 per cent for the 28-hour sample to 195 per cent for the 140-hour sample

The chief factor in the reduction of errors in estimating traffic appears to be the number of separate and independent observations made rather than the number of hours composing each observation

INTRODUCTORY

The methods of estimating highway traffic volume may be divided into two general classes One of these embraces all types of systematic sampling of the traffic variable at every point estimated This is the method that the Bureau of Public Roads has used in most of its surveys of traffic volume, and it includes both short and long observations ranging from one hour or less to 24 hours The other broad class has been called the short-count method, and it is based largely on the establishment of traffic patterns and the selection of a small number (usually only 2) of observations at each station as a basis for the estimate This second class includes a wide variety of methods of estimate but is sharply distinguished from all systematic methods of random sampling The present paper is confined to the comparison of the results obtained from two types of systematic sampling, those of (1) eight-hour and (2) one-hour observations

Something of the sampling method used and results obtained from traffic records of Holland Tunnel in 1933 was presented at the annual meeting of the Highway Research Board two years ago [1] Only a few comments will be made on (1) the results of the tests of Holland Tunnel and (2) the methods used in sampling the traffic variable over George Washington Bridge The wider variation of the traffic volume over George Washington Bridge than through Holland Tunnel will be shown in the results obtained

Our tests have been confined to these two Hudson River crossings, and to the problem of what is the most efficient sample from the standpoint of precision of estimate The comparisons have been restricted to the estimates of mean daily volume of traffic made from samples of watches of eight hours and those made from several types of samples of watches of one hour The test of comparative cost is reserved for an actual survey

The results of sampling tests for Holland Tunnel are shown in Figures 1 and 2 Samples of 70 separate and independent hours yield estimates of which

[1] Proceedings, Highway Research Board, vol 14, 1934, pp 398–409

the error, as measured by standard deviation, is only 63 percent of that for samples of 128 hours taken in 16 watches of 8 hours each Samples of 62 separate hours yield estimates of which the error is only 70 percent of that for samples of 128 hours taken in 16 watches of 8 hours each The curves of Figure 1 show that the range of variation is also greater for samples of 16 watches of 8 hours than for either of the other two types of sample

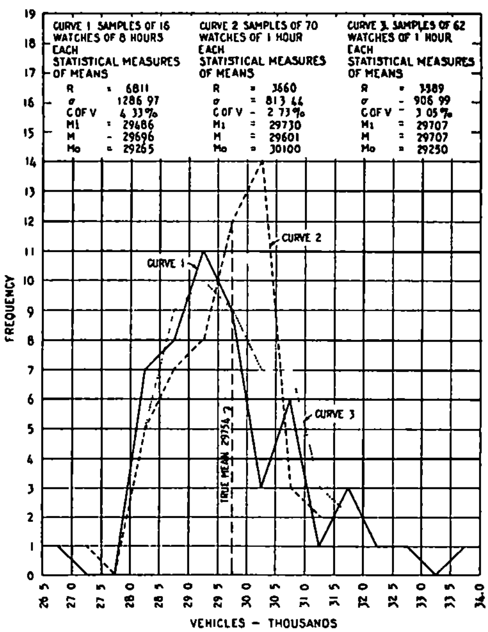

Figure 1 Frequency Distribution of Estimated Means of Daily Volume of Traffic Through Holland Tunnel for 16, 70, and 62 Observations

Samples of 35 separate hours yield estimates of mean daily volume of traffic of which the error is only 64 percent of that for 128 hours taken in 16 watches of 8 hours each, and those of 28 separate hours result in estimates of which the error is only 91 percent as great as for samples of 16 watches of 8 hours each Figure 2 shows that as the number of separate hours is reduced, the concentration of the estimates around the central group (true mean) decreases and the range of variation and the standard deviation both increase For all samples of 28, 35, 62, and 70 separate and independent hours the standard deviation and the range of variation are smaller than for 128 hours taken in 16 watches of 8 hours each It is a striking fact that 28 separate hours yield estimates of which the error is 9 percent less than for the

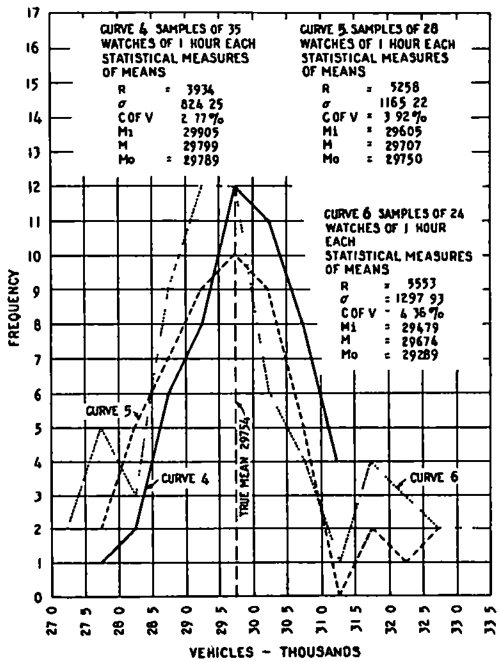

Figure 2 Frequency Distribution of Estimated Means of Daily Volume of Traffic Through Holland Tunnel for 35, 28, and 24 Observations.

estimates from 128 hours taken in clusters of 8 hours each

GEORGE WASHINGTON BRIDGE

For George Washington Bridge the samples were drawn in the same general manner as for Holland Tunnel except that the number of hours in the samples was extended up to 140 and down to 8 Two sets of estimates were also computed from samples of 35 watches for the purpose of testing the effect of using certain hours of less than average variation in comparison with those taken systematically from the natural cycle of variation, and the results show smaller error from the hours of greater than average constancy

Since the mean daily volume of traffic through Holland Tunnel is greater than that over George Washington Bridge by roughly two to one, the standard deviation of estimates is greater for the former, although the relative range and the coefficient of variation are much greater for George Washington Bridge In brief the relative dispersion is much greater for George Washington Bridge The two populations therefore supplement each other as a field for testing sampling methods

For George Washington Bridge the range of variation and the standard deviation are smaller for the several types of sample from 140 hours to 28 hours, inclusive, than for samples of 16 watches of 8 hours each For samples of 140 separate hours the range of variation of estimates is only 40 percent of that for 128 hours taken in clusters of 8 hours (Figure 3) The figure shows that the base of the curve covers only five groups, which indicates that if the number of hours is large the error is small The scatter of the estimates made from 16 watches of 8 hours each extends over 12 groups, and that of estimates of 70 separate hours covers 7 groups The standard deviation of estimates made from 140 separate hours is only 51 percent as great as from 16 watches of 8 hours each, and that for 70 separate hours is only 73 percent of that of 16 watches of 8 hours

The error of estimates for samples of 35 hours taken systematically is 90 percent of that of estimates from 16 watches of 8 hours each, and that of samples of 35 selected hours is only 79 percent as great as that for 128 hours taken in watches of 8 hours each. Even for 28 separate hours the error is only 99 percent as great as for 128 taken in clusters of 8 For 24 separate hours the error of estimate is 107 percent of that for 16 watches of 8 hours, and for 8 separate hours, it is only 180 percent of that for 128 hours taken in watches of 8 hours each These per-

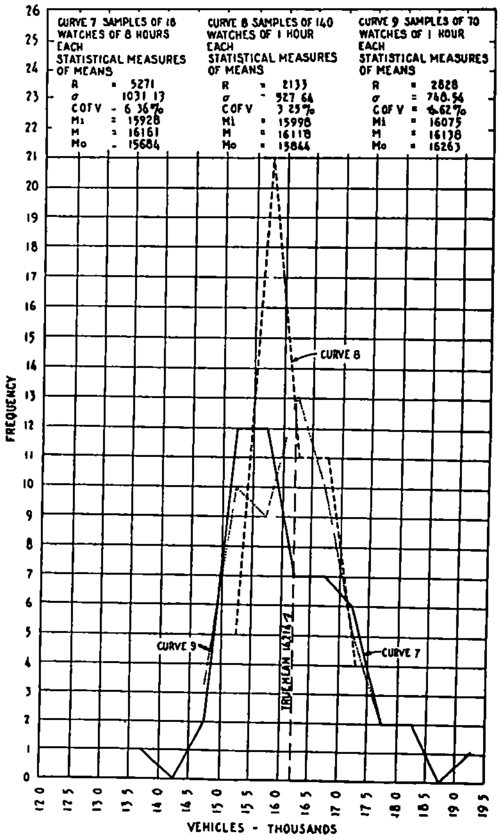

Figure 3. Frequency Distribution of Estimated Means of Daily Volume of Traffic Over George Washington Bridge for 16, 140, and 70 Observations.

centages of error show that for traffic over George Washington Bridge, which has a wide dispersion, the chief factor in the reduction of error is the number of observations and not their size The computations here made involve no theory whatever except that the simple processes of arithmetic are true

The form of the scatter of the estimates is shown in Figure 4 for the two sets of estimates made from 35 separate hours and the one made from 28 hours Curve 10 is scattered through 9 groups, curve 11 through only 7, and curve 12 through 12 groups Curve 11 shows the advantage gained from selecting some of the hours for less than average variation

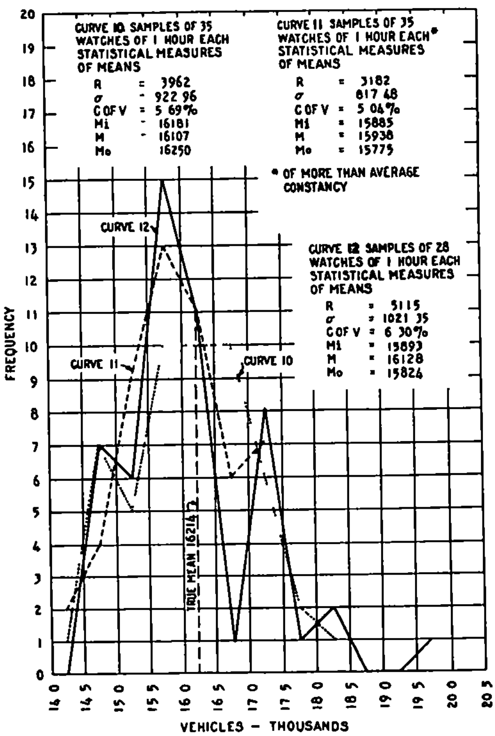

Figure 4 Frequency Distribution of Estimated Means of Daily Volume of Traffic Over George Washington Bridge for 2 Sets of 35 Observations and 1 of 28.

The error of estimate increases as the number of observations decreases, as shown in Figure 5 and indicated by both the range of variation and the standard deviation The standard deviation for 24 hours is slightly greater than for 16 watches of 8 hours each, and that for 16 independent hours is materially greater than for 128 hours in 16 watches of 8 hours each This indicates that the advantage in taking long watches is not as great as in taking more observations

When the number of separate hours in samples is fewer than 35, the width of the frequency curve is greater, and the length of the tails is greater as manifest in curves 12–14, which are based on samples of 28 hours, 24 hours, and 16, respectively The curve for 28 hours is in better form, however, than that for samples of 128 hours taken in 16 watches of 8 hours each, and that for 24 hours

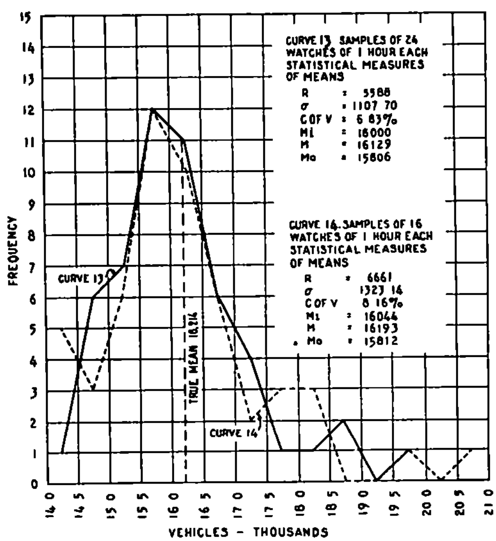

Figure 5. Frequency Distribution of Estimated Means of Daily Volume of Traffic over George Washington Bridge for 24 and 16 Observations.

is almost as good The curve for 16 separate hours is also in surprisingly good form, though the range of variation and the standard deviation increase at a rapid pace as the sample decreases If the number of separate hours is further reduced from 16 to 8, the error of estimate is increased by 41 percent, and the precision of estimate is reduced correspondingly Such a large increase in error should be avoided, of course, in surveys in which precision of estimate is important But the precision of estimate from 8 separate hours is, of course, much greater than from a very few watches of 8 hours each or from any number of hours below 8 taken in a single cluster

PRECISION OF ESTIMATE

The precision of estimate is the inverse of the error of estimate, as the standard deviation of estimates decreases, the precision increases The index of precision relative to that for samples of 16 watches of 8 hours each is shown to be greater than 100 percent for samples of separate hours of from 140 to 28, inclusive For 140 hours the precision is 195 percent of that for 16 watches of 8 hours each, for 70 hours, 138 percent, for 35 hours, 112 percent, for 35 hours of greater than average constancy, 126 percent, for 28 hours, 101 percent, for 24 hours, 93 percent, for 16 hours, 78 percent, and for 8 hours, 55 percent

For samples of 140 separate hours the precision is 195 percent of that for samples of 16 watches of 8 hours each This means that the error of estimates made from 140 separate hours is only roughly one half of that of estimates made from samples of 16 watches of 8 hours each As the number of separate and independent hours is reduced, the precision decreases For samples of 8 separate hours the precision relative to that for samples of 16 watches of 8 hours each is only 55, which is a decline to almost a fourth of the precision of estimate from samples of 140 separate hours

This increase in precision as the number of separate hours increases for a station whose hourly traffic varies as widely as does that of George Washington Bridge is convincing that precision is governed largely by the scatter of the population and the number of independent observations rather than by their magnitude While the measures here presented were computed from tests of the natural data, the general law can be roughly deduced directly from the means Since the relative error is greater for George Washington Bridge than for Holland Tunnel, the precision is correspondingly less for the Bridge than for the Tunnel for the same type of sample.

CENTRAL TENDENCY OF MEANS

If a measure of close estimates be desired, a rough one may be obtained by taking the number of estimates that fall within the central group (the one centering on the true mean) and those lying near it (Tables 1–2) For the 16 watches of 8 hours each of George Washington Bridge (Table 2) only 7 means fall within the central group, 26, within 3 intervals, and 44, within five For the 140 separate hours, however, 11 means (instead of 7) fall within the central group, 43 (instead of 26), within 3 groups, and the entire 52, within 5 groups For 70 hours 13 means of estimate fall within the central group, 32, within 3 groups, and 47, within five groups While there is some chance variation in the measures computed from the samples of moderate size here used, it is clear that the number of separate and independent hours of observation determines largely the number of estimates that fall near the true mean The degree of dispersion of the observations in a given population is a constant, and the only other factor is the size of the observation, which is not a large one The results are also relatively good for both sets of samples of 35 hours and relatively equally so for samples of 28 and 24 hours respectively Since the relative variation is greater for traffic over George Washington Bridge than for that through Holland Tunnel, the smaller samples yield results of less constancy than those derived from the same types of sample for Holland Tunnel That this difference is not great, however, is an indication that the larger the number of observations the greater is the precision and reliability of estimate

As in the case of Holland Tunnel so for the Bridge the range of variation, the standard deviation, and the central tendency are all consistent in showing of the best method of computation of means of estimate

The great concentration and the small deviation and range of variation of means

TABLE 1

COMPARISON OF SOME STATISTICAL MEASURES OF THE VARIATION OF 52 MEANS COMPUTED FROM SAMPLES DRAWN FROM TRAFFIC THROUGH HOLLAND TUNNEL

(Calendar year 1933)

Number of watches	Hours each	Range of variation		Standard deviation[1] (vehicles)	Coefficient of variation (percent)	Central tendency			Relative precision[2] (percent)
		Number of vehicles	Percent of true mean			Number of means in central groups: One	Three	Five	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
16	8	6,811	22 89	1,287	4 33	9	23	37	100
70	1	3,660	12 30	813	2 73	12	34	44	158
62	1	3,589	12 06	907	3 05	9	26	42	142
35	1	3,934	13 22	824	2 77	12	31	45	156
28	1	5,258	17 67	1,165	3 92	10	28	40	110
24	1	5,553	18 66	1,298	4 36	11	27	38	99
19	1	5,840	19 63	1,490	5 01	7	17	30	86

[1] Of the means from the true mean daily volume for the year

[2] Relative to precision of estimate for 16 observations of 8 hours each

TABLE 2

COMPARISON OF SOME STATISTICAL MEASURES OF THE VARIATION OF 52 MEANS COMPUTED FROM SAMPLES DRAWN FROM TRAFFIC OVER GEORGE WASHINGTON BRIDGE

(Calendar year 1933)

Number of watches	Hours each	Range of variation		Standard deviation[1] (vehicles)	Coefficient of variation (percent)	Central tendency			Relative precision[2] (percent)
		Number of vehicles	Percent of true mean			Number of means in central groups: One	Three	Five	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
16	8	5,271	32 51	1,031	6 36	7	26	44	100
140	1	2,133	13 16	528	3 25	11	43	52	195
70	1	2,828	17 44	749	4 62	13	32	47	138
35	1	3,962	24 43	923	5 69	10	30	41	112
35[3]	1	3,182	19 63	817	5 04	11	30	46	126
28	1	5,115	31 54	1,021	6 30	11	27	41	101
24	1	5,588	34 46	1,108	6 83	11	29	40	93
16	1	6,661	41 08	1,323	8 16	10	28	36	78
8	1	8,197	50 55	1,859	11 46	8	18	29	55

[1] Of the means from the true mean daily volume for the year

[2] Relative to precision of estimate for 16 observations of 8 hours each

[3] Of more than average constancy

that samples of 28 separate hours and larger ones are superior to those of 128 hours taken in 16 clusters of 8 hours each These are the three fundamental tests computed from samples of 140 separate hours are shown in curve (polygon) 8 of Figure 3 in contrast with the smaller central tendency and larger deviation and range of variation of means computed from 16 watches of 8 hours each (Figure 3, curve 7) The tails of curve 7 show rather wide error for a few estimates, but those of curve 8 are so nearly vertical that it is clear that the error of even the worst estimates is small The peak of the modal group of curve 8 is also so sharp as to include 21 of the 52 means The concentration around the true mean is not so great for samples of 70 hours as for 140, but it is far greater than for samples of 16 watches of 8 hours each (Figure 3, curves 7–9) The width of the base of the curves shows graphically the type of sample that is most reliable, and the central group and those adjoining it indicate the number of close estimates

It might be thought that the sampling results for neither Holland Tunnel nor George Washington Bridge would be applicable to the various types of roads of a State or the nation, especially to those of small volume of traffic The method, however, is one of random sampling of hours used for the purpose of estimating the mean annual daily volume of traffic The traffic pattern is not involved, it is simply the problem of sampling a variable for the purpose of estimating the mean (not the pattern) through the day and the year. The variable for hours for George Washington Bridge is much more widely dispersed than for Holland Tunnel, but the results from samples indicate that the same law of precision of estimate applies to both populations with only slight variation in degree While the two populations have very different patterns for the year in both average density and degree of variation from hour to hour, day to day, and month to month, still an increase in the number of observations in a sample results in an increase in the precision of estimate in both cases and to roughly the same degree (Tables 1 and 2).

The variation in the hourly volume of traffic over George Washington Bridge ranges from less than a half of one percent of that of the maximum hour to 100 percent This means that the dispersion ranges from almost zero to roughly 5,000 cars per hour Since the average hour is less than 700 cars, the coefficient of variation, or relative dispersion, is rather great, and that of a very small station would scarcely be greater The bimodal hourly variable obtains in all hourly populations of traffic, the problem of sampling at random for the purpose of estimating the mean daily volume of traffic is therefore similar for volumes of traffic from the smallest to the greatest The error of estimate is determined by (1) the natural dispersion of the observations, (2) the number of observations, and (3) to a moderate extent the length of them In brief, the method of sampling at random applies to all degrees of density of traffic The dispersion of any traffic population is determined by the functions that control traffic volume, but the number and size of observations and the scheme of their selection can be controlled The sampling tests for both Holland Tunnel and George Washington Bridge show that the variation in number and size of observations affects the results in the same direction and roughly to the same extent, and this is likely to be true for roads of all types of traffic and of all volumes with only moderate variation in the error of estimate

TRAFFIC POPULATIONS AND VARIABLES

The results of the sampling tests set forth in this paper are based directly on the ordinary established mathematical methods and calculations and do not rest on the assumption of the normal curve or any assumptions of the form of the frequency distributions of the two populations The results were obtained by computing estimates from samples taken systematically at random, and the error of estimate was obtained by comparison with the true measures of the populations It seems desirable, however, for the purposes of highway traffic surveys to present the frequency curves and other measures of the total populations of (1) the hours, (2) the watches of eight hours, and (3) the days for both Holland Tunnel and George Washington Bridge A brief statement of the population statistics of these populations as variables and of the variation of traffic through the day, the week, and the year is, therefore, deemed an essential part of this study

POPULATION OF HOURLY TOTALS

For Holland Tunnel the frequency distribution of the 8,760 hours for the year 1933 is distinctly a bimodal one The peak on the left of Figure 6 includes most of the night hours and a considerable number of the small traffic hours of the day period The one on the right includes the major portion of the day hours and a few of the larger traffic hours of the night period of both holidays and the larger traffic Sundays The arithmetic mean of the entire hourly universe is slightly smaller than the median, much smaller than the mode computed by the usual group method, and materially smaller than the one computed by the method of moments This order is the reverse of the more common form of frequency curves, and the factor that has drawn the mean to the left of the other measures is the concentration of the small night hours rather far to the left as manifest in the night peak The average of this peak group is low, and the number of the small hours is large pulling the mean, the median, and the night mode to the left of the day mode The form of the curve is not only flat-topped but also slightly gap-topped, or bimodal, (Figure 6) with a kurtosis (B_2) of 2 23 compared with that of the normal of 3 The small kurtosis signifies that the average dispersion is relatively small, which means that the population is more stable than one in the form of a normal curve Measures computed from the population of hours of Holland Tunnel are therefore less dispersed than those computed from a variable of the normal form The heights of the two peaks (Figure 6), which can be smoothed into a plateau, indicate that the concentration of the hours is greater than that for a normal distribution

These statements are not true for George Washington Bridge, as is indicated by the long tail to the right (Figure 7) and a kurtosis of 6 47, or more than twice that of a normal curve Notwithstanding this greater than average variation in the population of the Bridge, it has been found that estimates of the mean hourly volume of traffic can be made with ample precision provided the sample is not too small There are only a small number of large and widely scattered hours, and this extreme instability is not sufficient to prevent a fair estimate of the mean

Although the extreme range of variation for the hourly volume of traffic is greater for George Washington Bridge than for Holland Tunnel, 4,818 vehicles compared with 3,564, or 713 percent of the mean compared with 287 percent, the standard deviation for Holland Tunnel is 699 vehicles in comparison with 524 for George Washington Bridge The cause of the larger (absolute) standard deviation is the greater average hourly flow of traffic for the Tunnel than for the Bridge Nevertheless, the coefficient of variation shows that the relative variation is 78 percent of the mean for the Bridge compared with 56 percent for the Tunnel (Figures 6–7) The narrow absolute range for most of the hours is shown for the Bridge in the narrow part of the curve of Figure 7 in comparison with the wider range of the relatively few observations that appear in the tail on the right of the same curve Of the 8,760 hours 165 exceed 3 times the standard deviation compared with only 9 for Holland Tunnel, and only 2 exceed 6 times the standard deviation with none for Holland Tunnel The coefficient of variation, however, for the Bridge is 38 percent greater than for the Tunnel, and this greater average dispersion is due largely to the wide scatter to the right of a few large hours

It is the few large hours (1 percent of the total) in the population of George

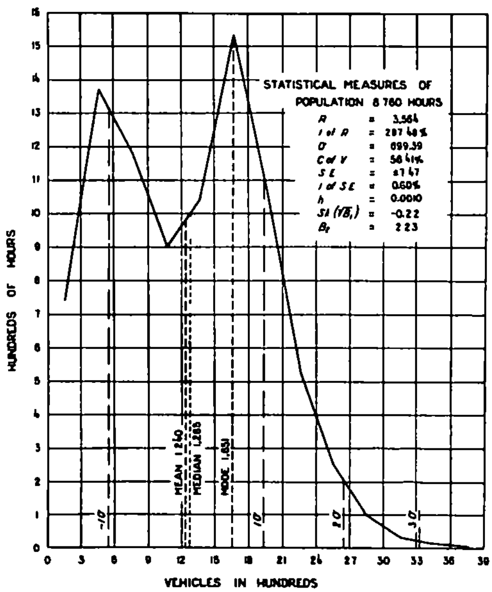

Figure 6. Frequency Distribution of Hourly Traffic Through Holland Tunnel for 1933 (8,760 Separate Hours).

Washington Bridge and the night peak of small hours that have set the coefficient of variation rather high, and the wide variation of a few large hours also causes large error from small samples drawn from the Bridge universe This is the reason that the coefficient of variation is greater for the Bridge than for the Tunnel despite the narrow vertical area of Figure 7 in comparison with the broader one of Figure 6 In the face of the wide scatter in size of hours it is significant that the estimates of the daily volume of traffic for the Bridge do not contain greater error than shown by the measures of dispersion in Tables 2–3

For the Bridge the mean is greater than either the mode or the median, as is the more common order in economic data, but the day mode is greater than the median, which is not true of the order in the case of the normal curve It is the large hours in the tail on the right of the curve (Figure 7) that have offset the night peak on the left and have held the

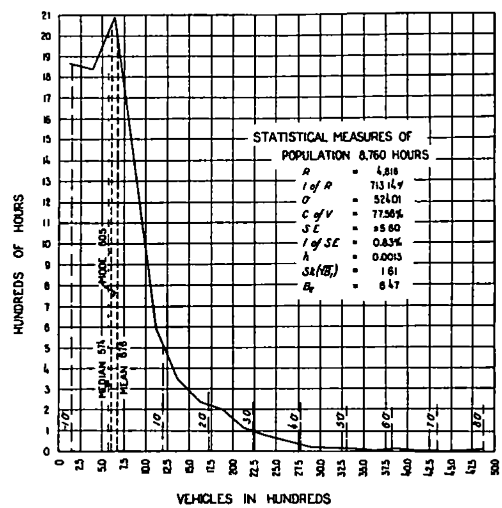

Figure 7. Frequency Distribution of Hourly Traffic Over George Washington Bridge for 1933 (8,760 Separate Hours).

true mean in its usual position The mode computed by the method of moments, however, would fall far below both the mean and the median It is the inclusion of almost 90 percent of all the observations in the first five groups of the distribution of the Bridge that causes the estimates to be very stable for rather large samples, e g, 140 hours (Figure 3, curve 8)

Estimates from a very few observations, however, would be much less dependable because of the inclusion or exclusion on chance of some of the large hours. Even estimates based on 28 separate hours are more dependable (Figure 4, curve 12) than those computed from 16 observations of 8 hours each, a total of 128 hours (Figure 3, curve 7) This is an example of greater dependability of estimates made from samples composed of a large number of observations than from those of a small number, and an analysis of the form of this frequency curve makes it clear that the result conforms to what should be expected

Most records of volume of traffic take the form of the curve of Figure 7 rather than of Figure 6, that is, a large proportion of the hours are small and only a few are very large The estimates made from the data of the Bridge therefore are representative of those that can be made from the records of the majority of the stations, while the results from Holland Tunnel are representative of what may be expected at points of very dense and rather constant traffic The distributions of the two populations (Figures 6–7) extend almost through the entire range of variation of hourly traffic, the Bridge records extending from near zero to a very large volume of 4,842 vehicles per hour

The skewness for the Holland Tunnel population is −0 22, a small negative one compared with a rather large positive one of 1 61 for George Washington Bridge The negative skewness for Holland Tunnel is due to the cluster of night hours on the left combined with the lack of a sufficient scatter to the right to offset it The night mode for the Bridge is offset by the rather long tail to the right resulting in large positive skewness and wide relative dispersion The forms of the two distribution curves indicate that a fair degree of precision of estimate can be obtained from each of the universes from samples of moderate size, and the results of the tests shown in Tables 2–3 confirm this presumption Although the average traffic volume at a majority of the points in a State survey is smaller than that of George Washington Bridge, and although this tends toward a greater coefficient of variation, still the extension of the hourly range of variation for the Bridge is from near zero to 4,842 vehicles, which is almost the entire range from zero to the maximum record for all but the largest stations The results from samples drawn from the population of George Washington Bridge should approach very close therefore to the widest range of error that would

TABLE 3

SOME STATISTICAL MEASURES OF POPULATIONS OF SEPARATE HOURS, EIGHT HOURS, AND 24 HOURS FOR TRAFFIC OF HOLLAND TUNNEL AND GEORGE WASHINGTON BRIDGE

	Units	Watches of Hours			Watches of Hours		
		1	8	24	1	8	24
		Holland Tunnel			George Washington Bridge		
Range	Vehicles	3,564	22,102	30,501	4,818	22,917	33,521
I of R	Percent	287	223	103	713	424	207
σ	Vehicles	699	4,340	5,794	524	3,604	7,544
C of V	Percent	56 4	43 8	19 5	77 6	66 7	46 5
SE	Vehicles	±7 5	±131 1	±303 3	±5 6	±108 9	±394 9
I of SE	Percent	0 60	1 32	1 02	0 83	2 02	2 44
Precision	$\frac{1}{\sigma\sqrt{2}}$	0 001	0 00016	0 00012	0 0013	0 0002	0 00009
Skewness	$\sqrt{B_1}$	−0 22	−0 09	0 56	1 61	1 53	1 39
Kurtosis	B_2	2 23	2 52	3 22	6 47	6 22	4 31

arise from any traffic population, although the coefficient of variation for very small stations might well be considerably greater due to the small means of traffic flow at those points. For estimating traffic volume at even small stations an increase in the number of observations is more important than an increase in their magnitude, and the greater the dispersion the greater the desirability of a large number of observations in the sample.

As the total hourly traffic at a point approaches zero, the maximum hour and the range are also usually reduced. The records of sampling tests for George Washington Bridge show that some 30 to 50 separate and independent hourly records are sufficient to reduce the error of the estimates of daily volume of traffic to narrow limits, and the frequency charts of the two populations combined with the facts stated in this paragraph indicate that satisfactory estimates for stations of all densities of traffic and of all degrees of variation can be made by the methods used in this study. The methods here used may be applied therefore to state-wide surveys and even to those of National scope with a probability that the range of error of estimate will not greatly exceed that found in sampling tests from the records of George Washington Bridge.

STABILITY OF HOURLY RELATIVES AMONG STATIONS

The percentage that the average traffic for each hour of the day is of the hourly average for the year is quite constant as shown in Figure 8 for Holland Tunnel and George Washington Bridge. The percentage curves follow each other so closely that the curve for either crossing could be applied to the other with only moderate error. For the period 9 a.m. to 3 p.m. a larger percentage is shown for Holland Tunnel, and for that from 4 to 11 p.m. the traffic is relatively larger for George Washington Bridge. The curves are computed from averages for each hour and the annual average, they do not show the dispersion of individual hours. The relative constancy of the averages for each hour of the day indicates that if samples are made large enough, relatives may be established for use at a given station or at a group of stations of rather large variation with moderate though not negligible error.

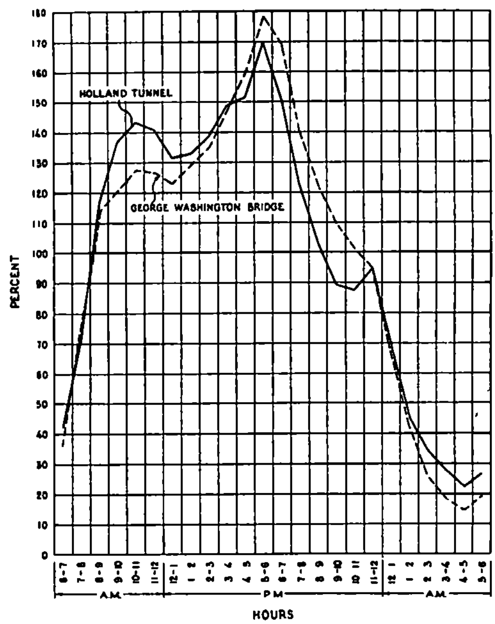

Figure 8. Percentage That the Average Traffic for Each Hour of the Day Is of the Hourly Average for the Year, for Holland Tunnel and George Washington Bridge, 1933

POPULATION OF WATCHES OF EIGHT HOURS

The curve of frequency distribution for the watches of 8 hours for the population of Holland Tunnel is bimodal in form to a more marked degree than is the curve of the frequency of the population for the hours, as is manifest in Figures 6 and 9. For George Washington Bridge, however, the night mode is almost smoothed out by the overlapping of the large night periods and the smaller day ones (Fig. 10).

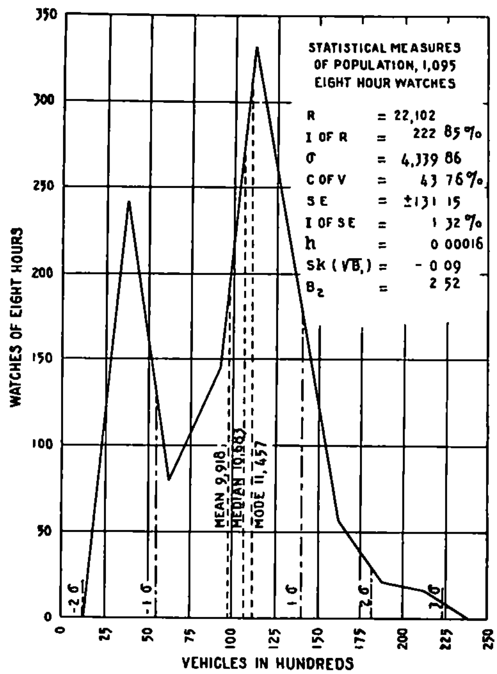

Figure 9 Frequency Distribution of Eight-hour Watches of Traffic Through Holland Tunnel for 1933 (1,095 Eight-hour Watches).

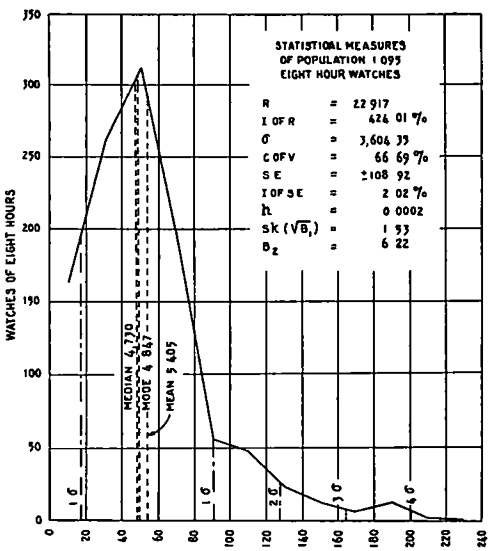

Figure 10 Frequency Distribution of Eight-hour Watches of Traffic Over George Washington Bridge for 1933 (1,095 Eight-hour Watches).

The relative dispersion is not as great for the populations of watches of 8 hours as for those of one hour, but the difference is not very great For Holland Tunnel the coefficient of variation is 43 8 percent compared with 56 4 percent, for George Washington Bridge it is 66 7 percent against 77 6 percent The skewness is also slightly less for the watches of 8 hours than for hours, and the kurtosis is a trifle nearer the normal The index of the range of variation is also only moderately less for watches of 8 hours than for hours These comparisons show only small advantage in sampling from populations of watches of 8 hours over that from hours, while the time required is much greater for the longer watches (Table 3 and Figures 6–7 and 9–10)

POPULATION OF DAILY TOTALS

When the entire day is taken as a unit for the watch, the bimodal form disappears from the frequency curve of the population (Figs 11–12), but the dispersion of the populations of even these larger totals is rather large The index of range is 102 5 percent for Holland Tunnel compared with 287 5 for the index of hours, and the coefficient of variation is 19 5 percent compared with 56 4 percent for the coefficient of hours The index of range for watches of eight hours is 222 9 percent for Holland Tunnel and 424 percent for George Washington Bridge If the size of the watch is changed to 24 hours, the indexes of range are reduced to 102 5 percent and 207 percent, respectively In brief, the increase in size of watch to 24 hours eliminates the bimodal effect contained in watches of one hour and those of 8 hours, but still embodies wide dispersion A watch of 24 hours includes almost as much counting time as 28 separate and independent hours scattered through the hours of the day, the days of the week, and the months of the year, but the error of an estimate made from a single day's record is far greater than one made from 28 separate and independent hours A considerable number of days would be necessary to obtain as great precision of estimate as from 28 separate hours taken at random

The comparative dispersions of the daily totals for the Tunnel and the Bridge are shown in Figures 11 and 12 Even the range of variation is greater for the Bridge, and the index of the range is more than twice as great as for the Tunnel ` The coefficient of variation for the Bridge is also more than twice as great as that for the Tunnel The long tail in Figure 12 contains 40 Sundays, 11 Saturdays, 5 holidays, and 7 other days, each of which exceeds 22,000 vehicles, although the true mean is only 16,214 These numbers indicate the proportion of days of dense traffic for each class named They are the days that constitute the wide scatter to the right in Figure 12 and are also a large factor in determining the index of range and the coefficient of variation for the days for the Bridge, which are more than twice as great as those relatives for Holland Tunnel, and these dense traffic days also add to the positive skewness and the kurtosis (B_2 in Figure 12) The 63 days of very dense traffic are due largely to the diversion of traffic from Holland Tunnel and other crossings on days of congestion and to the traffic capacity and popularity of the Bridge on Sundays, Saturdays, and holidays

The measures of the populations considered here have shown how the populations differ and why the sampling results of the first part of the paper were so divergent The sampling results had shown that the increase in the number of separate and independent observations was the largest factor in the precision of estimate and that the length of the watch and the dispersion of the observations were factors of less importance The population measures have thrown additional light on the forms of the population curves and the problems of estimating the mean daily volume of traffic from the several populations

The measures of the populations of watches of one hour, eight hours, and

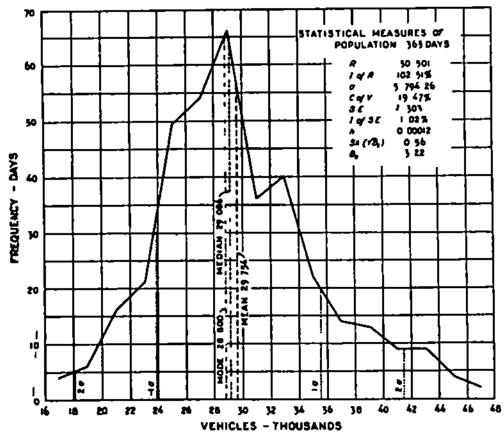

Figure 11 Frequency Distribution of Daily Traffic Through Holland Tunnel for 1933 (365 Days)

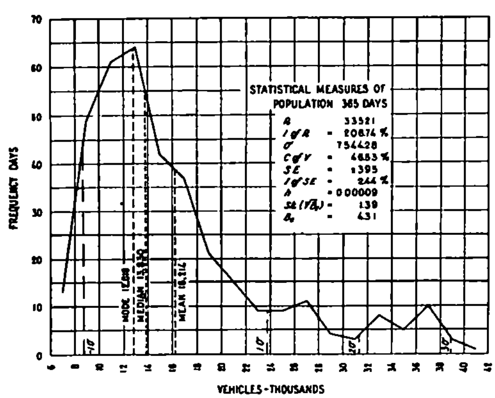

Figure 12. Frequency Distribution of Daily Traffic Over George Washington Bridge for 1933 (365 Days).

24 hours have made it clear that even with the longer watches it is necessary to include a considerable number of watches to obtain a satisfactory estimate, and of course any considerable number of days includes a large amount of time

It is clear that of the populations here analyzed the most efficient one from which to sample is the one of the shortest watch

A further study of watches of less than an hour is now being undertaken The basic data are being collected by automatic traffic counters in periods of five minutes, from which watches of all combinations of five-minute periods can be constructed From an analysis of the data for periods between five minutes and one hour it is hoped that the length of optimum watch can be determined on the basis of time required and precision of estimate obtained

APPENDIX 1 GLOSSARY OF STATISTICAL MEASURES

The following symbols of statistical measures appear on the charts.

M	Arithmetic mean
Mi	median
Mo	mode
R	range of variation
I of R	index of range of variation
σ	standard deviation, the root-mean-square deviation about the arithmetic mean
C of V	coefficient of variation
S E	standard error
I of S E	index of the standard error
Sk	skewness, based on Pearson's formula $sk = \frac{M - Mo}{\sigma}$
$\sqrt{B_1}$	skewness, computed by the method of moments
B_2	kurtosis, the relative flatness or peakedness of the curve
h	index of the precision of estimate $h = \frac{1}{\sigma\sqrt{2}}$