NTIA REPORT 79-21

Statistical Methods for Estimating Time and Rate Parameters of Digital Communication Systems

Edwin L. Crow

U.S. DEPARTMENT OF COMMERCE
Juanita M. Kreps, Secretary

Henry Geller, Assistant Secretary
for Communications and Information

June 1979

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

STATISTICAL METHODS FOR ESTIMATING TIME AND RATE PARAMETERS OF DIGITAL COMMUNICATION SYSTEMS

Edwin L. Crow*

Statistical methods are provided for estimating time and rate parameters of digital communication systems according to the specifications of the Proposed Federal Standard 1033. The methods may be applied to delay, disengagement, transfer and service times, and to transfer rates. Some ARPANET and Bell System data are examined for the form of distributions of times. The properties of various types of estimators are discussed, and the sample mean is recommended. Approximate confidence limits are given for time and rate parameters. A step-by-step procedure is provided for designing an experiment to estimate time and rate parameters with prescribed accuracy.

Key words: ARPANET; autocorrelation; confidence limits; design of experiments; distribution-free estimation; robust estimation

1. INTRODUCTION

The transmission of data over digital communication systems can be characterized in part by various time intervals: delay times, disengagement times, service times between outages, outage durations, and bit, block, and message transfer times. Transfer rates are derived from the latter times. Times of a particular type may depend on various factors or conditions (e.g., a busy hour on weekdays versus Sunday morning), but within specified conditions they vary in an apparently random fashion and hence can be represented by a random (stochastic) process or even a single theoretical probability distribution. However, the theoretical or long-term distribution can be estimated only from a finite amount of data. Criteria for specifying or judging systems should include methods for measuring all time intervals of interest and analyzing the data, and the methods should yield sufficiently accurate results to conclude satisfactory or unsatisfactory performance.

This report considers various possible approaches to the analysis of time interval data and recommends one that amounts in essence simply to

*The author is with the Institute for Telecommunication Sciences, National Telecommunications and Information Administration, U.S. Dept. of Commerce, Boulder, CO 80303.

calculating the sample mean and confidence limits for it. The recommended procedure is outlined in the Summary (Sec. 9). It is a procedure based on practicality as well as theory. All time intervals (or times, as we shall often state for brevity) have the common property of being nonnegative and hence of having more or less asymmetrical or skewed distributions. In Section 4 we examine some data on their shapes but find no information to distinguish the distributions of the various times mentioned above with respect to shape. Our discussions and results therefore apply equally well to all of the time and rate performance parameters, but the required amount of measurement (sample sizes) to achieve the specified accuracy may differ from one parameter to another.

The Proposed Federal Standard 1033 takes the average (or mean) time to characterize each time distribution. That is probably the most readily understood and calculated criterion but conceivably could be improved upon in theory. The question is discussed briefly in Section 2.

The classical methods of data analysis are based upon the assumption of particular forms of probability distributions, such as normal (Gaussian), lognormal, and gamma, and they tend to be optimum only for a particular form. Ideally methods should have properties independent of the form of distribution. The problem of whether to use such "distribution-free" methods or distribution-dependent methods is discussed in Section 3. Some of the data available on time interval distributions of communication systems are examined in Section 4.

We end up recommending use of the sample mean as the estimator of the population or theoretical distribution mean. The sample mean has been severely criticized by mathematical statisticians, primarily because of its sensitivity to wild or contaminating outlying observations, that is, observations not belonging to the distribution of interest. This disadvantage is disposed of for delay times, disengagement times, and block and message transfer times by the definition in the proposed Federal Standard 1033 that the maximum time is three times the nominal (presumably a specification) time; any trial resulting in a larger time is counted as a failure and enters into the performance assessment through the estimation of a probability. (The estimation of probabilities in digital communication systems, especially confidence limits, is discussed by Crow and Miles, 1977.) The statistical properties of the sample mean as an estimator of the population mean are discussed more fully in Section 5.

The more concrete results are given in Sections 6, 7, and 8: confidence limits for time parameters, that is, for mean times, given the data, are discussed in Section 6; confidence limits for transfer rates derived from mean times are shown in Section 7; and the way to go about designing a test or experiment to estimate such times or rates with a specified accuracy is discussed in Section 8. It would be more logical to give the design material first, but the design depends on the formulas for confidence limits.

2. CHOICE OF CRITERION

Is the mean, median, or 90% point of the distribution of times (such as access times) of primary interest? Since it is often desirable that such times be short, one might wish to require that almost all times are shorter than a specified tolerance and therefore take the 90%, 95%, or 99% point as his parameter. The proposed Federal Standard partially blunts such a desire by cutting off most time distributions at three times the nominal value. Recent work on data communication systems seems to have emphasized means (Bell System, cf. Duffy and Mercer, 1978; ARPANET, cf. Kleinrock, 1976, pp. 456, 480-482; and Payne, 1978), although histograms of entire distributions are commendably also presented. The histograms give full information (except possible time dependence), but that is probably not needed for a specification, and it is desirable to have a specification no more complicated than necessary. Since the total time wasted over the life of a communication system is simply a multiple of the mean time and longer delays may not be individually crucial, it may well be justified to take the easily understood mean time as the distribution parameter to be estimated. However, to do this with error bars or confidence limits usually requires estimating one or more other parameters of the distribution also.

3. DISTRIBUTION-FREE VERSUS DISTRIBUTION-DEPENDENT ESTIMATION

If one knows the analytic form of a population distribution, he may be able to estimate the ("true") mean of the population more accurately (e.g., with less variance) than if he doesn't know it. The sample (arithmetic) mean is a minimum variance unbiased estimator of the population mean if the distribution is normal (or gamma also for large sample sizes) but not if the distribution is lognormal. An alternative optimal estimator is available in the lognormal case. Since one does not know the distribution form with certainty, what should one do?

There are distribution-free, or nonparametric, estimation methods. The sample mean is always an unbiased estimator of the population mean but not always efficient (i.e., of minimum variance) and hence is not considered a nonparametric estimator. The sample median is a median-unbiased estimator of the population median, and distribution-free confidence limits for the population median are easily obtained in the form of particular ordered observations, but they do not apply to the population mean unless it coincides with the median. Furthermore, the median is rather inefficient if the population is normal (only 64% efficiency as measured by ratio of variances). (This raises the question of its efficiency for the lognormal and gamma distributions; it is doubtful that this question has been answered.)

If individual measurements are not expensive - and they may not be, the set-up time perhaps being so much larger than the individual measurement time - then the statistical inefficiency of the mean for some distributions (or of the median either for that matter) is not practically important. Plenty of measurements can be taken, and the Central Limit Theorem can be invoked to state that 95% (say) confidence limits for the population mean are (approximately, but closely enough) $\overline{x} \pm 2s/\sqrt{n}$, where $\overline{x}$ is the sample mean, s is the sample standard deviation, and n is the sample size. (It has been tacitly assumed that essentially independent observations can be taken within the specified conditions and that the communication system and its environment have been compartmentalized for characterization and consequent separate testing so that this essential independence holds. The procedure for departures from independence is discussed in Sections 6 and 8.)

A type of estimation method intermediate to the nonparametric and parametric types is the "robust" type. The properties of these methods are not independent of the form of distribution but are relatively insensitive to it, especially to outlying observations that may not belong to the population of interest. For example, the "midmean," or 50% trimmed mean, defined as the mean of the middle 50% of the ordered measurements, has efficiency of 80% over quite a range of symmetrical distributions.

If there is a substantial background of data and experience with the type of time measurements to which the Federal Standard is to be applied and all of the data are consistent with one form of distribution, then it is reasonable to consider that the form of distribution is known and proceed to use the optimal estimation methods for that form. An interesting discussion by Cox (1958) on the comparative merits of parametric and nonparametric inference favors this sort of conclusion but suggests the auxiliary calculation of the nonparametric statistic as a check. The next section will therefore be devoted to examining the immediately available data for the form or forms of distribution.

4. EXAMINATION OF BELL SYSTEM AND ARPANET HISTOGRAMS OF TIMES FOR APPROXIMATE FORM

Duffy and Mercer's study (1978) of direct-distance-dialing call attempts by telephone customers presents histograms of distributions of several different random times:

(a) from end of dialing to abandonment without a system response (their Fig. 4),

(b) from end of dialing to the first system response after dialing a valid number (Fig. 5),

(c) call attempt time for incomplete attempts (Fig. 6),

(d) call attempt time for successfully completed attempts (Fig. 7).

The order of decreasing asymmetry or skewness is a, c, b, and d; (a) is quite J-shaped, probably approximately exponential, judged from visual inspection, (b) and (c) are distinctly asymmetrical but with a single central mode aside from sampling fluctuations, and (d) is almost normal in shape, with a slightly longer tail on the right.

Duffy and Mercer also present means, medians, standard deviations, and 10% and 90% points of 16 to 22 time distributions (Table II). One can judge the degree of asymmetry of these in three different ways: comparing mean and median, comparing mean and standard deviation, and comparing the 10%-to-50% and 50%-to-90% intervals. Based on these three criteria jointly, one would roughly classify three of Duffy and Mercer's distributions as symmetrical, 15 as asymmetrical, and four as of doubtful classification.

Kleinrock (1976, Fig. 6.16) shows three histograms of round-trip delay time in ARPANET, for 1 hop, 5 hops, and 9 hops. The latter two are distinctly asymmetrical while the 1-hop distribution is approximately symmetrical despite having much shorter times.

Payne (1978) presents a histogram of 200 successful access times which appears roughly symmetrical except for one large outlying time.

The general conclusion is that most of the group of communications time distributions examined have the asymmetrical shape expected of a random variable bounded to the left by zero and unbounded on the right, but an appreciable proportion are symmetrical or almost so. The form is generally simple, with a single maximum, and would seem from visual inspection to be satisfactorily represented by a gamma or lognormal model with two, or occasionally three, parameters, some even by the normal model. It does not seem worthwhile at this point to pursue the precise fitting of models further.

An important influence on the effective distribution of several time variables is the decision in the proposed Federal Standard 1033 to regard all times greater than three times the "nominal value" as unacceptable and to declare a failure on any trial resulting in such a time (Sec. A.3.3.1 for access time, Sec. A.3.3.2 for block transfer time, and Sec. A.3.3.3 for disengagement time). It follows that there are no "outlying observations." Hence there is no substantial need to consider robust estimators.

5. ADVANTAGES AND DISADVANTAGES OF THE SAMPLE MEAN AS AN ESTIMATOR OF THE POPULATION MEAN

The sample mean is now proposed as the estimator of mean time parameters of communications systems regardless of the shape of the distribution. It has the following advantages:

(a) It is an <u>unbiased</u> estimator and a <u>consistent</u> estimator (i.e., converging in probability to the population mean as the sample size increases), regardless of the distribution, under the assumptions that the variance of the distribution is finite, which is satisfied if the distribution is bounded, as it surely must be in practice, and that successive observations are not completely correlated.

(b) It is an <u>efficient</u> estimator (i.e., contains all the sample information) if the distribution is either normal or gamma, which may be sufficient to describe the data satisfactorily.

(c) It is very well known and simple.

(d) Its standard deviation is $\sigma/\sqrt{n}$ if the observations are uncorrelated, regardless of the distribution, where n is the sample size and σ is the population standard deviation, which can be estimated by the sample standard deviation s.

(e) It is asymptotically (as n becomes infinite) - and hence approximately for finite sample size - normally distributed. (This is, in simple form, the Central Limit Theorem of probability theory and mathematical statistics.) Approximate confidence limits for the population mean can therefore be given as mentioned earlier and are stated more fully in Section 6.

With respect to advantage (a) it should be noted that the sample median, the best-known nonparametric estimator of location, is not an unbiased estimator of the distribution mean if the distribution is asymmetric. Neither are the various "robust" estimators proposed, such as the trimmed mean. In fact, robust estimation of location has been developed very little for asymmetric distributions except for small deviations from asymmetry that are regarded as contamination of a symmetric distribution the mean of which is the parameter of interest.

The sample mean has the following disadvantages:

(i) It is sensitive to outlying observations. If such outlying observations are not part of the population of interest, then this is a disadvantage but not necessarily otherwise.

(ii) It is not an efficient estimator for any distribution other than the normal or gamma, though the loss is probably moderate for moderate departures from those shapes. (Bondesson (1977) showed recently that if $\bar{x}$ has high efficiency then the distribution function must be "close" to the normal or to the gamma distribution function.) Bury (1975, p. 290, Fig. 8.4) shows the efficiency of $\bar{x}$ as a function of the shape parameter σ^2 of the lognormal distribution; the efficiency decreases from 100% to 63% as σ^2 increases from 0 to 2.

(iii) Except for the normal distribution the sample standard deviation is not an efficient estimator of the population standard deviation, so the confidence limits on the population mean are then not estimated as well as might be if the distribution shape were known. For example, Bury's Fig. 8.4 shows the efficiency of the sample standard deviation decreases rapidly as σ^2 increases in the lognormal case.

These disadvantages are not considered sufficient to preclude use of the sample mean in light of present knowledge.

6. CONFIDENCE LIMITS FOR TIME PARAMETERS

It is assumed that a sample of n uncorrelated measurements $w_1, w_2, \ldots, w_n$ of the time intervals of interest, say access time, is available and that the long-term system or population mean or average time, W, is to be estimated. For the reasons discussed in the preceding sections W is estimated by the sample mean

$$\bar{w} = \frac{1}{n} \sum_{i=1}^{n} w_i . \tag{6.1}$$

The scatter of the sample is measured by the sample standard deviation, which is the square root of the sample variance,

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (w_i - \bar{w})^2$$

$$= \frac{1}{n-1}[\sum w_i^2 - \frac{1}{n}(\sum w_i)^2]. \tag{6.2}$$

This is immediately available simultaneously with $\bar{w}$ on many pocket calculators. The sample standard deviation is an estimate of the corresponding population standard deviation, σ.

The variability of $\bar{w}$ is often measured by its "standard error",

$$s_{\bar{w}} = s/\sqrt{n}. \tag{6.3}$$

The corresponding population characteristic, $\sigma_{\overline{w}} = \sigma/\sqrt{n}$, is the standard deviation of the theoretical sampling distribution of all possible sample means of which the above $\overline{w}$ is the one observed. That theoretical distribution of sample means is approximately normal (arbitrarily closely for n sufficiently large) even if the distribution of measurements w_i is not.

It is desired to provide limits that contain W with high probability in repeated application. Such limits are called confidence limits. If σ is known and the distribution of measurements w_i is known to be normal, then $100(1-2\alpha)$ confidence limits for W are given by

$$\overline{w} \pm z_\alpha \sigma/\sqrt{n}, \tag{6.4}$$

where z_α is the upper 100α percent point of the standard normal distribution. For

50% confidence,	α = .25,	z_α = 0.6745
80%	.10	1.282
90%	.05	1.645
95%	.025	1.960
99%	.005	2.576

The 95% "confidence level" (or confidence coefficient) is most often used (though it is an arbitrary choice).

If σ is not known but the distribution of measurements is known to be normal, then $100(1-2\alpha)$ percent confidence limits for W are given by

$$\overline{w} \pm t_{n-1,\alpha} s/\sqrt{n}, \tag{6.5}$$

where $t_{n-1,\alpha}$ is the upper 100α percent point of the Student t distribution with n-1 degrees of freedom (d.f.). A table of $t_{n-1,\alpha}$ is given in most statistics books and is abstracted here:

n	90% confidence	95% confidence
2	6.314	12.706
3	2.920	4.303
4	2.353	3.182
5	2.132	2.776
10	1.833	2.262
20	1.729	2.093
40	1.685	2.023
∞	1.645	1.960

In practice the distribution of measurements cannot be known to be normal, and time intervals, being nonnegative, cannot be normally distributed. However, since the distribution of sample means is approximately normal for large n whatever the distribution of measurements, the confidence limits (6.5) hold approximately. Because of the approximation, the three decimal places listed above are unnecessary or even inappropriate, and "two sigma" limits are often close enough for 95% limits, though they are on the optimistic side for small samples. Examples of the use of (6.4) and (6.5) are given at the end of Section 7.

If the observations are correlated, the situation is much more complicated. The one simple solution offered is outlined a little more fully in Section 8 and is based on equally spaced time separations of the measured time intervals, which are assumed to be short relative to the separations between measurements. Then the half-length of the confidence interval in (6.4) and (6.5) must be multiplied by the factor $(1+\rho_1)/(1-\rho_1)$, where ρ_1 is the autocorrelation between successive measurements and is estimated by (8.10).

7. CONFIDENCE LIMITS FOR BIT AND BLOCK TRANSFER RATES

"Block transfer rate" is defined in Section A.3.3.2 of Proposed Federal Standard 1033 as the total number, B, of successful block transfers counted during a performance measurement period divided by the duration, w, of the period. (A similar definition applies to bit transfer rate.) The rate calculated from a particular performance period will be denoted by the lower case letter r, while the time rate of the system, which would be obtained from an infinitely long performance measurement period, will be denoted by R. Thus

$$r = B/w, \tag{7.1}$$

whereas

$$R = \lim_{w\to\infty} (B/w)$$

Both numerator and denominator of this fraction could be random variables in general, but a performance measurement period has been defined as the time interval to transfer a "message" (Seitz and McManamon, 1978, page 137), and a message is a specified set of blocks. Hence the numerator would seem to be a specified constant, but there may be slight variations because of inability to end the period instantaneously, i.e., at a specified number of blocks; this reason would seem to apply more strongly to bits. In any case the random variation of the denominator is so much greater than that of the numerator that the latter can be neglected.

In any case since the estimate of block transfer rate is a nonlinear function of the duration of the period, it is a biased estimate of the true rate, R, of the system. A heuristic demonstration of this may be worthwhile. In (7.1) we take a constant number B_1 of block transfers and consider the theoretical distribution of all possible times w_1 to transfer B_1 blocks. Let W_1 be the mean, or expected value, of that distribution; that is, $W=E(w)$. Since R is the true transfer rate, we assume that $W_1=B_1R$. Any one random time w_1 is an unbiased estimate of W_1 by definition. More generally, the mean of any finite sample of w_1's, say $\bar{w}_1$, is also an unbiased estimate of W; that is, $E(\bar{w}_1)=W_1$. What follows is true for any unbiased estimate of W_1, but we shall use the notation $\bar{w}_1$. Let $\bar{w}_1=W_1+d$. Then $E(d)=0$ and $E(d^2)$ is the variance of $\bar{w}_1$, denoted by V_1. We can write

$$r_1 = \frac{B_1}{\bar{w}_1} = \frac{B_1}{W_1+d} = \frac{B_1}{W_1(1+d/W_1)}$$

$$= \frac{B_1}{W_1}\left(1 - \frac{d}{W_1} + \frac{d^2}{W_1^2} - \ldots\right).$$

Then

$$E(r_1) = R\left(1 - 0 + \frac{V_1}{W_1^2} - \ldots\right)$$

$$\neq R$$

if it is accepted that the infinite series does not equal 1, and that seems plausible since the second-order approximation is not equal to 1. Hence r_1 is not an unbiased estimate of R.

Of course, the longer the period is the less the bias is. In practice, periods from several different days should be used, and a combined period can be obtained by adding together the several numbers, B_i, of block transfers and by adding together the several durations, w_i. Then the estimate of transfer rate is

$$\hat{R} = \frac{\sum_{i=1}^{m} B_i}{\sum_{i=1}^{m} w_i} = \frac{m\overline{B}}{m\overline{w}} = \frac{\overline{B}}{\overline{w}} . \tag{7.2}$$

Under the assumption that all B_i are equal, confidence limits for the true rate R can be derived from the confidence limits W_L and W_U for the population mean duration W, which are given in Section 6. They are simply

$$R_L = \overline{B}/W_U \quad \text{and} \quad R_U = \overline{B}/W_L . \tag{7.3}$$

In the same way it follows that $\hat{R}$ is a "median-unbiased" estimate of R; i.e., it will be larger than R in 50% of the cases it is calculated (infinitely many cases being assumed), and less than R in 50% of such cases, even though it is biased in the usual sense that its mean over such cases will not be R.

"Block rate efficiency" is defined in Section A.3.3.2 as the ratio of the product of block transfer rate and average block length, denoted by L here, to the signalling rate of the communication service, R(max):

$$Q = RL/R(\max). \tag{7.4}$$

As implied by the notation, Q, like R, is a characteristic of the communication system (or service), and is merely estimated by a measurement from a single performance period. If L and R(max) are known constants of the system, then confidence limits for Q follow directly from those for R in (7.3):

$$Q_L = R_L L/R(\max), \quad Q_U = R_U L/R(\max). \tag{7.5}$$

Example 1. Consider the nonswitched private line system in Table 1.1 of Kimmett and Seitz (1978), for which the block transfer rate is given as 5 blocks per second. Suppose that a message is specified as 10 blocks and that in a single performance measurement period the time to transfer the message is measured to be 2.13 seconds. The estimated block transfer rate from this measurement would be $\hat{R}$=r=10/2.13=4.69 blocks per second. We can not say how accurate an estimate this is without further information. If it were known that the standard deviation of such transfer times is 0.4 seconds and a normal distribution is assumed, then 95% confidence limits for the mean transfer time W from the single measurement would be, by (6.4),

$$2.13 \pm 1.96(0.4) = \begin{Bmatrix} 2.91 \\ 1.35 \end{Bmatrix} \text{ seconds,}$$

and by (7.3) the corresponding limits on R would be 3.43 and 7.43 blocks per second. If in fact individual transfer times are gamma distributed with 50 degrees of freedom (d.f.) (which corresponds to the combination of a time mean of 2 seconds and σ=0.4 seconds) rather than normally distributed, then the correct 95% confidence limits for W based on that knowledge (with 2-1/2% chance of being wrong in each direction) are 1.49 and 3.29 seconds (derived from a table of chi-squared percentage points, details not being shown), and those for R are 3.04 and 6.71 blocks per second. The change from the previous limits shows the effect of even a small amount of asymmetry of the distribution of times. (Neither one of these calculations takes into account the convention that trials that result in transfer times greater than three times the nominal are counted as failures in the Proposed Federal Standard 1033.)

Example 2. Same as Example 1 except that measurements are taken in 9 independent performance measurement periods. The total time to transfer the 90 blocks is found to be 19.17 seconds (taken artificially as 9x2.13 so that the results will be centered the same as in Example 1 for comparison). Then by (7.2) $\hat{R}$=90/19.17=4.69 blocks per second, but now the 95% confidence limits on W from the normal assumption are

$$2.13 \pm 1.96(0.4)/\sqrt{9} = \begin{Bmatrix} 2.39 \\ 1.87 \end{Bmatrix} \text{ seconds.}$$

and the corresponding limits on R are 4.18 and 5.35 blocks per second. Making use of knowledge of the gamma distribution with 50 d.f. would now change the limits on W only to 1.90 and 2.47 seconds and those on R only to 4.04 and 5.26 blocks per second. Not only are the limits narrower than in Example 1, but they are less dependent on knowledge of the shape of the distribution, a result of the Central Limit Theorem.

Example 3. Same as Example 2 but instead of relying on a given prior value for the standard deviation σ of transfer times, it is calculated from the 9 observed times. For the purpose of comparison assume it turns out to be 0.400 seconds. Then the 95% confidence limits on W are somewhat wider than before because of the uncertainty in the estimate of σ. Based on the assumption of the normal distribution of transfer times we need to use the Student t percentage point with 8 d.f. rather than the normal percentage point, that is, 2.306 rather than 1.96. Hence the limits on W are, by (6.5),

$$2.13 \pm 2.306(0.400)/\sqrt{9} = \begin{Bmatrix} 2.44 \\ 1.82 \end{Bmatrix} \text{ seconds}$$

and those on R are 4.10 and 5.49 blocks per second.

8. DESIGN OF EXPERIMENT TO ESTIMATE TIME AND RATE PARAMETERS

To be specific, we shall consider a distribution of access times, but the discussion will apply equally well to any other time parameters. The purpose of the experiment is taken to be the estimation of the mean access time with specified accuracy and confidence for a specified time period of the day, week, and year and a specified class of source-destination user pairs of the communication system. All user pairs may be of interest, but if they can be classified into subsets with predictably different access times, such as close neighbors and distant users, it would be poor design to lump them together. Likewise, traffic and hence the distribution of access times may be different for different hours of the day, week days and weekends, and perhaps months of the year. Of most interest may be the busiest hour of a week day. If the busiest hour is not known, a preliminary experiment to determine that should be conducted. Such an experiment may be necessary anyway to estimate the standard deviation σ of access times, because design of the final experiment to achieve specified accuracy requires prior knowledge of σ.

The preliminary experiment may be taken to consist of n_1 (say about 10, more if they are not expensive) access time measurements $w_1, w_2, \ldots, w_{n_1}$ at equal intervals within each hour of a day. For each hour the sample mean and standard deviation are calculated:

$$\bar{w}_1 = \frac{1}{n_1}\sum_{i=1}^{n_1} w_i, \quad s_1 = \left[\frac{1}{n_1-1}\sum_{i=1}^{n_1}(w_i-\bar{w}_i)^2\right]^{1/2}. \tag{8.1}$$

The hour (or hours) of interest may be selected as that one with the longest $\bar{w}_1$, or the largest s_1, or the most messages transmitted if that is available, or by some other criterion. Such selection introduces some bias, since the random variable used may be larger than average for the hour picked just by chance. Hence it is desirable to have prior knowledge of the time period of interest.

We wish to estimate the true mean access time W with error less than or equal to a specified value A with confidence level $1-2\alpha$. A rigorous two-stage sampling plan for doing this was given by Stein (1945) if the distribution sampled is normal and the observations are independent. We assume that the distribution is sufficiently close to normal to apply the plan, and we can rely on the Central Limit Theorem to claim a good approximation even if the sample distribution is not normal. We will consider alternatives to the independence assumption later. We follow Seelbinder's (1953) statement of Stein's method. A confidence interval for W can be calculated from the above initial sample. The half-width of this interval is $t_\alpha s_1/\sqrt{n_1}$, where t_α is the Student t 100α upper percentage point with n_1-1 degrees of freedom (d.f.). If

$$t_\alpha s_1/\sqrt{n_1} \leq A, \tag{8.2}$$

then no further observations need be taken, as the desired accuracy has already been attained with the initial n_1 observations. If $t_\alpha s_1/\sqrt{n_1} > A$, then n_2 additional observations w_i are taken, where n_2 is the smallest integer satisfying

$$n_2 \geqq t_\alpha s_1{}^2/A^2 - n_1. \tag{8.3}$$

The true mean W is estimated by

$$\bar{w} = \frac{1}{n}\sum_{i=1}^{n} w_i, \quad n = n_1 + n_2, \tag{8.4}$$

and $100(1-2\alpha)$ percent confidence limits on μ are $\bar{w} \pm A$ (or negligibly narrower by virtue of the inequality in (8.3)).

The second sample size n_2 may turn out to be impractically large, especially if n_1 is chosen too small. Seelbinder discusses the choice of n_1 so as to minimize n on the average, but such choice depends on guessing what the true standard deviation σ is and is not discussed further in this report.

If σ is known, then no preliminary sample is needed. From a sample of size n, a $100(1-2\alpha)$ percent confidence interval for W is given by

$$\bar{w} \pm z_\alpha \sigma/\sqrt{n} \tag{8.5}$$

where z_α is the upper 100α percentage point of the normal distribution (1.96 for 95% confidence level). Hence we can achieve accuracy A by taking

$$n = z_\alpha^2 \sigma^2 / A^2. \tag{8.6}$$

If the observations are not independent, it is necessary to know or estimate from a preliminary sample the degree of dependence in order to determine the sample size needed to attain a prescribed accuracy. Since only an approximate solution can be hoped for in practice, only a simple first order approximation is proposed. We assume that the process can be modeled, within the extent of our likely knowledge, as a stationary first order Markov process. Then the autocorrelation function ρ_t between access time measurements made at a time separation t is equal to $\rho_1^{\,t}$, where t=1 is a convenient unit time separation, large relative to the largest access time. If a sample of size n is taken at unit time separations, then its mean $\bar{x}$ has variance

$$\sigma_{\bar{x}}^2 \cong \frac{\sigma^2}{n} \cdot \frac{1+\rho_1}{1-\rho_1}\ . \tag{8.7}$$

If ρ_1 as well as σ is known, then a $100(1-2\alpha)$ confidence interval for μ is given approximately by

$$\bar{x} \pm z_\alpha \frac{\sigma}{\sqrt{n}} \left(\frac{1+\rho_1}{1-\rho_1}\right)^{1/2} . \tag{8.8}$$

Hence, we can achieve accuracy A by taking sample size

$$n = \frac{z_\alpha^2 \sigma^2}{A^2} \cdot \frac{1+\rho_1}{1-\rho_1} . \tag{8.9}$$

It is unlikely that ρ_1 and σ are known, but they can be estimated from a preliminary sample at the same unit intervals. We estimate σ from (8.1) and ρ_1 by

$$r_1 = \frac{\text{cov}}{s' s''} \tag{8.10}$$

where

$$\text{cov} = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1 - 1} (x_i - \bar{x}')(x_{i+1} - \bar{x}''),$$

$$\bar{x}' = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1 - 1} x_i, \quad s' = \left[\frac{1}{n_1 - 1} \sum_{i=1}^{n_1 - 1} (x_i - \bar{x}')^2 \right]^{1/2}, \tag{8.11}$$

$$\bar{x}'' = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1 - 1} x_{i+1}, \quad s'' = \left[\frac{1}{n_1 - 1} \sum_{i=1}^{n_1 - 1} (x_{i+1} - \bar{x}'')^2 \right]^{1/2} .$$

Substitutions of s_1 and r_1 for σ and ρ_1 in (8.9) results in further approximation, part of which might be avoided by substituting instead in Stein's two-stage formula (6.3) (adjoining the factor $(1+r_1)/(1-r_1)$ to s_1^2), but to no great point since several other approximations are already incorporated.

Still another model seems possible for the measurement of access times up to a prescribed accuracy for their mean value. There may be a common component of variation for all measurements taken on the same day, as well as a component of variation that varies randomly from one measurement to another on the same day. Thus a biased picture would be obtained by taking measurements on only one day. The problem of economically taking n measurements on each of m days is solved in Appendix A. It is necessary to have prior knowledge of the cost of adding another day of measurement relative to another measurement on each day and of the relative sizes of the two component standard deviations. Charts are given for reading off m and n for a wide range of values of these inputs, and examples of their use are given. Since the required inputs are probably not known and cannot be estimated from preliminary samples accurately enough unless those samples are quite large, Appendix A may have more an instructional than an operational value. Nevertheless, the possibility of differences between days seems substantial and should not be neglected. An example of the use of the components-of-variance model is given in Appendix A itself.

Example 1. Standard deviation of access times is known. We consider the example in Proposed Federal Standard 1033, Vol. II, Table 1.1, of a nonswitched private line. The (mean) access time is given as 5.2 seconds. We assume that the standard deviation σ of individual access times is 1.0 second, that they are independent, that we don't know the mean, and that we want to sample enough times to estimate it within ± 0.2 second with 95% confidence. We apply (8.6) to obtain

$$n = (1.96)^2(1)^2/(0.2)^2 = 96.$$

Example 2. Same as Example 1 but we don't know σ. We take a preliminary sample of 10 observations (using a table of normally distributed random numbers) and calculate $\bar{x}_1=5.413$ seconds, $s_1=1.195$ seconds. From a table of Student t percentage points we find $t_{.025}=2.262$ for 9 d.f. Substituting in the left-hand side of (8.2) gives

$$2.262(1.195)/\sqrt{10} = 0.8546 \text{ second.}$$

Hence we must make n_2 further observations, where, by (8.3),

$$n_2 \geq (2.262)^2(1.195)^2/(0.2)^2 - 10 = 172.6.$$

Thus in these examples at least, ignorance of the standard deviation leads to needing almost twice as many observations.

Example 3. Same as Example 1 except that we assume there is autocorrelation between equally spaced observations obeying the Markov model with $\rho_1=0.5$. Then by (8.9) the required number of observations is three times what it was in Example 1, that is, 288.

9. SUMMARY OF PROCEDURE

It has been the purpose of this report to provide methods of designing tests or experiments on a digital communication system and of analyzing the resulting data so that the time and rate parameters are estimated with prescribed accuracy. An outline of procedure is given here, with necessary references to parts of the report.

1. If nothing is known about the various time parameters and measurements on them, it is necessary to carry out a preliminary experiment to determine the classes of source-destination user pairs and the hour or hours of the day to be considered. There may be only one class of user pairs, but it should be reasonably homogeneous. The preliminary experiment may be taken as 10 or more measurements at equal intervals within each hour of the day. On the basis of the hourly sample means $\bar{w}_1$ it should be decided which hour or hours is of interest. See the first part of Section 8.
2. It must also be decided whether there are significant differences between time measurements from day to day. Hence the preliminary experiment needs to be expanded to include several days. The decision is made on the basis of an "F test" as stated briefly after equation (A.9) in Appendix A. If there is a significant between-day component of variance, then Figures 1-3 may be used to design an experiment to estimate the mean time parameter with prescribed accuracy, as illustrated in the Appendix A example; the design requires inputs that may be estimated from the preliminary experiment, but it must be realized that these estimates may have poor accuracy.
3. In the main body of the report the between-day variance is ignored. It is assumed either that the standard deviation σ of the time measurements is known beforehand or that it is estimated as s_1 from (8.1) and a preliminary experiment. If the measurements are uncorrelated and σ is known, the sample size necessary to achieve accuracy A is determined from equation (8.6), where z_α is given after (6.4); after the data are obtained, the mean time parameter W is estimated by $\bar{w}$ as in (6.1), and confidence limits are calculated by (6.4) or (8.5).
4. If the measurements are uncorrelated and σ is estimated from s_1, then a further sample of measurements may be needed in accordance with (8.2) and (8.3). The mean time parameter W is then estimated by the <u>combined</u> sample mean $\bar{w}$ as in (8.4), and confidence limits are $\bar{w} \pm A$ where A is the prescribed accuracy A.

5. If an experiment has not been designed to achieve prescribed accuracy, and one simply has a sample of data to analyze, then confidence limits for the mean time parameter W can be calculated from (6.4) if the true population σ is known, or, in the more likely situation that σ is estimated by the sample standard deviation s as in (6.2), from (6.5).

6. Whether the time measurements are uncorrelated within an hour (or longer period of the day) can be tested by taking the measurements at equal time separations and calculating the autocorrelation function as in (8.10). Even if it is not zero, it may not be significantly different from zero. This can be tested approximately by using a table of percentage points of the ordinary correlation coefficient to be found in many statistics books. The correlation problem can be avoided by taking essentially independent observations only, i.e., by taking them sufficiently separated. If it is decided to use observations with correlation included, then the sample size for a prescribed accuracy must be larger in accordance with (8.9), and the confidence limits on W are calculated by an approximation as in (8.8) or as described at the end of Section 6.

7. Confidence limits for transfer rates and for block rate efficiency are derived from those for transfer times as stated in equations (7.3) and (7.5).

10. ACKNOWLEDGMENTS

The work has been conducted as part of the data communications project of the National Telecommunications and Information Administration under the supervision of Neal B. Seitz, project leader, whose problem statement, encouragement, and discussions are gratefully acknowledged. The report was typed by Patricia Sanchez.

11. REFERENCES

Bondesson, L. (1977), When has the sample mean as an estimator of a location or scale parameter high efficiency? Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete, Vol. 38, pp. 127-146.

Bury, K.V. (1975), Statistical Methods in Applied Science (John Wiley and Sons, New York, NY).

Cox, D.R. (1958), Some problems connected with statistical inference, Annals of Mathematical Statistics, Vol. 29, pp. 357-372.

Crow, E.L., and M.J. Miles (1977), Confidence limits for digital error rates from dependent transmissions, OT Report 77-118, National Telecommunications and Information Administration, (Available thru NTIS, 5285 Port Royal Road, Springfield, VA 22161, Accession No. PB267777).

Duffy, F.P., and R.A. Mercer (1978), A study of network performance and customer behavior during direct-distance-dialing call attempts in the U.S.A., Bell System Technical Journal, Vol. 57, pp. 1-33.

Guenther, W.C. (1964), Analysis of Variance, Prentice-Hall, Inc., Englewood Cliffs, N.J.

Kimmett, F.G., and N.B. Seitz (1978), Digital communication performance parameters for proposed Federal Standard 1033, Volume II: Application examples, NTIA Report 78-4, National Telecommunications and Information Administration (Available from NTIS, 5285 Port Royal Road, Springfield, VA 22161, Accession No. PB 284235/AS).

Kleinrock, L. (1976), Queueing Systems, Vol. II: Computer Applications (John Wiley and Sons, New York, NY).

Payne, J.A. (1978), ARPANET host to host access and disengagement measurements, NTIA Report 783, National Telecommunications and Information Administration (Available from NTIS, 5285 Port Royal Road, Springfield, VA 22161, Accession No. PB 283554/AS).

Pearson, K., ed. (1965), Tables of the Incomplete Γ-Function (Cambridge University Press, New York, NY).

Proposed Federal Standard 1033 (1978), Telecommunications: Digital communication performance parameters, Version 10, September 15.

Seitz, N.B., and P.M. McManamon (1978), Digital communication performance parameters for proposed Federal Standard 1033, Volume I: Standard parameters, NTIA Report 78-4, National Telecommunications and Information Administration (Available from NTIS, 5285 Port Royal Road, Springfield, VA 22161, Accession No. PB 283580/AS).

Seelbinder, B.M. (1953), On Stein's two-stage sampling scheme, Annals of Mathematical Statistics, Vol. 24, pp. 640-649.

Stein, C. (1945), A two-sample test for a linear hypothesis whose power is independent of the variance, Annals of Mathematical Statistics, Vol. 16, pp. 243-258.

APPENDIX A. DESIGN OF EXPERIMENTS WHEN THERE ARE WITHIN-DAY AND BETWEEN-DAY COMPONENTS OF VARIATION

The design of a test or experiment is at least as important as its analysis. Indeed, the complete design will include specifications of the method of analysis, so that in theory the actual analysis will involve only routine substitution in formulas or running of a computer program. In practice, unforeseen difficulties often arise that require some adjustment. The difficulties can be classified as various nonconformances with the assumed mathematical statistical model, and of course the data should be used to test the model as well as to make the minimal calculations needed if the system meets the primary specification.

The design of the experiment will follow from the precise statement of its purpose. The purpose will be taken as the following:

> To determine with minimum cost the mean delay (or other) time for the busiest hour and busiest pair of terminals of a specified data communications system within specified accuracy with 95% confidence.

It may not be known what the busiest hour and the busiest pair of terminals are, and other hours and pairs may be of interest. However, it is assumed that the time period of interest is known and that time measurements of the type of interest can be made within this period. It is assumed in this Appendix that the measurements are mutually independent within each day but may contain a daily component of error. The modifications necessitated by autocorrelations between measurements on the same day are discussed briefly in Secs. 6 and 8. In practice, the measurements should be taken at equal time intervals each day such that the interval is not a period or quasi-period of the system. Subsets of the state-time space of the system other than days, busiest hours, and busiest terminal pairs may be of interest and may be substituted for those terms, especially days, in the sequel.

Since it is recommended that the true mean delay time be estimated by the sample mean, the accuracy will be specified in terms of its standard deviation $\sigma_{\overline{X}}$. (It is assumed that the measurements are unbiased, so that we can speak of accuracy and precision interchangeably.) Thus more specifically the experiment should be designed so as:

> to determine with minimum cost the number of measurements per day, n, and the number of days of measurements, m, to make $2\sigma_{\overline{X}} < B$, where B is specified.

It is assumed that there are random fluctuations within each day, to be denoted by e_{ij}, which may include measurement error as well as system variations, and that there may be additional random variations from day to day, to be denoted by a_i. Thus the measurements x_{ij} are assumed to be of the form

$$x_{ij} = \mu + a_i + e_{ij};\ i = 1,2,\ldots,m; \quad j = 1,2,\ldots,n; \tag{A.1}$$

where the a_i and e_{ij} are mutually independent (uncorrelated is sufficient for most purposes) with standard deviations σ_a and σ_e.

Before discussing the determination of m and n it is desirable to present the method of analysis to be used after the data have been obtained. Ideally the same number of measurements should be made each day, but inevitably some variations will occur, so we let j vary from 1 to n_i. We estimate the true mean μ by

$$\overline{x} = \frac{1}{N} \sum_{i=1}^{m} \sum_{j=i}^{n_i} x_{ij} \tag{A.2}$$

where

$$N = \sum_{i=1}^{m} n_i . \tag{A.3}$$

Approximate 95% confidence limits for μ are given by

$$\overline{x} \pm 2s_{\overline{x}} \tag{A.4}$$

where

$$s_{\overline{x}}^2 = s_a^2 \frac{\sum n_i^2}{N^2} + \frac{s_e^2}{N}, \tag{A.5}$$

$$s_e^2 = \frac{1}{N-m} \sum_{i=1}^{m} \sum_{j=1}^{n_i} (x_{ij} - x_{i.})^2, \tag{A.6}$$

$$S_a^2 = \frac{m-1}{N-\sum n_i^2/N} (s_1^2 - s_e^2), \tag{A.7}$$

$$s_1^2 = \frac{1}{m-1} \sum_{i=1}^{m} n_i \, (x_{i\cdot} - \bar{x})^2, \tag{A.8}$$

and

$$x_{i\cdot} = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{ij}. \tag{A.9}$$

(See, for example, Guenther, 1964, pp. 59-60.) Whether the true between-day variance σ_a^2 is in fact positive or zero is tested by calculating the ratio $F = s_1^2/s_e^2$ and comparing it with the 5 percent (say) point of the F distribution with m-1 and N-m degrees of freedom, tables of which are available in most statistics books.

With all $n_i = n$ the variance of $\bar{x}$ is

$$\sigma_{\bar{x}}^2 = \frac{1}{m}(\sigma_a^2 + \frac{\sigma_e^2}{n}). \tag{A.10}$$

It is helpful to have the bound B specified as a fraction b of the true mean μ, so that $B = b\mu$. Also σ_a and σ_e are represented as fractions of μ:

$$\sigma_a = k_a\mu, \; \sigma_e = k_e\mu. \tag{A.11}$$

Then the accuracy requirement becomes

$$\frac{2}{m} (k_a^2 + \frac{k_e^2}{n})^{1/2} \leq b. \tag{A.12}$$

Then m could be specified as

$$m = \frac{4}{b^2} (k_a^2 + \frac{k_e^2}{n}) \tag{A.13}$$

if k_a and k_e are known and n is given. In fact k_a and k_e have to be known or guessed at prior to testing if m and n are to be determined.

Even with knowledge of k_a and k_e the one equation (A.13) is not sufficient to determine the two unknowns m and n. The further condition needed will result from minimizing the total cost, but first let us simply minimize the total number of observations N = mn. From (A.13)

$$N = \frac{4}{b^2}(k_a^2 n + k_e^2). \tag{A.14}$$

Since N increases linearly with n, N is minimized by taking n=1, i.e., only one measurement per day. This results because increasing m decreases both within-day and between-day components of variance while increasing n decreases only the within-day component. If there is no between-day component, $k_a=0$, and it does not matter whether the measurements are made on N days or all on one day. However, the possible existence of a between-day component should not be ignored.

The above optimal solution is not attractive in practice because it surely is less costly to make N observations by taking many each day. If all were taken on one day, the between-day component of variance could not be estimated. Likewise, if n=1, the within-day component could not be estimated. Hence $m \geq 2$ and $n \geq 2$ in practice. It will be assumed that the total cost depends linearly on the number of days and also on the total number of observations:

$$\text{Total Cost} = C_o + C_1 m + C_2 mn. \tag{A.15}$$

However, this is minimized if we minimize the simpler expression

$$C = m + cmn. \tag{A.16}$$

Substituting for m from (A.13), differentiating with respect to n, setting the derivative equal to zero, and solving for n gives the solution

$$n_{min} = c^{-1/2} k_e / k_a. \tag{A.17}$$

Substitution in (A.13) yields

$$m_{min} = \frac{4k_a^2}{b^2}\left(1 + c^{1/2}\frac{k_e}{k_a}\right). \tag{A.18}$$

Also

$$C_{min} = m_{min}\left(1 + c^{1/2} k_e / k_a\right). \tag{A.19}$$

Since m and n should be integers, we must modify the above continuous solution. We calculate C for the two integers nearest n_{min} and let

$$n'_{min} = [n_{min}] \text{ or } n'_{min} = [n_{min}+1], \tag{A.20}$$

whichever yields the smaller C, using [y] to denote the largest integer less than or equal to y. Then

$$m'_{min} = [\frac{4}{b^2}(k_a^2+\frac{k_e^2}{n'_{min}})+1]. \tag{A.21}$$

These will surely achieve the accuracy requirement (A.12) but still will not necessarily be the integers that minimize the cost; if m'_{min} is small (roughly smaller than n'_{min}), it may be possible to decrease C by allowing a slight increase in m and obtaining a substantial decrease in n. However, this gain is illusory in practice because the balance shifts rapidly with change in k_e and k_a, and the integer pair (m,n) absolutely minimizing C with the assumed k_e and k_a will not achieve the accuracy bound (A.12) for some true k_e and k_a quite nearby. Hence (A.20) and (A.21) yield the practical solution in integers and form the theoretical solution also unless m'_{min} is small.

Since n_{min} depends on k_e and k_a only through their ratio, it is convenient to use k_e and k_a/k_e as parameters in m'_{min} rather than k_e and k_a. We also neglect taking the next larger integer than m_{min} in calculations. Hence we let

$$m''_{min} = \frac{4k_e^2}{b^2}\left[\left(\frac{k_a}{k_e}\right)^2 + \frac{1}{n'_{min}}\right]. \tag{A.22}$$

<u>Example</u>. $b=c=0.1$, $k_a/k_e=1$. Then $n_{min}=\sqrt{10}=3.162$, $m_{min}=526.491\ k_e^2$, $C_{min}=692.982\ k_e^2$. But the integral solution for n is 3 or 4. From (A.13) and (A.16)

$$m_3 = 533.333\ k_e^2,\ C_3 = 693.333\ k_e^2,$$

$$m_4 = 500.000\ k_e^2,\ C_4 = 700.000\ k_e^2.$$

Hence $n'_{min}=3$, $m''_{min}=533.333\ k_e^2$. If $k_e=0.1$, then the most economical design is to take 3 measurements each day for 6 days. If $k_e=1$, then $m''_{min}=533$ and the design is much less practical. Thus a criterion no doubt entering into the design beyond accuracy and cost is the maximum number of days permissible for the experiment.

Charts of the values of m and n that provide the most economical experiments for determining mean delay time with specified accuracy are given in Figures 1, 2, and 3. The results depend on the confidence level (already limited to 95%), the relative accuracy desired (b), the relative cost (c), and the relative standard deviations k_a and k_e. Since n'_{min} (called n on the charts for simplicity) depends on k_a and k_e only through their ratio, it is convenient to use k_a/k_e as a parameter on curves (which are straight lines) giving m''_{min} (called m_{min} on the charts) as a function of k_e. The only relative accuracy value used is b=0.1, i.e., 10% of the true mean. Three different relative costs of each measurement in a day, relative to the incremental cost of another day are considered, c=0.02, 0.1, 1. Values of k_a/k_e from 1/128 to 4 by factors of 2 are offered. In a few cases these correspond to n=1, just one measurement per day, an impractical solution, and in any case at least two measurements should be taken per day in order to estimate σ_e by (A.6).

The ranges of values of m and n on the charts are enormous. It is emphasized that in order to use the charts one must accept the relative accuracy b=0.1 as adequate, must decide which, if any, relative cost c fits his situation, and must know or guess the relative standard deviations k_a and k_e. If Figures 1-3 do not cover the situation at hand, the values of $n=n'_{min}$ and $m_{min}=m''_{min}$ can easily be calculated from (A.17), (A.20), (A.16), and (A.22).

As an example of the use of the charts, we consider the above example. We enter Figure 2 and find the line labeled $k_a/k_e=1$. It is also labeled n=3, so 3 measurements are to be made each day irrespective of k_e. With k_e given as 0.1, we find on the line the solution $m_{min}=5.3$, so the experiment should continue for 6 days to achieve the desired accuracy (on the safe side).

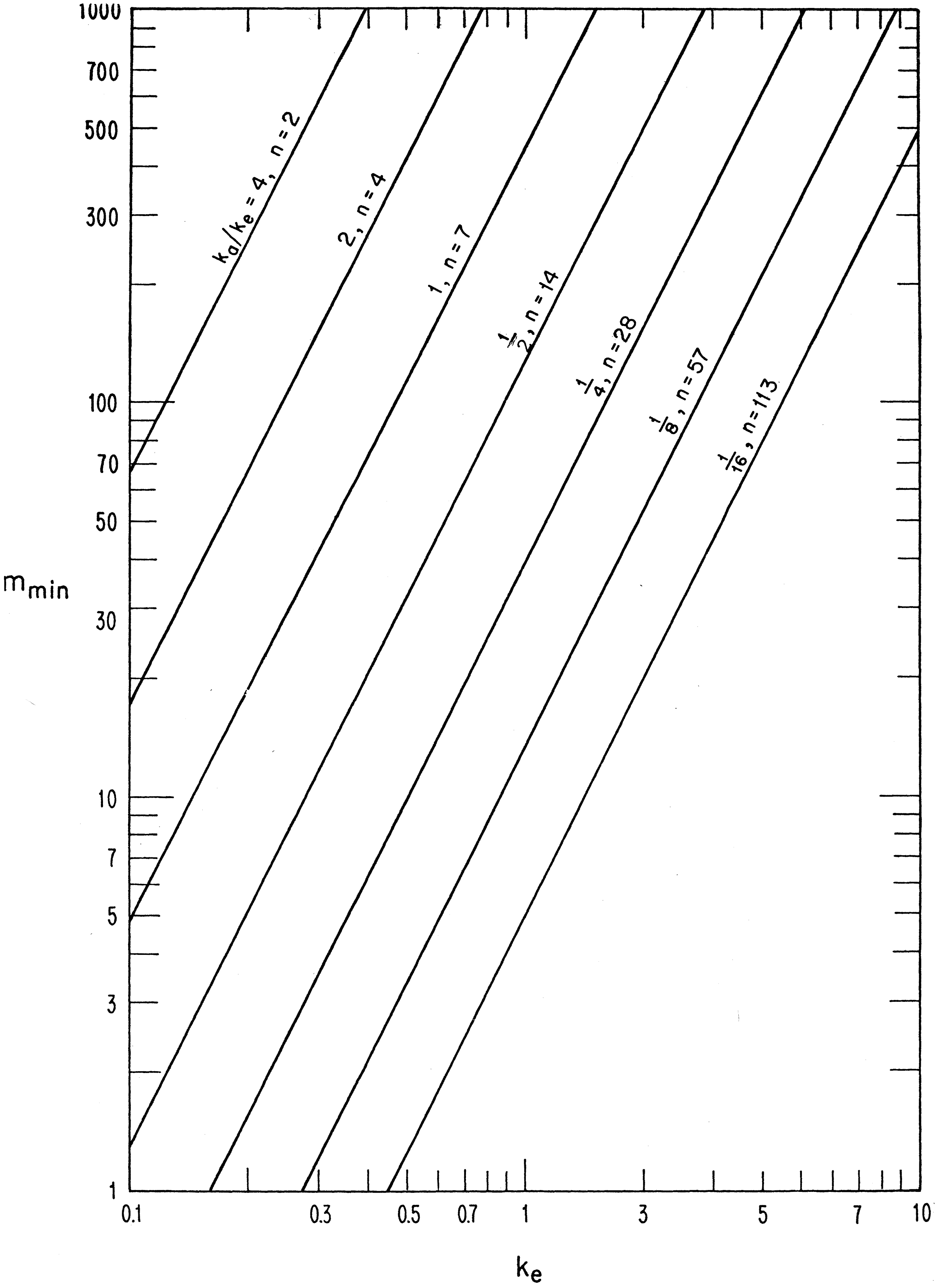


Figure 1. Minimum Cost Sample Sizes for b=0.1, c=0.02

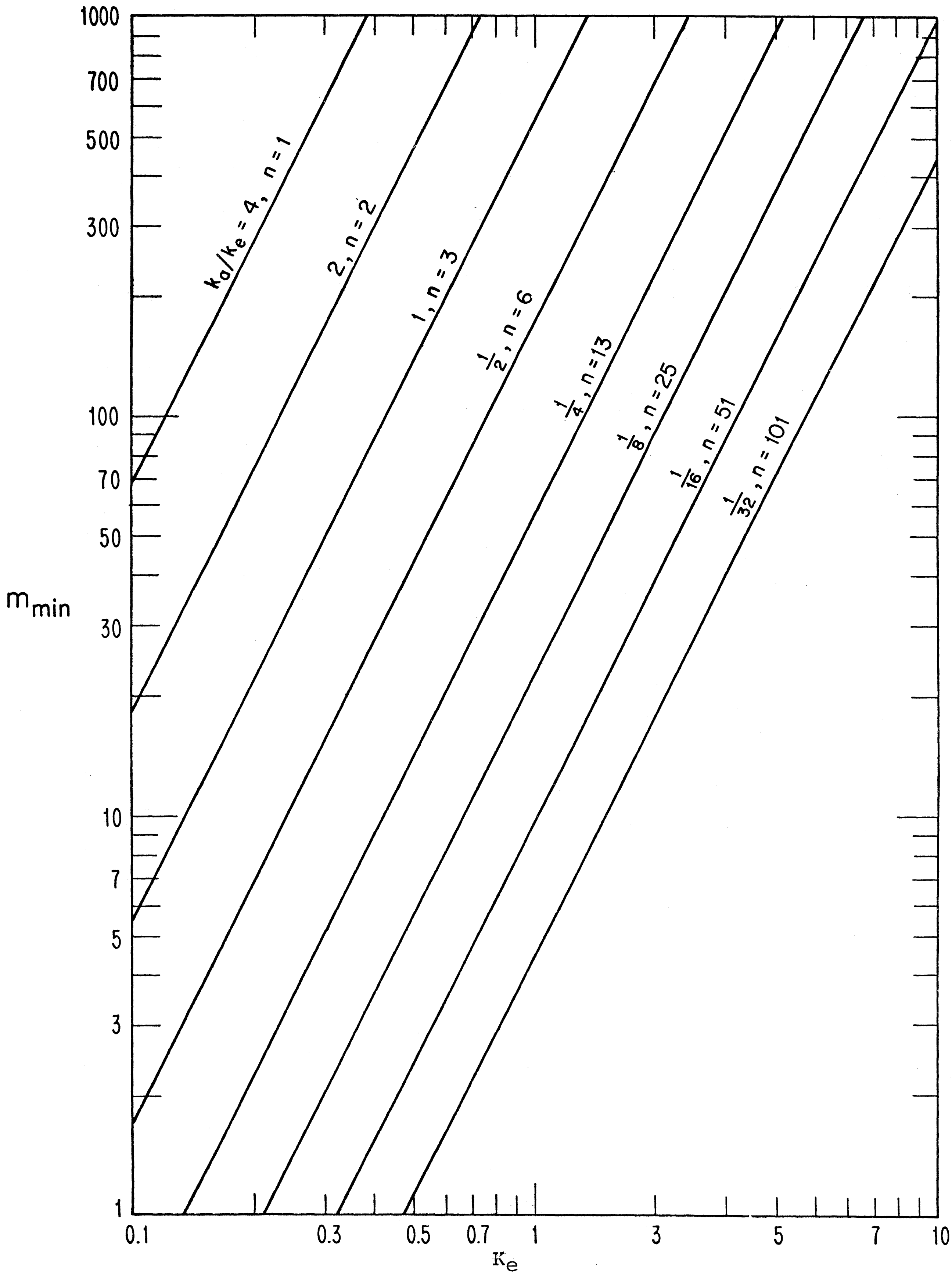


Figure 2. Minimum Cost Sample Sizes for b=0.1, c=0.1

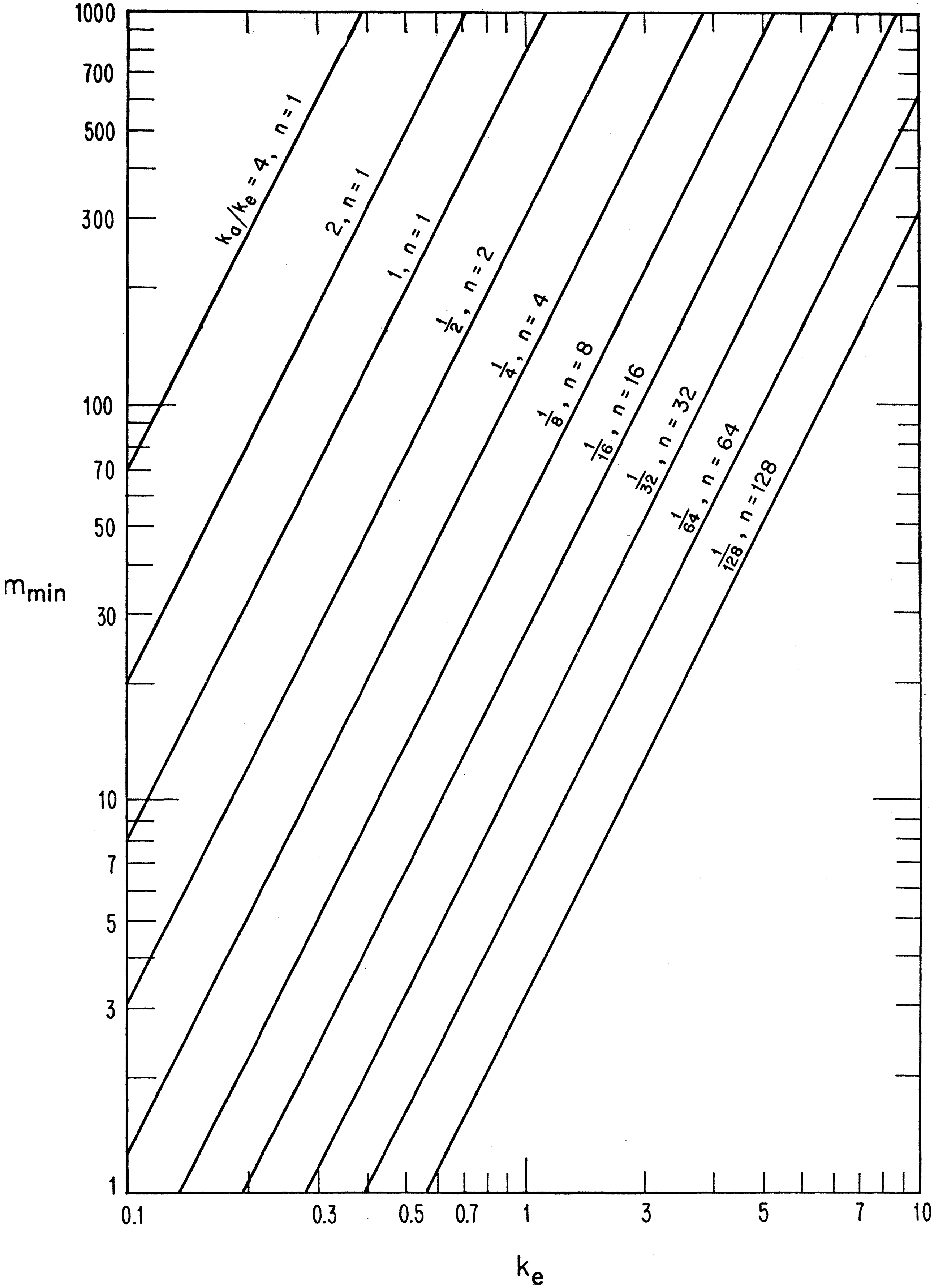


Figure 3. Minimum Cost Sample Sizes for b=0.1, c=1.

APPENDIX B. THE EFFECT OF TRUNCATING A DISTRIBUTION AT THREE TIMES ITS MEAN

Since the proposed Federal Standard 1033 considers any trial that results in a delay, disengagement, or block or message transfer time more than three times the "nominal" time as a failure, it seems of interest to consider the effect on an asymmetrical type of distribution of truncating it at three times its mean. That is, the nominal time is assumed to be the mean prior to truncation. A convenient type of asymmetrical distribution to use is the gamma family of distributions, all members of which extend from 0 to ∞ with continuous probability density functions (pdf's) having a single maximum and varying asymmetry. The particular effects of truncation considered are: (1) the probability that a time that follows a gamma distribution will be discarded by the truncation and (2) the change in mean due to the truncation.

A well-known subfamily of the gamma distributions is the set of chi-squared (χ^2) distributions, which have the single parameter called the "degrees of freedom," a positive integer, denoted here by ν. It is used to specify the particular pdf and thus its asymmetry, which ranges from very large for $\nu=1$ to 0 in the limit for $\nu\to\infty$. The mean of a chi-squared distribution is ν. In order to maintain the mean value constant, at unity, we consider the variable

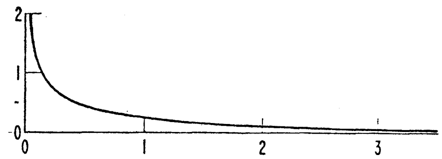
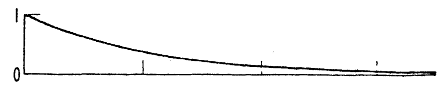
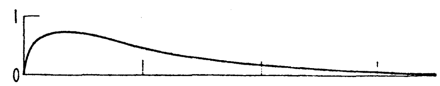
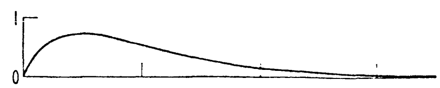
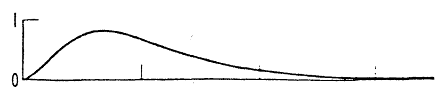
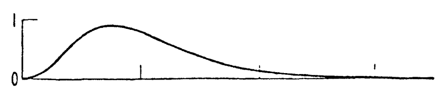
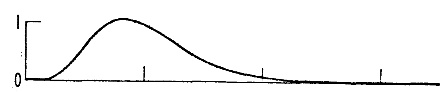
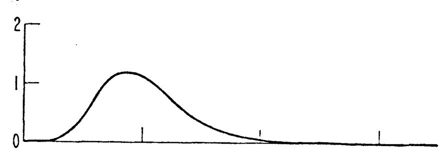
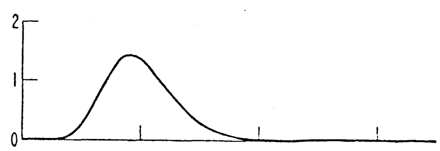
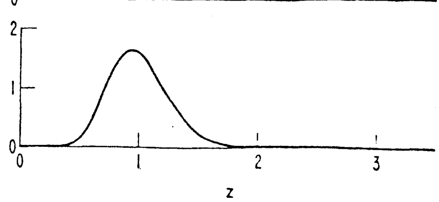
$$Z = \chi^2/\nu.$$

It is easy to calculate the probability that a gamma variable is three times its mean value, that is, that $Z>3$, from Karl Pearson's Tables of the Incomplete Γ-Function (1965). After a change of variable we find (details being omitted) that

$$P(Z>3) = 1-I(3(\nu/2)^{1/2}, \nu/2 - 1)$$

where $I(u,p)$ is the incomplete Γ-function tabulated by Pearson. Table 1 gives this probability for ν ranging from 1 to 32. The pdf's of the distributions are also plotted in Table 1. The probability of considering a trial a failure is appreciable for $\nu=1$, about 8%, and $\nu=2$ (the exponential distribution), about 5%, but decreases towards 0 as ν increases. Two further remarks are in order: (1) More extreme cases than $\nu=1$ are quite possible, from fractional values of ν, yielding much larger values of

Table 1. Probability of a gamma variable exeeding three times its mean and the mean of a gamma variable truncated at three times its mean

ν (degrees of freedom)	P(Z>3)	μ_t/μ	Probability density function
1	.0833	.663	
2	.0498	.843	
3	.0293	.918	
4	.1074	.9545	
6	.0062	.9849	
8	.0023	.9947	
12	.00033	.99928	
16	.000048	.999897	
24	.0000011	.9999977	
32	.0000000	.9999999	

P(Z>3); the table has a limited range. (2) Although the mean value of Z remains constant, its standard deviation steadily decreases as ν increases, as suggested by the pdf plots.

It is also easy to calculate the mean of the truncated distribution from Pearson's Tables. If μ denotes the mean of the gamma distribution and μ_t the mean of the same distribution after truncation at 3μ, then

$$\frac{\mu_t}{\mu} = \frac{I(3\nu/(2\nu+4)^{1/2}, \nu/2)}{I(3(\nu/2)^{1/2}, \nu/2 - 1)} .$$

This truncated mean is also tabulated in Table 1. It is about 34% less than the original mean for $\nu=1$ and 16% less for $\nu=2$, but rapidly approaches the original mean as ν increases. (The last digit in the ratio of means, as well as in the probability, may be in error by about one unit because of linear interpolation in Pearson's Tables.)

It may be useful to note that the mode (abscissa of the maximum of the pdf) of the gamma distribution with mean unity is $1-2/\nu$ for $\nu \geq 2$ and that the median is very close to $1-2/(3\nu)$ for $\nu \geq 2$. Thus, for $\nu=3$ the mode is 1/3 and the median is 0.789. For $\nu=24$ the mode is 11/12 and the median is 0.9725.

These illustrations using the gamma distribution do not imply any firm conclusions for actual time distributions in digital communication systems because the gamma distribution may not be the precise form and the degrees of freedom parameter no doubt varies with the system and the type of time interval. It is believed that the gamma distribution is a plausible model, but the second uncertainty is a considerable one, as indicated by Table 1. It is hoped that some intuitive feeling may be gained by associating the probabilities and truncated means with the distribution shapes.

The examination of Bell System and ARPANET time interval distributions in Section 4 revealed shapes varying over roughly the entire range shown in Table 1. Despite this wide range of shapes and the corresponding probabilities and means shown in the table, there is no suggestion that truncation should occur at any other point than at three times the mean.

FORM OT-29
(3-73)

U.S. DEPARTMENT OF COMMERCE
OFFICE OF TELECOMMUNICATIONS

BIBLIOGRAPHIC DATA SHEET

	1. PUBLICATION OR REPORT NO.	2. Gov't Accession No.	3. Recipient's Accession No.
	NTIA Report 79-21		

4. TITLE AND SUBTITLE	5. Publication Date
STATISTICAL METHODS FOR ESTIMATING TIME AND RATE PARAMETERS OF DIGITAL COMMUNICATION SYSTEMS	June, 1979
	6. Performing Organization Code: NTIA/ITS
7. AUTHOR(S): Edwin L. Crow	9. Project/Task/Work Unit No.
8. PERFORMING ORGANIZATION NAME AND ADDRESS: U.S. Dept. of Commerce, National Telecommunications and Information Administration, Institute for Telecommunication Sciences, 325 Broadway, Boulder, CO 80303	10. Contract/Grant No.
11. Sponsoring Organization Name and Address: U.S. Dept. of Commerce, National Telecommunications and Information Administration, 1325 G Street, NW, Washington, DC 20005	12. Type of Report and Period Covered
	13.

14. SUPPLEMENTARY NOTES

15. ABSTRACT *(A 200-word or less factual summary of most significant information. If document includes a significant bibliography or literature survey, mention it here.)*

Statistical methods are provided for estimating time and rate parameters of digital communication systems according to the specifications of the Proposed Federal Standard 1033. The methods may be applied to delay, disengagement, transfer and service times, and to transfer rates. Some ARPANET and Bell System data are examined for the form of distribution of times. The properties of various types of estimators are discussed, and the sample mean is recommended. Approximate confidence limits are given for time and rate parameters. A step-by-step procedure is provided for designing an experiment to estimate time and rate parameters with prescribed accuracy.

16. Key Words *(Alphabetical order, separated by semicolons)*

ARPANET; autocorrelation; confidence limits; design of experiments; distribution-free estimation; robust estimation

17. AVAILABILITY STATEMENT	18. Security Class *(This report)*	20. Number of pages
[X] UNLIMITED.	Unclassified	41
[ ] FOR OFFICIAL DISTRIBUTION.	19. Security Class *(This page)*: Unclassified	21. Price:

☆ U.S. GOVERNMENT PRINTING OFFICE: 1979-0-680-338/498 USCOMM-DC 29716-P73