Probability and Statistics Notes

Note 11

Some Practical Aspects of Sampling in a Finite Population

by

John W. Williams

April 1982

Mission Research Corporation

Abstract

This paper explores some of the practical implications of the formalism developed for calculations of confidence and reliability in a finite population (Ref. 1 and 2). Because evaluations of the operational capabilities of military systems are often expensive, it is usually crucial to achieve a given level of confidence with a minimum number of tests. In this note, we first discuss the sensitivity of the number of required tests to the number of failures observed during the test program. It is assumed throughout the analysis that the attribute of success or failure for each item is unknown prior to testing. Comparisons of test requirements are provided for partitioned and nonpartitioned populations of the same cardinality. Expressions for the maximum attainable confidence are derived for cases in which the testing of items from one partition is highly restricted. Results are illustrated with examples.

Introduction

Often the evaluation of military systems is subject to budgetary or time constraints which prohibit tests of every item in the population of interest. This problem can be especially acute in the estimation of operational capabilities for fleets of military systems when equipment must be removed from operational status in order to accommodate the test requirements. In addition, tests which are necessary to establish operational capabilities may be too expensive or time consuming to allow an exhaustive evaluation of all critical components for even a single system. For example, an examination of every crucial electrical circuit in a complex aircraft or missile might be impractical due to the large number of such circuits in each system. Hence, most test programs must confront the problem of quantifying the confidence in reliability based upon limited sampling of the total population.

Test Requirements for a Nonpartitioned Population

Mathematical techniques have been developed by Ashley (Ref. 1 and 2) for the treatment of confidence in a finite population. Following the notation in these references, we assume that a finite population of N items can be dichotomized through testing with respect to some attribute. One might designate items which are found to possess the attribute as failures or defectives, while items lacking the attribute are called successes or non-defective. We assume that the experimentalist has no knowledge of the success or failure of items prior to testing (Ref. 3). The number of items selected at random from the population for testing is denoted by L, while M represents the number of observed defectives. Allowing K to symbolize the unknown number of defective items in the population, the confidence that no more than K_0 items are defective is given by (Ref. 1):

$$C(K \leq K_o) = \frac{\sum_{I=M}^{K_o} \binom{I}{M}\binom{N-I}{L-M}}{\sum_{J=M}^{N-(L-M)} \binom{J}{M}\binom{N-J}{L-M}} \tag{1}$$

The denominator in equation (1) is the total number of outcomes consistent with the test data, including those outcomes in which the number of defectives may be greater than K_o. Contributions to the numerator include only the outcomes which are consistent with both the test data and the assumption that no more than K_o items are defective.

A calculation of confidence $C(K \leq 20)$ is shown in Figure 1 for N = 100 and N = 1000. Notice that while curves representing constant values of M are fairly insensitive to the cardinality of the populations, the minimum reliability

$$R = 1 - K/N \tag{2}$$

varies noticeably for the two populations. For example, under the assumption that $K \leq 20$ and N = 100, the minimum reliability is 0.80, while the minimum reliability is 0.98 for N = 1000. The number of items which must be tested in order to establish a fixed level of confidence depends markedly upon the number of observed defectives. If few failures are found among the items randomly selected for tests, high confidence can be achieved from an examination of a relatively small fraction of the population. On the other hand, as the number of observed failures approaches the value of K_o, nearly all of the items in the population must be tested to attain high confidence that K does not exceed K_o.

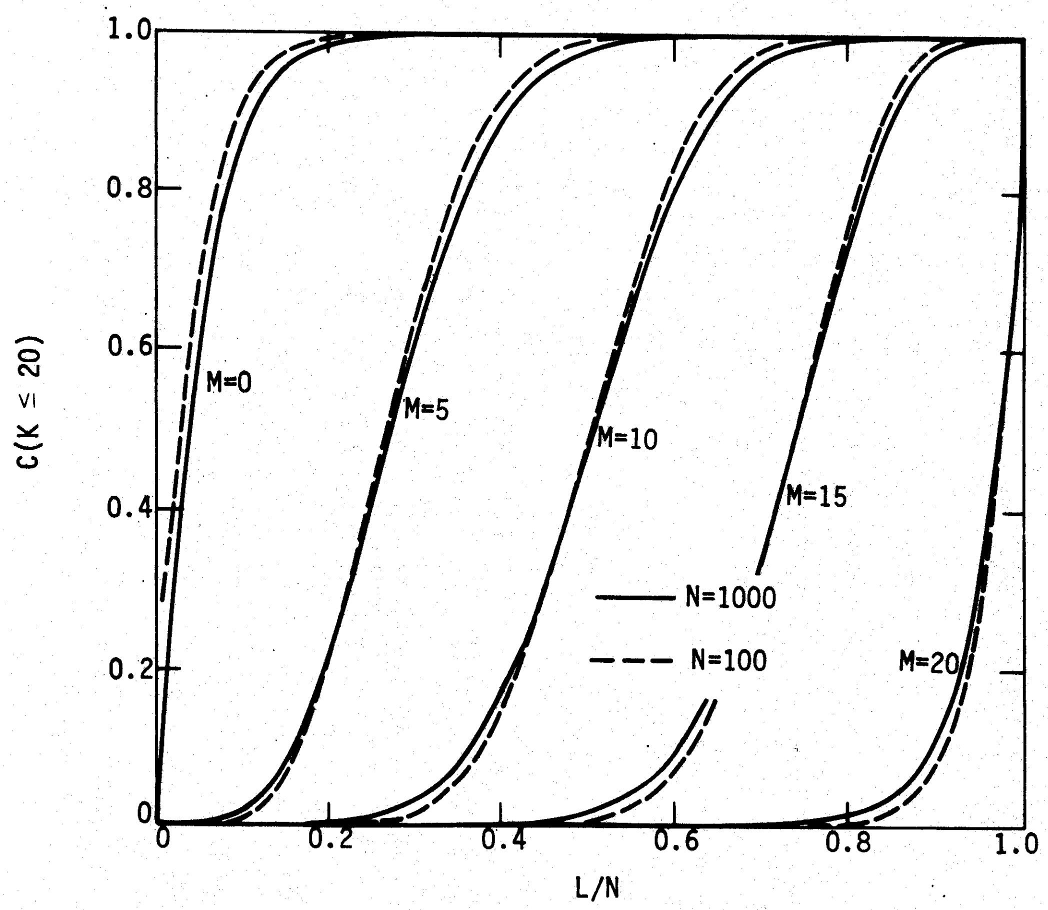

Figure 1. Confidence that no more than 20 items are defective for populations of cardinality N = 100 (R ≥ 80%) and N = 1000 (R ≥ 98%). The number of items randomly selected and tested is denoted by L, and M represents the number of failures.

Test Requirements for a Partitioned Population

Now consider a finite population which has been divided into two distinguishable partitions such that data from one partition may not give information about the reliability of another partition (Ref. 2). Such a division is usually considered when limited test resources force the experimentalist to exclude some items as candidates for testing or when operational demands remove items from the testable portion of the population. For example, some of the aircraft in a fleet might be assigned to flight duties which preclude their participation in the test program. Since only part of the entire fleet would be available for testing, the total population comprised of all aircraft in the fleet could be divided into two partitions with testing in one of the partitions prohibited or highly restricted. In this case the sample may not be random from the population as a whole. Specifically, if the sample is taken only from aircraft which are not assigned to essential flight duties ("shelf or replacement items"), the sample may not accurately reflect changes in reliability consequent upon the constant use and wear expected for aircraft in the unsampled "flight duty" partition.

The mathematical formalism appropriate for calculations of confidence for a finite population with random sampling from distinguishable partitions has been developed in reference 2. In the case of two partitions:

$$C(K \leq K_o) = \sum_{I_1=M_1}^{K_o} \left[\binom{I_1}{M_1} \binom{N_1-I_1}{L_1-M_1} \sum_{I_2=M_2}^{K_o-I_1} \binom{I_2}{M_2} \binom{N_2-I_2}{L_2-M_2} \right] \Bigg/ D, \qquad (3)$$

where

$$D = \left[\sum_{J_1=M_1}^{N_1-(L_1-M_1)} \binom{J_1}{M_1} \binom{N_1-J_1}{L_1-M_1} \right] \left[\sum_{J_2=M_2}^{N_2-(L_2-M_2)} \binom{J_2}{M_2} \binom{N_2-J_2}{L_2-M_2} \right] . \qquad (4)$$

Subscripts denote the partition, e.g. N_1 is the number of items in the first partition, K_1 is the unknown number of defective items in the first partition, etc. Although the total population has been divided into two distinct parts, our objective is to evaluate confidence in the reliability of the total population.

Consider first the case in which items from the second partition are excluded from testing (L_2=0). In this situation equations (3) and (4) yield:

$$C(K \leq K_o) = \sum_{I_1=M_1}^{K_o} \left(\frac{K_o - I_1 + 1}{N_2 + 1} \right) \overline{C}_{N_1}(I_1) \qquad (5)$$

where $\overline{C}_{N_1}(I_1)$ is the confidence that exactly I_1 items in a population of N_1 are defective (based upon tests of L_1 items selected randomly from the N_1). Notice that if the number of observed defectives in the first partition, M_1, is close to K_o, contributions of $\overline{C}_{N_1}(I_1)$ to the summation are limited to a few terms weighted by a factor proportional to $1/(N_2 + 1)$, so that confidence tends to increase relatively slowly with the number of items tested.

From a practical viewpoint, it is interesting to examine the penalty in the required number of tests to achieve a given confidence when a population is partitioned and the items in one partition are excluded from testing. Figure 2 shows values of confidence computed for partitioned and nonpartitioned populations of the same cardinality. Solid lines display the confidence that no more than twenty items are defective in a nonpartitioned population of 200 items. Dashed lines in the figure show the corresponding values of confidence computed with equation (5) for a partitioned set with N_1 = 180, N_2 = 20, and L_2 = 0. Notice that if five or more failures are observed in the first partition, then confidence values of 80% or higher cannot be achieved for the total population (even with exhaustive testing of the first partition). Figures 3 and 4 display the results of similar calculations for partitioned populations with N_1 = 190, N_2 = 10 and N_1 = 160, N_2 = 40, respectively. For the partitioned case shown in Figure 3, the number of items in the second partition

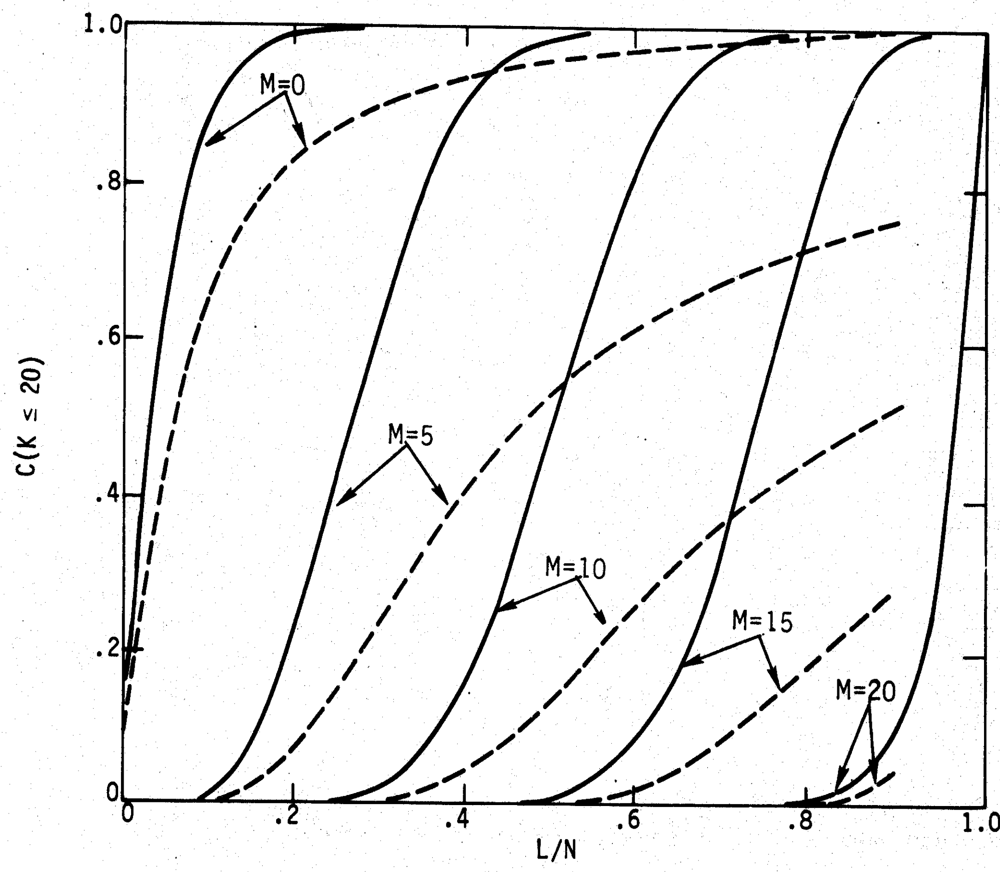

Figure 2. Confidence that no more than 20 items are defective in a partitioned (N_1 = 180, N_2 = 20) population and a nonpartitioned (N = 200) population. Items in the second partition are excluded from the random test selection (L_2 = 0). Dashed (solid) lines show confidence values for the partitioned (nonpartitioned) population.

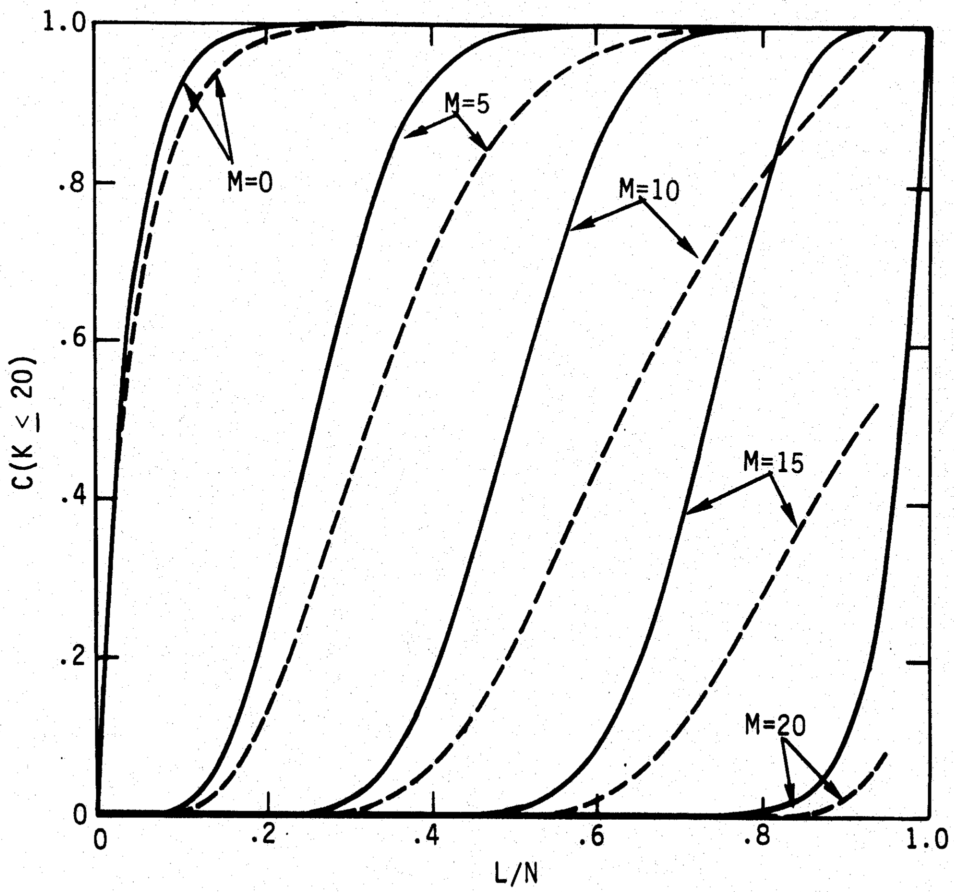

Figure 3. Confidence that no more than 20 items are defective in a partitioned (N_1 = 190, N_2 = 10) population and a nonpartitioned (N = 200) population. Items in the second partition are excluded from the random test selection (L_2 = 0). Dashed (solid) lines show confidence values for the partitioned (nonpartitioned) population.

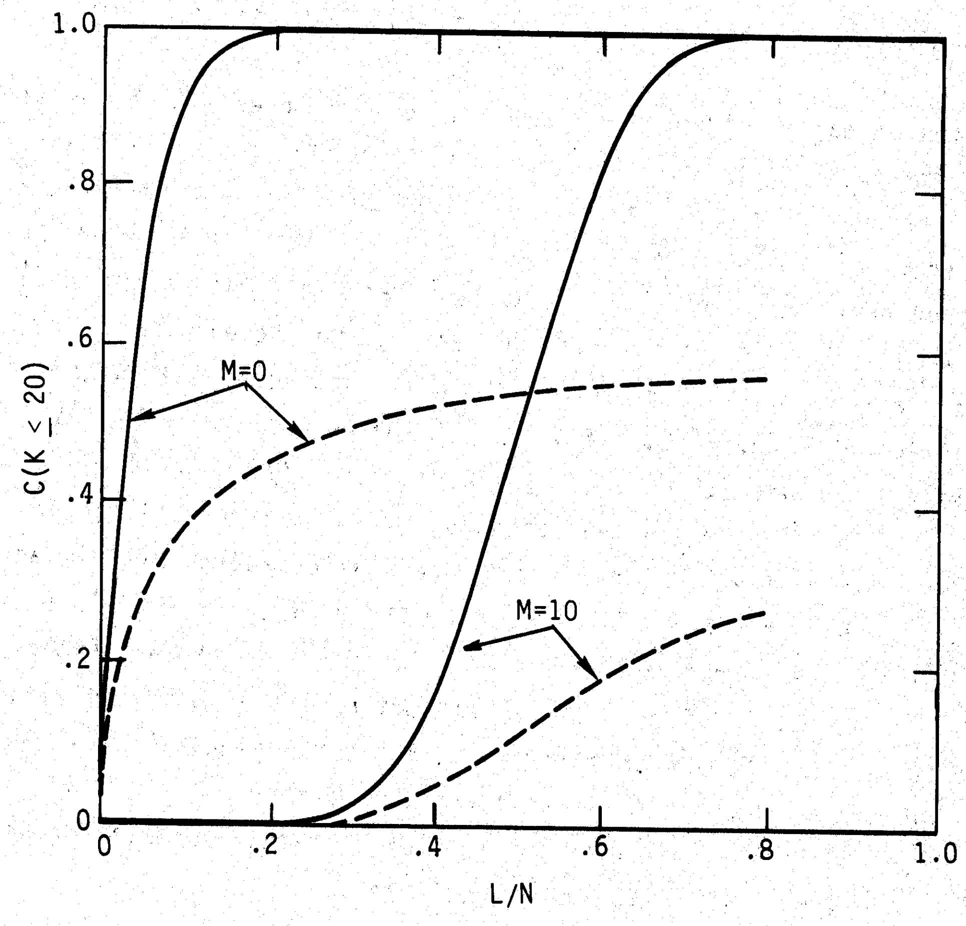

Figure 4. Confidence that no more than 20 items are defective in a partitioned (N_1 = 160, N_2 = 40) population and a nonpartitioned (N = 200) population. Items in the second partition are excluded from the random test selection. Dashed (solid) lines show confidence values for the partitioned (nonpartitioned) population.

is less than the maximum value of K, while Figure 4 illustrates confidence values for a case in which N_2 exceeds K.

The noticeable penalty in maximum attainable confidence which can be incurred by partitioning has a simple interpretation in terms of the assumption of maximum ignorance (Ref. 3). For example, let us reexamine the information displayed in Figure 2 for the partitioned population. Suppose that all N_1 = 180 items in the first partition are tested and M_1 = 5 items are observed to be defective. All of the N_2 = 20 items in the second partition were excluded as candidates for the test program. Recall we assumed that information concerning the success or failure of any item is available only from the experiment. This means that with regard to the fraction of items in the second partition which succeed or fail, all outcomes are presumed to be equally likely. If a fraction, 0/20, 1/20, 2/20,..., 15/20 of items in the second partition are defective, then the condition $K \leq 20$ is valid for the total population (since $M_1=K_1=5$ and $L_1=N_1$). On the other hand, if a fraction 16/20, 17/20,...20/20, of the items are failures, then the number of failures in the total population exceeds 20. Thus, 16 of the 21 possible outcomes are in accord with the condition $K \leq 20$. Notice that 16/21 = .762 is the peak value of confidence displayed in Figure 2 for the partitioned population with M_1 = 5. We shall now formalize expressions for the maximum attainable confidence when testing in one partition is highly restricted.

For a population divided into two partitions, the maximum confidence attainable from L_2 tests of items in the second partition is found from equations (3) and (4) with $L_1 = N_1$:

$$C(K \leq K_0 | L_1 = N_1) = \frac{\displaystyle\sum_{I_2=M_2}^{K_0-M_1} \binom{I_2}{M_2} \binom{N_2-I_2}{L_2-M_2}}{\displaystyle\sum_{J_2=M_2}^{N_2-(L_2-M_2)} \binom{J_2}{M_2} \binom{N_2-J_2}{L_2-M_2}} \qquad (6)$$

Comparing equations (1) and (6) we see that $C(K < K_o \mid N_1 = L_1)$ can be identified with the confidence that the reliability of items in the second partition is no less than

$$R_2 = 1 - (K_o - M_1)/N_2 \tag{7}$$

Reliability R_2 is the minimum reliability for items in the second partition which will assure an overall reliability of $R = 1 - K_o/(N_1 + N_2)$ for the entire population. For the special case with $L_2 = 0$, equation (6) reduces to

$$C(K \leq K_o \mid L_1 = N_1, L_2 = 0) = \text{Min}\left[\frac{K_o - M_1 + 1}{N_2 + 1}, 1\right] , \tag{8}$$

where the upper limit of unity is included in equation (8) for cases in which K_o-M_1 exceeds the number of items in the second partition. Equation (8) can be used to compute the maximum values of confidence displayed by dashed lines in Figures 2 through 4. For convenience in the discussion below, we use C_o to denote $C(K \leq K_o \mid L_1 = N_1, L_2 = 0)$, while C_1 will represent $C(K \leq K_o \mid L_1 = N_1, L_2 = 1)$.

If a single item from the second partition is selected at random, tested and found not to fail, then the maximum confidence found from equation (6) is given by

$$C_1(M_2 = 0) = \text{Min}\left[\left(\frac{K_o - M_1 + 1}{N_2 + 1}\right)\left(\frac{2N_2 - K_o + M_1}{N_2}\right), 1\right] \tag{9a}$$

$$= \text{Min}\left[C_o(1 + R_2), 1\right] , \tag{9b}$$

where the confidence appearing in brackets on the right side of the equation is computed for L_2=0, and R_2 is the reliability for the second partition defined in equation (7). On the other hand, if the single item randomly selected from the second partition is observed to fail, then

$$C_1(M_2 = 1) = \text{Min}\left[\left(\frac{K_o - M_1 + 1}{N_2 + 1}\right)\left(\frac{K_o - M_1}{N_2}\right),\ 1\right] \tag{10a}$$

$$= \text{Min}\left[\ C_o(1 - R_2),\ 1\ \right]. \tag{10b}$$

Figure 5 shows the maximum attainable confidence computed from equations (8) through (10) with $N_1 = 180$, $N_2 = 20$, and $K_o = 20$. If items in the second partition are exempted from testing, then the maximum confidence decreases linearly with the number of observed failures in the first partition. However, if a single item in the second partition is selected at random and found to fail, then the decrease in maximum confidence is quadratic in M_1. Finally, we note that if only one item from the second partition is selected at random and observed to pass the test, then the decrease in attainable confidence with increasing values of M_1 is appreciably less rapid than the corresponding case with $M_2 = 1$.

Conclusion

Calculations presented above demonstrate that the number of tests required to achieve a given confidence in the reliability is quite sensitive to the number of failures observed during the test program. If no defective items are found among the randomly selected sample, high confidence can be achieved from tests of a relatively small percentage of the total population. However, as the number of defective items observed in the test program approaches the maximum allowed number of failures consistent with the desired reliability, then practically the entire population must be tested to achieve high confidence.

Partitioning a population is usually considered when limited resources or other practical considerations restrict the test candidates for some

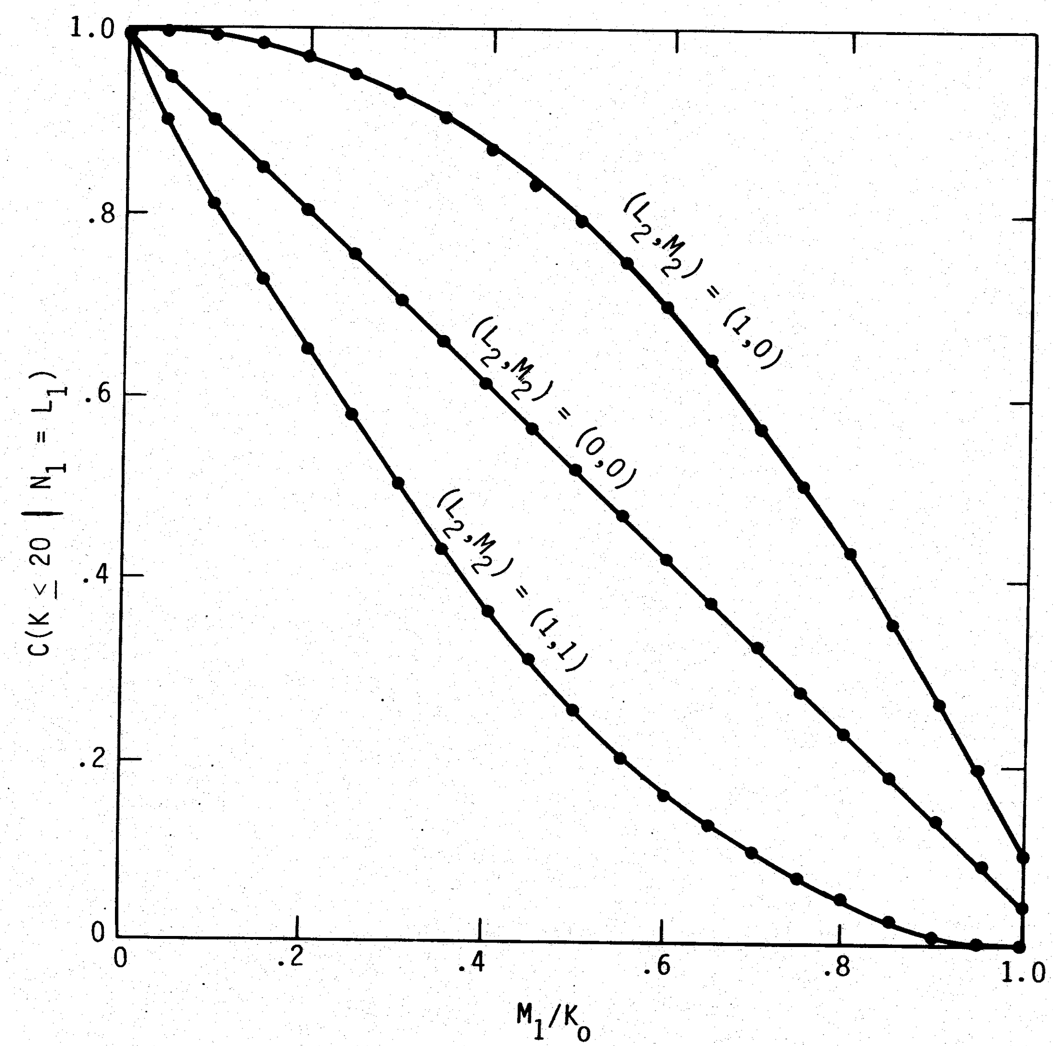

Figure 5. Maximum attainable confidence that no more than $K_o = 20$ items are defective in partitioned population with $N_1 = 180$ and $N_2 = 20$. All items in the first partition are tested ($N_1 = L_1$), and M_1 items in the first partition are observed to fail.

subset of the population. If the success or failure of items is unknown prior to testing, and if even a modest portion of the population is excluded from the random test selection process, then (1) the maximum achievable confidence can be markedly decreased, and (2) the number of tests required to achieve a given confidence can noticeably increase. Achievable confidence, like the number of required tests, is highly sensitive to the number of observed failures in a partitioned population.

References

1. Chris Ashley, "Confidence and Reliability in a Finite Population," Probability and Statistics Notes, Note 1, 18 February 1971.

2. Chris Ashley, "Confidence and Reliability in a Partitioned Finite Population," Probability and Statistics Notes, Note 5, 22 December 1972.

3. Chris Ashley, "Effects of Supplemental Information on Confidence and Reliability," Probability and Statistics Notes, Note 7, 22 September 1973.