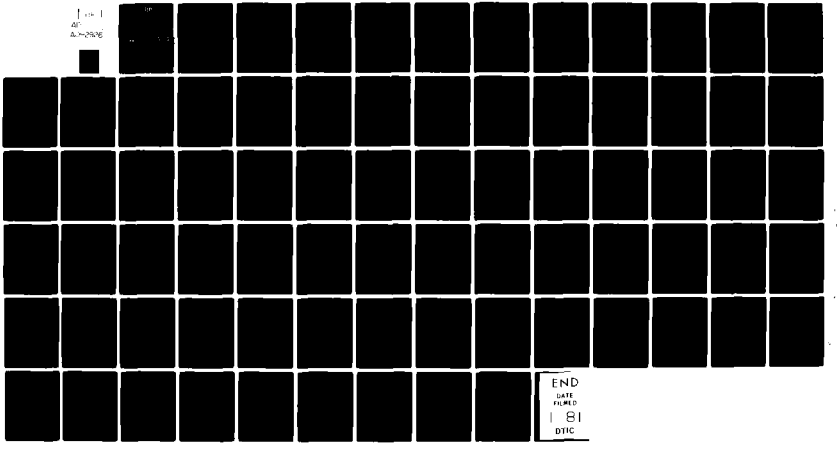
AD-A092 926 WASHINGTON UNIV SEATTLE F/G 13/12
APPLICATION OF THE BINOMIAL DISTRIBUTION TO FAILURE MODES FOR S--ETC(U)
NOV 80 A S HUND F29650-79-M-4047
UNCLASSIFIED AFWL-TR-80-90 NL



END
DATE FILMED
81
DTIC

AD-E200615

AFWL-TR-80-90

(2) LEVEL III

AFWL-TR-80-90

AD A092926

APPLICATION OF THE BINOMIAL DISTRIBUTION TO FAILURE MODES FOR SAFETY ANALYSIS

Amy Stocking Hund

University of Washington
Seattle, WA 98195

November 1980

Final Report

AIR FORCE WEAPONS LABORATORY
Air Force Systems Command
Kirtland Air Force Base, NM 87117

AFWL-TR-80-90

This final report was prepared by the University of Washington, College of Engineering, Seattle, Washington, under Contract F29650-79-M-4047, Job Order ILIR 7909 with the Air Force Weapons Laboratory, Kirtland Air Force Base, New Mexico. Major Fredrick W. Alles, Jr. (NTSS) was the Laboratory Project Officer-in-Charge.

When US Government drawings, specifications, or other data are used for any purpose other than a definitely related Government procurement operation, the Government thereby incurs no responsibility nor any obligation whatsoever, and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings,specifications, or other data, is not to be regarded by implication or otherwise, as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

This report has been authored by an employee of the United States Government. Accordingly, the United States Government retains a nonexclusive, royalty-free license to publish or reproduce the material contained herein, or allow others to do so, for the United States Government purposes.

This report has been reviewed by the Public Affairs Office and is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public, including foreign nations.

This technical report has been reviewed and is approved for publication.

Fredrick W. Alles Jr.

FREDRICK W. ALLES, JR.
Major, USAF
Project Officer

FOR THE DIRECTOR

William Chynoweth

WILLIAM CHYNOWETH
Lt Col, USAF
Chief, Surety Branch

HERBERT M. FERNANDEZ
Chief, Nuclear Systems Engineering Division

DO NOT RETURN THIS COPY. RETAIN OR DESTROY.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFWL-TR-80-90	2. GOVT ACCESSION NO. AD-A092 926	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) APPLICATION OF THE BINOMIAL DISTRIBUTION TO FAILURE MODES FOR SAFETY ANALYSIS		5. TYPE OF REPORT & PERIOD COVERED Final Report 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Amy Stocking Hund		8. CONTRACT OR GRANT NUMBER(s) F29650-79-M-4047
9. PERFORMING ORGANIZATION NAME AND ADDRESS University of Washington Seattle, Washington 98195		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61101F/ILIR 7909
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Weapons Laboratory (NTS) Kirtland Air Force Base, NM 87117		12. REPORT DATE November 1980 13. NUMBER OF PAGES 76
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Reliability Testing
Nuclear Safety Testing
Component Testing
Component Failure Modes
Safety Analysis
Binomial Distribution

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This study was performed to answer the question: Can tests to determine safety-related failure rates of typical nuclear weapon delivery system components be performed which will result in failure rate data, with acceptable confidence levels, at reasonable cost.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

CONTENTS

1.0 INTRODUCTION AND PROBLEM DEFINITION

The inherent safety of high technology functional systems is receiving increasing attention in an effort to maintain or improve system safety, while responding to pressure to streamline system designs. Cost is an ever-present constraint on the design and construction of safe and reliable hardware. In the case of aerospace applications, space and weight are additional high priority constraints contributing pressure to streamline designs. In the interest of optimizing these constraints in addition to reliability and safety requirements, a great deal of emphasis is placed on the collection and utilization of test data, particularly in the area of electrical/electronic reliability.

Although both safety and reliability deal with system or component failure, their relationship is often bittersweet. For many systems, safety improvements result in reduced reliability and vice versa. Optimizing both safety and reliability factors, therefore, becomes a complex problem. In very critical systems, such as nuclear weapons, no slack in these trade-off considerations can be tolerated. It must be possible to determine to a high degree of accuracy not only system reliability: the probability of failure; but also the safety of the system: the probability that a failure will occur which will result in an unsafe condition.

System components are capable of experiencing several types or modes of failures. Identifying the possible failure modes of electronic components, and their effects on system safety, is a complex issue. One reliability study, for example, identified eight distinct failure modes for relays: No operation, hang-up, failure to open circuit, excessive contact bounce, abnormal operate time, "chatter", abnormally high contact resistance, and contact short circuit. In a critical system some, but not all, failure modes of a component could result in a dangerous, perhaps even catastrophic situation, while other failure modes may simply render a component useless, but harmless. Commonly, if the failure of the component in question results in system unreliability, then the component will be tested simply for failure. No data would be accumulated on safety-related failure modes. The cost of accumulating failure mode data becomes a major factor in the assessment of system safety.

Strict hardware screening has been proposed and applied for determination of failure mode data. The result is limited, specific data, albeit critical, acquired at considerable cost. This type of data, however, is crucial to the design for safety in weapons systems.

Since it is not possible to eliminate the need for these data, or to influence significantly the cost of testing, an approach is proposed in this paper which will maximize the use of the data. By observing the behavior of the data as certain parameters are varied, it is possible, to establish test criteria which will minimize the cost of data collection. Better yet, by establishing and mapping the

approximate behavior of the data "a priori", it becomes possible not only to optimize test costs, but also to apply the data to the pre-established behavioral patterns, and extract accurate and valuable information about the safety and reliability of that component. The technique proposed here is applicable to all components and is based on probability distributions and confidence intervals. The derivation of the technique is discussed in the next section.

2.0 THE APPLICATION OF THE BINOMIAL DISTRIBUTION TO RELIABILITY TESTING

A binomial condition exists in reliability testing. That is, in testing for failure, all final outcomes for the components can be expressed by two states, "failed" or "not failed". It is also possible to establish a binomial test for component safety. Those failure modes which are inherently safe (eg: relay fail open) are identified. All other failure modes are classified "unsafe". (This principle is further discussed in section 3.0.) If all components are tested to failure, we now have two possible outcomes of the test - "failed-safe" and "failed unsafe". The behavior of the population of the tested components, can be approximated by the binomial distribution. The binomial distribution is characterized by:

1) An experiment of exactly "n" trials. In the application presented here, "n" will represent the number of components tested to failure.

2) Each trial results in one of two possible outcomes; in this case, "fail-safe" or "fail-unsafe".

3) The probability of the event of interest remains constant from trial to trial. The probability that any component fails "safe", given that it fails, is expressed in this application by "p". The proportion of tested components failing "unsafe" is represented by "k". "k" is used as the

estimate of component behavior for "a priori" analysis. For the purposes of this study, it is assumed that "k" remains constant throughout the test period. Assumptions are discussed in detail in Section 3.0.

4) All trials are mutually independent. This is a reasonable assumption for the testing of "n" identical, separate components.

The use of this distribution, particularly for reliability testing, is well documented. Even under conditions where all these assumptions are not met, the binomial distribution often provides an accurate approximation for component and system analysis. The distribution lends itself well, therefore, to this application and to the manipulation of the data which permits us to analyze the behavior of the subject populations parametrically.

The technique applied in this proposal is an adaptation of a well-known technique first presented by Clopper and Pearson in an article published in Biometrika in 1934.[1] This article presented a method of establishing an estimate of an unknown parameter of a

---

[1] E. S. Pearson, C. J. Clopper, "The Use of Confidence or Fudicial Limits Illustrated in the Case of the Binomial", Biometrika Vol. 26, 1934, pp 404-13.

population from a random sample of that population, and then establishing confidence intervals for that parameter, using the binomial distribution. "Confidence intervals" express the probability that the parameter of interest falls between some upper and lower limit, which form an interval. That is, limits p_1 and p_2 can be established between which we can state that $p_1 < p < p_2$, at some given confidence. The confidence that "p" lies within this interval depends upon the proportion of times that this prediction is correct in the long run of statistical experience. Figure 1 is reproduced from Pearson's article. The limits shown in Figure 1 are chosen so that the error is symmetrical above and below the limits.

We will now apply this technique to the failure modes of system components. By testing "n" components to failure, and by establishing the proportion which failed unsafe (k), we can draw some conclusions about the behavior of these components in general. In addition, we will be able to establish, through application of the binomial distribution, some level of confidence about the assumptions made concerning the population, and even the range of parameter values to which that confidence level applies. It will be shown also that, in general, the greater the number of components we test, the higher the confidence we can have in our parameter estimate. As we will see, application of this technique can provide us with a great deal of useful information about population parameters, including comparison data for values of n, k, and confidence level, C.

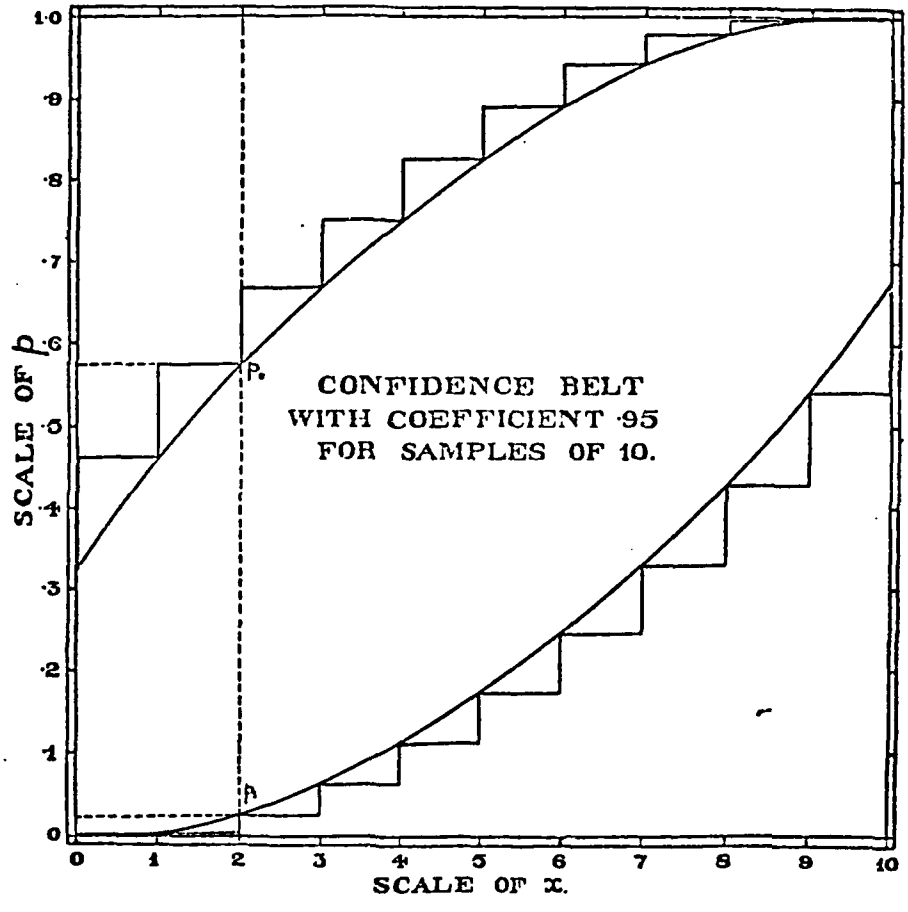

CONFIDENCE BELT
WITH COEFFICIENT ·95
FOR SAMPLES OF 10.
SCALE OF p.
SCALE OF x.
P.
A
0
1
2
3
4
5
6
7
8
9
10
·1
·2
·3
·4
·5
·6
·7
·8
·9
1·0

FIGURE 1

The confidence interval described by Pearson is a two-sided, symmetrical confidence interval. That is, if we have the 95 percent confidence interval $p_1 < p < p_2$, 2½ percent of actual values of p will lie above p_2 and 2½ percent will lie below p_1 (see Figure 1). When considering component reliability, the probability of reliability being above the upper confidence limit is of little or no interest. We are primarily interested in the probability that the reliability is less than some threshold level, which corresponds to the lower confidence limit. We can, therefore, remove the upper limit from the confidence interval, and express our example in terms of a 97½ percent, one-sided confidence interval, i.e., $p_1 < p$ with 97½ percent confidence.

2.1 USE OF POISSON DISTRIBUTION FOR SMALL VALUES OF K

In order to create a technique applicable to a broad range of components, parameters such as proportion of components failed unsafe (k) and sample size (n) will be varied over a reasonable range. Values of k to .001 will be examined in the discussion that follows. Testing for values of k this small (.001) will require large values of n. In order to expect the minimum number of failures to verify k=.001, a value for n of 1000 must be chosen. Direct calculation of binomial terms for large n becomes cumbersome. For large n, and small k, the binomial variable will be approximately normally distributed, and therefore the poisson distribution is used to approximate binomial probabilities. Values for k=.001 on Table 1, (discussed in the following section) are taken from the poisson distribution.

In order to provide useable values for threshold acceptance numbers, large values of n are graphed for k=.001. The graphs start at n=500 for reference and extend to n=5000.

2.2 BINOMIAL PARAMETRIC STUDY

We have now defined the concepts necessary to establish a simple and valuable technique for failure mode test data analysis. Using readily available binomial tables, graphs of parametric behavior have been prepared. Cost/confidence trade-offs, optimum sample size, threshold test numbers and attainable levels of confidence are easily read from these graphs, even before tests begin. After testing is completed, graphs are used to establish the confidence levels for the population parameter. Section 2.3 demonstrates the application of these graphs. This technique can be applied either to component test for failure to establish initially a component's failure mode characteristics, or it can be applied to acceptance testing of hardware to ensure a component lot meets the required safety criteria. The paragraphs that follow describe the derivation, application and interpretation of these charts. Parameter sensitivity, cost trade-offs and possibilities for expanding the information available from the charts will also be discussed.

In order to present a broad spectrum of applicability, ranges for each parameter were selected which seemed appropriate to the high technology, safety critical application, and to the components to which this technique will most likely be applied. The result is a set of graphs which display in a variety of ways information about test, component and population character. Of particular interest is the sensitivity information concerning sample size and confidence

level. The graphs were derived from the values shown in Table 1. Table 1 is a summary of the applicable portions of the cumulative binomial probability table (except, as previously noted, for k = .001). Its derivation is discussed in Pearson[2]. Each parameter and its range are discussed below.

The parameters compared in Figures 2 through 27 are k, p, n and C.

k The proportion of components in the sample which failed "unsafe". The value of k for a component or family of components may be determined either from extensive hardware tests to failure, or from historical or field component data. The values of k selected for this study are .001, .01, .1, .2, .3, .4, and .5

p The probability that a component from the subject population will fail "safe", given that it fails. This corresponds to 1-k in the sample. For this application, therefore, "p" is no longer the "probability of not failure", but now becomes the "probability of fail safe", and must, therefore, be used in conjunction with hardware failure rates for component safety analysis.

n The number of components tested to failure to determine "k". Values selected for n are 10, 25, 50, 100, 250 and 500. For k=.001, values of 1000,2500, and 5000 are also included.

---

[2] op. cit.

TABLE 1 CUMULATIVE PROBABILITY

			CONFIDENCE LEVELS			
k	n	r	85 %	90 %	95 %	99 %
.001	500	.5	.9983	.9980	.9969	.9953
	1000	1.0	.9984	.9981	.9975	.9964
	2500	2.5	.9985	.9983	.9980	.9974
	5000	5.0	.9986	.9985	.9983	.9979
.01	10	.1	.810	.781	.728	.617
	25	.25	.911	.897	.871	.814
	50	.5	.947	.940	.925	.893
	100	1.0	.966	.962	.953	.935
	250	2.5	.978	.976	.972	.964
	500	5.0	.983	.982	.979	.974
.1	10	1.0	.696	.663	.606	.496
	25	2.5	.795	.776	.744	.678
	50	5.0	.834	.822	.801	.758
	100	10.0	.858	.850	.836	.809
	250	25.0	.876	.871	.863	.847
	500	50.0	.884	.881	.875	.865
.2	10	2.0	.585	.550	.493	.388
	25	5.0	.681	.660	.625	.556
	50	10.0	.723	.709	.684	.637
	100	20.0	.749	.739	.723	.691
	250	50.0	.770	.763	.753	.735
	500	100.0	.779	.775	.768	.755

TABLE 1 (cont'd)

			CONFIDENCE LEVELS			
k	n	r*	85	90	95	99
.3	10	3.0	.482	.448	.393	.297
	25	7.5	.575	.553	.517	.449
	50	15.0	.617	.602	.576	.528
	100	30.0	.644	.634	.616	.582
	250	75.0	.666	.660	.649	.628
	500	150.0	.677	.672	.664	.659
.4	10	4.0	.387	.354	.304	.218
	25	10.0	.474	.452	.417	.352
	50	20.0	.515	.499	.474	.426
	100	40.0	.542	.531	.513	.479
	250	100.0	.565	.558	.546	.525
	500	200.0	.576	.571	.563	.547
.5	10	5.0	.297	.267	.222	.150
	25	12.5	.377	.356	.323	.265
	50	25.0	.416	.401	.376	.331
	100	50.0	.442	.427	.409	.377
	250	125.0	.465	.458	.446	.425
	500	250.0	.475	.470	.462	.447

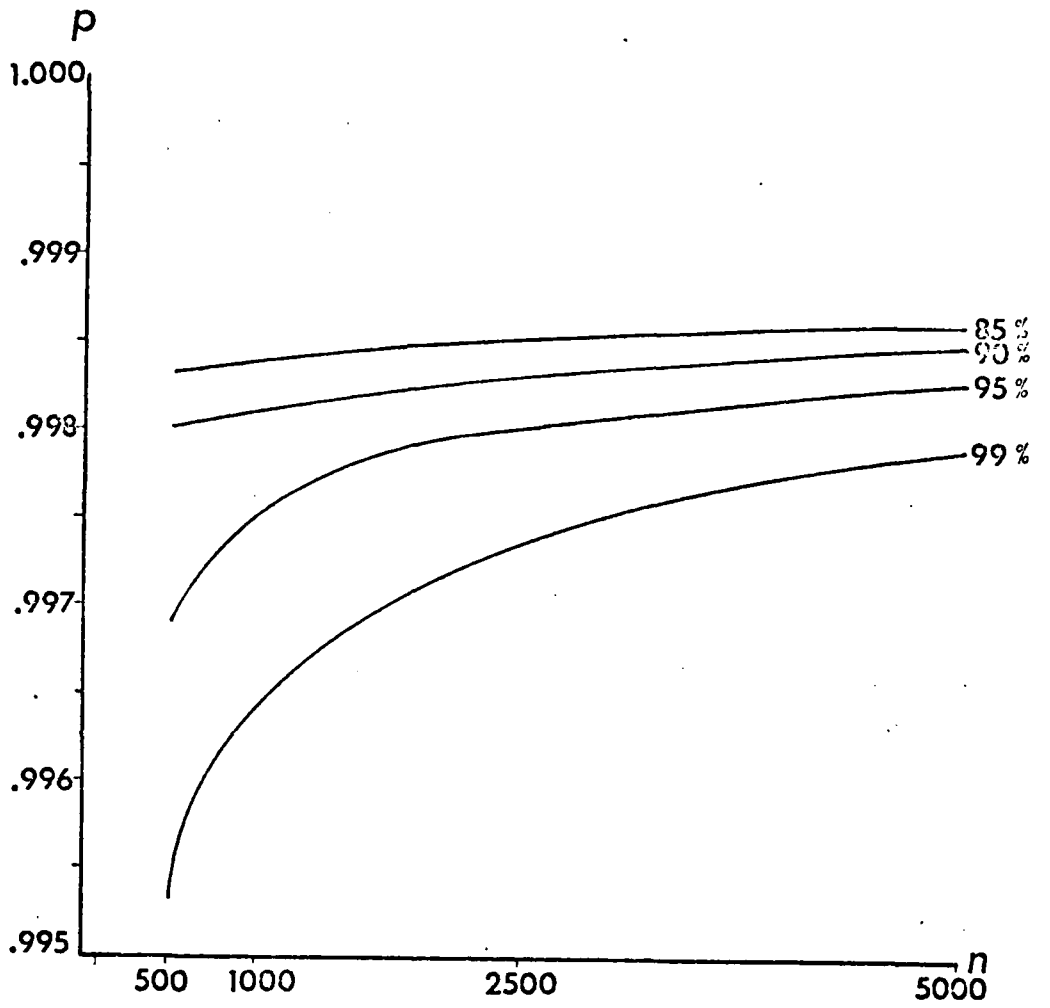
* r=(n)(k)

For sample size n and established k, table shows probability that r or fewer units fail "unsafe".

C The confidence limit. Values of 85, 90,95 abd 99 percent are shown on the graphs.

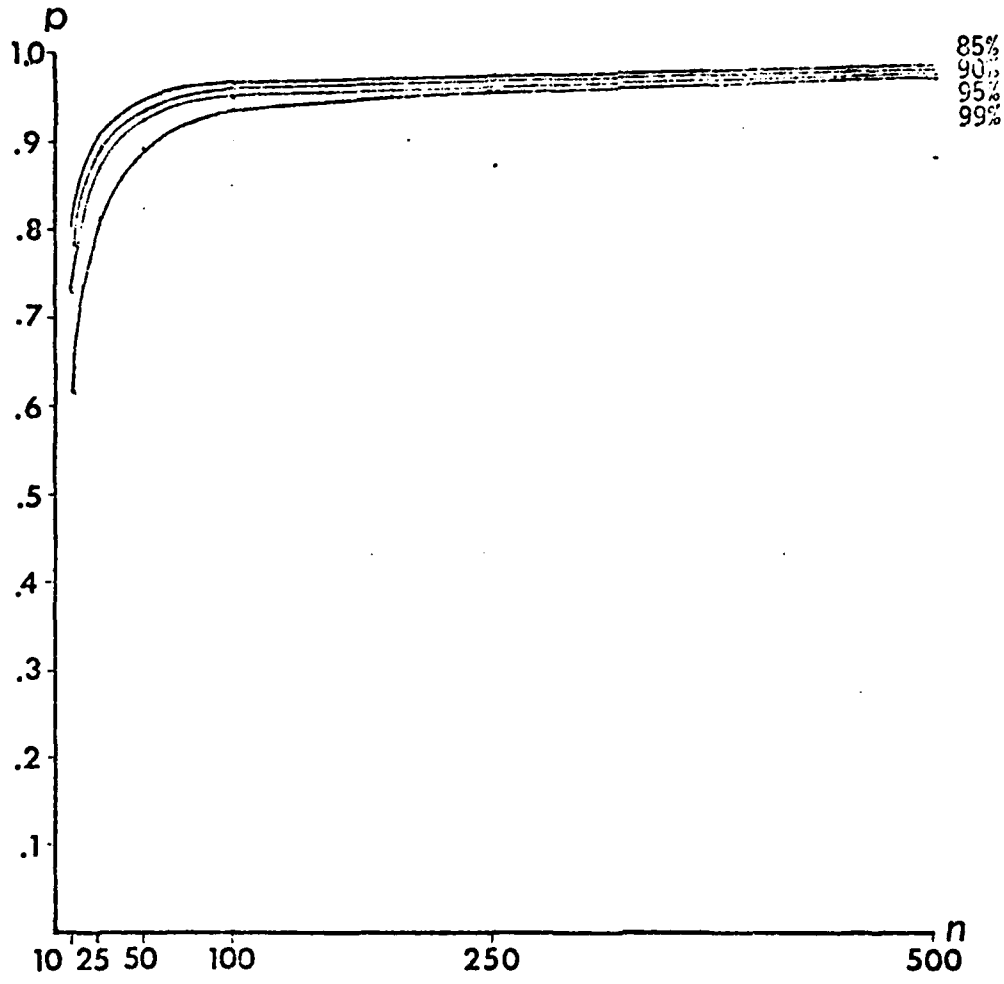
Figures 2 through 8 compare the confidence levels for p at different sample sizes for each value of k. These curves demonstrate a distinct trade-off for sample size (and, therefore, test cost) versus potential improvement in confidence in the population. From these graphs, optimum sample size for component test can be determined. For example, note that the higher the expected safety (fail-safe reliability), p, of the hardware (ie; the lower values of k), the smaller the optimum sample size (n). Note the change in scale for n in Figure 1 (k = .001).

A reasonable cost value for n can be established from the preceding figures. The next set of graphs, Figures 9 through 13, may also be used to minimize the cost of failure mode testing. These graphs may first be checked to establish compatibility of the minimum sample size which can be tested to verify p at the preferred confidence level, with the optimum value of n as previously determined. It is then possible to compare, by sample size, the required value of p and its corresponding value of k. For example, assume a requirement has been established, that a component must exhibit a fail-safe reliability of at least 80 percent at a 90 percent confidence level. From the preceeding graphs we determine that the optimum sample size (n) for this test is 100. From Figure 11, representing 90 percent confidence curves, we can determine that the threshold level for k to meet the prescribed safety requirement, is .14. That is, the maximum



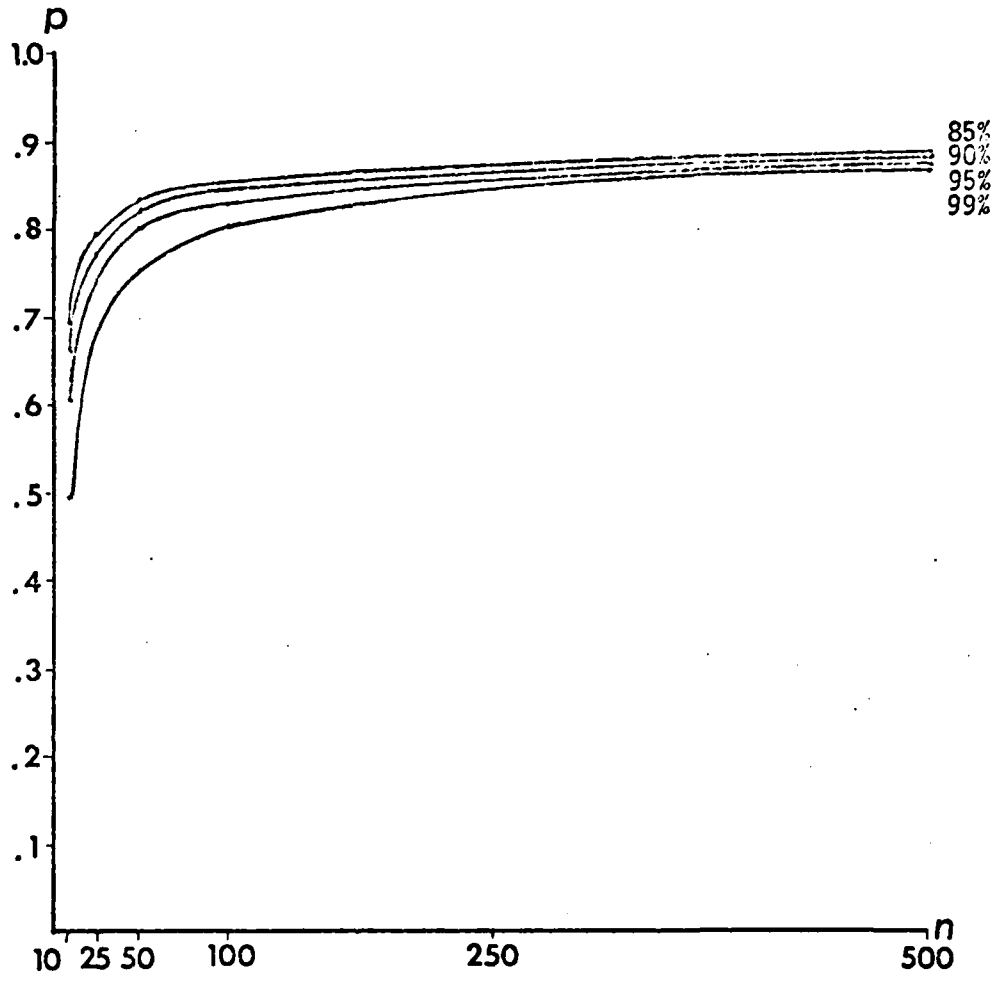
$k = .001$

FIGURE 2



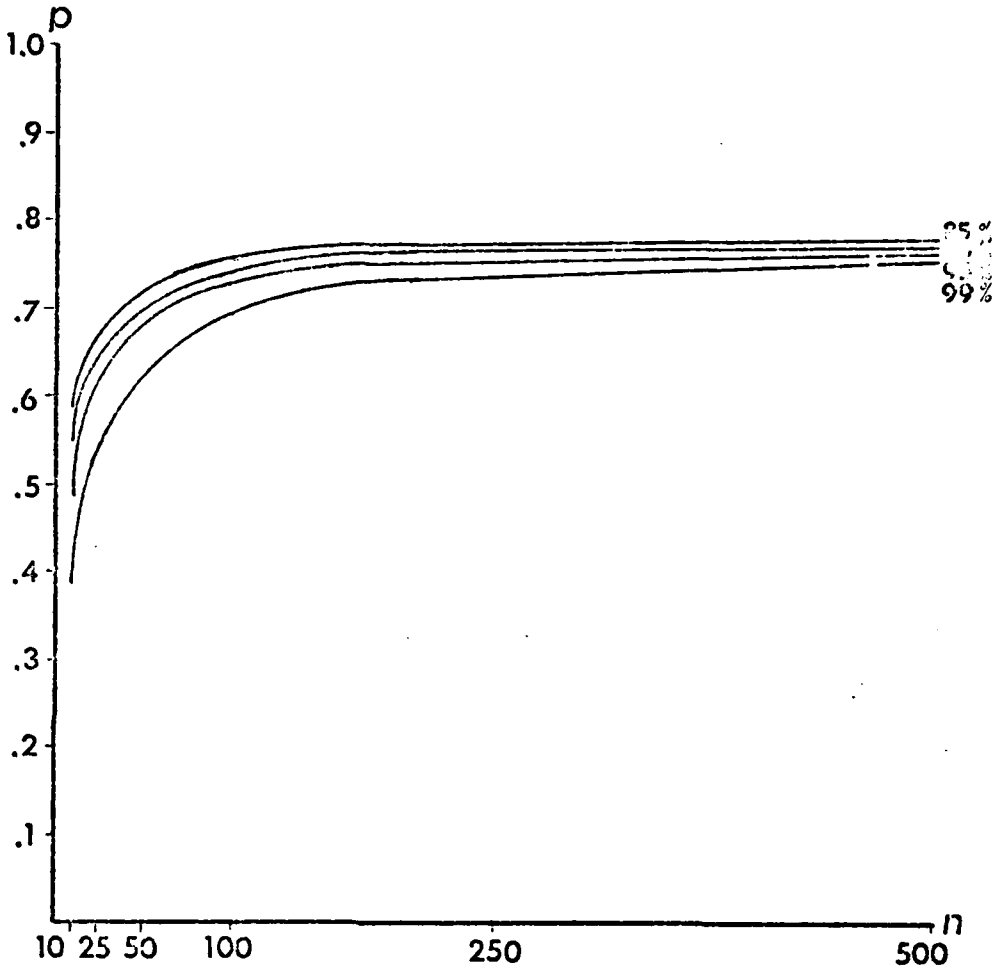
$k = .01$

FIGURE 3



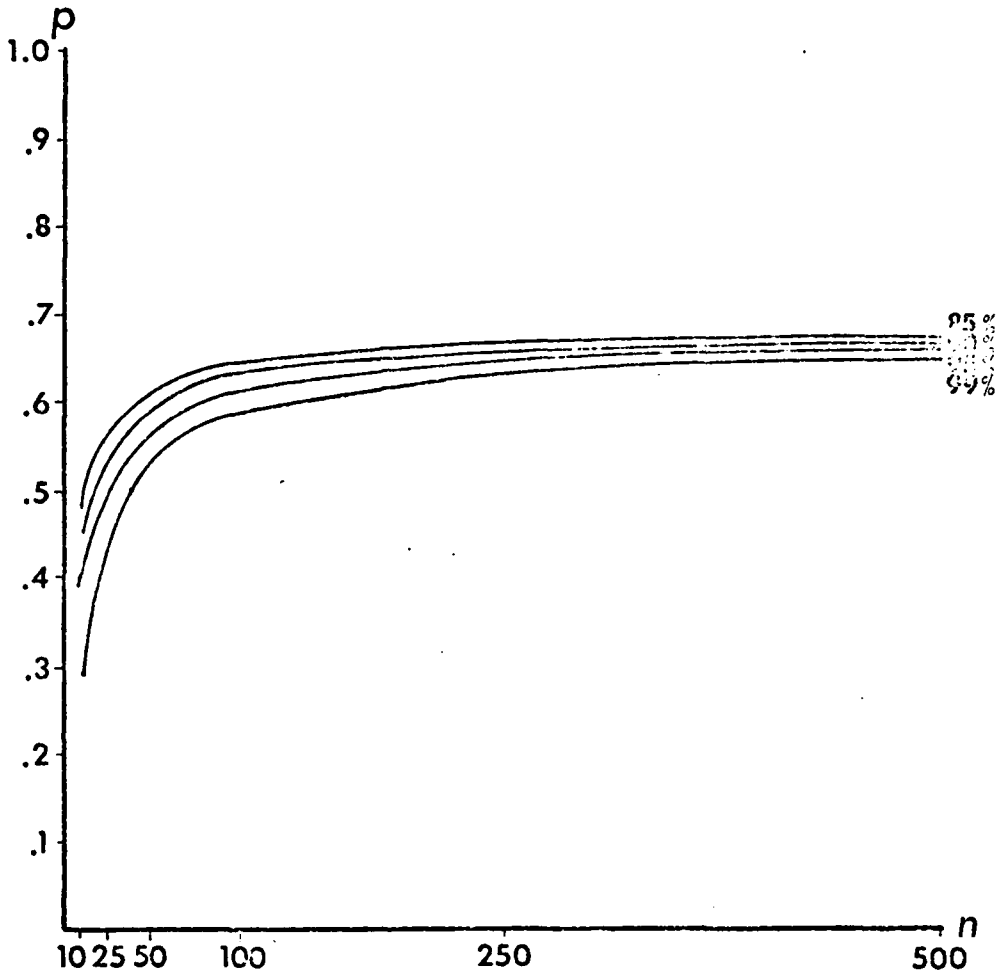
$k = .1$

FIGURE 4



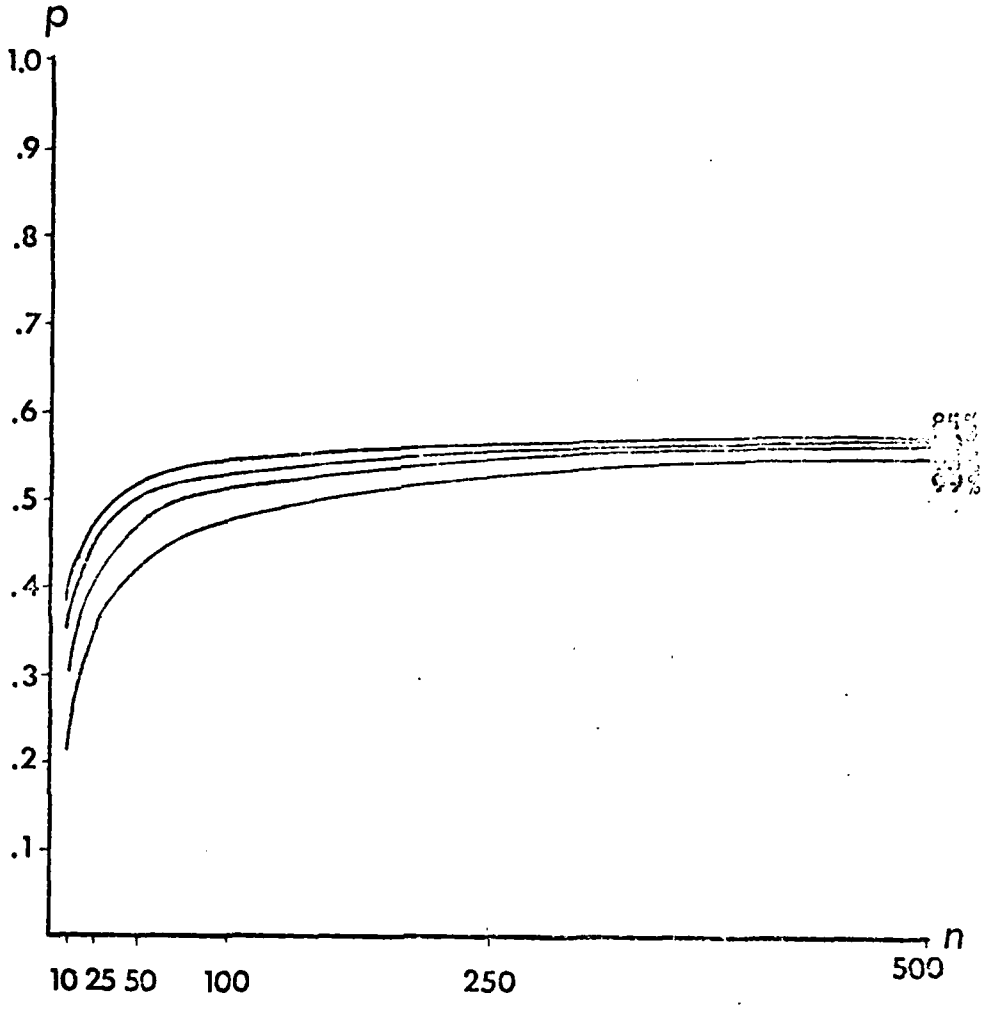
$k = .2$

FIGURE 5



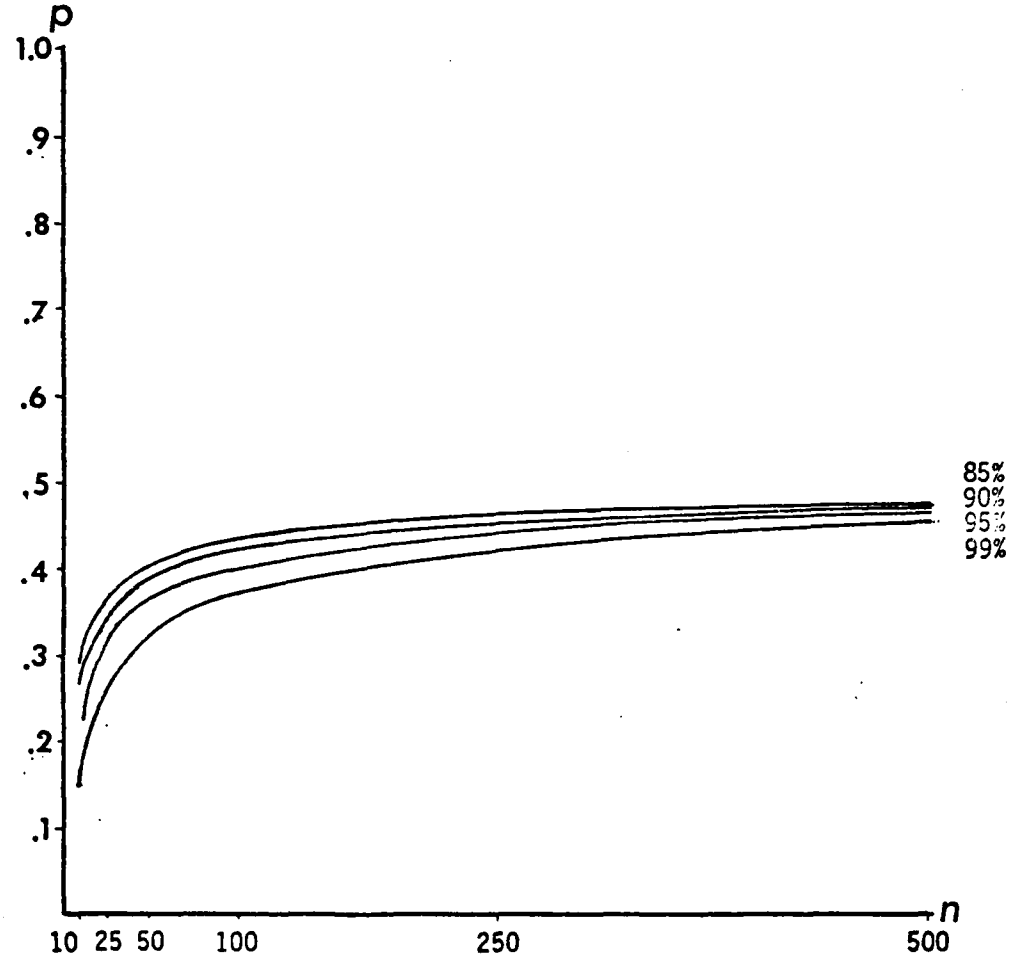
$k = .3$

FIGURE 6



$k = .4$

FIGURE 7



$k = .5$

FIGURE 8

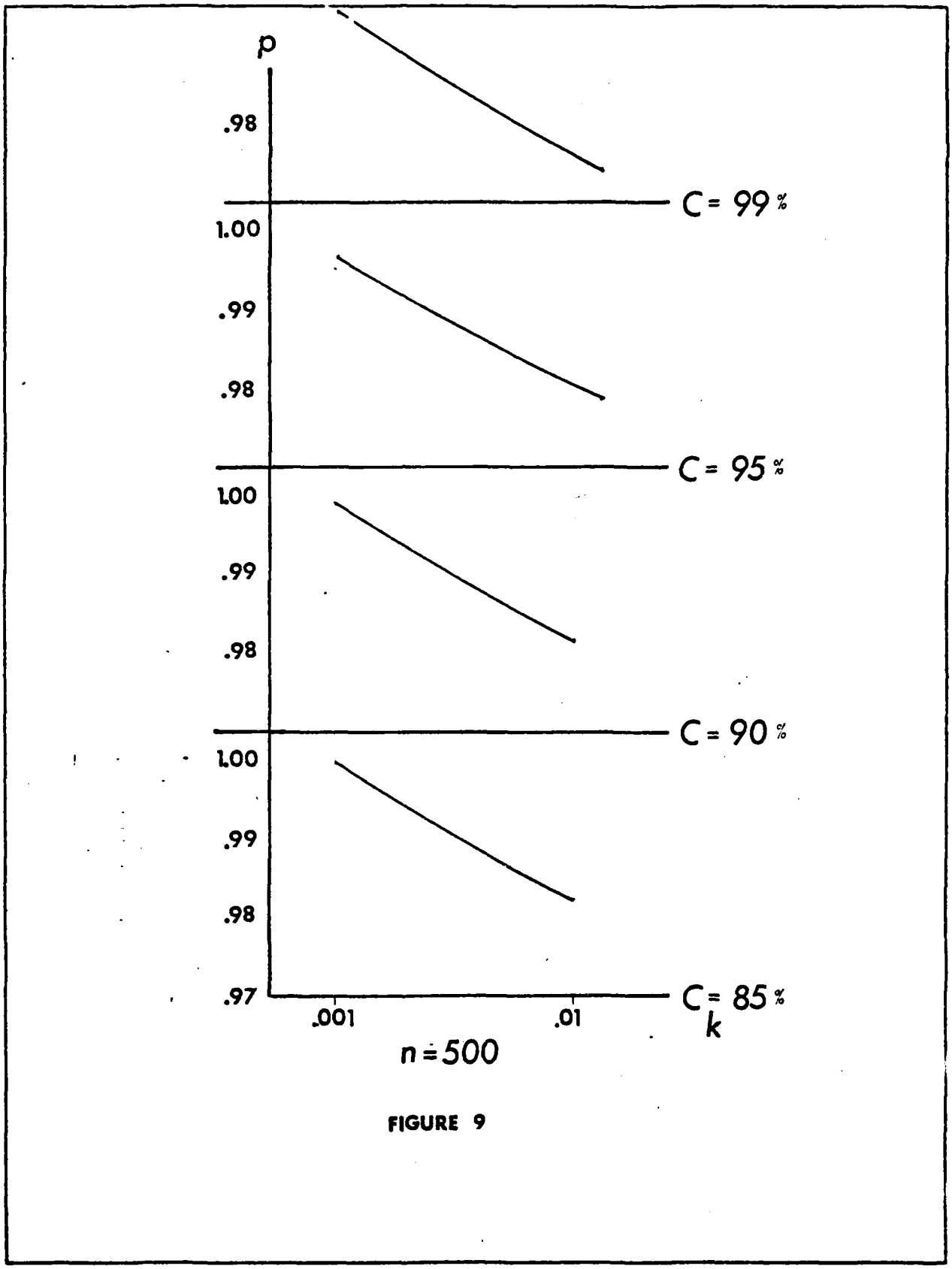
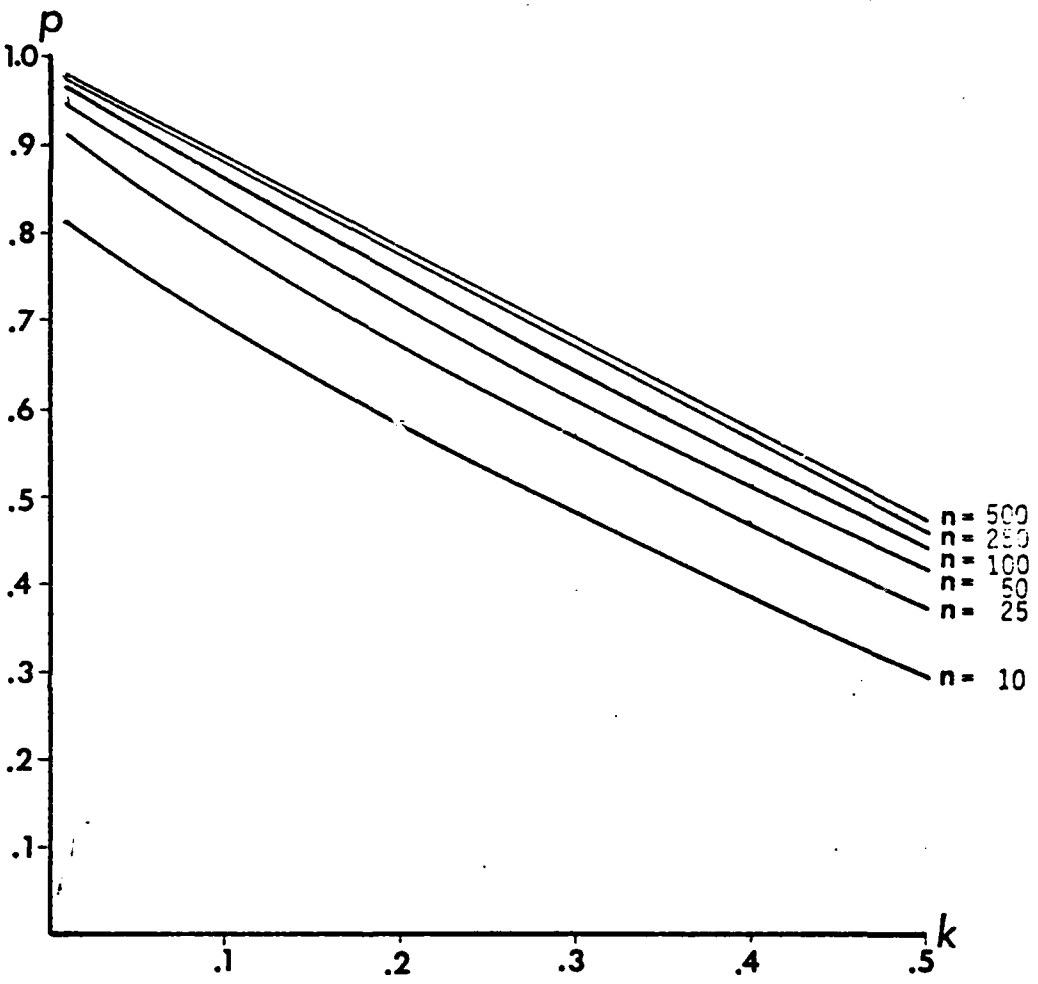
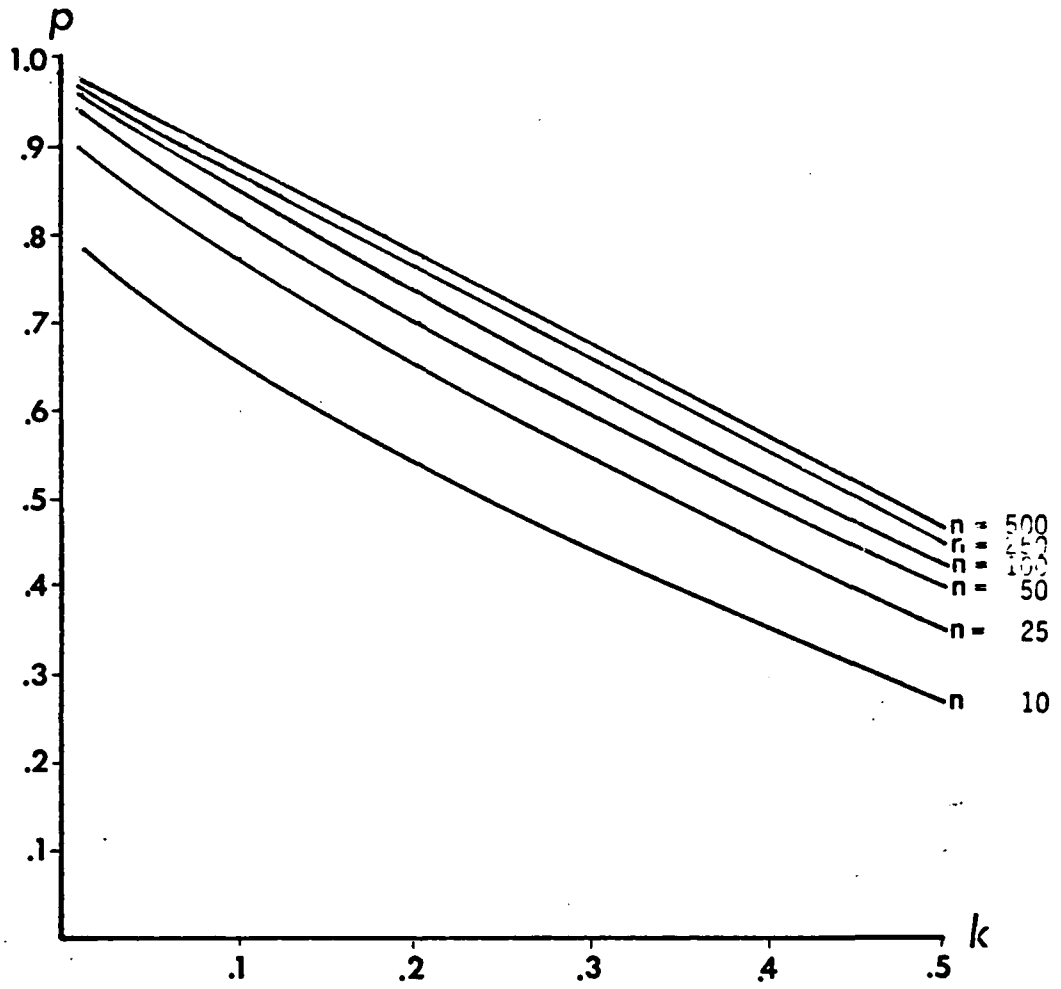

p
.98
1.00
.99
.98
C = 99 %
1.00
.99
.98
C = 95 %
1.00
.99
.98
C = 90 %
1.00
.99
.98
.97
C = 85 %
.001
.01
k
n = 500

FIGURE 9



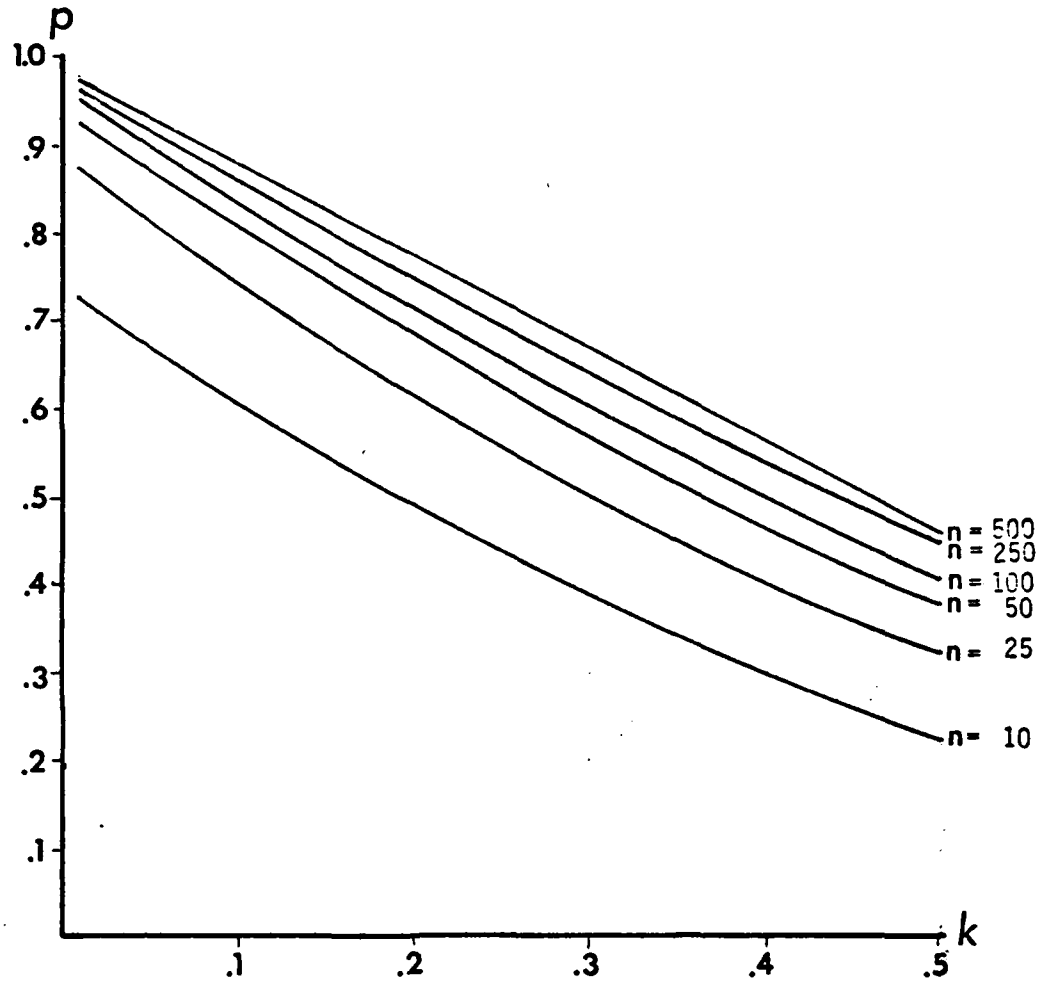
$\hat{C} = 85\%$

FIGURE 10



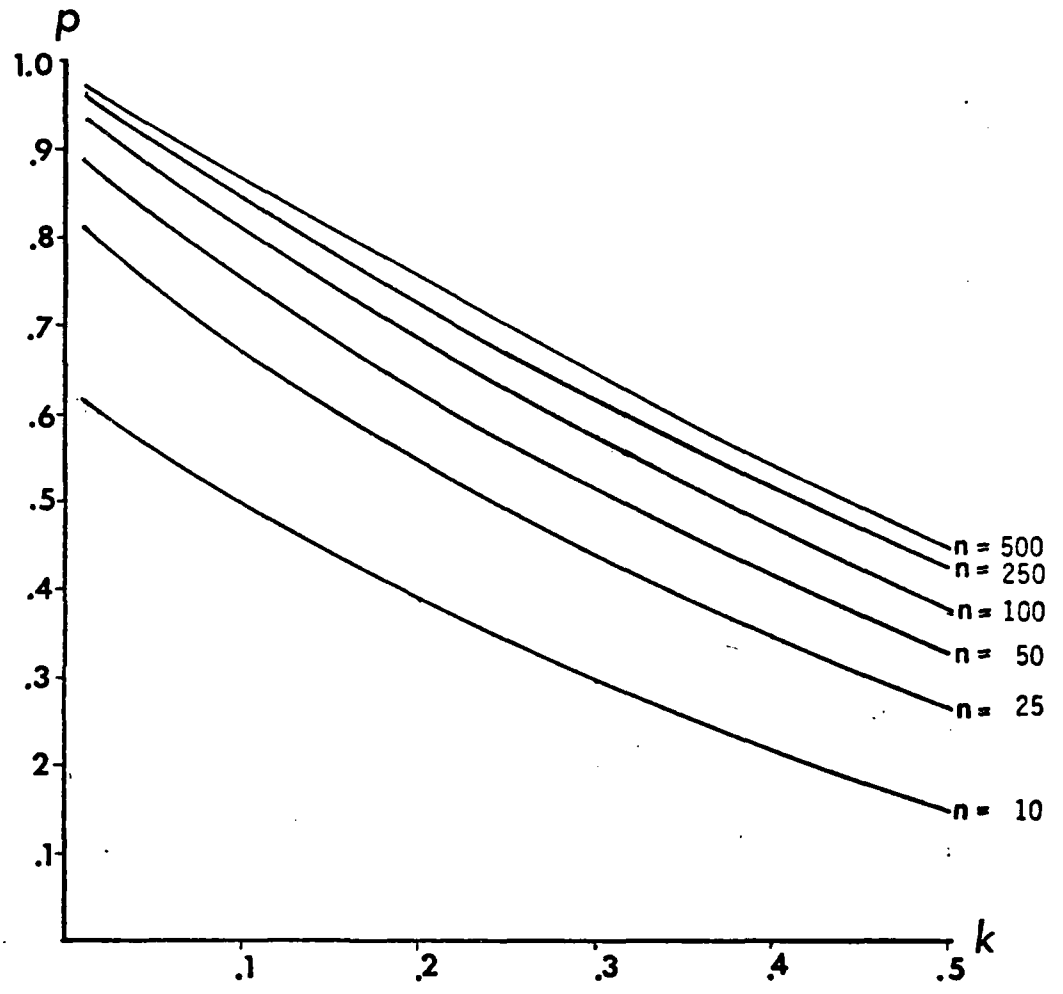
C = 90%

FIGURE 11



C = 95%

FIGURE 12



C = 99%

FIGURE 13

acceptable k for this lot is 14. Out of a sample of 100, 14 can be permitted to fail unsafe (100 x .14 = 14) Therefore, if the number of components having failed "unsafe" reaches 15 during failure mode testing, the test can be terminated because the lot is unacceptable. Section 2.3 contains a more complete description of this process. The result of establishing this threshold can be significant test savings. This is the same concept applied commonly to reliability testing when failure mode data is not required. Figure 9 reflects the graphs for k=.001 in this series. All four confidence levels have been shown here. In this case, the scale of k is an enlarged section of the inner-most portion of the k scale on the other graphs.

The third set of graphs, Figures 14 through 20, uses the same axis parameters as the previous set, but confidence lines are compared by sample size. Given a cost ceiling, or having established a preferred sample size, it is possible to evaluate confidence/safety trade-offs for threshold values of k. These graphs, therefore, enable the analyst to assess the safety penalty he may pay for relaxing threshold requirements on the sample size. Likewise, by imposing more stringent requirements on k, it is possible to see the improvement in confidence for the same reliability (p) requirement, or the improvement in reliability for the same confidence level.

The final set of graphs, Figures 21 through 27, are among the most interesting, and present the most complete picture of population behavior. Once test data (values of k) have been established for a component or family of components, these graphs can be used to

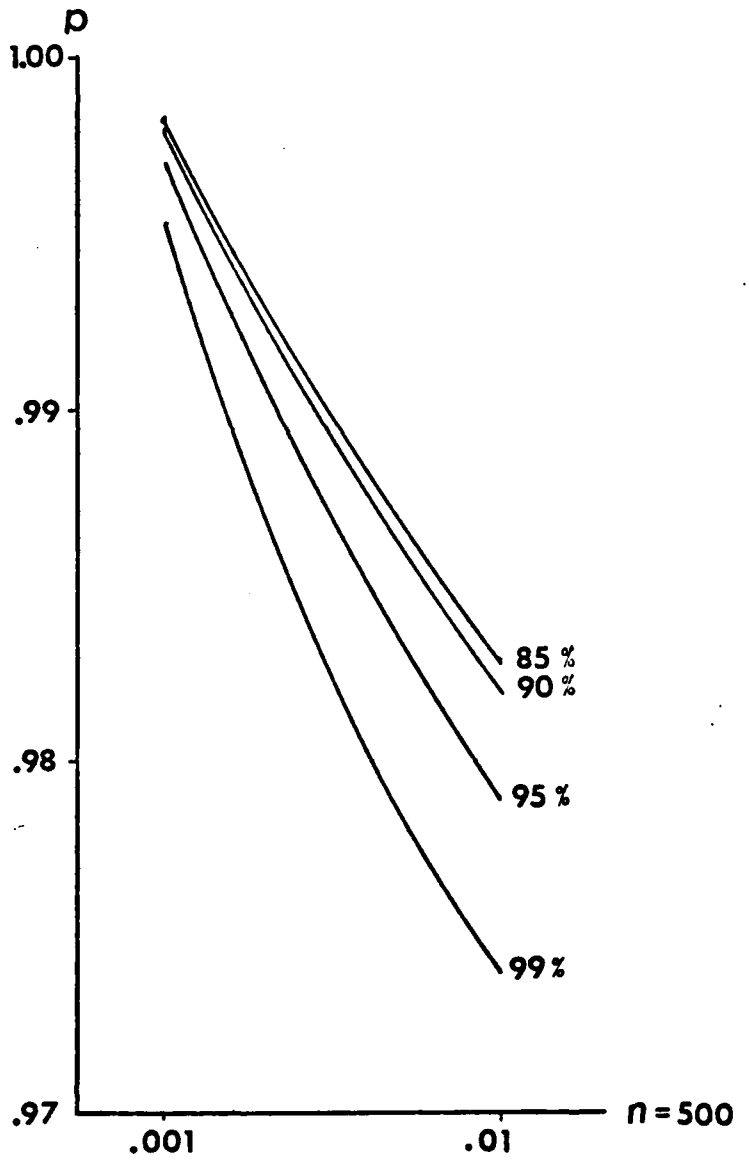


FIGURE 14

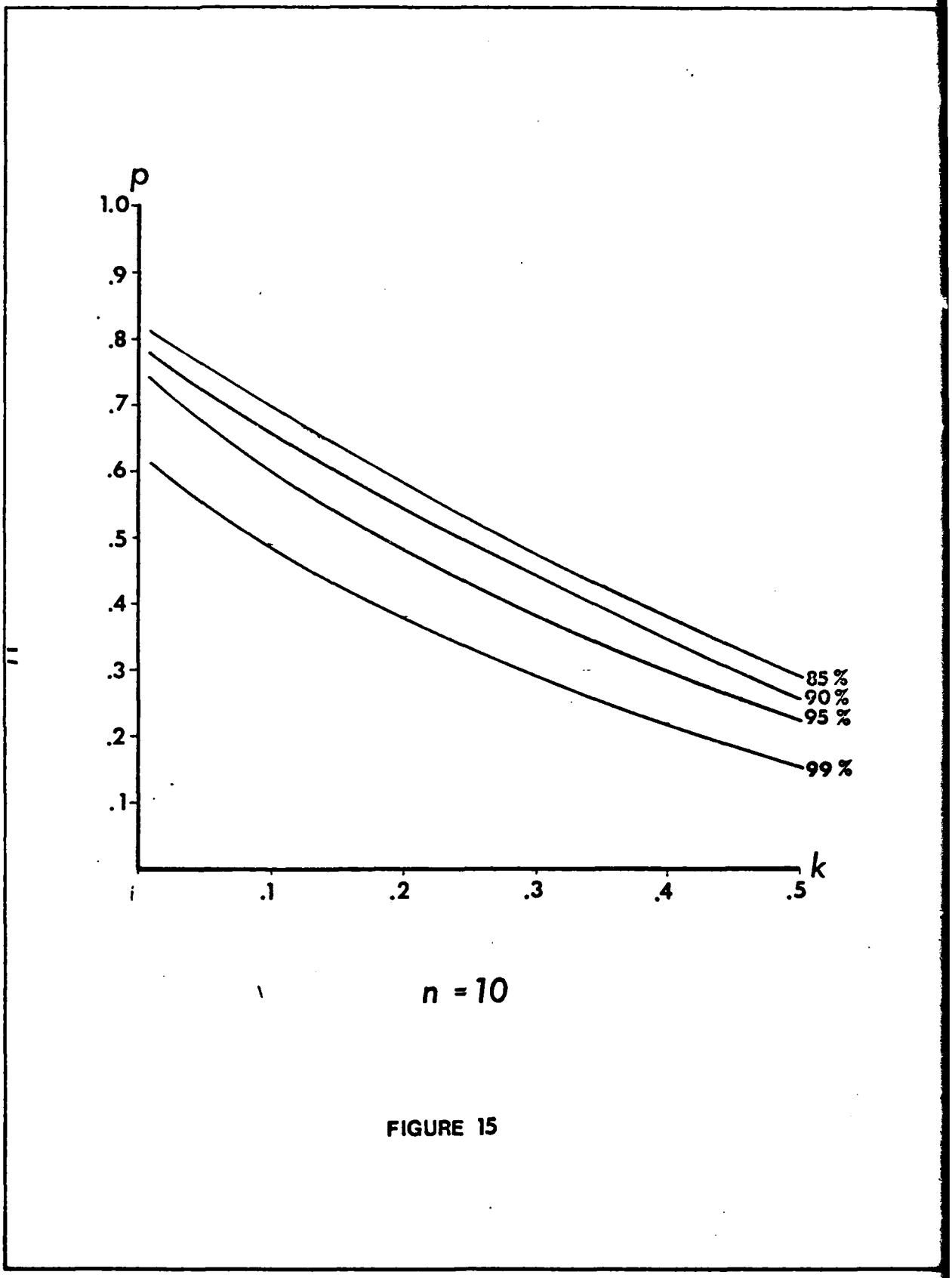
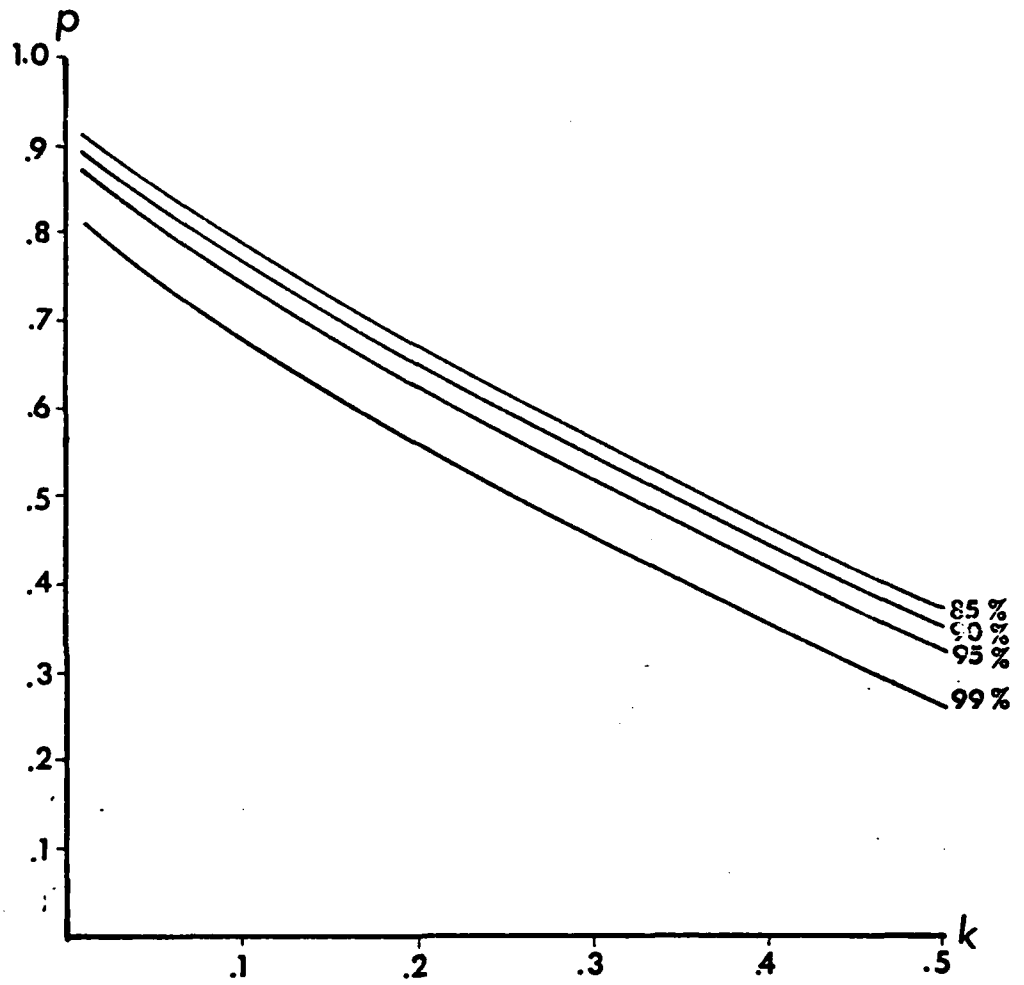
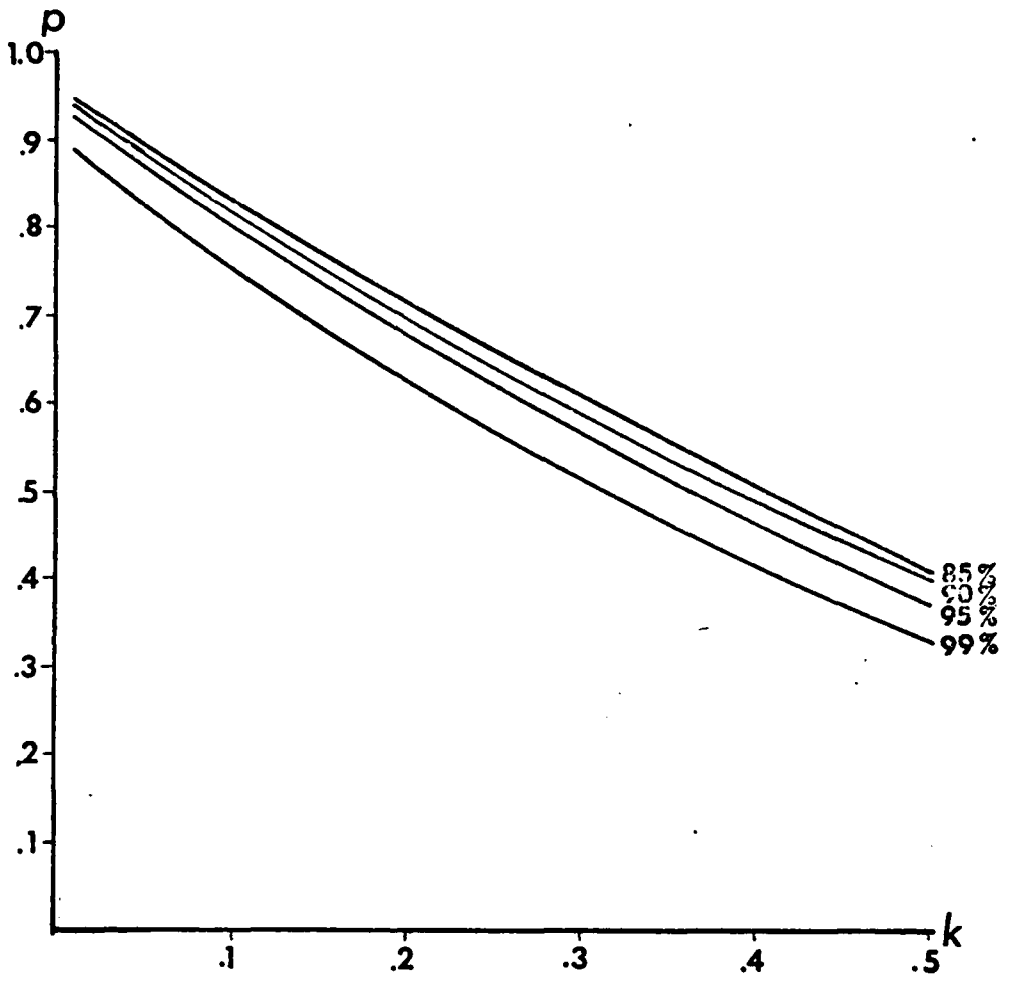


FIGURE 15



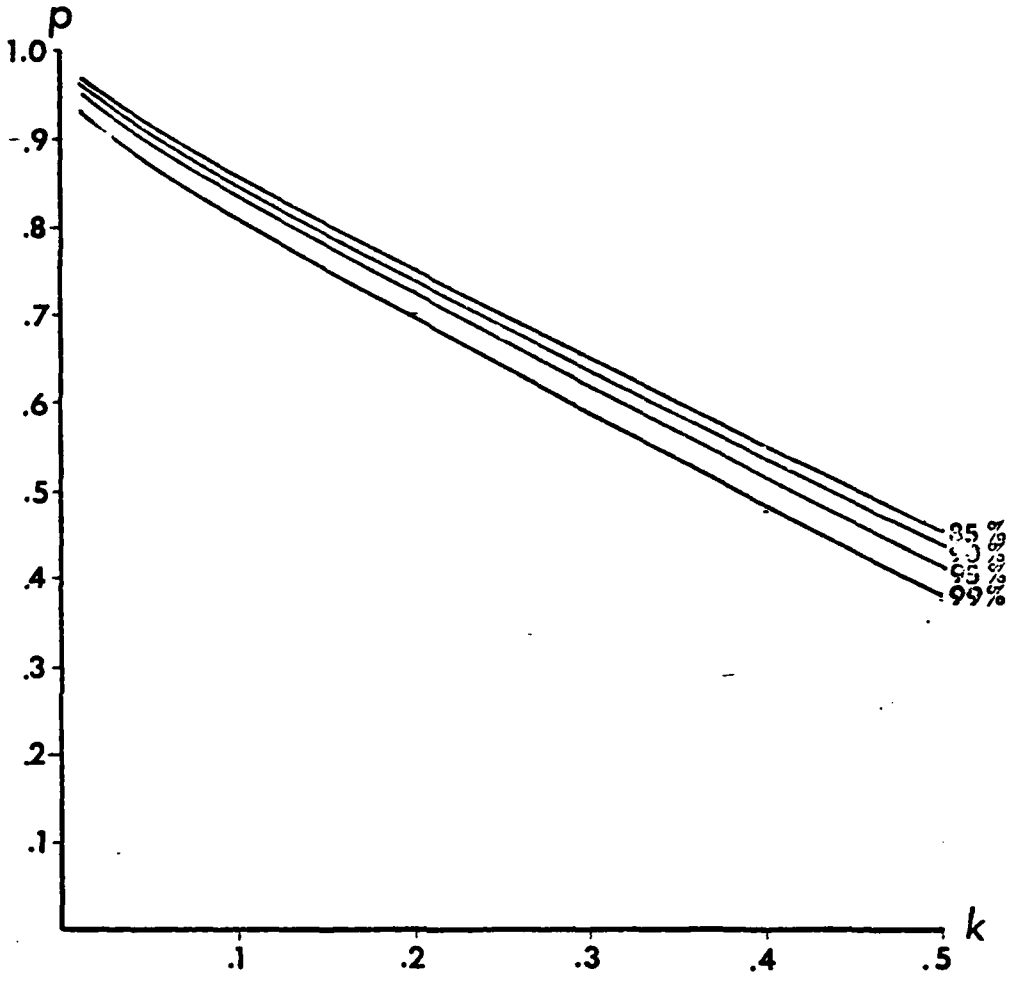
$n = 25$

FIGURE 16



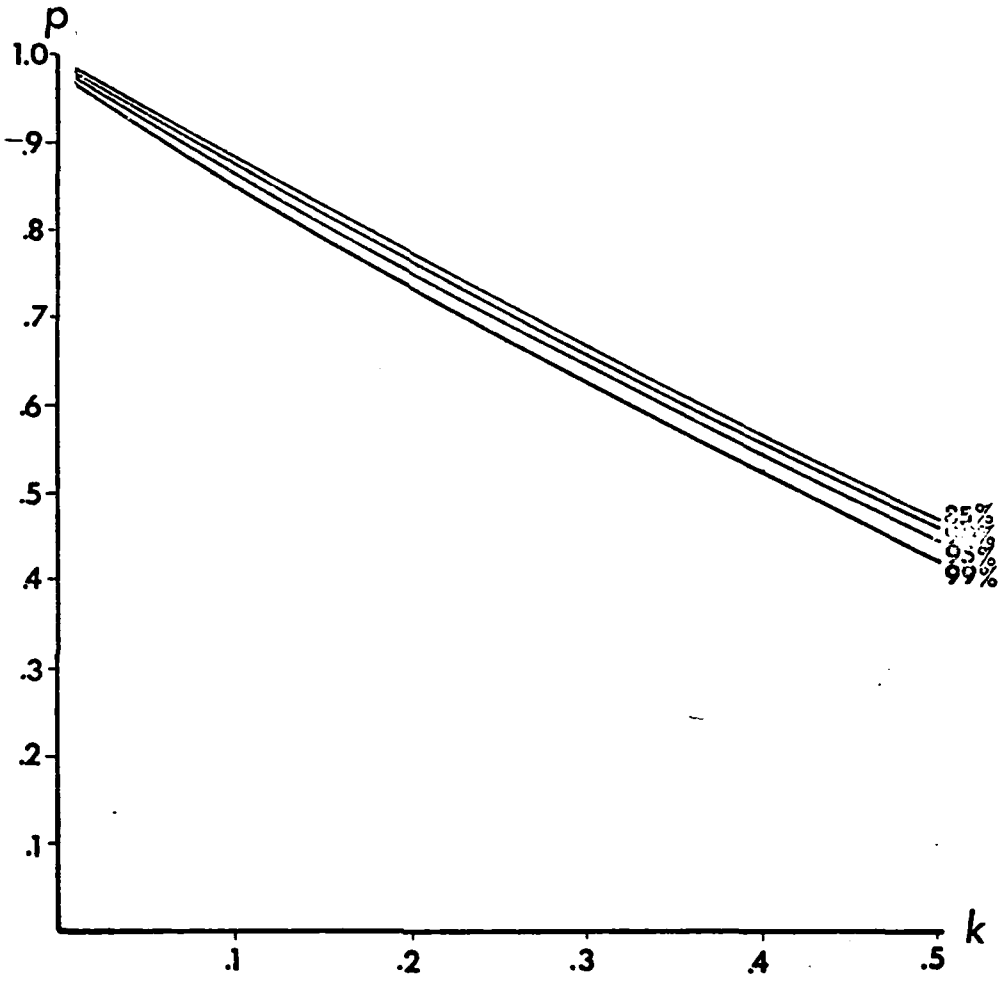
$n = 50$

FIGURE 17



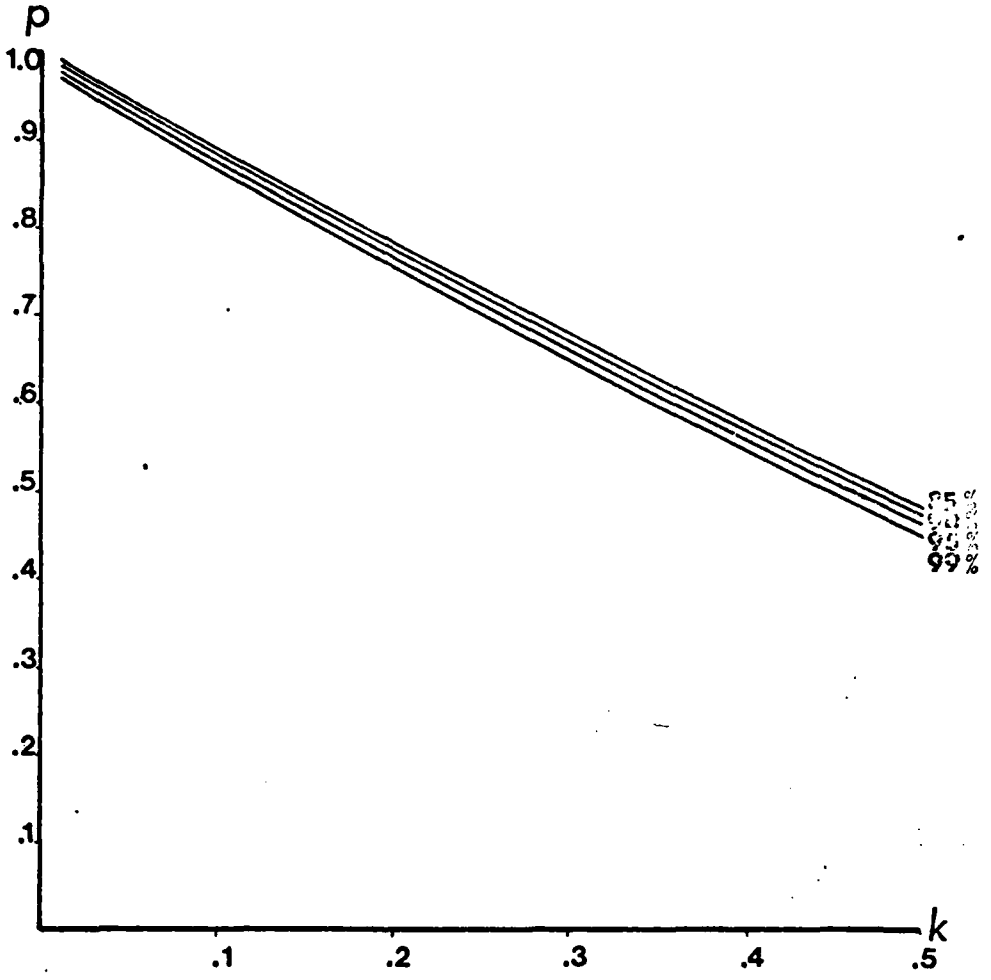
$n = 100$

FIGURE 18



$n = 250$

FIGURE 19



$n = 500$

FIGURE 20

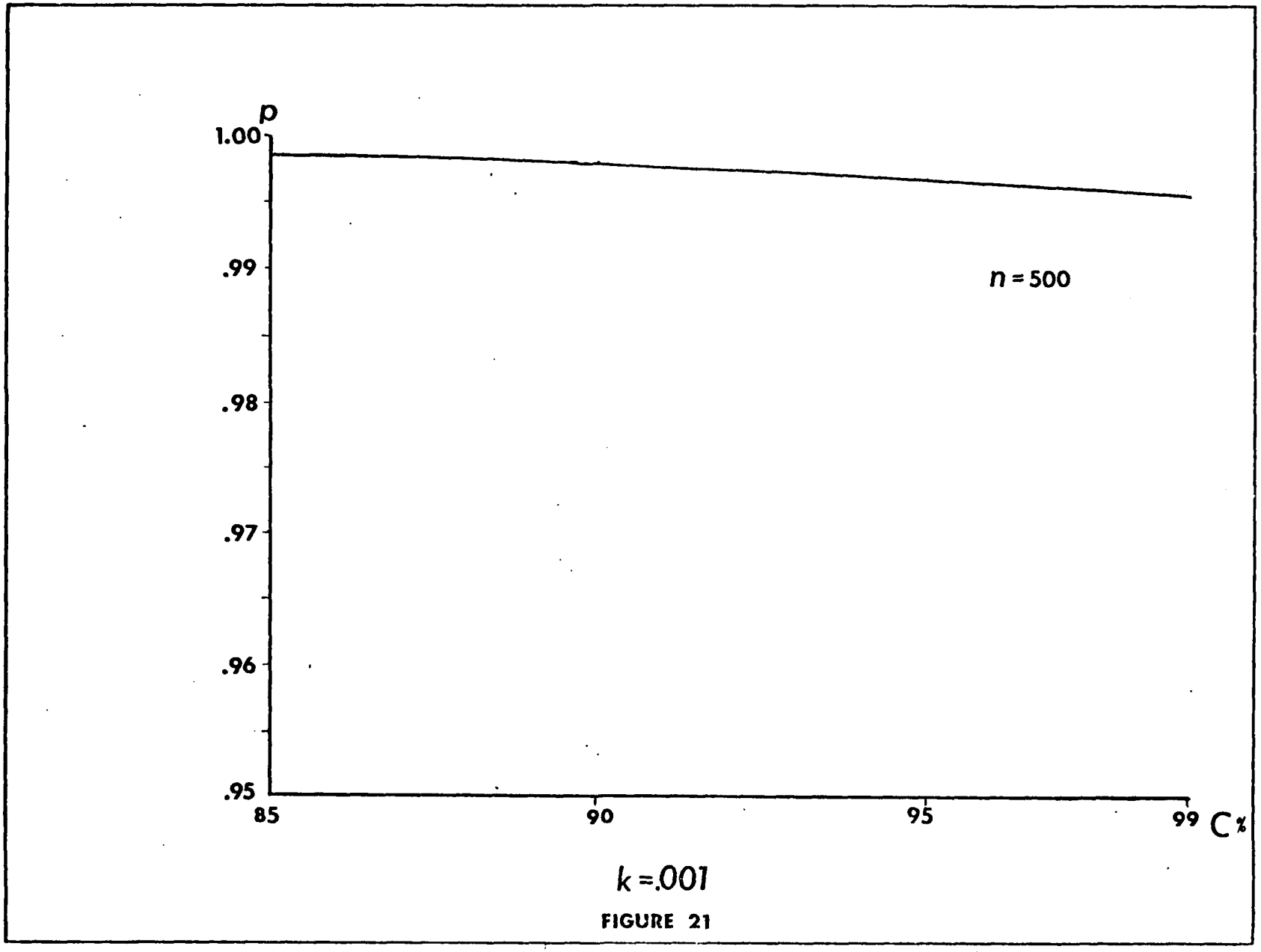
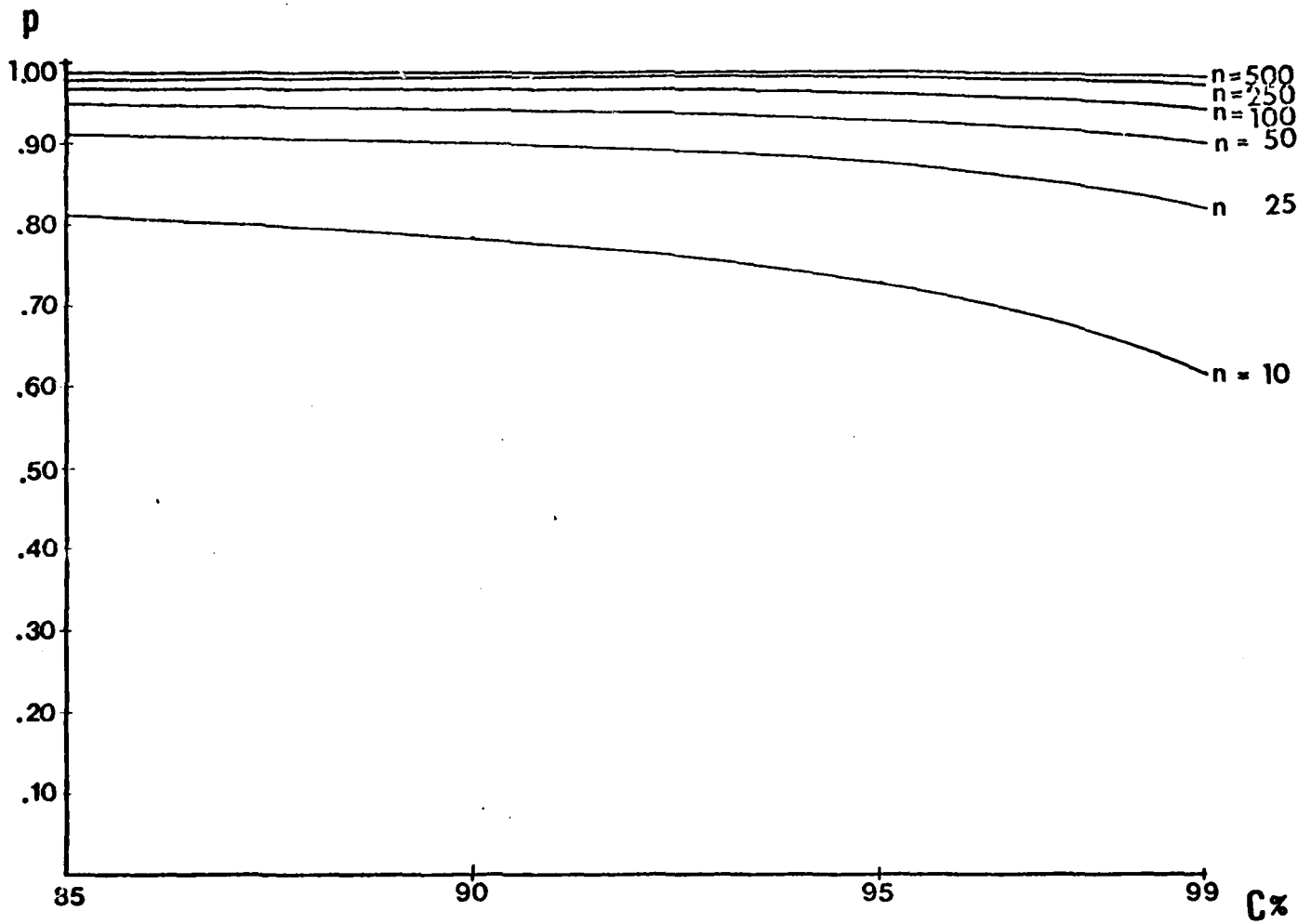
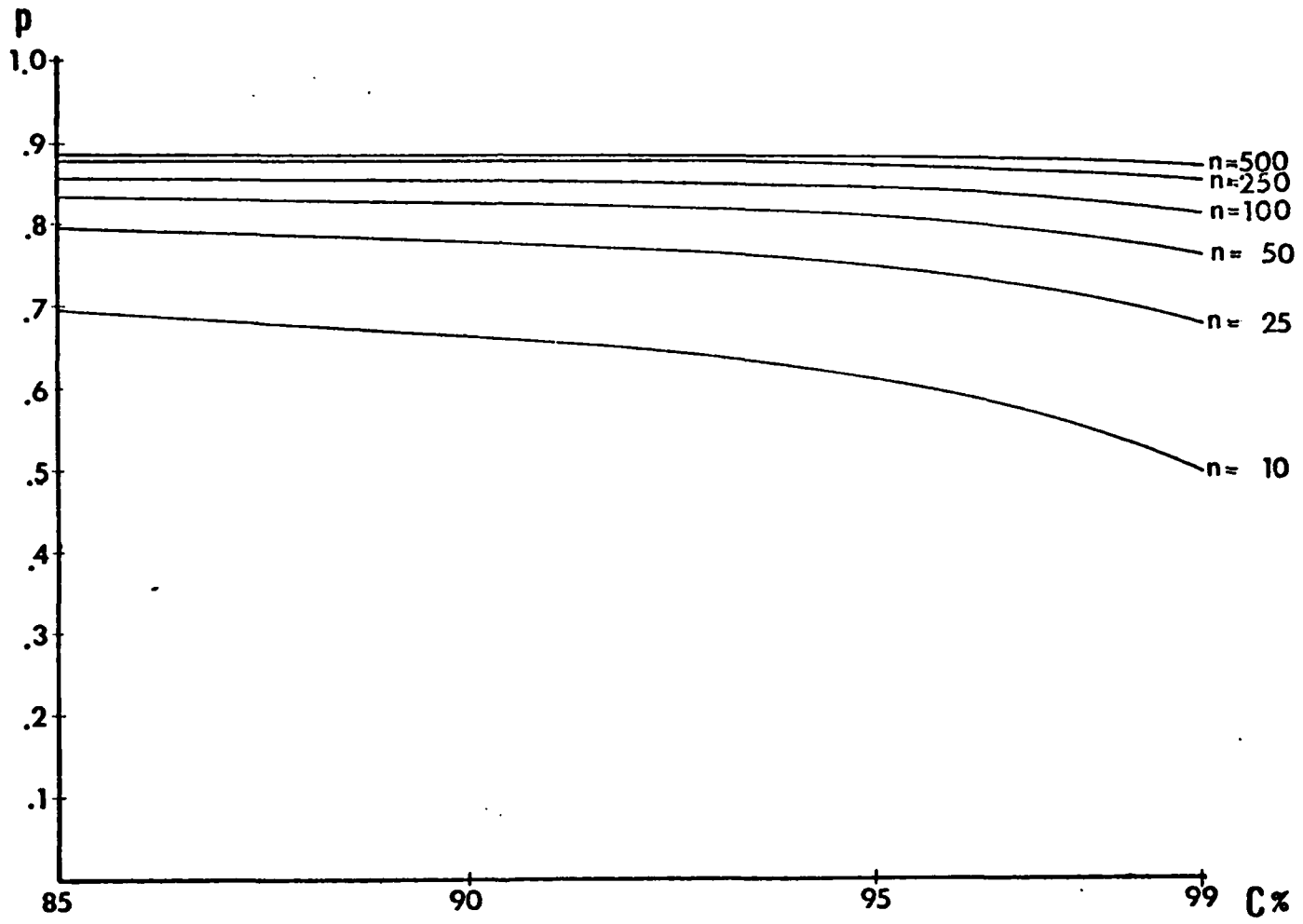


FIGURE 21



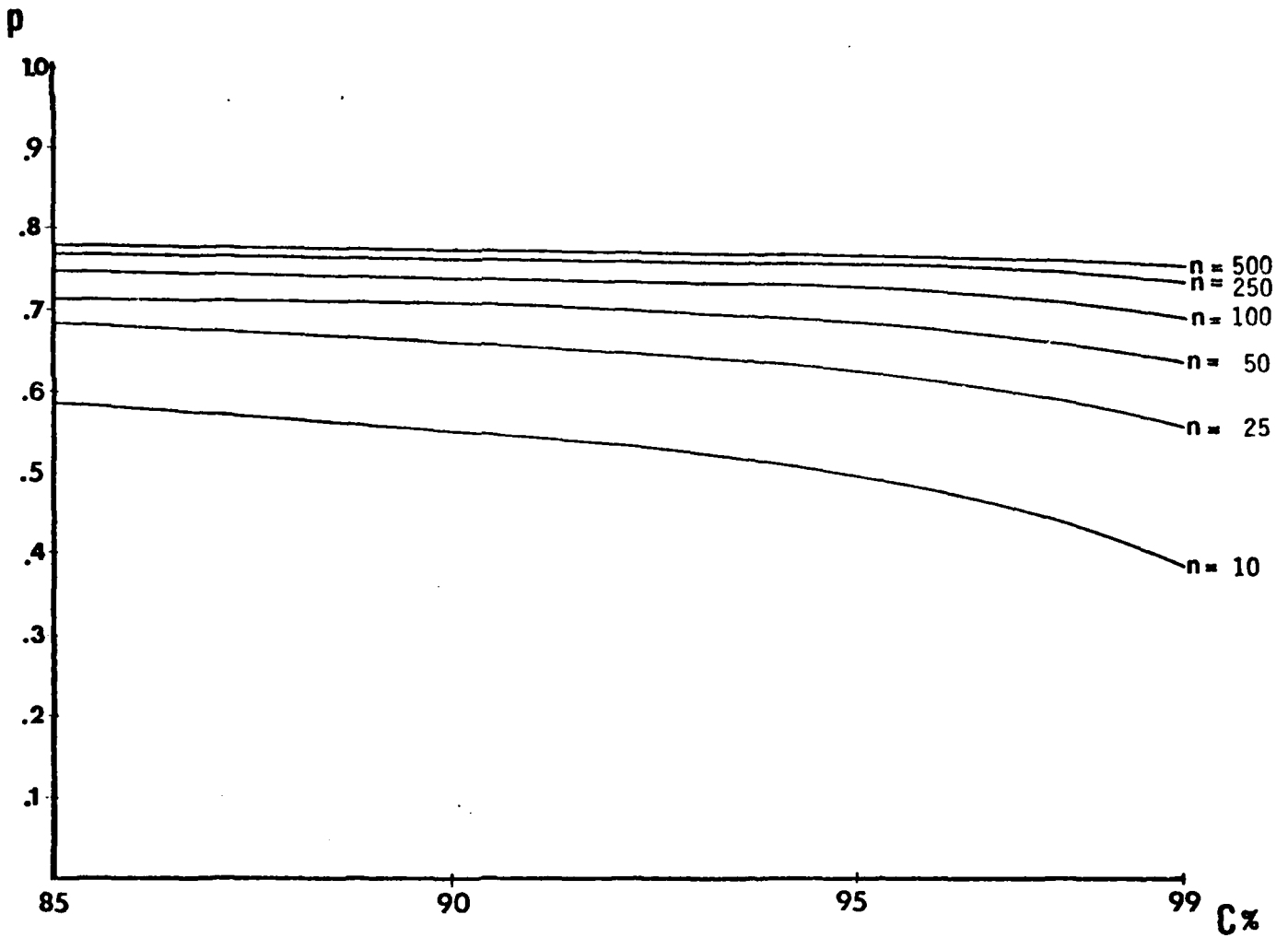
k = .01

FIGURE 22



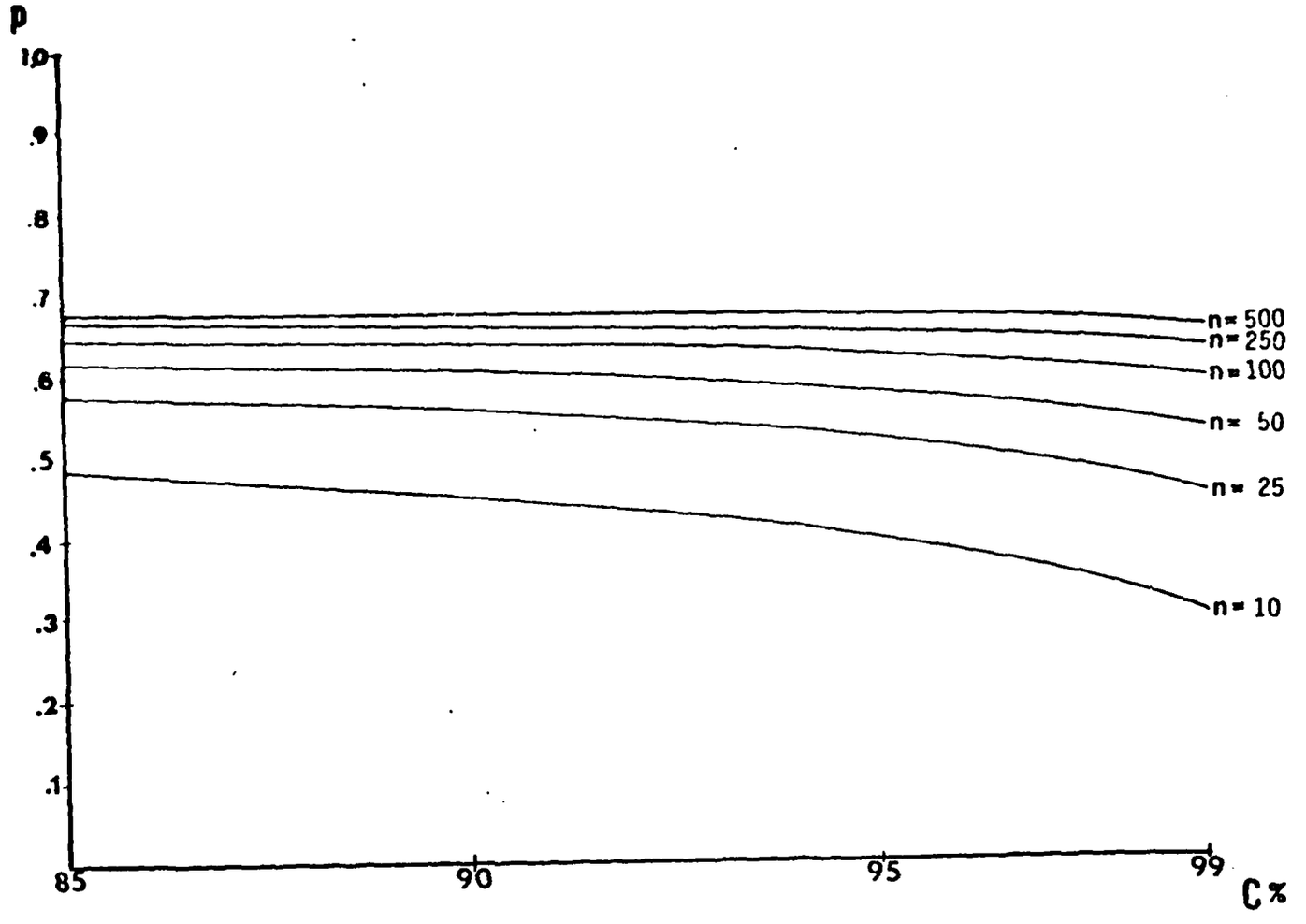
k = .10

FIGURE 23



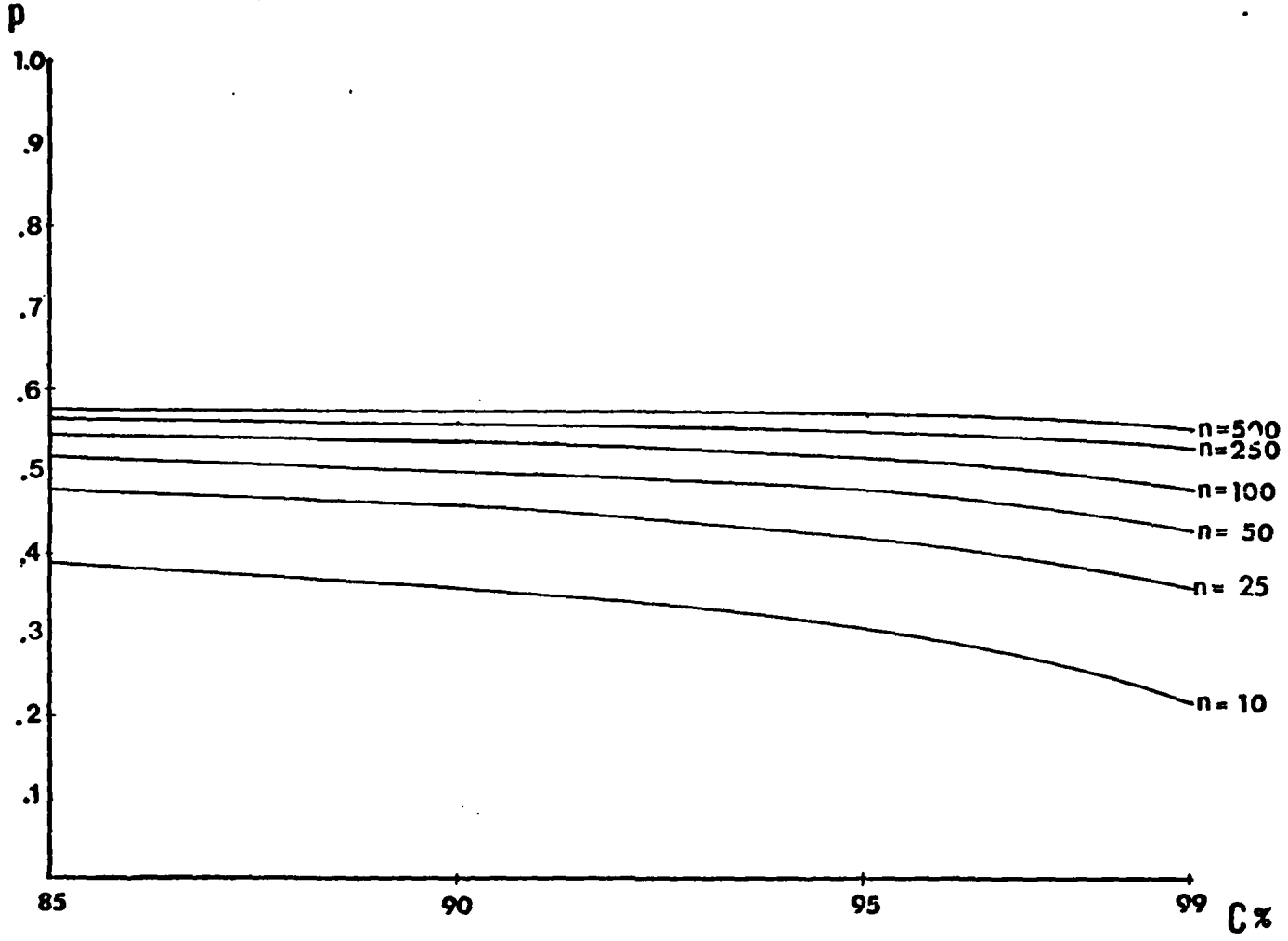
k = .20

FIGURE 24



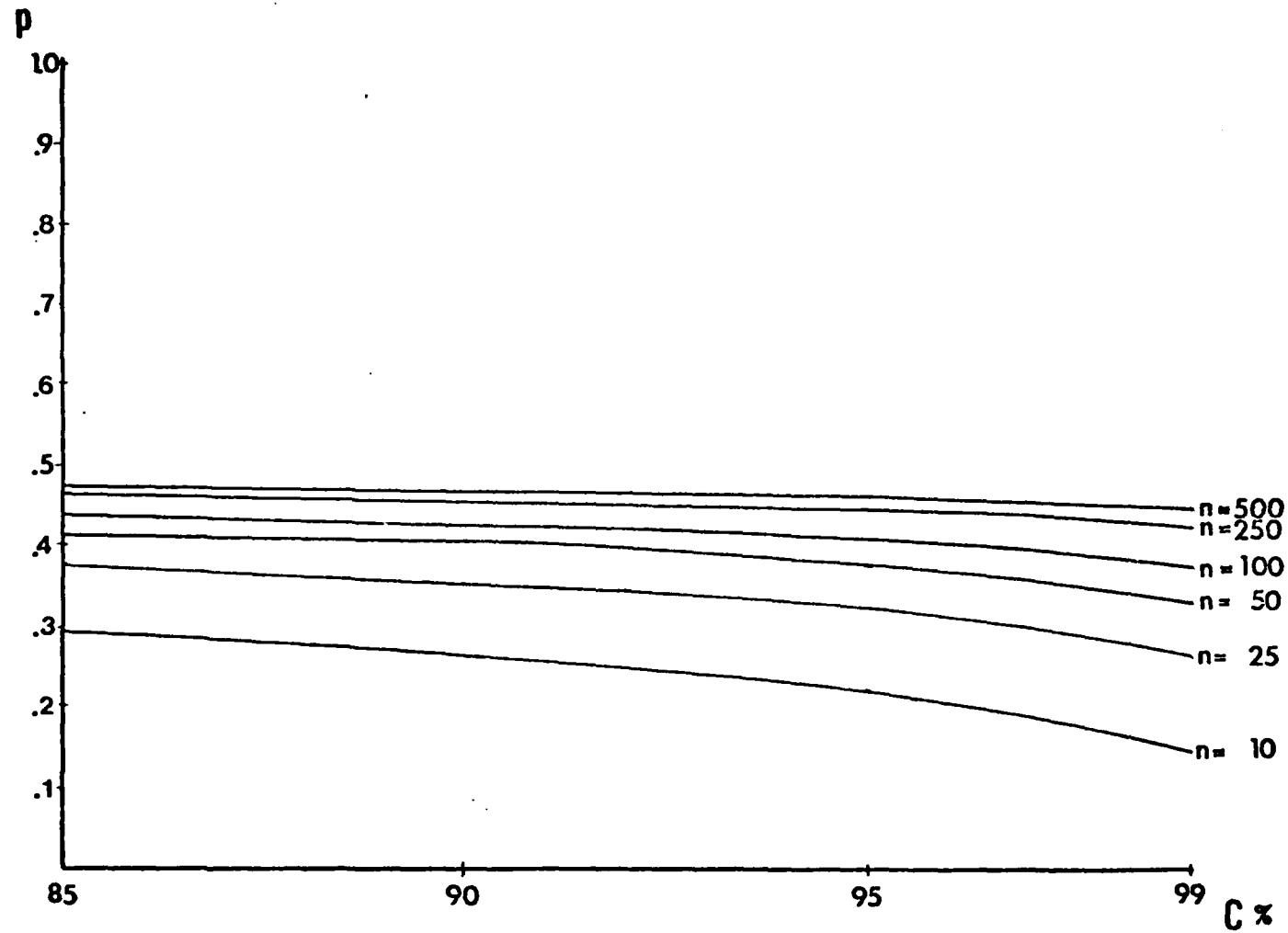
k = .30

FIGURE 25



k = 40

FIGURE 26



k = .50

FIGURE 27

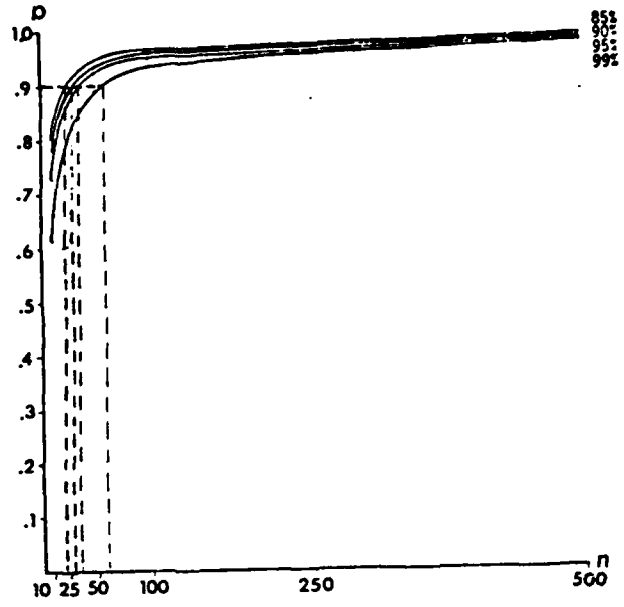
establish test requirements for threshold safety criteria. For example, assume it is determined that a certain family of relays experiences, in general, a proportion of unsafe failures (k) = .10 (90 percent safety reliability in the event of failure), and a requirement exists for that component in a system to meet safety criteria of 85 percent reliability with 99 percent confidence. From Figure 23, we can determine that a sample size of 250 would be required to verify hardware safety reliability at that level of "k". This represents a high test cost. We can see that if the confidence requirement can be reduced to 90 percent, the sample size can be reduced to 100. Conversely, it becomes obvious that, up to about 95-96 percent confidence, confidence can be increased significantly with very little increase in "n". If it were possible to find components with an order of magnitude higher reliability (k=.01), sample size (n) can be further reduced for either confidence level. It would be possible to achieve 97 percent confidence of 85 percent reliability with only 25 components. By testing only ten percent of the original number of components to failure, test costs are significantly reduced. It follows, then, that even if the higher reliability components were ten times as expensive, cost of component purchase for test has not been affected, and actual test costs have been reduced. This, of course, is only one of the cost trade-off considerations involved, but it demonstrates the potentially powerful impact of the application of these graphical techniques.

2.3 USE OF THE CONFIDENCE GRAPHS

A few examples of the use of the probability graphs follow, As described in the previous section, each set of graphs offers a particular type of information. However, because all the graphs include all parameters, any single parameter can be singled out in any graph. A value of interest, therefore, is available regardless of which parameters relative to a component or application are known.

Minimizing Test Costs

The curves on graphs 2-8 are very conducive to determining optimum sample sizes when "k" values are known. That is, when experience or previous testing yields a value of "k" for a component, minimum sample sizes required to verify the desired probability can be taken from these figures.



$k = .01$

FIGURE 3

1. Determine required probability of fail-safe and corresponding confidence level.
2. Look to graph for "k" value corresponding to that associated with the component.
3. Find probability (p) on vertical axis.
4. Read across to intersection with desired confidence level.
5. Read down for minimum sample size.

The example is for k=,01, p=.90, C=95%.

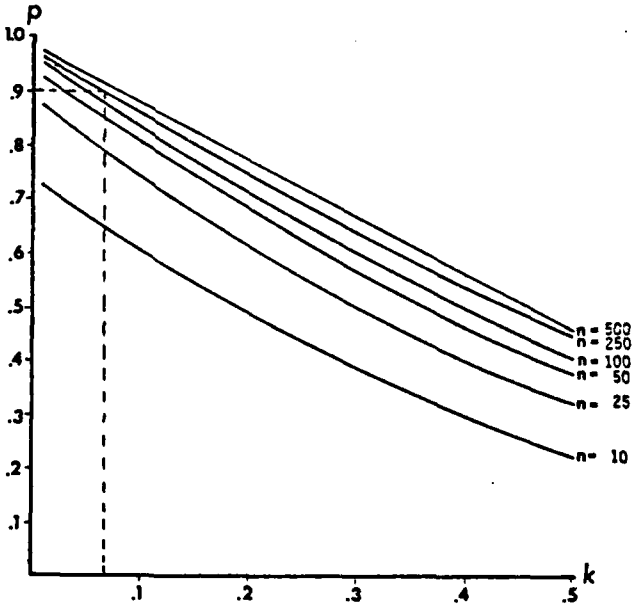
Reliability of Available Components

The reliability of components with a known value of "k" can be taken from the previous graph by reading the curves across from the probability values. Without further testing, for example, it can be shown that the components can be expected to be up to 97% reliable at 99% confidence assuming "k" was established from a large sample.

Determining Threshold Levels

Figures 9-13 are useful for determining test thresholds when the required probability and confidence level are predetermined and of high priority.

1. On the graph for the appropriate confidence level, read across curves from required p.
2. Note reasonable sample sizes. In the example below, C=95%, p=90%, curves for n=10 and n=25 do not intersect line. Minimum sample size can be taken from graphs 2-8 as described above.
3. If testing is required, available components may limit achievable levels of reliability. Eg., if less than approximately 40 components are available, requirements for either C or p or both will have to be relaxed.
4. Once the sample size is determined (optimum sample sizes taken from Figures 2-8) the threshold value of "k" can be read off the horizontal axis.



C - 95%

FIGURE 12

Example shown is for n=250. k is .075. (k)(n)=sampling number. That is, if in the course of testing more than (250)(.075) or 18.75 components failed unsafe, the components do not meet the established requirements. Likewise, once 250-18=232 components have failed safe, the components have been proven as qualified.

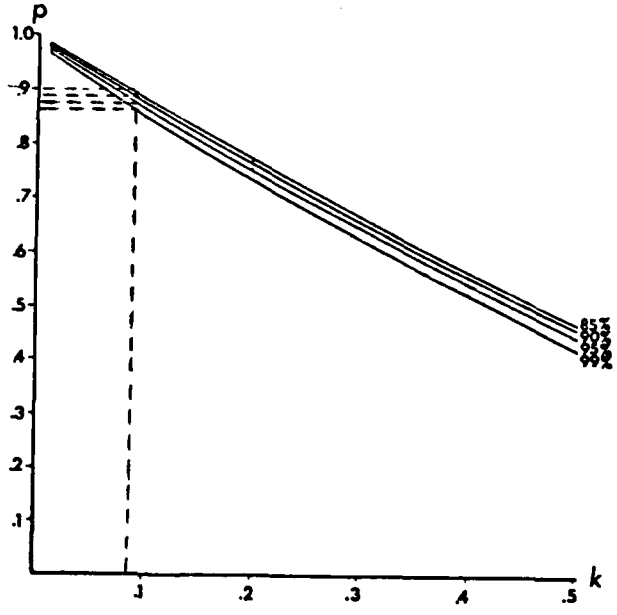
The reliability trade-offs for sample size reduction are readily apparent from these graphs. Assume th requirement for 95% confidence is enforced, but test costs for 250 components are prohibitive. From the sample graph, we can see that 25 components can be tested for probabilities as high as 87%.

Confidence/Reliability Trade-offs

For a known "k" based on a known sample size from testing or field experience, reliability/confidence trade-offs are immediately available.

1. From the graph for the appropriate sample size, read up from component "k" value.
2. The reliability associated with each confidence level is read off the vertical axis.

These graphs are useful in the presence of a cost ceiling (maximum test sample size) or preferred sample size.



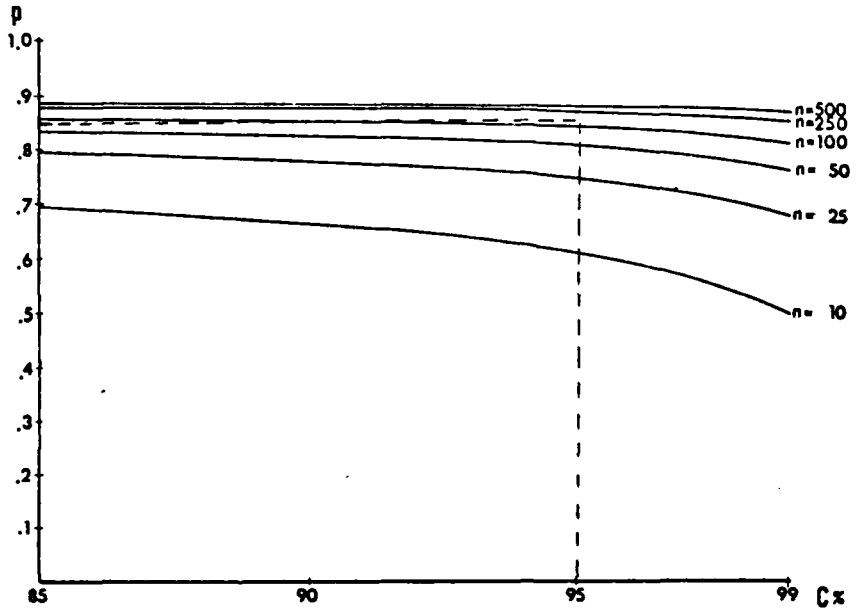
n = 250

FIGURE 19

Parameter Comparisons for Known k

Figures 21-27 make especially clear the relationship between sample size and reliability for the same test results (k). From the sample graph, we can see that the larger the value of n, the less sensitive the reliability to confidence level. Criteria for acceptance testing of components of known k can be established as follows:

1. On graph for appropriate k value, scan across from required reliability to desired confidence.
2. The intersection indicates minimum sample size to verify hardware compliance with reliability requirements.



k = .10

FIGURE 23

In the example, just over 100 components would be required. The actual number required can be taken from Figures 2-8. On the other hand, exactly 100 components can be tested if confidence can be relaxed to 92-93%. Or, if reliability requirements can be relaxed to 80%, 95% confidence can be maintained at n=50.

3.0 ASSUMPTIONS

The use of these graphs, and the application of these principles to failure mode reliability testing, is based on several important assumptions. These assumptions are discussed below.

1) The primary assumption which has been made is that failure mode proportions can be expressed by a constant, "k". Preliminary reviews of test data tend to verify this assumption.

2) One of the most fundamental assumptions made in reliability testing is that the probability of component failure remains constant throughout the interval of interest. This assumption must also be made about failure mode proportions when expanding the concept of reliability distributions to this application. It is assumed, therefore, that "k" is not a function of time and the propensity of a component to fail to a particular mode is constant throughout the life of that component.

3) Any failure mode which is not specifically defined as "fail safe" is considered to be "fail-unsafe". Using relays as an example, fail-open may be determined to be the fail-safe failure mode. Fail-closed, then, obviously becomes the primary fail-unsafe mode. However, chatter, improper operat-

ing time, contact bounce, etc., will all be labelled "unsafe" because their effect on system safety cannot be readily determined. By including all failure modes in the two categories, the binomial requirement of two possible outcomes is met.

4.0 ADVANTAGES OF PARAMETRIC STUDY

It cannot be disputed that a better handle is required for the failure mode distribution of safety related components used in critical applications. Although a wide variety of methods can be proposed which respond to this deficiency, most are dependent upon a complete modal analysis during reliability testing. This requirement represents significant expansion of existing reliability test procedures, resulting in higher testing costs. Of highest priority, then, is development of a method which maximizes the use of data already available. In addition, any concept which would minimize the costs of actual test, where it is required, is most attractive. The technique proposed here accommodates both these needs readily and comprehensively. The advantages which this parametric study offers are:

1) Simplicity: The use of a simple, graphical approach lends itself to immediate application to a broad range of data for any value of n, p, k or C, within the ranges provided.

2) Reduced requirements for special modal testing: By making maximum use of modal data which may already be available, from field experience, for example, the requirement for special testing for modal data is minimized. Where modal testing is required, the same technique applicable to acceptance testing (described in section 2.2) minimizes those test costs.

3) Graphical ease of parametric comparisons: The use of several graphs, each displaying the data from a different point of reference, means that at a glance it is possible to make judgements about the effects of altering parameters in the test environment on test costs and data confidence.

4) Establishes limits for acceptance testing: From these graphs, acceptance numbers, or failure mode thresholds, become obvious. These acceptable maximum values minimize the expense of testing a lot of components which do not meet pre-established criteria.

5) Adjustment for criticality: The range of confidence limits provided simplifies adjustment of hardware requirements to mission or component criticality.

5.0 TESTING FOR FAILURE MODE DATA

The preceding sections describe a technique which maximizes the information available from failure rate/failure mode data, and minimizes the cost of testing to accumulate that data. This section discusses the implications of testing for failure mode data.

Ideally, it would be possible to predict the failure mode of a component by inspection or test prior to use. Second best would be the ability to detect a propensity to failure mode at some point during the use of that component, or, for example, during burn-in. Unfortunately, neither of these capabilities exist to date. Therefore, it is necessary to test a part to failure to determine its failure mode. In most aerospace applications we are working with very high reliability components. Ball Brothers Research Corporation Preferred Parts Handbook for Space Systems[3] references all failure rates of preferred parts in 10^9 hours. Typical failure rates for capacitors and resistors are less than 10 failures per 10^9 hours. Diodes and transistors are less than 100 failures in 10^9 hours. Failure rates for microswitches are quoted at 250 failures per 10^9 hours. Relays vary from 250 to 900 failures per 10^9 hours, depending on the type of relay.

---

[3]Ball Brothers Research Corp., Preferred Parts Handbook for Space Systems, Rev. B, February 1971, pg. A-7.

It becomes readily apparent that a great many hours of testing will be required to test such high reliability hardware to failure. Equally apparent are the cost implications of this testing.

5.1 THE COSTS OF COMPONENT TESTING

A rule of thumb for component test is $1 per test socket for life test, $1 per month of test, and $1 per examination to determine if the component has failed. Realistically, of course the cost for testing various components varies with the type and complexity of the component, the type of information to be gained from the test, the time required to complete the test, the amount of failure analysis to be completed on a part, the confidence requirement on the statistics to be gathered, and the cost of the component itself.

Testing to failure of a battery of components whose failure rates are 10^{-9} hours will require a great deal of patience. 10^{9} hours is over 100,000 years! Obviously, accelerated testing is required. Costs for accelerated testing are significantly higher than those for normal life testing. Because of the stresses parts are subjected to in accelerated testing, test sockets must be more complex. The cost per socket can easily go as high as $35, as compared to $1 for simple test*. Likewise, the design time for the test is much higher, increasing engineering costs for the test. Unfortunately, another major factor of accelerated life testing is reduced confidence in the test results, particularly when failure mode is the parameter of concern. It is impossible to predict how the accelerated test environment has altered the components failure mode propensity.

---

*Costs for testing are those used for test estimates by the Boeing Aerospace Failure Analysis Laboratory.

Another factor affecting the cost, accuracy and confidence of failure mode testing is the level of technology available at the testing laboratory. For example, the difference in the rejection rate of a highly technical, well set up laboratory with all the proper equipment, and a less qualified lab might be as much as an order of magnitude. Frequently it is found that the manufacturer himself is best equipped to test his own components.

Components commonly found in safety related circuits include transistors, diodes, resistors, capacitors, and relays. Using these components as examples, some cost comparisons have been made.

As mentioned above, one of the variables in component test cost is the test socket. Costs for ordinary test sockets for resistors, capacitors and diodes are in the area of $1, $2 for transistors. Testing of relays, however, requires more complexity both in the test set-up and the test itself. Engineering design time and the cost of the test socket go up significantly. These sockets may be more on the order of $25 each, and may increase by an order of magnitude or more for accelerated testing.

Another variable in test cost if the number of times the components are examined for failure. If the ability to monitor the components is built into the test socket, the cost is increased again.

The cost of failure mode analysis is another significant cost of accumulating safety data. On the average, test labs charge eight hours to analyze the first component, and one hour per part thereafter. It is possible to design a test socket which will monitor the part for a particular failure mode, as well as for failure. This greatly reduces analysis cost and simplifies the accumulation of statistics.

However, it again increases the cost of engineering design time for test and socket. Another factor in mode/monitoring test sockets is the reliability of the socket itself. If each reported component failure is assumed to be actual failure of the part, and is not verified through failure analysis, the test data is likely to be on the conservative side.

5.1.1 COST EFFECTIVE COMPONENT MODEL

One important consideration in test cost, feasibility and hardware reliability is the trade-off between high reliability components and the cost to test them. A method to calculate the unit cost of component test as a function of component quality appears in a 1975 report by Goddard Space Flight Center[4]. This report compares, for the sake of example, three quality levels of transistors: JANTX, JAN, and commercial grade, with the intent of determining the economy of the hardware as a function of hardware cost and test cost.

A cost model was developed which compares the cost per unit of accepted components:

$$C = \frac{Q+S}{R}$$

where C is the accepted unit cost
Q is the purchase price of one part
S is the cost to screen the hardware , and
R is the ratio of accepted screened hardware to total units.

C, therefore, is generated as a function of transistor quality. To apply this technique to failure mode analysis, we convert S to the cost of life test and failure mode analysis, and R to the ratio of parts failing safe vs. unsafe. C, then, will now represent the unit cost of safety testing the hardware. Q and S will both increase with increasingly reliable parts. The behavior of R will depend on the failure mode propensity of that hardware. The most valuable

---

[4]Goddard Space Flight Center, Failure Analysis Section, Evaluation of GSFC Testing, Relative Quality Levels, and Cost Effective Selection of JANTX, JAN, and Commercial Grade Transistors, Parts and Components Evaluation Report #09-012, May 1975.

application of this model is in the determination of unit crossover points. It is possible to establish the most economical component category. For example, two components: X_1 and X_2 both meet the reliability criteria for a specific application. Their actual modal distribution and failure rates are different, however. This model can determine for us whether the more reliable of the two components is actually more economical in application. The utility of this model assumes that the variation in the quantity of components purchased for the test has a significant impact on test costs. This is a valid assumption when testing an optimum number of parts to attain a predetermined confidence level of reliability, as described in section 2.0 of this study.

To determine the most economical component category, compare component 1 and component 2 by determining the accepted unit cost for component 2 which is the same as the accepted unit cost of component 1.

$$\frac{Q_1+S_1}{R_1} = \frac{Q_2+S_2}{R_2}$$

Solve for Q_2:

$$Q_2 = \frac{R_2}{R_1}(Q_1) + \frac{R_2}{R_1}(S_1) - S_2$$

If the actual purchase price of X_2 is less than the calculated Q_2 ($X_2 < Q_2$), then the accepted unit cost of X_2 will be lower than that for X_1. If $X_2 = Q_2$, there is no cost advantage for either component. If $X_2 > Q_2$, then X_1 is the more economical component. In this way it is possible to order component quality levels according to relative accepted unit cost and select the most cost effective component category.

5.2 COST/CONFIDENCE TRADE-OFFS

The most flexible variable in safety analysis is confidence, as revealed by the parametric study in Section 2.0. This study would not be complete, therefore, without a discussion of the cost of confidence. Sample components have been selected from those safety related components listed previously: transistors, diodes, resistors, capacitors and relays. Actual field experience failure data has been found for each component, as well as a ballpark figure for component cost. A cost comparison for various confidence levels at a constant reliability value is shown in Figure 28. The sample size required and corresponding component costs, approximate test costs based on those estimated costs discussed previously, and test time by component failure rate have all been considered. Actual test time for each component under accelerated test has not been calculated. The comparisons, therefore, provide relative information by component type and confidence cost.

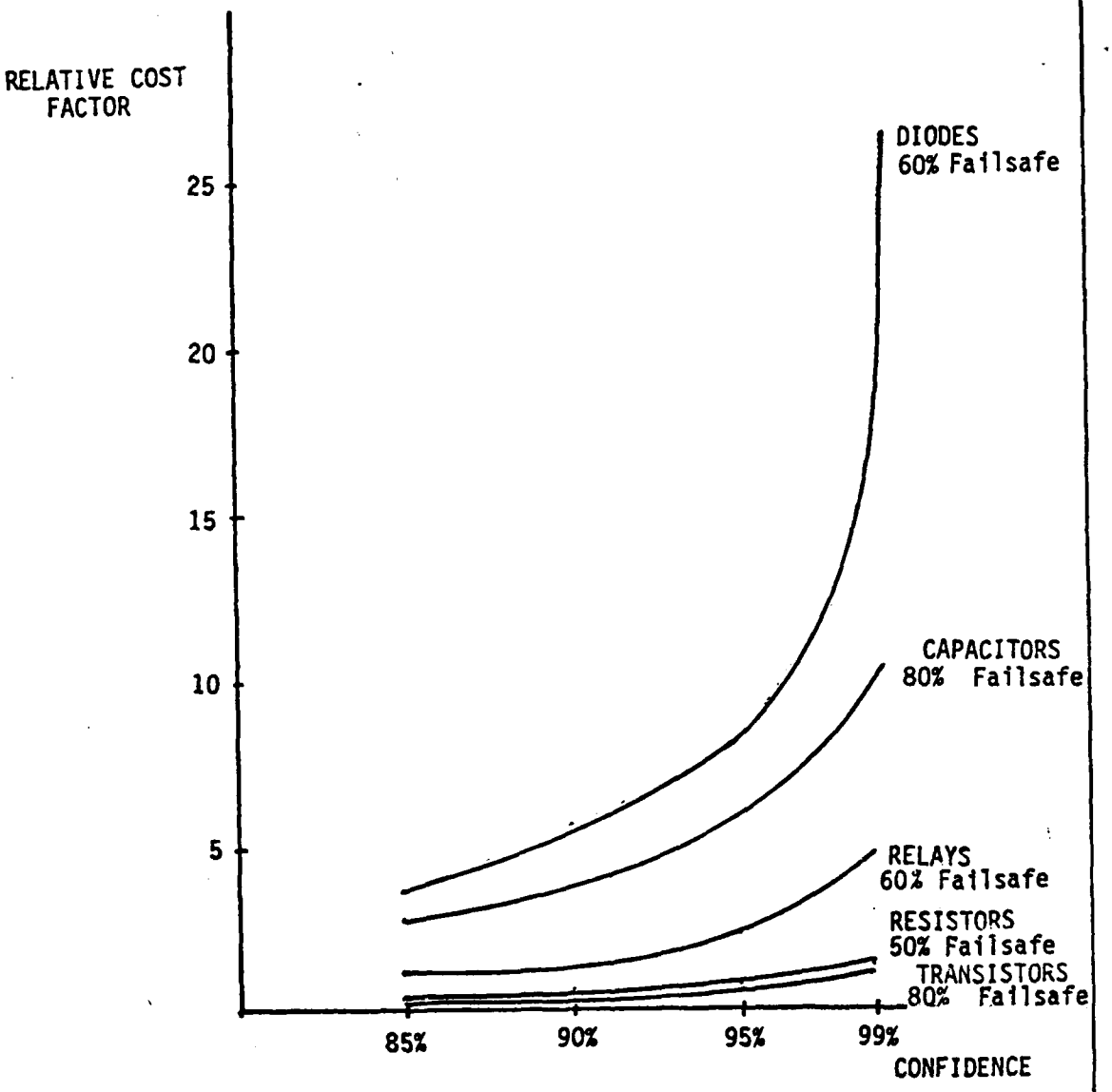
The safety reliability level selected for each component in this comparison was selected based on actual failure mode data available for each component type*, and is shown by each curve. Reliability levels are compared at 85, 90, 95, and 99% confidence levels. The "k"

*Data taken from Boeing Aerospace experience data

value corresponding to each component (as determined from experience data) was consulted on Figures 2 through 8, and required sample sizes determined. Cost of hardware and test set-up were calculated as a function of the required sample size. Test costs were determined by time for test based on failure rates, number of parts, and component cost. The data used in each case were part and application specific and for the purposes of this study were used for comparison only. Actual test costs in dollars would have to be calculated for specific tests, and compared within hardware type as discussed in Section 5.1.1.

The relative cost of confidence for various hardware types is shown in Figure 28. The greatest difference in test cost between components is due to component life and test set-up cost. Approximate costs for accelerated test sockets were used in all cases. Each component is assumed to have been tested for the reliability level which was just below their exhibited "k". The diodes, for example, have exhibited a "k" value of .3 (70% failsafe), and were assumed to be tested for 60% reliability.

The common diodes selected for the example (1N4101, $.28 @) exhibit a failure rate of 10×10^{-9}. The long life of these components and the low reliability (hence large sample size requirement) drove the test cost up. Note that the resistors also have low reliability but the associated failure rate is two orders of magnitude higher, and considerably shorter test time is assumed, thus lower test cost. The same factors influence the relative cost of test for the other components shown in the figure. Note that although the relays have very high test socket cost and component cost, their



COST/CONFIDENCE TRADE
Relative cost of confidence by component and reliability level

Figure 28

(relatively) short life drives the overall test cost to below that for capacitors and diodes. The transistors selected were 2N222A, $.15@. The failure mode split from experience is k=.10. The failure rate was 30×10^{-9}. The resistors were RC056 series, $.15@, k=.4, failure rate of $.5 \times 10^{-9}$. The capacitors selected for the example were M39003/01-2247, $1.05@, k=.1, 3×10^{-9}. The relays were by far the most expensive component: M39016/6-109L, $7.50@, k=.2. The cost of test for relays, however, was lowest because the failure rate for the relays was considerably higher than that for the other components: 150×10^{-9}.

What can be readily seen from the graph is that, at least among this selection of components, the cost driver is primarily the component life and not the component cost, anticipated reliability level, or the test set-up cost. In actual dollars, the costs involved for these test would be in the millions, even for transistor test. In order to determine specific test costs, component specific tests would have to be designed, sockets designed and built, and accelerated testing costed. The graph does, however, provide and easily visible comparison for the cost of confidence. The confidence cost accelerates most quickly for those components with long life (diodes, capacitors), and those with high test and component costs (relays). The curve is relatively flat for the medium life, low cost components (transistors, resistors) regardless of the reliability requirement.

6.0 CONCLUSIONS

This study has been performed in response to a request from the U.S. Air Force. The question to be answered, as initially posed, is this:

> "Can tests to determine safety-related failure rates of typical nuclear weapon delivery system components be performed which will result in failure rate data, with acceptable confidence levels, at reasonable cost?"

The approach selected was to conduct a parametric study which would produce a simple technique to minimize the confidence levels attainable. The technique developed is broadly applicable. Applying this technique to the components of interest, costs, cost-confidence relationships, and additional test cost minimization methods were examined.

It becomes obvious that, regardless of the component in question, failure testing for failure mode data of highly reliable components is a very expensive proposition. If the cost can be tolerated, and appropriate tests can be designed, very high confidence in the resulting data can be achieved. The question of the traceability of the data generated by accelerated testing of individual components in test sockets of questionable reliability, however, becomes a major source of uncertainty. While a great deal of data at very high confidence levels may be accumulated at great cost, there is no guarentee, nor any way of determining, the relationship of

the resulting component and system level failure predictions to the actual behavior of the weapon system in its operating environment. Although accelerated testing is widely used and well accepted, it is also well accepted that the resultant data do not necessarily accurately reflect the expected behavior of the component in actual use. This characteristic of accelerated testing would be likely to be compounded in the case of testing for failure mode data. Likewise, component failure data proliferates and is widely used in reliability predictions. It is also accepted, however, that the whole is not necessarily the sum of the parts in electronics subsystems. When these components are assembled and processed into circuit boards, for example, the processing itself, and other external influences may radically change the behavior characteristics of that subsystem.

There can be little doubt that the accumulation of failure mode data by component is valuable. When all the sources of potential inaccuracy of the data (technical level of the testing laboratory, test and test socket design, socket reliability, accelerated test characteristics) and the questionable applicability of the data to system behavior prediction are considered, and compared to the extraordinary cost to acquire the data, actual testing for component failure mode data becomes less attractive.

The most obvious alternative is the prudent but extensive use of field experience data which already exists. Reviews of documents of aerospace programs revealed vast amounts of failure rate/failure mode data which has been well documented by component type, subsystem application, and operating environment. Data sources exist which represent concentrated efforts to accumulate failure rate/failure

mode data. A recent example is the failure data available in WASH 1400 Reactor Safety Study[5]. The Military Electronics Laboratory in Stockholm, Sweden recently published a paper presenting a computerized method for analysis of data from a data bank[6]. Using the hazard plotting technique, failure rates, mode data, and confidence limits are generated. Such a system could greatly improve the flexibility and availability of the data already collected.

The second consideration is subsystem testing - emphasis on data collection at the functional assembly level. It must be noted that reliability requirements are most often allocated on this level. While costs of test and analysis would almost certainly increase, a variety of components would be tested concurrently, reducing overall test costs, and, most importantly, the data generated would be directily applicable to the weapons system safety/reliability analysis. This kind of reliability testing has become common on critical hardware and is generally referred to as AGREE or PRVT testing. It is a concept which requires verification of hardware reliability at the systems level. These alternatives are summarized in the next section.

---

[5]U.S. Nuclear Regulatory Commission, Reactor Safety Study, Appendices III and IV, October, 1975.

[6]Brobergg, Henrik, et al, Failure Rate Functions From Test Data, Military Electronics Laboratory (FTL), Stockholm, Sweden.

6.1 ALTERNATIVES

Availability of safety related failure data is critical to the analysis and design of high technology hardware such as weapons delivery systems. Various means of accumulating, analyzing and applying this type of data have been discussed throughout this study. These methods, their advantages, disadvantages and corresponding feasibility are summarized briefly below. Each is discussed in further detail in sections 5.0 and 6.0.

Testing for failure mode data:

The high cost of testing high reliability components for failure mode data has been well documented here. Special techniques such as accelerated testing and test set-ups which provide failure/mode monitoring would certainly be required. These special requirements have tremendous influence on the cost of hardware test. High levels of confidence are attainable by this method, but at great cost as shown in section 5.2. Use of the cost-minimization technique provided in section 2.0, cost/confidence trade-offs as discussed in section 5.2, and selection of the most economical component as discussed in section 5.1.1 will all serve to minimize the cost of hardware test. Only evaluation of the application and the risk involved can determine if the tremendous expense is justified.

Use of field experience data:

Years of aerospace experience and a concious effort to assemble

failure data for highly technical systems has resulted in a vast amount of available data. All techniques discussed in section 2.0 and 5.2 are applicable to the selection of components based on experience data and the minimization of the cost of verifying that data.

The confidence attainable from these data is, of course, a function of the source of the data. As emphasis on assembling highly reliable data in a central source is increased, the range and confidence of the data would also increase.

Subsystem testing:

If actual testing is to be performed for hardware data, testing at the subsystem level would be recommended. This would reduce the overall cost of testing and improve the traceability of the failure data to the system application. The system reliability requirements themselves are most often expressed on the functional subsystem level.

Cost/confidence relationships for subsystems would be largely dependent upon the components involved. Models could easily be devised which would minimize the cost of assembly testing through careful subassembly definition and optimum component selection. The unit cost model in section 5.1.1 would be very appropriate to this application.

This study has concluded that although testing for failure mode data is feasible at high confidence, the high cost associated with it renders it impractical except in those applications where high risk

justifies the expense. Other alternatives are available, however, and have been outlined here. In an effort to minimize those costs of testing, a parametric study was conducted which resulted in the sample size/confidence/reliability relationships shown in figures 2 through 27. Although this method is applicable to all types of hardware for all applications, it offers the greatest potential for cost savings in high cost, high reliability applications such as nuclear weapons systems.

BIBLIOGRAPHY

1. Ball Brothers Research Corp., Preferred Parts Handbook for Space Systems, Rev. B, February 1971, pg. A-7.

2. Brobergg, Henrik, et al, Failure Rate Functions From Test Data, Military Electronics Laboratory (FTL), Stockholm, Sweden.

3. Goddard Space Flight Center, Failure Analysis Section, Evaluation of GSFC Testing, Relative Quality Levels, and Cost Effective Selection of JANTX, JAN, and Commercial Grade Transistors, Parts and Components Evaluation Report #09-012, May 1975.

4. Pearson, E. S., C. J. Clopper, "The Use of Confidence or Fudicial Limits Illustrated in the Case of the Binomial", Biometrika Vol. 26, 1934, pp 404-13.

5. U.S. Nuclear Regulatory Commission, Reactor Safety Study, Appendices III and IV, October, 1975.

DISTRIBUTION

AUL (LDE)
DTIC (DDA)
AFWL
  (SUL)
  (HO)
  (NTSSA)
  (NTYC)
  (NTSW)
  (NTSA)
  (NTS)
AFSC (DLWM)
University of Washington, Seattle
Sandia National Laboratories, Albuquerque
Det 1, AFISC/DNS (SNA), Kirtland
Off Record Cpy (AFWL/NTSSM/Maj Alles)