1969

Statistical inferences for the Cauchy distribution based on maximum likelihood estimators

Gerald Nicholas Haas

Department: Mathematics and Statistics

STATISTICAL INFERENCES FOR THE CAUCHY DISTRIBUTION

BASED ON MAXIMUM LIKELIHOOD ESTIMATORS

by

GERALD NICHOLAS HAAS

A DISSERTATION

Presented to the Faculty of the Graduate School of the

UNIVERSITY OF MISSOURI - ROLLA

In Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

in

MATHEMATICS

1969

Charles E. Antle
Advisor

B E Gillett

Lee J. Bain

R E Douglas Jones

Henry J Hauser Jr

W A Thompson Jr

ABSTRACT

Various estimators of the location and scale parameters in the Cauchy distribution are investigated, and the superiority of the maximum likelihood estimators is established. Tables based on maximum likelihood estimators are presented for use in making statistical inferences for the Cauchy distribution. Those areas considered include confidence intervals, tests of hypothesis, power of the tests, and tolerance intervals. Both one- and two-sample problems are considered. Tables for testing the hypothesis of whether a sample came from a normal distribution or a Cauchy distribution are presented. The problems encountered in finding maximum likelihood estimators for the Cauchy parameters are discussed, and a computer program for obtaining the estimates is included.

ACKNOWLEDGEMENT

The author wishes to express his sincere appreciation to Dr. Charles E. Antle and Dr. Lee J. Bain of the Department of Mathematics for their help in the selection of this thesis subject and for their assistance and advice.

Also, the author would like to thank the staff of the University of Missouri at Rolla Computer Center for their assistance, the use of their facilities and computer time made available which was so vital to this study.

Finally, deepest gratitude is expressed to my wife and son for their patience and understanding during these years of study and to my wife for diligently typing this final copy. In all nontechnical aspects, this paper is truly a joint effort.

TABLE OF CONTENTS

LIST OF TABLES

TABLE | Page

LIST OF FIGURES

I INTRODUCTION

The Cauchy distribution $f(x;m,b) = \dfrac{1}{\pi b[1 + (\frac{x-m}{b})^2]}$, $-\infty < x < \infty$, $-\infty < m < \infty$, $b > 0$ has long appeared in texts on mathematical statistics [1, 2, 3]. Frequently it appears as an example of a distribution whose moments do not exist. The Cauchy distribution is symmetric about its location parameter, m, and is similar in appearance to the familiar normal distribution. The Cauchy distribution, however, has a larger area in the tails of the distribution. The cumulative distribution for the standard Cauchy, $m = 0$, $b = 1$, appears in Table A1. The difference in the tails of the distribution can be observed by comparison with the cumulative of the standard normal distribution.

There are a number of situations in which the Cauchy distribution arises as the appropriate probability model. Feller [4] discusses situations in the study of Brownian motion where the Cauchy distribution arises. Hald [5] points out that the tangent of an angle having a uniform distribution has a Cauchy distribution. Hald then gives an example of a physical situation where this could occur. It is not difficult to envision other physical situations of this sort. Consider, for example, a scattering of particles occurring at some height, h, above a flat surface. Let the origin of the coordinate system lie on the surface directly below the point of scattering. If we assume that the particles move in straight lines and are equally likely to scatter in any downward direction, then the marginal distributions of the x and y coordinates of impact are Cauchy distributions with $m = 0$, $b = h$.

The Cauchy distribution also arises as the ratio of two independent normally distributed random variables with zero means [6]. Thus, consider two random signals being received by a communications system where the recorded observation is the ratio of the two inputs. In the absence of signals, the noise will have a Cauchy distribution under the assumption that the individual errors have independent normal distributions with zero means. Therefore, the Cauchy distribution would arise in testing for the presence or absence of signals.

The Cauchy distribution should also be considered by the applied statistician as a possible error model. The normal distribution is used almost exclusively as an error model due to its ease of application and because of certain theoretical arguments derived from the Central Limit Theorem. It is questionable, however, whether the normal distribution is adequate for all experimental situations. In particular, if, in a given set of observations, one or more observations differ considerably from the remainder, then the assumption of normality is suspect. Dixon and Massey [7] consider this question of normality and suggest the Windsor test to determine whether or not an outlying observation should be discarded from the sample. While this procedure has the advantage of retaining the normal distribution for those observations not discarded, it would seem difficult in given samples, particularly for small sample sizes, to determine which set of observations to discard and which to retain. In other words, in certain samples it may be difficult to decide what constitutes the outliers and what constitutes the representative sample.

Dixon and Massey point out that the sample average, the minimum variance unbiased estimator of the mean of a normal population, may be greatly influenced by these outliers. An alternate procedure to discarding the outliers is to use as an error model a probability distribution giving greater probability to extreme observations. With such a model, it would be hoped that the estimates of the parameters would be more properly influenced by the extremes. The Cauchy distribution seems to fit the

the recorded observation is the ratio of the two inputs. In the absence of signals, the noise will have a Cauchy distribution under the assumption that the individual errors have independent normal distributions with zero means. Therefore, the Cauchy distribution would arise in testing for the presence or absence of signals.

The Cauchy distribution should also be considered by the applied statistician as a possible error model. The normal distribution is used almost exclusively as an error model due to its ease of application and because of certain theoretical arguments derived from the Central Limit Theorem. It is questionable, however, whether the normal distribution is adequate for all experimental situations. In particular, if, in a given set of observations, one or more observations differ considerably from the remainder, then the assumption of normality is suspect. Dixon and Massey [7] consider this question of normality and suggest the Windsor test to determine whether or not an outlying observation should be discarded from the sample. While this procedure has the advantage of retaining the normal distribution for those observations not discarded, it would seem difficult in given samples, particularly for small sample sizes, to determine which set of observations to discard and which to retain. In other words, in certain samples it may be difficult to decide what constitutes the outliers and what constitutes the representative sample.

Dixon and Massey point out that the sample average, the minimum variance unbiased estimator of the mean of a normal population, may be greatly influenced by these outliers. An alternate procedure to discarding the outliers is to use as an error model a probability distribution giving greater probability to extreme observations. With such a model, it would be hoped that the estimates of the parameters would be more properly influenced by the extremes. The Cauchy distribution seems to fit the desired category. Later, a comparison of estimators using the normal and Cauchy distributions will be given. Therefore, in addition to arising in certain physical situations, the Cauchy distribution might be considered as a possible error model for experimental observations. In fact, the Cauchy distribution has been considered as a noise model in certain signal detection problems arising in communication theory [8].

When working with the Cauchy distribution, the statistician has encountered difficulties in obtaining satisfactory estimates of the unknown parameters, m and b. The problem of the moments not existing has already been mentioned. In addition, it is known that the sample average is not a consistent estimator of m, there are no functions of the sample sufficient for m or b (other than the entire sample itself), and the maximum likelihood estimators cannot be obtained in closed form [9].

The objective of this paper is to present methods of statistical analysis when the assumed probability model is the Cauchy distribution. Various point estimators of the parameters are investigated and compared. Means of setting confidence intervals on the parameters are developed along with tests of hypothesis on the parameters. Tolerance intervals for the Cauchy distribution are determined. One- and two-sample problems are considered. Methods of discriminating between Cauchy and normal samples are also investigated. It is hoped that with the development of these procedures, the Cauchy distribution will now be considered as a possible model by the applied statistician.

II REVIEW OF THE LITERATURE

Due to the difficulties mentioned in Chapter I, most of the work dealing with the Cauchy distribution has been limited to point estimators of the location parameter. In 1947, Bhattacharyya [10] determined the Cramer-Rao lower bound for unbiased estimators of m and b. The Cramer-Rao bound is $2b^2/n$ for estimators of both parameters.

One of the earliest used estimators of m was the sample median. The median is known to be a consistent estimator of m with large sample variance equal to $\pi^2 b^2/4n$ [9]. Therefore, the large sample efficiency of the median is $8/\pi^2 = .81$. Rider [11] investigated the variance of the median of a Cauchy sample for small to moderate sample sizes. For sample sizes n = 1 to n = 4 the variance of the median is not defined. The variance of the median is tabulated for n = 5(2)31 with efficiencies ranging from .3275 for n = 5 to .74 for n = 31.

In 1964, Rothenberg, Fisher and Tilanus [12] proposed an estimator which offered an improvement in efficiency over the sample median. They considered an arithmetic average of a central subset of the sample order statistics. They determined that, roughly, the mean of the central 24 per cent of a Cauchy sample gives minimum asymptotic variance. The asymptotic variance of the Rothenberg et al estimator is $2.28b^2/n$ with efficiency of .88.

Bloch [13] presents a linear combination of 5 order statistics from a Cauchy sample as an estimator of m. He considers the order statistics $x_{(r)}$ (r = .13n, .4n, .5n, .6n, .87n) with weights -.052, .3485, .407, .3485, -.052. The asymptotic variance of this estimator is $2.102b^2/n$ with efficiency of .952.

It is well known [9] that under certain regularity conditions the variance of maximum likelihood estimators asymptotically attain the

Cramer-Rao lower bound. Barnett [14] investigates the maximum likelihood estimator of the location parameter in the Cauchy distribution for sample sizes 5 to 20. He obtains, by Monte Carlo simulation of $\hat{m}$ with b known, the small sample efficiencies of $\hat{m}$. The efficiencies range from 42.11% for n = 5 to 90.44% for n = 20. It is interesting to note that the efficiency of the median compared to the MLE remains at approximately 80% for all sample sizes, which is nearly the asymptotic efficiency. Barnett also points out the difficulties in maximum likelihood estimation due to the presence of multiple roots of the likelihood equation.

Barnett [15] determines the coefficients required to obtain the BLUE of the Cauchy location parameter (based on order statistics) for sample sizes n = 5(1)16, 18, 20. It is known that the BLUE is asymptotically efficient [16], and Barnett determines the small sample efficiency. The efficiency of the BLUE ranges from 32.75% for n = 5 to 79.56% for n = 20. The efficiency of the BLUE compared to the MLE ranges from 77.8% to 88.0% for the same sample sizes. Barnett also determines the small sample efficiency of the Rothenberg et al estimator. For the same sample sizes, the Rothenberg estimator achieves, at most, a 4% increase over the median.

Antle and Bain [17] have discovered a property of maximum likelihood estimators of location and scale parameters that makes them quite useful for obtaining confidence intervals on the parameters. They observe that $\frac{\hat{m} - m}{b}$, $\frac{\hat{b}}{b}$, and $\frac{\hat{m} - m}{\hat{b}}$ are distributed independently of the parameters, (m and b being location and scale parameters respectively, while $\hat{m}$ and $\hat{b}$ are their MLE's). This property makes the maximum likelihood estimators of location and scale parameters quite suitable for Monte Carlo simulation studies if the distributions of the estimators cannot be obtained analytically. Thoman [18] observed directly that a similar property exists for the parameters in the Weibull distribution.

Kale [19,20] investigates the solution of the likelihood equation(s) where iterative processes are required to obtain the maximum likelihood estimate.

Kendall and Stuart [9] consider a test of hypothesis on the location parameter of the Cauchy distribution against a simple alternative hypothesis. To simplify the computation, they consider a sample size of 1. Outside of this one elementary example, there has been little work reported on the Cauchy distribution outside of the above-mentioned point estimators of m. The material presented in this paper on point estimation of the Cauchy scale parameter, confidence intervals on both parameters, tests of hypothesis, tolerance intervals, two-sample problems, and discrimination appears to be new.

III INFERENCES BASED ON A SINGLE SAMPLE

A. Location and Scale Parameters Unknown

1. Point Estimation of the Scale Parameter

As was pointed out in the review of the literature, there has apparently been no work reported concerning point estimation of the scale parameter in the Cauchy distribution. Some point estimators of b will be presented here. Let $x_1, \ldots, x_n$ be a sample of size n from a Cauchy distribution with parameters m and b. Let $x_{(1)}, \ldots, x_{(n)}$ be the ordered sample. Barnett [15] points out that the expected values of $x_{(1)}$ and $x_{(n)}$ do not exist. Furthermore, the variances of $x_{(1)}$, $x_{(2)}$, $x_{(n-1)}$, and $x_{(n)}$ do not exist. The following two estimators of b will be restricted to linear functions of $x_{(3)}, \ldots, x_{(n-2)}$.

It seems reasonable that a function of the restricted range $x_{(n-2)} - x_{(3)}$ might be useful as a simple point estimator of b. Consider as a possible estimator a constant times the restricted range, i.e., $b^* = k(x_{(n-2)} - x_{(3)})$. The constant, k, such that $E(b^*) = b$, can be determined as follows:

We seek k such that

$$b = E(b^*) = kE[x_{(n-2)} - x_{(3)}]$$

$$= kbE\left[\frac{x_{(n-2)} - m}{b} - \frac{x_{(3)} - m}{b}\right].$$

Due to the symmetry of the Cauchy distribution,

$$-E\left[\frac{x_{(3)} - m}{b}\right] = E\left[\frac{x_{(n-2)} - m}{b}\right]$$

Therefore, $k = \dfrac{1}{2E\left[\frac{x_{(n-2)} - m}{b}\right]}$

Let $z = \dfrac{x_{(n-2)} - m}{b}$. Now z represents the third largest ordered random variable from a Cauchy sample with m = 0, b = 1. Hence, the density of z is given by

$$g(z) = \frac{n!}{2(n-3)!}\left[\frac{1}{2}+\frac{1}{\pi}\tan^{-1}z\right]^{n-3}\left[\frac{1}{2}-\frac{1}{\pi}\tan^{-1}z\right]^{2}\frac{1}{\pi(1+z^2)},$$

and E(z) is given by

$$E(z) = \int_{-\infty}^{\infty} zg(z)dz.$$

By letting $y = \tan^{-1}(z)$, this integral becomes

$$E(z) = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{n(n-1)(n-2)}{2}\tan y\left[\frac{1}{2}+\frac{y}{\pi}\right]^{n-3}\left[\frac{1}{2}-\frac{y}{\pi}\right]^{2}\frac{\sec^2 y\, dy}{\pi[1+\tan^2 y]}$$

This integral was evaluated using Simpson's Method of numerical integration [21]. Table 1 presents the values of k for sample sizes n = 6(1)55.

Table 1

Constants k such that $E[k(x_{(n-2)} - x_{(3)})] = b$.

n	k	n	k	n	k	n	k	n	k
6	1.382	16	.212	26	.124	36	.089	46	.069
7	.790	17	.198	27	.120	37	.086	47	.067
8	.578	18	.186	28	.115	38	.084	48	.0661
9	.464	19	.175	29	.111	39	.082	49	.0647
10	.392	20	.165	30	.107	40	.080	50	.0634
11	.341	21	.156	31	.103	41	.078	51	.0621
12	.303	22	.149	32	.100	42	.076	52	.0609
13	.273	23	.142	33	.097	43	.074	53	.0598
14	.249	24	.135	34	.094	44	.072	54	.0586
15	.229	25	.130	35	.091	45	.071	55	.0576

The next point estimator of b to be presented is the BLUE of b based on the ordered observations $x_{(3)}, \ldots, x_{(n-2)}$. The BLUE of a scale parameter for a symmetric density is given by

$$b^* = \frac{z'V^{-1}x}{z'V^{-1}z} \qquad (1)$$

where z is a vector of the expected values of the standardized ordered random variables and V is the corresponding covariance matrix [16]. Barnett [15] has calculated z and V for n = 6(1)16, 18, 20. Using these data, the BLUE coefficients for the Cauchy scale parameter were calculated and appear in Table 2.

Table 2

BLUE Coefficients c_i for the Ordered Variable x_i

n	i	c_i	n	i	c_i	n	i	c_i
6	4	1.3819	13	7	0.0	16	12	.1560
7	4	0.0		8	.2529		13	.0719
	5	.7898		9	.2933		14	.0172
8	5	.5780		10	.1615	18	10	.0633
	6	.4135		11	.0427		11	.1597
9	5	0.0	14	8	.1204		12	.1940
	6	.5611		9	.2654		13	.1645
	7	.2338		10	.2397		14	.1028
10	6	.3007		11	.1231		15	.0444
	7	.4291		12	.0308		16	.0104
	8	.1417	15	8	0.0	20	11	.0489
11	6	0.0		9	.1815		12	.1273
	7	.5717		10	.2484		13	.1672
	8	.1367		11	.1940		14	.1593
	9	.1361		12	.0933		15	.1190
12	7	.1801		13	.0228		16	.0693
	8	.3460	16	9	.0828		17	.0289
	9	.2258		10	.2055		18	.0067
	10	.0613		11	.2213			

Due to symmetry in the coefficients, only half of the coefficients appear in Table 2. Those not appearing are the negative of those presented. As an example, the BLUE of b for n = 9 would be

$$b^* = -.2338x_{(3)} - .5611x_{(4)} + .5611x_{(6)} + .2338x_{(7)}.$$

The third point estimator of b considered is the maximum likelihood estimator (MLE). The MLE's $\hat{m}$, $\hat{b}$ of m, b respectively are those functions of the observations such that

$L(x;\hat{m},\hat{b}) = \max L(x;m,b)$ (where $L(x;m,b)$ is the likelihood function) for all m,b such that $-\infty < m < \infty$, $b > 0$.

A discussion of the numerical solution of the likelihood equations is given in Chapter VII. A comparison of these three point estimators of b is given in part E of this chapter.

2. Confidence Intervals

As has already been pointed out, the distributions of $\frac{\hat{m}-m}{b}$, $\frac{\hat{b}}{b}$, and $\frac{\hat{m}-m}{\hat{b}}$ do not depend upon the parameters. The maximum likelihood estimators of the Cauchy parameters cannot be obtained in closed form; hence, the distributions of the above pivotal functions were obtained by a Monte Carlo simulation. A discussion of the difficulties in obtaining the MLE's will be presented in Chapter VII. It is assumed here that the MLE's are obtainable. The following theorem proved useful in obtaining the distributions of $\frac{\hat{m}-m}{b}$ and $\frac{\hat{m}-m}{\hat{b}}$.

THEOREM 1

Let $X \sim \frac{1}{b}f(\frac{x-m}{b})$ with the density of X symmetric about m. Let $\hat{m}$ and $\hat{b}$ be the maximum likelihood estimators of m and b. Then $\hat{m}$ is an unbiased estimator of m, and $\frac{\hat{m}-m}{b}$ and $\frac{\hat{m}-m}{\hat{b}}$ are distributed symmetrically about zero.

Proof:

Let $x_1, \ldots, x_n$ be a random sample from $\frac{1}{b} f(\frac{x-m}{b})$. Let $\hat{m}_x, \hat{b}_x$ be the MLE's based on this set of observations. Now

$$\prod_{i=1}^{n} \frac{1}{\hat{b}_x} f\left(\frac{x_i - \hat{m}_x}{\hat{b}_x} \right) \geq \prod_{i=1}^{n} \frac{1}{b^*} f\left(\frac{x_i - m^*}{b^*} \right)$$

for all m^* and b^* in the parameter space $\Omega = \{m,b \mid -\infty < m < \infty, \quad b > 0\}$. Let $z_i = \frac{x_i-m}{b}$, $i = 1, \ldots n$. Substituting for x_i yields

$$\prod_{i=1}^{n} \frac{1}{\hat{b}_x} f\left(\frac{bz_i+m-\hat{m}_x}{\hat{b}_x}\right) \geq \prod_{i=1}^{n} \frac{1}{b^*} f\left(\frac{bz_i+ m-m^*}{b^*} \right). \tag{2}$$

Let $y_1, y_2, \ldots y_n$ be the reflections of $x_1, \ldots x_n$ about m; i.e.,

$y_i = m - bz_i$, $i = 1, \ldots, n$. Let $\hat{m}_y$ and $\hat{b}_y$ be the MLE's of m and b based on the sample $y_1, \ldots, y_n$. Now

$$\prod_{i=1}^{n} \frac{1}{\hat{b}_y} f\left(\frac{y_i - \hat{m}_y}{\hat{b}_y} \right) \geq \prod_{i=1}^{n} \frac{1}{b^*} f\left(\frac{y_i - m^*}{b^*} \right)$$

for all m^* and b^* in Ω. Substituting for y_i in terms of z_i yields

$$\prod_{i=1}^{n} \frac{1}{\hat{b}_y} f\left(\frac{-bz_i + m - \hat{m}_y}{\hat{b}_y} \right) \geq \prod_{i=1}^{n} \frac{1}{b^*} f\left(\frac{-bz_i + m - m^*}{b^*} \right).$$

But since f is symmetric in its argument,

$$\prod_{i=1}^{n} \frac{1}{\hat{b}_y} f\left(\frac{bz_i - m + \hat{m}_y}{\hat{b}_y} \right) \geq \prod_{i=1}^{n} \frac{1}{b^*} f\left(\frac{bz_i - m + m^*}{b^*} \right). \quad (3)$$

By comparing (2) with (3), we observe that by letting $\hat{b}_y = \hat{b}_x$ and $\hat{m}_y = 2m - \hat{m}_x$, we maximize the likelihood function for the sample of y_i. Hence, the symmetry of the density of X carries over to the density of $\hat{m}$. Therefore, $\frac{\hat{m}-m}{b}$, $\frac{\hat{m}-m}{\hat{b}}$ are distributed symmetrically about 0, and $E(\hat{m}) = m$.

Q.E.D.

Due to this theorem, the simulated distribution of $\frac{\hat{m}-m}{\hat{b}}$ was forced to be symmetric about zero by averaging the number of estimates lying in cells equidistant from zero. The simulated distribution of $\frac{\hat{m} - m}{\hat{b}/\sqrt{n}}$ appears in Table A2. The simulated distribution of $\frac{\hat{b}}{b}$ appears in Table A3.

Confidence intervals on m with b unknown may be obtained as follows: From Table A2, for sample size n, one can obtain the percentage point k such that

$$P\left[-k \leq \frac{\hat{m} - m}{\hat{b}/\sqrt{n}} \leq k\right] = 1 - \alpha$$

so that

$$P\left[\hat{m} - \frac{k\hat{b}}{\sqrt{n}} \leq m \leq \hat{m} + \frac{k\hat{b}}{\sqrt{n}}\right] = 1 - \alpha.$$

Therefore, $\left[\hat{m} - \frac{k\hat{b}}{\sqrt{n}}, \hat{m} + \frac{k\hat{b}}{\sqrt{n}}\right]$ forms a $1-\alpha$ confidence interval for m.

Confidence intervals on b with m unknown may be obtained as follows: From Table A3 one can obtain the $\frac{\alpha}{2}$, $1-\frac{\alpha}{2}$ percentage points $k_{\frac{\alpha}{2}}$, $k_{1-\frac{\alpha}{2}}$ such that

$$P[k_{\frac{\alpha}{2}} \leq \frac{\hat{b}}{b} \leq k_{1-\frac{\alpha}{2}}] = 1 - \alpha$$

so that

$$p[\frac{\hat{b}}{k_{1-\frac{\alpha}{2}}} \leq b \leq \frac{\hat{b}}{k_{\frac{\alpha}{2}}}] = 1 - \alpha.$$

Therefore, $[\frac{\hat{b}}{k_{1-\frac{\alpha}{2}}}, \frac{\hat{b}}{k_{\frac{\alpha}{2}}}]$ forms a $1 - \alpha$ confidence interval for b. An example is given in part F of this chapter.

It is also possible to obtain confidence intervals for m and b based on the BLUE's. It would be extremely difficult to analytically obtain the distributions of the BLUE's; however, the following theorem makes the Monte Carlo simulation of the distributions feasible.

THEOREM 2

Let $x_1, \ldots, x_n$ be an independent set of values of a random variable $X \sim \frac{1}{b} f(\frac{x-m}{b})$. Let $m^* = \sum_{i=1}^{n} \alpha_i x_i$ and $b^* = \sum_{i=1}^{n} \lambda_i x_i$ be estimators of m and b. If $\sum_{i=1}^{n} \alpha_i = 1$ and $\sum_{i=1}^{n} \lambda_i = 0$, then $\frac{m^*-m}{b}$, $\frac{b^*}{b}$, and $\frac{m^*-m}{b^*}$ are distributed independently of the parameters.

Proof:

Let $y_1, \ldots, y_n$ be an independent set of values of a random variable $Y \sim f(y)$; i.e., $m = 0$, $b = 1$ so that $x_i = by_i + m$. Let

$$m_o^* = \sum_{i=1}^{n} \alpha_i y_i \text{ and } b_1^* = \sum_{i=1}^{n} \lambda_i y_i, \; i = 1, \ldots, n.$$

$$\text{Now } m^* = \sum_{i=1}^{n} \alpha_i x_i = \sum_{i=1}^{n} \alpha_i (by_i + m)$$

$$= b \sum_{i=1}^{n} \alpha_i y_i + m$$

$$= bm_o^* + m$$

so that $\frac{m^*-m}{b} = m_o^*$

Also, $b^* = \sum_{i=1}^{n} \lambda_i x_i$

$= \sum_{i=1}^{n} \lambda_i (by_i + m)$

$= b \sum_{i=1}^{n} \lambda_i y_i$

$= b\ b_1^*$

so that $\frac{b^*}{b} = b_1^*$

But the distributions of m_o^* and b_1^* do not depend on any unknown parameters. Therefore, the distributions of $\frac{m^*-m}{b}$ and $\frac{b^*}{b}$ do not depend on any unknown parameters; and, hence, the distribution of $\frac{m^*-m}{b^*}$ does not depend on any unknown parameters.

Q.E.D.

It is immediately obvious that the median and $k[x_{(n-2)}-x_{(3)}]$ are suitable random variables for setting confidence intervals on m and b. Kendall and Stuart [9] point out that for the BLUE of m and b in a density symmetric about m, $\sum_{i=1}^{n} \alpha_i = 1$, and $\sum_{i=1}^{n} \lambda_i = 0$. Thus, the BLUE's also come under Theorem 2. The simulated distributions of $\frac{m^*-m}{b^*/\sqrt{n}}$ and $\frac{b^*}{b}$ using order statistics were obtained for selected sample sizes. These tables appear in section E of this chapter. The superiority of the MLE's in setting confidence intervals is demonstrated in that section.

3. Tests of Hypotheses

Procedures will be developed to test hypotheses on m and b. The tests will be in terms of the MLE's, and the power of the tests will be determined.

Consider the following test of hypothesis:

$H_o: m = m_o \qquad H_1: m > m_o.$

Under the null hypothesis, the critical value c for a significance level α is given in Table A2 such that

$$P\left[\frac{\hat{m}-m_o}{\hat{b}/\sqrt{n}} \le c \right] = 1 - \alpha.$$

The null hypothesis is then rejected if $\frac{\hat{m}-m_o}{\hat{b}/\sqrt{n}} > c$. The power of the test is given by $P\left[\frac{\hat{m}-m_o}{\hat{b}/\sqrt{n}} > c \mid m = m_1 \right]$ where $m_1 > m_o$.

The following theorem is useful in determining the power of the test:

<u>THEOREM 3</u>

Let $X \sim \frac{1}{b} f(\frac{x-m}{b})$, and let m* and b* be estimators of m and b. If the critical region of the test $H_o : m = m_o$, $H_1 : m > m_o$ is given by (c,∞) where $P\left[\frac{\hat{m}-m_o}{\hat{b}/\sqrt{n}} \le c \right] = 1 - \alpha$ under H_o, then the power of the test is obtained from the distribution of $\frac{m^*-m_1}{b} - \frac{c}{\sqrt{n}} \frac{b^*}{b}$ where m_1 is the true value of m.

Proof:

$$\text{Power} = P\left[\frac{m^* - m_o}{b^*/\sqrt{n}} > c \mid m = m_1 \right]$$

$$= P\left[m^* - m_o > \frac{cb^*}{\sqrt{n}} \mid m = m_1\right]$$

$$= P\left[m^* - m_1 + m_1 - m_o > \frac{cb^*}{\sqrt{n}} \mid m = m_1\right]$$

$$= P\left[\frac{m^* - m_1}{b} - \frac{c}{\sqrt{n}} \frac{b^*}{b} > \frac{m_o - m_1}{b} \mid m = m_1\right].$$

Q.E.D.

Now if we use the MLE's as our estimators, then we see that the distribution of $\frac{\hat{m}-m_1}{b} - \frac{c}{\sqrt{n}}\frac{\hat{b}}{b}$ depends only on c and n. The distribution of $\frac{\hat{m}-m_1}{b} - \frac{c}{\sqrt{n}}\frac{\hat{b}}{b}$ was obtained by Monte Carlo methods for critical values, c, corresponding to Type 1 errors of .1, .05, .025, .01. The power curves appear in Figure 1a, 1b, 1c, and 1d. Due to symmetry, these power curves may also be used for $H_1 : m < m_o$.

The approximate power for the two-sided alternative may also be determined from Figure 1 for Type 1 errors of .2, .1, .05, .02. An example will be given in section F.

Consider the test of hypothesis $H_o : b = b_o$, $H_1 : b > b_o$. The critical region is of the form (c,∞) where c is determined from Table A3 such that

$$P[\frac{\hat{b}}{b_o} \le c] = 1 - \alpha$$

under the null hypothesis.

The power of the test for $b = b_1$ can also be obtained using Table A3 as follows:

$$\begin{aligned} \text{Power} &= P[\frac{\hat{b}}{b_o} > c \mid b = b_1] \\ &= P[\hat{b} > cb_o \mid b = b_1] \\ &= P[\frac{\hat{b}}{b_1} > c\frac{b_o}{b_1} \mid b = b_1]. \end{aligned}$$

The power curves for Type I errors of .1, .05, .025, .01 are given in Figures 2a, 2b, 2c, and 2d. The use of these curves is illustrated in the example in section F.

4. Asymptotic Convergence

From [9] it is seen that $\hat{m}$ and $\hat{b}$ have asymptotic normal distributions and are asymptotically efficient. Therefore, $\frac{\sqrt{n}(\hat{m}-m)}{b} \overset{\sim}{\to} N(0,2)$, and $\frac{\hat{b}}{b} \overset{\sim}{\to} N(1,\frac{2}{n})$. Since $\hat{b}$ is asymptotically unbiased and $\lim_{n\to\infty} Var(\hat{b}) = 0$, we have that $\hat{b}$ converges stochastically to b. Hence, from [22], $\frac{\sqrt{n}(\hat{m}-m)}{b}(\frac{b}{\hat{b}})$ has the same limiting distribution as $\frac{\sqrt{n}(\hat{m}-m)}{b}$. Therefore, $\frac{\sqrt{n}(\hat{m}-m)}{\hat{b}} \overset{\sim}{\to} N(0,2)$. In Table A2, the last line $(n = \infty)$ represents the

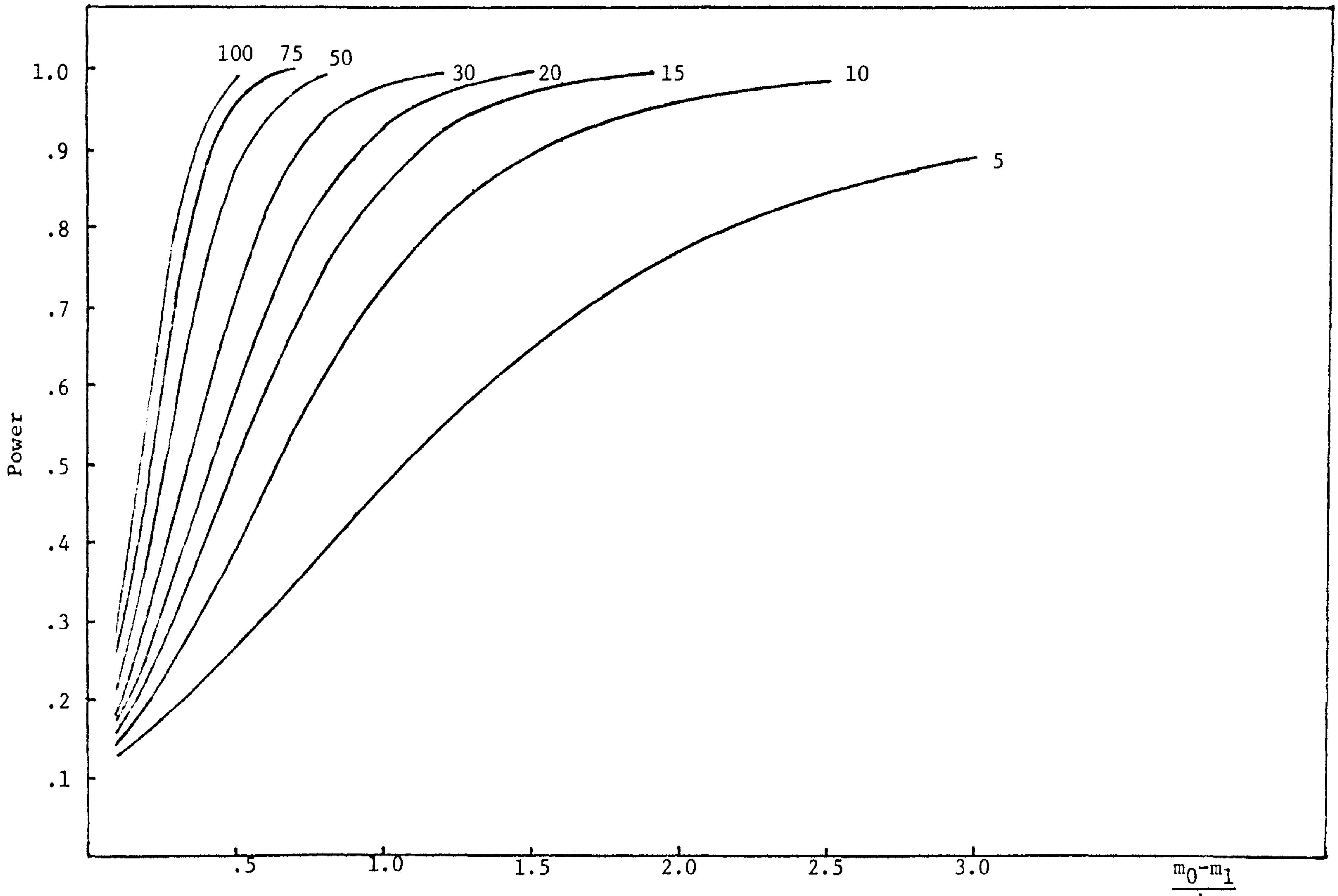

Figure 1a
Power of one-sided .10 level test of m (b unknown)

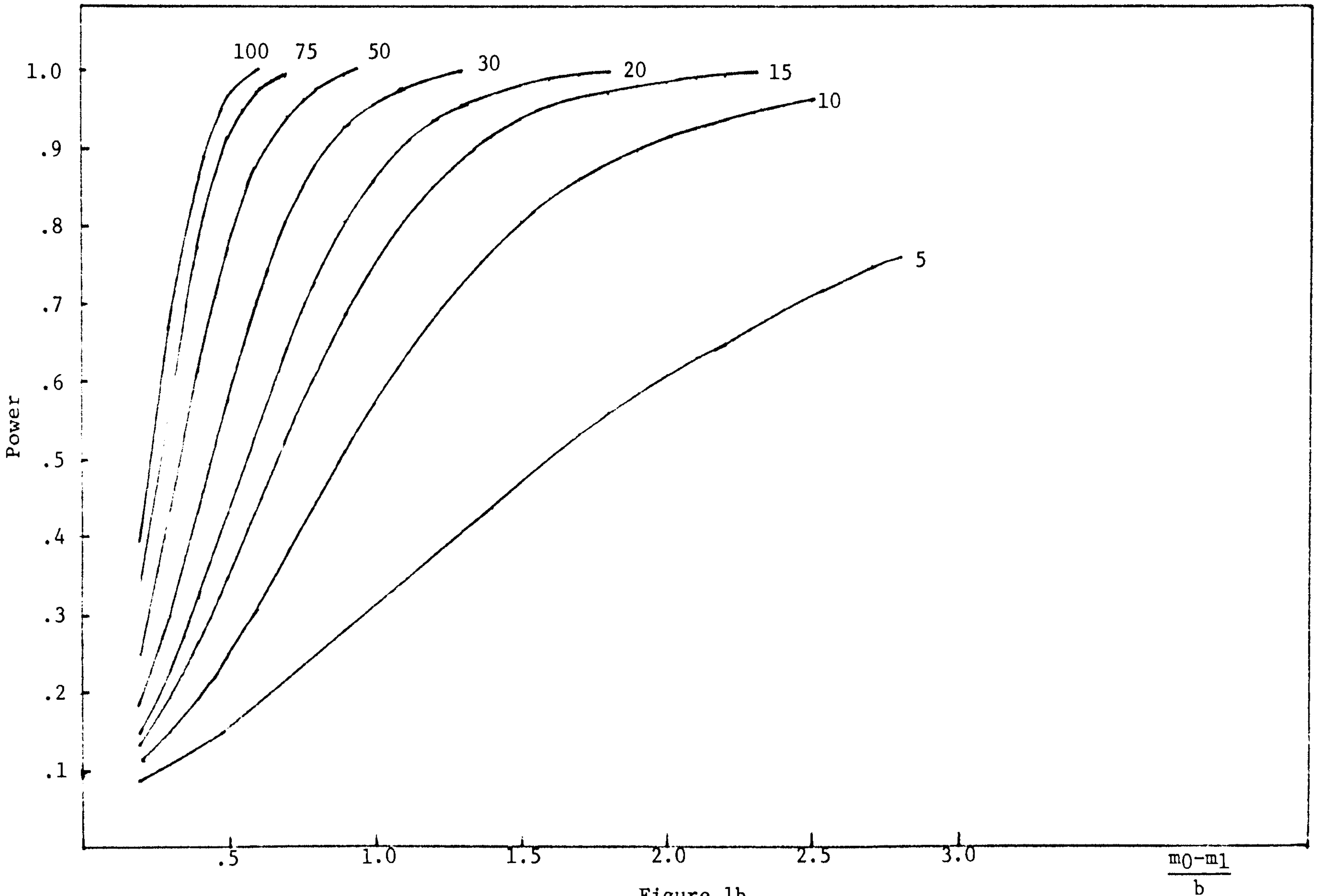

Figure 1b

Power of one-sided .05 level test of m (b unknown)

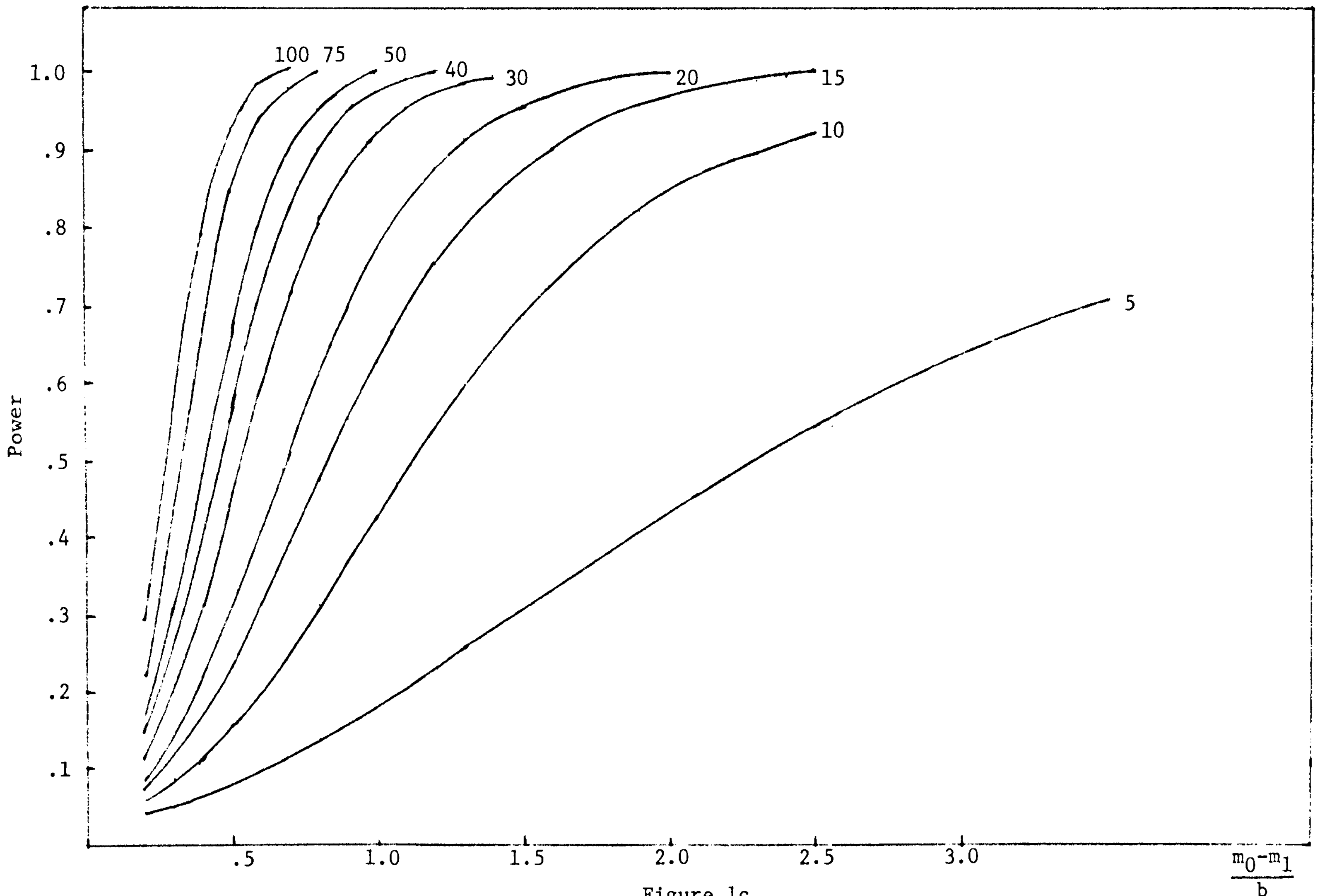

Figure 1c
Power of one-sided .025 level test of m (b unknown)

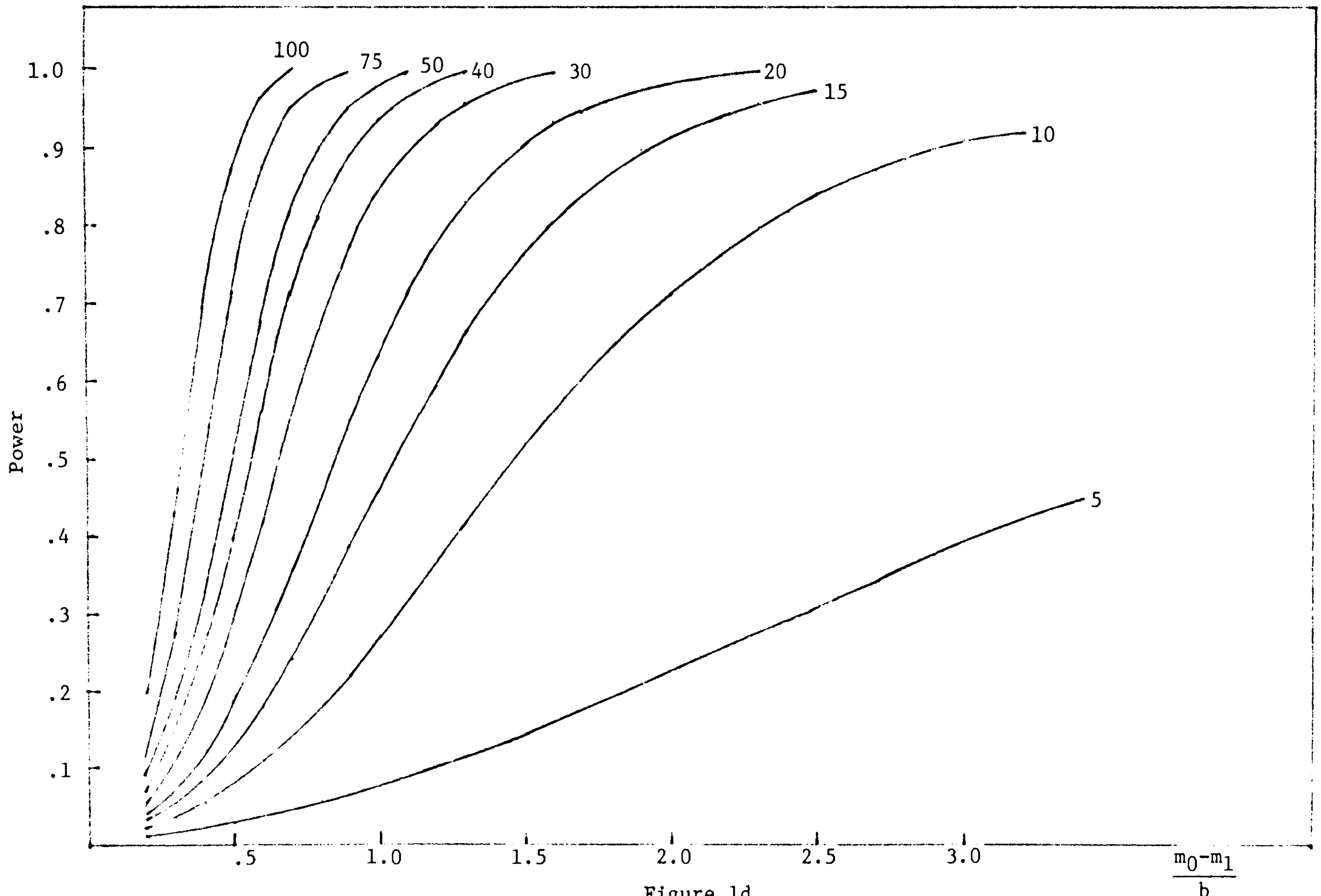

Figure 1d
Power of one-sided .01 level test of m (b unknown)

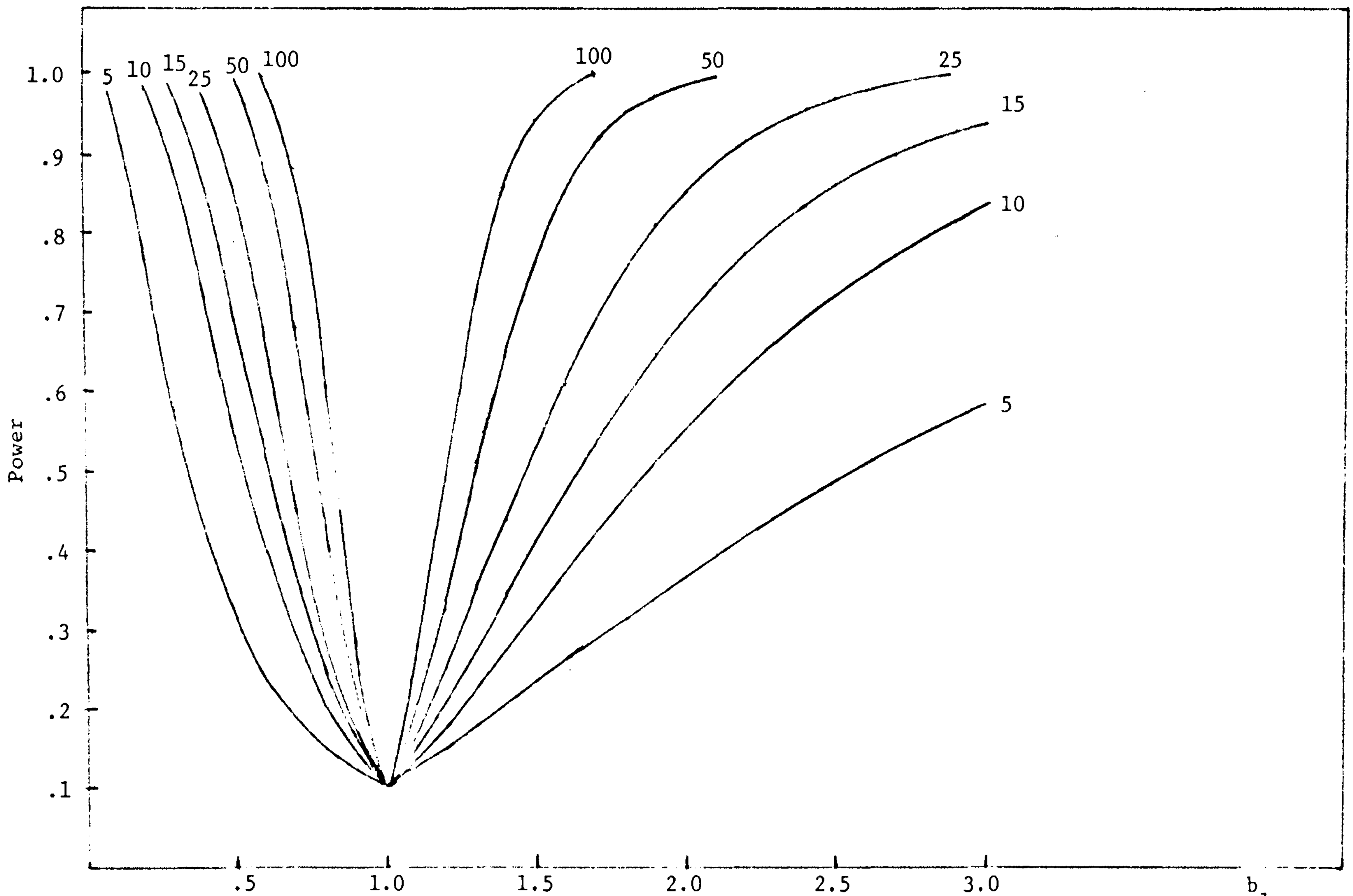

Figure 2a
Power of one-sided .10 level test on b (m unknown)

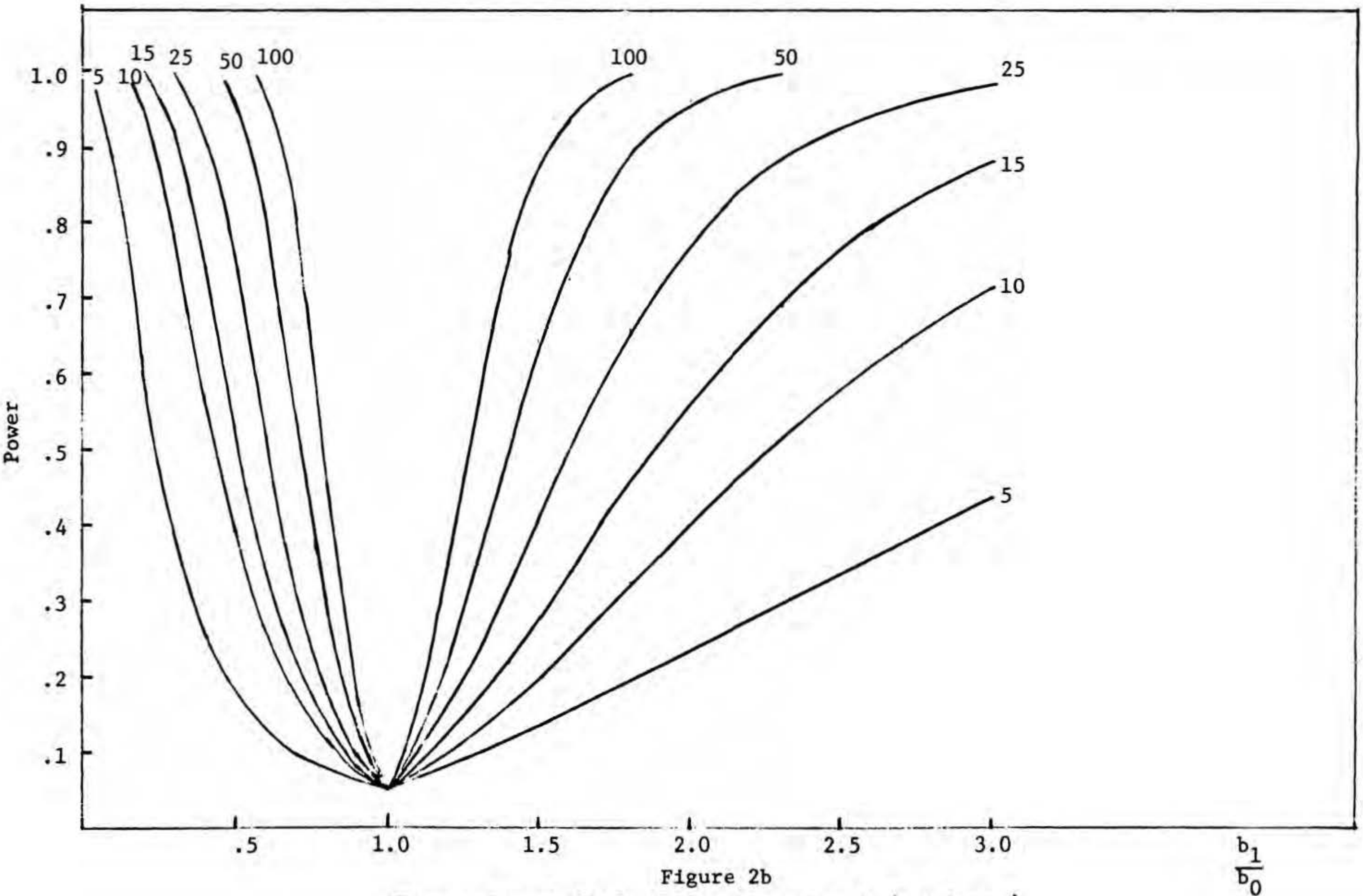

Figure 2b
Power of one-sided .05 level test on b (m unknown)

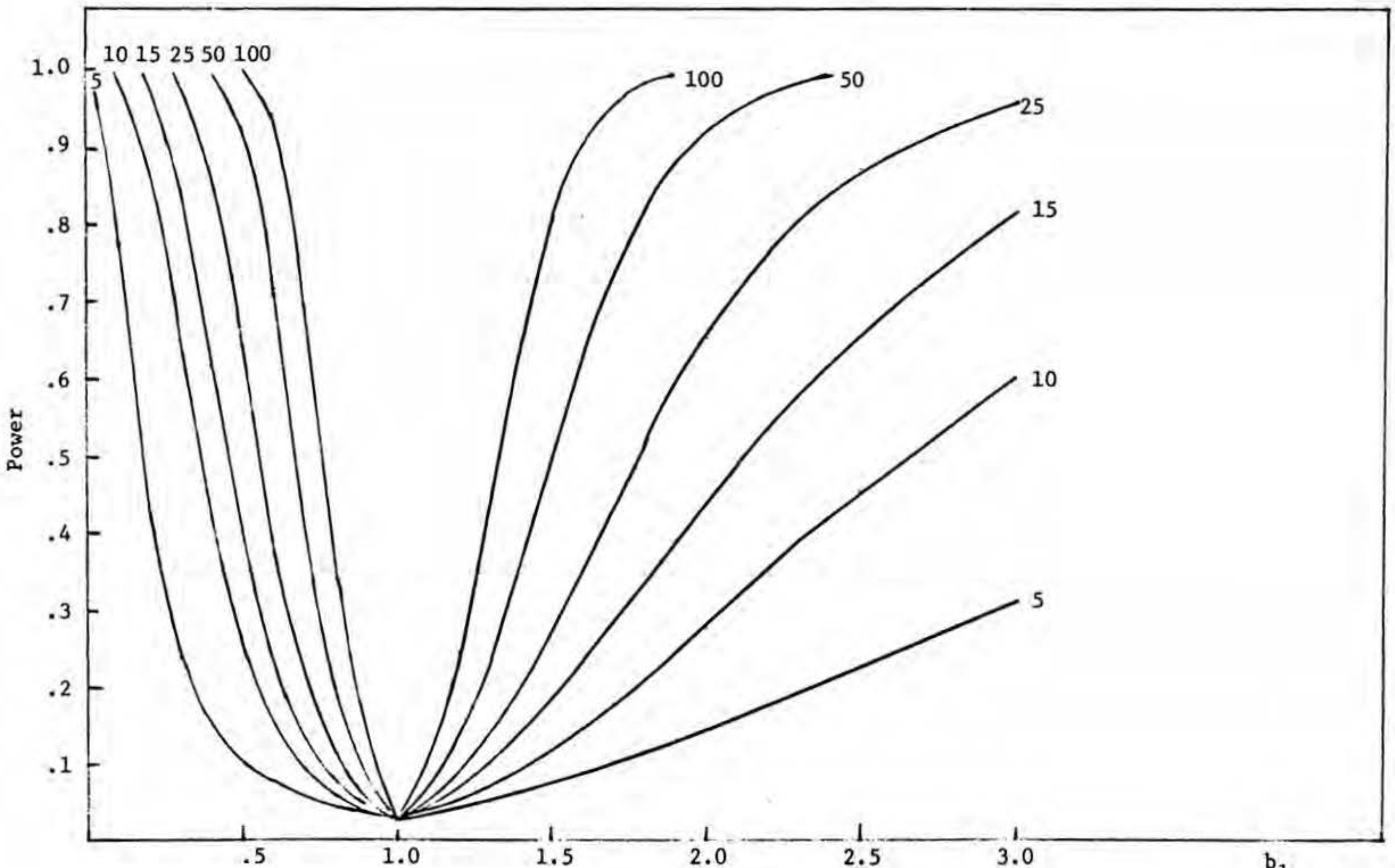

Figure 2c
Power of one-sided .025 level test on b (m unknown)

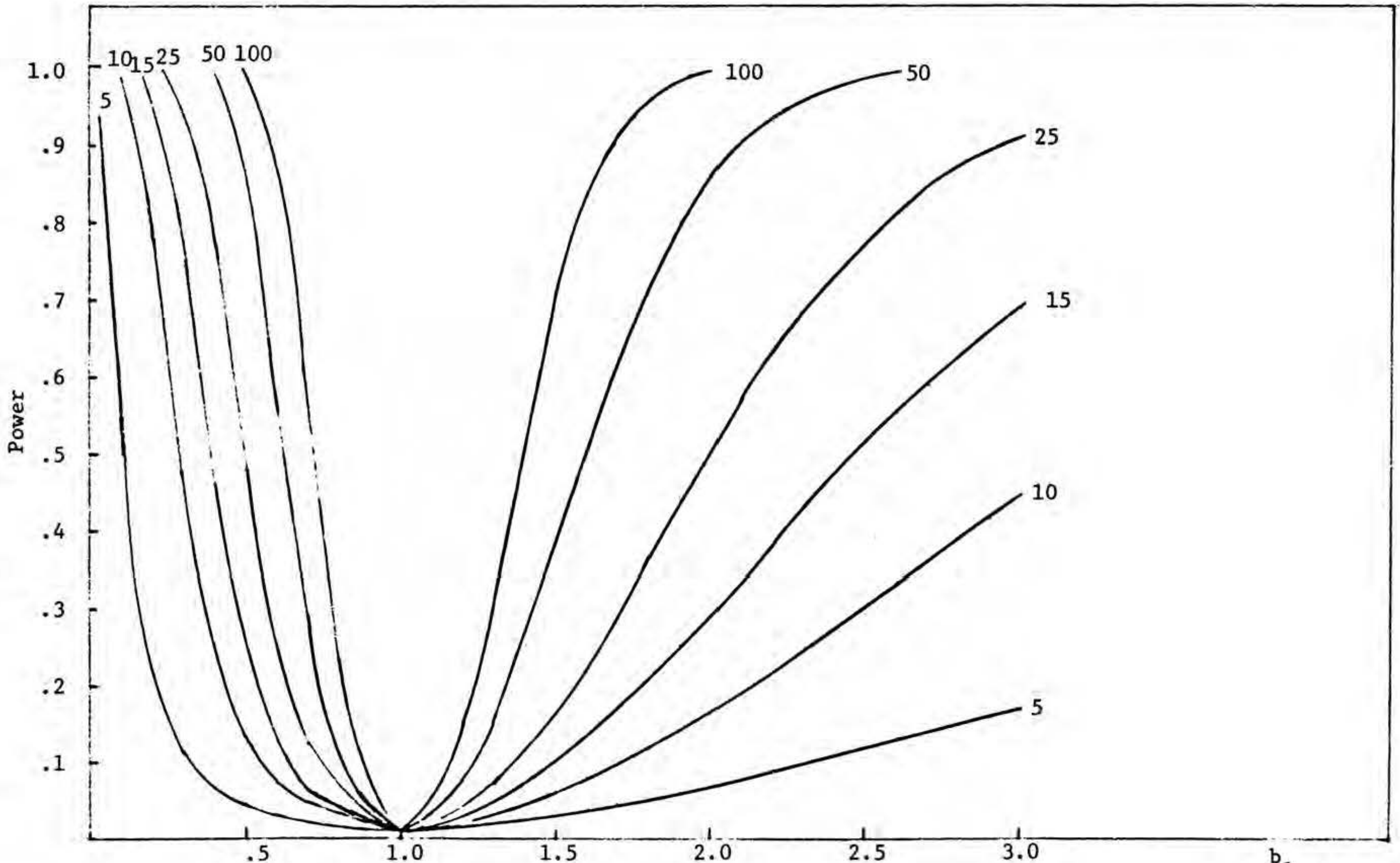

Figure 2d
Power of one-sided .01 level test on b (m unknown)

asymptotic limits as derived from the normal distribution. A comparison of the simulated distribution of $\frac{\hat{b}}{b}$ and the distribution of $\frac{\hat{b}}{b}$ using the normal approximation appears in Table A3 for n = 100.

B. Scale Parameter Known

1. Confidence Intervals

The material presented here is similar to section A, differing in the table required for critical values. From [17] we again have $\frac{\hat{m}-m}{b}$ distributed independent of m and b. Again, it will be assumed that the MLE of m can be found. The discussion of maximum likelihood estimation with b known is found in Chapter VII. The distribution of $\frac{\hat{m}-m}{b}$ was determined by a Monte Carlo simulation, and critical values of $\frac{\hat{m}-m}{b/\sqrt{n}}$ appear in Table A4. The distribution of $\frac{\hat{m}-m}{b/\sqrt{n}}$ is symmetric about zero.

From Table A4, for sample size n, one can obtain the $1-\frac{\alpha}{2}$ cumulative percentage point k such that

$$P[-k \le \frac{\hat{m} - m}{b/\sqrt{n}} \le k] = 1 - \alpha$$

so that

$$P[\hat{m} - \frac{kb}{\sqrt{n}} \le m \le \hat{m} + \frac{kb}{\sqrt{n}}] = 1 - \alpha.$$

2. Tests of Hypotheses

Consider the test of hypothesis (b known)

$$H_o : m = m_o, \; H_1 : m > m_o.$$

Under the null hypothesis, the critical value, c, such that $P[\frac{\hat{m} - m_o}{b/\sqrt{n}} \le c] = 1 - \alpha$ may be obtained from Table A4. The power of the test is given by $P[\frac{\hat{m} - m_o}{b/\sqrt{n}} > c]$ where the true value of the location parameter is m_1. The expression for the power may be easily manipulated to yield

$$\text{Power} = P[\frac{\hat{m} - m_1}{b/\sqrt{n}} \ge c - \frac{(m_1 - m_o)}{b/\sqrt{n}}].$$

Hence, the power may also be obtained from Table A4. The power curves appear in Figures 3a, 3b, 3c, and 3d.

3. Asymptotic Convergence

As has already been pointed out, $\frac{\hat{m} - m}{b/\sqrt{n}} \overset{\mathcal{D}}{\to} N(0,2)$. The last line in Table A4 represents the asymptotic values derived from the normal distribution.

C. Location Parameter Known

1. Confidence Intervals

Again $\frac{\hat{b}}{b}$ is distributed independently of b. Critical values of $\frac{\hat{b}}{b}$ were obtained by a Monte Carlo simulation and appear in Table A5. To obtain a two-sided $1 - \alpha$ confidence interval for b, we may obtain from Table A5, the critical values $k_{\frac{\alpha}{2}}$, $k_{1-\frac{\alpha}{2}}$ such that

$$P[k_{\frac{\alpha}{2}} \le \frac{\hat{b}}{b} \le k_{1-\frac{\alpha}{2}}] = 1 - \alpha$$

so that

$$P[\frac{\hat{b}}{k_{1-\frac{\alpha}{2}}} \le b \le \frac{\hat{b}}{k_{\frac{\alpha}{2}}}] = 1 - \alpha.$$

2. Tests of Hypotheses

The procedure is identical to that in III-A-3. The critical values, however, come from Table A5. The power curves appear in Figures 4a, 4b, 4c, and 4d.

3. Asymptotic Convergence

As has already been indicated, $\frac{\hat{b}}{b} \overset{\mathcal{D}}{\to} N(1,2/n)$. The last line of Table A5 gives the critical values derived from the normal distribution.

D. Tolerance Intervals

1. Both Parameters Unknown

Let X be a random variable with cumulative distribution function F(x).

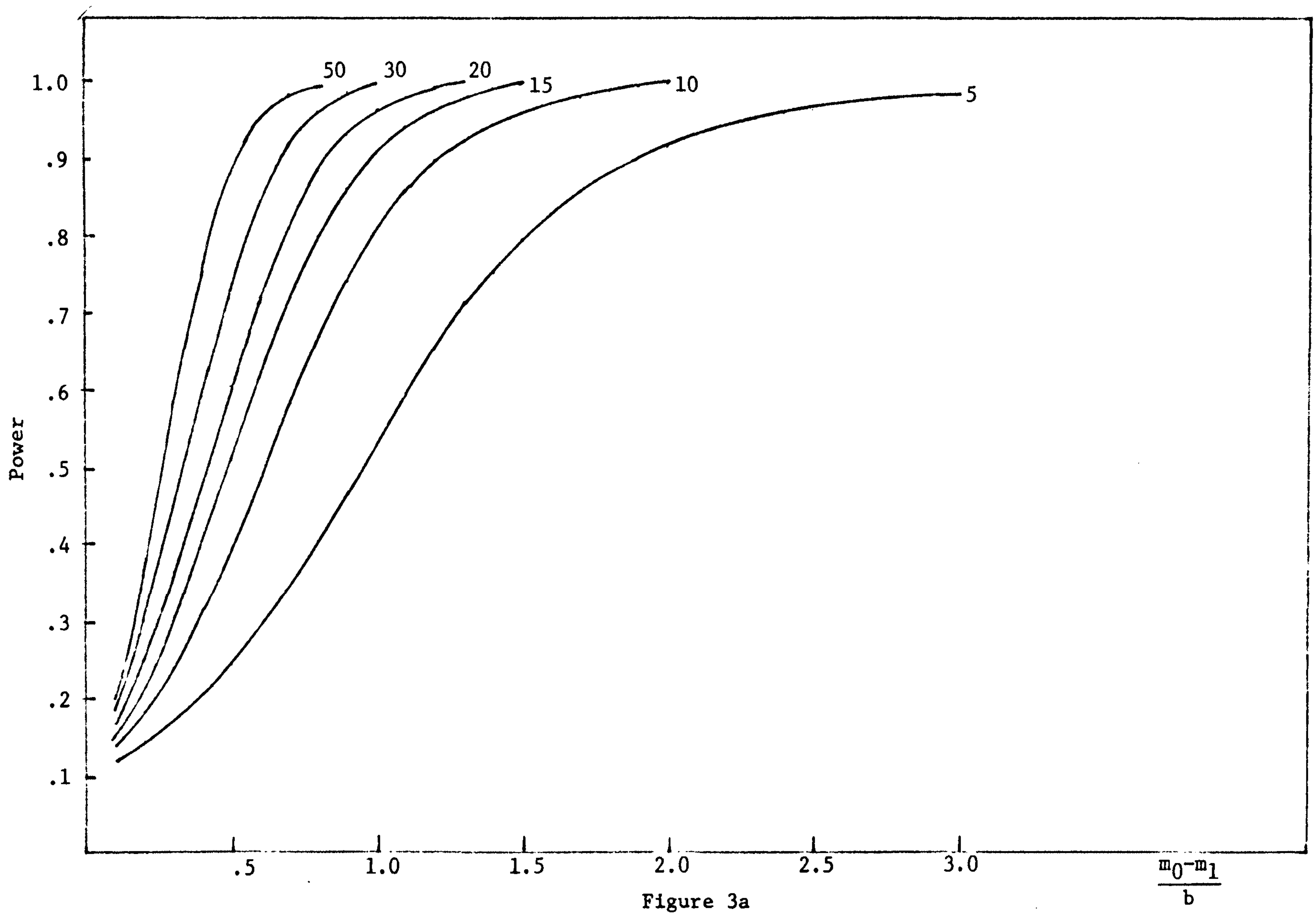

Figure 3a
Power of one-sided .10 level test of m (b known)

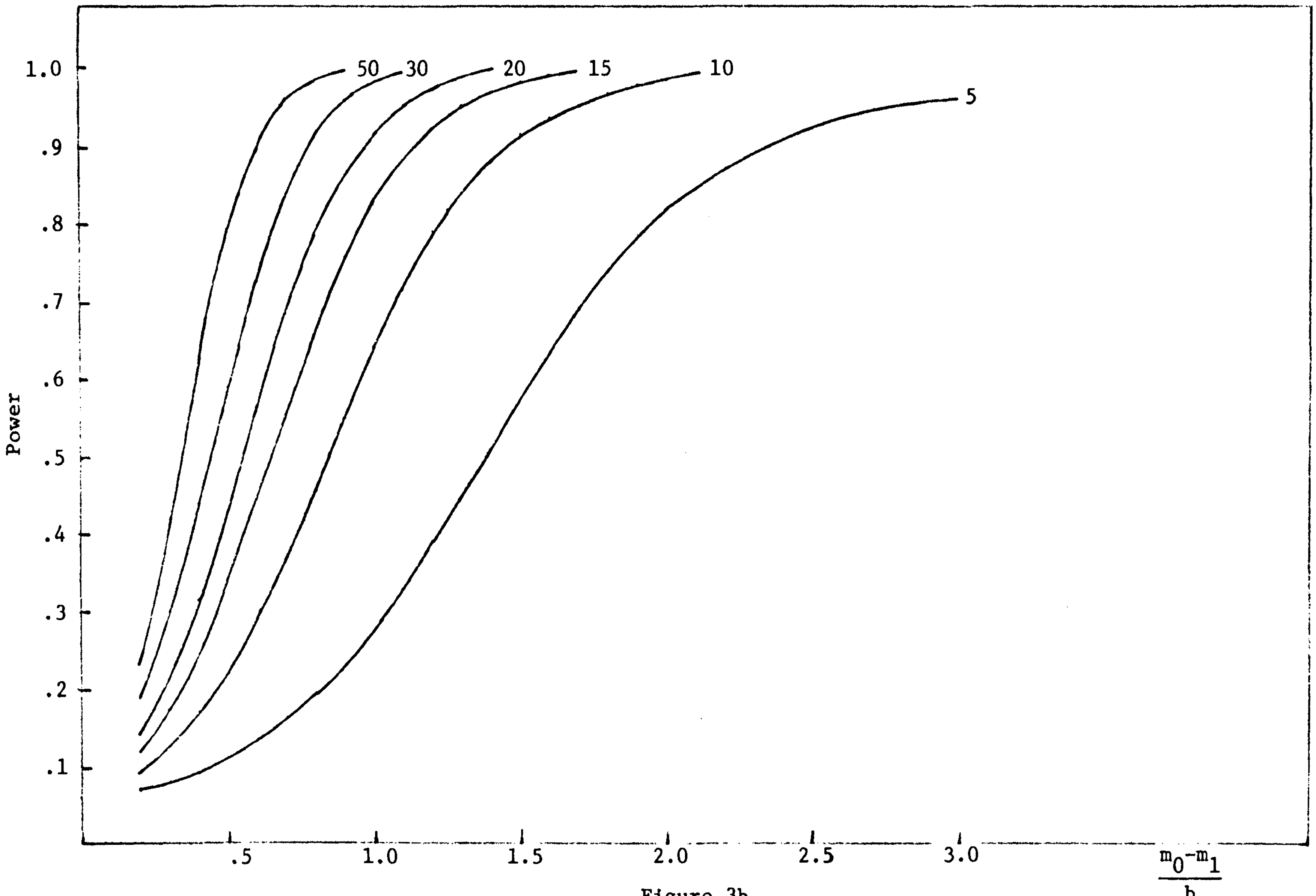

Figure 3b
Power of one-sided .05 level test of m (b known)

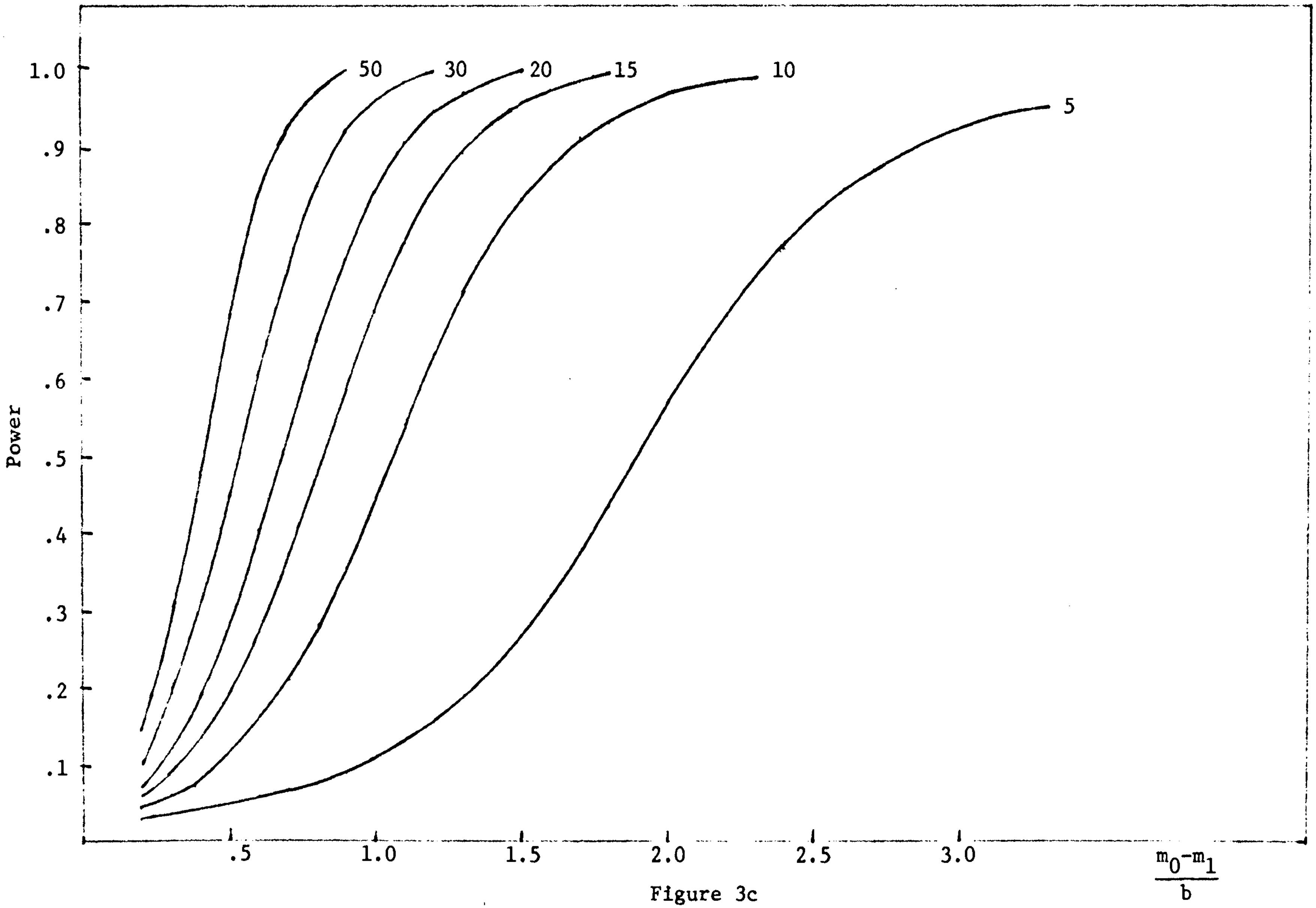

Figure 3c
Power of one-sided .025 level test of m (b known)

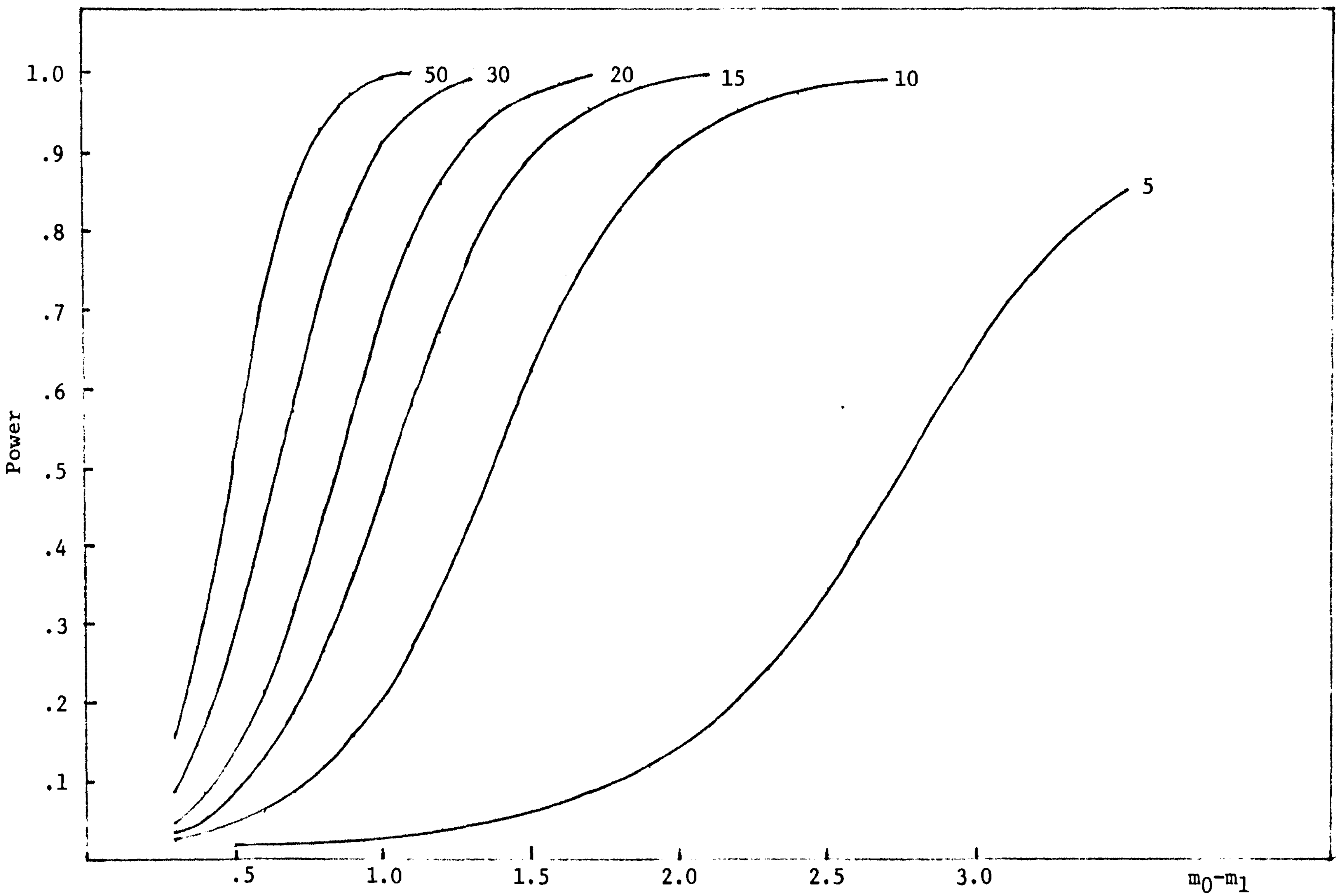

Figure 3d
Power of one-sided .01 level test of m (b known)

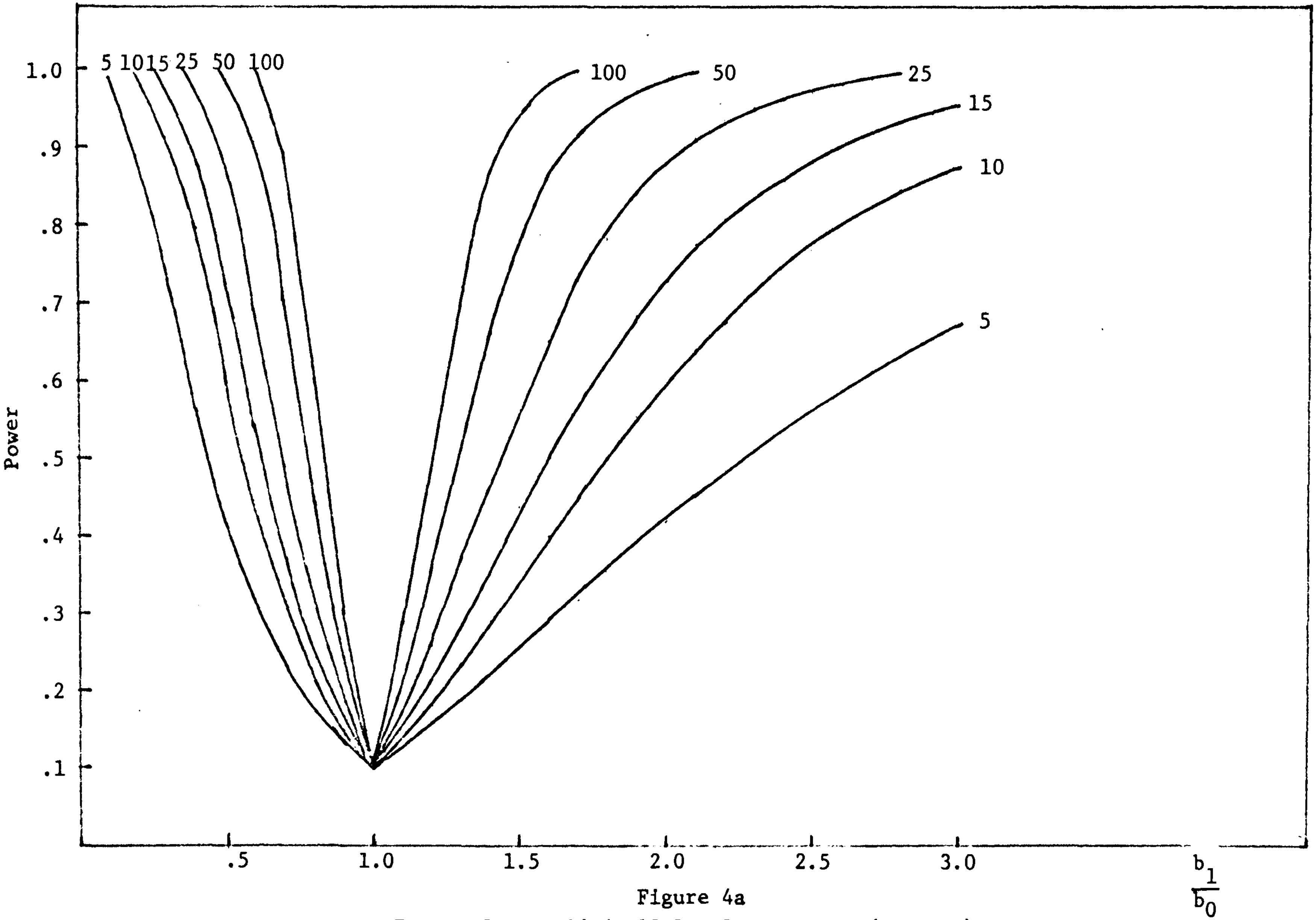

Figure 4a
Power of one-sided .10 level test on b (m known)

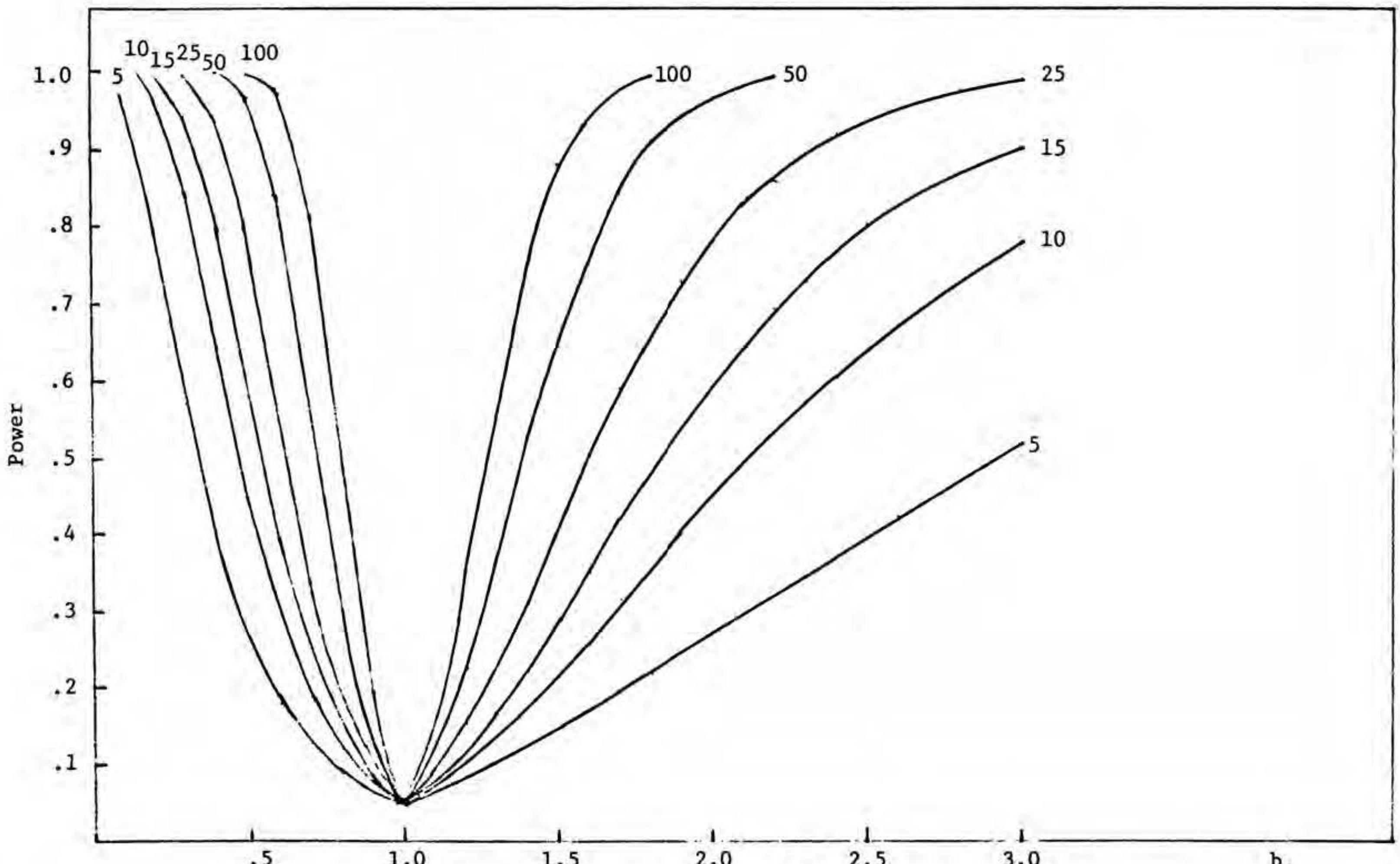

Figure 4b
Power of one-sided .05 level test on b (m known)

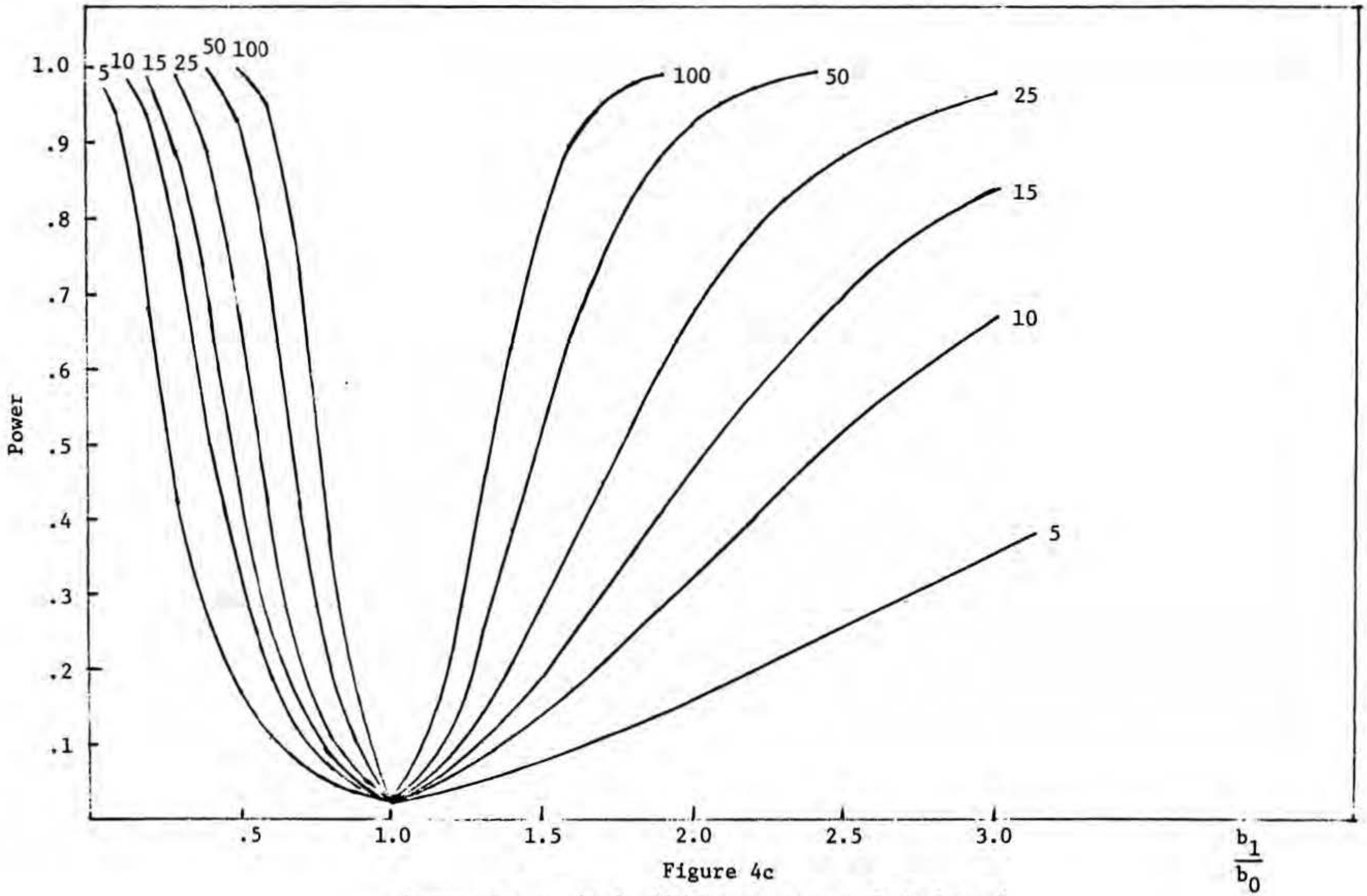

Figure 4c

Power of one-sided .025 level test on b (m known)

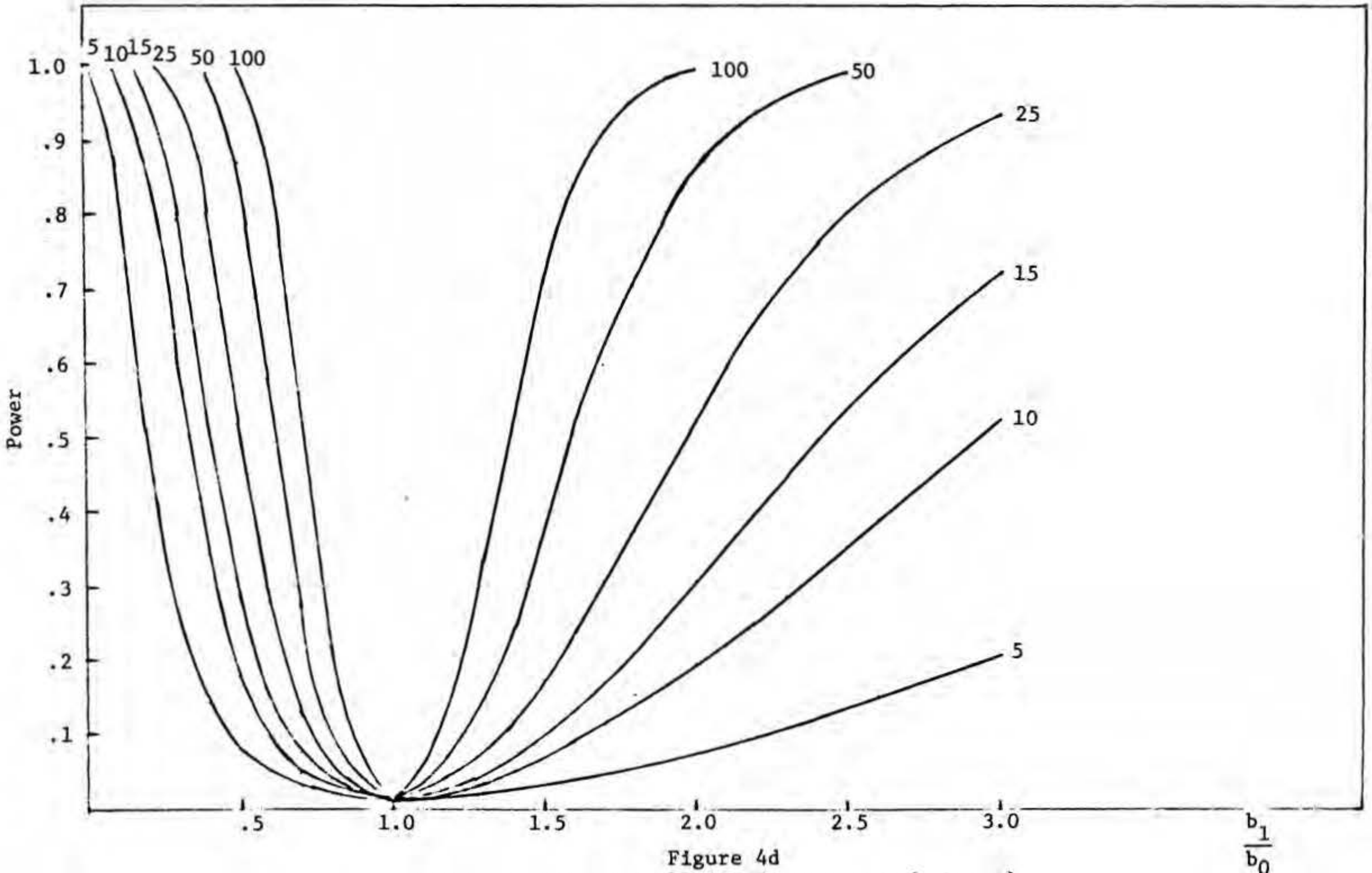

Figure 4d
Power of one-sided .01 level test on b (m known)

Let ξ_β be the point such that $F(\xi_\beta) = 1 - \beta$, and let $x_1, \ldots, x_n$ be a random sample from X. A function, $L(x_1, \ldots, x_n)$ is a lower one-sided β,γ tolerance limit if $P[L(x_1, \ldots, x_n) \le \xi_\beta] = \gamma$.

If the distribution of the random variable X depends on unknown location and scale parameters, then ξ_β may be expressed in the form $m - k(\beta)b$. The following theorem indicates that one-sided tolerance limits may always be obtained for distributions dependent on unknown locations and scale parameters.

<u>THEOREM 4</u>

There exists a function $z(\beta,\gamma,n)$ such that

$$P[\hat{m} - z(\beta,\gamma,n)\hat{b} \le m - k(\beta)b] = \gamma$$

for all m and b.

Proof: Consider the random variable $\frac{\hat{m}-m}{\hat{b}} + k(\beta)\frac{b}{\hat{b}}$. There exists a $G(\beta,\gamma,n)$ such that

$$P\left[\frac{\hat{m}-m}{\hat{b}} + k(\beta)\frac{b}{\hat{b}} \le G(\beta,\gamma,n)\right] = \gamma.$$

A simple algebraic simplification yields

$$P[\hat{m} - G(\beta,\gamma,n)\hat{b} \le m - k(\beta)b] = \gamma$$

Therefore, choose $z(\beta,\gamma,n) = G(\beta,\gamma,n)$.

Q.E.D.

The importance of this theorem lies in the fact that the distribution of $\frac{\hat{m}-m}{\hat{b}} + k(\beta)\frac{b}{\hat{b}}$ depends only on $k(\beta)$ and n. Hence, if its distribution cannot be determined analytically, it can be conveniently determined through a Monte Carlo simulation. Once the distribution of $\frac{\hat{m}-m}{\hat{b}} + k(\beta)\frac{b}{\hat{b}}$ has been determined, one can, for a given β,γ and n, determine $z(\beta,\gamma,n)$ and, hence, determine the tolerance limit $\hat{m} - z(\beta,\gamma,n)\hat{b}$.

For the two parameter Cauchy distribution, the constant $k(\beta)$ is determined such that

$$\frac{1}{2} + \frac{1}{\pi}\tan^{-1}[-k(\beta)] = 1 - \beta$$

and hence

$$k(\beta) = -\tan[\pi(.5 - \beta)] = \tan[\pi(\beta - .5)].$$

The simulated distribution of $\frac{\hat{m}-m}{\hat{b}} + \tan[\pi(\beta - .5)]\frac{b}{\hat{b}}$ was determined for $\beta = .8, .9, .95, .99$. Table A6 presents the required $z(\beta,\gamma,n)$ to form lower one-sided β,γ tolerance intervals for the Cauchy distribution for $\gamma = .90, .95, .99$.

Due to the symmetry of the Cauchy distribution, upper one-sided tolerance intervals are of the form $(-\infty, \hat{m} + z(\beta,\gamma,n)\hat{b})$.

Two-sided β,γ tolerance intervals may be chosen in the form

$$P[F(\hat{m} + w(\beta,\gamma,n)\hat{b}) - F(\hat{m} - w(\beta,\gamma,n)\hat{b}) \geq \beta] = \gamma \qquad (4)$$

where F is the cumulative distribution function. Hence, if one could find $w(\beta,\gamma,n)$, then two-sided β,γ tolerance intervals could be obtained. The following theorem allows one to find $w(\beta,\gamma,n)$ for the Cauchy distribution.

<u>THEOREM 5</u>

The function $w(\beta,\gamma,n)$ that satisfies (4) is defined by

$$P[\frac{b}{\hat{b}}\{-1-\sqrt{1+\tan^2\pi\beta(1+(\frac{\hat{m}-m}{b})^2)}\}/\tan\pi\beta \leq w(\beta,\gamma,n)] = \gamma.$$

Proof: from (4) we have

$$P[\tan^{-1}[(\hat{m} + w(\beta,\gamma,n)\hat{b} - m)/b] - \tan^{-1}[(\hat{m} - w(\beta,\gamma,n)\hat{b} - m)/b] \geq \pi\beta] = \gamma. \qquad (5)$$

Now if A, B and C are such that $-\frac{\pi}{2} < B \leq A \leq \frac{\pi}{2} < C\pi < \pi$, then $A-B \geq C\pi$ if, and only if, $\tan A - \tan B \leq (\tan C)(1+\tan A\tan B)$. Applying this to (5) with $A = \tan^{-1}[(\hat{m} + w(\beta,\gamma,n)\hat{b} - m)/b]$, $B = \tan^{-1}[(\hat{m} - w(\beta,\gamma,n)\hat{b} - m)/b]$, and $C = \beta$ subject to $.5 < \beta < 1$ we have

$$P[2w(\beta,\gamma,n)\frac{\hat{b}}{b} \leq (\tan\pi\beta)[1 + (\frac{\hat{m}-m}{b})^2 - (w(\beta,\gamma,n)\frac{\hat{b}}{b})^2]] = \gamma.$$

Noting that $w(\beta,\gamma,n) \geq 0$ we obtain

$$P[w(\beta,\gamma,n) \geq \frac{b}{\hat{b}}\{-1-\sqrt{1 + \tan^2\pi\beta(1 + (\frac{\hat{m}-m}{b})^2)}\}/\tan\pi\beta] = \gamma.$$

Q.E.D.

Now the distribution of

$$\frac{b}{\hat{b}}\{-1-\sqrt{1+\tan^2\pi\beta(1+(\frac{\hat{m}-m}{b})^2)}\}/\tan\pi\beta \qquad (6)$$

depends upon n and β but not upon m and b. The distribution of (6) was obtained by simulation for $\beta = .8, .9, .95, .99$. The values of $w(\beta,\gamma,n)$ such that $(\hat{m} - w(\beta,\gamma,n)\hat{b},\ \hat{m} + w(\beta,\gamma,n)\hat{b})$ forms a two-sided β,γ tolerance interval for the Cauchy distribution may be found in Table A7 for $\gamma = .90, .95, .99$.

Dumonceaux [23] has since observed that the cumulative function that satisfies (4) is distributed independent of m and b for all densities dependent upon scale and location parameters.

2. One Parameter Known

The question of tolerance intervals is more easily answered when one of the two parameters is known. Consider first the case where the scale parameter b is known. For a lower one-sided β,γ tolerance interval we seek a function $L(x_1, \ldots, x_n)$ such that $P[L(x_1,..x_n) \le m-k(\beta)b] = \gamma$. Consider the probability statement

$$P[\frac{\hat{m}-m}{b} \le z(\beta,\gamma,n) - k(\beta)] = \gamma$$

or

$$P[\hat{m} - z(\beta,\gamma,n)b \le m - k(\beta)b] = \gamma.$$

Hence $L(x_1, \ldots, x_n) = \hat{m} - z(\beta,\gamma,n)b$.

Now the distribution of $\frac{\hat{m} - m}{b/\sqrt{n}}$ is given in Table A4. Hence, for a given value of β, γ, and n, $z(\beta,\gamma,n)$ can be determined. The values of $z(\beta,\gamma,n)$ appear in Table A8.

Again, due to symmetry, $\hat{m} + z(\beta,\gamma,n)b$ forms an upper one-sided β,γ tolerance limit.

For two-sided β,γ tolerance limits with b known, we seek $w(\beta,\gamma,n)$ such that

$$P[\tan^{-1}[\frac{\hat{m} + w(\beta,\gamma,n)b - m}{b}] - \tan^{-1}[\frac{\hat{m} - w(\beta,\gamma,n)b - m}{b}] \geq \pi\beta] = \gamma.$$

or

$$P[2w(\beta,\gamma,n) \leq \tan\pi\beta\ [1 + (\frac{\hat{m} - m}{b})^2 - w^2(\beta,\gamma,n)] = \gamma$$

or

$$P[\frac{\sqrt{n}\ (\hat{m} - m)}{b} \geq \sqrt{[w^2(\beta,\gamma,n) + \frac{2w(\beta,\gamma,n)}{\tan\pi\beta} - 1]n}\ ] = \frac{\gamma + 1}{2}.$$

Hence, from Table A4 we can determine for given β,γ and n, the value of

$$\sqrt{[w^2(\beta,\gamma,n) + \frac{2w(\beta,\gamma,n)}{\tan\pi\beta} - 1]n}.$$

The solution for $w(\beta,\gamma,n)$ can now be obtained. The values of $w(\beta,\gamma,n)$ may be obtained from Table A9.

One-sided β,γ tolerance limits with m known may be obtained in a similar manner. For a lower limit we seek a $L(x_1, \ldots, x_n)$ such that

$$P[L(x_1, \ldots, x_n) \leq m - k(\beta)b] = \gamma.$$

The constant $z(\beta,\gamma,n)$ may be determined from Table A5 such that

$$P[\ \frac{\hat{b}}{b} \geq \frac{k(\beta)}{z(\beta,\gamma,n)}\ ] = \gamma.$$

This expression may be manipulated to yield

$$P[m - z(\beta,\gamma,n)\hat{b} \leq m - k(\beta)b] = \gamma.$$

Hence, let

$$L(x_1, \ldots, x_n) = m - z(\beta,\gamma,n)\hat{b}.$$

Table A10 tabulates the factors $z(\beta,\gamma,n)$.

For two-sided tolerance limits with m known, we seek factors $w(\beta,\gamma,n)$ such that

$$P[\tan^{-1}(w(\beta,\gamma,n)\frac{\hat{b}}{b}) - \tan^{-1}(-w(\beta,\gamma,n)\frac{\hat{b}}{b}) \geq \pi\beta] = \gamma.$$

Simplifying, we have

$$P[2w(\beta,\gamma,n)\frac{\hat{b}}{b} \leq (\tan \pi\beta)(1 - w^2(\beta,\gamma,n)(\frac{\hat{b}}{b})^2] = \gamma$$

or

$$P[w^2(\beta,\gamma,n) \tan \pi\beta(\frac{\hat{b}}{b})^2 + 2w(\beta,\gamma,n)\frac{\hat{b}}{b} - \tan \pi\beta \leq 0] = \gamma$$

which yields

$$P[\frac{\hat{b}}{b} \geq (\frac{-1 - \sqrt{1 + \tan^2\pi\beta}}{w(\beta,\gamma,n)\tan\pi\beta})] = \gamma \qquad (7)$$

for $.5 < \beta < 1$.

Critical values for (7) can be obtained from Table A5. Hence, the value of $w(\beta,\gamma,n)$ can be found. Values of $w(\beta,\gamma,n)$ are given in Table A11.

The following result relating one- and two-sided tolerance limits with m known is of interest.

THEOREM 6

Let $X \sim \frac{1}{b} f(\frac{x - m}{b})$, $-\infty < x < \infty$ where m is known and f is symmetric about m. If b^* is a function of a random sample and k a positive real number such that $(m - kb^*, m + kb^*)$ forms a two-sided β,γ tolerance interval, then $(-\infty, m + kb^*)$ forms an upper one-sided $\frac{\beta + 1}{2},\gamma$ tolerance interval and $(m - kb^*, \infty)$ forms a lower one-sided $\frac{\beta + 1}{2},\gamma$ tolerance interval.

Proof:

Let $m = 0$ so that $X \sim \frac{1}{b} f(\frac{x}{b})$, $-\infty < x < \infty$.

By hypothesis,

$$P[\int_{-kb^*}^{kb^*} \frac{1}{b} f(\frac{x}{b})\, dx \geq \beta] = \gamma.$$

Let k_1 be such that

$$P[\int_{-\infty}^{k_1 b^*} \frac{1}{b} f(\frac{x}{b})\, dx \geq \frac{\beta + 1}{2}] = \gamma.$$

and let

$$\int \frac{1}{b} f(\frac{x}{b})\, dx = F(\frac{x}{b}) + C.$$

Now we have

$$P[F(\frac{k_1 b^*}{b}) \geq \frac{\beta + 1}{2}] = \gamma$$

and

$$P[F(\frac{kb^*}{b}) - F(\frac{-kb^*}{b}) \geq \beta] = \gamma.$$

But due to symmetry,

$$F(\frac{-kb^*}{b}) = 1 - F(\frac{kb^*}{b}).$$

Hence,

$$P[2F(\frac{kb^*}{b}) - 1 \geq \beta] = \gamma$$

or

$$P[F(\frac{kb^*}{b}) \geq \frac{\beta + 1}{2}] = \gamma.$$

Hence, $k_1 = k$.

The lower one-sided result follows from the symmetry of the density function.

Q.E.D.

If we let $b^* = \hat{b}$ and choose the constant k as previously indicated, then the hypotheses of Theorem 6 are satisfied. Therefore, for example, the constant from Table A11 for $\beta = .90$ will be the same as the constant from Table A10 for $\beta = .95$.

3. Asymptotic Results

For the one-sided tolerance limits (m,b unknown) we require a $G = G(\beta, \gamma, n)$ such that

$$P[\frac{\hat{m} - m}{\hat{b}} + k(\beta)\frac{b}{\hat{b}} \leq G] = \gamma$$

or

$$P[\frac{\hat{m} - m}{b} - G\frac{\hat{b}}{b} \leq -k(\beta)] = \gamma.$$

Now

$$\frac{\hat{m} - m}{b} \overset{\sim}{\to} N(0, \frac{2}{n})$$

$$\frac{\hat{b}}{b} \overset{\sim}{\to} N(1, \frac{2}{n})$$

and $\hat{m}, \hat{b}$ are asymptotically independent [9].

Therefore,

$$\frac{\hat{m} - m}{b} - G\frac{\hat{b}}{b} \overset{\sim}{\rightarrow} N\left[-G, \frac{2}{n}(1 + G^2)\right]$$

so that for large n we have

$$-k(\beta) \doteq -G + z_\gamma\sqrt{2(1 + G^2)/n}$$

where z_γ is the γ percentage point for the standard normal.

Solving for G yields

$$G \doteq \frac{-k(\beta) - \sqrt{k^2(\beta) - \left(\frac{2z_\gamma^2}{n} - 1\right)\left(\frac{2z_\gamma^2}{n} - k^2(\beta)\right)}}{\frac{2z_\gamma^2}{n} - 1}$$

The last line in Table A6 gives the asymptotic approximation for $n = 100$. The close agreement gives confidence in the approximation for $n > 100$.

Since the values for two-sided tolerance limits for proportion β are nearly equal to the one-sided tolerance limits for proportion $\frac{\beta + 1}{2}$, equation (8) yields an approximation to the large sample two-sided tolerance limits, also.

The tolerance limits with one parameter known are derived directly from the distributions of $\frac{\hat{m} - m}{b}$ or $\frac{\hat{b}}{b}$. Since these two random variables are distributed asymptotically normal, large sample tolerance limits can be obtained quite easily.

For one-sided tolerance limits with b known, we have

$$z(\beta,\gamma,n) \doteq \sqrt{\frac{2}{n}}z_\gamma + k(\beta).$$

For two-sided tolerance limits with b known,

$$w \doteq \frac{-1 - \sqrt{1 + \tan^2\pi\beta\left[1 + 2z^2_{\frac{\gamma+1}{2}}/n\right]}}{\tan \pi\beta}$$

For one-sided tolerance limits with m known we have

$$z(\beta,\gamma,n) \doteq \frac{k(\beta)}{1 - \sqrt{\frac{2}{n}}\, z_\gamma}$$

For two-sided tolerance limits with m known,

$$w(\beta,\gamma,n) \doteq \frac{-1 - \sqrt{1 + \tan^2 \pi\beta}}{\tan \pi\beta(1 - \sqrt{\frac{2}{n}}\, z_\gamma)}.$$

E. Comparison of Estimators

1. Comparison of Point Estimators

Barnett [15] compares the variances of estimators of the location parameter. He considers the median, BLUE, Rothenberg et al estimator [12], and the MLE (b known). In Table 3 the comparisons from this study are presented.

Table 3

Variances of Unbiased Estimators of m (Divided by b^2)

n	BLUE	MLE (b known)	MLE (b unknown)*	MVB	Efficiency (Using*)
5	1.211	1.116	1.128	.4	.355
10	.326	.2804	.2873	.2	.695
15	.182	.1650	.1659	.1333	.805
20	.126	.1156	.1175	.1	.851
25		.0885	.0904	.08	.885
30		.0727	.0735	.0667	.907
40		.0535	.0542	.05	.924
50			.0424	.04	.945
75			.0273	.0267	.973
100			.0201	.02	.995

It can be seen that the maximum likelihood estimators have smaller variances than the BLUE for all sample sizes. It is also known that the MLE are asymptotically efficient. Therefore, using the minimum variance unbiased criteria, the MLE is a better estimator than the BLUE. Hence, the MLE is superior to all estimators discussed in the review of the literature since they are all linear functions of the ordered sample.

To investigate the accuracy of the simulated variances, the variance of the median for sample sizes 5, 9 and 15 was obtained by simulation and compared to the true variances [11]. The variance of the median for n = 5, 9, 15 is 1.2213, .4087, and .2048, respectively. For the simulated variance of the median, 20,000 samples were obtained for n = 5 and 10,000 samples for n = 9 and n = 15. The variances obtained by simulation were 1.266, .4163, and .2073. It is reasonable to assume that the variances of the maximum likelihood estimators are at least as accurate as that obtained for the median, since a larger simulation was conducted for the MLE. See Table 14 for the number of estimates used in this study. There is very little disagreement between the variances obtained by Barnett and the variances obtained in this study. Since a larger simulation was performed for this study, it seems reasonable to assume that the results presented here are at least as accurate as Barnett's.

The maximum likelihood estimator of b is not unbiased. However, from the simulation it was observed that the expected value of $\hat{b}$ differed from b by less than .0067 for all sample sizes considered. This bias could be removed; however, it hardly seems worth the effort, since it is so small. For all practical purposes, $\hat{b}$ is an unbiased estimator of b. Table 4 tabulates the variance of various estimators of b.

Table 4

Variances of Estimators of b (Divided by b^2)

n	$k[x_{(n-2)} - x_{(3)}]$	BLUE	MLE (m known)	MLE (m unknown)*	MVB	efficiency (using*)
5	----	--	1.143	1.017	.4	.394
10	.603	.468	.3015	.2954	.2	.679
15	.502	.207	.1715	.1664	.1333	.802
20	.482	.135	.1204	.1187	.1	.844
25			.0932	.0922	.08	.858
30			.0753	.0755	.0667	.883
40			.0552	.0541	.05	.924
50			.0434	.0426	.04	.940
75			.0277	.0284	.0267	.942
100			.0211	.0207	.02	.967

Again it is seen that the MLE is superior to the BLUE.

2. Comparison of Confidence Intervals

To compare confidence intervals using the MLE and the BLUE, the distributions of $\frac{m^* - m}{b/\sqrt{n}}$, $\frac{m^* - m}{b^*/\sqrt{n}}$, $\frac{b^*}{b}$ were obtained by simulation. The simulated distributions were obtained for n = 10, 15, and 20 (10,000 samples for each n). These distributions appear in Tables 5, 6, and 7.

Table 5

Critical Values for $\frac{m^* - m}{b/\sqrt{n}}$

Cumulative

n	.8	.9	.95	.975	.99
10	1.300	2.095	2.862	3.620	4.790
15	1.270	2.001	2.680	3.378	4.270
20	1.250	1.935	2.572	3.160	3.920

Table 6

Critical Values for $\frac{m^* - m}{b^*/\sqrt{n}}$

Cumulative

n	.8	.9	.95	.975	.99
10	1.520	2.465	3.435	4.380	5.800
15	1.397	2.222	2.962	3.740	4.640
20	1.330	2.065	2.755	3.422	4.260

Table 7

Critical Values for b^*/b

Cumulative

n	.01	.025	.05	.1	.9	.95	.975	.99
10	.214	.278	.331	.412	1.720	2.090	2.490	3.240
15	.332	.394	.454	.530	1.548	1.814	2.088	3.513
20	.407	.465	.530	.598	1.460	1.674	1.885	2.180

By comparing Tables 5, 6, and 7 with Tables A2, A3, A4, and A5, one can observe that the expected width of the confidence interval is shorter for the MLE. For example, for $n = 10$, $\alpha = .05$, the expected width of the two-sided confidence interval for m(b unknown) based on the MLE is $2.500 \cdot b$. The corresponding expected width based on the BLUE is $2.776 \cdot b$.

Similarly, the expected width of a two-sided .95 confidence interval for b(m unknown, $n = 10$) is $2.705 \cdot b$ using the MLE and $3.218 \cdot b$ using the BLUE.

F. Example

The following example is presented to illustrate the use of the tables discussed. Consider the following ordered sample of size 5 from a Cauchy distribution (assume m and b unknown): -2.07, -1.64, -1.03,

.154, 4.02. The maximum likelihood estimators are $\hat{m} = -1.288$, $\hat{b} = .793$. (A computer program is given in Appendix B which will obtain $\hat{m}$ and $\hat{b}$.)

Therefore, from Tables A2 and A3 we find that

$$(-1.288 \pm (6.780)(.793/\sqrt{5}) = (-1.116, 3.692)$$

forms a .95 confidence interval for m, and

$$(.793/3.279,\ .793/.132) = (.242, 6.100)$$

forms a .95 confidence interval for b.

Consider the test of hypothesis $H_0 : m = 1$, $H_1 : m < 1$. The hypothesis is rejected for values of $\frac{\hat{m} - 1}{\hat{b}/\sqrt{5}} < -4.771$ at the .05 level. For our sample we have $\frac{\hat{m} - 1}{b/\sqrt{5}} = \frac{-1.288 - 1}{.793/\sqrt{5}} = -6.49$. The hypothesis would be rejected at the .05 level. The power of the test ($\alpha = .05$) for $\frac{m_0 - m_1}{b} = 2$ is .61. Similarly, $H_0 : b = 1$, $H_1 : b \neq 1$ would have critical values of $\hat{b} = .132$ and $\hat{b} = 3.279$. Therefore, since $\hat{b} = .793$, we would not reject H_0 at the .05 level. The power of the test would be .235 for $b_1/b_0 = 2$. For $\beta = .90$, $\gamma = .95$, we have, from Table A6, $-1.288 - (17.314)(.793) = -15.018$ as a lower one-sided tolerance limit.

IV INFERENCES BASED ON TWO INDEPENDENT SAMPLES

A. Scale Parameters Known

1. Confidence Intervals on the Difference in Location Parameters

Consider two independent Cauchy samples; $X_1, \ldots, X_{n_1}$ with parameters m_1 and b_1, and $Y_1, \ldots, Y_{n_2}$ with parameters m_2 and b_2. Assume that b_1 and b_2 are known. Confidence intervals on $m_1 - m_2$ may be obtained as follows: The distribution of the random variable $\frac{\hat{m}_1 - m_1}{b_1}$ depends only on n_1, and the distribution of the random variable $\frac{\hat{m}_2 - m_2}{b_2}$ depends only on n_2. Hence, the distribution of the random variable

$$\frac{\hat{m}_1 - m_1}{b_1} - \frac{b_2}{b_1}\left(\frac{\hat{m}_2 - m_2}{b_2}\right)$$

depends only on n_1, n_2 and the ratio b_2/b_1. The distribution of

$$\frac{\hat{m}_1 - m_1}{b_1} - \frac{b_2}{b_1}\left(\frac{\hat{m}_2 - m_2}{b_2}\right)$$

was obtained by simulation for selected values of n_1, n_2, b_2/b_1, and the critical values appear in Table A12. From Table A12 one can obtain c such that (interpolation may be required)

$$P\left[-c < \frac{\hat{m}_1 - m_1}{b_1} - \frac{b_2}{b_1}\left(\frac{\hat{m}_2 - m_2}{b_2}\right) < c\right] = 1 - \alpha;$$

and, hence,

$$P[\hat{m}_1 - \hat{m}_2 - cb_1 < m_1 - m_2 < \hat{m}_1 - \hat{m}_2 + cb_1] = 1 - \alpha,$$

which establishes the desired confidence interval.

2. Tests of Hypotheses

A test of hypothesis on the equality of the location parameters will be presented now. As an illustration, consider $H_o : m_1 = m_2$, $H_1 : m_1 > m_2$. Under the null hypothesis, one can find a constant c from Table A12 such that

$$P[\hat{m}_1 - \hat{m}_2 > cb_1] = \alpha.$$

Thus, H_o is rejected if the observed $\hat{m}_1 - \hat{m}_2 > cb_1$. The power may be obtained as follows:

$$\text{Power} = P[\hat{m}_1 - \hat{m}_2 \geq cb_1 \mid m_1 \geq m_2]$$
$$= P[\hat{m}_1 - m_1 - \hat{m}_2 + m_2 \geq -m_1 + m_2 + cb_1]$$
$$= P[\frac{\hat{m}_1 - m_1}{b_1} - \frac{(\hat{m}_2 - m_2)}{b_1} \geq \frac{-(m_1 - m_2)}{b_1} + c]$$
$$= P[\frac{\hat{m}_1 - m_1}{b_1} - \frac{b_2}{b_1} (\frac{\hat{m}_2 - m_2}{b_2}) \geq c - \frac{(m_1 - m_2)}{b_1}].$$

Therefore, the power of the test may also be obtained from Table A12.

B. Location Parameter Known

1. Confidence Intervals on the Ratio of the Scale Parameters

As in IV-A, consider two independent Cauchy samples. In this section, assume m_1 and m_2 are known. Although the calculation of $\hat{b}_1$ and $\hat{b}_2$ depends on m_1 and m_2, the distribution of $\hat{b}_1/b_1$ depends on n_1 only; and the distribution of $\hat{b}_2/b_2$ depends only on n_2. Hence, the distribution of $(\hat{b}_2/b_2)(b_1/\hat{b}_1)$ depends only on n_1 and n_2. The distribution of $(\hat{b}_2/b_2)(b_1/\hat{b}_1)$ was obtained by a Monte Carlo simulation and appears in Table A13. Confidence intervals on b_1/b_2 may be obtained as follows: From Table A13 we may find the constants c_1, c_2 such that

$$P[c_1 < \frac{\hat{b}_2}{b_2} \cdot \frac{b_1}{\hat{b}_1} < c_2] = 1 - \alpha;$$

and, hence,

$$P[c_1 \frac{\hat{b}_1}{\hat{b}_2} < \frac{b_1}{b_2} < c_2 \frac{\hat{b}_1}{\hat{b}_2}] = 1 - \alpha,$$

which forms the desired confidence interval.

2. Tests of Hypotheses

A test of hypothesis on the equality of the scale parameters (location parameters known) will now be presented. As an example, consider

$H_o : b_1 = b_2$, $H_1 : b_1 > b_2$. Under the null hypothesis, one can find from Table A13 the constant c such that

$$P[\frac{\hat{b}_2}{\hat{b}_1} > c] = \alpha.$$

Hence, H_o will be rejected when $\frac{\hat{b}_2}{\hat{b}_1} > c$. The power of the test may be obtained as follows:

$$\text{Power} = P[\frac{\hat{b}_2}{\hat{b}_1} > c \mid b_1 > b_2]$$

$$= P[\frac{\hat{b}_2}{b_2} \cdot \frac{b_1}{\hat{b}_1} > c \frac{b_1}{b_2}].$$

Hence, the power also may be obtained from Table A13.

C. Scale Parameters Unknown but Assumed Equal

1. Confidence Intervals on the Difference in Location Parameters

Here we assume that $b_1 = b_2 = b$ where b is unknown. Confidence intervals on $m_1 - m_2$ may be obtained as follows: We know that $\frac{\hat{m}_1 - m_1}{b}$, $\frac{\hat{b}_1}{b}$, $\frac{\hat{m}_2 - m_2}{b}$, and $\frac{\hat{b}_2}{b}$ have distributions dependent on n_1 and n_2. Hence,

$$\frac{\frac{\hat{m}_1 - m_1}{b} - \frac{\hat{m}_2 - m_2}{b}}{\frac{\hat{b}_1}{b} + \frac{\hat{b}_2}{b}}$$

or

$$\frac{\hat{m}_1 - m_1 - (\hat{m}_2 - m_2)}{\hat{b}_1 + \hat{b}_2}$$

depends only on n_1 and n_2. The distribution of the above statistic was obtained by simulation and appears in Table A14. From Table A14 we may find c such that

$$P[-c \leq \frac{(\hat{m}_1 - m_1) - (\hat{m}_2 - m_2)}{\hat{b}_1 = \hat{b}_2} \leq c] = 1 - \alpha;$$

and, hence,

$$P[(\hat{m}_1 - \hat{m}_2) - c(\hat{b}_1 + \hat{b}_2) \le m_1 - m_2 \le (\hat{m}_1 - \hat{m}_2) + c(\hat{b}_1 + \hat{b}_2)] = 1 - \alpha,$$

which forms the desired confidence interval.

2. Tests of Hypotheses

Consider the test $H_o : m_1 = m_2$, $H_1 : m_1 > m_2$. The null hypothesis is rejected if

$$P[\frac{\hat{m}_1 - \hat{m}_2}{\hat{b}_1 + \hat{b}_2} > c] = \alpha$$

where c comes from Table A14.

Under the alternative hypothesis, the power can be obtained as follows:

$$\text{Power} = P[\frac{\hat{m}_1 - \hat{m}_2}{\hat{b}_1 + \hat{b}_2} > c \mid m_1 > m_2]$$

$$= P[\hat{m}_1 - \hat{m}_2 > c(\hat{b}_1 + \hat{b}_2)]$$

$$= P[\hat{m}_1 - m_1 - \hat{m}_2 + m_2 > -m_1 + m_2 + c(\hat{b}_1 + \hat{b}_2)]$$

$$= P[\frac{\hat{m}_1 - m_1}{b} - \frac{\hat{m}_2 - m_2}{b} - c(\frac{\hat{b}_1}{b} + \frac{\hat{b}_2}{b}) > \frac{m_2 - m_1}{b}].$$

Therefore, the power of the test may be determined from the critical values of $\frac{\hat{m}_1 - m_1}{b} - \frac{(\hat{m}_2 - m_2)}{b} - c(\frac{\hat{b}_1 + \hat{b}_2}{b})$ which depends upon c, n_1, and n_2. The distribution of this random variable was not obtained.

D. Scale and Location Parameters Unknown

1. Confidence Intervals on the Difference in Location Parameters

This problem, analagous to the Behrens-Fisher problem in normal theory, may be handled in a similar manner. Let $X_1, \ldots, X_n$ be a random sample from a Cauchy distribution with parameters m_1 and b_1. Let $Y_1, \ldots, Y_n$ (same sample size) be a random sample from a Cauchy distribution with parameters m_2 and b_2. Let $D_i = X_i - Y_i$, $i = 1, \ldots, n$. It is known [1] that D_i has a Cauchy distribution with parameters m_1-m_2

and $b_1 + b_2$. Hence, the methods of Chapter III-A-2 may be applied to the paired variable D_i to obtain a confidence interval for $m_1 - m_2$.

2. Tests of Hypotheses on the Location Parameters.

Tests of hypotheses may be carried out using the results of III-A-3.

3. Confidence Intervals on the Ratio of the Scale Parameters

It has been pointed out in IV-B-1 that $\frac{\hat{b}_2}{b_2} \cdot \frac{b_1}{\hat{b}_1}$ has a distribution dependent only on the sample sizes n_1 and n_2. The equations for obtaining $\hat{b}_2$ and $\hat{b}_1$ are, of course, different when m_1 and m_2 are unknown. The distribution of $\frac{\hat{b}_2}{b_2} \cdot \frac{b_1}{\hat{b}_1}$ with m_1, m_2 unknown has been obtained by simulation and appears in Table A15. The method of setting confidence intervals on $\frac{b_1}{b_2}$ is identical to that of IV-B-1 using, however, critical values from Table A15.

4. Tests of Hypotheses on the Scale Parameters

Tests of hypotheses on the scale parameters are performed in a manner similar to that in IV-B-2. The critical values come from Table A15.

E. <u>Example</u>

Consider two samples of size ten where $b_1 = b_2 = 2$. Suppose $m_1 = 3.5$ and $m_2 = 3.0$. From Table A12 we find the constant c = 1.550 such that $(3.5 - 3.0 \pm 1.55(2)) = (-2.6, 3.6)$ forms a .95 confidence interval for $m_1 - m_2$. If b_1 and b_2 were not known but rather $\hat{b}_1 = 2$, $\hat{b}_2 = 2$, then from Table A14 we find the constant c = .792 such that $(3.5 - 3.0 + .792(2 + 2)) = (-2.668, 3.668)$ forms a .95 confidence interval on $m_1 - m_2$.

V TRUNCATED AND CENSORED SAMPLES

A. Unbiased Estimators of the Location and Scale Parameters

Antle and Bain [17] point out that the distributions of $\frac{\hat{m}-m}{b}$, $\frac{\hat{b}}{b}$, and $\frac{\hat{m}-m}{\hat{b}}$ depend only on the sample size n and the point of censoring for type II censored samples and truncated distributions when the parent distribution is dependent on a location parameter and a scale parameter. In general, the likelihood equations must be solved by iterative methods [9] and the distributions of the pivotal functions determined by simulation.

It has been observed in III-A-2 that if the distribution of the random variable is symmetric about its location parameter, m, then $E(\hat{m}) = m$. In general, truncated and censored distributions do not have the required symmetry property so that $\hat{m}$ is not an unbiased estimator of m. The following theorem is useful in finding an unbiased estimator of m.

THEOREM 7

If m* and b* are functions of a random sample of size n from f(x; m,b) such that $E(\frac{m^* - m}{b}) = K(n)$, and $E(\frac{b^*}{b}) = L(n)$, then $m^* - \frac{K(n)b^*}{L(n)}$ is an unbiased estimator of m(m and b are not necessarily location and scale parameters).

Proof: From the hypothesis, we have

$$E(m^*) = bK(n) + m \qquad E(b^*) = L(n)b.$$

Now consider $w^* = m^* - \frac{K(n)}{L(n)} b^*$

$$E(w^*) = E(m^*) - \frac{K(n)}{L(n)} E(b^*) = m.$$

Q.E.D.

Since the distributions of $\frac{\hat{m} - m}{b}$, $\frac{\hat{b}}{b}$ depend on n alone, we have $E(\frac{\hat{m} - m}{b}) = K(n)$, $E(\frac{\hat{b}}{b}) = L(n)$; and, hence, $\hat{m} - \frac{K(n)}{L(n)} \hat{b}$ is an unbiased estimator of m. Since $E(\frac{\hat{b}}{b}) = L(n)$, it follows that $\frac{\hat{b}}{L(n)}$ is an unbiased estimator of b.

B. Confidence Intervals for the Location and Scale Parameters

If the distributions of $\frac{\hat{m} - m}{b}$, $\frac{\hat{b}}{b}$, and $\frac{\hat{m} - m}{\hat{b}}$ are known, it follows that confidence intervals for m and b could be obtained. Now the distributions of these pivotal functions depend upon the truncation or level of type II censoring involved. Consequently, a general Monte Carlo simulation would be quite lengthy. For this reason, a general program for generating the distributions of $\frac{\hat{m} - m}{b}$, $\frac{\hat{b}}{b}$, and $\frac{\hat{m} - m}{\hat{b}}$ along with $E(\frac{\hat{m} - m}{b})$ and $E(\frac{\hat{b}}{b})$ is provided for the Cauchy distribution (type II censoring) in Appendix II. The level of censoring is an input parameter to the program. A users guide is provided within the program. Hence, it is possible, by using this program, to find unbiased point estimators of m and b and set confidence intervals for m and b for type II censored samples from Cauchy distributions.

VI DISCRIMINATION BETWEEN CAUCHY AND NORMAL SAMPLES

A. The Chi Square and Related Tests

It would be useful for an applied statistician to have a method available for choosing between the normal distribution and the Cauchy distribution. In this chapter, ways of performing the test of hypothesis H_o : Sample from Normal Distribution, H_1 : Sample from Cauchy Distribution will be investigated. The first test considered is the Chi Square Goodness of Fit Test.

Let

$$X_n^2 = \sum_{i=1}^{k} \frac{(E_i - O_i)^2}{E_i}.$$

The k cells were determined in the following manner: The sample mean, $\bar{x}$, and the sample variance, s^2, were calculated; and five intervals $(-\infty, \bar{x}-k_2s)$, $(\bar{x}-k_2s, \bar{x}-k_1s)$, $(\bar{x}-k_1s, \bar{x}+k_1s)$, $(\bar{x}+k_1s, \bar{x}+k_2s)$, $(\bar{x}+k_2s, \infty)$ were determined by choosing k_1 and k_2 such that if the random variable X had a normal distribution with $\mu = \bar{x}$, $\sigma^2 = s^2$, then each interval would contain 20% of the area under the normal curve. Therefore, $E_i = 1/5$ (total number of observations), $i = 1, \ldots, 5$, and O_i = the number of observations in the i^{th} cell, $i = 1, \ldots, 5$.

Now, under H_o, X_n^2 has an approximate Chi Square distribution with three degrees of freedom. Since this is only an approximation, and since one cannot determine the power of the test in this manner, the distribution of X_n^2 was determined by a Monte Carlo simulation.

The distribution of X_n^2 is independent of the scale and location parameters of the distribution of the random sample, as seen from the following argument: Let $x_1, \ldots, x_n$ be a set of observations. Let $y_1, \ldots, y_n$ be the corresponding sample formed by letting $y_i = \frac{x_i - a}{c}$, $c > 0$, $i = 1, \ldots, n$. It is well known that $\bar{y} = \frac{\bar{x} - a}{c}$, $s_y = \frac{s_x}{c}$.

Hence, if x_i satisfies the inequality $\bar{x} + d_1 s_x < x_i < \bar{x} + d_2 s_x$, then $\bar{y} + d_1 s_y < y_i < \bar{y} + d_2 s_y$. Therefore, if x_i is in the k^{th} cell based on the sample of x's, then y_i is in the k^{th} cell based on the sample of y's.

To determine the power of the test, the distribution of X_n^2, when the observations are from a Cauchy distribution, is required. A comparison of this test with others is presented in Part C of this chapter.

Another test, related to the Chi Square Test, will now be discussed. Let

$$X_c^2 = \sum_{i=1}^{5} \frac{(E_i - O_i)^2}{E_i}$$

where E_i and O_i are defined as in the definition of X_n^2 except that: (1) the intervals are formed using $\hat{m}$ and $\hat{b}$, the maximum likelihood estimators of m and b in the Cauchy distribution, and (2) k_1 and k_2 are determined such that E_i = 1/5 (total number of observations) if the sample really came from a Cauchy distribution with parameters $m=\hat{m}$, $b=\hat{b}$.

From the following result, a lemma to the theorem in VI-B, it is observed that X_c^2 has a distribution independent of the scale and location parameters of the distribution of the random sample.

Lemma: Let $x_1, \ldots, x_n$ be a set of n observations. Let $y_1, \ldots, y_n$ be the corresponding set formed by letting $y_i = \frac{x_i - a}{c}$, $c > 0$, $i = 1, \ldots, n$. Let $\hat{m}_x, \hat{b}_x$ be the maximum likelihood estimators of the Cauchy parameters based on the sample of x's. Let $\hat{m}_y$, $\hat{b}_y$ be the MLE's based on the y sample. Then $\hat{m}_y = \frac{\hat{m}_x - a}{c}$ and $\hat{b}_y = \frac{\hat{b}_x}{c}$.

Proof:

$$\prod_{i=1}^{n} \frac{1}{\pi \hat{b}_x [1 + (\frac{x_i - \hat{m}_x}{\hat{b}_x})^2]} = \max_{m,b} \prod_{i=1}^{n} \frac{1}{\pi b[1 + (\frac{x_i - m}{b})^2]}$$

Let $x_i = cy_i + a$, and divide both sides by c. Then

$$\prod_{i=1}^{n} \frac{1}{\pi(\frac{\hat{b}_x}{c})\{1 + [\frac{y_i - (\frac{\hat{m}_x - a}{c})}{(\frac{\hat{b}_x}{c})}]^2\}} = \max_{m,b} \prod_{i=1}^{n} \frac{1}{\pi(\frac{b}{c})\{1 + [\frac{y_i - (\frac{m-a}{c})}{(\frac{b}{c})}]^2\}}.$$

Therefore, by letting $\hat{m}_y = \frac{\hat{m}_x - a}{c}$ and $\hat{b}_y = \frac{\hat{b}_x}{c}$, the likelihood function for the y sample is maximized.

Q.E.D.

Therefore, X_c^2 is distributed independently of the scale and location parameters of the distribution of the random sample. Consequently, it is feasible to obtain the distribution of $R = X_n^2/X_c^2$ by a Monte Carlo Simulation. The distribution of R was obtained for normal samples and Cauchy samples. Comparisons are made in part C of this chapter.

B. The Likelihood Ratio Test

As a third way of performing the desired test of hypothesis, consider a test based on the likelihood ratio

$$\lambda = \frac{\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\, s} e^{\frac{1}{-2s^2}(x_i - \bar{x})^2}}{\prod_{i=1}^{n} \frac{1}{\pi\hat{b}[1 + (\frac{x_i - \hat{m}}{\hat{b}})^2]}}$$

The following theorem is useful in considering a test based on λ.

THEOREM 8

The distribution of λ does not depend upon the location and scale parameters of the distribution of the x's.

Proof: Let $x_1, \ldots, x_n$ be a set of observations. Let $y_1, \ldots, y_n$ be a new set formed by letting $y_i = \frac{x_i - a}{c}$, $c > 0$, $i = 1, \ldots, n$. Now $\bar{y} = \frac{\bar{x} - a}{c}$, $s_y = \frac{s_x}{c}$, and from the previous lemma, $\hat{m}_y = \frac{\hat{m}_x - a}{c}$, $\hat{b}_y = \frac{\hat{b}_x}{c}$.

Let λ_x be determined from the x sample and λ_y from the y sample. In the expression for λ_x, divide numerator and denominator by c, and replace x_i with $cy_i + a$ which yields $\lambda_x = \lambda_y$.

Q.E.D.

Dumonceaux [23] has generalized this result to arbitrary scale and location parameter densities.

The distribution of λ was determined for samples from a normal distribution to determine the critical values of the test and from samples from a Cauchy distribution to determine the power of the test. The tables are presented in the next section.

C. Comparison of Methods

The following tables were prepared using a Monte Carlo simulation of the statistics X_n^2, R, and λ for testing the hypothesis H_0 : Sample from Normal Distribution, H_1 : Sample from Cauchy Distribution. L_c is the critical value for the test (the test is rejected for $\lambda < L_c$, $X_n^2 > L_c$, or $R > L_c$), α is the probability of a Type I error, and β is the power of the test. Due to the discrete nature of X_n^2 and R, the α's are approximate.

Table 8

Discrimination Based on X_n^2

n	$\alpha = .01$		$\alpha = .05$		$\alpha = .1$		$\alpha = .2$	
	L_c	β	L_c	β	L_c	β	L_c	β
15	4.8	.53	2.8	.61	1.6	.72	1.2	.82
25	9.6	.79	6.8	.86	5.2	.90	3.6	.95
35	14.4	.90	9.6	.945	8.0	.96	6.4	.98
50	--	--	14.0	.98	12.4	.987	10.0	.995

Table 9

Discrimination Based on R

n	α = .01		α = .05		α = .1		α = .2	
	L_c	β	L_c	β	L_c	β	L_c	β
15					9.0	.54		
25			5.75	.73	3.67	.83	2.0	.91
35	6.0	.77	3.0	.91	2.0	.95	1.2	.98
50	2.3	.962	1.21	.991	.95	.995	.7	.999

Table 10

Discrimination Based on λ

n	α = .01		α = .05		α = .1		α = .2	
	L_c	β	L_c	β	L_c	β	L_c	β
5	.086	.19	.363	.32	.628	.40	1.203	.53
10	.107	.51	.534	.65	1.118	.73	2.340	.81
15	.198	.74	1.010	.83	2.176	.88	5.045	.92
25	.67	.93	3.45	.96	8.98	.98	---	--
35	2.20	.986	15.6	.993	37.8	.995	121.5	.999
50	13.0	.997	125.0	.998	---	--	---	--

As is clearly indicated by Tables 8, 9, and 10, the likelihood ratio test is superior to the tests based on X_n^2 and R.

D. Example

The following data were obtained by Michelson in 1926 in determining the velocity of light by reflecting a light beam by mirrors between Mount Wilson and Mount San Antonio [24]. The data have been coded so that only the last two figures appear. The data are as follows:

47, 38, 29, 92, 41, 44, 47, 62, 59, 44, 47, 41

Using these 12 observations, the following results are obtained: $\bar{x} = 49.25$, $s = 16.03$, $\hat{m} = 44.46$, $\hat{b} = 4.39$. The value of the likelihood ratio is $\lambda = .072$. From Table 10 we see that even at the $\alpha = .01$ level, it is concluded that the sample comes from a Cauchy distribution. The power of the test may also be determined by interpolation from Table 10 for each of the four values of α.

The estimate of the velocity of light based on the Cauchy distribution is $c = 44.46$ with a .90 confidence interval of (40.62, 48.30).

VII NUMERICAL SOLUTION TO THE LIKELIHOOD EQUATIONS

A. Scale Parameter Known

The methods for solving the likelihood equations are significantly different depending upon whether b is known, m is known, or both m and b are unknown. In this section, the case where b is known will be considered.

For convenience, assume b = 1. The likelihood function L(x;m) is given by

$$L(x;m) = \prod_{i=1}^{n} \frac{1}{\pi[1 + (x_i - m)^2]}$$

and log L(x;m) by

$$\log L(x;m) = -n \log \pi - \sum_{i=1}^{n} \log[1 + (x_i - m)^2].$$

Hence, the likelihood equation $\frac{\partial \log L(x;m)}{\partial m} = 0$ is given by

$$\frac{\partial \log L(x;m)}{\partial m} = \sum_{i=1}^{n} \frac{2(x_i - m)}{1 + (x_i - m)^2} = 0. \tag{1}$$

This equation is obviously nonlinear in the unknown m; and, hence, an iterative technique must be used to find $\hat{m}$, a root of the above equations.

Barnett [14] points out that approximately 30% of the time there exist multiple roots to (1). Therefore, one must use caution in solving for a root. Barnett observes that for $n \geq 13$, the root maximizing L(x;m) was always the root nearest the median.

The Newton-Raphson Method was used with the sample median as a starting value in an attempt to find $\hat{m}$. For n = 5 and n = 10, frequent samples arose such that this iterative method either diverged or converged to a root other than the solution yielding the absolute maximum of L(x;m). However, for n = 15, the Newton-Raphson Method converged 1000 consecutive times to $\hat{m}$, the MLE. This result is consistent with Barnett's observation concerning the MLE being the root nearest the median. Therefore, for

$n \geq 15$, the Newton-Raphson Method, with the median as a starting value, appears to be a very safe method for finding the maximum likelihood estimator. It should be noted, however, that it is still possible for the Newton-Raphson method to diverge. The following table presents the number of times the Newton-Raphson Method diverged during the Monte Carlo simulation.

Table 11

Number of Divergences in Newton-Raphson Iteration (b known)

Sample Size	Number of Estimates	Number of Divergences
15	10,000	4
20	10,000	4
25	10,000	0
30	10,000	0
40	10,000	0
50	10,000	0

Therefore, for $n \geq 25$, there seems to be very little chance of the method diverging.

However, for $15 \leq n < 25$ an alternate method is required to handle the rare case when the Newton-Raphson iteration diverges. Also, for $n < 15$, an alternate procedure is required to simply find the root which actually maximizes $L(x;m)$.

The alternate procedure used was the method of false position, as suggested by Barnett. The m axis was scanned at an increment of .25, and the sign of $\frac{\partial \log L(x;m)}{\partial m}$ was noted. Each time a change in sign occurred, the method of false position was used to locate the root. When this scanning procedure terminated, the root of (1) that maximized $L(x;m)$ was selected. Barnett notes that all roots to (1) lie between min x_i and max x_i. Hence, we need scan only over this interval. It is always possible to miss a root in the scanning procedure. Barnett indicates that a scanning interval of .25 appears sufficiently small (when b = 1) to locate all roots.

The distribution of $\frac{(\hat{m} - m)}{b/\sqrt{n}}$ was obtained by a Monte Carlo simulation. For $n \geq 15$, the Newton-Raphson method was used with the method of false position as a backup in case of divergence. For n = 5 and n = 10, the method of false position was used exclusively. The Newton-Raphson Method is, of course, much more efficient in terms of computer time; and, hence, should be used whenever possible.

B. Location Parameter Known

The problem of determining $\hat{b}$, the MLE of b from the Cauchy model with m known, is a considerably simpler problem, as witnessed from the following theorem. (For convenience, let m = 0.)

THEOREM 9

Let L(x;b) be the likelihood function for the Cauchy distribution with m known. There exists a unique $\hat{b} > 0$ such that

$$\left.\frac{\partial \log L(x;b)}{\partial b}\right|_{b = \hat{b}} = 0.$$

Proof: L(x;b) is given by

$$L(x;b) = \prod_{i=1}^{n} \frac{1}{\pi b[1 + \frac{x_i^2}{b^2}]}$$

and

$$\log L(x;b) = \sum_{i=1}^{n} - \log[\pi b(1 + \frac{x_i^2}{b^2})],$$

or

$$\log L(x;b) = -n \log \pi - n \log b - \sum_{i=1}^{n} \log (1 + \frac{x_i^2}{b^2}).$$

Hence,

$$\frac{\partial \log L(x;b)}{\partial b} = -\frac{n}{b} + \frac{2}{b}\sum_{i=1}^{n} \frac{x_i^2}{b^2 + x_i^2}.$$

Therefore, we seek the solution of

$$g(b) = -n + 2\sum_{i=1}^{n} \frac{x_i^2}{b^2 + x_i^2} = 0.$$

Now $g(0) = n$, $g(\infty) = -n$, and

$$g'(b) = -2b \sum_{i=1}^{n} \frac{x_i^2}{(b^2 + x_i^2)^2} < 0 \text{ for all } b > 0.$$

Since $g(b)$ is continuous, it follows that there is a unique $\hat{b}$ such that $g(\hat{b}) = 0$; and, hence

$$\frac{\partial \log L(x;b)}{\partial b}\Big|_{b=\hat{b}} = 0.$$

Q.E.D.

Therefore, the problem of finding $\hat{b}$ when m is known is easier than finding $\hat{m}$ with b known. The equation

$$-n + 2\sum_{i=1}^{n} \frac{x_i^2}{b^2 + x_i^2}$$

is nonlinear in b. The Newton-Raphson Method was used to find $\hat{b}$. Any starting value that causes convergence is suitable since there is a single positive root.

C. Both Parameters Unknown

In the case where both parameters are unknown, we seek values for m and b that maximize

$$L(x;m,b) = \prod_{i=1}^{n} \frac{1}{\pi b[1 + (\frac{x_i - m}{b})^2]}.$$

For relative maxima we seek roots of

$$\frac{\partial \log L(x;m,b)}{\partial m} = \sum_{i=1}^{n} \frac{(x_i - m)}{b^2 + (x_i - m)^2} = 0 \tag{2}$$

$$\frac{\partial \log L(x;m,b)}{\partial b} = -n + 2\sum_{i=1}^{n} \frac{(x_i - m)^2}{b^2 + (x_i - m)^2} = 0. \tag{3}$$

It is possible for multiple pairs of roots to exist for (2) and (3). For instance, for $n = 2$, a sample of x_1 and $-1/x_1$ yield ($\hat{m} = 0$, $\hat{b} = 1$) and ($\hat{m} = (x_1 - \frac{1}{x_1})$, $\hat{b} = 1$) for solutions. It would appear that an iterative solution to (2) and (3) could converge to a solution that did not yield the absolute maximum of $L(x;m,b)$. If this were the situation, one would probably have to resort to a two-dimensional search procedure that would be so exorbitant in computer usage that maximum likelihood estimation would not be feasible for a Monte Carlo simulation. Fortunately, it appears that in actual numerical situations there is a unique solution to (2) and (3) for $n \geq 5$.

A study was made to compare the roots of (2) and (3) obtained from a Newton-Raphson iteration with the actual maximum likelihood estimators as obtained from a search procedure. For 107 samples of size 5, the Newton-Raphson Method converged to the MLE each time. In another study, 500 samples of size 5 were obtained. Twenty-five pair of starting values symmetrically located about the true parameter values were used to initiate the Newton-Raphson iteration. In 314 of the samples, more than one pair of starting values caused convergence of the iteration. In each instance the solution obtained was unique.

While it appears as if, in actual numerical samples, a unique solution exists to (2) and (3), the results of the previous paragraph indicate the strong dependence of the iterative method upon the starting values.

In the Monte Carlo simulation conducted, the starting values were obtained in the following manner: For the initial estimate of m, the median was selected. From the theorem of the preceding section, it follows that, for a given value of m, there exists a unique positive root of $\frac{\partial \log L}{\partial b} = 0$. Therefore, along the line m = median, b was incremented until $\partial \log L / \partial b$ changed signs. The first value of b after the change in sign was selected as the initial value of b. For reasons to be explained later, the increment size was the range of the sample divided by 1000. Choosing the initial estimates of m and b in this manner yielded the following results:

Table 12

Number of Divergences in Newton-Raphson Iteration (b unknown)

Sample Size	Number of Samples	No. of Times Divergence Occurred
5	40,000	102
10	30,000	68
15	20,000	47
20	20,000	70
25	20,000	70
30	20,000	90
40	20,000	131
50	16,000	146
75	12,000	158
100	9,000	160

In those cases where the method diverged, the MLE were obtained by a search technique.

The following theorem is of value in connection with the question of multiple roots:

THEOREM 10

If $X \sim \frac{1}{b} f(\frac{x-m}{b})$ and $\hat{m}^{(0)}$, $\hat{b}^{(0)}$ are initial estimates of the roots of $\frac{\partial \log L}{\partial m} = 0$, $\frac{\partial \log L}{\partial b} = 0$ such that $\frac{\hat{m}^{(0)}-m}{b}$, $\frac{\hat{b}^{(0)}}{b}$ are distributed independently of m and b; then, provided f is twice differentiable with respect to m and b, each successive approximation $\hat{m}^{(i)}$, $\hat{b}^{(i)}$ given by the Newton-Raphson iterative method possesses the property that $\frac{\hat{m}^{(i)}-m}{b}$, $\frac{\hat{b}^{(i)}}{b}$, and $\frac{\hat{m}^{(i)}-m}{\hat{b}^{(i)}}$ are distributed independently of m and b.

Proof: Let $y_1, y_2, \ldots, y_n$ be a sample of size n from a standardized distribution (m = 0, b = 1), and let $\hat{m}_0^{(0)}$, $\hat{b}_1^{(0)}$ be the initial estimate of the two parameters. Let $x_1, x_2, \ldots, x_n$ be the corresponding sample of size n from a distribution with location parameter m and scale parameter b obtained by letting $x_i = by_i + m$. Let $\hat{m}^{(0)}$, $\hat{b}^{(0)}$ be initial estimates of m,b such that

$$\frac{\hat{m}^{(0)}-m}{b} = \hat{m}_0^{(0)}, \quad \frac{\hat{b}^{(0)}}{b} = \hat{b}_1^{(0)}. \tag{4}$$

The first estimates of m and b given by the Newton-Raphson Iterative Method are

$$\begin{aligned} \hat{m}^{(1)} &= \hat{m}^{(0)} + \Delta m \\ \hat{b}^{(1)} &= \hat{b}^{(0)} + \Delta b. \end{aligned} \tag{5}$$

where Δm, Δb are given by

$$\frac{\partial \log L}{\partial m} + \frac{\partial^2 \log L}{\partial m^2}(\Delta m) + \frac{\partial^2 \log L}{\partial m \partial b}(\Delta b) = 0$$

$$\frac{\partial \log L}{\partial b} + \frac{\partial^2 \log L}{\partial b \partial m}(\Delta m) + \frac{\partial^2 \log L}{\partial^2 b}(\Delta b) = 0.$$

where the partial derivatives are evaluated at $\hat{m}^{(0)}$, $\hat{b}^{(0)}$. For convenience let $a_{11}=\partial^2 \log L/\partial m^2$; $a_{12}=\partial^2 \log L/\partial m \partial b$, $a_{21}=\partial^2 \log L/\partial b \partial m$, $a_{22}=\partial^2 \log L/\partial b^2$,, $b_1= -\partial \log L/\partial m$, $b_2= -\partial \log L/\partial b$. Again these functions are evaluated at $\hat{m}^{(0)}$, $\hat{b}^{(0)}$. Therefore, we have $a_{11}\Delta m + a_{12}\Delta b = b_1$, $a_{21}\Delta m + a_{22}\Delta b = b_2$, or

$$\begin{bmatrix}\Delta m\\ \Delta b\end{bmatrix} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{bmatrix} a_{22} & -a_{21} \\ -a_{12} & a_{11}\end{bmatrix}\begin{bmatrix} b_1 \\ b_2\end{bmatrix} \tag{6}$$

Let b_1^x denote b_1 for the x sample and b_1^y denote b_1 for the y sample. Now

$$b_1^x = -\sum_{i=1}^{n} \frac{1}{f(\frac{x_i - \hat{m}^{(0)}}{\hat{b}^{(0)}})} \frac{\partial f(\frac{x_i - \hat{m}^{(0)}}{\hat{b}^{(0)}})}{\partial m} (-\frac{1}{\hat{b}^{(0)}}) \tag{7}$$

$$b_1^y = -\sum_{i=1}^{n} \frac{1}{f(\frac{y_i - \hat{m}_0^{(0)}}{\hat{b}_1^{(0)}})} \frac{\partial f\left(\frac{y_i - \hat{m}_0^{(0)}}{\hat{b}_1^{(0)}}\right)}{\partial m} (-\frac{1}{\hat{b}_1^{(0)}}) \tag{8}$$

Substituting $x_i = by_1 + m$, we have

$$b_1^x = bb_1^y$$

Similarly,

$$\begin{aligned} b_2^x &= bb_2^y & a_{21}^x &= b^2 a_{21}^y \\ a_{11}^x &= b^2 a_{11}^y & a_{22}^x &= b^2 a_{22}^y \\ a_{12}^x &= b^2 a_{12}^y \end{aligned} \tag{9}$$

Now

$$\begin{aligned} \hat{m}^{(1)} &= \hat{m}^{(0)} + \Delta m \\ \hat{b}^{(1)} &= \hat{m}^{(0)} + \Delta b \end{aligned} \tag{10}$$

Substituting (9) and (6) into (10) yields

$$\hat{m}_x^{(1)} = b\hat{m}_y^{(1)} + m$$

$$\hat{b}_x^{(1)} = b\hat{m}_y^{(1)}$$

where the subscript x and y refer to the sample from which the estimate was generated.

Therefore, the distributional property has been maintained. Hence, by induction,

$$\hat{m}_x^{(k)} = b\hat{m}_y^{(k)} + m$$

$$\hat{b}_x^{(k)} = b\hat{b}_y^{(k)}$$

Q.E.D.

Now, by selecting the starting values as described earlier, we have the desired properties of this theorem. Therefore, even if, in some rare sample, one obtained a solution to (2) and (3) that was not the MLE, the solution m* and b* obtained would have the property that

$$\frac{m^* - m}{b}, \quad \frac{m^* - m}{b^*}, \quad \frac{b^*}{b}$$

are distributed independent of m and b. Hence, if one finds the roots to (2) and (3) in the same manner that was used to generate Tables A2 and A3, then these tables are still valid for setting confidence intervals, testing hypothesis, etc. A computer program is provided in Appendix B for use in finding the roots to (2) and (3).

It should be noted that the previous theorem does not hold for the Method of Steepest Descent.

VIII THE MONTE CARLO SIMULATION

A. Generation of Cauchy Random Variables

It is well known [22] that if X is a continuous random variable with cumulative distribution function F(x), then the random variable Z = F(X) has a uniform distribution over the interval (0,1). If X has a Cauchy distribution with m = 0, b = 1, then the cumulative distribution function is given by $F(x) = \frac{1}{2} + \frac{1}{\pi} \tan^{-1} x$. Hence, $Z = \frac{1}{2} + \frac{1}{\pi} \tan^{-1} X$ has a uniform distribution.

The UMR Computer Center provides a subroutine on their IBM 360 Model 50 computer for generating random samples from a uniform distribution over the interval (0,1). If z_i, i = 1, . . .,n, represents a sample of n observations from the uniform distribution, then it follows that $x_i = \tan[\pi(z_i - 1/2]$, i = 1, . . .,n represents a random sample from a standard Cauchy distribution. Consequently, $y_i = m + bx_i$ represents a sample from a Cauchy distribution with parameters m and b. All random samples used in this study were generated in this manner.

B. Comparison of an Exact Distribution with a Simulated Distribution

Random samples from a standard Cauchy distribution were generated in the manner described in part A. The maximum likelihood estimators were then calculated as discussed in Chapter VII. This process was repeated a large number of times, and the cumulative distributions of various required functions of the estimators were tabulated.

The simulated distributions might differ from the true (but unknown) distributions because of two kinds of errors. They are: (1) the random samples generated on the computer were not representative of a uniform distribution, and (2) not enough estimates of the parameters were obtained to give a precise simulated distribution.

It is possible to gain some insight into the magnitude of these errors by obtaining the simulated distribution of a random variable whose true distribution is known. For a sample of size 2k + 1, the median of a Cauchy distribution, which is the k + 1 ordered observation, has a probability density function given by

$$g(y_{k+1}) = \frac{(2k + 1)!}{(k)!(k)!} [\frac{1}{2} + \frac{1}{\pi} \tan^{-1} y_{k+1}]^k [\frac{1}{2} - \frac{1}{\pi} \tan^{-1} y_{k+1}]^k \frac{1}{\pi(1 + y_{k+1}^2)}.$$

The cumulative distribution can be obtained by numerical integration.

The cumulative distribution of the median for samples of size 5 and 9 were obtained by simulation and by numerical integration. These results appear in Table 13.

From the excellent agreement displayed in Table 13, there can be little doubt concerning the randomness of the random number generator. Since the variance of the MLE of m is less than the variance of the median, there is no reason to suspect that the simulated distributions of $\hat{m}$ are not at least as precise as the simulated distributions of the median. While it is difficult to make exact statements about the other random variables, it is felt that the errors present are of about the same order of magnitude.

Table 14 presents the number of estimates used for all random variables whose distributions were obtained by Monte Carlo simulation.

C. Smoothing of Critical Values

The sample sizes that were used in the Monte Carlo simulation appear in Table 14. For the distributions of $\frac{\hat{m} - m}{\hat{b}/\sqrt{n}}$, $\frac{\hat{b}}{b}$ (m unknown), $\frac{\hat{m} - m}{b/\sqrt{n}}$, and $\frac{\hat{b}}{b}$ (m known) continuous curves were fit by least squares to the critical values for the purpose of smoothing the critical values and then interpolating for sample sizes not run in the simulation.

Table 13

P[X < x] for the Sample Median of a Standard Cauchy Distribution

	2k + 1 = 5 10,000 Estimates		2k + 1 = 5 20,000 Estimates		2k + 1 = 5 10,000 Estimates	
x	Simulated	Exact	Simulated	Exact	Simulated	Exact
.1	.5617	.5593	.5609	.5593	.5792	.5777
.2	.6214	.6166	.6168	.6166	.6519	.6514
.3	.6707	.6700	.6702	.6700	.7179	.7181
.4	.7207	.7184	.7186	.7184	.7782	.7759
.5	.7647	.7611	.7593	.7611	.8234	.8242
.6	.8000	.7980	.7972	.7980	.8638	.8632
.7	.8311	.8294	.8282	.8294	.8930	.8942
.8	.8583	.8559	.8541	.8559	.9158	.9183
.9	.8798	.8780	.8762	.8780	.9340	.9386
1.0	.8982	.8965	.8949	.8965	.9497	.9511
1.1	.9131	.9118	.9113	.9118	.9630	.9619
1.2	.9256	.9246	.9244	.9246	.9714	.9702
1.3	.9364	.9352	.9338	.9352	.9771	.9766
1.4	.9454	.9440	.9423	.9440	.9819	.9815
1.5	.9526	.9515	.9506	.9515	.9859	.9852
1.6	.9593	.9577	.9570	.9577	.9892	.9881
1.7	.9847	.9630	.9620	.9630	.9907	.9904
1.8	.9687	.9675	.9671	.9675	.9930	.9922
1.9	.9720	.9713	.9712	.9713	.9941	.9936
2.0	.9752	.9745	.9742	.9745	.9955	.9948
2.1	.9788	.9774	.9769	.9774		
2.2	.9813	.9798	.9791	.9798		
2.3	.9832	.9819	.9815	.9819		
2.4	.9848	.9837	.9838	.9837		
2.5	.9862	.9853	.9855	.9853		
2.6	.9876	.9867	.9870	.9867		
2.7	.9887	.9879	.9881	.9879		
2.8	.9879	.9890	.9891	.9890		
2.9	.9906	.9899	.9900	.9899		
3.0	.9913	.9908	.9909	.9908		

Table 14

Number of Estimates Obtained in Monte Carlo Simulation

RANDOM VARIABLE	Sample Size									
	5	10	15	20	25	30	40	50	75	100
$\frac{\hat{m} - m}{\hat{b}/\sqrt{n}}$	40,000	30,000	20,000	20,000	20,000	20,000	20,000	16,000	12,000	9,000
$\frac{\hat{b}}{b}$(m unknown)	40,000	30,000	20,000	20,000	20,000	20,000	20,000	16,000	12,000	9,000
$\frac{\hat{m} - m}{b/\sqrt{n}}$	20,000	10,000	10,000	10,000	10,000	10,000	10,000	10,000	---	---
$\frac{\hat{b}}{b}$ (m known)	20,000	10,000	10,000	10,000	10,000	10,000	10,000	8,000	6,000	4,500
$\frac{\hat{m} - m}{\hat{b}} + k(\beta)\ \frac{b}{\hat{b}}$	40,000	20,000	20,000	20,000	20,000	20,000	20,000	16,000	12,000	9.000
$\frac{b(-1-\sqrt{1+\tan^2\pi\beta(1+(\frac{\hat{m}-m}{b})^2)})}{\hat{b}\tan\pi\beta}$	40,000	30,000	20,000	20,000	20,000	20,000	20,000	16,000	12,000	9,000
$\frac{\hat{m} - m}{b} - \frac{c}{\sqrt{n}}\frac{\hat{b}}{b}$	20,000	10,000	10,000	10,000	10,000	10,000	10,000	8,000	6,000	4,500
λ (Likelihood Ratio Test)	10,000	8,000	6,000	---	5,000	n = 35 2,000	---	1,500	---	---
$\frac{\hat{b}_2}{b_2} \cdot \frac{b_1}{\hat{b}_1}$ m_1, m_2 unknown	15,000	10,000	8,000	8,000	---	6,000	n = 45 4,000	---	---	---

Table 14 (continued)

RANDOM VARIABLE	Sample Size									
	5	10	15	20	25	30	40	50	75	100
$\dfrac{(\hat{m}_1-m_1) - (\hat{m}_2-m_2)}{\hat{b}_1 + \hat{b}_2}$	15,000	10,000	8,000	8,000	----	6,000	n = 45 4,000	----	----	----
$\dfrac{(\hat{m}_1-m_1)}{b_1} - \dfrac{b_2}{b_1}\left(\dfrac{\hat{m}_2-m_2}{b_2}\right)$	20,000	15,000	10,000	10,000	----	9,000	6,800	----	----	----

Random Variable $\dfrac{\hat{b}_2}{b_2} \cdot \dfrac{b_1}{\hat{b}_1}$ m_1, m_2 known

n_1 \ n_2	5	10	15	20	25	30	40
5	20,000	20,000	20,000				
10	20,000	10,000	10,000	10,000			
15	10,000	10,000	10,000	10,000	10,000		
20		10,000	10,000	10,000	10,000	10,000	
25			10,000	10,000	10,000	10,000	8,400
30				10,000	10,000	9.000	7,800
40					8,400	7,800	6,800

Let y_γ be the γ cumulative percentage point of the statistic being considered. For the critical values of $\frac{\hat{m} - m}{\hat{b}/\sqrt{n}}$ and $\frac{m - m}{b/\sqrt{n}}$ the following smoothing model was used

$$y_\gamma = \beta_o + \frac{\beta_1}{\beta_2 + n} + \varepsilon$$

where β_o is the asymptotic value of the γ percentage point derived from the normal distribution. The parameters β_1 and β_2 were estimated using the least squares criteria. The model is nonlinear in β_1 and β_2, and the Gauss-Newton Method was used to estimate β_1 and β_2. This model was then used to obtain the critical values of Tables A2 and A4.

For the statistics $\frac{\hat{b}}{b}$ (m unknown), $\frac{\hat{b}}{b}$ (m known), two different smoothing models were used. For $\gamma < .5$, the following model was used:

$$y_\gamma = \beta_o + \beta_1 n + \beta_2 \ln(n-1) + \varepsilon.$$

For $.5 < \gamma < 1$, the model $y_\gamma = \beta_o + \beta_1 n + \beta_1 n^{-\frac{1}{2}} + \beta_3 n^{-1} + \varepsilon$ was used. Both models are linear in the unknown parameters, and the parameters were estimated using standard linear least squares procedures.

These smoothing models were selected among several tried because they appeared to adequately fit the unsmoothed data. For purposes of comparison of the smoothed and unsmoothed critical values, the unsmoothed data from the simulation are now presented. The tables have the same format as Tables A2, A3, A4, and A5.

Table 15

Unsmoothed Critical Values of $\frac{\hat{m} - m}{\hat{b}/\sqrt{n}}$

n \ 1−α	.75	.80	.85	.90	.95	.975	.99
5	1.370	1.78	2.34	3.15	4.77	6.78	11.0
10	1.120	1.42	1.79	2.27	3.11	3.95	5.19
15	1.060	1.330	1.66	2.10	2.79	3.46	4.40
20	1.030	1.29	1.61	2.00	2.63	3.27	4.04
25	1.020	1.28	1.58	1.99	2.60	3.13	3.85
30	1.010	1.26	1.56	1.93	2.51	3.05	3.76
40	.987	1.25	1.54	1.91	2.48	2.97	3.63
50	.976	1.22	1.51	1.88	2.45	2.96	3.59
75	.969	1.20	1.46	1.85	2.40	2.89	3.47
100	.95	1.20	1.48	1.82	2.38	2.85	3.34

Table 16

Unsmoothed Critical Values of $\frac{\hat{b}}{b}$ (m unknown)

n \ 1−α	.01	.025	.05	.1	.9	.95	.975	.99
5	.084	.130	.181	.259	1.928	2.560	3.277	4.525
10	.256	.320	.387	.467	1.662	2.005	2.353	2.838
15	.357	.418	.479	.563	1.522	1.746	1.970	2.279
20	.424	.488	.546	.619	1.444	1.628	1.811	2.067
25	.473	.533	.583	.651	1.397	1.536	1.708	1.911
30	.507	.568	.621	.685	1.357	1.498	1.635	1.796
40	.573	.622	.670	.727	1.305	1.412	1.525	1.656
50	.610	.656	.702	.755	1.269	1.366	1.463	1.566
75	.669	.710	.751	.797	1.217	1.289	1.358	1.438
100	.712	.746	.779	.822	1.187	1.251	1.305	1.372

Table 17

Unsmoothed Critical Values of $\frac{\hat{m} - m}{b/\sqrt{n}}$

n \ 1−α	.8	.85	.90	.95	.975	.99
5	1.24	1.60	2.09	3.03	4.18	6.07
10	1.22	1.53	1.94	2.67	3.36	4.33
15	1.21	1.50	1.91	2.52	3.18	3.97
20	1.20	1.50	1.86	2.45	3.03	3.78
25	1.20	1.50	1.86	2.41	2.95	3.66
30	1.20	1.48	1.85	2.40	2.92	3.55
40	1.21	1.49	1.85	2.38	2.88	3.50
50	1.18	1.47	1.84	2.37	2.87	3.50

Table 18

Unsmoothed Critical Values of $\frac{\hat{b}}{b}$ (m known)

n \ 1−α	.01	.025	.05	.10	.90	.95	.975	.99
5	.203	.268	.337	.440	2.273	2.932	3.777	5.035
10	.333	.406	.471	.560	1.794	2.129	2.455	2.945
15	.408	.473	.541	.621	1.598	1.834	2.062	2.410
20	.463	.532	.597	.669	1.402	1.689	1.867	2.117
25	.508	.567	.623	.694	1.436	1.607	1.760	1.958
30	.542	.599	.656	.718	1.387	1.536	1.673	1.826
40	.590	.646	.694	.752	1.333	1.435	1.543	1.685
50	.625	.674	.721	.776	1.288	1.389	1.484	1.604
75	.674	.724	.756	.805	1.219	1.299	1.370	1.460
100	.713	.761	.790	.829	1.191	1.267	1.328	1.383

IX SUMMARY, CONCLUSIONS AND FURTHER PROBLEMS

Various estimators of the location and scale parameters in the Cauchy distribution have been investigated, and the superiority of the maximum likelihood estimators has been established. It has been shown that, with the aid of a digital computer program, it is not difficult to obtain the maximum likelihood estimators. A digital computer program has been provided for use in finding these estimators. Tables based on maximum likelihood estimators have been prepared by Monte Carlo simulation to aid in performing basic statistical analysis for the Cauchy distribution.

The applied statistician may now consider using the Cauchy distribution as a probability model. The Cauchy distribution should be an appealing model for experimental situations where outlying observations are frequently encountered. Test procedures and tables are provided for discriminating between the normal distribution and the Cauchy distribution for given samples. An example is provided based on real data where the test indicates that the Cauchy distribution should be used in preference to the normal distribution. The difference in the estimates of the location parameter is observed.

The work in this paper has dealt specifically with the Cauchy distribution; however, most of the theorems presented apply to any probability distribution involving location and scale parameters. Hence, the methods apply to distributions such as the log-normal and censored normal, and also to distributions such as the Weibull where a transformation on the random variable will produce a distribution involving a location and scale parameter.

The distributions required in the two-sample problem with both parameters unknown have not been extensively determined. A computer program is provided to generate these distributions should the occasion arise. The same holds for type II censored samples.

The question of robustness has not been investigated thoroughly enough to be presented here. Part of the problem in comparing the Cauchy distribution versus the normal distribution lies in the fact that if the distribution is Cauchy, then the normal estimators are meaningless. Perhaps a study applying both the Cauchy methods and normal methods to some intermediate distribution such as the t distribution would prove meaningful. It would also be of interest to see how the Cauchy estimators compared to the normal estimators in those cases where the sample was normal but the likelihood ratio test selected the Cauchy distribution.

REFERENCES

1. Cramer, H. (1946), Mathematical Methods of Statistics, Princeton University Press, Princeton, New Jersey.

2. Mood, A.M. and Graybill, F.A. (1963), Introduction to the Theory of Statistics, McGraw-Hill Book Co., Inc., New York, New York.

3. Lindgren, B.W. (1962), Statistical Theory, The Macmillan Company, New York, New York.

4. Feller, W. (1966), An Introduction to Probability Theory and Its Applications, Volume II, John Wiley & Sons, Inc., New York, New York.

5. Hald, A. (1952), Statistical Theory with Engineering Applications, John Wiley & Sons, New York, New York.

6. Weatherburn, C.E. (1946), A First Course in Mathematical Statistics, Cambridge University Press, Cambridge, England.

7. Dixon, W.J. and Massey, F.J. (1969), Introduction to Statistical Analysis, McGraw-Hill Book Co., New York, New York.

8. Fowler, J. (1968) Personal communication. April 1968, Rolla, Missouri.

9. Kendall, M.G. & Stuart, A. (1967), The Advanced Theory of Statistics, Vol. II, Second Edition, Hafner Publishing Co., New York, New York.

10. Bhattacharyya, A. (1947), On Some Analogues of the Amount of Information and Their Use in Statistical Estimation, Sankhyā, 8, 205.

11. Rider, P.R. (1960), Variance of the Median of Samples from a Cauchy Distribution, J. American Statistical Association, 55, 322-323.

12. Rothenberg, T.J., Fisher, F.M., and Tilanus, C.F. (1964), A Note on Estimation from a Cauchy Sample, J. American Statistical Association, 59, 460-463.

13. Bloch, D. (1966), A Note on the Estimation of the Location Parameter of the Cauchy Distribution, J. American Statistical Association, 61, 852-855.

14. Barnett, V.D. (1966), Evaluation of the Maximum Likelihood Estimator Where the Likelihood Equation has Multiple Roots, Biometrika, 53, 151.

15. Barnett, V.D. (1966), Order Statistics Estimators of the Location of the Cauchy Distribution, J. American Statistical Association, 61, 1205-1218.

16. Sarhan & Greenberg (1962), Contributions to Order Statistics, John Wiley & Sons, Inc., New York, New York.

17. Antle, C.E. and Bain, L.J. (1969), A property of Maximum Likelihood Estimators of Location and Scale Parameters, SIAM Review, (To Appear, April 1969).

18. Thoman, D.R. (1968), Inferences on the Parameters of the Weibull Distribution, Thesis, University of Missouri at Rolla, 89 p. (with seven figures, 32 tables).

19. Kale, B.K. (1961), On the Solution of the Likelihood Equation by Iterative Processes, Biometrika, 48, 452-456.

20. Kale, B.K. (1962), On the Solution of Likelihood Equations by Iteration Processes. The Multiparametric Case., Biometrika, 49, 479-486.

21. Hildebrand, F.B. (1956), Introduction to Numerical Analysis, McGraw-Hill Book Co., Inc., New York, New York.

22. Hogg, R.V. and Craig, A.T. (1965), Introduction to Mathematical Statistics, The Macmillan Co., New York, New York.

23. Dumonceaux, R.H. (1969), Estimation of Parameters in Censored Samples, Thesis, University of Missouri at Rolla.

24. Brownlee, K.A. (1960), Statistical Theory and Methodology in Science and Engineering, John Wiley & Sons, Inc., New York, New York.

APPENDIX A

TABLES

Table A1

Cumulative Values of the Standard Cauchy Distribution

$$F(x) = \int_{-\infty}^{x} \frac{dy}{\pi(1+y^2)} = P[X \leq x]*$$

x	.00	.20	.40	.60	.80
0	.50000	.56283	.62112	.67202	.71478
1	.75000	.77886	.80257	.82219	.83858
2	.85242	.86420	.87433	.88312	.89081
3	.89758	.90359	.90895	.91375	.91809
4	.92202	.92560	.92886	.93186	.93462
5	.93717	.93952	.94171	.94375	.94565
6	.94743	.94910	.95066	.95212	.95352
7	.95483	.95607	.95724	.95836	.95941
8	.96042	.96137	.96228	.96315	.96398
9	.96478	.96554	.96626	.96696	.96763
10	.96827	.96889	.96949	.97006	.97061
11	.97114	.97165	.97215	.97263	.97309
12	.97353	.97397	.97438	.97479	.97518
13	.97556	.97593	.97629	.97664	.97697
14	.97730	.97762	.97793	.97823	.97852
15	.97881	.97909	.97936	.97962	.97988
16	.98013	.98038	.98061	.98085	.98108
17	.98130	.98151	.98173	.98193	.98214
18	.98233	.98253	.98272	.98290	.98308
19	.98326	.98344	.98361	.98377	.98394
20	.98410	.98425	.98441	.98456	.98471
21	.98485	.98500	.98514	.98527	.98541
22	.98554	.98567	.98580	.98592	.98605
23	.98616	.98629	.98640	.98652	.98663
24	.98674	.98685	.98696	.98707	.98717
25	.98727	.98737	.98747	.98757	.98767
26	.98776	.98786	.98795	.98804	.98813
27	.98822	.98830	.98839	.98847	.98855
28	.98864	.98872	.98880	.98887	.98895
29	.98903	.98910	.98918	.98925	.98932
30	.98939	.98946	.98953	.98960	.98967
31	.98973	.98980	.98987	.98993	.98999
32	.99006	.99012	.99018	.99024	.99030
33	.99036	.99041	.99047	.99053	.99058
34	.99064	.99069	.99075	.99080	.99085
35	.99091	.99096	.99101	.99106	.99111
36	.99116	.99121	.99126	.99130	.99135
37	.99140	.99144	.99149	.99154	.99158
38	.99162	.99167	.99171	.99175	.99180
39	.99184	.99188	.99192	.99196	.99200
40	.99204	.99208	.99212	.99216	.99220
41	.99224	.99227	.99231	.99235	.99239
42	.99242	.99246	.99249	.99253	.99256
43	.99260	.99263	.99267	.99270	.99273
44	.99277	.99280	.99283	.99286	.99290
45	.99293	.99296	.99299	.99302	.99305

* t distribution with 1 degree of freedom.

Table A2

Cumulative Percentage Points of $\frac{(\hat{m} - m)}{\hat{b}/\sqrt{n}}$

Table of $c_{1-\alpha,n}$ such that $P[\frac{(\hat{m} - m)}{\hat{b}/\sqrt{n}} \leq c_{1-\alpha,n}] = 1 - \alpha$.

n \ 1-α	.75	.80	.85	.90	.95	.975	.99
5	1.370	1.780	2.341	3.150	4.771	6.780	11.000
6	1.275	1.639	2.112	2.782	4.030	5.474	8.075
7	1.215	1.552	1.979	2.572	3.634	4.810	6.759
8	1.174	1.494	1.891	2.437	3.388	4.408	6.011
9	1.144	1.451	1.829	2.343	3.220	4.138	5.529
10	1.121	1.419	1.783	2.273	3.098	3.944	5.191
12	1.089	1.374	1.719	2.177	2.933	3.685	4.751
15	1.059	1.332	1.661	2.090	2.786	3.458	4.374
17	1.045	1.314	1.635	2.051	2.722	3.360	4.215
20	1.030	1.293	1.607	2.010	2.654	3.256	4.048
22	1.023	1.283	1.593	1.990	2.620	3.205	3.967
25	1.014	1.271	1.577	1.966	2.581	3.146	3.873
27	1.009	1.265	1.569	1.953	2.561	3.115	3.824
30	1.003	1.257	1.558	1.938	2.535	3.077	3.764
32	.999	1.252	1.552	1.929	2.521	3.056	3.730
35	.996	1.247	1.545	1.918	2.503	3.029	3.689
37	.993	1.243	1.540	1.912	2.493	3.014	3.665
40	.990	1.239	1.535	1.904	2.480	2.994	3.634
45	.986	1.233	1.527	1.893	2.462	2.967	3.593
50	.983	1.229	1.521	1.884	2.448	2.946	3.560
55	.980	1.225	1.516	1.877	2.437	2.929	3.534
60	.978	1.222	1.512	1.871	2.428	2.916	3.513
65	.976	1.220	1.509	1.866	2.420	2.904	3.495
70	.974	1.218	1.506	1.862	2.413	2.894	3.479
75	.973	1.216	1.504	1.859	2.407	2.885	3.466
80	.972	1.214	1.501	1.856	2.402	2.898	3.455
85	.971	1.213	1.500	1.853	2.398	2.871	3.445
90	.970	1.211	1.498	1.850	2.394	2.865	3.436
95	.969	1.210	1.496	1.848	2.391	2.860	3.428
100	.968	1.209	1.495	1.846	2.388	2.856	3.421
∞	.954	1.190	1.470	1.810	2.330	2.770	3.290

Table A3

Cumulative Percentage Points of $\frac{\hat{b}}{b}$ (m unknown)

Values of $c_{1-\alpha,n}$ such that $P[\frac{\hat{b}}{b} \leq c_{1-\alpha,n}] = 1 - \alpha$.

n \ 1-α	.01	.025	.05	.10	.90	.95	.975	.99
5	.080	.132	.189	.255	1.928	2.563	3.279	4.524
6	.131	.184	.242	.324	1.849	2.383	2.971	3.953
7	.172	.227	.284	.375	1.786	2.249	2.747	3.551
8	.206	.262	.320	.415	1.734	2.144	2.576	3.253
9	.236	.293	.351	.447	1.691	2.058	2.441	3.023
10	.262	.320	.378	.474	1.654	1.988	2.332	2.842
12	.306	.365	.423	.516	1.594	1.877	2.166	2.572
15	.359	.418	.475	.563	1.526	1.760	1.994	2.307
17	.387	.447	.504	.587	1.491	1.701	1.912	2.184
20	.424	.483	.540	.617	1.449	1.632	1.816	2.048
22	.444	.504	.560	.633	1.426	1.595	1.766	1.978
25	.472	.531	.586	.654	1.396	1.548	1.705	1.894
27	.488	.547	.602	.666	1.379	1.522	1.670	1.848
30	.510	.568	.622	.682	1.357	1.489	1.627	1.792
32	.523	.581	.634	.691	1.344	1.470	1.602	1.759
35	.541	.598	.650	.704	1.327	1.445	1.569	1.718
37	.552	.608	.660	.712	1.317	1.430	1.550	1.693
40	.567	.622	.673	.722	1.303	1.410	1.525	1.661
45	.589	.643	.692	.737	1.284	1.382	1.489	1.616
50	.608	.660	.708	.750	1.267	1.360	1.459	1.579
55	.624	.674	.721	.762	1.254	1.341	1.434	1.547
60	.639	.687	.732	.772	1.242	1.325	1.412	1.519
65	.652	.698	.741	.781	1.232	1.311	1.393	1.494
70	.663	.707	.749	.789	1.223	1.299	1.377	1.471
75	.674	.715	.755	.797	1.215	1.289	1.362	1.450
80	.683	.722	.761	.803	1.209	1.280	1.348	1.431
85	.691	.728	.765	.809	1.203	1.272	1.336	1.412
90	.698	.733	.768	.815	1.197	1.265	1.325	1.395
95	.705	.738	.771	.821	1.193	1.259	1.314	1.379
100	.711	.741	.773	.825	1.189	1.254	1.305	1.362
Asym. 100	.671	.723	.768	.818	1.182	1.232	1.277	1.329

Table A4

Cumulative Percentage Points of $\frac{(\hat{m} - m)}{b/\sqrt{n}}$

Table of $c_{1-\alpha,n}$ such that $P[\frac{(\hat{m} - m)}{b/\sqrt{n}} \leq c_{1-\alpha,n}] = 1 - \alpha$.

n \ 1−α	.8	.85	.90	.95	.975	.99
5	1.241	1.600	2.090	3.036	4.182	6.069
6	1.233	1.573	2.039	2.884	3.871	5.386
7	1.228	1.555	2.004	2.786	3.672	4.973
8	1.223	1.543	1.979	2.717	3.534	4.696
9	1.220	1.533	1.959	2.666	3.433	4.497
10	1.217	1.526	1.943	2.627	3.355	4.347
12	1.213	1.516	1.920	2.572	3.244	4.137
15	1.209	1.506	1.898	2.518	3.139	3.943
17	1.206	1.501	1.887	2.494	3.091	3.856
20	1.204	1.496	1.875	2.468	3.039	3.762
22	1.203	1.494	1.869	2.454	3.013	3.715
25	1.201	1.491	1.862	2.439	2.982	3.660
27	1.201	1.489	1.858	2.430	2.966	3.630
30	1.200	1.487	1.853	2.420	2.945	3.594
32	1.199	1.486	1.850	2.414	2.934	3.574
35	1.198	1.485	1.847	2.406	2.919	3.548
37	1.198	1.484	1.845	2.402	2.911	3.533
40	1.197	1.483	1.842	2.397	2.900	3.514
45	1.196	1.481	1.839	2.389	2.885	3.488
50	1.196	1.480	1.836	2.383	2.873	3.467
55	1.195	1.479	1.833	2.378	2.863	3.451
60	1.195	1.478	1.831	2.374	2.855	3.437
65	1.194	1.478	1.830	2.370	2.849	3.425
70	1.194	1.477	1.828	2.367	2.843	3.415
75	1.194	1.477	1.827	2.365	2.838	3.407
80	1.194	1.476	1.826	2.363	2.834	3.399
85	1.193	1.476	1.825	2.361	2.830	3.393
90	1.193	1.476	1.824	2.359	2.826	3.387
95	1.193	1.475	1.823	2.357	2.823	3.382
100	1.193	1.475	1.823	2.356	2.820	3.377
∞	1.190	1.470	1.810	2.330	2.770	3.290

Table A5

Cumulative Percentage Points of $\frac{\hat{b}}{b}$ (m known)

Values of $c_{1-\alpha,n}$ such that $P[\frac{\hat{b}}{b} \le c_{1-\alpha,n}] = 1 - \alpha$

n \ 1−α	.01	.025	.05	.10	.90	.95	.975	.99
5	.196	.266	.337	.441	2.273	2.930	3.773	5.026
6	.236	.305	.375	.474	2.119	2.670	3.333	4.324
7	.268	.336	.405	.501	2.005	2.480	3.023	3.835
8	.295	.363	.431	.524	1.917	2.335	2.793	3.478
9	.319	.386	.453	.543	1.846	2.221	2.616	3.206
10	.339	.406	.472	.560	1.788	2.129	2.475	2.992
12	.374	.440	.505	.589	1.698	1.987	2.267	2.681
15	.416	.480	.544	.623	1.604	1.842	2.062	2.383
17	.439	.502	.565	.641	1.557	1.771	1.967	2.247
20	.468	.530	.591	.664	1.502	1.690	1.861	2.099
22	.484	.546	.606	.677	1.473	1.648	1.806	2.025
25	.507	.567	.626	.695	1.436	1.595	1.741	1.937
27	.520	.580	.637	.705	1.416	1.566	1.706	1.890
30	.538	.597	.653	.718	1.390	1.529	1.661	1.831
32	.548	.607	.662	.726	1.374	1.508	1.636	1.799
35	.563	.621	.675	.737	1.355	1.481	1.603	1.756
37	.572	.630	.683	.743	1.343	1.465	1.584	1.732
40	.585	.642	.693	.752	1.327	1.443	1.558	1.699
45	.603	.659	.708	.765	1.304	1.413	1.522	1.654
50	.619	.674	.721	.776	1.286	1.388	1.492	1.615
55	.634	.688	.732	.785	1.270	1.367	1.466	1.582
60	.647	.699	.742	.793	1.256	1.349	1.444	1.552
65	.658	.708	.750	.799	1.244	1.333	1.423	1.525
70	.668	.719	.757	.805	1.234	1.319	1.405	1.499
75	.679	.728	.764	.810	1.224	1.307	1.387	1.476
80	.686	.735	.769	.814	1.216	1.296	1.371	1.453
85	.694	.742	.774	.818	1.208	1.286	1.356	1.431
90	.702	.749	.778	.821	1.201	1.277	1.342	1.410
95	.708	.754	.782	.824	1.195	1.269	1.328	1.389
100	.714	.759	.785	.826	1.189	1.262	1.315	1.368
Asym. 100	.671	.723	.768	.818	1.182	1.232	1.277	1.329

Table A6

One-sided Tolerance Limits (both parameters unknown)

Factors $z(\beta,\gamma,n)$ such that $\hat{m} - z(\beta,\gamma,n)\hat{b}$ is a lower one-sided β,γ tolerance limit or $\hat{m} + z(\beta,\gamma,n)\hat{b}$ is an upper one-sided β,γ tolerance limit.

n	β = .8			β = .9			β = .95			β = .99		
	γ			γ			γ			γ		
	.9	.95	.99	.9	.95	.99	.9	.95	.99	.9	.95	.99
5	5.595	8.169	17.553	12.073	17.314	36.870	24.706	35.42	74.34	124.60	176.10	---
10	3.191	3.964	5.810	6.744	8.244	12.283	13.686	16.72	24.79	68.50	83.50	123.0
15	2.632	3.106	4.078	5.583	6.492	8.571	11.320	13.13	17.35	56.73	65.75	87.9
20	2.377	2.737	3.466	5.066	5.756	7.261	10.276	11.64	14.73	51.57	58.38	74.6
25	2.226	2.525	3.142	4.763	5.335	6.577	9.667	10.80	13.36	48.57	54.19	66.5
30	2.125	2.383	2.931	4.560	5.055	6.137	9.260	10.24	12.47	46.55	51.42	61.9
40	1.995	2.202	2.654	4.298	4.694	5.574	8.737	9.53	11.33	43.95	47.88	55.9
50	1.914	2.088	2.472	4.134	4.465	5.210	8.410	9.09	10.59	42.32	45.64	52.1
75	1.800	1.926	2.192	3.900	4.135	4.661	7.945	8.45	9.46	39.98	42.44	47.9
100	1.739	1.838	2.028	3.775	3.954	4.343	7.694	8.10	8.81	38.71	40.68	44.9
Asym. 100	1.741	1.870	2.160	3.789	4.048	4.640	7.728	8.25	9.44	38.87	41.47	47.4

Table A7

Two-sided Tolerance Limits (both parameters unknown)

Factors $w(\beta,\gamma,n)$ such that $(\hat{m} - w(\beta,\gamma,n)\hat{b},\ \hat{m} + w(\beta,\gamma,n)\hat{b})$ forms a two-sided β,γ tolerance interval

n	β = .8			β = .9			β = .95			β = .99		
	γ .9	γ .95	γ .99	γ .9	γ .95	γ .99	γ .9	γ .95	γ .99	γ .9	γ .95	γ .99
5	12.296	17.686	38.320	24.756	35.51	76.86	49.41	71.21	153.90	245.7	357.3	---
10	6.730	8.172	12.365	13.654	16.57	25.17	27.34	33.35	50.50	136.2	166.5	---
15	5.535	6.418	8.563	11.299	13.07	17.51	22.67	26.32	35.18	113.2	131.5	178.0
20	5.023	5.698	7.243	10.263	11.63	14.83	20.62	23.42	29.80	103.1	117.1	150.0
25	4.725	5.289	6.556	9.659	10.81	13.42	19.41	21.76	26.98	97.2	108.8	134.8
30	4.525	5.016	6.112	9.254	10.26	12.52	18.60	20.65	25.16	93.2	103.3	125.8
40	4.267	4.663	5.538	8.732	9.54	11.35	17.56	19.21	22.81	88.0	96.1	111.0
50	4.104	4.438	5.163	8.405	9.08	10.58	16.90	18.29	21.27	84.7	91.5	104.4
75	3.872	4.110	4.588	7.938	8.42	9.42	15.96	16.95	18.93	79.9	84.8	95.4
100	3.746	3.928	4.251	7.685	8.05	8.74	15.46	16.21	17.57	77.3	81.1	89.4

Table A8

One-Sided Tolerance Limits (b known)

Factors $z(\beta,\gamma,n)$ such that $\hat{m} - z(\beta,\gamma,n)b$ is a lower one-sided β,γ tolerance limit

or $\hat{m} + z(\beta,\gamma,n)b$ is an upper one-sided β,γ tolerance limit

n	β = .8			β = .9			β = .95			β = .99		
	γ .9	γ .95	γ .99	γ .9	γ .95	γ .99	γ .9	γ .95	γ .99	γ .9	γ .95	γ .99
5	2.311	2.734	4.091	4.012	4.435	5.792	7.248	7.672	9.028	32.756	33.179	34.535
10	1.991	2.207	2.751	3.692	3.908	4.452	6.928	7.145	7.688	32.435	32.652	33.196
15	1.866	2.027	2.394	3.568	3.728	4.096	6.804	6.969	7.332	32.311	32.471	32.839
20	1.796	1.928	2.218	3.497	3.630	3.919	6.733	6.866	7.155	32.240	32.373	32.662
25	1.749	1.864	2.108	3.450	3.566	3.810	6.686	6.802	7.046	32.193	32.309	32.553
30	1.715	1.818	2.033	3.416	3.520	3.734	6.652	6.756	6.970	32.159	32.263	32.477
40	1.668	1.755	1.932	3.367	3.457	3.633	6.605	6.693	6.869	32.112	32.200	32.377
50	1.636	1.713	1.867	3.337	3.415	3.568	6.573	6.651	6.804	32.081	32.158	32.311
75	1.587	1.649	1.770	3.289	3.351	3.471	6.525	6.587	6.707	32.032	32.094	32.214
100	1.559	1.612	1.714	3.260	3.313	3.415	6.496	6.549	6.652	32.003	32.057	32.159
Asym. 100	1.558	1.609	1.705	3.259	3.310	3.407	6.495	6.546	6.643	32.002	32.053	32.149

Table A9

Two-sided Tolerance Limits (b known)

Factors $\omega(\beta,\gamma,n)$ such that $(\hat{m} - \omega(\beta,\gamma,n)b,\ \hat{m} + \omega(\beta,\gamma,n)b)$ forms a two-sided β,γ tolerance interval.

n	β = .8			β = .9			β = .95			β = .99		
	γ .9	γ .95	γ .98	γ .9	γ .95	γ .98	γ .9	γ .95	γ .98	γ .9	γ .95	γ .98
5	3.553	3.905	4.580	6.587	6.815	7.301	12.849	12.975	13.259	63.678	63.704	63.765
10	3.270	3.381	3.564	6.419	6.483	6.594	12.760	12.794	12.853	63.660	63.669	63.679
15	3.198	3.261	3.359	6.378	6.414	6.470	12.740	12.758	12.787	63.656	63.659	63.665
20	3.165	3.208	3.274	6.360	6.384	6.421	12.730	12.743	12.762	63.654	63.656	63.660
25	3.146	3.179	3.228	6.350	6.368	6.396	12.725	12.734	12.748	63.653	63.655	63.658
30	3.134	3.161	3.200	6.344	6.358	6.380	12.722	12.729	12.740	63.652	63.654	63.656
40	3.119	3.138	3.166	6.336	6.346	6.361	12.718	12.823	12.731	63.651	63.652	63.654
50	3.111	3.126	3.147	6.331	6.339	6.351	12.715	12.719	12.725	63.651	63.652	63.653
75	3.099	3.109	3.123	6.325	6.330	6.338	12.712	12.715	12.719	63.650	63.651	63.652
100	3.094	3.101	3.111	6.322	6.326	6.331	12.711	12.713	12.715	63.650	63.650	63.651
Asym. 100	3.094	3.100	3.109	6.322	6.327	6.330	12.711	12.712	12.715	63.658	63.659	63.659

Table A10

One-sided Tolerance Limits (m known)

Factors $z(\beta,\gamma,n)$ such that $m - z(\beta,\gamma,n)\hat{b}$ is a lower one-sided β,γ tolerance limit

or $m + z(\beta,\gamma,n)\hat{b}$ is an upper one-sided β,γ tolerance limit

n	β = .8			β = .9			β = .95			β = .99		
	γ .90	γ .95	γ .99	γ .90	γ .95	γ .99	γ .90	γ .95	γ .99	γ .90	γ .95	γ .99
5	3.128	4.084	6.780	6.995	9.133	15.161	14.349	18.735	31.102	72.32	94.42	156.75
10	2.458	2.922	4.133	5.496	6.534	9.242	11.275	13.405	18.960	56.82	67.56	95.56
15	2.216	2.544	3.373	4.956	5.689	7.543	10.167	11.670	15.475	51.24	58.82	77.89
20	2.057	2.305	2.973	4.600	5.155	6.647	9.438	10.576	13.637	47.56	53.30	68.73
25	1.983	2.209	2.709	4.435	4.940	6.058	9.098	10.135	12.429	45.85	51.08	62.64
30	1.917	2.098	2.539	4.286	4.692	5.678	8.794	9.625	11.649	44.32	48.51	58.71
40	1.830	1.983	2.333	4.093	4.435	5.216	8.396	9.098	10.701	42.31	45.85	53.93
50	1.774	1.909	2.202	3.966	4.269	4.924	8.136	8.757	10.102	41.01	44.13	50.91
75	1.710	1.821	2.042	3.823	4.071	4.566	7.843	8.352	9.368	39.53	42.09	47.21
100	1.660	1.742	1.930	3.713	3.896	4.317	7.616	7.992	8.855	38.38	40.28	44.63
Asym. 100	1.681	1.794	2.052	3.760	4.011	4.587	7.713	8.229	9.411	38.87	41.47	47.43

Table A11

Two-sided Tolerance Limits (m known)

Factors $\omega(\beta,\gamma,n)$ such that $(m - \omega(\beta,\gamma,n)\hat{b},\ m + \omega(\beta,\gamma,n)\hat{b})$ forms a two-sided β,γ tolerance interval.

n	β = .8			β = .9			β = .95			β = .99		
	γ .90	γ .95	γ .99	γ .90	γ .95	γ .99	γ .90	γ .95	γ .99	γ .90	γ .95	γ .99
5	6.995	9.133	15.161	14.349	18.735	31.102	28.88	37.70	62.59	144.7	188.9	313.6
10	5.496	6.534	9.242	11.275	13.405	18.960	22.69	26.98	38.16	113.7	135.2	191.2
15	4.956	5.689	7.543	10.167	11.671	15.475	20.46	23.49	31.14	102.5	117.7	156.0
20	4.600	5.155	6.647	9.438	10.576	13.637	18.99	21.28	27.44	95.2	106.6	137.5
25	4.435	4.940	6.058	9.098	10.134	12.429	18.31	20.40	25.01	91.7	102.2	125.3
30	4.286	4.692	5.678	8.794	9.625	11.649	17.70	19.37	23.44	88.7	97.0	117.4
40	4.093	4.435	5.216	8.396	9.098	10.701	16.90	18.31	21.54	84.7	91.7	107.9
50	3.966	4.269	4.924	8.136	8.757	10.102	16.37	17.62	20.33	82.0	88.3	101.9
75	3.823	4.071	4.566	7.843	8.352	9.368	15.78	16.81	18.85	79.1	84.2	94.4
100	3.713	3.896	4.317	7.616	7.992	8.855	15.33	16.08	17.82	76.8	80.6	89.3
Asym. 100	3.760	4.011	4.587	7.713	8.228	9.411	15.52	16.56	18.94	77.8	82.9	94.9

Table A12

Cumulative Percentage Points of $\frac{\hat{m}_1 - m_1}{b_1} - \frac{b_2}{b_1}\left(\frac{\hat{m}_2 - m_2}{b_2}\right)$

Values of $c_{1-\alpha,n_1,n_2}$ such that

$$P\left[\frac{\hat{m}_1 - m_1}{b_1} - \frac{b_2}{b_1}\left(\frac{\hat{m}_2 - m_2}{b_2}\right) \le c_{1-\alpha,n_1,n_2}\right] = 1 - \alpha$$

$$n_1 = n_2 = n$$

The distribution is symmetric about zero.

$\frac{b_2}{b_1} = 1/4$

n \ $1-\alpha$	.8	.9	.56	.975	.99
5	.609	1.003	1.446	1.967	2.850
10	.410	.649	.873	1.133	1.503
15	.323	.505	.673	.845	1.055
20	.281	.434	.566	.689	.854
25	.249	.386	.501	.607	.734
30	.227	.346	.447	.548	.670
40	.196	.295	.389	.470	.564

$\frac{b_2}{b_1} = 1/3$

n \ $1-\alpha$	.8	.9	.95	.975	.99
5	.632	1.037	1.493	2.020	2.920
10	.422	.672	.895	1.165	1.542
15	.335	.521	.689	.863	1.070
20	.289	.445	.578	.704	.867
25	.255	.396	.511	.618	.747
30	.231	.353	.460	.557	.685
40	.201	.304	.398	.480	.577

$\frac{b_2}{b_1} = 1/2$

n \ $1-\alpha$	.8	.9	.95	.975	.99
5	.688	1.131	1.630	2.200	3.120
10	.450	.717	.968	1.229	1.643
15	.356	.556	.732	.918	1.146
20	.309	.474	.620	.747	.914
25	.274	.423	.539	.656	.791
30	.249	.376	.490	.589	.720
40	.215	.325	.424	.510	.605

$\frac{b_2}{b_1} = 2/3$

n \ $1-\alpha$	.8	.9	.95	.975	.99
5	.748	1.230	1.762	2.400	3.330
10	.489	.776	1.039	1.324	1.757
15	.385	.598	.792	.988	1.214
20	.334	.512	.670	.807	.975
25	.296	.455	.579	.702	.851
30	.268	.405	.524	.635	.768
40	.233	.354	.458	.546	.648

$\frac{b_2}{b_1} = 1.0$

n \ $1-\alpha$	.8	.9	.95	.975	.99
5	.889	1.460	2.074	2.865	3.990
10	.577	.924	1.218	1.550	2.110
15	.456	.704	.936	1.142	1.443
20	.392	.603	.788	.956	1.145
25	.353	.531	.684	.830	.993
30	.318	.482	.621	.742	.902
40	.274	.420	.537	.635	.770

TABLE A13

Cumulative Percentage Points of $\frac{\hat{b}_2}{b_2} \cdot \frac{b_1}{\hat{b}_1}$ (m_1, m_2 Known)

Values of $c_{1-\alpha,n}$ such that $P[\frac{\hat{b}_2}{b_2} \cdot \frac{b_1}{\hat{b}_1} \le c_{1-\alpha,n}] = 1-\alpha$

n_1 observations for estimating b_1, n_2 observations for estimating b_2

$1-\alpha$	$n_1 \downarrow$	$n_2 \rightarrow$ 5	10	15	20	25	30	40
.99	5	9.270	6.579	6.137				
.975		6.230	4.785	4.564				
.95		4.555	3.759	3.459				
.90		3.231	2.762	2.578				
.99	10	6.801	4.610	4.054	3.559			
.975		4.870	3.580	3.141	2.985			
.95		3.773	2.880	2.622	2.500			
.90		2.760	2.263	2.098	2.026			
.99	15	5.905	4.007	3.360	3.093	2.967		
.975		4.383	3.161	2.793	2.618	2.487		
.95		3.423	2.640	2.375	2.252	2.159		
.90		2.543	2.117	1.963	1.881	1.825		
.99	20		3.785	3.057	2.810	2.770	2.570	
.975			3.033	2.593	2.399	2.333	2.253	
.95			2.527	2.236	2.099	2.027	1.972	
.90			2.015	1.866	1.760	1.736	1.704	
.99	25			2.985	2.695	2.595	2.423	2.309
.975				2.453	2.316	2.198	2.130	2.041
.95				2.123	2.018	1.937	1.875	1.829
.90				1.795	1.720	1.681	1.619	1.597
.99	30				2.617	2.465	2.364	2.179
.975					2.262	2.133	2.097	1.931
.95					1.969	1.896	1.823	1.751
.90					1.704	1.646	1.597	1.546
.99	40					2.260	2.209	2.047
.975						2.015	1.977	1.834
.95						1.811	1.766	1.666
.90						1.591	1.549	1.507

Table A14

Cumulative Percentage Points of $\frac{(\hat{m}_1 - m_1) - (\hat{m}_2 - m_2)}{\hat{b}_1 + \hat{b}_2}$.

Table of $c_{1-\alpha,n}$ such that $P[\frac{(\hat{m}_1 - m_1) - (\hat{m}_2 - m_2)}{\hat{b}_1 + \hat{b}_2} \le c_{1-\alpha,n}] = 1 - \alpha$

The sample sizes are equal, i.e., $n_1 = n_2 = n$.

The density is symmetric.

n \ $1-\alpha$	.8	.85	.90	.95	.975	.99
5	0.544	0.691	0.891	1.245	1.613	2.100
10	0.311	0.391	0.493	0.650	0.792	0.978
15	0.244	0.303	0.376	0.498	0.595	0.730
20	0.202	0.250	0.315	0.404	0.497	0.611
30	0.164	0.201	0.246	0.321	0.383	0.467
45	0.132	0.161	0.201	0.261	0.310	0.374

Table A15

Cumulative Percentage Points of $\frac{\hat{b}_2}{b_2} \cdot \frac{b_1}{\hat{b}_1}$ (m_1, m_2 unknown).

Table of $c_{1-\alpha,n}$ such that $P[\frac{\hat{b}_2}{b_2} \cdot \frac{b_1}{\hat{b}_1} \leq c_{1-\alpha,n}] = 1 - \alpha$.

The sample sizes are equal, i.e., $n_1 = n_2 = n$.

n \ 1−α	.01	.025	.05	.1	.90	.95	.975	.99
5	.062	.101	.152	.237	4.283	6.690	---	---
10	.189	.251	.316	.410	2.448	3.195	4.010	5.170
15	.271	.349	.408	.489	2.034	2.511	3.017	3.720
20	.336	.400	.456	.543	1.810	2.133	2.503	2.960
30	.418	.473	.539	.617	1.606	1.843	2.089	2.450
45	.489	.549	.603	.679	1.499	1.656	1.847	2.205

APPENDIX B

COMPUTER PROGRAMS

COMPUTER PROGRAM B1

ESTIMATION OF CAUCHY PARAMETERS

The maximum likelihood estimates of the Cauchy parameters are obtained for a given sample.

INPUT DATA

NS	Number of Samples
NOBS	Number of Observations
X	Observations
EP	Accuracy of Estimate(s)
ITEST = 1	If b is known
= 2	If m is known
= 3	If both parameters are unknown
B	Scale Parameter if Known
FM	Location Parameter if Known

EXAMPLE

Suppose there are three different samples. In the first sample the location parameter is known to be zero. There are five observations: -2.3, 1.8, 2.0, -1.1, .7. In the second sample the scale parameter is known to be two. The six observations are as follows: 12.1, 13.6, 5.4, 10.0, 8.9, 6.7. In the third sample, neither parameter is known. The eight observations are as follows: 1.1, 1.3, 2.1, -3.0, -7.1, 4.0, 0.1, 8.2. Three place accuracy is required for all estimates. The data for this example follows.

The card deck for this computer program may be obtained from the Department of Mathematics, University of Missouri at Rolla.

C FOR COMMENT

STATEMENT NUMBER	Cont	FORTRAN STATEMENT

```
     3
     5
      -2.3            1.8                 2.0                -1.1
         .7
         .0005
     2
       0.0
     6
       12.1           13.6                5.4                10.0
       8.4            6.7
         .0005
     1
       2.0
     3
       1.1            1.3                 2.1                -3.0
      -7.1            4.0                 0.1                 8.2
         .0005
     3
```

COMPUTER PROGRAM B2

SIMULATED DISTRIBUTIONS FOR TWO-SAMPLE PROBLEMS

The distributions of $\frac{(\hat{m}_1 - m_1) - (\hat{m}_2 - m_2)}{\hat{b}_1 + \hat{b}_2}$ and $\frac{\hat{b}_2}{b_2} \cdot \frac{b_1}{\hat{b}_1}$ are obtained by a Monte Carlo Simulation. There are n_1 observations of the random variable having parameters m_1 and b_1, and n_2 observations of the random variable having parameters m_2 and b_2.

INPUT DATA

N1 - NUMBER OF OBSERVATIONS IN FIRST SAMPLE (n_1)

N2 - NUMBER OF OBSERVATIONS IN SECOND SAMPLE (n_2)

NI - NUMBER OF ESTIMATES IN MONTE CARLO SIMULATION

EP - ACCURACY OF ESTIMATES

EXAMPLE

Suppose there are five observations in the first sample, 12 observations in the second sample, 15,000 estimates are to be obtained in the simulation, and three place accuracy is desired. The data would appear as follows:

CARD 1

Column 6	Column 12	Column 18
↓	↓	↓
5	12	15000

CARD 2

Column 10
↓
.0005

The card deck for this computer program may be obtained from the Department of Mathematics, University of Missouri at Rolla.

COMPUTER PROGRAM B3

SIMULATED DISTRIBUTIONS FOR TYPE II CENSORED SAMPLES

The simulated distributions of $\frac{\hat{m} - m}{b}$, $\frac{\hat{b}}{b}$, $\frac{\hat{m} - m}{\hat{b}}$, and $E(\hat{m}), E(\hat{b})$ are obtained for Cauchy samples where the r_1 smallest and r_2 largest observations are not available.

INPUT DATA

N	Number of Observations (Observed and Unobserved)
NR1	The Number of Missing Observations on the Left
NR2	The Number of Missing Observations on the Right
NI	Number of Estimates to be Obtained in Simulation
EP	Accuracy of Estimates

EXAMPLE

Suppose the above distributions are desired for samples of size 20 where the two smallest and five largest observations are missing. 10,000 estimates are to be obtained with three place accuracy. The data would be as follows:

CARD 1

Column 6	Column 12	Column 18	Column 24
↓	↓	↓	↓
20	2	5	10000

CARD 2

Column 10
↓
.0005

The card deck for this computer program may be obtained from the Department of Mathematics, University of Missouri at Rolla.

VITA

Gerald Nicholas Haas was born on August 2, 1940 in Buffalo, New York. He received his first four and one-half years of elementary education there. In December, 1950 he moved with his family to Kansas City, Missouri, where he finished his elementary education. He attended North Kansas City High School, North Kansas City, Missouri, from September, 1954 to May, 1958 when he received his diploma.

In September, 1958 he entered the Missouri School of Mines and Metallurgy, Rolla, Missouri. In May, 1962 he received a Bachelor of Science Degree in Physics. From that time to January, 1963 he was employed with McDonnell Aircraft Corporation, St. Louis, Missouri. In February, 1963 he re-entered the Missouri School of Mines and Metallurgy. He graduated with a Master of Science Degree in Applied Mathematics in May, 1964. He served as a graduate assistant in the Computer Center from September, 1963 to May, 1964.

From June, 1964 to August, 1966 he was employed with North American Aviation, Inc., Anaheim, California. In September, 1966 he entered the University of Missouri at Rolla as a Ph.D. candidate in Mathematics. He has served as an Instructor of Mathematics since this time.

On May 30, 1964 he was married to the former Judith Joyce Percival of St. Louis, Missouri. They have one son, Jeffrey Eric, born on December 31, 1966.