A COMPARISON OF NORMAL THEORY AND BOOTSTRAP CONFIDENCE INTERVALS ON THE PARAMETERS OF NONLINEAR MODELS

by

MARY MARGARET ELLING

B.S., Kansas State University, 1979

A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Statistics

KANSAS STATE UNIVERSITY

1981

Approved by:

K. E. Kemp

Major Professor

Table of Contents

THIS BOOK
CONTAINS
NUMEROUS PAGES
WITH THE ORIGINAL
PRINTING BEING
SKEWED
DIFFERENTLY FROM
THE TOP OF THE
PAGE TO THE
BOTTOM.

THIS IS AS RECEIVED
FROM THE
CUSTOMER.

I. Introduction

Nonlinear models have a variety of applications in many fields of research. Usually, the postulated model is of the form

$$y_i = f(x_i, \underline{\theta}) + \varepsilon_i, \qquad i = 1, 2, \ldots, n \quad ,$$

where $y_1, y_2, \ldots, y_n$ are independent observable random variables, $\underline{\theta}$ is a vector of unknown parameters, and $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ are unobservable random errors with zero mean, variance σ^2, and some statistical distribution, F. $f(x, \underline{\theta})$ is a known function of x_i and $\underline{\theta}$. In model fitting situations one wants to obtain estimates of $\underline{\theta}$, denoted by $\hat{\underline{\theta}}$, and confidence intervals about the components of $\underline{\theta}$. Conventional methods for constructing confidence intervals rely on the assumptions that $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ are independent, normally distributed, and that the sample size is large. However, these assumptions are not usually satisfied; thus, alternative methods for constructing confidence intervals are also of interest.

One alternative to the normal theory method of obtaining confidence intervals for the parameters is the bootstrap technique. Work done by Efron (2) suggests that bootstrap methods may be useful in constructing confidence intervals for parameters of nonlinear models. This report consists of an empirical study of the performances of the bootstrap technique and comparisons to the normal theory method in interval estimation for the parameters of three nonlinear models.

II. The Procedure

A Monte Carlo simulation was conducted to evaluate the properties of bootstrap and normal theory confidence intervals for the parameters of three different nonlinear models. The three models are

1. the Michaelis - Menten enzyme kinetic model,

   $y_i = \theta x_i/(\mu + x_i) + \varepsilon_i, \quad i = 1, 2, \ldots, 20$

   where $\theta = 9.61023$ and $\mu = 4.49$,

2. the exponential growth model,

   $y_i = A(1 - C*\exp(-Bx_i^2)) + \varepsilon_i, \quad i = 1, 2, \ldots, 20$

   where $A = 90.56989$, $B = .125$, and $C = 1.00025$, and

3. the exponential decay model,

   $y_i = \alpha\exp(-\beta x_i) + \varepsilon_i, \quad i = 1, 2, \ldots, 20$

   where $\alpha = 255.45$ and $\beta = .0926$.

The normal theory method and three bootstrap procedures, the percentile method, the bias-corrected percentile method, and a smoothed percentile method were used to construct 95% confidence intervals on the parameters of the model. For all models the errors, $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$, were generated using Marsaglia's (6) random number generator (superduper) for a normal distribution with mean 0 and variance, $\sigma^2 = .25$. The sample size, n, was taken to be 20. Observed values of the independent variable, x_i, are listed in Table 1.

enzyme kinetic model	exponential growth model	exponential decay model
1.00	.120	1.0
1.30	.170	3.0
1.60	.196	8.0
2.50	.230	9.0
3.10	.275	11.0
4.50	.300	13.0
5.00	.450	15.0
7.50	.675	20.0
8.50	.750	23.0
15.50	.957	24.0
18.50	1.175	25.0
22.25	1.550	27.0
26.00	2.335	32.0
29.00	2.700	35.0
33.00	3.425	38.0
35.00	3.900	40.0
37.00	4.500	43.0
39.00	4.882	46.0
42.00	5.270	50.0
45.00	6.000	55.0

Table 1 Observed values of the independent variable, x_i, $i = 1, 2, \ldots, 20$, for each of the three models used.

To allow for a comparison of the normal theory and bootstrap methods when normality does not hold, F also assumed the Student's t, with 4 degrees of freedom, and the negative exponential distribution centered at zero.

Then for all cases least squares estimates of the parameters were obtained using the modified Gauss-Newton least squares estimation procedure. A simulation with 100 trials was conducted and the average width of the intervals and confidence levels were recorded.

III. Nonlinear Estimation - Normal Theory Method

1. Parameter Estimation

Given the postulated model

$$y_i = f(x_i, \underline{\theta}) + \varepsilon_i, \qquad i = 1, 2, \ldots, n \quad ,$$

we assume that $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ are uncorrelated, that Var $(\varepsilon_i) = \sigma^2$, and, usually, that $\varepsilon_i \sim N(0, \sigma^2)$. The modified Gauss-Newton method, as described by Draper and Smith (1) was used in this report to provide least squares estimates of the parameters.

Initial values of the parameters, $\underline{\theta}_o$, are generally obtained from whatever information is at hand, to be improved upon by successive iterations. A Taylor series expansion of the first derivative about the point $\underline{\theta}_o$ is used to approximate $f(x_i, \underline{\theta})$. So when $\underline{\theta}$ is close to $\underline{\theta}_o$

$$f(x_i, \underline{\theta}) \doteq f(x_i, \underline{\theta}_o) + \sum_{k=1}^{p} \left[\frac{\partial f(x_i, \underline{\theta})}{\partial \theta_k}\right]_{\underline{\theta} = \underline{\theta}_o} (\theta_k - \theta_{ko}) \quad .$$

Then letting

$$f_i = f(x_i, \underline{\theta}_o) \quad ,$$

$$d_k = \theta_k - \theta_{ko} \quad , \quad \text{and}$$

$$Z_{ik} = \left[\frac{\partial f(x_i, \underline{\theta})}{\partial \theta_k} \right]_{\underline{\theta} = \underline{\theta}_o} \qquad (3.1.1)$$

the model can be rewritten as

$$y_i - f_i = \sum_{k=1}^{p} Z_{ik} d_k + \varepsilon_i, \qquad i = 1, 2, \ldots, n \ .$$

The model is now linear in d_i; therefore, linear least squares can be used to obtain the components of the correction vector $\underline{d}$. The solution vector which minimizes the sum of squares residual,

$$SS(\underline{\theta}) = \sum_{i=1}^{p} (y_i - f(\underline{x}_i, \underline{\theta}_o) - \sum_{k=1}^{p} d_k Z_{ik})^2 \ ,$$

can be written as

$$\hat{\underline{d}} = (\underline{Z}'\underline{Z})^{-1} \underline{Z}' (\underline{y} - \underline{f})$$

where

$$\underline{Z} = \begin{bmatrix} Z_{11} \cdots Z_{1p} \\ \vdots \\ Z_{n1} \cdots Z_{np} \end{bmatrix} \ ,$$

$$\underline{y} = (y_1, y_2, \ldots, y_n)' \ ,$$

$$\underline{f} = (f_1, f_2, \ldots, f_n)' \ , \quad \text{and}$$

$$\hat{\underline{d}} = (\hat{d}_1, \hat{d}_2, \ldots, \hat{d}_p)' \ .$$

An updated estimate of $\underline{\theta}$, $\underline{\theta}_1$ is computed as $\underline{\theta}_1 = \underline{\theta}_0 + \lambda^*\underline{d}$. In general, $\underline{\theta}_j$ is updated to $\underline{\theta}_{j+1}$, where $\underline{\theta}_{j+1} = \underline{\theta}_j + \lambda^*\underline{d}$, and $\underline{d}$ is the correction vector $\underline{\theta} - \underline{\theta}_0$. Hartley (4) suggests selecting λ to minimize

$$SS(\underline{\theta}_{j+1}) = SS(\underline{\theta}_j + \lambda\, \underline{d}). \qquad (3.1.2)$$

This can be approximated by evaluating $SS(\underline{\theta}_{j+1})$ at $\lambda = 0$, $\lambda = \frac{1}{2}$, and $\lambda = 1$. First evaluate $SS(\underline{\theta}_{j+1})$ with $\lambda = 1$. If $SS(\underline{\theta}_{j+1})|_{\lambda=1} < SS(\underline{\theta}_{j+1})|_{\lambda=0}$ then use $\underline{\theta}_{j+1} = \underline{\theta}_j + \underline{d}$. If $SS(\underline{\theta}_{j+1})|_{\lambda=1} \geq SS(\underline{\theta}_{j+1})|_{\lambda=0}$ evaluate (3.1.2) at $\lambda = \frac{1}{2}$. If $SS(\hat{\underline{\theta}}_{j+1})|_{\lambda=\frac{1}{2}} < SS(\underline{\theta}_{j+1})|_{\lambda=0}$ then use $\underline{\theta}_{j+1} = \theta_j + \frac{1}{2}\underline{d}$; otherwise, evaluate $SS(\underline{\theta}_{j+1})$ at $\lambda = (\frac{1}{2})^2$. The computations continue in this manner until the final value of λ is $\lambda^* = (\frac{1}{2})^k$, where

$$SS(\underline{\theta}_{j+1})|_{\lambda=\lambda^*} < SS(\underline{\theta}_{j+1})|_{\lambda=0} \ .$$

Thus, $SS(\underline{\theta}_{j+1})|_{\lambda=\lambda^*}$ will approximate the minimum of the sum of squares residual. Then $\underline{\theta}_{j+1}$ is updated to $\underline{\theta}_{j+2}$ where $\underline{\theta}_{j+2} = \underline{\theta}_{j+1} + \lambda^*\underline{d}$. The iterative cycle begins again and continues until λ approaches 0 or $\underline{d} = \underline{0}$, where zero is taken to be some small number, say, 1.0×10^{-10}.

The least squares estimator of $\underline{\theta}$, denoted by $\hat{\underline{\theta}}$, is defined to be that value of $\underline{\theta}$ which minimizes $SS(\underline{\theta})$. It should be noted that under the assumption $\varepsilon_i \sim N(0, \sigma^2)$, $i = 1, 2, \ldots, n$, that $\hat{\underline{\theta}}$ can also be shown to be the maximum likelihood estimator of $\underline{\theta}$.

2. Confidence Intervals

Approximate confidence intervals about individual parameters can be obtained from the asymptotic distribution of the maximum likelihood estimator of $\underline{\theta}$. It can be shown that the maximum likelihood estimator, $\hat{\underline{\theta}}$, is asymptotically distributed as

$$\hat{\underline{\theta}} \sim \text{MVN}\,(\underline{\theta}, \sigma^2 C^{-1}) \quad ,$$

where

$$C = \left[\left(\sum_{i=1}^{n} \frac{\partial f(x_i, \underline{\theta})}{\partial \theta_s} \frac{\partial f(x_i, \underline{\theta})}{\partial \theta_r} \right) \right] \quad .$$

Here, $\sigma^2 C^{-1}$ is the inverse of the information matrix, see Milliken (7). This leads to a $(1 - \alpha)$ 100% asymptotic approximate confidence interval about $\underline{\ell}'\underline{\theta}$ which is

$$\underline{\ell}'\underline{\theta} \pm Z_{\frac{\alpha}{2}} \sqrt{\underline{\ell}'\sigma^2 C^{-1}\underline{\ell}} \quad ,$$

where

$$\hat{\sigma}^2 = \frac{1}{n-p} \sum_{i=1}^{n} (y_i - f(x_i, \hat{\underline{\theta}}))^2 \quad .$$

IV. The Bootstrap Technique

The bootstrap technique is applicable to a variety of statistical problems. For this reason a general description of the problem will be given first. Let $\underline{Y} = (Y_1, Y_2, \ldots, Y_n)'$ be a random sample drawn from the unknown distribution F. Given its observed values $\underline{y} = (y_1, y_2, \ldots, y_n)'$, the problem is to estimate the distribution of some prespecified random variable $R(\underline{Y}, F)$, which is a function of F and $\underline{Y}$. For example, $R(\underline{Y}, F) = (\hat{\theta} - \theta)/(\text{Var } (\hat{\theta}))^{\frac{1}{2}}$, where θ is an unknown parameter of F and $\hat{\theta}$ is an estimate of θ.

Let F_n be the sample distribution of $(y_1, y_2, \ldots, y_n)$, putting mass $1/n$ at each y_i. Draw a random sample $\underline{Y}^* = (Y_1^*, Y_2^*, \ldots, Y_n^*)'$ from F_n. Then the distribution of $R(\underline{Y}^*, F_n)$ based on $\underline{Y}$ can be used to approximate the distribution of $R(\underline{Y}, F)$. Hence, if $R(\underline{Y}, F)$ is a pivotal quantity of some unknown parameter, then the distribution of $R(Y^*, F_n)$ can be used to construct confidence intervals about θ.

1. Bootstrap Confidence Intervals

Let $\underline{Y} = (Y_1, Y_2, \ldots, Y_n)'$ be a random sample drawn from F, some unknown distribution, dependent on θ, an unknown parameter. Given $\underline{y} = (y_1, y_2, \ldots, y_n)'$, the observed values of $\underline{Y}$ and $\hat{\theta}$, an estimator of θ obtained by a method such as least squares or maximum likelihood estimation, the objective is to construct confidence intervals about θ.

1.1 The Percentile Method

Construct the sample distribution, F_n, by putting mass 1/n at each point $(y_1, y_2, \ldots, y_n)$. Then draw the random sample, $\underline{Y}^* = (Y_1^*, Y_2^*, \ldots, Y_n^*)$, of size n, with replacement from F_n. $\underline{Y}^*$ is called the bootstrap sample. Then the bootstrap estimator of θ will be $\theta^* = \hat{\theta}(Y_1^*, Y_2^*, \ldots, Y_n^*)$.

Let $\hat{CDF}$ (t) be the cumulative distribution function of θ^*. Then the lower and upper bounds of a bootstrap $(1 - \alpha)$ 100% confidence interval are

$$\hat{\theta}_L (\alpha/2) = \hat{CDF}^{-1} (\alpha/2) \quad ,$$

and

$$\hat{\theta}_u (1- \alpha/2) = \hat{CDF}^{-1}(1 - \alpha/2) \quad ,$$

respectively. Thus the $(1 - \alpha)$ 100% bootstrap confidence interval about θ can be constructed as

$$[\hat{\theta}_L(\alpha/2), \ \hat{\theta}_u(1 - \alpha/2)] \quad . \tag{4.1.1}$$

In practice $\hat{CDF}$ (t) cannot be expressed explicitly, however, it can be evaluated by the following simulation technique. Independently, draw samples of size n from F_n a large number of times, say "N" times, and obtain $\theta^*(1), \theta^*(2), \ldots, \theta^*(N)$, the bootstrap replications[+] of $\hat{\theta}$. Then the proportion of the times $\theta^*(i) \leq t$ $i = 1, 2, \ldots, N$, will be a good

---

+ "Replication" refers to the selection of the bootstrap sample, $Y_1^*, Y_2^*, \ldots, Y_n^* \sim F$, whereas "Trial" refers to a new choice of the data $Y_1, Y_2, \ldots, Y_n \sim F$.

approximation of $\hat{CDF}$ (t) if N is large. That is,

$$\hat{CDF}\ (t) \doteq \frac{\#\{\theta^*(i) \leq t\}}{N} \qquad \text{for N large.}$$

Therefore, if $\theta^*(1) \leq \theta^*(2) \leq \ldots\ \theta^*(N)$, the lower bound of the bootstrap $(1 - \alpha)$ 100% confidence interval about θ is the bootstrap replication with rank IL = $[N \times \alpha/2]$, the greatest integer less than or equal to $N \times \alpha/2$, and the upper bound is the bootstrap replication with rank, IU = N - IL + 1.

1.2 The Bias-corrected Percentile Method

A correction for bias in the bootstrap replications may be called for if the replications of $\hat{\theta}$ are skewed rather than symmetrically distributed about the value of $\hat{\theta}$. Efron (3) describes a bias-correction method for constructing confidence intervals in the following manner.

Let

$$Z_o = \Phi^{-1}[\hat{CDF}\ (\hat{\theta})] \quad ,$$

where Φ is the cumulative distribution function for a standard normal variate. The $(1 - \alpha)$ 100% bias-corrected bootstrap confidence interval about θ is

$$\left[\hat{CDF}^{-1}(\Phi(2Z_o - Z_{(1 - \frac{\alpha}{2})})),\ \ \hat{CDF}^{-1}(\Phi(2Z_o + Z_{(1 - \frac{\alpha}{2})}))\right] \quad .$$

As before, $Z_{(1 - \frac{\alpha}{2})}$ is the $(1 - \alpha/2)$ percentile point for the standard

normal, i.e. $\Phi(Z_1 - \frac{\alpha}{2}) = 1 - \alpha/2$. It can be seen that if

$$\frac{\#\{\theta^*(i) \leq \hat{\theta}\}}{N} = .5 \qquad \text{then}$$

$$Z_0 = \Phi^{-1}(\hat{CDF}(\hat{\theta})) = \Phi^{-1}(.5) = 0.$$

Then the bias-corrected bootstrap confidence interval is the percentile confidence interval in (4.1.1), only using

$$\hat{\theta}_L(\alpha/2) = \hat{CDF}^{-1}(\Phi(2Z_0 - Z_{(1 - \frac{\alpha}{2})}))$$

and

$$\hat{\theta}_u(1 - \alpha/2) = CDF^{-1}(\Phi(2Z_0 + Z_{(1 - \frac{\alpha}{2})})) \quad .$$

1.3 The Percentile Method Using a Smoothed Estimate of F

If one is dealing with continuous random variables it may be helpful to generate bootstrap observations from a smoother distribution than F_n, see Efron (2). A bootstrap sample is a random sample from the discrete distribution F_n. Some smoothness may be attributed to F_n by convoluting F_n with F_s, where F_s is a known continuous distribution, to obtain a smoothed bootstrap sample.

The procedure goes as follows: generate $W_1, W_2, \ldots, W_n$ from some known distribution having mean zero and variance σ_w^2. Possible distributions include the uniform on the interval $-\frac{1}{2}$ to $\frac{1}{2}$ or the standard normal. Then the $i^{\underline{th}}$ data point of the bootstrap sample is taken to be

$$y_i^* = \bar{y} + c[y(I_i) - \bar{y} + \hat{\sigma} W_i] \quad ,$$

where the $y(I_i)$ are randomly selected from the observed values $y_1, y_2, \ldots, y_n$, $\bar{y}$ and $\hat{\sigma}$ are the sample mean and sample standard deviation respectively, and c is a constant equal to $(1 + \sigma_w^2)^{-\frac{1}{2}}$. Thus, the Y_i^*, $i = 1, 2, \ldots, n$, have mean zero and variance $\hat{\sigma}^2$. Letting F* represent the distribution of Y_i^* we are replacing F_n with F*, where F* has the same mean and variance as F_n. Then we can obtain $\theta^* = \hat{\theta}(Y_1^*, Y_2^*, \ldots, Y_n^*)$, the estimate of $\hat{\theta}$ based on Y_i^* and x_i. A $(1 - \alpha)100\%$ confidence interval about θ will be

$$[\hat{\theta}_L(\alpha/2), \hat{\theta}_U(1 - \alpha/2)]$$

and can be constructed using the percentile method.

2. Bootstrap Confidence Intervals for the Parameters in a Nonlinear Model

The bootstrap technique is applicable to confidence interval construction for the parameters in a nonlinear model. Given

$$y_i = f(x_i, \theta) + \varepsilon_i, \quad i = 1, 2, \ldots, n \qquad (4.2.1)$$

where $\varepsilon_i \sim$ iid F, some unknown distribution, and $\hat{\theta}$ is the least squares estimate of θ.

The bootstrap technique takes the following steps.

Step 1. After obtaining the least squares estimator, $\hat{\theta}$, substitute $\hat{\theta}$ for θ in (4.2.1) to obtain $\hat{\varepsilon}_i$, defined by

$$\hat{\varepsilon}_i = y_i - f(x_i, \hat{\theta}), \quad i = 1, 2, \ldots, n \ .$$

Step 2. Construct the sample distribution F_n, by putting mass 1/n at each point $(\hat{\varepsilon}_1, \hat{\varepsilon}_2, \ldots, \hat{\varepsilon}_n)$. Then draw the random sample, $\underline{\varepsilon}^* = (\varepsilon_1^*, \varepsilon_2^*, \ldots, \varepsilon_n^*)$ of size n from F_n by the following selection process. Let

$$\varepsilon_i^* = \hat{\varepsilon}(I_i), \qquad i = 1, 1, \ldots, n \tag{4.2.2}$$

where I_i is a randomly selected integer between 1 and n. Random selection of I_i is accomplished using Marsaglia's random number generator (6). Select U_i, a uniform (0, 1) random number, multiply U_i by n, the number of observations, then find the largest integer less than or equal to $U_i \times n$. Thus, I_i is computed to be

$$I_i = [U_i \times n] + 1, \qquad i = 1, 2, \ldots, n \tag{4.2.3}$$

$\underline{\varepsilon}^* = (\varepsilon_1^*, \varepsilon_2^*, \ldots, \varepsilon_n^*)$ is called the bootstrap sample of $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$.

Step 3. Compute

$$y_i^* = f(x_i, \hat{\theta}) + \varepsilon_i^*, \qquad i = 1, 2, \ldots, n \ . \tag{4.2.4}$$

Then by the least squares method utilized in Step 1, obtain θ^* based on x_i and y_i^*, $i = 1, 2, \ldots, n$.

2.1 The Percentile Method

Construction of the $(1 - \alpha)100\%$ bootstrap confidence interval for θ can now be accomplished. Independently, repeat the procedure described in Steps 2 and 3 N times to obtain bootstrap replications of $\hat{\theta}$, denoted by $\theta^*(1), \theta^*(2), \ldots, \theta^*(N)$. Rank the bootstrap replications so that $\theta^*(1) \leq \theta^*(2) \leq \ldots \leq \theta^*(N)$. Define

$$IL = [N \times (\alpha/2)] \qquad (4.2.5)$$

and

$$IU = N - IL + 1 \quad ,$$

where IL is the greatest integer less than or equal to $N \times \alpha/2$. Then a $(1 - \alpha)100\%$ bootstrap confidence interval about θ is

$$[\theta^*(IL), \theta^*(IU)] \quad .$$

2.2 The Bias-corrected Percentile Method

A $(1 - \alpha)100\%$ confidence interval for θ using the percentile method could be too wide if the bootstrap replications, $\theta^*(1), \theta^*(2), \ldots, \theta^*(N)$, are not symmetrically distributed about $\hat{\theta}$. In this case a modification of the percentile method may be called for. Obtain bootstrap replications of $\hat{\theta}$ by the bootstrap technique. Then define

$$Z_0 = \Phi^{-1}(\hat{CDF}(\hat{\theta})),$$

where $\hat{CDF}(\hat{\theta}) = \frac{\#\{\theta^* \leq \hat{\theta}\}}{N}$.

Next, define the lower and upper percentile points adjusted for bias as

$$\alpha_b = \Phi(2Z_o - Z_{(1 - \alpha/2)}) \quad ,$$

$$1 - \alpha_b = \Phi(2Z_o + Z_{(1 - \alpha/2)}) \quad ,$$

respectively. Here, $Z_{(1 - \alpha/2)}$ is the $(1 - \alpha/2)$ percentile point for the standard normal variate, that is $\Phi(Z_{(1 - \alpha/2)}) = (1 - \alpha/2)$. Now, substitute α_b for α in (4.2.5) to find the bootstrap replication with rank IL. Then $\theta^*(IL)$ will be the lower bound for a $(1 - \alpha)100\%$ confidence interval about θ. Similarly, substitute $(1 - \alpha_b)$ for α in (4.2.5) to find the bootstrap replication with rank IU. $\theta^*(IU)$ will be the upper bound. Then as in the percentile method the $(1 - \alpha)100\%$ bootstrap confidence interval using a bias-correction procedure will be

$$[\theta^*(IL), \theta^*(IU)] \quad ,$$

where IL and IU have been adjusted for bias.

2.3 The Percentile Method Using a Smoothed Estimate of F

The smoothed version of the percentile method modifies the distribution F_n so that ε_i^* in (4.2.2) comes from a smoothed version of F_n. Then substitute ε_i^*, selected from the smoothed distribution, F_n, into

(4.2.4). The procedure continues as in the unsmoothed confidence interval construction. The smoothed ε_i^*, $i = 1, 2, \ldots, n$, are obtained in the following manner. At (4.2.2) calculate the mean and variance of $\hat{\varepsilon}_i$, denoted by $\bar{\hat{\varepsilon}}$ and $\hat{\sigma}_\varepsilon^2$, respectively, where

$$\bar{\hat{\varepsilon}} = \frac{\sum_{i=1}^{n} \hat{\varepsilon}_i^2}{n}$$

and

$$\hat{\sigma}_\varepsilon^2 = \frac{\sum_{i=1}^{n} \hat{\varepsilon}_i^2}{n-p}.$$

Then the smoothed estimate of $\hat{\varepsilon}_i$ is

$$\varepsilon_i^* = c(\hat{\varepsilon}(I_i) - \bar{\hat{\varepsilon}} + \hat{\sigma}_\varepsilon W_i) + \bar{\hat{\varepsilon}},$$

where W_i is a continuous random variable following some known distribution, and c is a constant such that ε_i^* have the same mean and variance as $\hat{\varepsilon}_i$. c is taken to be $(1 + \sigma_W^2)^{-\frac{1}{2}}$, where σ_W^2 is the variance of W_i. For example, if $W_i \sim N(0, 1)$ then $c = 1/\sqrt{2}$. The $\varepsilon^*(I_i)$ are selected using the procedure described in (4.2.3). By substituting ε_i^* into (4.2.4) one has values of ε_i^* drawn from a smoothed version of F_n.

3. Determining the Number of Bootstrap Replications Needed

Due to the magnitude of the calculations involved in even a single bootstrap trial coming primarily from the bootstrap replications, it was

deemed necessary to find the smallest number of bootstrap replications which would yield satisfactory results. To determine the number of replications, N, needed for adequate confidence intervals, the average width of the confidence interval from three trials was plotted against values of N. Three different nonlinear models were used: the enzyme kinetic model, the exponential growth model, and the exponential decay model. Trials using sample sizes of 12, 20, and 30 were run for the kinetic and exponential growth models. The exponential decay model was tested with only a sample size of 20. The central 95% confidence interval widths were generated by using both the percentile method and the bias-corrected percentile method. These widths were then plotted against N going from 100 through 1000 in increments of 100. The choice of N was determined by selecting the point where the curves seemed to level off.

For the kinetic and the exponential growth models 300 to 500 bootstrap replications appeared to be sufficient and not excessive. The decay model required more replications; 500 to 700 replications seemed necessary. The graphs (see Appendix A) point out the need for a larger number of bootstrap replications when using a smaller sample size. A sample size of 30 stablized the line more rapidly than the smaller sample sizes. The line representing n = 12 was the least stable. Use of a bias-correction technique also helped to stablize the line for a smaller number of bootstrap replications. Plots D1 and D2 show the effect of a bias-correction procedure.

V. Comparisons and Conclusions

1. Comparisons for Normal and Nonnormal Error Structures

Using only the enzyme kinetic model, F assumed the following distributions: the normal with mean 0 and variance .25, the Student's t with 4 degrees of freedom, and the standard exponential centered about zero. Tables 1.1 and 1.2 provide a summary of the average widths and confidence levels for each distribution. Table 1.3 records the average upper and lower bounds for all trials. It should be noted that with a sample of 100 trials the level of confidence is subject to an error of $\pm$ 2 percentage points.

The normal theory confidence levels were very close to the expected 95% level of coverage for each distribution, F. The widths of the intervals when F was the standard exponential centered about zero were the shortest yet the coverage was as good as for the other distributions using the normal theory method.

The widths of the intervals shortened for all distributions when applying the bootstrap technique, and the confidence level dropped slightly. For the normal distribution the percentile method performed as well as the normal theory method, but under the other distributions the level of coverage was less than satisfactory. Coverage using the bias-corrected percentile method was not as good as the normal or the percentile method for any distribution tested. One reason for the low

level of coverage may have been that the widths of the intervals were shorter when using the bias-corrected method.

F	Parameter	Normal Theory	Percentile Method	Bias-corrected Percentile Method
N(0, .25)	μ	1.87696	1.88462	1.73016
	θ	.983306	.97568	.90061
t(4)	μ	1.84091	1.79157	1.75885
	θ	.95965	.93205	.91661
exp(1)-1	μ	1.79407	1.73016	1.74247
	θ	.93083	.90061	.87606

Table 1.1 Average width of central 95% confidence intervals for parameters of the enzyme kinetic model based on 100 trials. Here, $y_i = \theta x_i/(\mu + x_i) + \varepsilon_i$, where $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{20} \sim F$.

F	Parameter	Normal Theory	Percentile Method	Bias-corrected Percentile Method
N(0, .25)	μ	94%	93%	93%
	θ	94%	94%	91%
t(4)	μ	94%	92%	90%
	θ	91%	90%	88%
exp(1)-1	μ	93%	90%	90%
	θ	96%	96%	95%

Table 1.2 The confidence level of the central 95% confidence intervals for the parameters of the enzyme kinetic model based on 100 trials. Here, $y_i = \theta x_i/(\mu + x_i) + \varepsilon_i$, where $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{20} \sim F$. The coverage is subject to an error of $\pm$ 2 percentage points.

Error Structure	Parameter	Normal Theory		Percentile Method		Bias-corrected Percentile Method	
N(0, .25)	μ	(3.55925	5.45360)	(3.47634	5.36096)	(3.45513	5.35368)
	θ	(8.89820	9.90753)	(9.06441	10.04009)	(9.07167	10.05770)
t(4)	μ	(3.517049	5.35796)	(3.43340	5.22497)	(3.41338	5.17224)
	θ	(9.01019	9.96984)	(9.27656	10.20861)	(9.28043	10.19704)
exp(1) - 1	μ	(3.69735	5.1553)	(3.67659	5.40675)	(3.57313	5.31560)
	θ	(9.12149	10.05232)	(9.08584	9.98645)	(9.10179	9.97785)

Table 1.3 Normal Theory and bootstrap 95% confidence intervals based on the average of 100 trials. Data was fitted to the enzyme kinetic model where the true parameters were $\mu = 4.49$ and $\theta = 9.61023$.

2. Comparisons Between Models

From each of the three models used, the enzyme kinetic, the exponential growth, and the exponential decay, data was generated such that the $\varepsilon_i \sim N(0, .25)$, $i = 1, 2, \ldots, n$ and the sample size, n, was 20. Table 2.1 exhibits the average widths of 100 95% confidence intervals and Table 2.2 records the corresponding coverage of the true parameters. Table 2.3 records the average of the upper and lower bounds of the intervals for 100 trials.

When using the normal theory method, confidence levels for all three models were either at 95% or were within 1 percentage point for all parameters of the models. However, the level of coverage was not as good when applying the percentile method. In both the kinetic and exponential decay models the confidence levels were recorded to be within 2 percentage points of 95%, but when using the 3-parameter growth model the confidence levels varied considerably between the different parameters of the model. Use of the bias-correction factor did not improve on the percentile method as expected, instead the coverage was lower, due to the shortened lengths of the confidence intervals.

Model	Parameter	Normal Theory	Percentile Method	Bias-corrected Percentile Method
Kinetic	μ	1.87696	1.88462	1.73016
	θ	.983306	.97568	.90061
Growth	A	1.322913	1.24958	1.24276
	B	.0058757	.005831	.005447
	C	.006928	.006916	.006825
Decay	α	1.82443	1.73698	1.69408
	β	.00095	.00091	.00086

Table 2.1 Average width of central 95% confidence intervals based on 100 trials. $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{20} \sim N(0, .25)$.

Model	Parameter	Normal Theory	Percentile Method	Bias-corrected Percentile Method
Kinetic	μ	94%	93%	93%
	θ	94%	94%	91%
Growth	A	94%	92%	91%
	B	94%	92%	91%
	C	95%	95%	91%
Decay	α	96%	97%	96%
	β	95%	93%	91%

Table 2.2 Confidence level of the central 95% confidence intervals based on 100 trials. $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{20} \sim N(0, .25)$

Model	Parameter	Normal Theory	Percentile Method	Bias-corrected Percentile Method
Kinetic	μ	(3.55925 5.45360)	(3.47634 5.36096)	(3.45513 5.35368)
	θ	(8.89820 9.90753)	(9.06441 10.04009)	(9.07167 10.05770)
		true parameters: $\mu = 4.49$ $\theta = 9.61023$		
Growth	A	(89.88182 91.26953)	(89.92912 91.17870)	(89.91617 90.40930)
	B	(.125888 .132075)	(.121765 .127595)	(.121675 .127122)
	C	(.996386 1.00367)	(.996565 1.003480)	(.996561 1.003386)
		true parameters: A = 90.56989 B = .125 C = 1.00025		
Decay	α	(254.5573 256.38173)	(254.5984 256.3354)	(254.5987 256.2928)
	β	(.09214 .093091)	(.092158 .093068)	(.09217 .093033)
		true parameters: $\alpha = 255.45$ $\beta = .09260$		

Table 2.3 Normal theory and bootstrap 95% confidence intervals based on the average of 100 trials for three models.

3. Smoothing the Distribution F

In order to add some smoothness to the discrete distribution of the bootstrap sample, ε_1^*, ε_2^*, ..., ε_n^* were chosen from a smoothed verson of F_n. Taking advantage of the fact that $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n \sim N(0, .25)$ the smoothed version of F_n, was taken to be the standard normal distribution. However, the resulting confidence intervals were too wide in the enzyme kinetic model - the only model tested in this manner. Table 3.1 shows that the average length of the intervals for both parameters was much longer than their unsmoothed counterparts, resulting in coverage closer to 99%. A decrease in the width of the interval and a subsequent improvement on the coverage may be made by decreasing $\hat{\sigma}_\varepsilon$. Selecting W_i from a distribution having a smaller variance will decrease the width of the interval and consequently yield a confidence level closer to the nominal 95% level.

	μ		θ	
	Percent Coverage	Averaged width	Percent Coverage	Averaged width
Normal theory	94%	1.87696	94%	.983306
Percentile Method	93%	1.88462	94%	.97568
Smoothed Percentile Method	98%	3.48378	100%	1.77670
Bias-Corrected Percentile Method	93%	1.89717	91%	.96847
Smoothed Bias-Corrected Percentile Method	99%	3.45961	99%	1.76491

Table 3.1 Average widths and percent coverage of the true parameter for normal theory and bootstrap 95% confidence intervals based on 100 trials. The kinetic enzyme model with normal error structure was used.

4. Conclusion

Information on the average widths and confidence levels for the three models shows that the confidence intervals based on normal theory give satisfactory coverage even when the underlying distribution is not normal. Confidence levels dropped slightly when using the bootstrap percentile and bias-corrected percentile methods presumably due to the shortened interval widths. Confidence intervals constructed using the smoothed percentile method were too wide in the case tested, however modifications could easily be tried that would give satisfactory results.

References

1. Draper, N.R., and Smith, H., Applied Regression Analysis, John Wiley and Sons, New York, 1966.

2. Efron, B., "Bootstrap Methods: Another Look at the Jackknife," The Annals of Statistics, Vol. 7, 1 - 26 (1979).

3. Efron, B., "Nonparametric Standard Errors and Confidence Intervals," Technical Report No. 67, Department of Statistics, Stanford University, 1981.

4. Hartley, H.O., "The Modified Gauss-Newton Method for the fitting of Nonlinear Regression Functions by Least-squares," Technometrics, Vol. 3, 269-280 (1961).

5. Goodnight, J.H., "A Tutorial on the Sweep Operator," The American Statistician, Vol. 33, 149-158 (1979).

6. Marsaglia, G., K. Ananthanarayanan, and N.J. Paul, "Improvement on Fast Methods for Generating Normal Random Variables," Information Processing Letters, Vol. 5, No. 2 (1976).

7. Milliken, G.A., Nonlinear Statistical Models, unpublished class notes, Kansas State University, Manhattan, Kansas, 1977.

8. Stilson, M., "A Study of Three Algorithms for Nonlinear Least Squares Parameter Estimation," M.S. Report, Kansas State University, Manhattan, Kansas, 1976.

Appendix A

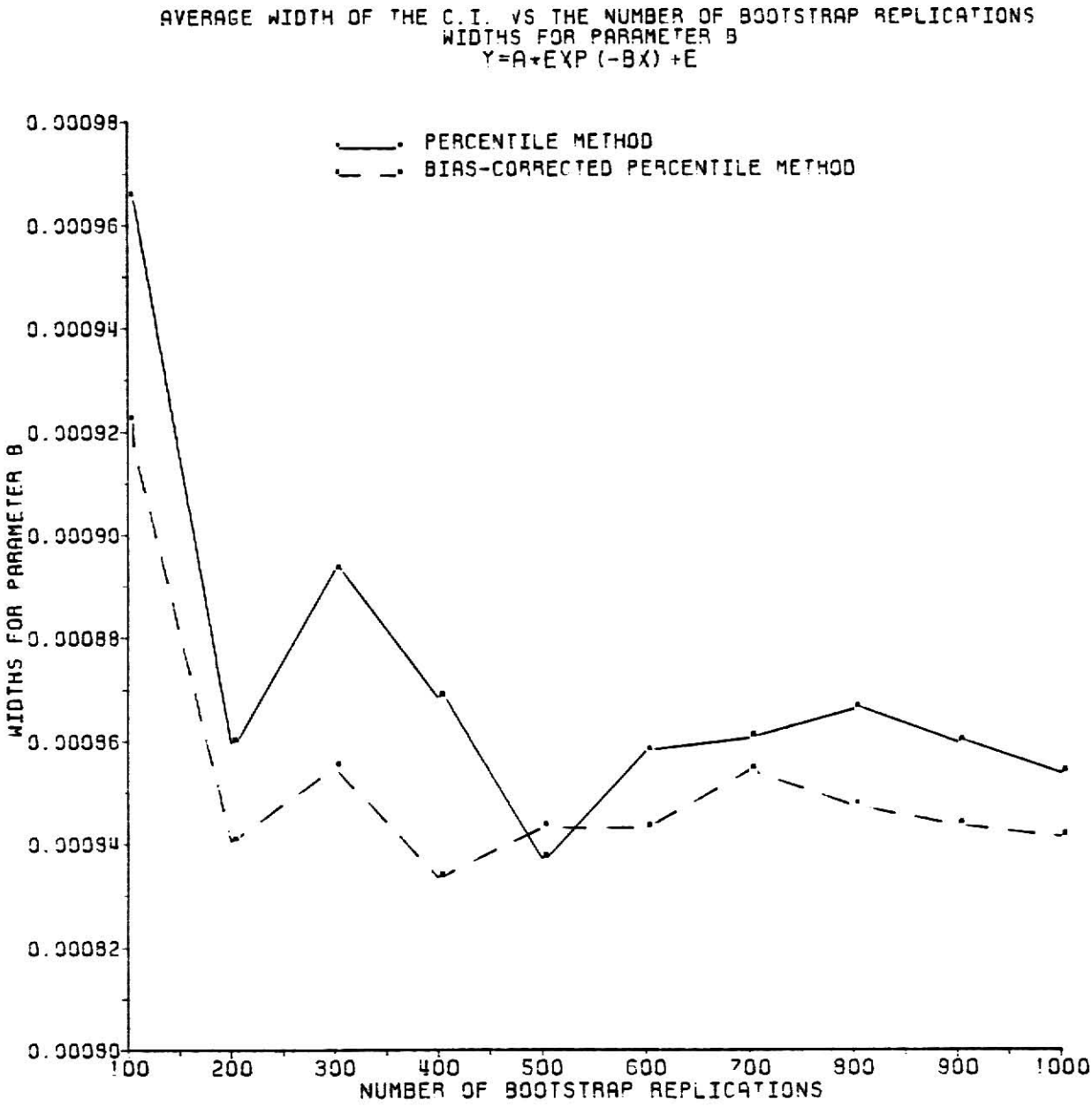
The following graphs were made in an effort to determine the minimum number of bootstrap replications needed per trial. Three models were investigated

(1) an exponential decay model, Plots D1 - D2

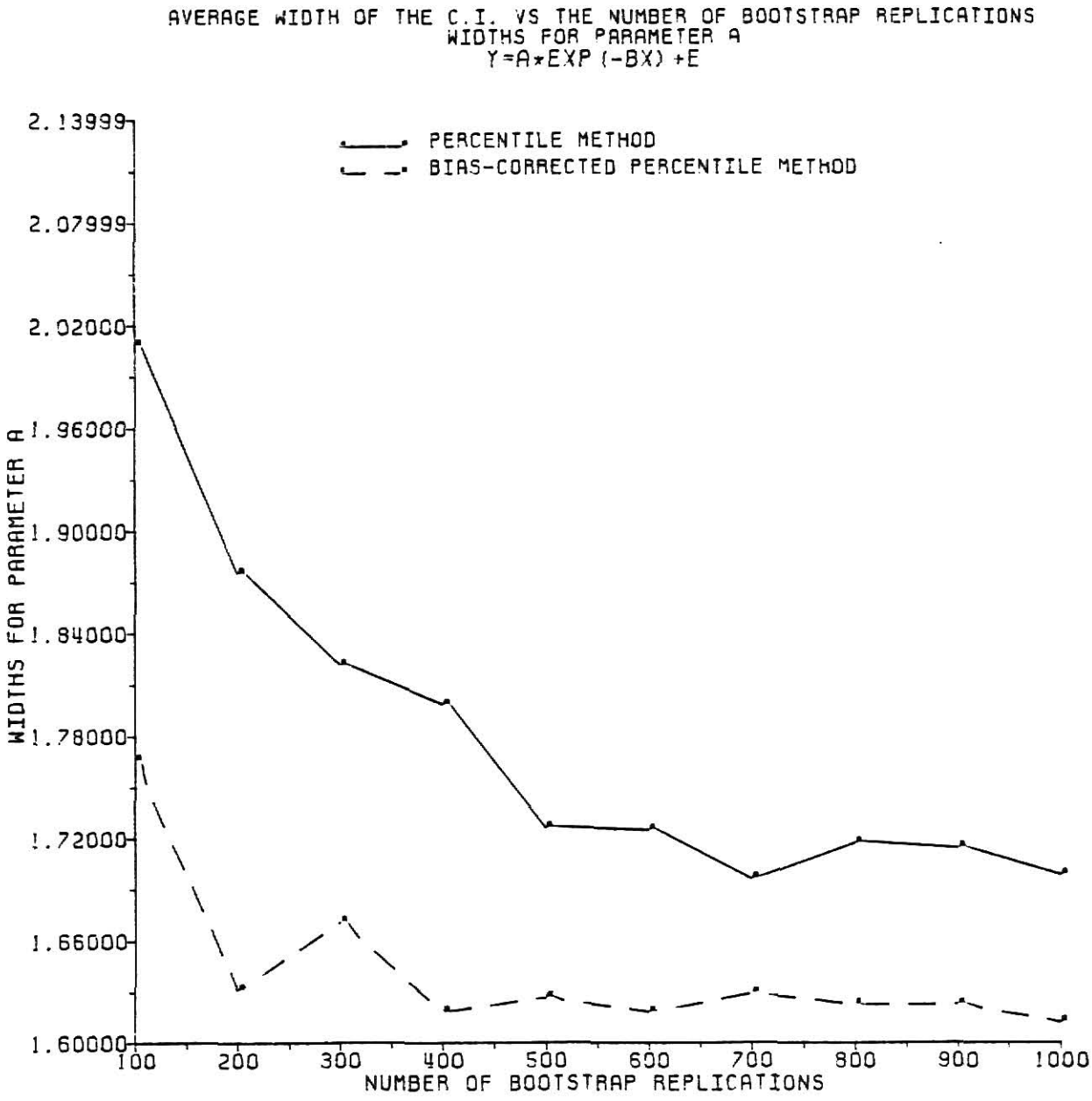
(2) an exponential growth model, Plots G1 - G6

(3) an enzyme kinetic model, Plots K1 - K4.

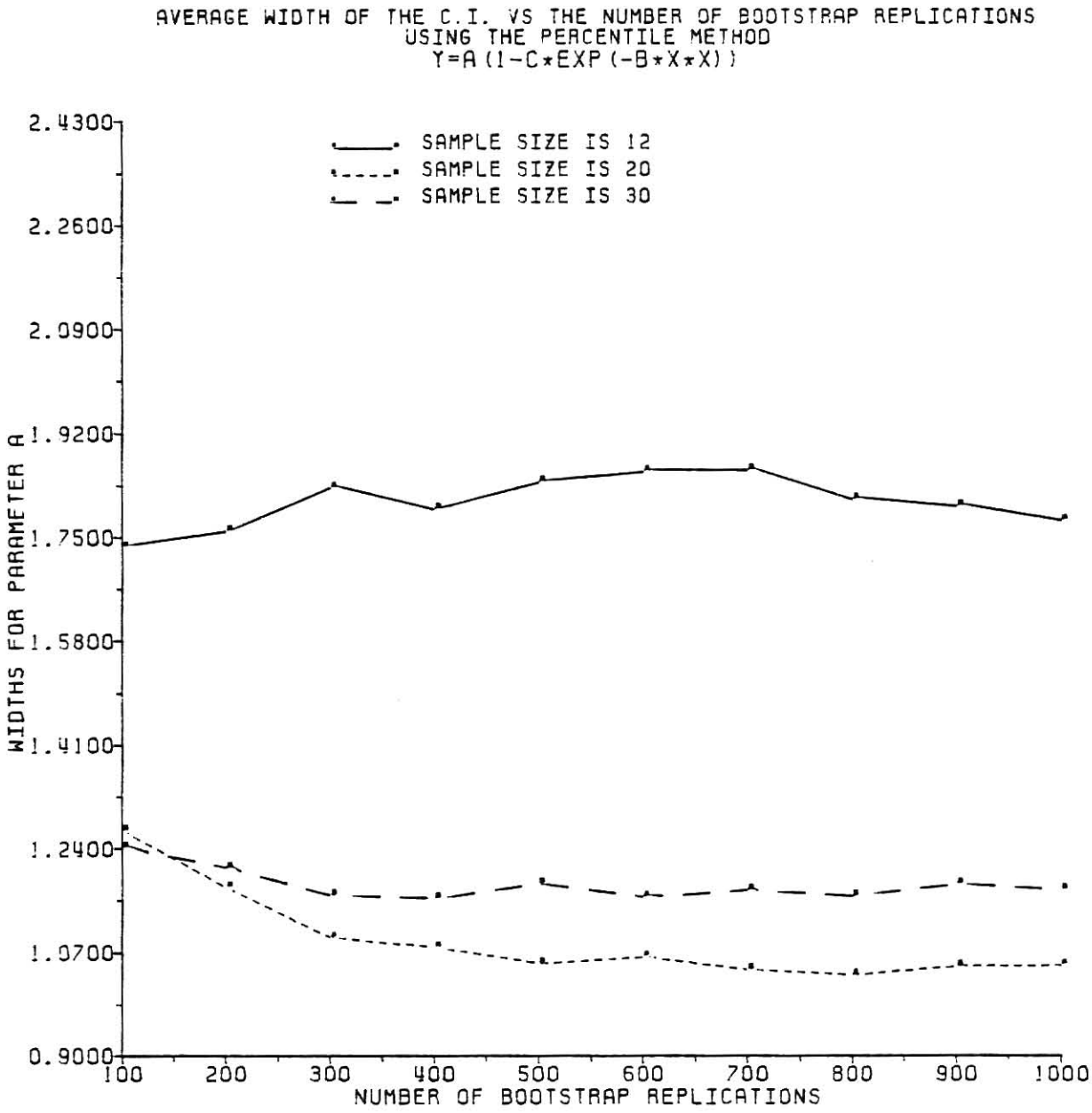
Plots G1 - G6 and K1 - K4 were made to represent each parameter for both the percentile and bias-correction methods. Each plot shows sample sizes of 12, 20 and 30. Plots D1 - D2 were made for each parameter at sample size 20. Each plot shows both percentile and bias-correction methods.


AVERAGE WIDTH OF THE C.I. VS THE NUMBER OF BOOTSTRAP REPLICATIONS
WIDTHS FOR PARAMETER B
Y=A+EXP(-BX)+E
PERCENTILE METHOD
BIAS-CORRECTED PERCENTILE METHOD
WIDTHS FOR PARAMETER B
0.00098
0.00096
0.00094
0.00092
0.00090
0.00088
0.00086
0.00084
0.00082
0.00080
100
200
300
400
500
600
700
800
900
1000
NUMBER OF BOOTSTRAP REPLICATIONS

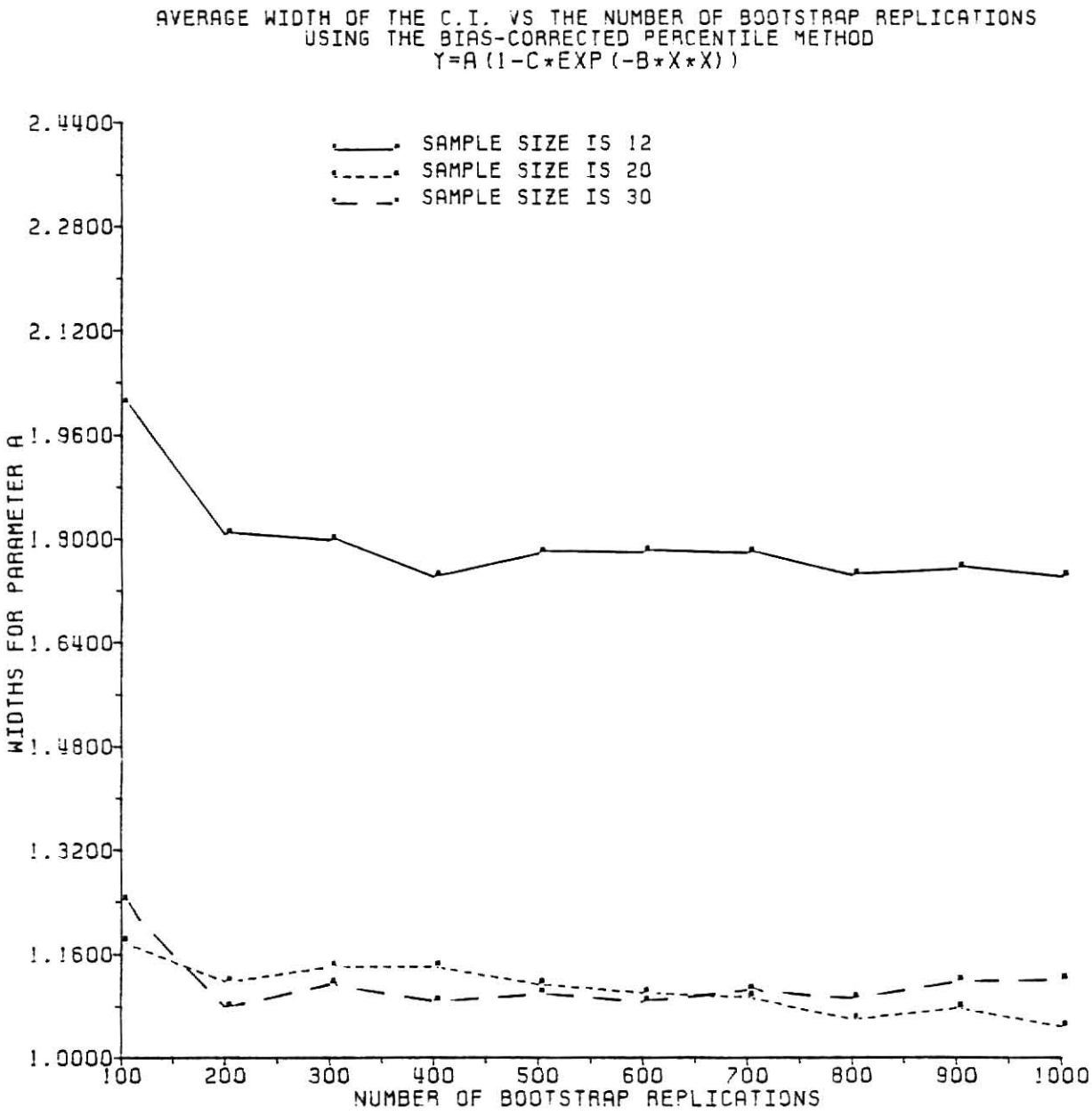
PLOT 01


AVERAGE WIDTH OF THE C.I. VS THE NUMBER OF BOOTSTRAP REPLICATIONS
WIDTHS FOR PARAMETER A
Y=A*EXP(-BX)+E
PERCENTILE METHOD
BIAS-CORRECTED PERCENTILE METHOD
WIDTHS FOR PARAMETER A
2.13999
2.07999
2.02000
1.96000
1.90000
1.84000
1.78000
1.72000
1.66000
1.60000
100
200
300
400
500
600
700
800
900
1000
NUMBER OF BOOTSTRAP REPLICATIONS

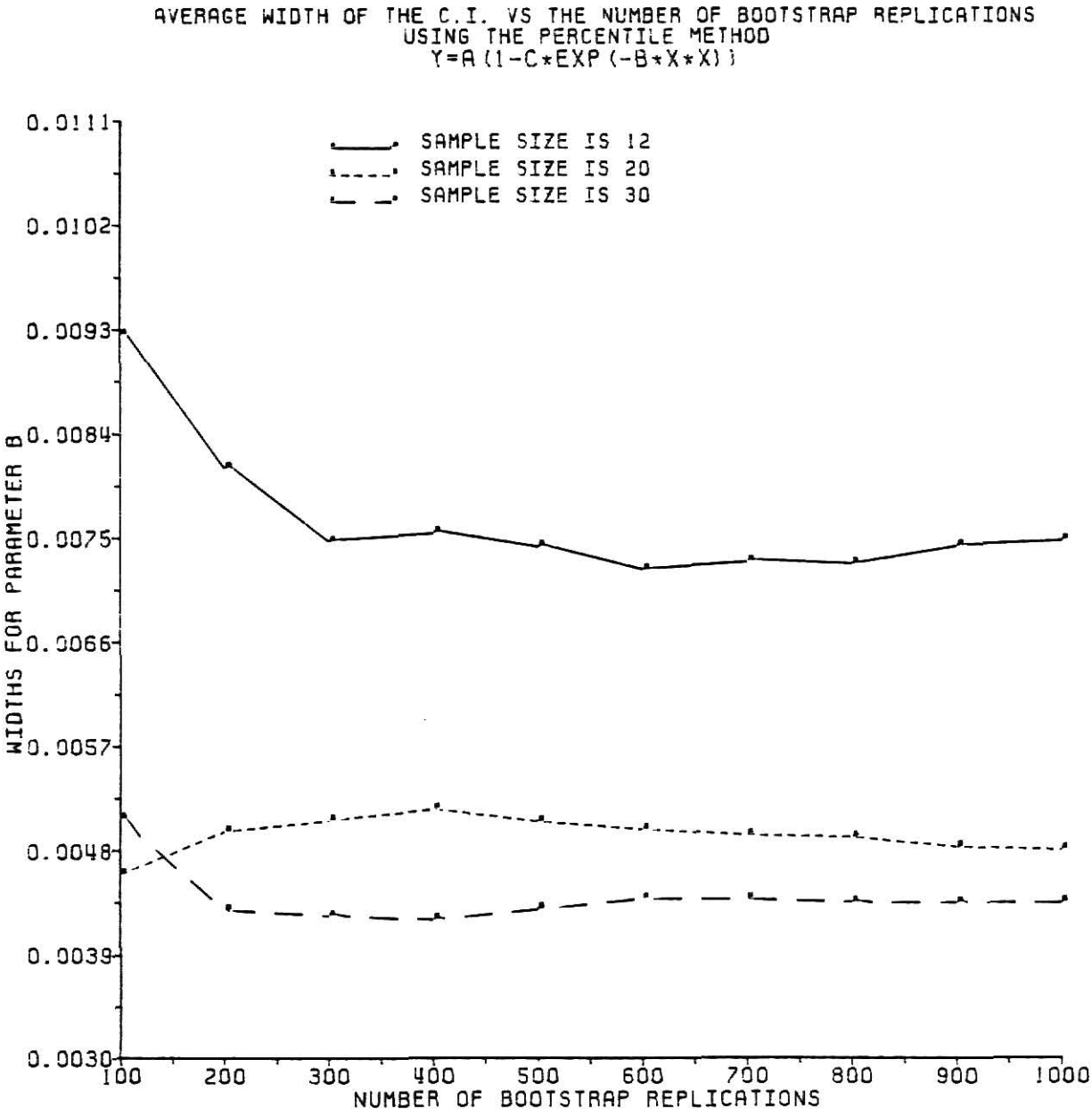
PLOT D2


AVERAGE WIDTH OF THE C.I. VS THE NUMBER OF BOOTSTRAP REPLICATIONS
USING THE PERCENTILE METHOD
Y=A(1-C*EXP(-B*X*X))
SAMPLE SIZE IS 12
SAMPLE SIZE IS 20
SAMPLE SIZE IS 30
WIDTHS FOR PARAMETER A
2.4300
2.2600
2.0900
1.9200
1.7500
1.5800
1.4100
1.2400
1.0700
0.9000
100
200
300
400
500
600
700
800
900
1000
NUMBER OF BOOTSTRAP REPLICATIONS

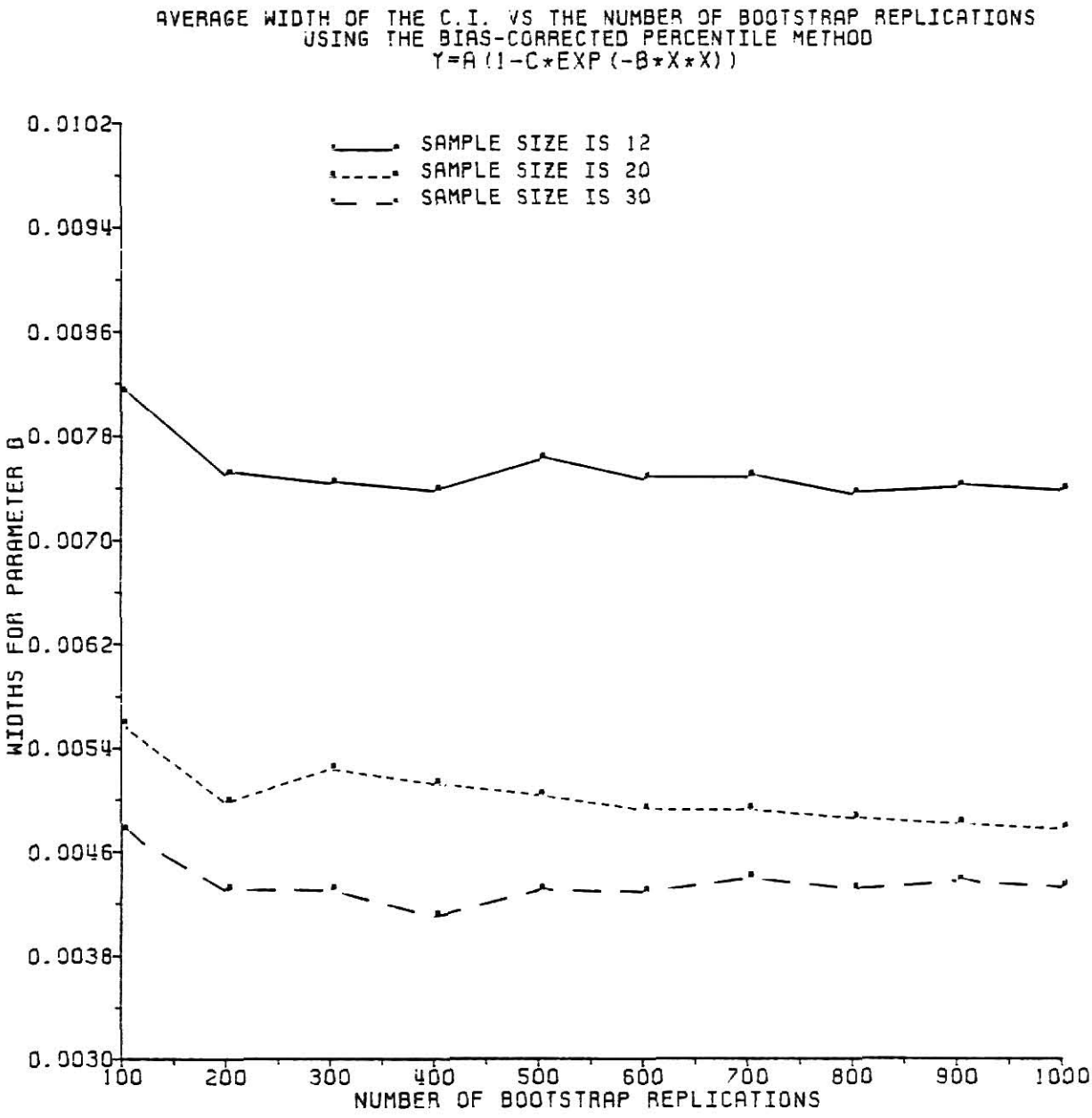
PLOT G1


AVERAGE WIDTH OF THE C.I. VS THE NUMBER OF BOOTSTRAP REPLICATIONS
USING THE BIAS-CORRECTED PERCENTILE METHOD
Y=A(1-C*EXP(-B*X*X))
SAMPLE SIZE IS 12
SAMPLE SIZE IS 20
SAMPLE SIZE IS 30
WIDTHS FOR PARAMETER A
2.4400
2.2800
2.1200
1.9600
1.8000
1.6400
1.4800
1.3200
1.1600
1.0000
100
200
300
400
500
600
700
800
900
1000
NUMBER OF BOOTSTRAP REPLICATIONS

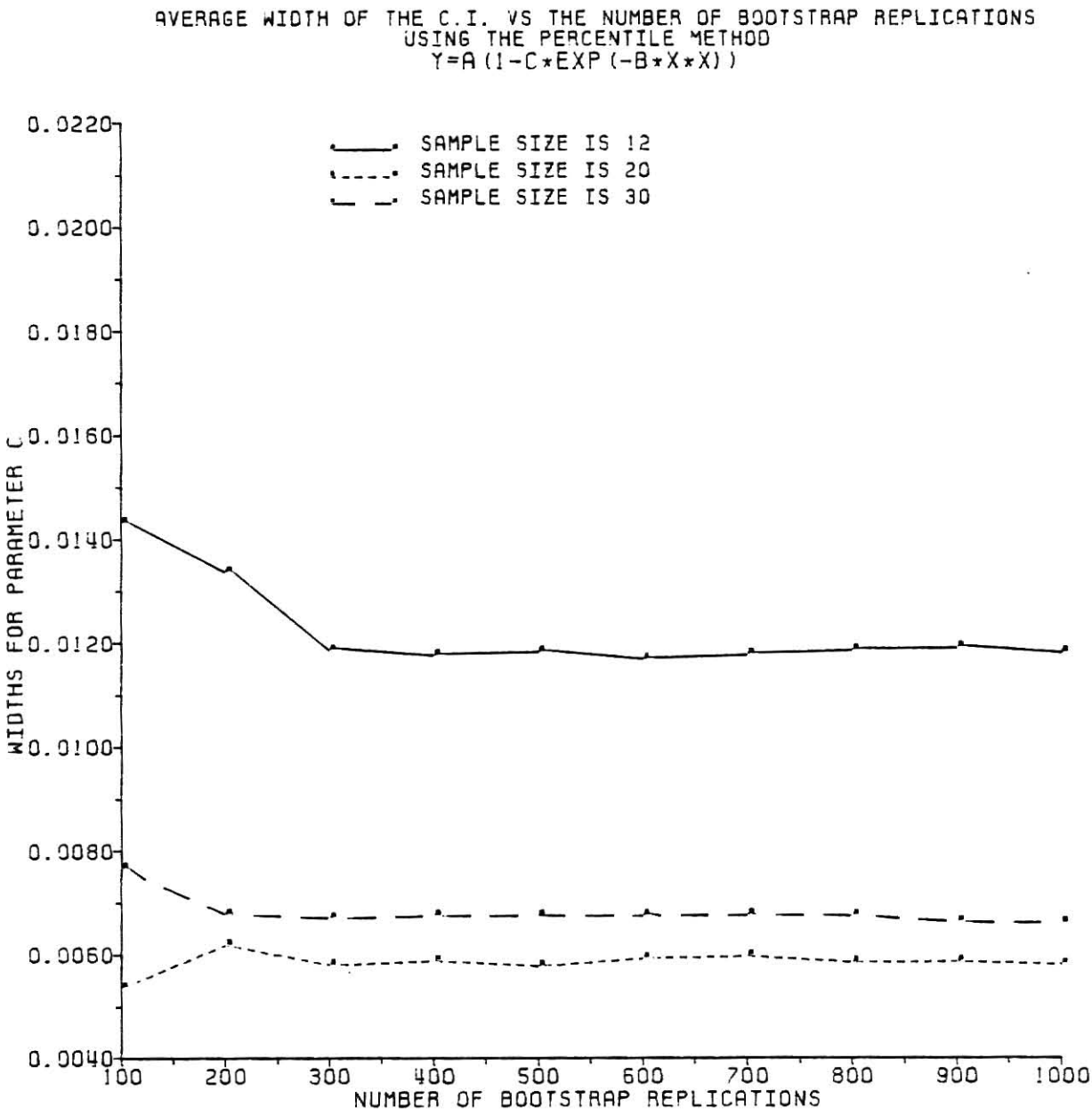
PLOT G2


AVERAGE WIDTH OF THE C.I. VS THE NUMBER OF BOOTSTRAP REPLICATIONS
USING THE PERCENTILE METHOD
Y=A(1-C*EXP(-B*X*X))
SAMPLE SIZE IS 12
SAMPLE SIZE IS 20
SAMPLE SIZE IS 30
WIDTHS FOR PARAMETER B
0.0111
0.0102
0.0093
0.0084
0.0075
0.0066
0.0057
0.0048
0.0039
0.0030
100
200
300
400
500
600
700
800
900
1000
NUMBER OF BOOTSTRAP REPLICATIONS

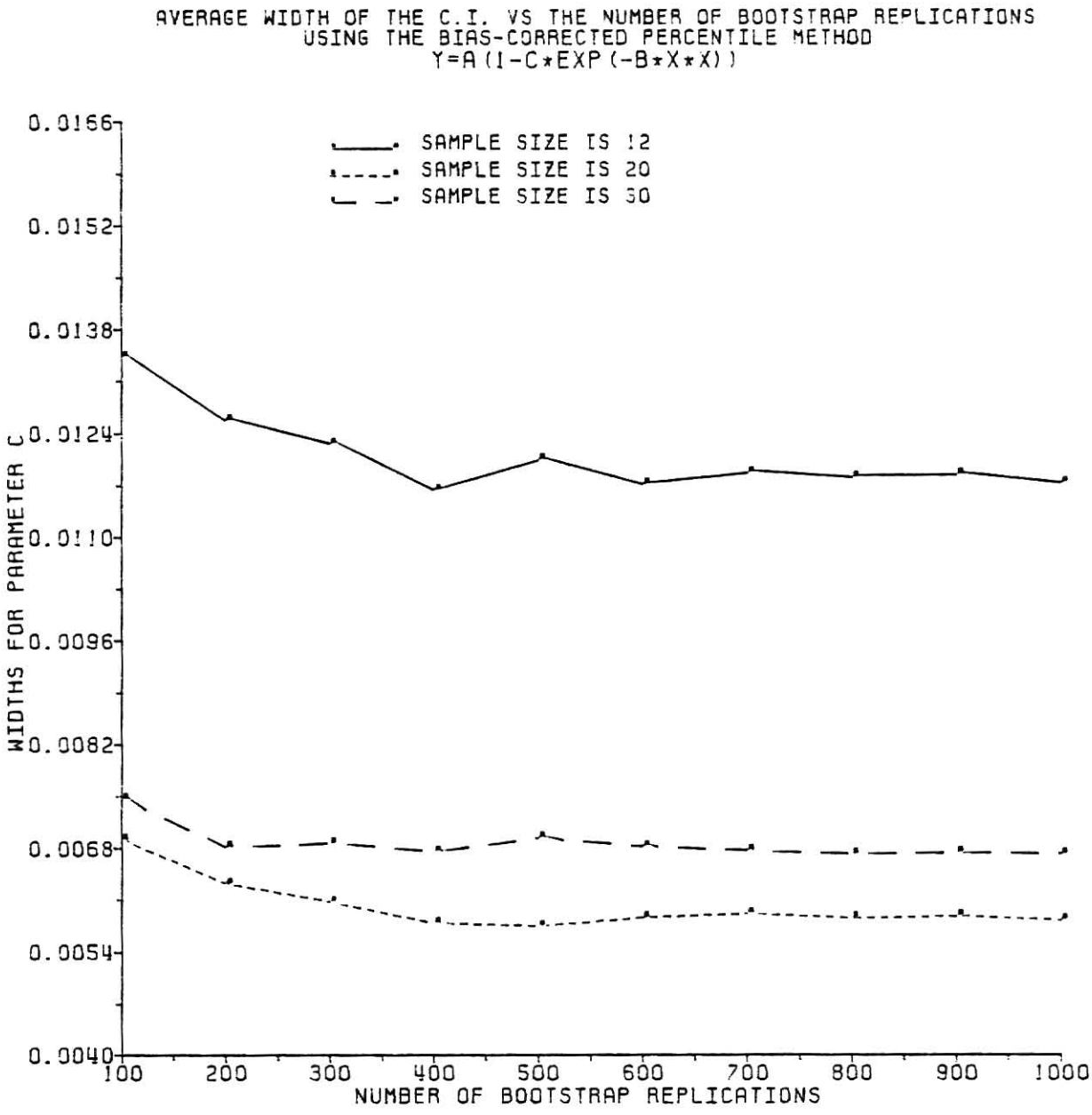
PLOT G3


AVERAGE WIDTH OF THE C.I. VS THE NUMBER OF BOOTSTRAP REPLICATIONS
USING THE BIAS-CORRECTED PERCENTILE METHOD
Y=A(1-C*EXP(-B*X*X))
SAMPLE SIZE IS 12
SAMPLE SIZE IS 20
SAMPLE SIZE IS 30
WIDTHS FOR PARAMETER B
0.0102
0.0094
0.0086
0.0078
0.0070
0.0062
0.0054
0.0046
0.0038
0.0030
100
200
300
400
500
600
700
800
900
1000
NUMBER OF BOOTSTRAP REPLICATIONS

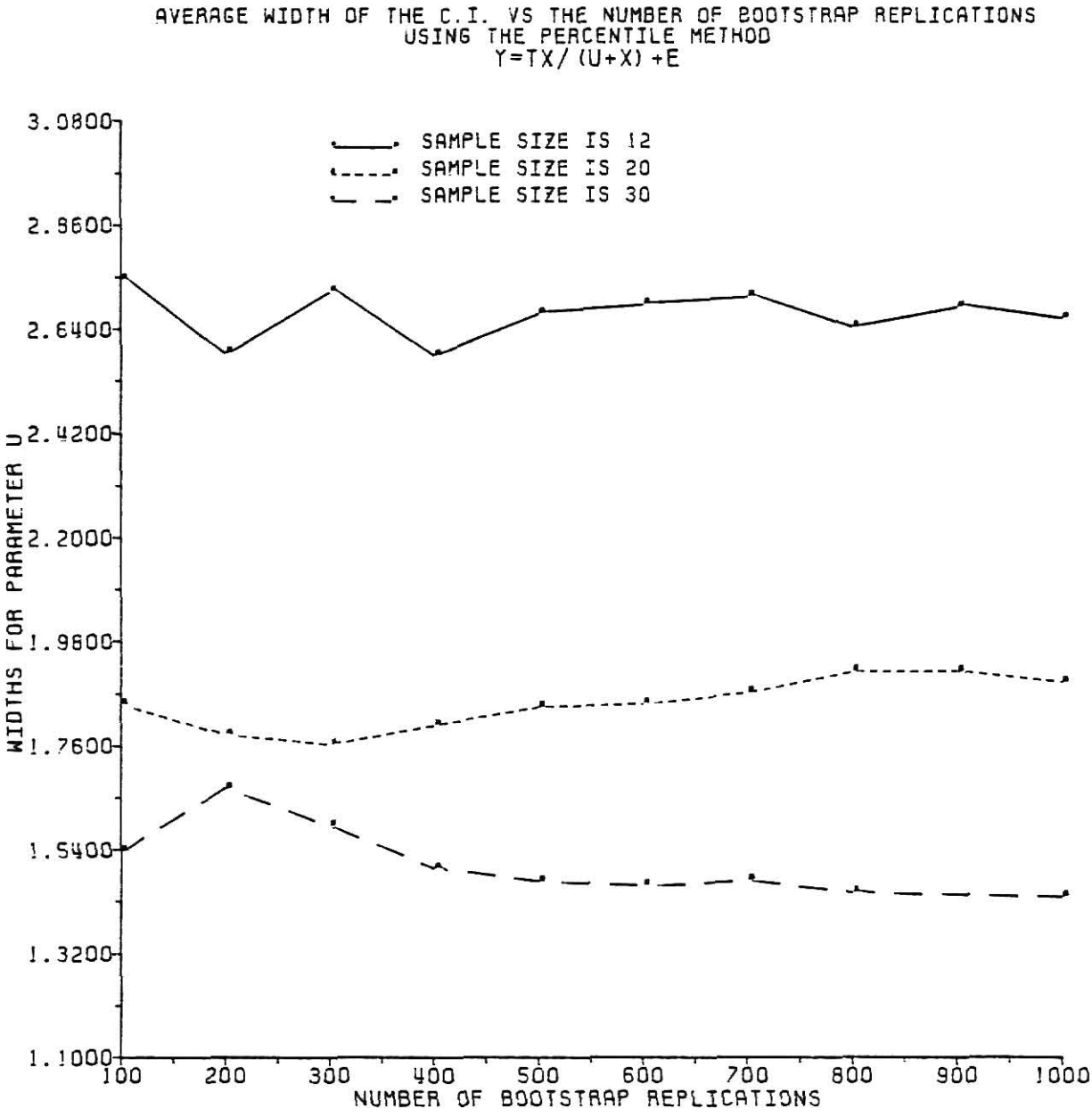
PLOT G4


AVERAGE WIDTH OF THE C.I. VS THE NUMBER OF BOOTSTRAP REPLICATIONS
USING THE PERCENTILE METHOD
Y=A(1-C*EXP(-B*X*X))
SAMPLE SIZE IS 12
SAMPLE SIZE IS 20
SAMPLE SIZE IS 30
WIDTHS FOR PARAMETER C
0.0220
0.0200
0.0180
0.0160
0.0140
0.0120
0.0100
0.0080
0.0060
0.0040
100
200
300
400
500
600
700
800
900
1000
NUMBER OF BOOTSTRAP REPLICATIONS

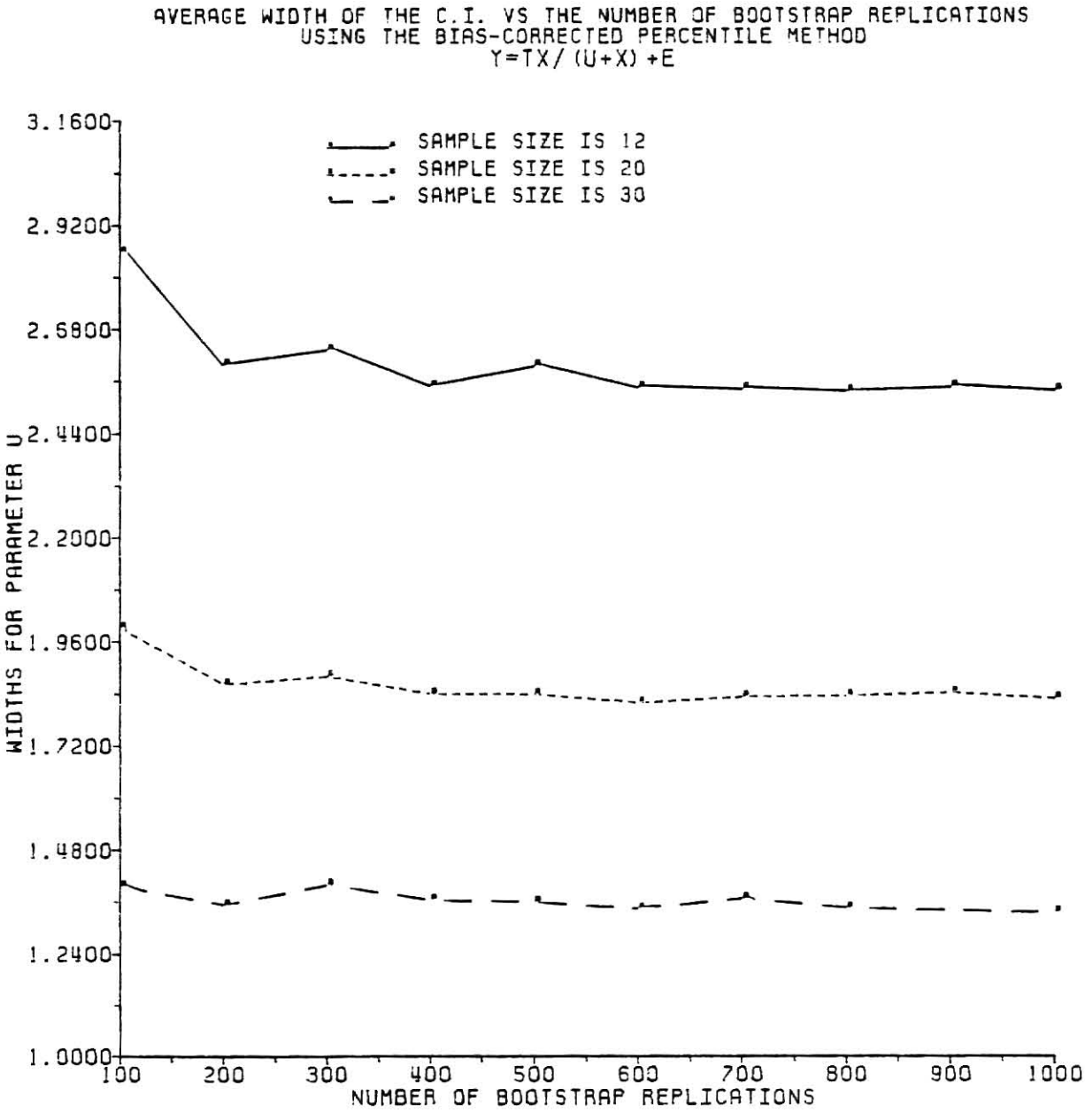
PLOT G5


AVERAGE WIDTH OF THE C.I. VS THE NUMBER OF BOOTSTRAP REPLICATIONS
USING THE BIAS-CORRECTED PERCENTILE METHOD
Y=A(1-C*EXP(-B*X*X))
SAMPLE SIZE IS 12
SAMPLE SIZE IS 20
SAMPLE SIZE IS 30
WIDTHS FOR PARAMETER C
0.0166
0.0152
0.0138
0.0124
0.0110
0.0096
0.0082
0.0068
0.0054
0.0040
100
200
300
400
500
600
700
800
900
1000
NUMBER OF BOOTSTRAP REPLICATIONS

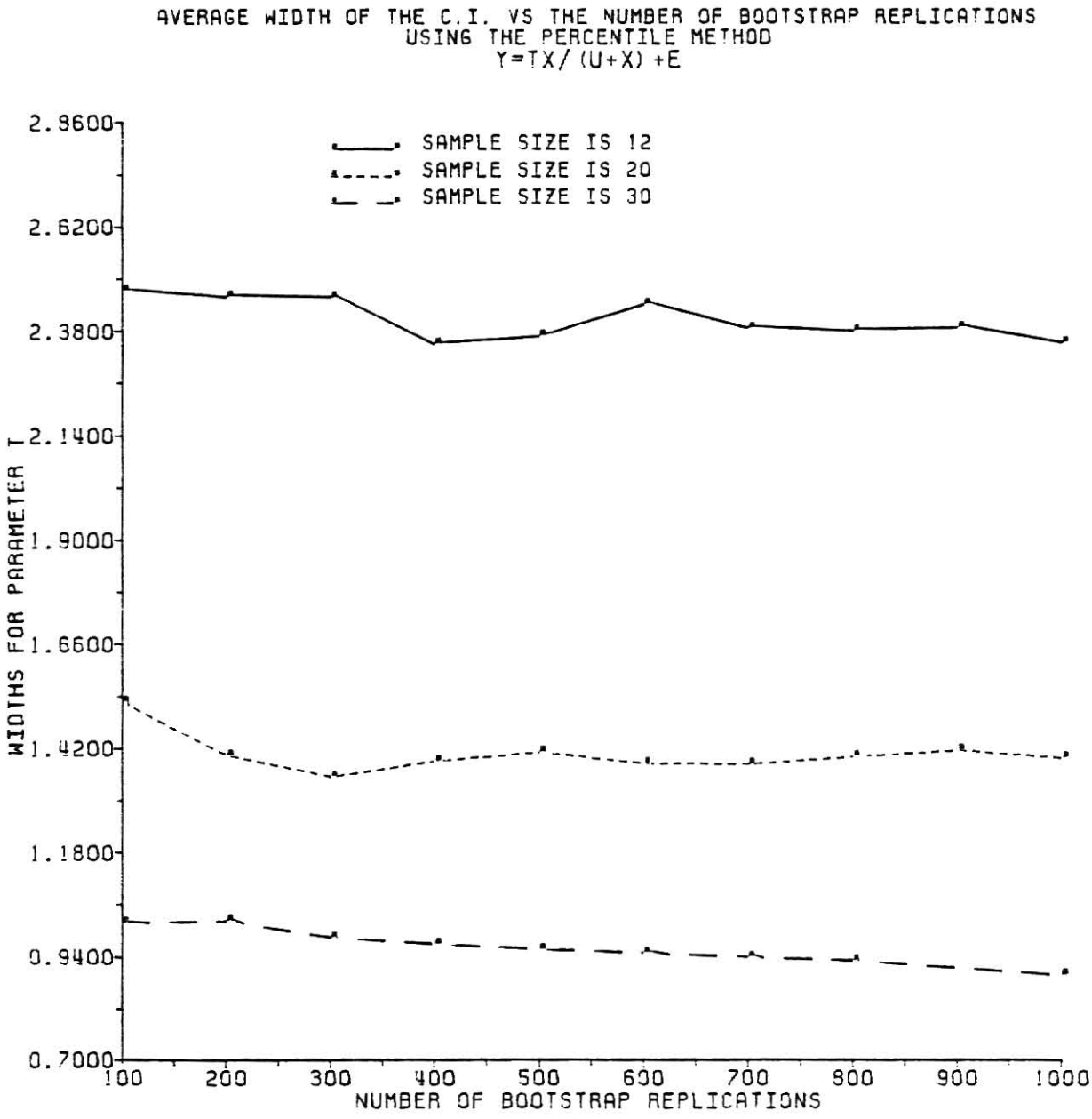
PLOT G6


AVERAGE WIDTH OF THE C.I. VS THE NUMBER OF BOOTSTRAP REPLICATIONS
USING THE PERCENTILE METHOD
Y=TX/(U+X)+E
SAMPLE SIZE IS 12
SAMPLE SIZE IS 20
SAMPLE SIZE IS 30
3.0800
2.8600
2.6400
2.4200
2.2000
1.9800
1.7600
1.5400
1.3200
1.1000
WIDTHS FOR PARAMETER U
100
200
300
400
500
600
700
800
900
1000
NUMBER OF BOOTSTRAP REPLICATIONS

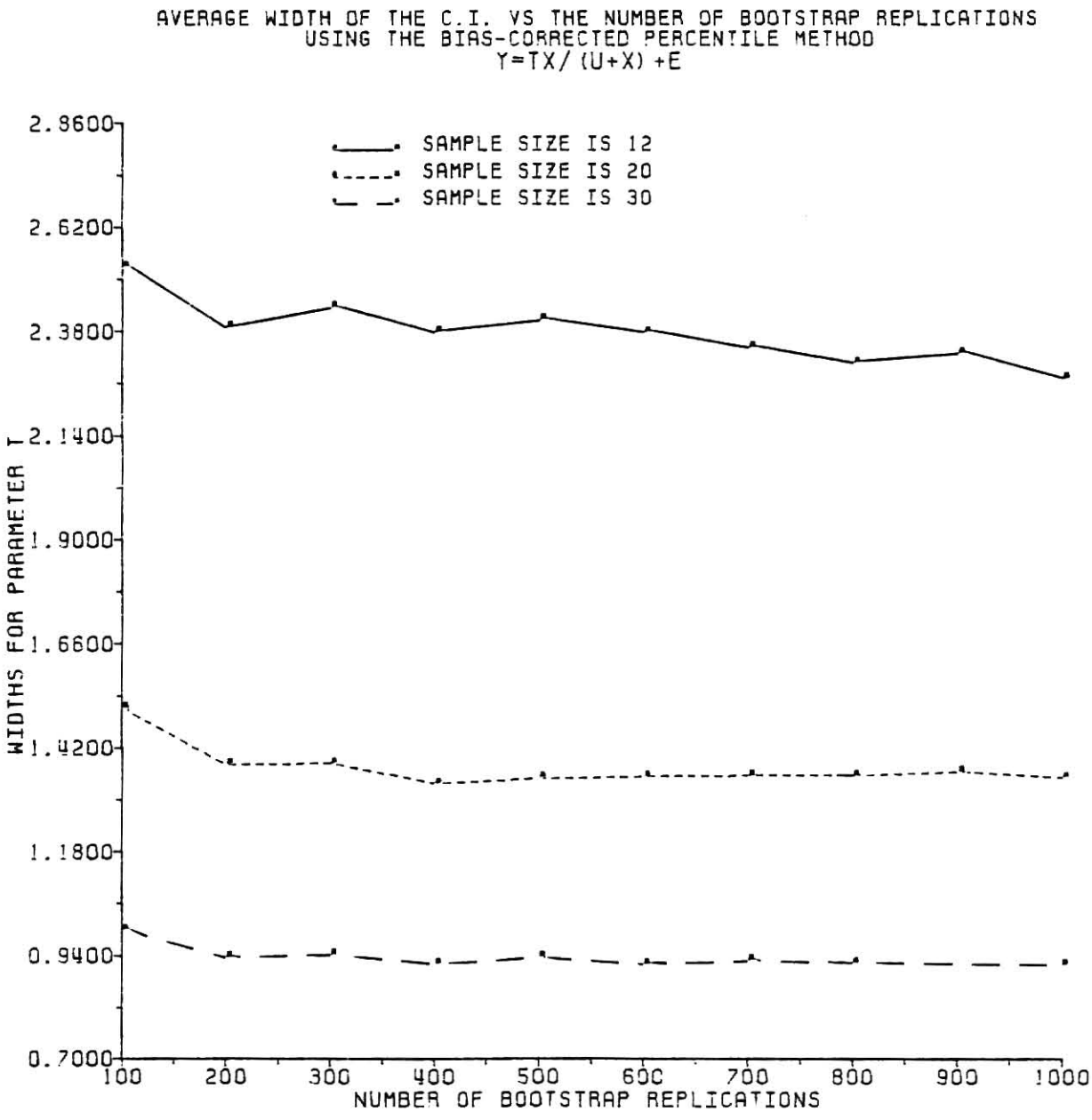
PLOT K1


AVERAGE WIDTH OF THE C.I. VS THE NUMBER OF BOOTSTRAP REPLICATIONS
USING THE BIAS-CORRECTED PERCENTILE METHOD
Y=TX/(U+X)+E
SAMPLE SIZE IS 12
SAMPLE SIZE IS 20
SAMPLE SIZE IS 30
WIDTHS FOR PARAMETER U
3.1600
2.9200
2.6800
2.4400
2.2000
1.9600
1.7200
1.4800
1.2400
1.0000
100
200
300
400
500
600
700
800
900
1000
NUMBER OF BOOTSTRAP REPLICATIONS

PLOT K2


AVERAGE WIDTH OF THE C.I. VS THE NUMBER OF BOOTSTRAP REPLICATIONS
USING THE PERCENTILE METHOD
Y=TX/(U+X)+E
SAMPLE SIZE IS 12
SAMPLE SIZE IS 20
SAMPLE SIZE IS 30
WIDTHS FOR PARAMETER T
2.8600
2.6200
2.3800
2.1400
1.9000
1.6600
1.4200
1.1800
0.9400
0.7000
100
200
300
400
500
600
700
800
900
1000
NUMBER OF BOOTSTRAP REPLICATIONS

PLOT K3


AVERAGE WIDTH OF THE C.I. VS THE NUMBER OF BOOTSTRAP REPLICATIONS
USING THE BIAS-CORRECTED PERCENTILE METHOD
Y=TX/(U+X)+E
SAMPLE SIZE IS 12
SAMPLE SIZE IS 20
SAMPLE SIZE IS 30
WIDTHS FOR PARAMETER T
2.8600
2.6200
2.3800
2.1400
1.9000
1.6600
1.4200
1.1800
0.9400
0.7000
100
200
300
400
500
600
700
800
900
1000
NUMBER OF BOOTSTRAP REPLICATIONS

PLOT K4

Appendix B

The following program and subroutines were used to generate nonlinear confidence intervals and estimates of the standard deviation of the parameters. Subroutines NONLIN, GAUSS, DERIV, SSQS, MMULT and FUNC were obtained from (8). The Sweep subroutine for inverting a matrix was suggested by (5). Subroutine NDTR and NDTRI are part of IBM's Scientific Subroutine Package.

```
      IMPLICIT REAL*8(A-H,O-Z)
      INTEGER OBS
      COMMON DELTA(4),BHAT(4),OBS,NVAR,NPARM,IHALT,N,NCONST,NCAP,NITER
      REAL*8 LCB95(4),MSE,LCBB95(4)
      DIMENSION FRMT(20), F(40), B(4), P(160), PT(160), VAR(4), SE(4), B
     1ETA(4), SIGHAT(4), BSSQ(4), BSUM(4), CDF(4), Z(4), BMEAN(4), D(500
     2), BSTAR(500,4), AWB95(4), UCBB95(4), SSCP(4), SSCPC(4), A(2,2), L
     3(40), M(40), XMODEL(10), X(40), E(40), BSUMC(4), Y(40), YSTAR(40),
     4 IRB95(4), IGB95(4), IBCT95(4), WB95(4), WID95(4), ICNT95(4), AWB8
     595(4), R95(4), P95(4), WBB95(4), IBPC95(4), AWID95(4), UCB95(4), G
     695(4), PG95(4)
C
      DATA ICNT95,AWID95,BSUMC,SSCPC,AWB95,IBCT95,AWBB95,IBPC95/4*0,4*0.
     10D0,12*0.0D0,4*0,4*0.0D0,4*0/,DJ1,DJ2,SIGMA2/1.0D0,2.0D0,.25D0/
      JY = 845245731
      IY = 254368425
      C = .7071067800
      IHALT = 0
C
      CALL RSTART (IY,JY)
C
C               READ IN AND WRITE OUT PARAMETER AND THE DATA CARDS.
C
      READ (5,1) XMODEL
    1 FORMAT (10A4)
    2 CONTINUE
      READ (5,3,END=51) OBS,NPARM,NVAR,NITER,NCONST,ITERB,NCAP
    3 FORMAT (I2,I2,I2,I2,I2,I4,I4)
      WRITE (6,4) OBS,XMODEL,NITER
    4 FORMAT ('ONUMBER OF OBSERVATIONS= ',I2//' FORM OF THE MODEL IS: ',
     110A4//' MAXIMUM NUMBER OF ITERATIONS POSSIBLE FOR CONVERGENCE= ',I
     22)
      WRITE (6,5) NCAP,ITERB
    5 FORMAT ('ONCAP=',I5,2X,'ITERB=',I5)
      READ (5,6) (B(I),I=1,NPARM)
      READ (5,6) (BETA(I),I=1,NPARM)
    6 FORMAT (8F10.5)
      WRITE (6,7) (B(I),I=1,NPARM)
    7 FORMAT ('0ORIGINAL PARAMETER ESTIMATES B(I)=',3(10X,F12.5))
      WRITE (6,8) (BETA(I),I=1,NPARM)
    8 FORMAT ('0TRUE PARAMETERS          BETA(I)=',3(10X,F12.5))
      READ (5,9) FRMT
    9 FORMAT (20A4)
      WRITE (6,10)
   10 FORMAT ('0INPUT DATA')
      DO 11 I=1,OBS
      NPARMN = NPARM*OBS
      N = (NVAR-1)*OBS+I
      READ (5,FRMT) X(I)
   11 WRITE (6,12) X(I)
   12 FORMAT (' ',10X,F10.5)
C
      IR95 = .025D0*NCAP
      IG95 = NCAP-IR95+1
      OBSINV = 1.0D0/DFLOAT(OBS)
      SIGMA = DSQRT(SIGMA2)
C
C               BEGIN THE NUMBER OF TRIALS
C
```

```
      DO 40 IT=1,ITERB
      ESUM = 0.0D0
      ESQ = 0.0D0
      DO 13 I=1,OBS
      ERROR = SIGMA*RNOR(0)
      Y(I) = BETA(1)*DEXP(-BETA(2)*X(I))+ERROR
   13 CONTINUE
C
C               USE GAUSS NEWTON METHOD TO GET ESTIMATES OF B(I).
C
      CALL NONLIN (X,B,P,SS,Y)
      NP = OBS-NPARM
      MSE = SS/NP
C
      DO 14 I=1,NPARM
      BSUMC(I) = BSUMC(I)+B(I)
      SQ = DJ1*B(I)-BSUMC(I)
      SSCPC(I) = SSCPC(I)+SQ*SQ/DJ2
   14 CONTINUE
      DJ1 = DJ1+1.000
      DJ2 = DJ1*(DJ1-1.0D0)
C



C               OBTAIN P TRANSPOSE,PT.
C
      CALL TRANSP (P,PT,OBS,NPARM,NPARMN)
C
C               MULTIPLY PT*P.  CALL IT A.
C               SUBROUTINE SWEEP TAKES THE INVERSE OF AN
C               IROW X IROW MATRIX.
C
      CALL MMULT (A,PT,P,NPARM,OBS,NPARM)
      DO 15 KK=1,NPARM
      CALL SWEEP (KK,NPARM,NPARM,A)
   15 CONTINUE
      K = (NPARM-1)*NPARM
C
C               USING CLASSICAL METHODS TO COMPUTE STANDARD DEVIATION
C               AND C.I.'S FOR B(I)
C
      DO 16 I=1,NPARM
      VAR(I) = DSQRT(A(I,I)*MSE)
      SE(I) = VAR(I)
      UCB95(I) = B(I)+1.9600*SE(I)
      LCB95(I) = B(I)-1.9600*SE(I)
      WID95(I) = 2.000*1.9600*SE(I)
      AWID95(I) = AWID95(I)+WID95(I)
   16 IF (LCB95(I).LT.BETA(I).AND.UCB95(I).GT.BETA(I)) ICNT95(I) = ICNT9
     15(I)+1
C
      WRITE (6,17)
   17 FORMAT ('-',132('*'))
      WRITE (6,18)
   18 FORMAT (' CLASSICAL METHOD')
      WRITE (6,19) IT,(VAR(J),J=1,NPARM)
   19 FORMAT (' ITER=',I4,5X,' SD(B(I))=',4(F15.10))
```

```
      DO 21 I=1,NPARM
      WRITE (6,20) LCB95(I),I,UCB95(I)
   20 FORMAT ('095*100% C.I. :',5X,'(',F17.7,' < B(',I1,') <',F15.7,')')
   21 CONTINUE
C
      DO 22 I=1,NPARM
      CDF(I) = 0.0D0
      BHAT(I) = B(I)
      BSSQ(I) = 0.0D0
   22 BSUM(I) = 0.0D0
      DN1 = 1.0D0
      DN2 = 2.0D0
C
C               FROM VECTOR OF EHAT VALUES, E(I), SELECT N EHAT TO
C               FORM VECTOR ESTAR(I)
C
      DO 23 I=1,OBS
      E(I) = Y(I)-(BHAT(1)*DEXP(-BHAT(2)*X(I)))
      ESUM = ESUM+E(I)
      ESQ = ESQ+E(I)*E(I)
   23 CONTINUE
      EMEAN = ESUM*OBSINV
      SIGMAE = ESQ/DSQRT(OBS-NPARM)
C
C               BEGIN BOOTSTRAP REPLICATIONS.
C
      DO 27 IB=1,NCAP
      DO 24 I=1,OBS
      U = UNI(0)
      IW = OBS*U
      JW = IW+1
      ESTAR = E(JW)
      YSTAR(I) = BHAT(1)*DEXP(-BHAT(2)*X(I))+ESTAR
   24 CONTINUE
C
      CALL NONLIN (X,B,P,SS,YSTAR)
C
C               INSERT THE BOOTSTRAP ESTIMATES, B(I) INTO
C               THE MATRIX BSTAR(IB,I)
C               - KEEPING A RUNNING SUM.
C               - CALCULATE THE SUMS OF SQUARES FOR EACH REPLICATION.
C
      DO 25 I=1,NPARM
      BSTAR(IB,I) = B(I)
      BSUM(I) = BSUM(I)+B(I)
      BSQ = DN1*B(I)-BSUM(I)
      BSSQ(I) = BSSQ(I)+BSQ*BSQ/DN2
   25 CONTINUE
      DN1 = DN1+1.0D0
      DN2 = DN1*(DN1-1.0D0)
C
      DO 26 I=1,NPARM
      IF (BSTAR(IB,I).LT.BHAT(I)) CDF(I) = CDF(I)+1.0D0
   26 CONTINUE
   27 CONTINUE
      WRITE (6,28) IHALT
   28 FORMAT ('0IHALT=',I7)
C
```

```
      DO 29 I=1,NPARM
      CDF(I) = CDF(I)/NCAP
      VAR(I) = DSQRT(BSSQ(I)/NCAP)
      BMEAN(I) = BSUM(I)/NCAP
   29 CONTINUE
C
C               MODIFICATIONS FOR THE BIAS-CORRECTED
C               PERCENTILE METHOD
C
      CALL NDTRI (CDF,Z,D,IE)
C
      DO 30 I=1,NPARM
      R95(I) = (2.0D0*Z(I))-1.96D0
      G95(I) = (2.0D0*Z(I))+1.96D0
   30 CONTINUE
C
      CALL NDTR (R95,P95,D,NPARM)
C
      DO 31 I=1,NPARM
      IRB95(I) = P95(I)*NCAP
      IGB95(I) = PG95(I)*NCAP
   31 CONTINUE
      CALL NDTR (G95,PG95,D,NPARM)
C
      DO 33 I=1,NPARM
      SESTAR = VAR(I)
      DO 32 IB=1,NCAP
      D(IB) = BHAT(I)-BSTAR(IB,I)
   32 CONTINUE
C
C               CONSTRUCT BOOTSTRAP CONFIDENCE INTERVALS.
C
      CALL DIS (D,NCAP,C)
C
      LCB95(I) = D(IR95)+BHAT(I)
      UCB95(I) = D(IG95)+BHAT(I)
      WB95(I) = D(IG95)-D(IR95)
      AWB95(I) = AWB95(I)+WB95(I)
      LCBB95(I) = D(IRB95(I))+BHAT(I)
      UCBB95(I) = D(IGB95(I))+BHAT(I)
      WBB95(I) = UCBB95(I)-LCBB95(I)
      AWBB95(I) = AWBB95(I)+WBB95(I)
      IF (LCB95(I).LT.BETA(I).AND.UCB95(I).GT.BETA(I)) IBCT95(I) = IBCT9
     15(I)+1
      IF (LCBB95(I).LT.BETA(I).AND.UCBB95(I).GT.BETA(I)) IBPC95(I) = IBP
     1C95(I)+1
   33 CONTINUE
C
      WRITE (6,34)
   34 FORMAT ('0BOOTSTRAP METHOD')
      WRITE (6,35) IT,(VAR(J),J=1,NPARM)
   35 FORMAT (' ITER=',I4,5X,' SD(B(I))=',4(F15.10))
      DO 36 I=1,NPARM
      WRITE (6,20) LCB95(I),I,UCB95(I)
   36 CONTINUE
      WRITE (6,37)
   37 FORMAT ('0BIAS - CORRECTION PERCENTILE METHOD')
      DO 38 I=1,NPARM
      WRITE (6,39) LCBB95(I),I,UCBB95(I)
```

```
   39 FORMAT ('095*100% C.I. :',5X,'(',F17.7,' < B(',I1,') <',F15.7,')')
   38 CONTINUE
   40 CONTINUE
      DO 42 I=1,NPARM
      SIGHAT(I) = SSCPC(I)/ITERB
      WRITE (6,41) SIGHAT(I)
   41 FORMAT ('0SIGHAT(I) IS THE SS OF THE DEVIATIONS OF ALL BHAT FROM T
     1HEIR MEAN   SIGHAT(I)=',F16.8)
      AWID95(I) = AWID95(I)/ITERB
      AWB95(I) = AWB95(I)/ITERB
      AWBB95(I) = AWBB95(I)/ITERB
   42 CONTINUE
      WRITE (6,17)
      DO 45 I=1,NPARM
      WRITE (6,43) I,ICNT95(I)
   43 FORMAT ('0FOR ALPHA = .05    B(',I1,')    IS IN THE CLASSICAL C.I.
     1    ',I4,'  TIMES.')
C
      WRITE (6,44) I,AWID95(I)
   44 FORMAT ('0FOR ALPHA = .05  THE AVERAGE WIDTH OF THE C.I. USING THE
     1 CLASSICAL METHOD  FOR    B(',I1,') IS ',F10.5)
   45 CONTINUE
      DO 48 I=1,NPARM
      WRITE (6,46) I,IBCT95(I)
   46 FORMAT ('0FOR ALPHA=.05   B(',I1,')   IS IN THE BOOTSTRAP C.I.
     1',I4,'   TIMES.')
      WRITE (6,47) I,AWB95(I)
   47 FORMAT ('0FOR ALPHA=.05  THE AVERAGE WIDTH OF THE C.I. USING THE P
     1ERCENTILE METHOD FOR  B(',I1,') IS  ',F10.5)
   48 CONTINUE
      WRITE (6,37)
      DO 50 I=1,NPARM
      WRITE (6,46) I,IBPC95(I)
      WRITE (6,49) I,AWBB95(I)
   49 FORMAT ('0FOR ALPHA=.05 THE AVERAGE WIDTH OF THE C.I. USING THE BI
     1AS - CORRECTION PERCENTILE METHOD FOR  B(',I1,')  IS ',F10.5)
   50 CONTINUE
      GO TO 2
   51 CONTINUE
C
      STOP
      END
```

```
      SUBROUTINE GAUSS (P,F,B,IERR3,SS,X,Y)
C
C               SUBROUTINE GAUSS PERFORMS THE GAUSS-NEWTON METHOD TO
C               OBTAIN PARAMETER ESTIMATES FOR NONLINEAR MODELS.
C
      IMPLICIT REAL*8(A-H,O-Z)
      INTEGER OBS
      COMMON DELTA(4),BHAT(4),OBS,NVAR,NPARM,IHALT,N,NCONST,NCAP,NITER
      DIMENSION P(160), F(40), B(4), A(2,2), PT(160), YF(40), TEMP(4), L
     1(40), M(40), B1(4), B2(4), W(24), X(40), Y(40)
C
C               OBTAIN PT, TRANSPOSE OF P.  P IS THE VECTOR
C               OF PARTIALS EVALUATED AT X AND AT BO.
C
      NOBS = OBS
      CALL TRANSP (P,PT,OBS,NPARM,NPARMN)
C
C               YF = Y-F.
C
      DO 1 I=1,NOBS
    1 YF(I) = Y(I)-F(I)
C
C               MULTIPLY PT*YF.  CALL IT TEMP.
C
      CALL MMULT (TEMP,PT,YF,NPARM,NOBS,1)
C
C               MULTIPLY PT*P.  CALL IT A.
C
      CALL MMULT (A,PT,P,NPARM,NOBS,NPARM)
C
C
C               OBTAIN INVERSE(PT*P).  USED TO CALCULATE
C               DELTA = INVERSE(PT*P)*PT*(Y-F)
C
      DO 2 KK=1,NPARM
      CALL SWEEP (KK,NPARM,NPARM,A)
    2 CONTINUE
C
C               MULTIPLY INVERSE(PT*P)*PT*YF.  CALL IT DELTA.
C
      CALL MMULT (DELTA,A,TEMP,NPARM,NPARM,1)
      V = 1.00
    3 V1 = .5D0*V
      DO 4 I=1,NPARM
      B1(I) = B(I)+V1*DELTA(I)
    4 B2(I) = B(I)+V*DELTA(I)
      CALL FUNC (B1,F,X)
      CALL SSQS (SS1,F,Y)
      CALL FUNC (B2,F,X)
      CALL SSQS (SS2,F,Y)
      TEMPS = SS2-SS1-SS1+SS
      IF (TEMPS.LT.1.0D-16) GO TO 5
      V1 = (.5D0+.25D0*(SS-SS2)/TEMPS)*V
    5 DO 6 I=1,NPARM
    6 B1(I) = B(I)+V1*DELTA(I)
      CALL FUNC (B1,F,X)
      CALL SSQS (SS1,F,Y)
      IF (SS1.LT.SS) GO TO 7
      IF (V.LT..01D0) GO TO 9
      V = .5D0*V
```

```
      SUBROUTINE NONLIN (X,B,P,SS,Y)
      IMPLICIT REAL*8(A-H,O-Z)
      INTEGER OBS
      COMMON DELTA(4),BHAT(4),OBS,NVAR,NPARM,IHALT,N,NCONST,NCAP,NITER
      DIMENSION F(40), B(4), P(160), PT(160), A(2,2), L(40), M(40), X(40
     1), E(40), Y(40)
C               FUNC: RETRIEVES FUNCTION VALUES FOR INITIAL GUESSES
C               OF THE PARAMETERS,B(I).
C               SSQS: CALCULATES ESS FOR INITIAL GUESSES OF B(I).
C
      CALL FUNC (B,F,X)
      CALL SSQS (SS,F,Y)
C
      N = NPARM*NPARM
      ITER = 0
C
C               BEGIN ITERATIONS OF GAUSS-NEWTON METHOD TO OBTAIN
C               PARAMETER ISTIMATES.  MAXIMUM ITERATIONS IS NITER.
C
    1 ITER = ITER+1
      IF (ITER.GT.NITER) GO TO 6
C
C               OBTAIN VECTOR OF PARTIALS EVALUATED AT BO,CURRENT
C               PARAMETER ESTIMATES.
C
      CALL DERIV (P,F,B,X)
      IERR3 = 1
      ITERM = 0
      CALL GAUSS (P,F,B,IERR3,SS,X,Y)
C
C               CHECK FOR CONVERGENCE.  SET ITERM=1 IF CONVERGENCE
C               CONFIRMED.
C
      DO 2 I=1,NPARM
      IF ((DABS(DELTA(I))/(.001D0+DABS(B(I)))).GT.1.D0-6) GO TO 3
    2 CONTINUE
      GO TO 4
    3 ITERM = 1
    4 CONTINUE
C
      GO TO (5,8), IERR3
C
C               END GAUSS-NEWTON METHOD ITERATIONS
C
    5 IF (ITERM.GT.0) GO TO 1
      GO TO 9
    6 WRITE (6,7) NITER
    7 FORMAT ('-PROCEDURE DOES NOT CONVERGE AFTER ',I3,'ITERATIONS.')
      GO TO 9
    8 IHALT = IHALT+1
C
C               EVALUATE MODEL AT FINAL PARAMETER ESTIMATES AND OBTAIN
C               VECTOR OF PARTIALS,P.
C
    9 CALL FUNC (B,F,X)
      CALL DERIV (P,F,B,X)
      RETURN
      END
```

```
      GC TC 3
    7 DO 8 I=1,NPARM
    8 B(I) = B1(I)
      SS = SS1
      GC TO 10
    9 IERR3 = 2
   10 RETURN
      END
```

```
      SUBROUTINE SSQS (SS,F,Y)
C
C               SUBROUTINE SSQS CALCULATES THE ESS BASED ON GIVEN
C               VALUES OF THE PARAMETERS, B(I).
C
      IMPLICIT REAL*8(A-H,O-Z)
      INTEGER OBS
      COMMON DELTA(4),BHAT(4),OBS,NVAR,NPARM,IHALT,N,NCONST,NCAP,NITER
      DIMENSION F(40), Y(40)
      NOBS = OBS
      CALL ERRSET (207,256,-1,1)
      SS = 0.0D0
      DO 1 I=1,NOBS
      SS = SS+(Y(I)-F(I))*(Y(I)-F(I))
      CALL OVERFL (J)
      IF (J.EQ.1) GO TO 2
    1 CONTINUE
      GO TO 3
    2 SS = 1.D74
    3 CALL ERRSET (207,256,256,2)
      RETURN
      END
```

```
      SUBROUTINE DERIV (P,F,B,X)
C
C               SUBROUTINE DERIV COMPUTES THE MATRIX OF
C               PARTIAL DERIVATIVES, P.
C
      IMPLICIT REAL*8(A-H,O-Z)
      INTEGER OBS
      COMMON DELTA(4),BHAT(4),OBS,NVAR,NPARM,IHALT,N,NCONST,NCAP,NITER
      DIMENSION B(4), B0(4), F(40), F0(40), P(160), X(40)
      NOBS = OBS
      DO 3 I=1,NPARM
      DO 1 J=1,NPARM
    1 B0(J) = B(J)
      H = B(I)*.00100
      H1 = 1.00/H
      B0(I) = B(I)+H
      CALL FUNC (B0,F0,X)
      DO 2 J=1,NOBS
    2 P((I-1)*NOBS+J) = (F0(J)-F(J))*H1
    3 CONTINUE
      RETURN
      END
```

```
      SUBROUTINE FUNC (B,F,X)
C
C               SUBROUTINE FUNC EVALUATES THE MODEL FOR INPUT
C               MATRIX GIVEN PARAMETER VALUES,B(I).
C
      IMPLICIT REAL*8(A-H,O-Z)
      INTEGER OBS
      COMMON DELTA(4),BHAT(4),OBS,NVAR,NPARM,IHALT,N,NCONST,NCAP,NITER
      DIMENSION B(4), F(40), X(40)
      DO 1 I=1,OBS
      F(I) = B(1)*DEXP(-B(2)*X(I))
    1 CONTINUE
      RETURN
      END
```

```
      SUBROUTINE TRANSP (YV,YVT,IROW,JCOLM,NPARMN)
C
C               SUBROUTINE TRANSP TAKES THE TRANSPOSE OF A
C               IROW X JCOLM MATRIX Y.  CALL IT YT.
C
      IMPLICIT REAL*8(A-H,O-Z)
      DIMENSION YV(NPARMN), YVT(NPARMN)
      JI = 0
      DO 1 I=1,IROW
      IJ = I-IROW
      DO 1 J=1,JCOLM
      IJ = IJ+IROW
      JI = JI+1
      YVT(JI) = YV(IJ)
    1 CONTINUE
      RETURN
      END
```

```
      SUBROUTINE MMULT (AB,A,B,L,M,N)
C
C               SUBROUTINE MMULT MULTIPLIES 2 INPUT MATRICES,
C               A AND B, TOGETHER.  THE RESULT IS AB.
C
      IMPLICIT REAL*8(A-H,O-Z)
      DIMENSION AB(L,N), A(L,M), B(M,N)
      DO 3 I=1,L
      DO 2 J=1,N
      AB(I,J) = 0.0
      DO 1 K=1,M
      AB(I,J) = A(I,K)*B(K,J)+AB(I,J)
    1 CONTINUE
    2 CONTINUE
    3 CONTINUE
      RETURN
      END
```

```
      SUBROUTINE SWEEP (K,IROW,JCOLM,A)
      IMPLICIT REAL*8(A-H,O-Z)
      REAL*8 A(IROW,JCOLM),OFDIAG,DIAG
      N = IROW-1
      DIAG = A(K,K)
      DO 1 J=1,JCOLM
    1 A(K,J) = A(K,J)/DIAG
      DO 3 I=1,IROW
      IF (I.EQ.K) GO TO 3
      OFDIAG = A(I,K)
      DO 2 J=1,JCOLM
      A(I,J) = A(I,J)-OFDIAG*A(K,J)
    2 A(I,K) = -OFDIAG/DIAG
    3 CONTINUE
      A(K,K) = 1.000/DIAG
      RETURN
      END
```

```
      SUBROUTINE DIS (K,N,R)
      IMPLICIT REAL*8(A-H,O-Z)
      INTEGER T,S,H,H1,SS
      REAL*8 K(N),R(N),KK,RR
      XN = 2*N*1
      T = DLOG(XN)/DLOG(3.0D0)-1.0001D0
      IF (T.LT.1) T = 1
      DO 4 SS=1,T
      S = T-SS+1
      H = (3**S-1)/2
      H1 = H+1
      DO 3 J=H1,N
      I = J-H
      RR = R(J)
      KK = K(J)
    1 IF (KK.GE.K(I)) GO TO 2
      R(I+H) = R(I)
      K(I+H) = K(I)
      K(I) = KK
      R(I) = RR
      I = I-H
      IF (I.GT.0) GO TO 1
    2 K(I+H) = KK
      R(I+H) = RR
    3 CONTINUE
    4 CONTINUE
      RETURN
      END
```

```
      SUBROUTINE NDTRI (P,Z,D,IE)
      IMPLICIT REAL*8(A-H,O-Z)
      COMMON DELTA(4),BHAT(4),OBS,NVAR,NPARM,IHALT,N,NCONST,NCAP,NITER
      DIMENSION P(4), Z(4)
      DO 10 I=1,NCAP
      IE = 0
      Z(I) = .99999E+74
      D = Z(I)
      IF (P(I)) 1,3,2
    1 IE = -1
      GO TO 10
    2 IF (P(I)-1.0D0) 5,4,1
    3 Z(I) = -.999999E+74
    4 D = 0.0D0
      GO TO 10
    5 D = P(I)
      IF (D-0.5D0) 7,7,6
    6 D = 1.0D0-D
    7 T2 = DLOG(1.0D0/(D*D))
      T = DSQRT(T2)
      Z(I) = T-(2.515517D0+0.802853D0*T+0.010328D0*T2)/(1.0D0+1.432788D0
     1*T+0.189269D0*T2+0.001308D0*T*T2)
      IF (P(I)-0.5D0) 8,8,9
    8 Z(I) = -Z(I)
    9 D = 0.3989423D0*DEXP(-Z(I)*Z(I)/2.0D0)
   10 CONTINUE
      RETURN
      END
```

```
      SUBROUTINE NDTR (X,P,D,NPARM)
      IMPLICIT REAL*8(A-H,O-Z)
      DIMENSION X(NPARM), P(NPARM)
      DO 2 I=1,NPARM
      AX = DABS(X(I))
      T = 1.0D0/(1.0D0+.2316419*AX)
      D = 0.3989423*DEXP(-X(I)*X(I)/2.0D0)
      P(I) = 1.0D0-D*T*(((((1.330274D0*T-1.821256D0)*T+1.781478D0)*T-0.35
     1656380)*T+0.3193815D0)
      IF (X(I)) 1,2,2
    1 P(I) = 1.0D0-P(I)
    2 CONTINUE
      RETURN
      END
```

A COMPARISON OF NORMAL THEORY AND BOOTSTRAP CONFIDENCE INTERVALS ON THE PARAMETERS OF NONLINEAR MODELS

by

MARY MARGARET ELLING

B.S., Kansas State University, 1979

AN ABSTRACT OF A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Statistics

KANSAS STATE UNIVERSITY

1981

Abstract

Confidence intervals for the parameters of three nonlinear models were constructed using both the classical normal theory approach and a percentile method, which is an empirical method using the bootstrap technique. Comparisons were made based on the proportion of times the confidence interval contained the true parameter and the average width of the interval using a Monte Carlo simulation. Models with both normal and nonnormal error structure showed the normal theory method performing well with confidence levels within 1% of the nominal 95% level. However, the bootstrap confidence intervals had lower coverage presumably due to their shorter average width.