Development and Applications of the Quantile Regression Estimation (QRE) Procedure

Dale P. Lifson

Inst. of Statistics Mimeo Series # 1399

TABLE OF CONTENTS

LIST OF TABLES

LIST OF TABLES (Continued)

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

The primary intent of this work is to introduce the quantile regression estimation (QRE) procedure as a means of estimating the parameters of the 3-parameter lognormal distribution. Since 1963, when Hill uncovered difficulties with maximum likelihood estimation of this distribution, several alternative estimation procedures have been suggested. These include local maximum likelihood estimation [Harter and Moore (1966), Wingo (1975), and Cohen and Whitten (1980)], estimation by a regression on expectations of order statistics [Munro and Wixley 1970)], discrete maximum likelihood estimation [Giesbrecht and Kempthorne (1976)], and modified maximum likelihood and modified moments estimation [Harter and Moore (1966), and Cohen and Whitten (1980)]. The procedure proposed here is for application to selected sample quantiles, whereas many of the above procedures apply best to complete samples.

The QRE procedure is an extension of an idea presented in Cohen and Whitten. It is so named because estimates of the parameters of an assumed distributional form are obtained from a regression of sample quantiles on their asymptotic expected values. The regression is typically nonlinear and involves iteratively reweighted least squares estimation techniques. While the procedure was developed in connection with the 3-parameter lognormal distribution, it has similar applications to related distributions such as the 3-parameter log equicorrelated-normal and S_B (4-parameter lognormal) distributions. Each of these applications involves approximating the expected quantile values with their asymptotic values. The expected covariance matrix of the regression residuals is also computed using the asymptotic joint distribution of the quantiles. Finally, the procedure may be used in a somewhat more straightforward fashion to estimate the parameters of many continuous probability distributions whose inverse distribution functions are known in closed form.

The paper is organized into five chapters, including this introduction. Chapter 2 reviews the literature on estimation of the 3-parameter lognormal distribution so that the proposed quantile regression estimation procedure can be placed in perspective. In Chapter 3, the QRE procedure is developed as it applies to the 3-parameter lognormal distribution and the necessary equations for applying the QRE procedure to this and several other distributions are derived. In addition, a conditional version of the QRE procedure is developed as an alternative means of estimating the 3-parameter lognormal distribution. The performance of the QRE procedure is empirically investigated in Chapter 4, where estimation is performed on real and generated data from many of the distributions discussed in Chapter 3. The conditional version of the QRE procedure is also examined empirically in Chapter 4. A summary and conclusions are presented in Chapter 5.

CHAPTER 2

LITERATURE REVIEW--HISTORY OF ESTIMATION IN THE 3-PARAMETER LOGNORMAL DISTRIBUTION

In this chapter, the important contributions to estimation in the 3-parameter lognormal distribution are reviewed in chronological order. The purpose of the review is to enable the reader to evaluate the quantile regression estimation procedure proposed here in relation to previously established estimation methods.

In the following review there are several discussions that refer to either the density function or the likelihood function associated with the 3-parameter lognormal distribution. If the random variable X follows the 3-parameter lognormal distribution with location parameter γ, scale parameter μ, and shape parameter σ, then the density function of X is given by

$$f_X(x;\gamma,\mu,\sigma) = (2\pi\sigma^2)^{-\frac{1}{2}}(x-\gamma)^{-1} \exp\{-[\ln(x-\gamma) - \mu]^2/2\sigma^2\} \qquad (2.1)$$

$$\text{for } \sigma^2 > 0, \quad \gamma < x < \infty$$

$$= 0 \quad \text{otherwise.}$$

The logarithm of the likelihood function of a sample $x_1, \ldots, x_n$ from the 3-parameter lognormal distribution is given by

$$\ln L(x_1,\ldots,x_n;\ \gamma,\mu,\sigma) = \qquad (2.2)$$

$$-(n/2)\ln(2\pi\sigma^2) - (1/2\sigma^2) \sum_{i=1}^{n} [\ln(x_i-\gamma)-\mu]^2 - \sum_{i=1}^{n} \ln(x_i-\gamma)\ .$$

In the review that follows, the differing notations employed by the various authors are changed to match that of Eq. (2.1) and (2.2) above.

Yuan (1933). Yuan was the first to provide a thorough accounting of the properties of the 3-parameter lognormal distribution. In addition to deriving expressions for a variety of measures of central tendency, dispersion, skewness and kurtosis, he was the first to derive the method of moments estimators of the three parameters (γ, μ and σ). The moment estimators are obtained by solving the following system of equations for γ, μ and σ:

$$\begin{aligned} \bar{x} &= \gamma + \omega^{\frac{1}{2}}\exp(\mu) \\ s^2 &= \omega(\omega-1)\exp(2\mu) \\ a_3 &= (\omega+2)(\omega-1)^{\frac{1}{2}} \end{aligned} \tag{2.3}$$

where $\omega = \exp(\sigma^2)$, and $\bar{x}$, s^2, and a_3 are the sample mean, variance and third standard moment, respectively.

The solution is found by first solving for ω in the cubic equation

$$\omega^3 + 3\omega^2 - (a_3^2 + 4) = 0. \tag{2.4}$$

By Descartes's "rule of signs", only one of the roots of this equation is real, and this root will be taken as the estimate ω^* of ω. The other estimators are easily defined in terms of ω^* according to

$$\begin{aligned} \sigma^{2*} &= \ln\omega^*, \\ \mu^* &= \ln\{s[\omega^*(\omega^*-1)]^{-\frac{1}{2}}\}, \text{ and} \\ \gamma^* &= \bar{x} - s(\omega^*-1)^{-\frac{1}{2}} . \end{aligned} \tag{2.5}$$

Wilson and Worcester (1945) and Cohen (1951). These authors independently developed what they thought were maximum likelihood estimators of γ, μ and σ. However, Hill (1963) showed that the solution to the likelihood equations corresponded to the local rather than global maximum. Therefore, the estimators developed by Wilson and Worcester and also by Cohen were actually local maximum likelihood estimators. The logarithm of the likelihood equation is given in Eq. (2.2). These authors derived the (local) maximum likelihood estimators [(L)MLE's] in the usual way, setting the partial derivatives of lnL with respect to each of the three parameters equal to zero and then solving the resulting system of equations for the parameters. Each easily derived the relationships

$$\hat{\mu}(\gamma) = (1/n) \sum_{i=1}^{n} \ln(x_i-\gamma) \text{ and} \tag{2.6}$$

$$\hat{\sigma}^2(\gamma) = (1/n) \sum_{i=1}^{n} [\ln(x_i-\gamma) - \hat{\mu}(\gamma)]^2$$

so that the (L)MLE's of μ and σ could be calculated once the (L)MLE of γ was obtained. Substitution of Eqs. (2.6) into the equation $\partial \ln L/\partial\gamma = 0$, gives the expression

$$\sum_{i=1}^{n} (x_i-\gamma)^{-1}\{n \sum_{i=1}^{n} \ln(x_i-\gamma) - n \sum_{i=1}^{n} \ln^2(x_i-\gamma) + [\sum_{i=1}^{n} \ln(x_i-\gamma)]^2\}$$
$$- n^2 \Sigma [\ln(x_i-\gamma)/(x_i-\gamma)] = 0. \tag{2.7}$$

In each paper, it was recommended that $\hat{\gamma}$ be found by employing an iterative searching algorithm to solve the above equation for γ.

Each also computed what they thought were the asymptotic variances and covariances of the (L)MLE's by inverting the Fisher information matrix. The validity of this procedure in the present context is in doubt because the range of the random variable depends on the parameter γ. The information matrix is given by

$$-E\left[\frac{\partial^2 \ln L}{\partial\theta_i \partial\theta_j}\right] = I(\theta) = I(\gamma,\mu,\sigma)$$

$$= \begin{bmatrix} \frac{\omega^2(1+\sigma^2)}{\mu^2\sigma^2} & \frac{\omega^{\frac{1}{2}}}{\mu^2\sigma^2} & \frac{-2\omega^{\frac{1}{2}}}{\mu\sigma} \\ \frac{\omega^{\frac{1}{2}}}{\mu^2\sigma^2} & \frac{1}{\mu^2\sigma^2} & 0 \\ -\frac{2\omega^{\frac{1}{2}}}{\mu\sigma} & 0 & \frac{2}{\sigma^2} \end{bmatrix} . \quad (2.8)$$

Moreover, the convergence in probability of $I(\hat{\theta})$ to $I(\theta)$ has never been proven. Using the notation devised by Cohen and Whitten (1980), let $A = [\omega(1+\sigma^2) - (1+2\sigma^2)]^{-1}$ and $C = \exp(\mu)/\omega^{\frac{1}{2}}$. If the probability limit of $I(\hat{\theta})$ were equal to $I(\theta)$, then the asymptotic variances and covariances of the parameters, taken from the appropriate elements of $[I(\gamma,\mu,\sigma)]^{-1}$, would be given by

$$\begin{aligned} V(\hat{\gamma}) &= (\sigma^2/n)AC^2 \\ V(\hat{\mu}) &= (\sigma^2/n)(1+A) \\ V(\hat{\sigma}) &= (\sigma^2/n)(1+2\sigma^2 A) \\ Cov(\hat{\gamma},\hat{\mu}) &= -(\sigma^2/n)AC \\ Cov(\hat{\gamma},\hat{\sigma}) &= (\sigma^3/n)AC \\ \text{and} \quad Cov(\hat{\mu},\hat{\sigma}) &= -(\sigma^3/n)A. \end{aligned} \quad (2.9)$$

Aitchison and Brown (1957). Aitchison and Brown authored a book devoted to the lognormal distribution which consisted of a review of virtually everything that was then known about the distribution, including methods of estimation. Perhaps their most significant contribution to parameter estimation was their introduction of the method of quantiles (MQ) estimation procedure. They recommended that the three quantiles X_α, $X_{.5}$, and $X_{1-\alpha}$ be equated to their population values and the resulting system of equations be solved to obtain the parameter estimates. The choice of a symmetric pattern of quantiles permitted a direct, rather than iterative solution to the system of equations. Letting $c_{1-\alpha} = \Phi^{-1}(1-\alpha)$ where Φ is the inverse of the standard normal distribution function, the method of quantiles estimates corresponding to X_α, $X_{.5}$ and $X_{1-\alpha}$ are given by

$$\hat{\sigma} = c_{1-\alpha}^{-1}[\ln(X_{1-\alpha}-X_{.5}) - \ln(X_{.5}-X_\alpha)], \qquad (2.10)$$

$$\hat{\mu} = \ln(X_{.5}-X_\alpha) - \ln[1-\exp(-\hat{\sigma}c_{1-\alpha})], \text{ and}$$

$$\hat{\gamma} = X_{.5} - \exp(\hat{\mu}).$$

Aitchison and Brown recommended that α be set equal to .05 and noted that the method fails if $\hat{\sigma} < 0$.

Hill (1963). Hill discovered that the solution of the likelihood equations leads to local, rather than global maximum likelihood estimates. In his proof, reproduced in Appendix A, he shows that there exists a path in the parameter space along which the likelihood of any ordered sample $x_1, \ldots, x_n$ from the 3-parameter lognormal distribution approaches $+\infty$. By defining estimates of μ and σ in terms of γ [see Eq. (2.6)] the likelihood function is seen to approach $+\infty$ as $[\gamma,\hat{\mu}(\gamma;x_1, \ldots, x_n),\hat{\sigma}(\gamma;x_1, \ldots, x_n)]$ converges to $[x_1,-\infty,+\infty]$. Hill states on p. 75 "that in a meaningful sense $[x_1,\hat{\mu}(x_1),\hat{\sigma}^2(x_1)] = (x_1,-\infty,+\infty)$ is the maximum likelihood estimate of (γ,μ,σ^2) although in fact the likelihood at that point is zero."

To sidestep the issue of local vs. global MLE's, Hill used Bayesian techniques to obtain 3PLN parameter estimates. He was able to construct a posterior distribution for γ, μ, and σ which had its greatest probability density corresponding to a γ value near the local maximum of the likelihood function rather than at x_1. Hill's results are compared with those from other estimation procedures in Section 4.2 of this thesis.

Lambert (1964). Lambert computed LMLE's for 23 generated 3PLN samples of size 32 and compared them with method of quantiles (MQ) estimates. His contribution was to apply the method of scoring to the computation of LMLE's. The scoring method uses the fact that the partial derivatives of the likelihood function with respect to the parameters equal zero at the point where the likelihood attains a local maximum. Thus, given a set of initial parameter estimates, it is possible to iterate to the LMLE solution by successively calculating corrected parameter estimates which satisfy second order Taylor's series approximations to the likelihood equations. Lambert proposed three methods of obtaining initial values, each of which lead to nearly identical final estimates. Strangely, he did not publish the results of the estimation on the 23 generated samples nor did he comment on the relative performances of the MQ and LMLE procedures.

Harter and Moore (1966). Harter and Moore investigated local maximum likelihood estimation in censored and uncensored samples from the 3PLN distribution. The censoring was considered for the first r and last n-m order statistics. They introduced a modification to the maximum likelihood estimation procedure for use when the standard procedure led to the global rather than local maximum. They simply censored the smallest uncensored sample value, x_{r+1}, and modified the likelihood equations accordingly. By using x_1 as an upper bound on the possible values of γ, censoring x_1 removed the problem of the likelihood function becoming unbounded. They noted that for their distribution (γ=10, μ=4, σ=2), the modified procedure was only necessary for some of the doubly-censored samples when the sample size was small (n=100). They suggested that the modification might also be required for uncensored samples of even smaller sample sizes. In their Monte Carlo study the modified estimates seemed to have properties similar to those of the conventional LMLE's, but their study was not designed to efficiently compare the two estimation procedures.

Munro and Wixley (1970). Munro and Wixley were concerned with 3PLN parameter estimation in very small samples (n=10, 15 and 20). They developed an iteratively reweighted least squares regression procedure in which the sample order statistics were regressed on their approximate expected values. Using a different parameterization than that given in Eq. (2.1) they were able to express the expectation of each order statistic as a linear function of the expectation of its corresponding standard lognormal order statistics, which depended only on the shape parameter σ.

If the value of σ were known, the remaining two parameters could be readily estimated as the intercept and slope estimates from the weighted least-squares regression. In the more common case with σ unknown, two types of approximations were required before the estimation could be performed. A first order Taylor's series expansion about σ was used to approximate the expectations of the sample order statistics, and a procedure devised by Blom (1958) was used to approximate the means, variances, and covariances of the corresponding standard lognormal order statistics.

Their weighted least squares procedure is similar to the quantile regression estimation (QRE) procedure to be proposed in this thesis. The two procedures differ in that the QRE procedure is a weighted least squares regression on sample quantiles, rather than on the full set of order statistics. The use of quantiles permits the derivation of asymptotic expectations of the first and second moments of the quantiles, thereby eliminating the need for the moment approximation techniques required in Munro and Wixley's method.

O'Neill and Wells (1972). O'Neill and Wells used the scoring method [see Lambert (1964)] to compute (local) maximum likelihood estimates of 2PLN and 3PLN parameters from grouped samples of insurance claim payment data. They concluded that the scoring method worked very well.

They then computed the efficiencies of the estimates based on two different grouping strategies relative to those obtainable from ungrouped data. The relative efficiency of a particular grouping was defined as the ratio of the large sample variance of the estimate using that grouping to the large sample variance of the estimate obtained using ungrouped data. These variances were computed from the inverses of the appropriate Fisher information matrices. The validity of their calculations depends upon the convergence in probability of the matrix $I(\hat{\theta})$ to its true value $I(\theta)$. While this result has never been proven, it is supported by the empirical results of Lambert (1964) and others.

For their claim payment data, estimates of both 2PLN and 3PLN parameters were clearly more efficient if the data were grouped according to logarithmic increments rather than equal increments. The advantage of logarithmic grouping appeared greater when fewer groups were formed.

Calitz (1973). In this empirical study, Calitz compared the efficiency of various estimation procedures in estimating the shape parameter σ of the 3PLN distribution. Using 50 generated samples from each of five distributions in which σ varied from 0.3 to 1.1, he computed estimates using the methods of maximum likelihood, moments, quantiles of order 5% and 95%, and optimal quantiles. Based on the empirical variance of the estimates of σ, Calitz concluded that the local maximum likelihood estimator was far superior to the other three, and that when σ was large the method of moments was much more precise than either of the quantile methods.

Wingo (1975). The objective of Wingo's work was to find an algorithm that permitted local maximum likelihood parameter estimates to be found from any 3PLN or S_B[1] sample. That is, he sought an algorithm that would guarantee the attainment of the local rather than the anomalous global solution of the likelihood equations.

The method Wingo chose was to use interior penalty function techniques which transformed the constrained optimization problem (solve for $\hat{\gamma}$ subject to $\hat{\gamma} < x_1$) into an unconstrained problem. For the 3PLN estimation, the unconstrained (transformed) objective was to maximize the penalty function

$$P(\gamma,r) = \ln L(\gamma) - r[(\gamma+c)^{-1} + (x_1-\gamma)^{-1}] \qquad (2.11)$$

where $L(\gamma)$ was the likelihood function in terms of $\hat{\mu}(\gamma)$ and $\hat{\sigma}(\gamma)^2$, c was

[1]The S_B distribution is also known as the 4-parameter lognormal distribution. The density function is given in Section 3.3 of this thesis.

any large positive number (say, 10^{25}) and r was a small number (say ≤ 1). The maximization of Eq. (2.11) is accompanied by successive reductions in the value of r. When r is relatively large, the second term in Eq. (2.11) is relatively large which ensures that a solution of $\hat{\gamma}=x_1$ or $\hat{\gamma}\to-\infty$ will be avoided. Once the iterative process moves in the direction of attaining the local maximum of $\ln L(\gamma)$ the value of r is decreased to permit the value of $P(\gamma,r)$ to be dominated by the value of $\ln L(\gamma)$.

The penalty function used to estimate the two location parameters of the S_B (4-parameter lognormal) distribution was analogous to Eq. (2.11). Using data from Hill (1963), Tiku (1968), Lambert (1964) and Iwai (1950), Wingo found that the penalty function approach always converged to a reasonable solution and thus concluded that the algorithm was robust.

Giesbrecht and Kempthorne (1976). Giesbrecht and Kempthorne pointed out that previously suggested methods of 3PLN parameter estimation have ignored the possible effects on the parameter estimates of grouping error in the data. They felt that when data are recorded within class intervals of width δ the likelihood function should properly be defined as

$$L(\gamma,\mu,\sigma) = K \prod p_i^{f_i} \qquad (2.12)$$

---

2 $$\hat{\mu}(\gamma) = \Sigma \ln(x_i-\gamma)/n,$$

$$\hat{\sigma}(\gamma) = \Sigma [\ln(x_i-\gamma) - \hat{\mu}(\gamma)]^2/n, \text{ and}$$

$$\ln L(\gamma) = -n[\hat{\mu}(\gamma) + \ln\hat{\sigma}(\gamma)] + \text{a constant}.$$

where K is a constant, f_i is the frequency of observations in the ith interval and p_i is defined by

$$p_i = \int_a^b (2\pi\sigma^2)^{-\frac{1}{2}}(x-\gamma)^{-1}\exp\{-[\ln(x-\gamma)-\mu]^2/2\sigma^2\}dx \qquad (2.13)$$

for $a = \max[(i-.5)\delta,\gamma]$ and $b = (i+.5)\delta$.

The authors computed maximum likelihood estimates for simulated data and for Hill's (1963) data using Eq. (2.12) and compared them to estimates obtained by maximizing the usual likelihood equation [Eq. (2.2)]. They found that in the case of the simulated data, the parameter estimates obtained using Eq. (2.12) were closer to the true parameter values than were those obtained using Eq. (2.2). Also, it will be seen in Chapter 4 of this thesis that the estimates they obtained for the Hill data provided a better fit to the data than did those that were obtained by solving the conventional likelihood equations. However, estimates from the quantile regression estimation (QRE) procedure will be seen to provide an even better fit to this data. The QRE procedure implicitly accounts for "grouping error" by assigning asymptotic expected values to the quantiles defined by the grouping.

Cohen and Whitten (1980). Cohen and Whitten introduced two modifications to the conventional maximum likelihood and moment estimation procedures. The two modifications involved replacing one of the likelihood equations, or one of the moment equations with either of two equations involving a small sample order statistic. Details of their methodology are given in Section 4.1 of this thesis. It is concluded there that none of their modifications yield estimators superior to the conventional local maximum likelihood estimators. Still, the ideas used in the development of these modified procedures helped motivate the development of the QRE procedure proposed in this thesis.

CHAPTER 3

DEVELOPMENT OF THE QRE PROCEDURE AND DERIVATION OF ESTIMATING EQUATIONS FOR SEVERAL APPLICATIONS

In this chapter, the quantile regression estimation (QRE) procedure is derived as a means of estimating the parameters of the 3-parameter lognormal distribution, some related distributions, and finally, many continuous distributions whose inverse distribution functions are known in a simple closed form. It is hoped that the relative ease of its application will make it attractive, especially when it is felt that condensing the data would be prudent.

The chapter is organized as follows. In Section 3.1, the QRE procedure is developed as it applies to the 3-parameter lognormal distribution. Following this development, related applications are derived in Sections 3.2 and 3.3 for the 3-parameter log equicorrelated-normal and S_B (4-parameter lognormal) distributions, respectively. These three uses of the QRE procedure rely on a technique that uses the moment generating function of the normal random variable to derive the asymptotic expectations of the sample quantiles. Then, in Section 3.4 the QRE procedure is developed in the more straightforward case where the asymptotic expectations of the quantiles can be obtained from the inverse of the distribution function. Regression equations are derived for the income distribution proposed by Singh and Maddala (1976) as an example. At the end of Section 3.4 a table is provided listing the QRE equations that pertain to several other probability distributions. Finally, in Section 3.5 a conditional quantile regression estimation procedure is derived as it applies to the 3-parameter lognormal distribution. In this procedure, the sample quantiles are regressed on their asymptotic conditional expectations, where the conditioning variable is the value of the preceding quantile.

3.1 The 3-Parameter Lognormal Distribution

The distribution of the random variable X is a member of the 3-parameter lognormal family of distributions if $\ln(X - \gamma) \sim N(\mu, \sigma^2)$. In this case, the density function of X is

$$f_X(x;\gamma,\mu,\sigma) = (2\pi\sigma^2)^{-\frac{1}{2}} (x - \gamma)^{-1} \exp\{-[\ln(x - \gamma) - \mu]^2/2\sigma^2\} \quad \text{for } \sigma^2 > 0,\ \gamma < x < \infty, \qquad = 0, \text{ otherwise.} \tag{3.1}$$

The ideas leading to the development of the QRE procedure came out of some modified maximum likelihood and modified moment estimation procedures proposed by Cohen and Whitten (1980). Their modifications derived from the fact that the lognormal variate X is related to the standard normal variate Z by the transformation

$$Z = [\ln(X - \gamma) - \mu]/\sigma. \tag{3.2}$$

Thus, if X_r is the rth order statistic from a random sample from the 3-parameter lognormal (3PLN) distribution, then $Z_r = [\ln(X_r - \gamma) - \mu]/\sigma$ is the rth order statistic of a corresponding sample from the standard normal distribution. The modifications proposed by Cohen and Whitten were to replace one of the three usual equations of the maximum likelihood or moment estimation procedures with the equation

$$\ln(X_r - \gamma) = \mu + \sigma E(Z_r) \tag{3.3}$$

where X_r is the rth sample order statistic. In the likelihood equations, $\partial \ln L/\partial \gamma = 0$ was replaced, and in the moment estimation, the equation for the third moment was replaced. Cohen and Whitten tried Eq. (3.3) with r = 1, 2, and 3. Details of their investigation are given in Section 1 of Chapter 4.

Whereas Cohen and Whitten's procedures utilized just one of the order statistics (at a time), one might consider using all of the order statistics to estimate the 3PLN parameters. This was attempted by Munro and Wixley (1970) for small samples. Under a different 3PLN parameterization they expressed the expectations of the 3PLN order statistics as linear in the transformed location (τ) and scale (υ) parameters ($\tau = \gamma + \mu$ and $\upsilon = \sigma\mu$), but nonlinear in the shape parameter σ. The expectations involved the first two moments of standard [i.e., $(\tau,\upsilon,\sigma) = (0,1,\sigma)$] lognormal variables and depended on the value of σ. However, in addition to the problem of σ being unknown, no expressions existed for the moments of the standardized variables and therefore approximations had to be used. Munro and Wixley used iteratively reweighted least squares techniques to obtain the parameter estimates. Their results seemed to display biased estimates of σ for sample sizes of 10, 15, and 20.

Another idea that uses the order statistics might be to exploit the monotonic transformation

$$X_r = \gamma + \exp(\mu + \sigma Z_r). \tag{3.4}$$

If the expectation of the right-hand side of Eq. (3.4) could be computed as a function of γ, μ and σ, then minimization of an appropriately weighted sum of $[X_r - E(X_r)]^2$ using weighted least squares techniques might be a good way to obtain estimates of the parameters. The estimates could be obtained by finding the weighted least squares solution to the regression equation

$$X_r = E(X_r) + w_r \; ; \qquad r = 1, 2, \ldots, n \tag{3.5}$$

where the w_r are correlated residuals. However, formidable computational problems arise with this approach. First consider the expectation on the right hand side of Eq. (3.5):

$$\begin{aligned} E(X_r) &= E[\gamma + \exp(\mu + \sigma Z_r)] \\ &= \gamma + \beta \; E[\exp(\sigma Z_r)], \end{aligned} \tag{3.6}$$

where $\beta=\exp(\mu)$. A closed-form expression does not exist for $E[\exp(\sigma Z_r)]$. Thus to perform the regression of Eq. (3.5), either costly and time-consuming Monte Carlo studies or other numerical integration procedures would have to be performed to compute the empirical distributions of $\exp(\sigma Z_r)$ for numerous values of σ and r, or else $E(X_r)$ and $Cov(X_r, X_s)$ would have to be approximated for $r,s = 1,\ldots,n$. The Monte Carlo results would be specific to the sample size, n, so that the Monte Carlo experiments would have to be redone each time n changed.

Due to these complications the suggestion is put forth to consider the regression equation (3.5) with sample quantiles in place of the sample order statistics. When the sample size n is large the joint distributions of the sample quantiles and the corresponding standard normal quantiles are approximately multivariate normal. The expectations, variances and covariances of the sample quantiles can be computed in closed form using these limiting distributions, making the quantile regression estimation procedure practical.

Since the data required for implementation of the QRE procedure are sample quantiles, these quantiles must first be selected from the complete set of data. The following definitions of population and sample quantiles, given in Sarhan and Greenberg (1962), will be used.

Let g(x) be any absolutely continuous probability density function.

Definition 3.1

The pth population quantile of g(x) is the value $X = c_p$ such that

$$\int_{-\infty}^{c_p} g(t)dt = p$$

Definition 3.2

The pth sample quantile of a sample of size n from density g(x) is defined as

$$\begin{aligned} X_p &= X_{np} \quad \text{if np is an integer} \\ &= X_{([np] + 1)} \quad \text{if np is not an integer,} \end{aligned}$$

where [np] denotes the largest integer $\leq$ np.

In order to compute the expectations required for performing quantile regressions, the limiting distributional properties of the sample quantiles must be derived. The following theorems, taken from Sarhan and Greenberg, specify these asymptotic distributions.

Theorem 3.1

If g(x) is differentiable in the neighborhood of $X = c_p$ and if $g(c_p) \neq 0$, then the distribution of the variate

$$[n/p(1 - p)]^{\frac{1}{2}} \; g(c_p) \; (X_p - c_p)$$

tends to N(0, 1) as n tends to infinity.

<u>Theorem 3.2</u>

For k given real numbers for which

$$0 < p_1 < p_2 < \cdots < p_k < 1,$$

let the p_i-quantile of the population be c_i, that is

$$\int_{-\infty}^{c_i} g(t)\,dt = p_i; \qquad i = 1, 2, \ldots, k.$$

Assume that the frequency function g(x) of the population is differentiable in the neighborhoods of $X = c_i$, $i = 1, \ldots, k$, and $g_i^* = g(c_i) \neq 0$, for $i = 1, \ldots, k$.

Then the joint distribution of $\{n^{\frac{1}{2}}(X_{n_i} - c_i)\}$, where $n_i=[np_i]+1$, $i=1, \ldots, k$ tends to a k-dimensional normal distribution with zero means and with covariance matrix

$$\begin{bmatrix} \frac{p_1(1-p_1)}{g_1^{*2}} & \frac{p_1(1-p_2)}{g_1^*g_2^*} & \cdots & \frac{p_1(1-p_k)}{g_1^*g_k^*} \\ \frac{p_1(1-p_2)}{g_1^*g_2^*} & \frac{p_2(1-p_2)}{g_2^{*2}} & \cdots & \frac{p_2(1-p_k)}{g_2^*g_k^*} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{p_1(1-p_k)}{g_1^*g_k^*} & \frac{p_2(1-p_k)}{g_2^*g_k^*} & \cdots & \frac{p_k(1-p_k)}{g_k^{*2}} \end{bmatrix}$$

as n tends to infinity.

If $Z_{p;n}$ is the pth quantile from a random sample of n standard normal variables then

$$\lim_{n\to\infty} E(Z_{p;n}) = c_p = \Phi^{-1}(p) \tag{3.7}$$

where Φ is the standard normal distribution function, and

$$\lim_{n\to\infty} n \operatorname{Var}(Z_{p;n}) = \frac{p(1-p)}{[\phi(c_p)]^2} , \tag{3.8}$$

where ϕ is the standard normal probability density function. Finally,

$$\lim_{n\to\infty} n \operatorname{Cov}(Z_p, Z_r) = \frac{p(1-r)}{\phi(c_p)\phi(c_r)}, \quad \text{for } p < r . \tag{3.9}$$

Now the regression of the 3-parameter lognormal quantiles on their asymptotic expected values may be derived, the result being closed form expressions for the regression equation and covariance matrix.

Let $X_{p1}, X_{p2}, \ldots, X_{pk}$ be k sample quantiles from a random sample of size n from the 3-parameter lognormal distribution, with $0 < p_1 < \ldots < p_k < 1$. Let $Z_{p1}, Z_{p2}, \ldots, Z_{pk}$ be defined by

$$Z_{pi} = \frac{\log(X_{pi}-\gamma)-\mu}{\sigma} . \tag{3.10}$$

Replacing the order statistics in Eq. (3.6) with the sample quantiles gives the regression equation

$$X_{pi} = \gamma + \beta\, E[\exp(\sigma Z_{pi})] + w_i . \tag{3.11}$$

Now, from Eq. (3.7) and (3.8) and noting that the expectation in Eq. (3.11) is the moment generating function of the asymptotically normal variate Z_{pi}, it follows that

$$E[\exp(\sigma Z_{pi})] \doteq \exp(\sigma c_i + \sigma^2 k_{ii}/2), \tag{3.12}$$

where

$$c_i = \Phi^{-1}(p_i) \text{ and } k_{ii} = \frac{p_i(1-p_i)}{n[\phi(c_i)]^2} .$$

Substituting Eq. (3.12) into Eq. (3.11) gives the nonlinear quantile regression equation

$$X_{pi} = \gamma + \beta \exp(\sigma c_i + \sigma^2 k_{ii}/2) + w_i . \tag{3.13}$$

The variances and covariances of the errors, w_i, may be computed as well, using similar techniques. The variances are computed from the standard computation formula

$$\begin{aligned} Var(w_i) &= Var(X_{pi}) \\ &= E(X_{pi})^2 - [E(X_{pi})]^2. \end{aligned} \tag{3.14}$$

An asymptotic expression for $E(X_{pi})$ is given by subtracting w_i from the right-hand side of Eq. (3.13). $E(X_{pi})^2$ is computed as

$$\begin{aligned} E(X_{pi})^2 &= E[\gamma + \beta\exp(\sigma Z_{pi})]^2 \\ &= E[\gamma^2 + 2\gamma\beta\exp(\sigma Z_{pi}) + \beta^2\exp(2\sigma Z_{pi})] \\ &= \gamma^2 + 2\gamma\beta E[\exp(\sigma Z_{pi})] + \beta^2 E[\exp(2\sigma Z_{pi})]. \end{aligned} \tag{3.15}$$

Asymptotic expressions for the two expectations in the last line of Eq. (3.15) may be obtained by again recognizing them as moment generating functions of asymptotically normal random variables, Z_{pi}. Thus,

$$E[\exp(\sigma Z_{pi})] \doteq \exp(\sigma c_i + \sigma^2 k_{ii}/2), \tag{3.16}$$

and

$$E[\exp(2\sigma Z_{pi})] \doteq \exp(2\sigma c_i + 2\sigma^2 k_{ii}) . \tag{3.17}$$

Substituting Eq. (3.16) and (3.17) into Eq. (3.15) gives

$$E(X_{pi})^2 \doteq \gamma^2 + 2\gamma\beta\exp(\sigma c_i + \sigma^2 k_{ii}/2) + \beta^2\exp(2\sigma c_i + 2\sigma^2 k_{ii}). \quad (3.18)$$

Subtracting the square of $E(X_{pi})$ in Eq. (3.13) from Eq. (3.18) gives the desired expression for the asymptotic variance of the sample quantiles:

$$\text{Var}(X_{pi}) \doteq \beta^2\exp(2\sigma c_i)\exp(\sigma^2 k_{ii})[\exp(\sigma^2 k_{ii}) - 1]. \quad (3.19)$$

The covariance between any two quantiles X_{pi} and X_{pj} is computed similarly, as follows.

$$\text{Cov}(X_{pi}, X_{pj}) = E(X_{pi}X_{pj}) - E(X_{pi})E(X_{pj}). \quad (3.20)$$

The first expectation in Eq. (3.20) is computed as

$$E(X_{pi}X_{pj}) = E\{[\gamma + \beta\exp(\sigma Z_{pi})][\gamma + \beta\exp(\sigma Z_{pj})]\} \quad (3.21)$$
$$= \gamma^2 + \gamma\beta E[\exp(\sigma Z_{pi}) + \exp(\sigma Z_{pj})] + \beta^2 E\{\exp[\sigma(Z_{pi} + Z_{pj})]\}.$$

In the last term, $Z_{pi} + Z_{pj}$ is a linear combination of asymptotically normal, correlated random variables. Thus,

$$Z_{pi} + Z_{pj} \,\dot{\sim}\, N(c_i + c_j, k_{ii} + k_{jj} + 2k_{ij}), \quad (3.22)$$

where

$$k_{ij} = [p_i(1-p_j)]/[n\phi(c_i)\phi(c_j)] \text{ for } p_i < p_j.$$

Again, making use of the normal moment generating function and Eq. (3.22) gives

$$E\{\exp[\sigma(Z_{pi} + Z_{pj})]\} \doteq \exp[\sigma(c_i + c_j) + \sigma^2(k_{ii} + k_{jj} + 2k_{ij})/2]. \quad (3.23)$$

Substituting Eq. (3.16) and Eq. (3.23) into Eq. (3.21) gives

$$E(X_{pi}X_{pj}) \doteq \gamma^2 + \gamma\beta\{\exp[\sigma(c_i + c_j)] + \exp[(\sigma^2 k_{ii} + \sigma^2 k_{jj})/2]\} + \beta^2 \exp\{\sigma(c_i + c_j) + \sigma^2[(k_{ii} + k_{jj} + 2k_{ij})/2]\}. \quad (3.24)$$

Expanding the product of the expectations in Eq. (3.20) gives

$$\begin{aligned} E(X_{pi})E(X_{pj}) &\doteq [\gamma + \beta\exp(\sigma c_i + \sigma^2 k_{ii}/2)][\gamma + \beta\exp(\sigma c_j + \sigma^2 k_{jj}/2)] \\ &= \gamma^2 + \gamma\beta\{\exp[\sigma(c_i + c_j)] \\ &\quad + \exp[(\sigma^2 k_{ii} + \sigma^2 k_{jj})/2]\} \\ &\quad + \beta^2 \exp[\sigma(c_i + c_j) + \sigma^2(k_{ii} + k_{jj})/2]. \end{aligned} \quad (3.25)$$

Finally, subtracting Eq. (3.25) from Eq. (3.24) gives

$$\mathrm{Cov}(X_{pi}, X_{pj}) \doteq \beta^2\{\exp[\sigma(c_i + c_j)]\exp[\sigma^2(k_{ii} + k_{jj})/2] \times [\exp(\sigma^2 k_{ij}) - 1]\}. \quad (3.26)$$

Defining, in matrix form for a given parameter vector $\theta=(\gamma, \mu, \sigma)$,

$$x = (x_{p1}, x_{p2}, \cdots, x_{pk})', \quad (3.27)$$

$$u(\theta) = [E_\theta(X_{p1}), E_\theta(X_{p2}), \cdots, E_\theta(X_{pk})]',$$

$$w(\theta) = x - u(\theta), \text{ and}$$

$$V(\theta) = E\ \{[w(\theta)]\ [w(\theta)]'\},$$

the solution vector $\hat{\theta} = (\hat{\gamma}, \hat{\mu}, \hat{\sigma})$ is that vector θ for which the corresponding weighted sum of squares of the errors [WSSE(θ)] is a minimum, where

$$\mathrm{WSSE}(\theta) = [x - u(\theta)]'\ [V(\theta)]^{-1}\ [x - u(\theta)]\ . \quad (3.28)$$

Since the asymptotic expression for the covariance matrix $V(\theta)$ depends on the parameters (θ) being estimated, the solution to Eq. (3.28) must be found via an iterative process. The estimation algorithm, described in detail in Section 1 of Appendix B, initially sets $V(\theta)$ equal to the identity matrix and uses the one-step Gauss-Newton procedure to compute an approximate ordinary least squares estimate of θ. This estimate of θ is used to compute a new estimate of $V(\theta)$ which is then used in the one-step Gauss-Newton procedure to re-estimate θ using weighted least squares. The process of recomputing an estimate of $V(\theta)$ and obtaining the corresponding weighted least squares estimate of θ continues until convergence in the parameters is attained. No guarantees that the algorithm will converge to the solution vector $\hat{\theta}$ have been discovered, although the algorithm performed very for well for the data herein.

The QRE procedure described above depends upon the first and second order moments of the distributions of the sample quantiles being in close agreement with the corresponding moments of the asymptotic distributions of the quantiles. The rate of convergence of the sample moments to their asymptotic expectations was examined in a Monte Carlo experiment reported in Appendix D. The relevant distribution to examine was that of $\exp(\sigma Z_{pi})$. This was done for various values of σ and p_i, and for various sample sizes n. The results of the experiment show close agreement between the sample moments and their limiting expectations for most values of σ and p_i when n is as small as 59.

3.2 The 3-Parameter Log Equicorrelated-Normal Distribution

Let Z_o, Z_1, ..., Z_n be n+1 independent standard normal random variables and consider the transformation

$$Z_i' = \rho^{\frac{1}{2}} Z_o + (1-\rho)^{\frac{1}{2}} Z_i;\ i=1, 2, \ldots, n \tag{3.29a}$$

for some ρ in the interval (0,1). Also, let Y_1, ..., Y_n be independent standard normal random variables and let Y_0 be distributed N(0,1) but with $E(Y_0 Y_i) = -(-\rho)^{\frac{1}{2}}/(1-\rho)^{\frac{1}{2}}$, and consider the transformation

$$Y_i' = (-\rho)^{\frac{1}{2}} Y_o + (1-\rho)^{\frac{1}{2}} Y_i;\ i=1, 2, \ldots, n \tag{3.29b}$$

for some ρ in the interval $[-(n-1)^{-1},0)$.

Each of these transformations induces the formation of n equicorrelated N(0, 1) random variables. The discussion that follows pertains to the case where $\rho > 0$. The random variables Z_i' may be shown to be equicorrelated and N(0,1) distributed as follows:

$$\begin{aligned} E(Z_i') &= E[\rho^{\frac{1}{2}} Z_o + (1-\rho)^{\frac{1}{2}} Z_i] \\ &= \rho^{\frac{1}{2}} E(Z_o) + (1-\rho)^{\frac{1}{2}} E(Z_i) = 0\ , \end{aligned} \tag{3.30a}$$

$$\begin{aligned} E(Z_i')^2 &= E[\rho^{\frac{1}{2}} Z_o + (1-\rho)^{\frac{1}{2}} Z_i]^2 \\ &= E[\rho Z_o^2 + 2\rho^{\frac{1}{2}}(1-\rho)^{\frac{1}{2}} Z_o Z_i + (1-\rho) Z_i^2] \\ &= \rho E(Z_o^2) + 2\rho^{\frac{1}{2}}(1-\rho)^{\frac{1}{2}} E(Z_o Z_i) + (1-\rho) E(Z_i^2) \\ &= \rho \cdot 1 + 0 + (1-\rho) \cdot 1 = 1\ , \end{aligned} \tag{3.30b}$$

and

$$E(Z_i' Z_j') = E\{[\rho^{\frac{1}{2}} Z_o + (1-\rho)^{\frac{1}{2}} Z_i]\ [\rho^{\frac{1}{2}} Z_o + (1-\rho)^{\frac{1}{2}} Z_j]\} \tag{3.30c}$$

$$= E[\rho Z_o^2 + \rho^{\frac{1}{2}}(1-\rho)^{\frac{1}{2}} Z_o(Z_i + Z_j) + (1-\rho)Z_i Z_j]$$

$$= \rho E(Z_o^2) + \rho^{\frac{1}{2}}(1-\rho)^{\frac{1}{2}}[E(Z_o Z_i) + E(Z_o Z_j)] + (1-\rho)E(Z_i Z_j)$$

$$= \rho \cdot 1 + \quad 0 \qquad\qquad + 0$$

$$= \rho \, .$$

Now consider a sample X_i, i=1, ..., n where the X_i are defined by the transformation

$$\frac{\log(X_i - \gamma) - \mu}{\sigma} = Z_i' \, . \tag{3.31}$$

The random variables X_i will be said to be distributed as 3-parameter log equicorrelated-normal (3PLEN) random variables. The QRE procedure is applicable in this equicorrelated situation, as will now be shown.

Just as in the uncorrelated case, in the equicorrelated case, Eq. (3.20) represents a monotonic transformation between the X_i and the original, uncorrelated Z_i. Indeed for any two values x_i and x_j

$$\begin{aligned} x_i < x_j &\rightarrow \log(x_i - \gamma) < \log(x_j - \gamma) \\ &\rightarrow \frac{\log(x_i - \gamma) - \mu}{\sigma} < \frac{\log(x_j - \gamma) - \mu}{\sigma} \quad \text{(since } \sigma > 0) \\ &\rightarrow z_i' < z_j' \\ &\rightarrow p^{\frac{1}{2}} z_o + (1-\rho)^{\frac{1}{2}} z_i < \rho^{\frac{1}{2}} z_o + (1-\rho)^{\frac{1}{2}} z_j \\ &\rightarrow z_i < z_j \quad [\text{since } (1 - p)^{\frac{1}{2}} > 0] \, . \end{aligned} \tag{3.32}$$

Therefore the rth order statistic X_r can be related to the rth standard normal order statistic Z_r by

$$\frac{\log(X_r-\gamma)-\mu}{\sigma} = Z_r' = \rho^{\frac{1}{2}}Z_o + (1-\rho)^{\frac{1}{2}}Z_r \tag{3.33}$$

or equivalently

$$X_r = \gamma + \exp[\mu + \sigma\rho^{\frac{1}{2}}Z_o + \sigma(1-\rho)^{\frac{1}{2}}Z_r]. \tag{3.34}$$

It follows then, that the same relationship exists between the two sets of quantiles:

$$X_{pi} = \gamma + \exp[\mu + \sigma\rho^{\frac{1}{2}}Z_o + \sigma(1-\rho)^{\frac{1}{2}}Z_{pi}]. \tag{3.35}$$

Thus, the quantile regression equation

$$X_{pi} = E(X_{pi}) + w_i \tag{3.36}$$

can be estimated provided that closed-form expressions for $E(X_{pi})$ and $E(w\ w')$ can be computed, where w is the vector $(w_1, \ldots, w_k)'$.

By making use of Theorems 3.1 and 3.2 and again using the expectation given by the normal moment generating function, these expectations may be computed with respect to the asymptotic distributions of the quantiles, giving

$$E(X_{pi}) \doteq \gamma + \beta \exp(\sigma^2\rho/2) \exp[\sigma(1-\rho)^{\frac{1}{2}}c_i] \exp[\sigma^2(1-\rho)k_{ii}/2] , \tag{3.37}$$

and

$$E(w\ w') = \begin{bmatrix} Var(X_{p1}) & Cov(X_{p1},X_{p2}) & \cdots & Cov(X_{p1},X_{pk}) \\ Cov(X_{p1},X_{p2}) & Var(X_{p2}) & \cdots & Cov(X_{p2},X_{pk}) \\ - - - & - - - & - - - & - - - \\ Cov(X_{p1},X_{pk}) & Cov(X_{p2},X_{pk}) & \cdots & Var(X_{pk}) \end{bmatrix}$$

where

$$\mathrm{Var}(X_{pi}) \doteq \beta^2 \exp(\sigma^2\rho)\ \exp[2\sigma(1-\rho)^{\frac{1}{2}}c_i]\ \exp[\sigma^2(1-\rho)k_{ii}] \tag{3.38}$$

$$\times\ \{\exp(\sigma^2\rho)\ \exp[\sigma^2(1-\rho)k_{ii}]-1\}$$

and

$$\mathrm{Cov}(X_{pi},X_{pj}) \doteq \beta^2\exp(\sigma^2\rho)\exp[\sigma(1-\rho)^{\frac{1}{2}}(c_i+c_j)]\exp[\sigma^2(1-\rho)(k_{ii}+k_{jj})/2] \tag{3.39}$$

$$\times\ \{\exp(\sigma^2\rho)\exp[\sigma^2(1-\rho)k_{ij}] - 1\}.$$

The asymptotic variances and covariances given above are derived as follows

$$\begin{aligned} \mathrm{Var}(X_{pi}) &= E(X_{pi})^2 - [E(X_{pi})]^2 \\ &= E[\gamma+\beta\exp(\sigma Z'_{pi})]^2 - \{E[\gamma+\beta\exp(\sigma Z'_{pi})]\}^2 \\ &= E[\gamma^2 + 2\gamma\beta\exp(\sigma Z'_{pi}) + \beta^2\exp(2\sigma Z'_{pi})] \\ &\qquad - \{\gamma + \beta E[\exp(\sigma Z'_{pi})]\}^2 \\ &= \gamma^2 + 2\gamma\beta E[\exp(\sigma Z'_{pi})] + \beta^2 E[\exp(2\sigma Z'_{pi})] \\ &\qquad - \{\gamma^2 + 2\gamma\beta E[\exp(\sigma Z'_{pi})] + \beta^2 E^2[\exp(\sigma Z'_{pi})]\}, \end{aligned} \tag{3.40}$$

which simplifies to

$$\mathrm{Var}(X_{pi}) = \beta^2\{E[\exp(2\sigma Z'_{pi})] - E^2[\exp(\sigma Z'_{pi})]\}, \tag{3.41}$$

After using Eq. (3.29) to substitute for Z'_{pi}, the expectations in Eq. (3.41) may be derived as

$$\begin{aligned} E[\exp(2\sigma Z'_{pi})] &= E\{\exp[2\sigma\rho^{\frac{1}{2}}Z_o + 2\sigma(1-\rho)^{\frac{1}{2}}Z_{pi}]\} \\ &= E\{\exp[2\sigma\rho^{\frac{1}{2}}Z_o]\exp[2\sigma(1-\rho)^{\frac{1}{2}}Z_{pi}]\} \\ &= E\{\exp[2\sigma\rho^{\frac{1}{2}}Z_o]\}E\{\exp[2\sigma(1-\rho)^{\frac{1}{2}}Z_{pi}]\}, \end{aligned} \tag{3.42}$$

the last step following due to the independence of Z_o and each Z_{pi}. In the last line of Eq. (3.42) the first expectation is the moment generating function of the normal random variable $2\sigma\rho^{1/2}Z_o$. Thus,

$$E\{\exp[2\sigma\rho^{1/2}Z_o]\} = \exp(2\sigma^2\rho). \tag{3.43a}$$

The second expectation in the last line of Eq. (3.42) may be approximated by treating it as the moment generating function of the asymptotically normal random variable $2\sigma(1-\rho)^{1/2}Z_{pi}$, giving

$$E\{\exp[2\sigma(1-\rho)^{1/2}Z_{pi}]\} \doteq \exp[2\sigma(1-\rho)^{1/2}c_i]\exp[2\sigma^2(1-\rho)k_{ii}]. \tag{3.43b}$$

Substituting Eqs. (3.43) into Eq. (3.42) gives

$$E[\exp(2\sigma Z'_{pi})] \doteq \exp(2\sigma^2\rho)\exp[2\sigma(1-\rho)^{1/2}c_i]\exp[2\sigma^2(1-\rho)k_{ii}]. \tag{3.44}$$

Taking the second expectation on the right-hand side of Eq. (3.41) gives

$$\begin{aligned} E^2[\exp(\sigma Z'_{pi})] &= E^2\{\exp[\sigma\rho^{1/2}Z_o + \sigma(1-\rho)^{1/2}Z_{pi}]\} \\ &= E^2\{\exp[\sigma\rho^{1/2}Z_o]\exp[\sigma(1-\rho)^{1/2}Z_{pi}]\} \\ &\doteq \exp(\sigma^2\rho/2)\exp[\sigma(1-\rho)^{1/2}c_i]\exp[\sigma^2(1-\rho)k_{ii}/2]. \end{aligned} \tag{3.45}$$

Finally, substituting Eqs. (3.44) and (3.45) into Eq. (3.41) gives

$$\begin{aligned} \mathrm{Var}(X_{pi}) \doteq\ & \beta^2\exp(\sigma^2\rho)\exp[2\sigma(1-\rho)^{1/2}c_i]\exp[\sigma^2(1-\rho)k_{ii}] \\ & \times \{\exp(\sigma^2\rho)\exp[\sigma^2(1-\rho)k_{ii}] - 1\}. \end{aligned} \tag{3.46}$$

Similar methodology is used to derive the covariances.

$$\begin{aligned} \mathrm{Cov}(X_{pi}, X_{pj}) &= E(X_{pi}X_{pj}) - E(X_{pi})E(X_{pj}) \\ &= E\{[\gamma + \beta\exp(\sigma Z'_{pi})][\gamma + \beta\exp(\sigma Z'_{pj})]\} \\ &\quad - E[\gamma + \beta\exp(\sigma Z'_{pi})]E[\gamma + \beta\exp(\sigma Z'_{pj})] \end{aligned} \tag{3.47}$$

$$= \gamma^2 + \gamma\beta E[\exp(\sigma Z'_{pi})] + E[\exp(\sigma Z'_{pj})]$$
$$+ \beta^2 E[\exp(\sigma Z'_{pi})\exp(\sigma Z'_{pj})]$$
$$- \gamma^2 - \gamma\beta\{E[\exp(\sigma Z'_{pi}) + E[\exp(\sigma Z'_{pj})]\}$$
$$- \beta^2 E[\exp(\sigma Z'_{pi})]E[\exp(\sigma Z'_{pj})]$$
$$= \beta^2\{E[\exp(\sigma Z'_{pi} + \sigma Z'_{pj})] - E[\exp(\sigma Z'_{pi})]E[\exp(\sigma Z'_{pj})].$$

Substituting Eq. (3.29) into this last expression and simplifying gives

$$\mathrm{Cov}(X_{pi},X_{pj}) = \beta^2\{E[\exp(2\sigma\rho^{\frac{1}{2}}Z_o)]E[\exp(\sigma(1-\rho)^{\frac{1}{2}}(Z_{pi} + Z_{pj}))] \quad (3.48)$$
$$- E[\exp(\sigma\rho^{\frac{1}{2}}Z_o + \sigma(1-\rho)^{\frac{1}{2}}Z_{pi})]E[\exp(\sigma\rho^{\frac{1}{2}}Z_o + \sigma(1-\rho)^{\frac{1}{2}}Z_{pj})]\}.$$

The four expectations in Eq. (3.48) are computed as follows. When necessary, the expectations are approximated as moment generating functions of normal random variables.

$$E[\exp(2\sigma\rho^{\frac{1}{2}}Z_o)] = \exp(2\sigma^2\rho). \quad (3.49a)$$
$$E\{\exp[\sigma(1-\rho)^{\frac{1}{2}}(Z_{pi}+ Z_{pj})]\} \quad (3.49b)$$
$$\doteq \exp[\sigma(1-\rho)^{\frac{1}{2}}(c_i + c_j) + \sigma^2(1-\rho)(k_{ii} + k_{jj} + 2k_{ij})/2].$$
$$E\{\exp[\sigma\rho^{\frac{1}{2}}Z_o + \sigma(1-\rho)^{\frac{1}{2}}Z_{pi}]\} \quad (3.49c)$$
$$\doteq \exp(2\sigma^2\rho)\exp[\sigma(1-\rho)^{\frac{1}{2}}c_i + \sigma^2(1-\rho)k_{ii}/2].$$
$$E\{\exp[\sigma\rho^{\frac{1}{2}}Z_o + \sigma(1-\rho)^{\frac{1}{2}}Z_{pj}]\} \quad (3.49d)$$
$$\doteq \exp(2\sigma^2\rho)\exp[\sigma(1-\rho)^{\frac{1}{2}}c_j + \sigma^2(1-\rho)k_{jj}/2].$$

Finally, substituting Eqs. (3.49) into Eq. (3.48) and simplifying gives

$$\mathrm{Cov}(X_{pi},X_{pj}) \doteq \beta^2\exp(\sigma^2\rho)\exp[\sigma(1-\rho)^{\frac{1}{2}}(c_i + c_j)] \quad (3.50)$$
$$\times \exp[\sigma^2(1-\rho)(k_{ii} + k_{jj})/2]\{\exp(\sigma^2\rho)\exp[\sigma^2(1-\rho)k_{ij}] -1\}.$$

In theory, the 3PLEN parameters could be estimated in the same fashion as that described for the 3PLN parameters in the previous section; i.e., by employing an iteratively reweighted generalized nonlinear least squares algorithm. One necessary modification of the procedure would be to restrict the estimate of ρ to lie between zero and one. This procedure was tried for 3PLEN samples generated from distributions with ρ=0.5 and (γ,μ,σ) corresponding to each of the four 3PLN distributions of Cohen and Whitten (see previous section). In these analyses, parameter estimates could not be obtained because the matrix of derivatives of the regression equation with respect to the parameters was nearly singular. Thus, the Gauss-Newton algorithm broke down when an attempt was made to invert this nearly singular matrix. However, the estimation worked well when ρ was fixed at its true value. Perhaps a promising alternative estimation procedure would be to introduce a further iteration on fixed values of ρ.

3.3 The S_B Distribution

The QRE procedure may also be readily applied to the S_B distribution (sometimes referred to as the 4-parameter lognormal distribution). The S_B distribution was first described by Johnson (1940) as one of a trio of distributional families that together spanned all possible combinations of skewness and kurtosis. As in the case of the 3-parameter lognormal distribution, an S_B distributed random variable X can be transformed easily to the standard normal random variable Z. In this case, the transformation has the form

$$Z = \frac{\ln[(X-\alpha)/(\beta-X)]-\mu}{\sigma}, \quad \text{for } \sigma > 0,\ \alpha < X < \beta. \qquad (3.51)$$

The possible values of an S_B random variable X are bounded below by α and above by β. The parameter μ determines the skewness. If μ is zero, the density function is symmetric, otherwise the sign of μ determines the sign of the skewness. Finally, the parameter σ determines the kurtosis, which decreases as σ increases. The density function of X is given by

$$\begin{aligned} f_X(x;\alpha,\beta,\mu,\sigma) &= (2\pi\sigma^2)^{-\frac{1}{2}}(\beta - \alpha)[(x - \alpha)(\beta - x)]^{-1} \\ &\quad \times \exp-(1/2\sigma^2)\{\ln[(x - \alpha)/(\beta - x)] - \mu\}^2, \qquad (3.52) \\ &\quad\quad \text{for } \sigma > 0,\ \alpha < x < \beta, \\ &= 0, \text{ otherwise.} \end{aligned}$$

[Further information about the S_B family is given in Johnson (1949), and Johnson and Kotz (1970).]

The QRE procedure cannot be applied to the S_B distribution unless the S_B quantiles X_{pi} are related to the standard normal quantiles Z_{pi} by a monotonic transformation. This may be proven as follows.

Defining $Y = \ln[(X-\alpha)/(\beta-X)]$, Eq. (3.51) can be rewritten as

$$Y = \mu + \sigma Z. \tag{3.53}$$

Since Eq. (3.53) is linear in Z, the transformation from Y to Z is monotonic. Thus if Y can be shown to be a monotonic transformation of of X, it will follow that Z is also a monotonic transformation of X. The transformation from X to Y is monotonic since, for $\alpha < x_i < x_j < \beta$,

$$x_i < x_j \rightarrow x_i - \alpha < x_j - \alpha$$

$$\rightarrow \frac{x_i-\alpha}{\beta-x_i} < \frac{x_j-\alpha}{\beta-x_i} < \frac{x_j-\alpha}{\beta-x_j}$$

$$\rightarrow \ln\left(\frac{x_i-\alpha}{\beta-x_i}\right) < \ln\left(\frac{x_i-\alpha}{\beta-x_j}\right)$$

$$\rightarrow y_i < y_j \ .$$

Since X is a monotonic transformation of Z, the quantiles Y_{pi} and Z_{pi} are related by

$$Y_{pi} = \mu + \sigma Z_{pi}. \tag{3.54}$$

The corresponding quantile regression equation is

$$Y_{pi} = E(\mu + \sigma Z_{pi}) + w_i \ . \tag{3.55}$$

This formulation does not allow the range parameters, α and β to be estimated directly. If α and β are known there is no problem. If not, the estimation may proceed iteratively, via a grid-search in the (α, β) plane for that quartet $[\alpha, \beta, \hat{\mu}(\alpha, \beta), \hat{\sigma}(\alpha, \beta)]$ that minimizes the weighted sum of squared deviations of the X_{pi} from their predicted values. An algorithm to perform this type of estimation is described in Section 3 of Appendix B.

When α and β are held constant, the following expressions define the asymptotic expectations necessary to perform the regression of Eq. (3.55):

$$E(\mu + \sigma Z_{pi}) = \mu + \sigma E(Z_{pi}) \doteq \mu + \sigma\Phi^{-1}(p_i), \tag{3.56a}$$

$$\mathrm{Var}(w_i) \doteq \sigma^2 k_{ii}, \text{ and} \tag{3.56b}$$

$$\mathrm{Cov}(w_i, w_j) \doteq \sigma^2 k_{ij}, \tag{3.56c}$$

where Φ^{-1} is the standard normal distribution function, and k_{ii} and k_{ij} denote the asymptotic variances and covariances, respectively, of the standard normal quantiles.

A small Monte Carlo study is given in section 4.3 in which the sampling behavior of quantile regression estimates is compared to that of local maximum likelihood estimates for the six S_B distributions given in Johnson and Kotz. A proof that the solution to the likelihood equations results in local maximum likelihood estimates, rather than global maximum likelihood estimates, is presented in Appendix A.

3.4 The Income Distribution of Singh and Maddala

In the three previous sections of this chapter, the QRE procedure has been used in connection with density functions whose integrals do not exist in closed form. The QRE procedure is much easier to apply when the distribution function is explicitly known and its inverse is expressible in a concise form. The purpose of this section is to present the methodology for applying the QRE procedure in such cases. Then, as an illustration, the QRE equations pertaining to the income distribution proposed by Singh and Maddala (1976) will be derived. Estimation of this distribution using U.S. family income data is presented in Section 4.4 of Chapter 4.

Let $x_1, \ldots, x_n$ be a random sample of size n from a continuous distribution with density f(x) and distribution function F(x). Also, let $x_{p1} < \ldots < x_{pk}$ be k sample quantiles and $c_1 < \ldots < c_k$ be k corresponding population quantiles, where $0 < p_1 < \ldots < p_k < 1$. If F^{-1} is a concise mathematical expression, and if the first two sample moments of the quantiles of X converge to the corresponding moments of their limiting multivariate normal distribution, then the QRE procedure may be applied straightforwardly. In such cases, $E(X_{pi}) \doteq F^{-1}(p_i) = c_i$ giving the quantile regression equation

$$X_{pi} = E(X_{pi}) + w_i \doteq c_i + w_i \tag{3.57a}$$

where

$$E(w_i) = 0, \tag{3.57b}$$

$$Var(w_i) \doteq [p_i(1-p_i)]/[n\ f^2(c_i)], \tag{3.57c}$$

and $$Cov(w_i, w_j) \doteq [p_i(1-p_j)]/[n\ f(c_i)f(c_j)]. \tag{3.57d}$$

Application to Income Distribution of Singh and Maddala

In their 1976 paper, Singh and Maddala proposed the following distribution function for income (X):

$$F_X(x) = 1 - (1 + ax^b)^{-c}, \qquad X > 0,\ a > 0,\ c > 0. \tag{3.58}$$

Upon differentiating Eq. (3.58), the density of X is found to be

$$f_X(x) = [abcx^{(b-1)}]/[(1 + ax^b)^{(c-1)}]. \tag{3.59}$$

The inverse of the distribution function is derived by solving the distribution function for X as follows:

$$\begin{aligned}
F_X(x) &= p \qquad (3.60)\\
\rightarrow 1 - (1 + ax^b)^{-c} &= p\\
\rightarrow (1 + ax^b)^{-c} &= 1 - p\\
\rightarrow 1 + ax^b &= (1 - p)^{-(1/c)}\\
\rightarrow ax^b &= [(1 - p)^{-(1/c)} - 1]\\
\rightarrow x^b &= (1/a)[(1 - p)^{-(1/c)} - 1]\\
\rightarrow x &= \{(1/a)[(1 - p)^{-(1/c)} - 1]\}^{(1/b)}\\
&= F^{-1}(p).
\end{aligned}$$

Defining $a^* = 1/a$, $b^* = 1/b$ and $c^* = 1/c$, gives

$$F^{-1}(p) = \{a^*[(1 - p)^{-c^*} - 1]\}^{b^*} \tag{3.61}$$

Thus, from Eq. (3.57a) the quantile regression equation for the "Singh-Maddala" distribution is

$$X_{pi} = \{a^*[(1-p_i)^{-c^*} - 1]\}^{b^*} + w_i. \tag{3.62}$$

The variances and covariances of the errors are computed according to Eqs. (3.57c) and (3.57d) where f(x) is given by Eq. (3.59).

In summary, this straightforward applicaton of the QRE procedure involves a regression of sample quantiles on their expected values, where the functional form for the expectations is the inverse of the distribution function. Thus, the only independent variable for this type of regression is the vector of quantile points, p_i. For many distributions, the quantile regression is nonlinear. Examples of distributions for which the regression is linear are the uniform, logistic, Gumbel and exponential. The quantile regression equations corresponding to these and several other well-known density functions are listed in Table 3.1.

Since the covariance structure always depends on the estimated parameters of the density function [see Eqs. (3.57c) and (3.57d)] the estimation should proceed recursively, i.e., the parameters and the covariance matrix should be estimated in an alternating fashion until the parameter estimates from successive iterations differ by less than a desired convergence criterion. In general, there is no guarantee that such an estimation algorithm will produce the global least squares solution. For most distributions there will be a neighborhood around the solution vector within which parameter start-values will lead to the desired least squares solution.

Table 3.1. Quantile Regression Equations for Selected Probability Distributions

Distribution	f(x)	F(x)	$E(X_p) \doteq F^{-1}(p)$
Uniform	$(b-a)^{-1} I_{[a,b]}(x)$	$[(x-a)/(b-a)] I_{[a,b]}(x)$	$a + (b-a)p$
Logistic	$\frac{e^{-(x-a)/b}}{b[1+e^{-(x-a)/b}]^2} I_{(-\infty,\infty)}(x)$	$[1 + e^{-(x-a)/b}]^{-1} I_{(-\infty,\infty)}(x)$	$a - b \ln[(1-p)/p]$
Pareto (2-parameter)	$\frac{a\, x_o^a}{x^{a+1}} I_{(x_o,\infty)}(x)$ for $a > 1$	$[1 - (x_o/x)^a] I_{(x_o,\infty)}(x)$	$x_o/(1-p)^{1/a}$
Pareto (3-parameter)	$\frac{a x_o}{(x+b)^{a+1}} I_{(x_o,\infty)}(x)$	$[1 - x_o(x+b)^{-a}] I_{(x_o,\infty)}(x)$	$(a x_o/p)^{1/(a+1)} - b$
Gumbel	$b^{-1}\, y e^{-y}$ for $b > 0$, where $y = e^{(x-a)/b} I_{(-\infty,\infty)}(x)$	e^{-y} where $y = e^{(x-a)/b} I_{(-\infty,\infty)}(x)$	$a - b \ln(-\ln p)$
Exponential	$a\, e^{-ax} I_{(0,\infty)}(x)$	$(1 - e^{-ax}) I_{(0,\infty)}(x)$	$a^{-1} \ln (1-p)^{-1}$

Table 3.1. (Continued)

Distribution	f(x)	F(x)	$E(X_p) \doteq F^{-1}(p)$
Truncated Exponential	$[1-e^{-aT}]^{-1} a e^{-ax} I_{(0,T)}(x)$	$a[1-e^{-aT}]^{-1} [a(1-e^{-ax})] \times I_{(0,T)}(x)$	$a^{-1} \ln[1-pa^{-2}(1-e^{-aT})]^{-1}$
Weibull	$abx^{b-1} \exp(-ax^b) I_{(0,\infty)}(x)$	$[1-\exp(-ax^b)] I_{(0,\infty)}(x)$	$[-(1/a)\ln(1-p)]^{1/b}$
von Bertalanffy	$\frac{3bce^{-cx}(1-be^{-cx})-(1-b)^3}{1-(1-b)^3} \times I_{(0,\infty)}(x)$ for b, c > 0.	$\frac{(1-be^{-cx})^3-(1-b)^3}{1-(1-b)^3} I_{(0,\infty)}(x)$	$-\frac{1}{c}\ln \frac{1}{b}\{1-[p-(p+1)(1-b)^3]^{1/3}\}$

3.5 The Conditional QRE Procedure Applied to the 3-Parameter Lognormal Distribution

Let $X_{p1}, X_{p2}, \ldots, X_{pk}$ be k quantiles from a sample of size n from a 3-parameter lognormal distribution. That is, $\ln(X - \gamma) \sim N(\mu,\sigma^2)$. In Section 3.1 it was shown that

$$E(X_{pi}) \doteq \gamma + \exp(\mu + \sigma c_i + \sigma^2 k_{ii}/2) \tag{3.63}$$

$$\text{where } c_i = \Phi^{-1}(p_i) \text{ and } k_{ii} = \frac{p_i(1-p_i)}{n[\phi(c_i)]^2} .$$

Expressions for the asymptotic values of $\mathrm{Var}(X_{pi})$ and $\mathrm{Cov}(X_{pi}, X_{pj})$ were also given, thus permitting the generalized nonlinear least squares estimation of the regression equation

$$X_{pi} = E(X_{pi}) + w_i \qquad \text{where } w_i \text{ is an error term.} \tag{3.64}$$

For some applications the conditioning of the covariance matrix of the quantiles may be too poor to permit meaningful regression results to be obtained. In these instances, one may wish to obtain unweighted, or ordinary QRE's. A second alternative, for which the covariance matrix is diagonal is proposed here. The approach will be called "conditional quantile regression estimation" (CQRE).

The CQRE approach regresses the sample quantiles on their asymptotic expectations conditional on the values of the preceding quantiles. The reason for adopting a conditional regression is to reduce the need for generalized nonlinear least squares by incorporating the correlation among the quantiles into the regression model itself. It is hoped that by absorbing the correlations among the quantiles into the regression equation, ordinary nonlinear least squares techniques will provide estimates nearly as precise as those given by the generalized nonlinear least squares regression of Eq. (3.64).

The conditional QRE application to the 3PLN distribution is derived as follows. The regression model is

$$\begin{aligned} X_{pi} &= E(X_{pi}) + u_i, \qquad i = 1 \\ &= E(X_{pi} \mid X_{p,i-1}, X_{p,i-2}, \ldots, X_{p1}) + u_i, \quad i = 2, 3, \ldots, k. \end{aligned} \tag{3.65}$$

Since the quantiles of a sample form a Markov chain, the regression Eq. (3.65) can be written in an equivalent, simpler form as

$$X_{pi} = E(X_{pi} \mid X_{p,i-1}) + u_i.$$

The u_i are uncorrelated residuals since, for any sequence of random variables $\{X_i\}$, the transformed random variables $Y_i = X_i - E(X_i \mid X_{i-1}, \ldots, X_1)$ are independent.

The conditional expectation in this regression may be approximated by treating the associated pairs of standard normal quantiles as having a bivariate normal (BVN) distribution. Defining $Z_{pi} = [\ln(X_{pi}-\gamma)-\mu]/\sigma$, Theorem 3.2 says that

$$(Z_{pi}, Z_{pj}) \mathrel{\dot{\sim}} BVN(c_i, k_{ii}, c_j, k_{jj}, \rho_{ij}) \tag{3.66}$$
$$\text{where } \rho_{ij} = k_{ij}/(k_{ii}\, k_{jj})^{1/2}$$

With this result, the asymptotic expression for $E(X_{pi} \mid X_{pj})$ is computed as follows:

$$E(X_{pi} \quad X_{pj}) = E[\gamma + \exp(\mu + \sigma Z_{pi}) \mid X_{pj}] \tag{3.67}$$
$$= \gamma + \exp(\mu)E[\exp(\sigma Z_{pi}) \mid X_{pj}]$$
$$= \gamma + \exp(\mu)E[\exp(\sigma Z_{pi}) \mid Z_{pj}],$$

the last step following because Z_{pj} is a monotonic transformation of X_{pj} and therefore knowledge of X_{pj} is equivalent to knowledge of Z_{pj}. Now,

$$E[\exp(\sigma Z_{pi}) \mid Z_{pj}] \doteq \exp[\sigma E(Z_{pi} \mid Z_{pj}) + .5\sigma^2 \mathrm{Var}(Z_{pi} \mid Z_{pj})] \tag{3.68}$$

since the expectation is the moment generating function of the asymptotically normal random variable Z_{pi} conditional on Z_{pj}. The conditional expectations on the right-hand side of Eq. (3.68), for large samples, are given by

$$E(Z_{pi} \mid Z_{pj}) \doteq c_i + \rho_{ij}(k_{ii}/k_{jj})^{\frac{1}{2}}(Z_{pj} - c_j) \tag{3.69}$$
$$= c_i + \alpha_{ij}(Z_{pj} - c_j) \text{ where } \alpha_{ij} = \rho_{ij}(k_{ii}/k_{jj})^{\frac{1}{2}}$$

and

$$\mathrm{Var}(Z_{pi} \mid Z_{pj}) \doteq k_{ii}(1 - \rho_{ij}^{\;2}) = \beta_{ij} \text{ (say)}. \tag{3.70}$$

Substituting Eq. (3.69) and (3.70) into Eq. (3.68) and then substituting Eq. (3.68) into Eq. (3.67) gives

$$E(X_{pi} \mid X_{pj}) = \gamma + \exp[\mu + \sigma c_i + \sigma\alpha_{ij}(Z_{pj} - c_j) + .5\sigma^2\beta_{ij}]. \tag{3.71}$$

Eq. (3.71) still cannot be used in the estimation of the CQRE equation because the Z_{pj} are unknown. Substituting $\ln(X_{pj}-\gamma) - \mu$ for σZ_{pj} takes care of this problem and gives, after rearranging,

$$E(X_{pi} \mid X_{pj}) = \gamma + \exp[(1 - \alpha_{ij})\mu + (c_i - \alpha_{ij}c_j)\sigma + .5\beta_{ij}\sigma^2] \tag{3.72}$$
$$\times (X_{pj} - \gamma)^{\alpha_{ij}}.$$

Eq. (3.66) may now be estimated by ordinary or weighted least squares by replacing $E(X_{pi} | X_{pi-1})$ with the right hand side of Eq. (3.72) (with j=i-1). For weighted least squares estimation, the weight matrix would be diagonal since the u_i are uncorrelated.

A Monte Carlo study of the sampling behavior of 3PLN parameter estimates given by the CQRE procedure using ordinary least squares is presented in Section 4.5. The study is formatted identically to that used to investigate the properties of the unconditional quantile regression estimation of the 3PLN distribution in Section 4.1. To provide an additional means of comparison, Section 4.5 also includes estimation results of the unweighted (or ordinary) QRE procedure applied to the same generated data. The latter procedures is identical to the QRE procedure except that the known asymptotic form of the residual covariance matrix $V(\theta)$ is replaced by the appropriately dimensioned identity matrix. These estimation results were obtained because the ordinary QRE procedure would be the logical competitor of the CQRE procedure for investigators preferring not to use the weighted least squares techniques.

CHAPTER 4

APPLICATIONS OF THE QRE PROCEDURE

In this chapter, five empirical analyses are performed which should permit assessments of the performance of the QRE procedure in many of the applications discussed in Chapter 3. In each case at least one other procedure is applied to the data so that the QRE results may be evaluated on a relative basis.

The most thorough of the following analyses is that pertaining to the 3-parameter lognormal (3PLN) distribution. For this distribution a large-scale Monte Carlo study is presented in Section 4.1 in which quantile regression estimates are compared with local maximum likelihood estimates for four 3PLN distributions and over a wide range of sample sizes. Then, in section 4.2, the QRE procedure is applied to Hill's (1963) epidemic data and the results are compared to those obtained by Hill and others who used different estimation methods.

A smaller Monte Carlo study, presented in Section 4.3, compares quantile regression estimates and local maximum likelihood estimates from each of six S_B distributions depicted in Johnson and Kotz (1970). Then, illustrating a more straightforward application of the QRE procedure, the income distribution proposed by Singh and Maddala (1976) is estimated in Section 4.4 using United States family income data. These results are compared with those of a simpler estimation procedure used by Singh and Maddala.

Finally, the conditional QRE procedure is investigated in Section 4.5 using the same generated 3PLN data used in Section 4.1. Unweighted QRE results are also computed for these data and each set of results is compared with the weighted QRE results of Section 4.1.

4.1 Monte Carlo Study I--Four 3-Parameter Lognormal Distributions of Cohen and Whitten (1980)

In chapter 3 the theory was developed to allow the estimation of the parameters of the 3-parameter lognormal (3PLN) distribution via a regression of selected sample quantiles on their asymptotic expected values. Now the sampling behavior of the estimates given by this quantile regression estimation (QRE) procedure will be examined in a Monte Carlo experiment.

Cohen and Whitten (1980) have already conducted a Monte Carlo study to compare the sampling behavior of 3PLN estimates from 14 estimation procedures. They chose four 3PLN distributions from which to generate random samples. Each of the four 3PLN random variables studied had a mean of zero and a variance of one, but the skewness, α_3, was varied from low ($\alpha_3 = 0.301$) to high ($\alpha_3 = 2.475$)[3]. The probability density functions of these four random variables are plotted in Figure 4.1. These same four distributions will also be employed in this Monte Carlo study to facilitate comparisons of the performance of the QRE procedure with those of the 14 estimation procedures studied by Cohen and Whitten.

Before discussing the details of the present Monte Carlo study it will be useful to review the design of Cohen and Whitten's study, describe the 14 methods they investigated and assess the results they obtained. Once this has been done, the design of the present study will be outlined, and its objectives stated. Then the results will be pre-

---

[3]The four values for skewness (α_3) corresponded to setting $\omega = 1.01$, 1.10, 1.25, and 1.50 where $\omega = \exp(\sigma^2)$.

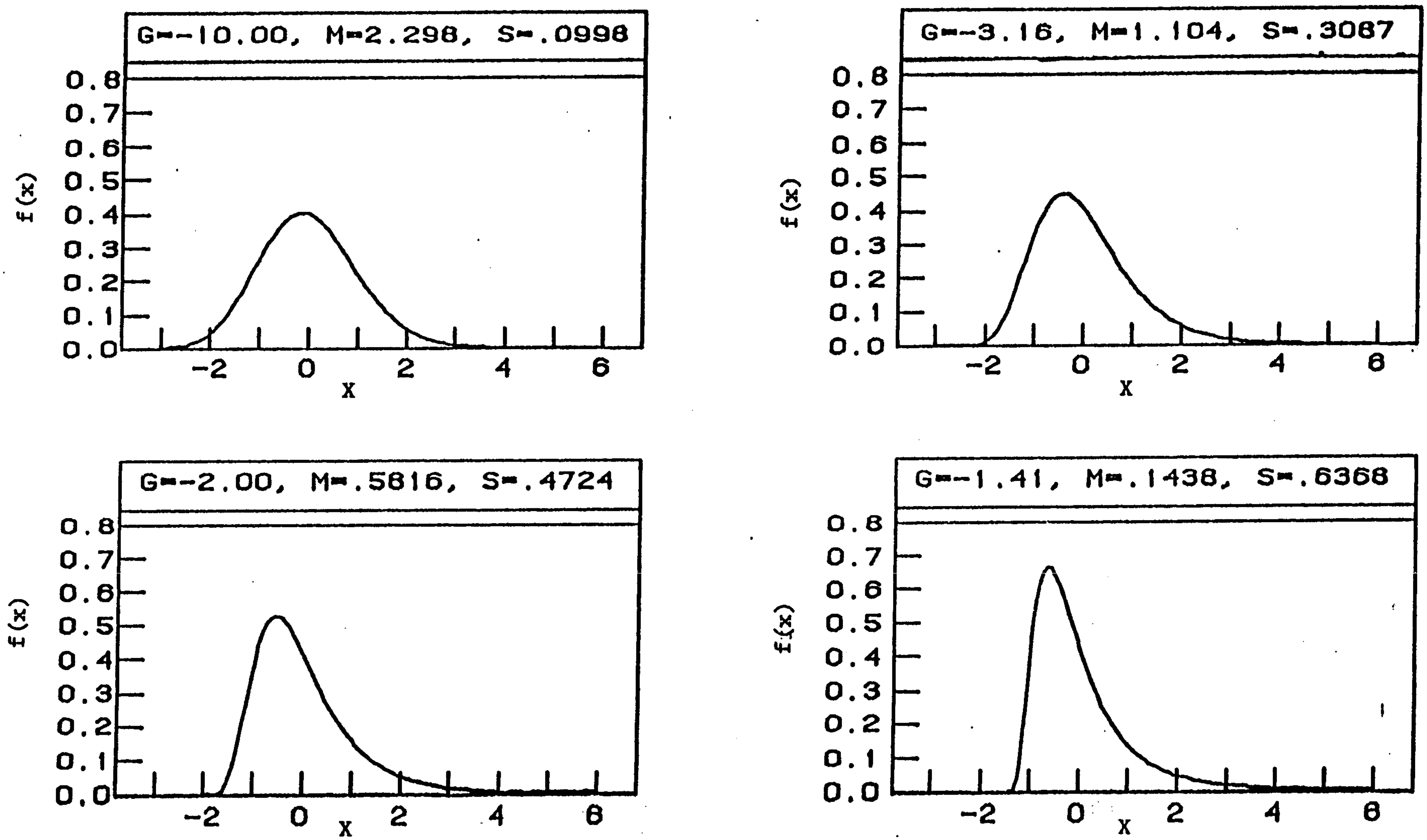

Figure 4.1. Probability density functions of Cohen and Whitten's four 3-parameter lognormal distributions.

sented and discussed in light of these objectives. Finally, general conclusions will be drawn concerning the usefulness of the QRE procedure as a means of estimating the parameters of the 3-parameter lognormal distribution.

A Review of Cohen and Whitten's Monte Carlo Study

Cohen and Whitten conducted their Monte Carlo study in an effort to assess the properties of twelve different estimation methods proposed in their paper as alternatives to the moment estimation (ME) and local maximum likelihood estimation (LMLE) procedures. Improvements were sought for the former because of the well-documented evidence of the inefficiency of its estimates [see Aitchison and Brown (1957) or Calitz (1973)] and for the latter because of uncertainty about the behavior of LMLE's compared to that of conventional maximum likelihood estimates (MLE's).

The twelve modified estimation methods of Cohen and Whitten consist of three variations each of two basic modifications of the ME and LMLE procedures. Recall that the traditional ME's and MLE's are obtained via the simultaneous solution of m equations where m is the number of unknown parameters (m=3 in the case of the 3PLN distribution). Both of the modifications proposed by Cohen and Whitten called for the replacement of one of the three equations with an auxiliary equation involving a small order statistic.[4] The same two types of auxiliary equations were used to modify both the ME and MLE procedures.

---

[4]In the case of LMLE, the equation $\partial \ln L/\partial\gamma=0$ was replaced. In the case of ME, the equation of the third sample moment to its expectation was replaced.

The first modification corresponded to the auxiliary equation that equated the expectation of the cumulative distribution function of the rth sample order statistic to its observed value, i.e.,

$$E(F(X_r)) = F(X_r). \tag{4.1}$$

Cohen and Whitten point out that $F(X_r)=\Phi(Z_r)$ where $Z_r=[\ln(X_r-\gamma)-\mu]/\sigma$, and Φ is the standard normal distribution function. Defining $k_r = \Phi^{-1}[r/(n+1)]$ and using $E[F(X_r)] = r/(n+1)$, Eq. (4.1) may be reexpressed as

$$\ln(X_r-\gamma) = \mu + \sigma k_r. \tag{4.2}$$

The second modification corresponded to the auxiliary equation in which k_r of Eq. (4.2) is replaced by $E(Z_r)$, the expectation of the rth standard normal order statistic. Therefore, the auxiliary equation corresponding to Cohen and Whitten's second modification is

$$\ln(X_r-\gamma) = \mu + \sigma E(Z_r). \tag{4.3}$$

Values of $E(Z_r)$ were obtained in tables provided by Harter (1961).

Cohen and Whitten investigated the performance of these two modifications using each of the first three sample order statistics (i.e., r = 1, 2 and 3). Thus, their twelve modified methods consisted of two modifications [Eq. (4.2) and Eq. (4.3)] using three order statistics in turn and applied to two estimation methods (ME and MLE).

As mentioned earlier, the twelve modifications along with the conventional moment and local maximum likelihood estimation procedures were applied to samples generated from four different 3PLN distributions. Each of the estimation procedures was applied to 100 independently generated samples of size 100 from each of the four distributions. In generating the samples, the sample skewness was checked and if it fell outside the interval 0.13 to 14.0 the sample was discarded. This was done to avoid known difficulties in applying the method of moments procedure to samples of very high or low skewness. Indeed, when the sample skewness is negative, moment estimators do not even exist.[5]

Cohen and Whitten's algorithms for their modified estimation methods involved iterative searching routines that did not always converge. For some samples, the search for the estimate of the location parameter γ diverged toward $-\infty$ or converged to the first sample order statistic, X_1. When this occurred the estimation procedure was terminated and was said to have failed for that sample.

The results of Cohen and Whitten's study were presented in two tables. The first contained the means of the parameter estimates over the successful applications of each method. The second contained the corresponding standard deviations of these estimates. The information in these tables permits the 14 estimation methods to be compared according to three important criteria: accuracy, precision and applicability.

The accuracy of a procedure may be investigated by comparing the means of its parameter estimates to their true values. The precision is estimated by the standard deviations of the parameter estimates, while

---

[5]A negative sample skewness will yield moment estimates identical to those given by a sample with a positive skewness of the same magnitude. However, as Cohen and Whitten point out, such estimates are inadmissible because there are no negatively skewed 3PLN distributions.

the applicability of a procedure will be defined here as the proportion of the time that the estimate of γ does not diverge toward $-\infty$ or X_1.

When all three criteria are considered jointly it will be seen that the unmodified LMLE procedure outperformed the ME and twelve modified procedures. This is not the conclusion drawn by Cohen and Whitten, but their evaluation of the results did not appear to be carefully or systematically conducted. The simple, but informative analysis of their results given below will illustrate the superiority of the performance of the LMLE method.

Perhaps the most popular empirical means of comparing the relative goodness of different estimators is to use the mean-squared error criterion. This criterion permits the accuracy and precision of the estimators to be jointly evaluated. Formally, if $T_\theta = t(X_1, \ldots, X_n)$ is an estimator of θ, then the mean-squared error (MSE) of T_θ is defined as

$$MSE[T_\theta] = E[T_\theta - \theta]^2. \tag{4.4}$$

Here, the mean-squared error criterion will be used to compare the goodness of the fourteen procedures' estimates of γ, since this is the parameter for which it is most critical to obtain good estimates. To use this criterion empirically, the expectation on the right-hand side of Eq. (4.4) must by approximated. An obvious choice is given by

$$E[T_\gamma - \gamma]^2 \doteq (1/n) \sum_i (\hat{\gamma}_i - \gamma)^2, \tag{4.5}$$

where $\hat{\gamma}_i$ is the estimate given by T_γ for the ith sample. The right hand side of Eq. (4.5) may be reexpressed as

$$(1/n) \sum_i (\hat{\gamma}_i - \gamma)^2 = (1/n) \sum_i (\hat{\gamma}_i - \bar{\gamma})^2 + (\gamma - \bar{\gamma})^2 \tag{4.6}$$

Table 4.1 Estimated Mean-Squared Errors (EMSE) of 14 Estimation Procedures Compared in Cohen and Whitten (1980)

Estimation Procedure (T_γ)*	Distribution 1 α_3= .301		Distribution 2 α_3= .980		Distribution 3 α_3= 1.625		Distribution 4 α_3=2.475	
	EMSE	Rank	EMSE	Rank	EMSE	Rank	EMSE	Rank
ME	32.92	3	5.236	9	1.1218	10	.4419	10
MME-I_1	531.41	14	11.524	11	.4194	8	.0399	7
MME-I_2	160.79	10	15.184	12	.2462	6	.0386	6
MME-I_3	206.98	11	6.312	10	.3057	7	.0467	8
MME-II_1	271.03	12	1.727	1	.1623	3	.0214	2
MME-II_2	71.66	7	2.236	7	.1338	2	.0244	3
MME-II_3	131.41	9	61.195	14	.1859	4	.0328	5
LMLE	21.35	1	1.931	4	.1165	1	.0168	1
MMLE-I_1	274.18	13	26.597	13	.7054	9	.0597	9
MMLE-I_2	118.58	8	1.913	3	-†	-	-	-
MMLE-I_3	61.73	5	2.012	5	-	-	-	-
MMLE-II_1	28.75	2	2.955	8	.2384	5	.0289	4
MMLE-II_2	61.79	6	1.884	2	-	-	-	-
MMLE-II_3	61.73	4	2.012	6	-	-	-	-

* ME: method of moments
MME: modified method of moments
LMLE: local maximum likelihood estimation
MMLE: modified maximum likelihood estimation

$-I_r$ or $-II_r$: first or second modification type using the rth order statistics where r = 1, 2, or 3

† a dash (-) signifies that the estimation algorithm failed to converge for at least 50 percent of the samples.

the modification. For the modified maximum likelihood estimation algorithms, this was true only if the first order statistic was used. When the second or third was used, convergence failed often, the failure rate increasing with the skewness of the underlying distribution from around 25 percent for the third most skewed distribution to well over 90 percent for the most skewed distribution. For the least skewed distribution, the failure rate varied between 8 and 27 percent across the 12 modified estimation methods.

Cohen and Whitten recommended $MME\text{-}II_1$ (with $MME\text{-}I_1$ and ME as alternatives) when skewness is less than one. For skewness greater than one, they recommended $MMLE\text{-}II_1$ (with $MMLE\text{-}I_1$ and LMLE as alternatives). It is difficult to agree with their recommendations based on the results in Table 4.1. Perhaps they felt that their modifications had other advantages over the LMLE procedure, such as ease of implementation.

The remainder of this section will be devoted to assessing the performance of the quantile regression estimation (QRE) procedure in estimating the parameters of the same four 3PLN distributions used by Cohen and Whitten. LMLE estimates will also be computed to provide a convenient reference for use in evaluating the goodness of the quantile regression estimates.

Monte Carlo Study Comparing QRE and LMLE Procedures

In the Monte Carlo experiment conducted for the present study, the sampling behavior of estimates given by the QRE and LMLE procedures were compared. The format of Cohen and Whitten's Monte Carlo study was duplicated in that 100 samples were generated from each of the four 3PLN distributions shown in Figure (4.1). However in this experiment, samples were not discarded if the skewness lay outside the interval 0.13

to 14.0. Rather, all samples were retained regardless of skewness so that unbiased estimates of the means and variances of the parameter estimates could be obtained. Also, the scope of this study was increased beyond Cohen and Whitten's in that the properties of the estimators were investigated over a wide range of sample sizes.

An important consideration in investigating the behavior of quantile regression estimates for varying sample sizes is the coordination of the choices of the quantile points and the sample sizes. Since the vector of chosen quantile points $p = (p_1, p_2, \ldots, p_k)$ has an influence on the properties of the QRE's, the same choice of p should be used for all sample sizes. This avoids confounding the effects on the behavior of the QRE's of sample size and quantile selections. But once p has been chosen, only a limited set of sample sizes will contain quantiles that correspond to p. Therefore, in this study the following criterion for sample size selection was used:

<u>Necessary Condition for Sample Size Selection n:</u>

Given a vector of quantile points $p = (p_1, p_2, \ldots p_k)$, n is an acceptable sample size for the Monte Carlo study if for each i = 1, 2, ..., k there exists an integer r_i such that $r_i/(n+1) = p_i$.

Satisfying this criterion ensures that for each quantile point p_i the sample will contain an order statistic X_{ri} for which

$$E[F(X_{ri})] = r_i/(n+1) = p_i, \tag{4.7}$$

and therefore the p_ith quantile X_{pi} is defined by $X_{pi} = X_{ri}$. In the present study, all p_i were chosen to be multiples of 0.01. In this case, all sample sizes one less than multiples of 100 would satisfy the necessary condition mentioned above. Specifically, the following ten quantile points were chosen:

$$p = (.03, .07, .15, .25, .35, .50, .65, .80, .90, .98).$$

These were chosen somewhat arbitrarily although an effort was made to space them fairly evenly and to capture information near the extremes of the data, especially at the low end. Further consideration of how the quantiles might be spaced is given in Appendix C.

When comparing the performance of the QRE and LMLE procedures it must be kept in mind that only 10 quantiles are used to generate a quantile regression estimate while the entire sample is used to generate a local maximum likelihood estimate. Under the assumption that the properties of the LMLE's are nearly the same as those normally possessed by MLE's[6] (consistent and asymptotically efficient), the QRE's cannot be expected to be more accurate or to have smaller variances than the LMLE's. Rather, the objectives of this Monte Carlo study are to assess how quickly the bias of QRE estimates disappears as sample size increases, and to estimate their precision relative to the LMLE's. Additionally, any failures of the algorithms to converge will be noted.

The lognormal samples of size n were created by generating n independent standard normal values $z_1, z_2, \ldots, z_n$ and transforming them to

---

[6]This assumption about the behavior of the LMLE's is supported by the results of Cohen and Whitten's Monte Carlo study, where the sampling variances of the LMLE's were shown to compare similarly to their asymptotic expectations given by the inverse of the Fisher information matrix. A very small bias was indicated in the estimates of γ when the skewness was less than one, but this could have been partially due to the elimination of samples having skewness of less than 0.13.

lognormal values $x_1, x_2, \ldots x_n$ by the 3-parameter lognormal transformation

$$x_i = \gamma + \exp(\mu + \sigma z_i). \tag{4.8}$$

For a given sample size the same 100 samples of standard normal variates were used in the formation of each of the four sets of lognormal samples. This was done to improve the efficiency of comparisons of parameter estimates across the four distributions.

The estimation algorithms used to obtain the QRE's and LMLE's are documented in Section 1 of Appendix B. To perform the nonlinear quantile regressions, start values for the parameter estimates were required. Two separate computer programs were written which differed only in the method of obtaining start values. The first program computes pseudo-moment estimators (described in Appendix B) for use as start values, and thus, requires positively skewed samples. The second program, intended only for use in the Monte Carlo study, uses the true parameter values as start values and thus does not require that samples be positively skewed. Convergence was slightly faster when the true values were used as start values, but the parameter estimates from positively skewed samples were always the same regardless of which method of start value selection was used.

In searching for the local maximum likelihood estimates, it is important to have knowledge of the shape of the likelihood function. The shape of the likelihood function of a typical 3-parameter lognormal sample has been discussed in Hill (1963) and Griffiths (1980). Unless skewness and sample size are both small, the likelihood function as a function of $\hat{\gamma}$ should have a local maximum followed by a minimum and then approach $+\infty$ at $\hat{\gamma} = x_1$. Much of the motivation for searching for methods of finding the local maximum has come out of a concern that merely solving the likelihood equations would yield the undesirable MLE solution $(\hat{\gamma},\hat{\mu},\hat{\sigma}) = (x_1,-\infty,\infty)$. Fortunately, using the true values of γ as start values in the searching algorithm used here always led to the local maximum rather than to the inadmissible solution. However, for samples from the least skewed 3PLN distribution, both the QRE's and LMLE's of γ occasionally diverged toward $-\infty$, especially for the smallest sample size (n=99). In actual applications of these procedures, large negative estimates of γ should suggest to the investigator that the sample may not have come from a lognormal population. In these cases, it may be better to assume a symmetric distribution, such as the normal. In this study, the estimation algorithms of the QRE and LMLE procedures were terminated if the estimate of γ fell below -50.

For the least skewed distribution, when the sample size was 99, the estimation of γ was terminated for 15 of 100 samples using the LMLE procedure and for 17 of 100 samples using the QRE procedure. For larger sample sizes, the estimation algorithms had to be terminated for fewer than 10 percent of the samples. Among the total of 1500 samples from the three more skewed distributions, the QRE estimation algorithm had to be terminated only once, for a sample of size 99, while the LMLE algorithm converged for all samples.

The mean parameter estimates, e.g., $\bar{\hat{\gamma}} = n^{-1}\Sigma\hat{\gamma}_i$, given by each procedure are shown in Table 4.2. The standard deviations of the estimates, e.g., $s.d.(\hat{\gamma}) = (n-1)^{-1}\Sigma(\hat{\gamma}_i-\bar{\hat{\gamma}})^2$, are given in Table 4.3. Each table also indicates, under the columns labeled "Runs", the number of samples (out of 100) for which the estimation algorithm converged to

Table 4.2. Means of 3PLN Parameter Estimates from LMLE and QRE Procedures

Sample Size	Method = LMLE				Method = QRE			
n	γ	μ	σ	Runs	γ	μ	σ	Runs
Distribution 1 (γ = -10, μ = 2.2976, σ = .0998, skewness = 0.301)								
99	-10.943	2.190	.1297	85	-11.431	2.172	.1414	83
299	-13.273	2.414	.1006	98	-13.181	2.390	.1073	92
499	-10.698	2.283	.1076	98	-11.092	2.317	.1065	96
699	-10.926	2.318	.1029	99	-11.398	2.327	.1046	97
899	-10.533	2.300	.1035	100	-11.504	2.367	.0987	100
Distribution 2 (γ = -3.1623, μ = 1.1036, σ = .3087, skewness = 0.980)								
99	-3.249	1.093	.3236	100	-3.423	1.127	.3283	99
299	-3.216	1.108	.3089	100	-3.334	1.135	.3077	100
499	-3.106	1.077	.3191	100	-3.192	1.103	.3147	100
699	-3.144	1.091	.3129	100	-3.209	1.108	.3110	100
899	-3.118	1.083	.3147	100	-3.194	1.107	.3092	100
Distribution 3 (γ = -2, μ = .5816, σ = .4724, skewness = 1.625)								
99	-1.981	.5568	.4924	100	-2.079	.5988	.4929	100
299	-1.998	.5759	.4743	100	-2.041	.5966	.4725	100
499	-1.963	.5557	.4860	100	-2.000	.5768	.4809	100
699	-1.982	.5671	.4779	100	-2.008	.5812	.4754	100
899	-1.973	.5615	.4801	100	-2.004	.5797	.4742	100
Distribution 4 (γ = -1.4142, μ = .1438, σ = .6368, skewness = 2.475)								
99	-1.393	.1174	.6616	100	-1.440	.1547	.6614	100
299	-1.409	.1376	.6383	100	-1.428	.1532	.6380	100
499	-1.395	.1223	.6513	100	-1.412	.1375	.6478	100
699	-1.406	.1337	.6407	100	-1.416	.1419	.6406	100
899	-1.403	.1295	.6426	100	-1.414	.1394	.6397	100

Table 4.3. Standard Deviations of 3PLN Parameter Estimates from LMLE and QRE Procedures*

Sample Size n	Method = LMLE γ	μ	σ	Runs	Method = QRE γ	μ	σ	Runs
Distribution 1	(γ = -10, μ = 2.2976, σ = .0998, skewness = 0.301)							
99	7.758	.6057	.0665	85	9.662	.6828	.0779	83
299	9.103	.5520	.0489	98	8.777	.5961	.0591	92
499	5.480	.3718	.0338	98	5.103	.3968	.0394	96
699	4.536	.3533	.0324	99	6.303	.4174	.0362	97
899	3.583	.3073	.0290	100	5.033	.3601	.0323	100
Distribution 2	(γ = -3.1623, μ = 1.1036, σ = .3087, skewness = 0.980)							
99	.9558	.2768	.0815	100	1.2684	.3361	.0978	99
299	.4988	.1681	.0524	100	.7481	.2236	.0678	100
499	.3334	.1131	.0375	100	.4147	.1367	.0452	100
699	.3087	.1058	.0337	100	.4598	.1418	.0421	100
899	.2580	.0927	.0286	100	.3102	.1082	.0328	100
Distribution 3	(γ = -2, μ = .5816, σ = .4724, skewness = 1.625)							
99	.2837	.1752	.0826	100	.4653	.2356	.1049	100
299	.1673	.1104	.0539	100	.2405	.1459	.0698	100
499	.1146	.0740	.0390	100	.1449	.0893	.0472	100
699	.1049	.0673	.0340	100	.1550	.0915	.0435	100
899	.0896	.0613	.0284	100	.1137	.0739	.0335	100
Distribution 4	(γ = -1.4142, μ = .1438, σ = .6368, skewness = 2.475)							
99	.1161	.1354	.0842	100	.1820	.1792	.1084	100
299	.0679	.0846	.0541	100	.1040	.1139	.0724	100
499	.0468	.0567	.0388	100	.0633	.0699	.0496	100
699	.0415	.0485	.0334	100	.0672	.0699	.0452	100
899	.0348	.0451	.0260	100	.0510	.0591	.0345	100

*Table entries are the standard deviations of the associated set of parameter estimates. For example, for Distribution 1, the 85 estimates of γ had a standard deviation of 7.758.

an estimate of γ greater than -50. It is difficult to draw precise inferences from the raw results of these tables without performing additional analyses of these results. Casual inspection of these results does reveal some broad findings, however. First, both methods tended to estimate the parameters fairly accurately, even for sample sizes as small as 99. Second, as expected, the variances of the QRE's were nearly always greater than those of the LMLE's.

Some further manipulations of the results shown in Tables 4.2 and 4.3 reveal more detailed comparisons. The accuracy of the procedures over the various sample sizes and distributions is examined in Table 4.4, with regards to the estimates of the location parameter γ. This table shows the percentage differences between the mean estimates and the true values of γ, the t-statistics corresponding to these differences, and the probability of a greater absolute value of t assuming the estimators are unbiased. An index of the relative precision of the QRE procedure compared to the LMLE procedure in estimating the 3PLN parameters will also be discussed.

Looking first at the tests for bias in Table 4.4, note that in general, the percentage differences between γ and the mean estimates of γ from both procedures were very small. Over the three most skewed distributions, the means of the LMLE's of γ were different from the true value of γ by less than three percent. For small sample sizes, the estimated bias of the QRE's tended to be greater than that of the LMLE's over these same three distributions, but it was still less than 10 percent. For the least skewed distribution, the mean estimates of γ from both procedures ranged from 5 to 33 percent less than the true value of γ. These relatively large absolute differences from the true

Table 4.4. Tests for Bias of LMLE and QRE Procedures in Estimating the 3PLN Location Parameter γ

Sample Size n	Method = LMLE			Method = QRE		
	$\Delta(\bar{\hat{\gamma}},\gamma)$*	$t^{\dagger}$	Prob>|t|	$\Delta(\bar{\hat{\gamma}},\gamma)$*	$t^{\dagger}$	Prob>|t|
<u>Distribution 1</u>	(γ = -10, μ - 2.2976, σ = .0998, skewness = 0.301)					
99	-9.43	-1.12	.2624	-14.31	-1.35	.1772
299	-32.73	-3.56	.0004	-31.81	-3.48	.0005
499	-6.98	-1.26	.2073	-10.92	-2.10	.0361
699	-9.26	-2.03	.0422	-13.98	-2.19	.0289
899	-5.33	-1.49	.1368	-15.04	-2.99	.0028
<u>Distribution 2</u>	(γ = -3.1623, μ = 1.1036, σ = .3087, skewness = 0.980)					
99	-2.76	-0.91	.3620	-8.25	-2.05	.0406
299	-1.70	-1.08	.2799	-5.42	-2.29	.0219
499	1.80	1.70	.0882	-0.93	-0.71	.4805
699	0.59	0.61	.5429	-1.47	-1.01	.3135
899	1.39	1.71	.0877	-0.99	-1.01	.3106
<u>Distribution 3</u>	(γ = -2, μ = .5816, σ = .4724, skewness = 1.625)					
99	0.95	0.67	.5011	-3.97	-1.71	.0881
299	0.09	0.10	.9140	-2.05	-1.71	.0881
499	1.83	3.20	.0014	-0.01	-0.01	.9893
699	0.92	1.76	.0778	-0.41	-0.52	.6005
899	1.35	3.02	.0025	-0.21	-0.38	.7075
<u>Distribution 4</u>	(γ = -1.4142, μ = .1438, σ = .6368, skewness = 2.475)					
99	1.52	1.85	.0640	-1.85	-1.44	.1495
299	0.35	0.73	.4650	-0.98	-1.34	.1806
499	1.34	4.05	<.0001	0.19	0.42	.6714
699	0.56	1.92	.0542	-0.12	-0.25	.8004
899	0.80	3.23	.0012	0.02	0.06	.9521

* $\Delta(\bar{\hat{\gamma}},\gamma) = (\bar{\hat{\gamma}} - \gamma)/\gamma$; i.e., the percentage difference between $\bar{\hat{\gamma}}$ and γ, where $\bar{\hat{\gamma}}$ is the mean of the $\hat{\gamma}$'s.

† t = t-statistic corresponding to H_o: $E(\hat{\gamma}) = \gamma$.

value of γ were significant at the 5 percent level for n = 299 and n = 699 among the LMLE's and for $n \geq 299$ among the QRE's. This apparent negative bias may be explained by the fact that the skewness of the empirical distribution of $\hat{\gamma}$ from each procedure was negative. Indeed, while the true value of γ was -10, the estimates ranged from -49.5 to -2.9. Such a highly skewed distribution of estimates will cause the mean of the distribution to be smaller; hence, the negative bias.

On the whole, however, the potential for bias in either estimation procedure for $n \geq 99$ seemed to pose little cause for concern. For any given sample from a 3PLN distribution having low skewness, the error that occurs in estimating γ is likely to be more due to the high variance of $\hat{\gamma}$ than to any bias in the estimation procedure. The standard deviations of $\hat{\gamma}$ shown in Table 4.3 reveal that when skewness is low, the variance of $\hat{\gamma}$ is disappointingly high for both the LMLE and QRE procedures. For samples from more skewed 3PLN distributions, the variance of $\hat{\gamma}$ becomes much smaller, but so does the estimated bias of $\hat{\gamma}$. While several of the tests for bias of the LMLE's in Table 4.4 reject unbiasedness, the percentage differences between the mean of $\hat{\gamma}$ and γ are small enough to be ignored for most purposes (less than 2 percent).

Since bias for sample sizes greater than 99 is not a problem in either procedure, the precision of the two procedures becomes the important point of comparison. Figure 4.2 plots the standard deviations of the estimates of γ from each procedure as a function of the sample size n. The four panels shown correspond to the four 3PLN distributions. Between each pair of curves is a shaded region indicating the difference in the estimated precision of the two procedures. When the differences are a small percentage of the standard deviations of the LMLE's, the

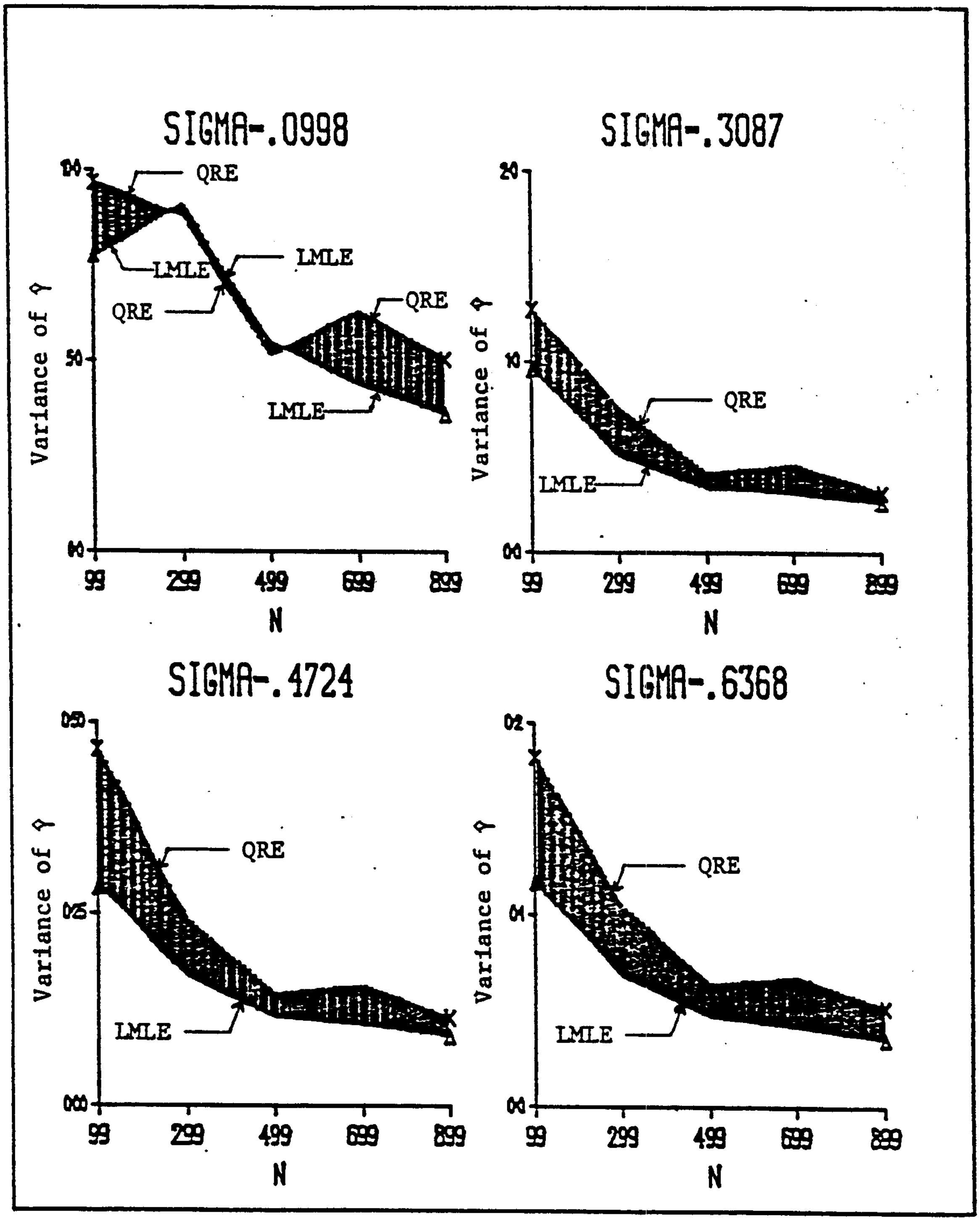

Figure 4.2. Variance of estimates of γ vs. sample size for QRE and LMLE procedures.

relative precision of the QRE procedure to the LMLE procedure is high. Figure 4.2 reveals that the percentage differences in precision are not that large, and remain fairly constant for sample sizes between 99 and 899. In addition to graphically comparing the precision of each method, the plots reveal that the standard errors of $\hat{\gamma}$ rapidly decrease as the sample size decreases or as the skewness increases.

Another useful means of comparing the precisions is to form an index of relative precision (R.P.). This was done by forming the ratio of the standard deviations. That is, for a given parameter θ,

$$\text{R.P.}(\theta) = \frac{\text{s.d.}(\hat{\theta})_{\text{LMLE}}}{\text{s.d.}(\hat{\theta})_{\text{QRE}}} \tag{4.9}$$

where s.d. = standard deviation. Since the true variance of QRE's is larger than that of LMLE's, the population mean of the ratio R.P. must lie between zero and one.

The R.P. estimates, shown in Table 4.5, indicate that over a wide range of σ-values and sample sizes, the QRE procedure estimated the location parameter γ with at least 60 percent of the precision of the LMLE procedure. The relative precision does not seem to depend very much on the sample size for $n \geq 99$, but does seem to depend on the skewness of the parent distribution. The precision of the QRE's relative to the LMLE's tended to decrease as skewness increased. The average values (over all sample sizes) of R.P.(γ) for example, were 0.86, 0.75, 0.71, and 0.67, respectively, across the four distributions.

Also, the QRE procedure compared more favorably to the LMLE procedure in terms of the precision of the estimates of μ and σ than in terms of the estimates of γ. Estimates of R.P. averaged 0.806 and 0.813

Table 4.5. Empirical Relative Precision (R.P.) of Quantile Regression Estimates to Local Maximum Likelihood Estimates of 3PLN Parameters*

Sample Size n	R.P.(γ)	R.P.(μ)	R.P.(σ)
Distribution 1 ($\gamma = -10$, $\mu = 2.2976$, $\sigma = .0998$, skewness = 0.301)			
99	0.8029	0.8871	0.8541
299	1.0371	0.9261	0.8282
499	1.0737	0.9370	0.8587
699	0.7196	0.8465	0.8962
899	0.7120	0.8533	0.8957
Distribution 2 ($\gamma = -3.1623$, $\mu = 1.1036$, $\sigma = .3087$, skewness = 0.980)			
99	0.7535	0.8234	0.8331
299	0.6667	0.7518	0.7721
499	0.8039	0.8275	0.8313
699	0.6714	0.7458	0.8000
899	0.8316	0.8564	0.8733
Distribution 3 ($\gamma = -2$, $\mu = .5816$, $\sigma = .4724$, skewness = 1.625)			
99	0.6098	0.7437	0.7875
299	0.6954	0.7570	0.7725
499	0.7910	0.8283	0.8268
699	0.6764	0.7359	0.7825
899	0.7877	0.8296	0.8502
Distribution 4 ($\gamma = -1.4142$, $\mu = .1438$, $\sigma = .6368$, skewness = 2.475)			
99	0.6380	0.7556	0.7773
299	0.6530	0.7430	0.7464
499	0.7386	0.8112	0.7819
699	0.6173	0.6937	0.7394
899	0.6820	0.7640	0.7532

* Empirical relative precision of QRE procedure to LMLE procedure with respect to estimates of a parameter θ is defined as

$$R.P.(\theta) = \frac{s.d.(\hat{\theta})_{LMLE}}{s.d.(\hat{\theta})_{QRE}},$$

where s.d. = standard deviation.

for μ and σ, respectively, but only 0.748 for γ (estimates taken over all sample sizes and σ-values).

Clearly, there is not a great sacrifice to be made in estimating 3PLN parameters from grouped, rather than raw data. Still, there is a loss in efficiency of up to 40 percent. It is hoped that the information in Table 4.5 will be helpful to those faced with the decision of whether or not to condense large datasets.

Thus far, the QRE results have been compared only to LMLE results. Since the Monte Carlo study presented here was patterned after that of Cohen and Whitten, in which fourteen estimation methods (including LMLE) were compared, it is sensible to compare the present QRE results with the results of their thirteen alternatives to the LMLE procedure. The mean-squared error criterion has already been used to make comparisons among the methods presented in Cohen and Whitten (see Table 4.1). These comparisons presented strong evidence for the superiority of the LMLE procedure to the other thirteen in estimating the location parameter γ.

Now, the estimated mean-squared error (MSE) of the QRE procedure will be computed and compared with those of Cohen and Whitten's methods. To facilitate the interpretations of the results, the estimated MSE's of the QRE and thirteen other procedures will be divided into the estimated MSE of the LMLE procedure to form estimated indices of efficiency (EIE). Thus, higher estimated indices of efficiency correspond to more efficient estimation procedures according to the MSE criterion.

The results of the computations of the estimated indexes of efficiency are shown in Table 4.6. The denominator of each index was the estimated MSE for Cohen and Whitten's LMLE results rather than for the LMLE results from the present study. In this way, all estimates of MSE

Table 4.6. Estimated Indexes of Efficiency (EIE) of QRE, ME, and 12 Modified Estimation Methods of Cohen and Whitten.

	Distribution			
Procedure*	1	2	3	4
QRE	.22	1.15	.52	.50
ME	.65	.37	.10	.04
MME-I_1	.04	.17	.28	.42
MME-I_2	.13	.13	.47	.44
MME-I_3	.10	.31	.38	.36
MME-II_1	.08	1.12	.72	.79
MME-II_2	.30	.86	.87	.69
MME-II_3	.16	.03	.63	.51
MMLE-I_1	.08	.07	.17	.28
MMLE-I_2	.18	1.01	-†	-
MMLE-I_3	.35	.96	-	-
MMLE-II_1	.74	.65	.49	.58
MMLE-II_2	.35	1.02	-	-
MMLE-II_3	.35	.96	-	-

* QRE: quantile regression estimation
ME: method of moments estimation
MME: modified method of moments estimation
MMLE: modified maximum likelihood estimation
-I_r or -II_r: first or second modification type using rth order statistic where r=1, 2, or, 3.

† a dash (-) signifies that the estimation algorithm converged for fewer than 50% of the samples.

used to compile Table 4.6 were associated with the same data base except those computed from the QRE's. Thus, it should be understood that the EIE's of Cohen and Whitten's 13 alternative methods can be compared amongst each other with somewhat more control than is possible in comparing them with the EIE's of the QRE procedure. It should also be noted that in computing the EIE's of the QRE procedure, only results for n=99 were used, since Cohen and Whitten's sample sizes were of this same magnitude (n=100).

The estimates in Table 4.6 reveal that the efficiency of the QRE procedure compares well in most instances to those of the thirteen other methods. Over the four 3PLN distributions the estimated index for the QRE procedure ranked seventh, first, fourth and fifth-best, respectively. While these comparisons would be more meaningful had the QRE results pertained to Cohen and Whitten's actual data, they still offer strong evidence that the efficiency of the QRE procedure is comparable to those of many alternatives to the LMLE procedure.

4.2 Hill's (1963) Epidemic Data (3-Parameter Lognormal Distribution)

In Hill's (1963) paper, he showed that the conventional solution to the likelihood equations of the 3-parameter lognormal distribution leads to local rather than global maximum likelihood estimates. The global maximum likelihood estimates of (γ,μ,σ) were shown to be $(x_1,-\infty,+\infty)$ where x_1 is the first sample order statistic. To avoid this unwanted solution, Hill estimated the parameters using Bayesian techniques. The data used by Hill have subsequently been used by other researchers in applications of alternative estimation methods [Wingo (1975), Giesbrecht and Kempthorne (1976)]. In this section, Hill's data will be applied to the QRE procedure and the results will be compared to those of the alternative methods used by others. These alternative methods were reviewed in Chapter 2.

Hill's data are well-suited for quantile regression estimation because they are grouped. The data consist of 310 observations on the incubation period (rounded off to the nearest day) of inoculated small-pox virus. The data are given in Table 4.7. The value of γ, if the distribution were actually 3-parameter lognormal, was considered by Hill to be -4, corresponding to the day of inoculation. In reality, γ should be somewhat greater than -4, since the onset of smallpox could not possibly occur immediately upon inoculation.

Although the data are grouped, a more aggregated grouping at the extremes seemed to be preferable for use in the QRE procedure. Four sets of groupings were tried corresponding to the choice of 8, 9, 10, and 11 quantiles. These groupings are shown in Table 4.8. Note that in each case the last observation (X = 19 days) was dropped from

Table 4.7. Hill's (1963) Data on Incubation Period of Inoculated Smallpox

Day (X)	Observed Frequency	Cumulative Frequency
-4		
-3		
-2		
-1		
0		
1	2	2
2	6	8
3	17	25
4	77	102
5	96	198
6	73	271
7	22	293
8	8	301
9	3	304
10	3	307
11		
12	1	308
13	1	309
14		
15		
16		
17		
18		
19	1	310

Table 4.8. Additional Groupings of Hill's Data into Quantiles

Day (X_{pi})	Cumulative Frequency			
	8 quantiles	9 quantiles	10 quantiles	11 quantiles
1.5		2	2	2
2.5	8	8	8	8
3.5	25	25	25	25
4.5	102	102	102	102
5.5	198	198	198	198
6.5	271	271	271	271
7.5	293	293	293	293
8.5	301	301	301	301
9.5				304
10.5			307	307
11.5				
12.5	308	308	308	308
13.5*	309	309	309	309

* This quantile corresponds to $p_i = 1.0$ and was therefore discarded from the regression.

the data. The assumption was that such a large outlier would introduce unwanted distortion into the estimation process.

The results of applying the QRE procedure to each of the four aggregated datasets are presented in the upper portion of Table 4.9. This table gives the parameter estimates, and their asymptotic standard errors as well as χ^2 values and the probability of greater χ^2 values corresponding to a goodness-of-fit test for 3-parameter lognormality. The lower portion of Table 4.9 presents the estimates and goodness-of-fit test results that apply to analyses of these data found elsewhere in the literature.

The algorithm described in Section 1 of Appendix B was used to perform the quantile regression estimation. Pseudo-moment estimates (see Appendix B) were used as start values. The standard error estimates were taken as the square roots of the diagonal elements of the estimated asymptotic covariance matrix A of the parameters. The matrix A was computed according to

$$A = (F'V^{-1}F)^{-1}\, s^2, \tag{4.10}$$

where F is the estimated matrix of partial derivatives of the regression equation with respect to the parameters, V is the estimated covariance matrix of the quantiles, and s^2 is the variance of the regression residuals.

The χ^2 value for the goodness-of-fit test was computed according to

$$\chi^2_{k-4} = \Sigma[(f_i - \hat{f}_i)^2/\hat{f}_i] \tag{4.11}$$

where k was the number of quantiles, f_i was the observed frequency in the ith grouping interval, and $\hat{f}_i$ was the predicted frequency of the ith

Table 4.9. Summary Statistics for Various 3PLN Estimations of Hill Data.*

Estimation Method	γ	μ	σ	χ^2	Prob>χ^2
QRE: 8 quantiles	-2.09 (4.11)	1.96 (.589)	.192 (.113)	18.10	.0012
9 quantiles	-4.02 (6.31)	2.20 (.706)	.154 (.109)	16.48	.0024
10 quantiles	-4.11 (6.10)	2.21 (.674)	.152 (.103)	16.41	.0025
11 quantiles	-3.79 (5.61)	2.18 (.645)	.159 (.103)	16.79	.0021
Other Methods (using 8 quantiles)					
"True" from Hill	-4	2.20	.167	21.57	.0002
Bayesian LMLE from Hill	-2.41	2.01	.204	23.03	.0001
Discrete MLE from Giesbrecht & Kempthorne	-1.65	1.90	.222	21.75	.0002
LMLE via penalty function from Wingo	-2.45	2.01	.203	22.76	.0001

*Standard Errors in parentheses; all χ^2-values computed with respect to 8-quantile grouping.

grouping interval according to the estimated parameters. In all of the χ^2 calculations of Table 4.9, k was set equal to eight, corresponding to the observed frequencies in the 8-quantile grouping.

The most striking information to be learned from Table 4.9 is that Hill's data was almost certainly not lognormal! All procedures gave fairly similar parameter estimates and χ^2 values in the goodness-of-fit tests. The QRE procedure using 10 quantiles gave the smallest χ^2 value. However, even this value rejected the hypothesis of 3-parameter lognormality at the .01 level of significance.

It is unfortunate that several researchers have used this data to apply 3PLN estimation procedures. Still, the data do appear to come from a distribution similarly shaped to the 3-parameter lognormal, and for these data the QRE procedure certainly compared favorably to the procedures espoused by Hill (Bayesian MLE), Wingo (MLE with penalty function) and Giesbrecht and Kempthorne (discrete MLE). In fact, the χ^2 values corresponding to each set of QRE results were smaller than any of those given by the various MLE procedures. Thus, the QRE procedure would seem to be quite useful as long as it performed this well on data that were truly 3PLN-distributed.

4.3 Monte Carlo Study II--Six S_B Distributions from Johnson and Kotz (1970)

In this section, the sampling behavior of quantile regression estimates of the parameters of the S_B distribution is examined in comparison with that of local maximum likelihood estimates. The data consisted of 100 independently generated random samples of size 299 from each of six S_B distributions.

Both the QRE and LMLE procedures are most easily implemented conditionally on fixed values of the range parameters, α and β. To estimate all four parameters, a grid search must be conducted in the α, β plane to find that pair $(\hat{\alpha}, \hat{\beta})$ for which the global minimum (or maximum) of the appropriate objective function occurs. The details of the QRE and LMLE algorithms are given in Sections 3 and 4 of Appendix B, respectively.

The six S_B distributions investigated in this small Monte Carlo study were taken from Johnson and Kotz [1970, pp. 24-25]. Each distribution is bounded below by zero and above by one, but the six vary considerably in skewness and kurtosis. The shape of the density functions of each are plotted in Figure 4.3. The parameter values of μ and σ are given at the top of each plot.

The samples of size 299 were created by generating vectors of 299 independent standard normal values, z, and then forming S_B-distributed values, x, via the transformation

$$x = (\alpha + \beta y)/(1 + y), \tag{4.12}$$

where

$$y = \exp(\mu + \sigma z). \tag{4.13}$$

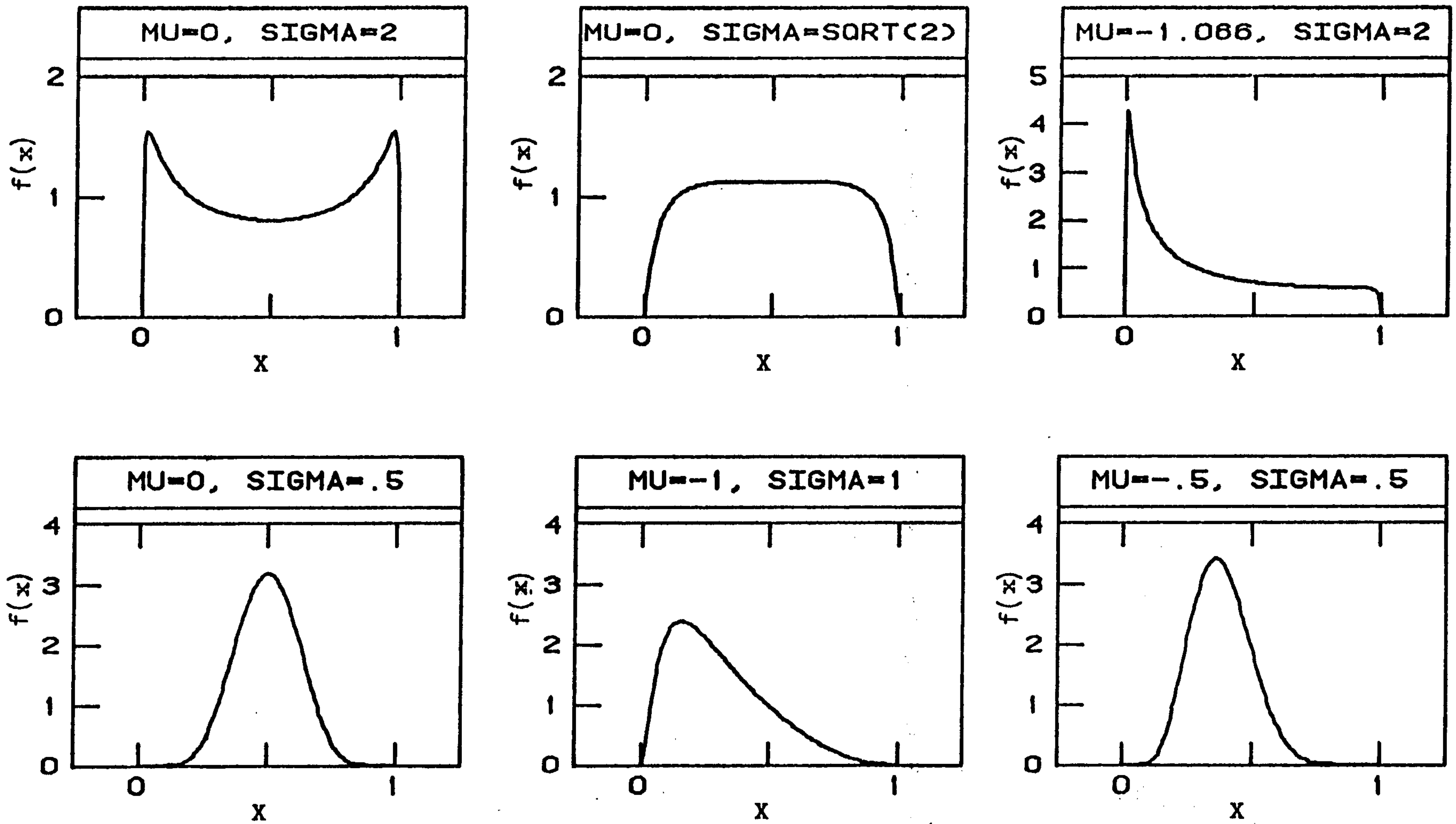

Figure 4.3. Probability density functions of six S_B distributions.

The sample size of 299 was decided upon after preliminary exploration revealed that the estimation algorithms frequently could not converge for smaller sample sizes (n=99) from the fourth and sixth of the six distributions.

In forming the quantiles, the following vector of quantile points, p, was used:

$$p = (.02, .08, .16, .26, .40, .60, .74, .84, .92, .98).$$

As in the case of the 3-parameter lognormal distribution, p was chosen somewhat arbitrarily, but with the intent of capturing information near the extremes of the data. In this case, however, p was chosen to be symmetric about .50, since three of the six distributions were symmetric.

The means and standard deviations of the S_B parameter estimates from the QRE and LMLE procedures are shown in Table 4.10. This table also shows the true parameter values of each distribution and the number of "runs" out of 100 for which the estimation algorithm converged. In this study, the estimation was halted if the estimate of α fell below -2 or if the estimate of β rose above 3. Without imposing these bounds the estimates would occasionally diverge toward plus or minus infinity.

Inspection of Table 4.10 reveals that for most of the distributions, both estimation procedures were reasonably accurate and precise. Clearly, the QRE procedure was somewhat less precise than the LMLE procedure, but in these cases the inferior precision did not lead to less accurate mean estimates. Both procedures did appear to give significantly biased estimates of σ, however, for distributions 4 and 6. In each case the true value of σ was 0.5 while the estimates of σ centered

Table 4.10. Means and Standard Deviations of S_B Parameter Estimates from LMLE and QRE Procedures

Parameter	True Value	Method = LMLE Mean	Method = LMLE Standard Deviation	Method = QRE Mean	Method = QRE Standard Deviation
Distribution 1					
α	0	0.0014	.0028	-0.0016	.0150
β	1	1.0061	.0029	1.0012	.0120
μ	0	-0.0404	.1096	0.0041	.1535
σ	2	1.9424	.0591	2.0236	.1643
		(Runs = 100)		(Runs = 100)	
Distribution 2					
α	0	0.0082	.0108	0.0003	.0326
β	1	0.9921	.0105	0.9973	.0301
μ	0	0.0055	.0882	0.0139	.1380
σ	1.4142	1.4631	.0711	1.4610	.1606
		(Runs = 100)		(Runs = 100)	
Distibution 3					
α	0	-0.0009	.0013	-0.0021	.0069
β	1	0.9980	.0083	1.0032	.0318
μ	-1.066	1.0351	.1093	-1.0456	.1680
σ	2	1.9614	.0581	2.0051	.1817
		(Runs = 100)		(Runs = 100)	
Distribution 4					
α	0	0.0254	.2388	0.0018	.2338
β	1	0.9736	.2377	0.9921	.2213
μ	0	0.0007	.1656	0.0080	.2186
σ	.5	0.5968	.1455	0.5998	.1927
		(Runs = 98)		(Runs = 85)	
Distribution 5					
α	0	0.0069	.0109	-0.0036	.0284
β	1	0.9552	.0731	1.0331	.1983
μ	-1	-0.9527	.1167	-1.0046	.1994
σ	1	1.0612	.0869	1.0198	.1729
		(Runs = 100)		(Runs = 99)	
Distribution 6					
α	0	0.0368	.0733	0.0185	.1200
β	1	0.9290	.2115	0.9881	.3389
μ	-.5	-0.4461	.2067	-0.4697	.2811
σ	.5	0.6063	.1304	0.6042	.1800
		(Runs = 96)		(Runs = 83)	

around 0.6. For all other parameters, though, the mean estimates were very close to the true parameter values. The precision of both procedures, especially that of the LMLE procedure, was remarkably good in the first three distributions, but declined for the latter three (more kurtotic) distributions. These subjective observations will now be refined by a more detailed analysis of the results.

Table 4.11 provides further insight into evidence of bias from each procedure. This table lists, for each procedure and each distribution, the mean parameter estimates, the value of the t-statistic corresponding to a test for unbiasedness and the probability of obtaining a greater absolute value of t. As expected, the procedures were significantly biased in the estimation of σ in distributions 4 and 6. But whereas there was little evidence of significant bias among the remaining quantile regression estimates, there was very strong evidence of bias in the majority of the remaining local maximum likelihood estimates. Since the degree of bias in the remaining LMLE's was small, this statistical significance only indicates that the estimation procedure is precise enough to detect a small bias for sample sizes of 299. The QRE's did not appear to be much more accurate than the LMLE's, but their variances were usually too large to render these inaccuracies statistically significant.

Since the results show no evidence that one procedure is more accurate than the other, the precision of the QRE procedure relative to the LMLE procedure would seem to serve as a valid criterion for evaluating the relative merits of the two procedures. Estimates of the relative precision associated with estimates of each of the four S_B parameters are given in Table 4.12.

Table 4.11. Tests for Bias in S_B Parameter Estimates from LMLE and QRE Procedures (Sample Size = 299)

		Method = LMLE			Method = QRE		
Parameter	True Value	Mean Estimate	t	Prob>|t|	Mean Estimate	t	Prob>|t|
Distribution 1							
α	0	0.0014	5.13	<.0001	-0.0016	-1.07	.2831
β	1	1.0061	20.78	<.0001	-1.0012	0.99	.3216
μ	0	-0.0404	-3.69	.0002	0.0041	0.27	.7870
σ	2	1.9424	-9.75	<.0001	2.0236	1.44	.1510
Distribution 2							
α	0	0.0082	7.61	<.0001	0.0003	0.10	.9203
β	1	0.9921	-7.58	<.0001	0.9973	-0.89	.3755
μ	0	0.0055	0.62	.5355	0.0139	1.01	.3149
σ	1.414	1.4631	6.88	<.0001	1.4610	2.91	.0036
Distribution 3							
α	0	-0.0009	-6.86	<.0001	-0.0021	-3.01	.0026
β	1	0.9980	-2.42	.0151	1.0032	1.01	.3145
μ	-1.066	-1.0351	2.83	.0047	-1.0456	1.21	.2245
σ	2	1.9614	-6.63	<.0001	2.0051	0.28	.7798
Distribution 4							
α	0	0.0254	1.05	.2930	0.0018	0.07	.9435
β	1	0.9736	-1.12	.2637	0.9921	-0.33	.7434
μ	0	0.0007	0.04	.9668	0.0080	0.34	.7348
σ	.5	0.5968	6.59	<.0001	0.5998	4.77	<.0001
Distribution 5							
α	0	0.0069	6.30	<.0001	-0.0036	-1.25	.2114
β	1	0.9552	-6.13	<.0001	1.0331	1.66	.0965
μ	-1	-0.9527	4.06	<.0001	-1.0046	-0.23	.8201
σ	1	1.0612	7.05	<.0001	1.0198	1.14	.2546
Distribution 6							
α	0	0.0368	4.92	<.0001	0.0185	1.41	.1598
β	1	0.9290	-3.29	.0010	0.9881	-0.32	.7490
μ	-.5	-0.4461	2.56	.0106	-0.4697	0.98	.3260
σ	.5	0.6063	7.99	<.0001	0.6042	5.27	<.0001

Table 4.12. Estimated Relative Precision of QRE's to LMLE's for Six S_B Distributions*

Distribution ID	Parameter			
	α	β	μ	σ
1	0.187	0.244	0.714	0.360
2	0.332	0.347	0.639	0.443
3	0.183	0.261	0.651	0.320
4	1.021	1.056	0.757	0.755
5	0.384	0.369	0.585	0.503
6	0.610	0.624	0.735	0.724

* Estimated relative precision of QRE procedure to LMLE procedure with respect to estimates of a parameter θ is defined as

$$R.P.(\theta) = \frac{s.d.(\hat{\theta})_{LMLE}}{s.d.(\hat{\theta})_{QRE}},$$

where s.d. = standard deviation.

The relative precision (R.P.) estimates in Table 4.12 vary considerably across the four parameters and down the six distributions. Some patterns do exist, however. First, down the distributions, the relative precision varied most for α and β, and least for μ. Second, for a given distribution, the relative precision was usually lowest for α and β and highest for μ. Finally, the precision of the QRE procedure was closer to that of the LMLE procedure when the kurtosis parameter σ was small. Since smaller values of σ also corresponded to poorer precision in both procedures, there was a negative correlation between the precision and the relative precision of the quantile regression estimates.

For distributions that could be estimated more precisely (i.e., when σ was high) the QRE procedure had poor relative precision, sometimes as low as 18 percent. Yet, since this low relative precision corresponded with high absolute precision, it may be said that the QRE procedure performed well, but the LMLE procedure performed extraordinarily well.

4.4 1960-1972 U.S. Family Income Data--Income Distribution Proposed by Singh and Maddala (1976)

The income distribution proposed by Singh and Maddala (1976) was presented in Section 4 of Chapter 3. The distribution function was given as

$$F_X(x) = 1 - (1 + ax^b)^{-c}, \quad X > 0,\ a > 0,\ c > 0. \tag{4.1}$$

Upon differentiation of the above function, the density function was found to be

$$f_X(x) = [abc\ x^{(b-1)}] \ / \ [(1 + ax^b)^{(c-1)}]. \tag{4.2}$$

Finally the quantile regression equation was derived as

$$X_{pi} = \{a^*[(1 - p_i)^{-c^*} - 1]\}^{b^*} + w_i, \tag{4.3}$$

where $a^* = 1/a$, $b^* = 1/b$, and $c^* = 1/c$.

In their 1976 paper, Singh and Maddala estimated the parameters of this distribution by employing the Davidon-Fletcher-Powell algorithm in a nonlinear regression program to minimize the quantity

$$\sum_{i=1}^{k} \{\ln[1 - F(x_{pi})] + c\ \ln(1 + ax_{pi}^b)\}^2 \tag{4.4}$$

This estimation procedure ignores the fact that the X_{pi} are correlated and heteroskedastic. Thus, the results given by this procedure are not expected to be as efficient as those obtainable via the QRE procedure. The purpose of the following analysis is to compare the performance of the QRE procedure [and also its unweighted (ordinary QRE) counterpart] with that of the procedure used by Singh and Maddala [Eq. (4.4)].

The income data used in this analysis are shown in Table 4.13. Notice that as nominal incomes rise, the number of income classes gradually increases from eleven in 1960-1967 to twelve in 1968-1970 to thirteen in 1971 and finally to fourteen in 1972. While the data were taken from the sources referenced by Singh and Maddala, there is no guarantee that the numbers given in Table 4.13 are identical to those used by Singh and Maddala. A telephone conversation with Professor Maddala revealed that it would be virtually impossible to recover the exact data used in his article.

The first analysis performed on the data in Table 4.13 consisted of an attempt to duplicate Singh and Maddala's results by estimating Eq. (4.4) using Marquardt's algorithm in the NLIN procedure of SAS. In the process of doing so, it was deduced that Singh and Maddala must have scaled their data by dividing each income by the median income for that year. This deduction was verified by the finding that the parameter estimates given in Singh and Maddala's paper predicted a median income very near to 1.0 for each of the thirteen years.

The results of the estimation of Eq. (4.4) together with those given in Singh and Maddala's paper are shown in Table 4.14. The estimates for each year are accompanied by a χ^2 goodness-of-fit statistic corresponding to the quantity

$$\chi^2_{k-4} = \Sigma[(f_j-\hat{f}_j)^2/\hat{f}_j]$$

where f_j and $\hat{f}_j$ are the observed and predicted frequencies, respectively, of the jth income class. For all χ^2 calculations, frequencies were predicted for the eleven income intervals given in Table 4.13 that

Table 4.13. 1960-1972 U.S. Family Income Data* (Table entries are number of families in corresponding income range.)

Income Range	1960	1961	1962	1963	1964	1965	1966	1967
0- 999	2285	2316	1950	1791	1532	1459	1149	1031
1000- 1999	3613	3573	3469	3250	2988	2956	2635	2189
2000- 2999	3970	4037	3901	3792	3864	3583	3197	2981
3000- 3999	4456	4387	4325	4142	4001	3806	3341	3155
4000- 4999	4773	4845	4669	4287	4113	3883	3474	3243
5000- 5999	5839	5439	5424	5253	4738	4502	4108	3879
6000- 6999	4889	4714	5100	4844	4714	4477	4574	4145
7000- 7999	3973	4231	4023	4300	4458	4683	4542	4414
8000- 9999	5135	5375	5804	6335	6635	6952	7408	7661
10000-14999	4795	5219	6019	6857	7761	8342	10008	11147
>15000	1707	2205	2314	2585	3031	3636	4486	5989
Total	45435	46431	46998	47436	47835	48279	48922	49834
Median	5620	5737	5956	6249	6569	6957	7532	7933

Income Range	1968	1969	1970
0- 999	909	804	829
1000- 1999	1717	1600	1535
2000- 2999	2756	2371	2237
3000- 3999	3081	2705	2613
4000- 4999	3031	2752	2728
5000- 5999	3485	3033	3026
6000- 6999	3839	3281	3122
7000- 7999	4142	3726	3294
8000- 9999	7678	7389	7054
10000-14999	12625	13682	13925
15000-24999	6112	8005	9193
>25000	1313	1889	2392
Total	50510	51237	51948
Median	8632	9433	9867

Income Range	1971
0- 999	784
1000- 1999	1339
2000- 2999	2242
3000- 3999	2574
4000- 4999	2888
5000- 5999	3027
6000- 6999	2955
7000- 7999	3327
8000- 9999	6560
10000-11999	6686
12000-14999	7674
15000-24999	10399
>25000	2841
Total	53296
Median	10285

Income Range	1972
0- 999	683
1000- 1999	1183
2000- 2999	2018
3000- 3999	2494
4000- 4999	2670
5000- 5999	2735
6000- 6999	2828
7000- 7999	3030
8000- 9999	6063
10000-11999	6245
12000-14999	7947
15000-19999	8597
20000-24999	3899
>25000	3982
Total	54373
Median	10650†

* Source: U.S. Bureau of the Census _Current Population Reports_, Series P-60 and P-20, 1960-1972.

† Projection, no data available.

Table 4.14. Parameter Estimates of Income Distribution of Singh and Maddala Obtained from Singh and Maddala's Estimation Procedure

Year	Parameter Estimates a	b	c	$\chi^{2\dagger}$
1. Results Reported in Singh and Maddala's Article				
1960	.2931	1.992	2.803	1798
1961	.2735	1.972	3.009	1620
1962	.3079	2.063	2.609	1635
1963	.3084	2.051	2.597	1557
1964	.3184	2.080	2.550	1332
1965	.3082	2.127	2.624	1926
1966	.3109	2.197	2.558	2167
1967	.3120	2.012	2.552	1035
1968	.3071	2.111	2.712	953
1969	.3101	2.131	2.611	1225
1970	.3102	2.121	2.546	1428
1971	.3125	2.139	2.544	1448
1972	.3070	2.064	2.538	846
2. Results Obtained in Present Estimation				
1960	.3385	2.130	2.459	2603
1961	.3578	2.081	2.364	2280
1962	.2760	2.023	2.942	1363
1963	.2037	1.961	3.846	1036
1964	.1515	1.899	5.059	565
1965	.1568	1.899	5.027	511
1966	.0848	1.882	8.862	572
1967	.0701	1.872	10.204	414
1968	.3928	2.261	2.181	1686
1969	.3228	2.144	2.576	1241
1970	.2715	2.033	2.997	875
1971	.2134	1.946	3.719	514
1972	.1863	1.927	3.894	397

† All χ^2 values are significant at the .01 percent level indicating that the assumed functional form of the income distribution was at best a useful approximation to the true distribution.

correspond to the years 1960-1967. Therefore, there are seven degrees of freedom associated with each χ^2 statistic.

Notice that for each year, the results obtained here differ from those reported by Singh and Maddala. Also note that for nine of the thirteen years, the estimates obtained here provided a better fit to the data than did those given in Singh and Maddala's paper. This suggests that either Singh and Maddala's data differed somewhat from that presented in Table 4.13, or that Singh and Maddala modified their estimation procedure somewhat, perhaps to enable themselves to obtain estimates that remained relatively stable across years. That the latter may have been done is suggested by the much smaller across-year variation in their parameter estimates relative to those obtained here. The computed χ^2 statistics in all cases were very highly significant, indicating that U.S. family income is not distributed according to this distribution. However, as Singh and Maddala point out, this distribution does provide a better fit to the data than do either of the Pareto or lognormal distributions, and thus represents an improvement over these two commonly assumed distributional forms.

Next, the QRE procedure [Eq. (4.3)] was applied to the data using both iteratively reweighted, and ordinary nonlinear least squares. For each year, the dependent variables in the regression, X_{pi}, consisted of the upper endpoints of the income classes listed in Table 4.13 normalized by the median income for that year. The "independent" variables, p_i, were computed as

$$p_i = N^{-1} \sum_{j=1}^{i} f_j$$

where N was the total number of families in the population and f_j was the number of families falling in the jth income class. The number of observations in each year's regression was one less than the number of available income classes since no quantile point could be obtained for the last class. A Gauss-Newton algorithm similar to that documented in Section 1 of Appendix B was used to perform the iteratively reweighted nonlinear least-squares (IRNLS) regressions. The parameter estimates and χ^2 statistics from each group of regressions are presented in Table 4.15.

For two of the years of data, 1966 and 1967, the IRNLS failed to converge. For these two years, $\hat{a}$ converged toward zero while $\hat{c}$ was diverging toward infinity. Inspection of the data in Table 4.13 and the estimation results in Table 4.15 reveals that there is a positive correlation between the magnitude of $\hat{c}$ and the proportion of the population in the highest income group. Since this proportion was largest in 1966 and 1967, it may be that the IRNLS algorithm failed for reasons relating to the excessive size of this omitted group.

The correlation between the size of $\hat{c}$ and the portion of families in the highest income class is disturbing because the class boundaries were defined arbitrarily, and therefore should not be related to the form of the distribution. A possible explanation for the apparent relationship is that there might exist paths in the (a,c) plane approaching the point $(0,\infty)$ along which the shape of the distribution changes very little. Perhaps, when no quantile points are available near the upper end of the distribution, the fit tends to be better at a point further down one of the paths.

Table 4.15. Parameter Estimates of Income Distribution of Singh and Maddala Obtained from Weighted and Ordinary QRE Procedures

Year	Parameter Estimates			$\chi^{2\dagger}$
	a	b	c	
1. Weighted QRE Procedure				
1960	.0623	1.720	12.155	1034
1961	.0809	1.695	9.382	870
1962	.0815	1.733	9.187	624
1963	.0456	1.714	16.072	505
1964	.0364	1.718	20.033	315
1965	.0253	1.704	29.377	418
1966	--	--	--	--
1967	--	--	--	--
1968	.1275	1.904	5.983	393
1969	.1297	1.874	5.877	451
1970	.1191	1.822	6.360	372
1971	.1184	1.813	6.416	299
1972	.0896	1.788	7.724	184
2. Ordinary (Unweighted) QRE Procedure				
1960	.1996	1.870	4.072	1252
1961	.2120	1.833	3.815	1069
1962	.1626	1.827	4.815	717
1963	.0985	1.780	7.621	554
1964	.0652	1.757	11.294	330
1965	.0476	1.736	15.671	430
1966	--	--	--	--
1967	--	--	--	--
1968	.2631	2.017	3.114	711
1969	.1954	1.928	4.025	590
1970	.1594	1.857	4.832	447
1971	.1396	1.830	5.488	330
1972	.1065	1.804	6.556	200

† All χ^2 values are significant at the .01 percent level indicating that the assumed functional form of the income distribution was at best a useful approximation to the true distribution.

To summarize, the QRE procedure seems to work very well as long as the objective of estimating this income distribution is to approximate the shape of the distribution rather than to obtain precise estimates of the individual parameters. In every year for which estimates could be obtained, the χ^2 statistics for the QRE's were smaller than those obtained using Singh and Maddala's estimation procedure. Furthermore, the χ^2 values corresponding to the weighted QRE's were always smaller than those corresponding to the ordinary (unweighted) QRE's.

In future applications of the QRE procedure, it would seem important that the investigator try to obtain quantile points within the extreme deciles of the distribution. In the present analysis, difficulties in obtaining estimates arose when 9.1 percent and 12.0 percent of the population, respectively, fell in the final income group. For the eleven years in which estimates could be obtained, this uppermost income group numbered less than 8 percent of the population.

4.5 Monte Carlo Study III--Application of Conditional QRE and Ordinary QRE Procedures to the Four 3-Parameter Lognormal Distributions of Cohen and Whitten

In Section 3.5, the conditional quantile regression estimation (CQRE) procedure was introduced as a possible alternative to the unconditional, or weighted QRE procedure outlined in Section 3.1. The usefulness of the CQRE procedure will depend heavily on how the sampling characteristics of its estimates compare with those of the weighted QRE procedure. Also, since the CQRE is especially recommended when weighted least squares is deemed undesirable, it is important to compare the properties of unweighted (or ordinary) CQRE's with those of ordinary quantile regression estimates (OQRE's).

The Monte Carlo study of this section examines the sampling behavior of conditional QRE's and ordinary QRE's. Each procedure was applied to the same 3PLN samples used in the Monte Carlo study of Section 4.1. Recall that the data are generated samples of size 99, 299, 499, 699, and 899 from each of four 3PLN distributions. Details of the data generation are given in Section 4.1.

The CQRE results are summarized in Table 4.16. This table presents the means and standard deviations of the CQRE parameter estimates and the number of samples (Runs) for which the estimation algorithm converged. In comparing these results with the LMLE and QRE results of Tables 4.2 and 4.3, it is found that the CQRE's for these samples are more biased and are estimated less precisely than either the LMLE's or the QRE's. For distributions 2 through 4 the bias of $\tilde{\gamma}$ was always negative and ranged between three and twenty percent of the true value

Table 4.16. Means and Standard Deviations of 3PLN Parameter Estimates Using Conditional QRE Procedure

Sample Size n	Mean γ	Mean μ	Mean σ	Standard Deviation γ	Standard Deviation μ	Standard Deviation σ	Runs
Distribution 1 (γ = -10, μ = 2.2976, σ = .0998, skewness = 0.301)							
99	-9.482	2.009	.1675	7.491	.6552	.0986	76
299	-12.094	2.291	.1182	8.881	.5984	.0620	87
499	-11.269	2.298	.1115	6.052	.4737	.0488	88
699	-11.410	2.323	.1064	6.163	.4410	.0416	94
899	-12.858	2.439	.0938	7.458	.4395	.0359	99
Distribution 2 (γ = -3.1623, μ = 1.1036, σ = .3087, skewness = 0.980)							
99	-3.754	1.129	.3472	2.1376	.5287	.1640	98
299	-3.459	1.144	.3110	1.1052	.3199	.0967	100
499	-3.454	1.159	.3032	.9356	.2711	.0804	100
699	-3.309	1.129	.3072	.6811	.2081	.0623	100
899	-3.341	1.150	.2967	.4803	.1575	.0480	100
Distribution 3 (γ = -2, μ = .5816, σ = .4724, skewness = 1.625)							
99	-2.244	.5701	.5179	.9067	.5027	.2158	100
299	-2.102	.5965	.4765	.4413	.2782	.1240	100
499	-2.121	.6220	.4645	.3959	.2343	.1029	100
699	-2.057	.5975	.4701	.2832	.1739	.0780	100
899	-2.082	.6206	.4561	.2158	.1382	.0614	100
Distribution 4 (γ = -1.4142, μ = .1438, σ = .6368, skewness = 2.475)							
99	-1.539	.0710	.6974	.4397	.5783	.2749	100
299	-1.469	.1365	.6442	.2443	.3111	.1586	100
499	-1.485	.1762	.6269	.2245	.2546	.1308	100
699	-1.446	.1540	.6341	.1589	.1841	.0979	100
899	-1.462	.1813	.6161	.1246	.1513	.0784	100

of γ. The precision of the parameter estimates for these three distributions was usually only about half that of estimates given by the QRE procedure. For Distribution 1, the bias and precision of the parameter estimates was not perceptibly different among the CQRE, LMLE and QRE procedures. However, in only 76 samples out of 100 could the CQRE procedure converge when the sample size was 99. The LMLE and QRE procedures converged more often (99 and 83 times, respectively) for these samples.

The CQRE procedure might still be of some value if it could provide estimates superior to those given by the OQRE procedure. Ordinary quantile regression estimation would be an obvious alternative whenever the weighted QRE procedure failed to converge. Therefore, it is important to determine which estimates are second best, CQRE's or OQRE's.

The OQRE results are summarized in Table 4.17. For Distributions 2-4, the OQRE's for all three parameters were always more accurate and more precise than were the CQRE's. This pattern was less pronounced for Distribution 1, but was still somewhat evident. As an interesting aside, comparisons of the OQRE's with the weighted QRE's of Tables 4.2 and 4.3 reveal strong evidence for the superiority of the weighted results in both accuracy and precision.

Therefore, the results of this Monte Carlo study lead to the recommendation that the CQRE procedure not be applied to samples from distributions similar to the four 3PLN distributions studied by Cohen and Whitten. If the conditioning of the covariance matrix or other considerations preclude the use of weighted QRE, then ordinary (unweighted) QRE should be attempted rather than conditional QRE.

Table 4.17. Means and Standard Deviations of 3PLN Parameter Estimates Using Ordinary QRE Procedure

Sample Size	Mean			Standard Deviation			
n	γ	μ	σ	γ	μ	σ	Runs
Distribution 1 (γ = -10, μ = 2.2976, σ = .0998, skewness = 0.301)							
99	-9.502	2.0238	.1589	7.877	.6161	.0802	77
299	-12.490	2.3385	.1115	8.668	.5787	.0564	89
499	-12.146	2.3614	.1047	7.408	.4828	.0441	95
699	-11.286	2.3274	.1045	5.535	.4066	.0367	96
899	-12.136	2.4101	.0950	5.771	.3834	.0323	100
Distribution 2 (γ = -3.1623, μ = 1.1036, σ = .3087, skewness = 0.980)							
99	-3.5039	1.1136	.3409	1.7132	.4188	.1281	99
299	-3.3977	1.1442	.3075	.8915	.2656	.0788	100
499	-3.3363	1.1398	.3060	.6521	.2021	.0619	100
699	-3.2616	1.1211	.3080	.5519	.1695	.0503	100
899	-3.3010	1.1409	.2994	.3891	.1296	.0386	100
Distribution 3 (γ = -2, μ = .5816, σ = .4724, skewness = 1.625)							
99	-2.1127	.5780	.5110	.6767	.3480	.1608	100
299	-2.0880	.6056	.4714	.3615	.2119	.0956	100
499	-2.0763	.6118	.4677	.2862	.1659	.0763	100
699	-2.0371	.5938	.4710	.2233	.1317	.0602	100
899	-2.0660	.6142	.4593	.1716	.1062	.0468	100
Distribution 4 (γ = -1.4142, μ = .1438, σ = .6368, skewness = 2.475)							
99	-1.4623	.1266	.6842	.3370	.3360	.1966	100
299	-1.4542	.1611	.6360	.1997	.2063	.1165	100
499	-1.4593	.1733	.6300	.1643	.1632	.0938	100
699	-1.4347	.1546	.6346	.1224	.1256	.0728	100
899	-1.4538	.1762	.6197	.0977	.1044	.0572	100

CHAPTER 5

SUMMARY AND CONCLUSIONS

The quantile regression estimation (QRE) procedure has been developed principally as a means of estimating the parameters of the 3-parameter lognormal (3PLN) and related distributions such as the 3-parameter log equicorrelated -normal (3PLEN) and S_B distributions. Parameters are estimated by regressing sample quantiles on their asymptotic expected values using iteratively reweighted (often nonlinear) least squares (IRLS) techniques. Since the data requirements are minimal (say, on the order of ten sample quantiles), the procedure is especially useful for preliminary analyses on large datasets that have yet to be keypunched, or for analyses of data that are only available in a condensed form. IRLS is necessary because the form of the asymptotic covariance matrix of the regression residuals depends upon the values of the unknown parameters.

The asymptotic means and variances of the quantiles can be obtained as long as the moments of the quantiles converge to the population moments of their limiting multivariate normal distribution. For the 3PLN and 3PLEN distributions, the asymptotic expectations of the sample quantiles were expressed in terms of the asymptotic moment generating functions of corresponding standard normal quantiles, which posses the necessary convergence properties. The S_B quantiles were espressed directly in terms of standard normal quantiles, which also exhibit moment convergence.

A more general application of the QRE procedure pertains to random variables whose inverse distribution functions (F^{-1}) are expressible in closed form. Provided moment convergence holds for quantiles from these distributions, the QRE procedure applies straightforwardly since the asymptotic expectations of the quantiles are directly obtainable from F^{-1}. In a recent application, Koutrouvelis (1981) independently derived the straightforward version of the QRE procedure for use in estimating the 2-parameter Pareto distribution.

Ultimately, judgements of the value of the QRE procedure must be based on the accuracy and precision of the estimates it produces. In Chapter 4 of this thesis, the sampling properties of QRE's from four 3-parameter lognormal and six S_B distributions were examined and compared to those of estimates from the local maximum likelihood estimation (LMLE) procedure. In these Monte Carlo studies LMLE's were obtained using whole samples, while of course, QRE's were derived using selected sample quantiles.

The results showed that the QRE procedure performed well. For moderately and heavily skewed 3-parameter lognormal distributions using sample sizes of 99 and greater, the estimated bias of the QRE's was small (less than 10 percent), and the relative precision of the QRE's to the LMLE's was never less than 60 percent. Among the six S_B distributions, the QRE's showed little bias except in estimating the kurtosis parameter σ when kurtosis was fairly high. The relative precision of the QRE's to the LMLE's varied substantially among the six distributions and four parameters, ranging from 18 to 106 percent. Fortunately, the low relative precisions corresponded to instances where the absolute precision of the QRE's was very high.

Twice in this thesis, the QRE procedure was applied to real-world grouped data. The 3-parameter lognormal distribution was estimated using data from Hill (1963) and the income distribution of Singh and

Maddala (1976) was estimated using 1960-1972 U.S. family income data. For each of these applications the estimates from the QRE procedure provided a superior fit to the data according to the χ^2-criteria than did estimates from alternative estimation procedures reported in the literature.

Amongst the various applications, the iteratively reweighted nonlinear least squares algorithm converged most of the time. In the 3PLN Monte Carlo studies, convergence was attained nearly always for samples of size 299 or greater. For samples of size 99, the algorithm failed to converge for 17 of 100 samples from the least skewed of the four distributions, but it converged in 299 of 300 samples from the three more skewed distributions. In the S_B Monte Carlo study, where all sample sizes were 299, convergence was a slight problem for two of the six distributions. In a preliminary analysis that used sample sizes of 99, the algorithm frequently failed to converge. It should be noted that these two "problem" S_B distributions were noticeably bell-shaped, whereas the remaining four had shapes that intersected the X-axis at a more severe angle.

In the applications to real data, convergence was attained for the Hill data and for eleven of the thirteen years of U.S. family income data. The two years for which convergence failed had larger proportions of families in the highest income category than did the other eleven years, suggesting that the procedure performs best when the uppermost quantile point is close to 1.0.

A conditional version of the QRE procedure was also developed in this thesis as an alternative means of estimating the 3-parameter lognormal distribution. The conditional version has the advantage of not requiring iteratively reweighted least squares in the estimation since much of the correlation among the quantiles is accounted for in the regression equation itself. Unfortunately, the results of this procedure were quite poor in comparison with the unconditional QRE results and even in comparison with unweighted unconditional QRE results.

Finally, a Monte Carlo study was conducted to examine how quickly the moment generating functions of standard normal quantiles, and equivalently, the moments of 3PLN quantiles converge to their asymptotic expected values. While the rate of convergence depends on both the quantile point, p, and the value of the 3PLN shape parameter, σ, it was found that on the whole, agreement between the actual and asymptotic moments of the quantiles was good for sample sizes of 59 or greater. Convergence was found to be fastest when p was near 0.5 and when σ was small.

In conclusion, the QRE procedure appears to be a valuable method of estimation to use on quantile data from a large class of distributions. The application is particularly attractive for the 3-parameter lognormal and S_B distributions, for which the regression equations may be expressed in terms of moment generating functions and expectations of standard normal quantiles, respectively, each of which have well-behaved asymptotic properties. When ungrouped samples are available from these distributions, local maximum likelihood estimation will tend to yield more precise estimates, and is therefore recommended over the QRE procedure unless data reduction (selection of quantiles) is desired for other reasons. Further work on the determination of optimal quantile spacings for condensing data might lead to significant improvements in the precision of quantile regression estimates.

LIST OF REFERENCES

Aitchison, J. and J. A. C. Brown. 1957. The Lognormal Distribution. Cambridge University Press, London.

von Bertalanffy, L. 1938. A quantitative theory of organic growth (inquiries on growth laws, II). Human Biology 10:181-213.

Blom, G. 1958. Statistical Estimates and the Transformed Beta-Variable. John Wiley and Sons, Inc., New York.

Bofinger, E. 1975. Optimal condensation of distributions and optimal spacing of order statistics. Journal of the American Statistical Association 70(349):151-154.

Calitz, F. 1973. Maximum likelihood estimation of the parameters of the three-parameter lognormal distribution--a reconsideration. Australian Journal of Statistics 3:185-190.

Cheng, S. W. 1975. A unified approach to choosing optimum quantiles for the ABLE's. Journal of the American Statistical Association 70(349):155-159.

Cohen, A. C. 1951. Estimating parameters of logarithmic-normal distributions by maximum likelihood. Journal of the American Statistical Association 46:206-212.

Cohen, A. C. and B. J. Whitten. 1980. Estimation in the three-parameter lognormal distribution. Journal of the American Statistical Association 75(370):399-404.

Cramér, H. 1946. Mathematical Methods of Statistics. Princeton University Press, Princeton, N. J.

David, H. A. 1970. Order Statistics. John Wiley and Sons, Inc., New York.

Eubank, R. L. 1981. A density-quantile function approach to optimal spacing selection. Annals of Mathematical Statistics 9(3):494-500.

Finney, D. J. 1941. On the distribution of a variate whose logarithm is normally distributed. Supplement to the Journal of the Royal Statistical Society VII:155-161.

Giesbrecht, F. and O. Kempthorne. 1976. Maximum likelihood estimation in the three-parameter lognormal distribution. Journal of the Royal Statistical Society, Series B, 38:257-264.

Griffiths, D. A. 1980. Interval estimation for the three-parameter lognormal distribution via the likelihood function. Journal of Applied Statistics 29(1):58-68.

Harter, H. L. 1961. Expected values of normal order statistics. Biometrika 48:151-166.

Harter, H. L., and A. H. Moore. 1966. Local-maximum-likelihood estimation of the parameters of three-parameter lognormal populations from complete and censored samples. Journal of the American Statistical Association 61:842-851.

Helwig, J. T., and K. A. Council (ed.). 1979. SAS User's Guide--1979 Edtion. SAS Institute, Inc., Cary, North Carolina.

Hill, B. M. 1963. The three-parameter lognormal distribution and Bayesian analysis of a point-source epidemic. Journal of the American Statistical Association 58:72-84.

Iwai, S. 1950. Duration curves of logarithmic normal distribution type and their application. Memoirs of the Faculty of Engineering, Kyoto University 12:1-19.

Johnson, N. L. 1949. Frequency curves generated by methods of translation. Biometrika 36:149-176.

Johnson, N. L., and S. Kotz. 1970. Continuous Univariate Distributions-1. Houghton Mifflin, Boston.

Kale, B. K. 1962. On the solution of the likelihood equations by iterative processes. Biometrika 49:479-486.

Kane, V. E. (forthcoming). Standard and goodness-of-fit parameter estimation methods for the three-parameter lognormal distribution. Union Carbide Corporation, Nuclear Division, Oak Ridge, Tennessee.

Kendall, M. G. and A. Stuart. 1958. The Advanced Theory of Statistics, Volume 1. Hafner, New York.

Koutrouvelis, I. A. 1981. Large sample quantile estimation in Pareto laws. Communications in Statistical Theory and Methods A10(2): 189-201.

Lambert, J. A. 1964. Estimation of parameters in the three-parameter lognormal distribution. Australian Journal of Statistics 6(2): 29-32.

Lambert, J. A. 1970. Estimation of parameters in the four-parameter lognormal distribution. Australian Journal of Statistics 12:33-43.

Michelini, C. 1972. The convergence pattern of the scoring method in estimating the parameters of a lognormal function. Journal of the American Statistical Association 67:319-323.

Mood, A. M., F. A. Graybill, and D. C. Boes. 1974. Introduction to the Theory of Statistics. McGraw-Hill Book Company, New York.

Mosteller, F. 1946. On some useful 'inefficient' statistics. Annals of Mathematical Statistics 17:377-402.

Munro, A. H., and R. A. J. Wixley. 1970. Estimation based on order statistics of small samples from a three-parameter lognormal distribution. Journal of the American Statistical Association 65:212-225.

Neter, J. and W. Wasserman. 1974. Applied Linear Statistical Models. Richard D. Irwin, Inc. Homewood, Illinois.

Ogawa, J. 1951. Contributions to the theory of systematic statistics, I. Osaka Mathematical Journal 3(2):175-213.

O'Neill, B., and W. T. Wells. 1972. Some recent results in lognormal parameter estimation using grouped and ungrouped data. Journal of the American Statistical Associateion 67(337):76-80.

Rao, C. R. 1952. Advanced Statistical Methods in Biometric Research. John Wiley and Sons, Inc., New York.

Sarhan, A. E., and B. G. Greenberg. 1962. Contributions to Order Statistics. John Wiley and Sons, Inc., New York.

Sarkadi, K. 1974. On the convergence of the expectation of the sample quantile, from Limit Theorems of Probability Theory, P. Révéz (ed.). North-Holland Publishing Company, Amsterdam, 1975.

Selby, S. L. (ed.). 1971. Standard Mathematical Tables--Nineteenth Edition. The Chemical Rubber Company, Cleveland.

Tiku, M. L. 1968. Estimating the parameters of the lognormal distribution from censored samples. Journal of the American Statistical Association 63:135-140.

U.S. Bureau of the Census. 1960-1972. Current Population Reports, Series P-60 and P-20, Washington, D.C.

Walker, A. M. 1968. A note on the asymptotic distribution of sample quantiles. Journal of the Royal Statistical Society, Series B, 30:570:575.

Wilson, E. B., and J. Worcester. 1945. The normal logarithmic transform. Review of Economic Statistics 27:17-22.

Wingo, D. R. 1975. The use of interior penalty functions to overcome lognormal distribution parameter estimation anomalies. Journal of Statistical Computation and Simulation 4:49-61.

Yuan, P. T. 1933. On the logarithmic frequency distribution and the semilogarithmic correlation surface. Annals of Mathematical Statistics 4:30-74.

APPENDIXES

APPENDIX A: MAXIMUM LIKELIHOOD ESTIMATION IN THE 3-PARAMETER LOGNORMAL AND S_B DISTRIBUTIONS

The 3-parameter lognormal (3PLN) distribution with parameters γ, μ and σ and the S_B distribution with parameters α, β, μ and σ are closely related in that each may be readily transformed to the normal (μ, σ^2) distribution. Indeed, X has the 3PLN distribution if $\ln(X - \gamma) \sim N(\mu, \sigma^2)$ and X has the S_B distribution if $\ln[(X - \alpha)/(\beta - X)] \sim N(\mu, \sigma^2)$. Therefore, it comes as no surprise to find that the same peculiarities that Hill (1963) discovered in 3PLN maximum likelihood estimation (see Chapter 2) also arise in S_B maximum likelihood estimation.

In this appendix Hill's proof that there exists a path in the parameter space along which the likelihood function of any ordered sample $x_1, \ldots, x_n$ from the 3PLN distribution approaches $+\infty$ will be presented followed by an analogous proof for the S_B distribution. In the second proof it will be shown that paths exist in the S_B parameter space along which the likelihood function approaches $+\infty$ as either $\alpha \to x_1$, or $\beta \to x_n$. Fortunately, unless either skewness is near zero or the sample size is very small, the likelihood functions from 3PLN and S_B samples usually have just one local maximum. The parameter estimates corresponding to this local maximum are called local maximum likelihood estimates. As the results of estimation in Section 4.1 of this thesis suggest, there is no reason to treat these local maximum likelihood estimates any differently than maximum likelihood estimates are normally treated.

Maximum Likelihood Estimation in the 3-Parameter Lognormal Distribution

Let $\ln(X - \gamma) \sim N(\mu, \sigma^2)$. The likelihood function of n independent observations from this distribution is

$$L(x_1,\ldots,x_n;\gamma,\mu,\sigma^2) = L(\gamma,\mu,\sigma^2)$$
$$= (2\pi\sigma^2)^{-n/2}\Pi(x_i - \gamma)^{-1}\exp-(1/2\sigma^2)\Sigma[\ln(x_i - \gamma) - \mu]^2. \quad (A.1)$$

Defining

$$\hat{\mu}(\gamma) = n^{-1}\Sigma\ln(x_i-\gamma) \quad (A.2)$$

and

$$\hat{\sigma}^2(\gamma) = n^{-1}\Sigma[\ln(x_i - \gamma) - \hat{\mu}(\gamma)]^2 \quad (A.3)$$

it is easily seen that

$$L^{**}(\gamma) = \sup_{\mu,\sigma^2} L(\gamma,\mu,\sigma^2) \propto [\hat{\sigma}(\gamma)]^{-n}\Pi(x_i - \gamma)^{-1}. \quad (A.4)$$

It will now be shown that

$$\lim_{\gamma\to x_1} L^{**}(\gamma) = +\infty. \quad (A.5)$$

Proof:

$$\hat{\sigma}^2(\gamma) = n^{-1}\Sigma[\ln(x_i - \gamma) - \hat{\mu}(\gamma)]^2 \quad (A.6)$$
$$\le n^{-1}\Sigma\ln^2(x_i - \gamma)$$
$$\le \ln^2(x_1 - \gamma), \text{ for } \gamma \text{ sufficiently near } x_1.$$

Hence,

$$L^{**}(\gamma) = [\hat{\sigma}(\gamma)]^{-n}\, \Pi\, (x_i - \gamma)^{-1} \quad (A.7)$$
$$\ge \left|\ln(x_1 - \gamma)\right|^{-n}\, \Pi\, (x_i - \gamma)^{-1}$$

for γ near x_1. Let $x_1 - \gamma = \varepsilon(\gamma)$. Then

$$L^{**}(\gamma) \geq \left| \ln\varepsilon(\gamma) \right|^{-n} \varepsilon^{-1}(\gamma) K(\gamma), \qquad \text{(A.8)}$$

where

$$K(\gamma) = \prod_{i=2}^{n} (x_i - \gamma)^{-1}.$$

The right hand side of Eq. (A.8) approaches infinity as ε approaches zero. Indeed, ignoring the constant term $K(\gamma)$ and noting that

$$\left|\ln\varepsilon(\gamma)\right| = -\ln\varepsilon(\gamma) \quad \text{for } 0 < \varepsilon(\gamma) < 1, \qquad \text{(A.9)}$$

permits the limit to be written as

$$\lim_{\varepsilon(\gamma)\to 0} \left| \ln\varepsilon(\gamma) \right|^{-n} \varepsilon^{-1}(\gamma) = \lim_{x\to 0} [f(x)/g(x)], \qquad \text{(A.10)}$$

where $x = \varepsilon(\gamma)$, $f(x) = x^{-1}$, and $g(x) = [-\ln(x)]^n$. L'Hospital's Rule may be used to evaluate the limit in Eq. (A.10):

$$\lim_{x\to 0} [f(x)/g(x)] = \lim_{x\to 0} [f^n(x)/g^n(x)] \qquad \text{(A.11)}$$

$$= \lim [1/(n!x)] = +\infty,$$

where $f^n(x)$ and $g^n(x)$ refer to the nth derivatives of $f(x)$ and $g(x)$, respectively.

By applying the result in Eq. (A.11) to Eq. (A.8) it is clear that $L^{**}(\gamma)$ approaches $+\infty$ as γ approaches x_1. Note from Eqs. (A.2) and (A.3) that if $\gamma = x_1$, then $\hat{\mu}(\gamma) = -\infty$ and $\hat{\sigma}(\gamma) = +\infty$.

Maximum Likelihood Estimation in the S_B Distribution

Let $\ln[(X - \alpha)/(\beta - X)] \sim N(\mu, \sigma^2)$. The density function of X is then

$$f_X(x) = (2\pi\sigma^2)^{-\frac{1}{2}}(\beta - \alpha)[(x - \alpha)(\beta - x)]^{-1} \times \exp-(1/2\sigma^2)\ln\{[(x - \alpha)/(\beta - x)] - \mu\}^2, \tag{A.12}$$

and for n independent observations ordered as $x_1, \ldots, x_n$ the likelihood function is

$$\begin{aligned} L(x_1,\ldots,x_n;\ \alpha,\beta,\mu,\sigma^2) &= L(\alpha,\beta,\mu,\sigma^2) \\ &= (2\pi\sigma^2)^{-n/2}(\beta - \alpha)^n\Pi[(x_i - \alpha)(\beta - x_i)]^{-1} \\ &\quad \times \exp-(1/2\sigma^2)\ \Sigma\ \ln\{[(x_i - \alpha)/(\beta - x_i)] - \mu\}^2. \end{aligned} \tag{A.13}$$

Defining

$$\hat{\mu}(\alpha,\beta) = n^{-1}\ \Sigma\ \ln[(x_i - \alpha)/(\beta - x_i)] \tag{A.14}$$

and

$$\hat{\sigma}^2(\alpha,\beta) = n^{-1}\ \Sigma\ \{\ln[(x_i - \alpha)/(\beta - x_i)] - \hat{\mu}(\alpha,\beta)\}^2 \tag{A.15}$$

it is seen that

$$\begin{aligned} L^{**}(\alpha,\beta) &= \sup_{\mu,\sigma^2} L(\alpha,\beta,\mu,\sigma^2) \\ &\propto [\hat{\sigma}(\alpha,\beta)]^{-n}\ \Pi\ [(x_i - \alpha)(\beta - x_i)]^{-1}. \end{aligned} \tag{A.15}$$

It will now be shown that

$$\lim_{\alpha \to x_1} L^{**}(\alpha,\beta) = +\infty \tag{A.16}$$

and

$$\lim_{\beta \to x_n} L^{**}(\alpha,\beta) = +\infty. \tag{A.17}$$

Proof:

<u>Case 1: $\alpha \to x_1$</u>

$$\begin{aligned}\hat{\sigma}^2(\alpha,\beta) &= n^{-1} \Sigma\{\ln[(x_i - \alpha)/(\beta - x_i)] - \hat{\mu}(\alpha,\beta)\}^2 \\ &\le n^{-1} \Sigma \ln^2[(x_i - \gamma)/(\beta - x_i)] \\ &\le \ln^2[(x_1 - \alpha)/(\beta - x_1)] \doteq \ln^2(x_1 - \alpha)\end{aligned} \tag{A.18}$$

for α sufficiently close to x_1. Hence,

$$\begin{aligned}L^{**}(\alpha,\beta) &= [\hat{\sigma}(\alpha,\beta)]^{-n} \Pi [(x_i - \alpha)(\beta - x_i)]^{-1} \\ &\ge \left| \ln(x_1 - \alpha) \right|^{-n} \Pi [(x_i - \alpha)(\beta - x_i)]^{-1}\end{aligned} \tag{A.19}$$

for α near x_1. Let $x_1 - \alpha = \varepsilon(\alpha)$. Then

$$L^{**}(\alpha,\beta) \ge \left| \ln\varepsilon(\alpha) \right|^{-n} \varepsilon^{-1}(\alpha)(\beta - x_1)^{-1}K(\alpha), \tag{A.20}$$

where

$$K(\alpha) = \prod_{i=2}^{n} [(x_i - \alpha)(\beta - x_i)]^{-1}.$$

Clearly the limit of the right-hand side of Eq. (A.20) as $\alpha \to x_1$ is equivalent to the limit of the right-hand side of Eq. (A.8) as $\gamma \to x_1$. Since this limit is $+\infty$, Eq. (A.16) has been proven. Note that if $\alpha = x_1$, then $\hat{\mu}(\alpha,\beta) = -\infty$ and $\hat{\sigma}(\alpha,\beta) = +\infty$ for all $x_n < \beta < \infty$.

<u>Case 2: $\beta \to x_n$</u>

$$\begin{aligned}\hat{\sigma}^2(\alpha,\beta) &= n^{-1} \Sigma \{\ln[(x_i - \alpha)/(\beta - x_i)] - \hat{\mu}(\alpha,\beta)\}^2 \\ &\le n^{-1} \Sigma \ln^2[(x_i - \gamma)/(\beta - x_i)] \\ &\le \ln^2[(x_n - \alpha)/(\beta - x_n)] \doteq \ln^2(\beta - x_n)\end{aligned} \tag{A.21}$$

for β sufficiently close to x_n. Hence,

$$L^{**}(\alpha,\beta) = [\hat{\sigma}(\alpha,\beta)]^{-n} \Pi [(x_i - \alpha)(\beta - x_i)]^{-1} \tag{A.22}$$
$$\geq \left| \ln(\beta - x_n) \right|^{-n} \Pi [(x_i - \alpha)(\beta - x_i)]^{-1}$$

for β near x_n. Let $\beta - x_n = \varepsilon(\beta)$. Then

$$L^{**}(\alpha,\beta) \geq \left| \ln\varepsilon(\beta) \right|^{-n} (\alpha - x_n)^{-1} \varepsilon^{-1}(\beta) K(\beta), \tag{A.23}$$

where

$$K(\beta) = \prod_{i=1}^{n-1} [(x_i - \alpha)(\beta - x_i)]^{-1}.$$

Again, the limit on the right hand side of Eq. (A.23) as $\beta \to x_2$ is clearly equivalent to the limit on the right hand side of Eq. (A.8) which was shown to be $+\infty$. Therefore Eq. (A.17) has been proven. Note that if $\beta = x_n$, then $\hat{\mu}(\alpha,\beta) = +\infty$ and $\hat{\sigma}(\alpha,\beta) = +\infty$ for all $-\infty < \alpha < x_1$.

If both $\alpha \to x_1$ and $\beta \to x_n$ simultaneously, the likelihood function also approaches $+\infty$. However, in this case $\lim_{\substack{\alpha \to x_1 \\ \beta \to x_n}} \hat{\mu}(\alpha,\beta)$ is undefined.

APPENDIX B: SELECTED ESTIMATION ALGORITHMS

Each of the analyses presented in Chapter 4 required substantial software development. All of the analysis programs were written in SAS (Statistical Analysis System) and executed on an IBM 3033 computer. In addition, several control programs written in SUPERWYLBUR were utilized. Documented listings of the source codes of these programs are available from the author upon request. In this appendix, representative algorithms used in the iteratively reweighted least squares and local maximum likelihood estimations in Chapter 4 are given.

B.1 Algorithm to Find Quantile Regression Estimates from 3PLN Samples

Given that a vector of k quantiles $x_{p1}, x_{p2}, \ldots, x_{pk}$ has been obtained, the following steps were followed in obtaining quantile regression estimates of the parameters. The notation used matches that of Chapter 3.

1. Compute the "independent variables" for the regression-- $c_i = \Phi^{-1}(p_i)$, $k_{ij} = [p_i(1-p_j)]/[n\phi(c_i)\phi(c_j)]$, for $i,j=1,2,\ldots,k$, $i<j$.
2. Compute pseudo-moment estimates estimates $\theta^* = (\gamma^*, \mu^*, \sigma^*)$ of $\theta = (\gamma, \mu, \sigma)$ for use as start values in the nonlinear estimation (see below).
3. Set $m = 0$, $\hat{\theta}_m = \theta^*$, V_m = identity matrix.
4. Set $m=m+1$. Compute $\hat{\theta}_m$ = 1-step Gauss-Newton estimate of θ based on $\hat{\theta}_{m-1}$ and $\hat{V}_{m-1}$. Compute $\hat{V}_m$ using $\hat{\theta}_m$.

5. If $\hat{\gamma}_m$ is less than user-specified minimum permissible estimate, then stop iteration and note that estimation has failed for this sample.

6. If the sum of the elements of $\left|\hat{\theta}_m - \hat{\theta}_{m-1}\right|$ is greater than .001 then return to step 4.

7. Output $\hat{\theta}_m$ as estimate of θ.

The Gauss-Newton procedure requires that starting values be supplied for the parameter estimates. Also, in this generalized nonlinear least squares situation, start values are required for the elements of the variance-covariance matrix V of the quantiles.

Regarding the choice of start values for the parameters, it is best if little work is required to obtain them, as long as they are close enough to the solution for the estimation algorithm to converge. The method of moments provides a relatively easy means of obtaining unbiased (but inefficient) parameter estimates which might serve as start values. Unfortunately, in the strictest sense, moment estimates are not obtainable from grouped data because the grouping precludes the computation of the sample moments. However, a pseudo-moment estimation procedure was developed whose estimates worked extremely well as start values.

This pseudo-moment estimation procedure is so named because a pseudo-dataset of n observations is created from the quantiles and then the method of moments is applied to the pseudo-data. To form the pseudo-data from a set of k quantiles, it is assumed that all of the observations in the intervals between each pair of consecutive quantiles fell exactly at the midpoints of these intervals. The observations less than the first quantile were assumed to equal an amount 50 percent less than the first quantile.

To form the initial estimate of the variance-covariance matrix V, the identity matrix was used. This too, worked extremely well. Since it worked well, there did not seem to be any point in using a more complicated expression. A further justification for using the identity matrix is that it gives unbiased parameter estimates.

B.2 Algorithm to Find Local Maximum Likelihood Estimates from 3PLN Samples

Given an estimate of γ, the conditional maximum likelihood estimates of μ and σ^2 are simply

$$\hat{\mu}(\gamma) = (1/n) \sum_{i=1}^{n} \ln(x_i - \gamma) \text{ and} \tag{B.1}$$

$$\hat{\sigma}^2(\gamma) = (1/n) \sum_{i=1}^{n} [\ln(x_i - \gamma) - \hat{\mu}(\gamma)]^2. \tag{B.2}$$

Traditionally, the method of scoring has been successfully used to compute maximum likelihood estimates from lognormal samples [see Lambert (1964), Harter and Moore (1966), and O'Neill and Wells (1972)]. Here the local maximum likelihood estimate of $\theta=(\gamma, \mu, \sigma)$ is found by searching over the permissible range of γ for that value γ^* for which the triplet $[\gamma^*, \hat{\mu}(\gamma^*), \hat{\sigma}(\gamma^*)]$ maximizes the value of the likelihood function. This method was easy to program and worked quite well.

The likelihood function can be shown to attain a local maximum for that value of γ for which $\{\ln[\hat{\sigma}(\gamma)] + \hat{\mu}(\gamma)\}$ attains a local minimum. The logarithm of the likelihood function of a sample from the 3-parameter lognormal distribution is

$$L(x_1,\ldots,x_n;\ \gamma,\ \mu,\ \sigma) =$$

$$-(n/2)\ln(2\pi\sigma^2) - (1/2\sigma^2)\sum_{i=1}^{n}[\ln(x_i-\gamma)-\mu]^2 - \sum_{i=1}^{n}\ln(x_i-\gamma). \quad \text{(B.3)}$$

When the likelihood equation is evaluated at the point $[\gamma,\ \hat{\mu}(\gamma),\ \hat{\sigma}(\gamma)]$ it simplifies to

$$L[x_1,\ldots,x_n;\ \gamma,\ \mu(\hat{\gamma}),\ \sigma(\hat{\gamma})] = -(n/2)\ln[2\pi\hat{\sigma}^2(\gamma)] - (n/2) - n\hat{\mu}(\gamma). \quad \text{(B.4)}$$

This expression can be rearranged to

$$L[x_1,\ldots,x_n;\ \gamma,\ \mu(\hat{\gamma}),\ \sigma(\hat{\gamma})] = -(n/2)[\ln(2\pi)+1] - n[\ln\hat{\sigma}(\gamma)+\hat{\mu}(\gamma)]. \quad \text{(B.5)}$$
$$= K_1 - n[\ln\hat{\sigma}(\gamma)+\hat{\mu}(\gamma)],$$

where the constant K_1 is independent of the parameter estimates. Thus the likelihood function attains a local maximum when $[\ln\hat{\sigma}(\gamma)+\hat{\mu}(\sigma)]$ attains a minimum over a restricted range of γ-values. (As γ is allowed to get very close to the first order statistics x_1, the value of the likelihood function corresponding to γ tends to infinity).

In the algorithm that follows, the maximum of the likelihood function over the range of γ-values in the interval $(-\beta,\ x_1-.10\ x_1)$ is found and the triplet $[\gamma^*,\hat{\mu}(\gamma^*),\hat{\sigma}(\gamma^*)]$ corresponding to this maximum is considered to be the local maximum likelihood estimate.

1. Compute $x_1 = \min(x_i;\ i=1,2,\ldots,n)$.
2. Set initial limits within which the search for γ^* will take place:

$$\gamma_L = x_1 - 4|x_1| \quad \text{(lower limit).} \quad \text{(B.6)}$$

$$\gamma_U = x_1 - .10|x_1| \quad \text{(upper limit).} \quad \text{(B.7)}$$

Set r=1.

3. Set $D = (\gamma_U - \gamma_L)/10$. Compute $L^*[\hat{\mu}(\gamma), \hat{\sigma}(\gamma)]$ corresponding to $\gamma = \gamma_L, \gamma_L + D, \gamma_L + 2D, \ldots, \gamma_U$; where $\hat{\mu}(\gamma)$ and $\hat{\sigma}(\gamma)$ are defined by Eq. (B.1) and (B.2) and where

$$L^*[\hat{\mu}(\gamma), \hat{\sigma}(\gamma)] = \ln\hat{\sigma}(\gamma) + \hat{\mu}(\gamma). \tag{B.8}$$

Define

$$\tilde{\gamma} = \mathrm{Min}\{L^*[\hat{\mu}(\gamma), \hat{\sigma}(\gamma)]\}. \tag{B.9}$$

4. If $\gamma_L < \hat{\gamma} < \gamma_U$ then go to step 5. If $\hat{\gamma} = \gamma_L$ then set $r=2r$. If $\hat{\gamma} = \gamma_U$ then set $r = (1/2)r$;
Set $\gamma_L = x_1 - 4r|x_1|$. Return to step 3.
5. Set $\delta_L = \tilde{\gamma} - D$ and $\gamma_U = \tilde{\gamma} + D$. If $(\gamma_U - \gamma_L) > .001$ then return to step 3.
6. Take local maximum likelihood estimates to be $\tilde{\gamma}$, $\hat{\mu}(\tilde{\gamma})$, $\hat{\sigma}(\tilde{\gamma})$.

Of course the barrier of Eq. (B.7) can be adjusted as the user sees fit. For the parent distributions used in the Monte Carlo study of Chapter 4, the barrier given in Eq. (B.7) worked well. In all cases a maximum was found less than this barrier value. Also the initial establishment of a lower bound for $\tilde{\gamma}$ can be adjusted. This initial bound will not affect the final estimate of $\tilde{\gamma}$, but only the speed at which the estimate is found. For small sample sizes the likelihood surface tends to be more irregular and several local maxima may exist. In such cases, the global maximum within the permissible range of γ may be more easily found if the stepsize D is decreased.

B.3 Algorithm to Find Quantile Regression Estimates from S_B Samples

This algorithm implements a combined grid search and least squares estimation technique. Of the four S_B parameters (α, β, μ and σ), μ and σ are estimated via generalized least squares conditional on values of α and β. A grid search is conducted in the (α, β) plane for that quartet of values α, β, $\hat{\mu}(\alpha, \beta)$, $\hat{\sigma}(\alpha, \beta)$ which yields predicted quantiles that minimize the weighted sum of squared residuals from the observed quantiles.

The regression equation is given in Eq. (3.34) of Chaper 3. The estimation is iterative in two ways. First, the region of the (α, β) plane within which the grid search is conducted must be successively narrowed. Second, the estimated covariance matrix of the quantiles, being a function of σ, must be successively updated by the most recently obtained estimate of σ.

The following algorithm presumes a vector of k quantiles x_{p1}, ..., x_{pk} from a random sample of size n from an S_B distribution.

1. Define: $x = (x_{p1}\ x_{p2}\ \ldots\ x_{pk})'$.

   $p = (p_1\ p_2\ \ldots\ p_k)'$.

   $z = \Phi^{-1}(p)$.

   $Z = [1 \mid z]$, where 1 is a $k \times 1$ vector of ones.

   K = k by k asymptotic covariance matrix of z with typical element

$$k_{ij} = [p_i(1-p_j)]/[n\phi(z_i)\phi(z_j)] \text{ for } p_i < p_j$$

   where n is the sample size and ϕ is the standard normal probability density function.

Set the initial value of V, the covariance matrix of the quantiles, equal to the k-dimensional identity matrix.

2. Define the minimum and maximum allowable estimates of the range parameters, α and β. This permits the algorithm to be stopped whenever estimates of α are diverging toward $-\infty$ or x_{p1}, or whenever estimates of β are diverging toward $+\infty$ or x_{pk}. For the six S_B distributions (all with $\alpha = 0$, $\beta = 1$) estimated in Chapter 4, the bounds on the estimates were set as follows.

   Lower bound on $\hat{\alpha}$: $\alpha_L^* = -2$,

   Upper bound on $\hat{\alpha}$: $\alpha_U^* = x_{p1} - .01|x_{p1}|$,

   Lower bound on $\hat{\beta}$: $\beta_L^* = x_{pk} + .01|x_{pk}|$,

   Upper bound on $\hat{\beta}$: $\beta_U^* = 3$.

3. Set initial lower and upper limits of the grid in the (α, β) plane. The grid search will begin within these limits, but the limits will be allowed to change if they are later found not to enclose the least squares solution. For the six S_B distributions investigated in Chapter 4, the initial limits of the grid were

   <u>α-dimension</u>

   Initial lower limit: $\alpha_L = \alpha_U^* - .1(\text{RANGE})$

   Initial upper limit: $\alpha_U = \alpha_U^*$

   <u>β-dimension</u>

   Initial lower limit: $\beta_L = \beta_L^*$

   Initial upper limit: $\beta_U = \beta_L^* + .1(\text{RANGE})$,

where RANGE = $\beta_L^* - A_U^*$.

Also, choose the number of gridsteps, S. The size of each grid step is, in the α-dimension

$$d\alpha = (\alpha_U - \alpha_L)/(S - 1),$$

and in the β-dimension

$$d\beta = (\beta_U - \beta_L)/(S - 1).$$

The algorithm cannot work with $S < 3$. It is recommended, however, that S be at least 5. For the estimations of Chapter 4, $S = 7$ was used.

4. For each grid point $(\tilde{\alpha}, \tilde{\beta})$ form the dependent regression variable

$$\tilde{y} = \ln[(x - \tilde{\alpha})/(\tilde{\beta} - x)]$$

and compute $\tilde{\theta} = [\mu(\tilde{\alpha}, \tilde{\beta}), \sigma(\tilde{\alpha}, \tilde{\beta})]$ where

$$\tilde{\theta} = (Z'V^{-1}Z)^{-1} Z'V^{-1}\tilde{y}.$$

Then compute

$$\begin{aligned} \hat{y}(\tilde{\alpha}, \tilde{\beta}) &= \mu(\tilde{\alpha}, \tilde{\beta}) + \sigma(\tilde{\alpha}, \tilde{\beta})z \\ &= Z\tilde{\theta}, \end{aligned}$$

and

$$\hat{x}(\tilde{\alpha}, \tilde{\beta}) = \{\tilde{\beta} \exp[\hat{y}(\tilde{\alpha}, \tilde{\beta})]\}/\{\exp[\hat{y}(\tilde{\alpha}, \beta)] + 1\}.$$

Finally, compute

$$SSE(\tilde{\alpha}, \tilde{\beta}) = [x - \hat{x}(\tilde{\alpha}, \tilde{\beta})]' V^{-1} [x - \hat{x}(\tilde{\alpha}, \tilde{\beta})].$$

5. Let (α^*, β^*) equal the grid point $(\tilde{\alpha}, \tilde{\beta})$ for which SSE $(\tilde{\alpha}, \tilde{\beta})$ was a minimum. If (α^*, β^*) is an interior point on the grid then go to step 6. If (α^*, β^*) is on the border of the grid then go to step 7.

6. If $d\alpha + d\beta < .001$ then go to step 9. Otherwise recompute V and narrow grid as follows:

$$V = [\sigma^2(\alpha^*, \beta^*)]K,$$

$$\alpha_L = \alpha^* - d\alpha, \qquad \beta_L = \beta^* - d\beta,$$

$$\alpha_U = \alpha^* + d\alpha, \qquad \text{and} \quad \beta_U = \beta^* + d\beta.$$

Return to step 4.

7. If $\alpha^* < \alpha^*_L$ then stop. Note that estimation failed.
If $\alpha^*_U - \alpha^* < .00001$ then stop. Note that estimation failed.
If $\beta^* - \beta^*_L < .00001$ then stop. Note that estimation failed.
If $\beta^* < \beta^*_U$ then stop. Note that estimation failed.
If $\alpha^* = \alpha_L$ then $\alpha_L = \alpha_L - .2(\text{RANGE})$. Otherwise $\alpha_L = x^* - d\alpha$.
If $\alpha^* = \alpha_U$ then $\alpha_U = \text{Min}(\alpha^* + d\alpha, \alpha^*_U)$. Otherwise $\alpha_U = \alpha^* + d\alpha$.
If $\beta^* = \beta_L$ then $\beta_L = \text{Max}(\beta^* - d\beta, \beta^*_L)$. Otherwise $\beta_L = \beta^* - d\beta$.
If $\beta^* = \beta_U$ then $\beta_U = \beta_U + .2(\text{RANGE})$. Otherwise $\beta_U = \beta^* + d\beta$.

8. $V = [\sigma^2(\alpha^*, \beta^*)]K$. Return to step 4.

9. Take estimates of S_B parameters to be

$$\hat{\alpha} = \alpha^*, \ \hat{\beta} = \beta^*, \ \hat{\mu} = \mu(\alpha^*, \beta), \ \hat{\sigma} = \sigma(\alpha^*, \beta^*).$$

B.4 <u>Algorithm to Find Local Maximum Likelihood Estimates from S_B Samples</u>

The likelihood function of a sample $x_1, \ldots, x_n$ from an S_B distribution is

$$L(x;\ \alpha,\ \beta,\ \mu,\ \sigma) = \tag{B.10}$$

$$(2\pi\sigma^2)^{-n/2} \exp\{-(1/2\sigma^2) \sum_i [\ln \frac{x_i - \alpha}{\beta - x_i} - \mu]^2\} \prod_i \frac{(\beta - \alpha)}{(x_i - \alpha)(\beta - x_i)} .$$

When α and β are known, the random variable $y = \ln[(X - \alpha)/(\beta - X)] \sim N(\mu,\ \sigma^2)$. Therefore, conditional on α and β, the maximum likelihood estimates of μ and σ are

$$\hat{\mu}(\alpha,\ \beta) = n^{-1} \sum_i y_i = n^{-1} \sum_i \ln\ [(x_i - \alpha)/(\beta - x_i)] \tag{B.11}$$

and

$$\hat{\sigma}(\alpha,\ \beta) = n^{-1} \sum_i [y_i - \hat{\mu}(\alpha,\ \beta)]^2 \tag{B.12}$$

$$= n^{-1} \sum_i \{\ln[(x_i - \alpha)/(\beta - x_i)] - \hat{\mu}(\alpha,\ \beta)\}^2 .$$

In logarithmic form the conditional likelihood function may be written

$$\ln L(\alpha,\ \beta) = -(n/2)[\ln(2\pi\hat{\sigma}^2) + 1] + n \ln(\beta-\alpha) - \sum_i \ln[(x_i-\alpha)(\beta-x_i)] . \tag{B.13}$$

The local maximum likelihood parameter estimates may be found by maximizing Eq. (B.13) via a grid search in the $(\alpha,\ \beta)$ plane. The algorithm follows the QRE algorithm fairly closely, except that the conditional objective functions differ. Thus, the algorithm will be given in terms of the QRE algorithm to save time.

Let $x_1,\ \ldots,\ x_n$ be a random sample from an S_B distribution.

1. Perform step 2 of QRE algorithm.
2. Perform step 3 of QRE algorithm.
3. For each grid point $(\tilde{\alpha},\ \tilde{\beta})$ compute

$$\tilde{\mu} = n^{-1} \sum_i [\ln(x_i - \tilde{\alpha})/(\tilde{\beta} - x_i)]$$

and

$$\tilde{\sigma}^2 = n^{-1} \sum \{[\ln(x_i - \tilde{\alpha})/(\tilde{\beta} - x_i)] - \tilde{\mu}\}^2.$$

Evaluate Eq. (B.13) using $\tilde{\alpha}$, $\tilde{\beta}$, $\tilde{\mu}$, and $\tilde{\sigma}^2$ to obtain ln $L(\tilde{\alpha}, \tilde{\beta})$.

4. Let (α^*, β^*) equal the grid point $(\tilde{\alpha}, \tilde{\beta})$ for which ln $L(\tilde{\alpha}, \tilde{\beta})$ was a maximum. If (α^*, β^*) is an interior point on the grid then go to step 5. If (α^*, β^*) is on the border of the grid then go to step 6.

5. If $d\alpha + d\beta < .001$ then go to step 7. Otherwise narrow grid as follows:

$$\alpha_L = \alpha^* - d\alpha, \qquad \beta_L = \beta^* - d\beta,$$
$$\alpha_U = \alpha^* + d\alpha, \qquad \text{and} \quad \beta_U = \beta^* + d\beta.$$

Return to step 3.

6. Perform step 7 of QRE algorithm. Return to step 3.

7. Take estimates of S_B parameters to be

$$\hat{\alpha} = \alpha^*, \ \hat{\beta} = \beta^*, \ \hat{\mu} = \mu(\alpha^*, \beta^*), \ \hat{\sigma} = \sigma(\alpha^*, \beta^*).$$

APPENDIX C: SOME INITIAL WORK ON DETERMINING THE OPTIMAL SPACING OF 3-PARAMETER LOGNORMAL QUANTILES

Whenever data are available in ungrouped, or raw form, use of the quantile regression estimation procedure must be preceded by the selection of quantiles from the raw data. Immediately, the question arises of how many quantiles to select, and which ones? Recall that choosing the largest possible set of quantiles, i.e., the full set of order statistics, invalidates the QRE procedure.[7] Thus, the number of quantiles should be small relative to the sample size. Once the number of quantiles k has been decided, there should be at least one spacing of the k quantile points for which the information retained from the sample about a particular parameter or parameters is a maximum. If the QRE procedure were applied to such optimally spaced quantiles, the precision of the parameter estimates would be optimized as well.

In the following discussion, a spacing S_k is defined as $S_k = \{p_i\} = p_1, p_2, \ldots, p_k$ where $0 < p_1 < p_2 < \ldots < p_k < 1$. When there is no chance of ambiguity the k-subscript will be dropped to give a cleaner notation.

Ogawa (1951) has outlined the procedure for determining an optimal spacing of k quantiles for an arbitrary continuous random variable X with density g(x) and unknown parameter θ. He defined the information about θ in a sample as I(θ), and in a set of quantiles as $I_S(\theta)$, where the quantiles have been formed according to the spacing S_k. Ogawa

---

[7]The QRE procedure relies on the asymptotic multivariate normal distribution of the quantiles. Order statistics can not be used in the QRE procedure because their joint distribution does not tend towards a multivariate normal.

defined the optimal quantile spacing for a given k as that spacing for which the relative information of the quantiles, $I_S(\theta)/I(\theta)$ was maximized. Equivalently, the optimal spacing is that S_k which maximizes $I_S(\theta)$.

In order to compute $I_S(\theta)$ for a given spacing of quantiles, the joint density function of the quantiles must be computed. While the exact joint distribution of the quantiles is not known, their asymptotic distribution, assuming g(x) is absolutely continuous, is known to be multivariate normal (see Theorem 3.2). Specifically, the joint density function of the k quantiles for a given spacing S_k is

$$h(x_{p1},\ldots,x_{pk}) = (2\pi^{-k/2} g_1 g_2 \ldots g_k [p_1(p_2-p_1)\ldots(p_k-p_{k-1})(1-p_k)]^{-\frac{1}{2}} n^{k/2}$$
$$\times \exp\{-\frac{n}{2} \;[\sum_{i=1}^{k} \frac{p_{i+1}-p_{i-1}}{(p_{i+1}-p_i)(p_i-p_{i-1})} \; g_i^2(x_{pi}-c_i)^2$$
$$-2\sum_{i=2}^{k} \frac{g_i g_{i-1}}{p_i-p_{i-1}}(x_{pi}-c_i)(x_{p,i-1}-c_{i-1})]\},$$

where, as in Chapter 3, c_i is the p_ith population quantile. The information in $\{X_{pi}\}_S$, the set of quantiles determined by the spacing S, about a parameter θ is then defined, in the Fisher sense as

$$I_S(\theta) = E[\partial \ln h/\partial\theta]^2 = -E[\partial^2 \ln h/\partial\theta^2].$$

In general, however, finding $I_S(\theta)$ is very difficult and often will depend upon the value of the unknown parameter θ. Then, the simultaneous solution to the k equations $\partial I_S(\theta)/\partial p_i = 0 \quad i=1,\ldots,k$ must be found and these equations have not been derived for the 3-parameter lognormal distribution. Ogawa has derived the system of equations for the normal distribution when either μ, or σ, or both are unknown. Cheng

(1975) developed a procedure to follow for obtaining the system of equations for any density function of the form $g(x) = \sigma^{-1}f[(x-\mu)/\sigma]$. However, the 3-parameter lognormal distribution does not fit into this class of distributions.

Thus far, the only contribution that has been made towards the determination of optimal spacings of 3PLN quantiles was the empirical study of O'Neill and Wells (1972). They estimated the information about each of the three parameters contained in grouped samples formed according to two types of spacings, equal and logarithmic. Their results indicated that the asymptotic variances of the parameter estimates were smaller when the data were grouped in logarithmic rather than equal intervals. These results were obtained from samples having considerable skewness ($\hat{\sigma} > .65$). The advantage of the logarthmic grouping might not be as great for less-skewed distributions.

A small empirical study was conducted here to help indicate a suitable value for p_1, the first quantile point, in estimating the location parameter γ for the four 3PLN distributions of Cohen and Whitten (see Section 4.1). The study consisted of selecting 10 quantiles from samples of size 299 from these four distributions. For each sample, five sets of quantiles were formed corresponding to five spacings that differed only in the first quantile point. The first quantile point p_1 was varied from 0.01 to 0.05 in increments of 0.01, while the remaining quantile points were held fixed at the values given in Section 4.1. Ten samples from each distribution were generated and the means and standard deviations of the QRE's corresponding to each of the five spacings were computed. The results, given in Table C.1, show that when $p_1 = 0.01$, the parameters were estimated with the greatest precision, and usually with the greatest accuracy.

Table C.1. Means and Standard Deviations of 3PLN Parameter Estimates for Various Definitions of the First Sample Quantile (p_1)*

p_1	Mean γ	Mean μ	Mean σ	Standard Deviation γ	Standard Deviation μ	Standard Deviation σ	Runs
Distribution 1 (γ = -10, μ = 2.2976, σ = .0998, skewness = 0.301)							
.01	-12.183	2.337	.1061	8.101	.5605	.0467	10
.02	-10.481	2.260	.1100	4.677	.4363	.0458	9
.03	-11.800	2.370	.0996	5.773	.4542	.0432	9
.04	-14.872	2.445	.1000	13.628	.6887	.0508	9
.05	-13.308	2.426	.0989	8.243	.5852	.0486	9
Distribution 2 (γ = -3.1623, μ = 1.1036, σ = .3087, skewness = 0.980)							
.01	-3.129	1.083	.3134	.4511	.1520	.0466	10
.02	-3.233	1.112	.3063	.5831	.1908	.0541	10
.03	-3.342	1.148	.2963	.5854	.1858	.0515	10
.04	-3.317	1.140	.2989	.5878	.1905	.0537	10
.05	-3.345	1.148	.2969	.6035	.1926	.0545	10
Distribution 3 (γ = -2, μ = .5816, σ = .4724, skewness = 1.625)							
.01	-1.976	.5647	.4756	.1576	.0986	.0479	10
.02	-2.011	.5830	.4685	.1991	.1259	.0542	10
.03	-2.053	.6082	.4584	.1979	.1228	.0522	10
.04	-2.043	.6020	.4610	.2079	.1243	.0545	10
.05	-2.054	.6079	.4590	.2137	.1263	.0561	10
Distribution 4 (γ = -1.4142, μ = .1438, σ = .6368, skewness = 2.475)							
.01	-1.400	.1305	.6382	.0677	.0762	.0505	10
.02	-1.415	.1437	.6312	.0859	.0987	.0553	10
.03	-1.436	.1633	.6211	.0857	.0973	.0539	10
.04	-1.430	.1584	.6237	.0921	.0955	.0564	10
.05	-1.436	.1630	.6216	.0955	.0976	.0589	10

*For each distribution, 10 3PLN samples of size 299 were generated.

Unfortunately, this analysis was conducted subsequent to the large Monte Carlo study reported in Section 4.1. Had it been conducted sooner, the first quantile point used in the creation of grouped data sets in Section 4.1 would have been set to 0.01 rather than 0.03.

APPENDIX D: A MONTE CARLO STUDY OF THE SMALL SAMPLE DISTRIBUTION OF EXP(σZ_p)

An important step in deriving the quantile regression estimation (QRE) equations for the 3-parameter lognormal (3PLN) distribution involves the recognition that 3PLN quantiles are related by a monotonic transformation to standard normal quantiles. Thus, their expectations may be computed in terms of the expectations of corresponding standard normal quantiles. While the exact expectations could not be expressed in closed form, it was shown in Chapter 3 that a closed-form expression did exist for the asymptotic expectations of the 3PLN quantiles. These were computed in terms of the moments of the limiting multivariate normal distribution of the standard normal quantiles. The purpose of this Monte Carlo study is to assess how accurately the asymptotic moments represent the exact moments of 3PLN quantiles for small sample sizes.

Let $0 < p_1 < p_2 < \ldots < p_k < 1$, and let $X_{p1}, X_{p2}, \ldots, X_{pk}$ be k corresponding sample quantiles from a 3PLN distribution. Using the relationship between the 3PLN and standard normal quantiles

$$X_{pi} = \gamma + \exp(\mu)\exp(\sigma Z_{pi}), \tag{D.1}$$

it can easily be shown how the expectations of X_{pi} depend upon the first two moments of $\exp(\sigma Z_{pi})$. Indeed,

$$\begin{aligned} E(X_{pi}) &= E[\gamma + \exp(\mu)\exp(\sigma Z_{pi})] \\ &= \gamma + \exp(\mu)E[\exp(\sigma Z_{pi})] \end{aligned} \tag{D.2}$$

and

$$\mathrm{Var}(X_{pi}) = \mathrm{Var}[\gamma + \exp(\mu)\exp(\sigma Z_{pi})] \tag{D.3}$$
$$= \exp(2\mu)\mathrm{Var}[\exp(\sigma Z_{pi})].$$

Since the limiting distributions of the Z_{pi} are normal with means $c_i = \Phi^{-1}(p_i)$ and variances $k_{ii} = [p_i(1-p_i)]/n\phi^2(c_i)$, the limiting expectations and variances of the $\exp(\sigma Z_{pi})$ can be derived by treating them as moment generating functions of normal random variables. The resulting asymptotic expectations are

$$\lim_{n\to\infty} E[\exp(\sigma Z_{pi})] = \lim_{n\to\infty} \exp(\sigma c_i + 1/2\sigma^2 k_{ii}) \tag{D.4}$$

$$= \exp(\sigma c_i) \tag{D.5}$$

since $\lim_{n\to\infty} k_{ii} = 0.$

The variances, of course, converge to zero. However,

$$\lim_{n\to\infty} \mathrm{Var}[n^{\frac{1}{2}}\exp(\sigma Z_{pi})] = \lim_{n\to\infty} n \exp(2\sigma c_i)\exp(\sigma^2 k_{ii})[\exp(\sigma^2 k_{ii})-1], \tag{D.6}$$

tends to a finite limit. Rather than computing this limit, it was evaluated numerically by computing the right-hand side of Eq. (D.6) for larger and larger values of n until near convergence was attained.

The purpose of this Monte Carlo study is to compare the empirical means and variances of $\exp(\sigma Z_{pi})$ with their asymptotic expectations when the sample size is small (between 19 and 99). The divergence of the empirical means and variances from their asymptotic expectations will, apart from sampling variation, depend upon the value of σ, the quantile point p_i, and the sample size n. Therefore, these three factors are varied in this study. The various values of σ, p_i and n for which the sampling behavior of $\exp(\sigma Z_{pi})$ was investigated are shown in Table D.1. The four values of σ shown in Table D.1 are those that were used to define the four 3PLN distributions studied by Cohen and Whitten [1980] (and again in Chapter 4 of this paper).

The Monte Carlo study was conducted as follows. For each combination of n and p_i, 10,000 random samples of size n were generated from the standard normal distribution. Values of n were selected in such a way that for each p_i in Table D.1 there would be an order statistic in the sample whose cumulative distribution function had p_i as its expected value.[8] Next, the p_ith quantile, z_{pi}, was computed for each sample and then the corresponding four values of $\exp(\sigma z_{pi})$ were computed for each of the four values of σ. In this way, random samples of size 10,000 on $\exp(\sigma Z_{pi})$ were constructed for each of 140 combinations of σ, p_i and n. For a given combination of p_i and n, the samples corresponding to the four σ-values were formed from the same sample of Z_{pi} so that comparisons of the sampling behavior of $\exp(\sigma Z_{pi})$ across the range of σ would be more controlled.

To complete the Monte Carlo study, the first four sample moments were calculated for each of the 140 samples of quantiles and these moments were compared to their asymptotic expectations. The mean of the asymptotic distribution of $\exp(\sigma Z_{pi})$ is given by Eq. (D.5). The expression for the asymptotic variance of $n^{\frac{1}{2}}\exp(\sigma Z_{pi})$ is given in Eq. (D.6).

---

[8]A sufficient condition for such a value of n given a particular p_i is given in section 1 of Chapter 4.

Table D.1. Values of σ, p_i, and n Used in the Monte Carlo Study of $\mathrm{Exp}(\sigma Z_{pi})$

σ	p_i	n
.0998	.05	19
.3087	.10	39
.4724	.20	59
.6368	.50	79
	.80	99
	.90	
	.95	

While all of the first four moments are examined in this study, only the first two moments must closely approximate their asymptotic expectations for the QRE procedure to be accurate. Results for the third and fourth moments are given only to assist the curious reader in determining how nearly normal are the small sample quantiles. The raw results of the Monte Carlo study are shown in Tables D.2-D.5, in which the means, variances, skewness, and kurtosis of the 140 samples are given, respectively. To aid in appreciating the effects of n, p_i and σ on the approximation, two sets of plots are given. In Figure D.1, for the seven values of p_i, the percentage differences between the empirical and asymptotic values of $E[\exp(\sigma Zp_i)]$ are plotted as a function of the sample size for each value of σ. Figure D.2 presents analogous plots for $Var[n^{\frac{1}{2}}\exp(\sigma Z_{pi})]$.

Several patterns emerge in each of the figures. The notable patterns are summarized by the following observations:

(a) The empirically determined percentage differences between $E[\exp(\sigma Z_{pi})]$ and $\lim_{n\to\infty} \exp(\sigma Z_{pi})$ were, holding other factors constant, closer to zero when

1. n was larger,
2. p was nearer .20, and
3. σ was smaller.

(b) The empirically determined percentage differences between $Var[n^{\frac{1}{2}}\exp(\sigma Z_{pi})]$ and $\lim_{n\to\infty} Var[n^{\frac{1}{2}}\exp(\sigma Z_{pi})]$ tended to be closer to zero, holding other factors constant, when

1. n was larger,
2. p was near .50, and
3. σ was smaller.

Table D.2. Means of $\mathrm{Exp}(\sigma Z_{pi})$

Sample Size n	p_i=.05	p_i=.10	p_i=.20	p_i=.50	p_i=.80	p_i=.90	p_i=.95
<u>Distribution 1</u> (σ = .0998)							
19	0.8340	0.8716	0.9159	1.0001	1.0936	1.1492	1.2018
39	0.8412	0.8752	0.9168	1.0002	1.0910	1.1425	1.1903
59	0.8439	0.8769	0.9174	1.0000	1.0902	1.1414	1.1864
79	0.8451	0.8777	0.9176	1.0003	1.0895	1.1401	1.1844
99	0.8456	0.8778	0.9178	1.0001	1.0895	1.1393	1.1833
∞	0.8486	0.8799	0.9195	1.0000	1.0876	1.1365	1.1784
<u>Distribution 2</u> (σ = .3087)							
19	0.5750	0.6573	0.7648	1.0032	1.3239	1.5458	1.7815
39	0.5879	0.6639	0.7657	1.0020	1.3115	1.5137	1.7212
59	0.5931	0.6673	0.7668	1.0011	1.3078	1.5081	1.7011
79	0.5953	0.6688	0.7671	1.0016	1.3047	1.5019	1.6913
99	0.5961	0.6689	0.7674	1.0009	1.3046	1.4986	1.6856
∞	0.6018	0.6732	0.7712	1.0000	1.2966	1.4854	1.6618
<u>Distribution 3</u> (σ = .4724)							
19	0.4329	0.5294	0.6661	1.0081	1.5431	1.9598	2.4457
39	0.4455	0.5358	0.6659	1.0046	1.5175	1.8913	2.3068
59	0.4509	0.5394	0.6669	1.0027	1.5099	1.8787	2.2615
79	0.4531	0.5411	0.6672	1.0032	1.5039	1.8663	2.2398
99	0.4539	0.5411	0.6675	1.0020	1.5035	1.8593	2.2273
∞	0.4597	0.5458	0.6720	1.0000	1.4881	1.8321	2.1754
<u>Distribution 4</u> (σ = .6368)							
19	0.3276	0.4278	0.5816	1.0153	1.8052	2.4982	3.3882
39	0.3382	0.4329	0.5797	1.0083	1.7594	2.3702	3.1063
59	0.3431	0.4363	0.5803	1.0051	1.7459	2.3457	3.0165
79	0.3450	0.4378	0.5803	1.0055	1.7358	2.3236	2.9744
99	0.3456	0.4376	0.5805	1.0036	1.7346	2.3109	2.9501
∞	0.3508	0.4421	0.5852	1.0000	1.7089	2.2619	2.8510

Table D.3. Variances of $n^{\frac{1}{2}}\text{Exp}(\sigma Z_{p_i})$

Sample Size n	p_i=.05	p_i=.10	p_i=.20	p_i=.50	p_i=.80	p_i=.90	p_i=.95
Distribution 1 (σ = .0998)							
19	0.0333	0.0231	0.0171	0.0156	0.0253	0.0401	0.0733
39	0.0313	0.0225	0.0170	0.0155	0.0243	0.0375	0.0675
59	0.0316	0.0218	0.0170	0 0158	0.0242	0.0366	0.0638
79	0.0309	0.0229	0.0169	0.0158	0.0242	0.0390	0.0624
99	0.0314	0.0224	0.0174	0.0156	0.0241	0.0382	0.0613
∞	0.0321	0.0225	0.0172	0.0156	0.0240	0.0376	0.0618
Distribution 2 (σ = .3087)							
19	0.1481	0.1245	0.1136	0.1510	0.3605	0.7097	1.6203
39	0.1439	0.1233	0.1133	0.1495	0.3390	0.6397	1.3959
59	0.1483	0.1206	0.1137	0.1524	0.3363	0.6180	1.2785
79	0.1441	0.1268	0.1132	0.1526	0.3333	0.6515	1.2342
99	0.1482	0.1243	0.1165	0.1502	0.3329	0.6358	1.2079
∞	0.1543	0.1263	0.1157	0.1497	0.3270	0.6147	1.1765
Distribution 3 (σ = .4724)							
19	0.1942	0.1881	0.2017	0.3589	1.1659	2.7326	7.5076
39	0.1918	0.1878	0.2005	0.3527	1.0698	2.3696	6.0513
59	0.1999	0.1844	0.2018	0.3585	1.0565	2.2638	5.3829
79	0.1964	0.1940	0.2007	0.3592	1.0402	2.3664	5.1331
99	0.2003	0.1905	0.2065	0.3531	1.0382	2.3032	5.0008
∞	0.2108	0.1944	0.2057	0.3505	1.0085	2.1900	4.7214
Distribution 4 (σ = .6368)							
19	0.2010	0.2230	0.2800	0.6661	2.9584	8.2942	27.7200
39	0.1997	0.2227	0.2761	0.6477	2.6346	6.8673	20.6585
59	0.2098	0.2192	0.2784	0.6559	2.5861	6.4733	17.7509
79	0.2062	0.2306	0.2762	0.6571	2.5275	6.7027	16.6891
99	0.2103	0.2266	0.2841	0.6447	2.5200	6.5019	16.1706
∞	0.2230	0.2317	0.2834	0.6370	2.4168	6.0652	14.7363

Table D.4. Skewness of $Exp(\sigma Z_{pi})$

Sample Size n	p_i=.05	p_i=.10	p_i=.20	p_i=.50	p_i=.80	p_i=.90	p_i=.95
Distribution 1 (σ = .0998)							
19	-0.1864	-0.1190	-0.0358	0.0719	0.2560	0.3020	0.5180
39	-0.1940	-0.0684	-0.0530	0.0426	0.1508	0.2634	0.4591
59	-0.1261	-0.0532	0.0255	0.0320	0.1907	0.1971	0.3396
79	-0.1531	-0.0862	0.0072	0.0567	0.0965	0.1232	0.2918
99	-0.1621	-0.0251	-0.0100	0.0514	0.1065	0.1509	0.3286
Distribution 2 (σ = .3087)							
19	0.1203	0.1339	0.1642	0.2541	0.4823	0.5657	0.8678
39	0.0141	0.0984	0.0906	0.1671	0.2991	0.4418	0.7207
59	0.0488	0.0843	0.1470	0.1337	0.3107	0.3397	0.5278
79	-0.0059	0.0406	0.1049	0.1467	0.1997	0.2461	0.4531
99	-0.0226	0.0820	0.0814	0.1286	0.1984	0.2631	0.4786
Distribution 3 (σ = .4724)							
19	0.3559	0.3289	0.3204	0.3983	0.6673	0.7857	1.1685
39	0.1741	0.2284	0.2025	0.2649	0.4182	0.5862	0.9507
59	0.1843	0.1920	0.2425	0.2138	0.4064	0.4537	0.6841
79	0.1077	0.1401	0.1817	0.2176	0.2813	0.3437	0.5841
99	0.0853	0.1661	0.1532	0.1894	0.2707	0.3518	0.6002
Distribution 4 (σ = .6368)							
19	0.5930	0.5250	0.4783	0.5457	0.8622	1.0232	1.5015
39	0.3336	0.3590	0.3148	0.3639	0.5413	0.7366	1.2130
59	0.3199	0.3006	0.3391	0.2948	0.5045	0.5709	0.8507
79	0.2210	0.2406	0.2591	0.2892	0.3644	0.4432	0.7207
99	0.1931	0.2509	0.2256	0.2509	0.3439	0.4421	0.7267

Table D.5. Kurtosis of $\text{Exp}(\sigma Z_{pi})$

Sample Size n	p_i=.05	p_i=.10	p_i=.20	p_i=.50	p_i=.80	p_i=.90	p_i=.95
Distribution 1 (σ = .0998)							
19	2.9999	3.0435	2.9531	3.0249	3.2014	3.1568	3.3730
39	3.0234	2.9474	3.0170	2.9866	3.0670	3.1416	3.5048
59	3.0609	3.0016	3.0907	3.9734	3.0780	3.1039	3.2368
79	3.0383	3.0905	2.9482	3.0226	3.0470	3.0164	3.2269
99	3.1452	2.9946	3.0142	2.9548	3.0490	3.0886	3.3554
Distribution 2 (σ = .3087)							
19	2.9494	3.0175	2.9797	3.1196	3.5033	3.5806	4.1762
39	2.9440	2.9425	3.0105	3.0258	3.2084	3.3670	4.2032
59	3.0182	3.0085	3.1239	3.0052	3.1934	3.2438	3.5787
79	2.9814	3.0890	2.9695	3.0559	3.1081	3.0986	3.4568
99	3.0897	3.0042	3.0276	2.9835	3.0930	3.1660	3.5853
Distribution 3 (σ = .4724)							
19	3.1246	3.1499	3.0941	3.2781	3.8873	4.1355	5.1891
39	2.9846	3.0043	3.0546	3.0943	3.3819	3.6344	5.0863
59	3.0614	3.0591	3.1867	3.0559	3.3222	3.4074	3.9658
79	2.9881	3.1319	3.0091	3.1024	3.1846	3.2043	3.7163
99	3.0937	3.0402	3.0594	3.0208	3.1486	3.2600	3.8327
Distribution 4 (σ = .6368)							
19	3.4898	3.4192	3.2935	3.5164	4.4264	4.9473	6.6608
39	3.1134	3.1248	3.1419	3.1983	3.6193	3.9949	6.4287
59	3.1705	3.1502	3.2829	3.1301	3.4885	3.6242	4.4845
79	3.0385	3.2153	3.0698	3.1675	3.2881	3.3493	4.0576
99	3.1385	3.1019	3.1104	3.0715	3.2235	3.3856	4.1485

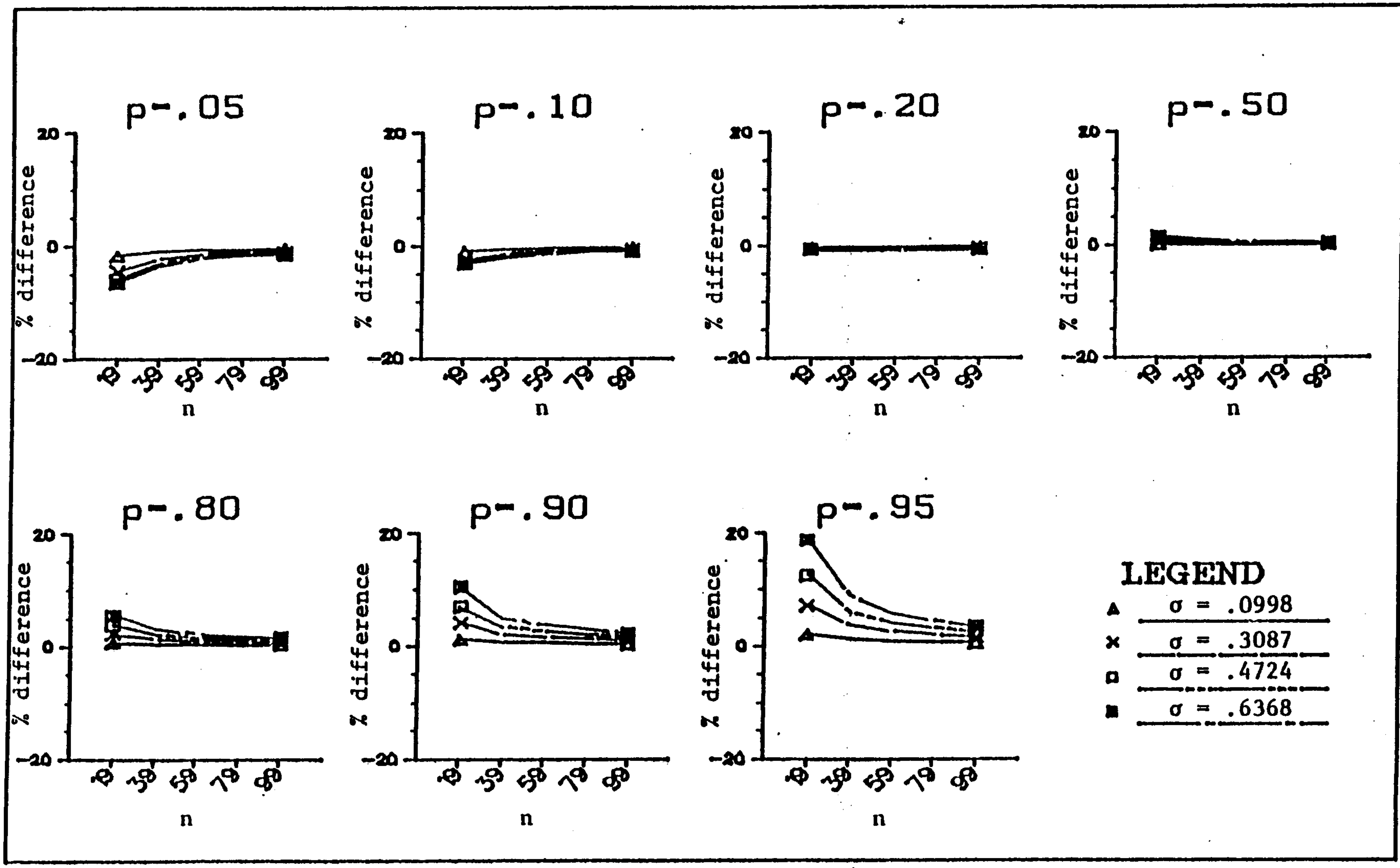

Figure D.1. Percentage differences between empirical and limiting values of $E[\exp(\sigma Z_{pi})]$ for varying values of n, p_i, and σ.

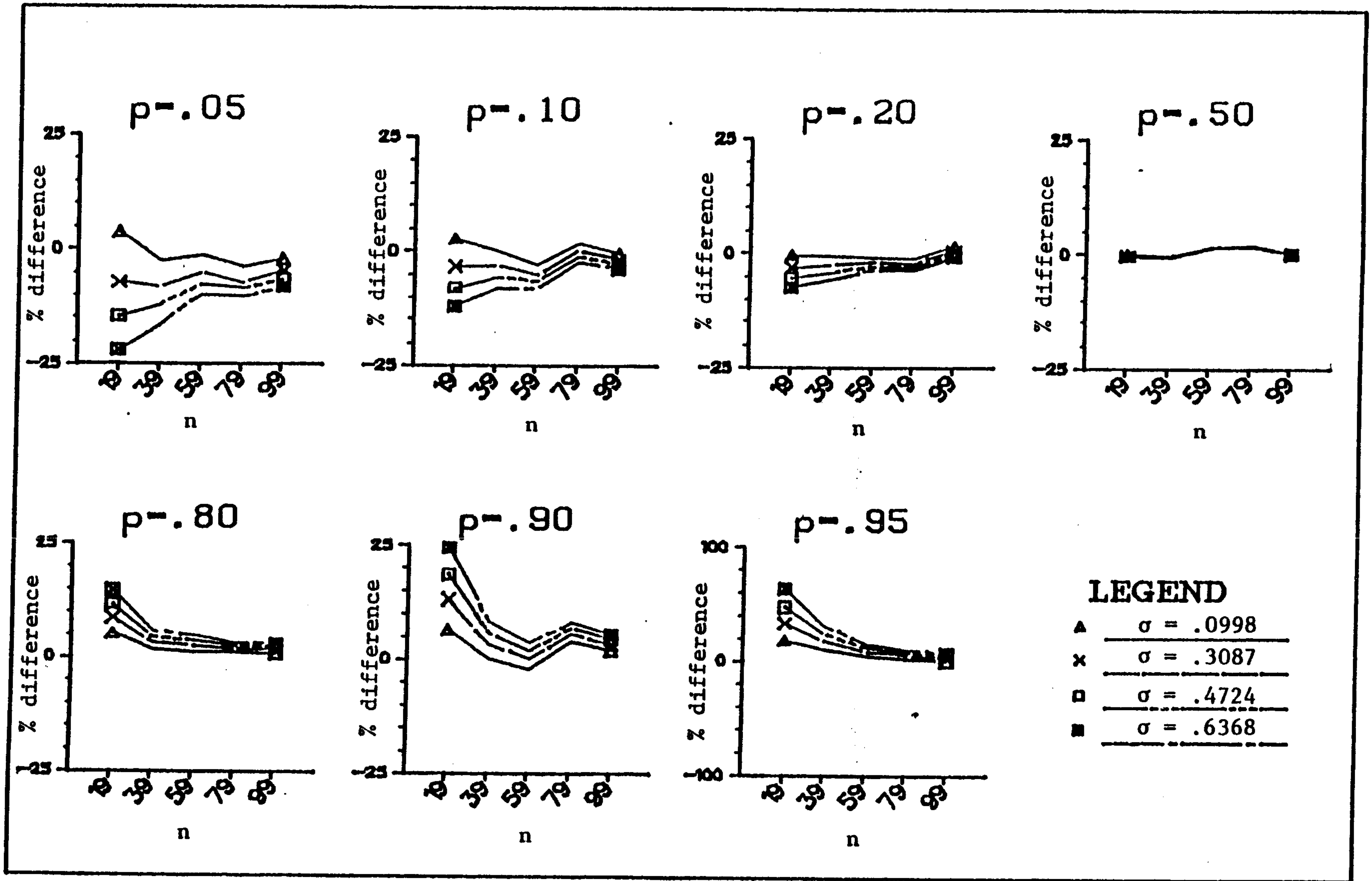

Figure D.2. Percentage differences between empirical and limiting values of Var $[n^{\frac{1}{2}}\exp(\sigma Z_{pi})]$ for varying values of n, p_i, and σ.