AN INVESTIGATION OF L-MOMENTS AND THE GENERALIZED LOGISTIC DISTRIBUTION: APPLIED AS A NEW WAY TO MODEL ICE STRENGTH

LINDA BARTHOLOMEW

An Investigation of L-Moments and the Generalized Logistic Distribution: Applied as a New Way to Model Ice Strength

by

Linda Bartholomew

A practicum report submitted to the
School of Graduate Studies in partial fulfilment
of the requirements for the degree of
Master of Applied Statistics

Department of Mathematics and Statistics
Faculty of Science
Memorial University of Newfoundland
May, 1995

St.John's Newfoundland Canada

ISBN 0-612-06106-X

Abstract

In cold ocean research the development of statistical techniques useful in the analysis of cold ocean data such as ice strength is of important practical concern. A central interest is the identification of and fitting of suitable models to the data, for the analysis of such data. In this practicum we study the theory of L-Moments as a method of distributional identification and parameter estimation. In particular, the Generalized Logistic Distribution (GLD) is fitted to nine data sets consisting of breaking strength measurements of different types of ice using the method of L-Moments. The results compare favorably to the original analysis of the data based on Maximum Likelihood fitting of the Weibull distribution. The asymptotic distribution of the L-Moment estimators is derived, and a test for the symmetry of the GLD, based on these asymptotic results, is developed. A Monte Carlo simulation study demonstrates the performance of the method of L-Moments for the estimation of the parameters of the GLD and compares it to the method Maximum Likelihood and the method of Moments. L-Moment estimators are easy to compute and perform consistently well across a wide range of parameter values. The method was found to be a simple and reliable method for estimation and distributional identification and thus it provides an attractive alternative method to the standard techniques. The application of this method to real data illustrates the implementation of the method and the contexts in which the method is useful.

Acknowledgements

I wish to thank my supervisor, Dr. Roy Bartlett (Faculty of Mathematics and Statistics, Memorial University of Newfoundland) for helpful comments and suggestions about my practicum. I also wish to thank Dr. Leonard Lye (Faculty of Engineering and Applied Science, Memorial University of Newfoundland) for the idea and the background material for this research. I acknowledge NSERC, Government of Canada, for financial support.

Contents

List of Tables

Chapter 1

Theory of L-Moments

1.1 Introduction

Parameter estimation and model selection for data analysis are two important areas in statistics. There are many methods suggested in the literature for parameter estimation, including the method of Maximum Likelihood and the Method of Moments. The technique used for a given problem is selected based on ease of implementation and reliability. To choose a probability distribution to best describe a set of data, researchers rely on experience, assumptions, and a variety of descriptive statistics and statistically based goodness-of-fit tests. The method of **L-Moments**, developed by J.R.M. Hosking in a technical report [8] as an alternative method of estimation and distribution identification, has become prevalent in hydrology and engineering journals (see [5],[10],[13],[16],[25],[32] for examples), but has been overlooked in the statistical literature. The development and advancement of the theory by hydrologists and engineers is due to its simplicity and effectiveness in many hydrological applications. Although the theory of L-Moments has been developed extensively for some selected statistical distributions, many aspects and distributions have yet to be studied in detail.

The method of conventional moments is frequently used for summarizing and describing a data set, as well as for parameter estimation. For the purpose of estimation, sample moments are equated to population moments and the resulting equations are solved for estimates of the parameters of interest. The method of L-Moments considered in this report follows a similar derivation, the L-Moments are derived and equated to their sample analogues. L-Moments are special linear combinations of order statistics, giving the sample L-Moments the immediate advantage of a level of robustness not found with conventional sample moments, the calculation of which involves raising observations to powers of two or more. The performance of L-Moments for estimating parameters of some selected distrihutions has been examined in the literature, primarily in hydrological and engineering journals and most notably in Hosking's technical report [8]. Hosking states in this report that, in his experience, L-Moments often yield more accurate parameter estimates than the method of Maximum Likelihood (MLE) and the Method of Moments (MOM), especially in small samples.

L-estimators, estimators based on linear combinations of order statistics, are not new to statisticians [4]; however, the specific linear combinations that define L-Moments had not been considered as part of a unified theory of distribution description, parameter estimation and hypothesis testing until Hosking's technical report [8], and later in Hosking [9].

1.2 Definitions and Results

To understand the theory of L-Moments, we first consider quantities called Probability Weighted Moments (PWMs) proposed by Greenwood et. al. [5]. It was the results of this paper that led to the theory of L-Moments developed by Hosking [8]. L-Moments are special linear combination of PWMs, which are themselves linear combinations of expected values of order statistics.

1.2.1 Probability Weighted Moments

Definition: Let X be a real valued random variable with distribution function $F(x)$. The probability weighted moments of X are:

$$M_{p,r,s} = E[X^p\{F(X)\}^r\{1-F(X)\}^s] = \int x^p\{F(x)\}^r\{1-F(x)\}^s dF(x) \quad (1.1)$$

where p, r and s are real numbers. If r and s are positive integers, then:

$$M_{p,r,s} = \frac{r!s!}{(r+s+1)!}E(X^p_{r+1:r+s+1})$$

where $X_{k:n}$ is defined as the k^{th} order statistic of a sample of size n. This definition holds for both discrete and continuous random variables. In the continuous case:

$$M_{p,r,s} = \int x^p\{F(x)\}^r\{1-F(x)\}^s f(x)dx$$

Then, by the transformation $F = F(x)$

$$M_{p,r,s} = \int_0^1 \{x(F)\}^p F^r(1-F)^s dF$$

where $x(F)$ is the *inverse distribution function.*

Note that when $r = s = 0$, the PWMs reduce to ordinary conventional moments. In the paper of Greenwood et. al. [5], two particular PWMs were defined as follows:

$$\alpha_r = M_{1,0,r} = E[X\{1 - F(X)\}^r],\ r = 0, 1, 2.... \tag{1.2}$$

and

$$\beta_r = M_{1,r,0} = E[X\{F(X)\}^r],\ r = 0, 1, 2... \tag{1.3}$$

Note that $r\alpha_{r-1} = E(X_{1:r})$ and $r\beta_{r-1} = E(X_{r:r})$ are the expected values of the smallest and largest order statistics respectively in a sample of size r.

Hosking [8] states and proves in his report that these PWMs are sufficient to characterize a probability distribution. Further, characterizations by either α or β PWMs are interchangeable due to the following relationship between the two:

$$\alpha_r = \sum_{k=0}^{r}(-1)^k \binom{r}{k} \beta_k \tag{1.4}$$

or equivalently

$$\beta_r = \sum_{k=0}^{r}(-1)^k \binom{r}{k} \alpha_k \tag{1.5}$$

The first relationship is easy to derive by writing $(1 - F(X))^r$ in powers of $F(X)$, as follows:

$$\alpha_r = E(X\{1 - F(X)\}^r) = \sum_{k=0}^{r} \binom{r}{k} (-1)^k E(X\{F(X)\}^k)$$

using the binomial theorem. The expectation in the expression above is simply the definition of β_k. Thus:

$$\alpha_r = \sum_{k=0}^{r} \binom{r}{k} (-1)^k \beta_k$$

as required. Similarly, to show the second of the two relationships, expand $\{F(X)\}^r$ in powers of $\{1 - F(X)\}$ again using the binomial theorem.

$$\beta_r = \sum_{k=0}^{r} \binom{r}{k} (-1)^k E(X\{1 - F(X)\}^k)$$

The expectation in the expression above is the definition of α_k.

$$\beta_r = \sum_{k=0}^{r} \binom{r}{k} (-1)^k \alpha_k$$

as required. The first few such equalities are:

$$\alpha_0 = \beta_0$$

$$\alpha_1 = \beta_0 - \beta_1$$

$$\alpha_2 = \beta_0 - 2\beta_1 + \beta_2$$

$$\alpha_3 = \beta_0 - 3\beta_1 + 3\beta_2 - \beta_3$$

for the first of the two relationships and:

$$\beta_0 = \alpha_0$$

$$\beta_1 = \alpha_0 - \alpha_1$$

$$\beta_2 = \alpha_0 - 2\alpha_1 + \alpha_2$$

$$\beta_3 = \alpha_0 - 3\alpha_1 + 3\alpha_2 - \alpha_3$$

for the second of the two relationships.

1.2.2 Examples of $r\alpha_{r-1}$ and $r\beta_{r-1}$ Calculation for Some Common Distributions

Here, three examples of PWM calculation are given for illustration. $r\alpha_{r-1}$ and $r\beta_{r-1}$ are derived instead of simply α_r and β_r, since the former have a more straightforward meaning; they are the expectations of the smallest and largest order statistics respectively in a sample of size r. α_r and β_r can be found easily from these quantities.

i. **UNIFORM (a,b)** The p.d.f. is given by:

$$f(x) = \frac{1}{b-a}\ , a < x < b\ ,\ -\infty < a, b < \infty$$

The C.D.F. is given by:

$$F(x) = \frac{x-a}{b-a}$$

Then:

$$r\alpha_{r-1} = E(X_{1:r}) = \int_a^b rx(1-F(x))^{r-1}f(x)dx$$

$$= \int_a^b rx\left(\frac{b-x}{b-a}\right)^{r-1}\frac{1}{b-a}dx = \frac{r}{(b-a)^r}\int_a^b x(b-x)^{r-1}dx$$

Using integration by parts gives:

$$= \frac{r}{(b-a)^r}\left[\frac{a(b-a)^r}{r} + \frac{(b-a)^{r+1}}{r(r+1)}\right] = \frac{ra+b}{r+1}$$

Similarly, $r\beta_{r-1}$ is given by:

$$r\beta_{r-1} = E(X_{r:r}) = \int_a^b rx(F(x))^{r-1}f(x)dx$$

$$= \int_a^b rx \left(\frac{x-a}{b-a}\right)^{r-1} \frac{1}{b-a} dx = \frac{a+rb}{r+1}$$

ii. **EXPONENTIAL(θ)** The p.d.f. is given by:

$$f(x) = \frac{1}{\theta} e^{-x/\theta} \; , \; 0 < x < \infty \; , \; \theta > 0$$

The C.D.F. is given by:

$$F(x) = 1 - e^{-x/\theta}$$

Thus:

$$r\alpha_{r-1} = E(X_{1:r}) = \int_0^\infty rx(1 - F(x))^{r-1} f(x) dx$$

$$= \int_0^\infty rxe^{-(r-1)x/\theta} \frac{1}{\theta} e^{-x/\theta} dx = \frac{r}{\theta} \int_0^\infty xe^{-rx/\theta} dx = \frac{\theta}{r}$$

using integration by parts. For this example, calculation of $r\beta_{r-1}$ is not straightforward. It requires a knowledge of some specialized functions (see [8]). The exponential distribution is an example of a distribution where the calculation of one type of PWM may be easier than the other. Since a distribution can be characterized by either type of PWM, and because α and β PWMs are related by the relationship given previously in this section, knowledge of both is not necessary. The next example illustrates a case where the β PWMs are easily derived while the α PWMs are not.

iii. **GENERALIZED EXTREME VALUE {GEV(ε, α, k)}**

The p.d.f. of this distribution is given by:

$$f(x) = \frac{1}{\alpha}\left(1 - \frac{k}{\alpha}(x-\varepsilon)\right)^{1/k-1} \exp\left[-\left(1 - \frac{k}{\alpha}(x-\varepsilon)\right)^{1/k}\right]$$

$$\varepsilon + \alpha/k \le x < \infty \qquad \text{if } k < 0$$

$$-\infty < x \le \varepsilon + \alpha/k \qquad \text{if } k > 0$$

$$-\infty < \varepsilon < \infty \quad , \quad 0 < \alpha < \infty$$

The C.D.F. is given by:

$$F(x) = \exp\left[-\{1 - k(x-\varepsilon)/\alpha\}^{1/k}\right]$$

The ***inverse distribution function*** is given by:

$$x(F) = \varepsilon + \frac{\alpha}{k}\left(1 - (-\ln F)^k\right)$$

The inverse distribution function is considered here because making the transformation $F = F(x)$ often makes computation of expectations of the extreme order statistics easier, as will be seen in the following calculations. As mentioned above, in this example, $r\beta_{r-1}$ may be found with little difficulty, while $r\alpha_{r-1}$ may not.

$$r\beta_{r-1} = E(X_{r:r}) = \int_x rx\{F(x)\}^{r-1} f(x)dx$$

Make the transformation $F = F(x)$. Then $dF = f(x)dx$.

$$r\beta_{r-1} = r\int_0^1 x(F)F^{r-1}dF = r\int_0^1 \left[\varepsilon + \frac{\alpha}{k}(1 - (-\ln F)^k\right] F^{r-1}dF$$

$$= \left(\varepsilon + \frac{\alpha}{k}\right) - r\frac{\alpha}{k}\int_0^1 (-\ln F)^k F^{r-1} dF$$

Let $u = -\ln F$. Then $F = e^{-u}$ and $dF = -e^{-u}du$. Then

$$r\beta_{r-1} = \left(\varepsilon + \frac{\alpha}{k}\right) - r\frac{\alpha}{k}\int_0^\infty u^k e^{-ru} du$$

Note that the expression under the integral is a gamma distribution except for a multiplying constant. Thus one can easily evaluate this integral.

$$r\beta_{r-1} = \left(\varepsilon + \frac{\alpha}{k}\right) - r\frac{\alpha}{k} r^{-(k+1)}\Gamma(k+1)$$

$$= \varepsilon + \frac{\alpha}{k}\left(1 - r^{-k}\Gamma(k+1)\right), \qquad k > -1$$

The properties of the L-Moment estimators of the parameters of the GEV have been extensively studied in Hosking et. al.[14].

These examples and many others are described in detail in [8].

1.3 L-Moments

While PWMs characterize a distribution, they have no easily interpretable descriptive meaning. Hosking [8], therefore, proposed functions of PWMs that give a descriptive summary of the location, scale, skewness and kurtosis of a probability distribution.

Definition: Given a real valued random variable X, the L-Moments of X are defined to be the quantities:

$$\lambda_r = r^{-1} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} E(X_{r-k:r}) \qquad r = 1, 2..., n \tag{1.6}$$

These L-Moments can be written in terms of the previously defined α and β PWMs as follows:

$$\lambda_{r+1} = (-1)^r \sum_{k=0}^{r} p_{r,k}\, \alpha_k = \sum_{k=0}^{r} p_{r,k}\, \beta_k \ , r = 0, 1, 2..., n-1 \tag{1.7}$$

where

$$p_{r,k} = (-1)^{r-k} \binom{r}{k} \binom{r+k}{k}$$

Using this relationship, we can easily write the first few L-Moments in terms of the α and the β PWMs.

$$\lambda_1 = \alpha_0 = \beta_0$$

$$\lambda_2 = \alpha_0 - 2\alpha_1 = 2\beta_1 - \beta_0$$

$$\lambda_3 = \alpha_0 - 6\alpha_1 + 6\alpha_2 = 6\beta_2 - 6\beta_1 + \beta_0$$

$$\lambda_4 = \alpha_0 - 12\alpha_1 + 30\alpha_2 - 20\alpha_3 = 20\beta_3 - 30\beta_2 + 12\beta_1 - \beta_0$$

Hosking notes that λ_r exists for $r = 1, 2, ..., n$ iff $E\,|X|$ exists. Thus a distribution may be fully described by L-Moments even if some of its conventional moments do not exist. Furthermore, Hosking asserts that such a description is meaningful because a distribution whose *mean exists* is characterized by its L-Moments. As with conventional moments, it is convenient to standardize some of the higher order L-Moments.

Hosking introduced the L-Moment ratios defined to be the quantities:

$$\tau_r = \lambda_r / \lambda_2 \qquad r = 3, 4..., n$$

Hosking states and proves the following assertion that gives numerical bounds for these L-Moment ratios:

- If X is a non-degenerate random variable whose mean exists, then the L-Moment ratios of X satisfy $|\tau_r| < 1$.

1.3.1 Describing a Probability Distribution with L-Moments

The L-Moments and L-Moment ratios are meaningful quantities for describing the features of a distribution. In fact, they are in some ways analogous to conventional moments. λ_1 is simply the mean, a standard measure of location. Further, Hosking proposes that λ_2 is a measure of scale or variation, and τ_3 and τ_4 are measures of skewness and kurtosis respectively. These measures are called L-location, L-scale, L-skewness and L-kurtosis. In his report, Hosking first presents an intuitive discussion to verify that these measures are sensible and also uses some known definitions of "scale", "skewness" and "kurtosis" to provide a theoretical basis. (Note that the quantity λ_2 has been seen in the literature before. Aside from a scalar multiple, it is the expectation of Gini's Mean Difference Statistic [9].)

Distribution	L-Skewness	Skewness	L-Kurtosis	Kurtosis
Normal(0,1)	0.0	0.0	.12	3.0
Logistic(0,1)	0.0	0.0	.17	4.2
Exponential(0,1)	.33	2.0	.17	9.0
Gumbel(0,1)	.17	1.1	.15	5.4
Rayleigh(0,1)	.11	.63	.11	2.0

Table 1.1: Comparison of L-Skewness/Skewness and L-Kurtosis/Kurtosis for some specific distributions

L-skewness and L-kurtosis have some immediately obvious advantages over conventional skewness and kurtosis.

- L-skewness and L-kurtosis are not as sensitive to the extreme tails of the distribution. Calculation of conventional skewness and kurtosis from a sample involves raising data values to the third and fourth powers, making them extremely sensitive to outliers. In small samples, the conventional sample skewness and kurtosis may be quite different from the true underlying values due to this sensitivity.

- L-skewness and L-kurtosis exist on bounded intervals making comparisons between skewness and kurtosis of different data sets easier. Some examples for some simple distributions are given in Table 1.1. Hosking has compared Moments and L-Moments for measuring distributional shape in [11].

Because of this boundedness, it is possible to identify a distribution by calculating the L-skewness and L-kurtosis of the data set and plotting the point on an **L-Moment diagram** . Hosking has created such a diagram which includes L-skewness/ L-kurtosis curves for many distributions common in hydrological applications (see Appendix F). In the data analysis chapter of this practicum, the diagram is used to determine whether a given set of data appears to be best described by the distribution considered as a model in this report.

1.4 Estimation of PWMs and L-Moments

In order to formulate a method of parameter estimation based on L-Moments or PWMs, one must first consider how to estimate the population L-Moments and PWMs, which can then be used to derive parameter estimators. Hosking proposes the use of **U-Statistics**, which were introduced by Hoeffding [7]. Since the PWMs are linear combinations of expected values of order statistics of a sample of size $r+1$, it is natural to estimate them by U-Statistics. The appropriate U-Statistics are the corresponding functions of the sample order statistics averaged over all subsamples of size $r+1$ that can be constructed from a sample of size n. U-Statistics have good properties such as high efficiency, asymptotic normality and robustness, which make them good candidates for the estimation of PWMs. The unbiased U-Statistic estimators of the PWMs (called sample PWMs) are:

$$\hat{\alpha}_r = a_r = n^{-1} \sum_{i=1}^{n} \frac{\binom{n-i}{r}}{\binom{n-1}{r}} x_i, \qquad r = 0, 1, \ldots n-1 \qquad (1.8)$$

$$\hat{\beta}_r = b_r = n^{-1} \sum_{i=1}^{n} \frac{\binom{i-1}{r}}{\binom{n-1}{r}} x_i, \qquad r = 0, 1, \ldots n-1 \tag{1.9}$$

where x_i is the i^{th} sample order statistic in a sample of size n. The a_r and the b_r are related in the same way as their population analogues:

$$a_r = \sum_{k=0}^{r} (-1)^k \binom{r}{k} b_k$$

and

$$b_r = \sum_{k=0}^{r} (-1)^k \binom{r}{k} a_k$$

Greenwood et al. [5] have given a detailed proof of the unbiasedness of a_r and b_r. Because the L-Moments are linear combinations of either the α or β PWMs, we can construct estimators of the L-Moments which are the corresponding linear combinations of the a_r or b_r. The sample L-Moments are thus given by:

$$l_{r+1} = (-1)^r \sum_{k=0}^{r} p_{r,k}\, a_k = \sum_{k=0}^{r} p_{r,k}\, b_k \tag{1.10}$$

where l_r is an unbiased estimator of λ_r. The L-Moment ratios are estimated by the sample L-Moment ratios. Although these ratio statistics are not necessarily unbiased for the population ratios, Hosking [8] states that they are *consistent* estimators.

Remark: Sample L-Moments can be used to summarize the features of a data set in the same way as conventional moments, often in a simpler manner. L-Moments are frequently preferable, because, being linear combinations of the data points, they tend to be less sensitive to variability and outliers in the data. They may, therefore, prove to be more accurate and robust.

1.4.1 Asymptotic Results

The derivation of the asymptotic distributions of vectors of sample PWMs and sample L-Moments is given in Hosking [8]. The results are based on Stigler's form of a theorem giving the asymptotic distribution of order statistics [35].

Let X be a real valued random variable with C.D.F. $F(x)$, with population PWMs and L-Moments as defined previously. a_r, b_r and l_{r+1} , $r = 0, 1, \ldots n-1$ are the sample PWMs and sample L-Moments respectively, calculated from a random sample of size n. $\underline{a}_r, \underline{\alpha}_r, \underline{b}_r, \underline{\beta}_r, \underline{l}_{r+1}$,and $\underline{\lambda}_{r+1}$ are defined as vectors of length $r+1$ of sample/population PWMs or L-Moments. For example, $\underline{a}_2 = (a_0, a_1, a_2)$ and $\underline{\lambda}_3 = (\lambda_1, \lambda_2, \lambda_3)$. (Thus, the following results may be used to find the asymptotic distribution of whatever subset of the sample quantities is desired). The basic results are:

1. $\sqrt{n}(\underline{a}_r - \underline{\alpha}_r) \rightarrow N(0, A)$, where the $(i,j)^{th}$ element of the matrix A is $(i, j = 0, 1, \ldots r)$:

$$A_{ij} = I_{ij} + I_{ji}$$

   where

$$I_{ij} = \iint_{x<y} \{1 - F(x)\}^i \{1 - F(y)\}^j F(x)\{1 - F(y)\} dx dy$$

2. $\sqrt{n}(\underline{b}_r - \underline{\beta}_r) \rightarrow N(0, B)$, where the $(i,j)^{th}$ element of the matrix B is $(i, j = 0, 1, \ldots r)$:

$$B_{ij} = J_{ij} + J_{ji}$$

where

$$J_{ij} = \iint_{x<y} \{F(x)\}^i \{F(y)\}^j F(x)\{1 - F(y)\} dx dy$$

3. $\sqrt{n}(\underline{l_s} - \underline{\lambda_s}) \rightarrow N(0, \Lambda)$, where the $(i, j)^{th}$ element of the matrix Λ is $(i, j = 1, 2, ...s)$:

$$\Lambda_{ij} = \iint_{x<y} \{P^*_{i-1}(F(x))P^*_{j-1}(F(y)) + P^*_{j-1}(F(x))P^*_{i-1}(F(y))\} F(x)\{1-F(y)\} dx dy$$

where $P^*_i(x)$ is the i^{th} shifted Legendre Polynomial as defined in Appendix B.

1.5 Parameter Estimation By L-Moments

Estimation of distribution parameters is very important in statistical data analysis. The goal is to accurately estimate the parameters of the underlying distribution using a random sample from that distribution. Many techniques are available in the literature for parameter estimation as mentioned previously, and in this section we introduce a method based on the PWMs and L-Moments of the previous sections.

The method is the same as that for conventional moments. If p parameters are to be estimated, equate the first p population PWMs (L-Moments) to the first p sample PWMs (L-Moments), then solve the resulting p equations to derive estimators of the parameters. The decision to use PWMs or L-Moments depends on the distribution: for some distributions, the PWMs may provide neater solutions or more tedious ones than the L-Moments. The decision should be based on ease of computation.

Hosking has also suggested two refinements to the technique which, if judged necessary, may improve the estimators. The first concerns distributions with an end point that is a function of the parameters; it is sometimes efficient to estimate the end point by the appropriate extreme order statistic. Using the PWM approach, this would involve equating $x_{1:n} = na_{n-1} = n\alpha_{n-1}$ or $x_{n:n} = nb_{n-1} = n\beta_{n-1}$ and using one or both of these equations in combination with equations obtained from lower order PWMs. The second refinement involves making use of more than p of the PWMs (L-Moments). This can be a means to achieving a smaller estimator variance. In the next chapter we shall see an example of this technique.

As an illustration, consider estimating the mean of a symmetric distribution. For this type of distribution, the L-skewness $\lambda_3 = 0$ and consequently the expected value of the third sample L-Moment, l_3, is zero. Instead of estimating the mean by just l_1 , we could estimate it using $l_1 + al_3$, where a is chosen to minimize the variance of the estimator.

Exact distributions of the estimators are usually difficult to find, but their asymptotic distributions follow from the results of Hosking [8]. For most standard distributions the multivariate δ-method [3] can be used to show L-Moment based estimators are asymptotically normally distributed.

To justify the use of these types of estimators, they must perform comparably to established methods of estimation. Maximum Likelihood estimators (MLEs) are consistent and asymptotically efficient, making them commonly used estimators among statisticians and researchers. However, MLEs are sometimes difficult to compute, requiring recourse to numerical methods to solve complex systems of non-linear equations. L-Moment estimators are often more tractable. For many distributions they produce simple closed form expressions for the estimators, or estimators that are solutions of simple equations. L-Moment estimators are likely to be preferable if a distribution function (C.D.F.) can be expressed in inverse form: that is, there is a closed form expression for $F(x)$ that can be inverted. The reason for this is obvious from the form of the equations for the population L-Moments: expectations of order statistics are required. However, the method of L-Moments is by no means restricted to this case, rather it is more likely to produce tractable estimates if this is true. (Finding PWMs when $F(x)$ does not have a closed form is considered for some specific distributions in [19]). It is clear why these estimators are used by hydrologists and engineers. Many of the distributions used in their research are of the family of distributions called survival or reliability type distributions. Examples of these types of distributions include the Weibull, Generalized Extreme Value, Lognormal, Exponential, Gumbel and Pareto. Many such survival distributions have C.D.F.s which can be written in *inverse* form, making the method of L-Moment estimation especially appealing in such applications.. Hosking states that through his experience with some common distributions, L-Moment estimators have been shown to give reasonably efficient estimates, and, with small samples, often providing more reliable estimates then the method of Maximum Likelihood. Although maximum likelihood estimation is considered to be asymptotically best, its good asymptotic properties may not be evident with small data sets.

It is also of interest, besides comparing L-Moment based estimators to MLEs, to consider a comparison between the Method of L-Moments and the Metbod of Moments (MOM), since the two methods are clearly very similar.

As mentioned previously, we might expect the L-Moment estimators to be more robust, especially in small samples, since they are linear combinations of order statistics and require no calculation of powers of the data values as do conventional moments.

1.6 Hypothesis Testing

L-Moment based estimators, for most standard distributions, are asymptotically normally distributed. This allows for construction of simple test statistics for parametric hypotheses based on the asymptotic results. For example, a test of:

$$H_o : \theta = \theta_0 \text{ vs } H_a : \theta \neq \theta_0$$

for a parameter θ may be based on the test statistic:

$$\frac{(\hat{\theta} - \theta_0)}{\sigma(\hat{\theta})}$$

where $\hat{\theta}$ is the L-Moment estimator of θ and the denominator of the test statistic is the square root of the asymptotic variance of $\hat{\theta}$. Under some general conditions, this statistic will have a limiting N(0,1) distribution under H_o. An example of such a test is given in Chapter 6.

1.7 Discussion

The theory of L-Moments sets the foundation for the study of a viable alternative to conventional methods of estimation, especially with respect to applications requiring the modeling of survival or reliability type data found in the hydrological and engineering sciences. L-Moment estimators are often simple to calculate, while retaining a sufficient degree of accuracy and reliability for some distributions. In the next chapter, a distribution called the Generalized Logistic Distribution will be introduced as an alternative model for engineering applications. In Chapter 3, the L-Moment, conventional Moment and Maximum Likelihood estimators will be derived for this specific distribution. In Chapter 4, the three methods will be compared using a simulation study, and in Chapter 5 the three methods will be used to fit the distribution to some specific engineering data sets. Finally, in Chapter 6, a test of symmetry for the Generalized Logistic Distribution will be derived based on the results given in this chapter.

Chapter 2

THE GENERALIZED LOGISTIC DISTRIBUTION

2.1 Introduction

The Generalized Logistic Distribution (GLD) studied in this practicum was introduced in Hosking's technical report [8]. Although there have been other distributions discussed in the literature by this name, so far there has been no other mention of this version of the GLD. The GLD is considered in this practicum for two reasons. First, although Hosking has given the L-Moments and associated parameter estimators for this distribution, the properties of these estimators have not been investigated. Secondly, due to its similarity in shape to the GEV (Generalized Extreme Value) distribution, currently one of the most useful distributions for describing extreme phenomena, it offers an alternative choice for hydrological applications, as will be seen when it is used for the analysis of the data considered in Chapter 5. The fact that it has three parameters is also of importance to researchers, as it provides a wide variety of possible shapes for a diverse number of data sets.

2.2 Probability Distribution Function

The **probability distribution function** (p.d.f.) of the GLD is given as follows:

$$f(x) = \frac{1}{\alpha}\frac{(1-\frac{k(x-\varepsilon)}{\alpha})^{1/k-1}}{(1+(1-\frac{k(x-\varepsilon)}{\alpha})^{1/k})^2}, \quad k \neq 0 \qquad (2.1)$$

$$f(x) = \frac{1}{\alpha}\frac{e^{-\frac{(x-\varepsilon)}{\alpha}}}{(1+e^{-\frac{(x-\varepsilon)}{\alpha}})^2}, \qquad k = 0 \qquad (2.2)$$

where

$$\varepsilon + \frac{\alpha}{k} \leq x \leq \infty, \qquad \text{if } k < 0$$

$$-\infty < x < \infty, \qquad \text{if } k = 0$$

$$-\infty < x \leq \varepsilon + \frac{\alpha}{k} \qquad \text{if } k > 0$$

ε = location parameter

α = scale parameter

k = shape parameter

Note that when $k = 0$, the GLD reduces to the well known Logistic distribution. In this report the primary interest is the case where $k \neq 0$, but we consider the special case as well, for completeness. The **cumulative distribution function** (C.D.F.) is given by:

$$F(x) = \frac{1}{(1 + (1 - \frac{k(x-\varepsilon)}{\alpha})^{1/k})}, \quad k \neq 0 \tag{2.3}$$

and

$$F(x) = \frac{1}{(1 + e^{-\frac{(x-\varepsilon)}{\alpha}})}, \qquad k = 0 \tag{2.4}$$

Note that since we have an explicit form for the C.D.F., we can easily write the inverse of the C.D.F. for this distribution. The **inverse distribution function**, denoted $x(F)$, is given by:

$$x(F) = \varepsilon + \frac{\alpha}{k}\left[1 - \left(\frac{1-F}{F}\right)^k\right], \quad k \neq 0 \tag{2.5}$$

and

$$x(F) = \varepsilon - \alpha \ln\left(\frac{1-F}{F}\right), \qquad k = 0 \tag{2.6}$$

The mean and variance of the distribution, when $k \neq 0$ are:

$$\mu_x = \varepsilon + \frac{\alpha}{k}\left[1 - \Gamma(1-k)\Gamma(1+k)\right], \quad |k| < 1$$

$$\sigma_x^2 = \frac{\alpha^2}{k^2}\left[\Gamma(1-2k)\Gamma(1+2k) - \Gamma^2(1-k)\Gamma^2(1+k)\right], \quad |k| < 1/2$$

When $k = 0$, the mean and variance are given by:

$$\mu_x = \varepsilon$$

$$\sigma_x^2 = \frac{\pi^2}{3}\alpha^2$$

Because the inverse distribution function can be defined in closed form, the GLD is a good candidate for L-Moment parameter estimation. We shall see in the next chapter that the estimates derived with this method are considerably more manageable than either maximum likelihood estimates (MLEs) or Method of Moments (MOM) estimates, and can be written in closed form.

2.3 PWMs and L-Moments

The PWMs and the L-Moments for the GLD are stated below and are followed by a derivation of each. As in Chapter 1, we consider the form of $r\alpha_{r-1}$ and $r\beta_{r-1}$ as opposed to α_r and β_r :

$$r\alpha_{r-1} = \varepsilon + \frac{\alpha}{k}\left[1 - \frac{\Gamma(1-k)\Gamma(r+k)}{\Gamma(r)}\right] \ , \ \mid k \mid < 1$$

$$r\beta_{r-1} = \varepsilon + \frac{\alpha}{k}\left[1 - \frac{\Gamma(1+k)\Gamma(r-k)}{\Gamma(r)}\right] \ , \ \mid k \mid < 1$$

When $k = 0$ (simple Logistic), the PWMs are given by:

$$r\alpha_{r-1} = \varepsilon - \alpha \sum_{s=1}^{r-1} s^{-1}$$

$$r\beta_{r-1} = \varepsilon + \alpha \sum_{s=1}^{r-1} s^{-1}$$

The first four L-Moments of the GLD are:

$$\lambda_1 = \varepsilon + \frac{\alpha}{k}\left[1 - \Gamma(1-k)\Gamma(1+k)\right]$$

$$\lambda_2 = \alpha\Gamma(1-k)\Gamma(1+k)$$

$$\lambda_3 = -k\alpha\Gamma(1-k)\Gamma(1+k)$$

$$\lambda_4 = \frac{1}{6}(1+5k^2)\alpha\Gamma(1-k)\Gamma(1+k)$$

and the L-skewness and L-kurtosis are (as defined in Chapter 1):

$$\tau_3 = -k$$

$$\tau_4 = \frac{(1+5k^2)}{6}$$

When $k = 0$, the first four L-Moments are:

$$\lambda_1 = \varepsilon$$

$$\lambda_2 = \alpha$$

$$\lambda_3 = 0$$

$$\lambda_4 = \frac{\alpha}{6}$$

and the L-skewness and L-kurtosis are given by:

$$\tau_3 = 0$$

$$\tau_4 = \frac{1}{6}$$

2.3.1 Derivation of Results

In this section, the derivations of the results of the previous section are given.

$$r\alpha_{r-1} = E(X_{1:r}) = \int_x rx\{1-F(x)\}^{r-1}f(x)dx$$

Make the transformation $F = F(x)$.

$$= r\int_0^1 x(F)(1-F)^{r-1}dF$$

where $x(F)$ is the inverse function as defined in Chapter 1.

$$= r\int_0^1 \left[\varepsilon + \frac{\alpha}{k}(1 - \left(\frac{1-F}{F}\right)^k)\right](1-F)^{r-1}dF$$

$$= \left(\varepsilon + \frac{\alpha}{k}\right) - r\frac{\alpha}{k}\int_0^1 F^{-k}(1-F)^{r+k-1}dF$$

Note that the integrand is a Beta distribution, apart from a multiplying constant. Hence, it can be easily evaluated.

$$= \left(\varepsilon + \frac{\alpha}{k}\right) - r\frac{\alpha}{k}\frac{\Gamma(1-k)\Gamma(r+k)}{\Gamma(r+1)}$$

$$= \varepsilon + \frac{\alpha}{k}\left[1 - \frac{\Gamma(1-k)\Gamma(r+k)}{\Gamma(r)}\right]$$

as required. The derivation of the β PWMs is similar.

$$r\beta_{r-1} = E(X_{r:r}) = \int_x rx\{F(x)\}^{r-1}f(x)dx$$

$$= r\int_0^1 x(F)F^{r-1}dF$$

$$= r\int_0^1 \left[\varepsilon + \frac{\alpha}{k}(1 - \left(\frac{1-F}{F}\right)^k)\right]F^{r-1}dF$$

$$= \left(\varepsilon + \frac{\alpha}{k}\right) - r\frac{\alpha}{k}\int_0^1 (1-F)^k F^{r-k-1}dF$$

$$= \varepsilon + \frac{\alpha}{k}\left[1 - \frac{\Gamma(k+1)\Gamma(r-k)}{\Gamma(r)}\right]$$

as required. The first four L-Moments of the GLD are found using the relationship between them and either the α or β PWMs. In this case, the β PWMs will be used.

$$\lambda_1 = \beta_0 = \varepsilon + \frac{\alpha}{k}\left(1 - \Gamma(1-k)\Gamma(1+k)\right)$$

$$\begin{aligned}
\lambda_2 &= 2\beta_1 - \beta_0 \\
&= \varepsilon + \frac{\alpha}{k}\left(1 - \Gamma(1+k)\Gamma(2-k)\right) - \varepsilon - \frac{\alpha}{k}\left(1 - \Gamma(1-k)\Gamma(1+k)\right) \\
&= \frac{\alpha}{k}\left(\Gamma(1+k)\Gamma(1-k) - \Gamma(1+k)\Gamma(2-k)\right) \\
&= \frac{\alpha}{k}\Gamma(1+k)\left(\Gamma(1-k) - (1-k)\Gamma(1-k)\right) \\
&= \alpha\Gamma(1+k)\Gamma(1-k)
\end{aligned}$$

$$\begin{aligned}
\lambda_3 &= 6\beta_2 - 6\beta_1 + \beta_0 \\
&= 2\left[\varepsilon + \frac{\alpha}{k}\left(1 - \frac{\Gamma(1+k)\Gamma(3-k)}{2}\right)\right] - 3\left[\varepsilon + \frac{\alpha}{k}\left(1 - \Gamma(1+k)\Gamma(2-k)\right)\right] \\
&\qquad + \left[\varepsilon + \frac{\alpha}{k}\left(1 - \Gamma(1-k)\Gamma(1+k)\right)\right]
\end{aligned}$$

$$\begin{aligned}
&= \frac{\alpha}{k}\Gamma(1+k)\Gamma(1-k)\left(3(1-k) - (2-k)(1-k) - 1\right) \\
&= -\alpha k\Gamma(1-k)\Gamma(1+k)
\end{aligned}$$

$$\lambda_4 = 20\beta_3 - 30\beta_2 + 12\beta_1 - \beta_0$$
$$= 5\left[\varepsilon + \frac{\alpha}{k}\left(1 - \frac{\Gamma(1+k)\Gamma(4-k)}{6}\right)\right] - 10\left[\varepsilon + \frac{\alpha}{k}\left(1 - \frac{\Gamma(1+k)\Gamma(3-k)}{2}\right)\right]$$
$$+6\left[\varepsilon + \frac{\alpha}{k}(1 - \Gamma(1+k)\Gamma(2-k))\right] - \left[\varepsilon + \frac{\alpha}{k}(1 - \Gamma(1-k)\Gamma(1+k))\right]$$
$$= \frac{\alpha}{k}\Gamma(1-k)\Gamma(1+k)\frac{1}{6}(5k^3 + k)$$
$$= \frac{\alpha}{6}(1 + 5k^2)\Gamma(1-k)\Gamma(1+k)$$

$$\tau_3 = \frac{\lambda_3}{\lambda_2} = -k$$
$$\tau_4 = \frac{\lambda_4}{\lambda_2} = \frac{(1+5k^2)}{6}$$

For $k = 0$, the simple Logistic Distribution, the α and β PWMs are found by taking the limit of the PWMs for the GLD as $k \to 0$. Hence for this special case;

$$r\alpha_{r-1} = \lim_{k\to 0}\left(\varepsilon + \frac{\alpha}{k}\left[1 - \frac{\Gamma(1-k)\Gamma(r+k)}{\Gamma(r)}\right]\right)$$
$$= \varepsilon + \frac{\alpha}{\Gamma(r)}\lim_{k\to 0}\left(\frac{\Gamma(r) - \Gamma(1-k)\Gamma(r+k)}{k}\right)$$

Using L'Hôpital's Rule gives:

$$= \varepsilon + \frac{\alpha}{\Gamma(r)}\lim_{k\to 0}(-\Gamma(1-k)\Gamma(r+k)\Psi(r+k) + \Gamma(r+k)\Gamma(1-k)\Psi(1-k))$$
$$= \varepsilon + \frac{\alpha}{\Gamma(r)}(\Gamma(r)\Psi(1) - \Gamma(r)\Psi(r))$$

$$= \varepsilon - \alpha\left(\Psi(r) - \Psi(1)\right) = \varepsilon - \alpha \sum_{s=1}^{r-1} s^{-1}$$

Similarly,

$$r\beta_{r-1} = \varepsilon + \frac{\alpha}{\Gamma(r)} \lim_{k\to 0} \left(\frac{\Gamma(r) - \Gamma(1+k)\Gamma(r-k)}{k} \right)$$

$$= \varepsilon + \frac{\alpha}{k} \lim_{k\to 0} \left(\Gamma(1+k)\Gamma(r-k)\Psi(r-k) - \Gamma(r-k)\Gamma(1+k)\Psi(1+k)\right)$$

$$= \varepsilon + \alpha\left(\Psi(r) - \Psi(1)\right) = \varepsilon + \alpha \sum_{s=1}^{r-1} s^{-1}$$

as required, where Ψ is the Digamma function (Appendix B). The first four L-Moments then follow from either the α or β PWMs.

$$\lambda_1 = \beta_0 = \varepsilon + \alpha(\Psi(1) - \Psi(1)) = \varepsilon$$

$$\lambda_2 = 2\beta_1 - \beta_0 = \varepsilon + \alpha(\Psi(2) - \Psi(1)) - \varepsilon = \alpha$$

$$\lambda_3 = 6\beta_2 - 6\beta_1 + \beta_0$$

$$= 2\alpha(\Psi(3) - \Psi(1)) - 3\alpha(\Psi(2) - \Psi(1)) = 0$$

$$\lambda_4 = 20\beta_3 - 30_2 + 12\beta_1 - \beta_0$$

$$= 5\alpha\left[1 + \frac{1}{2} + \frac{1}{3}\right] - 10\alpha\left[1 + \frac{1}{2}\right] + 6\alpha = \frac{\alpha}{6}$$

It then easily follows that $\tau_3 = 0$ and $\tau_4 = 1/6$.

In Appendix C, some graphs of the p.d.f. are given for various values of k. For negative values of k, the distribution is skewed right, and for positive values of k, it is skewed left. When $k = 0$, the distribution is symmetric.

2.4 Discussion

The Generalized Logistic Distribution appears to be a good possibility for modeling the types of data that will be discussed in this practicum. As we have seen in the previous sections, it has tractable expressions for the L-Moments. By inspection of these expressions, we can see there are simple relationships between them and the three parameters. Further, its shape is similar to that of the GEV, a widely used model for hydrological data sets. In the following chapters, there will be further justification for the use of this particular distribution. Although the GLD has not yet been exploited by statisticians, its versatility for reliability and engineering data sets makes it an ideal distribution for the analysis of such data.

Chapter 3

LMOM, MLE AND MOM ESTIMATORS

3.1 Introduction

In this chapter, the three types of estimators being considered in this report are derived for the GLD. They will be compared on the basis of mathematical tractability, and in Chapter 4 they will be compared for accuracy and precision with a simulation study. It will be seen that the L-Moment estimators have the simplest form for the purposes of computation, while the maximum likelihood estimators are solutions of a complicated set of nonlinear equations. Although one cannot choose a set of estimators based solely on simplicity of form, it is at least a consideration in combination with other required properties. Each method will be considered in turn and the resulting estimators will be presented. The special case of the simple Logistic distribution ($k = 0$) is considered separately.

3.2 L-Moments (LMOM)

Estimates are chosen to be based on L-Moments as opposed to PWMs due to the simple relationships between the population L-Moments and the three parameters. However, for some distributions, using PWMs will yield simpler estimators (see [24], [14]). Since L-Moments are simple linear combinations of PWMs, the two approaches are essentially the same. The simple relationships suggest that the GLD is a natural choice for the method of L-Moments. They are given by:

$$k = -\tau_3$$

$$\alpha = \frac{\lambda_2}{\Gamma(1-k)\Gamma(1+k)}$$

$$\varepsilon = \lambda_1 - \frac{\alpha}{k}\left(1 - \Gamma(1-k)\Gamma(1+k)\right)$$

To find estimators, substitute sample L-Moments for the population L-Moments. Hosking [8] has stated these results, but he notes that the properties of the estimators have not been investigated. He suggests some possible modifications of these estimators because the distribution has an end point that is a function of the parameters, a situation that sometimes causes estimation difficulties. However, after experimenting with these modifications, it appears that the estimators as they are given are the best. Note that the sample L-Moments are easy to calculate from the data, making these estimators appealing in terms of simplicity and tractability.

3.3 MLE

Maximum Likelihood estimates are those values of the parameters that maximize the log likelihood function. No investigation of the MLEs for the GLD is given in the literature; below we give a derivation of the MLEs in a form suitable for use in this report. Consider a random sample of size n from a GLD; $x_1, x_2, \ldots x_n$. The log likelihood for the GLD is given as follows:

$$\ln L = -n\ln\alpha + \frac{(1-k)}{k}\sum \ln\left(1-\frac{k(x_i-\varepsilon)}{\alpha}\right) - 2\sum \ln\left(1+\left(1-\frac{k(x_i-\varepsilon)}{\alpha}\right)^{1/k}\right) \tag{3.1}$$

where Σ represents summation from $i = 1$ to n.

To maximize this function for the three parameters, we need to find the partial derivatives of the function with respect to each parameter, then set the resulting three equations equal to 0 and solve the equations for the three parameters. The partial derivatives are:

$$\frac{\partial \ln L}{\partial \varepsilon} = (1-k)\sum \frac{1}{\left(1-\frac{k(x_i-\varepsilon)}{\alpha}\right)} - \frac{2}{\alpha}\sum \frac{\left(1-\frac{k(x_i-\varepsilon)}{\alpha}\right)^{1/k-1}}{\left(1+\left(1-\frac{k(x_i-\varepsilon)}{\alpha}\right)^{1/k}\right)}$$

$$\frac{\partial \ln L}{\partial \alpha} = -\frac{n}{\alpha} + \frac{(1-k)}{\alpha}\sum \frac{(x_i-\varepsilon)}{(\alpha-k(x_i-\varepsilon))} - \frac{2}{\alpha^2}\sum \frac{(x_i-\varepsilon)\left(1-\frac{k(x_i-\varepsilon)}{\alpha}\right)^{1/k-1}}{\left(1+\left(1-\frac{k(x_i-\varepsilon)}{\alpha}\right)^{1/k}\right)}$$

$$\frac{\partial \ln L}{\partial k} = \frac{(k-1)}{k}\sum \frac{(x_i-\varepsilon)}{(\alpha-k(x_i-\varepsilon))} - \frac{1}{k^2}\sum \ln\left(1-\frac{k(x_i-\varepsilon)}{\alpha}\right)$$

$$+\frac{2}{k}\sum\frac{(x_i-\varepsilon)\left(1-\frac{k(x_i-\varepsilon)}{\alpha}\right)^{1/k}}{(\alpha-k(x_i-\varepsilon))\left(1+\left(1-\frac{k(x_i-\varepsilon)}{\alpha}\right)^{1/k}\right)}$$

$$+\frac{2}{k^2}\sum\frac{\ln\left(1-\frac{k(x_i-\varepsilon)}{\alpha}\right)\left(1-\frac{k(x_i-\varepsilon)}{\alpha}\right)^{1/k}}{\left(1+\left(1-\frac{k(x_i-\varepsilon)}{\alpha}\right)^{1/k}\right)}$$

The MLEs are the values of the parameters that satisfy:

$$\frac{\partial \ln L}{\partial \varepsilon}=0$$

$$\frac{\partial \ln L}{\partial \alpha}=0$$

$$\frac{\partial \ln L}{\partial k}=0$$

Inspecting the three equations we note that we must solve three nonlinear equations for which no closed form solution exists. Consequently, the solutions will have to be determined by numerical methods. In a later chapter of this report, we show that achieving convergence to a solution is difficult and very sensitive to choice of starting values.

3.4 MOM

Method of Moments estimators are derived by finding p (number of parameters) population conventional moments and equating them to corresponding sample moments. The resulting p equations are then solved for the parameters. For the GLD we have three parameters, so we require the first three sample moments: sample mean, variance and skewness. These sample moments are equated to their population analogues, and the resulting equations are:

$$\varepsilon + \frac{\alpha(1-g_1)}{k} = \widehat{\mu}_x$$

$$\frac{\alpha^2(g_2 - g_1^2)}{k^2} = \widehat{\sigma}_x^2$$

$$(-\text{sign}(k))\frac{(g_3 - 3g_2g_1 + 2g_1^3)}{(g_2 - g_1^2)^{3/2}} = \frac{\sum(x_i - \widehat{\mu}_x)^3}{(\widehat{\sigma}_x^2)^{3/2}}$$

where

$$g_r = \Gamma(1 - rk)\Gamma(1 + rk)$$

and $\widehat{\mu_x}$ and $\widehat{\sigma_x}^2$ are the sample mean and sample variance respectively.

The MOM estimates are the solutions to these three equations. In order to find estimates, we solve the third equation for k, which is a function of k alone. Given an estimate of k, the other two parameters may be solved for explicitly. However, numerical methods are necessary to solve the third equation.

Before doing a comparison by use of a simulation study, it is clear from inspection of the equations that the L-Moment based estimators are the easiest to compute, requiring only the evaluation of the Gamma (Γ) function. This is an obvious advantage of the L-Moment estimators over both MLEs and MOM estimates. Furthermore, the MLEs appear to be the most difficult to work with, requiring an iterative computer algorithm for solving three simultaneous nonlinear equations - a difficult task especially if the function is not well behaved or has multiple maxima and/or minima in close proximity to one another.

In Chapter 4 a simulation study is done to see if the L-Moment based estimators have comparable accuracy and precision with the two other established methods.

3.5 Special Case: k=0.0, the Simple Logistic Distribution

For completeness, we also consider the properties of the three types of estimation for the simple Logistic Distribution, which is the special case of the GLD when the shape parameter, k, is zero.

3.5.1 LMOM

The L-Moment estimators are given by Hosking [8] using the relationships:

$$\lambda_1 - \lambda_3 = \varepsilon$$

$$\lambda_2 = \alpha$$

where the estimates are found by substituting sample L-Moments for population L-Moments. ε is estimated by $l_1 - l_3$ instead of l_1 alone to reduce variance. This is an example of a modification suggested in Chapter 1. Since the simple Logistic Distribution is symmetric, $E(l_3) = \lambda_3 = 0$ ([8]).

3.5.2 MLE

The log likelihood function is given by:

$$\ln L = -n\ln\alpha - \sum\frac{(x_i-\varepsilon)}{\alpha} - 2\sum\ln\left(1+e^{-\frac{(x_i-\varepsilon)}{\alpha}}\right) \tag{3.2}$$

To find the maximum likelihood estimates of ε and α we need to solve:

$$\frac{\partial\ln L}{\partial\varepsilon} = 0$$

$$\frac{\partial\ln L}{\partial\alpha} = 0$$

where

$$\frac{\partial\ln L}{\partial\varepsilon} = \frac{n}{\alpha} - \frac{2}{\alpha}\sum\frac{e^{-\frac{(x_i-\varepsilon)}{\alpha}}}{\left(1+e^{-\frac{(x_i-\varepsilon)}{\alpha}}\right)}$$

and

$$\frac{\partial\ln L}{\partial\alpha} = -\frac{n}{\alpha} + \sum\frac{(x_i-\varepsilon)}{\alpha^2} - 2\sum\frac{e^{-\frac{(x_i-\varepsilon)}{\alpha}}}{\left(1+e^{-\frac{(x_i-\varepsilon)}{\alpha}}\right)}\frac{(x_i-\varepsilon)}{\alpha^2}$$

3.5.3 MOM

To find the MOM estimators, we equate the first two conventional population moments to the sample mean and sample variance respectively,

$$\varepsilon = \widehat{\mu}_x$$

$$\alpha = \frac{\sqrt{3\hat{\sigma}^2}}{\pi}$$

.

3.6 Asymptotic Distribution of the L-Moment Estimators

It is suggested by Hosking [8] that the variances of the L-Moment-based estimators in finite samples are quite close to the asymptotic variances. Thus, it is of interest to find the asymptotic variance-covariance matrix for the estimators for different parameter values. A two-step procedure is required to determine the form of this matrix. First, recall that the parameter estimates are functions of sample L-Moments, l_rs, which themselves are linear combinations of the sample β PWMs, the b_rs. (We consider the β as opposed to the α PWMs since the covariance matrix for the sample β PWMs of the GLD is readily available in [8]). The sample PWMs are linear combinations of order statistics. Therefore the vector of $r+1$ sample β PWMs has an asymptotic multivariate normal distribution with mean $\underline{\beta}= (\beta_0, \beta_1, \beta_2 ... \beta_r)$, where r can take any value from 0 to $n-1$, and covariance matrix $n^{-1}V$, where, for the GLD, the $(i,j)^{th}$ element of V is given in [8] by $(i, j = 0, 1 ... r)$:

$$V_{ij} = J_{ij} + J_{ji}$$

where

$$J_{ij} = \frac{\alpha^2}{1+k} \frac{\Gamma(1+2k)\Gamma(i+j+1-2k)}{\Gamma(i+j+2)} \, {}_3F_2\begin{vmatrix} 1, j+1, 1+2k \\ i+j+2, 2+k \end{vmatrix}, \quad | k |< 1/2$$

Since we are interested in deriving the asymptotic distribution for our three parameter estimators, we consider the specific case of $r = 2$, ($i, j = 0, 1, 2$) since the β PWMs involved are β_0, β_1 and β_2. ${}_3F_2$ is the Generalized Hypergeometric Function of unit argument (See Appendix B for definition). Let $\theta = (\varepsilon, \alpha, k)$, the vector of parameters. Then the distribution of the vector of L-Moments estimators can be found using the previous result and the multivariate δ-method. The vector of estimators is:

$$\hat{\theta} = (\hat{\varepsilon}, \hat{\alpha}, \hat{k}) = f(b_0, b_1, b_2)$$

Although we have written our parameters in terms of L-Moments, it is easy to write them in terms of the β PWMs by simply substituting the appropriate linear combination of β PWMs for each of the L-Moments in the expression for each parameter.

$$k = -\frac{(6\beta_2 - 6\beta_1 + \beta_0)}{(2\beta_1 - \beta_0)}$$

$$\alpha = \frac{(2\beta_1 - \beta_0)}{\Gamma(1-k)\Gamma(1+k)}$$

$$\varepsilon = \beta_0 - \frac{\alpha}{k}(1 - \Gamma(1-k)\Gamma(1+k))$$

Define the 3x3 matrix $G = (g_{ij})$ by $g_{ij} = \partial f_i / \partial b_j$. Asymptotically the vector of estimators will then have a multivariate normal distribution with mean vector $f(\underline{\beta}) = (\varepsilon, \alpha, k)$ and variance-covariance matrix $n^{-1}GVG^T$. The matrix has the form:

$$n^{-1}GVG^T = n^{-1}\begin{bmatrix} \alpha^2 w_{11} & \alpha^2 w_{12} & \alpha w_{13} \\ \alpha^2 w_{12} & \alpha^2 w_{22} & \alpha w_{23} \\ \alpha w_{13} & \alpha w_{23} & w_{33} \end{bmatrix} \tag{3.3}$$

The w_{ij} are functions of k alone, and have complicated algebraic forms. However they can be evaluated numerically and are given in Table 3.1, for a range of typical values of k (Since ε is a location parameter, without loss of generality, let $\varepsilon = 0$). The elements of the matrix G can be derived explicitly and are as follows:

$$g_{11} = 1 + \frac{1}{k}\left[\frac{(k-1)(\Psi(1+k) - \Psi(1-k))}{\Gamma(1+k)\Gamma(1-k)} + \frac{(2k-1)}{k}\left(\frac{1}{\Gamma(1+k)\Gamma(1-k)} - 1\right)\right]$$

$$g_{12} = \frac{2(3-k)}{k}\left[\frac{\Psi(1+k) - \Psi(1-k)}{\Gamma(1+k)\Gamma(1-k)}\right] + \frac{(6-4k)}{k^2}\left[\frac{1}{\Gamma(1+k)\Gamma(1-k)} - 1\right]$$

$$g_{13} = -\frac{6}{k}\left[\frac{\Psi(1+k) - \Psi(1-k)}{\Gamma(1+k)\Gamma(1-k)} + \frac{1}{k}\left(\frac{1}{\Gamma(1+k)\Gamma(1-k)} - 1\right)\right]$$

$$g_{21} = \frac{1}{\Gamma(1+k)\Gamma(1-k)}\left[(1-k)(\Psi(1+k) - \Psi(1-k)) - 1\right]$$

$$g_{22} = \frac{2}{\Gamma(1+k)\Gamma(1-k)}\left[1 - (3-k)(\Psi(1+k) - \Psi(1-k)\right]$$

$$g_{23} = \frac{6(\Psi(1+k) - \Psi(1-k))}{\Gamma(1+k)\Gamma(1-k)}$$

k	w_{11}	w_{12}	w_{13}	w_{22}	w_{23}	w_{33}
-.4	4.1221	1.0525	1.3538	1.2220	-0.4761	1.7862
-.3	3.4202	0.8026	0.5376	1.0040	-0.3384	0.7743
-.2	3.2174	0.5416	0.3009	0.8435	-0.2170	0.4553
-.1	3.1427	0.2741	0.2125	0.7457	-0.1054	0.3259
0.0	3.2836	0.0016	-0.2867	0.7122	0.0002	0.2894
.1	3.1449	-0.2706	0.2145	0.7443	0.1054	0.3260
.2	3.2226	-0.5396	0.0305	0.8410	0.2153	0.4556
.3	3.4259	-0.8018	0.5422	0.9995	0.3343	0.7718
.4	4.1208	-1.0548	1.3512	1.2156	0.4667	1.7710

Table 3.1: Values of w_{ij} for typical values of k

$$g_{31} = -\frac{(1-k)}{\alpha\Gamma(1-k)\Gamma(1+k)}$$

$$g_{32} = \frac{2(3-k)}{\alpha\Gamma(1+k)\Gamma(1-k)}$$

$$g_{33} = -\frac{6}{\alpha\Gamma(1+k)\Gamma(1-k)}$$

where Ψ is the Digamma function. The $w_{ij}s$ are comp ted using a computer program, however it can easily be seen that α factors out in each term in the form given in the matrix.

3.6.1 Special Case: $k = 0$

Asymptotic results for this special case have been derived in [8] in detail. The asymptotic variance covariance matrix for the L-Moment estimators is given by:

$$n\text{var}\begin{bmatrix}\hat{\varepsilon}\\ \hat{\alpha}\end{bmatrix} \sim \alpha^2 \begin{bmatrix} 3 & 0 \\ 0 & .7101 \end{bmatrix}$$

Further, Hosking states the asymptotic variance-covariance matrix for the MLEs, given by Johnson and Kotz, [20]. For comparison, the matrix is:

$$n\text{var}\begin{bmatrix}\hat{\varepsilon}_{ml}\\ \hat{\alpha}_{ml}\end{bmatrix} \sim \alpha^2 \begin{bmatrix} 3 & 0 \\ 0 & 9/(3+\pi^2) \end{bmatrix} \doteq \alpha^2 \begin{bmatrix} 3 & 0 \\ 0 & .6993 \end{bmatrix}$$

Thus, from these two results, a comparison of the two sets of estimators may be made by calculating the relative efficiency of the L-Moment estimates to the MLE estimates.

$$\text{eff }(\hat{\varepsilon}) = 1$$

$$\text{eff }(\hat{\alpha}) = .9848$$

$$\text{eff }\begin{bmatrix}\hat{\varepsilon}\\ \hat{\alpha}\end{bmatrix} = .9848$$

(Note: The relative efficiency of the *vector* of estimators is defined as the ratio of the determinants of the two variance-covariance matrices.)

Unfortunately, the asymptotic variance-covariance matrix for the MLEs in the general case could not be found so that a similar comparison could be done for $k \neq 0$. However, given the high efficiency demonstrated in the special case of $k = 0$, the possibility exists that such a property holds in the general case.

3.7 Discussion

In this chapter, the three sets of estimators we wish to compare have been derived. From their theoretical forms, one can see that the L-Moment estimators are the easiest to calculate; each of the parameter estimators have an explicit form, and only evaluation of the required Gamma function may pose some difficulty. The maximum likelihood estimates are solutions of a complex set of nonlinear equations that will require a sophisticated numerical algorithm to determine the solution. The Method of Moments estimators require numerical techniques for the solution of one nonlinear equation, to find the shape parameter estimate, and have explicit expressions for the location and scale parameters. At this point, the L-Moment estimators certainly appear to be good estimators. However, the quality of all three methods must be determined in terms of precision and accuracy, before a choice among the three can be made. Such a comparison is the subject of the next chapter, in which a detailed simulation study is carried out to compare the methods for a range of small sample sizes, and for a range of reasonable parameter values.

In section 3.6 of this chapter, the asymptotic variance-covariance matrix was derived for the set of L-Moment estimators for the GLD parameters. These results will be used in Chapter 6 to derive a test of symmetry of the GLD based on the shape parameter estimator and its asymptotic normal distribution.

Chapter 4

SIMULATION STUDY

4.1 Introduction

In this chapter, the quality and viability of each of the three sets of estimators for the GLD are compared based on two criteria: bias and root mean square error [2]. The bias of an estimate $\hat{\theta}$ of a parameter θ is defined as

$$\text{Bias}(\hat{\theta}) = \hat{\theta} - \theta \tag{4.1}$$

and the Root Mean Square Error is defined as

$$\text{RMSE}(\hat{\theta}) = \sqrt{\text{Var}(\hat{\theta}) + (\text{Bias}(\hat{\theta}))^2} \tag{4.2}$$

These two quantities effectively measure the precision and accuracy of the estimators. Attention is restricted here to small sample sizes, since, for many practical applications in engineering, the data sets being analyzed are small due to the nature of the experiments. Further, it is well-known that the method of maximum likelihood provides asymptotically efficient estimators.

Thus, for very large sample sizes, the maximum likelihood estimators will provide the best parameter estimates. However, the good asymptotic properties of MLEs frequently do not hold in the small sample case. Therefore, in this section, we attempt to ascertain which of the three estimation methods performs best for small samples.

The properties of the three sets of estimators are examined using a simulation study based on 10,000 simulations, with sample sizes 15, 25, 50 and 100 and for values of k, the shape parameter, from -.4 to .4 in increments of .1. All methods of estimation are invariant under linear transformations of the data, so, without loss of generality, the location and scale parameters are held constant at $\varepsilon = 0$ and $\alpha = 1.0$ throughout.

4.2 Computation of Estimates

The MLEs posed some computational difficulties as they required the use of numerical methods to find the estimates, which concerned solving a system of three simultaneous nonlinear equations. Newton's Method [21] for three variables was the logical choice for an algorithm, but it failed to yield convergent estimates, regardless of the quality of initial values. Hence, the algorithm devised for this simulation study uses a somewhat inefficient, yet successful technique composed of the Bisection Method and Newton's Method for single variables. Each of the three equations is solved for one parameter (while the other two are held constant) in turn. At each successive step, the updated estimates are used to solve the appropriate equation. The algorithm cycles through the three equations in this fashion until the

n	k								
	-.4	-.3	-.2	-.1	0.0	.1	.2	.3	.4
15	35.1	22.4	22.4	8.2	6.3	11.5	5.7	8.0	11.8
25	23.5	10.1	9.0	2.3	6.1	1.4	1.1	2.0	4.1
50	22.4	8.7	2.1	1.1	5.1	0.1	0.3	0.5	0.6
100	17.2	14.2	1.1	0.6	1.1	0.1	0.1	0.2	0.3

Table 4.1: Failure Rate in Percent of MLE Algorithm

differences between successive estimates of each of the three parameters are less then a specified tolerance. At that point, the solution to the three equations is found (a more detailed algorithm is given in Appendix D). Naturally, any algorithm that depends on such an iterative numerical scheme is prone to failure during a run of simulations; that is, convergent estimates cannot be found. Such failures are more likely for small sample sizes and highly skewed data, when outliers can have a significant effect on the course of the iterations. In the case of a failure, the sample is discarded. For each set of simulations, the failure rate in percent is given in Table 4.1. The likelihood of failure is a significant drawback of the MLEs for this distribution.

The L-Moment estimates require no numerical methods as they are given in closed form, and only evaluation of the Gamma function (Γ) is required. For the MOM estimates, the shape parameter k must be found using an iterative scheme, and use of the Bisection Method proves a successful approach with no incidence of failure. The primary drawback of the MOM estimates is that they do not exist for $|k| > 1/3$, since population skewness does not exist for such values of k.

4.3 Results of the Simulation Study

4.3.1 Simple Logistic Distribution

Before we examine the general case we will consider a simulation study of the three methods for the simple Logistic Distribution (the special case of the GLD when the shape parameter $k = 0$). The MLEs of ϵ and α were found using Newton's Method for two variables to solve the two simultaneous equations given in Chapter 2, with negligible incidence of failure. Both the L-Moment estimates and the MOM estimates were found explicitly with no need for recourse to numerical methods.

From the simulation study results of Table 4.2, the following observation is made:

- The three methods perform similarly, except that the MLEs, in comparison to the other two methods, perform poorly in terms of bias and RMSE of the estimate of α for $n = 15$ and 25.

4.3.2 Generalized Logistic Distribution

We now consider the general case of estimating the three parameters of the GLD. The simulation study results are presented for three separate categories: $k < 0$, $k = 0$, $k > 0$. The results are given in Tables 4.3 to 4.7, where the column headings are the values of k. (In Table 4.5, $k = 0$).

n	Method	BIAS($\hat{\varepsilon}$)	BIAS($\hat{\alpha}$)	RMSE($\hat{\varepsilon}$)	RMSE($\hat{\alpha}$)
15	MLE	-.01	-.05	.44	.36
	LMOM	-.01	.00	.46	.23
	MOM	-.01	-.02	.47	.23
25	MLE	.00	-.04	.34	.26
	LMOM	.00	.00	.35	.17
	MOM	.00	-.02	.36	.18
50	MLE	.00	-.01	.25	.14
	LMOM	.00	.00	.25	.12
	MOM	.00	-.01	.26	.13
100	MLE	.00	-.01	.17	.08
	LMOM	.00	.00	.17	.08
	MOM	.00	.00	.18	.09

Table 4.2: Estimation of the 2 parameters of the Simple Logistic Distribution

n	Method	BIAS($\hat{\varepsilon}$)				BIAS($\hat{\alpha}$)				BIAS($\hat{k}$)			
		-.4	-.3	-.2	-.1	-.4	-.3	-.2	-.1	-.4	-.3	-.2	-.1
15	MLE	.12	-.03	-.02	.05	.14	-.07	-.03	.03	-.04	.08	-.03	-.06
	LMOM	.04	.02	.02	.00	-.05	-.04	-.04	-.04	.07	.04	.03	.01
	MOM	—	.07	.14	.08	—	-.23	.03	-.02	—	-.03	.09	.05
25	MLE	.13	-.05	-.01	.00	.14	-.05	-.03	-.03	.07	-.02	-.03	-.01
	LMOM	.03	-.01	.02	-.01	-.03	-.03	-.03	-.02	.05	.03	.02	.01
	MOM	—	.04	.13	.06	—	-.20	.04	-.01	—	-.03	.08	.04
50	MLE	.12	.01	.00	.00	.14	-.02	-.02	-.02	.09	-.01	-.01	-.01
	LMOM	.02	.01	.00	.00	-.02	-.02	-.02	.01	.03	.02	.01	.00
	MOM	—	.03	.10	.06	—	-.18	.04	.00	—	-.03	.07	.03
100	MLE	.13	.01	.00	.00	.14	-.01	-.01	-.01	.14	.00	-.01	.00
	LMOM	.01	.01	.00	.00	-.02	-.01	-.01	.00	.02	.01	.01	.00
	MOM	—	.02	.08	.04	—	-.15	.04	.00	—	-.03	.06	.03

Table 4.3: Bias of Estimates for GLD, k = -0.4 to -0.1

		BIAS($\hat{\varepsilon}$)				BIAS($\hat{\alpha}$)				BIAS($\hat{k}$)			
n	Method	.1	.2	.3	.4	.1	.2	.3	.4	.1	.2	.3	.4
15	MLE	.36	.35	.44	.74	.16	-.02	-.12	-.29	.27	.21	.20	.24
	LMOM	-.01	-.01	-.03	-.05	-.04	-.04	-.04	-.05	-.01	-.03	-.05	-.07
	MOM	-.05	.00	-.08	—	-.03	-.15	-.20	—	-.03	.05	.03	—
25	MLE	.22	.20	.36	.55	.09	-.01	-.10	-.23	.17	.11	.12	.15
	LMOM	-.01	-.01	-.02	-.03	-.02	-.02	-.03	-.03	-.01	-.02	-.03	-.05
	MOM	-.05	-.01	-.07	—	-.02	-.11	-.15	—	-.03	.04	.02	—
50	MLE	.10	.09	.36	.54	.05	-.02	-.10	-.20	.08	.04	.07	.10
	LMOM	.00	.00	-.01	-.02	-.01	-.02	-.02	-.02	.00	-.01	-.02	-.03
	MOM	-.04	-.01	-.06	—	-.01	-.07	-.10	—	-.03	.02	.01	—
100	MLE	.04	.04	.49	.85	.02	-.01	-.12	-.26	.04	.01	.06	.12
	LMOM	.00	.00	-.01	-.01	.01	-.01	-.01	-.01	.00	-.01	-.01	-.02
	MOM	-.03	-.03	-.05	—	.01	-.02	-.05	—	.00	-.01	.00	—

Table 4.4: Bias of Estimates for GLD, $k = 0.1$ to 0.4

<u>Case 1: $k < 0$</u> **(Tables 4.3, 4.6)**

- For k = -.4, MLEs suffer large positive bias for estimates of ε and α.

- For k = -.3 to -.1, the MLEs and LMOM estimates for all three parameters perform comparably in terms of bias, with the LMOM estimators performing slightly better for the smaller sample sizes.

- For k = -.3 to -.1, the MOM estimate of ε has the largest bias of the three methods. For k = -.3, the MOM estimate of α has large bias. For the estimate of k, the bias of the MOM estimate is larger than the other two methods in almost all cases.

- For k = -.4, the RMSE of all three MLE estimators is at least as large as that of the LMOM estimators. For the estimate of ε , for k = -.3 to -.1, the three methods perform comparably for all sample sizes.

- For k = -.3 to - .1, the MLE and LMOM estimators of α have comparable RMSE, but MOM estimates of α have larger RMSE is all cases.

- For k = -.3 to -.1, the MLE and LMOM estimates of k have comparable RMSE for all sample sizes, whereas the MOM estimates of k seem to be the best.

<u>Case 2: $k = 0$</u> **(Table 4.5)**

- In terms of bias, the LMOM estimates appear to perform the best for all sample sizes and for all three parameters, while the MOM estimates are the worst.

		$\hat{\varepsilon}$		$\hat{\alpha}$		$\hat{k}$	
n	Method	BIAS	RMSE	BIAS	RMSE	BIAS	RMSE
15	MLE	.09	.51	-.03	.28	.02	.25
	LMOM	.01	.47	-.04	.22	.00	.15
	MOM	-.04	.46	-.04	.22	-.03	.05
25	MLE	.02	.38	-.03	.14	.01	.18
	LMOM	-.01	.36	-.02	.17	.00	.11
	MOM	-.05	.36	-.02	.17	-.03	.05
50	MLE	.00	.26	-.02	.12	.00	.09
	LMOM	.00	.25	-.01	.12	.00	.08
	MOM	-.04	.26	-.01	.13	-.02	.04
100	MLE	.00	.18	-.01	.05	.00	.06
	LMOM	.00	.18	.00	.08	.00	.05
	MOM	- .03	.18	-.01	.09	-.02	.03

Table 4.5: Bias and RMSE of Estimates for GLD when $k = 0$

		RMSE($\hat{\varepsilon}$)				RMSE($\hat{\alpha}$)				RMSE($\hat{k}$)			
n	Method	-.4	-.3	-.2	-.1	-.4	-.3	-.2	-.1	-.4	-.3	-.2	-.1
15	MLE	.49	.42	.44	.47	.49	.23	.23	.23	.25	.17	.17	.24
	LMOM	.49	.48	.47	.47	.29	.26	.24	.23	.20	.18	.16	.15
	MOM	—	.50	.51	.47	—	.50	.36	.25	—	.03	.09	.07
25	MLE	.47	.33	.35	.37	.72	.19	.18	.17	.21	.13	.13	.16
	LMOM	.37	.36	.37	.36	.22	.20	.18	.18	.16	.14	.13	.12
	MOM	—	.37	.40	.37	—	.49	.29	.21	—	.03	.09	.07
50	MLE	.34	.24	.25	.25	.51	.14	.13	.12	.12	.10	.09	.08
	LMOM	.27	.26	.26	.25	.16	.14	.13	.12	.12	.10	.09	.08
	MOM	—	.27	.28	.26	—	.21	.21	.15	—	.03	.08	.06
100	MLE	.30	.18	.18	.18	.37	.10	.09	.09	.25	.07	.06	.06
	LMOM	.19	.18	.18	.18	.11	.10	.09	.09	.09	.08	.07	.06
	MOM	—	.19	.21	.19	—	.12	.15	.10	—	.03	.07	.05

Table 4.6: Root Mean Square of Estimates for GLD, k = -0.4 to -0.1

- In terms of RMSE, for the estimates of ε and α, the three methods perform comparably. For the estimates of k, the MOM performs the best, and the MLEs perform the worst, especially for $n = 15$.

<u>Case 3: $k > 0$</u> **(Tables 4.4, 4.7)**

- For the smaller samples, the MLEs are very biased in all cases for all parameter estimates. They are much more biased than either the LMOM or MOM estimators.

		RMSE($\hat{\varepsilon}$)				RMSE($\hat{\alpha}$)				RMSE($\hat{k}$)			
n	Method	.1	.2	.3	.4	.1	.2	.3	.4	.1	.2	.3	.4
15	MLE	.62	.64	.70	.82	.41	.29	.28	.35	.35	.31	.30	.32
	LMOM	.46	.47	.48	.49	.23	.24	.26	.29	.15	.16	.18	.19
	MOM	.48	.49	.51	—	.26	.44	.64	—	.07	.09	.04	—
25	MLE	.49	.50	.61	.74	.28	.21	.22	.29	.24	.20	.22	.24
	LMOM	.36	.36	.37	.38	.17	.18	.20	.22	.12	.13	.14	.16
	MOM	.37	.38	.39	—	.20	.37	.60	—	.07	.09	.04	—
50	MLE	.32	.35	.55	.72	.17	.13	.17	.25	.13	.11	.14	.17
	LMOM	.25	.25	.26	.27	.12	.13	.14	.16	.08	.09	.11	.12
	MOM	.26	.28	.28	—	.14	.29	.53	—	.06	.08	.05	—
100	MLE	.20	.24	.59	.70	.10	.09	.15	.23	.07	.07	.10	.11
	LMOM	.18	.18	.18	.19	.09	.09	.10	.11	.06	.07	.08	.09
	MOM	.19	.21	.21	—	.10	.22	.44	—	.05	.07	.05	—

Table 4.7: Root Mean Square of Estimates for GLD, k = 0.1 to 0.4

- MOM estimates have large bias for the estimate of α for smaller sample sizes.

- For positive k, the LMOM estimates of all three parameters seem to be uniformly better in terms of bias than either of the other two methods.

- For small samples, $k = .3$ and $.4$, the RMSE of the MLEs for all parameter estimates is much larger than that of either the LMOM or MOM estimators.

- For the estimates of ε, $k = .1$ to $.3$, the LMOM and MOM estimators perform comparably in terms of RMSE for all sample sizes.

- For the estimates of α, the MOM estimator has high RMSE, especially for $k = .2$ and $.3$. The MOM estimators of k perform slightly better than either LMOM estimators or MLEs.

- Overall for $k > 0$, in terms of RMSE, LMOM estimators seem to perform the most consistently for all parameters.

4.4 Discussion

In conclusion, it would appear from this study that the LMOM estimators often perform better than the MLEs: in the cases where this is not true, the improvement achieved by MLEs is negligible. In terms of practical research, it is desirable to have a method of estimation that performs consistently for a variety of situations, as do the LMOM estimates in this study. The same cannot be said for the MLE or MOM estimates as performance varies greatly depending of the amount of skewness of the data and the sample size. Further, the MLE method is prone to failure, making it an even less reliable method to use for real data. The results of this study and the simplicity of the computation makes the LMOM estimators a good choice.

Chapter 5

MODELING ICE BREAKAGE STRENGTH

5.1 Introduction

In this section, we consider nine data sets, five of which have been analyzed previously by Lal and Parsons [29], and all of which have been analyzed by Lal [22]. These data sets contain flexural strength and fracture toughness measurements of various samples of ice, which are important engineering parameters. The strength or toughness of ice is measured by the amount of pressure or compressive stress required to break or fracture a specimen of a specific type of ice. Each data set represents a different type of ice having varying micro-structure, temperature, salinity, volume, etc. These data sets have been accumulated over several years by a variety of workers. For details on how the data were collected, see Parsc s et al. [30]. It is of considerable interest to analyze the collected data using a straightforward method of statistical analysis. Due to the diversity in varieties of ice, a different model from a family of models is required for each set of data. Two distributions commonly used to model this type of data are the three-parameter Weibull distribution and the two-parameter Gumbel distribution (Extreme Value Distribution Type I). In the analysis of Lal and Parsons [29], a three-parameter Weibull distribution (a reparametrization of the Generalized Extreme Value distribution) was fitted using maximum likelihood estimation to thirteen different data sets measuring flexural strength, five of which are considered in this practicum. The Kolmogorov-Smirnov goodness-of-fit test was performed to see how well the fitted model described the data, and in all cases the conclusion was that the model was adequate. Although the Weibull model was proven useful, in [29] the authors mention that maximum likelihood parameter estimation for the Weibull distribution can be difficult. For engineering and hydrological purposes, it is desirable to seek simpler distributions, and for this reason they considered the Gumbel distribution as an alternative. In Lal [22], the analysis was ex.ended further to include twelve more data sets measuring fracture toughness, four of which are considered in this practicum. In his report, the author stresses the use of the Gumbel distribution for all of the data sets to simplify analysis.

This desire for simplicity and ease of computation strengthens the case for L-Moment estimation. As we have seen, the MLEs for the GLD are extremely difficult to calculate, while the L-Moment estimates are easy to compute. The GLD was considered as a good candidate model for these types of data sets because of the GLD's similarity to the Generalized Extreme Value distribution in shape. Since a reparametrization of the GEV distribution (Weibull) was previously successful for modeling the data, it is reasonable to hypothesize that the using the GLD may provide a good alternative model and possibly an improved analysis. Further, we can use L-Moment estimation to fit the model, providing a great simplification over fitting the three-parameter Weibull distribution by maximum likelihood. (There is also the possibility of fitting the Weibull distribution to the data using L-Moment estimation, but the L-Moment estimates for this distribution are not as tractable as those for the GLD).

The purpose of this chapter is first to introduce a new distribution as a candidate model for this type of data, and second, to demonstrate the performance of L-Moment estimation for fitting the GLD model in comparison with MLE and MOM.

5.2 Selection of Data Sets

Out of twenty-five available data sets, nine were chosen to be modeled by the GLD on the basis of their sample L-skewness and L-kurtosis. (The data sets are given in Appendix A). In previous analysis of the twenty-five data sets, all were fitted by the same two distributions, either three-parameter Weibull or Gumbel. Although all of the twenty-five data sets could probably be fitted reasonably by the GLD, we have selected only those data sets which seem more suited to be fitted by the GLD than any other distribution. For the remaining data sets, the researcher may have to make a subjective decision between the GLD and the Weibull distribution. The L-Moment diagram

Data Set	Sample Size	L-Skewness	L-Kurtosis
CDAT2-Freshwater Ice -4C (kN $m^{-3/2}$)	44	-0.0780	0.1786
CDAT4-Freshwater Ice -24C (kN $m^{-3/2}$)	44	-0.0294	0.2185
CDAT8 -Small Size Specimen Fresh Water -2C (MPa)	55	0.1625	0.2040
CDAT9-Sea Ice (Horizontal) (MPa)	50	0.1212	0.1944
CDAT12-Sea Ice -20C (MPa)	20	0.1401	0.2464
CDAT13-Sea Ice (Old) (MPa)	19	0.0972	0.1591
CDAT14-Sea Ice -5C (MPa)	19	0.1274	0.2807
CDAT19-Resolute 87 (kN $m^{-3/2}$)	80	-0.0795	0.2886
CDAT23-Finegrained Columnar Freshwater Ice -20C (kN $m^{-3/2}$)	59	0.2329	0.1812

Table 5.1: L-Skewness and L-Kurtosis for the 9 Data Sets

(see Appendix F) would be helpful in making such a choice. The L-Moment diagram provides an easy way to judge which models may be appropriate for the data. The sample L-skewness and L-kurtosis for each of the data sets are given in Table 5.1. Note that the observations in each data set are either measurements of Critical Stress Intensity Factor (kN $m^{-3/2}$) — fracture toughness, or Compressive Stress (MPa) — flexural strength.

Plotting the L-skewness and L-kurtosis for each data set on the diagram (Appendix F) suggests that the GLD model is appropriate for these data sets.

5.3 Fitting the Model

For each data set, the GLD was fitted by Maximum Likelihood (MLE), L-Moments (LMOM) and conventional Moments (MOM). The Kolmogorov-Smirnov goodness-of-fit test statistic was then calculated to see which of the estimation procedures yielded a reasonable model. The results are given in Tables 5.2 and 5.3. Further, in Appendix E, three graphs are given for each of the nine data sets displaying the empirical C.D.F. and the fitted C.D.F. plotted on the same graph, for each method of estimation. This allows a visual inspection of the goodness-of-fit for each fitted model. In order to calculate the MLEs, the LMOM estimates were used as initial values for the iterative program. In all cases, these initial values failed to converge to a solution. Starting with the MOM estimates also failed to provide convergence. It was necessary to grid search over intervals around the L-Moment estimates to find suitable initial values.

As we can see from the results, the LMOM and MOM procedures give fairly similar estimates, while the MLE procedure yields estimates that are quite different from both of the other methods. The most obvious difference in the estimators is that the sign of $\hat{k}$, the estimate of the shape parameter, is usually **reversed** in the MLE case. The fact that the MLE of k is almost always positive for these data sets is a concern, since a positive value means

Data Set	Method	$\hat{\varepsilon}$	$\hat{\alpha}$	$\hat{k}$	K-S Stat
CDAT2	MLE	116.556	18.285	-0.053	0.072
(44)	LMOM	116.789	18.275	-0.042	0.073
	MOM	118.287	17.650	0.008	0.084
CDAT4	MLE	109.770	16.803	0.275	0.149
(44)	LMOM	106.023	11.804	-0.116	0.073
	MOM	107.323	11.616	-0.052	0.071
CDAT8	MLE	2.510	0.638	0.128	0.123
(55)	LMOM	2.410	0.480	-0.162	0.061
	MOM	2.432	0.489	-0.014	0.051
CDAT9	MLE	0.408	0.129	0.095	0.464 *
(50)	LMOM	0.548	0.070	-0.121	0.067
	MOM	0.553	0.071	-0.080	0.033
CDAT12	MLE	0.928	0.138	0.285	0.218
(20)	LMOM	0.900	0.080	-0.140	0.128
	MOM	0.908	0.082	-0.081	0.120
CDAT13	MLE	0.498	0.191	0.164	0.464 *
(19)	LMOM	0.689	0.103	0.097	0.092
	MOM	0.698	0.101	-0.037	0.075
CDAT14	MLE	0.874	0.216	0.269	0.202
(19)	LMOM	0.852	0.122	-0.127	0.118
	MOM	0.857	0.123	-0.099	0.111

Table 5.2: Fitting the Model by 3 Methods: * Significant at 5 %

Data Set	Method	$\widehat{\varepsilon}$	$\widehat{\alpha}$	$\widehat{k}$	K-S Stat
CDAT19	MLE	110.792	23.447	0.109	0.102
(80)	LMOM	107.582	19.566	-0.124	0.049
	MOM	108.123	19.692	-0.107	0.042
CDAT23	MLE	122.082	26.981	0.153	0.149
(59)	LMOM	116.112	18.379	-0.233	0.048
	MOM	119.273	19.474	-0.133	0.042

Table 5.3: Fitting the Model by 3 Methods: * Significant at 5 % (cont.)

that the distribution is skewed left with an upper bound (See Chapter 2). However, if one considers a quantity such as breaking strength or fracture toughness of ice, it is clear that there should be a *lower* bound. Although the MLE model is rejected only twice for lack of fit out of the nine data sets, this tendency towards positive estimates of k when they should clearly be negative is a definite disadvantage.

The three methods produce similar values for the estimate of the location parameter ε. For the scale parameter α, the LMOM estimates and the MOM estimates are quite similar, with the MLEs providing quite different values for almost all of the nine data sets. The rejection of the MLE model for CDAT9 and CDAT13 is most likely a result of poor estimation jointly for all parameters. Both the MOM and LMOM estimates perform well for all data sets, providing adequate models according to the Kolmogorov-Smirnov goodness-of-fit test. The slightly lower value of the test statistic for the MOM model may be attributed to the superior performance, in some cases, of the

MOM estimator of k in terms of RMSE, as seen in Chapter 4.

5.4 Discussion

The poor performance of the MLEs in comparison with the LMOM and MOM estimators should not be surprising. The difficulty in choosing initial values for the maximum likelihood algorithm suggests that the algorithm is not a reliable one. The presence of other maxima or minima of the log likelihood function, or simply outliers in the data, are likely to cause the resulting estimates to be of questionable accuracy. The data sets considered in this chapter were relatively small, and in some cases were quite skewed. Neither of these conditions provide an ideal environment for the numerical methods required to find the maximum likelihood estimates.

There are two important conclusions to be made based on the analysis of this chapter:

- The GLD should be considered as a model by hydrologists and engineers working with data of the type analyzed in this chapter. The GLD provides an alternative to the commonly used GEV distribution (or the equivalent Weibull distribution), while maintaining a similar range of shapes to accommodate different data sets.

- L-Moment estimation provides a quick, easy and reliable way to fit the GLD model.

Chapter 6

A TEST OF $k = 0$

6.1 Introduction

The simple Logistic distribution is a special case of the GLD when $k = 0$. A test of $k = 0$ could be viewed as a test of whether a set of data, fitted by the GLD, came from a symmetric simple Logistic distribution or from a skewed GLD (i.e. a test of symmetry of the underlying distribution). A test of this hypothesis can be based on the L-Moment estimator of k. In this chapter, such a test will be developed and illustrated with examples. To determine the reliability of the test, two simulation studies, based on 10,000 simulations, are performed. The first is used to determine how well the new test statistic controls the nominal size of the test. The second simulation study compares the power of the new test to the Likelihood Ratio Test [31] for some specific alternative hypotheses.

6.2 Symmetry Test

6.2.1 The L-Moment Test Statistic

In Chapter 3, the asymptotic distribution of the L-Moment parameter estimators was derived for some specific values of k. Under $H_o : k = 0$, the L-Moment estimator of k, $\hat{k}$, was found to be asymptotically $N(0, .2894/n)$ (See Table 3.1). A simple Z test statistic can be constructed by dividing the estimator by the square root of its asymptotic variance as follows:

$$Z = \frac{\sqrt{n}\,\hat{k}}{\sqrt{.2894}} \tag{6.1}$$

One must then compare the value of the statistic to the critical values of the standard normal distribution. A significant positive Z implies rejection of H_o in favor of the alternative $k > 0$, and similarly significant negative values of Z imply a rejection of H_o in favor of the alternative $k < 0$.

6.2.2 Likelihood Ratio Test (LR)

The Likelihood Ratio (LR) Test is a well-known test based on maximum likelihood estimation of the parameters under both the null and alternative hypotheses. The Likelihood Ratio test statistic for testing the hypothesis $H_o : \theta = \theta_o$, where θ is the vector of parameters of interest, and θ_o the restricted vector of parameter values under the null hypothesis, is given by:

$$\Lambda = \frac{L(\hat{\theta}_0/H_o)}{L(\hat{\theta}/H_a)} \tag{6.2}$$

The numerator is the Likelihood function evaluated at the value of the parameter estimates under the null hypothesis H_o. The denominator is the

Likelihood function evaluated under the alternative hypothesis H_a. Exact tests based on this statistic require knowledge of the distribution of Λ, which is usually not known. However, an asymptotic test can be constructed based on the following result: [31]:

$$-2\ln\Lambda \longrightarrow \chi_p^2$$

under H_o, where p is the number of restrictions under H_o and χ_p^2 is a chi-square random variable with p degrees of freedom.

For the test we want to consider in this section, we wish to test the hypothesis $H_o : k = 0$, against both one and two-sided alternatives. Thus, for this test, $\theta = (\varepsilon, \alpha, k)$ and $\theta_o = (\varepsilon, \alpha)$. The numerator of Λ is the Likelihood function of the simple Logistic distribution (recall that when $k = 0$, the GLD reduces to this distribution) evaluated at the maximum likelihood estimates of ε and α. The denominator is the likelihood function of the GLD evaluated at the maximum likelihood estimates of ε, α and k (see Chapter 3 for the maximum likelihood equations). For the two-tailed alternative, the statistic $-2\ln\Lambda$ is calculated and compared to the appropriate χ^2 critical value with one degree of freedom (there is one restriction made under H_o). For the one-tailed alternative, we refer to [12], where a similar test is given for testing symmetry of a Generalized Extreme Value distribution. The square root of the test statistic is taken, and given the sign of the estimate of k under H_a. Since it is known that the square root of a χ^2 variable with 1 degree of freedom is a N(0,1) variable, the square root quantity is compared to the appropriate N(0,1) critical value.

ALTERNATIVES						
	$k<0$		$k>0$		$k\neq 0$	
Sample Size	5%	10%	5%	10%	5%	10%
15	6.2	12.2	6.3	11.5	6.5	12.1
25	5.7	11.0	5.9	11.4	5.8	11.1
50	5.4	9.5	5.3	10.5	5.4	10.1
100	5.2	10.1	5.2	10.4	5.3	10.0
200	5.2	10.1	5.3	10.2	4.9	10.6
500	5.1	10.8	5.0	10.3	5.2	10.4

Table 6.1: Empirical Size for Nominal 5%, 10% Significance Level

6.2.3 Comparison of the Two Tests

It can be seen from the results in Table 6.1 that for sample sizes of 25 and greater, the size of the Z test is adequately controlled. (The results in this table were found by simulating 10,000 samples from the simple Logistic Distribution, calculating the Z test statistic for each and calculating the proportion of samples rejected). In Table 6.2 we see that the power of the Z test is almost as high as that of the Likelihood Ratio test, and in some cases it has improved power. Note that in Table 6.2, $n = 50$, and the significance level is 5%. (The values in the table were found by simulating 10,000 samples from each of the specified alternatives, calculating the two test statistics for each sample and calculating the proportion of samples rejected).

The LR test requires numerical evaluation of the MLEs of the parameters, and it can be quite difficult to compute. If convergence cannot be achieved

POWER UNDER SPECIFIED ALTERNATIVES						
k	Z Test			LR Test		
	$k < 0$	$k > 0$	$k \neq 0$	$k < 0$	$k > 0$	$k \neq 0$
-.4	.99	*	.98	.99	*	.96
-.3	.95	*	.92	.96	*	.93
-.2	.77	*	.67	.79	*	.69
-.1	.35	*	.25	.36	*	.26
0.0	.05	.05	.05	.06	.06	.06
.1	*	.35	.26	*	.47	.26
.2	*	.76	.68	*	.83	.69
.3	*	.95	.92	*	.96	.72
.4	*	.99	.99	*	.90	.68

Table 6.2: Power Comparison Between Two Tests

for the estimates, then the statistic cannot be used. For real data sets, this is a serious concern. Further, due to the tendency toward large MSE of the MLEs when k is close to .4, it can be seen from Table 6.2 that, for practical purposes, the power of the LR test decreases at this point. Strong bias for positive values of k also affects the power of the LR test.

6.3 Examples

To demonstrate the Z Test based on L-Moments, we shall select two of the data sets considered in the previous chapter. From the results of Table 5.2 and 5.3 it is clear that some of the data sets may be fitted by a simpler model, namely the simple Logistic distribution. The estimates of k for such data sets tend to be quite small. For example, consider the data set CDAT2. The L-Moment estimate of k is $\hat{k} = -0.042$. Since this value is very close to 0, it is reasonable to test whether the reduced model adequately describes the data. Our hypotheses of interest are $H_o : k = 0$ vs. $H_a : k \neq 0$. The Z test statistic is:

$$\frac{\sqrt{44}(-0.042)}{\sqrt{.2894}} = -0.529$$

This value is not significant at any reasonable significance level, so we may conclude that, for this data set, the simpler model is adequate.

Next, consider a smaller data set, CDAT13. Hosking [8] asserts that the asymptotic variances of L-Moment estimators are good approximations even for small sample sizes. Further, this Z test controls the size of the test reasonably well for a sample size as small as 15. Hence, we can feel confident using this Z statistic for CDAT13, with a sample size of 19. The L-Moment estimate of k is $\hat{k} = -.097$. The same hypotheses as above are considered and the test statistic is:

$$\frac{\sqrt{19}(-.097)}{\sqrt{.2894}} = -.786$$

Again, this value is not significant for any reasonable significance level.

For both of the above data sets, we can conclude that the simple Logistic model may be used to fit these data sets (i.e. $k = 0$).

6.4 Discussion

The purpose of this chapter has been to provide a simple alternative to the well known Likelihood Ratio test for testing symmetry of the GLD. Although the LR test possesses good properties, this is not always evident when dealing with highly skewed and/or small data sets. The test statistic requires the calculation of the maximum likelihood estimates of the model parameters under both the null and alternative hypotheses. As mentioned in previous chapters, estimation requires complicated numerical methods that, for small sample sizes, are prone to convergence failure. Furthermore, as seen in Chapter 5, when using real data sets, this estimation procednre becomes even more difficult due to the problem of finding successful initial starting values for the maximum likelihood algorithm. For highly skewed or small data sets, the MLEs tend to be severely biased which may cause a misleading test statistic value. On the other hand, the Z Test statistic is both simple and powerful and its computation is straightforward. It appears that although this is an asymptotic test, it may be used for even small sample sizes. In conclusion, this test provides a good test of symmetry, and is powerful enough to replace the LR statistic as a test statistic for this distribution.

Chapter 7

Conclusion

In this practicum, a relatively new and under-utilized method of estimation based on quantities called L-Moments has been reviewed. A distribution called the Generalized Logistic Distribution (GLD), first given by Hosking [8] in his technical report, has been suggested as a good alternative to the commonly used Generalized Extreme Value distribution for some engineering data. A model for a given set of data may be chosen by using the L-Moment Ratio diagram (see Appendix F). As an estimation technique for the fitting of the GLD, L-Moments work extremely well, performing consistently and with good accuracy and precision even in small samples, as illustrated by the simulation study of Chapter 4. They are frequently superior to the maximum likelihood estimates (MLEs), which for the GLD are difficult to calculate due to convergence problems in the maximum likelihood algorithm. The application to some real engineering data shows that although all three methods may be used with reasonable security (goodness-of-fit hypothesis is not rejected), the MLEs often give an estimate of the shape parameter of the GLD inconsistent with the nature of the data. The L-Moment estimators have a simple explicit form, making them the preferred choice of the three for fitting the GLD model.

Finally, a test was proposed for testing if the shape parameter of the GLD is equal to zero based on the L-Moment estimator of the shape parameter and its asymptotic normal distribution discussed in Chapter 3. The proposed test was simple and easy to compute, and performed well with respect to controlling the size of the test. Further, it had power comparable to that of the Likelihood Ratio Test.

In conclusion, it is the opinion of the author that for the GLD, and likely for many other models, L-Moments provide simple and reliable estimates. The method of L-Moments could be used by engineers and hydrologists instead of maximum likelihood estimation, often without need for recourse to complicated numerical methods.

Bibliography

[1] Ahmad, M.I., Sinclair, C.D., Spurr, B.D. (1988) "Assessment of Flood Frequency Models Using Empirical Distribution Function Statistics." *Water Resources Research,* 24(8): 1323-1328.

[2] Bickel, P.J., Doksum, K.A. (1977) *Mathematical Statistics: Basic Ideas and Selected Topics.* Holden-Day Inc., San Francisco.

[3] Bishop, Y.M.M., Fienberg, S.E., Holland, P.W. (1991) *Discrete Multivariate Analysis: Theory and Practice.*, MIT Press, Cambridge.

[4] David, H.A. (1981) *Order Statistics.* , Wiley, New York.

[5] Greenwood, J.A., Landwehr, J.M., Matalas, N.C., Wallis, J.R. (1979) "Probability Weighted Moments: Definition and Relation to Parameters of Several Distributions Expressible in Inverse Form." *Water Resources Research,* 15(5):1049-1054.

[6] Haktanir, T. (1992) "Comparison of Various Flood Frequency Distributions Using Annual Flood Peaks Data of Rivers in Anatolia." *Journal of Hydrology,* 136:1-31.

[7] Hoeffding, W. (1948) "A Class of Statistics with Asymptotically Normal Distribution." *Annals of Mathematical Statistics.* 19 : 293-325.

[8] Hosking, J.R.M. (1986). *The Theory of Probability Weighted Moments*, Technical Report RC12210, IBM. Research Division , New York.

[9] Hosking, J.R.M. (1990) "L-Moments: Analysis and Estimation of Distributions using Linear Combinations of Order Statistics." *Journal of the Royal Statistical Society, Series B*, 32(1):105-121.

[10] Hosking, J.R.M. (1987) "Parameter and Quantile Estimation for the Generalized Pareto Distribution." *Technometrics*, 29(3):339-348.

[11] Hosking, J.R.M. (1992) "Moments or L-Moments? An Example Comparing Two Measures of Distributional Shape." *The American Statistician*, 46(3):186-189.

[12] Hosking J.R.M. (1984) "Testing Whether the Shape Parameter is Zero in the Generalized Extreme Value Distribution." *Biometrika*, 71(2):367-374.

[13] Hosking, J.R.M. (1993) "Some Statistics Useful in Regional Frequency Analysis." *Water Resources Research*, 29(2): 271-281.

[14] Hosking, J.R.M. (1985) "Estimation of the Generalized Extreme Value Distribution by the Method of Probability Weighted Moments." *Technometrics*, 27(3) : 251-261.

[15] Jakubowski, W. (1992) "The Application of Probability Weighted Moments in Estimating the Parameters of a Pearson Type 3 Distribution - Comment." *Journal of Hydrology,* 136:395-399.

[16] Jing, D., Dedun, S. (1988) "The Application of Probability Weighted Moments in Estimating the Parameters of the Pearson Type 3 Distribution." *Journal of Hydrology,* 101:47-61.

[17] Jing, D., Dedun, S., Ronfu, Y. (1989) "Further Research on Application of Probability Weighted Moments in Estimating Parameters of the Pearson Type 3 Distribution." *Journal of Hydrology.* 110:239-257.

[18] Jing, D., Rongfu, Y. (1988) "The Determination of Probability Weighted Moments with the Incorporation of Extraordinary Values into Sample Data and Their Application to Estimating Parameters for the Pearson Type 3 Distribution." *Journal of Hydrology,* 101: 63-81.

[19] Jing, D., Dedun, S. Ronfu, Y.,Yu, H. (1989) "Expressions Relating Probability Weighted Moments to Parameters of Several Distributions Inexpressible in Inverse Form." *Journal of Hydrology,* 110:271-294.

[20] Johnson, N.L., Kotz, S. (1970) *Distributions in Statistics: Continuous Univariate Distributions.,* Wiley, New York.

[21] Kahaner, D., Moler, C., Nash, S. (1989) *Numerical Methods and Software.* Prentice Hall Series in Computational Mathematics, New Jersey.

[22] Lal, M. (1988) "Statistical Analysis of Fracture Toughness and Flexural Strength." Technical Report LM-AVR-25, Memorial University of Newfoundland.

[23] Lal, M. (1989) "Extremal Analysis of Flexural Strength of Ice and Crack Size Distribution." Technical Report cr-1989-04, Memorial University of Newfoundland.

[24] Landwehr, J.M., Matalas, N.C., Wallis, J.R. (1979) "Probability Weighted Moments Compared with Some Traditional Techniques in Estimating Gumbel Parameters and Quantiles." *Water Resources Research,* 15(5):1055-1064

[25] Landwehr, J.M., Matalas, N.C., Wallis, J.R. (1979) "Estimation of Parameters and Quantiles of Wakeby Distributions. 1. Known Lower Bounds." *Water Resources Research,* 15(6) :1361-1372.

[26] Manoukian, E.B. (1986) *Modern Concepts and Theorems of Mathematical Statistics.* Springer Verlag, New York.

[27] Mathai, A.M., Saxena, R.K. (1973) *Generalized Hypergeometric functions with Applications in Statistics and Physical Sciences.,* Springer, Berlin.

[28] Muirhead, R.J. (1982) *Aspects of Multivariate Statistical Theory.* Wiley, New York.

[29] Parsons, B.L., Lal, M. (1991) "Distribution Parameters for Flexural Strength of Ice." *Cold Regions Science and Technology,* 19: 285-293

[30] Parsons, B.L., Lal, M., Williams, F.M., Dempsey, J.P., Snellen, J.B., Everard, J., Slade, T., Williams, J. (1992) "The Influence of Beam Size on the Flexural Strength of Sea Ice, Freshwater Ice, and Iceberg Ice." *Philosophical Magazine A.* 66(6) : 1017-1036.

[31] Rao, C.R. (1973) *Linear Statistical Inference and its Applications.* Wiley, New York.

[32] Shoukri, M.M., Mian, I.U.H., Tracy, D.S. (1988) "Sampling Properties of Estimators of the Log-logistic Distribution with Application to Canadian Precipitation Data." *The Canadian Journal of Statistics,* 16(3) :223-236.

[33] Singh, V.P., Gou, H., Yu, F.X. (1993) "Parameter Estimation for Three-Parameter Log-logistic Distribution (LLD3) by POME." *Stochastic Hydrology and Hydraulics,* 7:163-177.

[34] Singh, V.P., Cruise, J.F.,Ma. Ming (1990) "A Comparative Evaluation of the Estimators of the Weibull Distribution by Monte Carlo Simulation". *Journal of Statistical Computation and Simulation.* 36:229-241.

[35] Stigler, S.M. (1974) "Linear Functions of Order Statistics with Smooth Weight Functions." *Annals of Statistics,* 2: 676-693.

[36] Wang, Q.J. (1991) "The POT Model Described by the Generalized Pareto Distribution with Poisson Arrival Rate." *Journal of Hydrology,* 129:263-280.

Appendix A

Data Sets

CDAT2 : (kN$m^{-3/2}$) 43.6 65.2 66.4 79.8 83.9 84.5 85.4 85.8 88.8 95.3 95.6 96.9 98.5 99.4 100.8 102.5 105.7 107.1 109.9 111 113.9 115.5 118.7 120.6 122.2 122.2 122.6 123.8 124.1 127.1 132.7 135.7 139.2 139.4 144 148.2 157.1 159.3 163.3 164.5 166.3 168.3 174.9 184.9

CDAT4: (kN$m^{-3/2}$) 72.3 73.4 78.7 80.4 82.7 83.7 85.3 85.3 90.5 93.3 93.5 95.8 98.3 98.6 98.9 100.2 101.2 101.3 101.9 102.3 102.3 102.7 103.5 103.6 105.2 107.4 107.5 107.7 115 116.1 116.9 118.1 123.8 124 125.4 128.1 132.8 136 139.2 140.3 140.8 142.4 148 161.4

CDAT8 : (MPa) 1.05 1.21 1.22 1.23 1.45 1.47 1.60 1.63 1.67 1.71 1.77 1.86 1.89 1.89 1.94 1.94 2.01 2.05 2.06 2.11 2.18 2.26 2.32 2.37 2.38 2.39 2.40 2.47 2.50 2.50 2.55 2.56 2.65 2.65 2.67 2.72 2.79 2.89 2.90 2.96 2.99 3.01 3.05 3.05 3.08 3.21 3.29 3.32 3.37 3.46 3.50 3.82 4.29 5.48 6.09

CDAT9 : (MPa) .27 .35 .37 .42 .43 .43 .44 .45 .45 .46 .46 .46 .47 .48 .48 .49 .49 .50 .51 .52 .52 .53 .55 .55 .55 .56 .56 .57 .57 .57 .58 .59 .60 .61 .61 .61 .61

.62 .63 .63 .63 .67 .70 .71 .72 .73 .74 .79 .90 .97

CDAT12 : (MPa) .63 .73 .81 .82 .83 .83 .84 .84 .86 .90 .92 .92 .95 .96 .97 1.03 1.04 1.05 1.11 1.34

CDAT13 : (MPa) .36 .51 .52 .55 .57 .57 .58 .67 .68 .69 .74 .74 .75 .76 .78 .87 .94 1.02 1.08

CDAT14 : (MPa) .49 .55 .65 .74 .76 .79 .79 .80 .80 .842 .88 .91 .95 .97 1.00 1.00 1.10 1.10 1.55

CDAT19 : (kN$m^{-3/2}$) 41 42 50 54 55 69 70 71 73 74 78 79 79 80 80 80 84 85 86 88 88 88 89 89 89 91 93 96 96 98 98 98 98 103 103 103 104 105 108 109 110 112 112 112 112 113 114 115 116 118 119 120 122 123 125 127 127 128 129 130 131 134 135 136 138 139 143 144 144 148 149 151 155 156 156 158 195 199 215 258

CDAT23 : (kN$m^{-3/2}$) 68 69 75 77 80.5 84 90 91 92.4 93 96.5 96.5 97 97.5 98 98.5 98.5 99 102 103 103 106 106 106 106 107 107 107 112 116 116 120 120 124 126 126 128 129 132 133 134 136 136 137 139 141 146 149 151 153 166 167 170 173 180 181 221 222 255

Appendix B

Special Functions

1. The Generalized Hypergeometric Function with Unit Argument [28]:

$$ {}_3F_2(a, b, c; d, e) = \sum_{k=0}^{\infty} \frac{(a)_k (b)_k (c)_k}{(d)_k (e)_k} \frac{1}{k!} $$

where

$$ (x)_k = x(x+1)\cdots(x+k-1) $$

2. The Digamma Function

$$ \Psi(x) = \frac{d \ln \Gamma(x)}{dx} = \frac{\frac{d}{dx}\Gamma(x)}{\Gamma(x)} $$

3. The n^{th} shifted Legendre Polynomial

$$ P_n^*(x) = \sum_{k=0}^{n} p_{n,k}^*(x) x^k $$

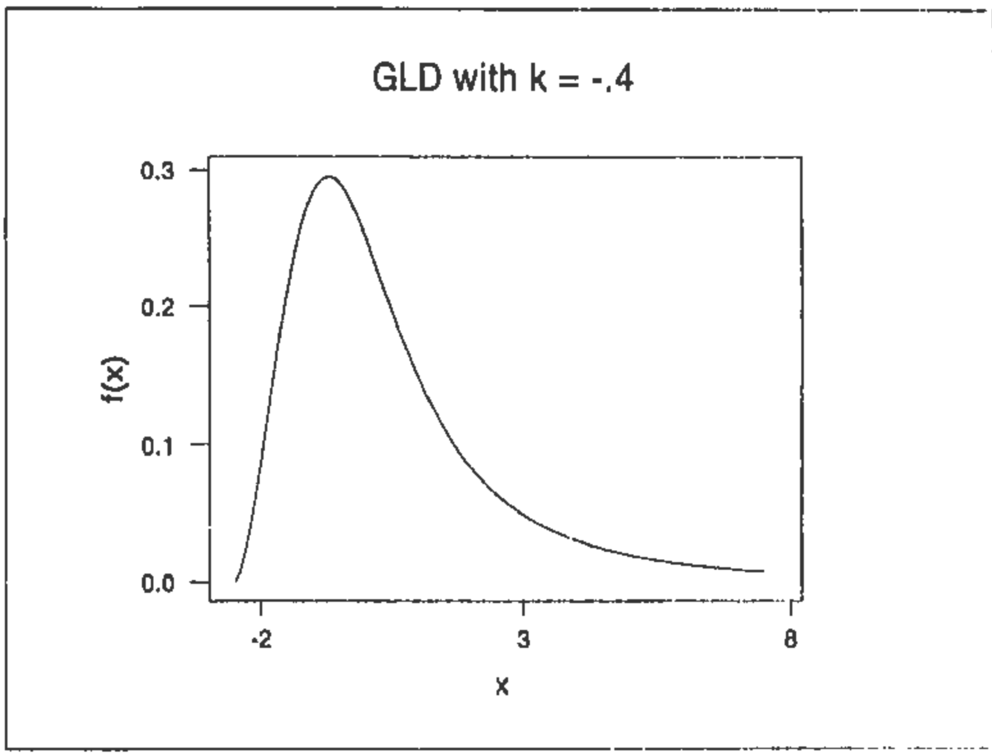
where

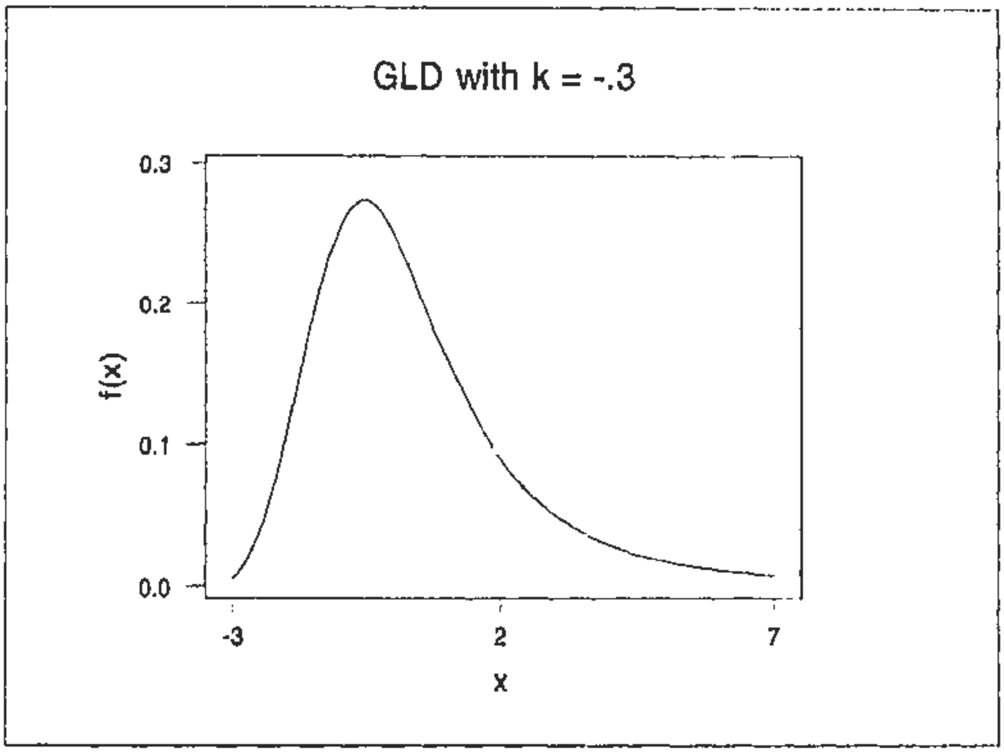
$$ p_{n,k}^*(x) = (-1)^{n-k} \binom{n}{k} \binom{n+k}{k} $$

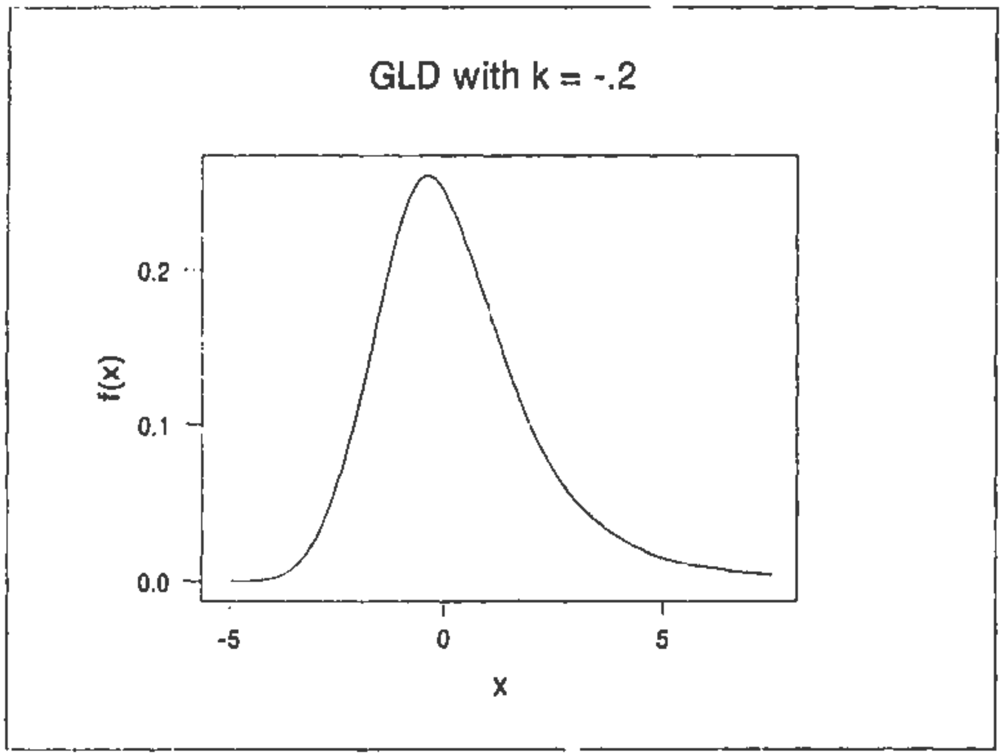
Appendix C

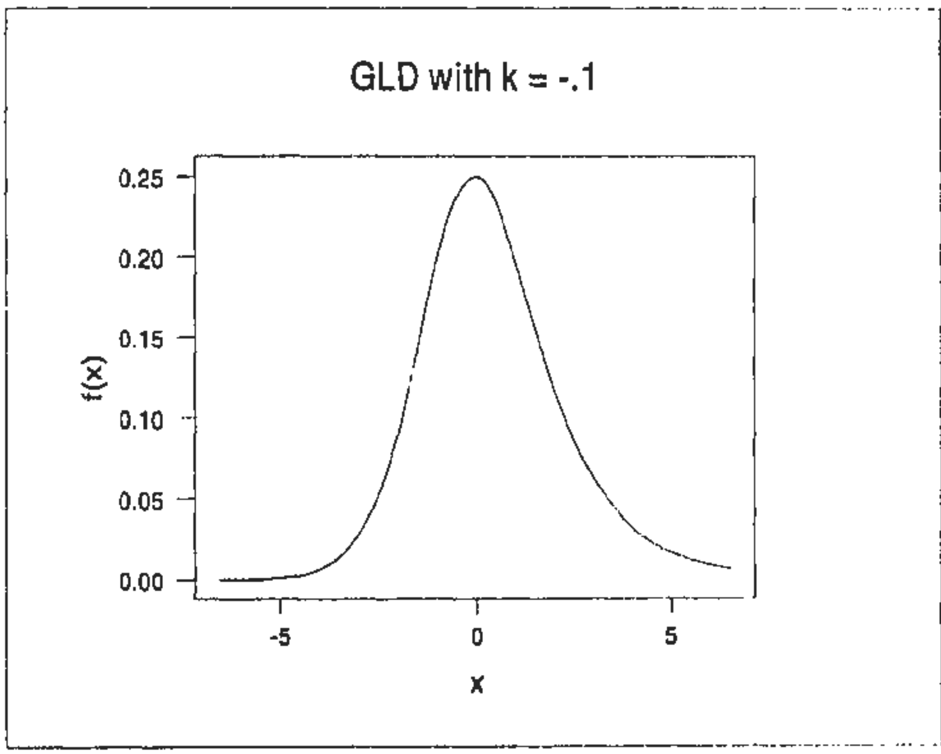
GLD for Different Values of the Shape Parameter

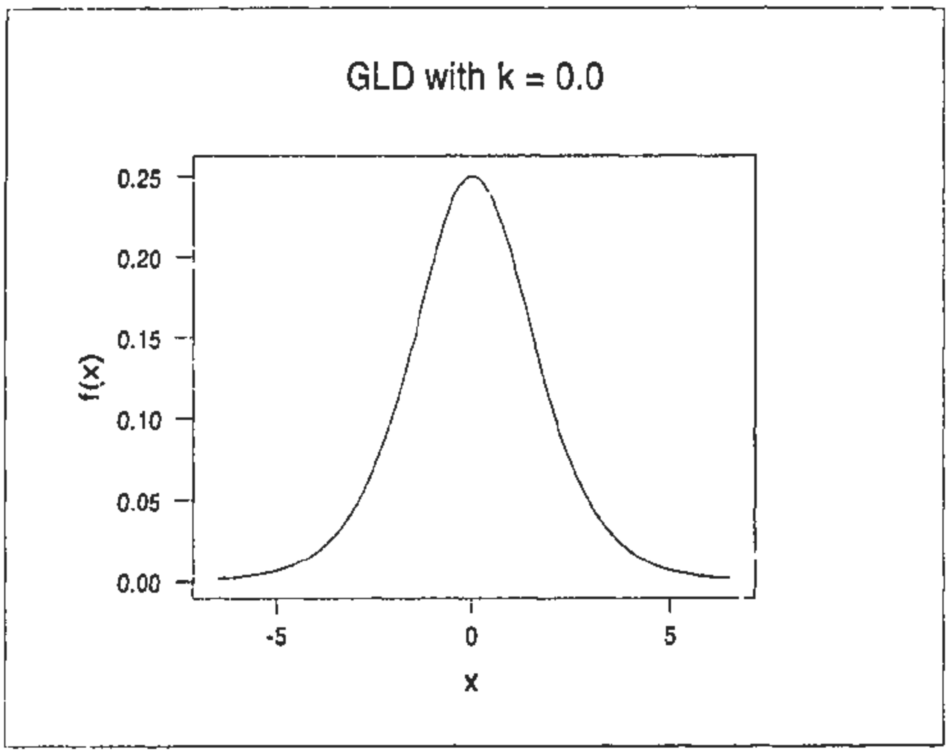
In this appendix, the probability distribution function of the GLD is graphed for nine different values of the shape parameter: $k = -.4, -.3, -.2, -.1, 0.0, .1, .2, .3$, and $.4$. $\varepsilon = 0.0$ and $\alpha = 1.0$ for all graphs.

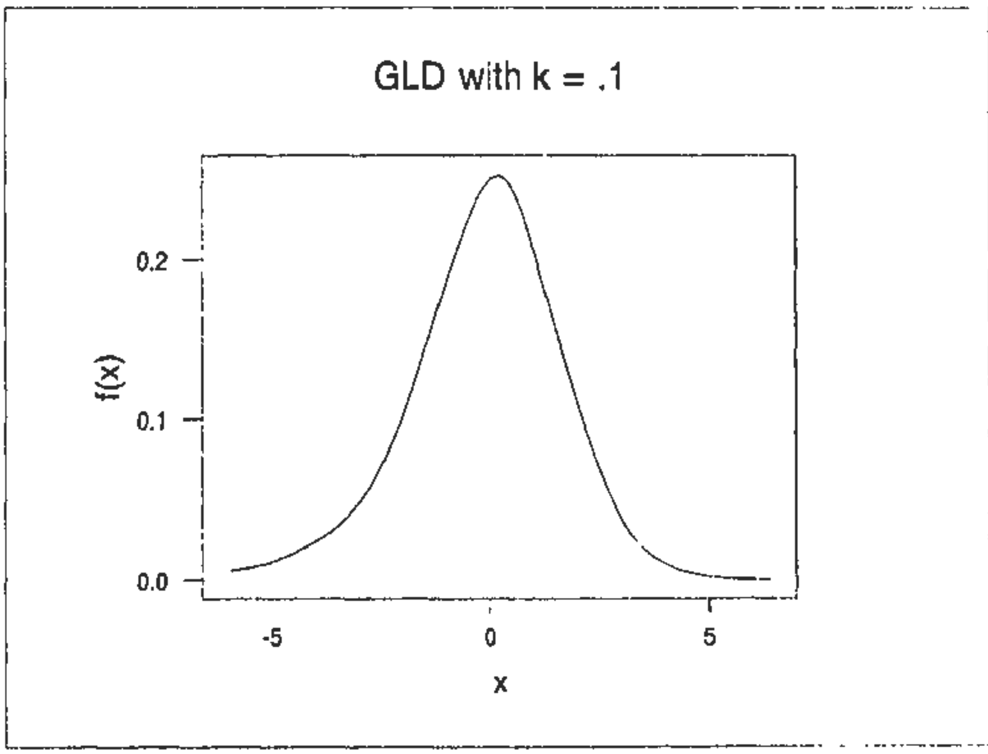

GLD with k = -.4
0.3
0.2
0.1
0.0
f(x)
-2
3
8
x

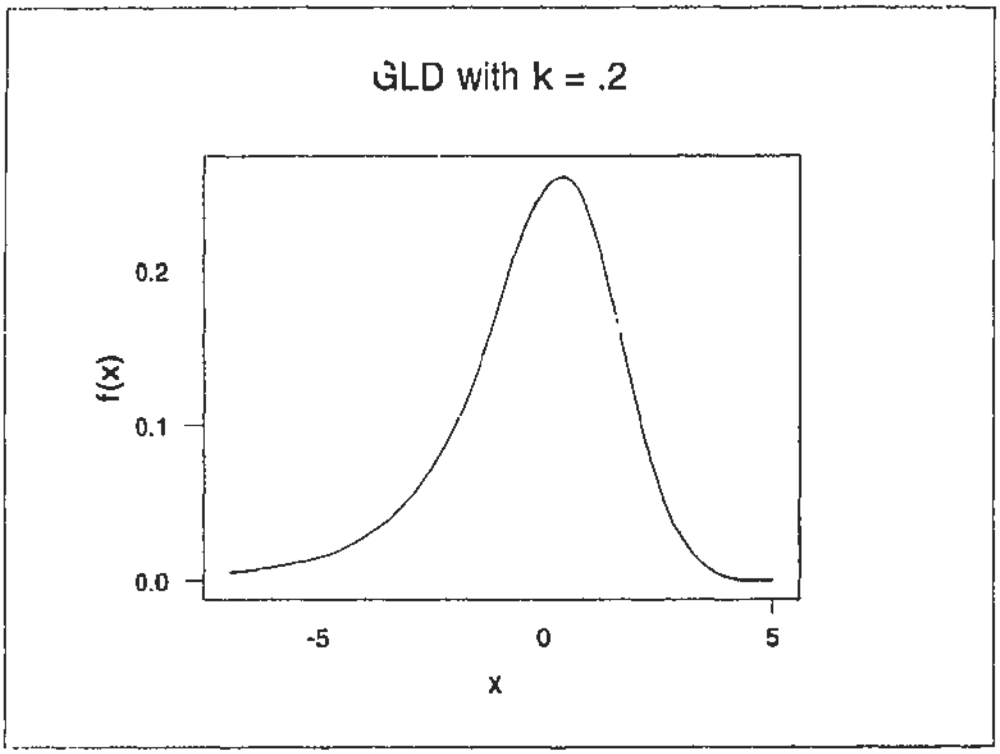

GLD with k = -.3
0.3
0.2
0.1
0.0
f(x)
-3
2
7
x

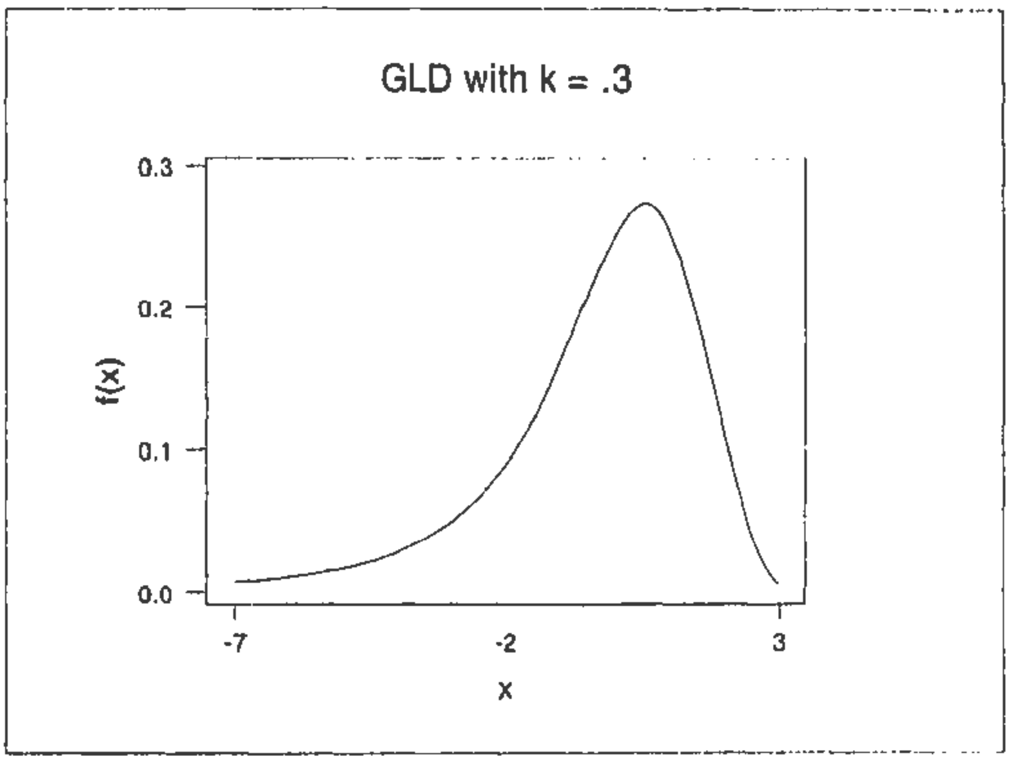

GLD with k = -.2
f(x)
0.0
0.1
0.2
-5
0
5
x

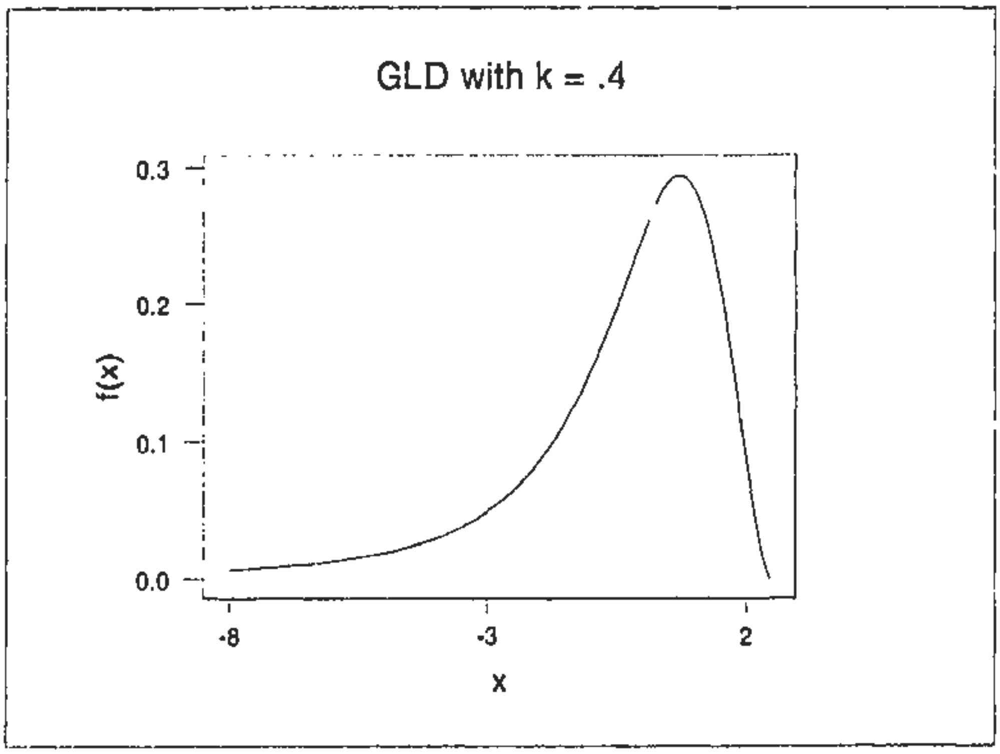

GLD with k = -.1
0.25
0.20
0.15
0.10
0.05
0.00
f(x)
-5
0
5
x


GLD with k = 0.0
0.25
0.20
0.15
0.10
0.05
0.00
f(x)
-5
0
5
x


GLD with k = .1
f(x)
0.0
0.1
0.2
-5
0
5
x


GLD with k = .2
f(x)
0.2
0.1
0.0
-5
0
5
x


GLD with k = .3
f(x)
0.3
0.2
0.1
0.0
-7
-2
3
x


GLD with k = .4
0.3
0.2
0.1
0.0
f(x)
-8
-3
2
x

Appendix D

MLE Program Algorithm

This appendix contains the algorithm of the program used in the Monte Carlo simulation study of Chapter 4.

PROGRAM ALGORITHM

(10,000 simulations.)

1. A random sample from the Generalized Logistic Distribution is generated. This is done by using the Inverse Cumulative Distribution given in Chapter 2. The values of ε and α are 0 and 1 respectively for reasons specified in Chapter 4, and the value of k is supplied by the user.

2. Subroutine is called to calculate L-Moment estimates.

3. Subroutine using numerical methods is called to calculate MLEs. Initial values for the iterative procedure are the true values used to generate the sample. In order to solve the three nonlinear equations,

$$\frac{\partial \ln L}{\partial \varepsilon} = 0, \qquad \frac{\partial \ln L}{\partial \alpha} = 0, \qquad \frac{\partial \ln L}{\partial k} = 0$$

the following steps are followed:

Step 1: Hold α and k fixed at initial values and solve first equation for ε using Bisection Method.

Step 2: Hold ε and k fixed; ε at the value found in Step 1, solve second equation for α using Newton's Method. If Newton's Method fails, Bisection Method is used.

Step 3: Hold ε and α fixed at values from Steps 1 and 2 and solve the third equation for k using the Bisection Method.

Step 4: Repeat Steps 1-3 using updated estimates for initial values. Continue until change in the three estimates in successive iterations <.001. If solution not found after 40 iterations, a failure of the algorithm is recorded.

Care is taken in the above procedure to insure that none of the estimates violate parameter restrictions imposed by the distribution at any point during the calculations.

4. Subroutine is called to calculate MOM estimates. Bisection Method is used to find estimate of k.

5. Bias and Root Mean Square of the estimates for each method are found by calculating these two quantities for the 10 000 estimates of each parameter.

Appendix E

Goodness of Fit Graphs for the Data

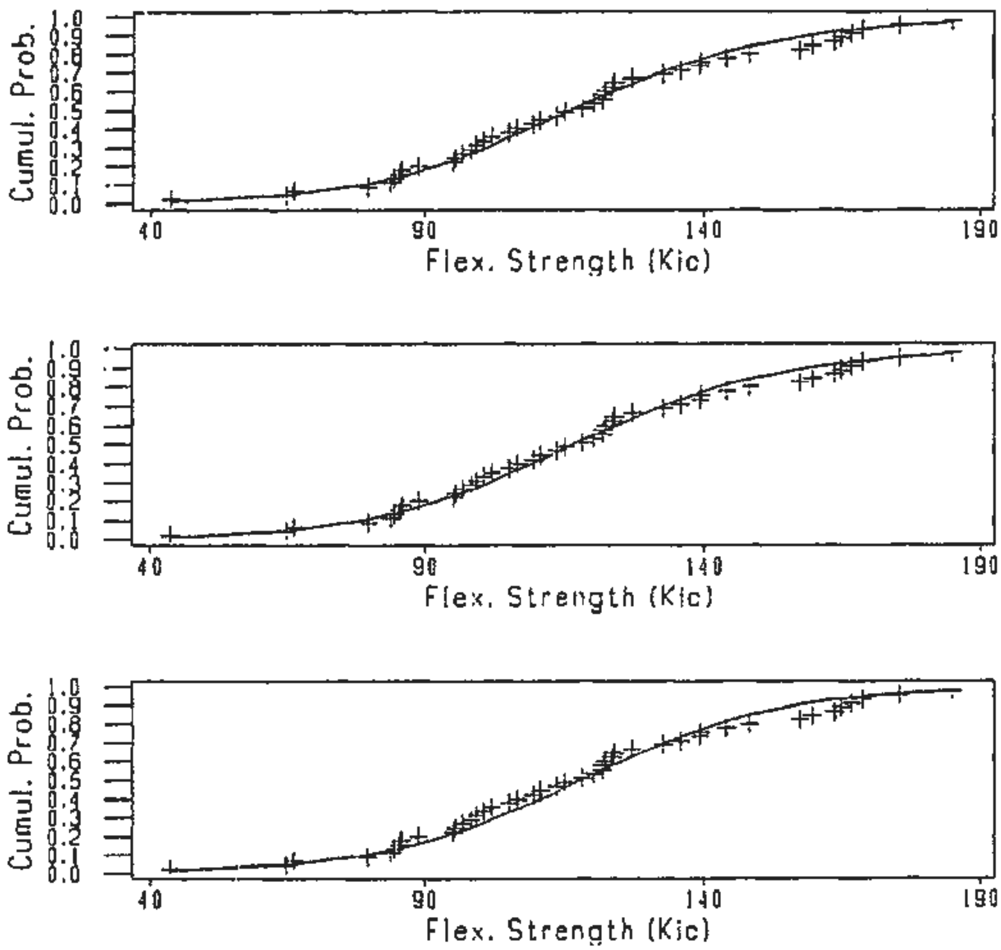
In this appendix, plots of the empirical distribution function (denoted by +), compared to the fitted distribution function, are given for each of the three methods used to fit the GLD model for the nine data sets. These plots allow a visual comparison of the goodness of fit of each of the the three fitted models. The plots are in the following order:
CDAT2, CDAT4, CDAT8, CDAT9, CDAT12, CDAT13, CDAT14, CDAT19, CDAT23.

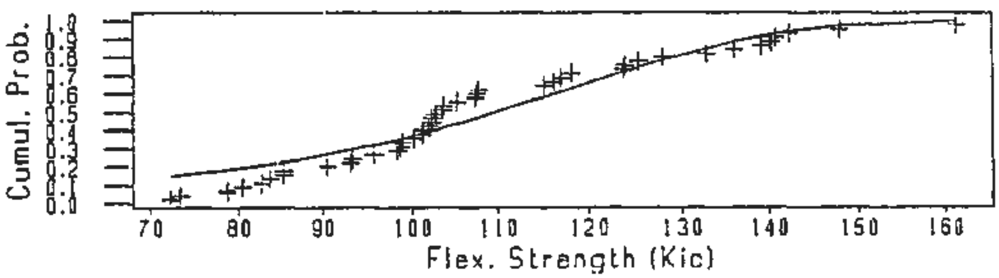
For each data set, the three fitted distribution functions are given in the order:

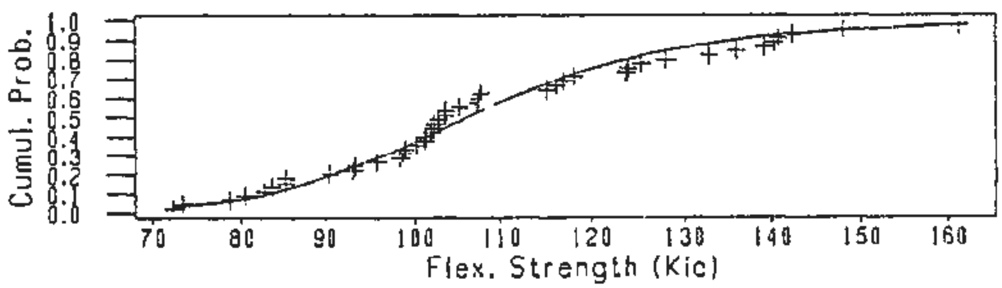
Fitted MLE

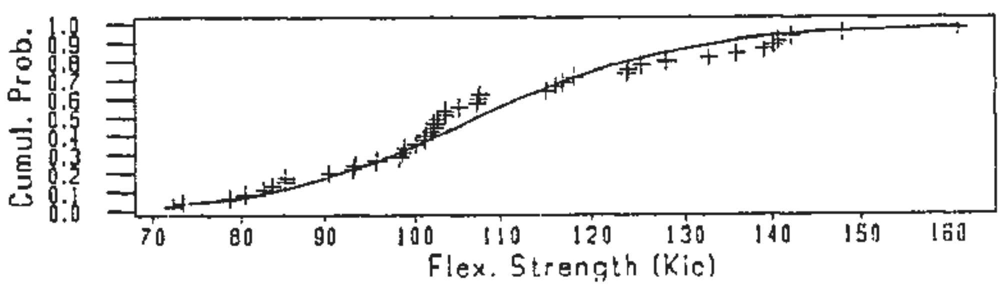
Fitted LMOM

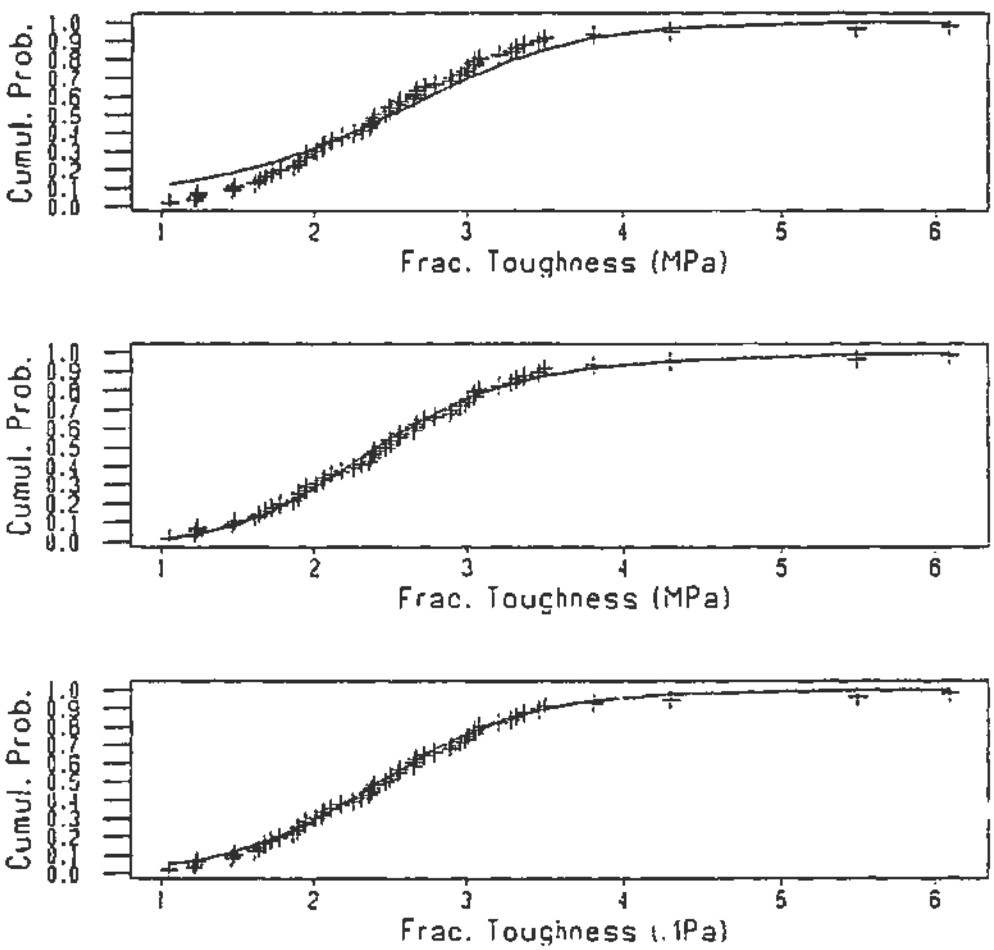
Fitted MOM


Cumul. Prob.
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
40
90
140
190
Flex. Strength (Kic)
Cumul. Prob.
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
40
90
140
190
Flex. Strength (Kic)
Cumul. Prob.
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
40
90
140
190
Flex. Strength (Kic)

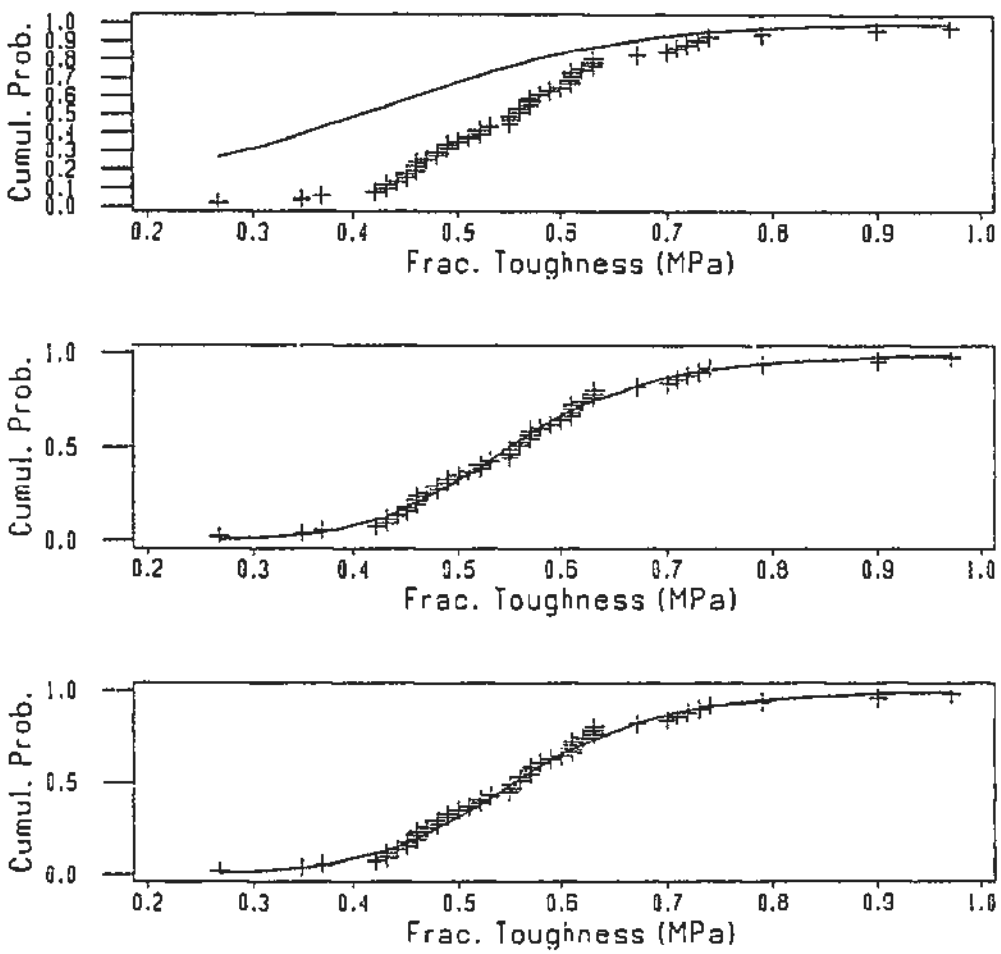

Cumul. Prob.
0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
70 80 90 100 110 120 130 140 150 160
Flex. Strength (Kic)

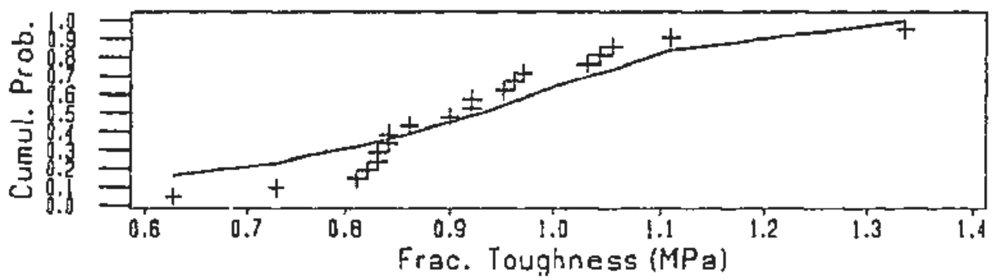

Cumul. Prob.
0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
70 80 90 100 110 120 130 140 150 160
Flex. Strength (Kic)

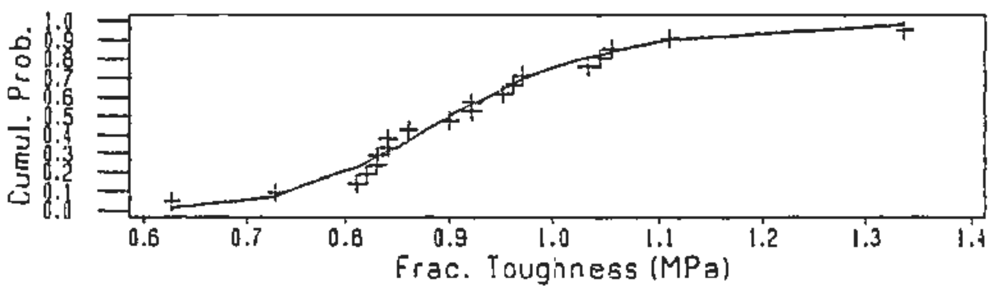

Cumul. Prob.
0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
70 80 90 100 110 120 130 140 150 160
Flex. Strength (Kic)

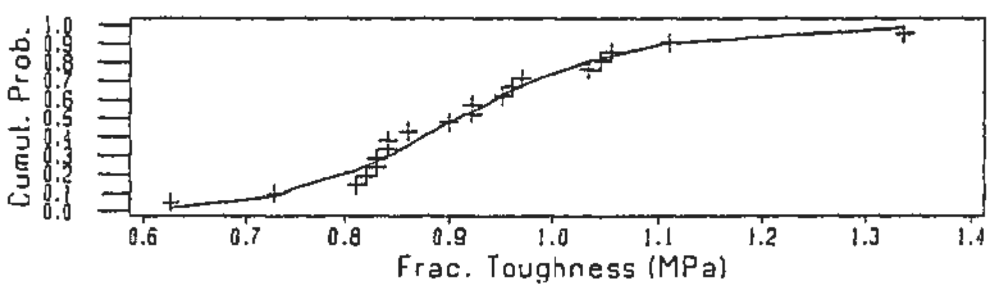


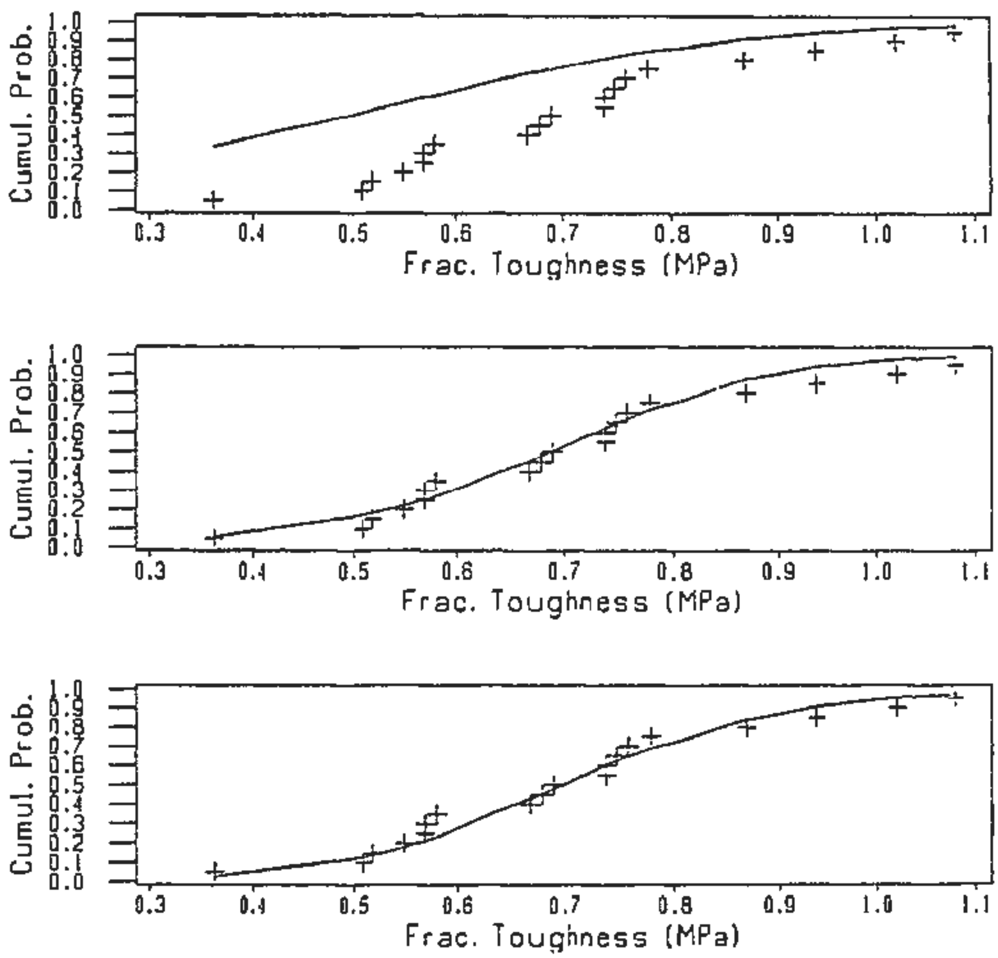
Cumul. Prob.
1
2
3
4
5
6
Frac. Toughness (MPa)
Cumul. Prob.
1
2
3
4
5
6
Frac. Toughness (MPa)
Cumul. Prob.
1
2
3
4
5
6

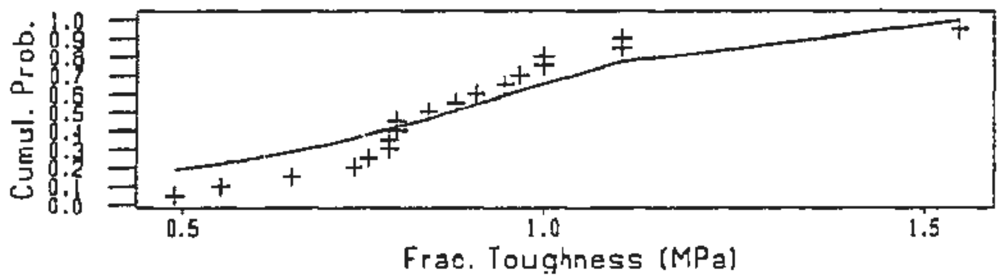

Cumul. Prob.
Frac. Toughness (MPa)
Cumul. Prob.
Frac. Toughness (MPa)
Cumul. Prob.
Frac. Toughness (MPa)

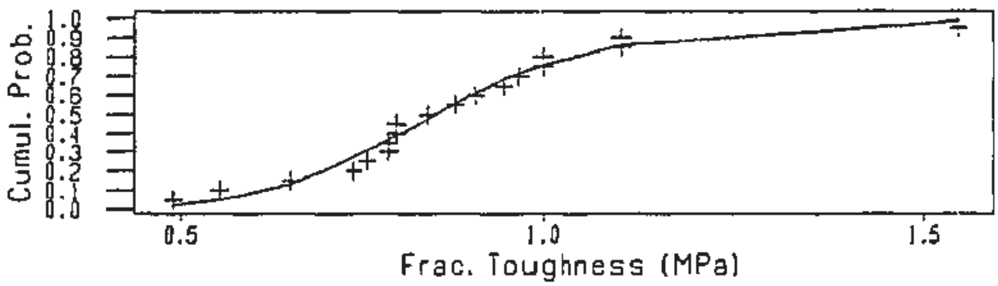

Cumul. Prob.
0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
0.6 0.7 0.8 0.9 1.0 1.1 1.2 1.3 1.4
Frac. Toughness (MPa)

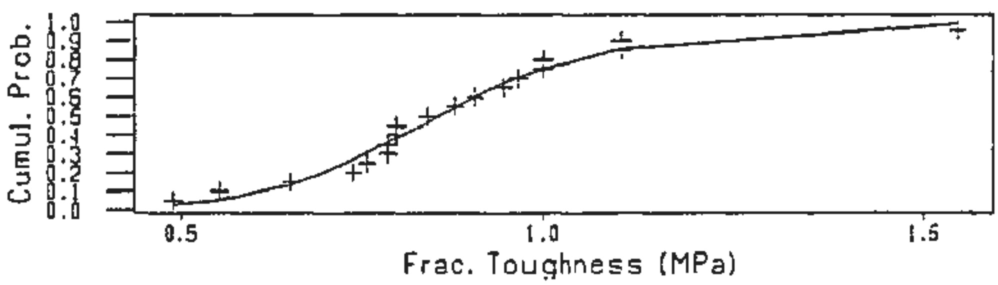

Cumul. Prob.
0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
0.6 0.7 0.8 0.9 1.0 1.1 1.2 1.3 1.4
Frac. Toughness (MPa)

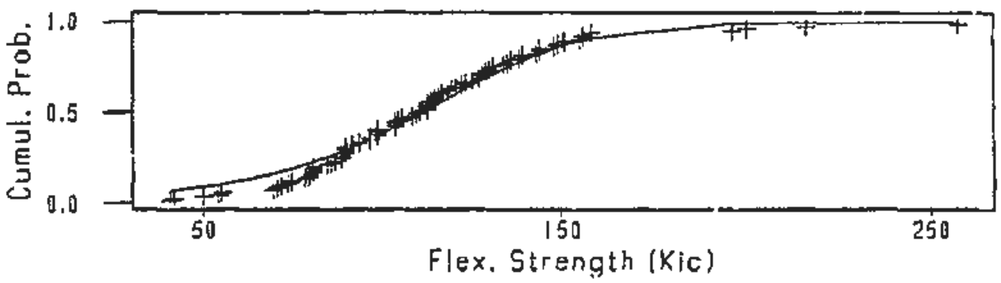

Cumul. Prob.
0.0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0
0.6 0.7 0.8 0.9 1.0 1.1 1.2 1.3 1.4
Frac. Toughness (MPa)

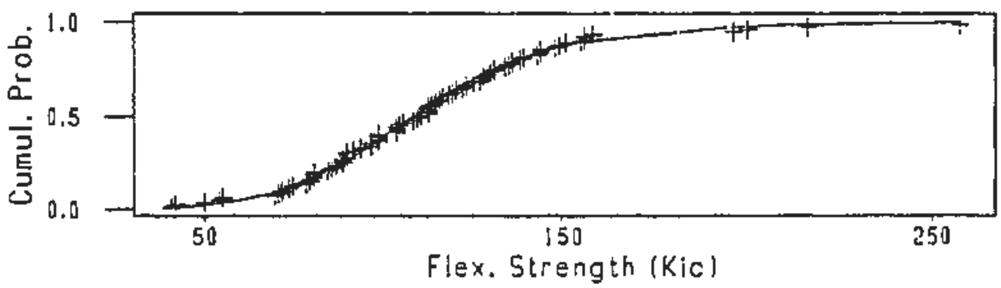

Cumul. Prob.
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
1.1
Frac. Toughness (MPa)

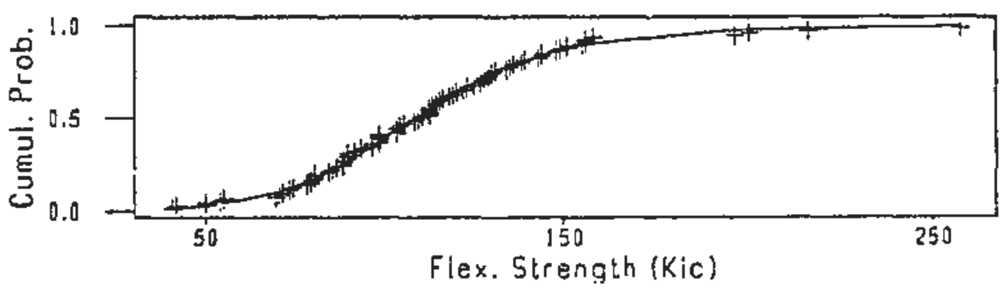

Cumul. Prob.
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
0.5
1.0
1.5
Frac. Toughness (MPa)

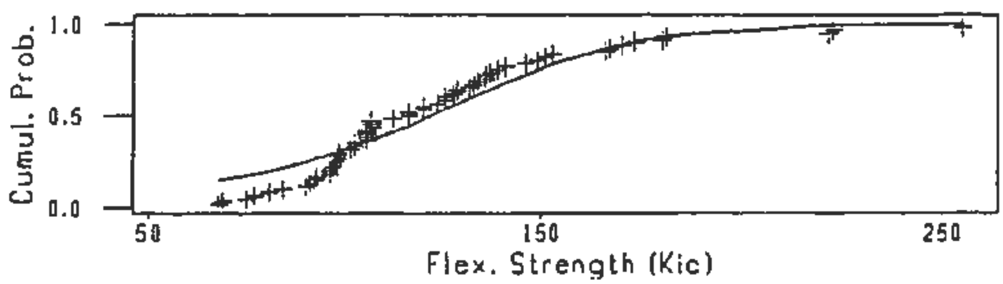

Cumul. Prob.
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
0.5
1.0
1.5
Frac. Toughness (MPa)

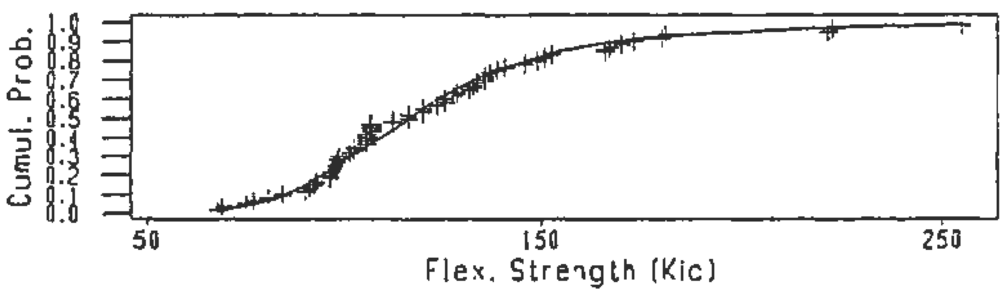

Cumul. Prob.
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
0.5
1.0
1.5
Frac. Toughness (MPa)

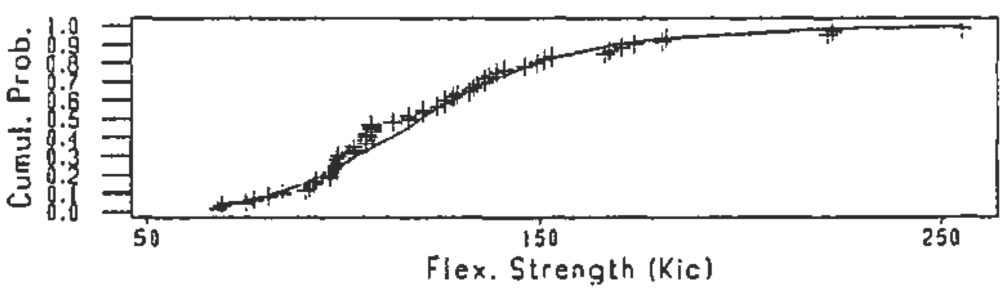

1.0
0.5
0.0
Cumul. Prob.
50
150
250
Flex. Strength (Kic)


1.0
0.5
0.0
Cumul. Prob.
50
150
250
Flex. Strength (Kic)


1.0
0.5
0.0
Cumul. Prob.
50
150
250
Flex. Strength (Kic)


Cumul. Prob.
1.0
0.5
0.0
50
150
250
Flex. Strength (Kic)


Cumul. Prob.
50
150
250
Flex. Strength (Kic)

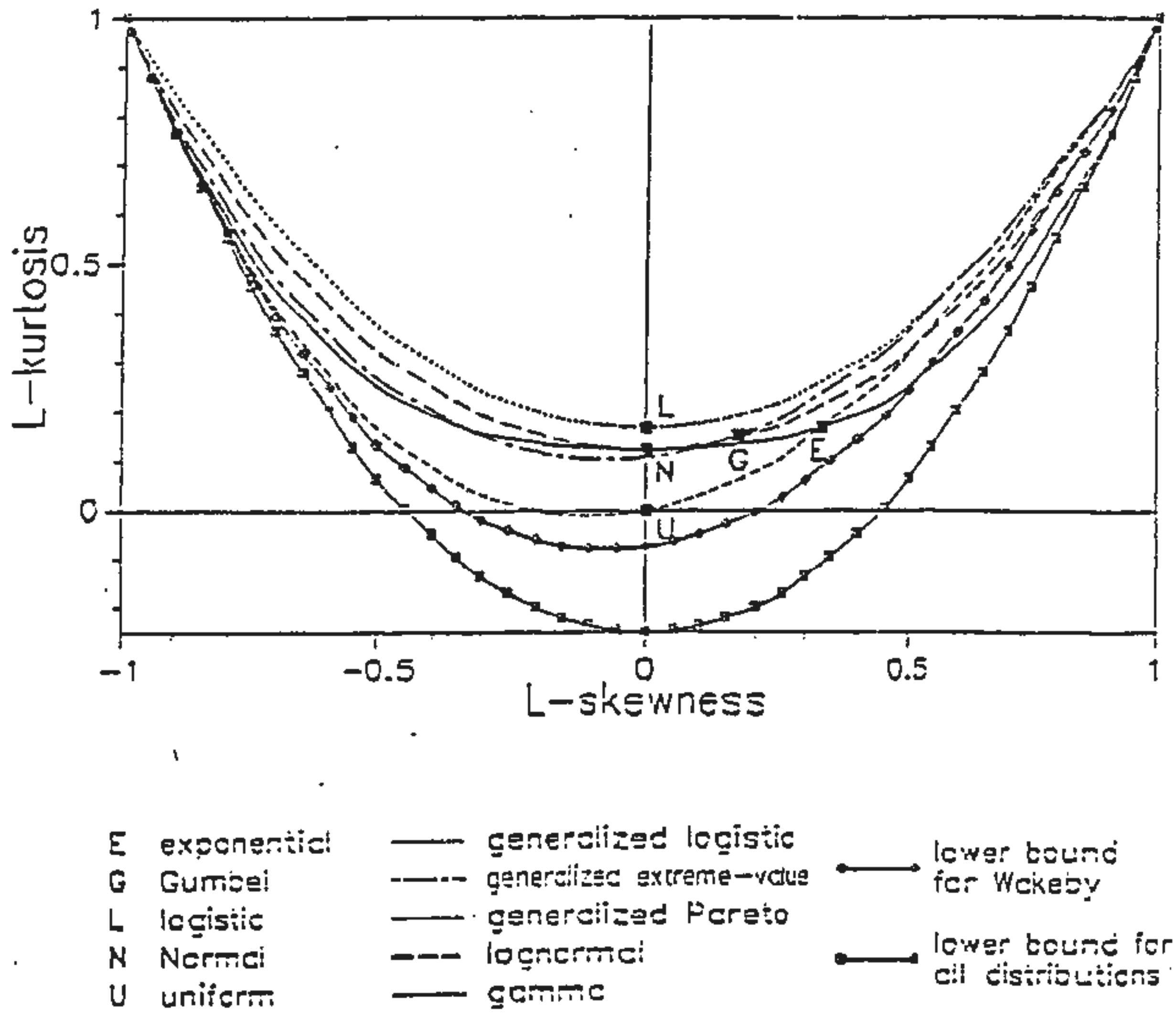

Cumul. Prob.
50
150
250
Flex. Strength (Kic)

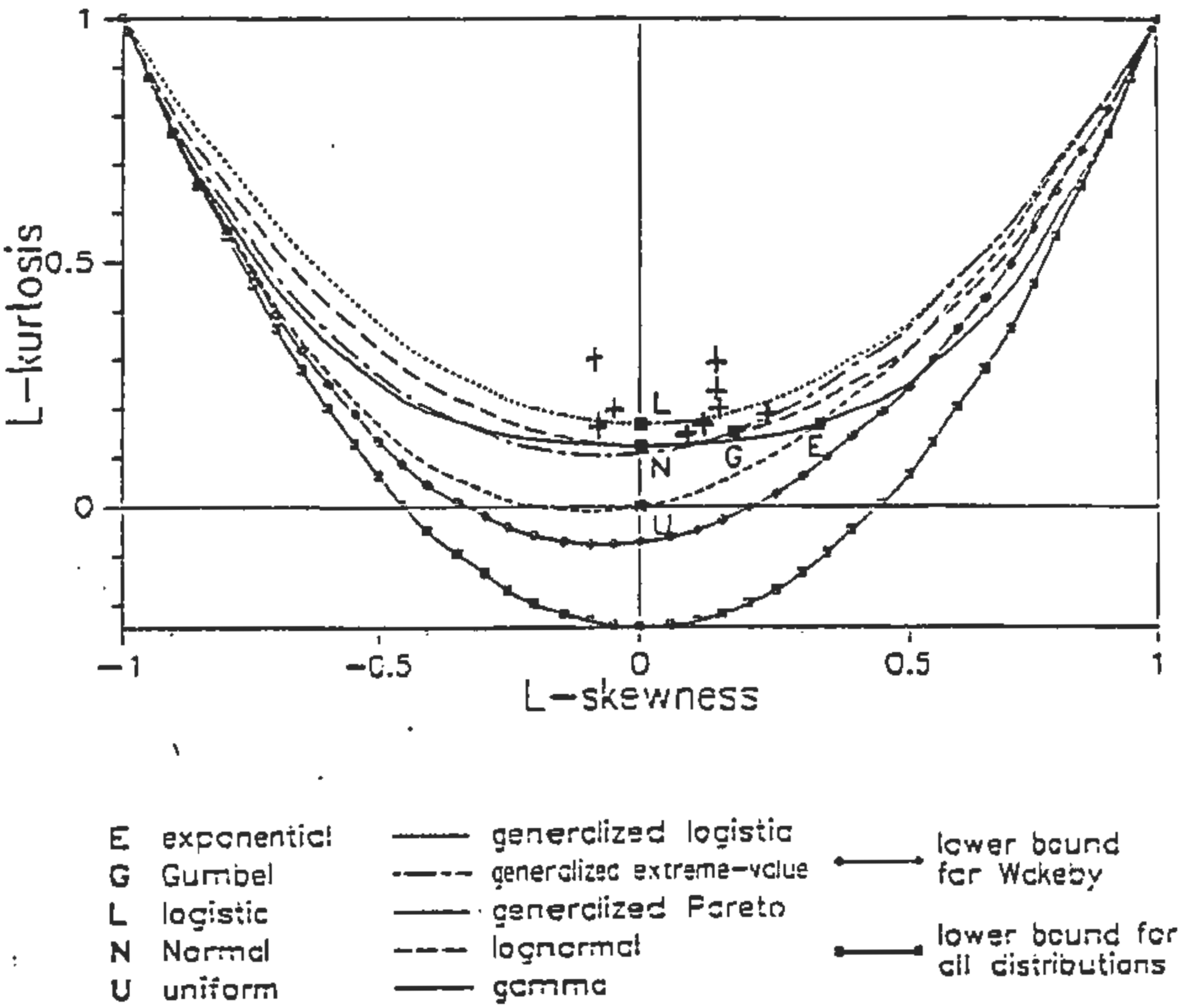
Appendix F

L-Moment Diagram

This appendix contains the L-Moment diagram (reprinted with permission of the author) as given in Hosking [8]. This diagram is used for choosing a suitable distribution to model the data. For a given set of data, the sample L-skewness and L-kurtosis are calculated and plotted on the diagram. The lines on the diagram represent L-kurtosis as a function of L-skewness for a number of common hydrological and engineering distributions. The distribution which best describes the data is determined by the line which is closest to the plotted point.

Two copies of the diagram are given — the first is a reproduction from Hosking [8], and the second one contains the plotted L-skewness vs. L-kurtosis points for each of the nine data sets analyzed in Chapter 5. From the position of the points on the graph, we see that these data sets are suited for modeling by the GLD.


L-kurtosis
L-skewness
1
0.5
0
-1
-0.5
0
0.5
1
L
N
G
E
U
E exponential
G Gumbel
L logistic
N Normal
U uniform
generalized logistic
generalized extreme-value
generalized Pareto
lognormal
gamma
lower bound for Wakeby
lower bound for all distributions


1
0.5
0
L-kurtosis
-1
-0.5
0
0.5
1
L-skewness
L
N
G
E
U
E exponential
G Gumbel
L logistic
N Normal
U uniform
generalized logistic
generalized extreme-value
generalized Pareto
lognormal
gamma
lower bound for Wakeby
lower bound for all distributions

Appendix G

Box-Plots of the Data

This appendix contains some summary statistics and box-plots of the data analyzed in this practicum. For each of the nine data sets, the mean, standard deviation (St. Dev.) and skewness is given, then a box-plot is given to illustrate the distribution of the observations.

A box-plot consists of a box, whiskers, and outliers. A line is drawn across the box at the median. The bottom of the box is at the first quartile (Q1) and the top is at the third quartile (Q3). Thus the box contains roughly 50% of the observations. The whiskers are the lines that extend from the top and bottom of the box to the adjacent values, the lowest and highest observations still inside the region defined by the lower limit Q1 - 1.5 (Q3 - Q1) and the upper limit Q1 + 1.5 (Q3 - Q1). Outliers are points outside the lower and upper limits, plotted with asterisks (*).

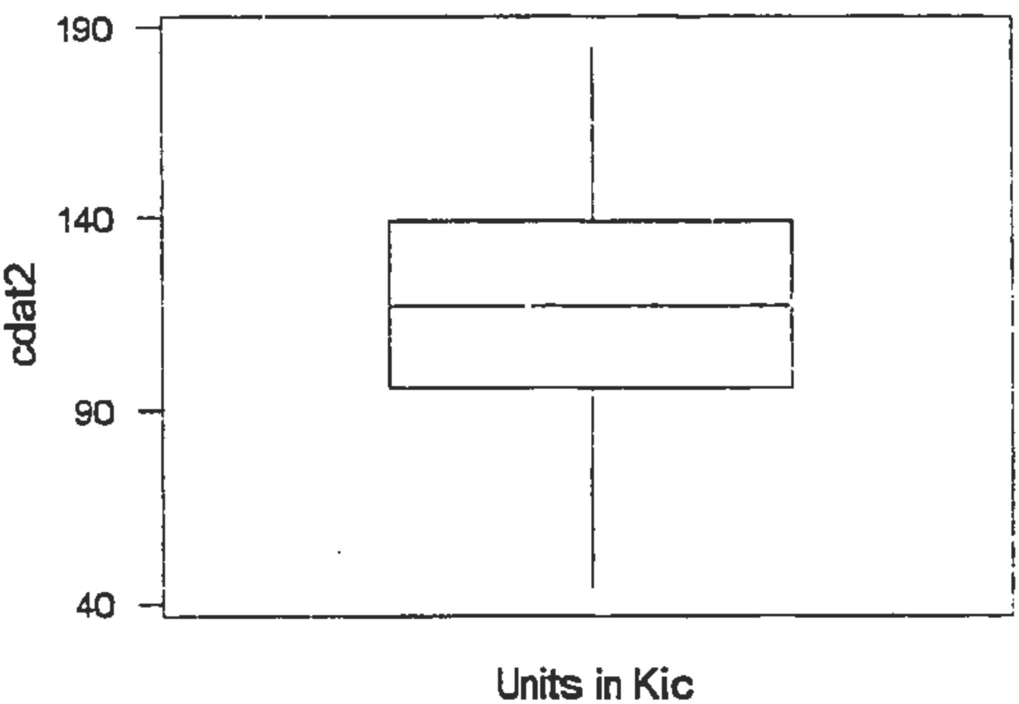
CDAT2: ($n = 44$)

Mean : 118.06

St. Dev. : 32.02

Skewness : 0.10

Box-Plot



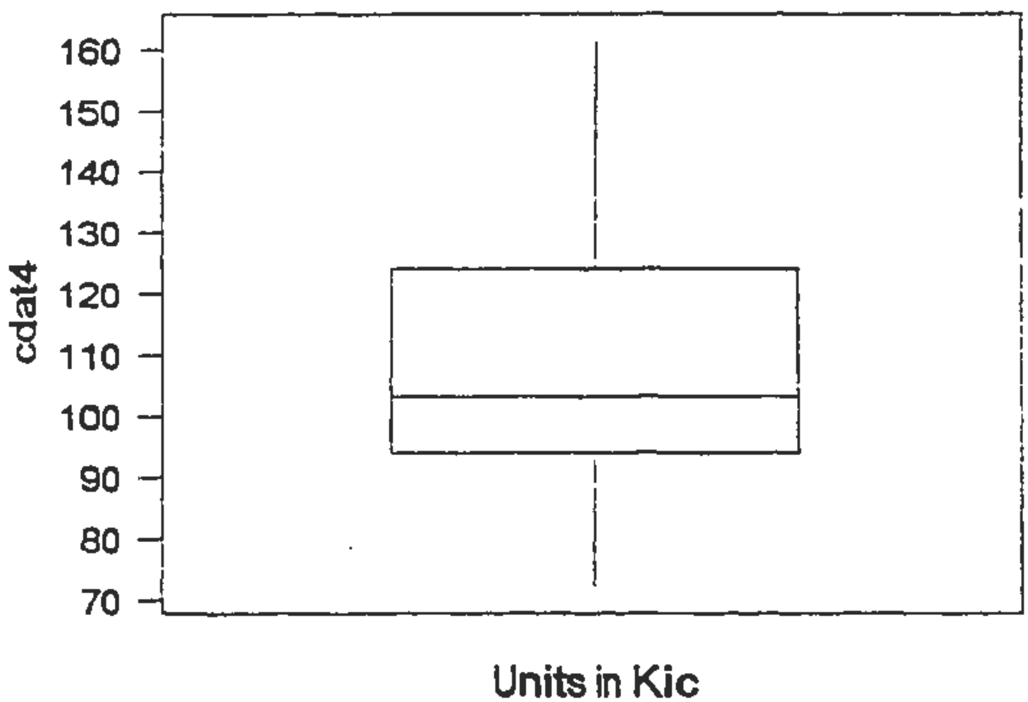
<u>CDAT4:</u> ($n = 44$)

Mean : 108.31

St. Dev. : 21.27

Skewness : 0.49

<u>Box-Plot</u>



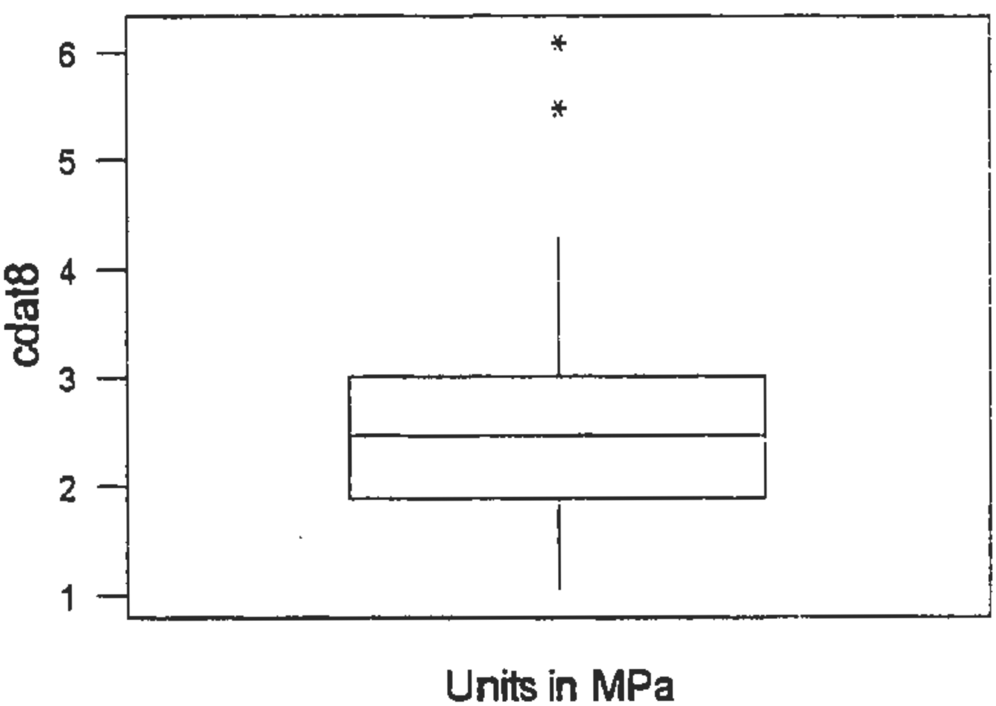
CDAT8: ($n = 55$)

Mean : 2.54

St. Dev. : 0.95

Skewness : 1.43

Box-Plot



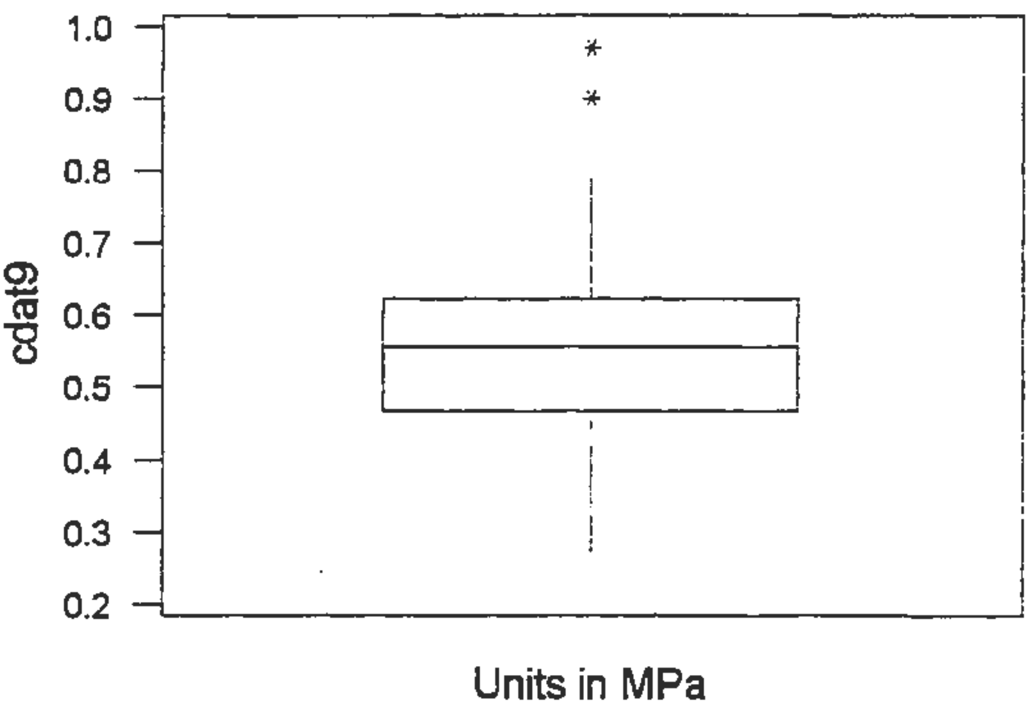
CDAT9: ($n = 50$)

Mean : 0.56

St. Dev. : 0.13

Skewness: 0.77

Box-Plot



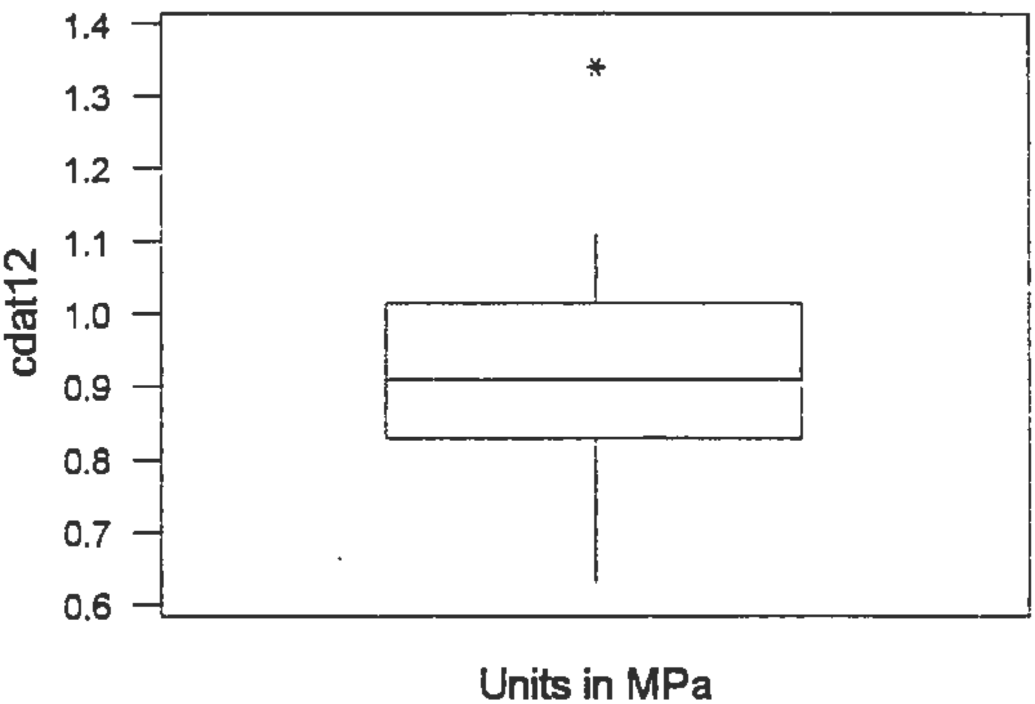
CDAT12: ($n = 20$)

Mean : 0.92

St. Dev. : 0.15

Skewness: 0.86

Box-Plot



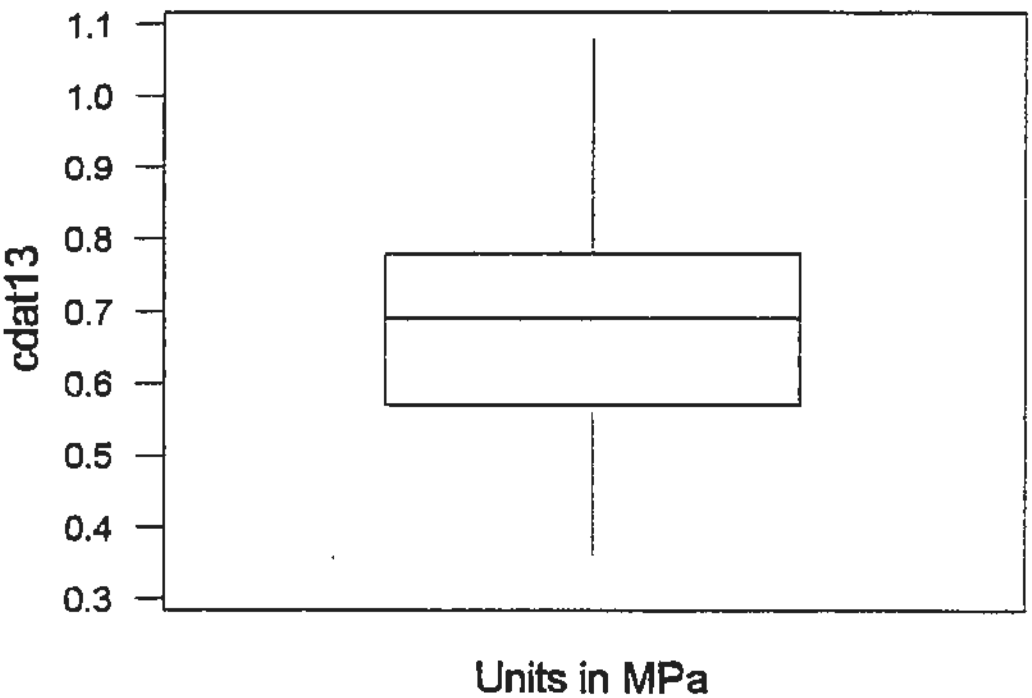
<u>CDAT13:</u> ($n = 19$)

Mean : 0.70

St. Dev. : 0.18

Skewness: .38

<u>Box-Plot</u>



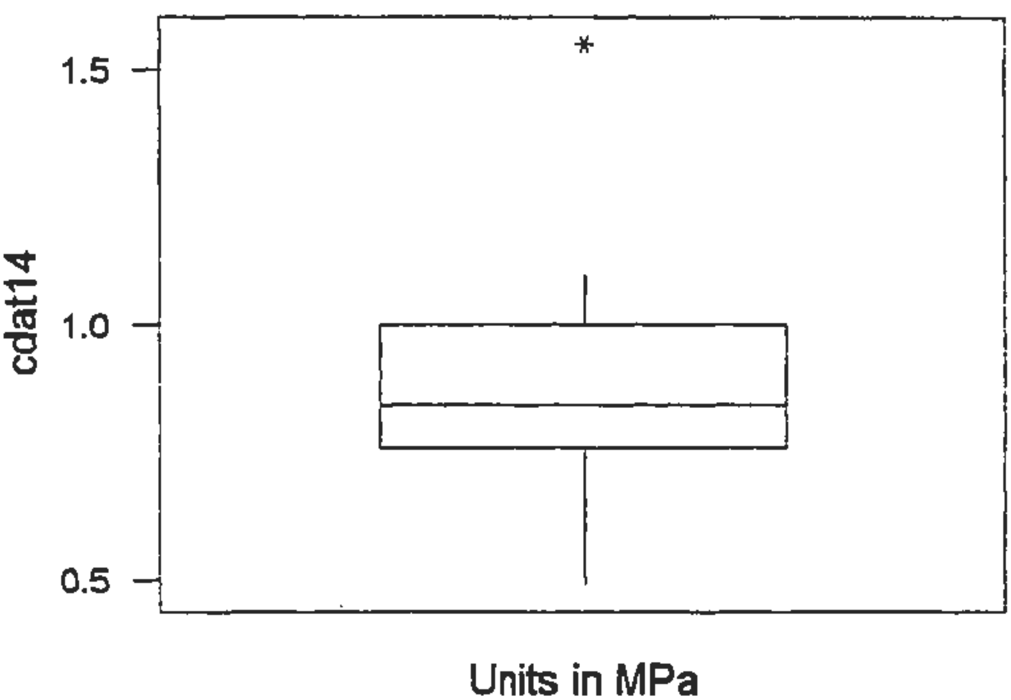
CDAT14: ($n = 19$)

Mean : 0.88

St. Dev. : 0.23

Skewness: 1.10

Box-Plot



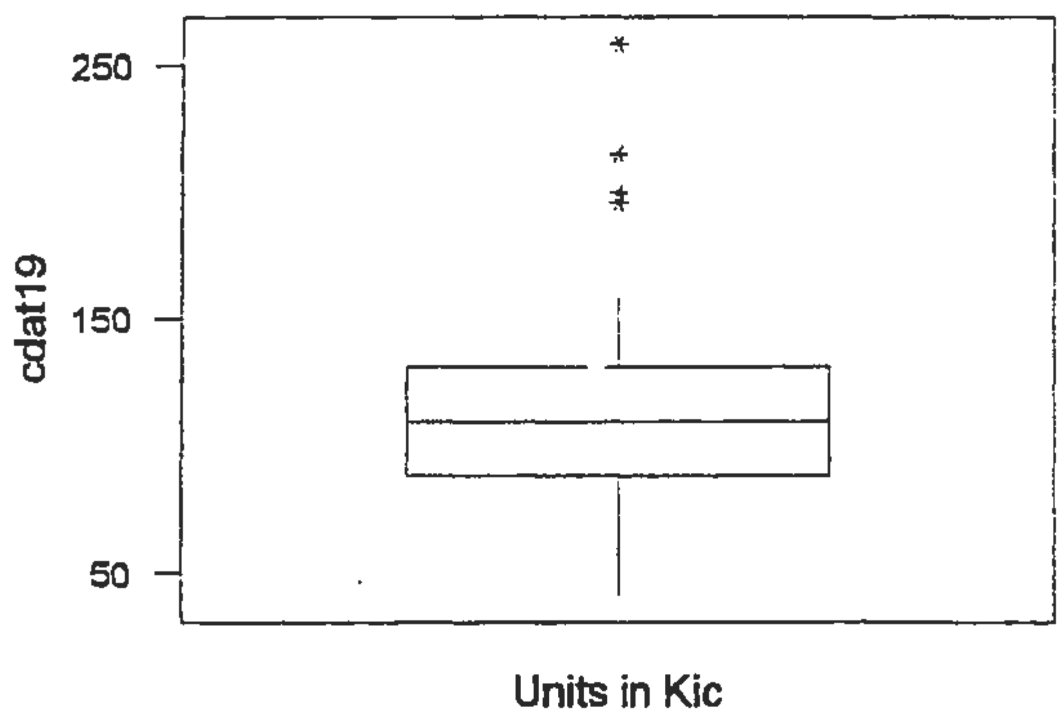
CDAT19: ($n = 80$)

Mean : 111.65

St. Dev. : 37.26

Skewness: 1.06

Box-Plot



<u>CDAT23:</u> ($n = 59$)

Mean : 123.62

St. Dev. : 37.71

Skewness: 1.31

<u>Box-Plot</u>