ENTROPY-BASED PARAMETER ESTIMATION IN HYDROLOGY

Water Science and Technology Library

VOLUME 30

The titles published in this series are listed at the end of this volume.

ENTROPY-BASED PARAMETER ESTIMATION IN HYDROLOGY

by

VIJAY P. SINGH
Water Resources Program,
Department of Civil and Environmental Engineering,
Louisiana State University,
Baton Rouge
U.S.A.

SPRINGER-SCIENCE+BUSINESS MEDIA, B.V.

A C.I.P. Catalogue record for this book is available from the Library of Congress.

ISBN 978-90-481-5089-2 ISBN 978-94-017-1431-0 (eBook)
DOI 10.1007/978-94-017-1431-0

Printed on acid-free paper

To:

Anita

Vinay, and

Arti

TABLE OF CONTENTS

PREFACE

Since the pioneering work of Shannon in the late 1940's on the development of the theory of entropy and the landmark contributions of Jaynes a decade later leading to the development of the principle of maximum entropy (POME), the concept of entropy has been increasingly applied in a wide spectrum of areas, including chemistry, electronics and communications engineering, data acquisition and storage and retreival, data monitoring network design, ecology, economics, environmental engineering, earth sciences, fluid mechanics, genetics, geology, geomorphology, geophysics, geotechnical engineering, hydraulics, hydrology, image processing, management sciences, operations research, pattern recognition and identification, photogrammetry, psychology, physics and quantum mechanics, reliability analysis, reservoir engineering, statistical mechanics, thermodynamics, topology, transportation engineering, turbulence modeling, and so on. New areas finding application of entropy have since continued to unfold. The entropy concept is indeed versatile and its applicability widespread.

In the area of hydrology and water resources, a range of applications of entropy have been reported during the past three decades or so. This book focuses on parameter estimation using entropy for a number of distributions frequently used in hydrology. In the entropy-based parameter estimation the distribution parameters are expressed in terms of the given information, called constraints. Thus, the method lends itself to a physical interpretation of the parameters. Because the information to be specified usually constitutes sufficient statistics for the distribution under consideration, the entropy method provides a quantitative way to express the information contained in the distribution. Furthermore, it also provides a way to derive a distribution from specified constraints. In other words, the hypotheses underlying the distribution have an easy interpretation. These advantages notwithstanding, the entropy-based parameter estimation has received comparatively little attention from the hydrologic community. It is hoped that this book will stimulate interest in this fascinating area of entropy and its application in hydrology, environmental engineering, and water resources.

The subject matter of the book spans 22 chapters. The entropy concept and the principle of maximum entropy are introduced in Chapter 1. Also introduced therein are the entropy-based method of parameter estimation and parameter-space expansion method. The chapter is concluded with a brief account of the use of entropy as a criterion for goodness of fit and the dependence of entropy on the sample size. A short discussion of five other popular methods of parameter estimation, including the methods of moments, probability weighted moments, L-moments, maximum likelihood estimation, and least squares, is presented in Chapter 2. Also presented there is a brief account of errors and statistical measures of performance evaluation, including bias, consistency, efficiency, sufficiency, resilience, standard error, root mean square error, robustness, and relative mean error.

The next two chapters present the base distributions-uniform and exponential. The uniform distribution is discussed in chapter 3. Beginning with specification of the constraint, the chapter goes on to discuss construction of the zeroth Lagrange multiplier, estimation of the parameter, and the entropy of the distribution. Chapter 4 presents parameter estimation for the exponential distribution. The discussion on the entropy-based parameter estimation method is divided into the ordinary entropy method and the parameter-space expansion method. The first method includes specification of constraints, construction of the zeroth Lagrange multiplier, relation between Lagrange multipliers and constraints, relation between Lagrange multipliers and parameter, relation between parameter and constraint and distribution entropy. The second method discusses specification of constraints, derivation of the entropy function and the relation between distribution parameter and constraint. The chapter is concluded with a discussion of the methods of moments, maximum likelihood estimation, probability weighted moments and L-moments.

The three succeeding chapters are devoted to normal and lognormal distributions. The organization of presentation, comprising both the ordinary entropy method and the parameter space expansion method, is similar in each chapter. The ordinary entropy method includes specification of constraints, construction of the zeroth Lagrange multiplier, relation between Lagrange multipliers and constraints, relation between Lagrange multipliers and parameters, relation between parameters and constraints, and distribution entropy. The parameter space expansion method includes specification of constraints, derivation of the entropy function, and relation between distribution parameters and constraints. Chapter 5, devoted to parameter estimation for the normal distribution, is concluded with a discussion of the methods of moments, maximum likelihood estimation and probability weighted moments. Chapter 6 treats the two-parameter lognormal distribution and is concluded with a short account of the methods of moments, maximum likelihood estimation and probability weighted moments. Chapter 7 is devoted to parameter estimation for the three-parameter lognormal distribution. In addition to discussing the methods of moments (regular and modified), maximum likelihood estimation (regular and modified) and probability weighted moments, the chapter is concluded with a comparative evaluation of these parameter estimation methods, including the entropy method, using Monte Carlo experimentation and field data.

The next four chapters are devoted to the extreme-value distributions. The organization of discussion of the entropy-based parameter estimation is the same as in the preceding chapters. Chapter 8 discusses parameter estimation for extreme-value type I or Gumbel distribution. The other methods of parameter estimation included in the chapter are the methods of moments, maximum likelihood estimation, least squares, incomplete means, probability weighted moments, and mixed moments. The chapter is concluded with a comparative evaluation of the parameter estimation methods using both field data and Monte Carlo simulation experiments. Chapter 9 discusses parameter estimation for log-extreme-value type I or log-Gumbel distribution. It goes on to discuss the methods of moments and maximum likelihood estimation, and is concluded with a comparative evaluation of the parameter estimation methods using annual flood data. Chapter 10 discusses parameter estimation for extreme-value type III distribution. It also includes a discussion of the methods of moments and maximum likelihood estimation and a comparative evaluation of the parameter estimation methods. Chapter 11 discusses parameter estimation for the generalized extreme-value distribution, and is concluded with a discussion of the methods of moments, probability weighted moments, L-moments and maximum likelihood estimation.

The next five chapters discuss parameter estimation for the Weibull distribution and Pearsonian distributions. The organization of presentation of the entropy-based parameter estimation is the same as in the preceding chapters. Chapter 12 discusses parameter estimation for the Weibull distribution. It goes on to discuss the methods of moments, maximum likelihood estimation, probability weighted moments and least squares, and is concluded with a comparative evaluation of the methods using rainfall data and Monte Carlo experiments. Chapter 13 discusses parameter estimation for gamma distribution. It also presents a discussion of the methods of moments, cumulants, maximum likelihood estimation, least squares, and probability weighted moments. By applying these parameter estimation methods to unit hydrograph estimation and flood frequency analysis, the chapter is ended with a comparative assessment of the methods. Chapter 14 discusses parameter estimation for Pearson type III distribution. It also includes a discussion of the methods of moments, maximum likelihood estimation and probability weighted moments; as well as a comparative assessment of the estimation methods using annual maximum discharge data. Chapter 15 discusses parameter estimation for log-Pearson type III distribution. It also includes a treatment of the methods of moments (direct as well as indirect), maximum likelihood estimation, and mixed moments. Finally, it presents a comparative evaluation of the estimation methods using field data as well as Monte Carlo experiments. Chapter 16 discusses parameter estimation for beta distribution (or Pearson type I) distribution and is concluded with a discussion of the methods of moments and maximum likelihood estimation.

The next two chapters present two and three parameter log-logistic distribution. The organization of presentation of the entropy-based parameter estimation is similar to that of the preceding chapters. Chapter 17 discusses parameter estimation for two-parameter log-logistic distribution. It also treats the methods of moments, probability weighted moments and maximum likelihood estimation; and is concluded with a comparative evaluation of the estimation methods using Monte Carlo simulation experiments. Chapter 18 discusses parameter estimation for three-parameter log-logistic distribution. It also presents the methods of moments, maximum likelihood estimation and probability weighted moments; as well as a comparative assessment of the parameter estimation methods.

The next three chapters present Pareto distributions. The organization of presentation of the entropy-based parameter estimation is the same as in the preceding chapters. Chapter 19 discusses parameter estimation for the 2-parameter Pareto distribution. It also discusses the methods of moments, maximum likelihood estimation and probability weighted moments. The chapter is ended with a comparative assessment of parameter estimation methods using Monte Carlo simulation. Chapter 20 discusses parameter estimation for the 2-parameter generalized Pareto distribution. It also discusses the methods of moments, maximum likelihood estimation and probability weighted moments, and presents a comparative assessment of the parameter estimation methods. Chapter 21 discusses parameter estimation for the 3-parameter generalized distribution. It also includes a discussion of the methods of moments, maximum likelihood estimation and probability weighted moments; as well as a comparative evaluation of the parameter estimation methods.

Chapter 22 discusses parameter estimation for the two-component extreme-value distribution. The first method is the ordinary entropy method that includes a discussion of specification of constraints, construction of zeroth Lagrange multiplier, relation between parameters and constraints, and estimation of parameters, including point estimation and regional estimation. The chapter also includes a discussion of the methods of maximum likelihood estimation and probability weighted moments, and is concluded with a comparative evaluation of the parameter estimation methods.

V.P. Singh
Baton Rouge, Louisiana

ACKNOWLEDGMENTS

Much of the material reported in the book was developed while I was working on research projects funded by various organizations, including the United States Geological Survey, Louisiana Water Resources Research Institute, and the U.S. Department of the Army, Corps of Engineers. I worked with several of my graduate students and colleagues on entropy-related projects. Of particular mention are: Dr. A. K. Rajagopal, Naval Research Laboratory, with whom I worked for about 4 years and our collaborative work led to the development of the parameter-space expansion method; Professor M. Fiorentino and Dr. P. Claps of the Department of Environmental Engineering and Physics of the University of Basilicata at Potenza, Italy, with whom I worked on the two-component extreme value distribution and geomorphologic laws; Dr. Donald Barbe of the Department of Civil and Environmental Engineering, University of New Orleans, Louisiana, and Professor J.F. Cruise of the Department of Civil and Environmental Engineering, University of Alabama, Huntsville, Alabama, with whom I worked on hydraulic applications of entropy. Messrs. K. Arora, S.K. Jain, K. Singh, and H. Guo; Dr. P.F. Krstanovic, and Dr. F.X. Yu were some of my students who helped me continue to work on entropy. I am grateful to all of these individuals. Indeed this book draws significantly from the fruits of their labor.

My family (wife, Anita; son, Vinay; and daughter, Arti) allowed me to work during evenings, weekends, and holidays. They supported me and were always there when I needed them. Without their love and support this book would not have been completed. As a small token of my gratitude, I dedicate the book to them.

CHAPTER 1

ENTROPY AND PRINCIPLE OF MAXIMUM ENTROPY

Clausius coined the term 'entropy' from the Greek meaning transformation. Thus, entropy originated in physics and occupies an exceptional position among physical quantities. It does not appear in the fundamental equations of motion. Its nature is, rather, a statistical or probabilistic one, for it can be interpreted as a measure of the amount of chaos within a quantum mechanical mixed state. It is an extensive property like mass, energy, volume, momentum, charge, number of atoms of chemical species, etc., but, unlike these quantities, it does not obey a conservation law. Entropy is not an observable; rather it is a function of state. For example, if the state is described by the density matrix, its entropy is given by the van Neumann formula. In physical sciences, entropy relates macroscopic and microscopic aspects of nature and determines the behavior of macroscopic systems in equilibrium.

Entropy can be considered as a measure of the degree of uncertainty or disorder associated with a system. Indirectly it also reflects the information content of space-time measurements. Entropy is viewed in three different but related contexts and is hence typified by three forms: thermodynamic entropy, statistical-mechanical entropy, and information-theoretical entropy. In environmental and water resources, the most frequently used form is the information-theoretical entropy. Shannon (1948a, b) developed the theory of entropy and Jaynes (1957a, b) the principle of maximum entropy. The works of Shannon and Jaynes form the basis for a wide range of applications of entropy in recent years in hydrology and water resources. Singh and Rajagopal (1987) discussed advances in application of the principle of maximum entropy (POME) in hydrology. Rajagopal et al. (1987) presented new perspectives for potential applications of entropy in water resources research. Singh (1989) reported on hydrologic modeling using entropy. A historical perspective on entropy applications in water resources was presented by Singh and Fiorentino (1992). Harmancioglu et al. (1992b) discussed the use of entropy in water resources. Alpaslan et al. (1992) discussed the role of entropy, and Harmancioglu et al. (1992a) its application in design and evaluation of water quality monitoring networks. Singh (1997) reported on the use of entropy in hydrology and water resources. These surveys have discussed the state-of-art of entropy-based modeling in environmental and water resources. One of the most useful applications of entropy is parameter estimation. Before discussing this particular application, a brief discussion of entropy theory and POME is in order.

1.1 Entropy Theory

We consider a random variable whose behavior is described by a probability distribution. There is some uncertainty associated with this distribution and, for that matter, with any distribution employed to describe the random variable. The concept of entropy provides a quantitative measure of this uncertainty. To that end, consider a probability density function (pdf) f(x) associated with a dimensionless random variable X. The dimensionless random variable may be constructed by dividing the observed quantities by its mean value, e.g., annual flood maxima divided by mean annual flood. As usual, f(x) is a positive function for every x in some interval (a, b) and is normalized to unity such that

$$\int_a^b f(x)\, dx = 1 \tag{1.1}$$

We often make a change of variable from X to Z, based on physical or mathematical considerations, as

$$X = W(Z) \tag{1. 2}$$

where W is some function. Under such a transformation, quite generally we have the mapping

$$X\colon (a,b) \to Z\colon (L,U) \tag{1.3}$$

where a = W(L) and b = W(U). Thus, L and U stand for lower and upper limits in the Z-variable. Then,

$$\begin{aligned} f(x)\, dx &= f(x = W(z))\, \left|\frac{dx}{dz}\right| dz \\ &\cong g(z)\, dz \end{aligned} \tag{1.4}$$

in which

$$g(z) = f(x = W(z))\, \left|\frac{dx}{dz}\right| \tag{1.5}$$

Here g(z) is, again, a pdf but is in the Z-variable, and has positivity as well as normalization properties:

$$\int_L^U g(z)\, dz = 1 \tag{1.6}$$

Often f (x) is not known beforehand, although some of its properties (or constraints) may be known, e.g., moments, lower and upper bounds, etc. These constraints and the condition in equation (1.1) are generally insufficient to define f(x) uniquely but may delineate a set of feasible distributions. Each of these distributions contains a certain amount of uncertainty which can be expressed by employing the concept of entropy.

The most popular measure of entropy was first mathematically given by Shannon (1948a, b) and has since been called the Shannon entropy functional, SEF in short, denoted as I[f] or I[x]. It is a numerical measure of uncertainty associated with f(x) in describing the random variable X, and is defined as

$$I[f] = I[x] = -k \int_a^b f(x) \, ln \, [f(x)/m(x)] \, dx \tag{1.7}$$

where $k > 0$ is an arbitrary constant or scale factor depending on the choice of measurement units, and m(x) is an invariant measure function guaranteeing the invariance of I[f] under any allowable change of variable, and provides an origin of measurement of I[f]. Scale factor k can be absorbed into the base of the logarithm and m(x) may be taken as unity so that equation (1.7) is often written as

$$I[f] = - \int_a^b f(x) \, ln \, f(x) \, dx; \int_a^b f(x) \, dx = 1 \tag{1.8}$$

We may think of I[f] as the mean value of -lnf(x). Actually, -I measures the strength, +I measures the weakness. SEF allows choosing that f(x) which minimizes the uncertainty. Note that f(x) is conditioned on the constraints used for its derivation. Verdugo Lazo and Rathie (1978), and Singh et al. (1985, 1986) have given expressions of SEF for a number of probability distributions.

SEF with the transformed function g(z) is written accordingly as

$$I[g] = - \int_L^U g(z) \ln g(z) \, dz \tag{1.9}$$

It can be shown that

$$\begin{aligned} I[f] &= I[g] + \int_L^U g(z) \, ln \, \left|\frac{dx}{dz}\right| dz \\ &= I[g] + \int_a^b f(x) \, ln \, \left|\frac{dx}{dz}\right| dx \end{aligned} \tag{1.10}$$

In practice, we usually have a discrete set of data points x_i, i = 1, 2, ..., N, instead of a continuous variable X. Therefore, the discrete analog of equation (1.8) can be expressed as

$$I[f] = - \sum_{i=1}^{i=N} f_i \, ln \, f_i; \sum_i f_i = 1 \tag{1.11}$$

in which f_i denotes the probability of occurrence of x_i , and N is the sample size. Here $0 \leq f_i \leq 1$ for all i. The passage from the continuous mode to the discrete one and vice versa is subtle because f_i in equation (1.11) is a probability and f(x) in equation (1.8) is a probability density. The use of m(x), as in equation (1.2), facilitates, to some extent, the understanding of these transformations from the discrete mode to the continuous one and vice versa. Except for mentioning this point, we shall not discuss this aspect further. Mostly, we will use the form in equation (1.8) in formal analysis but in actual numerical work, the discrete version in equation (1.11) is employed. For a clear discussion of continuous random variables, their transformations, and probability distributions, one may refer to Rohatgi (1976).

Shannon (1948a, 1948b) showed that I is unique and the only functional that satisfies the following properties: (1) It is a function of the probabilities f_1, f_2, ..., f_N. (2) It follows an additive law, i.e., I[xy] = I[x] + I[y]. (3) It monotonically increases with the number of outcomes when f_i are all equal. (4) It is consistent and continuous.

1.2 Principle of Maximum Entropy

The principle of maximum entropy (POME) was formulated by Jaynes (1961, 1982). According to this principle, "the minimally prejudiced assignment of probabilities is that which maximizes the entropy subject to the given information." Mathematically, it can be stated as follows: Given m linearly independent constraints C_i in the form

$$C_i = \int_a^b y_i(x)\, f(x) dx, \; i = ,2, ..., m \tag{1.12}$$

where y_i (x) are some functions whose averages over f(x) are specified, then the maximum of I, subject to the conditions in equation (1.12), is given by the distribution

$$f(x) = exp\, [- \lambda_0 - \sum_{i=1}^{m} \lambda_i\, y_i\, (x)] \tag{1.13}$$

where λ_i, i = 0, 1, . . . , m, are Lagrange multipliers and can be determined from equations (1.12) and (1.13) along with the normalization condition in equation (1.1).

The Lagrange multipliers can be determined as follows.According to POME, we maximize equation (1.8), subject to equation (1.12), that is,

$$\delta\, (-I) = \int_a^b [1 + ln\, f\, (x)]\; \delta\, f\, (x) dx \tag{1.14}$$

The function I can be maximized using the method of Lagrange multipliers. This introduces parameters (λ_0 - 1), λ_1, λ_2, . . . , λ_n, which are chosen such that variations in a functional of f(x)

$$F(f) = - \int_a^b f(x)\, [ln\, f(x) + (\lambda_0 - 1) + \sum_{i=1}^{m} \lambda_i\, y_i\, (x)]\; dx \tag{1.15}$$

vanish:

$$\delta\, F(f) = - \int_a^b [ln\, f(x) + 1 + (\lambda_0 - 1) + \sum_{i=1}^{m} \lambda_i\, y_i\, (x)]\; \delta\, f(x)\; dx = 0 \tag{1.16}$$

Equation (1.16) produces

$$f(x) = exp\ [\ -\ \lambda_0\ -\ \sum_{i=1}^{m}\ \lambda_i\ y_i\ (x)] \tag{1.17}$$

which is the same as equation (1.13).

The value of I for such f(x), as given by equation (1.13), is

$$I_m[f] = -\int_a^b f(x)\ ln\ f(x)\ dx = \lambda_0 + \sum_{i=1}^{m} \lambda_i\ C_i \tag{1.18}$$

Subscript m attached to I is to emphasize the number of constraints used. This, however, raises an important question: How does I change with the changing number of constraints? To address this question, let us suppose that g(x) is some other pdf such that $\int_a^b g(x)dx = 1$ and is found by imposing n constraints (n > m) which include the previous m constraints in equation (1.12). Then

$$I_n[g] \leq I_m[f]\ for\ n \geq m \tag{1.19}$$

where

$$I_n[g] = -\int_a^b g(x) \ln g(x)\ dx \tag{1.20}$$

and

$$I_m[f] - I_n[g] \geq \frac{1}{2}\int_a^b g(x)\ (\frac{f(x) - g(x)}{g(x)})^2\ dx \geq 0 \tag{1.21}$$

In order to prove these statements, we consider

$$I[g\,|f] = \int_a^b g(x)\ ln\ [\frac{g(x)}{f(x)}]\ dx \tag{1.22}$$

Because of Jensen's inequality,

$$ln\ x \geq 1 - \frac{1}{x} \tag{1.23}$$

we have, upon normalization of f(x) and g(x),

$$I\,[g|\,f\,] \geq 0 \tag{1.24}$$

From equation (1.22), this relation may be written as

$$-\int_a^b g(x)\ \ln\ g(x)\ dx \leq -\int_a^b g(x)\ ln\ f(x)\ dx \tag{1.25}$$

Inserting equation (1.17) for f(x) in the right side of this inequality and the definitions given by equations (1.18) and (1.20), we get equation (1.19). To obtain equation (1.21), we note that

$$I[g|f] = \int_a^b g(x)\, ln\, [\frac{g(x)}{f(x)}]\, dx$$
$$= -\frac{1}{2}\int_a^b g(x)\, ln\, [1 + \frac{f(x) - g(x)}{g(x)}]^2\, dx \tag{1.26}$$
$$\geq +\frac{1}{2}\int_a^b g(x)\, (\frac{f(x) - g(x)}{g(x)})^2\, dx$$

Since $-\int_a^b g(x)\, ln\, f(x)\, dx = -\int_a^b f(x)\, \ln f(x) dx$ in this problem, because the first m constraints are the same, we have

$$I[g|f] = I_m[f] - I_n[g] \tag{1.27}$$

and hence we obtain equation (1.21). The significance of this result lies in the fact that the increase in the number of constraints leads to less uncertainty as to the information concerning the system. Since equation (1.27) defines the gain in information or reduction in uncertainty due to increased number of constraints, an average rate of gain in information I_r can be defined as

$$I_r = \frac{I_m[f] - I_n[g]}{n - m} \tag{1.28}$$

1.3 Entropy-Based Parameter Estimation

The general procedure for deriving an entropy-based parameter estimation method for a frequency distribution involves the following steps: (1) Define the given information in terms of constraints. (2) Maximize the entropy subject to the given information. (3) Relate the parameters to the given information. More specifically, let the available information be given by equation (1.12). POME specifies f(x) by equation (1.13). Then inserting equation (1.13) in equation (1.8) yields

$$I[f] = \lambda_0 + \sum_{i=1}^{m} \lambda_i\, C_i \tag{1.29}$$

In addition, the potential function or the zeroth Lagrange multiplier λ_0 is obtained by inserting equation (1.13) in equation (1.1) as

$$\int_a^b exp\, [-\lambda_0 - \sum_{1=m}^{m} \lambda_i\, y_i]\, dx = 1 \tag{1.30}$$

resulting in

$$\lambda_0 = ln \int_a^b exp\,[-\sum_{i=1}^{m} \lambda_i\, y_i]\, dx \tag{1.31}$$

The Lagrange multipliers are related to the given information (or constraints) by

$$-\frac{\partial \lambda_0}{\partial \lambda_i} = C_i \tag{1.32}$$

It can also be shown that

$$\frac{\partial^2 \lambda_0}{\partial \lambda_i^2} = var\,[y_i(x)]; \frac{\partial^2 \lambda_0}{\partial \lambda_i \partial \lambda_j} = cov\,[y_i(x),\, y_j(x)], i \neq j \tag{1.33}$$

With the Lagrange multipliers estimated from equations (1.32) and (1.33), the frequency distribution given by equation (1.13) is uniquely defined. It is implied that the distribution parameters are uniquely related to the Lagrange multipliers. Clearly, this procedure states that a frequency distribution is uniquely defined by specification of constraints and application of POME.

Quite often, we anticipate a certain structure of pdf, say in the form [this is normalized according to equation (1.1)],

$$f(x) = A\, x^k\, exp\,[-\sum_{i=1}^{m} \lambda_i\, y_i(x)] \tag{1.34}$$

where $y_i(x)$ are known functions and k may not be known explicitly but the form x^k is a guess. Then we may apply POME as follows. We explicitly construct the expression for I[f] in the form.

$$I[f] = -\, ln\, A - k\, E[ln\, x] + \sum_{i=1}^{m} \lambda_i\, E[y_i(x)] \tag{1.35}$$

We may then seek to maximize I[f] subject to the constraints, E[ln x], $E[y_i(x)]$, which can be evaluated numerically by means of experimental data. In this fashion, we arrive at an estimation of the pdf which is least biased with respect to the specified constraints and is of the surmised form based upon our intuition. This provides a method of deducing the constraints, given a "form" for the pdf.

This procedure can be applied to derive any probability distribution for which appropriate constraints can be found. The hydrologic import of constraints for every distribution, except a few, is not clear at this point. This procedure needs modification, however, if the distribution is expressed in inverse form as for example the Wakeby distribution.

The above discussion indicates that the Lagrange multipliers are related to the constraints on one hand and to the distribution parameters on the other hand. These two sets of relations are used to eliminate the Lagrange multipliers and develop, in turn, equations for estimating parameters in terms of constraints. For example, consider the gamma distribution. The Lagrange multipliers λ_1 and λ_2 are related to the constraints E(x) and E(ln x), and independently to the two distribution parameters. Finally, the relation between parameters and the specified constraints is found. Thus, POME leads tc a method of parameter estimation.

1.4 Parameter-Space Expansion Method

The parameter-space expansion method was developed by Singh and Rajagopal (1986). This method is different from the previous entropy method in that it employs enlarged parameter space and maximizes entropy subject to both the parameters and the Lagrange multipliers. An important implication of this enlarged parameter space is that the method is applicable to virtually any distribution, expressed in direct form, having any number of parameters. For a continuous random variable X having a probability density function f (x,θ) with parameters θ, SEF can be expressed as

$$I[f] = \int_{-\infty}^{\infty} f(x;\theta) \ln f(x,\theta)\, dx \tag{1.36}$$

The parameters of this distribution, θ, can be estimated by achieving the maximum of I[f]. The method works as follows: For the given distribution, the constraints (to be obtained from data) are first defined. Using the method of Lagrange multipliers (as many as the number of constraints), the POME formulation of the distribution is obtained in terms of the parameters to be estimated and the Lagrange multipliers. This formulation is used to define SEF whose maximum is then sought. If the probability distribution has N parameters, θ_i, I=1,2,3,.....,N, and (N-1) Lagrange multipliers, λ_i, I=1,2,3,.......,(N-1), then the point where I[f] is maximum is a solution of (2N-1) equations:

$$\frac{\partial I[f]}{\partial \lambda_i} = 0\,,\ i = 1, 2, 3, \ldots, N-1 \tag{1.37}$$

and

$$\frac{\partial I[f]}{\partial \theta_i} = 0\,,\ i = 1, 2, 3, \ldots\ldots, N \tag{1.38}$$

Solution of equations (1.37) and (1.38) yields distribution parameter estimates.

1.5 Entropy as a Criterion for Goodness of Fit

It is plausible to employ entropy to evaluate goodness of fit and consequently delineate the best parameter estimates of a fitted distribution. This can be accomplished as follows. For a given sample, compute entropy and call it observed entropy. To this end, we may use an appropriate plotting position formula. Then, compute parameters of the desired distribution by different methods (moments, maximum likelihood, least squares, POME, etc.). Calculate the entropy for each of these methods, and call it computed entropy. The method providing the computed entropy closest to the observed entropy is deemed the best method.

1.6 Dependence of Entropy on Sample Size

In practice, we usually employ a discrete set of data points, x_i, i = 1, 2, ..., N, to determine the constraints the representativeness and accuracy of which depend upon the sample size. To emphasize the dependence of I on N, we write equation (1.11) as

$$I_N[f] = -\sum_{i=1}^{N} f(x_i;\, a)\, ln\, f(x_i;\, a),\ with \sum_{i=1}^{N} f(x_i;\, a) = 1 \qquad (1.39)$$

where a is a parameter set. Using the inequality

$$f(x) - f^2(x) \leq f(x) \ln f(x) \leq 1 - f(x) \qquad (1.40)$$

we obtain

$$1 - \sum_{i=1}^{N} f^2(x_1;a) \leq I_N[f] \leq N - 1 \qquad (1.41)$$

If however, $f_i = 1/N$ (uniform distribution) then

$$0 \leq I_N[f] \leq \ln N \qquad (1.42)$$

References

Alpaslan, N., Harmancioglu, N. B. and Singh, V.P., 1992. The role of entropy concept in design and evaluation of water quality montoring networks. In: *Entropy and Energy Dissipation in Water Resources*, edited by V.P. Singh and M. Fiorentino, pp. 261-282, Kluwer Academic Publishers, Dordrecht, The Netherlands.

Harmancioglu, N.B., Singh, V.P., and Alpaslan, N. 1992a. Versatile uses of the entropy concept in water resources. In: *Entropy and Energy Dissipation in Water Resources*, edited by V.P. Singh and M. Fiorentino, pp. 261-282, Kluwer Academic Publishers, Dordrecht, The Netherlands.

Harmancioglu, N.B., Alpaslan, N. and Singh, V.P., 1992b. Application of the entropy concept in design of water quality monitoring networks. In: *Entropy and Energy Dissipation in Water Resources*, edited by V.P. Singh and M. Fiorentino, pp. 2831-302, Kluwer Academic Publishers, Dordrecht, The Netherlands.

Jaynes, E.T., 1957a. Information theory and statistical mechanics, I. *Physical Review*, Vol. **106**, pp. 620-630.

Jaynes, E.T., 1957b. Information theory and statistical mechanics, II. *Physical Review*, Vol. **108**, pp. 171-190.

Jaynes, E.T., 1961. *Probability Theory in Science and Engineering*. McGraw-Hill Book Company, New York.

Jaynes, E.T., 1982. On the rationale of maximum entropy methods. *Proceedings of IEEE*, Vol. **70**, pp. 939-0952.

Rajagopal, A.K. Teitler, S. And Singh, V.P., 1987. Some new perspectives on maximum entropy techniques in water resources research. In: *Hydrologic Frequency Modeling*, edited by V.P. Singh, pp. 347-366, D.Reidel Publishing Company, Dordrecht, The Netherlands.

Rohatgi, V.K., 1976. *An Introduction to Probability Theory and Mathematical Statistics.* John Wiley & Sons, New York.

Shannon, C.E., 1948a. The mathematical theory of communications, I and II. *Bell System Technical Journal*, Vol. **27**, pp. 379-423.

Shannon, C.E., 1948b. The mathematical theory of communications, III and IV. *Bell System Technical Journal*, Vol. **27**, pp. 623-656.

Singh, V.P., 1989. Hydrologic modeling using entropy. *Journal of the Institution of Engineers (India) Engineering Division*, Vol. **70**, pp. 55-60.

Singh, V.P., 1997. The use of entropy in hydrology and water resources. *Hydrological Processes*, Vol. **11**, pp. 587-626.

Singh, V.P. and Fiorentino, M., 1992. A historical perspective of entropy applications in water resources. In: *Entropy and Energy Dissipation in Water Resources*, edited by V.P. Singh and M. Fiorentino, pp. 21-62, Kluwer Academic Publishers, Dordrecht, The Netherlands.

Singh, V.P. and Rajagopal, A.K., 1986. A new method of parameter estimation for hydrologic frequency analysis. *Hydrological Science and Technology*, Vol. **2**, No. 3, pp. 33-40.

Singh, V.P. and Rajagopal, A.K., 1987. Some recent advances in application of the principle of maximum entropy (POME) in hydrology. *IAHS Publication* No. **164**, pp. 353-364.

Singh, V.P., Rajagopal, A.K. and Singh, K., 1986. Derivation of some frequency distributions using the principle of maximum entropy (POME). *Advances in Water Resources*, Vol. **9**, pp. 91-106.

Singh, V.P., Singh, K. and Rajagopal, A..K., 1985. Application of the principle of maximum entropy (POME) to hydrologic frequency analysis. *Completion Report* **06**, 144p., Louisiana Water Resources Research Institute, Louisiana State University, Baton Rouge, Louisiana.

Verdugo Lazo, A.C. G. and Rathie, P.N., 1978. On the entropy of continuous probability distributions. *IEEE Transactions, Information Theory*, Vol **IT-24**, No. 1, pp. 120-122.

CHAPTER 2

METHODS OF PARAMETER ESTIMATION

There is a multitude of methods for estimating parameters of hydrologic frequency models. Some of the popular methods used in hydrology include (1) method of moments (Nash, 1959; Dooge, 1973; Harley, 1967; O'Meara, 1968; Van de Nes and Hendriks, 1971; Singh, 1988); (2) method of probability weighted moments (Greenwood, et al., 1979); (3) method of mixed moments (Rao, 1980, 1983; Shrader, et al., 1981); (4) L-moments (Hosking, 1986, 1990, 1992); (5) maximum likelihood estimation (Douglas, et al., 1976; Sorooshian, et al., 1983; Phien and Jivajirajah, 1984); and (6) least squares method (Jones, 1971; Snyder, 1972; Bree, 1978a, 1978b). A brief review of these methods is given here.

2.1 Method of Moments for Continuous Systems

The method of moments is frequently utilized to estimate parameters of linear hydrologic models (Nash, 1959; Diskin, 1967; Diskin and Boneh, 1968; Dooge, 1973; Singh, 1988). Nash (1959) deveoped the theorem of moments which relates the moments of input, output and impulse response functions of linear hydrologic models. Diskin (1967) and Diskin and Boneh (1968) generalized the theorem. Moments of functions are amenable to use of standard methods of transform, such as the Laplace and Fourier transforms. Numerous studies have employed the method of moments for estimating parameters of frequency distributions. Wang and Adams (1984) reported on parameter estimation in flood frequency analysis. Ashkar et al. (1988) developed a generalized method of momemts and applied it to the generalized gamma distribution. Kroll and Stedinger (1996) estimated moments of a lognormal distribution using censored data.

2.1.1 DEFINITION AND NOTATION

Let X be a continuous variable and f(x) its function satisfying some necessary conditions. The r-th moment of f(x) about an arbitrary point is denoted as M_r^a. This notation will be employed throughout the chapter. Here M denotes the moment, $r \geq 0$ is the order of the moment, the subscript denotes the order of the moment, the superscript denotes the point about which to take the moment, and the quantity within the parentheses denotes the function, in normalized form, whose moment to take. Then, the r-th moment of the function f(x) can be defined as

$$M_r^a(f) = \int_{-\infty}^{\infty} (x-a)^r f(x)\, dx \tag{2.1}$$

This is the definition used normally in statistics. In engineering, however, the area enclosed by the function f(x) may not always be one. Then, the definition of equation (2.1) becomes

$$M_r^a(f) = \frac{\int_{-\infty}^{\infty} (x-a)^r f(x)\,dx}{\int_{-\infty}^{\infty} f(x)\,dx} \qquad (2.2)$$

As the denominator in equation (2.2) defines the area under the curve which is usually unity or made to unity by normalization, the two definitions are numerically the same. In this text we use the definition of equation (2.1) with f(x) normalized beforehand. The variable X may or may not be a random variable.

Analogous to equation (2.1), the absolute moment of order r about a, W_r^a, can be defined as

$$W_r^a = \int_{-\infty}^{\infty} |x-a|^r f(x)\,dx \qquad (2.3)$$

Here W stands for the absolute moment. Clearly, if r is even, then the absolute moment is equal to the ordinary moment. Furthermore, if the range of the function is positive, then the absolute moments about any point to the left of the start of the function are equal to the ordinary moments of corresponding order.

It is, of course, assumed here that the integral equation (2.1) converges. There are some functions which will possess moments of lower order, and some will not possess any except the moment of zero order. However, if a moment of higher order exists, moments of all lower order must exist. Moments are statistical descriptors of a distribution and reflect on its qualitative properties. For example, if r=0 then equation (2.1) yields

$$M_0^a = \int_{-\infty}^{\infty} (x-a)^0 f(x)\,dx = \int_{-\infty}^{\infty} f(x)dx = 1 \qquad (2.4)$$

Thus, the zero-order moment is the area under the curve defined by f(x) subject to $-\infty<x<\infty$.

If r=1, then equation (2.1) yields

$$M_1^a = \int_{-\infty}^{\infty} (x-a)^1 f(x)\,dx = \mu - a \qquad (2.5)$$

where μ is the centroid of the area or mean. Thus, the first moment is the weighted mean about the point a. If a=0, the first moment gives the mean. When $a=\mu$, then the r-th moment about the mean is

$$M_r^\mu = \int_{-\infty}^{\infty} (x-\mu)^r f(x)\,dx \qquad (2.6)$$

Henceforth we will, for simplicity of notation, drop the superscript if the moment is taken about

0. The descriptive properties of the moments with respect to a specific function can be summarized as follows:

$M_0 =$ Area

$M_1 =$ Lag or Mean

$M_2^\mu =$ Variance, a measure of dispersion of the function about the mean

$M_3^\mu =$ Measure of skewness of the function

$M_4^\mu =$ Kurtosis, a measure of the peakedness of the function

2.1.2 MOMENT GENERATING FUNCTION

The moments of a function can be determined either directly by using their definition in equation (2.1) or by using generating functions. This latter approach is more viable, more frequently used and has other advantages. One generating function is the moment generating function (MGF), $G(\theta)$. The MGF of f(x) can be defined as

$$G(\theta) = \int_{-\infty}^{\infty} e^{\theta x} f(x)\, dx = E[e^{\theta x}] \tag{2.7}$$

where E is the expectation operator and θ is the transform variable of MGF. We can write equation (2.7) as

$$G(\theta) = \int_{-\infty}^{\infty} \exp(\theta x) f(x)\, dx \tag{2.8}$$

If $G(\theta)$ exists then it is continuously differentiable in some neighborhood of the origin. If $G(\theta)$ in equation (2.8) is differentiated r times with respect to θ, we get

$$\frac{d^r}{d\theta^r} G(\theta) = \int_{-\infty}^{\infty} x^r [e^{\theta} x] f(x) dx \tag{2.9}$$

If $\theta = 0$, we obtain

$$\frac{d^r}{d\theta^r} G(\theta)|_{\theta=0} = \int_{-\infty}^{\infty} x^r f(x) dx = E[x^r] = M_r \tag{2.10}$$

Equation (2.10) shows generation of moments by taking the derivatives of MGF and evaluating the derivatives at $\theta = 0$. The symbol on the left hand side is to be interpreted to mean the r-th derivative of $G(\theta)$ evaluated at $\theta = 0$. Thus, the moments of a function can be obtained by differentiation of its moment generating function. Furthermore, if we expand $G(\theta)$ about $\theta=0$ by Taylor series then

$$G(\theta)=\sum_{r=0}^{\infty}\left[\frac{d^rG(\theta)}{d\theta^r}\Big|_{\theta=0}\right]\frac{\theta^r}{r!} \tag{2.11a}$$

Equation (2.11a) can be written as

$$G(\theta)=\sum_{r=0}^{\infty}M_r\frac{\theta^r}{r!} \tag{2.11b}$$

Equation (2.11b) shows that M_r's are nothing but coefficients of θ^r in the Taylor series expansion of MGF. This same result can be obtained by expanding the exponential term in equation (2.7):

$$e^{\theta x}=\sum_{r=0}^{\infty}\frac{(\theta x)^r}{r!} \tag{2.12}$$

On taking expectation of both sides, we obtain

$$G(\theta)=E\left[\sum_{r=0}^{\infty}\frac{(\theta x)^r}{r!}\right]=\sum_{r=0}^{\infty}\frac{\theta^r}{r!}E[x^r]=\sum_{r=0}^{\infty}M_r\frac{\theta^r}{r!} \tag{2.13}$$

Equation (2.13) shows that M_r's are the coefficients of θ^r in the exponential expansion.

2.1.3 CHARACTERISTIC FUNCTION

Unfortunately, the moment generating functions do not always exist. Often it is better to work with the characteristic function (CF) which always exists. A characteristic function can be defined as

$$C(\theta)=E[e^{i\theta x}]=\int_{-\infty}^{\infty}e^{i\theta x}f(x)dx \tag{2.14}$$

where $C(\theta)$ can be viewed as an integral transform of f(x). The integral of $C(\theta)$ converges absolutely and uniformly for the functions of our interest in hydrology.

If $C(\theta)$ is differentiated r times with respect to θ, then

$$\frac{d^r}{d\theta^r}C(\theta)=(i)^r\int_{-\infty}^{\infty}e^{i\theta x}x^rf(x)dx \tag{2.15}$$

Evaluating equation (2.15) at $\theta = 0$, one gets

$$\frac{d^r}{d\theta^r} C(\theta)\,|_{\theta=0} = (i)^r \int_{-\infty}^{\infty} x^r f(x)\, dx = (i)^r E[x^r] = (i)^r M_r \tag{2.16}$$

Therefore, equation (2.16) shows that

$$M_r = (i)^{-r} \frac{d^r}{d\theta^r} C(\theta)\,|_{\theta=0} \tag{2.17}$$

Thus, the moments of a function can be obtained by differentiation of the corresponding characteristic function. Hence, if $C(\theta)$ is known explicitly, the moments about the origin can be obtained.

Moreover, if we expand $C(\theta)$ by Taylor series about the origin, we get

$$C(\theta) = \sum_{r=0}^{\infty} \frac{(\theta)^r}{r!} [\frac{dr}{d\theta^r} C(\theta)|_{\theta=0}] = \sum_{r=0}^{\infty} \frac{(i\theta)^r}{r!} M_r \tag{2.18}$$

We find that the moments about the origin are the coefficients of θ^r in the expansion of $C(\theta)$. This same result is obtained by expanding the exponential term in equation (2.14):

$$C(\theta) = E[\sum_{r=0}^{\infty} \frac{(i\theta x)^r}{r!}] = \sum_{r=0}^{\infty} \frac{(i\theta)^r}{r!} E[x^r] = \sum_{r=0}^{\infty} \frac{(i\theta)^r}{r!} M_r \tag{2.19}$$

It is interesting to contrast $C(\theta)$ with the Fourier transform which can be defined for a function f(x) as

$$F(w) = \int_{-\infty}^{\infty} f(x) e^{iwx} dx \tag{2.20}$$

where F(w) is the Fourier transform of f(x). A comparison of equations (2.20) and (2.14) shows that the Fourier transform is equivalent to the characteristic function. In terms of the Fourier transform, the function f(x) can be defined as

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(w) e^{-iwx} dw \tag{2.21}$$

Likewise,

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} C(\theta)e^{-i\theta x}d\theta \qquad (2.22)$$

Thus, it is seen that moments of a function can be generated using Fourier transform as well.

2.1.4 LAPLACE TRANSFORM

For certain functions the integral represented by $G(\theta)$ or F(w) may not exist. It is then necessary to use the Laplace transform. For functions that are zero for x<0, the ordinary Laplace transform, with Laplacian variable s, can be used in which we have

$$F(s) = L[f(x)] = \int_{0}^{\infty} f(x)e^{-sx}\,dx \qquad (2.23)$$

For functions which have values for x<0, we must use the bilateral Laplace transform given by

$$F(s) = L[f(x)] = \int_{-\infty}^{\infty} f(x)\,e^{-sx}\,dx \qquad (2.24)$$

Differentiating r times the function F(s) with respect to s and evaluating the derivatives at s=0,

$$\frac{d^r}{ds^r}F(s)\Big|_{s=0} = \int_{0}^{\infty}(-1)^r x^r f(x)dx = (-1)^r M_r \qquad (2.25)$$

Thus, we see that once the Laplace transform of a function is known, its moments can be obtained by differentiation.

2.1.5 FOURIER TRANSFORM

The moments of a function f(x) can also be obtained by employing the Fourier transform expressed as

$$F(w) = T[f(x)] = \int_{-\infty}^{\infty} e^{iwx} f(x)\,dx \qquad (2.26)$$

in which T denotes the Fourier transform. Differentiating r times the function F(w) with respect to w and evaluating the derivatives at w=0, one obtains

$$\frac{d^r}{dw^r}F(w)\Big|_{w=0} = \int_{-\infty}^{\infty}(-ix)^r f(x)dx = (-i)^r M_r \qquad (2.27)$$

Thus, if the Fourier transform of a function is known, its moments about the origin can be obtained by differentiation.

2.1.6 CHANGE OF REFERENCE POINTS FOR MOMENTS

Let a and b be two constants and let c=b-a. Let us denote the r-th moment of a function about a and b respectively by M_r^a and M_r^b. Expanding $(x-a)^r$ binomially, one gets

$$(x-a)^r = (x-b+b-a)^r = (x-b+c)^r = \sum_{j=0}^{r} \binom{r}{j} (x-b)^{r-j} c^j \tag{2.28}$$

Then we write the r-th moment about a:

$$M_r^a = \int_{-\infty}^{\infty} (x-a)^r f(x) dx \tag{2.29}$$

Substituting equation (2.28) in equation (2.29) one obtains

$$\begin{aligned} M_r^a &= \int_{-\infty}^{\infty} \sum_{j=0}^{r} \binom{r}{j} (x-b)^{r-j} c^j f(x)\, dx \\ &= \sum_{j=0}^{r} \binom{r}{j} c^j M_{r-j}^b \\ &= \sum_{j=0}^{r} \binom{r}{j} c^j \int_{-\infty}^{\infty} (x-b)^{r-j} f(x)\, dx \end{aligned} \tag{2.30}$$

Equation (2.30) gives the r-th moment of a function about a in terms of its r-th moment and lower moments about b. Writing the suffixes as power indices (without, of course, interpreting them as such except for the purpose of expansion), the symbolic mnemonic form of the above relationship becomes

$$M_r^a = \left[M^b + c\right]^r \text{ , for all } r \tag{2.31}$$

in which $\left[M^b\right]^r$ is interpreted as M_r^b for all r. If we specialize by taking a as the origin and b the centroid or the first moment μ, then we get

$$M_r = \sum_{j=0}^{r} \binom{r}{j} (\mu)^j M_{r-j}^{\mu} \tag{2.32}$$

The symbolic mnemonic form of the above relationship can be expressed as

$$M_r = \left[M^{\mu} + \mu\right]^r \text{ ,for all } r \tag{2.33}$$

In particular, this expression leads to

$$\begin{aligned} M_0 &= M_0^{\mu} = 1 \\ M_1 &= M_1^{\mu} + \mu\, M_0^{\mu} = \mu \\ M_2 &= M_2^{\mu} + \mu^2 \end{aligned} \tag{2.34}$$

$$M_3 = M_3^{\mu} + 3\mu\, M_2^{\mu} + \mu^3$$

$$M_4 = M_4^{\mu} + 4\mu\, M_3^{\mu} + 6M_2^{\mu}\mu^2 + \mu^4$$

and so on. Equation (2.33) can be manipulated to express moments about the controid in terms of the moments about the origin. In particular, this yields

$$\begin{aligned} M_0^{\mu} &= 1 \\ M_1^{\mu} &= 0 \\ M_2^{\mu} &= M_2 - \mu^2 \\ M_3^{\mu} &= M_3 - 3M_2\mu + 2\mu^3 \\ M_4^{\mu} &= M_4 - 4M_3\mu + 6\mu^2 M_2 - 3\mu^4 \end{aligned} \tag{2.35}$$

Equation (2.35) can also be derived directly. Writing the r-th moment about point b, one gets

$$M_r^{\,b} = \int_{-\infty}^{\infty} (x-b)^r f(x)\, dx \tag{2.36}$$

Expanding $(x-b)^r$ binomially, one obtains

$$\begin{aligned} (x-b)^r &= (x-a-b+a)^r = (x-a-c)^r \\ &= \sum_{j=0}^{r} \binom{r}{j} (x-a)^{r-j}(-c)^j \end{aligned} \tag{2.37}$$

Inserting equation (2.37) into the general expression for $M_r^{\,b}$ in equation (2.36), one gets

$$\begin{aligned} M_r^{\,b} &= \int_{-\infty}^{\infty} \sum_{j=0}^{r} \binom{r}{j} (x-a)^{r-j}(-c)^j f(x) dx \\ &= \sum_{j=0}^{r} \binom{r}{j} (-c)^j M_{r-j}^{\,a} \end{aligned} \tag{2.38}$$

The symbolic mnemonic form of this relationship is

$$M_r^{\,b} = [\, M^{\,a} - c\,]^r, \textit{ for all } r \tag{2.39}$$

In particular, if a=0 and b=μ, then

$$M_r^{\mu} = \sum_{j=0}^{r} (-1)^j \binom{r}{j} \mu^j M_{r-j} \tag{2.40}$$

Likewise,

$$M_r^{\mu} = [m-\mu]^r, for\ all\ r \tag{2.41}$$

Again, the coefficients of the expansions in equation (2.41) are those of the binomial expansion. The powers of terms containing the quantity μ are real powers, the powers of terms containing M are not real powers but only indices of the moments concerned.

2.1.7 INVARIANCE PROPERTY

The moments have an invariance property which states that when the variate-values are multiplied by a constant, the r-th moment M_r is multiplied by a^r. This is evident at once from its definition.

2.1.8 INVERSION OF MOMENTS

Moments can be used as parameters to represent distribution functions. A question arises: Can the distribution function be derived from a knowledge of the values of the moments? Let us recall that

$$M_r = (-1)^r [\frac{d^r}{ds^r} F(s)]|_{s=0} \tag{2.42}$$

If the moments about the origin are known for a well-behaved function, the Laplace transform can be expressed in terms of these moments by means of Taylor series as

$$F(s) = \sum_{r=0}^{\infty} (-s)^r \frac{M_r}{r!} \tag{2.43}$$

Therefore, a knowledge of the Laplace transform gives certain information about the behavior of the original function. If, however, we wish to explicitly know the complete function exactly, it would be necessary to invert the Laplace transform numerically. This can perhaps be best done by using orthogonal functions.

2.1.9 DIMENSIONLESS MOMENTS

It is often convenient to use dimensionless moments, which are independent of one another, in model calibration. Nash (1959) used dimensionless moments about the mean defined as

$$m_2 = \frac{M_2^{\mu}}{(M_1)^2 (M_0)} \tag{2.44}$$

$$m_3 = \frac{M_3^\mu}{(M_1)^3(M_0)} \tag{2.45a}$$

$$m_4 = \frac{M_4^\mu}{(M_1)^4(M_0)} \tag{2.45b}$$

where m_2, m_3, and m_4 are generally called shaped factors. The objective of dividing by the first moment is to remove the time scale effect from higher moments and thus make them dimensionless. These shape factors can be used to compare distribution functions by constructing an m_3 versus m_2 diagram as suggested by Nash (1959) and done by Harley (1967), O'Meara (1968), and Dooge (1973), among others.

A more popular way of obtaining dimensionless moments is to use the second central moment as the divisor. Thus, the rth dimensionless moment is obtained as

$$s_r = \frac{M_r^\mu(x)}{[M_2^\mu(x)]^{r/2}} \tag{2.46}$$

where s_r (x) is the rth moment of x. It should be noted that higher order moments can be expressed as functions of lower moments. For two-parameter distributions, M_3^μ can be expressed as a function of M_2^μ. As an example, the coefficient of skewness C_S = $M_3^\mu / (M_2^\mu)^{3/2}$ can be expressed as a unique function of the coefficient of variation. Likewise, the coefficient of kurtosis, $C_k = M_4^\mu / (M_2^\mu)^2$ can be expressed as a unique function of C_S. Thus, the C_S - C_k relationship defines a moment ratio diagram. Sometimes, the moments ratios are squared, as done by Bobee et al. (1993). The moments ratio diagrams are a very useful tool in selection of a distribution, comparing shapes of distributions, etc. Johnson and Kotz (1985) provided a comprehensive account of these diagrams and their usefulness.

2.2 Method of Moments for Discrete Systems

2.2.1 DEFINITION

If the function is discrete, represented as f_j, j, $= -\infty, \ldots, -1, 0, 1, 2, \ldots, \infty$, then its r-th moment about the origin or any other arbitrary point can be defined in a manner analogous to that for continuous functions. For convenience, let the arbitrary point be the origin. Then, the r-th moment is defined as

$$M_r = \sum_{m=-\infty}^{\infty} m^r f_m \tag{2.47}$$

It is assumed here that f_m is normalized, that is,

$$\sum_{m=-\infty}^{\infty} f_m = 1 \tag{2.48}$$

Otherwise,

$$M_r = \sum_{m=-\infty}^{\infty} m^r f_m / \sum_{m=-\infty}^{\infty} f_m \tag{2.49}$$

It is thus seen that equations (2.47) and (2.49) are analogous to equations (2.1) and (2.2).

2.2.2 MOMENT GENERATING FUNCTION

The moments of a discrete function can be determined either directly from their definition or by using generating functions. If x is discrete and takes the value j with probability p_j then MGF is

$$G(\theta) = \sum_j e^{\theta j} P_j = E[e^{\theta j}] = P[e^{\theta}] \tag{2.50}$$

in which P is probability generating function. The r-th moment can be determined by differentiating r times $G(\theta)$ with respect to θ and then equating to zero:

$$\frac{d^r G(\theta)}{d\theta^r} \Big|_{\theta=0} = M_r \tag{2.51}$$

Therefore, equation (2.10) also holds for discrete functions.

For discrete functions the Z-transform can be used as a moment generating function in the same way as the Laplace transform is used for continuous functions. The Z-transform of the function f_j is defined for the bilateral case as

$$Z(f_j) = F(z) = \sum_{j=-\infty}^{\infty} f_j z^{-j} \tag{2.52}$$

and for the unilateral case as

$$Z(f_j) = F(z) = \sum_{j=0}^{\infty} f_j z^{-j} \tag{2.53}$$

in which F(z) signifies the Z-transform of the function f_j. In most hydrologic cases $f_j = 0$, $j < 0$ so the bilateral transform reduces to the unilateral case. The moments of f_j about the origin can be obtained from the Z-transform in the following manner:

$$M_0 = \sum_{j=-\infty}^{\infty} f_j = F(z)\big|_{z=1} \tag{2.54}$$

$$M_1 = \sum_{j=-\infty}^{\infty} j f_j = -z\frac{d}{dz}(F(z)|_{z=1} \tag{2.55}$$

$$M_2 = \sum_{j=-\infty}^{\infty} j^2 f_j = -z\frac{d}{dz}[-z\frac{d}{dz}F(z)]|_{z=1} \tag{2.56}$$

$$M_3 = \sum_{j=-\infty}^{\infty} j^3 f_j = [-z\frac{d}{dz}\left\{-z\frac{d}{dz}(-z\frac{d}{dz}F(z))\right\}]|_{z=1} \tag{2.57}$$

and so on. The moments of discrete functions can be obtained about any reference points. Consequently, the relationships of equations (2.31) and (2.33) hold. Likewise, the theorem of moments, derived for continuous functions, is also valid for discrete functions. It can be easily seen by noting the Z-transform of the convolution summation:

$$y_j = \sum_{i=0}^{j} x_i h_{j-i} = \sum_{i=0}^{j} h_i x_{j-i}, j=0,1,2,... \tag{2.58}$$

which yields

$$Z[y_j] = Z[\sum_{i=0}^{j} x_j h_{j-i}] = Z[\sum_{i=0}^{j} h_i x_{j-i}] \tag{2.59}$$

Equation (2.59) can be expressed as

$$Y(z) = X(z)H(z) = H(z)X(z) \tag{2.60}$$

Taking logarithm to the base e of equation (2.60), one gets

$$\ln Y(z) = \ln X(z) + \ln H(z) \tag{2.61}$$

Thus, in logarithmic domain the log of Y is a linear sum of the logs of X and H.

2.2.3 INVERSION OF Z-TRANSFORM

The inverse Z-transform can best be defined as

$$Z^{-1}[F(z)] = f_j = \frac{1}{2\pi i}\int F(z) z^{j-1} dz$$

$$= \sum_{poles} \mathit{Residues\ of}\,[F(z)z^{j-1}]\,\mathit{at\ the\ poles\ of}\,F(z) \tag{2.62}$$

For a multiple pole of order m at z = w, we get

$$\text{The Residue} = [\frac{1}{(m-1)!}\frac{d^{m-1}}{dz^{m-1}}\{(z-w)^m F(z) z^{j-1}\}]|_{z=w} \tag{2.63}$$

2.3 Method of Probability Weighted Moments

Greenwood, et al.(1979) introduced the method of probability weighted moments (PWM) and showed its usefulness in deriving explicit expressions for parameters of distributions whose inverse forms X=X (F) can be explicitly defined. They derived relations between parameters and PWMs for generalized lambda, Wakeby, Weibull, Gumbel, logistic and kappa distributions. Hosking (1986) developed the theory of probability weighted moments and applied to estimate parameters of several distributions. Landwehr et al. (1979a,b) developed inference procedures using PWMs. For flood frequency analysis, Haktanir (1996) modified the conventional method of probability-weighted moments for estimation of parameters of any distribution without the need to use a plotting position formula. Wang (1996) defined partial PWMs and derived them for extreme value type I and III distributions. He applied these moments to lower bound censored samples.

Let a probability distribution function be denoted as $F=F(X)=P[X \le x]$. The PWMs of this function can be defined as

$$M_{i,j,k} = E[x^i F^j (1-F)^k] = \int_0^1 [x(F)]^i F^j (1-F)^k dF \tag{2.64}$$

where $M_{i,j,k}$ is the probability weighted moment of order (i, j, k), E is the expectation operator and i, j and k are real numbers. If j=k=0 and i is a nonnegative integer then $M_{i,0,0}$ represents the conventional moment about origin of order i. If $M_{i,0,0}$ exists and X is a continuous function of F, then $M_{i,j,k}$ exists for all nonnegative real numbers j and k.

For nonnegative integers j, k, we can express

$$M_{i,0,k} = \sum_{j=0}^{k} \binom{k}{j} (-1)^j M_{i,j,0} \tag{2.65a}$$

$$M_{i,j,0} = \sum_{k=0}^{j} \binom{j}{k} (-1)^k M_{i,0,k} \tag{2.65b}$$

If $M_{i,0,k}$ exists and X is a continuous function of F then $M_{i,j,0}$ exists. When the inverse X=X (F) of the distribution F=F (X) cannot be analytically defined, it may in general be difficult to derive $M_{i,j,k}$ analytically.

When i, j, k are nonnegative integers, the probability weighted moment of order (i, j, k), $M_{i,j,k}$ is proportional to $E[x^i_{j+1,j+k+1}]$, the i-th moment about the origin of the (j+1)th order statistic for a sample of size n=k+j+1. Symbolically,

$$M_{i,j,k} \, \alpha \, E[x^i_{j+1,j+k+1}] \tag{2.66a}$$

where ∝ denotes "proportional." Specifically,

$$E[x^i_{j+1,j+k+1}] = M_{i,j,k}/B[j+1,k+1] \tag{2.66b}$$

where B[.,.] denotes the beta function. If j=0, then

$$E[x^i_{i,k+1}] = (k+1)M_{i,0,k} \tag{2.67a}$$

Here $(k+1)M_{i,0,k}$ represents the i-th moment about the origin of the first order statistic for a sample of size k+1. Likewise, if k=0, then

$$E[x^i_{j+1,j+1}] = (j+1)M_{i,j,0} \tag{2.67b}$$

where $(j+1)M_{k,j,0}$ represents the i-th moment about the origin of the (j+1)th order statistic for a sample of size j+1. We are usually interested in cases where j and k are positive integers.

The expected value of the range of X in a sample of size n=k+1=j+1 can be written as

$$E[x_{n,n} - x_{1,n}] = n(M_{1,n-1,0} - M_{1,0,n-1}) \tag{2.68}$$

The PWMs can be derived for a distribution which can be expressed as X=X(F). These can then be related to the distribution parameters. The resulting relations may be of simpler structure than those between the conventional moments and the parameters. The simpler structure may be due in part to X being taken to only the first power.

We normally work with the moments $M_{i,j,k}$ into which x enters linearly. In particular, what we refer to as PWM for hydrologic applications are defined as

$$a_r = M_{1,0,r} = E[x\{1 - F(x)\}^r], \quad r = 0,1,2,.......... \tag{2.69a}$$

$$b_r = M_{1,r,0} = E[x\{F(x)\}^r], \qquad r = 0,1,2,.............. \tag{2.69b}$$

Note that $a_{k-1} = E[x_{1:k}]$, and $b_k = E[x_{k:k}]$ are expected values of extreme order statistics.

In general, a_r and b_r are functions of each other as

$$a_r = \sum_{k=0}^{r} (-1)^k \binom{r}{k} b_k \tag{2.70a}$$

$$b_r = \sum_{k=0}^{r} (-1) \binom{r}{k} a_k \tag{2.70b}$$

Therefore,

$$\begin{aligned} a_0 &= b_0 & ,b_0 &= a_0 \\ a_1 &= b_0 - b_1 & ,b_1 &= a_0 - a_1 \\ a_2 &= b_0 - 2b_1 + b_2 & ,b_2 &= a_0 - 2a_1 + a_2 \\ a_3 &= b_0 - 3b_1 + 3b_2 - b_3 & ,b_3 &= a_0 - 3a_1 + 3a_2 - a_3 \end{aligned} \tag{2.71}$$

A complete set of the a or b probability-weighted moments characterizes a distribution.

2.4 Methods of Mixed Moments

Rao (1980, 1983) proposed a method of mixed moments (MIXM) for fitting log-Pearson type III distribution. The MIXM method is applicable to any log-probability distribution. As the name suggests, the MIXM method is based on mixing the moments of real and logarithmically transformed data. Thus, only the first two moments (mean and variance) of the data are used. For example, if it is desired to fit the log-Pearson type (LP) III distribution to a given set of data then its parameters can be estimated in two ways: (1) The first method uses the mean ($\bar{x}$) and variance S_x^2 of real data and mean of logarithmically transformed values (y =log x). (2) the second method uses the mean of the real data ($\bar{x}$) and the mean and variance S_y^2 of logarithmically transformed data (y=log x). Rao (1980) showed using Monte Carlo experimentation that the first method possessed superior statistical properties as compared to the second method.

2.5 Method of L-Moments

The method of L-moments was developed by Hosking (1986, 1990) and has since become quite popular for characterization of probability distributions, summarization of observed data samples, parameter estimation or fitting of probability distributions to data, interval estimation, and testing of hypotheses about distributional form. Hosking and Wallis (1991) extended the use of L-moments and developed statistics for use in regional frequency analysis to measure discordancy, regional homogeneity, and goodness-of-fit. They (Hosking and Wallis, 1995) compared unbiased and plotting position estimators of L moments. Vogel and Fennessey (1993) proposed replacing product moment diagrams by L-moment diagrams and used them to discriminate among alternate distributional hypotheses about daily streamflows in Massachusetts. L-moment diagrams have been employed by Hosking and Wallis (1987) for selecting the generalized extreme value (GEV) distribution over the gamma distribution for modeling annual maximum hourly rainfall data. Vogel, et al. (1993a) used them to show that flood flows at 383 sites in southwestern United States were equally well approximated by log-Pearson type 3 (LP3), lognormal 3 (LN3), and generalized extreme value (GEV) distributions. Vogel et al. (1993b) used them to show that flood flows were well represented by a GEV distribution in the region of Australia which received most rainfall during winter months and by a generalized Pareto (GPA) distribution in the regions of Australia which received most rainfall during summer months. Vogel and Wilson (1996) constructed L-moment diagrams for annual minimum, average, and maximum streamflows at more than 1455 river basins in the United States. They then found that the genearlized extreme value (GEV), three-parameter lognormal (LN3) and log Pearson type III (LP III) distributions provided good approximations to the distribution of annual maximum flood flows. Bobee et al. (1993) discussed two kinds of moment ratio diagrams and their application in hydrology and stressed their need to choose between distributions. Rao and Hamed (1994) used L-moments for frequency analysis of upper Cauvery River annual maximum flow data in India. The 3-parameter lognomal and the generalized extreme value distributions were selected for the analysis as a result. L-moments are preferable to product moments for evaluating the power of alternative hypothesis tests for the normal distribution. Rao and Hamed (1997) applied L-moments to regional frequency analysis of Wabash River flood data. Wang (1997) developed a generalization of L-moments, called LH moments based on linear combinations of higher-order statistics. He introduced them to characterize the upper part of distributions and larger events in data. Thus use of these moments reduced undesirable influences that small sample events might have had on the estimation of large return periods. He formulated the method of LH moments for the generalized extreme value distribution.

The probability-weighted moments characterize a distribution but are not meaningful by themselves. L-moments were developed by Hosking (1986) as functions of PWMs which provide a descriptive summary of the location, scale, and shape of the probability distribution. L-moments are analogous to ordinary moments and are expressed as linear combinations of order statistics. They can also be expressed by linear combinations of probability-weighted moments. Thus, the ordinary moments, the probability weighted moments and L-moments are related to each other. L-moments are known to have several important advantages over ordinary moments. L-moments have less bias than ordinary moments because they are always linear combinations of ranked observations. As an example, variance (second moment) and skewness (third moment) involve squaring and cubing of observations, respectively, which compel them to give greater weight to the observations far from the mean. As a result, they result in substantial bias and variance.

If X is a real value ordered random variate of a sample of size n, such that $x_{1:n} \le x_{2:n} \le \dots \le x_{n:n}$ with cumulative distribution function F(x) and quantile function x(F), then the r-th L-moment of X (Hosking 1990) can be defined as a linear function of expected order statistics as:

$$L_r = \frac{1}{r} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} E\{X_{r-k:r}\}, \quad r = 1,2,.... \tag{2.72}$$

where E{.} is the expectation of an order statistic and is equal to

$$E\{X_{j:r}\} = \frac{r!}{(r-j)!\,(j)!} \int x\{F(x)\}^{j-1} \{1-F(x)\}^{r-j} dF(x) \tag{2.73}$$

As noted by Hosking (1990), the natural estimator of L_r, based on an observed sample of data, is a linear combination of the ordered data values, i.e., an L-statistic. Substituting equation (2.73) in equation (2.72), expanding the binomials of F(x) and summing the coefficients of each power of F(x), one can write as

$$L_r = E[x\,P^*_{r-1}(F(x))] = \int_0^1 x(F) P^*_{r-1}(F)\,dF, \quad r=1,2,.... \tag{2.74}$$

where $P_r^*(F)$ is the r-th shifted Legendre polynomial expressed as

$$P_r^*(F) = \sum_{r}^{k} (-1)^{r-k} \binom{r}{k} \binom{r+k}{k} F^k \qquad (2.75)$$

Equation (2.75) can simply be written as

$$P_r^*(F) = \sum_{k=0}^{r} P_{r,k} F^k \qquad (2.76)$$

and

$$P_{r,k} = (-1)^{r-k} \binom{r}{k} \binom{r+k}{k} \qquad (2.77)$$

The shifted Legendre polynomials are related to the ordinary Legendre polynomials $P_r(u)$ as $P_r^*(u) = P_r(2u-1)$, and are orthogonal on the interval (0,1) with constant weight function.

The first four L moments are

$$L_1 = E(x) = \int x dF \qquad (2.78)$$

$$L_2 = \frac{1}{2} E(x_{2:2} - x_{1:2}) = \int x(2F-1) dF \qquad (2.79)$$

$$L_3 = \frac{1}{3} E(x_{3:3} - 2x_{2:3} + x_{1:3}) = \int x(6F^2 - 6F + 1) dF \qquad (2.80)$$

$$L_4 = \frac{1}{4} E(x_{4:4} - 3x_{3:4} + 3x_{2:4} - x_{1:4}) = \int x(20F^3 - 30F^2 + 12F - 1) dF \qquad (2.81)$$

Perhaps the simplest way to define L moments is through the use of the probability-weighted moments (PWMs). L-moments are linear functions of PWMs (Hosking, 1990) discussed in the previous section. Greenwood et al. (1979) defined PWM's, $M_{q,r,s}$, as

$$M_{q,r,s} = E[(x(F))^q \{1 - x(F)\}^s \{F(x)\}^r] \qquad (2.82)$$

With q=1 and s=0,

$$M_{1,r,0} = b_r = \int_0^1 x(F) F^r dF \qquad (2.83)$$

With q=1 and r=0,

$$M_{1,0,s} = a_s = \int_0^1 x(F)(1-F)^s dF \qquad (2.84)$$

Both a_s and b_r are linear in x and are related to each other as

$$a_s = \sum_{k=0}^{s} \binom{s}{k} (-1)^k b_k \qquad (2.85)$$

$$b_r = \sum_{k=0}^{r} \binom{r}{k} (-1)^k a_k \qquad (2.86)$$

When r=0, b_0 is the mean. All higher-order PWMs are simply linear combinations of the order statistics $x_{(n)} \le x_{(n-1)} \le .... \le x_{(1)}$.

Unbiased sample estimates of PWMs for any distribution can be computed as

$$b_r = \frac{1}{n} \sum_{j=1}^{n-r} \frac{\binom{n-j}{r}}{\binom{n-1}{r}} x(j) \qquad (2.87)$$

where x(j) represents the ordered data, with $x_{(1)}$ being the largest observation and $x_{(n)}$ the smallest.

L-moments, L_{r+1}, can be expressed in terms of PWM's, a_r and b_r, as

$$L_{r+1} = (-1)^r \sum_{k=0}^{r} P_{r,k} a_k = \sum_{k=0}^{r} P_{r,k} b_k \qquad (2.88)$$

In particular,

$$L_1 = a_0 = b_0 \qquad (2.89)$$

$$L_2 = a_0 - 2a_1 = 2b_1 - b_0 \qquad (2.90)$$

$$L_3 = a_0 - 6a_1 + 6a_2 = 6b_2 - 6b_1 + b_0 \qquad (2.91)$$

$$L_4 = a_0 - 12a_1 + 30a_2 - 20a_3 = 20b_3 - 30b_2 - 12b_1 + b_0 \qquad (2.92)$$

The sample PWM's can be calculated from plotting positions as

$$a_r = \frac{1}{n} \sum_{i=1}^{n} (1 - P_{i:n})^r x_i \qquad (2.93)$$

$$b_r = \frac{1}{n} \sum_{i=1}^{n} P_{i:n}^r x_i \qquad (2.94)$$

where $P_{i:n}$ is obtained from a plotting position. The use of $P_{i:n}$ = (i-0.35)/n usually gives good results (Cunnane 1989).

The first L-moment is equal to the mean μ and is hence a measure of location. Other L-moments are measures of the scale and of the shape of a probability distribution. Analogous to the conventional moment ratios, Hosking (1990) defined L-moment ratios, R, as

$$R_2 = \frac{L_2}{L_1} = L - C_v \tag{2.95}$$

$$R_r = \frac{L_r}{L_2}, r \geq 3 \tag{2.96}$$

where R_2 is a measure of scale or dispersion, called $L - C_v$, R_3 is L-skewness, and R_4 is L-kurtosis. Thus, R_2, R_3, and R_4 can be thought of as measures of a distribution's scale, skewness, and kurtosis, respectively. The ratios R_r are independent of the units of measurement.

2.6 Method of Maximum Likelihood Estimation

The method of maximum likelihood (ML) estimation is widely accepted as one of the most powerful parameter estimation methods. Asymptotically, ML parameter estimates are unbiased, minimum variance, and normally distributed, while in some cases these properties hold for small samples. The MLE method has been extensively used for estimating parameters of frequency distributions as well as fitting conceptual models. Douglas et al. (1976) used likelihood functions to fit conceptual models with more than one dependent variable. Sorooshian et al. (1983) evaluated ML parameter estimation techniques for conceptual rainfall-runoff models and evaluated the influence of data varaibility and length on model credibility. Gupta and Sorooshian (1985) discussed the relationship between data used for hydrologic model calibration and the precision of model parameters estimated by the maximum likelihood approach. Rao and Mao (1987) investigated instrumental variable-approximate maximum likelihood method for modeling and forecasting daily flows. This method eliminates bias in parameter estimates. Duan et al. (1988) developed an MLE criterion suitable for conceptual model calibration using data which are recorded at unequal time intervals and which contain autocorrelated errors. Clarke (1996) developed residual maximum likelihood (REML) methods for analyzing hydrological data series and applied them to estimate mean areal monthly rainfall in Amazonia, using incomplete records from 48 raingage sites, as well as to analyze annual flood data from 19 flow gaging sites in sub-basins of a large river system in southern Brazil. REML is useful for analysis of hydrological data sets with records of varying lengths, intersite correlations and year-to-year effects. Kitanidis and Lane (1985) applied the MLE method to estimate hydrologic spatial processes.

The use of MLE is even more extensive in flood frequency analysis. Dubey (1967) applied MLE method to estimate shape parameter of the Weibull distribution. Cohen and Whitten (1982) modified the ML method for 3-parameter Weibull distribution. Kappenman (1985) estimated parameters of three-parameter Weibull, lognormal, and gamma distributions using the ML method. Kline and Bender (1990) estimated parameters of shifted Weibull and lognormal distributions. Phien and Jivajirajah (1984) compared the method of moments and MLE for fitting the four-parameter Johson S_B curve. Hosking (1985) proposed a correction for the bias of ML estimators of Gumbel parameters. Koch (1991) investigated bias error in maximum likelihood estimation.

Let f(x; $a_1, a_2, ..., a_m$) be a probability density function (pdf) of the random variable X with parameters a_i, i=1, 2,..., m, to be estimated. For a random sample of data, $x_i, x_2, ...x_n$, drawn from this probability density, the joint pdf is defined as

$$f(x_i, x_2, x_3, ...x_n; a_1, a_2, ...a_m) = \overset{n}{\underset{i=1}{\pi}} f(x_i; a_1, a_1, ...a_m) \qquad (2.97)$$

Interpreted conceptually, the probability of obtaining a given value of X, say x_1, is proportional to $f(x; a_1, a_1, ..., a_m)$. Likewise, the probability of obtaining the random sample $x_1, x_2, ..., x_n$ from the population of X is proportional to the product of the individual probability densities or the joint pdf. This joint pdf is called the likelihood function, denoted by L,

$$L = \overset{n}{\underset{i=1}{\pi}} f(x_1; a_1, a_2, ...a_m) \qquad (2.98)$$

where the parameters a_1, i=1,2,...m, are unknown.

By maximizing the likelihood that the sample under consideration is the one that would be obtained if n random observations were selected from $f(x; a_1, a_2, ...a_m)$, the unknown parameters are determined, and hence the name the method of maximum likelihood estimation (MLE). The values of parameters so obtained are known as MLE estimators. Since the logarithm of L (ln L) attains its maximum for the same values of $a_1, i = 1,2,...,m$, as does L, the MLE function can also be expressed as

$$\ln L = L^* = \ln \overset{n}{\underset{i=1}{\pi}} f(x_1; a_1, a_2, ..., a_m) = \sum_{i=1}^{n} \ln f(x_1; a_1, a_2, ...a_m) \qquad (2.99)$$

Frequently ln [L] is maximized, for it is many times easier to find the maximum of the logarithm of the maximum likelihood function than that of the normal L.

The procedure for estimating the parameters or determining the point where the MLE function achieves its maximum involves differentiating L or ln L partially with respect to each parameter and equating each differential to zero. This results in as many equations as the number unknown parameters. For m unknown parameters, we get

$$\begin{aligned} \frac{\partial L(a_1, a_2, ...a_m)}{\partial a_1} &= 0 \\ \frac{\partial L(a_1, a_2, ...a_m)}{\partial a_2} &= 0 \\ &\vdots \\ \frac{\partial L(a_1, a_2, ...a_m)}{\partial a_m} &= 0 \end{aligned} \qquad (2.100)$$

These m equations in m unknowns are then solved for the m unknown parameters.

2.7 Method of Least Squares

The method of least squares (MOLS) is one of the most frequently used parameter estimation methods in hydrology. Natale and Todini (1974) presented a constrained MOLS for linear models in hydrology. Williams and Yeh (1983) described MOLS and its variants for use in rainfall-runoff models. Jones (1971) linearized weight factors for least squares (LS) fitting. Shrader et al. (1981) deveoped a mixed-mode version of MOLS and applied it to estimate parameters of log-normal distribution. Snyder (1972) repoted on fitting of distribution functions by non-linear least squares. Stedinger and Tasker (1985) performed regional hydrologic analysis using ordinary, weighted and generalized least squares.

Let there be a function $y=f(x;a_1,a_2,...,a_m)$ where $a_i, i=1,2,...,m$, are parameters to be estimated. The method of least squares (MOLS) involves estimating parameters by minimizing the sum of squares of all deviations between observed and computed values of y. Mathematically, this sum S can be expressed as

$$S=\sum_{i=1}^{n} d_1^2=\sum_{i=1}^{n} [y_0(i)-y_c(i)]^2$$

$$=\sum_{i=1}^{n} [y_o(i)-f(x;a_1,a_2,...,a_m)]^2 \tag{2.101}$$

where $y_0(i)$ is the i-th observed value of Y, $y_c(i)$ is the i-th computed value of Y, and n>m is the number of observations. The minimum of S in equation (2.101) can be obtained by differentiating S partially with respect to each parameter and equating each differential to zero:

$$\frac{\partial \sum_{i=1}^{n} [y_0(i)-f(x_i;a_1,a_2,...a_m)]^2}{\partial a_1}=0$$

$$\frac{\partial \sum_{i=1}^{n} [y_0(i)-f(x_i;a_1,a_2,...a_m)]^2}{\partial a_2}=0 \tag{2.102}$$

$$\frac{\partial \sum_{i=1}^{n} [y_0(i)-f(x_i;a_1,a_2,...,a_m)]^2}{\partial a_m}=0$$

Thus, m equations, usually called the normal equations, are obtained, which are then solved for estimation of m parameters.

2.7.1 MATRIX REPRESENTATION OF MOLS

Many of the computations performed in fitting multiple variable models to observed data can be expressed more efficiently using the matrix notation. The normal equations in MOLS can often be reduced to the form

$$XA = Y \qquad (2.103)$$

where X is a matrix of (n+1, m+1) dimensions, A is a vector of the parameters having (m+1,1) dimensions, and Y is a vector having (n+1,1) dimensions. In this section the upper case letters will signify matrices or vectors. The deviations between observed and computed values of y can be written as

$$D = Y - XA \qquad (2.104)$$

where D is a vector of deviations $d_i, i=1,2,..,n,$ between observed and computed values.

The sum of squares of deviations can be obtained by using the inner product which is obtained by multiplying D by its transpose D^T:

$$\begin{aligned}\sum d_i^2 &= D^T D \\ &= [Y^T - A^T X^T][Y - XA] \qquad (2.105) \\ &= Y^T Y - Y^T XA - A^T X^T Y + A^T X^T XA\end{aligned}$$

Since superscript T signifies the transpose, where A and Y are column vectors their transposes will be row vectors. Thus, the second and third terms on the right side of equation (2.105) will be scalar in form. Since a scalar and its transpose are the same, we get

$$\sum d_i^2 = Y^T Y - 2A^T X^T Y + A^T X^T XA \qquad (2.106)$$

Differentiating equation (2.106) with respect to A and equating to zero, we obtain

$$X^T XA = X^T Y \qquad (2.107)$$

The matrix $[X^T X]$ and the vector $[X^T Y]$ are calculated from observed values of the variates. The parameters can then be determined as

$$A = [X^T X]^{-1} X^T Y \qquad (2.108)$$

Equations (2.108) expresses in compact form the least squares solution.

2.8 Comparison with Entropy Method

It may be useful to briefly compare the POME method of parameter estimation with the method of moments (MOM) and method of maximum likelihood estimation (MLE), two of the most frequently used methods. To contrast the POME method with the MLE method, we consider the case of a general pdf f $(x;\theta)$ where θ represents a family of parameters $\lambda_1, i = 1, 2, ..., M$. In the MLE method we construct the likelihood function L

$$L = \prod_{i=1}^{N} f(x_i;\theta) \tag{2.109}$$

and maximize either L or ln L. Here N is the sample size. Taking logarithm of (2.109)

$$\ln L = \sum_{i=1}^{N} \ln f(x_i;\theta) \tag{2.110}$$

By differentiating ln L with respect to each of the parameters λ_i separately and equating to zero, we guarantee as many equations as the number of parameters. We solve these equations to obtain parameter estimates.

If, however, we multiply (2.110) by -(1/N) then

$$-\frac{1}{N}\ln L = -\frac{1}{N}\sum_{i=1}^{N} \ln f(x;\theta) = -\sum_{i=1}^{N} \frac{1}{N} f(x_i;\theta) \tag{2.111}$$

Recall that

$$I[f] = -\sum_{i=1}^{N} f(x_i;\theta) \ln f(x:\theta); \sum_{i=1}^{N} f(x_i;\theta) = 1 \tag{2.112}$$

On comparing equation (2.111) with equation (2.112) it is seen that

$$I[f] = -\frac{1}{N}\ln L \tag{2.113}$$

provided ln f $(x_i;\theta)$ is uniformly weighted over the entire sample. The POME method involves population expectations, whereas the MLE method involves sample averages. If population is replaced by a sample then the two methods would yield the same parameter estimates. To fully appreciate the signifance of equation (2.113), we consider the case of an exponential distribution

$$f(x_i) = \alpha \exp(-\alpha x_i) \tag{2.114}$$

Then

$$I[f] = -\sum_{k=1}^{N} \alpha \exp(-\alpha x_i) \ln [\alpha \exp(-\alpha x_i)] \tag{2.115}$$

$$= -\ln \alpha + \alpha E[x]$$

By maximizing I[f] with respect to α

$$\alpha = 1/E[x] \tag{2.116}$$

On the other hand, the log-likelihood function for the exponential distribution is

$$\ln L = N \ln \alpha - \alpha \sum_{i=1}^{N} x_i \tag{2.117a}$$

By maximizing ln L of equation (2.117a), we get

$$\alpha = 1/[\sum_{i=1}^{N} x_i/N] = 1/\bar{x} \tag{2.117b}$$

The difference in the two estimates of α given by equations (2.116) and (2.117) is that the POME method uses the expectation of X or the population mean, whereas the MLE method uses the average of X or sample mean.

This result can be extended to a very general case of f (x), written as

$$f(x) = A x^k \exp[-\sum_{i=1}^{m} \lambda_i y_i(x)] \tag{2.118}$$

The Shannon entropy functional (SEF) of this function is

$$I[f] = -\ln A - k E[\ln x] + \sum_{i=1}^{m} \lambda_i \sum_{j=1}^{N} E[y_i(x_j)] \tag{2.119}$$

On the other hand, the log-likelihood function of equation (2.118) is

$$\ln L = \sum_{i=1}^{N} \ln[A x_i^k \exp[-\sum_{j=1}^{m} \lambda_i y_i(x_j)]]$$

$$= \sum_{i=1}^{N} \ln A + k \sum_{i=1}^{N} \ln x_i - \sum_{i=1}^{N} \lambda_1 \sum_{j=1}^{m} y_i (x_j) \tag{2.120}$$

Multiplying equation (2.120) by -(1/N) throughout, we get

$$-\frac{1}{N} L = -\ln A - k \sum_{i=1}^{N} \frac{\ln x_i}{N} - \sum_{i=1}^{N} \frac{\lambda_i}{N} \sum_{j=1}^{m} y_i(x_j) \tag{2.121}$$

Equation (2.119) is the same as equation (2.121) if E [·] terms are replaced by corresponding averages.

To compare the POME method with MOM is not straightforward anf requires further research. The MOM is not variational in character, whereas the POME method is. If the constraints in the entropy method are ordinary moments then the parameter estimates by the two methods would be the same. This is, for example, true in the case of exponential and normal distributions. If the constraints are other than ordinary moments which is true of most distributions then the two methods would likely be expected to yield different parameter estimates and it is not known what conditions, if any, would there be for differences in the parameter estimates to vanish.

2.9 Problems of Parameter Estimation

We estimate parameters of a distribution function from sample values. There are, of course, myriad ways by which to obtain parameter estimates. The sample data may contain errors, the hypotheses underlying the method of parameter estimation may not yield accurate estimates, and there may be truncation and roundoff errors. These sources of errors may result in errors in parameter estimates. Each estimate of a parameter is a function of sample values which are observations of a random variable. Thus, the parameter estimate itself is a random variable having its own sampling distribution. An estimate obtained from a given set of values can be regarded as an observed value of the random variable. Thus, the goodness of an estimate can be judged from its distribution. A question then arises: How should we best use the data to form estimates? This immediately raises another question: What do we mean by the best estimates? Also, are these estimates unique? How do we select the best parameter estimator if there is one? A number of statistical properties are available by which to address the above questions. Troutman (1985a, b) investigated errors and their sources in complex conceptual rainfall-runoff models. By treating errors as random variables and defining the probabilistic structure of the errors he estimated bias in parameter estimates and related it to model error and input error. Kitanidis (1986) estimated parameter uncertainty in estimation of spatial functions using Bayesian analysis. Field (1985) described the concept of robustness. Kuczera (1982a, b, c) applied this concept to parameter estimation for conceptual catchment models. In a series of papers, Fiering (1982a, b, c,d) investigated and theorized a similar concept, called resilience. Excellent discussions on parameter uncertainty, errors, robustness, bias, resilience and the like have been reported by Kuczera (1982a,b,c, 1983a,1983b), Sorooshian and Gupta (1983), Gupta and Sorooshian (1983) among others.

2.9.1 BIAS

Let the parameter be a and its estimate a_c. The estimate a_c is called an unbiased estimate of a if E (a_c) = a. In general, an estimate will have a certain bias b(a) depending on a so that

$$E(a_c) = a + b(a) \tag{2.122}$$

Obviously, b(a)=0 for an unbiased estimate. It should, however, be noted that an individual a_c is not equal to or even close to a even if b(a)=0. It simply implies that the average of many independent estimates of a will be equal to a.

The bias in a given quantity is usually measured in dimensionless terms and is often referred to as standardized bias (or BIAS). Thus, BIAS is defined as

$$BIAS = \frac{E\ (\hat{a}) - a}{a} \tag{2.123}$$

where $\hat{a}$ is an estimate of parameter or quantile of a. In Monte Carlo experimentation, large numbers of samples of different sizes are generated from a given population. For each sample, then, an estimate of a is obtained. If there are, say, 1000 samples of a given size generated then there are 1000 values of parameter or quantile a. Thus, E(a) is the average of the 1000 estimates of a for a given sample size and is estimated as

$$E\ (\hat{a}) = \frac{\sum_{i=1}^{n} \hat{a}_i}{n}. \tag{2.124}$$

where n is the number of samples generated or number of values of the a estimate.The value of a in equation (2.123) is the true value of a or the value of parameter a of the population.

2.9.2 CONSISTENCY

Let there be a sample be of size n. The estimate a_c is called a consistent estimate of a if it converges to a with probability one as n tends to infinity. Because many unbiased estimates have variances of the type

$$Var(a_c) \cong C/(n)^{0.5} \tag{2.125}$$

the condition of consistency is satisfied in most cases. Here C is constant. What is, however, desirable in practice is to have $Var(a_c)$ as small as possible. This would imply that the probability density function of a_c would be more concentrated about a.

2.9.3 EFFICIENCY

An estimate a_c of a is said to be efficient if it is unbiased and its variance is at least as small as that of any other unbiased estimate of a. If there are two estimates of a, say a_1 and a_2, then the relative efficiency of a_1 with respect to a_2 is defined as

$$e = \frac{E[a_1 - a]^2}{E[a_2 - a]^2} \leq 1 \tag{2.126}$$

if $E[a_2 - a]^2 > E[a_1 - a]^2$, then e ≤ 1. An efficient estimate has e=1. If an efficient estimate exists, it may be approximately obtained by use of the MLE or entropy method.

2.9.4 SUFFICIENCY

An estimate a_c of a is said to be sufficient if it uses all of the information that is contained in the sample. More precisely, let a_1 and a_2 be two independent estimates of a. a_1 is considered a sufficient estimate if the joint probability distribution of a_1 and a_2 has the property

$$f(a_1, a_2) = f(a_1) f(a_2 \mid a_1) = f(a_1) K(x_1, x_2, \ldots x_n) \tag{2.127}$$

in which f(a_1) is the distribution of a_1, $f(a_2 \mid a_1)$ is the conditional distribution of a_2 given a_1, and $K(x_1, x_2, \ldots x_n)$ is not a function of a but only of x_i's. If equation (2.127) holds, then a_2 does not produce any new information about a which is not already contained in a_1. In this case a_1 is a sufficient estimate.

2.9.5 RESILIENCE

The concept of resilience (Fiering, 1982a, b, c, d) is analogous to the statistical notion of robustness, meaning that even if an unlikely event occurs, the decision has a high probability of being correct or at least good enough. Another approach to resilience and robustness is based on partial and total derivatives of the system response. The partial derivative of system reponse with respect to a decision variable measures the sensitivity of response to that variable alone, all other decision variables being held constant. If the partial derivative is small, the system is robust with respect to such changes. If the partial derivative is not small, the system response need not change significantly because changes in other decision variables might accommodate unanticipated change in the dependence of response on that variable. The total derivative is constituted by the sum of the products of partial derivatives of system response to decision variables and total derivatives of operating decisions with respect to decision variables. Thus, this is a measure of the system's ability to adjust, to utilize redundant capabilities or a measure of resilience of the given system design.

2.9.6 STANDARD ERROR

Another dimensionless performance measure frequently used in hydrology is the standard error (SE), defined as

$$SE = \frac{\sigma(\hat{a})}{a} \tag{2.128}$$

where σ (.) denotes the standard deviation of a and is computed as

$$\sigma(\hat{a}) = [\frac{1}{n-1} \sum_{i=1}^{n} \{\hat{a}_i - E(\hat{a}_i)\}^2]^{1/2} \tag{2.129}$$

where the summations are over n estimates $\hat{a}$ of a. In Monte carlo experiments, referred to above, for each sample size, a value of SE is obtained. Thus this measure is similar to the coefficient of variation.

2.9.7 ROOT MEAN SQUARE ERROR

The root mean square error (RMSE) is one of the most frequently employed performances measures. and is defined as

$$RMSE = \frac{E[(\hat{a} - a)^2]^{1/2}}{a} \tag{2.130}$$

where E[.] is the expectation of [.]. It can be shown that RMSE is related to BIAS and SE as

$$RMSE = [\frac{n-1}{n} SE^2 + BIAS^2]^{1/2} \tag{2.131}$$

2.9.8 ROBUSTNESS

Kuczera (1982a, b, c) defined a robust estimator as the one that is resistant and efficient over a wide range of population fluctuations. Two criteria for resistant estimator are mini-max and minimum average RMSE. According to the mini-max criteria, the maximum RMSE for all population cases should be minimum. Thus, for a resistant estimator the average RMSE as well as the maximum RMSE should be minimum.

2..9.9 RELATIVE MEAN ERROR

Another measure of error in assessing the goodness of fit of hydrologic models is the relative mean error (RME) defined as

$$RME = \frac{1}{N}(\sum_{i=1}^{N}[\frac{Q_0 - Q_c}{Q_0}]^2)^{0.5} \tag{2.232}$$

in which N is sample size, Q is observed quantity of a given probability and Q_c is computed quantity of the same probability. Also used sometimes is the relative absolute error defined as

$$RAE = \frac{1}{N}\sum_{i=1}^{N}|\frac{Q_0 - Q_c}{Q_c}| \tag{2.233}$$

References

Ashkar, F., Bobee, B., Leroux, D. and Morisette, D., 1988. The generalized method of moments as applied to the generalized gamma distribution. *Stochastic Hydrology and Hydraulics*, Vol. **2**, pp. 161-174.

Bobee, B., Perreault, L. and Ashkar, F., 1993. Two kinds of moment diagrams and their applications in hydrology. *Stochastic Hydrology and Hydraulics*, Vol. **7**, pp. 41-65.

Bree, T., 1978a. The stability of parameter estimation in the linear model. *Journal of Hydrology*, Vol. **37**, pp. 47-66.

Bree, T., 1978b. The general linear model with prior information. *Journal of Hydrology*, Vol. **39**, pp.113-127.

Clarke, R.T., 1996. Residual maximum likelihood (REML) methods for analysing hydrological data series. *Journal of Hydrology*, Vol. **182**, pp. 277-295.

Cohen, A.C. and Whitten, B., 1982. Modified maximum likelihood and modified moment estimators for the three-parameter Weibull distribution. *Communications in Statistics-Theoretical Methods*, Vol. **11**, o. 23, pp. 2631-2656.

Cunnane, C., 1989. Statistical distribution for flood frequency analysis. *WMO Operational Hydrology Report* No-**33**, WMO-No 718, Geneva, Switzerland.

Diskin, M.H., 1967. A Laplace transform proof of the theorem of moments for the instantaneous unit hydrograph. *Water Resources Research*, Vol. **3**, No. 2, pp. 385-388.

Diskin, M.H. and Boneh, A., 1968. Moments of the input, output, and impulse response functions of linear systems about arbitrary points. *Water Resources Research*, Vol. **4**, No. 4, pp. 727-735.

Dooge, J.C.I., 1973. *Linear theory of hydrologic systems*. Technical Bulletin No. **1468**, 327 pp., Agricultural Research Service, U.S. department of Agriculture, Washington, D.C.

Douglas, J.R., Clarke and Newton, S.G., 1976. The use of likelihood functions to fit conceptual models with more than one dependent variable. *Journal of Hydrology*, Vol. **29**, pp. 181-198.

Duan, Q., Sorooshian, S. And Gupta, V.K., 1988. A maximum likelihood criterion for use with data collected at unequal time intervals. *Water Resources Research*, Vol. **24**, No. 7, pp. 1163- 1173.

Dubey, S.D., 1967. Some percentile estimators for Weibull parameters. *Technometrics*, Vol. **9**, pp. 119-129.

Field, C., 1985. Concepts of robustness. Chapter 15 in *A Celebration of Statistics*, edited by A.C. Atkinson and S.E. Fienberg, pp. 369-375, Springer-Verlag, Heidelberg, Germany.

Fiering, M.B. 1982 a. A screening model to quantify resilience. *Water Resources Research*, Vol. **18**, No. 1, pp. 27-32.

Fiering, M.B. 1982b. Alternative indices of resilience. *Water Resources Research*, Vol. **18**, No. 1, pp. 33-39.

Fiering, M.B. 1982c. Estimates of resilience indices by simulation. *Water Resources Research*, Vol. **18**, No. 1, pp. 41-50.

Fiering, M.B. 1982d. Estimating resilience by canonical analysis. *Water Resources Research*, Vol. **18**, No. 1, pp. 51-57.

Greenwood, J.A., Landwehr, J.M., Matalas, N.C. and Wallis, J.R., 1979. Probability-weighted moments: definition and relation to parameters of several distributions expressible in inverse form. *Water Resources Research*, Vol. **15**, pp. 1049-1054.

Gupta, V.K. and Sorooshian, S., 1983. Uniqueness and observability of conceptual rainfall-runoff model parameters: The percolation process examined. *Water Resources Research*, Vol. **19**, No. 1, pp. 29-276.

Gupta, V.K. and Sorooshian, S., 1985. The rel;ationship between data and the precision of parameter estimates of hydrologic models. *Journal of Hydrology*, Vol. **81**, pp. 57-77.

Haktanir, T., 1996. Probability-weighted moments without plotting position formula. *Journal of Hydrologic Engineering*, Vol. **1**, No. 2, pp. 89-91.

Harley, B.M., 1967. *Linear routing in uniform open channels*. M. Eng. Sc. Thesis, University College, Cork, Ireland.

Hosking, J.R.M., 1985. A correction for the bias of maximum likelihood estimators of Gumbel parameters-comment. *Journal of Hydrology*, Vol. **78**, pp. 393-396.

Hosking, J.R.M., 1986. *The theory of probability-weighted moments*. Technical Report RC **12210**, Mathematics, 160 pp., IBM Thomas J. Watson Research Center, Yorktown Heights, New York.

Hosking, J.R.M., 1992. Moments or L-moments. An example comparing two measures of distributional shape. *American Statistician*, Vol, **46**, No. 3, pp. 186-189.

Hosking, J.R.M., 1990. L-Moments: Analysis and estimation of distribution using linear combination of order statistics. *Journal of Royal Statistical Society*, Series B, Vol. **52**, No. 1, pp. 105-124.

Hosking, J.R.M. and Wallis, J.R., 1991. *Some statistics useful in regional frequency analysis*. Research Report RC **1709623** AP, IBM Research Division, T.J. Watson Research Center, Yorkstown Heights, N.Y.

Hosking, J.R.M. and Wallis, J.R., 1987. An "index-flood" procedure for regional rainfall frequency analysis. *EOS Transactions*, AGU, Vol. **68**, pp. 312.

Hosking, J.R.M. and Wallis, J.R., 1995. A comparison of unbiased and plotting position estimators of L moments. *Water Resources Research*, Vol. **31**, No. 8, pp. 2019-2025.

Johnson, N.L. and Kotz, S., 1985. Moment ratios. *Encyclopedia of Statistical Sciences*, Vol. **1**, pp. 603-604.

Jones, L.E., 1971. Linearizing weight factors for least quares fitting. *Journal of the Hydraulics Division,* ASCE, Vol. **97**, No. HY5, pp. 665-675.

Kappenman, R.F., 1985. Estimation for the three-parameter Weibull, lognormal, and gamma distributions. *Computational Statistics and Data Analysis,* Vol. **3**, No. 1, pp. 11-23.

Kitanidis, P. K., 1986. Parameter uncertainty in estimation of spatial functions: Bayesian analysis. *Water Resources Research*, Vol. **22**, No. 4, pp. 499-507.

Kitanidis, P.K. and Lane, R.W., 1985. Maximum likelihood parameter estimation of hydrologic spatial processes by the Gauss-Newton method. *Journal of Hydrology*, Vol. **79**, pp. 53-71.

Kline, D.E. and Bender, D.A., 1990. Maximum likelihood estimation for shifted Weull and lognormal distributions. *Transactions in Agriculture*, Vol. **33**, No.1, pp. 330-335.

Koch, S.P., 1991. Bias error in maximum likelihood estimation. *Journal of Hydrology*, Vol. **122**, pp. 289-300.

Kroll, C.N. and Stedinger, J.R., 1996. Estimation of moments and quantiles using censored data. *Water Resources Research*, Vol. **32**, No. 4, pp. 1005-1012.

Kuczera, G., 198a. Combining site-specific and regional information: An empirical Bayes approach. *Water Resources Research*, Vol. **18**, No. 2, pp. 306-314.

Kucczera, G., 1982b. Robust flood frequency models. *Water Resources Research*, Vol. **18,** No. 2, pp. 315-324.

Kuczera, G., 1982c. On the relationship between the reliability of parameter estimates and hydrologic time series data used in calibration. *Water Resources Research*, Vol. **18**, No. 1, pp. 146-154.

Kuczera, G., 1983a. Improved parameter inference in catchment models: 1. Evaluating parameter uncertainty. *Water Resources Research*, Vol. **19**, No. 5, pp. 1151-1162.

Kuczera, G., 1983b. Improved parameter inference in catchment models: 2.Combining different kinds of hydrologic data and testing their compatibity. *Water Resources Research*, Vol. **19**, No. 5, pp. 1163-1172.

Landwehr, J.M., Matalas, N.C. and Wallis, J.R., 1979a. Probability-weighted moments compared with some traditional techniques in estimating Gumbel parameters and quantiles. Water Resources Research, Vol. 15, pp. 1055-1064.

Landwehr, J.M., Matalas, N.C. and Wallis, J.R., 1979b. Estimation of parameters and quantiles of Wakeby distributions. *Water Resources Research*, Vol. **15,** pp. 1361-1379; correction, p. 1672.

Nash, J.E., 1959. A systematic determination of unit hydrograph parameters. *Journal of Geophysical Research*, Vol. **64**, No. 1, pp. 111-115.

Natale, L. and Todini, E., 1974. *A constrained parameter estimation technique for linear models in hydrology*. Publication No. **13**, Institute of Hydraulics, University of Pavia, Pavia, Italy.

O'meara, W.A., 1968. *Linear routing of lateral inflow in uniform open channels*. M.Eng. Sc. Thesis, University College, Cork, Ireland.

Phien, H.N. and Jivajirajah, T., 1984. Fitting the S_B curve by the method of maximum likelihood. *Journal of Hydrology*, Vol. **67**, pp. 67-75.

Rao, A.R. and Hamed, K.H., 1994. Frequency analysis of upper Cauvert flood data by L-moments. *Water Resources Management*, Vol. **8**, pp. 183-201.

Rao, A.R. and Hamed, K.H., 1997. Regional frequency analysis of Wabash River flood data by L-moments. *Journal of Hydrologic Engineering*, Vol. **2**, No. 4, pp. 169-179.

Rao, A.R. and Mao, L.T., 1987. An investigation of the instrumental varaible-approximate maximum likelihood method of modeling and forecasting daily flows. *Water Resources Manageemnt*, Vol. **1**, pp. 79-106.

Rao, D.V., 1980. Log-Pearson type 3 distribution: method of mixed moments. *Journal of Hydraulics Division, ASCE*, Vol. **106**, No. HY6, pp. 999-1019.

Rao, D.V., 1983. Estimating log Pearson parameters by mixed moments. *Journal of Hydraulic Engineering*, Vol. **109**, No. 8, pp. 1118-1132.

Shrader, M.L., Rawls, W.J., Snyder, W.M. and McCuen, R.H., 1981. Flood peak regionalization using mixed-mode estimation of the parameters of the log-normal distribution. *Journal of Hydrology*, Vol. **52**, pp. 229-237.

Singh, V. P., 1988. *Hydrologic Systems, Vol.1: Rainfall-Runoff Modeling*. Prentice hall, Englewood Cliffs, New Jersey.

Snyder, W.M., 1972. Fitting of distribution functions by nonlinear least squares. *Water Resources Research*, Vol. **8**, No. 6, pp. 1423-1432.

Sorooshian, S. And Gupta, V.K., 1983. Automatic calibration of conceptual rainfall-runoff models: the question of parameter observability and uniqueness. *Water Resources Research*, Vol. **19**, No. 1, pp. 260-268.

Sorooshian, S., Gupta, V.K. and Fulton, J.L., 1983. Evaluation of maximum likelihood parameter estimation techniques for conceptual rainfall-runof models: Influence of calibration data variability and length on model credibility. *Water Resources Research*, Vol. **19**, No. 1, pp. 251-259.

Stedinger, J.R. and Tasker, G.D., 1985. Regional hydrologic analysis: 1. Ordinary, weighted and generalized least squares compared. *Water Resources Research*, Vol. **21**, No. 9, pp. 1421-1432.

Troutman, B.M., 1985a. Errors and parameter estimation in precipation-runoff modeling: 1.

Theory. *Water Resources Research*, Vol. **21**, No. 8, pp. 1195-1213.

Troutman, B.M., 1985b. Errors and parameter estimation in precipation-runoff modeling: 2. Case study. *Water Resources Research*, Vol. **21**, No. 8, pp. 1214-1222.

Van des Nes, T.J. and Hendriks, M.H., 1971. *Analysis of a linear distributed modelof surface model.* Report No. **1**, Laboratory of Catchment Hydraulics and Hydrology, Agricultural University of Wageningen, The Netherlands.

Vogel, R.M.S. and Fennessey, N.M., 1993. L-moment diagrams should replace product moment diagrams. *Water Resources Research*, Vol. **29**, No. 6, pp. 1745-1752.

Vogel, R.M. and Wilson, I., 1996. Probability distribution of annual maximum, mean, and minimum streamflows in the United States. *Journal of Hydrologic Engineering,* Vol. **1**, No. 2, pp. 69- 76.

Vogel, R.M.S., Thomas W.O., and McMahon, T.A., 1993a. Flood flow frequency model selection in southwestern U.S.A. *Journal of Water Resources Planning and Management*, ASCE, Vol. **119**, No. 3, pp. 353-366

Vogel, R.M.S., McMahon, T.A., and Chiew, F.H.S., 1993b. Flood flow model selection in Australia. *Journal of Hydrology,* Vol. 146, pp. 421-450.

Wang, S.X. and Adams, B.J., 1984. *Parameter estimation in floodfrequency analysis.* Publication **84-02,** Department of Civil Engineering, University of Toronto, Toronto, Canada.

Wang, Q.J., 1997. LH moments for statistical analysis of extreme events. *Water Resources Research,* Vol. **33**, No. 12, pp. 2841-2848.

Williams, B.J. and Yeh, W.W.-G., 1983. Parameter estimation in rainfall-runoff models. *Journal of Hydrology*, Vol. **63**, pp. 373-393.

CHAPTER 3

UNIFORM DISTRIBUTION

Uniform distribution is the simplest statistical distribution. Although there is hardly any hydrologic variable that follows a uniform probability distribution, it is invoked in a variety of applications. For example, in Bayesian statistical modeling in hydrology it is frequently used as a prior distribution. In systems hydrology, uniform distribution is the pulse function obtained by subtracting two step functions lagged by the length of the uniform distribution. The pulse function is a key to deriving the unit hydrograph theory. The instantaneous unit hydrograph of the rational method, used in urban hydrology, is a uniform distribution (Singh, 1988). Of all the statistical distributions, uniform distribution has the highest entropy. In river morphology, when a river approaches equilibrium or dynamic equilibrium, its characteristics tend to follow a uniform distribution. Under equilibrium, rivers follow the minimum rate of energy dissipation. Furthermore, a river constantly adjusts its cross-sectional geometry and longitudinal profile to accommodate the influx of water and sediment coming from its drainage basin, and this adjustment is in accordance with the principle of maximum entropy. Thus, there is a close link between equilibrium and uniform distribution and then between maximum entropy (uniform distribution) and minimum rate of energy dissipation. This link plays a fundamental role in river engineering and training works, river morphology, evolution of deltas, etc.

A random variable X is defined to have a uniform distribution if its probability density function (pdf) is given by

$$f(x) = \frac{1}{b-a}, \quad a \leq x \leq b \tag{3.1}$$

Equation (3.1) can also be expressed in terms of the unit step function or Heavyside function, u_1(x), as

$$f(x) = \frac{1}{b-a}[u_1(x-a) - u_1(x-b)] \tag{3.2}$$

By definition f(x) is a rectangular pulse of length (b-a) and height 1/(b-a). The cumulative distribution function (cdf) of the uniform distribution can be expressed as

$$F(x) = \frac{x-a}{b-a} \tag{3.3}$$

V. P. Singh, *Entropy-Based Parameter Estimation in Hydrology*

The uniform distribution does not have any parameter, for once one bound is known the other bound is fixed by virtue of the rectangularity of the distribution. Thus, the parameter estimation in this case is trivial. Nevertheless it is didactic to apply the entropy method.

3.1 Specification of Constraint

Taking logarithm of equation (3.1) to the base 'e', one gets

$$\ln f(x) = -\ln(b-a) \tag{3.4}$$

Multiplying equation (3.3) by [-f(x)] and integrating between a and b, one gets

$$-\int_a^b f(x)\ln x\, dx = \ln(b-a)\int_a^b f(x)dx \tag{3.5}$$

Following Singh et al. (1985,1986), the constraint appropriate for equation (3.1) can be written as

$$\int_a^b f(x)\,dx = 1 \tag{3.6}$$

which is the total probability law.

3.2 Construction of Zeroth Lagrange Multiplier

The least-biased pdf consistent with equation (3.6), determined by the principle of maximum entropy (POME), takes the form:

$$f(x) = \exp(-\lambda_0) \tag{3.7}$$

where λ_0 is the Lagrange multiplier.

Substitution of equation (3.7) in equation (3.6) yields

$$\int_a^b f(x)\,dx = \int_a^b \exp(-\lambda_0)\,dx = 1 \tag{3.8}$$

or

$$\exp(\lambda_0) = \int_a^b dx = b - a \tag{3.9}$$

which is the partition function.

The zeroth Lagrange multiplier λ_0 is given by equation (3.9) as

$$\lambda_0 = \ln(b-a) \tag{3.10}$$

Inserting equation (3.9) into equation (3.7), the result is

$$f(x) = \frac{1}{b-a} \tag{3.11}$$

which is the same as equation (3.1).

3.3 Estimation of Parameter

The uniform distribution is a none-parameter distribution. If the lower bound of X is known, the upper unknown parameter b is estimated such that

$$\int_a^b \frac{dx}{b-a} = 1 \tag{3.12}$$

If the lower limit of X, a, is zero, then

$$\int_0^b \frac{dx}{b} = 1 \tag{3.13}$$

The quantity 1/b specifies the height or intensity of the rectangular pulse over the interval (0, b).

3.4 Distribution Entropy

The entropy, I(x), of the uniform distribution can be expressed as

$$I(x) = -\int_a^b f(x) \ln f(x) dx = \ln(b-a) \tag{3.14}$$

For a continuous random variable bounded by a finite interval, the uniform probability density provides the maximum entropy.

References

Singh, V.P., 1988. *Hydrologic Systems, Vol.1: Rainfall-Runoff Modeling*. Prentice Hall, Inc., Englewood Cliffs, New Jersey, U.S.A.

Singh, V.P., Rajagopal, A.K. and Singh, K., 1986. Derivation of some frequency distributions using the principle of maximum entropy (POME). *Advances in Water Resources*, Vol. **9**, pp. 91-106.

Singh, V.P., Singh, K. and Rajagopal, A.K., 1985. Application of the principle of maximum entropy (POME) to hydrologic frequency analysis. *Completion Report* **06,** 144 p., Louisiana Water Resources Research Institute, Louisiana State University, Baton Rouge, Louisiana.

CHAPTER 4

EXPONENTIAL DISTRIBUTION

The exponential distribution is a basic distribution for constructing a number of other distributions. For example, the gamma distribution is obtained from the distribution of the sum of random variables where each variable follows an exponential distribution. Indeed, it is the simplest member of the gamma family of distributions and can be considered as a special case of the two-parameter gamma distribution. It is a one-parameter distribution and has found widespread application in hydrology and water resources. The instantaneous unit hydrograph of a linear reservoir, frequently used in systems hydrology, is exponential (Singh, 1988). The exponential distribution is often used for frequency analysis of rainfall depth, intensity and duration, and number of rainfall events (Eagleson, 1982). It is frequently used in biology, genetics, quantum mechanics, reliability engineering, to name but a few.

A random variable X is defined to have an exponential distribution if its probability density function (pdf) is given by

$$f(x) = a \exp(-ax), \ a > 0, \ x > 0 \tag{4.1}$$

where a is a parameter. The exponential distribution is a one-parameter distribution. Sometimes equation (4.1) is also referred to as negative exponential distribution. Its cumulative distribution function (cdf) can be expressed as

$$F(x) = 1 - \exp(-ax) \tag{4.2a}$$

The inverse form of equation (4.2a) is given as

$$x = -\frac{1}{a}\ln(1 - F) \tag{4.2b}$$

4.1 Ordinary Entropy Method

4.1.1 SPECIFICATION OF CONSTRAINTS

Taking logarithm of equation (4.1) to the base 'e', one obtains

$$\ln f(x) = \ln a - ax \tag{4.3a}$$

Multiplying equation (4.3a) by [-f(x)] and integrating between 0 and ∞, one gets

$$-\int_0^\infty f(x)\, ln f(x)\, dx = -ln\, a \int_0^\infty f(x)\, dx + a \int_0^\infty x f(x)\, dx \tag{4.3b}$$

Following Singh et al. (1985, 1986), the equations of constraints appropriate for equation (4.1) can be obtained from equation (4.3b) as:

$$\int_0^\infty f(x)\ dx = 1 \tag{4.4}$$

$$\int_0^\infty x\, f(x)\ dx = \bar{x} \tag{4.5}$$

where $\bar{x}$ is the mean or the first moment of the distribution about its origin.

4.1.2 CONSTRUCTION OF ZEROTH LAGRANGE MULTIPLIER

The least-biased pdf corresponding to the principle of maximum entropy (POME) and consistent with equations (4.4) and (4.5) takes the form:

$$f(x) = \exp(-\lambda_0 - \lambda_1 x) \tag{4.6}$$

where λ_0 and λ_1 are Lagrange multipliers. Substitution of equation (4.6) in equation (4.4) yields

$$\int_0^\infty f(x)\ dx = \int_0^\infty exp\ (-\lambda_0 - \lambda_1 x)\ dx = 1$$

or

$$exp\ (\lambda_0) = \int_0^\infty exp\ (-\lambda_1 x)\ dx \tag{4.7}$$

Equation (4.7) defines the partition function for the exponential distribution. The zeroth Lagrange multiplier λ_0 is given as

$$\lambda_0 = -\ln \lambda_1 \tag{4.8}$$

From equation (4.7) we also get the zeroth Lagrange multiplier as

$$\lambda_0 = ln \int_0^\infty exp\ (-\lambda_1 x)\ dx \tag{4.9}$$

4.1.3 RELATION BETWEEN LAGRANGE MULTIPLIERS AND CONSTRAINTS

Differentiating equation (4.8) with respect to (w.r.t.) λ_1, one gets

$$\frac{\partial \lambda_0}{\partial \lambda_1} = -\frac{1}{\lambda_1} \tag{4.10}$$

Differentiating equation (4.10) with respect to λ_1, one obtains

$$\begin{aligned} \frac{\partial \lambda_0}{\partial \lambda_1} &= -\frac{\int_0^\infty x\ exp\ (-\ \lambda_1 x)\ dx}{\int_0^\infty exp\ (-\ \lambda_1 x)\ dx} \\ &= -\int_0^\infty x\ exp\ (-\ \lambda_0\ - \lambda_1 x)\ dx \\ &= -\int_0^\infty x\, f(x)\ dx\ =\ -\ \bar{x} \end{aligned} \tag{4.11}$$

Equating equations (4.10) and (4.11) the result is

$$-\frac{1}{\lambda_1} = -\ \bar{x} \quad or \quad \bar{x} = \frac{1}{\lambda_1} \tag{4.12}$$

4.1.4 RELATION BETWEEN LAGRANGE MULTIPLIERS AND PARAMETER

Substitution of equation (4.8) in equation (4.6) yields

$$f(x) = \exp\ [-(-\ \ln \lambda_1) - \lambda_1 x] = \lambda_1 \exp\ (-\ \lambda_1 x) \tag{4.13}$$

Comparing equation (4.13) with equation (4.1), one gets

$$\lambda_1 = a \tag{4.14}$$

4.1.5 RELATION BETWEEN PARAMETER AND CONSTRAINT

With use of equation (4.12), we obtain

$$a = \frac{1}{\bar{x}} \tag{4.15}$$

Equation (4.14) shows that parameter a of the exponential distribution is related to the Lagrange multiplier λ_1 and equation (4.15) shows that the distribution parameter, in turn, is related to the constraint of equation (4.5). Thus, parameter a is found to be the inverse of the mean of X.

4.1.6 DISTRIBUTION ENTROPY

The entropy value of the exponential distribution is obtained by inserting equation (4.1) in the definition of entropy:

$$I(x) = -\int_0^{\infty} a\,exp(-ax)\,ln[a\,exp(-ax)]\,dx = -ln\,a + a\bar{x} \tag{4.16}$$

Because $\bar{x} = 1/a$,

$$I(x) = -ln\ a + 1 = ln\ (\bar{x}e) \tag{4.17}$$

4.2 Parameter-Space Expansion Method

4.2.1 SPECIFICATION OF CONSTRAINTS

Following Singh and Rajagopal (1986), the constraints for this method are specified by equation (4.4) and

$$\int_0^{\infty} (ax)f(x)dx = E[ax] \tag{4.18}$$

4.2.2 DERIVATION OF ENTROPY FUNCTION

The pdf corresponding to POME and consistent with equations (4.4) and (4.18) takes the form

$$f(x) = \exp(-\lambda_0 - \lambda_1 ax) \tag{4.19}$$

where λ_0 and λ_1 are Lagrange multipliers. Insertion of equation (4.19) in equation (4.4) leads to

$$\begin{aligned} exp(\lambda_0) &= \int_0^{\infty} exp\ (-\lambda_1 ax)dx \\ &= \frac{1}{a\lambda_1} \end{aligned} \tag{4.20}$$

which is the partition function. From equation (4.20), the zeroth Lagrange multiplier is expressed as

$$\lambda_0 = -\ln a - \ln \lambda_1 \tag{4.21}$$

Substitution of equation (4.21) in equation (4.19) yields

$$f(x) = a\lambda_1 \exp(-\lambda_1 ax) \tag{4.22}$$

A comparison with equation (4.1) shows that $\lambda_1 = 1$.

Taking a logarithm of equation (4.22) and multiplying by minus one, one gets

$$- \ln f(x) = - \ln a - \ln \lambda_1 + \lambda_1 ax \quad (4.23)$$

Using equation (4.23), the entropy function takes the form

$$I(f) = - \ln a - \ln \lambda_1 + \lambda_1 E[ax] \quad (4.24)$$

where E denotes the expectation operator.

4.2.3 RELATION BETWEEN PARAMETER AND CONSTRAINT

Taking the partial derivative of equation (4.24) with respect to a and λ_1 separately, and equating each derivative to zero, one obtains

$$\frac{\partial I}{\partial \lambda_1} = 0 = - \frac{1}{\lambda_1} + E[ax] \quad (4.25)$$

$$\frac{\partial I}{\partial a} = 0 = - \frac{1}{a} + \lambda_1 E[x] \quad (4.26)$$

Solution of equations (4.25) and (4.26) leads to:

$$E[x] = \frac{1}{a} \quad (4.27)$$

which is the parameter estimation equation. This is the same result as obtained earlier.

4.3 Other Methods of Parameter Estimation

Other popular methods of parameter estimation are the methods of moments, maximum likelihood estimation, probability-weighted moments, L-moments, and least squares. In the case of exponential distribution all methods lead to the same parameter estimate.

4.3.1 METHOD OF MOMENTS

Since equation (4.1) has one parameter, taking one moment of f(x) will suffice. To that end,

$$M_1 [f(x)] = \int_0^\infty x f(x) dx = a \int_0^\infty x \exp(-ax) dx \quad (4.28)$$

$$= \frac{1}{a}$$

where M_1 is the first moment of f(x) about the origin. Because $M_1 [f(x)] = E[x] = \bar{x}$, we obtain

$$a = \frac{1}{\bar{x}} \quad (4.29)$$

which is the same as equation (4.27).

4.3.2 METHOD OF MAXIMUM LIKELIHOOD ESTIMATION

The likelihood function for a sample of size N drawn from an exponential population given by equation (4.1) can be written as

$$L = \prod_{i=1}^{N} f(x_i) = \prod_{i=1}^{N} a^N \exp(-ax_i) \tag{4.30}$$

Taking logarithm of equation (4.30), one gets

$$\ln L = N \ln a - a \sum_{i=1}^{N} x_i \tag{4.31}$$

Differentiating equation (4.31) with respect to a and equating the derivative to zero, one obtains

$$\frac{\partial \ln L}{\partial a} = \frac{N}{a} - \sum_{i=1}^{N} x_i = 0 \tag{4.32}$$

This yields

$$a = \frac{1}{\bar{x}} \tag{4.33}$$

which is the same as equation (4.29).

4.3.3 METHOD OF PROBABILITY-WEIGHTED MOMENTS

The probability-weighted moment of order (1,0,0) or zero order, W_0, is calculated as

$$W_0 = \int_0^1 x\,dF = -\frac{1}{a}\int_0^1 \ln(1-F)\,dF \tag{4.34}$$

which yields

$$W_0 = \frac{1}{a} \tag{4.35}$$

But $W_0 = \bar{x}$. Therefore, $a = 1/\bar{x}$, which is the same as equation (4.29).

4.3.4 METHOD OF L-MOMENTS

For the exponential distribution, the first L-moment, L_1, is the same as W_0 given by equation (4.35). Therefore,

$$L_1 = \frac{1}{a} \tag{4.36}$$

But $L_1 = E[x] = \bar{x}$. Thus, $a = 1/\bar{x}$, which is the same as equation (4.29).

References

Eagleson, P.S., 1972. Dynamics of flood frequency. *Water Resources Research*, Vol. **8**, No. 4, pp. 878-898.

Singh, V.P., 1988. *Hydrologic Systems, Vol.1: Rainfall-Runoff Modeling*. Prentice Hall, Inc., Englewood Cliffs, New Jersey, U.S.A.

Singh, V.P. and Rajagopal, A.K., 1986. A new method of parameter estimation for hydrologic frequency analysis. *Hydrological Science and Technology*, Vol. **2**, No. 3, pp. 33-40.

Singh, V.P., Rajagopal, A.K. and Singh, K., 1986. Derivation of some frequency distributions using the principle of maximum entropy (POME). *Advances in Water Resources*, Vol. **9**, pp. 91-106.

Singh, V.P., Singh, K. and Rajagopal, A.K., 1985. Application of the principle of maximum entropy (POME) to hydrologic frequency analysis. *Completion Report* **06**, 144 p., Louisiana Water Resources Research Institute, Louisiana State University, Baton Rouge, Louisiana.

CHAPTER 5

NORMAL DISTRIBUTION

The normal distribution is probably the most popular statistical distribution. It is also known as the Gaussian distribution or error function. Many statistical parameters are found to be approximately normally distributed; therefore, the normal distribution is often used for statistical inferences. A variety of natural phenomena either approximately follow a normal distribution or can be transformed to follow a normal distribution. One of the earliest applications of the normal distribution in hydrology was made by Hazen (1914), who introduced the normal probability paper for ananlysis of hydrologic data. Markovic (1965) fitted the normal distribution to annual rainfall and runoff data. Slack et al. (1975) showed that when the information about the distribution of floods and economic losses associated with the design of flood retardation structures was lacking, it was better to use the normal distribution than other distributions such as extreme value, lognormal, Weibull, etc. The other advantages of the normal distribution are that it is extensively tabulated and the standardized normal variate is the same as the frequency factor.

A random variable X is defined to have a normal distribution if its probability density

$$f(x) = \frac{1}{b\sqrt{2\pi}} \exp\left[-\frac{(x-a)^2}{2b^2}\right] \tag{5.1a}$$

function (pdf) is given by
satisfying $-\infty < x < \infty$ and $b^2 > 0$. Here a and b are parameters which turn out to be the mean and standard deviation of the distribution. The cumulative distribution function (cdf) is given as

$$F(x) = \int_{-\infty}^{x} \frac{1}{b\sqrt{2\pi}} \exp\left[-\frac{(x-a)^2}{2b^2}\right] dx = \Phi(x; a, b) \tag{5.1a}$$

If the variable X is normalized as u = (x-a)/b, then equation (5.1a) becomes

$$f(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right) \tag{5.2a}$$

which is known as the standard normal distribution or error function. The variable u is known as the standard normal variate. Equation (5.2a) can be numerically approximated (Abramowitz and Stegun, 1965) as

V. P. Singh, *Entropy-Based Parameter Estimation in Hydrology*

$$f(u) = (b_0 + b_2 u^2 + b_4 u^4 + b_6 u^6 + b_8 u^8 + b_{10} u^{10})^{-1} \quad (5.2b)$$

where u varies between 0 and infinity, and

b_0 = 2.5052367	b_2 = 1.2831204	b_4 = 0.2264718
b_6 = 0.1306469	b_8 = -0.0202490	b_{10} = 0.0039132

The error in approximation of equation (5.2a) by equation (5.2b) is less than 2.3 $\times 10^{-4}$. Note that f(u) is an even function so that f(u) = f(-u).

The cdf of u, F(u), is given as

$$F(u) = \int_{-\infty}^{u} \frac{1}{\sqrt{2\Pi}} \exp(-t^2/2)\, dt \quad (5.2c)$$

Note that F (-u) = 1 - F (u). Equation (5.2c) can be numerically be approximated (Abramowitz and Stegun, 1965) as

$$F(u) = 1 - f(u)\left[b_1 q + b_2 q^2 + b_3 q^3 + b_4 q^4 + b_5 q^5\right] \quad (5.2d)$$

where q = 1/[1 +p u], p = 0.2316419, $0 \le u \le \infty$, and

b_1 = 0.319381530	b_2 = - 0.356563782	b_3 = 1.781477937
b_4 = - 1.821255978	b_5 = 1.330274429	

The error in approximation of equation (5.2c) by equation (5.2d) is less than 7.5 $\times 10^{-8}$.

5.1 Ordinary Entropy Method

5.1.1 SPECIFICATION OF CONSTRAINTS

Taking logarithm of equation (5.1) to the base 'e', one gets

$$\ln(x) = -\ln\sqrt{2\pi} - \ln b - \frac{(x-a)^2}{2b^2}$$

or

$$\ln f(x) = -\ln\sqrt{2\pi} - \ln b - \frac{x^2}{2b^2} - \frac{a^2}{2b^2} + \frac{2ax}{2b^2} \quad (5.3a)$$

Multiplying equation (5.3a) by [-f(x)] and integrating between - ∞ to ∞, one gets

$$I(x) = -\int_{-\infty}^{\infty} f(x) \ln f(x) dx = [\ln \sqrt{2\pi} + \ln b + \frac{a^2}{2b^2}] \int_{-\infty}^{\infty} f(x) dx + \\ + \frac{1}{2b^2} \int_{-\infty}^{\infty} x^2 f(x) dx - \frac{a}{b^2} \int_{-\infty}^{\infty} x f(x) dx \quad (5.3b)$$

From equation (5.3b), the constraints appropriate for equation (5.1a) can be written (Singh et al., 1985) as

$$\int_{-\infty}^{\infty} f(x)\, dx = 1 \quad (5.4)$$

$$\int_{-\infty}^{\infty} x f(x)\, dx = E[x] = \bar{x} \quad (5.5)$$

$$\int_{-\infty}^{\infty} x^2 f(x)\, dx = E[x^2] = s_x^2 + \bar{x}^2 \quad (5.6)$$

where $\bar{x}$ is the mean and s_x^2 is the variance of x.

5.1.2 CONSTRUCTION OF ZEROTH LAGRANGE MULTIPLIER

The least-biased probability density function f(x) consistent with equations (5.4) to (5.6) and based on the principle of maximum entropy (POME) takes the form (Singh et al., 1985, 1986):

$$f(x) = \exp(-\lambda_0 - \lambda_1 x - \lambda_2 x^2) \quad (5.7)$$

where λ_0, λ_1 and λ_2 are Lagrange multipliers. Substitution of equation (5.7) in the normality condition in equation (5.4) gives

$$\int_{-\infty}^{\infty} f(x) dx = \int_{-\infty}^{\infty} \exp(-\lambda_0 - \lambda_1 x - \lambda_2 x^2) dx = 1 \quad (5.8)$$

Equation (5.8) can be written as

$$exp\ (\lambda_0) = \int_{-\infty}^{\infty} exp\ (-\lambda_1 x - \lambda_2 x^2)\ dx \quad (5.9)$$

Equation (5.9) defines the partition function. Making the argument of the exponential as a square in equation (5.9), one obtains

$$exp\ (\lambda_0) = \int_{-\infty}^{\infty} exp\ (-\lambda_1 x - \lambda_2 x^2 + \frac{\lambda_1^2}{4\lambda_2} - \frac{\lambda_1^2}{4\lambda_2})\ dx$$

$$= exp\ (\frac{\lambda_1^2}{4\lambda_2}) \int_{-\infty}^{\infty} exp\ -[\lambda_1 x + \lambda_2 x^2 + \frac{\lambda_1^2}{4\lambda_2}]\ dx \qquad (5.10)$$

$$= exp\ (\frac{\lambda_1^2}{4\lambda_2}) \int_{-\infty}^{\infty} exp\ -[\sqrt{\lambda_2 x} + \frac{\lambda_1}{2\sqrt{\lambda_2}}]^2\ dx$$

Let

$$t = \sqrt{\lambda_2 x} + \frac{\lambda_1}{2\sqrt{\lambda_2}} \qquad (5.11)$$

then

$$\frac{dt}{dx} = \sqrt{\lambda_2} \qquad (5.12)$$

Making use of equations (5.11) and (5.12) in equation (5.10), we get

$$exp\ (\lambda_0) = \frac{exp\ (\frac{\lambda_1^2}{4\lambda_2})}{\sqrt{\lambda_2}} \int_{-\infty}^{\infty} \exp\ (-t^2)\ dt$$

$$= \frac{2\ \exp\ (\frac{\lambda_1^2}{4\lambda_2})}{\sqrt{\lambda_2}} \int_{0}^{\infty} exp\ (-t^2)\ dt \qquad (5.13)$$

Consider the expression:

$$\int_{0}^{\infty} exp\ (-t^2)\ dt \qquad (5.14)$$

Let $k = t^2$. Then $t = k^{1/2}$ and $[dk/dt] = 2t$. Hence, equation (5.14) can be simplified as Substituting equation (5.15) in equation (5.13), one gets

$$\int_0^{\infty} \exp(-t^2)\,dt = \int_0^{\infty} \exp(-k)\frac{dk}{2k^{1/2}} = \frac{1}{2}\int_0^{\infty} k^{-1/2}\exp(-k)\,dk$$

$$= \frac{1}{2}\int_0^{\infty} k^{[(1/2)-1]}\exp(-k)\,dk = \frac{\Gamma(1/2)}{2} = \frac{\sqrt{\pi}}{2} \qquad (5.15)$$

Substituting equation (5.15) in equation (5.13), one gets

$$exp\ (\lambda_0) = \frac{2\ exp\ (\frac{\lambda_1^2}{4\lambda_2})}{\sqrt{\lambda_2}} \frac{\sqrt{\pi}}{2} = \sqrt{\frac{\pi}{\lambda_2}}) \tag{5.16}$$

Equation (5.16) is another definition of the partition function. The zeroth Lagrange multiplier λ_0 is given by equation (5.16) as

$$\lambda_0 = \frac{1}{2}\ ln\ \pi - \frac{1}{2}\ ln\ \lambda_2 + \frac{\lambda_1^2}{4\lambda_2} \tag{5.17}$$

One also obtains the zeroth Lagrange multiplier from equation (5.9) as

$$\lambda_0 = ln \int_{-\infty}^{\infty} exp\ (-\lambda_1 x - \lambda_2 x^2)\ dx \tag{5.18}$$

5.1.3 RELATION BETWEEN LAGRANGE MULTIPLIERS AND CONSTRAINTS

Differentiating equation (5.18) with respect to λ_1 and λ_2, respectively, one gets

$$\begin{aligned} \frac{\partial\lambda_0}{\partial\lambda_1} &= - \frac{\int_{-\infty}^{\infty} x\ exp\ (-\ \lambda_1 x - \lambda_2 x^2)\ dx}{\int_{-\infty}^{\infty} exp\ (-\ \lambda_1 x - \lambda_2 x^2)\ dx} \\ &= - \int_{-\infty}^{\infty} x\ exp\ (-\ \lambda_0 - \lambda_1 x - \lambda_2 x^2)\ dx \\ &= - \int_{-\infty}^{\infty} x\ f(x)\ dx = -\ \bar{x} \end{aligned} \tag{5.19}$$

$$\begin{aligned} \frac{\partial\lambda_0}{\partial\lambda_2} &= - \frac{\int_{-\infty}^{\infty} x^2\ exp\ (-\ \lambda_1 x - \lambda_2 x^2)\ dx}{\int_{-\infty}^{\infty} exp\ (-\ \lambda_1 x - \lambda_2 x^2)\ dx} \\ &= - \int_{-\infty}^{\infty} x^2\ exp\ (-\ \lambda_0 - \lambda_1 x - \lambda_2 x^2)\ dx \\ &= - \int_{-\infty}^{\infty} x^2\ f(x)\ dx = -\ (s_x^2 + \bar{x}^2) \end{aligned} \tag{5.20}$$

Differentiating equation (5.17) with respect to λ_1 and λ_2, respectively, one obtains

$$\frac{\partial\lambda_0}{\partial\lambda_1} = \frac{2\lambda_1}{4\lambda_2} = \frac{\lambda_1}{2\lambda_2} \tag{5.21}$$

$$\frac{\partial \lambda_0}{\partial \lambda_2} = - \frac{1}{2\,\lambda_2} - \frac{\lambda_1^2}{4\,\lambda_2^2} \tag{5.22}$$

Equating equation (5.19) to equation (5.21) and equation (5.20) to equation (5.22) one gets

$$\frac{\lambda_1}{2\,\lambda_2} = - \bar{x} \tag{5.23}$$

$$\frac{1}{2\,\lambda_2} + \frac{1}{4} \left(\frac{\lambda_1}{\lambda_2}\right)^2 = s_x^2 + \bar{x}^2 \tag{5.24}$$

From equation (5.23), one gets

$$\lambda_1 = - 2\,\lambda_2\,\bar{x} \tag{5.25}$$

Substituting equation (5.25) in equation (5.24) one obtains

$$\begin{aligned} \frac{1}{2\,\lambda_2} + \frac{1}{4}\,\frac{4\,\lambda_2^2\,\bar{x}}{\lambda_2^2} &= s^2 + \bar{x}^2 \\ \frac{1}{2\,\lambda_2} &= s_x^2 \\ \lambda_2 &= \frac{1}{2\,s_x^2} \end{aligned} \tag{5.26}$$

Eliminating λ_2 in equation (5.23) yields

$$\lambda_1 = - 2\,\frac{1}{2\,s_x^2}\,\bar{x} = - \frac{\bar{x}}{s_x^2} \tag{5.27}$$

5.1.4 RELATION BETWEEN LAGRANGE MULTIPLIERS AND PARAMETERS

Substitution of equation (5.17) in equation (5.7) yields

$$f(x) = [-\frac{1}{2}\ln\ \pi + \frac{1}{2}\ln\lambda_2 - \frac{\lambda_1^2}{4\lambda_2} - \lambda_1 x - \lambda_2 x^2]$$

$$= \exp[\ln(\pi)^{(-1/2)} + \ln(\lambda_2)^{1/2} - \frac{\lambda_1}{4\lambda_2} - \lambda_1 x - \lambda_2 x^2] \quad (5.28)$$

$$= (\pi)^{-(1/2)}(\lambda_2)^{1/2}\exp(-\lambda_1 x - \lambda_2 x^2 - \frac{\lambda_1}{4\lambda_2})$$

A comparison of equation (5.28) with equation (5.1a) shows that

$$\lambda_1 = -a/b^2 \quad (5.29)$$

$$\lambda_2 = 1/(2b^2) \quad (5.30)$$

5.1.5 RELATION BETWEEN PARAMETERS AND CONSTRAINTS

The normal distribution has two parameters a and b which are related to the Lagrange multipliers by equations (5.29) and (5.30), which themselves are related to the constraints through equations (5.26) and (5.27) [and in turn through equations (5.5) and (5.6)]. Eliminating the Lagrange multipliers between these two sets of equations, we obtain

$$a = \bar{x} \quad (5.31)$$

$$b = s_x \quad (5.32)$$

5.1.6 DISTRIBUTION ENTROPY

Substitution of equation (5.31) and (5.32) in equation (5.36b) yields

$$I(x) = [\ln\sqrt{2\pi} + \ln s_x + \frac{\bar{x}^2}{2s_x^2}] \int_{-\infty}^{\infty} f(x)\, dx$$

$$+ \frac{1}{2s_x^2}\int_{-\infty}^{\infty} x^2 f(x)\, dx - \frac{\bar{x}}{s_x^2}\int_{-\infty}^{\infty} x f(x)\, dx \quad (5.33)$$

$$= [\ln\sqrt{2\pi} + \ln s + \frac{\bar{x}^2}{2s_x^2}]\ 1 + \frac{1}{2s_x^2}(\bar{x}^2 + s_x^2) - \frac{\bar{x}^2}{s_x^2}$$

$$= \ln[s_x(2\pi e)^{0.5}]$$

5.2 Parameter - Space Expansion Method

5.2.1 SPECIFICATION OF CONSTRAINTS

Following Singh and Rajagopal (1986), the constraints for this method are given by equation (5.4) and

$$\int_{-\infty}^{\infty} \left(\frac{xa}{b^2}\right) f(x)\, dx = E\left[\frac{xa}{b^2}\right] = \frac{a}{b^2} \tag{5.34}$$

$$\int_{-\infty}^{\infty} \left(\frac{x^2}{2b^2}\right) f(x)\, dx = E\left[\frac{x^2}{2b^2}\right] = \tag{5.35}$$

5.2.2 DERIVATION OF ENTROPY FUNCTION

The pdf corresponding to POME and consistent with equations (5.4), (5.34), and (5.35) takes the form:

$$f(x) = exp\left[-\lambda_0 - \lambda_1 \frac{xa}{b^2} - \lambda_2 \left(\frac{x^2}{2b^2}\right)\right] dx \tag{5.36}$$

where λ_0, λ_1, and λ_2 are Lagrange multipliers. Insertion of equation (5.36) into equation (5.4) yields

$$\begin{aligned} exp\,(\lambda_0) &= \int_{-\infty}^{\infty} exp\left[-\lambda_1 \frac{xa}{b^2} - \lambda_2 \left(\frac{x^2}{2b^2}\right)\right] dx \\ &= \frac{b\sqrt{2\pi}}{\sqrt{\lambda_2}}\, exp\left[\frac{a^2\lambda_1^2}{2\lambda_2 b^2}\right] \end{aligned} \tag{5.37}$$

Equation (5.37) is the partition function. Taking logarithm of equation (5.37) leads to the zeroth Lagrange multiplier which can be expressed as

$$\lambda_0 = ln\ b + \frac{1}{2}\, ln\,(2\pi) - \frac{1}{2}\, ln\ \lambda_2 + \frac{a^2\lambda_1^2}{2\lambda_2 b^2} \tag{5.38}$$

The zeroth Lagrange multiplier is also obtained from equation (5.37) as

$$\lambda_0 = ln \int_{-\infty}^{\infty} exp\ [-\ \lambda_1 \frac{xa}{b^2} - \lambda_2 (\frac{x^2}{2b^2})]\ dx \tag{5.39}$$

Introduction of equation (5.38) in equation (5.36) gives

$$f(x) = \frac{\sqrt{\lambda_2}}{b\sqrt{2\pi}}\ exp\ [-(\frac{a^2\lambda_1^2}{2\lambda_2 b^2} + \frac{\lambda_1 xa}{b^2} + \frac{\lambda_2 x^2}{2b^2})] \tag{5.40}$$

A comparison of equation (5.40) with equation (5.1) shows that $\lambda_2 = 1$ and $\lambda_1 = -1$.

Taking logarithm of equation (5.40) and multiplying by [-1], one gets

$$-lnf(x) = -\frac{1}{2} ln\lambda_2 + lnb + \frac{1}{2} ln(2\pi) + \frac{a^2\lambda_1^2}{2\lambda_2 b^2} + \frac{\lambda_1 xa}{b^2} + \frac{\lambda_2 x^2}{2b^2} \tag{5.41}$$

Multiplying equation (5.41) by f(x) and integrating from minus infinity to positive infinity, we get the entropy function which takes the form:

$$I(f) = -\frac{1}{2} ln\lambda_2 + lnb + \frac{1}{2} ln(2\pi) + \frac{a^2\lambda_1^2}{2\lambda_2 b^2} + \frac{\lambda_1 a}{b^2} E[x] + \frac{\lambda_2}{2b^2} E[x^2] \tag{5.42}$$

5.2.3 RELATION BETWEEN DISTRIBUTION PARAMETERS AND CONSTRAINTS

Taking partial derivatives of equation (5.42) with respect to λ_1, λ_2, a, and b individually, and then equating each derivative to zero, one obtains

$$\frac{\partial I}{\partial \lambda_1} = 0 = \frac{2a^2\lambda_1}{2\lambda_2 b^2} + \frac{a}{b^2} E[x] \tag{5.43}$$

$$\frac{\partial I}{\partial \lambda_2} = 0 = -\frac{1}{2\lambda_2} - \frac{a^2\lambda_1^2}{2\lambda_2^2 b^2} + \frac{1}{2b^2} E[x^2] \tag{5.44}$$

$$\frac{\partial I}{\partial a} = 0 = -\frac{2a\lambda_1^2}{2b\lambda_2} + \frac{\lambda_1}{b^2} E[x] \tag{5.45}$$

$$\frac{\partial I}{\partial b} = 0 = \frac{1}{b} - \frac{2a^2\lambda_1^2}{2\lambda_2 b^3} - \frac{2a\lambda_1}{b^3} E[x] - \frac{2\lambda_2}{2b^3} E[x^2] \tag{5.46}$$

Simplification of equation (5.43) through (5.46) results in

$$E[x] = a \tag{5.47}$$

$$E[x^2] = a^2 + b^2 \tag{5.48}$$

$$E[x] = a \tag{5.49}$$

$$E(x^2] = b^2 + a^2 \tag{5.50}$$

Equations (5.47) and (5.49) are the same, and so are equations (5.48) and (5.50). Thus, the parameter estimation equations are equations (5.47) and (5.48).

5.3 Other Methods of Parameter Estimation

Three other parameter estimation methods are briefly discussed: methods of moments, probability-weighted moments and maximum likelihood estimation.

5.3.1 METHOD OF MOMENTS

The normal distribution has two parameters, a and b, so the first two moments will suffice. The first moment about the origin is calculated as follows:

$$M_1 = \int_{-\infty}^{\infty} x \frac{1}{b\sqrt{2\Pi}} \exp[-\frac{(x-a)^2}{2b^2}] dx \tag{5.51}$$

Equation (5.51) is easily solved by transforming x to the standard normal variate z = (x-a)/b. A little algebraic manipulation shows that

$$M_1 = a \tag{5.52}$$

Thus, parameter a is the same as mean of X or the first moment. The second moment is computed about the centroid a as

$$M_2^a = \int_{-\infty}^{\infty} (x-a)^2 \frac{1}{b\sqrt{2\Pi}} \exp[-\frac{(x-a)^2}{2b^2}] dx \tag{5.53}$$

Using the transformation $y = (x - a)^2 / 2b^2$, and noting that y is an even function and the integral is twice the integral from 0 to infinity, we obtain

$$M_2^a = b^2 \tag{5.54}$$

Equation (5.54) shows that parameter b is the same as the second moment about the centroid, s^2, where s is the standard deviation.

5.3.2 METHOD OF MAXIMUM LIKELIHOOD ESTIMATION

The log-likelihood function of a sample of size n drawn from a normal distribution is given by

$$\log L(a, b) = -n \log b - n \log \sqrt{2\Pi} - \frac{1}{2b^2} \sum_{i=1}^{n} (x_i - a)^2 \quad (5.55)$$

Differentiating equation (5.55) with respect to a and equating the derivative to zero yield

$$a = \frac{1}{n} \sum_{i=1}^{n} x_i \quad (5.56)$$

Differentiating equation (5.55) with respect to b and equating the derivative to zero yield

$$b^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - a)^2 \quad (5.57)$$

Equation (5.57) shows that parameter b is the same as the standard deviation. Thus, the methods of moments and maximum likelihood estimation yield the parameter estimates.

5.3.3 METHOD OF PROBABILITY-WEIGHTED MOMENTS

The normal distribution cannot be expressed explicitly in terms of x and this makes evaluation of probability-weighted moments (PWMs) complicated. Hosking (1986) derived PWMs for normal distribution in terms of L-moments:

$$a_1 = b_0 = a \quad (5.58)$$

$$a_2 = 2b_1 - b_0 = \frac{b}{\sqrt{\pi}} \quad (5.59)$$

where a_1 and a_2 are first and second order L-moments, and b_0 and b_1 are PWMs of the zero and first order. The parameter estimates by the method of PWMs are given in terms of the sample moments:

$$a = \bar{x} \quad and \quad b = \sqrt{\pi}\, a_2 \quad (5.60)$$

References

Abramowitz, M. and Stegun, I.A., 1965. *Handbook of Mathematical Functions*. Dover Publications, New York.

Hazen, A., 1914. Discussion on "Flood flows" by W.E. Fuller. *Transactions of American Society of Civil Engineers*, Vol. **77**, pp. 626-632.

Hosking, J.R.M., 1986. The theory of probability weighted moments. *Research Report* **RC 12210 (#54860),** Thomas J.Watson Research Center, I.B.M., Yorktown Heights, New York.

Markovic, R.D., 1965. Probability Functions of best fit to distributions of annual precipitation and runoff. *Hydrology Paper* No. **8**, Colorado State University, Fort Collins, Colorado.

Slack, J.R., Wallis, J.R. and Matalas, N.C., 1975. On the value of information to flood frequency analysis. *Water Resources Research*, Vol. **23**, No. 8, pp. 629-647.

Singh, V.P. and Rajagopal, A.K., 1986. A new method of parameter estimation for hydrologic frequency analysis. *Hydrological Science and Technology*, Vol. **2,** No. 3, pp. 33-40.

Singh, V.P., Rajagopal, A.K. and Singh, K., 1986. Derivation of some frequency distributions using the principle of maximum entropy (POME). *Advances in Water Resources*, Vol. **9**, pp. 91-106.

Singh, V.P., Singh, K. and Rajagopal, A.K., 1985. Application of the principle of maximum entropy (POME) to hydrologic frequency analysis. *Completion Report* **06**, 144 p., Louisiana Water Resources Research Institute, Louisiana State University, Baton Rouge, Louisiana.

CHAPTER 6

TWO-PARAMETER LOGNORMAL DISTRIBUTION

The logarithmic normal probability law is widely used to describe the distribution of annual maximum values of hourly or daily precipitation (Weiss, 1957), flood flows (Chow, 1951, 1954), hydraulic conductivity (Freeze and Cherry, 1979), soil properties (physical, chemical and microbiological) (Parkin and Robinson, 1993), etc. Kalinske (1946) found that many times river discharge data and sand sizes followed the normal law if they were logarithmically transformed. Chow (1951, 1954, 1959) gave a historical background of the log-probability law and discussed its wide-ranging application in engineering, and exensively worked with the lognormal distribution. Aitchison and Brown (1957) presented a comprehensive statistical treatment of the lognormal distribution. Parkin et al. (1988) evaluated statistical methods for log-normally distributed variables, including the method of moments, maximum likelihood, and Finney's method. Parkin and Robinson (1993) evaluated soil properties using log-normal distribution. Brakensiek (1958) employed the least squares method for fitting the log-normal distribution to annual runoff. Moran (1957) fitted a log-normal distribution to fifty annual values of extreme monthly flow of the River Murray in Australia. Lewis (1979) applied log-normal distribution to maximum measured discharges of River Kafue in Africa.Weiss (1957) developed a nomogram for log-normal frequency analysis.Alexander et al. (1969) discussed statistical properties of lognormal distribution. Using mean square error of estimation as a criterion, Stedinger (1980) evaluated the efficiency of alternative methods of fitting the lognormal distribution. Charbeneau (1978) compared two- and three-parameter log-normal distributions for simulation of stream flow.

A random variable X with range $\{x: 0 < x < \infty\}$ is said to have a lognormal distribution if Y = ln X is normally distributed. That is, X = exp (Y) withY normal. Then the probability density function (pdf), f(x), of X is given by

$$f(x) = \frac{1}{xb\sqrt{2\pi}}\ exp\ [\frac{-(\ln x - a)^2}{2b^2}], x>0 \tag{6.1a}$$

$$= 0, \quad x \le 0 \tag{6.1b}$$

where a and b are parameters which, respectively, are the mean and the variance of Y, $\bar{y}$ and s_y^2. These are related to the mean $\bar{x}$ and the variance s_x^2 of X, denoted, respectively, as c and d:

$$c = \exp [\ a + (b^2/2)\ ] \tag{6.2a}$$

and

$$d^2 = \{\exp(2a + b^2)\}\{\exp(b^2) - 1\} \tag{6.2b}$$

The cumulative distribution function (cdf) of X can be written as

$$F(x) = \int_0^x \frac{1}{xb\sqrt{2\pi}} \exp\left[-\frac{(\ln x - a)^2}{2b^2}\right] dx \tag{6.3}$$

If the variate ln x is standardized as u = [ln x - a]/b, then the standard normal variate u will have the pdf given by equation (5.2).

6.1 Ordinary Entropy Method

6.1.1 SPECIFICATION OF CONSTRAINTS

Taking logarithm of equation (6.1a) to the base 'e', one gets

$$\ln f(x) = -\ln(b\sqrt{2\pi} - \ln x - \frac{(\ln x - a)^2}{2b^2} \tag{6.4}$$

Multiplying equation (6.4) by [-f(x)] and integrating between 0 and ∞, one obtains the entropy function as

$$I(x) = -\int_0^\infty f(x) \ln f(x)\, dx = \ln(b\sqrt{2\Pi}) \int_0^\infty f(x)\, dx + \int_0^\infty \ln x f(x)\, dx + \int_0^\infty \frac{(\ln x - a)^2}{2b^2} f(x)\, dx \tag{6.5}$$

From equation (6.5), the constraints appropriate for equation (6.1a) can be written (Singh et al., 1985, 1986) as:

$$\int_0^\infty f(x)\, dx = 1 \tag{6.6}$$

$$\int_0^\infty \ln x\, f(x)\, dx = E[\ln x] = E[y] = \bar{y} \tag{6.7}$$

$$\int_0^\infty (\ln x - a)^2 f(x)\, dx = E[\ln x - a)^2] \tag{6.8}$$

Recalling that $a = \bar{y}$, one can write equation (6.8) as

$$\int_0^\infty [(ln\ x)^2 + a^2 - 2a\ ln\ x]\ f(x)\ dx$$
$$= \int_0^\infty (ln\ x)^2 f(x)\ dx + a^2 \int_0^\infty f(x)\ dx$$
$$- 2a \int_0^\infty ln\ x\ f(x)\ dx \tag{6.9}$$
$$= \int_0^\infty (ln\ x)^2 f(x)\ dx - a^2$$

Making use of equation (6.9), equation (6.8) can be written as

$$\int_0^\infty (ln\ x)^2 f(x)\ dx = s_y^2 + a^2 \tag{6.10}$$

6.1.1 CONSTRUCTION OF ZEROTH LAGRANGE MULTIPLIER

The least-biased pdf based on POME and consistent with equations (6.6) to (6.8) takes the form:

$$f(x) = \exp(-\lambda_0 - \lambda_1 \ln x - \lambda_2 (\ln x)^2) \tag{6.11}$$

where λ_0, λ_1, and λ_2 are Lagrange multipliers. Substitution of equation (6.11) in equation (6.6) yields

$$\int_0^\infty f(x) dx = \int_0^\infty exp[-\lambda_0 - \lambda_1 lnx - \lambda_2 (lnx)^2] dx = 1 \tag{6.12}$$

Equation (6.12) yields the partition function given as

$$\exp(\lambda_0) = \int_0^\infty exp[-\lambda_1 lnx - \lambda_2 (lnx)^2] dx \tag{6.13a}$$

or

$$\exp(\lambda_0) = \int_0^\infty exp(lnx^{-\lambda_1}) exp[-\lambda_2 (lnx)(lnx)] dx$$
$$= \int_0^\infty x^{-\lambda_1} \exp[(\ln x^{-\lambda_2 \ln x})] dx \tag{6.13b}$$
$$= \int_0^\infty x^{-\lambda_1} (x^{-\lambda_2})^{\ln x} dx$$

Let z = ln x . Then

[dz/dx] = [1/x], x = exp (z); and dx = x dz = exp (z) dz. Hence,

$$\exp(\lambda_0) = \int_{-\infty}^{\infty} \exp(-z\lambda_1)\exp(-z^2\lambda_2)\exp(z)\,dz = \int_{-\infty}^{\infty} \exp[-(\lambda_1 - 1)z - \lambda_2 z^2]\,dz \tag{6.14}$$

Making the argument of the exponential as a square in equation (6.14), one gets

$$\exp(\lambda_0) = \int_{-\infty}^{\infty} exp\ [-(\lambda_1 - 1)z - \lambda_2 z^2 + \frac{(\lambda_1 - 1)^2}{4\lambda_2} - \frac{(\lambda_1 - 1)^2}{4\lambda_2}]\,dz$$
$$= exp\ [\frac{(\lambda_1 - 1)^2}{4\lambda_2} \int_{-\infty}^{\infty} exp - [(\lambda_1 - 1)z + \lambda_2 z^2 + \frac{(\lambda_1 - 1)^2}{4\lambda_2}]\,dz$$
$$= exp\ [\frac{(\lambda_1 - 1)^2}{4\lambda_2}] \int_{-\infty}^{\infty} exp - [\sqrt{\lambda_2}\,z + \frac{(\lambda_1 - 1)}{2\sqrt{\lambda_2}}]^2\,dz \tag{6.15}$$

Let

$$t = \sqrt{\lambda_2}\,z + \frac{(\lambda_1 - 1)}{2\sqrt{\lambda_2}};\ \text{then}\ \frac{dt}{dz} = \sqrt{\lambda_2}$$

Substituting these quantities in equation (6.15), we get

$$exp\ (\lambda_0) = \frac{exp\ [\frac{(\lambda_1 - 1)^2}{4\lambda_2}]}{\sqrt{\lambda_2}} \int_{-\infty}^{\infty} exp\ (-t^2)\,dt$$
$$= \frac{exp\ [\frac{(\lambda_1 - 1)^2}{4\lambda_2}]}{\sqrt{\lambda_2}}\ 2\int_{0}^{\infty} exp\ (-t^2)\,dt$$
$$= \frac{exp\ [\frac{(\lambda_1 - 1)^2}{4\lambda_2}]}{\sqrt{\lambda_2}}\ 2\,\frac{\sqrt{\pi}}{2} \tag{6.16}$$

$$exp\ (\lambda_0) = \sqrt{\frac{\pi}{\lambda_2}}\ exp\ [\frac{(\lambda_1 - 1)^2}{4\lambda_2}] \tag{6.17}$$

Equation (6.17) defines the partition function. Thus, the zeroth Lagrange multiplier λ_0 is obtained as

$$\lambda_0 = \frac{1}{2} \ln \pi - \frac{1}{2} \ln \lambda_2 + \frac{(\lambda_1 - 1)^2}{4 \lambda_2} \tag{6.18}$$

The zeroth Lagrange multiplier is also obtained from equation (6.12) as

$$\lambda_0 = \ln \int_0^\infty \exp [- \lambda_1 \ln x - \lambda_2 (\ln x)^2] \, dx \tag{6.19}$$

6.1.3 RELATION BETWEEN LAGRANGE MULTIPLIERS AND CONSTRAINTS

Differentiating equation (6.19) with respect to λ_1 and λ_2 ,respectively, one gets

$$\begin{aligned} \frac{\partial \lambda_0}{\partial \lambda_1} &= - \frac{\int_0^\infty \ln x \, [\exp (-\lambda_1 \ln x - \lambda_2 (\ln x))^2] \, dx}{\int_0^\infty \exp [-\lambda_1 \ln x - \lambda_2 (\ln x)^2] \, dx} \\ &= - \int_0^\infty \ln x \exp [-\lambda_0 - \lambda_1 \ln x - \lambda_2 (\ln x)^2] \, dx \\ &= - \int_0^\infty \ln x f(x) \, dx = E[\ln x] = - \bar{y} \end{aligned} \tag{6.20}$$

$$\begin{aligned} \frac{\partial \lambda_0}{\partial \lambda_2} &= - \frac{\int_0^\infty (\ln x)^2 \, [\exp (-\lambda_1 \ln x - \lambda_2 (\ln x)^2] \, dx}{\int_0^\infty \exp [- \lambda_1 \ln x - \lambda_2 (\ln x)^2] \, dx} \\ &= - \int_0^\infty (\ln x)^2 \exp [- \lambda_0 - \lambda_1 \ln x - \lambda_2 (\ln x)^2] \, dx \\ &= - \int_0^\infty (\ln x)^2 f(x) \, dx = - (s_y^2 + \bar{y}^2) \end{aligned} \tag{6.21}$$

Differentiating equation (6.18) with respect to λ_1 and λ_2 , respectively, one obtains

$$\frac{\partial \lambda_0}{\partial \lambda_1} = \frac{2 (\lambda_1 - 1)}{4 \lambda_2} = \frac{(\lambda_1 - 1)}{2 \lambda_2} \tag{6.22}$$

$$\frac{\partial \lambda_0}{\partial \lambda_2} = \frac{(\lambda_1 - 1)^2}{4 \lambda_2^2} - \frac{1}{2 \lambda_2} \tag{6.23}$$

Equating equations (6.20) and (6.22) as well as equations (6.21) and (6.23), one gets

$$\frac{\lambda_1 - 1}{2\,\lambda_2} = -\,\bar{y}$$
$$(\lambda_1 - 1)^2 = 4\,\lambda_2^2\,\bar{y}^2 \tag{6.24}$$

$$\frac{(\lambda_1 - 1)^2}{4\,\lambda_2^2} + \frac{1}{2\,\lambda_2} = s_y^2 + \bar{y}^2 \tag{6.25}$$

Substitution of equation (6.24) in equation (6.25) produces

$$\frac{4\,\lambda_2\,y^2}{4\,\lambda_2^2} + \frac{1}{2\,\lambda_2} = s_y^2 + \bar{y}^2$$
$$\lambda_2 = \frac{1}{2\,s_y^2} \tag{6.26}$$

Eliminating λ_2 from equation (6.24), one gets

$$(\lambda_1 - 1)^2 = 4\,\frac{1}{4\,s_y^4}\,\bar{y}^2 = \left(\frac{\bar{y}}{s_y^2}\right)^2$$
$$\lambda_1 - 1 = -\,\frac{\bar{y}}{s_y^2} \tag{6.27}$$
$$\lambda_1 = 1 - \frac{\bar{y}}{s_y^2}$$

6.1.4 RELATION BETWEEN LAGRANGE MULTIPLIERS AND PARAMETERS

Equation (6.18) can now be expressed as

$$\lambda_0 = \frac{1}{2}\,ln\,\pi - \frac{1}{2}\,ln\,\left[\frac{1}{2s_y^2}\right] + \frac{y^2}{s_y^4}\,\frac{1}{4\,(1/2/s_y^2)}$$
$$= ln\,\sqrt{\pi} - \frac{1}{2}\,ln\,(2s_y^2)^{-1} + \frac{\bar{y}^2}{2s_y^2} \tag{6.28}$$
$$= ln\,\sqrt{\pi} + ln\,(\sqrt{2}s_y) + \frac{\bar{y}^2}{2s_y^2}$$

The pdf in equation (6.11) can now be written as

$$f(x) = exp\ [-\ ln\ \sqrt{\pi}\ -\ ln\ (\sqrt{2}\ s_y)\ -\ \frac{\bar{y}^2}{2s_y^2}$$
$$-\ (1\ -\ \frac{\bar{y}}{s_y^2})\ ln\ x\ -\ \frac{(ln\ x)^2}{2s_y^2}]$$
$$= exp\ [ln\ (\sqrt{\pi})^{-1}]\ exp\ [ln\ (s_y\ \sqrt{2})^{-1}$$
$$\cdot\ exp\ [-\ \frac{\bar{y}^2}{2s_y^2}\ ln\ x\ +\ \frac{\bar{y}}{s_y^2}\ ln\ x\ -\ \frac{(ln\ x)^2}{2s_y^2}] \qquad (6.29)$$
$$= \frac{1}{s_y\sqrt{2\pi}}\ exp\ [ln\ (x)^{-1}]\ exp\ [-\ \frac{1}{2s_y^2}\ (ln\ x\ -\ \bar{y}^2)]$$
$$+\ \frac{1}{xs_y\sqrt{2\pi}}\ exp\ [-\ \frac{1}{2s_y^2}\ (ln\ x\ -\ \bar{y})^2]$$

A comparison of equation (6.29) with equation (6.1a) shows that $a = \bar{y}$ *and* $b = s_y$.

6.1.5 RELATION BETWEEN PARAMETERS AND CONSTRAINTS

The two-parameter lognormal distribution has 2 parameters a and b which, respectively, equal $\bar{y}$ and s_y as shown above. These parameters are related to the Lagrange multipliers through equations (6.26) and (6.27),which, in turn, are related to constraints in equations (6.7) and (6.8).Eliminating the Lagrange multipliers between these two sets of equations, we get distribution parameters in terms of constraints.

6.1.6 DISTRIBUTION ENTROPY

The entropy of the lognormal distribution can be derived in two ways.

$$I(x)\ =\ -\ \int_0^\infty f(x)\ ln\ f(x)\ dx\ =\ ln\ (s_y\sqrt{2\pi})\ \int_0^\infty f(x)\ dx$$
$$+\ \int_0^\infty ln\ x\ f(x)\ dx\ +\ \frac{1}{2s_y^2}\ \int_0^\infty (ln\ x\ -\ \bar{y})^2\ f(x)\ dx$$
$$=\ ln\ (s_y\sqrt{2\pi})\ +\ \bar{y}\ +\ \frac{1}{2s_y^2}\ s_y^2\ =\ ln\ (s_y\sqrt{2\pi})\ +\ \bar{y}\ +\ \frac{1}{2}\ ln\ e \qquad (6.30)$$
$$=\ ln\ (s_y\sqrt{2\pi e})\ +\ \bar{y} =\ I(y)\ +\ \bar{y}$$

where $I(y) = ln(s_y\sqrt{2\pi e})$.

Alternatively, since the transformation $x = e^y = g(y)$ is monotonic with the Jacobian $J(y/x) = l/x$, the above result follows immediately. From the general relationship, i.e.,

$$I(x)=I(y) - E\ \{\ ln|J(\ y\ /\ x\ )|]$$

$$J = \frac{\partial(x,y)}{\partial(u,v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix}$$

$$J(\frac{y}{x}) = \frac{\partial y}{\partial x} = \frac{\partial}{\partial x}(ln\ x) = \frac{1}{x} \tag{6.31}$$

$$I(x) = I(y) - E[\ ln\ |\frac{1}{x}|] = I(y) + E[ln\ x] + I(y) + \bar{y}$$

which is the same as before.

6.2 Parameter - Space Expansion Method

6.2.1 SPECIFICATION OF CONSTRAINTS

For this method, the constraints are given, following Singh and Rajagopal (1986), by equation (6.6) and

$$\int_0^\infty (-\ln x + \frac{a \ln x}{b^2}) f(x) dx = E[-\ln x + \frac{a \ln x}{b^2}] \tag{6.32}$$

and

$$\int_0^\infty [\frac{(\ln x)^2 + a^2}{2b^2}] f(x) dx = E[\frac{a^2 + (\ln x)^2}{b^2}] \tag{6.33}$$

6.2.2 DERIVATION OF ENTROPY FUNCTION

The pdf corresponding to POME and consistent with equations (6.6) and (6.32) and (6.33) takes the form

$$f(x) = exp\ [-\ \lambda_0\ -\lambda_1\ \frac{a\ ln\ x}{b^2} + \lambda_1\ ln\ x - \lambda_2\ (\frac{(ln\ x)^2 + a^2}{2b^2})] \tag{6.34}$$

where λ_0, λ_1, and λ_2 are Lagrange multipliers. Insertion of equation (6.34) into equation (6.6) yields

$$exp\ (\lambda_0) = \int_0^\infty exp\ [-\frac{\lambda_1 a}{b^2}\ ln\ x + \lambda_1\ ln\ x - \frac{\lambda_2}{2b^2}\ ((ln\ x)^2 + a^2)]\ dx$$
$$= \frac{b\sqrt{2\pi}}{\sqrt{\lambda_2}}\ exp\ [-\frac{a^2\lambda_2}{2b^2} + \frac{b^2}{2\lambda_2} + \frac{b^2\lambda_1^2}{2\lambda_2} - \frac{a\lambda_1^2}{\lambda_2} + \frac{a^2\lambda_1^2}{2b^2\lambda_2} - a\ \frac{\lambda_1}{\lambda_2} + \frac{b^2\lambda_1}{\lambda_2}] \tag{6.35}$$

Equation (6.35) defines the partition function. Taking logarithm of equation (6.35) yields the zeroth Lagrange multiplier given as

$$\lambda_0 = ln\ (\frac{b\sqrt{2\pi}}{\sqrt{\lambda_2}}) - \frac{a^2\lambda_2}{2b^2} + \frac{b^2}{2\lambda_2} + \frac{b^2\lambda_1^2}{2\lambda_2} - \frac{a\lambda_1^2}{\lambda_2} + \frac{a^2\lambda_1^2}{2b^2\lambda_2} - \frac{a\lambda_1}{\lambda_2} + \frac{b^2\lambda_1}{\lambda_2} \tag{6.36}$$

The zeroth Lagrange multiplier is also obtained from equation (6.35) as

$$\lambda_0 = ln \int_0^\infty exp\ [-\frac{\lambda_1 a}{b^2}\ lnx + \lambda_1\ ln\ x - \frac{\lambda_2}{2b^2}\ ((lnx)^2 + a^2)]\ dx \tag{6.37}$$

Introduction of equation (6.35) in equation (6.34) produces

$$f(x) = \frac{\sqrt{\lambda_2}}{b\sqrt{2\pi}}\ exp\ [-\frac{b^2}{2\lambda_2} - \frac{b^2\lambda_1^2}{2\lambda_2} + \frac{a\lambda_1^2}{\lambda_2} - \frac{a^2\lambda_1^2}{2b^2\lambda_2} + \frac{a\lambda_1}{\lambda_2} - \frac{b^2\lambda_1}{\lambda_2} + \lambda_1\ lnx - \lambda_1 a\ \frac{lnx}{b^2} - \lambda_2\ \frac{(lnx)^2}{b^2}] \tag{6.38}$$

A comparison of equation (6.38) with equation (6.1a) shows that $\lambda_1 = -1$ and $\lambda_2 = 1$. Taking -logarithm of equation (6.38), one obtains

$$-\ln f(x) = -\frac{1}{2}\ln\lambda_2 + \ln b + \frac{1}{2}\ln(2\pi) + \frac{b^2}{2\lambda_2} + \frac{b^2\lambda_1^2}{2\lambda_2} - a\frac{\lambda_1^2}{\lambda_2} + \frac{a^2\lambda_1^2}{2b^2\lambda_2} - \frac{a\lambda_1}{\lambda_2} + \frac{b^2\lambda_1}{\lambda_2} - \lambda_1 \ln x + \frac{a\lambda_1}{b^2}\ln x + \frac{\lambda_2}{2b^2}(\ln x)^2 \tag{6.39}$$

Multiplying equation (6.39) by f(x) and integrating from $-\infty$ to $+\infty$, we get the entropy function:

$$I(f) = -\frac{1}{2}\ln\lambda_2 + \ln b + \frac{1}{2}\ln(2\pi) + \frac{b^2}{2\lambda_2} + \frac{b^2\lambda_1^2}{2\lambda_2} - \frac{a\lambda_1^2}{\lambda_2} + \frac{a^2\lambda_1^2}{2b^2\lambda_2} - \frac{a\lambda_1}{\lambda_2} + \frac{b^2\lambda_1}{\lambda_2} - \lambda_1 E[\ln x] + \frac{a\lambda_1}{b^2}E[\ln x] + \frac{\lambda^2}{2b^2}E[(\ln x)^2] \tag{6.40}$$

6.2.3 RELATION BETWEEN PARAMETERS AND CONSTRAINTS

Taking partial derivatives of equation (6.40) with respect to λ_1, λ_2, a and b, and equating each derivative to zero, one obtains

$$\frac{\partial I}{\partial \lambda_1} = 0 = \frac{2b^2\lambda_1}{2\lambda_2} - 2a\frac{\lambda_1}{\lambda_2} + \frac{2a^2\lambda_1}{2b^2\lambda_2} - \frac{a}{\lambda_2} + \frac{b^2}{\lambda_2} - E[\ln x] + \frac{a}{b^2}E[\ln x] \tag{6.41}$$

$$\frac{\partial I}{\partial \lambda_2} = 0 = -\frac{1}{2\lambda_2} - \frac{b^2}{2\lambda_2^2} - \frac{b^2\lambda_1^2}{2\lambda_2^2} + \frac{a\lambda_1^2}{\lambda_2^2} - \frac{a^2\lambda_1^2}{2b^2\lambda_2^2} + \frac{a\lambda_1}{\lambda_2^2} - \frac{b^2\lambda_1}{\lambda_2^2} + \frac{1}{2b^2}E[(\ln x)^2] \tag{6.42}$$

$$\frac{\partial I}{\partial a} = 0 = -\frac{\lambda_1^2}{\lambda_2} + \frac{2a\lambda_1^2}{2b^2\lambda_2} - \frac{\lambda_1}{\lambda_2} + \frac{\lambda_1}{b^2} E\,[lnx] \tag{6.43}$$

$$\frac{\partial I}{\partial b} = 0 = \frac{1}{b} + \frac{2b}{2\lambda_2} + \frac{2b\lambda_1^2}{2\lambda_2} - \frac{2a^2\lambda_1^2}{2b^3\lambda_2} + \frac{2b\lambda_1}{\lambda_2} - \frac{2a\lambda_1}{b^3} E\,[lnx] - \frac{2\lambda_2}{2b^3} E\,[(lnx)^2] \tag{6.44}$$

Simplification of equations (6.41) through (6.44) leads, respectively, to

$$E\,[\,\ln x\,] = a \tag{6.45}$$

$$E\,[\,(\,\ln x\,)^2\,] = a^2 + b^2 \tag{6.46}$$

$$E\,[\,\ln x\,] = a \tag{6.47}$$

$$E\,[\,(\,\ln x\,)^2\,] = a^2 + b^2 \tag{6.48}$$

Thus, the parameter estimation equations are equations (6.45) and (6.46).

6.3 Other Methods of Parameter Estimation

6.3.1 METHOD OF MOMENTS

Chow (1954, 1959) developed a graphical method to determine the frequency factor of the lognormal (LN2) distribution. The LN2 distribution has two parameters; therefore two moments will suffice. Kite (1978) has given the first two moments of X as

$$M_1 = \exp[a + (b^2/2)] \tag{6.49}$$

$$M_2^c = (M_1)^2\,[\exp(b^2) - 1\,] \tag{6.50}$$

where M_1 and M_2^c are, respectively, the first moment about the origin and the second moment about the centroid in the x domain. Taking natural logarithm of equations (6.49) and (6.50).

$$\ln M_1 = a + \frac{b^2}{2} \tag{6.51}$$

$$\ln M_2^c = 2\ln M_1 + \ln[\exp(b^2) - 1] \tag{6.52}$$

Equations (6.51) and (6.52) are solved with M_1 and M_2^c replaced by their sample estimates. The parameter estimation equations are

$$b^2 = \ln[\,(M_2^c / M_1{}^2) + 1] \tag{6.53}$$

$$a = \ln M_1 - \frac{b^2}{2} \tag{6.54}$$

Equations (6.53) and (6.54) relate b and a to the sample mean M_1 and sample variance, M_2^c.

6.3.2 METHOD OF MAXIMUM LIKELIHOOD ESTIMATION

The log-likelihood function of a sample of size n drawn from an LN2 distribution is given as

$$\ln L = \sum_{i=1}^{n} \ln[\frac{1}{x_i}] - \frac{n}{2}\ln(2\Pi) - \frac{n}{2}\ln b^2 - \frac{1}{2b^2}\sum_{i=1}^{n}[\ln x_i - a]^2 \tag{6.55}$$

Differentiating equation (6.55) with respect to parameter a and equating the derivative to zero give

$$a = \frac{1}{n}\sum_{i=1}^{n} \ln x_i \tag{6.56}$$

Differentiating equation (6.55) with respect to b and equating the derivative to zero yield

$$b^2 = \frac{1}{n}\sum_{i=1}^{n}[\ln x_i - a]^2 \tag{6.57}$$

Equations (6.56) and (6.57) correspond to the estimates given by the method of moments when the transformed observations y = ln x are used which, in turn, are normally distributed.

6.3.3 METHOD OF PROBABILITY-WEIGHTED MOMENTS

The LN2 distribution cannot be explicitly expressed in terms of x and therefore complex algebra is needed to derive the probability-weighted moments (PWMs). Hosking (1990) derived PWMs of the LN2 distribution in terms of L-moments:

$$L_1 = \exp[a + (b^2/2)] \tag{6.58}$$

$$L_2 = \exp[a + \frac{b^2}{2}]\, erf\,(\frac{b}{2}) \tag{6.59}$$

where L_1 and L_2 are first and second order L-moments, and erf (.) is the error function defined as

$$erf\,(x) = \frac{2}{\sqrt{\Pi}} \int_0^x \exp(-u^2)\, du = 2\ F\,(x\sqrt{2}) - 1 \tag{6.60}$$

where F (.) Is the normal distribution. Equations (6.58) and (6.59) are solved for parameters a and b by replacing L_1 and L_2 by their sample estimates:

$$b = 2\, erf^{-1}(\frac{L_2}{L_1}) = 2\, erf^{-1}(t) \tag{6.61}$$

$$a = \ln\ L_1 - \frac{b^2}{2} \tag{6.62}$$

In equation (6.61) erf^{-1} (t) is evaluated using equation (6.60) as u/ $(2)^{0.5}$, where u is the standard normal variate corresponding to F = (t + 1)/2.

References

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

Alexander, G.N., Karoly, A. and Susts, A.B., 1969a. Equivalent distributions with application to rainfall as an upper bound to flood distributions. *Journal of Hydrology*, Vol.**79**, No. 4, pp. 322-344.

Alexander, G.N., Karoly, A. and Susts, A.B., 1969b. Equivalent distributions with application to rainfall as an upper bound to flood distributions. *Journal of Hydrology*, Vol. **9**, No.3, pp. 345-373.

Brakensiek, D.L., 1958. Fitting a generalized log-normal distribution to hydrologic data. *Transactions, American Geophysical Union*, Vol. **39**, No.3, pp. 469-473.

Charbeneau, R.J., 1978. Comparison of the two- and three-parameter lognormal distributions used in streamflow sysnthesis. *Water Resources Research*, Vol. **14**, No. 1, pp. 149-150.

Chow, V.T., 1951. A general formula for hydrologic frequency analysis. *Transactions, American Geophysical Union*, Vol. **32**, pp. 231-237.

Chow, V.T., 1954. The log-probability law and its engineering application. *Proceedings,* *American Society of Civil Engineers*, Vol. **80,** pp. 536-1 to 536-14.

Chow, V.T., 1959. Determination of hydrologic frequency factor. *Journal of the Hydraulics Division, ASCE,* Vol. **85**, pp. 93-98.

Freeze, A.R. and Cherry, J.C., 1979. *Ground Water.* Prentice Hall, Inc., Englewood Cliff, New Jersey.

Hosking, J.R.M., 1990. L-moments: Analysis and estimation of distributions using linear combinations of order statistics. *Journal of Royal Statistical Society*, Series B, Vol. **52,** No. 1, pp. 105-124.

Kalinske, A.A., 1946. On the log-probability law. *Transactions, American Geophysical Union,* Vol. **27,** No. V, pp. 709-710.

Kite, G. W., 1978. *Frequency and Risk Analyses in Hydrology.* Water Resources Publications, Fort Collins, Colorado.

Lewis, G., 1979. A statistical estimation of flood flows. *Proceedings, Institution of Civil Engineers*, Part **2**, Vol. **67**, pp. 841-844.

Moran, P.A.P., 1957. The statistical treatment of flood flows. *Transaction, American Geophysical Union*, Vol. **38**, No. 4, pp. 519-523.

Parkin, T.B., Meisinger, J.J., Chester, S.T., Starr, S.T., and Robinson, J.A., 1988. Evaluation of statistical estimation methods for lognormally distributed variables. *Soil Science Society of America Journal*, Vol. **52,** pp. 323-329.

Parkin, T.B. and Robinson, J.A., 1993. Statistical evaluation of median estimators for lognormally distributed variables. *Soil Society of America Journal*, Vol. **57**, pp. 317-323.

Singh, V.P. and Rajagopal, A.K., 1986. A new method of parameter estimation. *Hydrological Science and Technology*, Vol. **2,** No. 3, pp. 33-40.

Singh, V.P., Rajagopal, A.K. and Singh, K., 1986. Derivation of some frequency distributions using the principle of maximum entropy. *Advances in Water Resources*, Vol. **9**, pp. 91-106.

Singh, V.P., Singh, K. and Rajagopal, A.K., 1985. Application of the principle of maximum entropy. *Completion Report* **06**, Louisiana Water Resources Research Institute, Louisiana State University, Baton Rouge, Louisiana.

Stedinger, J.R., 1980. Fitting log normal distributions to hydrologic data. *Water Resources Research,* Vol. **16,** No. 3, pp. 481-490.

Weiss, L.L., 1957. A nomogram for log-normal frequency analysis. *Transactions, American Geophysical Union*, Vol. 38, No. 1, pp. 33-37.

CHAPTER 7

THREE-PARAMETER LOGNORMAL DISTRIBUTION

The three-parameter lognormal (TPLN)distribution is frequently used in hydrologic analysis of extreme floods, seasonal flow volumes, duration curves for daily streamflow, rainfall intensity-duration, soil water retention, etc. It is also popular in synthetic streamflow generation. Properties of this distribution are discussed by Aitchison and Brown (1957), and Johnson and Kotz (1970). Its applications are discussed by Slade (1936), Chow (1954), Matalas (1967), Sangal and Biswas (1970), Fiering and Jackson (1971), Snyder and Wallace (1974), Burges et al. (1975), Burges and hoshi (1978), Charbeneau (1978), Stedinger (1980), Singh and Singh (1987), Kosugi (1994), among others. Burges et al. (1975) discussed properties of the three-parameter lognormal distribution and compared two methods of estimation of the third parameter "a". Kosugi (1994) applied the three-parameter lognormal distribution to the pore radius distribution function and to the water capacity function which was taken to be the pore capillary distribution function. He found that three parameters were closely related to the statistics of the pore capillary pressure distribution function, including the bubbling pressure, the mode of capillary pressure, and the standard deviation of transformed capillary distribution function. Burges and Hoshi (1978) proposed approximating the normal populations with 3-parameter lognormal distributions to facilitate multivariate hydrologic disaggregation or generation schemes in cases where mixed normal and lognormal populations existed.

Several estimation techniques have been applied to estimate parameters of the three-parameter lognormal distribution. Sangal and Biswas (1970) used the median method, comprising the mean, median and standard deviation, to estimate the three parameters. Bates et al. (1974) applied the median method and the skew method to estimate the parameters and provided tables of parameters. Snyder and Wallace (1974) fitted a lognormal distribution using the method of least squares. Using the mean square error of selected quantiles, Stedinger (1980) evaluated the efficiency of alternative methods of fitting, including method of moments (using sample moment estimators), quantile method (using sample mean, variance, and quantile estimate of the lower bound), method of moments (using unbiased standard deviation and skew coefficient), and quantile method with moment estimates of the first two parameters. Hoshi et al. (1984) compared, using average bias and root mean square error, the maximum likelihood estimation (MLE) method, method of moments, and two quantile-lower bound estimators in combination with two moments in real or in log space. Singh and Singh (1987) applied the principle of maximum entropy to estimate the TPLN parameters and compared it with the method of moments and maximum likelihood estimation. Using Monte Carlo simulation, Singh et al. (1990) estimated parameters and quantiles of the three-parameter lognormal distribution using the method of moments, modified method of moments, maximum likelihood estimation,

V. P. Singh, *Entropy-Based Parameter Estimation in Hydrology*

modified maximum likelihood estimation and entropy. Stevens (1992) employed MLE in which historical data could also be included. Using Monte Carlo simulation he demonstrated that inclusion of historical data reduced the bias and variance of extreme flows.

For a random variable X, if Y=ln(X-a) has a normal distribution then X will have a lognormal distribution whose probability density function (pdf) can be expressed as

$$f(x) = \frac{1}{(x-a)\, c\sqrt{2\pi}} \exp\left[\frac{-[\ln(x-a)-b]^2}{2c^2}\right] \tag{7.1a}$$

where 'a' is a positive quantity defined as a lower boundary, and b and c^2 are the form and scale parameters of the distribution. It turns out that b and c^2 are equal to the mean $(\bar{y})$ and variance s_y^2 of ln (x-a). Thus, the TPLN distribution has three parameters: a, b, and c. The three-parameter lognormal (LN3) distribution is similar to the two-parameter lognormal (LN2) distribution, except that x is shifted by an amount a which represents a lower bound. Thus, (x-a) represents a shifted variable. The standardized variable u is obtained in the usual manner as

$$u = \frac{\ln(x-a)-b}{c} \tag{7.1b}$$

The cumulative distribution function (cdf) of the TPLN distribution can be written as

$$F(x) = \int_a^x \frac{1}{(x-a)c\sqrt{2\pi}} \exp\left[-\frac{(\ln(x-a)-b)^2}{2c^2}\right] dx \tag{7.2}$$

Because of the integral nature of equation (7.2), it is not possible to express the LN3 distribution in terms of x as a function of F.

7.1 Ordinary Entropy Method

7.1.1 SPECIFICATION OF CONSTRAINTS

Integrating equation (7.1a) we obtain:

$$\int_a^\infty f(x)dx = \frac{1}{c\sqrt{2\pi}} \int_a^\infty \frac{1}{(x-a)} \exp\left[\frac{-[\ln(x-a) - b]^2}{2c^2}\right] dx \tag{7.3}$$

Let

$$z = \frac{\ln(x-a) - \bar{y}}{c}; \quad \frac{dz}{dx} = \frac{1}{(x-a)c} \tag{7.4}$$

Substituting equation (7.4) in equation (7.3), we get

$$\int_a^\infty f(x)dx = \frac{1}{c\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{1}{(x-a)} \exp\,[(-z^2/2)]dz\,(x-a)c$$
$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\,[(-z^2/2)]dz = \frac{2}{\sqrt{2\pi}} \int_0^\infty exp\,[(-z^2/2)]dz \tag{7.5}$$

Let $\frac{z}{\sqrt{2}} = k$. Then equation (7.5) can be written as

$$\int_a^\infty f(x)dx = \frac{2}{\sqrt{2\pi}} \int_0^\infty \exp\,(-k^2)\,\sqrt{2}\,dk$$
$$= \frac{2\sqrt{2}}{\sqrt{2\pi}} \int_0^\infty exp\,(-k^2)\,dk = \frac{2\sqrt{2}}{\sqrt{2\pi}} \frac{\sqrt{\pi}}{2} = 1 \tag{7.6}$$

Taking logarithm of equation (7.1a) to the base 'e' results in

$$\ln f(x) = -\ln[c\sqrt{2\pi}] - \ln(x-a) - \frac{[\ln(x-a)-b]^2}{2c^2}$$
$$= -\ln[c\sqrt{2\pi}] - \ln(x-a) - \frac{[\ln(x-a)]^2}{2c^2} - \frac{b^2}{2c^2} + \frac{b\ln(x-a)}{c^2} \tag{7.7}$$

Multiplying equation (7.7) by [-f(x)] and integrating between a to ∞ , one obtains the entropy function:

$$I(f) = -\int_a^\infty f(x)\,ln\,f(x)\,dx = [ln(c\sqrt{2\pi}) + \frac{b^2}{c^2}] \int_a^\infty f(x)\,dx$$
$$+ (1-\frac{b}{c^2}) \int_a^\infty ln(x-a)\,f(x)\,dx + \frac{1}{2c^2} \int_a^\infty [ln(x-a)]^2 f(x)dx \tag{7.8}$$

From equation (7.8), the constraints appropriate for equation (7.1a) can be written (Singh et al., 1985, 1986) as

$$\int_a^\infty f(x)dx = 1 \tag{7.9}$$

$$\int_a^\infty ln(x-a) = E\,[ln(x-a)] = E\,[y] = \bar{y} \tag{7.10}$$

$$\int_a^\infty (ln(x-a))^2 f(x)dx = E[(ln(x-a))^2] = var[ln(x-a)] + \bar{y}^2 = s_y^2 + \bar{y} \qquad (7.11)$$

where var [.] is the variance of [.].

7.1.2 CONSTRUCTION OF ZEROTH LAGRANGE MULTIPLIER

The least-biased pdf consistent with equations (7.9) to (7.11) and corresponding to the principle of maximum entropy (POME) takes the form:

$$f(x) = \exp\ [-\lambda_0 - \lambda_1 \ln(x-a) - \lambda_2 (\ln(x-a))^2] \qquad (7.12)$$

where λ_0, λ_1, and λ_2 are Lagrange multipliers. Integrating equation (7.12) between a and ∞, one gets

$$\int_a^\infty f(x)dx = \int_a^\infty \exp(-\lambda_0 - \lambda_1 \ln(x-a) - \lambda_2 (\ln(x-a))^2]dx \qquad (7.13)$$

Because the left side of equation (7.13) equals one by virtue of equation (7.9), the partition function becomes

$$\exp(\lambda_0) = \int_a^\infty \exp[-\lambda_1 \ln(x-a) - \lambda_2 (\ln(x-a))^2]dx \qquad (7.14)$$

In order to evaluate the above integral, let

$$z = ln(x-a);\ [dz/dx] = [1/(x-a)] \qquad (7.15)$$
$$(x-a) = \exp(z);\ dx = \exp(z)\ dz$$

Substituting these quantities in equation (7.14), one gets:

$$exp\ (\lambda_0) = \int_{-\infty}^\infty exp\ [-\lambda_1 z - \lambda_2 z^2]\ e^z\ dz = \int_{-\infty}^\infty exp\ [-(\lambda_1 - 1)z - \lambda_2 z^2]dz = \frac{\sqrt{\pi}}{\lambda_2}\ exp\ [\frac{(\lambda_1 - 1)^2}{4\lambda_2}] \qquad (7.16)$$

Taking logarithm of equation (7.16) results in the zeroth Lagrange multiplier λ_0 given as

$$\lambda_0 = \frac{1}{2} \ln \pi - \frac{1}{2} \ln \lambda_2 + \frac{(\lambda_1 - 1)^2}{4\lambda_2} \tag{7.17}$$

One can also write the zeroth Lagrange multiplier from equation (7.14) as

$$\lambda_0 = \ln \int_a^\infty \exp[-\lambda_1 \ln(x-a) - \lambda_2 \{\ln(x-a)\}^2] dx \tag{7.18}$$

7.1.3 RELATION BETWEEN LAGRANGE MULTIPLIERS AND CONSTRAINTS

Differentiating equation (7.18) with respect to λ_1 and λ_2, one gets

$$\begin{aligned} \frac{\partial \lambda_0}{\partial \lambda_1} &= -\frac{\int_a^\infty ln(x-a) exp[-\lambda_1 ln(x-a) - \lambda_2 (ln(x-a))^2] dx}{\int_a^\infty \exp[-\lambda_1 ln(x-a) - \lambda_2 (ln(x-a))^2] dx} \\ &= -\int_a^\infty ln(x-a) exp[-\lambda_0 - \lambda_1 ln_1(x-a) - \lambda_2 (ln(x-a)^2)] dx \\ &= -\int_a^\infty ln(x-a) f(x) dx = -\bar{y} \end{aligned} \tag{7.19}$$

$$\begin{aligned} \frac{\partial \lambda_0}{\partial \lambda_2} &= -\frac{\int_a^\infty (ln(x-a))^2 exp[-\lambda_1 ln(x-a) - \lambda_2 (ln(x-a))^2] dx}{\int_a^\infty exp[-\lambda_1 ln(x-a) - \lambda_2 (ln(x-a))^2] dx} \\ &= -\int_a^\infty (ln(x-a))^2 exp[-\lambda_0 - \lambda_1 ln(x-a) - \lambda_2 (ln(x-a))^2] dx \\ &= -\int_a^\infty [ln(x-a)]^2 f(x) dx \\ &= -E[(ln(x-a))^2] = -(s_y^2 + \bar{y}^2) \end{aligned} \tag{7.20}$$

Furthermore, one can write

$$\frac{\partial^2 \lambda_0}{\partial \lambda_2^2} = Var\ [ln(x-a)]^2 \tag{7.21}$$

Also, differentiating equation (7.17) with respect to λ_1 and λ_2 we get

$$\frac{\partial \lambda_0}{\partial \lambda_1} + \frac{\lambda_1 - 1}{2\lambda_2} \tag{7.22}$$

$$\frac{\partial \lambda_0}{\partial \lambda_1} = -\frac{(\lambda_1 - 1)^2}{4\lambda_2} - \frac{1}{2\lambda_2} \tag{7.23}$$

$$\frac{\partial^2 \lambda_0}{\partial \lambda_2^2} = \frac{1}{2\lambda_2^2} + \frac{(\lambda_1 - 1)^2}{2\lambda_2^3} \tag{7.24}$$

Equating equation (7.19) to (7.22), equation (7.21) to (7.24), as well as equation (7.21) to (7.24), we obtain

$$\lambda_1 = 1 - \frac{\bar{y}}{S_y^2} \tag{7.25}$$

$$\lambda_2 = \frac{1}{2S_y^2} \tag{7.26}$$

$$\frac{1}{2\lambda_2^2} + \frac{(\lambda_1 - 1)^2}{2\lambda_2^3} = var\ [ln(x-a)]^2 \tag{7.27}$$

7.1.4 RELATION BETWEEN LAGRANGE MULTIPLIERS AND PARAMETERS

Inserting equations (7.25) and (7.26) into equation (7.17), one gets

$$\lambda_0 = ln\sqrt{\pi} + ln(\sqrt{2}\ s_y) + \frac{\bar{y}^2}{2s_y^2} \tag{7.28}$$

Thus, the pdf can be written as

$$
\begin{aligned}
f(x) &= exp\ [-ln\sqrt{\pi} - ln(\sqrt{2}s_y) - \frac{\bar{y}}{2S_y^2} \\
&- (1 - \frac{\bar{y}}{2s_y^2})ln(x-a) - \frac{(ln(x-a))^2}{2s_y^2}] \\
&= exp[ln(\sqrt{\pi})^{-1}]\exp[\ln(S_y\sqrt{2})^{-1}]\ exp\ [-\frac{\bar{y}}{2s_y^2} - ln(x-a) \\
&+ \frac{\bar{y}}{s_y^2}ln(x-a) - \frac{(ln(x-a))^2}{2s_y^2}] \\
&= \frac{1}{s_y\sqrt{2\pi}}exp\ [ln(x-a)^{-1}]exp\ [-\frac{1}{2s_y^2}(ln(x-a)-\bar{y})^2] \\
&= \frac{1}{(x-a)s_y\sqrt{2\pi}}\ exp[-\frac{1}{2s_y^2}(ln(x-a)-\bar{y})^2]
\end{aligned} \tag{7.29}
$$

A comparison of equation (7.29) with equation (7.1a) shows that b = $\bar{y}$ and c = s_y .

7.1.5 RELATION BETWEEN PARAMETERS AND CONSTRAINTS

The 3-parameter lognormal distribution has 3 parameters a, b, and c which are related to the Lagrange multipliers by equations (7.25) to (7.27), which, in turn, are related to the constraints by equations (7.9) to (7.11). Eliminating the Lagrange multipliers through these two sets of equations, we obtain parameters in terms of constraints. The third equation for parameter estimation is obtained by equating (7.21) to (7.24):

$$
E\ [ln(x-a)]^4 = 2s_y^2\ [2\bar{y}^2 + s_y^2] \tag{7.30}
$$

7.1.6. DISTRIBUTION ENTROPY

The entropy of LN3 distribution is obtained as follows:

$$
\begin{aligned}
I(x) &= -\int_a^\infty f(x)\ ln\ f(x)dx = [ln(s_y\sqrt{2\pi} + \frac{\bar{y}^2}{2s_y^2}]\int_a^\infty f(x)dx \\
&+ (1-\frac{\bar{y}}{s_y^2})\int_a^\infty ln(x-a)\ f(x)dx + \frac{1}{2s_y^2}\int_a^\infty (ln(x-a))^2 f(x)dx \\
&= ln(s_y\sqrt{2\pi}) + \frac{\bar{y}^2}{2s_y^2} + (1-\frac{\bar{y}}{s_y^2})\bar{y} + \frac{1}{2s_y^2}(s_y^2 + \bar{y}^2) \\
&= ln(s_y\sqrt{2\pi e}) + \bar{y}
\end{aligned} \tag{7.31}
$$

Therefore,

$$I(x) = I(y) + \bar{y} \tag{7.32}$$

where $I(y) = ln(s_y\sqrt{2\pi e})$. Alternatively,

$$I(x) = I(y) - E\,[ln\left|J(\frac{y}{x})\right|] \tag{7.33}$$

where J(y/x) is the Jacobian and can be expressed as

$$J(\frac{y}{x}) = \frac{\partial y}{\partial x} = \frac{\partial\, ln(x-a)}{\partial x} = \frac{1}{x-a} \tag{7.34}$$

$$\begin{aligned} I(x) &= I(y) - E\,[ln\left|\frac{1}{x-a}\right|] \\ &= I(y) + E\,[ln(x-a)] = I(y) + \bar{y} \end{aligned} \tag{7.35}$$

which is the same as equation (7.32).

7.2 Parameter - Space Expansion Method

7.2.1 SPECIFICATION OF CONSTRAINTS

Following Singh and Rajagopal (1986), the constraints for this method are specified by equation (7.9) and

$$\begin{aligned} &\int_a^\infty [-ln(x-a) + \frac{b}{c^2}ln(x-a)]\, f(x)dx \\ &= E\,[-ln(x-a) + \frac{b}{c^2}ln(x-a)] \end{aligned} \tag{7.36}$$

and

$$\begin{aligned} &\int_a^\infty [\frac{(ln(x-a))^2}{2c^2}]\, f(x)dx \\ &= E\,[\frac{(ln(x-a))^2}{2c^2}] \end{aligned} \tag{7.37}$$

7.2.2 DERIVATION OF ENTROPY FUNCTION

The pdf corresponding to the principle of maximum entropy (POME) and consistent with equations (7.9), (7.36), and (7.37) takes the form

$$f(x) = exp\,[-\lambda_0 - \lambda_1 \frac{b}{c^2} ln(x-a) + \lambda_1 ln(x-a) - \lambda_2 \frac{(ln(x-a))^2}{2c^2}] \tag{7.38}$$

where λ_0, λ_1, and λ_2 are Lagrange multipliers. Insertion of equation (7.38) into equation (7.9) yields

$$\int_a^\infty f(x) = 1 = \int_a^\infty exp\,[-\lambda_0 - \lambda_1 \frac{b}{c^2} ln(x-a) + \lambda_1 ln(x-a) - \frac{\lambda_2}{2c^2}(ln(x-a))^2]dx \tag{7.39}$$

This leads to

$$exp\,(\lambda_0) = \frac{c\sqrt{2\pi}}{\sqrt{\lambda_2}}\, exp\,(\frac{d^2c^2}{2\lambda_2}), \quad d = \frac{b\lambda_1}{c^2} - \lambda_1 - 1 \tag{7.40}$$

The zeroth Lagrange multiplier is given by

$$\lambda_0 = ln\ c + \frac{1}{2} ln(2\pi) - \frac{1}{2} ln \lambda_2 + \frac{d^2c^2}{2\lambda_2} \tag{7.41}$$

Also, from equation (7.22) we obtain

$$\lambda_0 = ln \int_a^\infty exp\,[(-\lambda_1 \frac{b}{c^2} ln(x-a) + \lambda_1 ln(x-a) - \frac{\lambda_2}{2c^2}(ln(x-a))^2]dx \tag{7.42}$$

Introduction of equation (7.41) in equation (7.39) gives

$$f(x) = \frac{\sqrt{\lambda_2}}{c\sqrt{2\pi}} \; exp \; [- \frac{d^2c^2}{2\lambda_2} - \lambda_1 \frac{b}{c^2} ln(x-a)$$
$$+ \lambda_1 ln(x-a) - \frac{\lambda_2}{2c^2}(ln(x-a))^2]dx \tag{7.43}$$

A comparison of equation (7.43) with equation (7.1a) shows that $\lambda_1 = -1$ and $\lambda_2 = 1$. Taking - logarithm of equation (7.43), one gets

$$-lnf(x) = -\frac{1}{2} ln\lambda_2 + ln\; c + \frac{1}{2} ln(2\pi)$$
$$+ \frac{d^2c^2}{2\lambda_2} + \frac{\lambda_1 b}{c^2} ln(x-a)$$
$$- \lambda_1 ln(x-a) + \frac{\lambda_2}{2c^2} (ln(x-a))^2 \tag{7.44}$$

The entropy function then becomes

$$I(f) = -\frac{1}{2} ln\lambda_2 + ln\; c + \frac{1}{2} ln(2\pi) + \frac{d^2c^2}{2\lambda_2}$$
$$+ \frac{\lambda_1 b}{c^2} E\,[ln(x-a)] - \lambda_1 E\,[ln(x-a)] + \frac{\lambda_2}{2c^2} E\,[(ln(x-a))^2] \tag{7.45}$$

7.2.3 RELATION BETWEEN PARAMETERS AND CONSTRAINTS

Taking partial derivatives of equation (7.45) with respect to λ_1, λ_2, a, b, and c, and equating each derivative to zero, one gets

$$\frac{\partial I}{\partial \lambda_1} = 0 = -\frac{2b}{c^2}(\frac{b}{c^2} - 1)\frac{c^2}{2\lambda_2} + \frac{b}{c^2} E\,[ln(x-a)] - E\,[ln(-a)] \tag{7.46}$$

$$\frac{\partial I}{\partial \lambda_2} = 0 = -\frac{1}{2\lambda_2} - \frac{d^2c^2}{2\lambda_2^2} + \frac{1}{2c^2} E\,[(ln(x-a))^2] \tag{7.47}$$

$$\frac{\partial I}{\partial a} = 0 = -\frac{\lambda_1 b}{c^2} E\left[\frac{1}{x-a}\right] + \lambda_1 E\left[\frac{1}{x-a}\right] - \frac{\lambda_2}{c^2} E\left[\frac{ln(x-a)}{x-a}\right] \tag{7.48}$$

$$\frac{\partial I}{\partial b} = 0 = \frac{\lambda_1}{c^2} ln(x-a) + \frac{\lambda_1}{\lambda_2}\left(\frac{\lambda_1 b}{c^2} - \lambda_1 - 1\right) \tag{7.49}$$

$$\frac{\partial I}{\partial c} = 0 = \frac{1}{c} + \frac{1}{2\lambda_2}\left[2c^2\left(\frac{\lambda_1 b}{c^2} - \lambda_1 - 1\right)\left(-\frac{2}{c^3}\right)\lambda_1 b + 2c\left(\frac{\lambda_1 b}{c^2} - \lambda_1 - 1\right)^2\right] - \frac{2\lambda_1 b}{c^3} E[ln(x-a)] - \frac{2\lambda_2}{2c^3} E[ln(x-a))^2] \tag{7.50}$$

Simplification of equation (7.46) to (7.50) produces

$$E[ln(x-a)] = b \tag{7.51}$$

$$E[(ln(x-a))^2] = b^2 + c^2 \tag{7.52}$$

$$E\left[\frac{ln(x-a)}{x-a}\right] = (b-c^2)\, E\left[\frac{1}{x-a}\right] \tag{7.53}$$

$$E[ln(x-a)] = b \tag{7.54}$$

$$c\, E[(ln(x-a))^2] - 2cb\, E[ln(x-a)] = 4b^3 - c^2 \tag{7.55}$$

Equations (51) and (54) are the same and equation (7.55) is an identity. Thus, the parameter estimation equations are equations (7.51), (7.52), and (7.53).

7.3 Other Methods of Parameter Estimation

Three of the most popular methods of parameter estimation are the methods of moments (MOM), probability-weighted moments (PWM) and maximum likelihood estimation (MLE). The variants of MOM and MLE have also been reported in the literature. To this end we briefly summarize these methods.

7.3.1 REGULAR METHOD OF MOMENTS

For the regular method of moments (RMOM), the r-th moment of equation (7.1a) about the lower bound `a` is

$$M_r^a = \int_0^\infty \frac{(x-a)^r}{(x-a)\,c\sqrt{2\pi}} \exp\left[-\frac{\{\ln(x-a)-b\}^2}{2c^2}\right] dx \tag{7.56}$$

Let y = ln (x - a). Equation (7.56) can be written as

$$M_r^a = \frac{1}{c\sqrt{2\pi}} \int_0^\infty \exp\left[-\frac{1}{2c^2}\{(y-b^2 - 2ryc^2\}\right] dy$$

$$= \exp\left(rb + \frac{r^2c^2}{2}\right) \tag{7.57}$$

$$M_1^a = \exp\left(b + \frac{c^2}{2}\right) \tag{7.58}$$

$$M_2^a = \exp(2b + 2c^2) \tag{7.59}$$

$$M_3^a = \exp\left(3b + \frac{9}{2}c^2\right) \tag{7.60}$$

The moments given by (7.57) can be converted to the moments about the origin by using the following expression:

$$M_r^0 = \sum_{j=0}^{r} \binom{r}{j} M_{r-j}^a\, a^j \tag{7.61}$$

Therefore,

$$M_1^0 = \exp\left(b + \frac{c^2}{2}\right) + a \tag{7.62}$$

$$M_2^0 = \exp(2b + 2c^2) + 2a\exp\left(b + \frac{c^2}{2}\right) + a^2 \tag{7.63}$$

$$M_3^0 = \exp\left(3b + \frac{9}{2}c^2\right) + 3a\exp(2b + 2c^2) + 3a^2\exp\left(b + \frac{c^2}{2}\right) \tag{7.64}$$

Furthermore,

$$M_1 = \bar{x} = \exp\left(b + \frac{c^2}{2}\right) + a \tag{7.65}$$

$$M_2 = var(x) = M_2^0 - (M_1^0)^2 = \exp(2b + c^2)[\exp(c^2) - 1] \tag{7.66}$$

For purposes of parameter estimation, it may be useful to recall the relationships between characteristics of X and Y. If μ_x, σ_x^2 and G_x denote mean, variance and skewness of the lognormal variate X, and μ_y and σ^2_y the mean and variance of Y then one can show that

$$\mu_x = a + \exp(\mu_y)\, w^{1/2} \tag{7.67}$$

$$\sigma_x^2 = \exp(2\mu_y)\; w(w-1) \tag{7.68}$$

$$G_x = (w+2)\;(w-1)^{1/2} \tag{7.69}$$

where w = $\exp(\sigma_y^2)$. If we let $\theta = (w-1)^{1/2}$, then

$$G_x = \theta^3 + 3\theta \tag{7.70}$$

Equation (7.70) can be solved for θ expressed (Kite, 1978) as

$$\theta = \frac{1-B^{2/3}}{B^{1/3}}\;,\; B = \frac{1}{2}\left[-G_x + (G_x^2+4)^{1/2}\right] \tag{7.71}$$

Therefore,

$$w = \theta^2 + 1 \tag{7.72}$$

$$\sigma_y^2 = \ln w \tag{7.73}$$

$$\mu_y = \frac{1}{2}\ln\left[\frac{\sigma_x^2}{w(w-1)}\right] \tag{7.74}$$

$$a = \mu_x - \frac{\sigma_x}{(w-1)^{1/2}} = \mu_x - \frac{\sigma_x}{\theta} \tag{7.75}$$

The quantities μ_x, σ_x^2 and G_x are for the population of X, but are estimated from a sample of size n as:

$$\hat{\mu}_x = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{7.76}$$

$$\hat{\sigma_x^2} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\hat{\mu}_x)^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^{\,2} - n\,\hat{\mu}_x^2\right) \tag{7.77}$$

$$\hat{G}_x = \frac{n}{(n-1)(n-2)}\,\frac{1}{\hat{\sigma}_x^3}\sum_{i=1}^{n}(x_1 - \hat{\mu}_x)^3$$

$$= \frac{n}{(n-1)(n-2)}\,\frac{1}{\hat{\sigma}_x^3}\left(\sum_{i=1}^{n} x_i^3 - 3\hat{\mu}_x\sum_{i=1}^{n} x_i^2 + 2n\,\hat{\mu}_x^3\right) \tag{7.78}$$

7.3.2 MODIFIED METHOD OF MOMENTS

The RMOM is modified to obtain the modified method of moments (MMOM) in estimation of w as suggested by Cohen and Whitten (1980) who derived the following relationship:

$$\frac{\sigma_x^2}{(\mu_x - x_1)} = \frac{w(w-1)}{[\sqrt{w} - \exp(\sigma_y E[Z_1])]^2} \tag{7.79}$$

where x_1 is the first order statistic, and $E[Z_1]$ denotes the expected value of the first order statistic of the standard normal variate for a random sample of size n (Singh et al., 1990). Here the "fixed point" method can be used to solve equation (7.79) for w. To that end, let $u = \sigma_x / (\mu_x - x_1)$. Rearranging equation (7.79),

$$(w - 1)^{1/2} = \frac{u(\sqrt{w} - \exp(\sigma_y E[Z_1]))}{\sqrt{w}} \tag{7.80}$$

or

$$\theta = \frac{u(\sqrt{w} - \exp(\sigma_y E[Z_1]))}{\sqrt{w}}, \quad w = \theta^2 + 1 \tag{7.81}$$

First, equation (7.80) is solved iteratively for w using a standard numerical method such as the Newton-Raphson method. Then equations (7.73) - (7.75) are used to estimate the parameters a, μ_y and σ_y.

7.3.3 REGULAR METHOD OF MAXIMUM LIKELIHOOD ESTIMATION

For the regular maximum lukelihood method (RMLE), the likelihood function of receiving the sample data D = $\{x_1, x_2, \ldots, x_n\}$ from a TPLN population, given the values of a, b, and c, is

$$L(D|a,b,c) = \prod_{i=1}^{n} L(x_i) \tag{7.82}$$

Therefore,

$$L(D|a,b,c) = \frac{1}{[x_1 - a)\ldots(x_n - a)c^n(\sqrt{2\pi})^n}$$

$$\exp[-\frac{1}{2c^2}\{(\ln(x_1 - a) - b^2 + \ldots + (\ln(x_n - a) - b)^2\}] \tag{7.83}$$

The MLE method involves finding the values of a, b, c^2 which together maximize the likelihood function. If $L(D|a,b,c)$ is maximal then so is ln $L(D|a,b,c)$ so estimates of a, b and c are sought which produce

$$\frac{\partial}{\partial a}\,[\ln L(D\,|\,a,b,c^2)] = 0 \tag{7.84}$$

$$\frac{\partial}{\partial b}\,[\ln L(D\,|\,a,b,c^2)] = 0 \tag{7.85}$$

$$\frac{\partial}{\partial c}\,[\ln L(D\,|\,a,b,c^2)] = 0 \tag{7.86}$$

Equations (7.84) - (7.86) lead to:

$$\hat{\mu}_y = \frac{1}{n}\sum_{i=1}^{n} \ln(x_1 - a) \tag{7.87}$$

$$\hat{\sigma}_y^2 = \frac{1}{n}\sum_{i=1}^{n} [\ln(x_i - a) - \hat{\mu}_y]^2 = \frac{1}{n}\sum_{i=1}^{n} [\ln(x_i - a)]^2 - [\frac{1}{n}\sum_{i=1}^{n} \ln(x_i - a)]^2 \tag{7.88}$$

$$f(a) = \sum_{i=1}^{n} (x_i = a)^{-1} (\hat{\mu}_y - \hat{\sigma}_y^2) - \sum_{i=1}^{n} [\frac{\ln(x_i - a)}{(x_i - a)}] \tag{7.89a}$$

$$= 0\,,\; a < x_i \tag{7.89b}$$

The function, f(a), can have multiple roots. In that case, the value of "a" closest to x_1 (the lowest value in the sample) is chosen except when n < 30 in which case experience has shown that the second root from x_1 is the MLE of "a." This is based upon the criterion that the value of "a" which results in the closest agreement between $\hat{\mu}_x$ and $\bar{x}$ should be chosen (Cohen and Whitten, 1980). Hence, the equations for parameter estimation are

$$\sum_{i=1}^{n} \frac{\ln(x_i - a)}{(x_i - a)} = \sum_{i=1}^{n} \frac{(b - c^2)}{(x_i - a)} \tag{7.90}$$

$$b = \frac{1}{n}\sum_{i=1}^{n} \ln\,(x_i - a) \tag{7.91}$$

$$c^2 = \frac{1}{n}\sum_{i=1}^{n} (\ln\,(x_i - a) - b)^2 \tag{7.92}$$

Equations (7.90) - (7.92) are nonlinear but can be easily solved numerically for a, b, and c.

7.3.4 MODIFIED MAXIMUM LIKELIHOOD ESTIMATION

The modified maximum likelihood estimation (MMLE) method differs from RMLE in estimation of the parameter a. Cohen and Whitten (1980) derived a function f(a) expressed as

$$f(a) = \ln(x_1 - a) - \frac{1}{n}\sum_{i=1}^{n} \ln(x_i - a) - E[Z_1]\,(\frac{1}{n}\sum_{i=1}^{n} [\ln(x_i - a)]^2$$

$$- [\frac{1}{n}\sum_{i=1}^{n} \ln(x_i - a)]^2)^{1/2} = 0 \tag{7.93}$$

or

$$f(a) = \ln(x_1 - a) - \mu_y - E[Z_1]\,\sigma_y = 0, a < x_1 \tag{7.94}$$

The other two equations (7.87) - (7.88) remain the same. The value of "a" is obtained using an iterative numerical method.

7.3.5 METHOD OF PROBABILITY-WEIGHTED MOMENTS

The PWM expressions for the LN3 distributions are difficult to get but are derived by Hosking (1990) in terms of L-moments given as follows:

$$L_1 = a + \exp[\mu_y + (\sigma_y^{\ 2}/2)] \tag{7.95}$$

$$L_2 = \exp[\mu_y + \frac{\sigma_y^2}{2}]\ erf\,(\frac{\sigma_y}{2}) \tag{7.96}$$

$$\lambda_3 = \frac{L_3}{L_2} = \frac{6}{\sqrt{\Pi}} \int_0^{\sigma_y/2} \frac{erf\,(x/\sqrt{3})\exp(-x^2)}{erf\,(\sigma_y/2)}\,dx \tag{7.97}$$

where L_i, i=1,2,3, are the L-moments, and erf (.) Is the error function. An approximate solution for parameter estimates (Hosking, 1990) follows:

$$\sigma_y = 0.999281z - 0.006118z^3 + 0.000127z^5 \tag{7.98}$$

$$\mu_y = \ln[L_2 / erf\,(\frac{\sigma_y}{2})] - \frac{\sigma_y^2}{2} \tag{7.99}$$

$$a = L_1 - \exp[\mu_y + \frac{\sigma_y^2}{2}] \tag{7.100}$$

where z is defined as

$$z = \sqrt{\frac{8}{3}}\ \Phi^{-1}\left(\frac{1+\lambda_3}{2}\right) \tag{7.101}$$

Φ (.) = F (.) is the standard normal distribution function. Therefore, Φ^{-1} (.) = F^{-1} = u = the standardized variable.

7.4 Comparative Evaluation of Estimation Methods

7.4.1 EXPERIMENTAL DESIGN

7.4.1.1 Monte Carlo Samples: To assess the performance of RMOM, MMOM, RMLE, MMLE, and POME parameter estimation methods, Singh et al. (1990) conducted Monte Carlo sampling experiments. Their work is followed here. Four TPLN cases, listed in Table 7.1, were considered. For each population case, 1,000, 1,500, and 2,000 random samples of size 10, 20, 30, 50, 75, 100, 200 and 400 were generated, and then parameters and quantiles were estimated. For quantile estimation, X(F) is not expressible in direct form and was obtained numerically, where F is the cumulative distribution function. The relative performance of the methods did not greatly depend on the number of samples generated.

Table 7.1 Lognormal population cases considered in the sampling experiments (μ = 10).

Lognormal Population	Coefficient of Variation (CV)	Coefficient of Skewness (G)
Case 1	0.5	0.5
Case 2	0.5	1.0
Case 3	0.5	1.5
Case 4	0.5	2.0

Singh et al. (1990) observed that for some Monte Carlo samples generated from the TPLN populations, the convergence was not obtained for some parameters and restrictions had to be imposed which generally involved parameter "a." For instance, for RMLE (or MMLE) and POME estimation "a" < x_1 . It has been found (Cohen and Whitten, 1980) that to attain proper convergence for the other parameters (μ_y, σ_y) in these cases, $\hat{a}$ must not be less than ($\hat{\mu}_y$ - $100\hat{\sigma}_x$) and not greater than ($x_1 - 10^{-8}$). Therefore, samples for which $\hat{a}$ did not fall within this restriction were rejected by the program for RMLE, MMLE and POME estimation. Likewise, for moments methods (RMOM, MMOM), the relationship $a = \mu_y - \sigma_x/\theta$ must hold.

From the convergence restriction $\hat{a} \geq (\hat{\mu}_x - 100\,\hat{\sigma}_x)$, we get $\theta = .01$. Since $w = \theta^2 + 1$, then $w \geq 1.0001$. Thus, samples where w > 1.0001 were rejected for the MMOM method. Similarly, since $G_x = \theta^3 + 3\theta = 0.3$, for samples with $G_x \leq .03$, the RMOM estimates could not be calculated.

7.4.1.2 Performance Indices: The 2,000 values of estimated parameters and quantiles for each sample size and population case were used to approximate standard bias (BIAS), standard error (SE), and root mean square (RMSE) for evaluating the performance the performance of the parameter estimation mmethods. Following Kuczera (1982a, 1982b), robustness was also used to evaluate the methods. Two criteria for identifying a resistant estimator are mini-max and minimum average RMSE. Based on the mini-max criterion, the preferred estimator is the one whose maximum RMSE for all population cases is minimum. The minimum average RMSE was used for the estimator whose RMSE average over the test cases was minimum. The number of samples of 2,000 may arguably not be large enough to produce the true values of BIAS, SE and RMSE, but was considered sufficient to compare the performance of the estimation methods.

7.4.2 RESULTS AND DISCUSSION OF MONTE CARLO EXPERIMENTS

7.4.2.1 Bias in Parameter Estimates: The results of the parameter bias analyses showed that in general, RMLE and POME performed consistently in estimating parameter "a" for all sample sizes of all population cases. As sample size increased, all methods, as expected, produced less bias in "a." Indeed the reduction in bias for each method was two orders of magnitude as sample size increased from 10 to 400. For small sample sizes of all population cases, the bias by RMOM was the highest, followed by MMLE, that by RMLE was the least, and that by MMOM and POME was comparable. For increasing sample sizes MMOM, RMLE and POME tended to be comparable. To summarize, RMLE performed very well across all population cases and thus appears to tbe the best estimator of "a" in terms of bias. For sample sizes $\geq$ 100, both POME and RMOM performed reasonably well and can be used for estimating "a."

The results of bias in estimators of μ_y varied with the sample size and skewness of the population considered. For small sample sizes ($n \leq 20$) of populations with $G \leq 0.5$, RMOM produced the least bias in μ_y and RMLE the highest bias, and the other methods were comparable. For the same sizes, the least bias was produced by POME for populations of higher G values. With increasing sample size, RMLE and POME took over and produced the least bias, and for n > 100, RMLE, was, in general, the best of all the methods. However, POME was the best for populations of $G \geq 2.0$.

The bias in the estimates of σ_y also varied with sample size and population skewness and somewhat mirrored the results of the bias in μ_y. For small sample sizes ($n \leq 100$) and $G_x \leq .5$, RMOM gave the least biased estimators of σ_y. POME produced the second best estimators of σ_y in terms of bias for small samples and the best for n > 100. MMLE produced the most biased estimators of σ_y for small skewed populations. However, as population skewness increased, MMLE and RMLE became the least biased estimators of σ_y, especially for larger sample sizes ($n \geq 50$). For the smaller samples, no method demonstrated consistently superior results; however MMOM, MMLE and POME generally performed well in these cases. It is noted that RMOM estimators of σ_y were negatively biased for all sample sizes for $G_x > .5$, and MMOE and POME resulted in negative bias for large samples for $G_x = 1.5$ and were negatively biased for all samples for $G_x = 2$.

7.4.2.2 Bias in Quantile Estimates: The results of the bias analysis for selected quantile estimates showed that for small population skewness($G_x \leq .5$) RMOM appeared to be the least biased estimators of all quantiles for all sample sizes with the exception of small quantiles ($T \leq 10$) for sample sizes larger than 100. Here T is the return period. For the larger sample sizes, MMOM performed well, while RMLE and POME performed similarly for this case. As population skewness increased, RMOM estimates became increasingly more biased until they were generally the most biased for all sample sizes for case 4. Meanwhile, the MMOM method did not deteriorate at the same rate as RMOM and performed well in estimating large quantiles (T > 50) for nearly all sample sizes even in case 4 ($G_x = 2.0$). In general, bias tends to increase with increasing quantiles and to decrease as sample size increases. In this regard, POME performed very well, particularly for the cases of high population skewness. In fact, POME generally showed the least deterioration for larger quantile estimation than any other method for all sample sizes in all population cases. RMLE and MMOM also performed well in this regard for the cases of high population skewness. In summary, for populations which exhibited small skewness ($G_x \leq .5$), ROME generally gave the least biased quantile estimates, while for populations with larger skewness, MMOM performed well for the estimation of large quantiles and POME performed consistently well for all sample sizes in these cases.

7.4.2.3 RMSE of Parameter Estimates: The results of the parameter RMSE analysis showed, in estimation of parameter "a," RMLE generally exhibited the smallest RMSE over all population cases for most sample sizes. In general, the RMSE of "a" decreased for all methods as population skewness and sample size increased. MMOM performed well in estimating "a" for sample sizes $\leq$ 100. RMOM was generally the worst estimator of "a" in terms of RMSE for small samples ($n \leq 50$) for small population skewness($G_x \leq 1.0$) and for all sample sizes for cases of $G_x > 1.0$.

The RMSE of the estimators of μ_y also varied according to population skewness and sample size. In the case of small skewness ($G_x \leq .5$), RMOM estimators of μ_y exhibited the lowest RMSE for samples of size n $\leq$ 20, while POME was superior for 30 $\leq$n$\leq$ 75 and RMLE was superior for cases of n > 75 with POME comparable. For increasing population skewness, MMOM exhibited superior RMSE for the smaller samples ($n \leq 100$), while RMLE was superior for cases of n > 100 for all population cases. However, POME was generally comparable as an estimator of μ_y in terms of RMSE for these cases. MMLE generally performed poorly for all sample sizes over all population cases.

As in the previous case, the RMSE results with respect to estimators of μ_y varied with population skewness and sample sizes. Again, for the small population skew case ($G_x \leq .5$), RMOM provided the lowest RMSE estimator of σ_y for sample sizes up to 75 where POME took over and was superior for the remaining samples. RMLE also performed well for the larger samples. Contrary to the previous two cases, RMSE of all estimators of σ_y increased slightly with increasing population skewness. Again, under increasing population skewness, MMOM estimators were superior in terms of RMSE for samples of sizes up to n = 100. As previously, POME and RMLE estimators of σ_y were superior for the largest sample sizes (n > 100) in all population cases. Also, as in the previous cases, MMLE performed relatively poorly in all cases.

7.4.2.4 RMSE of Quantile Estimates: The results of the RMSE analysis of selected quantile estimates showed that for small sample sizes (n $\leq$ 30), RMOM estimates provided the lowest RMSE for quantiles T $\geq$ 100 for all population cases. RMOM estimates were also superior in terms of RMSE for quantiles down to T = 50 under smaller population skewness ($G_x \leq .5$). The RMSE of quantile estimates generally increased with increasing population skewness and increased from the smaller quantiles to the larger. However, the increase in RMSE for case 1 to case 4 was generally less than 100% for all methods. As in previous cases, RMLE and POME generally exhibited the smallest RMSE for large sample sizes($n \geq 100$) for small quantiles

$(T \leq 10)$ for all population cases. However, MMOM performed well in terms of RMSE across all sample sizes and population cases as well. In fact, MMOM was compatible with the best quantile estimators in all cases of skewness and sample sizes.

7.4.2.5 Robustness Evaluation: The relative robustness of different methods of parameter and quantile estimation showed the method which performed the best and the worst for the smallest and largest sample sizes generated (n = 10, n = 400) for each parameter and selected quantiles for all population cases considered. Interestingly, the performance of the parameter estimation techniques was relatively insensitive to population skewness. The relative performance of the techniques for case 4 was virtually identical to that for case 1. The results clearly illustrated the generally superior performance of RMLE, POME, and MMOM for parameter estimation. As expected, RMLE performed in a superior manner in terms of BIAS, SE, and RMSE for large sample sizes for all parameters. POME, MMOM and RMOM generally performed in a superior manner for small samples depending on the parameter being estimated.

In contrast to the parameter results, the performance of different methods in quantile estimation varied somewhat with increasing population skewness. The performances also varied with the quantile being estimated and with sample size. For cases of small population skew $(G \leq 1.0)$, RMOM was generally superior for estimation of all quantiles under all indices for small sizes. As expected, RMLE performed in a superior manner in all cases for large samples but generally performed poorly for small samples. As population skewness increased, the performance of RMOM for small samples deteriorated at a somewhat faster rate than that of MMOM and POME which became superior in terms of BIAS for estimation of larger quantiles$(T \geq 100)$ under larger population skewness $(G \geq 2.0)$. However, RMOM continued to perform well in terms of SE and RMSE for the larger quantiles in these cases. MMOM performed in a superior manner for the smaller quantiles$(T \leq 10)$ for these large skew cases. In summary, RMLE performed in a clearly superior manner for quantile estimation under large sample sizes for all cases, while RMOM and MMOM were superior for small samples.

7.4.2.6 Concluding Remarks: The Monte Carlo study revealed that no one method performed in a superior manner in terms of bias and RMSE of parameter and quantile estimation across all population cases. In general, RMLE and POME performed well in terms of bias and RMSE of both parameter and quantile estimation for large sample sizes in all cases. In terms of bias, RMOM generally performed well in cases of small sample sizes in both parameters and quantile estimation for populations which exhibited small skewness. RMOM also performed well in small sample size estimation of μ_y and σ_y and large quantiles $(T \geq 100)$ in terms of RMSE in small skew cases. However, RMOM estimation generally deteriorated for larger sample sizes and smaller quantiles. RMOM also generally tended to deteriorate with increasing population skewness more rapidly than other methods. The alternative method recommended by Cohen and Whitten (1980) (MMOM) did represent an improvement over RMOM in many cases. MMOM estimation did not deteriorate as rapidly as RMOM under increasing population skewness and sample sizes in terms of both bias and RMSE. Thus, MMOM appeared to be a more robust estimator than RMOM. However, the computational difficulties encountered in using this technique were apparently increased relative to RMOM. For some sample sizes, a significant number of samples were rejected for this method. These rejections may reduce the reliability of the results of the methods for which significant numbers of samples were rejected.

The other alternative recommended by Cohen and Whitten (1980) (MMLE) did not generally represent an improvement over RMLE. In addition, the results demonstrated that this method represented the worst computational difficulties in terms of convergence of any of the techniques compared. For the cases of small sample sizes under large population skewness

($G_x > .5$), MMOM and POME appeared to be the superior estimators in terms of both bias and RMSE. In fact, in terms of consistent performance in the largest number of cases of both sample size and population skewness, these two methods appeared to be the most robust of the techniques compared.

7.4.3 APPLICATION TO FIELD DATA

Singh and Singh (1987) applied MOM, MLE and POME to annual peak discharge data of six selected rivers. Pertinent characteristics of the data are given in Table 7.2. These data were selected on the basis of length, completeness, homogeneity and independence of record. Each gaging station had a record of more than 30 years. The parameters estimated by the three methods are summarized in Table 7.3. For two sample gaging stations, a comparison of observed and computed frequency curves is shown in Figures 7.1 and 7.2. The observed frequency curve was obtained by using the Gringorton plotting position formula.

Table 7.2 Pertinent data characteristics of six selected rivers.

River gaging station	Drainage area (sq.km)	Length of record (years)	Mean discharge (cu. m/s)	St. dev. (cu. m/s)	Skewness (C_s)	Kurtosis (K_s)
Comite River at Comite, LA	1,896	38	315.7	166.8	0.54	2.77
Amite River at Amite, LA	4,092	34	745.1	539.5	0.71	3.03
St. Mary River at Still Water, Nova Scotia	1,653	59	409.5	147.9	1.42	6.25
St. John River at Ninemile Bridge, ME	1,890	32	699.0	223.7	0.41	3.01
Allagash River at Allagash, ME	1,659	51	438.8	159.8	0.71	3.30
Fish River near Fort Kent, ME	890	53	241.1	71.4	0.43	3.22

Although the differences between parameter estimates were not large in any case, the parameter estimates obtained by the POME and MLE methods were most similar. Consequently, their corresponding frequency curves were also closer. POME did not require the use of the coefficient of skewness, whereas MOM did. In this way, the bias was reduced when POME was used to estimate the parameters of TPLN distribution.

To compare these three methods further, relative mean error (RME) and relative absolute error (RAE) were computed as given in Table 7.4. With a couple of exceptions, notably the RME values associated with the Comite River and Amite River, RME and RAE values were essentially the same for the various data sets. This implies that the three procedures, MOM, MLE, and POME, provided equivalent results.

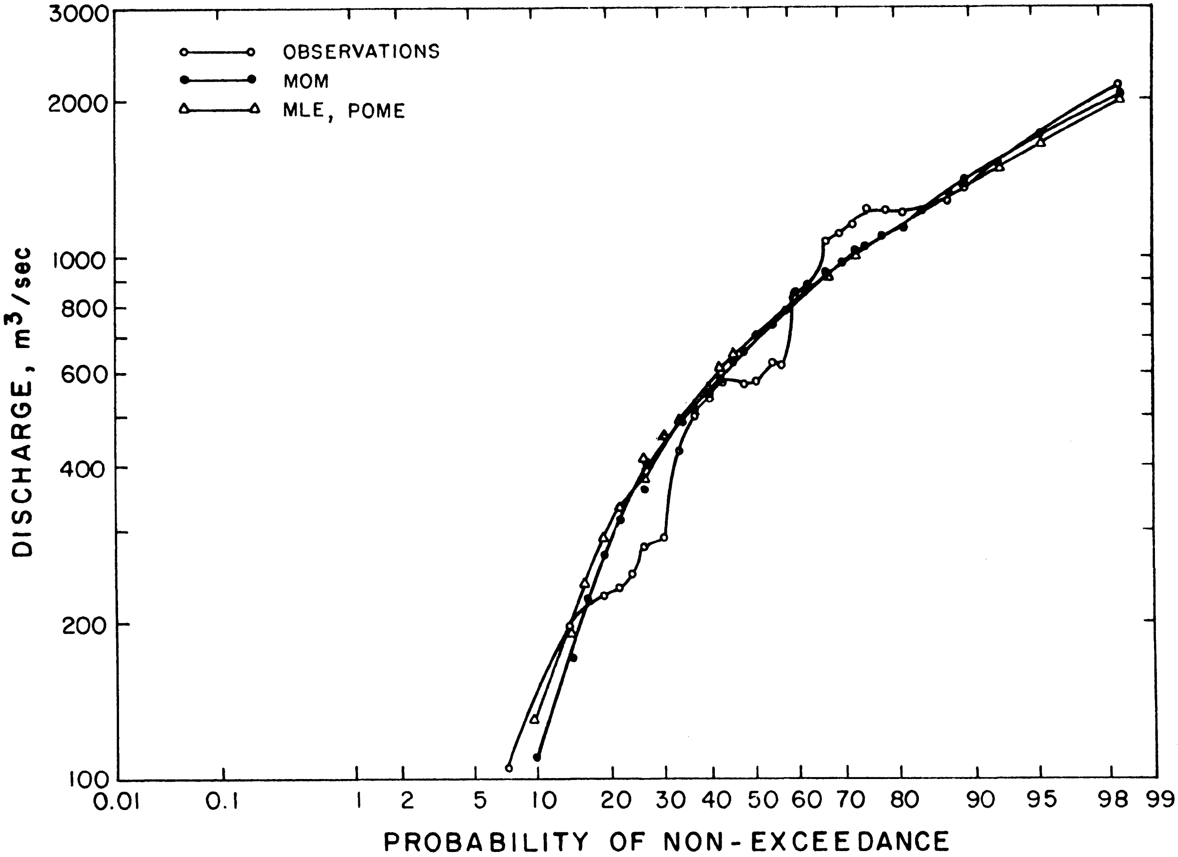

Figure 7.1 Comparison of observed and computed frequency curves for the Amite River basin at Amite, Louisiana.

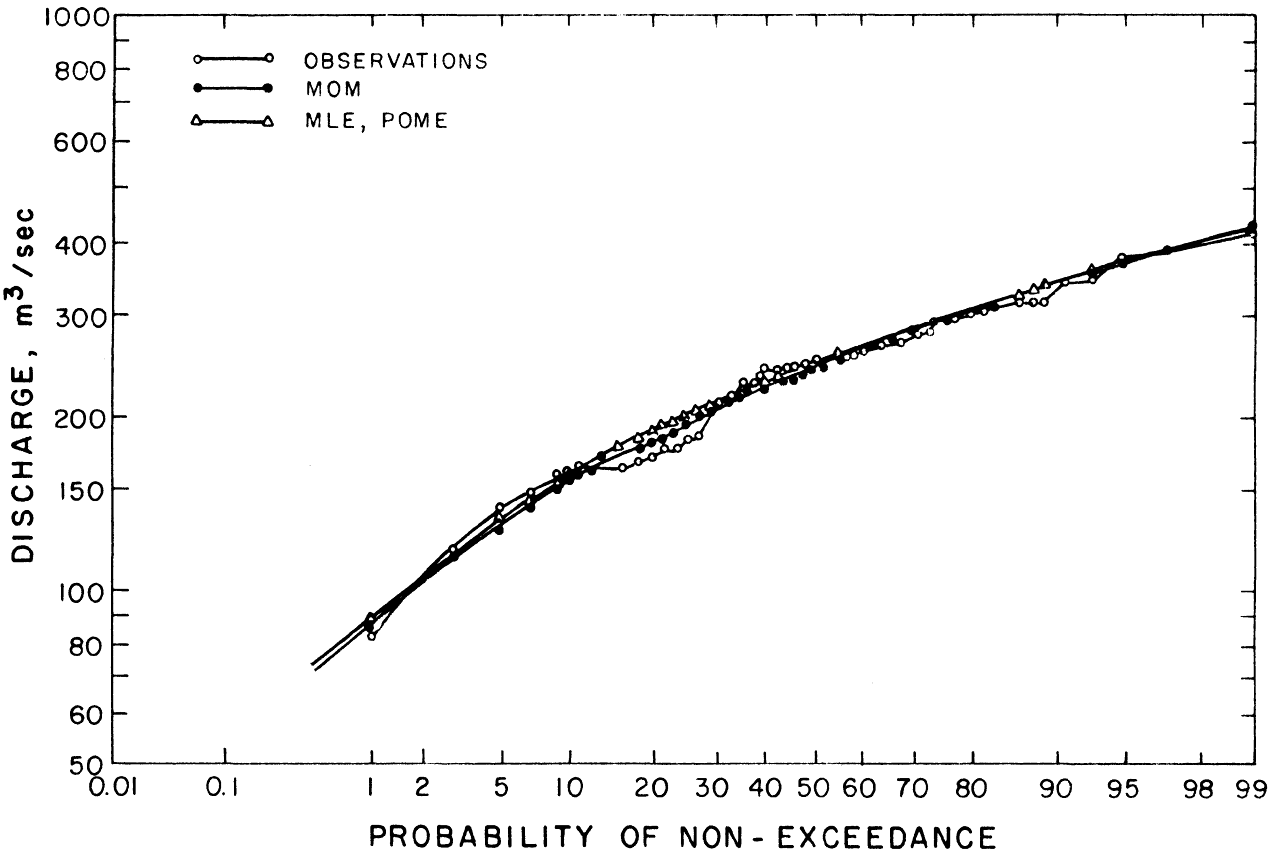

Figure 7.2 Comparison of observed and computed frequency curves for the Fish River basin near Fort Kent, Maine.

MLE, and POME, provided equivalent results.

Table 7.3 Parameter estimates by MOM, MLE and POME methods for six selected rivers ($b = \bar{y}, c^2 = s_y^2$)

River	MOM			MLE			POME		
Gaging Station	a	b	c^2	a	b	c^2	a	b	c^2
Comite River at Comite, LA	-692.1	6.90	0.027	-692.0	6.90	0.0256	-692.1	6.90	0.0256
Amite River at Amite, LA	-6879.3	8.93	0.0023	-6879.3	8.93	0.0022	-6879.8	8.93	0.0023
St. Mary River at Still Water, Nova Scotia	60.05	5.97	0.1648	60.05	5.78	0.1587	60.05	5.78	0.1591
St. John River at Ninemille Bridge, ME	-1123.8	7.50	0.0149	-1123.8	7.50	0.0143	-1123.8	7.50	0.0143
Allagash River at Allagash, ME	-294.3	6.57	0.0464	-294.32	6.57	0.0442	-294.3	6.57	0.0446
Fish River near Fort Kent, ME	-513.61	6.62	0.0089	-513.63	6.62	0.00872	-514.5	6.62	0.0090

Table 7.4 Relative mean error and relative absolute error by MOM, MLE and POME methods for six selected rivers.

River	RME			RAE		
Gaging station	MOM	MLE	POME	MOM	MLE	POME
Comite River at Comite, LA	2.82	2.35	2.35	7.40	7.54	7.54
Amite River at Amite, LA	7.29	6.60	7.19	13.09	13.32	13.12
St. Mary River at Still Water, Nova Scotia	0.12	0.12	0.12	2.68	2.59	2.59
St. John River at Ninemile Bridge, ME	0.23	0.22	0.22	3.67	2.54	3.56
Allagash River at Allagash, ME	0.18	0.18	0.18	3.20	3.20	3.20
Fish River near Fort Kent, ME	0.18	0.18	0.18	3.20	3.20	3.20

References

Aitchison, J. And Brown, J.A.C., 1957. *The lognormal Distribution with Special Reference to its Uses in Economics.* Cambridge University Press, London.

Bates, C.L., Lettenmaier, D.P. and Burges, S.J., 1974. Tables of parameters for the three parameter log normal probability distribution. *Technical Report* No. **41**, C.W. Harris Hydraulics Laboratory, Department of Civil Engineering, University of Washington, Seattle, Washington.

Burges, S.J. and Hoshi, K., 1978. Approximation of a normal distribution by a three-parameter lognormal distribution. *Water Resources Research*, Vol. **14**, No. 4, pp. 620-622.

Burges, S.J., Lettenmaier, D.P. and Bates, C.L., 1975. Properties of the three-parameter log normal probability distribution. *Water Resources Research*, Vol. **11**, No. 2, pp. 229-235.

Charbeneau, R.J., 1978. Comparison of the two- and three-parameetr log-normal distributions used in streamflow synthesis. *Water Resources Research*, Vol. **14**, No. 1, pp. 149-150.

Chow, V.T., 1954. The log probability law an dits engineering applications. *Proceedings, American Society of Civil Engineers*, Vol. **80**, No. 5, pp. 1-25.

Cohen, A.C. and Whitten, B.J., 1980. Estimation in the three parameter lognormal distribution. *Journal of American Statistical Association*, Vol. **70**, pp. 370-399.

Fiering, M.B. anf Jackson, B.B., 1971. *Synthetic Streamflows.* Water Resources Monograph No.**1**, American Geophysical Union, Washington, D.C.

Hoshi, K., Stedinger, J.R., and Burges, S.J., 1984. Estimation of lo g-normal quantiles: Monte Carlo results and first-order approximations. *Journal of Hydrology*, Vol. **71**, pp. 1-30.

Hosking, J.R.M., 1990. L-moments: Analysis and estimation of distributions using linear combinations of order statistics. *Journal of Royal Statistical Society*, Series B, Vol. **52**, No. 1, pp. 105-124.

Johnson, W.L. and Kotz, S., 1970. *Distributions in Statistics: Continuous Univariate Distributions* **1**. Houghton-Mifflin, Boston, Massachsetts.

Kite, G.W., 1978. *Frequency and Risk Analyses in Hydrology.* Water Resources Publications, Fort Collins, Colorado.

Kosugi, K., 1994. Three-parameter lognormal distribution model for soil water retention. *Water Resources Research*, Vol. **30**, No. 4, pp. 891-901.

Kuczera, G., 1982a. Combining site-specific and regional information: An empirical Bayes approach. *Water Resources Research*, Vol. **18**, pp. 306-314.

Kuczera, G., 1982b. Robust flood frequency models. *Water Resources Research*, Vol. **18,** pp.

315-324.

Matalas, N.C., 1967. Mathematical assessment of synthetic hydrology. *Water Resources Research*, Vol. **3,** No. 4, pp. 937-946.

Sangal, B.P. and Biswas, A.K., 1970. Three parameter lognormal distribution and its applications in hydrology. *Water Resources Research*, Vol. **6**, No. 2, pp. 505-515.

Singh, V.P. and Singh, K., 1987. Parameter estimation for TPLN distribution for flood frequency analysis. *Water Resources Bulletin*, Vol. **23**, No. 6, pp. 1185-1191.

Singh, V.P., Cruise, J.F. and Ma, M., 1990. A comparative evaluation of the estimators of the three-parameter lognormal distribution by Monte Carlo simulation. *Computational Statistics and data Analysis*, Vol. **10**, pp. 71-85.

Singh, V.P. and Rajagopal, A.K., 1986. A new method of parameter estimation for hydrologic frequency analysis. *Hydrological Science and Technology*, Vol. **2**, No. 3, pp. 33-40.

Singh, V.P., Rajagopal, A.K. and Singh, K., 1986. Derivation of some frequency distributions using the principle of maximum entropy. *Advances in Water Resources*, Vol. **9**, pp. 91-106.

Singh, V.P., Singh, K. and Rajagopal, A.K., 1985. Application of the principle of maximum entropy (POME) to hydrologic frequency analysis. *Completion Report* **06**, 144 p., Louisiana Water Resources Research Institute, Louisiana State University, Baton Rouge, Baton Rouge, Louisiana.

Slade, J.J., 1936. An asymmetric probability function. *Transactions, American Society of Civil Engineers*, Vol. **101**, pp. 481-490.

Snyder, W.M. and Wallace, J.R., 1974. Fitting a three-parameter lognormal distribution by least squares. *Nordic Hydrology*, Vol. **5**, pp. 129-145.

Stedinger, J.R., 1980. Fitting log nomal distributions to hydrologic data. *Water Resources Research*, Vol. **16**, No. 3, pp. 481-490.

Stevens, E.W., 1992. Uncertainty of extreme flood estimates incorporating historical flood data. *Water Resources Research*, Vol. **28**, No. 6, pp. 1057-1068.

CHAPTER 8

EXTREME VALUE TYPE 1 DISTRIBUTION

The extreme value type 1 (EV1) distribution is one of the most popularly used distributions for frequency analysis of extreme values of meteorologic or climatic and hydrologic variables, such as floods, rainfall, droughts, etc. This distribution was derived by Fisher and Tippett (1928) as a limiting form of the frequency distribution of the largest or smallest of a sample. In a series of papers Gumbel (1941a, b, 1942a, b, 1948) derived the EV1 distribution for flood flows and applied it to frequency analysis of floods, droughts, and meteorological data. Gumbel (1958) published a treatise on statistics of extremes, which contains a comprehensive treatment of EV1 distribution. Bardsley and Manly (1987) examined the transformations under which non-Gumbel distributions of annual flood flow maxima would converge to the Gumbel distribution. Smith (1986) presented a family of statistical distributions and estimators based on a fixed number (greater than one) of the largest annual events. Jenkinson (1955) found a general solution of the function equation derived by Fisher and Tippett (1928) for extreme values and showed that the Gumbel distribution was a special case of the general solution. Singh et al. (1986) derived this distribution using the principle of maximum entropy. Al-Mashidini et al. (1978) presented a simplified form of EV1 distribution for flood estimation.

Hershfield and Kohler (1960) made an empirical appraisal of the Gumbel distribution using thousands of station years of rainfall data. Their results showed that this was an acceptable distribution for predicting the probability of occurrence of extreme values of rainfall. Stol (1971) applied it to analyze daily, monthly and annual rainfall data. Lambert and Li (1994) applied EV1 distribution to evaluate risk of extreme events for strictly monotonically increasing univariate-loss functions in water resources, where uncertainties about the distributions of precipitation, runoff, and wind and wave magnitude are widely found. Coulson (1966) prepared tables for computing and plotting flood frequency curves using EV1 distribution. Weiss (1955) developed a nomogram for determining values for various return periods. Majumdar and Sawhney (1965) compared EV1 distribution with lognormal and Foster's type 3 distributions for extreme values. The distribution yielded good estimates for return periods up to 1000 years when the coefficient of variation was less than 0.5. Shen et al. (1980) compared EV1 and log-Pearson type 3 distributions. They noted that EV1 could lead to very large underestimates of extreme events in those cases where EV 2 distribution was appropriate. Reich (1970) analyzed annual flood peaks from 26 Pennsylvanian watersheds (smaller than 200 square miles) using EV1, log EV1 and log-Pearson type 3 distributions. Consistent overestimates of long return period extremes resulted from log EV1 distribution.

The EV1 distribution is a two parameter distribution. The two parameters have been estimated using a number of methods. Phien (1987) reviewed four methods of parameter

V. P. Singh, *Entropy-Based Parameter Estimation in Hydrology*

estimation, including methods of moments (MOM), maximum likelihood estimation (MLE), principle of maximum entropy (POME), and probability weighted moments (PWM). He noted that PWM was the best in terms of bias and MLE the best in terms of the root mean square error and efficiency. By all criteria, POME was the second best and followed MLE more closely than PWM. Fiorentino and Gabriele (1984) proposed modifications to MLE for reducing bias in parameter and quantile estimates. Lettenmaier and Burges (1982) showed that the parameter and quantile estimates were much improved if Gumbel's m was set to infinity, rather than the sample length. Using bias and efficiency, Lowery and Nash (1970) compared four parameter estimation methods: moments, regression, maximum likelihood, and Gumbel's fitting method. They found that next to the method of maximum likelihood the method of moments was the most accurate. Jowitt (1979) estimated the EV1 parameters using the principle of maximum entropy. Houghton (1978) proposed the method of incomplete means for parameter estimation.

Landwehr et al. (1979) compared PWM with MLE and MOM. While PWM produced unbiased estimates of Gumbel parameters, MLE produced minimum variance estimates. In general, MLE was the most efficient of the three methods. Raynal and Salas (1986) analyzed six estimation methods for EV1 and its generalized version: MOM, MLE, PWM, least squares (MOLS), and mode and interquartile range method (MIR). On the whole, PWM and MLE were judged to be the best methods. Jain and Singh (1987) compared seven methods of parameters: MOM, MLE, POME, MOLS, PWM, mixed moments (MIX), and incomplete means (MIM). They found that MLE, POME, MOM, and PWM were adequate for general use, although POME or MLE would be preferable. Arora and Singh (1987) made a statistical comparison of EV1 estimators using Monte Carlo experimentation: MOM, MLE, PWM, POME, MIX, MOLS and MIM. In addition, they made a bias correction to the MOM-quantile estimator. MLE provided the most efficient quantile estimates followed closely by POME. For small samples, PWM and MOM performed comparably in efficiency of estimating quantiles. PWM resulted in nearly unbiased quantile estimates. Serial correlation in samples resulted in deterioration of the performance of all estimators. Fill and Stedinger (1995) analyzed the power of two L-moment and probability plot correlation coefficient goodness-of-fit tests for the EV1 distribution and the impact of autocorrelation. They recommended use of unbiased L moment estimators for goodness-of-fit tests and distribution selection, as well as parameter estimation. Phien and Arbhabhirama (1980) evaluated the effect of the selection of plotting position formulae and class division schemes on goodness-of-fit tests for EV1 distribution using annual flood and annual maximum daily rainfall data. They found that the plotting position formulae had a minor influence on the tests while the class division had a pronounced effect on the chi-square test. They recommended MLE for parameter estimation.

A random variable X is said to have an extreme value type 1 (EV1) or Gumbel distribution if its probability density function (pdf) is given by

$$f(x) = a \exp[-a(x-b) - e^{-a(x-b)}] \qquad (8.1a)$$

where $a > 0$ and $-\infty < b < x$ are parameters. Parameter a ia a concentration parameter and parameter b is a measure of central tendency. Thus, the EV1 distribution is a two-parameter distribution. Its cumulative distribution function (cdf) can be expressed as

$$F(x) = \exp[-e^{-a(x-b)}] \qquad (8.1b)$$

Let y be the reduced variate defined as

$$y = a(x - b) \tag{8.2a}$$

Then equation (8.1a) an be expressed in terms of y as

$$f(x) = a \exp[-y - \exp(-y)] \tag{8.2b}$$

and equation (8.1b) as

$$F(y) = \exp[-\exp(-y)] \tag{8.2c}$$

The EVI distribution has a constant coefficient of skew of approximately 1.14 and a constant coefficient of kurtosis of approximately 5.40. For a return period T, the reduced variate y_T can be expressed from equation (8.2c) as

$$y_T = -\ln[-\ln\frac{T-1}{T}] \tag{8.2d}$$

where T is defined as 1/[1-F].

8.1 Ordinary Entropy Method

8.1.1 SPECIFICATION OF CONSTRAINTS

Taking logarithm of equation (8.1) to the base 'e', one gets

$$\ln f(x) = \ln a - a(x-b) - \exp[-a(x-b)] \tag{8.3}$$

Multiplying equation (8.3) by [- f(x)] and integrating from $-\infty$ to ∞ yield

$$-\int_{-\infty}^{\infty} f(x) \ln f(x)dx = -[\ln a + ab] \int_{-\infty}^{\infty} f(x)dx + a \int_{-\infty}^{\infty} x\, f(x)dx$$

$$+ \exp(ab) \int_{-\infty}^{\infty} \exp(-ax)\, f(x)dx \tag{8.4}$$

From equation (8.4), the constraints appropriate for equation (8.1a) can be written (Singh et al., 1985, 1986) as

$$\int_{-\infty}^{\infty} f(x)dx = 1 \tag{8.5}$$

$$\int_{-\infty}^{\infty} x\, f(x)dx = \bar{x} \tag{8.6}$$

$$\int_{-\infty}^{\infty} \exp(-ax)\, f(x)dx = E[\exp(-ax)] \tag{8.7}$$

Equation (8.5) can be verified as

$$\int_{-\infty}^{\infty} f(x)dx = a \int_{-\infty}^{\infty} \exp[-a(x-b) - e^{-a(x-b)}]dx \tag{8.8}$$

Let a(x-b) = y. Then dy = a dx. Hence,

$$\int_{-\infty}^{\infty} f(x)dx = a \int_{-\infty}^{\infty} \exp[-y - e^{-y}] \frac{dy}{a} = \int_{-\infty}^{\infty} \exp(-y) \exp(-e^{-y}) dy \tag{8.9}$$

Let z = exp (-y). Then

$$\frac{dz}{dy} = -\exp(-y); \; dy = -\frac{dz}{z} \tag{8.10}$$

$$\int_{-\infty}^{\infty} f(x)dx = -\int_{\infty}^{0} z \exp(-z) \frac{dz}{z} = \int_{\infty}^{0} z \exp(-z) dz = 1$$

8.1.2 CONSTRUCTION OF ZEROTH LAGRANGE MULTIPLIER

The least-biased pdf based on POME and consistent with equations (8.5) - (8.7) takes the form:

$$f(x) = \exp[-\lambda_0 - \lambda_1 x - \lambda_2 e^{-ax}] \tag{8.11}$$

where λ_0, λ_1, and λ_2 are Lagrange multipliers.

Inserting equation (8.11) into equation (8.5) gives

$$\int_{-\infty}^{\infty} f(x)dx = \int_{-\infty}^{0} \exp[-\lambda_0 - \lambda_1 x - \lambda_2 e^{-ax}] dx = 1 \tag{8.12}$$

From equation (8.12) the partition function can be written as

$$\exp(\lambda_0) = \int_{-\infty}^{0} \exp[-\lambda_1 x - \lambda_2 e^{-ax}] dx \tag{8.13}$$

Let $\lambda_2 e^{-ax} = y$. Then

$$\begin{aligned} & e^{-ax} = \frac{y}{\lambda_2}; -ax = ln(\frac{y}{\lambda_2}); x = -\frac{1}{a} ln \frac{y}{\lambda_2} \\ & x = -\frac{1}{a}[ln y - ln \lambda_2] = -\frac{ln y}{a} + \frac{ln \lambda_2}{a} \\ & \frac{dy}{dx} = \lambda_2(-a)e^{-ax}; -\lambda_1 x = \frac{\lambda_1 ln y}{a} - \frac{\lambda_1 ln \lambda_2}{a} \\ & dx = -\frac{dy}{\lambda_2 a e^{-ax}} = -\frac{dy}{ay} \end{aligned} \tag{8.14}$$

Substitution of the above quantities in equation (8.13) yields

$$exp(\lambda_0) = \int_{-\infty}^{\infty} exp[-\lambda_1 x] exp[-\lambda_2 e^{-ax}] dx$$
$$= -\int_{-\infty}^{0} exp[ln(y)^{\lambda_1/a} + ln(\lambda_2)^{-(\lambda_1/a)}] exp[-y] \frac{dy}{ay}$$
$$= \frac{(\lambda_2)^{-(\lambda_1/a)}}{a} \int_0^{\infty} y^{(\lambda_1/a)-1} exp(-y) dy = \frac{(\lambda_2)^{-(\lambda_1/a)}}{a} \Gamma(\frac{\lambda_1}{a}) \tag{8.15}$$

The zeroth Lagrange multiplier λ_0 is given as

$$\lambda_0 = -ln\, a - \frac{\lambda_1}{a} ln\, \lambda_2 + ln\, \Gamma(\frac{\lambda_1}{a}) \tag{8.16}$$

The zeroth Lagrange multiplier is also obtained from equation (8.13) as

$$\lambda_0 = \ln \int_{-\infty}^{\infty} \exp[-\lambda_1 x - \lambda_2 e^{-ax}] dx \tag{8.17}$$

8.1.3 RELATION BETWEEN LAGRANGE MULTIPLIERS AND CONSTRAINTS

Differentiating equation (8.16) with respect to λ_1 and λ_2, respectively, one gets

$$\frac{\partial \lambda_0}{\partial \lambda_1} = -\frac{ln\, \lambda_2}{a} + \frac{\partial}{\partial \lambda_1} ln\, \Gamma(\frac{\lambda_1}{a}) \tag{8.18}$$

$$\frac{\partial \lambda_0}{\partial \lambda_2} = -\frac{\lambda_1}{a \lambda_2} \tag{8.19}$$

Also, differentiation of equation (8.17) with respect to λ_1 yields

$$\frac{\partial \lambda_0}{\partial \lambda_1} = -\frac{\int_{-\infty}^{\infty} x\, exp[-\lambda_1 x - \lambda_2 e^{-ax}] dx}{\int_{-\infty}^{\infty} exp[-\lambda_1 x - \lambda_2 e^{-a}] dx}$$
$$= -\int_{-\infty}^{\infty} x f(x) dx = -\bar{x} \tag{8.20}$$

Similarly, differentiation of equation (8.17) with respect to λ_2 yields

$$\frac{\partial \lambda_0}{\partial \lambda_2} = - \frac{\int_{-\infty}^{\infty} e^{-ax} exp[-\lambda_1 x - \lambda_2 e^{-ax}]dx}{\int_{-\infty}^{\infty} exp[-\lambda_1 x - \lambda_2 e^{-a}]dx} \tag{8.21}$$
$$= - \int_{-\infty}^{\infty} e^{-ax} f(x) dx = - E[e^{-ax}]$$

Equating equations (8.18) and (8.20), as well as equations (8.19) and (8.21), one obtains

$$\bar{x} = \frac{ln\lambda_2}{a} - \frac{\partial}{\partial \lambda_1} ln\, \Gamma(\frac{\lambda_1}{a}) \tag{8.22}$$

$$E[e^{-ax}] = \frac{\lambda_1}{a\lambda_2} \tag{8.23}$$

8.1.4 RELATION BETWEEN PARAMETERS AND LAGRANGE MULTIPLIERS

Substitution of equation (8.16) in equation (8.11) gives

$$f(x) = exp[ln a + \frac{\lambda_1}{a} ln\lambda_2 - ln\Gamma(\frac{\lambda_1}{a}) - \lambda_1 x - \lambda_2 e^{ax}$$
$$= \frac{a(\lambda_2)^{\lambda_1/a}}{\Gamma(\frac{\lambda_1}{a})} exp[-\lambda_1 x - \lambda_2 e^{-ax}] \tag{8.24}$$

A comparison of equation (8.24) with equation (8.1a) shows that

$$\lambda_1 = a \tag{8.25}$$

$$\lambda_2 = e^{ab} \tag{8.26}$$

8.1.5 RELATION BETWEEN PARAMETERS AND CONSTRAINTS

The EV1 distribution has two parameters a and b which are related to the Lagrange multipliers by equations (8.25) and (8.26) which, in turn, are related to the constraints by equations (8.22) and (8.23). Eliminating the Lagrange multipliers between these two sets of equations yields relations between parameters and constraints,

$$\bar{x} = b - \frac{0.5772}{a} \tag{8.27}$$

$$E\ [e^{-(x-b)a}] = 1 \tag{8.28}$$

8.1.6 DISTRIBUTION ENTROPY

The entropy of the EV1 distribution can be expressed as follows:

$$\begin{aligned} I(x) &= -\ ln\,a\int_{-\infty}^{\infty} f(x)\,dx + a\int_{-\infty}^{\infty} x f(x)\,dx \\ &\quad - ab\int_{-\infty}^{\infty} f(x)\,dx + \int_{-\infty}^{\infty} exp[-a(x-b)] f(x)\,dx \\ &= -\ ln\,a + a\bar{x} - ab + \int_{-\infty}^{\infty} \exp[-a(x-b)]\,dx \end{aligned} \tag{8.29}$$

$$W = \int_{-\infty}^{\infty} exp[-a(x-b)]\,a\,exp[-a(x-b) - e^{-a(x-b)}]\,dx \tag{8.30}$$

Let a(x-b) = y. Then dy = a dx. Hence, equation (8.30) becomes

$$\begin{aligned} W &= \int_{-\infty}^{\infty} exp[-y]\ a\ exp[-y-e^{-y}]\frac{dy}{a} \\ &= \int_{-\infty}^{\infty} exp[-2y-e^{-y}]\,dy = \int_{-\infty}^{\infty} exp[-2y]\,exp[-e^{-y}]\,dy \end{aligned} \tag{8.31}$$

Let $e^{-y} = z$. Then,

$$z^2 = e^{-2y} = \frac{dz}{dy} = -\ e^{-y} = -\ z$$

Therefore, $dy = -\ dz/z$.

$$\begin{aligned} W &= -\int_{\infty}^{0} z^2 e^{-z}\frac{dx}{z} = \int_{0}^{\infty} z e^{-z}\,dx \\ &= \int_{0}^{\infty} z^{2-1} e^{-z}\,dx = \Gamma(2) = 1 \end{aligned} \tag{8.32}$$

8.2 Parameter - Space Expansion Method

8.2.1 SPECIFICATION OF CONSTRAINTS

$$I(x) = -\ ln a + a\bar{x} - ab + 1$$
$$= ln\frac{1}{a} + a\bar{x} - ab + lne = ln(\frac{e}{a}) + a(\bar{x} - b) \tag{8.33}$$

For this method, the constraints, following Singh and Rajagopal (1986), are given by equation (8.5) and

$$\int_{-\infty}^{\infty} a(x-b)f(x)dx = E\ [a(x-b)] \tag{8.34}$$

$$\int_{-\infty}^{\infty} e^{-a(x-b)} f(x)\,dx = E\ [e^{-a(x-b)}] \tag{8.35}$$

8.2.2 DERIVATION OF ENTROPY FUNCTION

The pdf corresponding to the principle of maximum entropy (POME) and consistent with equations (8.5), (8.34), and (8.35) takes the form

$$f(x) = exp[-\lambda_0 - \lambda_1 a(x-b) - \lambda_2 e^{-a(x-b)}] \tag{8.36}$$

where λ_0, λ_1, and λ_2 are Lagrange multipliers. Insertion of equation (8.36) into equation (8.5) yields

$$\exp(\lambda_0) = \int_{-\infty}^{\infty} \exp\{-\lambda_1 a(x-b) - \lambda_2 \exp[-a(x-b)]\}\,dx = \frac{\Gamma(\lambda_1)}{a\lambda_2^{\lambda_1}} \tag{8.37}$$

The zeroth Lagrange multiplier is given by equation (8.37) as

$$\lambda_0 = -ln a - \lambda_1 ln\lambda_2 + ln\Gamma(\lambda_1) \tag{8.38}$$

Also from equation (8.37), the zeroth Lagrange multiplier is given as

$$\lambda_0 = ln \int_{-\infty}^{\infty} exp[-\lambda_1 a(x-b) - \lambda_2 e^{-a(x-b)}] dx \tag{8.39}$$

Introduction of equation (8.38) in equation (8.36) gives

$$f(x) = \frac{a\lambda_2^{\lambda_1}}{\Gamma(\lambda_1)} exp[-\lambda_1 a(x-b) - \lambda_2 e^{a(x-b)}] \tag{8.40}$$

A comparison of equation (8.40) with equation (8.1a) shows $\lambda_1 = \lambda_2 = 1$.

Taking minus logarithm of equation (8.40) produces

$$\begin{aligned} -\ln f(x) = & -ln\ a - \lambda_1\ ln\ \lambda_2 + ln\ \Gamma(\lambda_1) \\ & + \lambda_1 a(x-b) + \lambda_2\ e^{-a(x-b)} \end{aligned} \tag{8.41}$$

The entropy function of the EV1 distribution is obtained by multiplying equation (8.41) by f(x) and integrating the product from $-\infty$ to $+\infty$:

$$\begin{aligned} I(f) = & -ln\ a - \lambda_1\ ln\ \lambda_2 + ln\ \Gamma(\lambda_1) \\ & + \lambda_1\ E[a(x-b)] + \lambda_2\ E[e^{-a(x-b)}] \end{aligned} \tag{8.42}$$

8.2.3 RELATION BETWEEN PARAMETERS AND CONSTRAINTS

Taking partial derivatives of equation (8.42) with respect to λ_1, λ_2, a, and b, separately, and equating each derivative to zero results in

$$\frac{\partial I}{\partial \lambda_1} = 0 = -ln\ \lambda_2 + \Psi(\lambda_1) + E[a(x-b)] \tag{8.43}$$

$$\frac{\partial I}{\partial \lambda_2} = 0 = -\frac{\lambda_1}{\lambda_2} + E[e^{-a(x-b)}] \tag{8.44}$$

$$\begin{aligned} \frac{\partial I}{\partial a} = 0 = & -\frac{1}{a} + \lambda_1\ E[(x-b)] \\ & - \lambda_2\ E[e^{a(x-b)}(x-b) \end{aligned} \tag{8.45}$$

$$\frac{\partial I}{\partial b} = 0 = -\lambda_1 a + E[e^{-a(x-b)} a] \tag{8.46}$$

where Ψ is the digamma function. Simplification of equation (8.43) to (8.46) leads to

$$E\,[a(x-b)] = -\,\Psi(1) \tag{8.47}$$

$$E\,[e^{-a(x-b)}] = 1 \tag{8.48}$$

$$E\,[a(x-b)] - E\,[a(x-b)e^{a(x-b)}] = 1 \tag{8.49}$$

$$E\,[e^{-a(x-b)}] = 1 \tag{8.50}$$

Equations (8.48) and (8.50) are the same and equation (8.49) is an identity. Thus, the parameter estimation equations are equations (8.47) and (8.48).

8.3 Other Methods of Parameter Estimation

The two parameters a and b in equation (8.1) and (8.2) can be estimated by using any one of the standard methods. Some of these methods are briefly outlned here. In practice, these parameters are estimated from sample statistics. These then lead to equations for fitting the EV1 distribution to the sample data.

8.3.1 METHOD OF MOMENTS

The method of moments (MOM) is one of the most popular methods for estimating parameters a and b (Lowery and Nash, 1970; Landwehr, et al., 1979). Since there are two parameters, only the first two moments of the distribution are needed. Gumbel (1958) has shown that the mean μ_y and standard deviation σ_y of y can be expressed as

$$\mu_y = \gamma \;\; (Euler's\ constant = 0.5772\ ) \tag{8.51a}$$

and

$$\sigma_y = \frac{\pi}{\sqrt{6}} \tag{8.51b}$$

If μ_x and σ_x are the mean and standard deviation of X, then from definition of Y we get:

$$\mu_x = b + \frac{\gamma}{a} \tag{8.52a}$$

$$\sigma_x = \frac{\pi}{a\sqrt{6}} \tag{8.52b}$$

or

$$b = \mu_x - \frac{\gamma \sigma_x \sqrt{6}}{\pi} \tag{8.53a}$$

$$a = \frac{\pi}{\sigma_x \sqrt{6}} \tag{8.53b}$$

The mean and standard deviation of X, μ_x and σ_x, are replaced by the sample mean and standard deviation (almost unbiased), $\bar{x}$ and S_x, which for a sample of size n are defined as

$$\bar{x} = \sum_{i=1}^{n} x_i / n \tag{8.54a}$$

$$S_x = [\Sigma(x_i - \bar{x})^2 / (n-1)]^{0.5} \tag{8.54b}$$

Then equations (8.53a) and (8.53b) become

$$a = 1.28255 / S_x \tag{8.55a}$$

$$b = \bar{x} - 0.450041\, S_x \tag{8.55b}$$

Thus, the fitting equation for the EV1 distribution is

$$x_T = \bar{x} + K(T)\, S_x \tag{8.56a}$$

$$K(T) = -\,[0.45 + 0.7797 \ln(-\ln(1 - \frac{1}{T}))] \tag{8.56b}$$

where K(T) is the frequency factor (Chow, 1953) and T is the recurrence interval. Clearly, K(T) is analogous to the reduced variate which is a function of T only. Thus, equation (8.56b) can also be expressed as

$$K = -\frac{\sqrt{6}}{\pi}(\gamma - y_T) \tag{8.56c}$$

where y_T is the value of y corresponding to T.

8.3.2 METHOD OF MAXIMUM LIKELIHOOD ESTIMATION

For the method of maximum likelihood estimation (MLE), the parameters are estimated such that they maximize the likelihood of obtaining the sample drawn from the EV1 population. Kimball

(1949) derived the MLE estimators as

$$\frac{1}{a} = \bar{x} - \frac{\sum_{i=1}^{n} x_i \exp(-ax_i)}{\sum_{i=1}^{n} \exp(-ax_i)} \tag{8.57}$$

$$b = \frac{1}{a} \ln \left[\frac{n}{\sum_{i=1}^{n} \exp(ax_i)} \right] \tag{8.58}$$

The fitting equation can then be obtained by inserting a and b through equation (8.2d).

8.3.3 METHOD OF LEAST SQUARES

The method of least squares (LEAS) estimates parameters a and b by considering equation (8.56a) as a linear regression of x on K where μ_x and σ_x are treated simply as parameters of the linear relationship. This requires an a priori estimate of T (and hence of K from equation (8.56b)) which is usually done by arranging the data in descending or ascending order of magnitude, choosing an appropriate plotting position formula, and assigning a return period to each event accordingly. By using equation (8.1b), these parameters can be estimated as

$$a^{-1} = \frac{n \sum_{i=1}^{n} s_i x_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} z_i}{(\sum_{i=1}^{n} x_i)^2 - n \sum x_i^2} \tag{8.59}$$

$$b = \bar{x} + \frac{\sum_{i=1}^{n} z_i}{n a} \tag{8.60}$$

where $z_i = \ln[-\ln(F(x_i)]$ is obtained for each data point from the plotting position formula which defines the cumulative probability of non-exceedance for each data point x_i.

8.3.4 METHOD OF INCOMPLETE MEANS

The incomplete means method (ICM) uses means calculated over only parts of the data range. By arranging the sample in ascending order $x_1, x_2, \ldots, x_n$, first the sample mean $(\bar{x})$ is calculated. $\bar{x}$ is then used to divide the sample into disjointed sets. The mean of the upper set having values greater than $\bar{x}$ is calculated and called first incomplete mean $\bar{x}_1$. Similarly, the mean of all observations above $\bar{x}_1$ is calculated and is the second incomplete mean $\bar{x}_2$, and so on. For the EV1 distribution, the first two imcomplete means are

$$\bar{x}_1 = b - \frac{n}{a(n-n_i)} [J \ln J - \frac{J^2 \ln J}{2} + \frac{J^3 \ln J}{6} - \frac{J^4 \ln J}{24}$$

$$- J + \frac{J^2}{4} - \frac{J^3}{18} + \frac{J^4}{96}] \ , \ i = 1,2 \qquad (8.61)$$

The sum of terms containing ln J can be simplified as ln $J(1 - e^{-J})$. Therefore,

$$\bar{x}_i = b - \frac{n}{a(n-n_i)} [\ln J(1 - e^{-J}) - J + \frac{J^2}{4} - \frac{J^3}{18} + \frac{J^4}{96}] \qquad (8.62)$$

where J = ln (n/n_i), n is the size of the sample and n_i the number of observations corresponding to the lower limit of the range on which the incomplete mean is calculated. $\bar{x}_1$ and $\bar{x}_2$ are then used to calculate parameters a and b.

8.3.5 METHOD OF PROBABLITY WEIGHTED MOMENTS

The probability weighted moments (PWM) method requires expressing the distribution function in inverse form which for the EV1 distribution is

$$x = b - \frac{1}{a} \ln[-\ln F(x)] \qquad (8.63)$$

Following Landwehr, et al. (1979), the parameters a and b can be given as

$$b = \bar{x} - \frac{.5772}{a} \qquad (8.64)$$

$$\frac{1}{a} = \frac{\bar{x} - 2M_{101}}{\ln 2} \qquad (8.65)$$

where M_{101} is the first probability weighted moment defined as

$$M_{101} = \frac{1}{n(n-1)} \sum_{i=1}^{n} x_i (n-i) \qquad (8.66)$$

This method produces unbiased estimates of parameters a and b if the sample is drawn from a purely random population, and less bias for a serially correlated case than the MOM and MLE methods.

8.3.6 METHOD OF MIXED MOMENTS

The method of mixed moments (MIX) uses the first moment of the EV1 distribution and the first moment of its logarithmic version. The parameters a and b are given by

$$a = \frac{1.28255}{S_x} \qquad (8.67)$$

$$\exp(ab) = 1 + a\bar{x} + \frac{a^2}{2} [S_x^2 + \bar{x}^2] \qquad (8.68)$$

where S_x^2 is the variance of x. Equation (26) is the same as equation (8.53) and equation (8.68) is derived Jain and Singh (1986).

8.4 Evaluation of Parameter Estimation Methods Using Field Data

8.4.1 ANNUAL FLOOD DATA

Jain and Singh (1987) compared the aforementioned methods of parameter estimation using instantaneous annual flood values for fifty-five river basins with drainage areas ranging from 104 to 2,500 km^2. Their work is followed here. These data were selected on the basis of length, completeness, homogeneity, and independence of record. The gaging stations were chosen because the length of record was continuous for more than 30 years. Each of the fifty-five data sets was tested for homogeneity by using the Kruskal-Wallis test (Seigel, 1956) and the Mann-Whitney test (Mann and Whitney, 1947), as well as for independence by the Wald-Wolfowitz test (Wald and Wolfowitz, 1943) and the Anderson test (Anderson, 1941). In each case, the sample was found homogeneous and independent. Parameters a and b were estimated by each method for each data set. Observed and computed frequency curves for two sample gaging stations are plotted in Figures 8.1 - 8.3.

8.4.2 PERFORMANCE CRITERIA

Two criteria were used for comparing the seven methods of parameter estimation. These have also been used by Benson (1968) and by Bobee and Robitaille (1977). The first criterion is defined as the average of relative deviations between observed and computed flood discharges for the entire sample, with algebraic sign ignored.

$$D_a = \frac{1}{n}\sum \mid G(T) \mid {}^{*}100 \tag{8.69}$$

in which

$$G(T) = (x_0(T) - x_c(T))/x_0(T) \tag{8.70}$$

where x_0 and x_c are observed and computed flood values respectively for a given value of return period T.

The other criterion is the average of squared relative deviations:

$$D_r = \frac{1}{n}\sum (G(T))^2 * 100 \tag{8.71}$$

Statistics D_a and D_r are objective indexes of the goodness of fit of each method to sample data. The observed flood discharge corresponds to a return period which was computed by using the Gringorton plotting position formula (Adamowski, 1981),

$$T = (n + 0.12)/(m - 0.44) \tag{8.72}$$

in which m is the rank assigned to each data point in the sample with a value for one for the highest flow, two for the second highest flow and so on, with n for the lowest value.

These criteria were computed using each method of parameter estimation for each sample.

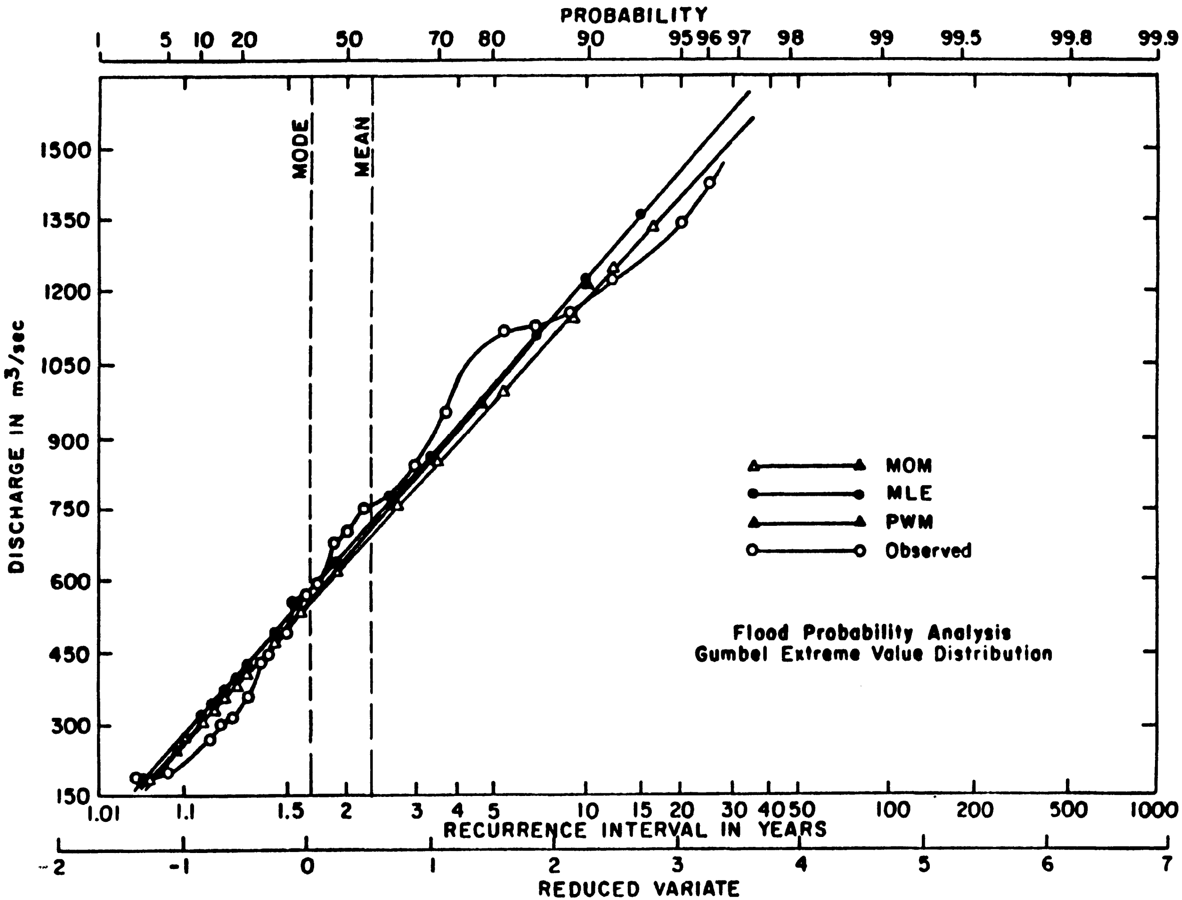

Figure 8.1 Comparison of observed and computed frequency curves for the Amite River basin near Magnolia, Louisiana.

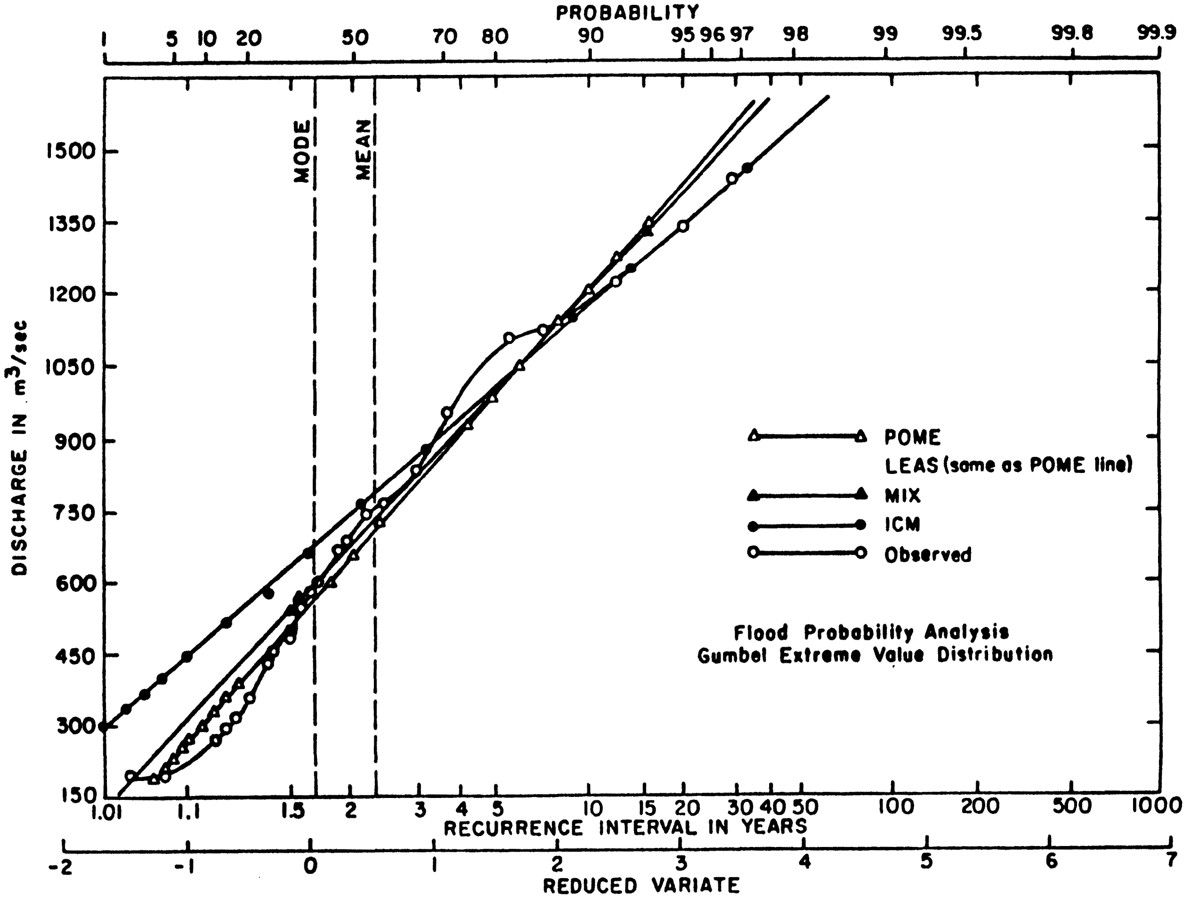

Figure 8.2 Comparison of observed and computed frequency curves for the Amite River basin near Magnolia, Louisiana.

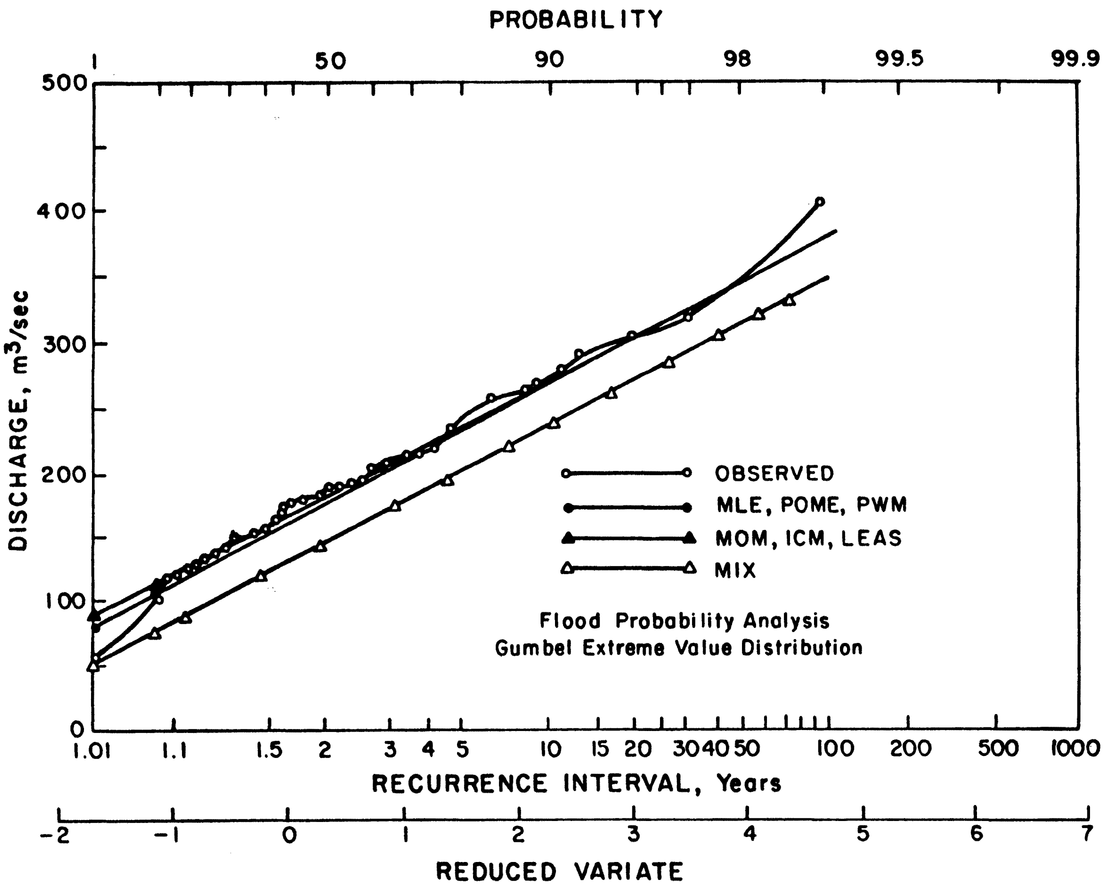

Figure 8.3 Comparison of observed and computed frequency curves for the Sebsticook River near Pittsfield, Maine.

The seven methods were ranked for all the gaging stations according to the values of D_a and D_r on a scale of 1 to 7, with one being the best method. Table 8.1 ranks the seven methods according to D_a. Clearly, the MLE method was the best of all, followed by the POME, PWM, MOM, MIX, and the ICM methods in descending order of the performance. The poor performance of the ICM is not surprising since the first moment is computed over parts of the sample only. The MIX method did not do well for the EV1 distribution because it is based on the first moment only. The MLE, POME and PWM methods were comparable. The MOM and LEAS methods worked well as long as the sample skewness was within one standard deviation of the distribution skewness.

Table 8.1 Ranking of the methods of parameter estimation for 55 gaging stations by absolute relative deviation D_a (on a scale of 1 to 7 with 1 being the best method).

Method	Number of Gaging Stations Receiving Ranking						
	1	2	3	4	5	6	7
MOM	6	5	4	28	12	0	0
MLE	38	5	4	5	1	2	0
PWM	4	9	28	9	3	0	2
POME	5	33	13	2	2	0	0
LEAS	2	2	4	8	18	21	0
MIX	0	1	1	2	17	19	15
ICM	0	0	1	1	2	13	38

Table 8.2 Ranking of the methods of parameter estimation for 55 gaging stations by mean square deviation D_r (on a scale of 1 to 7 with 1 being the best method)

Method	Number of gaging stations receiving ranking						
	1	2	3	4	5	6	7
MOM	2	2	6	15	30	0	0
MLE	43	7	4	0	0	1	0
PWM	4	5	32	8	4	0	2
POME	5	39	8	2	0	0	0
LEAS	1	1	3	10	8	31	1
MIX	0	0	2	19	11	11	11
ICM	0	0	0	1	12	12	40

The ranking of the seven methods according to D_r is given in Table 8.2. The MLE method is the best of all, followed by POME, PWM, MOM, MIX and ICM methods in descending order of their performance. Again, the previous conclusions hold. It should, however, be pointed out that the differences between MLE, POME, MOM and PWM were not too great and, therefore, these methods could be considered comparable for practical purposes.

8.5 Evaluation of Parameter Estimators Using Monte Carlo Experiments

Arora and Singh (1987) estimated extreme value type 1 distribution parameters and quantiles by methods of moments, maximum likelihood estimation, probability weighted moments, entropy, mixed moments, least squares and incomplete means for Monte Carlo samples generated from two sampling cases: purely random process and serially correlated process. The performance of these estimators was statistically inter-compared. Additionally, a bias correction was made to the method of moments-quantile estimator. The corrected estimator provided nearly unbiased quantile estimates even for small samples and high non-exceedance probabilities. The work of Arora and Singh (1987) is essentially an extension of the one by Landwehr and Matalas (1979), where they used sampling experiments to compare the method of probability weighted moments with the method of moments and maximum likelihood estimation in two cases: purely random samples and serially correlated samples. Arora and Singh (1987) made a comprehensive evaluation by including all methods that were apparently available to that date. Additionally, they also addressed the question of bias correction for method of moments-quantile estimation. The MOM method has been widely used, owing to its simplicity. However, as investigated by Matalas, et al. (1979) and Lettenmaier and Burges (1982) among others, and also corroborated by this work, this estimator yielded biased estimates of the quantile.

Usually, selection of the best estimator is governed by the type of the loss function which is a measure of the loss resulting from over or under-design. In certain situations of design, the loss function is minimized by least squares estimator. Moreover, the bias corrected moment estimator, if not accompanied by a significant worsening of the mean square error (MSE), can prove to be useful in regional estimation procedure where the possible increase in the variance is made insignificant by the larger sample size. Sampling experiments were used to arrive at a practically unbiased moment-quantile estimator. This bias corrected estimator yielded nearly unbiased estimates of quantiles, even for samples of small size. Moreover, it did not entail any practical worsening of MSE. Indeed, as is shown by simulation, the MSE values obtained from the bias-correction estimator were close to those from the uncorrected estimator.

8.5.1 ANALYSIS

The inverse form of equation (8.1b) is given by

$$x(F) = b - \frac{\ln(-\ln F)}{a} \tag{8.73}$$

where x (F) denotes the quantile of cumulative probability F. For sample sizes n = 5, 10, 15, 20, 30, 50, 100, parameters a and b were estimated by the methods of moments (MOM), maximum likelihood estimation (MLE), probability weighted moments (PWM), entropy (POME), mixed moments (MIX), least squares (LEAS), and incomplete means (ICM). The quantiles x (F), for F = 0.05, . . . , 0.99, were then calculated from equation (8.73).

Let $\hat{\theta}$ denote an estimate of $\theta \in (a, b, x(F))$. $\hat{\theta}$ is a random variable whose distribution function depends upon sample size, the method of parameter estimation, and of course, the distribution of the sample itself. The performance statistics of $\hat{\theta}$ are given as follows:

$$\text{Bias, BIAS } (\hat{\theta}) = \theta - E(\hat{\theta}) \tag{8.74}$$

$$\text{Variance, } STD^2(\hat{\theta}) = E[\hat{\theta} - E(\hat{\theta})]^2 \tag{8.75}$$

$$\text{Mean Square Error, MSE } (\hat{\theta}) = E(\theta - \hat{\theta})^2$$
$$= BIAS^2 (\hat{\theta}) + STD^2 (\hat{\theta}) \quad (8.76)$$

These statistics present a good picture of the relative performance of various estimators of θ. They were estimated through Monte Carlo sampling experiments, i.e., through generation of a large number of pseudo-random samples for: (1) purely random process (independent and identically distributed Gumbel random variables), and (2) serially correlated process (with the first order serial correlation coefficient of o.5).

On the basis of N Monte Carlo samples of size n, the statistics of $\hat{\theta}$, viz, equations (8.74) - (8.76) were estimated as:

$$E(\hat{\theta}) \cong \mu(\hat{\theta}) = \sum_{i=1}^{N} \frac{\hat{\theta}_i}{N} \quad (8.77)$$

$$STD^2(\hat{\theta}) \cong \sum_{i=1}^{N} \frac{[\hat{\theta}_i - \mu(\hat{\theta})]^2}{N-1} \quad (8.78)$$

$$\text{BIAS } (\hat{\theta}) \cong \theta - \mu(\hat{\theta}) \quad (8.79)$$

$$\text{MSE } (\hat{\theta}) \cong BIAS^2 (\hat{\theta}) + STD^2 (\hat{\theta}) \quad (8.80)$$

These estimates were expected to be very close approximations to the theoretical values owing to the large value of N (= 50,000 for n = 5, ... , 100, and, = 10,000 for n = 1000).

The mean square error of all methods relative to that of MLE was compared using the relative efficiency defined as:

$$EFF(\hat{\theta}) = \frac{MSE(\hat{\theta} \mid MLE)}{MSE(\hat{\theta} \mid other\ method)} \quad (8.81)$$

A value of EEF $(\hat{\theta}) < 1$ implied that the method under consideration was less efficient (i.e., had higher mean square error) compared to MLE and vice versa.

8.5.2 PERFORMANCE STATISTICS OF PARAMETER AND QUANTILE ESTIMATORS

8.5.2.1 Case 1: Parameter Estimates for a Purely Random Process: The MIX and ICM methods were prima-facie rejected as unreliable estimators of Gumbel parameters. MIX, while performing reasonably efficiently for estimation of a, failed in providing even a moderately biased estimate of b, and thus resulted in highly inefficient estimate of b. ICM failed to provide a satisfactory estimate for either a or b as reflected by its very high bias and standard deviation of a and high standard deviation of b.

Of the remaining five methods, the bias of a showed the following trend:

MLE > POME > MOM > PWM > LEAS, for n = 5, 10
MOM > MLE > POME > PWN > LEAS, for n = 15 - 50
MOM > MLE > POME > LEAS > PWM, for n = 75, 100
LEAS > others , for n = 1000

The standard deviation of a compared as:

MLE > POME > MOM > PWM > LEAS, for n = 5
MOM > PWM > MLE > POME > LEAS, for n = 10
MOM > PWM = LEAS > POME > MLE, for n = 15
MOM > LEAS > PWM > POME > MLE, for n > 20

The efficiency of a compared as:

LEAS > PWM > MOM > POME > MLE, for n = 5
LEAS > PWM > POME > MLE > MOM, for n = 10, 15
MLE > POME > PWM > LEAS > MOM, for n > 20

where '>' means that the method on the left hand side of > has a bigger statistic (bias, standard deviation, or efficiency) then the method on the right hand side.

From the above trends, it appeared that for rather small samples $(n \leq 15)$, LEAS was the preferred method for estimating a. The results also revealed that as n increased, the efficiency of estimating a by POME remained close to 1.00, while the efficiency from other methods reduced considerably. In estimating b, PWM provided practically unbiased estimates. Analyzing the bias, standard deviation and the efficiency in much the same way as for a, it was easily concluded that PWM provided superior estimates of b for the entire sample range considered.

8.5.2.2 Quantile Estimates: PWM provided unbiased quantile estimates for all n and F. MOM provided estimates with lower bias than MLE and POME. POME resulted in slightly less bias than MLE, while LEAS produced more bias than MLE for all n except for n = 5. MIX and ICM again failed to provide satisfactory estimates of quantiles compared to other estimators because their lower bias was deteriorated by high standard deviation and vice versa, thus resulting in low efficiencies of estimates compared to MLE estimates.

The standard deviation of quantile estimates, while decreasing for increasing n, increased for higher non-exceedance probabilities, F. MLE resulted in lowest standard deviation, closely followed by POME. MOM had slightly higher standard deviation than PWM for nearly all n and F. LEAS estimates showed higher standard deviation than MOM, although the difference reduced as n increased. MLE estimates were most efficient, closely followed by POME estimates. The PWM estimates proved to be more efficient than MOM estimates, though less coefficient than POME estimates.

8.5.2.3 Case 2: Parameter Estimates in a Serially Correlated Process: When the samples were generated from a serially correlated process but assumed to be random for the purpose of estimation, all the estimation methods produced significantly higher bias and standard deviation than the corresponding random process estimators of case 1. However, LEAS consistently produced least bias of a followed by PWM.

From sample size 10 onwards, MLE, followed closely by POME, gave least standard deviation of a. However, up to sample size n = 20, LEAS resulting in comparable standard deviation produced estimates of a with a higher efficiency than MLE. Hence, LEAS can be preferred for such sample sizes.

For b, PWM was without doubt the superior method resulting in less bias, least standard deviation and higher efficiency estimates. Although LEAS resulted in lower bias than PWM for n > 15, it showed less efficiency than PWM owing to its higher standard deviation. But it is significant to note that the effect of serial correlation was to markedly lower the performance deviation among the first five methods. In fact the first five methods performed to within 98 percent of the efficiency of MLE method for estimating b.

8.5.2.2.4 Quantile Estimates: The bias in quantile estimates also increased for serially correlated samples as compared to purely random samples. LEAS provided least biased estimates of the first five methods for sample sizes 5 to 100. PWM provided the next lowest bias. MLE and POME continued to provide very close bias estimates, although POME produced slightly lower bias. MOM provided lower bias than MLE for up to about n = 30, beyond which MOM produced slightly higher bias than MLE. Quite in contrast with other methods, the absolute bias resulting from MIX increased with n for any given F.

MLE resulted in least standard deviation of quantile estimates among the first five methods, even for n = 5, closely followed by POME. MOM gave lower standard deviation than PWM for n = 5 and 10 only, after which mostly PWM produced lower standard deviation estimates. LEAS provided estimates with rather high standard deviation.

Except for n = 5, and some quantiles less than 0.5 for other n, MLE turned out to be the most efficient method for serially correlated samples, followed closely again by POME. PWM provided the next higher efficiency estimates mostly at all quantile values for all n except at n = 5. MOM came next and MOLS provided the least efficient estimates.

8.5.3 BIAS CORRECTION IN QUANTILE ESTIMATES OF MOM

Owing to its simplicity and ease of calculation, MOM has been widely used as an estimator of EV1 distribution parameters. However, MOM results in biased estimates as shown previously. The bias resulting from MOM-quantile estimator can be corrected using simulation results as follows: From equation (8.73) with sample estimates of parameters a and b, we have

$$x - \hat{x} = b - \hat{b} - \ln(-\ln F) \cdot \left[\frac{1}{a} - \frac{1}{\hat{a}}\right] \tag{8.82}$$

$$E(x - \hat{x}) = E(b - \hat{b}) - \ln(-\ln F) \cdot E\left[\frac{1}{a} - \frac{1}{\hat{a}}\right] \tag{8.83}$$

But from equations (8.55a) and (8.55b),

$$E(b - \hat{b}) = -0.57721 \cdot E\left[\frac{1}{a} - \frac{1}{\hat{a}}\right] \tag{8.84}$$

Substituting equation (8.84) in (8.83) we obtain

$$E(x - \hat{x}) = -\frac{1}{a} \cdot E\left[1 - \frac{a}{\hat{a}}\right] \cdot \left\{(0.57721 + \ln(-\ln F))\right\} \tag{8.85}$$

$E[1 - (a/\hat{a})]$ is the bias of the scaled random variable $a/\hat{a}$. It is a dimensionless quantity.

To investigate the bias of $a/\hat{a}$ as a function of the sample size and the distribution parameters, three sets of sampling experiments were carried out using N = 25,000 Monte Carlo samples of sizes n (= 5, 7, 10, 15, 20, 30, 40, 50, 75, 100, 150, and 200). The samples were generated respectively from the following populations:

(1) a = 1.00, b = 0.0
(2) a = 0.05, b = 100.0
(3) a = 0.01, b = 200.0

The bias $E[1 - (a/\hat{a})]$ was computed for various of n in each parameter set. The results were plotted on a log-log plot and are shown in Figure 8.4. It is apparent from this figure that the bias of $a/\hat{a}$ is independent of the population parameters from which it was computed and depends on the sample size n. The regression line is closely fitted by the equation:

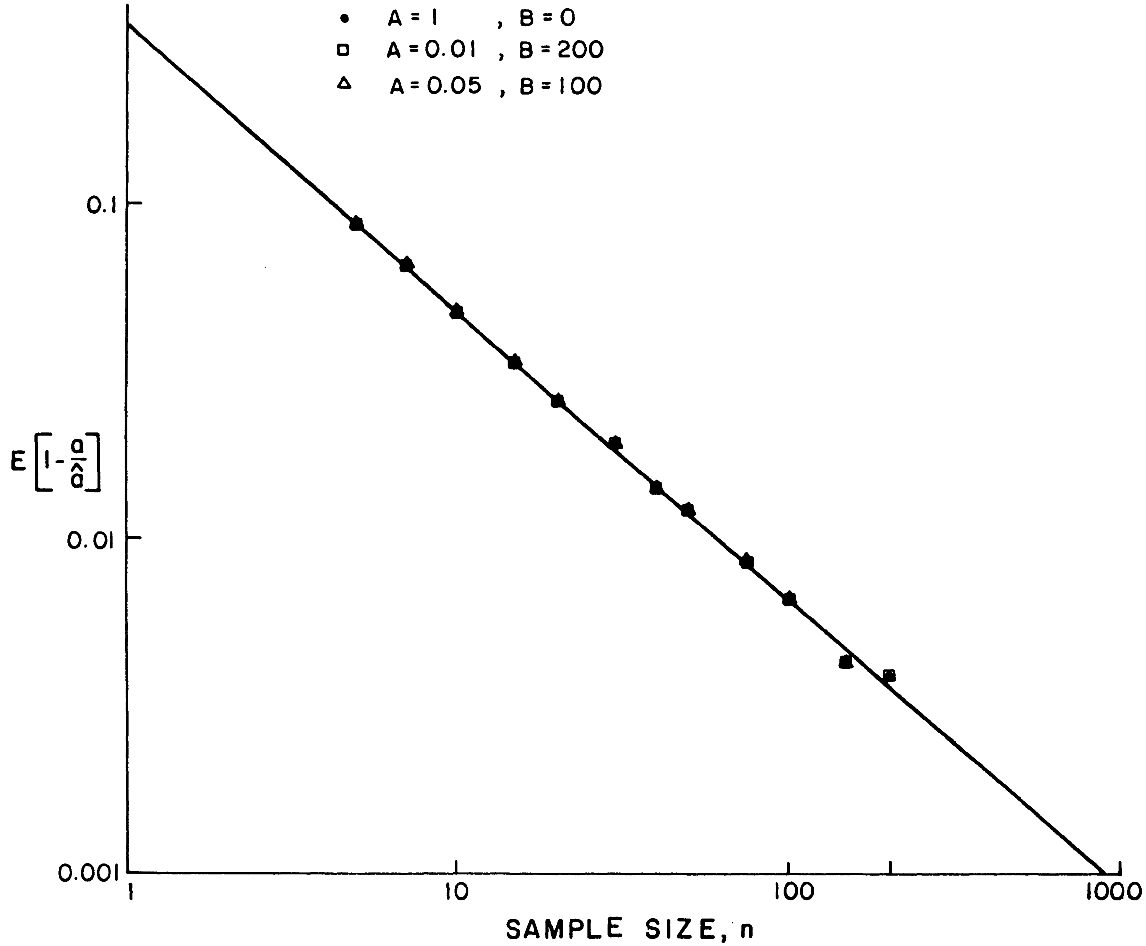

Figure 8.4 Bias $E[1-(a / \hat{a})]$ versus sample size n.

$$E(1-\frac{a}{\hat{a}}) = \frac{0.35}{n^{0.8589}} \cong f(n) \quad (8.86)$$

where f (n) denoting the "true correction" is used in subsequent discussion.

Using equation (8.86) in equation (8.85), we obtain

$$E(x-\hat{x}) = -\frac{1}{a} \cdot [0.57721 + \ln(-\ln F)] \quad (8.87)$$

From equation (8.86) we can write

$$E[\frac{1}{\hat{a}(1-f(n))}] = \frac{1}{a} \quad (8.88)$$

which implies that $1/[\hat{a}(1-f(n))]$ is an unbiased estimator of 1 / a. Substituting equations (8.88) in equation (8.87) and simplifying, we get

$$E[x-(\hat{x}-f(n) \cdot [0.57221 + \ln(-\ln F) \cdot \frac{1}{\hat{a}(1-f(n))})] = 0 \quad (8.89)$$

Hence, by definition

$$\hat{x} = \hat{x} - f(n) \cdot [0.57221 + \ln(-\ln F)] \cdot \frac{1}{\hat{a}(1-f(n))} \quad (8.90)$$

is an unbiased estimator of x. It is denoted as CMOM estimator of quantile. Simplifying equation (8.90) further, we get

$$\hat{x}=\hat{b}-\frac{1}{\hat{a}} \cdot [\ln(-\ln F) + \{0.57221 + \ln(-\ln F\} \cdot \frac{f(n)}{1-f(n)}] \quad (8.91)$$

and understandably enough the bias corrected quantile estimator is a function of not only 'F' but 'n' also.

The bias and standard deviation of equation (8.91), with f (n) substituted by the expression in equation (8.88), against the method CMOM showed that the variance of $\hat{x}$ was larger than the variance of $\hat{x}$. Figure 8.5 shows the bias of original and corrected quantile estimators for 99.9 percent non-exceedance probability. The results were typical of other probabilities too.

8.5.4 CONCLUDING REMARKS

Based on a statistical comparison of seven estimators of EV1 distribution parameters and quantiles, using Monte Carlo sampling experiments, performed on two cases: a purely random process and a serially correlated process, some of the important conclusions were as follows: (1) The methods of mixed moments and incomplete means resulted in poor estimation of the parameters and quantiles. (2) The method of least squares provided minimum bias and maximum efficiency estimate of parameter 'a' for very small samples and also provided competitive estimates of parameter 'b'. (3) The maximum likelihood method generally provided most efficient quantile estimates followed closely by the entropy method. In fact, POME performed practically in the same manner as MLE, and was relatively easier to solve. (4) For small samples, the method of probability weighted moments and the method of moments performed comparably in efficiency of estimating the quantiles. However, the efficiency of PWM improved relative to MLE with increasing sample size. PWM also resulted in nearly unbiased quantile estimates. (5) The incorporation of serial correlation in samples resulted in deterioration of the performance of

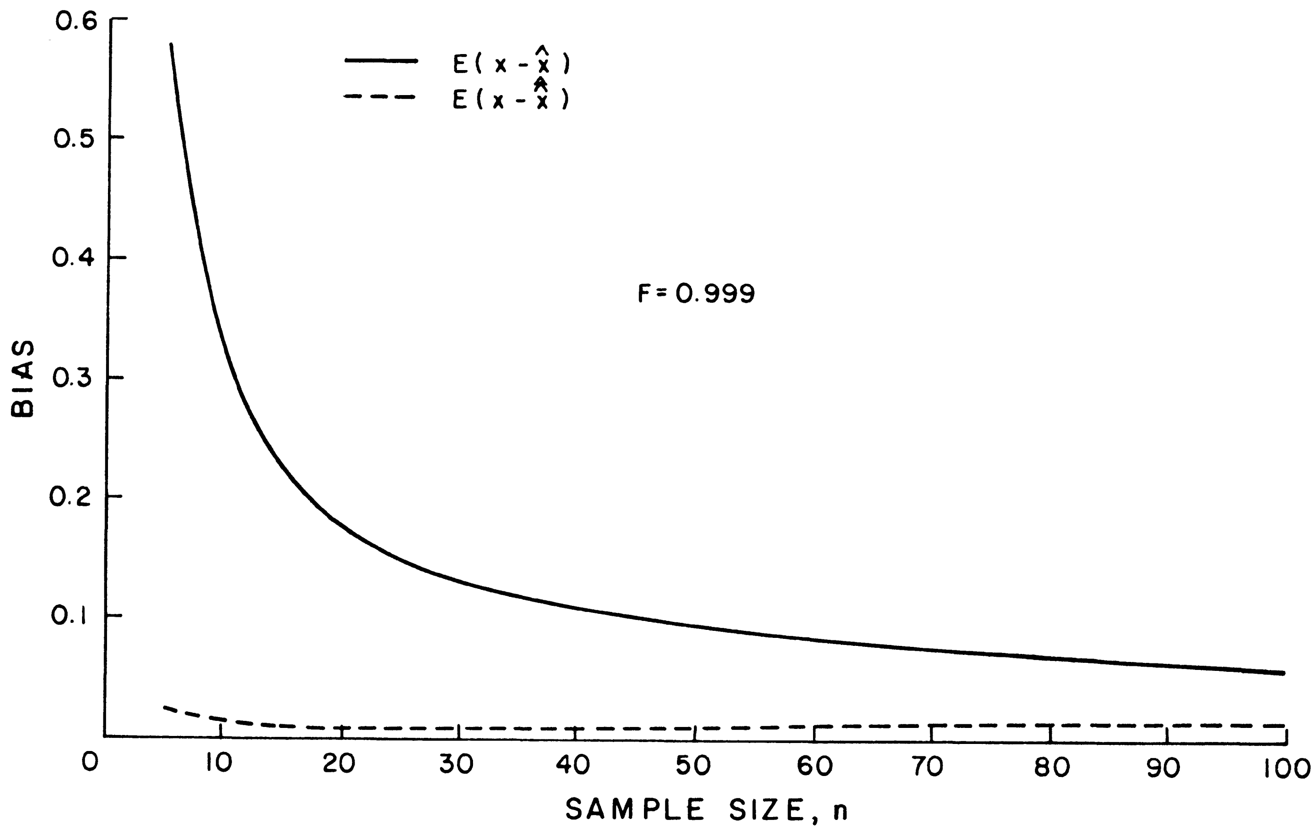

Figure 8.5 Bias in quantile versus sample size n.

all estimators. However, all the methods performed much more similarly in this case. (6) A bias corrected MOM estimator of quantile, developed for purely random process, resulted in practically unbiased quantiles even for very small sample sizes without causing any appreciable deterioration in the mean square error (MSE). It is clear that no method was uniformly superior across all sample sizes and quantiles.

References

Adamowski, K., 1981. Plotting formula for flood frequency. *Water Resources Bulletin*, Vol. **17**, No. 2, pp. 197-202.

Al-Mashidini, G., Pande, B.B.L. and Mujda, M.F., 1978. A simple version of Gumbel's method for flood estimation. *Hydrological Sciences Bulletin*, Vol. **23**, No. 3, pp. 373-380.

Anderson, R.L., 1941. Distribution of the serial correlation coefficient. *Annals of Mathematical Statistics*, Vol. **8,** pp. 1-13.

Arora, K. and Singh, V.P., 1987. On statistical intercomparison of EV1 estimators by Monte Carlo simulation. *Advances in Water Resources*, Vol. **10**, No. 2, pp. 87-107.

Bardsley, W.E. and Manly, B.F.J., 1987. Transformations for improved convergence of distribution of flood maxima to a Gumbel limit. *Journal of Hydrology*, Vol. **91**, pp. 137-152.

Benson, M.A., 1968. Uniform flood frequency esimating methods for federal agencies. *Water Resources Research*, Vol. **4**, No. , pp. 891-908.

Bobee, B. and Robitaille, 1977. The use of Pearson type 3 and log Pearson type 3 distributions revisited. *Water Resources Research*, Vol. **13,** No. 2, pp. 427-443.

Chow, V.T., 1953. Frequency analysis of hydrologic data. *Bulletin* **414**, Engineering Experiment Station, University of Illinois, Urbana, Illinois.

Coulson, A., 1966. Tables for computing and plotting flood frequency curves. *Technical Bulletin* No. **3**, Water Resources Branch, Department of Energy, Mines and Resources, Ottawa, Canada.

Fill, H.D. and Stedinger, J.R., 1995. L moment and probability plot correlation coefficient goodness-of-fit tests for the Gumbel distribution and impact of autocorrelation. *Water Resources Research*, Vol. 31, No. 1, pp. 225-229.

Fisher, R.A. and Tippett, L.H.C., 1928. Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Proceedings, Cambridge Philosophical Society*, Vol. **24**, pp. 48-55.

Fiorentino, M. and Gabriele, S., 1984. A correction for the bias of maximum likelihood estimators of Gumbel parameters. *Journal of Hydrology*, Vol. **73**, pp. 39-49.

Gumbel, E.J., 1941a. The return period of flood flows. *Annals of Mathematical Statistics*, Vol. **XII,** No. 2, pp. 163-190.

Gumbel, 1941b. The probability-interpretation of the observed return periods of floods. *Transactions, American Geophysical Union*, Vol. **22**, Part III, pp. 836-849.

Gumbel, E.J., 1942a. On the frequency distribution of extreme values in meteorological data. *Bulletin of the American Meteorological Society*, Vol. **23**, No.3, pp. 95-105.

Gumbel, E.J., 1942b. Statistical control-curves for flood discharges. *Transactions, American Geophysical Union*, Vol. **23**, pp. 489-500.

Gumbel, E.J., 1948. The statistical forecast of floods. *Technical Report*, Ohio Water Resources Board, Columbus, Ohio.

Gumbel, E.J., 1958. *Statistics of Extremes*. Columbia University Press, Columbia, New York

Hersfield, D.M. and Kohler, M.A., 1960. An empirical appraisal of the Gumbel extreme value procedure. *Journal of Geophysical Research*, Vol. **65**, No. 6, pp. 1737-1746.

Houghton, J.C., 1978. The incomplete means estimation procedure applied to flood frequency analysis. *Water Resources Research*, Vol. **14**, No. 6, pp. 1111-1115

Jain, D. and Singh, V.P., 1987. Estimating parameters of EV1 distribution for flood frequency analysis. *Bulletin, American Water Resources Association*, Vol. **23**, No.1, pp. 59-71.

Jenkinson, A.F., 1955. The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quarterly Journal of Royal Meteorological Society*, Vol. **81**, pp. 158-171.

Jowitt, P.W., 1979. The extreme value type 1 distribution and the principle of maximum entropy. *Journal of Hydrology*, Vol. **42**, pp. 23-38.

Kimball, B.F., 1949. An approximation to the sampling variances of an estimated maximum value of given frequency based on fit of doubly exponential distribution of maximum values. *Annals of Mathematical Statistics*, Vol. **20,** pp. 110-113.

Lambert, J.. And Duan, L., 1994. Evaluating risk of extreme events for univariate-loss functions. *Journal of Water Resources Planning and Management*, Vol. **120**, No. 3, pp. 382-399.

Landwehr, J.M., Matalas, N.C. and Wallis, J.R., 1979. Probability weighted moments compared with some traditional techniques in estimating Gumbel parameters and quantiles. *Water Resources Research*, Vol. **15**, No. 5, pp. 1055-1064.

Lattenmaier, D.P. and Burges, S.J., 1982. Gumbel's extreme value I distribution: A new look. *Journal of Hydraulics Division, ASCE*, Vol. **108**, No. HY4, pp. 502-514.

Lowery, M.D. and Nash, J.E., 1970. A comparison of methods of fitting the double exponential distribution. *Journal of Hydrology*, Vol. **10**, pp. 259-275.

Majumdar, K.C. and Sawhney, R.P., 1965. Estimates of extreme values by different distribution functions. *Water Resources Research*, Vol. **1**, No. 3, pp. 429-434.

Mann, H.B. and Whitney, D.R., 1947. On a test of whether one of two random variables is stochastically larger than the other. *Annals of Mathematical Statistics*, Vol. **18**, pp. 50-61.

Matalas, N.C., Slack, J.R. and Wallis, J.R., 1975. Regional skew in search of a parent. *Water Resources Research*, Vol. **11**, No. 6, pp. 815-826.

Phien, H.N., 1987. A review of methods of parameter estimation fro the extreme value type-1 distribution. *Journal of Hydrology*, Vol. **90**, pp. 251-268.

Phien, H.N. and Arbhabhirama, A., 1980. A comparison of statistical tests on the extreme-value type-1 distribution. *Bulletin, American Water Resources Association*, Vol. **16**, No. 5, pp. 856-861.

Raynal, J.A. and Salas, J.D., 1986. Estimation procedures for the type-1 extreme value distribution. *Journal of Hydrology*, Vol. **87**, pp. 315-336.

Reich, BM., 1970. Flood series compaed to rainfall extremes. *Water Resources Research*, Vol. **6**, No. 6, pp. 1655-1667.

Shen, H.W., Bryson, M.C. and Ochoa, I.D., 1980. The effect of tail assumption on flood predictions. *Water Resources Research*, Vol. **16**, pp. 361-364.

Siegel, S., 1956. *Nonparametric Statistics for the Behavioral Sciences.* Mc-Graw Hill Book Company, New York.

Singh, V.P. and Rajagopal, A.K., 1986. A new method of parameter estimation for hydrologic frequency analysis. *Hydrological Science and Technology*, Vol. **2**, No. 3, pp. 33-40.

Singh, V.P., Rajagopal, A.K. and Singh, K., 1986. Derivation of some frequency distributions using the principle of maximum entropy. *Advances in Water Resources*, Vol. **9**, pp. 91-106.

Singh, V.P., Singh, K. and Rajagopal, A.K., 1985. Application of the principle of maximum entropy (POME) to hydrologic frequency analysis. *Completion Report* **06**, 144 p., Louisiana Water Resources Research Institute, Louisiana State University, Baton Rouge, Baton Rouge, Louisiana.

Smith, R.L., 1986. Extreme value theory based on the r largest annual events. *Journal of Hydrology*, Vol. **86**, pp. 27-43.

Stol, P.T., 1971. On the decomposition of the extreme value distribution of daily rainfall depths and the derivation of probabilities of compound events. *Journal of Hydrology*, Vol. **14**, pp. 181-196.

Wald, A. and Walfowitz, J., 1943. An exact test for randomness in the nonparametric case based on serial correlation. *Annals of Mathematical Statistics*, Vol. **14**, pp. 378-388.

Weiss, L.L., 1955. A nomograph based on the theory of extreme values for determining values for various return periods. *Monthly Weather Review*, Vol. **83**, pp. 69-71.

CHAPTER 9

LOG-EXTREME VALUE TYPE 1 DISTRIBUTION

The logarithmically transformed extreme value type 1 (LEV1) distribution is the log-Gumbel distribution. The logarithmic version of EV1 distribution is not as popular as the original EV1 distribution. Reich (1970) employed log-Gumbel distribution to analyze annual series of maximum instantaneous flood peaks from 26 Pennsylvanian watersheds smaller than 200 square miles in area. He found consistently overestimation of long-period extremes from use of the log-Gumbel distribution. Using the principle of maximum entropy Singh (1985) derived the log-Gumbel distribution and its parameters. Heo and Salas (1996) estimated quantiles and confidence intervals for the log-Gumbel distribution. They used the methods of moments, maximum likelihood and probability weighted moments for parameter estimation.

Let Y = ln X, where X is a positive random variable. If Y has a Gumbel distribution, then X will have a log-Gumbel distribution with probability density function (pdf) given by

$$f(x) = \frac{a}{x} \, exp \, [-a(lnx - b) - e^{-(lnx-b)}] \tag{9.1}$$

where a and b are parameters, respectively, interpreted as scale and location parameters. Parameter b also specifies the lower bound for ln x. Thus, the log-Gumbel distribution is a two-parameter distribution. Its cumulative distribution function (cdf) can be expressed as
To verify if f(x), given by equation (9.1) is a pdf, we write

$$\begin{aligned} F &= \int_0^\infty f(x)dx \\ &= a \int_0^\infty \frac{1}{x} \, exp \, [-a(lnx - b) \\ &- e^{-a(lnx-b)}]dx \end{aligned} \tag{9.3}$$

Let z = ln x and dz = $\frac{dx}{x}$. Then equation (9.3) becomes

$$F = a\int_{-\infty}^{\infty} \frac{1}{x}\, exp\,[-a(z-b) - e^{-(z-b)}]xdz$$
$$= \int_{-\infty}^{\infty} a\, exp\,[-a(z-b) - e^{-(z-b)}]dz = 1 \tag{9.4}$$

This confirms that f (x) is a pdf.

9.1 Ordinary Entropy Method

9.1.1 SPECIFICATION OF CONSTRAINTS

Taking logarithm of equation (9.1) to the base 'e', we get

$$\log f(x) = \ln a - \ln x - a(\ln x-b) - \exp[-a(\ln x-b)] \tag{9.5}$$

Multiplying equation (9.5) by [-f(x)] and integrating between 0 and ∞ yield the entropy function:

$$I(x) = -\int_0^\infty f(x)\ln f(x)dx = -\ln a\int_0^\infty f(x)dx + \int_0^\infty \ln x\, f(x)dx$$
$$+ a\int_0^\infty \ln x\, f(x)dx - ab\int_0^\infty f(x)dx + \int_0^\infty \exp[-a(\ln x-b)]\, f(x)dx$$
$$= -[\ln a + ab]\int_0^\infty f(x)dx + (1+a)\int_0^\infty \ln x\, f(x)dx$$
$$+ \int_0^\infty \exp[-a(\ln x-b)]\, f(x)dx \tag{9.6}$$

From equation (9.6), the constraints appropriate for equation (9.1) can be written (Singh et al., 1985, 1986) as

$$\int_0^\infty f(x)dx = 1 \tag{9.7}$$

$$\int_0^\infty \ln x\, f(x)dx = E[\ln x] = E[y] = \bar{y} \tag{9.8}$$

$$\int_0^\infty \exp[-a(\ln x-b)]\, f(x)dx = E[\exp(-a(\ln x-b))] = 1 \tag{9.9}$$

The least-biased density function based on POME consistent with equations (9.7) - (9.9) takes the form:

$$f(x) = \exp[-\lambda_0 - \lambda_1 \ln x - \lambda_2 e^{-a(\ln x - b)}] \tag{9.10}$$

where λ_0, λ_1, and λ_2 are Lagrange multipliers.

9.1.2 CONSTRUCTION OF ZEROTH LAGRANGE MULTIPLIER

Substitution of equation (9.10) in equation (9.7) gives

$$\int_0^\infty f(x)dx = \int_0^\infty \exp[-\lambda_0 - \lambda_1 \ln x - \lambda_2 e^{-a(\ln x-b)}]dx = 1 \tag{9.11}$$

Equation (9.11) yields the partition function as

$$\exp(\lambda_0) = \int_0^\infty \exp[-\lambda_1 \ln x - \lambda_2 e^{-a(\ln x - b)}]dx \tag{9.12}$$

Equation (9.12) can be written as

$$W = \exp(\lambda_0) = \int_0^\infty \exp[\ln x^{\lambda_1} - \lambda_2 e^{-a(\ln x-b)}]dx$$

$$= \int_0^\infty x^{-\lambda_1} \exp[-\lambda_2 e^{-1(\ln x-b)}]dx \tag{9.13}$$

Equation (9.13) can be simplified as follows:

$$Let\ \lambda_2 e^{-a(lnx-b)} = y\,;\ e^{-a(lnx-b)} = \frac{y}{\lambda_2}\,;\ -a(lnx-b) = ln(\frac{y}{\lambda_2})$$

$$lnx-b = -\frac{1}{a}\,ln(\frac{y}{\lambda_2}) = \frac{1}{a}\,ln\,(\frac{\lambda_2}{y})\,;lnx = b + ln\,(\frac{\lambda_2}{y})^{1/a}$$

$$x = e^b\,exp\,[ln(\frac{\lambda_2}{y})^{1/a}] = e^b(\frac{\lambda_2}{y})^{1/a}\ or\ y^{1/a} = \frac{e^b\lambda_2^{1/a}}{x}$$

$$x = e^b\lambda_2^{1/a}y^{-(1/a)}\,;\ \frac{dx}{dy} = e^b\lambda_2^{1/a}(-\frac{1}{a})y^{(-(1/a)-1)} = -\frac{e^b}{a}\lambda_2^{(-(1/a)-1)}$$

Substitution of these quantities in equation (9.13) yields

$$W = exp\ (\lambda_0) = \int_\infty^0 [e^b\lambda_2^{1/a}y^{-(1/a)}]^{-\lambda_1}e^{-y}(-\frac{e^b}{a}\lambda_2^{1/a}y^{(-(1/a)-1)})dy$$

$$= \int_0^\infty \frac{1}{a}e^{-\lambda_1 b}\lambda_2^{-\lambda_1/a}y^{\lambda_1/a}e^b\lambda_2^{1/a}y^{(-(1/a)-1)}e^{-y}dy$$

$$= \int_0^\infty \frac{exp[-\lambda_1 b+b]}{a}\lambda_2^{(-(\lambda_1/a)+(1/a))}y^{((\lambda_1-1)/a)-1}dy \tag{9.14}$$

$$= \frac{exp[b(1-\lambda_1)]}{a}\lambda_2^{((1-\lambda_1)/a)}\int_0^\infty y^{((\lambda_1-1)/a)}e^{-y}dy$$

$$= \frac{exp[b(1-\lambda_1)]}{a}\lambda_2^{((1-\lambda_1)/a)}\,\Gamma(\frac{\lambda_1-1}{a})$$

The zeroth Lagrange multiplier λ_0 is given as

$$\lambda_0 = -\ lna + b - b\lambda_1 + (\frac{1-\lambda_1}{a})ln\lambda_2 + ln\ \Gamma(\frac{\lambda_1-1}{a}) \tag{9.15}$$

The zeroth Lagrange multiplier is also obtained from equation (9.12) as

$$\lambda_0 = ln \int_0^\infty exp\ [-\lambda_1 lnx - \lambda_2 e^{-a(lnx-b)}]dx \tag{9.16}$$

9.1.3 RELATION BETWEEN LAGRANGE MULTIPLIERS AND CONSTRAINTS

Differentiation of equation (9.15) with respect to λ_1 and λ_2 , respectively, yields

$$\frac{\partial\lambda_0}{\partial\lambda_1} = -b - \frac{\ln\lambda_2}{a} + \frac{\partial}{\partial\lambda_1}[ln\,\Gamma\,(\frac{\lambda_1-1}{a})] \tag{9.17}$$

$$\frac{\partial\lambda_0}{\partial\lambda_2} = \frac{1-\lambda_1}{\lambda_2 a} \tag{9.18}$$

Differentiation of equation (9.18) with respect to λ_1 gives

$$\begin{aligned}\frac{\partial\lambda_0}{\partial\lambda_1} &= -\frac{\int_0^\infty lnx exp[-\lambda_1 lnx - \lambda_2 e^{-a(lnx-b)}]dx}{\int_0^\infty exp[-\lambda_1 lnx - \lambda_2 e^{-a(lnx-b)}]dx}\\ &= -\int_0^\infty lnx exp[-\lambda_0 - \lambda_1 lnx - \lambda_2 e^{-a(lnx-b)}]dx\\ &= -\int_0^\infty lnx f(x)dx = -E\,[lnx] = -\bar{y}\end{aligned} \tag{9.19}$$

Differentiation of equation (9.18) with respect to λ_2 gives

$$\begin{aligned}\frac{\partial\lambda_0}{\partial\lambda_2} &= -\frac{\int_0^\infty e^{-a(lnx-b)} exp[-\lambda_1 lnx - \lambda_2 e^{-a(lnx-b)}]dx}{\int_0^\infty exp[-\lambda_1 lnx - \lambda_2 e^{-a(lnx-b)}]dx}\\ &= -\int_0^\infty e^{-a(lnx-b)} exp[-\lambda_0 - \lambda_1 lnx - \lambda_2 e^{-a(lnx-b)}]dx\\ &= -\int_0^\infty e^{-a(lnx-b)} f(x)dx = -E[e^{-a(lnx-b)}] = -1\end{aligned} \tag{9.20}$$

Equating equations (9.17) and (9.19), as well as equations (9.18) and (9.20), produces

$$b + \frac{ln\lambda_1}{a} - \frac{\partial}{\partial\lambda_1}[ln\ \Gamma\ (\frac{\lambda_1-1}{a})] = \bar{y} \tag{9.21}$$

$$\frac{\lambda_1 - 1}{\lambda_2 a} = 1 \tag{9.22}$$

9.1.4 RELATION BETWEEN PARAMETERS AND LAGRANGE MULTIPLIERS

Equation (9.22) can be cast as

$$\frac{\lambda_1 - 1}{a} = \lambda_2 \tag{9.23}$$

Inserting equation (9.23) in equation (9.21) gives

$$b + \frac{ln\lambda_1}{a} - \frac{\partial}{\partial\lambda_2}[ln\,\Gamma(\lambda_2)]\frac{\partial\lambda_2}{\partial\lambda_1} = \bar{y} \tag{9.24}$$

or

$$b + \frac{ln\lambda_1}{a} - \Psi(\lambda_2)\frac{1}{a} = \bar{y} \tag{9.25}$$

Equation (9.15) can be written as

$$\lambda_0 = -\,lna + b - b\lambda_1 + (\frac{1-\lambda_1}{a})\,ln\lambda_2 + ln\Gamma\,(\frac{\lambda_1 - 1}{a}) \tag{9.26}$$

Therefore, equation (9.10) can be expressed as

$$\begin{aligned} f(x) &= exp[lna - b + b\lambda_1 - (\frac{1-\lambda_1}{a})\,ln\lambda_2 \\ &\quad - ln\Gamma(\frac{\lambda_1 - 1}{a}) - \lambda_1 lnx - \lambda_2 e^{-a(lna-b)} \\ &= exp[lna]exp[b\lambda_1 - b]exp[ln(\lambda_2)^{((\lambda_1-1)/a)} \times \\ &\quad \times \exp[\ln(\Gamma(\frac{\lambda_1 - 1}{a})^{-1}] \times \exp\,[\ln x^{-\lambda_1}] - \exp\,[-\lambda_2 e^{-a(\ln x - b)}] \\ &= a\,\exp[b(\lambda_1 - 1)]\,(\lambda_2)^{(\lambda_1-1)/a)}\frac{x^{-\lambda_1}}{\Gamma(\frac{\lambda_1-1}{a})} \end{aligned} \tag{9.27}$$

A comparison of equation (9.27) with equation (9.1) shows

$$\lambda_1 = a + 1 \tag{9.28}$$

$$\lambda_2 = 1 \tag{9.29}$$

9.1.5 RELATION BETWEEN PARAMETERS AND CONSTRAINTS

The log-Gumbel distribution has two parameters a and b which are related to the Lagrange multipliers by equations (9.28) and (9.29), which, in turn, are related to the known constraints by equations (9.23) and (9.25). These two sets of equations are used to eliminate the Lagrange multipliers between them and directly relate the parameters to the constraints. Therefore,

$$E[e^{-a(\ln x-b)}] = 1 \tag{9.30}$$

$$b + \frac{1}{a}\ln(1+a) - \psi(1)\frac{1}{a} = \bar{y} \tag{9.31}$$

9.1.6 DISTRIBUTION ENTROPY

From the definition of entropy, we get

$$\begin{aligned} I(x) &= -\int_0^\infty f(x)\ln f(x)dx \\ &= -[\ln a + ab]\int_0^\infty f(x)dx + (1+a)\int_0^\infty \ln x f(x)dx + \int_0^\infty e^{-a(\ln x-b)}f(x)dx \\ &= -(\ln a + ab) + (1+a)\bar{y} + 1 = a(\bar{y}-b) + \ln(\frac{e}{a}) + \bar{y} = I(y) + \bar{y} \end{aligned} \tag{9.32}$$

where I(y) = $a(\bar{y}-b) + \ln(\frac{e}{a})$. Alternatively,

$$\begin{aligned} I(x) &= I(y) - E[\ln|J(\frac{y}{x})|] \\ J(\frac{y}{x}) &= \frac{\partial}{\partial x}(\ln x) = \frac{1}{x} \\ I(x) &= I(y) - E[\ln|\frac{1}{x}|] = I(y) + E[\ln x] = I(y) + \bar{y} \end{aligned} \tag{9.33}$$

which is the same as equation (9.32).

9.2 Parameter - Space Expansion Method

9.2.1 SPECIFICATION OF CONSTRAINTS

For this method the constraints, following Singh and Rajagopal (1986), are specified by equation (9.4) and

$$\int_0^\infty [\ln x + a(lnx-b)] f(x)dx = E[\ln x + a(lnx-b)] \tag{9.34}$$

$$\int_0^\infty \exp[-a(lnx-b)]f(x)dx = E[\exp(-a(lnx-b))] \tag{9.35}$$

9.2.2 DERIVATION OF ENTROPY FUNCTION

The pdf corresponding to the principle of maximum entropy (POME) and consistent with equations (9.7), (9.34) and (9.35) takes the form

$$f(x) = \exp[-\lambda_0 - \lambda_1 lnx - \lambda_1 a(lnx-b) - \lambda_2 \, e^{-a(lnx-b)}]dx \tag{9.36}$$

where λ_0, λ_1, and λ_2 are Lagrange multipliers. Insertion of equation (9.36) into equation (9.7) yields

$$\exp(\lambda_0) = \int_0^\infty \exp[-\lambda_l lnx - \lambda_1 a\,(lnx-b) - \lambda_2 \; e \; -a\,(lnx-b)]dx$$

$$= \frac{e^{b(1-\lambda_1)}}{a}\left(\frac{1}{\lambda_2}\right)^{\lambda_1 (1+\frac{1}{a}) - \frac{1}{a}} \Gamma[\lambda_1 (1+\frac{1}{a}) - \frac{1}{a}] \tag{9.37}$$

The zeroth Langrange multiplier is given by

$$\lambda_0 = -\ln a + b(1-\lambda_1) - [\lambda_1 (1+\frac{1}{a}) - \frac{1}{a}] \ln\lambda_2 + \ln\Gamma[\lambda_1 (1+\frac{1}{a}) - \frac{1}{a}] \tag{9.38}$$

The zeroth Lagrange multiplier is also obtained from equation (9.37) as

$$\lambda_0 = \ln \int_0^\infty \exp[-\lambda_1 lnx - \lambda_1 a(lnx-b) - \lambda_2 e^{-a(lnx-b)}]dx \tag{9.39}$$

Introduction of equation (9.37) in equation (9.36) yields

$$f(x) = \frac{a e^{b(\lambda_1 - 1)} \lambda_2^{\lambda_1(1+\frac{1}{a})-\frac{1}{a}}}{\Gamma[\lambda_1(1+\frac{1}{a})-\frac{1}{a}]} \exp[-\lambda_1 lnx - \lambda_1 a(lnx-b) - \lambda_2 e^{-a(lx-b)}] \tag{9.40}$$

A comparison of equation (9.40) with equation (9.1) shows that $\lambda_1 = \lambda_2 = 1$. Taking - logarithm of equation (9.40) yields

$$-\ln f(x) = -\ln a - b(\lambda_1 - 1) - [\lambda_1 (1 + \frac{1}{a}) - \frac{1}{a}]\ln\lambda_2$$

$$+ \ln\Gamma[\lambda_1 (1 + \frac{1}{a}) - \frac{1}{a}] + \lambda_1 lnx + \lambda_1 a(lnx - b) + \lambda_2 e^{-a(lnx-b)} \tag{9.40}$$

Multiplying equation (9.40) by f(x) and integrating from - ∞ to + ∞, we get the entropy function:

$$I(f) = -lna - b(\lambda_1 - 1) - [\lambda_1(1 + \frac{1}{a}) - \frac{1}{a}]\ln\lambda_2$$
$$+ \ln\Gamma[\lambda_1(1+\frac{1}{a}) - \frac{1}{a}] + \lambda_1 E[lnx] + \lambda_1 E[a(lnx-b)] \tag{9.41}$$
$$+\lambda_2 E[e^{-a(lnx-b)}]$$

9.2.3 RELATION BETWEEN PARAMETERS AND CONSTRAINTS

Taking partial derivatives of equation (9.41) with respect to λ_1, λ_2, a and b separately, and equating each derivative to zero, one obtains

$$\frac{\partial I}{\partial \lambda_1} = 0 = -b - (1+\frac{1}{a})\ln\lambda_2 + (1+\frac{1}{a})\psi(K) + E[lnx]$$
$$+ E[a(lnx-b)], K = \lambda_1(1+\frac{1}{a}) - \frac{1}{a} \tag{9.42}$$

where $\psi(.)$ Is the digamma function. Simplification of equations (9.42) to (9.45) leads to

$$a\, E[\ln x] = ab - \psi(1) \tag{9.46}$$

$$E[e^{-a(\ln x - b)}] = 1 \tag{9.47}$$

$$E[\ln x - b] - E[e^{-a(\ln x - b)} \ln x] = 1/a \tag{9.48}$$

$$\frac{\partial I}{\partial \lambda_2} = 0 = -\frac{K}{\lambda_2} + E[e^{-a(lnx-b)}] \tag{9.43}$$

$$\frac{\partial I}{\partial a} = 0 = -\frac{1}{a} - (\frac{-\lambda_1}{a^2} + \frac{1}{a^2}) \ln\lambda_2 + (\frac{1}{a^2} - \frac{\lambda_1}{a^2})\, \psi(k) + \lambda_1 E[lnx - b] - \lambda_2 E[e^{-a(lnx-b)} lnx] \tag{9.44}$$

$$\frac{\partial I}{\partial b} = 0 = -(\lambda_1 - 1) - a\lambda_1 + \lambda_2 E[a e^{-a(lnx-b)}] \tag{9.45}$$

$$E[e^{-a(\ln x - b)}] = 1 \tag{9.49}$$

Equations (9.47) and (9.49) are the same. The parameter estimation equations are equations (9.46) and (9.47).

9.3 Other Methods of Parameter Estimation

In the Y domain, all methods of parameter estimation described in Chapter 8 will apply and will thus not be repeated. Only the methods of moments (MOM) and maximum likelihood estimation (MLE) are briefly summarized.

9.3.1 METHOD OF MOMENTS

In the Y domain, Y = ln X , follows the EV1 distribution. Therefore, parameters a and b can be estimated from the first two moments of Y. The estimation equation for the method of moments (MOM) are summarized below:

$$a = 1.2825 / \sigma_y \tag{9.50}$$

$$b = \mu_y - 0.45\, \sigma_y \tag{9.51}$$

where μ_y and σ_y are respectively mean and standard deviation of y and are defined as

$$\mu_y = \frac{1}{n}\sum_{i=1}^{n} \ln x_i \tag{9.52}$$

$$\sigma_y = [\sum_{i=1}^{n} (\ln x_i - \mu_y)^2/(n-1)]^{0.5} \tag{9.53}$$

9.3.2 METHOD OF MAXIMUM LIKELIHOOD ESTIMATION

The method of maximum likelihood estimation (MLE) gives

$$E[\exp\{-a(\ln x - b)\}] = 1 \tag{9.54}$$

$$\sum_{i=1}^{n} \ln x_i x_i^{-a} = (\sum_{i=1}^{n} \ln x_i - \frac{1}{a}) \sum_{i=1}^{n} x_i^{-a} \tag{9.55}$$

Note that equations (9.54) and (9.55) are equivalent to equations (9.46) and (9.47). This implies that POME and MLE methods should yield comparable parameter estimates.

9.4 Comparative Evaluation of Parameter Estimation Methods

9.4.1 ANNUAL FLOOD DATA

Singh (1986) made a comparative evaluation of MOM, MLE and POME using instantaneous annual flood values for five rivers with drainage basins ranging in areas from 135 to 1,653 km^2. Some pertinent characteristics of the data are given in Table 9.1. These data were selected on the basis of length, completeness, homogeneity and independence of record. These stations had more than 30 years of record. Each of the five data sets was tested for homogeneity by using the Kruskal-Wallis test (Siegel, 1956) and the Mann-Whitney test (Mann and Whitney, 1947), as well as for independence by the Wald-Wolfowitz test (Wald and Wolfowitz, 1943) and the Anderson test (Anderson, 1941). In each case the sample was found homogeneous and independent.

9.4.2 PERFORMANCE CRITERIA

Two criteria were used for comparing the three methods of parameter estimation. These have been used by Benson (1698) and also by Bobee and Robitaille (1977). The first criterion is defined as the average of relative deviations between observed and computed flood discharge for the entire sample, with algebraic sign ignored,

$$D_a = \Sigma \mid G(T) \mid / n \tag{9.56}$$

in which

$$G(T) = [X_0(T) - X_c(T)] \cdot 100 / X_0(T) \tag{9.57}$$

where X_0 and X_c are the observed and computed flood values, respectively, for a given value of return period T, and n is the sample size.

The other criterion is the average of squares of relative deviations

$$D_r = \Sigma[G(T)]^2/n \qquad (9.58)$$

The statistics D_a and D_r are objective indexes of the goodness of fit of each method to sample data. The observed flood discharge corresponded to a return period which was computed by using the Gringorton plotting position formula,

$$T = (n + 0.12) / (m - 0.44) \qquad (9.59)$$

in which m is the rank assigned to each data point in the sample with one for the highest flow, two for the second highest and so on, with n for the lowest period.

Table 9.1 Pertinent characteristics of annual flood data of five USGS gaging stations. $\bar{x}$ = discharge (m^3/s) ; S_x = standard deviation; C_s = coefficient of skewness; C_k = coefficient of kurtosis.

River	Gaging Station Location	DrainageArea (sq. km)	Period of Record	$\bar{x}$	S_x	C_s	C_k
St. Mary	Still Water, Nova Scotia	1,653	1915-1974	409.5	147.9	1.42	6.25
Royal	Yarmouth	642	1950-82	110.7	54.0	2.24	10.67
Nezinscot	Turner Center, New Hampshire	733	1942-82	105.2	61.2	3.03	14.95
Pemigewasset	plymouth, New Hapmshire	1,243	1904-81	658.6	297.6	2.19	8.36
Smith	Bristol, New Hampshire	135	1919-81	57.5	33.5	3.04	14.94

9.4.3 EVALUATION OF METHODS

Parameters a and b were estimated using each method for each sample and are given in Table 9.2. Clearly, the three methods yielded comparable values of b for the five rivers, but produced values of a which differed at the first decimal place. Tables 9.3 and 9.4 give values of D_a and D_r. According to the values of D_a, MOM was the best of the three, followed by MLE and then POME. However, the results were mixed according to the values of D_r. For three rivers (Royal, Nezinscat and Pemigewasset) POME was the best of the three methods, followed by MOM and then MLE. For the remaining two rivers (St. Mary and Smith), however, MOM was the best, followed by MLE and POME. Since values of D_r and D_a are relatively small, three methods can be considered comparable.

References

Anderson, R.L., 1941. Distribution of the serial correlation coefficient. *Annals of Mathematical Statistics*, Vol. **8,** pp. 1-13.

Table 9.2 Parameter estimates obtained by MOM, MLE and POME

River	MOM		MLE		POME	
Gaging Station	a	b	a	b	a	b
St. Mary	3.8	5.8	3.7	5.8	3.4	5.8
Royal	2.9	4.4	2.9	4.4	2.4	4.4
Nezinscot	2.9	4.4	2.9	4.3	2.8	4.3
Pemigewasset	3.5	6.2	3.5	6.3	3.3	6.2
Smith	2.9	3.7	2.9	3.7	2.8	3.7

Table 9.3 Values of the mean absolute relative derivations, D_a.

River Gaging Station	MOM	MLE	POME
St. Mary	5.41	6.12	8.21
Royal	9.92	13.53	11.30
Nezinscot	3.82	4.11	9.87
Pemigewasset	3.72	4.66	4.29
Smith	4.32	4.69	7.21

Table 9.4 Values of the mean square deviations, D_r.

River Gaging Station	MOM	MLE	POME
St. Mary	0.439	0.632	1.196
Royal	2.21	3.23	2.45
Nezinscot	0.29	0.29	0.22
Pemigewasset	0.38	0.43	0.32
Smith	0.32	0.34	1.73

Heo, J.H. and Salas, J.D., 1996. Estimation of quantiles and confidence intervals for the log-Gumbel distribution. *Stochastic Hydrology and Hydraulics*, Vol. **10**, pp. 187-207.

Mann, H.B. and Whitney, D.R., 1947. On a test of whether one of two random variables is stochastically larger than the other. *Annals of Mathematical Statistics*, Vol. **18**, pp. 50-61.

Reich, B.M., 1970. Flood series compared to rainfall extremes. *Water Resources Research*, Vol. **6**, No. 6, pp. 1655-1667.

Siegel, S., 1956. *Nonparametric Statistics for Behavioral Sciences*. McGraw-Hill Book Company, New York.

Singh, V.P., 1985. On the log-Gumbel distribution. *Hydrology Journal of IAH*, Vol. **VIII**, No. 4, pp. 34-42.

Singh, V.P. and Rajagopal, A.K., 1986. A new method of parameter estimation for hydrologic frequency analysis. *Hydrological Science and Technology*, Vol. **2**, No. 3, pp. 33-40.

Singh, V.P., Rajagopal, A.K. and Singh, K., 1986. Derivation of some frequency distributions using the principle of maximum entropy. *Advances in Water Resources*, Vol. **9**, pp. 91-106.

Singh, V.P., Singh, K. and Rajagopal, A.K., 1985. Application of the principle of maximum entropy (POME) to hydrologic frequency analysis. *Completion Report* **06**, 144 p., Louisiana Water Resources Research Institute, Louisiana State University, Baton Rouge, Baton Rouge, Louisiana.

Wald, A. and Wolfowitz, J., 1943. An exact test for randomness in the nonparametric case based on serial correlation. *Annals of Mathematical Statistics*, Vol. **14**, pp. 378-388.

CHAPTER 10

EXTREME VALUE TYPE III DISTRIBUTION

The extreme value type (EV) III distribution has been employed for frequency analysis of low river flows (Gumbel, 1963; Matalas, 1963; Condie and Nix, 1975; Kite, 1978; Loganathan et al., 1985). Otten and Van Montefort (1978) discussed tests for the EV distributions. Gumbel (1963) estimated the EV III parameters using the method of moments (MOM). Matalas (1963) estimated them using MOM and the method of maximum likelihood estimation (MLE). Condie and Nix (1975) also used MOM and MLE. Kite (1978) described both MLE and MOM for the EV III distribution. Singh (1987) employed the principle of maximum entropy (POME) to estimate the EV III parameters and compared it with MOM and MLE.

A random variable X is said to have an extreme value type (EV) III distribution if its probability density function (pdf) is given by

$$f(x) = \frac{a}{b-c} \left(\frac{x-c}{b-c}\right)^{a-1} \exp\left[-\left(\frac{x-c}{b-c}\right)^{a}\right], \; a > 0 \, , \; b > 0 \qquad (10.1)$$

in which `a` is the scale parameter equal to the order of the lowest derivative of the cumulative distribution function (cdf) that is not zero at x = c, b is the location parameter, and c is the lower limit to x. The EV III distribution is a three-parameter distribution. Its cdf can be expressed as

$$F(x) = \exp\left[-\left(\frac{x-c}{b-c}\right)\right] \qquad (10.2)$$

10.1 Ordinary Entropy Method

10.1.1 SPECIFICATION OF CONSTRAINTS

Taking logarithm of equation (10.1) to the base 'e' results in

$$\ln f(x) = \ln a - \ln(b-c) + (a-1)\ln(x-c) - (a-1)\ln(b-c) - \frac{(x-c)^a}{(b-c)^a} \qquad (10.3)$$

Multiplying equation (10.3) by [-f(x)] and integrating between c and ∞ yield the entropy function:

$$-\int_c^\infty f(x)\ln f(x)dx = -\int_c^\infty [\ln a + \ln(b-c) - (a-1)\ln(b-c)]f(x)dx$$
$$+ (a-1)\int_c^\infty \ln(x-c)\, f(x)dx + \frac{1}{(b-c)^a}\int_c^\infty (x-c)^a f(x)dx \tag{10.4}$$

From equation (10.4) the constraints appropriate for equation (10.1) can be written (Singh et al., 1985, 1986) as

$$\int_c^\infty f(x)dx = 1 \tag{10.5}$$

$$\int_c^\infty \ln(x-c)\, f(x)dx = E[\ln(x-c)] \tag{10.6}$$

$$\int_c^\infty (x-c)^a f(x)dx = E[(x-c)^a] = (b-c)^a \tag{10.7}$$

Equations (10.5) and (10.7) can be verified as follows: Integrating equation (10.1) between c and ∞, one gets

$$\int_c^\infty f(x)dx = \frac{a}{b-c}\int_c^\infty \left(\frac{x-c}{b-c}\right)^{a-1} \exp\left[-\left(\frac{x-c}{b-c}\right)^a\right] dx \tag{10.8}$$

Let y = (x-c) /(b-c). Then dx = (b-c)dy. Therefore, equation (10.8) becomes

$$\int_c^\infty f(x)dx = \frac{a}{b-c}\int_c^\infty y^{a-1}\exp[-y^a]\,(b-c)dy = a\int_0^\infty y^{a-1}\exp[-y^a]dy \tag{10.9}$$

Let $z = y^a$. Then $dz = a\, y^{a-1}\, dy$. Therefore, equation (10.9) becomes

$$\int_c^\infty f(x)dx = a\int_0^\infty y^{a-1}e^{-z}\frac{dz}{ay^{a-1}} = \int_0^\infty e^{-z}dz = 1 \tag{10.10}$$

Likewise, equation (10.7) can be written as

$$\int_c^\infty (x-c)^a f(x)dx = \int_c^\infty (x-c)^a \left(\frac{a}{b-c}\right)\left(\frac{x-c}{b-c}\right)^{a-1} \exp\left[-\left(\frac{x-c}{b-c}\right)^a\right]dx \tag{10.11}$$

Let $\left(\frac{x-c}{b-c}\right)^a = y$. Then

$$\frac{dy}{dx} = \frac{a(x-c)^{a-1}}{(b-c)^a} = \frac{a}{x-c}\left(\frac{x-c}{b-c}\right)^a \tag{10.12}$$

$$dx = dy\left(\frac{x-c}{a}\right)\frac{1}{\left(\frac{x-c}{b-c}\right)^a} \tag{10.13}$$

Substituting equation (10.13) in equation (10.11), one obtains the following:

$$\begin{aligned}\int_c^\infty (x-c)^a f(x)dx &= \int_0^\infty (x-c)^a \left(\frac{a}{b-c}\right)\left(\frac{x-c}{b-c}\right)^{a-1} e^{-y}\, dy \left(\frac{x-c}{a}\right)\left(\frac{b-c}{x-c}\right)^a \\ &= \int_0^\infty (x-c)^a \left(\frac{a}{b-c}\right) y\left(\frac{x-c}{b-c}\right)^{-1} e^{-y}\left(\frac{x-c}{a}\right) y\, dy \\ &= \int_0^\infty (x-c)^a e^{-y}\, dy\end{aligned} \tag{10.14}$$

$$\begin{aligned}&= \int_0^\infty y\,(b-c)^a e^{-y}\, dy = (b-c)^a \int_0^\infty y e^{-y}\, dy = (b-c)^a\,\Gamma(2) \\ &= (b-c)^a\end{aligned} \tag{10.15}$$

10.1.2 CONSTRUCTION OF ZEROTH LAGRANGE MULTIPLIER

The least-biased density function f(x) based on the principle of maximum entropy (POME) and consistent with equations (10.5) - (10.7) takes the form:

$$f(x) = \exp[-\lambda_0 - \lambda_1 \ln(x-c) - \lambda_2 (x-c)^a] \tag{10.16}$$

where λ_0, λ_1, and λ_2 are Lagrange multipliers. Substitution of equation (10.16) in equation (10.5) yields

$$\int_c^\infty f(x)dx = \int_c^\infty \exp[-\lambda_0 - \lambda_1 \ln(x-c) - \lambda_2 (x-c)^a]dx = 1 \tag{10.17}$$

Equation (10.17) gives the partition function:

$$\begin{aligned} \exp(\lambda_0) &= \int_c^\infty \exp[\ln(x-c)^{-\lambda_1} - \lambda_2 (x-c)^a]dx \\ &= \int_c^\infty (x-c)^{-\lambda_1} \exp[-\lambda_2 (x-c)^a]\, dx \end{aligned} \tag{10.18}$$

Equation (10.18) can be simplified as follows: Let x-c = y. Then dy = dx. Hence, equation (10.18) becomes

$$\exp(\lambda_0) = \int_0^\infty y^{-\lambda_1} \exp[-\lambda_2 y^a]dy \tag{10.19}$$

Let $\lambda_2 y^a = z$. Then $\frac{dz}{dy} = \lambda_2 a y^{a-1}$; $dy = \frac{dz}{\lambda_2 a y^{a-1}}$; $y^a = \frac{z}{\lambda_2}$; and $y = (\frac{z}{\lambda_2})^{1/a}$

Substitution of the above quantities in equation (10.19) yields

$$\begin{aligned} \exp(\lambda_0) &= \int_0^\infty [(\frac{z}{\lambda_2})^{1/a}]^{-\lambda_1} e^{-z} \frac{dz}{\lambda_2 a [(\frac{z}{\lambda_2})^{1/a}]^{a-1}} \\ &= \int_0^\infty \frac{z^{-\lambda_1/a}}{\lambda_2^{-\lambda_1/a}} e^{-z} \frac{1}{\lambda_2 a} \frac{dz}{\frac{z^{(a-1)/a}}{\lambda_2^{(a-1)/a}}} \end{aligned} \tag{10.20}$$

Since

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx \tag{10.21}$$

equation (10.20) can be written as

$$\exp(\lambda_0) = \frac{1}{a} \int_0^\infty z^{(-(\lambda_1/a)-1+(1/a))} e^{-z} \lambda_2^{-[(-\lambda_1/a)-1+1+(1/a)]} dz$$

$$= \frac{1}{a} \frac{1}{((1-\lambda_1)/a)} \int_0^\infty z^{((1-\lambda_1/a)-1)} e^{-3} dz = \frac{\Gamma(\frac{1-\lambda_1}{a})}{a \lambda_2^{((1-\lambda_1)/a)}} \tag{10.22}$$

Taking logarithm of equation (10.22), we get the zeroth Lagrange multiplier:

$$\lambda_0 = \ln \Gamma(\frac{1-\lambda_1}{a}) - \ln a - (\frac{1-\lambda_1}{a}) \ln \lambda_2 \tag{10.23}$$

The zeroth Lagrange multiplier is also obtained from equation (10.18) as

$$\lambda_0 = \ln \int_c^\infty \exp[-\lambda_1 \ln(x-c) - \lambda_2 (x-c)^a] dx \tag{10.24}$$

10.1.3 RELATION BETWEEN LAGRANGE MULTIPLIERS AND CONSTRAINTS

Differentiating equation (10.24) with respect to λ_1 and λ_2, respectively, one obtains

$$\begin{aligned} \frac{\partial \lambda_0}{\partial \lambda_1} &= - \frac{\int_c^\infty \ln(x-c) \exp[-\lambda_1 \ln(x-c) - \lambda_2 (x-c)^a] dx}{\int_c^\infty \exp[-\lambda_1 \ln(x-c) - \lambda_2 (x-c)^a] dx} \\ &= - \int_c^\infty \ln(x-c) \exp[-\lambda_0 - \lambda_1 \ln(x-c) - \lambda_2 (x-c)^a] dx \\ &= - \int_c^\infty \ln(x-c) f(x) dx = -E[\ln(x-c)] \end{aligned} \tag{10.25}$$

$$\begin{aligned} \frac{\partial \lambda_0}{\partial \lambda_2} &= - \frac{\int_c^\infty (x-c)^a \exp[-\lambda_1 \ln(x-c) - \lambda_2 (x-c)^a] dx}{\int_c^\infty \exp[\lambda_1 \ln(x-c) - \lambda_2 (x-c)^a] dx} \\ &= - \int_c^\infty (x-c)^a \exp[\lambda_0 - \lambda_1 \ln(x-c) - \lambda_2 (x-c)^a] dx \\ &= - \int_c^\infty (x-c)^a f(x) dx = -E[(x-c)^a] = -(b-c)^a \end{aligned} \tag{10.26}$$

Differentiating equation (10.23) with respect to λ_2, one gets

$$\frac{\partial \lambda_0}{\partial \lambda_2} = -\left(\frac{1-\lambda_1}{a}\right)\frac{1}{\lambda_2} \tag{10.27}$$

Equating equations (10.26) and (10.27), we get

$$(b-c)^a = \frac{1-\lambda_1}{\lambda_2 a} \tag{10.28}$$

Differentiating equation (10.23) with respect to λ_1, we get

$$\frac{\partial \lambda_0}{\partial \lambda_1} = \frac{\partial}{\partial \lambda_1}\left[\ln \Gamma\left(\frac{1-\lambda_1}{a}\right)\right] = \frac{\ln \lambda_2}{a} \tag{10.29}$$

Equating equations (10.29) and (10.25), one obtains

$$\frac{\partial}{\partial \lambda_1}\left[\ln \Gamma\left(\frac{1-\lambda_1}{a}\right)\right] + \frac{\ln \lambda_2}{a} = -E[\ln(x-c)] \tag{10.30}$$

10.1.4 RELATION BETWEEN LAGRANGE MULTIPLIERS AND PARAMETERS

Substituting equation (10.23) in equation (10.16) yields

$$\begin{aligned} f(x) &= \exp\left[-\ln \Gamma\left(\frac{1-\lambda_1}{a}\right) + \ln a + \left(\frac{1-\lambda_1}{a}\right)\ln \lambda_2 \right. \\ &\qquad \left. - \lambda_1 \ln(x-c) - \lambda_2 (x-c)^a\right] \\ &= \exp\left[\ln\left(1/\left(\Gamma\left(\frac{1-\lambda_1}{a}\right)\right)\right) + \ln a + \ln \lambda_2^{((1-\lambda_1)/a)} \right. \\ &\qquad \left. + \ln(x-c)^{-\lambda_1} - \lambda_2 (x-c)^a\right] \\ &= \frac{1}{\Gamma\left(\frac{1-\lambda_1}{a}\right)} a \lambda_2^{(1-\lambda_1)/a} (x-c)^{-\lambda_1} \exp[-\lambda_2 (x-c)^a] \end{aligned} \tag{10.31}$$

A comparison of equation (10.31) with equation (10.1) shows that

$$\lambda_2 = \frac{1}{(b-c)^a} \tag{10.32}$$

10.1.5 RELATION BETWEEN PARAMETERS AND CONSTRAINTS

The EV III distribution has 3 parameters a, b, and c which are related to the Lagrange multipliers

$$\lambda_1 = 1-a \tag{10.33}$$

by equations (10.32) and (10.33) which, in turn, are related to the known constraints by equations (10.28) and (10.30). Eliminating the Lagrange multipliers between these two sets of equations we get equations which relate the parameters directly to the constraints as

$$(b-c)^a = E[(x-c)^a] \tag{10.34}$$

$$\psi(1) - \ln b = E[\ln(x-c)] \tag{10.35}$$

Since there are three parameters, equations (10.34) and (10.35) are not sufficient and another equation is needed. Recall that

$$\frac{\partial^2 \lambda_0}{\partial \lambda_1} = Var[\ln(x-c)] \tag{10.36a}$$

one obtains

$$\frac{\partial^2}{\partial \lambda_1^2}[\ln \Gamma(\frac{1-\lambda_1}{a})] = Var[\ln(x-c)] \tag{10.36b}$$

in which λ_1 =1-a, and Var[.] is the variance of the quantity [.].

10.1.6 DISTRIBUTION ENTROPY

The distribution entropy is given by equation (10.4) which is rewritten as

$$\begin{aligned} I(x) &= -\int_c^\infty f(x) \ln f(x) dx \\ &= [-\ln a + \ln(b-c) + (a-1)\ln(b-c)] \int_c^\infty f(x) dx \\ &\quad - (a-1)\int_c^\infty \ln(x-c) f(x) dx + \frac{1}{(b-c)^a}\int_c^\infty (x-c)^a f(x) dx \end{aligned} \tag{10.37}$$

For evaluating the last integral in equation (10.37), we write

$$W = \int_c^\infty (x-c)^a f(x)dx$$
$$= \int_c^\infty (x-c)^a \left(\frac{a}{b-c}\right) \left(\frac{x-c}{b-c}\right)^{a-1} \exp\left[-\left(\frac{x-c}{b-c}\right)^a\right]dx \tag{10.38}$$

Let $\left(\frac{x-c}{b-c}\right)^a = y$. Then

$$\frac{dy}{dx} = \frac{a(x-c)^{a-1}}{(b-c)^a} = \frac{a}{x-c}\left(\frac{x-c}{b-c}\right)^a \tag{10.39}$$

Therefore,

$$dx = \left(\frac{x-c}{a}\right)\left(\frac{x-c}{b-c}\right)^{-a} dy \tag{10.40}$$

$$W = \int_0^\infty (x-c)^a \left(\frac{a}{b-c}\right) \left(\frac{x-c}{b-c}\right)^{a-1} e^{-y} \left(\frac{x-c}{a}\right) \left(\frac{b-c}{x-c}\right)^a dx$$
$$= \int_0^\infty (x-c)^a \left(\frac{a}{b-c}\right) y \left(\frac{x-c}{b-c}\right) e^{-y} \left(\frac{x-c}{a}\right) \frac{1}{y} dy$$
$$= \int_0^\infty (x-c)^a e^{-y} dy = \int_0^\infty y(b-c)^a e^{-y} dy$$
$$= (b-c)^a \int_0^\infty y^{2-1} e^{-y} dy = (b-c)^a \Gamma(2) = (b-c)^a \tag{10.41}$$

Hence, the entropy function takes the form:

$$I(x) = 1n(b-c) + 1n(b-c)^{a-1} - 1n\ a - (a-1)\ E[1n(x-c)]$$
$$+ \frac{1}{(b-c)^a}(b-c)^a$$
$$= 1n\left\{\frac{(b-c)\,(b-c)^{a-1}}{a}\right\} - (a-1)\ E[1n(x-c)] + 1n\ e$$
$$= 1n\left\{\frac{(b-c)^a}{a} e\right\} - (a-1)\ E[1n(x-c)] \tag{10.42}$$

10.2 Parameter - Space Expansion Method

10.2.1 SPECIFICATION OF CONSTRAINTS

Following Singh and Rajagopal (1986), the constraints consistent with the POME method and appropriate for the EV III distribution are equation (10.5) and

$$\int_c^\infty \left(\frac{x-c}{b-c}\right)^a f(x)\,dx = E\left[\left(\frac{x-c}{b-c}\right)^a\right] \quad (10.43)$$

$$\int_c^\infty \ln\left(\frac{x-c}{b-c}\right) f(x)\,dx = E\left[\ln\left(\frac{x-c}{b-c}\right)\right] \quad (10.44)$$

10.2.2 DERIVATION OF ENTROPY FUNCTION

The least-biased pdf corresponding to POME and consistent with equations (10.5), (10.43) and (10.44) takes the form:

$$f(x) = \exp(\lambda_0)\exp\left[-\lambda_1\left(\frac{x-c}{b-c}\right)^a\right]\left(\frac{x-c}{b-c}\right)^{\lambda_2} \quad (10.45)$$

where λ_0, λ_1 and λ_2 are Lagrange multipliers corresponding to the normality condition in equation (10.5) and the constraints in equations (10.43) and (10.44).

Using equation (10.45) in the definition of the total probability, one obtains

$$1 = \exp(\lambda_0)\int_c^\infty \exp\left[-\lambda_1\left(\frac{x-c}{b-c}\right)^a\right]\left(\frac{x-c}{b-c}\right)^{\lambda_2} dx \quad (10.46)$$

Equation (10.46) yields the partition function as

$$\exp(-\lambda_0) = \int_c^\infty \exp\left[-\lambda_1\left(\frac{x-c}{b-c}\right)^a\right]\left(\frac{x-c}{b-c}\right)^{\lambda_2} dx \quad (10.47)$$

Let $y = [(x-c)/(b-c)]^a$. Then $dx = a^{-1}(b-c)\,y^{(1-a)/a}\,dy$, and equation (10.47) becomes

$$\exp(-\lambda_0) = \left(\frac{b-c}{a}\right)\Gamma\left(\frac{\lambda_2+1}{a}\right)/\lambda_1^{(\lambda_2+1)/a} \quad (10.48)$$

Inserting equation (10.48) into equation (10.45), one obtains

$$f(x) = \exp\left[-\lambda_1\left(\frac{x-c}{b-c}\right)^a\right]\left(\frac{x-c}{b-c}\right)^{\lambda_2}\left(\frac{a}{b-c}\right)\frac{\lambda_1^{(\lambda_2+1)/a}}{\Gamma\left(\frac{\lambda_2+1}{a}\right)} \quad (10.49)$$

Equation (10.49) can specialize into the following distributions for appropriate values of a, c, λ_1 and λ_2:

(1) $a = 1, c = 0, \lambda_1 = 1, \lambda_2 = 0$	exponential
(2) $a = 1, c = 0, \lambda_1 = 1, \lambda_2 = d$	Gamma
(3) $a = 1, \lambda_1 = 1, \lambda_2 = d$	Pearson type III

(4) $\lambda_1 = 1, \lambda_2 = a_1 - 1$ EV III
(5) $a = a_1, c = 0, \lambda_1 = 1, \lambda_2 = a_1 - 1$ Weibull
(6) $a = 2, c = 0, \lambda_1 = 1, \lambda_2 = 0$ (truncated) normal

To derive parameters a, b and c for the general formulation in equation (10.49), the entropy function in equation (10.4) is expressed as

$$I[f] = \ln(\frac{b-c}{a}) + \ln\Gamma(\frac{\lambda_2+1}{a}) - (\frac{1+\lambda_2}{a})\ln\lambda_1 + \lambda_1 E[(\frac{x-c}{b-c})^a]$$
$$-\lambda_2 E[\ln(\frac{x-c}{b-c})] \tag{10.50}$$

Taking partial derivatives of equation (10.50) with respect to λ_1, λ_2, a, b and c separately and equating each derivative to zero, respectively, yields:

$$\frac{\partial I}{\partial \lambda_1} = 0 = -\frac{\lambda_2+1}{a\lambda_1} + E[(\frac{x-c}{b-c})^a] \tag{10.51}$$

or

$$\frac{\lambda_2+1}{a\lambda_1} = E[(\frac{x-c}{b-c})^a] \tag{10.52}$$

$$\frac{\partial I}{\partial \lambda_2} = 0 = \frac{1}{a}\psi(\frac{\lambda_2+1}{a}) - \frac{1}{a}\ln\lambda_1 - E[\ln(\frac{x-c}{b-c})] \tag{10.53}$$

or

$$\frac{1}{a}\psi(\frac{\lambda_2+1}{a}) - \frac{1}{a}\ln\lambda_1 = E[\ln(\frac{x-c}{b-c}] \tag{10.54}$$

$$\frac{\partial I}{\partial a} = 0 = -\frac{1}{a} - (\frac{\lambda_2+1}{a^2})\psi(\frac{\lambda_2+1}{a}) + (\frac{\lambda_2+1}{a^2})\ln\lambda_1$$
$$+ \lambda_1 E[(\frac{x-c}{b-c})^a \ln(\frac{x-c}{b-c})] \tag{10.55}$$

$$\frac{\partial I}{\partial b} = 0 = \frac{1}{b-c} - \frac{\lambda_1 a}{b-c} E[(\frac{x-c}{b-c})^a] + \lambda_2 \frac{1}{b-c} \tag{10.56}$$

or

$$\lambda_2 + 1 = \lambda_1 a E[(\frac{x-c}{b-c})]^a \tag{10.57}$$

which is the same as equation (10.52).

$$\frac{\partial I}{\partial c} = 0 = -\frac{1}{b-c} - \frac{\lambda_1 a}{b-c} E[(\frac{x-c}{b-c})^{a-1}] + \frac{\lambda_1 a}{b-c} [(\frac{x-c}{b-c})^a]$$

$$+ \lambda_2 E[\frac{1}{x-c}] - \frac{\lambda_2}{b-c} \tag{10.58}$$

Equation (10.54) is simplified using equation (10.54) as

$$-\frac{1}{a} - (\frac{\lambda_2 + 1}{a}) E[\ln(\frac{x-c}{b-c})] + \lambda_1 E[(\frac{x-c}{b-c})^a \ln(\frac{x-c}{b-c})] = 0 \tag{10.59}$$

Using equation (10.52) in equation (10.59), we get

$$-\frac{1}{a} - \lambda_1 E[(\frac{x-c}{b-c})^a] E[\ln(\frac{x-c}{b-c})] + \lambda_1 E[(\frac{x-c}{b-c})^a \ln(\frac{x-c}{b-c})] = 0 \tag{10.60}$$

Rewriting equation (10.58), we get

$$-\frac{1+\lambda_2}{b-c} + \frac{\lambda_1 a}{b-c} E[(\frac{x-c}{b-c})^a] - \frac{\lambda_1 a}{b-c} E[(\frac{x-c}{b-c})^{a-1}] + \lambda_2 E[\frac{1}{x-c}] = 0 \tag{10.61}$$

The first two terms of equation (10.61) vanish as the result of equation (10.57); therefore,

$$\frac{\lambda_1 a}{b-c} E[(\frac{x-c}{b-c})^{a-1}] = \lambda_2 E[\frac{1}{x-c}] \tag{10.62}$$

Multiplying equations (10.52) and (10.62) and simplifying, we get

$$\frac{\lambda_2 + 1}{b-c} = E[(\frac{x-c}{b-c})^{a-1}] = \lambda_2 E[\frac{1}{x-c}] E[(\frac{x-c}{b-c})^a] \tag{10.63}$$

or

$$\frac{\lambda_2 + 1}{\lambda_2} = \frac{E[\frac{1}{x-c}] E[(x-c)^a]}{E[(x-c)^{a-1}]} \tag{10.64}$$

Equations (10.52), (10.54), (10.60) and (10.64) are utilized to estimate the parameters of the distributions outlined earlier and discussed below.

10.2.2.1 Case 1: Exponential Distribution: The exponential distribution can be obtained from equation (10.49) with a = 1, c = 0, $\lambda_1 = 1$ and $\lambda_2 = 0$ as

$$f(x) = b \exp(-x/b) \tag{10.65}$$

Then, one obtains from equations (10.52), (10.54), (10.60) and (10.64) respectively:

$$1 = E(x/b) \quad \text{or} \quad b = E[x] \tag{10.66}$$

$$\psi(1) = E[\ln(x/b)] \quad (10.67)$$

$$1 = E[(x/b)\ln(x/b)] - E[(x/b)]\ E[\ln(x/b)] \quad (10.68)$$

and equation (10.64) does not exist. Equation (10.67) and (10.68) are identities for the exponential distribution, and equation (10.66) is the equation for estimation of parameter b.

10.2.2.2 Case 2: Gamma Distribution: This distribution can be obtained from equation (10.49) with a = 1, c = 0, $\lambda_1 = 1$ and $\lambda_2 = d$ as

$$f(x) = \frac{1}{\Gamma(d+1)} \left(\frac{1}{b}\right) \left(\frac{x}{b}\right)^d \exp\left(-\frac{x}{b}\right) \quad (10.69)$$

Then, from equations (10.52), (10.53), (10.60) and (10.64), respectively, one obtains:

$$d + 1 = E\left[\frac{x}{b}\right] \quad (10.70)$$

$$\psi(d+1) = E\left[\ln\left(\frac{x}{b}\right)\right] \quad (10.71)$$

$$1 = E\left[\left(\frac{x}{b}\right)\ln\left(\frac{x}{b}\right) - E\left[\frac{x}{b}\right] E\left[\ln\left(\frac{x}{b}\right)\right]\right. \quad (10.72)$$

$$\frac{1}{D} = E\left[\frac{b}{x}\right] \quad (10.73)$$

Equation (10.72) is an identity, and in equation (10.73) E [1/ x] does not exist for negative integers less than 1. Equations (10.70) and (10.71) are the equations for estimation of b and d.

10.2.2.3 Case 3: Pearson Type (PT) III Distribution: By inserting a = 1, $\lambda_1 = 1$ and $\lambda_2 = d$ into equation (10.52), the PT III distribution can be written as

$$f(x) = \frac{1}{\Gamma(d+1)} \left(\frac{1}{b-c}\right) \left(\frac{x-c}{b-c}\right)^d \exp\left(-\frac{x-c}{b-c}\right) \quad (10.74)$$

From equations (10.52), (10.53), (10.60), and (10.64), respectively, one obtains

$$d + 1 = E\left[\frac{x-c}{b-c}\right] \quad (10.75)$$

$$\psi(d+1) = E\left[\ln\left(\frac{x-c}{b-c}\right)\right] \quad (10.76)$$

$$1 = E\left[\left(\frac{x-c}{b-c}\right)\ln\left(\frac{x-c}{b-c}\right)\right] - E\left[\frac{x-c}{b-c}\right] E\left[\ln\frac{x-c}{b-c}\right] \quad (10.77)$$

$$\frac{1}{x} = E\left[\frac{b-c}{x-c}\right] \quad (10.78)$$

Equation (10.77) is an identity, and equations (10.75), (10.76) and (10.78) are the equations for estimation of parameters a, b and c.

10.2.2.4 Case 4: Truncated Normal Distribution: If a = 2, c = 0, $\lambda_1 = 1$ and $\lambda_2 = 0$ are inserted into equation (10.49), a truncated normal distribution is obtained:

$$f(x) = \frac{2}{b\sqrt{\pi}} \exp(-(x/b)^2) \tag{10.79}$$

From equations (10.52), (10.54), (10.60), and (10.64), one obtains respectively:

$$\frac{1}{2} = E[(x/b)^2] \tag{10.80}$$

$$\frac{1}{2}\psi(\frac{1}{2}) = E[\ln(\frac{x}{b})] \tag{10.81}$$

$$-\frac{1}{2} - E[(x/b)^2]\, E[\ln(\frac{x}{b})] + E[(\frac{x}{b})^2 \ln(\frac{x}{b})] = 0 \tag{10.82}$$

Equations (10.64) does not exist, and equations (10.81) and (10.82) are identities. Thus, equation (10.80) is the equation for estimation of parameter b.

10.2.3 RELATION BETWEEN PARAMETERS AND CONSTRAINTS

Substitution of $\lambda_1 = 1$ and $\lambda_2 = a - 1$ in equation (10.47) results in the EV III distribution:

$$f(x) = \frac{a}{b-c}(\frac{x-c}{b-c})^{a-1} \exp[-(\frac{x-c}{b-c})^a] \tag{10.83}$$

Then, one obtains from equations (10.52), (10.53), (10.60) and (10.64) respectively:

$$1 = E[(\frac{x-c}{b-c})^a] \tag{10.84}$$

$$\frac{\psi(1)}{a} = E[\ln(\frac{x-c}{b-c})] \tag{10.85}$$

$$\frac{1}{a} = E[(\frac{x-c}{b-c})^a \ln(\frac{x-c}{b-c})] - E[(\frac{x-c}{b-c})^a]\, E[\ln(\frac{x-c}{b-c})] \tag{10.86}$$

$$\frac{a}{a-1} = \frac{E[\frac{1}{x-c}]\, E[(x-c)^a]}{E[(x-c)^{a-1}]} \tag{10.87}$$

The parameters a, b and c are estimated from equations (10.84), (10.85) and (10.87).

10.3 Other Methods of Parameter Estimation

The methods of moments (MOM) and maximum likelihood estimation (MLE) are briefly outlined below.

10.3.1 METHOD OF MOMENTS

The EV III distribution has three parameters so three moments are needed. The r-th moment about the lower bound c of the EV III distribution can be written as

$$M_r^c = (b-c)^r \, \Gamma\,[\,(a+r)/c\,], \quad r = 1, 2, \ldots . \tag{10.88}$$

These moments about c can be converted to the moments about the origin M_r^0 or moments about the centroid M_μ^r. To determine the parameters a, b and c, the first three moments can be specified as

$$M_1^0 = c(b-c)\,\Gamma[\frac{a+1}{a}] = \mu \tag{10.89}$$

$$M_2^\mu = \sigma_x^2 = (b-c)\,[\Gamma(\frac{a+2}{a}) - \Gamma^2(\frac{1+a}{a})] \tag{10.90}$$

$$M_3^\mu = (b-c)^3\,[\Gamma(\frac{a+3}{a}) - 3\Gamma(\frac{a+2}{a})\,\Gamma(\frac{a+1}{a}) + 2\Gamma^3(\frac{a+1}{a})] \tag{10.91}$$

where μ is the centrod and σ_x^2 is the variance of X. Equations (10.89)-(10.91) are solved iteratively to estimate parameters a, b, and c.

10.3.2 METHOD OF MAXIMUM LIKELIHOOD ESTIMATION

For the maximum likelihood estimation (MLE) method the parameter estimation equations are:

$$(a-1)\sum_{i=1}^{n}\frac{1}{(x_i-c)} = \frac{n\,a\sum_{i=1}^{n}(x_i-c)^{a-1}}{\sum_{i=1}^{n}(x_i-c)^a} \tag{10.92}$$

$$n + a\sum_{i=1}^{n}\ln(x_i-c) = \frac{n\,a\sum_{i=1}^{n}(x_i-c)^{a-1}\ln(x_i-c)}{\sum_{i=1}^{n}(x_i-c)^a} \tag{10.93}$$

$$(b-c)^a = \frac{1}{n}\sum_{i=1}^{n}(x_i-c)^a \tag{10.94}$$

A comparison of POME and MLE methods shows that equation (10.94) is equivalent to equation (10.84), equation (10.93) equivalent to equation (10.86) and equation (10.92) equivalent to equation (10.87). Thus, these two methods would yield comparable parameter estimates.

10.4 Comparative Evaluation of Estimation Methods

Singh (1987) compared MOM, MLE and POME methods of parameter estimation using annual minimum 7-day average flows for the Shoal Creek and Buffalo River in Tennessee, U.S.A. These data were for the period 1926 to 1969 and are given by Riggs (1972). Parameters a, b and c were estimated using the three methods were obtained and are given below:

Method	Shoal Creek a	b	c	Buffalo Creek a	b	c
MOM	1.672	111.77	60.55	2.105	149.97	75.27
MLE	2.25	113.98	47.50	2.75	151.85	57.50
POME	2.20	112.83	38.00	2.75	151.16	50.00

Table 10.1 Comparison of three methods of fitting the EV III distribution to annual 7-day low flow data of the Shoal Creek, Tennessee, U.S.A. Low flow values are computed for various return periods.

Return Period	Observed flow	Computed flow (cfs)					
T	(cfs)	POME		MLE		MOM	
(Years)		Computed	Error (%)	Computed	Error (%)	Computed	Error (%)
1.05	62.5	64	2.4	68	8.8	70	10.7
1.1	72	69	4.2	72	0	72	0
1.2	82	76	7.3	80	2.4	78	4.9
1.3	84	81	3.6	81	3.6	78	7.1
1.4	88	85	3.4	89	6.3	86	2.3
1.5	95	90	5.3	92	9.8	90	5.3
2	102	102	0	105	2.9	101	1.0
3	120	118	1.6	118	1.7	115	4.2
5	127	131	3.1	129	1.6	128	0.8
10	138	150	8.7	145	5.0	148	7.3

Using the parameter values, the EV III distribution was fitted to the 7-day flow data of the Shoal Creek and Buffalo River as shown in Figures 10.1 and 10.2. For various return periods, low flows were computed using the three methods for both rivers as shown in Tables 10.1 and 10.2. For the Buffalo River, the MLE and POME methods were almost the same and represented the EV III distribution reasonably well. MOM was also comparable. In case of the Shoal Creek, MLE and POME methods were closer for discharges exceeding 75 cfs; for lower discharges, MOM and MLE were closer. However, the differences between the three methods were marginal and were therefore considered comparable. It must be pointed out that these calculations offer no information on the sampling properties of the POME method which can be best accomplished using Monte Carlo simulations. The results show that the parameter estimates yielded by the principle of maximum entropy were comparable to those yielded by the methods of moments and maximum likelihood estimation.

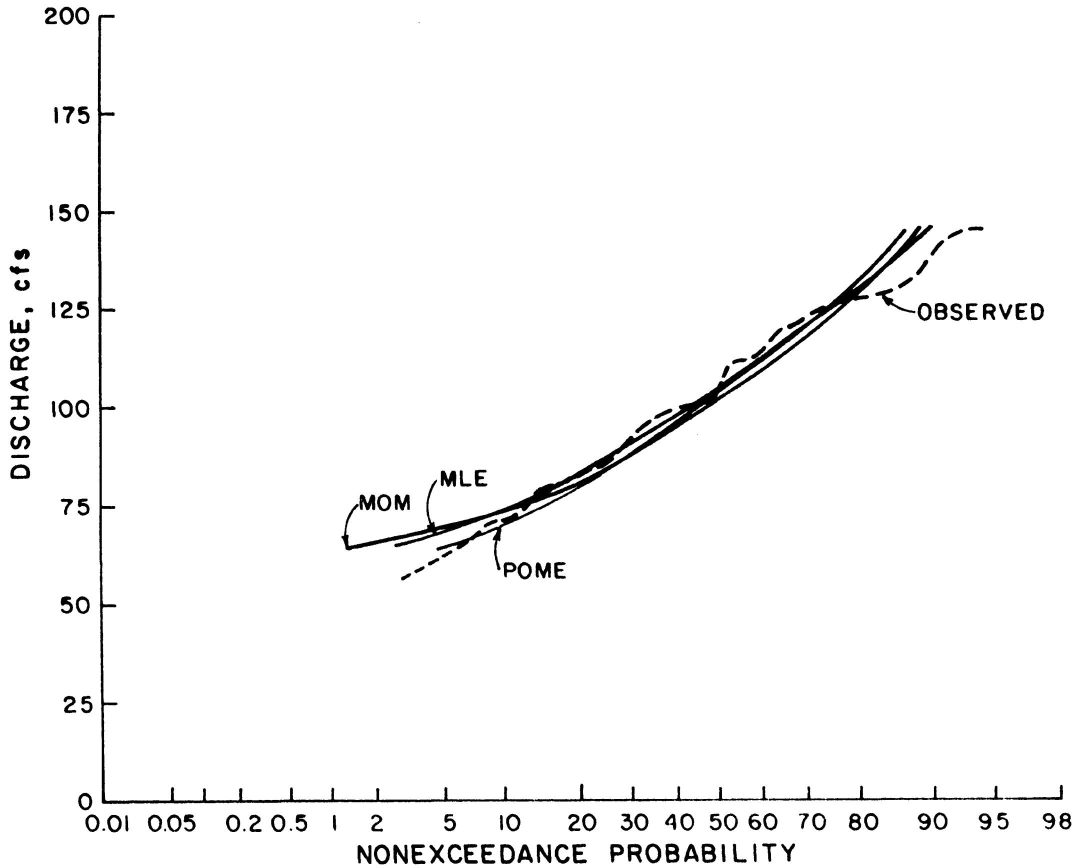

Figure 10.1 Comparison of the MOM, MLE and POME methods of fitting the EV III distribution to annual 7-day low flow data of the Shoal Creek, Tennessee, U.S.A.

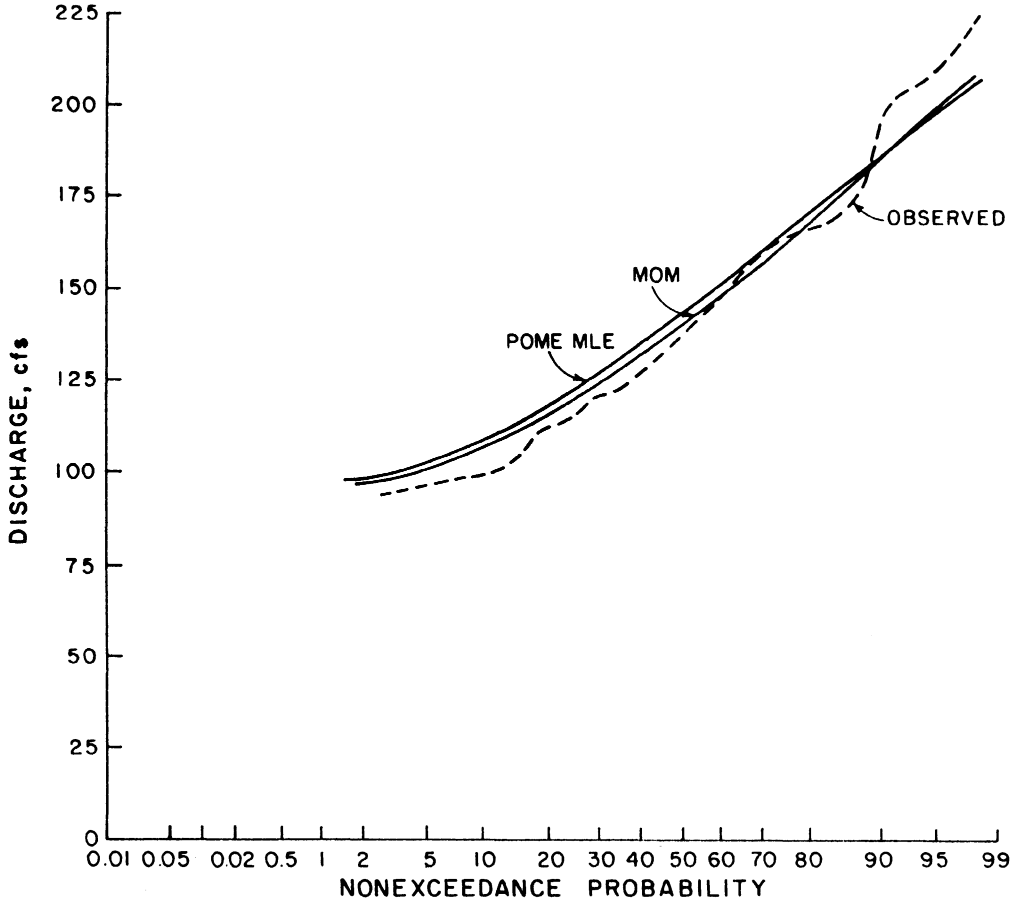

Figure 10.2 Comparison of the MOM, MLE and POME methods of fitting the EV III distribution to annual 7-day low flow data of the Buffalo River, Tennessee, U.S.A.

Table 10.2 Comparison of three methods of fitting the EV III distribution to annual 7-day low flow data of the Buffalo River, Tennessee, U.S.A. Low flow values are computed for various return periods.

Return Period	Observed Flow	Computed Flow					
T	(cfs)	POME		MLE		MOM	
(Years)		Computed	Error (%)	Computed	Error (%)	Computed	Error (%)
1.05	96	102	6.25	102	6.25	100	4.2
1.1	99	107.5	8.6	107.5	8.6	107	8.1
1.2	105	113	7.6	113	7.6	112	6.6
1.3	114	120	5.3	120	5.3	118	3.5
1.4	118	124	5.1	124	5.1	121	2.5
1.5	122	129	5.7	129	5.7	125	2.4
2	138	142	2.9	142	2.9	139	0.7
5	167	170	1.8	170	1.8	168	0.5
10	190	185	2.6	185	2.6	186	2.1
20	209	197	5.7	197	5.7	199	4.8
30	213	200	6.1	200	6.1	202	5.2
40	222	202	9.0	202	9.0	207	6.8

References

Condie, R. and Nix, G. (1975). Modeling of low flow frequency distributions and parameter estimation. *Proceedings, International IWRA Symposium on Water for Arid Lands*, Tehran, Iran.

Gumbel, E.J., 1963. Statistical forecast of droughts. *IAHS Bulletin*, Vol. **8,** No. 1, pp. 5-23.

Loganathan, G.V., Kuo, C.Y. and McCormick, T.C., 1985. Frequency analysis of low flows. *Nordic Hydrology*, Vol. **16**, pp. 105-128.

Matalas, N.C., 1963. Probability distribution of low flows. *Geological Survey Professional Paper* **434-A**, pp. A27, U.S. Geological Survey, Washington, D.C.

Otten, A. and Van Montfort, M.A.J., 1978. The power of two tests on the type of distributions of extremes. *Journal of Hydrology*, Vol. **37**, pp. 195-199.

Riggs, H.C., 1972. Low-flow investigations. *U.S. Geological Survey Techniques for Water Resources Investigations*, Book **4,** Chapter B1, Washington, D.C.

Singh, V.P. and Rajagopal, A.K., 1986. A new method of parameter estimation for hydrologic frequency analysis. *Hydrological Science and Technology*, Vol. **2**, No. 3, pp. 33-40.

Singh, V.P., Rajagopal, A.K. and Singh, K., 1986. Derivation of some frequency distributions using the principle of maximum entropy. *Advances in Water Resources*, Vol. **9**, pp. 91-106.

Singh, V.P., Singh, K. and Rajagopal, A.K., 1985. Application of the principle of maximum entropy (POME) to hydrologic frequency analysis. *Completion Report* **06**, 144 p., Louisiana Water Resources Research Institute, Louisiana State University, Baton Rouge, Baton Rouge, Louisiana.

CHAPTER 11

GENERALIZED EXTREME VALUE DISTRIBUTION

The generalized extreme-value (GEV) distribution was introduced by Jenkinson (1955, 1969) and recommended by Natural Environment Research Council (1975) of Great Britain. The GEV distribution is the most widely accepted distribution for describing flood frequency data from the United Kingdom (Sinclair and Ahmad, 1988) and has also become popular elsewhere (Otten and van Montfort, 1980; Prescott and Walden, 1980, 1983; Turkman, 1985; Hosking et al., 1985; Arnell et al., 1986). Sinclair and Ahmad (1988) introduced location-invariance in the context of using plotting positions in estimating parameters of the GEV distribution by the method of probability-weighted moments. They emphasized that this was an important factor in the selection of an appropriate plotting position, for otherwise the estimate of the shape parameter might not be independent of location. Tawn (1988) presented a method of filtering the original time series containing dependent data to obtain independent extremes. He then used the limiting joint generalized extreme value distribution for the r largest order statistics.

During the last two dacdes, the significance of using nonsystematic data in flood frequency analysis has been reognized. Nonsystematic data are the historical flood data recorded before the beginning of the systematic period and the paleo data resulting from the analysis of certain proxy data. Both these types of data contain information beyond that of the systematic period. Historical flood information prior to the systematic period is collected from high water marks left by extreme floods, writen accounts in news papers and books, damage reports and repair reports prepared by insurance companies and government agencies, unpublished written records, and verbal communications from the general public. Paleo data are generally collected from the botanical evidence left by past floods through corrosion scars, adventitious sprouts, ring anomalies, vegetation age distribution, etc. Both types of data provide information in various forms such as the date and magnitude of one or more floods greater than a certain threshold value. Together they provide perhaps the most accurate information on the magnitude and frequency of extreme floods occurring prior to the systematic period. Frances et al. (1994) considered flood frequency analysis with systematic and historical or paleoflood data based on the two-parameter GEV distribution. They found the value of historical and paleoflood data to depend on (1) the relative magnitudes of systematic period and historical period, (2) the return period of the flood quantile of interest, and (3) the return period of the threshold level of perception.

The GEV distribution has three parameters. The methods of moments (MOM), probability weighted moments (PWM), L-moments (LMOM), LH-moments (LHMOM), and maximum likelihood estimation (MLE) are some of the popular methods for estimation of GEV parameters. Hosking (1986) and Hosking et al. (1986) described the theory of PWMs and derived GEV parameters in terms of PWMs. The PWM method has since been a very popular method of parameter estimation. Haktanir (1996) proposed a modification to the conventional

V. P. Singh, *Entropy-Based Parameter Estimation in Hydrology*

PWM method wherein PWMs are computed from the probability of each element in the sample series using the distribution itself instead of a plotting position formula. Wang (1996) described a method of partial PWMs for fitting disributions to censored data. Hosking (1986, 1990) developed the theory of L-moments and used the L-moment ratio diagram to identify underlying parent distribution and L-moment ratios for testing hypotheses about forms of probability distributions. Hosking and Wallis (1991) extended the application of L-moments and derived statistics to measure discordancy, regional homogeneity, and goodness of fit statistics needed in regional frequency analyses. Using L-moments Rao and Hamed (1994) recommended 3-parameter log-normal distribution and GEV distribution for frequency analysis of data in the Upper Cauvery River basin in India. They extended the application of L-moments to regional frequency analysis of Wabash River flood data and recommended GEV as the regional parent distribution. Wang (1997) derived LH moments, a generalization of L-moments for frequency analysis of large return period events. He argued that LH moments reduced undesirable influences of small events on estimation of large return period events. Vogel and Fennessey (1993) suggested L-moment diagrams to replace product moment diagrams, for the latter exhibit substantial bias and variance for small samples. Vogel and Wilson (1996) constructed L-moment diagrams for annual maximum, average, and maximum streamflows at more than 1,455 river basins in the United States. Goodness-of-fit comparisons revealed that GEV, 3-parameter lognormal and log-Pearson type 3 distributions provided good approximations to the distribution of annual maximum flood flows. Otten and van Montfort (1980) modified the procedure of Jenkinson (1955) and estimated the GEV parameters using MLE. Phien and Emma (1989) employed the MLE method to estimate the GEV parameters and quantiles for censored samples.

A random variable X is said to have a generalized extreme value distribution if its probability density function (pdf) is given by

$$f(x) = \frac{1}{a}[1-\frac{b}{a}(x-c)]^{\frac{(1-b)}{b}} \exp[-(1-\frac{b}{a}(x-c))]^{\frac{1}{b}} \tag{11.1}$$

where $a > 0$ and c are respectively the scale and location parameters, and b is a shape parameter. The range of X depends on the value of b: it is bounded by $c+(a/b)$ from above for $b > 0$, i.e., $-\infty < x \le c+(a/b)$; and it is bounded from below for $b < 0$, i.e., $c+(a/b) \le x \le \infty$. The shape parameter b determines which extreme value distribution is represented. Depending on the value of b, equation (11.1) corresponds to the Fisher-Tippett distribution types I, II, and III: the Gumbel distribution (extreme value type I) for $b = 0$, the extreme value type II distribution for $b < 0$, and the extreme value type III distribution for $b > 0$. For $b = 2$, equation (11.1) gives rise to a reverse Raleigh distribution and for $b=1$ it becomes a reverse exponential distribution. It can also be shown that the Weibull dstribution is a reverse GEV distribution.

The cumulative distribution function (cdf) of the GEV distribution can be expressed as

$$F(x) = \exp[-(1-\frac{b}{a}(x-c))^{\frac{1}{b}}] \tag{11.2}$$

Sometimes equation (11.1) is also expressed as

$$f(x) = \frac{1}{a(1-z)} \exp[-y-\exp(-y)] \tag{11.3}$$

where

$$y = -\frac{1}{b}\ln(1-z) \tag{11.4}$$

$$z = \frac{b}{a}(x-c) \tag{11.5}$$

11.1 Ordinary Entropy Method

11.1.1 SPECIFICATION OF CONSTRAINTS

Taking logarithm to the base 'e' of equation (11.1), one gets

$$\ln f(x) = -\ln(a) + \frac{(1-b)}{b}\ln\left[1-\frac{b}{a}(x-c)\right] - \left[1-\frac{b}{a}(x-c)\right]^{\frac{1}{b}} \tag{11.6}$$

Multiplying equation (11.6) by [- f (x)] and integrating, the result is the entropy function, I(x), of the GEV distribution:

$$-\int f(x)\ln f(x)dx = -\int\left[-\ln a + \frac{(1-b)}{b}\ln\left(1-\frac{b}{a}(x-c)\right)\right]f(x)\,dx - \int\left[1-\frac{b}{a}(x-c)\right]^{\frac{1}{b}} f(x)dx \tag{11.7}$$

The constraints appropriate for equation (11.1), consistent with POME, are derived from equation (11.7) as

$$\int f(x)\,dx = 1 \tag{11.8}$$

$$-\int \ln\left[1-\frac{b}{a}(x-c)\right]f(x)\,dx = -E\left[\ln\left(1-\frac{b}{a}(x-c)\right)\right] \tag{11.9}$$

$$\int \left[1-\frac{b}{a}(x-c)\right]^{1/b} f(x)\,dx = E\left[1-\frac{b}{a}(x-c)\right]^{1/b} \tag{11.10}$$

11.1.2 CONSTRUCTION OF ZEROTH LAGRANGE MULTIPLIER

The least-biased pdf, based on POME, and consistent with equations (11.8)-(11.10), takes the form:

$$f(x) = \exp\{-\lambda_0 - \lambda_1 \ln[1 - \frac{b}{a}(x-c)] - \lambda_2 [1 - \frac{b}{a}(x-c)]^{1/b}\} \quad (11.11)$$

Substitution of equation (11.11) in equation (11.8) yields the partition function:

$$\exp(\lambda_0) = \int [1 - \frac{b}{a}(x-c)]^{\lambda_1} \exp\{-\lambda_2 [1 - \frac{b}{a}(x-c)]^{1/b}\} dx \quad (11.12)$$

Equation (11.12) can be simplified and expressed as

$$\exp(\lambda_0) = a\, \lambda_2^{\,b(\lambda_1+1)}\, \Gamma(b(\lambda_1 + 1) \quad (11.13)$$

Taking logarithm of equation (11.13) gives the zeroth Lagrange multiplier:

$$\lambda_0 = \ln a + b(\lambda_1 + 1) \ln \lambda_2 + \ln \Gamma[b(\lambda_1 + 1)] \quad (11.14)$$

The zeroth Lagrange multiplier is also obtained from equation (11.12) as

$$\lambda_0 = \ln \int \exp\{-\lambda_1 \ln[1 - \frac{b}{a}(x-c)] - \lambda_2 [1 - \frac{b}{a}(x-c)]^{1/b}\} dx \quad (11.15)$$

11.1.3 RELATION BETWEEN LAGRANGE MULTIPLIERS AND CONSTRAINTS

Differentiating equation (11.15) with respect to λ_1 and λ_2, respectively, one gets

$$\frac{\partial \lambda_0}{\partial \lambda_1} = \frac{\int \ln[1 - \frac{b}{a}(x-c)] f(x)\, dx}{\int f(x)\, dx} = E[\ln\{1 - \frac{b}{a}(x-c)\}] \quad (11.16)$$

$$\frac{\partial \lambda_0}{\partial \lambda_2} = \frac{\int [1 - \frac{b}{a}(x-c)]^{1/b} f(x) dx}{\int f(x)\, dx} = E[\{1 - \frac{b}{a}(x-c)\}^{1/b}] \quad (11.17)$$

where f(x) is given by equation (11. 11). Also, differentiating equation (11.14) we get

$$\frac{\partial \lambda_0}{\partial \lambda_1} = b \ln \lambda_2 + b\, \Psi(k), \quad k = b(\lambda_1 - 1) \quad (11.18)$$

$$\frac{\partial \lambda_0}{\partial \lambda_2} = b(\lambda_1 - 1) / \lambda_2 \quad (11.19)$$

Equating equation (11.16) to equation (11.18), we obtain

$$E[\ln\{1-\frac{b}{a}(x-c)\}]=b\ln\lambda_2+b\Psi(k) \tag{11.20}$$

$$E[\{1-\frac{b}{a}(x-c)\}^{1/b}]=b(1+\lambda_1)/\lambda_2 \tag{11.21}$$

Because the GEV distribution has three parameters, another equation is needed. This equation is obtained by recalling that

$$\frac{\partial^2\lambda_0}{\partial\lambda_2^2}=Var[\{1-\frac{b}{a}(x-c)\}^{1/b}] \tag{11.22}$$

Differentiating equation (11.19) with respect to λ_2, we get

$$\frac{\partial^2\lambda_0}{\partial\lambda_2^2}=-b(\lambda_1-1)/\lambda_2^2 \tag{11.23}$$

Equating equation (11.22) to equation (11.23), we get

$$Var[\{1-\frac{b}{a}(x-c)\}^{1/b}]=-b(\lambda_1-1)/\lambda_2^2 \tag{11.24}$$

11.1.4 RELATION BETWEEN LAGRANGE MULTIPLIERS AND PARAMETERS

Substituting equation (11.14) in equation (11.11), we get

$$f(x)=\frac{1}{a}(\lambda_2)^{b(\lambda_1+1)}\frac{1}{\Gamma(b[(\lambda_1+1)]}[1-\frac{b}{a}(x-c)]^{\lambda_1} \times\exp\{-\lambda_2[1-\frac{b}{a}(x-c)]^{1/b}\} \tag{11.25}$$

A comparison of equation (11.25) with equation (11.1) shows that $\lambda_1=(1-b)/b$ and $\lambda_2=1$.

11.1.5 RELATION BETWEEN PARAMETERS AND CONSTRAINTS

The relation between Lagrange multipliers and constraints is given by equations (11.20), (11.21) and (11.22) and that between Lagrange multipliers and parameters by equation (11.25). Eliminating the Lagrange multipliers between these two sets of equations yields the relation between parameters and constriants. Therefore, we obtain

$$E[1-\frac{b}{a}(x-c)]^{1/b}=1 \tag{11.26}$$

$$E[\ln\{1-\frac{b}{a}(x-c)\}]=b\Psi(k) \tag{11.27}$$

$$Var\,[1-\frac{b}{a}(x-c)]^{1/b}=1 \tag{11.28}$$

11. 2 Parameter-Space Expansion Method

11.2.1 SPECIFICATION OF CONSTRAINTS

The constraints for this method are specified by equation (11.8) and

$$\int \frac{b-1}{b}\ln[1-\frac{b}{a}(x-c)]f(x)dx = E[\frac{b-1}{b}\ln(1-\frac{b}{a}(x-c))] \tag{11.29}$$

$$\int [1-\frac{b}{a}(x-c)]^{\frac{1}{b}}f(x)\,dx = E[1-\frac{b}{a}(x-c)]^{\frac{1}{b}} \tag{11.30}$$

11.2.2 CONSTRUCTION OF ZEROTH LAGRANGE MULTIPLIER

The least biased-pdf, f(x), consistent with equations (11.8) and (11.29)-(11.30) and based on POME takes the form:

$$f(x) = \exp[-\lambda_0-\lambda_1\ln[1-\frac{b}{a}(x-c)]^{\frac{b-1}{b}}-\lambda_2[1-\frac{b}{a}(x-c)]^{\frac{1}{b}}] \tag{11.31}$$

where λ_0, λ_1, and λ_0 are Lagrange multipliers. Substituting equation (11.31) in equation (11.8) yields

$$\int \exp[-\lambda_0-\lambda_1\ln[1-\frac{b}{a}(x-c)]^{\frac{b-1}{b}}-\lambda_2[1-\frac{b}{a}(x-c)]^{\frac{1}{b}}]\,dx = 1 \tag{11.32}$$

$$\exp(\lambda_0) = \int [1-\frac{b}{a}(x-c)]^{\frac{-\lambda_1(b-1)}{b}}\exp[-\lambda_2[1-\frac{b}{a}(x-c)]^{\frac{1}{b}}] \tag{11.33}$$

Let y = 1-[b (x-c)/a]. Then dx = -a dy/b. Substituting in equation (11.33) and changing the limits of integration, we get

$$\exp(\lambda_0) = \frac{a}{b}\int y^{\frac{-\lambda_1(b-1)}{b}}\exp(-\lambda_2 y^{\frac{1}{b}})\,dy \tag{11.34}$$

Let $z = \lambda_2 y^{1/b}$. Then $dy = (b/\lambda_2)(z/\lambda_2)^{b-1} dz$. Equation (11.34) becomes

$$\exp(\lambda_0) = \frac{a}{\lambda_2}\int [\frac{z}{\lambda_2}]^{(b-1)(1-\lambda_1)} \exp(-z)\, dz$$

or

$$\exp(\lambda_0) = \frac{a}{\lambda_2^{1+(b-1)(1-\lambda_1)}}\int z^{[1+(b-1)(1-\lambda_1)]-1} \exp(-z)\, dz \qquad (11.35)$$

The integral in equation (11.35) is equal to $\Gamma(K)$, $K = 1+(b-1)(1-\lambda_1)$. Therefore,

$$\exp(\lambda_0) = \frac{a}{\lambda_2^K}\Gamma(K) \qquad (11.36)$$

This yields the zeroth Lagrange multiplier:

$$\lambda_0 = \ln a - K \ln \lambda_2 + \ln \Gamma(K) \qquad (11.37)$$

From equation (11.33), the zeroth Lagrange is obtained as

$$\lambda_0 = \ln \int [1 - \frac{b}{a}(x-c)]^{-\lambda_1 \frac{(b-1)}{b}} \exp[-\lambda_2 (1 - \frac{b}{a}(x-c))^{\frac{1}{b}}] \qquad (11.38)$$

11.2.3 DERIVATION OF ENTROPY FUNCTION

Introduction of equation (11.37) in equation (11.31) produces

$$f(x) = \exp[-\ln a + K \ln \lambda_2 - \ln \Gamma(K) - \lambda_1 \ln [1 - \frac{b}{a}(x-c)]^{\frac{b-1}{b}} - \lambda_2 [1 - \frac{b}{a}(x-c)]^{\frac{1}{b}}]$$

$$= \frac{\lambda_2^K}{a\Gamma(K)}[1 - \frac{b}{a}(x-c)]^{\frac{-\lambda_1(b-1)}{b}} \exp[-(1 - \frac{b}{a}(x-c))^{\frac{1}{b}}] \qquad (11.39)$$

A comparison of equation (11.39) with equation (11.1) shows that $\lambda_1 = 1$ and $\lambda_2 = 1$.
Taking logarithm of equation (11.39) yields

$$\ln f(x) = -\ln a + K\ln\lambda_2 - \ln\Gamma(K) - \frac{\lambda_1(b-1)}{b}\ln[1-\frac{b}{a}(x-c)] - \lambda_2[1-\frac{b}{a}(x-c))^{\frac{1}{b}}]$$

(11.40)

Making use of equation (11.40) the entropy function can be written as

$$I(f) = \ln a - K\ln\lambda_2 + \ln\Gamma(K) + \lambda_1 E[\frac{(b-1)}{b}\ln(1-\frac{b}{a}(x-c))]$$
$$+ \lambda_2 E[1-\frac{b}{a}(x-c)]^{\frac{1}{b}} \qquad (11.41)$$

11.2.4 RELATION BETWEEN PARAMETERS AND CONSTRAINTS

Taking partial derivatives of equation (11.41) with respect to a, b, c, λ_1 and λ_2 and equating each derivative to zero yields

$$\frac{\partial I}{\partial\lambda_1} = 0 = (b-1)\ln\lambda_2 - \frac{(b-1)d\Gamma(K)}{\Gamma(K)\,d\lambda_1} + E[\frac{(b-1)}{b}\ln(1-\frac{b}{a}(x-c))] \qquad (11.42)$$

$$\frac{\partial I}{\partial\lambda_1} = 0 = -\frac{1}{\lambda_2}[1+(b-1)(1-\lambda_1)] + E[1-\frac{b}{a}(x-c)]^{\frac{1}{b}} \qquad (11.43)$$

$$\frac{\partial I}{\partial a} = 0 = \frac{1}{a} + \lambda_1 E[\frac{(b-1)(x-c)}{a^2(1-\frac{b}{a}(x-c))}] + \lambda_2 E[(1-\frac{b}{a}(x-c))^{\frac{1-b}{b}}(\frac{x-c)}{a^2}] \qquad (11.44)$$

$$\frac{\partial I}{\partial b} = 0 = -(1-\lambda_1)\ln\lambda_2 + \frac{(1-\lambda_1)d\Gamma(K)}{\Gamma(K)d\,b} +$$

$$\lambda_1 E[\frac{1}{b^2}\ln(1-\frac{b}{a}(x-c)) - \frac{(b-1)}{b}\frac{(x-c)}{a(1-\frac{b}{a}(x-c))}] +$$

$$+ \lambda_2 E[(1-\frac{b}{a}(x-c))^{\frac{1}{b}}[-\frac{1}{b^2}\ln(1-\frac{b}{a}(x-c)) - \frac{x-c}{ab(1-\frac{b}{a}(x-c)}]] \quad (11.45)$$

$$\frac{\partial I}{\partial c} = 0 = \lambda_1 E[\frac{b-1}{a}\frac{1}{1-\frac{b}{a}(x-c)}] + \lambda_2 E[(1-\frac{b}{a}(x-c))^{\frac{1-b}{b}}\frac{1}{a}] \quad (11.46)$$

Simplification of equations (11.42)-(11.46) and recalling that λ_1 and $\lambda_2 = 1$ yields respectively:

$$b\frac{d\Gamma(K)}{d\lambda_1} = E[\ln(1-\frac{b}{a}(x-c))] \quad (11.47a)$$

$$1 = E[1-\frac{b}{a}(x-c)]^{\frac{1}{b}} \quad (11.47b)$$

$$E[\{1-\frac{b}{a}(x-c)\}^{1/b}]=1 \quad (11.47c)$$

$$b\frac{d\Gamma(k)}{dk}+b = E[\{1-\frac{b}{a}(x-c)\}^{1/b}\ln\{1-\frac{b}{a}(x-c)\}] \quad (11.48a)$$

$$(1-b)E[\frac{1}{1-\frac{b}{a}(x-c)}] = E[1-\frac{b}{a}(x-c)]^{(1-b)/b} \quad (11.48b)$$

Equations (11.47a) and (11.47b) are the same. Therefore, the estimation equations are equations (11.47a), (11.47b) and (11. 48a).

11. 3 Other Methods of Parameter Estimation

11.3.1 METHOD OF MOMENTS

The GEV distribution has three parameters, a, b, and c so three moments are needed for parameter estimation. For b less than zero (EV type II for flood frequency analysis) the first three moments using the transformation:

$$y = [1 - \frac{b}{a}(x - c)]^{\frac{1}{b}} \tag{11.49}$$

are found to be:

$$M_1^0 = c + \frac{a}{b}[1 - \Gamma(1 + b)] \tag{11.50}$$

$$M_2 = \frac{a^2}{b^2}[\Gamma(1 + 2b) - \Gamma^2(1 + b)] \tag{11.51}$$

$$M_3 = \frac{a^3}{b^3}[-\Gamma(1 + 3b) + 3\Gamma(1 + b)\Gamma(1 + 2b) - 2\Gamma^3(1 + b)] \tag{11.52}$$

where M_1^0, M_2 and M_3 are, respectively, the first moment about origin, and the second and third moments about the centroid. The value of parameter b is computed numerically from its relationship to the skewness coefficient C_S as

$$C_S = \frac{M_3}{M_2^{3/2}} \tag{11.53}$$

11.3.2 METHOD OF PROBABILITY WEIGHTED MOMENTS

The probability weighted moments (PWM) of the GEV distribution are given by Hosking (1986) and Hosking et al. (1985):

$$\beta_r = (r + 1)^{-1}[c + \frac{a}{b}(1 - (r + 1)^{-b}\Gamma(1 + b))] \tag{11.54}$$

where β_r is the r th PWM. The value of parameter b is given by Hosking et al. (1985) as the solution of:

$$\frac{3+t_3}{2} = \frac{(1-3^{-b})}{(1-2^{-b})} \tag{11.55}$$

which is approximated as

$$b = 7.8590\ C + 2.9554\ C^2 \tag{11.56}$$

where C is expressed as

$$C = \frac{2b_1 - b_0}{3b_2 - b_0} - \frac{\log 2}{\log 3} \tag{11.57}$$

where b_i , i=0,1,2, are sample estimates of β_i , I=0,1,2. With the value of b determined as above, parameters a and c are estimated as follows:

$$a = \frac{(2b_1 - b_0)b}{\Gamma(1+b)(1-2^{-b})} \tag{11.58}$$

$$c = b_0 + \frac{a}{b}[\Gamma(1+b) - 1] \tag{11.59}$$

11.3.3 METHOD OF L-MOMENTS

The GEV parameters are estimated using L-moments given by Hosking (1986, 1990) as:

$$b = 7.8590\ C + 2.9554\ C^2 \tag{11.60}$$

where C is given as

$$C = \frac{2}{3+t_3} - \frac{\log 2}{\log 3} \tag{11.61}$$

$$a = \frac{L_2 b}{\Gamma(1+b)(1-2^{-b})} \tag{11.62}$$

$$c = L_1 + \frac{a}{b}[\Gamma(1+b) - 1] \tag{11.63}$$

where t_3 is the L-moment ratio of order 3, and L_1 is the sample estimate of L-moment of order 1.

11.3.4 METHOD OF MAXIMUM LIKELIHOOD ESTIMATION

Jenkinson (1969) and NERC (1975) describe application of MLE to GEV parameter estimation. Jenkinson (1969) transforms equation (11.1)

$$f(x) = \frac{1}{a}\exp[-\exp(y)]\exp[-y(1-b)] \tag{11.64}$$

using the transformation

$$x = c + \frac{a}{b}[1 - \exp(-by)] \tag{11.65}$$

The log-likelihood function can be expressed as

$$\log L = -NLoga - (1-b)\sum_{i=1}^{N} y_i - \sum_{i=1}^{N} \exp(-y) \tag{11.66}$$

where

$$y_i = -\frac{1}{b}\log(1 - \frac{x_i - c}{a}b) \tag{11.67}$$

is obtained from equation (11.65). Differentiating equation (11.66) with respect to a, b, and c and equating each derivative to zero yields respectively:

$$\frac{Q}{a} = 0 \tag{11.68}$$

$$\frac{1}{a}\frac{P+Q}{b} = 0 \tag{11.69}$$

$$\frac{1}{b}(R - \frac{P+Q}{b}) = 0 \tag{11.70}$$

where

$$P = N - \sum_{i=1}^{N} \exp(y_i) \tag{11.71}$$

$$Q = \sum_{i=1}^{N} \exp(y_i + bY_i) - (1-b)\sum_{i=1}^{N} \exp(by_i) \tag{11.72}$$

$$R = N - \sum_{i-1}^{N} y_i + \sum_{i=1}^{N} y_i \exp(-y_i) \tag{11.73}$$

Equations (11.49)-(11.54) are solve numerically to obtain estimates of parameters a, b, and c.

References

Frances, F., Salas, J.D. and Boes, D.C., 1994. Flood frequency analysis with systematic and historical or paleoflood data based on the two-parameter geenral extreme value models. *Water Resources Research*, Vol. **30**, No. 6, pp. 1653-1664.

Hosking, J.R.M., 1986. The theory of probability weighted moments. *IBM Research Report* RC **12210** (#54860), 160 pp., IBM Research Division, T.J. Watson Research Station, Yorktown Heights, New York.

Hosking, J.R.M., 1989. Some theoretical results concerning L-moments. *IBM Research Report* RC **14492** (#64907), 8 pp., IBM Research Division, T.J. Watson Research Station, Yorktown Heights, New York.

Hosking, J.R.M., 1990. L-moments: Analysis and estimation of distributions using linear combinations of order statistics. *Journal of Royal Statistical Society* B, Vol. **52**, No. 1, pp. 105-124.

Hosking, J.R.M. and Wallis, J.R., 1995. A comparison of unbiased and plotting-position estimators of L moments. *Water Resources Research*, Vol. **31**, No. 8, pp. 2019-2025.

Hosking, J.R.M., Wallis, J.R. and Wood E.F., 1985. Esimation of the generalized extreme value distribution by the method of probability-weighted moments. *Technometrics*, Vol. **27**, pp. 251-261.

Jenkinson, A.F., 1955. The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quarterly Journal of Royal Meteorological Society*, Vol. **81**, pp. 58-171.

Jenkinson, A.F., 1969. Estimation of maximum floods. *World Meteorological Organization Technical Note* **98** (5), pp. 183-227, Geneva, Switzerland.

Natural Environment Research Council (NERC), 1975. *Flood Studies Report*. Vol. **1**, pp. 81-97, London, U.K.

Otten, A. and van Montfort, M.A.J., 1980. Maximum likelihood estimation of the general extreme-value distribution parameters. *Journal of Hydrology*, Vol. **47**, pp. 187-192.

Phien, H.N. and Emma, F.T., 1989. Maximum likelihood estimation of the parameters and quantiles of the general extreme value distribution from censored samples. *Journal of Hydrology*, Vol. **105**, pp. 139-155.

Rao, A.R. and Hamed, K.H., 1994. Frequency analysis of upper Cauvery flood data by L-moments. *Water Resources Management*, Vol. **8**, pp. 183-201.

Rao, A.R. and Hamed, K.H., 1997. Regional frequency analysis of Wabash River flood data by L-moments. *Journal of Hydrologic Engineering*, Vol. **2**, No. 4, pp. 169-179.

Sinclair, C.D. and Ahmad, M.I., 1988. Location-invariant plotting positions for PWM estimation of the parameters of the GEV distribution. *Journal of Hydrology*, Vol. **99,** pp. 271-279.

Singh, V.P., 1987. On derivation of the extreme value (EV) type III distribution for low flows using entropy. *Hydrological Sciences Journal*, Vol. **32**, No. 4, pp. 521-533.

Singh, V.P. and Rajagopal, A.K., 1996. A new method of parameter estimation for hydrologic frequency analysis. *Hydrological Science and Technology*, Vol. **2,** No. 3, pp. 33-40.

Tawn, J.A., 1988. An extreme-value theory model for dependent observations. *Journal of Hydrology*, Vol. **101**, pp. 227-250.

Vogel, R.M. and Fennessey, N.M., 1993. L moment diagram should replace product moment diagrams. *Water Resources Research*, Vol. **29,** No. 6, pp. 1745-1752.

Vogel, R.M. and Wilson, I., 1996. Probability distribution of annual maximum, mean, and minimum streamflows in the United States. *Journal of Hydrologic Engineering*, Vol. **1**, No. 1, pp. 69-76.

Wang, Q.J., 1996. Using partial probability weighted moments to fit the extreme value distributions to censored smples. *Water Resources Research*, Vol. **32**, No. 6, pp. 1767-1771.

Wang, Q.J., 1997. LH moments for statistical analysis of extreme events. *Water Resources Research*, Vol. **33**, No. 12, pp. 2841-2848.

CHAPTER 12

WEIBULL DISTRIBUTION

The Weibull distribution is commonly used for frequency analysis as well as risk and reliability analysis of the life times of systems and their components. Its applications have been reported frequently in hydrology and meteorology. Grace and Eagleson (1966) fitted this distribution to the wet and dry sequences and obtained satisfactory results. Rao and Chenchayya (1974) applied it to short-term increment urban precipitation characteristics in various parts of the U.S.A. and obtained satisfactory fit to the durations of wet and dry periods as well as other characteristics. Singh (1987) derived the Webull distribution and estimated its parameters using the principle of maximum entropy (POME). For the precipitation data used, he found POME-based parameter estimates to be either superior or at least comparable to those obtained with the methods of moments and maximum likelihood. Nathan and McMahon (1990) considered some practical aspects concerning the application of the Weibull distribution to low-flow frequency analysis on 134 catchments located in southeastern Australia. They examined the relative performance of the methods of moments, maximum likelihood, and probability weighted moments. They found that different estimation methods provided distinct sets of quantile estimates and the differences between estimation methods decreased as the sample size increased. While fitting the Weibull distribution to annual minimum low flows of different durations, Polarski (1989) found that occasionally the frequency distributions for different durations crossed, in which case the distribution parameters were constrained by adding to the likelihood function the conditions to prevent the curves from crossing. Vogel and Kroll (1989) developed probability-plot correlation coefficient (PPCC) tests for the Weibull distribution. He then used PPCC tests to discriminate among both competing distributional hypotheses for the distribution of fixed shape and competing parameter estimation methods for distributions with variable shape. Durrans (1996) applied the Weibull distribution to obtain estimates of low-flow quantiles, such as 7-day, 10-year low flow. For developing a stochastic flood model Eknayake and Cruise (1993) compared Weibull and exponentially-based models for flood exceedances. They found that the Weibull-based model possessed predictive properties to those of the exponential model when samples exhibited coefficients of variation less than 1.5 and sample sizes were on the order of two events per year. Using Monte Carlo simulation, Singh et al. (1990) made a comparative evaluation of different estimators of the Weibull distribution parameters, including the methods of Moments, probability-weighted moments, maximum likelihood (MLE), least squares, and POME, with the objective of identifying the most robust estimator. Their analysis showed that MLE and POME demonstrated the most robustness.

A random variable X is said to have a Weibull distribution if its probability density function (pdf) is given by

$$f(x) = \frac{a}{b} \left(\frac{x}{b}\right)^{a-1} \exp\left[-\left(\frac{x}{b}\right)^{a}\right], \; a > 0, \; b > 0 \tag{12.1}$$

V. P. Singh, *Entropy-Based Parameter Estimation in Hydrology*

Its cumulative distribution function (cdf) can be expressed as

$$F(x) = \exp[-(\frac{x}{b})^a] \tag{12.2}$$

The Weibull distribution is a two-parameter distribution and can be thought of as a reverse generalized extreme value (GEV) distribution.

12.1 Ordinary Entropy Method

12.1.1 SPECIFICATION OF CONSTRAINTS

Taking logarithm of equation (12.1) to the base 'e', one gets

$$\begin{aligned} \ln f(x) &= \ln a - \ln b + (a-1)[\ln x - \ln b] - \frac{x^a}{b^a} \\ &= \ln a - \ln b - (a-1)\ln b + (a-1)\ln x - \frac{x^a}{b^a} \end{aligned} \tag{12.3}$$

Multiplying equation (12.3) by [-f(x)] and integrating between 0 and ∞ yield the entropy function:

$$\begin{aligned} I(x) = -\int_0^\infty f(x)\ln f(x)dx &= -\int_0^\infty [\ln a - \ln b - (a-1)\ln b] f(x)dx \\ &- (a-1)\int_0^\infty \ln x\, f(x)dx + \frac{1}{b^a}\int_0^\infty x^a f(x)\, dx \end{aligned} \tag{12.4}$$

From equation (12.4) the constraints appropriate for equation (12.1) can be written as

$$\int_0^\infty f(x)dx = 1 \tag{12.5}$$

$$\int_0^\infty \ln x\, f(x)dx = E[\ln x] \tag{12.6}$$

$$\int_0^\infty x^a f(x)dx = E[x^a] = b^a \tag{12.7}$$

Equations (12.5) and (12.7) can be verified as follows. Substituting equation (12.1) in equation (12.5), one gets

$$\int_0^\infty f(x)dx = \frac{a}{b}\int_0^\infty (\frac{x}{b})^{a-1}\exp[-(\frac{x}{b})^a]dx \tag{12.8}$$

Let y = $\frac{x}{b}$. Then equation (12.8) becomes

$$\int_0^\infty f(x)dx = \frac{a}{b}\int_0^\infty y^{a-1} e^{-y^a} b\, dy = a \int_0^\infty y^{a-1} e^{-y^a} dy \tag{12.9}$$

Let x = y^a. Then dz = a y^{a-1} dy. Therefore,

$$\int_0^\infty f(x)dx = a \int_0^\infty y^{a-1} e^{-z} \frac{dz}{a\, y^{a-1}} = \int_0^\infty e^{-z} dz = 1 \tag{12.10}$$

Likewise, substituting equation (12.1) in equation (12.7) one gets

$$\int_0^\infty x^a f(x)dx = \frac{a}{b}\int_0^\infty x^a (\frac{x}{b})^{a-1} \exp[- (\frac{x}{b})^a] dx \tag{12.11}$$

Let $(\frac{x}{b})^a = y$. Then

$$\frac{dy}{dx} = \frac{a\, x^{a-1}}{b^a} \quad or \quad dx = \frac{b^a\, dy}{a\, x^{a-1}} \tag{12.12}$$

Therefore,

$$\begin{aligned}\int_0^\infty x^a f(x)dx &= \frac{a}{b}\int_0^\infty x^a (\frac{x}{b})^{a-1} e^{-y} \frac{b^a\, dy}{a\, x^{a-1}} \\ &= \frac{1}{b\, b^{a-1}} \int_0^\infty b^a\, y\, e^{-y}\, b^a\, dy \\ &= b^a \int_0^\infty e^{-y} y^{2-1} dy = b^a\, \Gamma(2) = b^a\end{aligned} \tag{12.13}$$

12.1.2 CONSTRUCTION OF ZEROTH LAGRANGE MULTIPLIER

The least biased pdf , f(x), consistent with equations (12.5) to (12.7) and based on the principle of maximum entropy (POME), takes the form:

$$f(x) = \exp[- \lambda_0 - \lambda_1 \ln x - \lambda_2 x^a] \tag{12.14}$$

where λ_0, λ_1 and λ_2 are Lagrangian muliplіers. Substitution of equation (12.14) in equation

(12.5) yields

$$\int_0^\infty f(x)dx = \int_0^\infty \exp[-\lambda_0 - \lambda_1 \ln x - \lambda_2 x^a]dx = 1 \tag{12.15}$$

Equation (12.15) yields the partition funciton as

$$\exp(\lambda_0) = \int_0^\infty \exp[(\ln x)^{-\lambda_1} - \lambda_2 x^a]dx = \int_0^\infty x^{-\lambda_1} \exp[-\lambda_2 x^a]\, dx \tag{12.16}$$

Let $\lambda_2 x^a = y$. Then

$$\frac{dy}{dx} = \lambda_2 a x^{a-1} \ ; dx = \frac{dy}{a \lambda_2 x^{a-1}} \ ; x^a = \frac{y}{\lambda_2} \ : x = (\frac{y}{\lambda_2})^{1/a}$$

Substituting the above quantities in equation (12.16) one obtains

$$\begin{aligned} \exp(\lambda_0) &= \int_0^\infty [(\frac{y}{\lambda_2})^{1/a}]^{-\lambda_1} e^{-y} \frac{dy}{a \lambda_2 [\frac{y}{\lambda_2}]^{(a-1)/a}} \\ &= \frac{1}{a} \int_0^\infty (\frac{y}{\lambda_2})^{-(\lambda_1/a)} \frac{1}{\lambda_2} \frac{e^{-y}}{[\frac{y}{\lambda_2}]^{1-(1/a)}} dy \\ &= \frac{1}{a} \int_0^\infty y^{(-(\lambda_1/a)-1+(1/a))} \frac{e^{-y}}{\lambda_2^{(-\lambda_1/a)+1-1+(1/a))}} dy \\ &= \frac{1}{a \lambda_2^{(1-\lambda_1)/a}} \int_0^\infty y^{(((1-\lambda_1)/a)-1)} e^{-y} dy = \frac{\Gamma(1-\frac{\lambda_1}{a})}{a \lambda_2^{(1-\lambda_1)/a}} \end{aligned} \tag{12.17}$$

Equation (12.17) yields the zeroth Lagrange multiplier:

$$\lambda_0 = \ln \Gamma(\frac{1-\lambda_1}{a}) - \ln a - (\frac{1-\lambda_1}{a}) \ln \lambda_2 \tag{12.18}$$

Equation (12.16) also gives the zeroth Lagrange multiplier as

$$\lambda_0 = \ln \int_0^\infty \exp[-\lambda_1 \ln x - \lambda_2 x^a]dx \tag{12.19}$$

12.1.3 RELATION BETWEEN LAGRANGE MULTIPLIERS AND CONSTRAINTS

Differentiating equation (12.19) with respect to λ_1 and λ_2, respectively, yields

$$\frac{\partial \lambda_0}{\partial \lambda_1} = -\frac{\int_0^\infty \ln x \exp[-\lambda_1 \ln x - \lambda_2 x^a]dx}{\int_0^\infty \exp[-\lambda_1 \ln x - \lambda_2 x^a]dx}$$

$$= -\int_0^\infty \ln x \exp[-\lambda_0 - \lambda_1 \ln x - \lambda_2 x^a]dx$$

$$= -\int_0^\infty \ln x f(x)dx = -E[\ln x] \tag{12.20}$$

$$\frac{\partial \lambda_0}{\partial \lambda_2} = -\frac{\int_0^\infty x^a \exp[-\lambda_1 \ln x - \lambda_2 x^a]dx}{\int_0^\infty \exp[-\lambda_1 \ln x - \lambda_2 x^a]dx}$$

$$= -\int_0^\infty x^a \exp[-\lambda_0 - \lambda_1 \ln x - \lambda_2 x^a]dx$$

$$= -\int_0^\infty x^a f(x)dx = -E[x^a] = -b^a \tag{12.21}$$

Differentiating equation (12.18) with respect to λ_2, one obtains

$$\frac{\partial \lambda_0}{\partial \lambda_2} = -\left(\frac{1-\lambda_1}{a}\right)\frac{1}{\lambda_2} \tag{12.22}$$

Equating equations (12.22) and (12.21) results in

$$E[x^a] = b^a = \frac{1 - \lambda_1}{a \lambda_2} \tag{12.23}$$

Differentiating equation (12.18), with respect to λ_1, one obtains

$$\frac{\partial \lambda_0}{\partial \lambda_1} = \frac{\partial}{\partial \lambda_1} [\ln \Gamma (\frac{1 - \lambda_1}{a})] + \frac{\ln \lambda_2}{a} \tag{12.24}$$

Equating equations (12.24) and (12.20) one gets

$$\frac{\partial}{\partial \lambda_1} [\ln \Gamma (\frac{1 - \lambda_1}{a})] + \frac{\ln \lambda_2}{a} = - E[\ln x] \tag{12.25}$$

2.1.4 RELATION BETWEEN LAGRANGE MULTIPLIERS AND PARAMETERS

Substituting equation (12.18) in equation (12.16), one obtains

$$f(x) = \frac{1}{\Gamma(\frac{1-\lambda_1}{a})} a \lambda_2^{(1-\lambda_1)/a} x^{-\lambda_1} \exp[-\lambda_2 x^a] \tag{12.26}$$

Comparing equation (12.26) with equation (12.1), one gets

$$\lambda_2 = \frac{1}{b^a} \tag{12.27}$$

$$\lambda_1 = 1-a \tag{12.28}$$

12.1.5 RELATION BETWEEN PARAMETERS AND CONSTRAINTS

The Weibull distribution has two parameters a and b which are related to the Lagrange multipliers by equations (12.27) and (12.28) which themselves are related to the known constraints by equation (12.22) and (12.25). These two sets of equations are used to eliminate the Lagrange multipliers between them and relate the parameters directly to the constraints as:

$$b^a = E[\ln x^a] \tag{12.29}$$

$$\psi(1) - \ln b = E[\ln x] \tag{12.30}$$

12.1.6 DISTRIBUTION ENTROPY

The distribution entropy is given by equation (12.4) which is rewritten as

$$\begin{aligned} I(x) &= -\int_0^\infty f(x)\ln f(x)dx = [-\ln a + \ln b + (a-1)\ln b]\int_0^\infty f(x)dx \\ &- (a-1\int_0^\infty \ln x f(x)dx + \frac{1}{b^a}\int_0^\infty x^a f(x)dx \end{aligned} \tag{12.31}$$

Evaluating the last integral, we get

$$W = \int_0^\infty x^a f(x)dx = b^a \tag{12.32}$$

Therefore,

$$\begin{aligned} I(x) &= -\ln a + \ln b + \ln b^{a-1} - (a-1)E[\ln x] + 1 \\ &= \ln(\frac{e b^a}{a}) - (a-1)E[\ln x] \end{aligned} \tag{12.33}$$

12.2 Parameter - Space Expansion Method

12.2.1 SPECIFICATION OF CONSTRAINTS

Following Singh and Rajagopal (1986) the constraints for this method are specified by equation (12.5) and

$$\int_0^\infty (\frac{x}{b})^a f(x)dx = E[(\frac{x}{b})^a] \tag{12.34}$$

$$\int_0^\infty \ln(\frac{x}{b})^{a-1} f(x)dx = E[\ln(\frac{x}{b})^{a-1}] \tag{12.35}$$

12.2.2 DERIVATION OF ENTROPY FUNCTION

The pdf corresponding to the principle of maximum entropy (POME) and consistent with equations (12.5), (12.34), and (12.35) takes the form:

$$f(x) = \exp[-\lambda_0 - \lambda_1 (\frac{x}{b})^a - \lambda_2 \ln(\frac{x}{b})^{a-1}] \tag{12.36}$$

where λ_0, λ_1, and λ_2 are Lagrange multipliers. Insertion of equation (12.36) into equation (12.5) yields

$$\begin{aligned} \exp(\lambda_0) &= \int_0^\infty \exp[-\lambda_1 (\frac{x}{b})^a - \lambda_2 \ln(\frac{x}{b})^{a-1}]dx \\ &= \frac{b}{a}\lambda_1^{-K}\Gamma(k), K = \frac{1}{a}[1 - \lambda_2 (a-1)] \end{aligned} \tag{12.37}$$

The zeroth Lagrange multiplier is given by

$$\lambda_0 = \ln b - \ln a - K \ln \lambda_1 + \ln \Gamma(K) \tag{12.38}$$

Also from equation (12.37), one gets the zeroth Lagrange multiplier:

$$\lambda_0 = \ln \int_0^\infty \exp[-\lambda_1 (\frac{x}{b})^a - \lambda_2 \ln(\frac{x}{b})^{a-1}]dx \tag{12.39}$$

Introduction of equation (12.39) in equation (12.36) produces

$$f(x) = \frac{a}{b}\frac{\lambda_1^K}{\Gamma(K)}\exp[-\lambda_1(\frac{x}{b})^a - \lambda_2 \ln(\frac{x}{b})^{a-1}] \tag{12.40}$$

A comparison of equation (12.40) with equation (12.1) shows that $\lambda_1 = 1$ and $\lambda_2 = -1$.
Taking logarithm of equation (12.40) yields

$$-\ln f(x) = -\ln a + \ln b - K \ln \lambda_1 + \ln \Gamma(K) + \lambda_1 (\frac{x}{b})^a + \lambda_2 \ln(\frac{x}{b})^{a-1} \tag{12.41}$$

Multiplying equation (12.41) by f(x) and then integrating from 0 to ∞, we get the entropy function which can be expressed as

$$I(f) = -\ln a + \ln b - K \ln \lambda_1 + \ln \Gamma(K) + \lambda_1 E[(\frac{x}{b})^a] + \lambda_2 E[\ln(\frac{x}{b})^{a-1}] \tag{12.42}$$

12.2.3 RELATION BETWEEN PARAMETERS AND CONSTRAINTS

Taking partial derivations of equation (12.42) with respect to λ_1, λ_2, a, and b separately and equating each derivative to zero, respectively, yields

$$\frac{\partial I}{\partial \lambda_1} = 0 = -\frac{K}{\lambda_1} + E[(\frac{x}{b})^a] \tag{12.43}$$

$$\frac{\partial I}{\partial \lambda_2} = 0 = \frac{(a-1)}{a}\ln\lambda_1 - \frac{(a-1)}{a}\psi(k) + E[\ln(\frac{x}{b})^{a-1}] \tag{12.44}$$

$$\frac{\partial I}{\partial a} = 0 = -\frac{1}{a} + \ln\lambda_1 \cdot \frac{(1+\lambda_2)}{a^2} - \frac{(1+\lambda_2)}{a^2}\psi(K) + \lambda_1 E[(\frac{x}{b})^a \ln(\frac{x}{b})] + \lambda_2 E[\ln(\frac{x}{b})] \tag{12.45}$$

$$\frac{\partial I}{\partial b} = 0 = \frac{1}{b} - \lambda_1 E[\frac{a}{b}(\frac{x}{b})^a] - \lambda_2(\frac{a-1}{b}) \tag{12.46}$$

Simplification of equations (12.44) to (12.46) yields

$$E[(\frac{x}{b})^a] = 1 \tag{12.47}$$

$$E[\ln(\frac{x}{b})] = \frac{\psi(1)}{a} \tag{12.48}$$

$$E[(\frac{x}{b})\ln(\frac{x}{b})] - E[\ln(\frac{x}{b})] = \frac{1}{a} \tag{12.49}$$

$$E[(\frac{x}{b})^a] = 1 \tag{12.50}$$

Equations (12.49) and (12.50) are the same. Equation (12.40) does not exist for all parameter and variate values. Therefore, equations (12.47) and (12.48) are the parameters estimation equations.

12.3 Other Methods of Parameter Estimation

12.3.1 METHOD OF MOMENTS

For the method of moments (MOM), the first two moments suffice to estimate parameters a and b. These two moments about the origin, $M_1(x)$ and $M_2(x)$, are

$$M_1(x) = b\Gamma(1 + \frac{1}{a}) \tag{12.51}$$

$$M_2(x) = b^2\Gamma(1 + \frac{2}{a}) \tag{12.52}$$

where $\Gamma(\cdot)$ is the gamma function. From these moments, the mean $\mu = M_1(x)$, and variance $\sigma^2 = M_2(x) - [M_1(x)]^2$, are

$$\mu = b\Gamma(1 + \frac{1}{a}) \tag{12.53}$$

$$\sigma^2 = b^2[\Gamma(1 + \frac{2}{a}) - \Gamma^2(1 + \frac{1}{a})] \tag{12.54}$$

The moment estimators of a and b, therefore, are

$$\hat{b} = \bar{x}/[\Gamma(1 + \frac{1}{\hat{a}})] \tag{12.55}$$

$$\Gamma(1 + \frac{2}{\hat{a}}) = \frac{S^2}{\hat{b}^2} + \frac{\bar{x}^2}{\hat{b}^2} \tag{12.56}$$

where S^2 is the sample estimate of the variance σ^2 and $\bar{x}$ is the sample mean estimate of μ.

12.3.2 METHOD OF MAXIMUM LIKELIHOOD ESTIMATION

For the method of maximum likelihood estimation (MLE), the log-likelihood function for a sample $x = \{x_1, x_2, \ldots, x_N\}$ drawn from a Weibull population is

$$\log L(x;a,b) = N\log(\frac{a}{b}) + (a-1)\sum_{i=1}^{N}\log(\frac{x_i}{b}) - \sum_{i=1}^{N}(\frac{x_i}{b})^a \tag{12.57}$$

where N is the sample size. The maximum likelihood estimators (MLE's) of a and b are taken to be the values $\hat{a}$ and $\hat{b}$, which yield the maximum of log L. This produces

$$\frac{1}{\hat{a}} = \frac{\hat{a}}{N(\hat{b})}\sum_{i=1}^{N} x_i^{\hat{a}}\log x_i - \overline{\log x} \tag{12.58}$$

$$\hat{b}^{\hat{a}} = \frac{1}{N}\sum_{i=1}^{N} x_i^{\hat{a}} \tag{12.59}$$

A comparison of POME and MLE methods shows that equation (12.59) is equivalent to equation (12. 47), and equations (12.58) and (12.48) have E [ln x] in common. Thus, intuitively, it appears that these methods would yield comparable parameter estimates.

12.3.3 METHOD OF PROBABILITY WEIGHTED MOMENTS

The probability weighted moments (PWM) of a random variable X with the distrubution function F(x) are defined (Greenwood, et al., 1979; Landwehr, et al., 1979) as

$$M_{r,s,t} = E[x^r\{F(x)\}^s \{1 - F(x)\}^t], \text{ for real r, s, and t} \tag{12.60}$$

Again, as in the case of moment estimation, the first two PWM's $(M_{1,0,0}, M_{1,0,1})$ are sufficient to obtain estimates of the parameters a and b. These moments are related to the parameters by (Greenwood, et al., 1979) :

$$M_{1,0,t} = \frac{a\Gamma(1 + 1/b)}{(1 + t)^{1 + 1/b}} \tag{12.61}$$

Thus, the PWM estimators of a and b are:

$$\hat{a} = \frac{M_0}{\Gamma[\ln(\frac{M_0}{M_1})/\ln(2)]} \tag{12.62}$$

$$\hat{b} = \frac{\ln(2)}{\ln(\frac{M_0}{2M_1})} \tag{12.63}$$

where $M_0 = M_{1,0,0}$ and $M_1 = M_{1,0,1}$.

12.3.4 METHOD OF LEAST SQUARES

The method of least squares (MOLS) is based on a linear regression of the observations x_i on the empirical probabilities of x_i estimated from a plotting position formula. In order to obtain these estimates, the data are ranked in descending order and the empirical exceedance probability (1 - F (x)) estimated by:

$$P_i = \frac{m_i}{N + 1} \tag{12.64}$$

where P_i = empirical exceedance probability of observation x_i, m_i = rank of observation x_i , and N = sample size.

From equation (12.2) one can obtain: $P_i = \hat{a} \ln x_i - \hat{a} \ln \hat{b}$. The least squares estimates of a and b are therefore:

$$\hat{a} = \frac{N \sum_{i=1}^{N} P_i \ln x_i - \sum_{i=1}^{N} \ln x_i \sum_{i=1}^{N} P_i}{n \sum_{i=1}^{N} (\ln x_i)^2 - (\sum_{i=1}^{N} \ln x_i)^2} \tag{12.65}$$

$$\hat{b} = \exp\left[\frac{\hat{a} \sum_{i=1}^{N} \ln x_i - \sum_{i=1}^{N} P_i}{N \hat{a}}\right] \tag{12.66}$$

12.4 Comparative Evaluation of Parameter Estimation Methods

12.4.1 APPLICATION TO FREQUENCY ANALYSIS OF PRECIPITATION CHARACTERISTICS

Singh (1987) evaluated MOM, MLE, and POME using rainfall depth values corresponding to given frequencies and durations, which are frequently used in design of urban storm drainage. He employed two sets of excessive precipitation data recorded at Chicago, Illinois, and compiled by the Chicago Weather Bureau Office. The objective of this application was fourfold: (1) to illustrate the POME method for the Weibull distribution, and (2) to examine the adequacy of this distribution for frequency analysis of precipitation characteristics. These data are for the period 1913 to 1935, and have been analyzed by Chow (1953). The first set of data was comprised of accumulated depths (originally given in inches but converted to centimeters) during excessive rates for individual rain storms. The second set of data was for durations of the rain storms considered in the first data set.

Parameters a and b of the Weibull distribution were estimated using MOM, MLE and POME for both data sets and were obtained as:

Data Set	Method	a	b
I Rainfall Depths (2.54 cm)	MOM	1.70	0.858
	MLE	1.83	0.867
	POME	1.99	0.888
II Rainfall Duration (min.)	MOM	1.56	33.08
	MLE	1.67	33.50
	POME	1.76	34.03

Using these parameter values, the Weibull distribution was fitted to the two data sets as shown in Figures 12.1 and 12.2. Clearly the three methods yielded comparable parameter values and consequently comparable agreements between observed and computed distributions. Thus, the POME method is a useful alternative for estimating parameters of the Weibull distribution.

The Weibull distribution did not represent well the probability density functions of rainfall

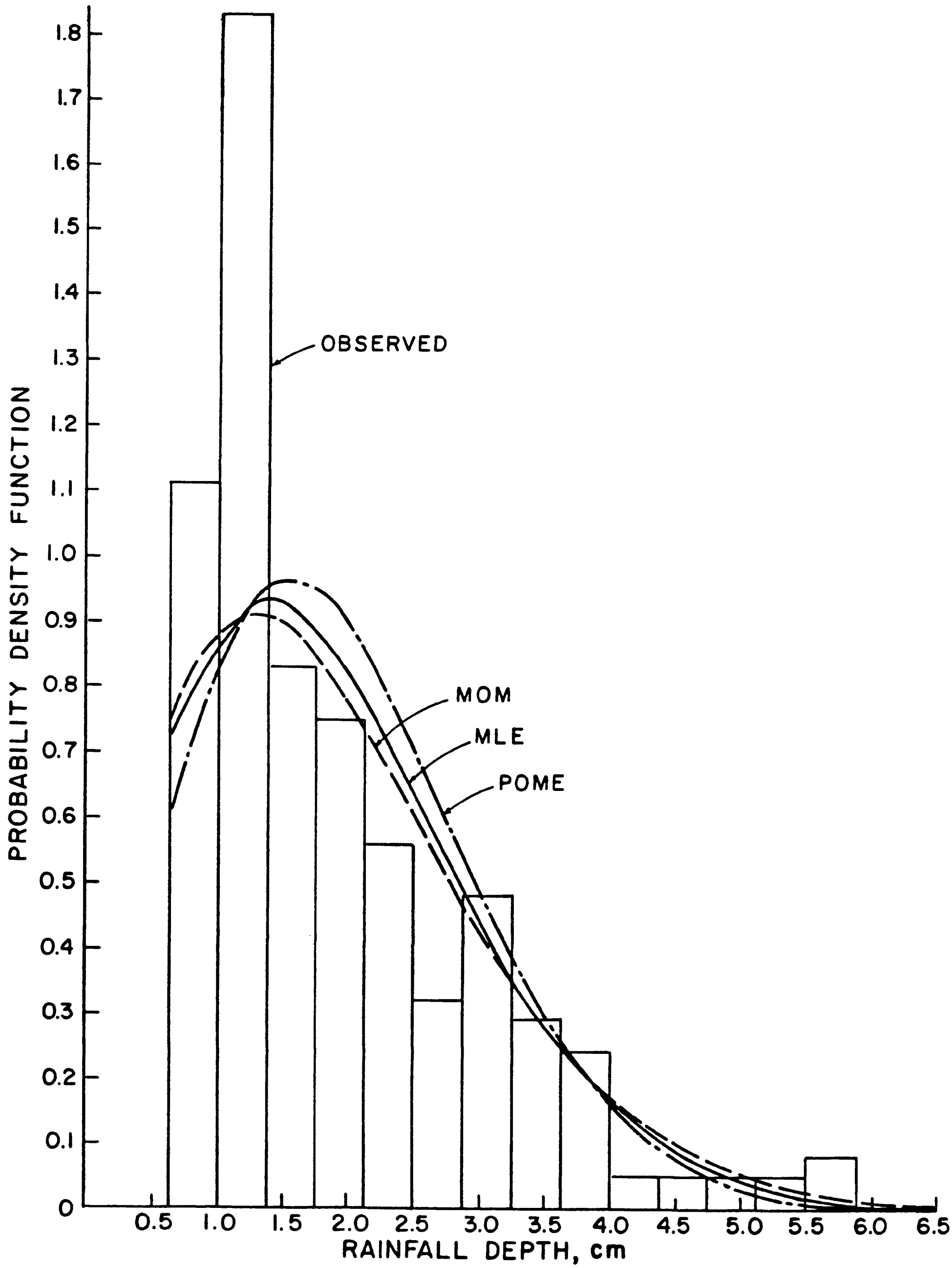

Figure 12.1 Fitting the Weibull distribution to rainfall depth (data set I) by the MOM, MLE and POME methods.

depths and durations for the data used in this study as seen from Figures 12.1 and 12.2. This is at variance with the findings of Rao and Chenchayya (1974). This then suggests that other probability distribution should be found which can more accurately represent probability distributions of rainfall characteristics. This is important in view of large expenditures involved in the design of urban drainage structures.

12.4.2 MONTE CARLO EXPERIMENTATION

12.4.2.1 Monte Carlo Samples: The inverse form of equation (12.2) is given by

$$x(F) = -b[1 - \log P(x)]^{1/a} \qquad (12.67)$$

where x(F) denotes the quantile of cumulative probability P or 1 - F(x). To assess the performance of the parameter estimation methods outlined above, Monte Carlo sampling experiments were performed by Singh et al. (1990). Their work is summarized here. Seven Weibull population cases, listed in Table 12.1, were considered. For each population case, 1,000, 1,500, and 2,000 random samples of size 10, 20, 30, 50, 75, 100, 500, and 1000 were generated, and then parameters and quantiles were estimated by the aforementioned methods. The relative performance of the methods did not significantly depend on the number of samples generated.

12.4.2.2 Performance Indices: The 2,000 estimated values of estimated parameters and quantiles for each sample size and population case were used to approximate the following performance indices for that case: standard Bias (BIAS), standard Error (SE), and root mean square Error (RMSE).

Table 12.1 Weibull population case considered in sampling experiments ($\mu = 1$).

Weibull Distribution	Coefficient of Variation	Parameters	
Population	COV	a	b
Case 1	0.30	3.7142	1.1079
Case 2	0.50	2.1014	1.1291
Case 3	0.70	1.4513	1.1030
Case 4	1.0	1.000	1.000
Case 5	1.5	0 .685	0.773
Case 6	2.0	0.543	0.575
Case 7	3.0	0.411	0.324

12.4.2.2 Bias in Parameter Estimates: The seven cases considered represent a wide variation in variance of the population data. The results of the parameter bias analyses showed that MLE and POME performed very consistently for all cases and all sample sizes in estimating parameter a. MOM demonstrated less bias for case 1 (COV = .30) than MLE; however, MLE still performed well for this case. MOM showed the least consistency of all the mothods, and MOLS, although performing consistently for the wide range in population variation, resulted in high negative bias in all cases. PWM performed very poorly for the cases of small population variance resulting in high negative bias, but performed very well for cases of high variance in the data. In fact, for the last two cases, PWM showed the smallest bias in estimation of a of any of the methods. In summary, it appears that POME performs very well across all population cases and thus appears to be the best estimator of a in terms of bias if the population variance is unknown. However, if a large degree of variance is suspected in the data, PWM may be the superior estimator.

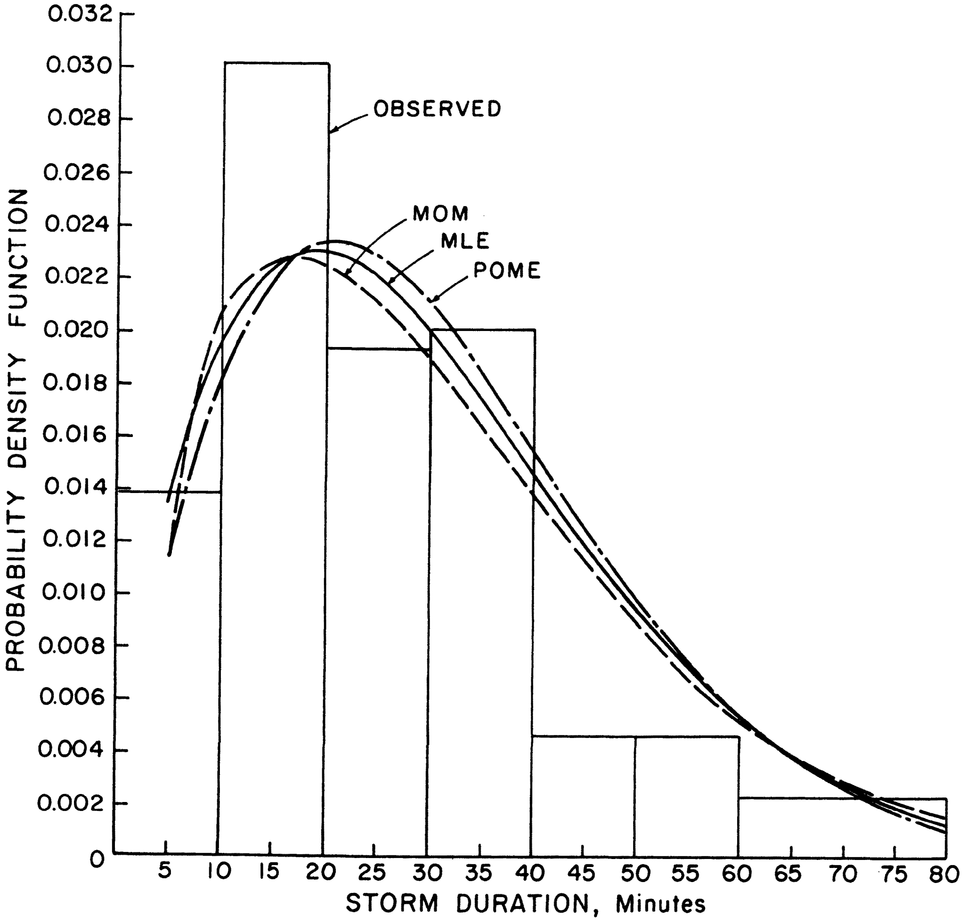

Figure 12.2 Fitting the Weibull distribution to rainfall depth (data set II) by the MOM, MLE and POME methods.

The results of bias in estimation of b were similar to the previous case. Again, MLE and POME preformed well across all cases with MOM deteriorating rapidly with large variance in the population. In fact, there was some deterioration in the performance of all the methods in contrast to the previous case. MOLS performed poorly for all the cases. PWM again offered an interesting case. PWM resulted in negative bias for all cases except one in estimation of b. While there was some deterioration in bias for this method up to case 6 (COV = 2.0), there was improvement in case 7. Therefore, again as in the previous case, PWM appeared to be superior in terms of bias for populations which exhibited a high degree of variance. However, POME and MLE performed most consistently relative to other methods across all fluctuations in population variance.

12.4.2.3 Bias in Quantile Estimates: The results of the quantile bias analyses showed that, in general, as sample size increased, bias for a given return period decreased. For a specified sample size, bias increased as the return period increased. As in the case of bias in parameter estimates, MLE and POME appeared to perform well for all cases and all sample sizes. They resulted in negative bias for all quantiles at small COV values. There was some deterioration in bias as COV increases for all methods. However, MLE and POME consistently performed well relative to all other methods. PWM did not perform as well for higher COV in this case as in the previous case. There did not appear to be any condition for which PWM estimators of quantiles were consistently superior. MOM performed well for small COV, but deteriorated for higher COV values and larger quantiles. MOLS performed very poorly for all cases. MLE appeared to perform better than POME for larger values of COV while the reverse was true of the smaller COV cases. Thus, MLE appeared to be the most resistant and robust estimator in terms of quantile bias.

12.4.2.4 RMSE of Parameter Estimates: The RMSE values of the estimates of Weibull parameters showed that in the case of parameter bias, MLE and POME performed very consistently in terms of RMSE of parameter a. There was no deterioration in RMSE as COV increased for these two methods. MOLS also performed consistently; however RMSE was larger for this method than MLE or POME. As in the previous case, PWM performed poorly for small COV and very well for larger values of COV. However, with one exception (COV= 2.0 for small sample sizes), the PWM estimate always exhibited larger RMSE than MLE or POME. MOM deteriorated rapidly as COV increased and thus was the least robust estimator of a. MLE and POME were very close with a slight edge to MLE except in the case of small sample sizes.

All methods showed some deterioration in RMSE for estimates of parameter b as COV increased. However, MLE and POME performed well in comparison to others in all cases. PWM exhibited less deterioration for increasing COV than any other method. For the cases of COV $\geq$ 2.0, PWM estimates of b were superior for small to medium sample sizes, with MLE and POME performing slightly better for larger sample sizes (N $\geq$ 50). Thus, MLE and POME appeared to be more consistent and robust estimators of b with a very slight advantage to MLE.

12.4.2.5 RMSE of Quantile Estimates: The results of the RMSE analyses for quantile estimation showed that as in the previous cases, MLE and POME again appeared to perform relatively well for most cases. MOM performed well for small sample sizes, but deteriorated somewhat for the larger samples. MOM appeared to perform particularly well for larger quantiles for small sample sizes. Since this is a case of interest to many engineers and physical scientists, MOM appeared attractive because of this characteristic. However, it was not as consistent as MLE or POME for other cases. As in previous cases, PWM performed poorly for small COV values and increasingly better relative to other methods for larger COV values. For the case (COV = 3.00), PWM was superior for the .90 quantile for all sample sizes except 1000 and was competitive for other quantiles. MOLS performed most poorly for all the methods. Although there was deterioration in RMSE for all methods as COV increased, POME and MLE performed most consistently for more cases than the other methods.

12.4.2.6 Robustness of 100 Year Quantile Estimates: The quantile with a return period of 100 (F(x) = .99) is of particular interest to engineers and hydrologists since the Weibull distribution has been applied to rainfall frequency studies. It is generally assumed in practice that the return period of the rainfall is equal to the return period of the peak of the runoff hydrograph resulting from that rainfall. Thus, the rainfall values with a return period of 100 (say, years) is of interest because it may represent a flood hydrograph of the same return period. The 100 year flood is used as the basis for the federal flood insurance program as well as the basis for design of many hydraulic structures. Thus, it is of particular interest to select the estimation method which would result in the most resistant (least RMSE) estimate of that quantile.

The results of the robustness evaluation of this quantile from the seven population cases used in the study for four sample sizes showed that for small N (10 - 20), MOM was the robust estimator by both criteria with PWM and MLE performing better than POME and MOLS. For the case of N = 5, MOM was still superior by the mini-max criterion; however, MLE and POME were better by the average criterion. For large samples (N = 100), MLE and POME were superior by both criteria, with MLE performing slightly better. Thus, if only small samples (10 - 20) were available, one would use MOM; while for large samples, MLE would be preferred by the RMSE criteria.

The results of the bias analysis for the 100 year quantile estimates for all seven test populations showed that for the cases of small population variance, MOM, MLE, and POME all resulted in negative bias for all sample sizes. MLE and POME became increasingly negative in bias for all sample sizes. MLE and POME became increasingly negative up to case 3 (COV = .7) after which POME became less negative and finally became positive in case 5 for small samples and case 6 for all samples except the largest. MLE became less negative after case 4 (COV = 1.0) and finally became positive for the last two cases. Bias in MOLS was positive for all sample sizes and became larger as COV increased. PWM bias was also positive for all cases; however it moved in the opposite direction of MOLS. The largest bias was for case 1 and the bias became less as COV increased until PWM was one of the least biased methods for the last two cases. However, overall MLE and POME consistently showed the least absolute bias, whether positive or negative, than other methods. POME showed the smallest absolute bias in all but the last two cases. However, MLE showed the least deterioration in bias over all seven test populations. Therefore, from an overall perspective, considering both RMSE and bias, MLE would be the most consistent and therefore robust estimator of the 100 year quantile for the Weibull distribution.

12.4.2.7 Concluding Remarks: The Monte Carlo experiments showed that the maximum likelihood estimation method and POME performed most consistently for the largest number of situations for both parameter and quantile estimation. Two exceptions can be noted, however. For very small sample sizes, MOM appeared to be superior in estimating the larger quantile ($F(x) \geq .99$) in terms of RMSE, but not in terms of bias. Also, in cases where large variance was expected in the population, PWM was superior for parameter estimation and for estimation of some quantiles in terms of RMSE and bias.

References

Chow, V.T., 1953. Frequency analysis of hydrologic data with special application to rainfall intensities. *Bulletin* No. **414**, Engineering Experiment Station, University of Illinois, Urbana, Illinois.

Durrans, S.R., 1996. Low-flow analysis with a conditional Weibull tail model. *Water Resources Research*, Vol. **32**, No. 6, pp. 1749-1760.

Eknayake, S.T. and Cruise, J.F., 1993. Comparisons of Weibull- and exponential-based partial duration stochastic flood models. *Stochastic Hydrology and Hydraulics*, Vol. **7**, pp. 283-297.

Grace, R.A. and Eagleson, P.S., 1966. The synthesis of short-term-increment rainfall sequences. *Hydromechanics Laboratory Report* No. **91**, Department of Civil Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts.

Greenwood, J.A., Landwehr, J.M., Matalas, N.C. and Wallis, J.R., 1979. Probability weighted moments: definition and relation to parameters of several distributions expressible in inverse form. *Water Resources Research*, Vol. **15**, pp. 1049-1054.

Landwehr, J.M., Matalas, N.C. and Wallis, J.R., 1979. Probability weighted moments compared with some traditional techniques in estimating Gumbel parameters and quantiles. *Water Resources Research*, Vol. **15,** No. 2, pp. 315-324.

Nathan, R.J., and McMahon, T.A., 1990. Practical aspects of low-flow frequency analysis. *Water Resources Research*, Vol. **26**, No. 9, pp. 2135-2141.

Polarski, M., 1989. Fitting distributions to annual minimum flows of different durations. *IAHS Publication* **187**, pp. 97-104.

Rao, A.R. and Chenchayya, B.T., 1974. Probabilistic analysis and simulation of the short-term increment rainfall process. *Technical Report* No. **55**, Indiana Water Resources Research Center, Purdue University, West Lafayette, Indiana.

Singh, V.P., 1987. On application of the Weibull distribution in hydrology. *Water Resources Management*, Vol. **1**, pp. 33-44.

Singh, V.P., Cruise, J.F. and Ma, M., 1990. A comparative evaluation of the estimators of the Weibull distribution by Monte Carlo simulation. *Journal of Statistical Computation and Simulation*, Vol. **36**, pp. 229-241.

Singh, V.P. and Rajagopal, A.K., 1986. A new mwthod of parameter estimation for hydrologic frequency analysis. *Hydrological Science and Technology*, Vol. **2**, No. 3, pp. 33-40.

Singh, V.P., Rajagopal, A.K. and Singh, K., 1986. Derivation of some frequency distributions using the principle of maximum entropy. *Advances in Water Resources*, Vol. **9**, pp. 91-106.

Singh, V.P., Singh, K. and Rajagopal, A.K., 1985. Application of the principle of maximum entropy (POME) to hydrologic frequency analysis. *Completion Report* **06**, 144 p., Louisiana Water Resources Research Institute, Louisiana State University, Baton Rouge, Baton Rouge, Louisiana.

Vogel, R.M. and Kroll, C.N., 1989. Low-flow frequency analysis using probability-plot correlation coefficients. *Journal of Water Resources Planning and Management*, Vol. **115**, No. 3, pp. 338-357.

CHAPTER 13

GAMMA DISTRIBUTION

The two-parameter gamma distribution is commonly employed for synthesis of instantaneous or finite-period unit hydrographs (Dooge, 1973) and also for flood frequency analysis (Haan, 1977; Phien and Jivajirajah, 1984; Yevjevich and Obseysekera, 1984). By making two hydrologic postulates, Edson (1951) was perhaps the first to derive it for describing a unit hydrograph (UH). Using the theory of linear systems Nash (1957, 1959, 1960) showed that the mathematical equation of the instantaneous unit hydrograph (IUH) of a basin represented by a cascade of equal linear reservoirs would be a gamma distribution. This also resulted as a special case of the general unit hydrograph theory developed by Dooge (1959). On the other hand, using statistical and mathematical reasoning, Lienhard and associates (Lienhard, 1964; Lienhard and Davis, 1971; Lienhard and Meyer, 1967) derived this distribution as a basis for describing the IUH. Thus, these investigators laid the foundation of a hydrophysical basis underlying the use of this distribution in synthesizing the direct runoff. There has since been a plethora of studies employing this distribution in surface water hydrology (Gray, 1961; Wu, 1963; DeCoursey, 1966; Dooge, 1973; Gupta and Moin, 1974; Gupta, et al., 1974; Croley, 1980; Aron and White, 1982; Singh, 1982a, 1982b, 1988; Collins, 1983).

If X has a gamma distribution then its probability density function (pdf) is given by

$$f(x) = \frac{1}{a\Gamma(b)}\left(\frac{x}{\alpha}\right)^{b-1} e^{-x/a} dx \qquad (13.1a)$$

where a > 0 and b > 0 are parameters. The gamma distribution is a two-parameter distribution. Its cumulative distribution function (cdf) can be expressed as

$$F(x) = \int_0^{\infty} \frac{1}{a\Gamma(b)}\left(\frac{x}{a}\right)^{b-1} e^{-x/a} dx \qquad (13.1b)$$

If y =x/a then equation (13.1b) can written as

$$f(y) = \frac{1}{\Gamma(b)} \int_0^y y^{b-1} \exp(-y)\, dy \qquad (13.2a)$$

Abramowitz and Stegun (1958) express F(y) as

$$F(y) = F(\chi^2 | \upsilon) \quad (13.2b)$$

where $F(\chi^2 | \nu)$ is the chi-square distribution with degrees of freedom as $\nu = 2b$ and $\chi^2 = 2y$. According to Kendall and Stuart (1963), for ν greater than 30, the following variable follows a normal distribution with zero mean and variance equal to one:

$$u = [(\frac{\chi^2}{\upsilon})^{1/3} + \frac{2}{9\upsilon} - 1] (\frac{9\upsilon}{2})^{1/2} \quad (13.2c)$$

This helps compute F(x) for a given x by first computing y=x/a and $\chi^2 = 2y$ and then inserting these values into equation (13.2c) to obtain u. Given a value of u, F(x) can be obtained from use of the normal distribution tables.

13.1 Ordinary Entropy Method

13.1.1 SPECIFICATION OF CONSTRAINTS

Taking logarithm of equation (13.1a) to the base e, one gets

$$\ln f(x) = -\ln a\Gamma(b) + (b-1)\ln x - (b-1)\ln a - \frac{x}{a}$$

$$= -\ln a\,\Gamma(b) + (b-1)\ln a + (b-1)\ln x - [x/a] \quad (13.3)$$

Multiplying equation (13.3) by [–f(x)] and integrating between 0 and ∞, one obtains the entropy function:

$$I(f) = -\int_0^\infty f(x)\ln f(x)\,dx = [\ln a\Gamma(b) + (b-1)\ln a]\int_0^\infty f(x)dx$$

$$-(b-1)\int_0^\infty \ln x f(x)\,dx + \frac{1}{a}\int_0^\infty x f(x)\,dx \quad (13.4)$$

From equation (13.4) the constraints appropriate for equation (13.1a) can be written (Singh et al., 1985, 1986) as

$$\int_0^\infty f(x) = 1 \quad (13.5)$$

$$\int_0^\infty x f(x)\,dx = \bar{x} \quad (13.6)$$

$$\int_0^\infty \ln x f(x)\, dx = E[\ln x] \tag{13.7}$$

13.1.2 CONSTRUCTION OF ZEROTH LAGRANGE MULTIPLIER

The least-biased pdf based on the principle of maximum entropy (POME) and consitent with equations (13.5) to (13.7) takes the form:

$$f(x) = \exp[-\lambda_0 - \lambda_1 x - \lambda_2 \ln x] \tag{13.8}$$

where λ_0, λ_1, and λ_2 are Lagrange multipliers.

Substitution of equation (13.8) in equation (13.5) yields

$$\int_0^\infty f(x) dx = \int_0^\infty \exp[-\lambda_0 - \lambda_1 x - \lambda_2 \ln x] dx = 1 \tag{13.9}$$

This leads to the partition function as

$$\exp(\lambda_0) = \int_0^\infty \exp[-\lambda_1 x - \lambda_2 \mathit{lnx}] dx = \int_0^\infty \exp[-\lambda_1 x] \exp[-\lambda_2 \mathit{lnx}] dx$$

$$= \int_0^\infty \exp[-\lambda_1 x] \exp[\mathit{lnx}^{-\lambda_2}] dx = \int_0^\infty \left(\frac{y}{\lambda_1}\right)^{-\lambda_2} e^{-y} \frac{dy}{\lambda_1} \tag{13.10}$$

Let λ_1 x = y. Then [dy/λ_1]/x. Therefore, equation (13.10) becomes

$$\exp(\lambda_0) = \int_0^\infty \left(\frac{y}{\lambda_1}\right)^{-\lambda_2} \exp(-y) \frac{dy}{\lambda_1} = \frac{1}{\lambda_1^{1-\lambda_2}} \int_0^\infty y^{-\lambda_2} e^{-y} dy = \frac{1}{\lambda_1^{1-\lambda_2}} \Gamma(1-\lambda_2) \tag{13.11}$$

Thus, the zeroth Lagrange multipliers λ_0 is given by equation (13.11) as

$$\lambda_0 = (\lambda_2 - 1) \ln \lambda_1 + \ln \Gamma(1 - \lambda_2) \tag{13.12}$$

The zeroth Lagrange multiplier is also obtained from equation (13.10) as

$$\lambda_0 = \ln \int_0^\infty \exp[-\lambda_1 x - \lambda_2 \ln x] dx \tag{13.13}$$

13.1.3 RELATION BETWEEN LAGRANGE MULTIPLIERS AND CONSTRAINTS

Differentiating equation (13.13) with respect to λ_1 and λ_2, respectively, produces

$$\begin{aligned}\frac{\partial \lambda}{\partial \lambda_1} &= -\frac{\int_0^\infty x \exp[-\lambda_1 x - \lambda_2 \ln x] dx}{\int_0^\infty \exp[-\lambda_1 x - \lambda_2 \ln x] dx} \\ &= -\int_0^\infty x \exp[-\lambda_0 - \lambda_1 x - \lambda_2 \ln x] dx = -\int_0^\infty x f(x)\, dx = -\bar{x}\end{aligned} \tag{13.14}$$

$$\begin{aligned}\frac{\partial \lambda_0}{\partial \lambda_2} &= -\frac{\int_0^\infty \ln x \exp[-\lambda_1 x - \lambda_2 \ln x] dx}{\int_0^\infty \exp[-\lambda_1 x - \lambda_2 \ln x] dx} \\ &= -\int_0^\infty \ln x \exp[-\lambda_0 - \lambda_1 x - \lambda_2 \ln x] dx = -\int_0^\infty \ln x f(x)\, dx = -E[\ln x]\end{aligned} \tag{13.15}$$

Also, differentiating equation (13.12) with respect to λ_1 and λ_2 gives

$$\frac{\partial \lambda_0}{\partial \lambda_1} = \frac{\lambda_2 - 1}{\lambda_1} \tag{13.16}$$

$$\frac{\partial \lambda_0}{\partial \lambda_2} = \ln \lambda_1 + \frac{\partial}{\partial \lambda_2} \Gamma(1 - \lambda_2) \tag{13.17}$$

Let $1 - \lambda_2 = k$. Then

$$\frac{\partial k}{\partial \lambda_2} = -1 \tag{13.18a}$$

$$\frac{\partial \lambda_0}{\partial \lambda_2} = \ln \lambda_1 + \frac{\partial}{\partial k} \Gamma(k) \frac{\partial k}{\partial \lambda_2} = \ln \lambda_1 - \psi(k) \tag{13.18b}$$

Equating equations (13.14) and (13.16) as well as equation (13.15) and (13.18), one gets

$$\frac{\lambda_2 - 1}{\lambda_1} = -\bar{x} \quad ; \quad \bar{x} = \frac{k}{\lambda_1} \tag{13.19}$$

$$\psi(k) - E[lnx] = \ln \lambda_1 \tag{13.20}$$

From equation (13.19), $\lambda_1 = k/\bar{x}$, and substituting λ_1 in equation (13.20), one gets

$$E[\ln x] - \ln \bar{x} = \psi(k) - \ln k \tag{13.21}$$

We can find the value of 'k' ($=1-\lambda_2$) from equation (13.21) and substitute it in equation (13.19) to get λ_1.

13.1.4 RELATION BETWEEN LAGRANGE MULTIPLIERS AND PARAMETERS

Substituting equation (13.11) in equation (13.8) gives the entropy-based pdf as

$$\begin{aligned} f(x) &= \exp[(1-\lambda_2)\ln\lambda_1 - \ln\Gamma(1-\lambda_2) - \lambda_1 x - \lambda_2 lnx] \\ &= \exp[\ln\lambda_1^{1-\lambda_2}]\exp[\ln(\frac{1}{\Gamma(1-\lambda_2)})]\exp[-\lambda_1 x]\exp[\ln x^{-\lambda_2}] \\ &= \lambda_1^{1-\lambda_2}\frac{1}{\Gamma(1-\lambda_2)}\exp[-\lambda_1 x]x^{-\lambda_2} \end{aligned} \tag{13.22}$$

If $\lambda_2 = 1-k$ then

$$f(x) = \frac{\lambda_1^k}{\Gamma(k)}\exp[-\lambda_1 x]x^{k-1} \tag{13.23}$$

A comparison of equation (13.23) with equation (13.1a) produces

$$\lambda_1 = 1/a \tag{13.24}$$

and

$$\lambda_2 = 1 - b \tag{13.25}$$

13.1.5 RELATION BETWEEN PARAMETERS AND CONSTRAINTS

The gamma distribution has two parameters a and b which are related to the Langrange mutlipliers by equations (13.24) and (13.25), which themselves are related to the known constraints by equations (13.19) and (13.21). Eliminating the Lagrange multipliers between these two sets of equations, we get parameters directly in terms of the constraints as

$$ba = \bar{x} \tag{13.26}$$

$$\psi(b) - \ln b = E[\ln x] - \ln \bar{x} \tag{13.27}$$

13.1.6 DISTRIBUTION ENTROPY

Equation (13.4) gives the distribution entropy. Rewriting it, one gets

$$I(x) = -\int_0^\infty f(x) \ln f(x)\, dx$$

$$= [\ln a\Gamma(b) + (b-1)\ln a] \int_0^\infty f(x) dx - (b-1) \int_0^\infty \ln x f(x)\, dx$$

$$+ \frac{1}{a} \int_0^\infty x f(x)\, dx$$

$$= [\ln a\Gamma(b) + \ln a^{b-1}] - (b-1) E[\ln x] + \frac{\bar{x}}{a}$$

$$= \ln(a\Gamma(b) a^{b-1}) + \frac{\bar{x}}{a} - (b-1) E[\ln x]$$

$$= \ln(\Gamma(b) a^{b}) + \frac{\bar{x}}{a} - (b-1) E[\ln x] \tag{13.28}$$

13.2 Parameter-Space Expansion Method

13.2.1 SPECIFICATION OF CONSTRAINTS

For this method the constraints, following Singh and Rajagopal (1986), are equation (13.5) and

$$\int_0^\infty \frac{x}{a} f(x)\, dx = E[\frac{x}{a}] \tag{13.29}$$

$$\int_0^\infty \ln(\frac{x}{a})^{b-1} f(x)\, dx = E[\ln(\frac{x}{a})^{b-1}] \tag{13.30}$$

13.2.2 DERIVATION OF ENTROPY FUNCTION

The least-biased pdf corresponding to POME and consistent with equations (13.5), (13.29) and (13.30) takes the form

$$f(x) = \exp\left[-\lambda_0 - \lambda_1 \left(\frac{x}{a}\right) - \lambda_2 \ln\left(\frac{x}{a}\right)^{b-1} \right] \tag{13.31}$$

where λ_0, λ_1, and λ_2 are Lagrange multipliers. Insertion of equation (13.31) into equation (13.5) yields the partition function:

$$\begin{aligned}\exp(\lambda_0) &= \int_0^\infty \exp\left[-\lambda_1 \left(\frac{x}{a}\right) - \lambda_2 \ln\left(\frac{x}{a}\right)^{b-1} \right] dx \\ &= a(\lambda_1)^{\lambda_2(b-1)-1}\Gamma(1-\lambda_2(b-1))\end{aligned} \tag{13.32}$$

The zeroth Lagrange multiplier is given by equation (13.32) as

$$\lambda_0 = \ln a - (1-\lambda_2(b-1))\ln\lambda_1 + \ln\Gamma(1-\lambda_2(b-1)) \tag{13.33}$$

Also, from equation (13.32) one gets the zeroth Lagrange multiplier:

$$\lambda_0 = \ln \int_0^\infty \exp\left[-\lambda_1 \left(\frac{x}{a}\right) - \lambda_2 \ln\left(\frac{x}{a}\right)^{b-1} \right] dx \tag{13.34}$$

Introduction of equation (13.32) in equation (13.31) produces

$$f(x) = \frac{1}{a}(\lambda_1)^{1-\lambda_2(b-1)} \frac{1}{\Gamma(1-\lambda_2(b-1))} \exp\left[-\lambda_1 \frac{x}{a} - \lambda_2 \ln\left(\frac{x}{a}\right)^{b-1} \right] \tag{13.35}$$

A comparison of equation (13.35) with equation (13.1a) shows that $\lambda_1 = 1$ and $\lambda_2 = -1$. Taking logarithm of equation (13.35) yields

$$\ln f(x) = -\ln a + (1-\lambda_2(b-1))\ln\lambda_1 - \ln\Gamma(1-\lambda_2(b-1)) - \lambda_1 \frac{x}{a} - \lambda_2 \ln\left(\frac{x}{a}\right)^{b-1} \tag{13.36}$$

Multiplying equation (13.36) by [-f(x)] and integrating from 0 to ∞ yield the entropy function of the gamma distribution which can be written as

$$I(f) = \ln a - (1-\lambda_2(b-1))\ln\lambda_1 + \ln\Gamma(1-\lambda_2(b-1)) + \lambda_1 E\left[\frac{x}{a}\right] + \lambda_2 E\left[\ln\left(\frac{x}{a}\right)^{b-1}\right] \tag{13.37}$$

13.2.3 RELATION BETWEEN PARAMETERS AND CONSTRAINTS

Taking partial derivatives of equation (13.37) with respect to λ_1 , λ_2 , a, and b separately and equating each derivative to zero, respectively, yields

$$\frac{\partial I}{\partial\lambda_1} = 0 = -(1-\lambda_2(b-1))\frac{1}{\lambda_1} + E\left(\frac{x}{a}\right) \tag{13.38}$$

$$\frac{\partial I}{\partial\lambda_2} = 0 = +(b-1)\ln\lambda_1 - (b-1)\psi(K) + E\left[\ln\left(\frac{x}{a}\right)^{b-1}\right], K = 1-\lambda_2(b-1) \tag{13.39}$$

$$\frac{\partial I}{\partial a} = 0 = +\frac{1}{a} - \frac{\lambda_1}{a}E\left[\frac{x}{a}\right] + \frac{(1-b)}{a}\lambda_2 \tag{13.40}$$

$$\frac{\partial I}{\partial b} = 0 = +\lambda_2\ln\lambda_1 + E\left[\ln\left(\frac{x}{a}\right)^{b-1}\right]^{-\lambda_2}\psi(K) \tag{13.41}$$

Simplification of equation (13.38) – (13.41), respectively, gives

$$E\left(\frac{x}{a}\right) = b \tag{13.42}$$

$$E\left[\ln\left(\frac{x}{a}\right)\right] = \psi(k) \tag{13.43}$$

$$E\left(\frac{x}{a}\right) = b \tag{13.44}$$

$$E[\ln(\frac{x}{a})] = \psi(k) \tag{13.45}$$

Equation (13.42) is the same as equation (13.44), and equation (13.43) as equation (13.45). Therefore, equations (13.42) and (13.43) are the parameter estimation equations.

13.3 Other Methods of Parameter Estimation

The parameters of gamma disribution can be estimated from known sample data D = (x_1, x_2, ..., x_n) using a number of methods which, in general, produce differing estimates and confidence intervals. Croley (1980) discussed some of these methods. Singh and Chowdhury (1985) compared 12 different methods by fitting gamma distribution to four experimentally observed runoff hydrographs and found that statistical methods of parameter estimation were superior to those based on point or planar boundary conditions. Some of these methods are discussed here: (1) method of moments (MOM), (2) method of cumulants (MOC), (3) method of maximum likelihood estimation (MLE), (4) probability weighted moments (PWM), and (5) method of least squares (MOLS). These methods are frequently employed for fitting the gamma distribution in hydrology. The entropy method is then compared with these methods.

13.3.1 METHOD OF MOMENTS

The r–th moment of equation (13.1a) about origin is given as

$$M_r = \frac{1}{a^b \Gamma(b)} \int_0^\infty x^{r+b-1} \exp(-\frac{x}{a})\, dx \tag{13.46}$$

Let (x/a) = y. Then

$$M_r = \frac{a^r}{\Gamma(b)} \int_0^\infty y^{r+b-1} \exp(-y)\, dy = \frac{a^r}{\Gamma(b)} \Gamma(r+b) \tag{13.47}$$

Since there are two parameters, it will suffice to determine the first two moments for the method of moments (MOM):

$$M_1 = ab \tag{13.48}$$

$$M_2 = a^2 b(b+1) \tag{13.49}$$

or

$$E[x] = ab \tag{13.50}$$

$$Var[x] = a^2 b \tag{13.51}$$

Hence, parameters a and b can be estimated by knowing the mean and variance of the variable X.

13.3.2 METHOD OF CUMULANTS

The method of cumulants (MOC) involves finding the first two cumulants C_1 and C_2, and solving for a and b. The r–th cumulant can be expressed as

$$C_r = \frac{d^r}{d\theta^r} \ln G(\theta)|_{\theta=0}; \qquad r = 1,2,... \tag{13.52}$$

in which G(θ) is the moment generating function of f(x) defined as

$$G(\theta) = \int_0^\infty \exp(\theta s) f(s) ds \tag{13.53}$$

Therefore,

$$C_1 = ab \tag{13.54}$$

$$C_2 = a^2 b \tag{13.55}$$

Since cumulants and moments are uniquely related, we get

$$C_1 = M_1 \tag{13.56}$$

$$C_2 = M_2 - M_1^2 \tag{13.57}$$

Cumulants C_1 and C_2 are obtained from M_1 and M_2, and then a and b are determined. It is then clear that the methods of moments and cumulants will yield the same values of a and b.

13.3.3 METHOD OF MAXIMUM LIKELIHOOD ESTIMATION

For the method of maximum likelihood estimation (MLE), the likelihood function L of receiving the sample data D ≅ $(x_1, x_2, ..., x_n)$ given the values of a and b is:

$$L(D|a,b) = \prod_{i=1}^{n} f(x_i) \tag{13.58}$$

Therefore,

$$L(D|a,b) = \frac{1}{a^n(\Gamma(b))^n} [(\frac{x_1}{a})^{b-1} \ldots (\frac{x_n}{a})^{b-1}] \times \exp [- (\frac{x_1}{a} + \ldots + \frac{x_n}{a})] \tag{13.59}$$

If L(D|a, b) is maximal, then so is ln L(D|a, b). Therefore, log-likelihood function is

$$\ln L = - n b \ln a - n \ln \Gamma(b) + (b-1) \sum_{i=1}^{n} \ln x_i - \frac{1}{a} \sum_{i=1}^{n} x_i \tag{13.60}$$

Thus values of a and b are sought which produce

$$\frac{\partial}{\partial a}[\ln L(D|a,b)] = 0 \tag{13.61a}$$

$$\frac{\partial}{\partial b}[\ln L(D|a,b)] = 0 \tag{13.61b}$$

Hence, the equations for estimating a and b are

$$\frac{1}{n} \sum_{i=1}^{n} x_i = \bar{x} = ab \tag{13.62a}$$

$$\psi(b) + \ln a = \frac{1}{n} \sum_{i=1}^{n} \ln x_i \tag{13.62b}$$

where $\psi(b) = d[\ln \Gamma(b)]/db$. Note that $\bar{x}$, in actual practice, will be weighted mean, not arithmetic mean.

Bobee and Ashkar (1991) proposed the following method of solving equations equations (13.62a) and (13.62b). Taking logarithm of equation (13.62a) yields

$$E(b) = W = \ln A - \ln Z; \quad A = \frac{1}{n} \sum_{i=1}^{n} x_i \ ; \quad Z = (x_1 x_2 \ldots\ldots x_n)^{1/n} \tag{13.63a}$$

where A is arithmetic mean and Z is geometric mean. An approximate value of b was obtained as:

$$b = \frac{1}{W}[0.5000876 + 0.1648852\,W - 0.054427\,W^2]\,,\ 0 \le W \le 0.5772 \tag{13.63b}$$

$$b = \frac{8.898919 + 9.059950\,W + 0.9775373\,W^2}{W\,(17.7928 + 11.968477\,W + W^2)}\,,\ 0.5772 \le W \le 17.0 \tag{13.63c}$$

The error in approximation by equation (13.63a) is less than 0.0088% and by equation (13.63b) is less than 0.0054%. The value of b obtained above is substituted in equation (13.62a) to obtain the value of a.

13.3.4 COMPARISON OF POME AND MLE METHODS

On comparing (13.42) and (13.43) with (13.62) and (13.64) it is clear that MLE equations involve sample averages whereas POME equations involve expectations. In other words, if E[x] is replaced by $\Sigma x_i/n$ and E[ln x] by $\Sigma \ln x_i/n$ then the two sets of equations become identical, and will produce identical parameter estimates. In practice, since sample values are used, the two methods would yield the same parameter values.

13.3.5 METHOD OF LEAST SQUARES

The method of least squares (MOLS) minimizes the sum of squares of deviations between observed values (f_0) and computed values (f_c) of the function f. To that end, it is more convenient to use ln f(x) than f(x),

$$\begin{aligned} E &= \sum_{i=1}^{n} [\ln f_0(x) - \ln f_c(x)]^2 \\ &= \sum_{i=1}^{n} [\,\ln f_0(x_i) + \ln\Gamma(b) + b\ln a - (b-1)\ln x_i + \frac{x_1}{a}\,]^2 \rightarrow \min \end{aligned} \tag{13.64}$$

where E is the error function. Differentiating E with respect to a and equating to zero, and doing likewise with respect to b results in two nonlinear equations which can be solved for a and b. However, the global minimum of E is more easily found by computing the surface of E, without logarithmically transforming f(x), in the a-b plane. This procedure has been found to be equally efficient and more instructive as it pictures evolution of the error surface with variations in a and b (Singh and Chowdhury, 1985).

13.3.6 METHOD OF PROBABILITY WEIGHTED MOMENTS

Hosking (1990) derived probability-weighted moments (PWM) for the gamma distribution. These are expressed in terms of L-moments from which parameters a and b are obtained as follows:

$$\lambda_1 = a\,b \tag{13.65}$$

$$\lambda_2 = \frac{1}{\sqrt{\pi}} a \frac{\Gamma(b+1/2)}{\Gamma(b)} \quad (13.66)$$

where λ_1 and λ_2 are first and second order L-moments. In practice these are replaced by their sample estimates. Hosking gave the following solution for a and b. Let $t = \lambda_2/\lambda_1$. For t between 0 and 0.5, $z = \pi t^2$ and b is obtained as

$$b = (1 - 0.3080 z)/(z - o.05812 z^2 + 0.01765 z^3) \quad (13.67)$$

and for t between 0.5 and 1, b is given as

$$b = (0.7213 z - 0.5947 z^2)/(1 - 2.1817 z + 1.2113 z^3) \quad (13.68)$$

With b obtained as above, a is got from

$$a = \lambda_1 / b \quad (13.69)$$

13.4 Comparative Evaluation of Estimation Methods

13.4.1 APPLICATION TO UNIT HYDROGRAPH ESTIMATION

The instantaneous unit hydrograph (IUH) of a drainage basin can be considered as a probability distribution. The IUH is the hydrograph of direct runoff occurring at the basin outlet due to the effective rainfall having unit volume, infinitesimally small duration, and occurring uniformly in the basin. If h(t) is the IUH ordinate at time t and Δt is time interval > 0, then

$$h(t) \geq 0, t \geq 0 \quad (13.70)$$

$$\int_0^{\infty} h(t)dt = 1 \quad (13.71)$$

$$\int_0^{\infty} h(t)dt \leq \int_0^{t+\Delta t} \leq 1, t \leq \infty \quad (13.72)$$

Clearly, h(t) satisfies the qualifications of a probalitiy density function of a random variable t. This can also be perceived from a hydrologic standpoint. When an instantaneous burst of effective rainfall of unit volume occurs uniformly in a basin, the direct runoff appears at the basin outlet. The time taken by a body of water to travel to the basin outlet depends upon the position where the travel is initiated and the path it follows. In a given basin there can be an infinite number of positions where raindrops will land and initiate their travel in association with topographic characteristics. Likewise, there can be an infinite number of paths of travel. These paths are carved by topographic slope configuration and channel network existing in the basin. The time of travel that water spends in following a given path depends on its composition.

Obviously, different paths may have different travel times. If it is assumed that the time of travel is a random variable then by dividing these times of travel into a finite number of class intervals, a frequency distribution of the time of travel can be constructed. Since the volume of effective rainfall is unity, the area occupied by this distribution will be unity. This distribution is then the same as the IUH. In other words, the basin IUH can be considered as a probability distribution of time of travel. Experimental and field experience suggests that this distribution can be represented reasonably well by a two-parameter gamma distribution.

Table 13.1 Some pertinent characteristics of four experimental rainfall-runoff events.

Rainfall				Runoff		
Event	Intensity (mm/h)	Duration (sec)	Depth (mm)	Peak (mm/h)	Peak time (sec)	Duration (sec)
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1	11.6	111	0.36	4.2	205	796
2	26.4	78	0.57	10.6	154	700
3	60.7	71	1.19	33.2	102	545
4	32.7	112	1.02	26.8	86	635

Table 13.2 Parameters a and b of two-parameter gamma distribution estimated by various methods of four experimental rainfall-runoff events.

Methods	Event 1		Event 2		Event 3		Event 4	
	a	b (sec)	a	b (sec)	a	b (sec)	a	b (sec)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
MOM	3.24	89.40	4.33	39.80	3.63	33.50	2.31	47.80
MOC	3.24	89.40	4.33	39.80	3.63	33.50	2.31	47.80
MLE	3.30	87.80	5.60	30.80	5.40	22.50	3.70	29.90
POME	3.30	87.80	5.60	30.80	5.40	22.50	3.70	29.90
MOLS	4.25	70.00	5.00	40.00	5.00	30.00	4.50	30.00

Singh and Chowdhury (1985) used data on four rainfall-runoff events observed at a large outdoor rainfall-runoff experimental facility located at Colorado State University, Fort Collins Colorado. The area covered by these events was approximately 296 m^2. The rainfall intensity was uniform in both space and time for each event. Because the surface of this facility was impervious, virtually the entire rainfall became runoff. Some pertinent characteristics of these events are given in Table 13.1. The unit hydrographs of these events were obtained by simply dividing the runoff hydrographs by their corresponding volumes. Furthermore, because the duration of each event was very small, these unit hydrogrpahs would approximate the instantaneous unit hydrographs. The 2-parameter gamma distribution was fitted to the unit hydrographs which was equivalent to fitting it to the observed runoff hydrographs. Parameters a and b of this distribution were estimated by POME, as well as by the methods of moments (MOM), cumulants (MOC), maximum likelihood estimation (MLE) and least squates (MOLS). The values of these parameters are given in Table 13.2. POME and MLE yielded identical parameter estimates, and so did MOM and MOC for all four events. The parameter estimates of POME were closer to MOM than MOLS for event 1, but the opposite was true for the remaining three events.

Table 13.3 Errors in fitting of the gamma distribution to four experimental rainfall-runoff events by different methods

Method	Relative Error (%) in		Mean Squared Deviation (Discharge)
(1)	Peak Discharge (2)	Time to Peak (3)	(4)
		Event 1	
MOM, MOC	11.86	0	0.181
MLE, POME	11.31	0	0.186
MOLS	5.45	-8.85	0.053
		Event 2	
MOM, MOC	-2.66	16.86	6.687
MLE, POME	-14.02	8.13	5.549
MOLS	6.45	-7.88	0.737
		Event 3	
MOM, MOC	8.16	13.62	76.32
MLE, POME	-6.91	4.21	55.21
MOLS	16.10	-22.33	11.76
Event 4			
MOM, MOC	7.49	23.83	61.66
MLE, POME	-6.36	9.70	46.28
MOLS	6.21	-26.84	6.07

[Relative error = [(observed quantity-computrd quantity)/observed quantity] x 100; Mean squared deviation = Sum of squares of differences between observed and computed discharges, divided by the number of discharge values]

With the parameters estimated as above, the unit hydrographs were used to regenerate the runoff hydrographs. Figures 13.1 to 13.4 compare observed runoff hydrographs with regenerated runoff hydrographs of different methods for the four events. In reproducing peak discharge MOM and MOLS had a slight edge over POME but the difference was so minor as to be negligible. The time to peak was reproduced by POME significantly more accurately than MOM and MOLS. This is seen from Tables 13.3 and 13.4. When the gamma distribtuion fit to the entire hydrograph was examined then MOLS was slightly better than POME which was better than MOM. This is clear from Table 13.3 as well as Figures 13.1 to 13.4. On the whole, the rising hydrograph was better reproduced by MOLS than POME and MOM but the opposite was true for the recession hydrograph. The differences between fits of these methods grew with steepness in rise and fall of the runoff hydrograph as seen from Figures 13.1 to 13.4.

Table 13.4 Average errors (ignoring algebraic sign) of different methods of fitting the gamma distribution

Method	RE in Peak (%)	RE in Peak Time (%)	Average Deviation in Real Time (sec)
MOM, MOC	7.54	13.58	11
MLE, POME	9.65	5.51	2
MOLS	8.55	16.48	19

RE = Relative Error

Table 13.5 Values of entropy showing goodness of fit of each method to experimental data

Method	Event 1	Event 2	Event 3	Event 4
MOM	6.39	5.76	5.48	5.54
MOC	6.39	5.76	5.48	5.54
MLE	6.38	5.65	5.32	5.38
POME	6.38	5.65	5.32	5.38
MOLS	6.31	5.85	5.56	5.50

To further evaluate the goodness of fit of each method to the experimental data the entropy (I) was computed for each event as shown in Table 13.5. The value of I was the smallest for POME except for the event I where MOLS had the smallest value. However, the differences in values of I computed by different methods were quite small, and hence the values were comparable. This means that these methods were comparable. Thus, it can be concluded that POME offered a promising alternative for parameter estimation. For these events the peak time was more accurately reproduced by POME than MOM and MOLS but the opposite was true for peak discharge. Furthermore, the recession hydrograph was better reproduced by POME than MOM and MOLS but the opposite was true for the rising hydrograph.

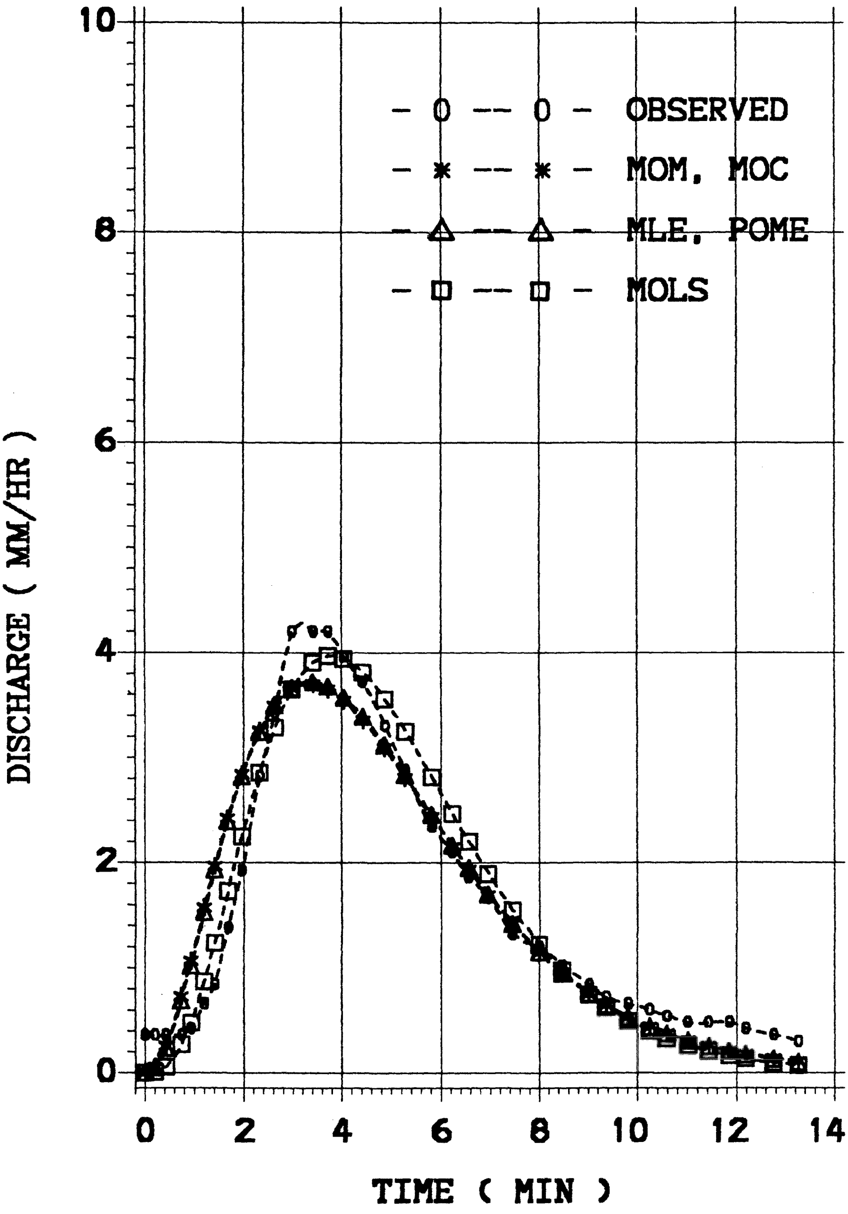

Figure 13.1 Comparison of observed and computed runoff hydrographs for event 1. The methods of computation are MOM, MOC, POME, MLE and MOLS.

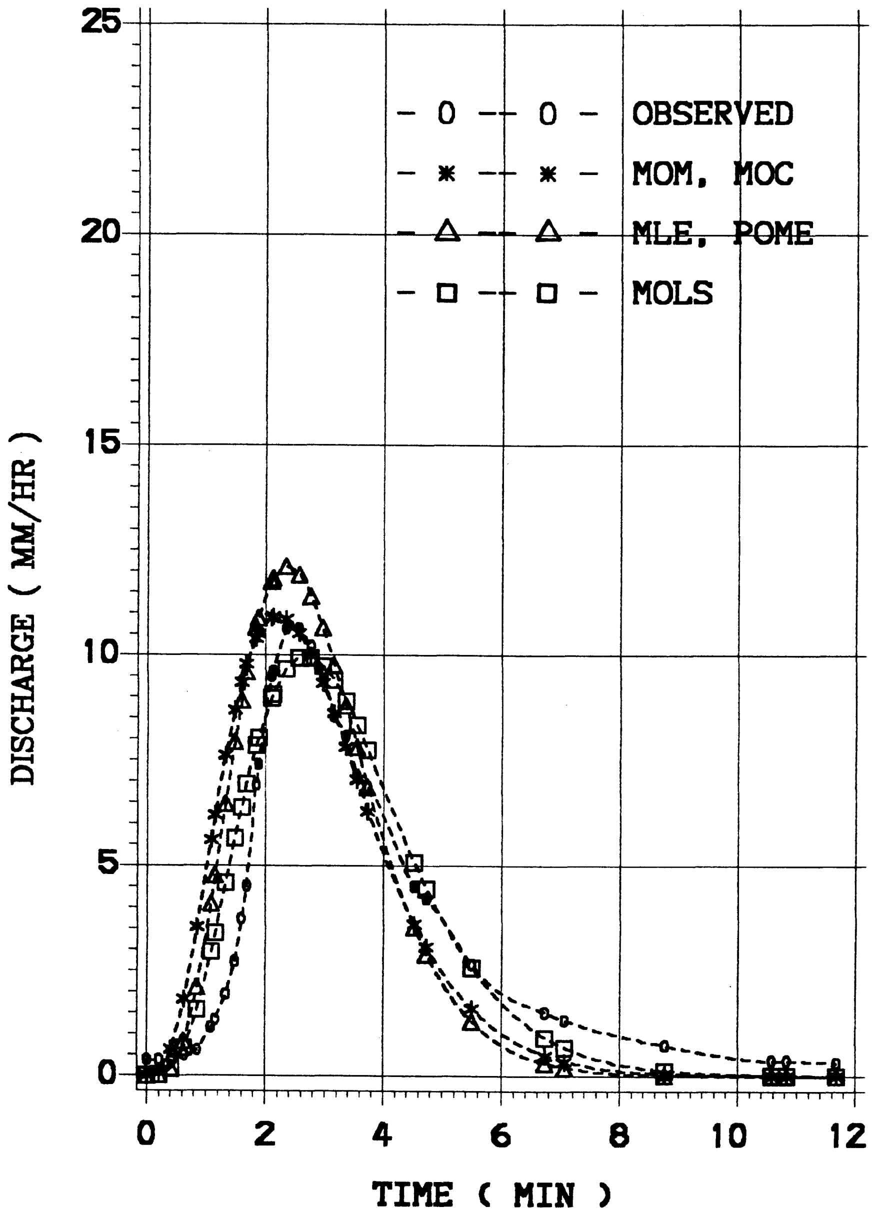

Figure 13.2 Comparison of observed and computed runoff hydrographs for event 2. The methods of computation are MOM, MOC, POME, MLE and MOLS.

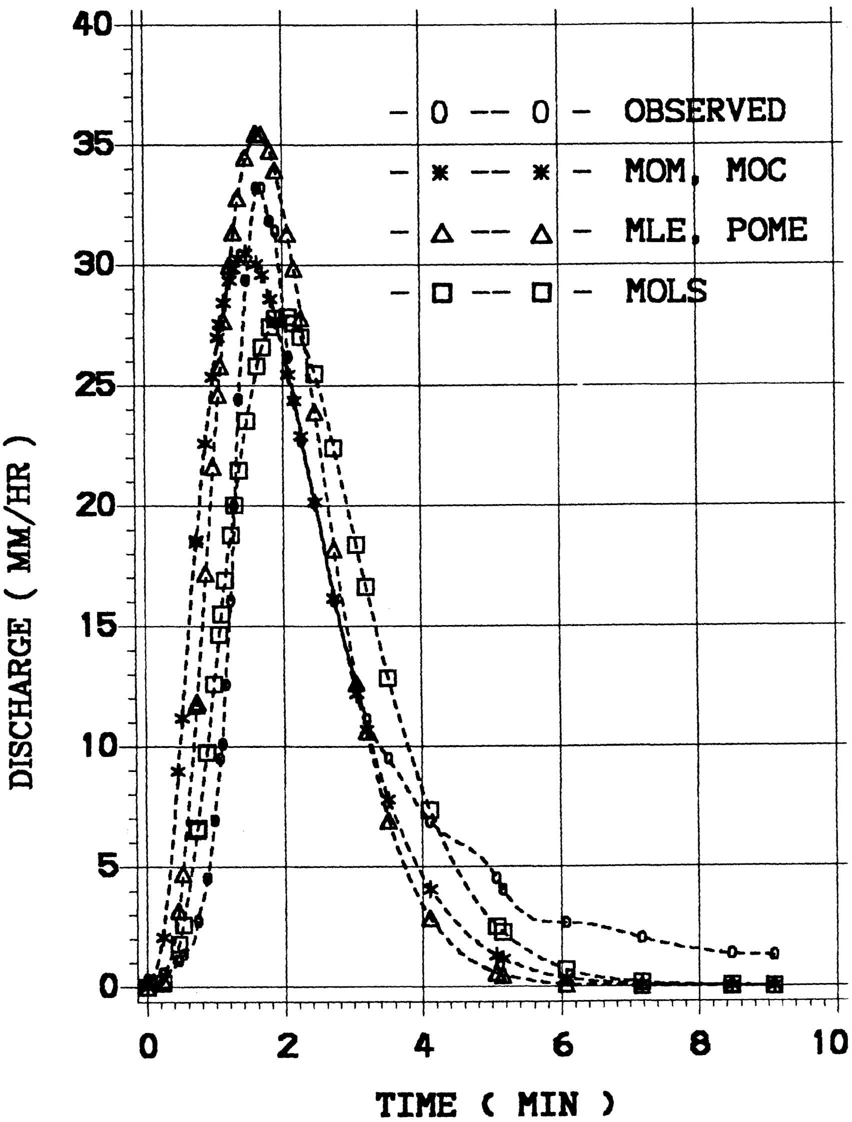

Figure 13.3 Comparison of observed and computed runoff hydrographs for event 3. The methods of computation are MOM, MOC, POME, MLE and MOLS.

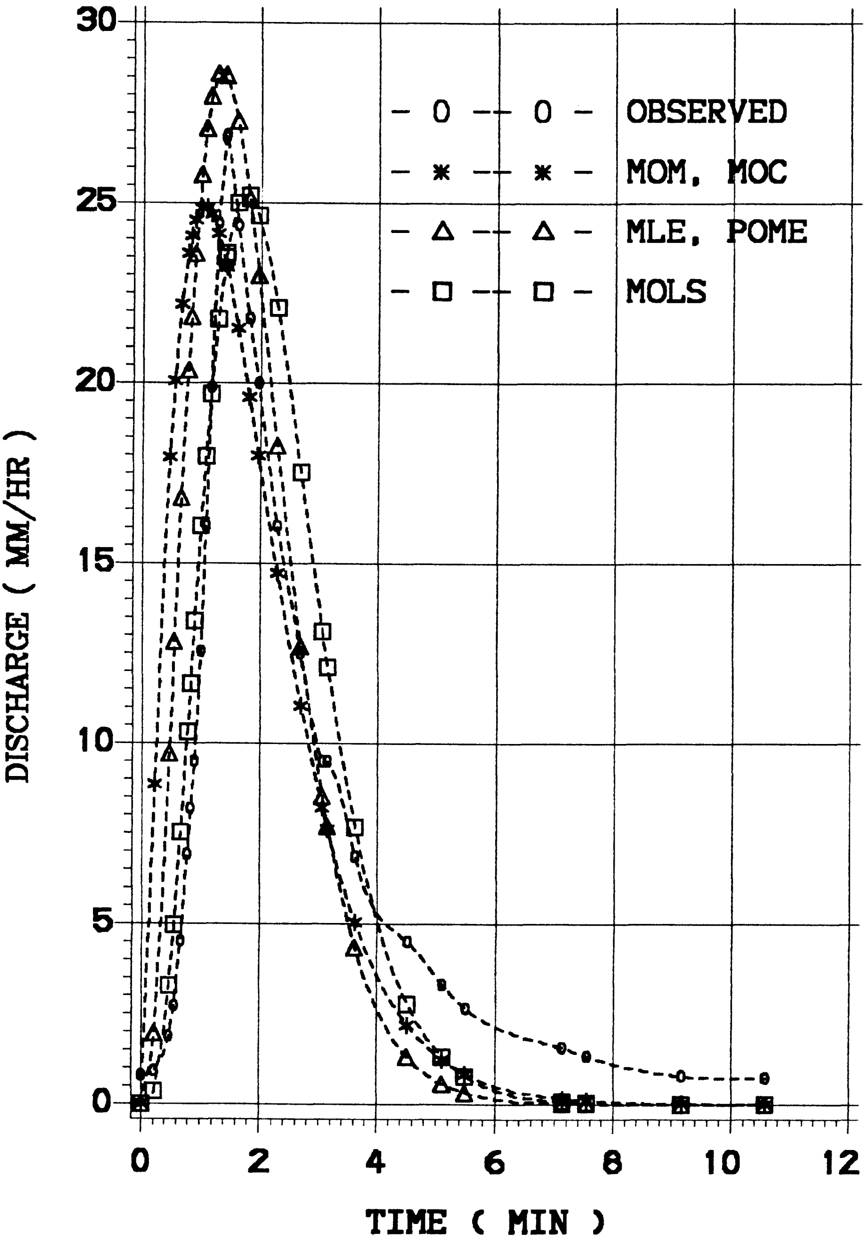

Figure 13.4 Comparison of observed and computed runoff hydrographs for event 4. The methods of computation are MOM, MOC, POME, MLE and MOLS.

Table 13.6 Some pertinent statistical characteristics of annual maximum discharge series for six selected river gaging stations.

River Gaging Station	N	Area (sq.Km.)	$\bar{x}$ (cu.m/s)	S_x	C_s	C_k
Comite River near Olive Branch, Louisiana	38	375.55	238.2	174.5	0.7	2.52
Comite River near Comite, Luoisiana	38	735.56	315.7	166.8	0.54	2.77
Amite River near Magnolia, Louisiana	34	1888.00	745.1	539.5	0.71	3.03
St. John River at Nine Mile Bridge, Maine	32	1670.0	699.0	223.7	0.41	3.01
St. John River at Dickey, Maine	36	3540.0	1449.7	517.7	0.35	2.55
Allagash River near Allagash, Maine	51	1240.0	438.8	159.8	0.71	3.30

Table 13.7 Parameters of the gamma distribution fitted to annual maximum discharge series by MOM, MLE and POME methods.

River	MOM		MLE		POME	
Gaging Station	a	b	a	b	a	b
Comite River near Olive Branch, Louisiana	127.85	1.86	131.82	1.81	131.82	1.81
Comite River near Comite, Louisiana	88.07	3.59	95.15	3.32	95.15	3.32
Amite River near Magnolia, Louisiana	390.62	1.91	445.72	1.67	445.72	1.67
St. John River at Nine Mile Bridge, Maine	71.61	9.76	70.98	9.85	70.98	9.85
St. John River at Dickey, Maine	184.91	7.84	187.62	7.73	187.62	7.73
Allagash River near Allagash, Maine	58.18	7.54	55.97	7.84	55.97	7.84

13.4.2 APPLICATION TO FLOOD FREQUENCY ANALYSIS

Singh and Singh (1985) used data on annual maximum discharge series for six selected river gaging stations for fitting the two-parameter gamma distribution. Some pertinent characteristics of the discharge series are given in Table 13.6. These gaging stations are selected on the basis of homogeneity, completeness, independence and length of record. Each station had more than 30 years of record. They evaluated and compared fitting of the gamma distribution to discharge series by the methods of moments (MOM), maximum likelihood estimation (MLE) and principle of maximum entropy (POME). Parameters a and b of the gamma distribution obtained by the

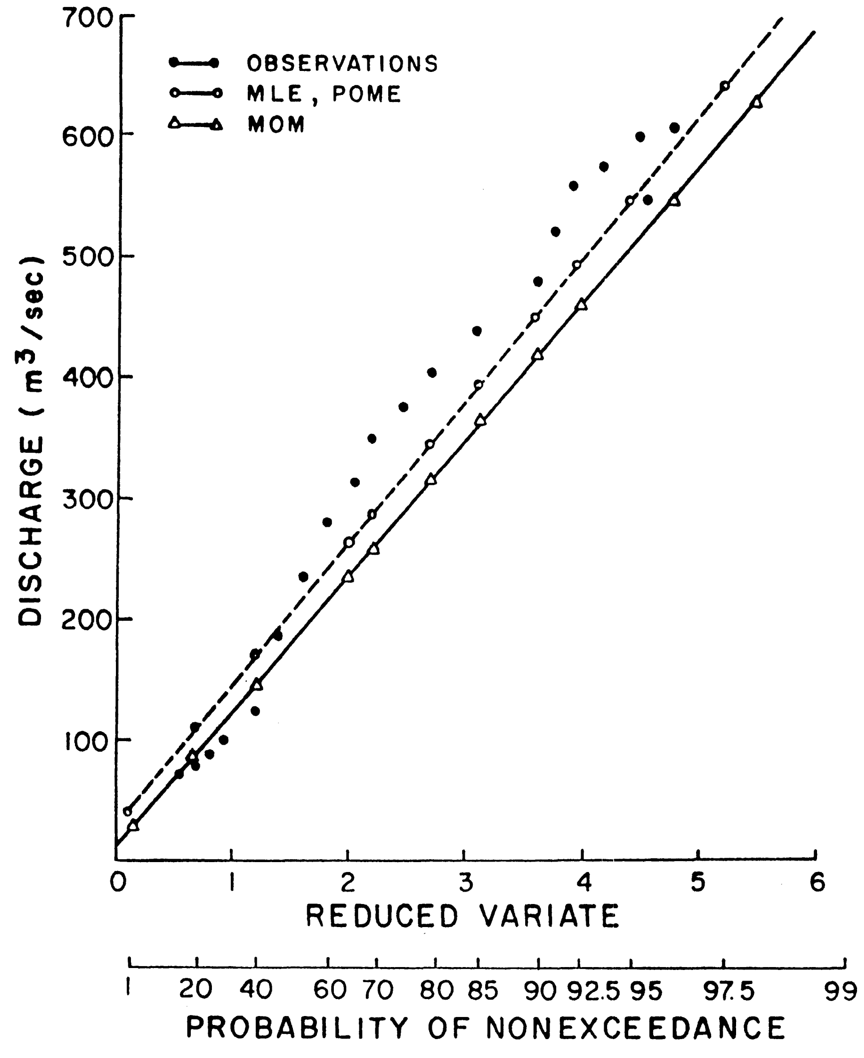

Figure 13.5 Comparison of observed and computed frequency curves for annual maximum discharge series for the Comite River basin near Olive Branch, Louisiana.

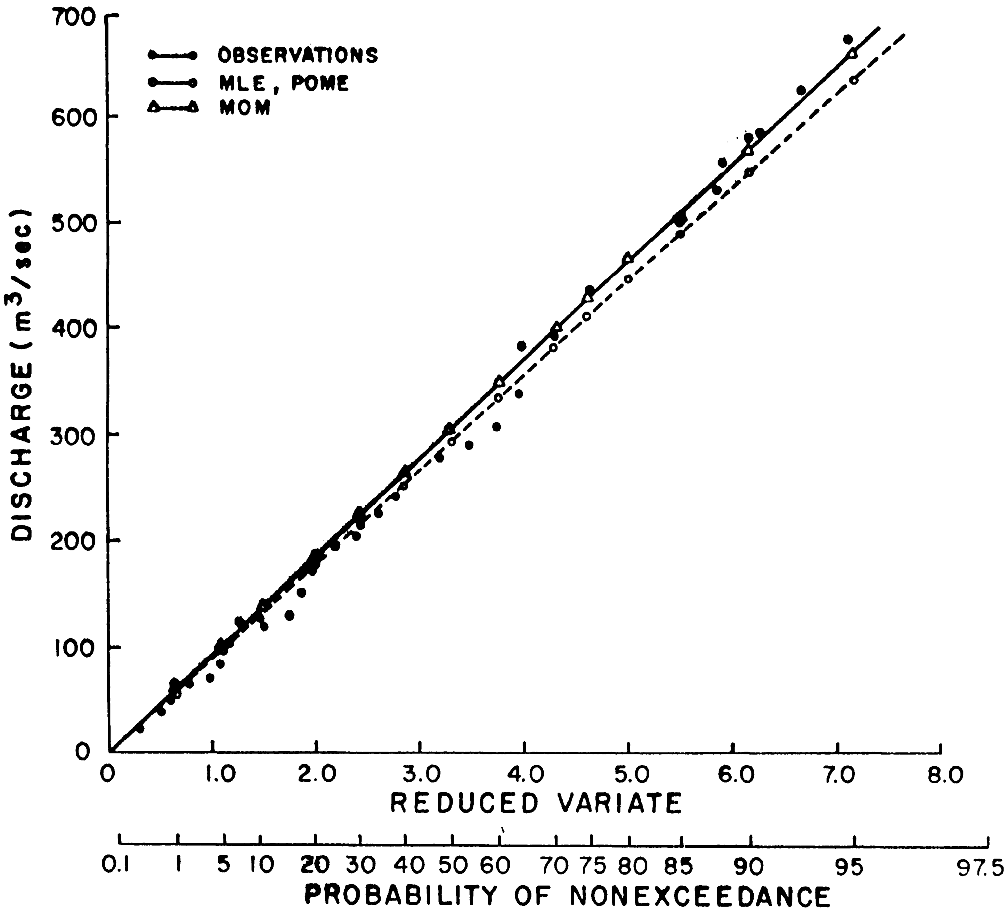

Figure 13.6 Comparison of observed and computed frequency curves for annual maximum discharge series for the Comite River basin near Comite, Louisiana.

three methods for each discharge series are given in Table 13.7. Clearly, the parameter estimates by MLE and POME methods were, as expected, identical. Also these were not greatly different from those by MOM. This is further illustrated by Figures 13.5 and 13.6 which compare frequency curves generated by these methods for two sample gaging stations. POME was found to be superior to MOM for the data used. The SEF was computed with parameters estimated by each method for each discharge series, as given in Table 13.8. Consistently, the SEF values obtained by POME were less than or equal to those by MOM. This implies that POME was a better parameter estimation method than MOM. Thus, it can be concluded that POME offered a promising alternative for parameter estimation.

Table 13.8 Values of entropy (I or SEF).

River gaging station	SEF of sample	SEF			SEF difference		
(1)	(2)	MOM (3a)	POME (3b)	MLE [(2)-(3a)]	MOM [(2)-(3a)]	POME [(2)-(3b)]	MLE [(2)-(3c)]
Comite River near Olive Branch, Luisiana	3.592	3.162	3.166	3.166	0.430	0.426	0.426
Comite River near Comite, Louisaian	3.664	3.343	3.349	3.349	0.321	0.315	0.315
Amite River near Magnolia, Louisiana	3.397	2.946	2.977	2.977	0.451	0.420	0.420
St John River at Nine Mile Bridge, Maine	3.412	3.144	3.143	3.143	0.268	0.269	0.269
St. John River at Dickey, Maine	3.611	3.325	3.326	3.326	0.286	0.285	0.285
Allagash River near Allagash, Maine	3.781	3.544	3.540	3.540	0.237	0.241	0.241

References

Abramowitz, M. and M. and Stegun, I.A., 1965. *Handbook of Mathematical Functions.* Dover Publications, New York.

Amorocho J. And Espildora, B., 1973. Entropy in the assessment of uncertainty in hydrologic systems and models. *Water Resources Research,* Vol. **9**, No. 6, pp. 151-1522.

Aron, G. and White, E.L., 1982. Fitting a gamma distribution over a synthetic unit hydrograph. *Water Resources Bulletin* , Vol. **18**, No. 1, pp. 95-98.

Benson, M.A., 1968. Uniform flood-frequency methods for federal agencies. *Water Resources Research,* Vol. **4**, No. 5, pp. 891-908.

Bobee, B. and Ashkar, F., 1991. *The Gamma Family and Derived Distributions Applied in*

Hydrology. Water Resources Publications, Littleton, Colorado.

Bobee, B. and Robitaille, R. 1975. Correction of bias in the estimation of the coefficient of skewness. *Water Resources Research*, Vol. **11**, No. 6, pp. 851-854.

Bobee, B. and Robitaille, R., 1977. The use of the Pearson type 3 and log-Pearson type 3 distributions revisited. *Water Resources Research*, Vol. **13**, No. 2, pp. 427-443.

Boltzmann, L., 1872. Weitere studien uber das warmegleichgewich unter gasmolekulen. *K. Acad. Wiss. (Wein) Sitzb.*, II Abt. Vol. **66**, pp. 275.

Bucket, J. and Oliver, F.R., 1977. Fitting the Pearson type 3 distribution in practice. *Water Resources Research,* Vol. **13**, No. 5, pp. 851-852.

Burges, S.J., Lettenmaier, D.P. and Bates, C.L., 1975. Properties of the three-parameter log normal probability distribution. *Water Resources Research*, Vol. **11**, pp. 299-235.

Chow, V.T., 1954. The log-probability law and its engineering applications. *Proceedings, American Society of Civil Engineers*, Vol. **80**, Paper No. 536, pp. 1-25.

Collins, M.A., 1983. Discussion of 'Fitting a gamma distributin over a synthetic unit hydrogrtaph' by G. Aron and E.L. White. Water Resources Bulletin, Vol. **19**, No. 2, pp. 303-304.

Crawford, N.H. and Linsley, R.K., 1966. Digital simulation in hydrology: Stanford watershed model IV. *Technical Report* No. **39**, Department of Civil Engineering, Stanford University, Palo Alto, California.

Croley, II, T.E., 1980. Gamma synthetic hydrographs. *Journal of Hydrology*, Vol. **47**, pp. 41-52.

Davy, B.W. and Davies, T.R.H., 1979. Entropy concepts in fluvial geomorphology: a reevaluation. *Water Resources Research*, Vol. **15**, No. 1, pp. 103-106.

DeCoursey, D.G., 1966. A runoff hydrograph equation. Agricultural Research Service, U.S. Department of Agriculture, Washington, D.C., *Report* **ARS-41-116**, 23 pp.

Domokos, M. and Szasz, D. (1968). Generation of fitting probability distribution functions of discharges by electronic computer. *Int. Ass. Sci. Hydrol. Publ.* **81**, pp. 535-545.

Dooge, J.C.I., 1959. A general theory of the unit hydrograph. *Journal of Geophysical Research*, Vol. **64**, No. 2, pp. 311-374.

Dooge, J.C.I., 1973. Linear theory of hydrologic systems. *Technical Bulletin* No. **1468**, 327pp., Agricultural Research Service, U.S. Department of Agriculture, Washington, D.C.

Eagleson, P.S., 1972. Dynamics of flood frequency. *Water Resources Research*, Vol. **8**, No. 4, pp. 878-898.

Edson, C.G., 1951. Parameters for relating unit hydrographs to watershed characteristics. *Transactions, American Geophysical Union*, Vol. **32**, pp. 591-596.

Fisher, R.A. (1922). On the mathematical foundations of theoretical statistics. *Phil. Trans.*, Vol. **A222**, pp. 309-368.

Gray, D.M., 1961. Synthetic unit hydrographs for small watersheds. *Journal of the Hydraulics Division, Proceedings of the American Society of Civil Engineers*, Vol. **87**, No. HY4, pp. 33-54.

Gumbel, E.J., 1958. *Statistics of Extremes*. 375pp., Columbia University Press, New York.

Gupta, V.L. and Moin, A., 1974. Surface runoff hydrograph equation. *Journal of the Hydraulics Division, Proceedings of the American Sociey of Civil Engineers*, Vol. **100**, No. HY10, pp. 1353-1368.

Gupta, V.L., Thongchareon, V. and Moin, S. M.A., 1974. Analytical modeling of surface runoff hydrographs for major streams in northeast Thailand. *Hydrological Sciences Bulletin*, Vol. **XIX**, pp. 523-540.

Haan, C.T., 1977. *Statistical Methods in Hydrology*. The Iowa State University Press, Ames, Iowa.

Hosking, J.R.M., 1990. L-moments: Analysis and estimation of distributions using linear combinations of order statistics. *Journal of Royal Statistical Society*, Series B, Vol. **52**, pp. 105-124.

Jaynes, E.T., 1957a. Information theory and statistical mechanics, I. *Physical Review*, Vol. **106**, pp. 620-630.

Jaynes, E.T., 1957b. Information theory and statistical mechanics, II. *Physical Review*, Vol. **108**, pp. 171-190.

Jaynes, E.T., 1961. *Probability Theory in Science and Engineering*. McGraw-Hill Book Co., New York.

Jaynes, E.T., 1982. On the rationale of entropy methods. *Proceedings of the IEEE*, Vol. **70**, No. 9, pp. 939-952.

Jowitt, P.W., 1979. The extreme-value type-I distribution and the principle of maxium entropy. *Journal of Hydrology*, Vol. **42**, pp. 23-38.

Kalinske, A.A., 1946. On the logarithmic-probability law. *Transactions, American Geophysical Union*, Vol. **27**, pp. 709-711.

Kendall, M.G. and Stuart, A., 1963. *The Advanced Theory of Statistics, Vol. 1: Distribution Theory*. Griffin, London.

Kirby, W. (1974). Algebraic boundedness of sample statistics. *Water Resources Research*, Vol.

10, No. 2, pp. 220-222.

Kite, G.W., 1977. *Frequency and Risk Analysis in Hydrology.* Water Resources Publications, Littleton, Colorado.

Kottegoda, N.T., 1980. *Stochastic Water Resources Technology.* pp. 208-263. John Wiley and Sons, New York.

Leopold, L.B. and Langbein, W.B., 1962. The concept of entropy in landscape evolution. *U.S. Geological Survey Professional Paper* **500A**, 20pp.

Lettenmeier, D.P. and Burges, S.J., 1982. Gumbel's extreme value 1 distribution: a new look. *Journal of the Hydraulics Division, Proceedings of the American Society of Civil Engineers*, Vol. **108**, No. HY4, pp. 502-514.

Levine, R.D. and Tribus, M., 1978. *The Maximum Entropy Formalism.* The MIT Press, Cambridge, Massachusetts.

Lienhard, J.H., 1964. A statistical mechanical prediction of the dimensionless unit hydrograph. *Journal of Geophysical Research*, Vol. **69**, No. 24, pp. 5231-5228.

Lienhard, J.H. and Davis, L.B., 1971. An extension of statistical mechanics to the description of a broad class of macroscopic systems. *Journal of Applied Mathematics and Mechanics*, Vol. **22**, pp. 85-96.

Lienhard, J.H. and Meyer, P.L., 1967. A physical basis for the generalized gamma distribution. *Quarterly of Applied Mathematics*, Vol. **25**, No. 3, pp. 330-334.

Markovic, R.D. (1965). Probability functions of best fit to distrbutions of annual precipitation and runoff. *Hydrol. Pap.* **8**, 33pp., Colorado State University, Fort Collins.

Matalas, N.C. (1963). Probability distribution of low flows. U.S. Geol. Surv. Prof. Pap. **434-A**, pp. 1-27.

Matalas, N.C. and Wallis, J.R., 1973. Eureka! It fits a Pearson type 3 distribution. *Water Resource Research*, Vol. **9**, No. 2, pp. 281-289.

Nash, J.E., 1957. The form of the instantaneous unit hydrograph. *International Association of Scientific Hydrology Publication* **42**, pp. 114-118.

Nash, J.E., 1959. Systematic determination of unit hydrograph parameters. *Journal of Geophysical Research*, Vol. 64, pp. 111-115.

Nash, J.E., 1960. A unit hydrograph study, with particular reference to British catchments. *Proceedings, Institution of Civil Engineers*, Vol. **17**, pp. 249-282.

Phien, H.N. and Jivajirajah, T., 1984. The transformed gamma distribution for annual streamflow frequency analysis. *Proceedings, Fourth Congress of IAHR-APD on Water Resources Management and Development*, Vol. **II**, pp. 1151-1166, Chiang Mai, Thailand.

Rao, D.V., 1980. Log Pearson type 3 distribution: method of mixed moments. *Journal of the Hydraulics Division, Proceedings of the American Society of Civil Engineers*, Vol. **106**, No. HY6, pp. 999-1019.

Rao, D.V., 1983. Estimating log Pearson parameters by mixed moments. *Journal of the Hydraulics Division, Proceedings of tohe American Society of Civil Engineers,* Vol. **109**, No. HY9, pp. 1118-1131.

Rao, A.R. and Chenchagya, B.T., 1974. Probabilistic analysis and simulation of the short-term increment rainfall process. *Technical Report* No. **55**, Indiana Water Resources Research Institute, Purdue University, West Lafayette, Indiana.

Rao, A.R. and Chenchagya, B.T., 1975. Comparative analysis of short time increment urban precipitation characteristics. *Proceedings, National AGU Symposium on Precipitation Analysis for Hydrologic Modeling*, Davis, California, pp. 90-98.

Richardson, C.W., 1982. A comparison of three distributions for the generation of daily rainfall amounts. In *Statistical Analysis of Rainfall and Runoff* edited by V.P. Singh, Water Resources Publication, Littleton, Colorado.

Rohatgi, V.K., 1976. An Introdouction to Probability Theory and Mathematical Statistics. John Wiley and Sons, New York.

Rosenkrantz, R.D., editor, 1983. *E.T. Jaynes: Papers on Probability, Statistics and Statistical Physics*. 434 pp., D. Reidel Publishing Company, Boston, Massachusetts.

Sangal, B.P. and Biswas, A.K., 1970. The 3-parameter lognormal distribution and its application in hydrology. *Water Resources Research*, Vol. **6,** pp. 505-515.

Shannon, C.E., 1948a. The mathematical theory of communications, I and II. *Bell System Technical Journal*, Vol. **27**, pp. 379-423.

Shannon, C.E., 1948b. The mathematical theory of communications, III and IV. *Bell System Technical Journal*, Vol. **27**, pp. 623-656.

Shannon, C.E. and Weaver, W., 1949. *The Mathematical Theory of Communication*. 117pp., University of Illinois Press, Urbana, Illinois.

Singh, V.P., editor, 1982a. *Rainfall-Runoff Relationship*. 581pp., Water Resources Publications, Littleton, Colorado.

Singh, V.P., editor, 1982b. *Statistical Analysis of Rainfall and Runoff.* 700pp., Water Resources Publications, Littleton, Colorado.

Singh, V.P., 1988. *Hydrologic Systems, Vol. 1: Rainfall-Runoff Modeling*. Prentice Hall, Englewood Cliffs, New Jersey.

Singh, V.P. and Chowdhury, P.K., 1985. On fitting gamma distribution to synthetic runoff hydrographs. *Nordic Hydrology*, Vol. **16**, pp. 177-192

Singh, V.P. and Singh, K., 1985. Derivation of the Gamma distribution by using the principle of maximum entropy (POME). *Water Resources Bulletin*, Vol. **21,** No. 6, pp. 941-952.

Singh, V.P. and Rajagopal, A.K., 1986. A new mwthod of parameter estimation for hydrologic frequency analysis. *Hydrological Science and Technology*, Vol. **2**, No. 3, pp. 33-40.

Singh, V.P., Rajagopal, A.K. and Singh, K., 1986. Derivation of some frequency distributions using the principle of maximum entropy. *Advances in Water Resources*, Vol. **9**, pp. 91-106.

Singh, V.P., Singh, K. and Rajagopal, A.K., 1985. Application of the principle of maximum entropy (POME) to hydrologic frequency analysis. *Completion Report* **06**, 144 p., Louisiana Water Resources Research Institute, Louisiana State University, Baton Rouge, Baton Rouge, Louisiana.

Sonuga, J.O., 1972. Principal of maximum entropy in hydrologic frequency analysis. *Journal of Hydrology*, Vol. **17**, pp. 177-191.

Sonuga, J.O., 1976. Entropy principle applied to rainfall-runoff process. *Journal of Hydrology,* Vol. **30**, pp. 81-94.

Todorovic, P., 1982. Stochastic modeling of floods. In *Statistical Analysis of Rainfall and Runoff*, edited by V. P. Singh, pp. 597-636, Water Resources Publications, Littleton, Colorado.

Tribus, M., 1969. *Rational Descriptors, Decisions and Designs.* Pergamon Press, New York.

Verdugo Lazo, A.C.Z. and Rathie, P.N., 1978. On the entropy of continuous probability distributions. *IEEE Transactions on Information Theory,* Vol. **IT-24**, No. 1, pp. 120-122.

Wu, I.P., 1963. Design hydrographs for small watersheds in Indiana. *Journal of the Hydraulics Division, Proceedings of the American Society of Civil Engineers*, Vol. **89**, No. HY6, pp. 35-66.

Yevjevich, V. and Obeysekera, J.T.B., 1984. Estimation of skewness of hydrologic variables. *Water Resources Research*, Vol. **20**, No. 7, pp. 935-943.

CHAPTER 14

PEARSON TYPE III DISTRIBUTION

The Pearson type (PT) III distribution is the generalized gamma distribution and is one of the most popular distributions for hydrologic frequency analysis. Bobee and Robitaille (1977) compared PT III and log PT III distributions using several long-term records of annual flood flows and found PT III distribution to be preferable, especially when the method of moments (MOM) was applied to observed sample data. Bobee (1973), Chang and Moore (1983), among others, used it for flood frequency analysis. Markovic found practically no difference in fitting of Pearson and lognormal distributions to annual precipitation and runoff data. Matalas (1963) found PT III distribution to be representative of low flows. Obeyesekera and Yevjevich (1985) presented a procedure for generation of samples of an autoregressive scheme that has an exact Pearson type III distribution with given mean, variance and skewness. Harter (1958) prepared tables for percentage points of the PT III distribution. Wilk et al. (1962) described a procedure for preparing probability plots for randon samples from an assumed PT III distribution. Haktanir (1991) developed a practical method for computation of PT III frequency factors. Shaligram and Lele (1978) analyzed hydrologic data using PT III distribution and showed that the confidence intervals for this distribution were larger than for the Gumbel distribution.

The Pearson type III distribution has three parameters, which have been estimated in various ways. Ribeiro-Correra and Rousselle (1993) employed a hierarchical approach for regional curve fitting with higher-order moment ratios estmated over large areas. They combined this approach with an empirical Bayes approach for estimation of the scale parameter of PT III distribution. Lall and Beard (1982) estimated PT III moments and investigated into the bias of moment estimates of skew. Ding and Yang (1988) estimated parameters of PT III distribution using the probability-weighted moments (PWM). They extended the PWM method to the samples with extraordinary values. Wu et al. (1991) developed the method of lower bound (MLB) for determining the design quantiles from PT III distribution. Durrans (1992) modified the classical method of moments (MOM) using mean, variance and an extreme order statistic. Singh and Singh (1985) derived the PT III parameter estimates using the principle of maximum entropy (POME). The method of maximum likelihood estimation (MLE) has been used for fitting the PT III distribution to streamflow data (Matalas, 1963; Markovic, 1965; Domokos and Szasz, 1968). The MLE estimates are not as mathematically tractable as moment estimates and are seemingly more difficult to use in operational programs such as the generation of synthetic streamflow sequences. Matalas and Wallis (1973) compared MOM and the method of maximum likelihood estimation (MLE) for parameter estimation and found MLE estimates to be less biased and less variable than MOM estimates. Buckett and Oliver (1977) compared MOM and MLE when fitting a PT III distribution to streamflow data. They concluded that the MLE method gave much more satisfactory estimates of percentiles. From a practical point of view, a comparison between MLE and MOM (considering in the latter method different corrections of the coefficient of skewness C_s) was made by Bobee and Robitaille (1977) on 18 samples of flood data and they

V. P. Singh, *Entropy-Based Parameter Estimation in Hydrology*

showed that: (1) corrections of C_s proposed by Bobee and Robitaille (1975) following the study of Kirby (1974) gave the most satisfactory results when using the method of moments, and (2) that MOM with correction of C_s generally gave better results than the MLE method. Using Monte Carlo experimentation, Hoshi and Leeyavanjia (1986) evaluated the performance of six parameter estimation procedures for different sample sizes and different combinations of population statistics. They found that there was little advantage to using unbiased skew estimates in MOM for estimating the upper quantiles. They showed that the quantile-MLE, quantile-moment and sextile methods performed, in general, accurately. Singh and Singh (1985) found POME to be comparable to MLE and superior to MOM.

If a random variable X has a Pearson type III distribution then its probability density function (pdf) is given by

$$f(x) = \frac{1}{a\Gamma(b)}\left(\frac{x-c}{a}\right)^{b-1}\exp\left[-\left(\frac{x-c}{a}\right)\right] \tag{14.1a}$$

where a > 0, b > 0 and 0 < c < x are parameters. In general, parameter a can take on negative or positive values, but for negative values of a the distribution becomes upper bounded and is therefore unsuitable for frequency analysis of floods. The cumulative distribution function (cdf) of the PT III distribution can be expressed as

$$F(x) = \frac{1}{a\Gamma(b)}\int_0^\infty \left(\frac{x-c}{a}\right)^{b-1}\exp\left(-\frac{x-c}{a}\right)dx \tag{14.1b}$$

The Pearson type III distribution is a three-parameter gamma distribution. If Y = (X-c)/a, then Y follows

$$f(y) = \frac{1}{\Gamma(b)} y^{b-1} \exp(-y) \tag{14.2a}$$

Then the cdf of Y is given by

$$F(y) = \frac{1}{\Gamma(b)} \int_{c0}^{\frac{x_0-c}{a}} y^{b-1} \exp(-y)\, dy \tag{14.2b}$$

The value of X can be computed analytically in the same way as for the gamma distribution, except that the value of Y is calculated from its definition given above.

14.1 Ordinary Entropy Method

14.1.1 SPECIFICATION OF CONSTRAINTS

Taking logarithm of equation (14.1) to the base 'e' one gets

$$\ln f(x) = -\ln a - \ln\Gamma(b) + (b-1)\ln a + \frac{c}{a} - \frac{x}{a} + (b-1)\ln(x-c) \tag{14.3}$$

Multiplying equation (14.3) by [–f(x)] and integrating between c to ∞ one obtains the entropy function I(f):

$$I(f) = -\int_c^{\infty} f(x)\ln f(x)\,dx = -\int_c^{\infty}[-\ln a - \ln\Gamma(b) - (b-1)\ln a + \frac{c}{a}]f(x)dx + \frac{1}{a}\int_c^{\infty} x f(x)dx - (b-1)\int_c^{\infty}\ln(x-c)f(x)\,dx \tag{14.4}$$

From equation (14.4) the constraints appropriate for equation (14.1a) can be written as

$$\int_c^{\infty} f(x)dx = 1 \tag{14.5}$$

$$\int_c^{\infty} x f(x)dx = E[x] \tag{14.6}$$

$$\int_c^{\infty}\ln(x-c)f(x)dx = E[\ln(x-c)] \tag{14.7}$$

Equation (14.5) can be verified as follows. Substituting equation (14.1) in equation (14.5), one gets

$$\int_c^{\infty} f(x)\,dx = \int_c^{\infty}\frac{1}{a\,\Gamma(b)}\left(\frac{x-c}{a}\right)^{b-1}\exp\left[-\left(\frac{x-c}{a}\right)\right]dx \tag{14.8}$$

Let $\frac{x-c}{a} = y$. Then $dy = \frac{dx}{a}$. Therefore, equation (14.8) becomes

$$\int_c^{\infty} f(x)\,dx = \int_0^{\infty}\frac{a}{a\,\Gamma(b)}y^{b-1}e^{-y}\,dy = \frac{}{\Gamma(b)}\int_c^{\infty} y^{b-1}e^{-y}dy = \frac{1}{\Gamma(b)}\Gamma(b) = 1 \tag{14.9}$$

14.1.2 CONSTRUCTION OF ZEROTH LAGRANGE MULTIPLIER

The least-biased pdf, f(x), consistent with equations (14.5) to (14.7) and corresponding to the principle of maximum entropy (POME), takes the form

$$f(x) = \exp[-\lambda_0 - \lambda_1 x - \lambda_2 \ln(x-c)] \tag{14.10}$$

where λ_0, λ_1 and λ_2 are Lagrange multipliers. Substituting equation (14.10) in equation (14.5), we get

$$\int_c^\infty f(x)dx = \int_c^\infty \exp[-\lambda_0 - \lambda_1 x - \lambda_2 \ln(x-c)]\, dx = 1 \tag{14.11}$$

Therefore, the partition function is given by equation (14.11) as

$$\begin{aligned}\exp(\lambda_0) &= \int_c^\infty \exp[-\lambda_1 x - \lambda_2 \ln(x-c)]\, dx = \int_c^\infty \exp[-\lambda_1 x] \exp[-\lambda_2 \ln(x-c)]\, dx \\ &= \int_c^\infty \exp[-\lambda_1 x]\exp[\ln(x-c)^{-\lambda_2}]\, dx = \int_c^\infty (x-c)^{-\lambda_2} \exp[-\lambda_1 x] dx\end{aligned} \tag{14.12}$$

Substituting y = x–c and dy = dx in equation (14.8), one gets

$$\begin{aligned}\exp(\lambda_0) &= \int_0^\infty y^{-\lambda_2} \exp[-\lambda_1 (y+c)]\, dy = \exp[-\lambda_1 c] \int_0^\infty y^{-\lambda_2} \exp[-\lambda_1 y]\, dy \\ &= \exp[-\lambda_1 c] \int_0^\infty \frac{\lambda_1^{-\lambda_2}}{\lambda_1^{-\lambda_2}} y^{-\lambda_2} \exp[-\lambda_1 y]\, dy = \frac{\exp[-\lambda_1 c]}{\lambda_1^{-\lambda_2}} \int_0^\infty (\lambda_1 y)^{-\lambda_2} \exp[-\lambda_1 y]\, dy \\ &= \frac{\exp[-\lambda_1 c]}{\lambda_1^{1-\lambda_2}} \int_0^\infty (\lambda_1 y)^{-\lambda_2} \exp[\lambda_1 y](\lambda_1 dy)\end{aligned} \tag{14.13}$$

Let λ_1 y = z. Then dz = λ_1 dy. Therefore, equation (14.13) becomes

$$\exp(\lambda_0) = \frac{\exp[-\lambda_1 c]}{\lambda_1^{1-\lambda_2}} \int_0^\infty z^{-\lambda_2} e^{-z} dz \tag{14.14}$$

Since

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx \tag{14.15}$$

$$\Gamma(1-\lambda_2) = \int_0^\infty x^{1-\lambda_2-1} e^{-x} dx = \int_0^\infty x^{-\lambda_2} e - x dx \tag{14.16}$$

equation (14.14) becomes

$$\exp(\lambda_0) = \frac{\exp[-\lambda_1 c]\,\Gamma(1-\lambda_2)}{\lambda_1^{1-\lambda_2}} \tag{14.17}$$

The zeroth Lagrange multiplier λ_0 is given as

$$\lambda_0 = -\lambda_1 c + (\lambda_2 - 1)\ln\lambda_1 + \ln\Gamma(1-\lambda_2) \tag{14.18}$$

The zeroth Lagrange multiplier is also obtained from equation (14.12) as:

$$\lambda_0 = \ln \int_0^\infty \exp[-\lambda_1 x - \lambda_2 \ln(x-c)]\, dx \tag{14.19}$$

14.1.3 RELATION BETWEEN LAGRANGE MULTIPLIERS AND CONSTRAINTS

According to Tribus (1969) the relation between Lagrange multipliers and constraints is obtained by taking partial derivatives of the zeroth Lagrange multipliers with respect to other multipliers and then equating these derivatives to the constraints given by equations (14.16) and (14.17). Differentiating equation (14.19) With respect to λ_1 and λ_2, respectively, one obtains

$$\begin{aligned}\frac{\partial\lambda_0}{\partial\lambda_1} &= -\frac{\int_c^\infty \exp[-\lambda_1 x - \lambda_2\ln(x-c)]dx}{\int_c^\infty \exp[-\lambda_1 x - \lambda_2\ln(x-c)]dx}\\ &= -\int_c^\infty x\exp[-\lambda_0 - \lambda_1 x - \lambda_2\ln(x-c)]\,dx = -\int_c^\infty xf(x)\,dx\\ &= -E[x]\end{aligned} \tag{14.20}$$

$$\begin{aligned}\frac{\partial\lambda_0}{\partial\lambda_2} &= -\frac{\int_c^\infty \ln(x-c)\exp[-\lambda_1 x - \lambda_2\ln(x-c)]dx}{\int_c^\infty \exp[-\lambda_1 x - \lambda_2\ln(x-c)]dx}\\ &= -\int_c^\infty \ln(x-c)\exp[-\lambda_0 - \lambda_1 x - \lambda_2\ln(x-c)]dx\\ &= -\int_c^\infty \ln(x-c)f(x)dx = -E[\ln(x-c)]\end{aligned} \tag{14.21}$$

Also differentiating equation (14.18) with respect to λ_1 one gets

$$\frac{\partial\lambda_0}{\partial\lambda_1} = -c + \frac{(\lambda_2 - 1)}{\lambda_1} \tag{14.22}$$

Equating equation (14.22) and (14.20), one gets

$$E[x] = c + \frac{1 - \lambda_2}{\lambda_1} \qquad (14.23)$$

Differentiating equation (14.18) with respect to λ_2, one obtains

$$\frac{\partial \lambda_0}{\partial \lambda_2} = \ln \lambda_1 + \frac{\partial}{\partial \lambda_2}[\ln \Gamma(1 - \lambda_2)] \qquad (14.24)$$

Equating equations (14.24) and (14.21), we get

$$\ln \lambda_1 + \frac{\partial}{\partial \lambda_2}[\ln \Gamma(1 - \lambda_2)] = -E[\ln(x - c)] \qquad (14.25)$$

The PT III distribution has three parameters. Therefore, equations (14.23) and (14.25) are not sufficient and another equation is needed. This is obtained by recalling that

$$\sigma^2(x) = \frac{\partial^2 \lambda_0}{\partial \lambda_1^2} = \frac{-(1 - \lambda_2)}{\lambda_1^2} \qquad (14.26a)$$

Therefore,

$$\sigma^2(x) = \frac{\lambda_2 - 1}{\lambda_1^2} \qquad (14.26b)$$

where $\sigma^2(x)$ is variance of x.

14.1.4 RELATION BETWEEN LAGRANGE MULTIPLIERS AND PARAMETERS

Let $1 - \lambda_2 = b$. Then $[\partial b/\partial \lambda_2] = -1$. In terms of b we find from equations (14.23) to (14.25),

$$E[x] = c + \frac{b}{\lambda_1} \qquad (14.27)$$

$$\sigma^2(x) = \frac{b}{\lambda_1^2} \qquad (14.28)$$

This leads to

$$E[x] = c + ab \qquad (14.29)$$

$$\sigma^2(x) = a^2 b \qquad (14.30)$$

$$-E[\ln(x - c)] = \ln \lambda_1 + \frac{\partial}{\partial b} \ln \Gamma(b) \frac{\partial b}{\partial \lambda_2}$$

$$E[\ln(x-c)] = \psi(b) - \ln \lambda_1 \qquad (14.31a)$$

The function, $\psi(z) = d[\ln \Gamma(z)]/dz$ is called the digamma function. Equation (14.31a) yields

$$E[\ln(x-c)] = -\psi(b) + \ln a \qquad (14.31b)$$

Substituting equation (14.18) in equation (14.10), one gets

$$f(x) = \exp[\lambda_1 c - (\lambda_2 - 1)\ln \lambda_1 - \ln\Gamma(1-\lambda_2) - \lambda_1 x - \lambda_2 \ln(x-c)]$$

$$= \exp[\lambda_1 c - \lambda_1 x + \ln \lambda_1^{1-\lambda_2} + \ln(x-c)^{-\lambda_2} - \ln\Gamma(1-\lambda_2)]$$

$$= \exp(\lambda_1 c - \lambda_1 x + \ln[\lambda_1^{1-\lambda_2}(x-c)^{-\lambda_2}/\Gamma(1-\lambda_2)]) = \frac{\lambda_1^{1-\lambda_2}(x-c)^{-\lambda_2}}{\Gamma(1-\lambda_2)}\exp[-\lambda_1(x-c)]$$

$$(14.32)$$

A comparison of equation (14.32) with equation (14.1a) shows that

$$1 - \lambda_2 = b \qquad (14.33)$$

$$\lambda_1 = \frac{1}{a} \qquad (14.34)$$

14.1.5 RELATION BETWEEN PARAMETERS AND CONSTRAINTS

The PT 3 distribution has 3 parameters: a, b and c. Equations (14.23), (14.25) and (14.26) relate the Lagrange multipliers to the known constraints and the variance of x, and equations (14.33) and (14.34) relate the Lagrange multipliers to parameters. Eliminating the Lagrange multipliers between these two sets of equations, we get parameters in terms of the constraints, as given by equations (14.29), (14.30) and (14.31). These equations are nonlinear but can be solved iteratively.

An example problem is given to illustrate the procedure. To summarize, the mean of X is

$$\bar{x} = ab + c \qquad (14.35)$$

The variance of X is

$$\sigma^2(x) = a^2 b \text{ or } \sigma(x) = ab^{0.5} \qquad (14.36)$$

Furthermore,

$$\frac{1}{n}\sum_{i=1}^{n} \ln(x_i - c) = \psi(b) + \ln a \qquad (14.37)$$

Therefore,

$$a = \left(\frac{\sigma_x^2}{b}\right)^{1/2} \qquad (14.38)$$

$$\bar{x} = (\frac{\sigma_x^2}{b})^{1/2} b + c = \sigma_x^2 b^{1/2} + c \qquad (14.39)$$

$$c = \bar{x} - \sigma_x b^{1/2} \qquad (14.40)$$

Thus,

$$\frac{1}{n} \sum_{i=1}^{n} \ln[x_i - c] = \psi(b) = \ln a \qquad (14.41)$$

$$\frac{1}{n} \sum_{i=1}^{n} \ln[x_i - \bar{x} + \sigma_x b^{1/2}] = \psi(b) + \ln [\frac{\sigma_x}{b^{1/2}}] \qquad (14.42)$$

$$\frac{1}{n} \sum_{i=1}^{n} \ln[x_i - \bar{x} + \sigma_x b^{1/2}] = \ln b - \frac{1}{2b} - \frac{1}{12b^2} + \frac{1}{120b^4} - \frac{1}{252b^6} - \frac{1}{2}\ln b + \ln \sigma_x \qquad (14.43)$$

Therefore,

$$\frac{1}{n} \sum_{i=1}^{n} \ln[x_1 - \bar{x} + \sigma_x b^{1/2}] = \frac{1}{2}\ln b + \ln \sigma_x - \frac{1}{2b} - \frac{1}{12b^2} + \frac{1}{120b^4} - \frac{1}{252b^6} \qquad (14.44)$$

From the data , it is known that $\bar{x} = 11151.842$ and $S_x = 5889.9675$. Using equations (14.41)-(14.43), the parameters are found to be:

a = 2038.3101; b = 8.3499768; c = – 5868.0003

To verify the accuracy of these parameter estimates, we compute

$$\sum_{i=1}^{n} \ln (x_i - c) = 367.9648$$

From equation (14.29), 17019.842 – 5868.0003 = 11151.842, which is the mean.

From equation (14.30), 2038.3101 x $(8.3499768)^{0.5}$ = 5889.9674, which is the standard deviation.

From equation (14.34), the left hand side (LHS) = (367.9648)/38 = 9.68

The right hand side (RHS) = $\psi(8.3499768) + \ln(2038.3101)$

$$= 2.0611878 + 7.618764 \quad = 9.6810642$$

The LHS is about the same as the RHS. This concludes estimation of a, b and c.

If the coefficient of skeweness of the data is computed then

$$C_s = 0.708; \; \overline{x} = 26317.65; \; S = 18773.359$$

$$C_s = C_s [1 + \frac{8.5}{n}] \frac{[n(n-1)]^{0.5}}{n-2}$$

$$= 0.708 [1 + \frac{8.5}{34}] \frac{[34 \; x \; 33]^{0.5}}{32} = 0.9263812$$

$$b = (\frac{2}{0.9263812})^2 = 4.6610153$$

$$a = \frac{18773.359}{(4.6610153)^{0.5}} = 8695.6434$$

$$c = 26317.65 - 18773.359 \text{ x } (4.6610153)^{0.5} = -14212.877$$

$$\frac{1}{n} \sum_{i}^{n} \ln (x_i - c) = 347.30754$$

$$\psi(4.6610153) + \ln(8695.6434) = 1.428142 + 9.0705774 = 10.498719$$

$$\frac{1}{n} \sum_{i}^{n} (x_i - c) = 10.214928$$

This concludes the computation of parameters.

14.1.6 DISTRIBUTION ENTROPY

Equation (14.4) gives the distribution entropy. It can be rewritten as

$$I(x) = - \int_c^{\infty} f(x) \ln f(x) dx$$

$$= [\ln a + \ln \Gamma(b) + (b-1) \ln a - \frac{c}{a}] \int_c^{\infty} f(x) dx + \frac{1}{a} \int_c^{\infty} x \, f(x) dx$$

$$- (b-1) \int_c^{\infty} \ln(x-c) f(x)dx$$

$$= [\ln a + \ln\Gamma(b) + \ln a^{b-1} - \frac{c}{a}] + \frac{1}{a}\bar{x} - (b-1) \; E[\ln(x-c)]$$

$$= \ln \; (a^b\Gamma(b)) - \frac{c}{a} + \frac{\bar{x}}{a} - \; (b-1) \; E[\ln(x-c)]$$

$$= \ln \; (a^b\Gamma(b)) - \frac{c}{a} + \frac{\bar{x}}{a} - - (b-1) \; E[\ln(x-c)] \tag{14.45}$$

14.2 Parameter - Space Expansion Method

14.2.1 SPECIFICATION OF CONSTRAINTS

For this method equation (14.5) holds and the other two constraints one defined somewhat differently and can be written, following Singh and Rajagopal (1986), as

$$\int_c^{\infty} (\frac{x-c}{a}) f(x) dx = E[\frac{x-c}{a}] \tag{14.46}$$

$$\int_c^{\infty} \ln(\frac{x-c}{a})^{b-1} f(x) dx = E \ln[\frac{x-c}{a}]^{b-1} \tag{14.47}$$

14.2.2 DERIVATION OF ENTROPY FUNCTION

The least-biased pdf corresponding to the principle of maximum entropy (POME) and consistent with equation (14.5), (14.46), and (14.47) takes the form

$$f(x) = \exp[-\lambda_0 - \lambda_1(\frac{x-c}{a}) - \lambda_2 \ln[\frac{x-c}{a}]^{b-1}] \tag{14.48}$$

where λ_0, λ_1, and λ_2 are Lagrange multipliers.

Insertion of equation (14.45) into equation (14.5) yields

$$\exp(\lambda_0) = \int_c^{\infty} \exp[-\lambda_1(\frac{x-c}{a}) - \lambda_2 \ln[\frac{x-c}{a}]^{b-1}] dx$$

$$= a(\lambda_1)^{\lambda_2(b-1)-1} \; \Gamma(1 - \lambda_2(b-1)) \tag{14.49}$$

The zeroth Lagrange multiplier is given as

$$\lambda_0 = \ln a + (\lambda_2(b-1)-1)\ln\lambda_1 + \ln\Gamma(1-\lambda_2(b-1)) \tag{14.50}$$

The zeroth Lagrange multiplier is also obtained from equation (14.49) as

$$\lambda_0 = \int_c^{\infty} \exp[-\lambda_1(\frac{x-c}{a}) - \lambda_2 \ln[\frac{x-c}{a}]^{b-1}]dx \tag{14.51}$$

Introduction of equation (14.49) in equation (14.48) yields

$$f(x) = \frac{\lambda_2^{1-(b-1)}}{a\Gamma(1-\lambda_2(b-1))}\exp[-\lambda_1(\frac{x-c}{a}) - \lambda_2 \ln(\frac{x-c}{a})^{b-1}] \tag{14.52}$$

A comparison of equation (14.52) with equation (14.1) yields $\lambda_1=1$ and $\lambda_2=-1$.
Taking logarithm of equation (14.52) results in

$$\ln f(x) = -\ln a + (1-\lambda_2(b-1))\ln\lambda_1 - \ln\Gamma(1-\lambda_2(b-1)) - \lambda_1(\frac{x-c}{a}) + \ln(\frac{x-c}{a})^{-\lambda_2(b-1)} \tag{14.53}$$

Thus, the entropy I(f) of the PT III distribution can be expressed as

$$I(f) = +\ln a - (1-\lambda_2(b-1))\ln\lambda_1 + \ln\Gamma(1-\lambda_2(b-1)) + \lambda_1 E[\frac{x-c}{a}]$$

$$+\lambda_2(b-1)E\ [\ln(\frac{x-c}{a})] \tag{14.54}$$

This is the entropy function of PT III distribution.

14.2.3 RELATION BETWEEN PARAMETERS AND CONSTRAINTS

According to Singh and Rajagopal (1986) the relation between distribution parameters and constraints is obtained by taking partial derivatives of the entropy function given by equation (14.54) with respect to Lagrange multipliers as well as distribution parameters and then equating these derivatives each to zero. To that end, taking partial derivative of equation (14.54) with respect to a, b, c, λ_1, and λ_2 separately, and equating each derivative to zero yields:

$$\frac{\partial I}{\partial \lambda_1} = 0 = -(1-\lambda_2(b-1))\frac{1}{\lambda_1} + E(\frac{x-c}{a}) \tag{14.55}$$

$$\frac{\partial I}{\partial \lambda_2} = 0 = +(b-1)\ln\lambda_1 + (b-1)E[\ln(\frac{x-c}{a})] - (b-1)\psi(1-\lambda_2(b-1)) \tag{14.56}$$

where ψ is a digamma function = d [*ln* Γ(x)]/dx.

$$\frac{\partial I}{\partial a} = 0 = +\frac{1}{a} - \frac{\lambda_1}{a}E\,[\frac{x-c}{a}] - \frac{\lambda_2(b-1)}{a} \tag{14.57}$$

$$\frac{\partial I}{\partial b} = 0 = +\lambda_2\ln\lambda_1 + \lambda_2 E[\ln(\frac{x-c}{a})] - \lambda_2\psi(1-\lambda_2(b-1)) \tag{14.58}$$

$$\frac{\partial I}{\partial c} = 0 = -\frac{\lambda_1}{a} - \frac{\lambda_2(b-1)}{a}E[\frac{a}{x-c}] \tag{14.59}$$

Note that $\lambda_1=1$ and $\lambda_2=-1$. Simplification of equations (14.55) to (14.59), respectively, yields

$$E[\frac{x-c}{a}] = b \tag{14.60}$$

$$E[\ln(\frac{x-c}{a})] = \psi(K), K = (1-\lambda_2(b-1)) \tag{14.61}$$

$$E[\frac{x-c}{b}] = b \tag{14.62}$$

$$E[\ln(\frac{x-c}{a})] = \psi(K) \tag{14.63}$$

$$E[\frac{a}{x-c}] = \frac{1}{b-1} \tag{14.64}$$

Equations (14.61) and (14.63) are the same, and so are equations (14.60) and (14.62). Therefore, the parameter estimation equations are equations (14.60), (14.63), and (14.64).

14.3 Other Methods of Parameter Estimation

14.3.1 METHOD OF MOMENTS

The r-th moment M_r^c of equation (14.1) about a point 'c' is

$$M_r^c = \int_c^\infty (x-c)^r \frac{1}{a\Gamma(b)} \left(\frac{x-c}{a}\right)^{b-1} \exp[-(x-c)/a]\, dx \tag{14.65}$$

Let z = (x–c)/a. Then equation (14.65) reduces to

$$M_r^c = \frac{a^r}{\Gamma(b)} \int_0^\infty z^{r+b-1} \exp(-z) dz = \frac{a^r}{\Gamma(b)} \Gamma(r+b) \tag{14.66}$$

Since the PT III distribution has three parameters, it will suffice to determine the first three moments for the method of moments (MOM).

$$M_1^c = a\,b \tag{14.67}$$

$$M_2^c = a^2\, b(b+1) \tag{14.68}$$

$$M_3^c = a^3\, b(b+1)(b+2) \tag{14.69}$$

The moments given by equation (14.66) can be converted to the moments M_r^0 about the origin by using the following expression:

$$M_r^0 = \sum_{j=0}^{r} \binom{r}{j} M_{r-j}^c c^j \tag{14.70}$$

Specifically,

$$M_1^0 = a\,b + c \tag{14.71}$$

$$M_2^0 = a^2 b^2 + a^2 b + 2\,a\,b\,c + c^2 \tag{14.72}$$

$$M_3^0 = a^3 b(b+2)(b+1) + 3\,a^2 b\,c(b+1) + 3\,a\,b\,c^2 + c^3 \tag{14.73}$$

Therefore,

$$M_1^0 = \bar{x} = a\,b + c \tag{14.74}$$

$$M_2 = \mathrm{Var}(x) = M_2^0 - (M_1^0)^2 = a^2 b \tag{14.75}$$

$$M_3 = M_3^0 - M_2^0 M_1^0 + 2(M_1^0)^3 = 2\,a^3\,b \tag{14.76}$$

where M_2 and M_3 are the second and third moments of the distribution about the centroid. Thus, the equations to determine a, b, and c are

$$\bar{x} = a\,b + c \tag{14.77}$$

$$S_x^2 = a^2\,b \tag{14.78}$$

$$C_s = M_3 / M_2^{3/2} = 2/b^{0.5} \tag{14.79}$$

Bobee and Robitaille (1975) have shown that the sample coefficient of skewness given by equation (14.79) should be corrected for bias as

$$C_s = C_s{*}\,[1 + (8.5)/n]\,[n(n-1)]^{0.5}/(n-2) \tag{14.80}$$

where n is the sample size and C_s* is the biased value of C_s.

14.3.2 METHOD OF MAXIMUM LIKELIHOOD ESTIMATION

For the method of maximum likelihood estimation (MLE), the likelihood function, L, of receiving the sample data D ≡ $\{x_1, x_2, \ldots, x_n\}$ from the PT III distribution, given the values of a, b, and c, is:

$$L(D \mid a,b,c) = \prod_{i=1}^{n} f(x) \tag{14.81}$$

Therefore,

$$L(D \mid a,b,c) = \frac{1}{a^n [\Gamma(b)]^n} \left(\frac{x_1 - c}{a} \cdots \frac{x_n - c}{a}\right)^{b-1}$$

$$\exp\left[-\left(\frac{x_1 - c}{a} + \ldots + \frac{x_n - c}{a}\right)\right] \tag{14.82a}$$

The MLE method involves finding the values of a, b, and c which simultaneously maximize the likelihood of observing the data from a PT III population. If L(D | a, b, c) is maximal, then so is ln L(D | a, b, c).

$$\ln L = -n \ln a - n \ln \Gamma(b) + (b-1) \sum_{i=1}^{n} \ln(x_i - c) - n\,(b-1) \ln a - \frac{1}{a} \sum_{i=1}^{n} (x_i - c) \tag{14.82b}$$

Thus, values of a, b, and c are sought which produce

$$\frac{\partial}{\partial a}[\ln L(D|a,b,c)] = 0 \tag{14.83}$$

$$\frac{\partial}{\partial b}[\ln L(D|a,b,c)] = 0 \tag{14.84}$$

$$\frac{\partial}{\partial c}[\ln L(D|a,b,c)] = 0 \tag{14.85}$$

Equations (14.83)-(14.85) lead to following estimation equations:

$$\frac{nb}{a} - \frac{1}{a^2}\sum_{i=1}^{n}(x_i - c) = 0 \tag{14.86}$$

$$n\psi(b) - \sum_{i=1}^{n}\ln(x_i - c) + n\ln a = 0 \tag{14.87}$$

$$(b-1)\sum_{i=1}^{n}\left(\frac{1}{x_i - c}\right) - \frac{n}{a} = 0 \tag{14.88}$$

where $\psi(b) = d[\ln \Gamma(b)]/db$ is the digamma function. Equations (14.86)-(14.88) are solved iteratively, as done by Matalas and Wallis (1973). It must be noted that a solution does not always exist for very small sample skew C_S . Furthermore, when b is less than 1, there is no solution; thus, the coefficient of skewness must not exceed the value of 2. When the skewness coefficient is greater than 2, the conditional MLE method of Bobee and Ashkar (1991) is recommended for use.

14.3.3 METHOD OF PROBABILITY-WEIGHTED MOMENTS

The PT III distribution cannot be expressed in inverse form and therefore its probability-weighted moments (PWMs) are not easy to obtain. However, Hosking (1986, 1990) derived these and are given in terms of L-moments as follows:

$$L_1 = c + ab \tag{14.89}$$

$$L_2 = \frac{1}{\sqrt{\pi}}\, a\, \frac{\Gamma(b+0.5)}{\Gamma(b)} \tag{14.90}$$

$$\tau_3 = \frac{L_3}{L_2} = 6\, I_{1/3}(b, 2b) - 3 \tag{14.91}$$

where L_i , i=1,2,3, are L-moments, and I_x (., .) is the incomplete beta function. Hosking (1991) gave an approximate solution of equation (14.91) which is reproduced here. For τ_3 greater than or equal to 1/3 and $t = 1 - \tau_3$,

$$b = \frac{0.36067\,t - 0.59567\,t^2 + 0.25361\,t^3}{1 - 2.78861\,t + 2.5096\,t^2 - 0.77045\,t^3} \tag{14.92}$$

and for τ_3 less than 1/3 and t = $3\,\pi\,\tau_3{}^2$,

$$b = \frac{1 + 0.2906\,t}{t + 0.1882\,t^2 + 0.0442\,t^3} \tag{14.93}$$

With the value of b obtained as above, the value of a is obtained from equation (14.90) and then the nvalue of c from equation (14.89).

14.4 Comparative Evaluation of Estimation Methods

The PT III distribution has found particular application in stochastic analysis of annual flood discharges. Although the distribution followed precisely by floods is unknown, the application of PT III distribution is justified by a combination of experience and its goodness of fit to empirical data. Singh and Singh (1985) applied the MOM, MLE and POME methods of parameter estimation to annual maximum discharge data for six selected rivers, and their work is summarized here. Pertinent characteristics of the data are given in Table 14.1. These data were selected on the basis of length, completeness, homogeneity, and independence of record. Each gaging station had a record length of more than 30 years.

Table 14.1 Pertinent characteristics of data of six selected rivers

River Gaging Station	N	Mean $\bar{x}$ (m^3/s)	Standard Deviation S_x	Skeweness C_s	Kurtosis K_s
St. Francis River near Connors, New Brunnswick	31	215.3	84.4	0.53	3.25
Fish River near Fort Kent, Maine	53	241.1	71.4	0.30	3.45
St. John River below Fish River, at Fort Kent, Maine	56	2405.2	754.1	0.43	3.22
St. John River at Ninemile Bridge, Maine	32	699.0	223.7	0.41	3.01
St. John River at Dickey, Maine	36	1449.7	517.7	0.35	2.55
Allagash River near Allagash, Maine	51	438.8	159.8	0.71	3.30

Table 14.2 Parameter estimates by MOM, MLE and POME methods.

River Gaging Station	Values of Parameters	Methods of Parameter Estimation		
		MOM	MLE	POME
St. Francis River near Connors, New Brunnswick	a	28.33	27.74	27.70
	b	8.86	9.10	9.03
	c	– 35.84	– 35.84	– 35.73
Fish River near Fort Kent, Maine	a	12.48	12.35	12.35
	b	32.69	33.03	33.00
	c	– 167.00	– 167.00	– 167.50
St. John River below Fish River, at Fort Kent, Maine	a	188.30	188.40	188.45
	b	16.03	16.02	16.01
	c	– 614.80	– 614.80	– 614.50
St. John River at Ninemile Bridge, Maine	a	57.69	55.35	55.35
	b	15.04	15.67	15.65
	c	– 168.60	– 168.60	– 168.60
St. John River at Dickey, Maine	a	113.20	108.00	108.00
	b	20.91	21.90	21.90
	c	–918.30	–918.30	918.30
Allagash River near Allagash, Maine	a	65.71	64.30	64.30
	b	5.91	6.03	6.03
	c	50.35	50.35	50.35

Table 14.3 Root mean square error and bias by MOM, MLE and POME methods for six selected rivers.

Station	RMSE			BIAS		
	MOM	MLE	POME	MOM	MLE	POME
St. Francis River near Connors, New Brunswick	0.0941	0.0995	0.0977	1.3667	1.3724	1.3927
Fish River near Fort Kent, Maine	0.1010	0.1029	0.1010	1.7459	1.7536	1.7603
St. John River below Fish River at Fort Kent, Maine	0.1600	0.1597	0.1594	1.5142	1.5939	1.5910
St. John River at Nine Mile Bridge, Maine	0.0667	0.0690	0.0685	1.1603	1.2021	1.2038
St. John River at Dickey, Maine	0.1065	0.1037	0.1037	1.5754	1.5710	1.5710
Allagash River near Allagash, Maine	0.0929	0.0929	0.0929	1.7115	1.7035	1.7305

The parameters estimated are summarized in Table 14.2. For two sample gaging stations a comparison of observed and computed frequency curves is shown in Figures 14.1 and 14.2. The observed frequency curve was obtained by using the Gringorton plotting position formula. The parameter estimates obtained by the POME and MLE methods were almost identical. Consequently, the frequency curves obtained from these methods were also identical. POME does not require the use of coefficient of skewness whereas MOM does. In this way the bias is reduced when POME is used to estimate the parameters of PT III distribution.

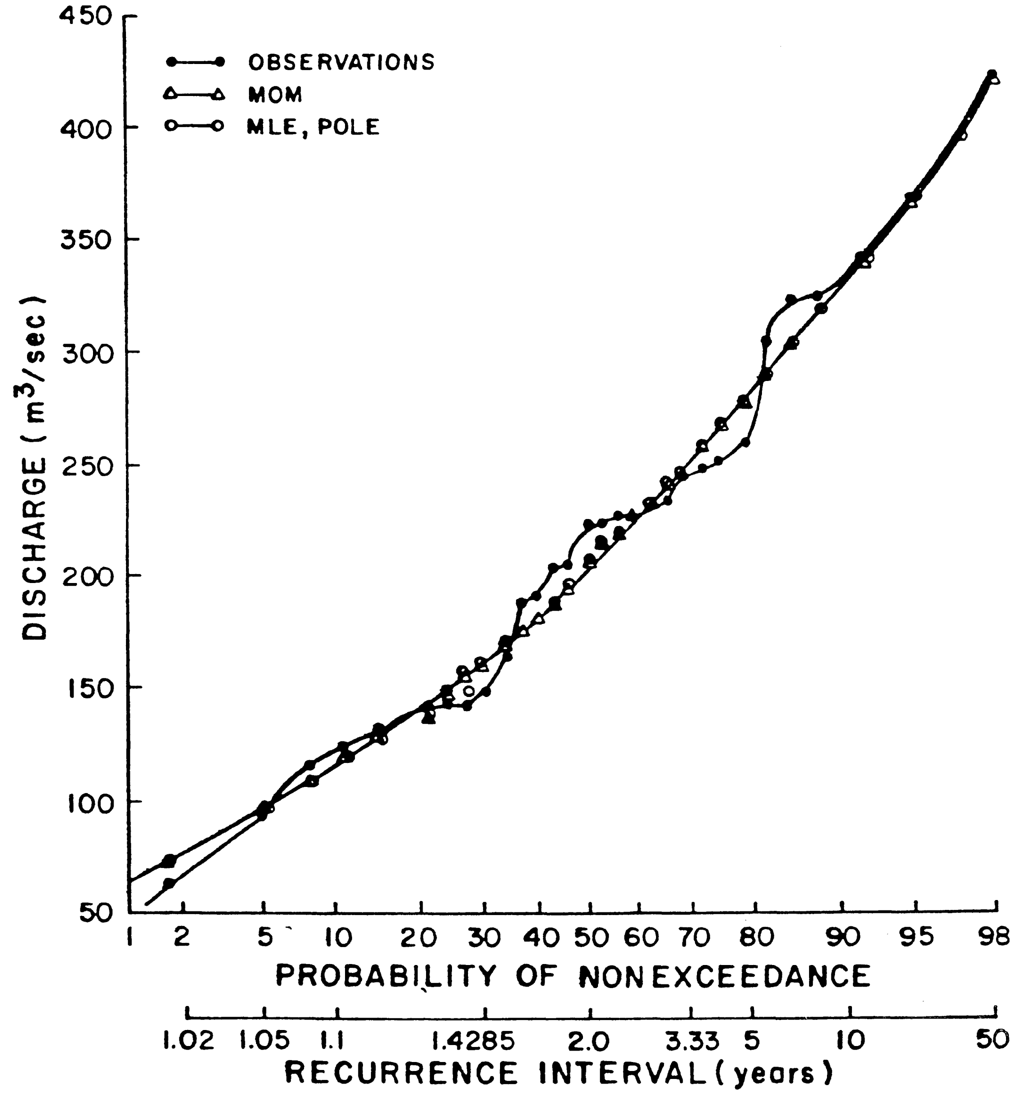

Figure 14.1 Comparison of observed and computed frequency curves for annual maximum discharge series for the St. Francis River basin near Connors, New Brunswick, Maine.

To compare these methods further, the root mean square error (RMSE) and bias were computed as given in Table 14.3. For a given river, the values of RMSE of different methods differed from one another at the second or third decimal place only. This means that the three methods were comparable for the data used. It should be emphasized that the smallest RMSE value does not necessarily lead to the best fit. A comparison of BIAS values of the three methods also reflected what was indicated by the RMSE values.

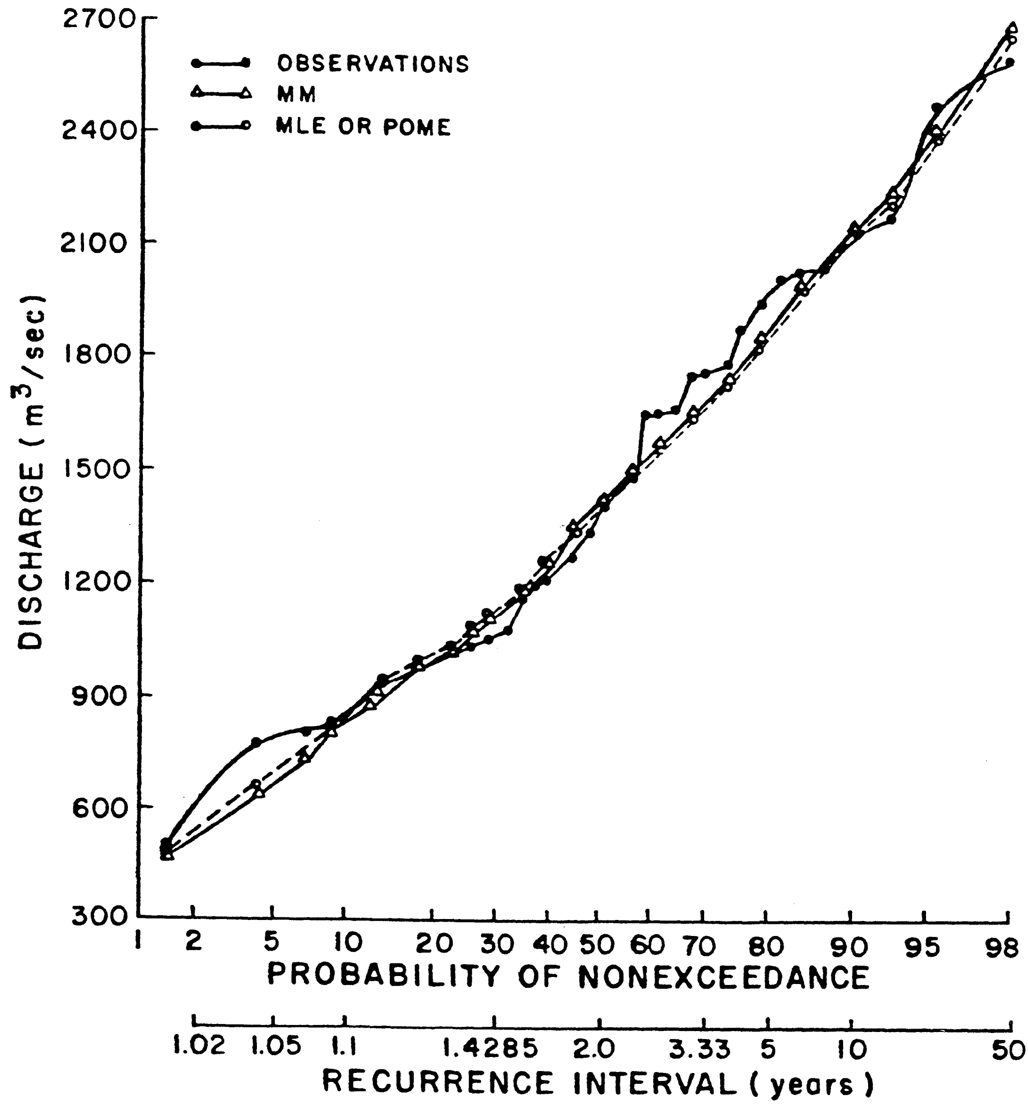

Figure 14.2 Comparison of observed and computed frequency curves for annual maximum discharge series for the St. John River basin at Dickey, Maine.

References

Bobee, B., 1973. Sample error of T-year events computed by fitting a Pearson type-3 distribution. *Water Resources Research*, Vol. **9**, No. 5, pp. 1264-1270.

Bobee, B. and Ashkar, F., 1991. *The Gamma Family and Derived Distributions Applied in Hydrology.* Water Resources Publications, Littleton, Colorado.

Bobee, B. and Robitaille, R., 1975. Correction of bias in estimation of the coefficient of skewness. *Water Resources Research*, Vol.**11**, No. 6, pp. 851-854.

Bobee, B.B. and Robitaille, R., 1977. The use of the Pearson type 3 and log Pearson type 3 distributions revisited. *Water Resources Research*, Vol. **13**, No. 2, pp. 427-443.

Buckett, J. and Oliver, F.R., 1977. Fitting the Pearson type 3 distribution in practice. *Water Resources Research*, Vol.**13**, No. 5, pp. 851-852.

Chang, S.K. and Moore, S.M., 1983. Flood frequency analysis for mall watersheds in southern Illinois. *Water Resources Research*, Vol. **19**, No. 2, pp. 277-282.

Ding, J. and Yang, R., 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 three distribution. *Journal of Hydrology*, Vol. **101**, pp. 63- 81.

Durrans, S.R., 1992. Parameter estimation for the Pearson type 3 distribution using order statistics. *Journal of Hydraulic Engineering*, Vol. **133**, pp. 215-232.

Haktanir, T., 1991. Practical computation of gamma frequency factors. *Hydrological Sciences Journal*, Vol. **36**, No. 6, pp. 599-610.

Harter, H.L., 1968. New tables of percentage points of the Pearson type III distribution. *Technical Release* No. **38**, Central Technical Unit, Soil Conservation Service, Engineering Division, U.S. Department of Agriculture, Washington, D.C.

Hoshi, K. and Leeyavanjia, U., 1986. A new approach to parameter estimations of gamma-type distributions. *Journal of Hydroscience and Hydraulic Engineering*, Vol. **4**, No. 2, pp. 79-95.

Hosking, J.R.M., 1986. The theory of probability weighted moments. *Research Report* RC **12210**, IBM Research Division, T.J. Watson Research Center, Yorktown Heights, New York.

Hosking, J.R.M., 1990. L-moments: Analysis and estimation of distributions using linear combinations of order statistics. *Journal of Royal Statistical Society*, Series B, Vol. **52**, pp. 105-124.

Hosking, J.R.M., 1991. Approximations for use in constructing L-moment ratio diagrams. *Research Report* RC **16635**, IBM Research Division, T.J. Watson Research Center, Yorktown Heights, New York.

Lall, U. and Beard, L.R., 1982. Estimation of Pearson type 3 moments. *Water Resources Research*, Vol. **18**, No. 5, pp. 1563-1569.

Markovic, R.D., 1963. Probability functions of best fit to distributions of annual precipitation and runoff. *Hydrology Paper* No. **8**, 33 pp., Colorado State University, Fort Collins, Colorado.

Matalas, N.C., 1963. Probability distribution of low flows. *U.S. Geological Survey, Professional Paper* **434-A**, pp. 1-27, Washington, D.C.

Matalas, N.C. and Wallis, J.R., 1973. Eureka! It fits a Pearson type 3 distribution. *Water Resources Research*, Vol. **9,** No. 2, pp.281-289.

Obeysekera, J.T. and Yevjevich, Y., 1985. A note on simulation of samples of gamma-autoregressive variables. *Water Resources Research*, Vol. **21**, No. 10, pp. 1569-1572.

Ribeiro-Correa, J. and Rousselle, J., 1993. A hierarchical and empirical Bayes approach for the regional Pearson type III distribution. *Water Resources Research*, Vol. **29**, No. 2, pp.435- 444.

Shaligram, V.M. and Lele, V.S., 1978. Analysis of hydrologic data using Pearson type III distribution. *Nordic Hydrology*, Vol. **9**, pp. 31-42.

Singh, V.P. and Rajagopal, A.K., 1986. A new method of parameter estimation for hydrologic frequency analysis. *Hydrological Science and Technology*, Vol. **2**, No. 3, pp. 33-40.

Singh, V.P. and Singh, K., 1985. Derivation of the Pearson type (PT) III distribution by using the principle of maximum entropy (POME). *Journal of Hydrology*, Vol. **80**, pp. 197-214.

Wilk, M.B., Gnanadesikan, R. and Huyett, M.J., 1962. Probability plots for the gamma distribution. *Technometrics*, Vol. **4,** No. 1, pp. 1-20.

Wu, B., Hou, Y. and Ding, J., 1991. The method of lower-bound to estimate the parameters of a Pearson type III distribution. *Hydrological Sciences Journal*, Vol. **36,** No. 3, pp. 271-280.

CHAPTER 15

LOG-PEARSON TYPE III DISTRIBUTION

The log-Pearson type 3 (LP3) distribution has been one of the most frequently used distributions for hydrologic frequency analyses since the recommendation of the Water Resources Council (1967, 1982) of the United States as to its use as the base method. The Water Resources Council also recommended that this distribution be fitted to sample data by using mean, standard deviation and coefficient of skewness of the logarithms of flow data [i.e., the method of moments (MOM)]. A large volume of literature on the LP3 distribution has since been published with regard to its accuracy and methods of fitting or parameter estimation. McMahon and Srikanthan (1981) and Srikanthan and McMahon (1981) examined the applicability of LP3 distribution to Australian rivers and questioned the assumption of setting to zero the coefficient of skewness of logarithms of peak discharges that were not statistically different from zero. They evaluated the effect of sample size, distribution parameters and dependence on peak annual flood estimates. Gupta and Deshpande (1974) applied LP3 distribution to evaluate design earthquake magnitudes. Phien and Jivajirajah (1984) applied LP3 distribution to annual maximum rainfall, annual streamflow and annual rainfall. Wallis and Wood (1985) found, based on Monte Carlo experiments, that the flood quantile estimates obtained by using an index flood type approach with either a generalized extreme value distribution or a Wakeby distribution fitted by PWM were superior to those obtained by LP3 distribution with MOM -based parameters. This finding was challenged later by several investigators (Beard, 1986; Landwehr et al., 1986).

The Water Resources Council recommended the use of a generalized skew coefficient. Bobee and Robitaille (1975) proposed a correction for bias in estimation of the coefficient of skewness. Tung and Mays (1981) investigated various methods of determining generalized skew coefficients. They introduced a method for determining generalized skew coefficients using a weighting procedure based on the varaince of regional map skew coefficients and variance of sample skew coefficients. Oberg and Mades (1987) evaluated several techniques of estimating generalized skew of LP3 distribution for Illinois rivers: (1) a generalized skew map of U.S., (2) an isoline map, (3) a prediction equation, (4) a regional mean skew. They found no appreciable difference between flood estimates computed using the variations of the regional mean technique and flood estimates using the skew map prepared by the Water Resources Council. Bobee (1975) showed that the method of moments recommended by the Water Resources Council (1967) would introduce bias in fitting LP3 distribution because the method of moments (MOM) used logarithms of observed data and not the moments of the observed values. He used a method which retained the moments of the original data. Ashkar and Bobee (1987) and Bobee and Ashkar (1988) used four different versions of MOM and obtained a generalized MOM (GMOM).

They concluded that one version of GMOM might be best for estimating high flows (flows above the range covered by the data) and another for estimating low flows.

Using simulation, Ouarda and Ashkar (1998) evaluated the effect of trimming on LP3 flood quantile estimates. They examined the effect of various proportions of symmetric trimming on the estimation of moments, distribution parameters, and quantiles. The influence of sample size and parent distribution parameters on the estimation performance was also investigated for several return periods and for three fitting methods. Cohen et al. (1997) developed the expected moments algorithm (EMA) for computing moments-based flood quantile estimates when historical flood information was available. EMA can utilize three types of at-site flood information: systematic stream gage record, information about the magnitude of historical flood, and knowledge of the number of years in the historical period when no large flood occurred. Based on Monte Carlo simulations, they showed that EMA was more efficient than MOM and nearly as efficient as MLE. Rao (1980a, b, 1981, 1983a,b 1988) evaluated the properties and results of LP3 distribution in a general fashion and proposed a method of mixed moments (MIX) to estimate LP3 parameters. He found that MIX using the mean and variance of real data and the mean of log data possessed superior statistical properties to MOM. Song and Ding (1988) applied the probability weighted moments (PWM) to estimate the parameters of LP3 distribution. Their results of Monte Carlo experiments showed that PWM compared favorably with MOM and was equally efficient when compared with MLE and other curve fitting methods. Condie (1977) derived LP3 parameters using the method of maximum likelihood estimation (MLE). By fitting 37 long-term unregulated flood data sets in Canada, he concluded that MLE was superior to MOM. However, Nozdryn-Plotnicki and Watts (1979) found in their simulation study of the standard error of the T-year flood that MLE and MOM were almost comparable, and hence they suggested the use of MOM because of its computational ease.

Phien and Hsu (1985) compared a number of techniques for LP3 parameter estimation. These were MOM and modified versions of MOM. Singh and Singh (1988) estimated LP3 parameters using the principle of maximum entropy (POME) and found it comparable to MOM and MLE for historical data used. Phien and Hira (1983) estimated LP3 parameters using four methods: MLE, direct and indirect MOM, MIX, method of Bobee (1975), and other versions of MOM. They found the MIX method, consisting of the first two moments of the original data and the variance of the log-transformed values, to be providing the best estimates. Arora and Singh (1987a, b, 1989) made a comparative evaluation of different estimators of LP3 distribution: Direct and indirect MOM, MIX, MLE, and POME. Using Monte Carlo experiments, they found MIX to be markedly superior to other methods in terms of both resistance and efficiency of estimation.

Benson (1967) reported on uniform flood-frequency estimating methods for federal agencies. Among 2-parameter gamma, Gumbel, log-Gumbel, log-normal, Hazen and LP3 distributions, the LP3 distrribution was selected as the base method, with provisions for departures where justified. Reich (1970) analyzed flood peaks from Pennsylavanian streams using Gumble, log-Gumbel, and LP3 distributions. He found the Gumbel distribution to be generally applicable. Shen et al. (1980) investigated the tail behavior in extreme events using LP3 and Gumbel distributions and found LP3 distribution to be a better description of field data. Rao (1981) compared 3-parameter distributions, including LP3, Pearson type 3 (P3), lognormal (LN3), and Weibull (W); and presented bounds, negative areas of distribution and selected quantiles.The choice of the best distribution was not clear and depended on the sample statistics and the choice of the T-year flood. Loganathan et al. (1986) analyzed frequencies of low flows using a mixed LP3, a double bounded probability density function, partial duration series, and a physically based approach. The results of LP3 model were consistent with other methods.

Tasker (1987) estimated 7-day, 10-year and 7-day 20-year low flows using bootstrap using the hypothetical LP3 and W distributions. The use of these distributions led to lower mean square error than did the Box-Cox transformation and the log-Boughton method. In statistical modeling of annual maximum flows of Turkish rivers, Haktanir (1991) compared LN3, P3, LP3, EV1, log-Boughton, log-logistic (LL), and smemax distributions at 112 sites representing 23 major basins and and did not find a single distribution performing consistently better. LP3 and LL performed better more times than others. Vogel et al. (1992) discussed flood-flow frequency model selection in southwestern United States. Using flood flow data at 383 sites, they found LP3, generalized extreme value (GEV) and two-parameter and three-parameter LN distributions to provide a good approximation to flood flow data in this region. Bobee et al. (1993) reported on a systematic approach to comapring distributions in flood frequency analysis.

Hoshi and Burges (1981a) investigated sampling covariance structures of estimated parameters for LP3 dsitribution from sample estimates of mean, coefficient of variation and skew coefficient in the natural domain. They showed that there was no justification for use of logarithmic skew coefficients or the regional skew estimates in log space. Hoshi and Burges (1981b) developed an approximate method for computing the derivative of a standard gamma quantile with respect to the distribution shape parameter necessary for estimating the sampling variance of a specified quantile. Ashkar and Bobee (1988) derived confidence intervals for flood events under LP3 distribution. Condie (1977) used MLE to derive the T-year event and its asymptotic standard error. Philon and Admowski (1993) derived the asymptotic standard error of estimate of the T-year flood.

Let Y = ln X where X is a positive random variable. If Y has a Pearson type (P) III distribution then X will have a log-Pearson type (LP) III distribution with probability density function (pdf) given by

$$f(x) = \frac{1}{a\,x\,\Gamma(b)} \left(\frac{\ln x - c}{a}\right)^{b-1} \exp\left[-\left(\frac{\ln x - c}{a}\right)\right] \tag{15.1}$$

where a > 0, b > 0 and 0 < c < ln x are the scale, shape and location parameters, respectively. The LP IIII distribution is a three-parameter distribution. Its cumulative distribution function (cdf) can be expressed as

$$F(x) = \frac{1}{a\,\Gamma(b)} \int_0^{\infty} \frac{1}{x} \left(\frac{\ln x - c}{a}\right)^{b-1} \exp\left(-\left(\frac{\ln x - c}{a}\right)\right] dx \tag{15.2}$$

One can verify if f(x) given by equation (15.1) is a pdf as follows:

$$\int_{e^c}^{\infty} f(x)\, dx = 1 \tag{15.3a}$$

Substituting equation (15.1) in equation (15.3), one gets

$$\frac{1}{a\,\Gamma(b)} \int_{e^c}^{\infty} \frac{1}{x} \frac{\ln x - c}{a}\Big)^{b-1} \exp\left[-\left(\frac{\ln x - c}{a}\right)\right] dx = 1 \tag{15.3b}$$

Let $(\ln x - c)/a = y$. Then $(dy/dx) = (a/x)$, and dx = xady.

If $x = e^c$ then y = $[(\ln e^c - c)/a] = [(c \ln e - c)/a] = [(c-c)/a]$. Substituting these quantities in equation (15.3b), one gets

$$f(x) = \frac{1}{a\Gamma(b)} \int_0^\infty \frac{1}{x} y^{b-1} e^{-y} x a \, dy = \frac{1}{\Gamma(b)} \int_0^\infty e^{-y} y^{b-1} \, dy = \frac{\Gamma(b)}{\Gamma}(b) = 1$$

(15.3c)

If y = [ln (x) - c]/ a is substituted in equation (15.2) the following is the result:

$$F(y) = \frac{1}{\Gamma(b)} \int_0^y y^{b-1} \exp(-y) \, dy \qquad (15.4)$$

Equation (15.4) can be approximated by noting that F (y) can be expressed as a Chi-square distribution with degrees of freedom as 2b and chi-square as 2y. This approximation is given in Chapter 13.

It may be useful to briefly discuss some of the characteristics of the LP III distribution. To that end, the mean, variance and skewness coefficients of both X and Y are given. For Y these, respectively, are:

Mean: $\mu_y = c + a\,b$ (15.5a)

Variance: $\sigma_y^2 = b\,a^2$ (15.5b)

Skew: $\gamma_y = \frac{|a|}{a} \frac{2}{b^{1/2}}$ (15.5c)

The moments of X about the origin can be written as

$$\mu'_r = \frac{\exp(rc)}{(1 - ra)^b}, \quad 1 - ra \succ 0, r = 0, 1, 2, \ldots\ldots.. \qquad (15.6a)$$

From equation (15.6a), the mean, coefficient of variation (CV), coefficient of skewness (skew), and kurtosis of X are given as

Mean: $\mu = \frac{\exp(c)}{(1-a)^b}$ (15.6b)

$$\text{Variance: } \sigma_x^2 = \exp(2c) \bullet A \tag{15.6c}$$

$$\text{Coefficient of Variation (CV): } \beta = (1-a)^b \bullet A^{1/2} \tag{15.6d}$$

$$\text{Skew: } \gamma = \left[\frac{1}{(1-3a)^b} - \frac{3}{(1-a)^b(1-2a)^b} + \frac{2}{(1-a)^{3b}}\right] / A^{3/2} \tag{15.6e}$$

Kurtosis:

$$\lambda = \left[\frac{1}{(1-4a)^b} - \frac{4}{(1-a)^b(1-3a)^b} + \frac{6}{(1-a)^{2b}(1-2a)^b} - \frac{3}{(1-a)^{4b}}\right] \bullet A^{-2} \tag{15.6f}$$

where

$$A = \left[\frac{1}{(1-2a)^b} - \frac{1}{(1-a)^{2b}}\right] \tag{15.6g}$$

It is to be noted that the coefficient of variation, skewness, and kurtosis in equations (15.6e) - (15. 6f) are independent of the location parameter c. It should also be noted that higher order moments of order r do not exist if the value of a is greater than 1/r (Bobee and Ashkar, 1991).

Consider equation (15.5a). If $a > 0$, then skew is greater than zero, implying that Y is positively skewed and $c < Y < \infty$. In this case, X is also positively skewed (Rao, 1980a), and $\exp(c) < X < \infty$. If $a < 0$, then skew is less than zero, implying that Y is negatively skewed and $-\infty < Y < c$. In this case , X is either positively skewed or negatively skewed depending on the values of parameters a and b, and $-\infty < X < \exp(c)$. For this case the density function f (x) = 0, and may be arbitrarily defined as zero.

The overall geometric shape of the LP III distribution is governed by parameters a and b (Rao, 1980a; Bobee, 1975). The pdf is capable of assuming diverse shapes, such as reverse J, U, J, and, of course, unimodal (skewed) bell shape. Hoshi and Burges (1981a) point out that if $\gamma < \beta^3 + 3\beta$, then $a < 0$, $0 < x < \exp(c)$, kurtosis of the LP III distribution is less than the kurtosis of the three-parameter lognrmal distribution and vice versa. The LP III distribution degenerates to the lognormal distribution when parameters a and b become zero and infintely, respectively (or equivalently, $\gamma = \beta^3 + 3\beta$, and $\gamma = 0$). For flood frequency analysis, only values of b greater than one and 1/a greater than zero are of interest. Negative coefficients of skew correspond to negative a values and the distribution would then become upper bounded. Under these conditions, this might be considered for low flow analysis but would be unsuitable for flood analysis.

15.1 Ordinary Entropy Method

15.1.1 SPECIFICATION OF CONSTRAINTS

Taking logarithm of equation (15.1.) to the base 'e' , one obtains

$$\ln f(x) = -\ln a\Gamma(b) - \ln x + (b-1)\ln[\frac{\ln x - c}{a}] - (\frac{\ln x - c}{a})$$
$$= -\ln a\,\Gamma(b) - \ln x + (b-1)\ln[\ln x - c] - (b-1)\ln a - \frac{\ln x}{a} + \frac{c}{a} \tag{15.7a}$$

Multiplying equation (15.7a) by -1, we get

$$-\ln f(x) = \ln a\,\Gamma(b) - \frac{c}{a} + (b-1)\ln a + (1 + \frac{1}{a})\ln x - (b-1)\ln[\ln x - c] \tag{15.7b}$$

Multiplying equation (15.7b) by f(x) and integrating between e^c and ∞, the result is the entropy function:

$$-\int_{e^c}^{\infty} f(x)\ln f(x)\,dx = [\ln a\,\Gamma(b) - \frac{c}{a} + (b-1)\ln a]\int_{e^c}^{\infty} f(x)\,dx$$
$$+ (\frac{a+1}{a})\int_{e^c}^{\infty} \ln x f(x)\,dx - (b-1)\int_{e^c}^{\infty} \ln(\ln x - c)f(x)\,dx \tag{15.7c}$$

From equation (15.7c) the constraints appropriate for equation (15.1) can be written as:

$$\int_{e^c}^{\infty} f(x)\,dx = 1 \tag{15.8}$$

$$\int_{e^c}^{\infty} \ln x f(x)\,dx = E[\ln x] = \bar{y} \tag{15.9}$$

$$\int_{e^c}^{\infty} \ln(\ln x - c)f(x)\,dx = E[\ln(\ln x - c)] \tag{15.10}$$

15.1.2 CONSTRUCTION OF ZEROTH LAGRANGE MULTIPLIER

The least-biased pdf f(x), consistent with equations (15.8) to (15.10) and based on the principle of maximum entropy (POME), takes the form:

$$f(x) = \exp[-\lambda_0 - \lambda_1 \ln x - \lambda_2 \ln(\ln x - c)] \tag{15.11}$$

where λ_0, λ_1, and λ_2 are Lagrange multipliers. Substitution of equation (15.11) in equation (15.8) yields

$$\int_{e^c}^{\infty} f(x)\,dx = \int_{e^c}^{\infty} \exp[-\lambda_0-\lambda_1 \ln x-\lambda_2 \ln(\ln x-c)]\,dx = 1 \qquad (15.12)$$

Equation (15.12) gives the partition function function as

$$\exp(\lambda_0) = \int_{e^c}^{\infty} \exp[-\lambda_1 \ln x-\lambda_2 \ln(\ln x-c)]dx \qquad (15.13)$$

Equation (15.13) is simplified as

$$\exp(\lambda_0) = \int_{e^c}^{\infty} \exp[\ln y^{-\lambda_1}] \exp[\ln(\ln x-c^{-\lambda_2})]\,dx = \int_{e^c}^{\infty} x^{-\lambda_1}(\ln x-c)^{-\lambda_2}\,dx \qquad (15.14)$$

Let ln x–c=y. Then, ln x = y + c; x = exp(y+c); (dy/dx) = (1/x); dx = xdy; and dx = exp(y+c) dy. Substituting these quantities in equation (15.14), we get

$$\begin{aligned}\exp(\lambda_0) &= \int_0^{\infty} [e^{y+c}]^{-\lambda_1} y^{-\lambda_2} e^{y+c}\,dy = \int_0^{\infty} (e^y e^c)^{-\lambda_1} y^{-\lambda_2} e^y e^c\,dy \\ &= \exp[c-c\lambda_1] \int_0^{\infty} \exp[-\lambda_1 y + y]\, y^{-\lambda_2}\,dy \\ &= \exp[-c(\lambda_1-1)] \int_0^{\infty} \exp[-y(\lambda_1-1)]\, y^{-\lambda_2}\,dy\end{aligned} \qquad (15.15)$$

Let $y(\lambda_1-1) = z$. Then $y = [z/(\lambda_1-1)]$, and $(dz/dy) = (\lambda_1-1)$. Therefore, equation (15.15) becomes

$$\begin{aligned}\exp(\lambda_0) &= \exp[-c(\lambda_1-1)] \int_0^{\infty} e^{-z} \left(\frac{z}{\lambda_1-1}\right)^{-\lambda_2} \frac{dz}{(\lambda_1-1)} \\ &= \frac{\exp[-c(\lambda_1-1)]}{(\lambda_1-1)^{1-\lambda_2}} \int_0^{\infty} z^{-\lambda_2} e^{-z}\,dz\end{aligned} \qquad (15.16)$$

Since

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x}\,dx \qquad (15.17)$$

equation (15.16) reduces to the partition function:

$$\exp(\lambda_0) = \frac{\exp[-c(\lambda_1-1)]}{(\lambda_1-1)^{1-\lambda_2}} \Gamma(1-\lambda_2) \qquad (15.18)$$

Therefore, the zeroth Lagrange multiplier is obtained from equation (15.18) as

$$\lambda_0 = -c(\lambda_1 - 1) + (\lambda_2 - 1)\ln(\lambda_1 - 1) + \ln\Gamma(1-\lambda_2) \tag{15.19}$$

The zeroth Lagrange multiplier is also obtained from equation (15.13) as

$$\lambda_0 = \ln \int_{e^c}^{\infty} \exp[-\lambda_1 \ln x - \lambda_2 \ln(\ln x - c)]\, dx \tag{15.20}$$

14.1.3 RELATION BETWEEN LAGRANGE MULTIPLIERS AND CONSTRAINTS

Differentiating equation (15.11) with respect to λ_1 and λ_2, one gets

$$\frac{\partial \lambda_0}{\partial \lambda_1} = -\frac{\int_{e^c}^{\infty} \ln x \exp[-\lambda_1 \ln x - \lambda_2 \ln(\ln x - c)]\, dx}{\int_{e^c}^{\infty} \exp[-\lambda_1 \ln x - \lambda_2 \ln(\ln x - c)]\, dx}$$

$$= -\int_{e^c}^{\infty} \ln x \exp[-\lambda_0 - \lambda_1 \ln x - \lambda_2 \ln(\ln x - c)]\, dx$$

$$= -\int_{e^c}^{\infty} \ln x f(x)\, dx = -E[\ln x] \tag{15.21}$$

$$\frac{\partial \lambda_0}{\partial \lambda_2} = -\frac{\int_{e^c}^{\infty} \ln(\ln x - c) \exp[-\lambda_1 \ln x - \lambda_2 \ln(\ln x - c)\, dx}{\int_{e^c}^{\infty} \exp[-\lambda_1 \ln x - \lambda_2 \ln(\ln x - c)]\, dx}$$

$$= -\int_{e^c}^{\infty} \ln(\ln x - c) \exp[-\lambda_0 - \lambda_1 \ln x - \lambda_2 \ln(\ln x - c)]\, dx$$

$$= -\int_{e^c}^{\infty} \ln(\ln x - c) f(x)\, dx = -E[\ln(\ln x - c)] \tag{15.22}$$

Also differentiating equation (15.19) with respect to λ_1 and λ_2, respectively, one gets

$$\frac{\partial \lambda_0}{\partial \lambda_1} = -c + \frac{\lambda_2 - 1}{\lambda_1 - 1} \tag{15.23}$$

$$\frac{\partial \lambda_0}{\partial \lambda_2} = \ln(\lambda_1 - 1) + \frac{\partial}{\partial \lambda_2} \ln \Gamma(1-\lambda_2) \tag{15.24}$$

Equation (15.24) can be simplified as

$$\frac{\partial\lambda_0}{\partial\lambda_2} = \ln(\lambda_1 - 1) + \frac{\partial}{\partial p}\ln\Gamma(p)\frac{\partial p}{\partial\lambda_2} = \ln(\lambda_1 - 1) - \psi(p) \tag{15.25}$$

where $p = 1 - \lambda_2$. Since the LP III distribution has three parameters, equations (15.23) and (15.25) are not sufficient and another equation is neded. This is obtained by recalling that

$$\frac{\partial^2\lambda_0}{\partial\lambda_1^2} = \frac{(\lambda_2 - 1)}{(\lambda_1 - 1)^2}(-1) = \frac{1-\lambda_2}{(\lambda_1 - 1)^2} = \sigma_y^2 \tag{15.26}$$

Equating equations (15.21) and (15.23), as well as equations (15.22) and (15.25), one obtains

$$\frac{p}{\lambda_1 - 1} = E[\ln x] - c \tag{15.27}$$

$$\psi(p) - \ln(\lambda_1 - 1) = E[\ln(\ln x - c)] \tag{15.28}$$

15.1.4 RELATION BETWEEN LAGRANGE MULTIPLIERS AND PARAMETERS

Inserting equation (15.19) into equation (15.11), one gets

$$f(x) = \exp[c(\lambda_1 - 1) - (\lambda_2 - 1)\ln(\lambda_1 - 1) - \ln\Gamma(1 - \lambda_2) - \lambda_1 - \ln x$$

$$-\lambda_2 \ln(\ln x - c) - \ln x + \ln x]$$

$$= \exp[-(\lambda_1 - 1)(\ln x - c) - \ln x + \ln(\lambda_1 - 1)^{-(\lambda_2 - 1)}$$

$$+ \ln[\Gamma(1 - \lambda_2))^{-1} + \ln(\ln x - c)^{-\lambda_2}]$$

$$= \exp[-(\lambda_1 - 1)(\ln x - c)\frac{1}{x}(\lambda_1 - 1)^{-(\lambda_2 - 1)}\frac{(\ln x - c)^{-\lambda_2}}{\Gamma(1 - \lambda_2)}] \tag{15.29}$$

Comparing equation (15.29) with equation (15.1), one gets

$$1 - \lambda_2 = b \tag{15.30}$$

$$\lambda_1 - 1 = \frac{1}{a} \tag{15.31}$$

Then

$$\lambda_1 = 1 + \frac{1}{a} \tag{15.32}$$

$$\lambda_2 = 1 - b \tag{15.33}$$

15.1.5 RELATION BETWEEN PARAMETERS AND CONSTRAINTS

The LP III distribution has 3 parameters a, b, and c. The known constraints are related to the Lagrange multipliers by equations (15.26), (15.27) and (15.28) which, in turn, are related to the parameters by equations (15.30) and (15.31). Eliminating the Lagrange multipliers between these two sets of equations, we obtain parameters directly in terms of the constraints as

$$ab + c = E[\ln x] \tag{15.34}$$

$$\psi(b) - \ln a = E[\ln\{\ln(x-c)\}] \tag{15.35}$$

$$ba^2 = \sigma_y^2 \tag{15.36}$$

15.1.6 DISTRIBUTION ENTROPY

Equation (15.7) gives the distribution entropy. Rewriting it,

$$I(x) = -\int_{e^c}^{\infty} f(x) \ln f(x)\, dx$$

$$= [\ln a\Gamma(b) - \frac{c}{a} + \ln a^{b-1}] \int_{e^c}^{\infty} f(x)\, dx + (\frac{a+1}{a}) \int_{e^c}^{\infty} \ln x\, f(x)\, dx$$

$$-(b-1) \int_{e^c}^{\infty} \ln(\ln x - c) f(x)\, dx$$

$$= \ln a^b \Gamma(b) - \frac{c}{a} + (\frac{a+1}{a}) \bar{y} - (b-1) E[\ln(\ln x - c)]$$

$$= \ln a^b \Gamma(b) - \frac{c}{a} + (\frac{a+1}{a}) \bar{y} - (b-1) E[\ln(y-c)] \tag{15.37}$$

Alternatively, since the transformation x = e^y is monotonic with the Jacobian J(y/x) = 1/x, we can write

$$I(x) = I(y) - E[\ln |J(\frac{y}{x})|] = I(y) + \bar{y}$$

$$= \ln(a^b \Gamma(b)) + \frac{\bar{y}}{a} - \frac{c}{a} - (b-1) E[\ln(y-c)] + \bar{y}$$

$$= \ln(a^b \Gamma(b)) + \frac{a+1}{a}) \bar{y} - \frac{c}{a} - (b-1) E[\ln(y-c)] \tag{15.38}$$

which is identical to equation (15.37).

15.2 Parameter-Space Expansion Method

15.2.1 SPECIFICATION OF CONSTRAINTS

Following Singh and Rajagopal (1986), the constraints for this method are given by equation (15.8) and

$$\int_{e^c}^{\infty} [\ln x + \frac{\ln x - c}{a}]^{b-1} f(x)\, dx = E[\ln(\frac{\ln x - c}{a})^{b-1}] \tag{15.39}$$

$$\int_{e^c}^{\infty} \ln [\frac{\ln x - c}{a}]^{b-1} f(x)\, dx = E[\ln(\frac{\ln x - c}{a})^{b-1}] \tag{15.40}$$

15.2.2 DERIVATION OF ENTROPY FUNCTION

The pdf corresponding to the principle of maximum entropy (POME)and consistent with equations (15.8), (15.39), and (15.40) takes form

$$f(x) = \exp[-\lambda_0 - \lambda_1 \ln x - \lambda_1 (\frac{\ln x - c}{a}) - \lambda_2 \ln (\frac{\ln x - c}{a})^{b-1}] \tag{15.44}$$

where, λ_0, λ_1 and λ_2 are Lagrange multipliers. Insertion of equation (15.44) into equation (15.8) yields the partition function:

$$\begin{aligned} \exp(\lambda_0) &= \int_{e^c}^{\infty} \exp[-\lambda_1 \ln x - \lambda_1 (\frac{\ln x - c}{a}) - \lambda_2 \ln (\frac{\ln x - c}{a})^{b-1}]\, dx \\ &= a e^{c(1-\lambda_1)} (\frac{1}{\lambda_1 (1+a) - a})^{1-\lambda_2(b-1)} \Gamma[1 - \lambda_2(b-1))] \end{aligned} \tag{15.42}$$

The zeroth Lagrange multiplier is given by equation (15.42) as

$$\lambda_0 = \ln a + c(1 - \lambda_1) - K \ln \alpha + \Gamma(K), K = 1 - \lambda_2(b-1), \alpha = \lambda_1 (1+a) - a \tag{15.43}$$

Also, one gets the zeroth Lagrange multiplier from equation (15.28) as

$$\lambda_0 = \ln \int_{e^c}^{\infty} \exp \left[-\lambda_1 \ln x - \lambda_1 \left(\frac{\ln x - c}{a}\right) - \lambda_2 \ln \left(\frac{\ln x - c}{a}\right)^{b-1} \right] dx \tag{15.44}$$

Introduction of equation (15.43) in equation (15.41) yields

$$f(x) = \frac{e^{c(1-\lambda_1)}(\alpha)^K}{a\Gamma(K)} \exp \left[-\lambda_1 \ln x - \lambda_1 \frac{\ln x - c}{a} - \lambda_2 \ln \left(\frac{\ln x - c}{a}\right)^{b-1} \right] \tag{15.45}$$

A comparison of equation (15.45) with equation (15.1) shows $\lambda_1 = 1$, and $\lambda_2 = -1$.
Taking - logarithm of equation (15.45) leads to

$$\begin{aligned} -\ln f(x) = {} & -c(\lambda_1 - 1) + \ln a + \ln\Gamma(K) - K \ln \alpha + \lambda_1 \ln x + \lambda_1 \left(\frac{\ln x - c}{a}\right) \\ & + \lambda_2 \ln \left(\frac{\ln x - c}{a}\right)^{b-1} \end{aligned} \tag{15.46}$$

Therefore, the entropy function of the LP III distribution becomes

$$\begin{aligned} I(f) = {} & -c(\lambda_1 - 1) + \ln a + \ln\Gamma(K) - K \ln \alpha + \lambda_1 E \ \ln x] + \lambda_1 E \ \left(\frac{\ln x - c}{a}\right) \\ & + \lambda_2 E \left[\ln \left(\frac{\ln x - c}{a}\right)^{b-1} \right] \end{aligned} \tag{15.47}$$

15.2.3 RELATION BETWEEN PARAMETERS AND CONSTRAINTS

Taking partial derivatives of equation (15.47) with respect to λ_1, λ_2, a, and b separately, and equating each derivative to zero, one gets

$$\frac{\partial I}{\partial \lambda_1} = 0 = -c - K(1 + a)\psi(\alpha) + E[\ln x] + E\left(\frac{\ln x - c}{a}\right) \tag{15.48}$$

$$\frac{\partial I}{\partial \lambda_2} = 0 = -(b-1)\psi(K) + (b-1)\ln \alpha + E\left[\ln \left(\frac{\ln x - c}{a}\right)^{b-1}\right] \tag{15.49}$$

$$\frac{\partial I}{\partial \lambda_2} = 0 = -(b-1)\psi(K) + (b-1)\ln\alpha + E[\ln(\frac{\ln x - c}{a})^{b-1}] \tag{15.49}$$

$$\frac{\partial I}{\partial a} = 0 = \frac{1}{a} - K(\lambda_1 - 1)\psi(\alpha)1 - \frac{\lambda_1}{a}E(\frac{\ln x - c}{a}) - \lambda_2\frac{b-1}{a} \tag{15.50}$$

Simplification of equations (15.48) to (15.51) gives

$$E[\ln x] = c + a b \psi(1) \tag{15.52}$$

$$E[\ln(\ln x - c)] = \ln a + \psi(b) \tag{15.53}$$

$$E[\ln x] = c + ab \tag{15.54}$$

$$E[\ln(\ln x - c)] = \ln a + \psi(b) \tag{15.55}$$

Equations (15.53) and (15.55) are identical. The parameter estimation equations are equations (15.52) to (15.54).

15.3 Other Methods of Parameter Estimation

15.3.1 METHOD OF MOMENTS (DIRECT)

The direct method of moments (MOMD) (Bobee, 1975) uses sample estimates of moments of untransformed (real) data. Using equation (15.6a) we can write

$$\ln \mu_1' = c - b\ln(1-a) \tag{15.56}$$

$$\ln \mu_2' = 2c - b\ln(1-2a) \tag{15.57}$$

$$\ln \mu_3' = 3c - b\ln(1-3a) \tag{15.58}$$

Equations (15.56)-(15.58) can be rearranged to yield

$$\frac{\ln \mu_3' - 3\ln \mu_1'}{\ln \mu_2' - 2\ln \mu_1'} = \frac{3\ln(1-a) - \ln(1-3a)}{2\ln(1-a) - \ln(1-2a)} \; (= B,\ say) \tag{15.59}$$

For a sample under consideration, $B = [\ln \mu_3' - 3\ln \mu_1'] / [\ln \mu_2' - 2\ln \mu_1']$ can be estimated from the sample estimates of the first three moments about the origin. The right side of equation(15.59), which is a function of parameter a only (say, B (a)), reveals that a is less than 1/3. In the limit, B(a) approaches infinity, 3, and 2, as a approaches 1/3, 0, and minus infinity, respectively. It should be possible to approximate the B(a) versus a relation by a series of polynomials, as for example discussed by Kite (1978). Then a good approximation of the sample estimate of a could directly be found from the sample estimate of B and should be good enough for most fitting problems. However, for purposes of simulation, a large number of (a-B(a)) points can be generated in the region a less than 1/3 (Bobee, 1975). Subsequently, a sample estimate of a can be interpolated corresponding to the sample estimate of the B value from the generated a-B(a) points, and refined using a method such as the Newton-Raphson method applied to equation (15.59). With the interpolated value of a being a good starting solution, the iterative scheme quickly converges to the true solution to a desired degree of significant accuracy. Parameters b and c can then be estimated using equations (15.56) and (15.57).

15.3.2 METHOD OF MOMENTS (INDIRECT)

The indirect method of moments (MOMI) is basically the method advocated by the U.S. Water Resource Council (1967). This method is applied to the log-transformed data. The method uses equations (15.6b) - (15.6d) and is described in Bulletin No. 15, 17A and 17 as well as by Rao (1980b) among others. Two variations of MOMI, designated as MOMI 1 and MOMI 2, were tested by Arora and Singh (1989a, b), which essentially differ in the sample skewness estimator used for the log-transformed data:

$$g_y = \frac{n}{(n-1)(n-2)} \sum_{i=1}^{n} (y_i - \overline{y})^3 / s_y^3 \tag{15.60}$$

$$g_y' = (1 + \frac{8.5}{n}) g_y \tag{15.61}$$

where n is the sample size, and $\overline{y}$ and S_y are the sample mean and standard deviation, respectively, of log-transformed data.

15.3.3 METHOD OF MIXED MOMENTS

Rao (1980b, 1983) proposed the method of mixed moments (MIX) for the LP III distribution, with the objective of obviating the use of the sample skewness coefficient in parameter estimation. After use of various combinations mixing the first two moments of the untransformed and log-transformed samples he found one particular combination to be preferable on the basis of sampling properties. This method conserves the sample mean and variance of the untransformed data and the sample mean of the log-transformed data. Thus, equations (15.6b), (15.6c), and (15.5a) are solved to estimate parameters a, b, and c. An improved method, as compared with the method described by Rao (1983), was developed by Arora and Singh (1989b), and will not be repeated here.

15.3.4 METHOD OF MAXIMUM LIKELIHOOD ESTIMATION

For the method of maximum likelihood estimation (MLE), the likelihood function of a sample of n observations drawn from a log-Pearson type III distribution can be expressed as

$$L = \frac{1}{\prod_{i=1}^{n}} [a\Gamma(b)]^{-n} \prod_{i=1}^{n} [\frac{\ln x_i - c}{a}]^{(b-1)} \exp[-\frac{1}{a}\sum_{i=1}^{n} (\ln(x_i - c)] \quad (15.62)$$

The log-transformed L becomes

$$\ln L = -n \ln a - n \ln \Gamma(b) - \sum_{i=1}^{n} x_i + (b-1) \sum_{i=1}^{n} ln\{[\ln x - c]/a\} - \frac{1}{a} \sum_{i=1}^{n} (\ln x_i - c)$$

(15.63)

Differentiating equation (15.63) with each respect to a, b, and c, and equating each derivative to zero produces the following:

$$\frac{\partial(\ln L)}{\partial a} = -nab + \sum_{i=1}^{n} (\ln x_i - c) = 0 \quad (15.64)$$

$$\frac{\partial(\ln L)}{\partial b} = -n\psi(b) + \sum_{i=1}^{n} [\frac{(\ln x_i - c)}{c}] = 0 \quad (15.65)$$

$$\frac{\partial(\ln L)}{\partial c} = \frac{n}{a} - (b-1)\sum_{i=1}^{n} \frac{1}{(\ln x_i - c)} = 0 \quad (15.66)$$

Equations (15.64) and (15.66) can be rearranged to give

$$a = \frac{s_1}{nb} \quad (15.67)$$

$$b = \frac{s_1 s_2}{(s_1 s_2 - n^2)}, \; s_1 = \sum_{i=1}^{n} (\ln x_i - c), \; s_2 = \sum_{i=1}^{n} \frac{1}{(\ln x_i - c)} \quad (15.68)$$

$$\ln L = -n \ln a - n \ln \Gamma(b) - \sum_{i=1}^{n} x_i + (b-1) \sum_{i=1}^{n} ln\{[\ln x - c]/a\} - \frac{1}{a} \sum_{i=1}^{n} (\ln x_i - c) \quad (15.69)$$

For a specified value of c, parameters a and b can be explicitly found from equations (15.67) and (15.68), respectively. Substitution of these values of a, b, and c in equation (15.65) yields $\partial(\ln L)/\partial b = R$ *(some residual value)*. The objective is to minimize R and this involves an iterative procedure. An efficient algorithm was developed by Arora and Singh (1988).

15.4 Comparative Evaluation of Estimation Methods

15.4.1 MONTE CARLO SIMULATION

Arora and Singh (1987a,b, 1989a, b) compared various methods of parameter esimation using Monte Carlo experiments. Noting that annual flood data generally lie in the area of the β-γ diagram delineated by $0.3 \prec \beta \prec 0.8$ and upward of 1 (Rossi et al., 1986; Wallis and Wood, 1985; Landwehr et al., 1978), they generated five cases of LP III population, representative of the real flood data, for Monte Carlo experiments. These cases are listed in Table 15.1. It is noted that $\lambda_1 \prec \lambda_{LN3} \prec \lambda_5 \prec \lambda_2 \prec \lambda_4 \prec \lambda_3$, where the subscripts of λ refer to the the LP III populations and subscript LN3 refers to the three-parameter lognormal population. For each of the population cases, 1000 random samples of size 10, 20, 30, 50, and 75 were generated, and parameters and quantiles were estimated using different parameter estimation methods. The 1000 estimated values of parameters and quantiles for each sample size and population cases were used to approximate the values of the standardized bias (BIAS), standard error (SE), and root mean square error (RMSE). Due to the limited number of random samples used, the results are not expected to reproduce the true values of BIAS, SE, RMSE, and robustness (Kuczera, 1982a, 1982b), but they do provide a means of comparing the performances of various estimation methods.

15.4.1.2 BIAS in Parameter Estimates: In general, unusually high BIAS was observed in estimates of parameters a, b, and c produced by all methods. MIX yielded considerably less bias than MOMD and was clearly superior to MOMD in terms of both mini-max BIAS and minimum average BIAS criteria. This was observed for most sample sizes and return periods.

Table 15.1 LP III population cases considered in sampling experiments (μ = 1) (after Arora and Singh, 1989a)

LPT III Population	Population Statistics			Parameters		
Cases	CV (β)	Skew (γ)	γ_y	a	b	c
Case 1	0.5	1	-0.45	-0.11832	19.82269	2.216713
Case 2	0.5	3	0.62	0.127683	10.30311	-1.407434
Case 3	0.5	5	1.12	0.205678	3.215257	-0.740366
Case 4	0.3	3	1.22	0.150978	2.681889	-0.438946
Case 5	0.7	3	0.20	0.059798	98.38009	-6.066213

15.4.1.2 RMSE in Parameter Estimates: There were wide differences in the RMSE performance of estimators, with the percent difference between the best and worst being as much as 425% for sample size of 10. Either MIX or MOMD provided the most favorable RMSE values. The MIX estimator was superior on the basis of the minimum-average RMSE criteria, and comparable to MOMD on the basis of mini-max RMSE criteria. Although MIX was expected to be the most resistant estimator, MOMD performed comparably. MIX and MOMD performed markedly superior to other methods. MOMI 1 performed poorly, as did MLE and POME.

15.4.1.3 Bias in Quantiles: In general, unusually high BIAS, SE and RMSE were observed for parameter estimates of a, b, and c of all methods. However, the intercorrelation among parameter estimates was such that reasonable quantile estimates were obtained. MOMD and MIX mostly underestimated the quantiles (negative bias), especially for T greater than 25. MIX consistently produced smaller BIAS than MOMD, and the difference became more pronounced at higher return periods. MOMI 1 and MOMI 2 mostly overestimated the quantiles (positive bias). Such trends were not discernible for MLE and POME. MOMI 1 mostly produced smaller absolute bias estimates than did MIX.

15.4.1.4 SE in Quantiles: In terms of standard error, MOMI 1 and MOMI 2 consistently produced higher standard error than other methods, especially MOMD and MIX. MOMI 2 fared worse than MOMI 1. MLE and POME seemed susceptible to smaller sample sizes, and in general produced higher standard error than other methods for such sample sizes. MIX and MOMD depicted remarkable stability even for smaller sample sizes when some of the other methods showed a deterioration in standard error. In general, MIX and MOMD outperformed other estimators in terms of SE for all population cases.

15.4.1.5 RMSE in Quantiles: As compared with other estimators, MOMI 1 and MOMI 2 performed poorly in terms of RMSE. While MLE and POME did perform well for some population cases and sample sizes, they depicted large deterioration in RMSE statistics for smaller sample sizes. MIX and MOMD consistently produced least or comparable RMSE estimates. MIX seemed to hold an edge over MOMD. Both of these estimators were remarkably stable for smaller sample sizes.

15.4.2 APPLICATION TO FIELD DATA

Singh and Singh (1988) compared POME, MOM, and MLE using annual maximum discharge data for six selected rivers. These data were selected on the basis of length, completeness, homogeneity, and independence of record. Each gaging station had a record length of more than 30 years. The methods were compared using relative mean error (RME) and relative absolute error (RAE).The parameter estimates obtained by POME and MLE were closer to each other than those for MOM. For two gaging stations observed and computed frequency curves are shown in Figures 15.1 and 15.2. The observed frequency curve was computed using the Gringorton plotting position formula. POME does not rquire the use of skewness whereas MOM does. In this way, bias is reduced when POME is used to estimate the LPT III parameters. For five of the six selected data sets, both RME and RAE yielded by POME were less than or equal to those of MLE. For only one data set, values of these measures were lower for MOM than those for POME, but the differences were marginal. For all six data sets, POME and MLE were found comparable.

References

Arora, K. and Singh, V.P., 1987a. A comparative evaluation of the estimators of commonly used flood frequency models: 1. Monte Carlo simulation. *Completion Report*, Louisiana Water Resources Research Institute, Louisiana State University, Baton Rouge, Louisiana.

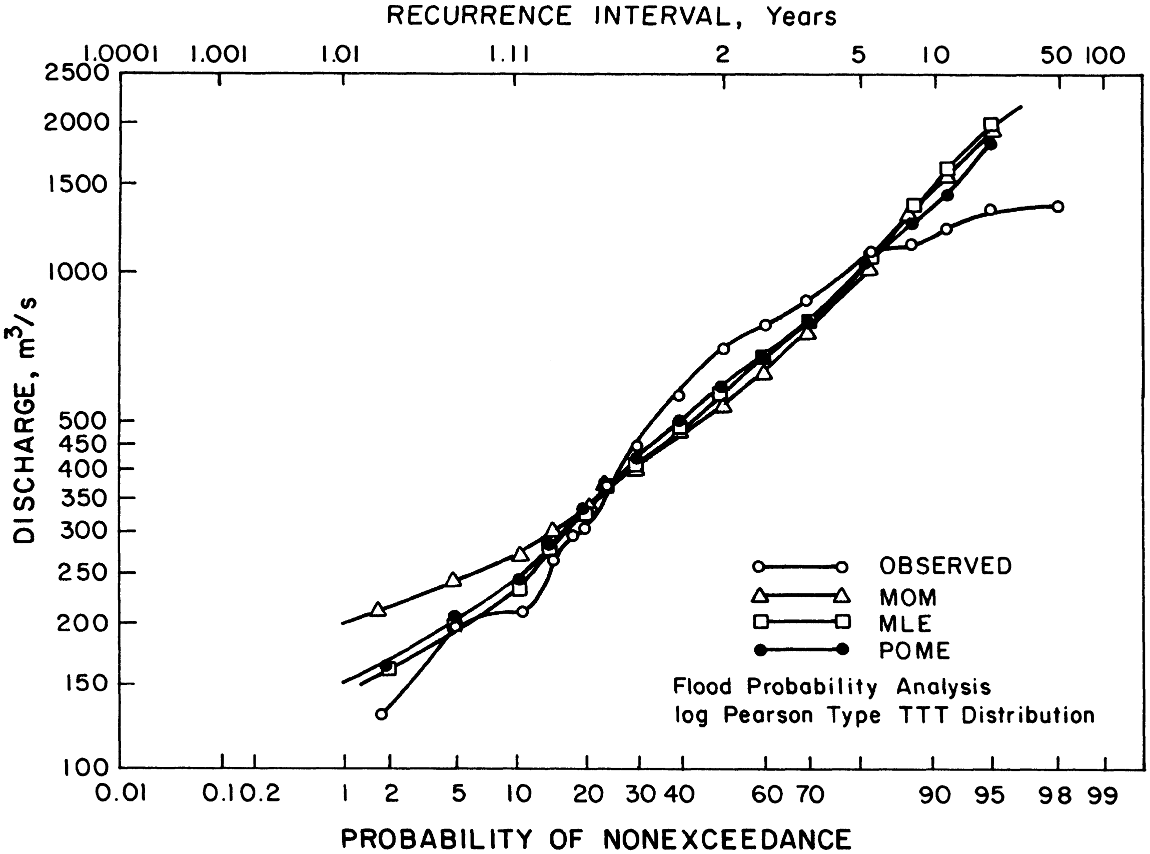

Figure 15.1 Comparison of observed and computed frequency curves using the POME, MLE and MOM methods for annual maximum discharge series for the Amite River basin at Magnoloia, Louisiana.

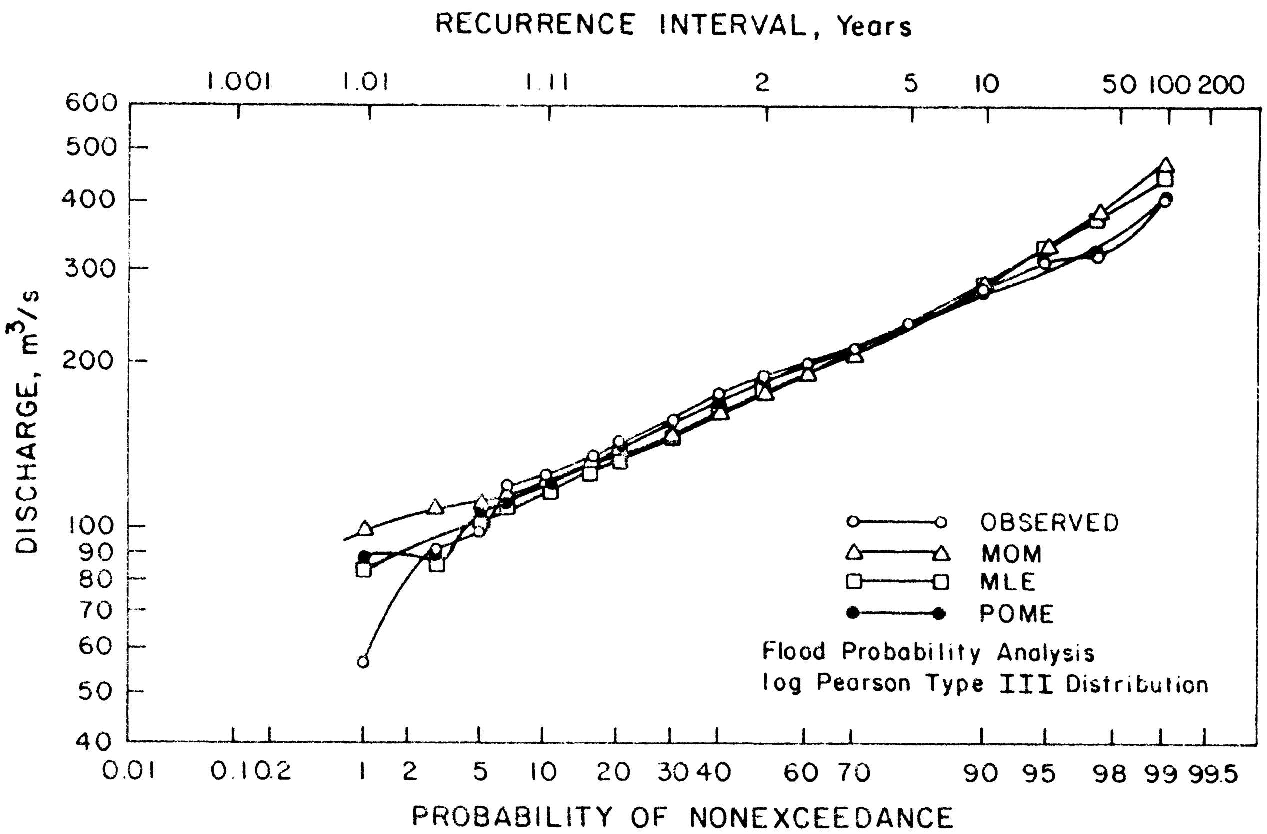

Figure 15.2 Comparison of observed and computed frequency curves using the POME, MLE and MOM methods for annual maximum discharge series for the Sebasticook River at Pittsfield, Maine.

Arora, K. and Singh, V.P., 1987b. A comparative evaluation of the estimators of commonly used flood frequency models: 2. Computer programs. *Completion Report*, Louisiana Water Resources Research Institute, Louisiana State University, Baton Rouge, Louisiana.

Arora, K. and Singh, V.P., 1988. On the method of maximum likelihood estimation for log-Pearson type 3 distribution. *Stochastic Hydrology and Hydraulics*, Vol. **2**, No. 2, pp. 156-160.

Arora, K. and Singh, V.P., 1989a. A comparative evaluation of the estimators of the log Pearson type (LP) 3 distribution. *Journal of Hydrology*, Vol. 105, pp. 9-37.

Arora, K. and Singh, V.P., 1989b. A note on the mixed moment estimation for the log-Pearson type 3 distribution. *Journal of the Institution of Engineers (India), Civil Engineering Division*, Vol. **69**, Part C15, pp. 298-301.

Ashkar, F. and Bobee, B., 1987. The generalized method of moments as applied to problems of flood frequency analysis: Some practical results for the log-Pearson type 3 distribution. *Journal of Hydrology*, Vol. **90**, pp. 199-217.

Ashkar, F. and Bobee, B., 1988. Confidence intervals for flood events under a Pearson 3 or log Pearson 3 distribution. Water Resources Bulletin, Vol. 24, No. 3, pp. 639-650.

Beard, L.R., 1986. Comment on "Relative accuracy of log Pearson III procedures by J.R. Wallis and E.F. Wood." *Journal of Hydraulic Engineering*, Vol. **112**, pp. 1205-1206.

Benson, M.A., 1968. Uniform flood-frequency estimating methods for federal agencies. *Water Resources Research*, Vol. **4,** No. 5, pp. 891-908.

Bobee, B., 1975. The log Pearson type 3 distribution and its application in hydrology. *Water Resources Research*, Vol. **11**, No. 5, pp. 681-689.

Bobee, B. and Ashkar, F., 1988. Generalized method of moments applied to LP3 distribution. *Journal of Hydraulic Engineering*, Vol. **114,** N. 8, pp. 899-909.

Bobee, B. and Ashkar, F., 1991. *The Gamma Family and Derived Distribution Applied in Hydrology*. Water Resources Publications, Littleton, Colorado.

Bobee, B., Cavadias, G., Ashkar, F., Bernier, J. And Rasmussen, P., 1993. Towards a systematic approach to comparing distributions used in flood frequency analysis. *Journal of Hydrology*, Vol. **142**, pp. 121-136.

Bobee, B. and Robitaille, R., 1975. Correction of bias in the estimation of the coefficient of skewness. *Water Resources Research*, Vol. **11**, No. 6, pp. 851-854.

Cohen, T.A., Lane, W.L. and Baier, W.G., 1997. An algorithm for computing moments-based flood quantile estimates when historical flood information is available. *Water Resources Research*, Vol. **33**, No. 9, pp. 2089-2096.

Condie, R., 1977. The log Pearson type 3 distribution: The T-year event and its asymptotic standard method by maximum likelihood theory. *Water Resources Research*, Vol. **13**, pp. 987-991.

Gupta, I.D. and Deshpande, V.C., 1994. Appliation of log-Pearson type-III distribution for evaluating design earthquake magnitudes. *Journal of the Institution of Engineers (India), Civil Engineering Division*, Vol. **75,** pp. 129-134.

Haktanir, T., 1991. Statistical modelling of annual maximum flows in Turkish rivers. *Hydrological Sciences Journal*, Vol. **36**, No. 4, pp. 367-389.

Hoshi, K. And Burges, S.J., 1981a. Approximate estimation of the derivative of a standard gamma quantile for use in cinfidence interval estimates. *Journal of Hydrology*, Vol. **53**, pp. 317-325.

Hoshi, K. and Burges, S.J., 1981b. Sampling properties of parameter estimates for the log Pearson type 3 distribution using moments in real space. *Journal of Hydrology*, Vol. **53**, pp. 305-316.

Kite, G.F. 1978. *Frequency and Risk Analyses in Hydrology*. Water Resources Publications, Littleton, Colorado.

Kuczera, G., 1982a. Combining site-specific and regional information: An empirical Bayes approach. *Water Resources Research*, Vol. **18**, No. 2, pp. 306-314.

Kuczera, G., 1982b. Robust flood frequency models. *Water Resources Research*, Vol. **18**, No. 2, pp. 315-324.

Landwehr, J.M., Matals, N.C. and Wallis, J.R., 1978. Some comparison of flood statistics in real and log space. *Water Resources Research*, Vol. **14**, No. 5, pp. 902-920.

Landwehr, J.M., Tasker, G.D. and Jarret, R.D., 1986. Comment on "Relative accuracy of log Pearson III procedures by J.R. Wallis and E.F. Wood." *Journal of Hydraulic Engineering*, Vol. **112**, pp. 1206-1210.

Loganathan, G.V., Mattejat, P., Kuo, C.Y. and Diskin, M.H., 1986. Frequency analysis of low flows: Hypothetical distribution methods and a physically based approach. *Nordic Hydrology*, Vol. **17**, pp. 129-150.

McMahon, T.A. and Srikanthan, R., 1981. Log Pearson III distribution-Is it applicable to flod frequency analysis of Australian streams? *Journal of Hydrology*, Vol. **52**, pp. 139-147.

Nozdryn-Plotinicki, M.J. and Watt, W.E., 1979. Assessment of fitting techniques for the log Pearson type 3 distribution using Monte Carlo simulation. *Water Resources Research*, Vol. **15**, No. 3, pp. 714-718.

Oberg, K.A. and Mades, D.M., 1987. Estimating generalized skew of the log-Pearson type III distribution. *Water Resources Investigations Report* **86-4009**, 42 pp., U.S. Geological Survey, Urbana, Illinois.

Ouarda, T.B.M. and Ashkar, F., 1998. Effect of trimming on LP III flood quantile estimates. *Journal of Hydrologic Engineering*, Vol. 3, No. 1, pp. 33-42.

Phien, H.N. and Hira, M.A., 1983. Log Pearson type-3 distribution: Parameter Estimation. *Journal of Hydrology*, Vol. **64**, pp. 25-37.

Phien, H.N. and Hsu, L.C., 1985. Variance of the T-year event in the log Pearson type 3 distribution. *Journal of Hydrology,* Vol. **77**, pp. 141-158.

Phien, H.N. and Jivajirajah, T., 1984. Application of the log Pearson type-3 distribution in hydrology. *Journal of Hydrology*, Vol. **73**, pp. 359-372.

Pilon, P.J. and Adamowski, K., 1993. Asymptotic variance of flood quantile in log Pearson type III distribution with historical information. *Journal of Hydrology*, Vol. **143**, pp. 481-503.

Rao, D.V., 1980a. Log Pearson type 3 distribution: Method of mixed moments. *Journal of Hydraulics Division, ASCE,* Vol. **106**, No. HY6, pp. 999-1019.

Rao, D.V., 1980b. Log Pearson type 3 distribution: A generalized evaluation. *Journal of Hydraulics Division, ASCE*, Vol. **106**, No. HY5, pp. 853-872.

Rao, D.V., 1981. Three-parameter probability distribution. Journal of Hydraulics Division, ASCE, Vol. **107**, No. HY3, pp. 339-358.

Rao, D.V., 1983a. Estimating log Pearson parameters by mixed moments. *Journal of Hydraulic Engineering*, Vol. **109**, No. 8, pp. 1118-1132.

Rao, D.V., 1983b. Three-parameter probability distributions. *Journal of the Hydraulics Division, ASCE*, Vol. **107**, No. HY3, pp. 339-357.

Rao, D. V.,1988. Estimating extreme events: Log Pearson type 3 distribution. Chapter 9 in *Civil Engineering Practice: Water Resources and Environmental Management*, edited by P.N. Cheremisinoff, N.C. Cheremisinoff and S.L. Cheng, Vol. **5**, pp. 291-304, Technomic Publishing Company, Inc., Lancaster, Basel.

Reich, B.M., 1970. Flood series compared to rainfall extremes. *Water Resources Research*, Vol. **6**, No. 6, pp. 1655-1667.

Rossi, F., Fiorentino, M. and Versace, P., 1984. Two-component extreme value distribution for flood frequency analysis. *Water Resources Research*, Vol. **20**, pp. 847-856.

Shen, H.W., Bryson, M.C. and Ochoa, I.D., 1980. Effect of tail behavior assumptions on flood predictions. *Water Resources Research*, Vol. **16**, No. 2, pp. 361-364.

Singh, V.P. and Singh, K., 1988. Parameter estimation for log-Pearson type III distribution by POME. *Journal of Hydraulic Engineering*, Vol. **114**, No. 1, pp. 112-122.

Song, D. and Ding, J., 1988. The application of probability weighted moments in estimating the parameters of the Pearson type three distribution. *Journal of Hydrology*, Vol. **101**, pp. 47-61.

Srikanthan, R. and McMahon, T.A., 1981. Log Pearson III distribution-Effect of dependence, distribution parameters and sample size on peak annual flood estimates. *Journal of Hydrology*, Vol. **52**, pp. 149-159.

Tasker, G.D., 1987. A comparison of methods for estimating low flow characteristics of streams. *Water Resources Bulletin*, Vol. **23**, No. 6, pp. 1077-1083.

Tung, K.T. and Mays, L.W., 1981. Generalized skew coefficients for flood frequency analysis. *Water Resources Bulletin*, Vol. **17**, No. 2, pp. 262-269.

Vogel, R.M., Thomas, W.O. and McMahon, T.A., 1992. Flood-flow frequency model selection in southwestern United States. *Journal of Water Resources Planning and Management,* Vol. **119**, No. 3, pp. 353-366.

Wallis, J.R. and Wood, E.F., 1985. Relative accuracy of log Pearson III procedures. *Journal of Hydraulic Engineering*, Vol. **111**, No. 7, pp. 1043-1056.

Water Resources Council, 1967. A uniform technique for determining flood flow frequencies. *Bulletin* No. **15,** 15 pp., Washington, D.C.

Water Resources Council, 1982. Guidelines for determining flood flow frequency. *Bulletin* **17 B**, Hydrology Subcommittee, Interagency Advisory Committee on Water Data, Washington, D.C.

CHAPTER 16

BETA DISTRIBUTION

In reliability safety analysis of civil engineering systems, we encounter parameters which are generally bounded and skewed random quantities. Exemplifying these parameters are factors of safety or safety indexes, variables representing strength of materials, intensity of loads, etc. Harr (1977) demonstrated the ability of the beta (or Pearson type 1) distribution to approximate most of the geotechnical parameters. Obini and Bourdeau (1985) simplified use of the beta distribution and investigated its sensitivity to the bound locations. Fielitz and Myers (1975) argued for the method of moments (MOM) to estimate the parameters of the beta distribution for ease of computation. Romesburg (1976) commented that formulation of the problem in terms of smallest order statistics would allow the use of the method of maximum likelihood estimation (MLE) to estimate the parameters of the beta distribution with little more effort than MOM. In multivariate cases, however, MOM would be the only practical method for parameter estimation.

If X has a beta distribution then its probability density function (pdf) is given by

$$f(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1} \tag{16.1}$$

where a > 0, b > 0 and 0 < x < 1. The beta distribution is a two-parameter distribution. Its cumulative distribution function (cdf) can be written as

$$F(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^{\infty} x^{a-1}(1-x)^{b-1}\, dx \tag{16.2}$$

16.1 Ordinary Entropy Method

16.1.1 SPECIFICATION OF CONSTRAINTS

Taking logarithm of equation (16.1) to the base 'e', we get

$$\ln f(x) = \ln\Gamma(a+b) - \ln\Gamma(a) - \ln\Gamma(b) + (a-1)\ln x + (b-1)\ln(1-x) \tag{16.3}$$

Multiplying equation (16.3) by [–f(x)] and integrating betwee 0 and 1, we get the entropy function:

V. P. Singh, *Entropy-Based Parameter Estimation in Hydrology*

$$I(f) = -\int_0^1 f(x)\ln f(x)dx = [-\ln\Gamma(a+b) + \ln\Gamma(a) + \ln\Gamma(b)]\int_0^1 f(x)dx - (a-1)\int_0^1 \ln x f(x)dx - (b-1)\int_0^1 \ln(1-x)f(x)dx \qquad (16.4)$$

From equation (16.4) the constraints appropriate for equation (16.1) can be written as

$$\int_0^1 f(x)dx = 1 \qquad (16.5)$$

$$\int_0^1 \ln x f(x) = E[\ln] \qquad (16.6)$$

$$\int_0^1 \ln(1-x)f(x)dx = E[\ln(1-x)] \qquad (16.7)$$

Equation (16.5) can be verified as follows. Substituting equation (16.1) in equation (16.5) one gets

$$\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\int_0^1 x^{a-1}(1-x)^{b-1}dx \qquad (16.8)$$

Because

$$\int_0^1 x^{a-1}(1-x)^{b-1}dx = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} = \beta(a,b) \qquad (16.9)$$

where $\beta(a,b)$ is a beta function. Therefore, we obtain

$$\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\,\frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} = 1$$

16.1.2 CONSTRUCTION OF ZEROTH LAGRANGE MULTIPLIER

The least-biased pdf f(x) consistent with equations (16.5) to (16.7) and corresponding to the principle of maximum entropy (POME) takes the form:

$$f(x) = \exp[-\lambda_0 - \lambda_1 \ln x - \lambda_2 \ln(1-x)] \qquad (16.10)$$

where λ_0, λ_1, and λ_2 are Lagrange multipliers. Substituting equation (16.10) in equation (16.5) yields

$$f(x) = \exp\int_0^1 \exp[-\lambda_0 - \lambda_1 \ln x - \lambda_2 \ln(1-x)]dx \qquad (16.11)$$

Equation (16.11) gives the partition function as

$$\exp(\lambda_0) = \int_0^1 \exp[-\lambda_1 \ln x - \lambda_2 \ln(1-x)]dx \tag{16.12}$$

Equation (16.12) can be simplified as follows:

$$\begin{aligned}\exp(\lambda_0) &= \int_0^1 \exp[\ln x^{-\lambda_1}]\exp[\ln(1-x)^{-\lambda_2}]dx = \int_0^1 x^{-\lambda_1}(1-x)^{-\lambda_2}dx \\ &= \int_0^1 x^{1-\lambda_1-1}(1-x)^{1-\lambda_2-1}dx = \frac{\Gamma(1-\lambda_1)\Gamma(1-\lambda_2}{\Gamma(2-\lambda_1-\lambda_2)}\end{aligned} \tag{16.13a}$$

The zeroth Lagrange multipliers λ_0 is got from equation (16.13a) as

$$\lambda_0 = \ln\Gamma(1-\lambda_1) + \ln\Gamma(1-\lambda_2) - \ln\Gamma(2-\lambda_1-\lambda_2) \tag{16.13}$$

The zeroth Lagrange multiplier is also obtained from equation (16.12) as

$$\lambda_0 = \ln \int_0^1 \exp[-\lambda_1 \ln x - \lambda_2 \ln(1-x)]dx \tag{16.14}$$

16.1.3 RELATION BETWEEN LAGRANGE MULTIPLIERS AND CONSTRAINTS

Differentiating equation (16.14) with respect to λ_1 and λ_2 respectively, one gets

$$\frac{\partial \lambda_0}{\partial \lambda_1} = -\frac{\int_0^1 \ln x \exp[-\lambda_1 \ln x - \lambda_2 \ln(1-x)]dx}{\int_0^1 \exp[-\lambda_1 \ln x - \lambda_2 \ln(1-x)]dx}$$

$$= -\int_0^1 \exp[-\lambda_0 - \lambda_1 \ln x - \lambda_2 \ln(1-x)]\ln x dx$$

$$= -\int_0^1 \ln x f(x)dx = -E[\ln x] \tag{16.15}$$

$$\frac{\partial \lambda_0}{\partial \lambda_1} = -\frac{\int_0^1 \ln(1-x)\exp[-\lambda_1 \ln x - \lambda_2 \ln(1-x)dx}{\int_0^1 \exp[-\lambda_1 \ln x - \lambda_2 \ln(1-x)]dx}$$

$$= -\int_0^1 \ln(1-x)\exp[-\lambda_0 - \lambda_1 \ln x - \lambda_2 \ln(1-x)]\ln x dx$$

$$= -\int_0^1 \ln(1-x) f(x)dx = -E[\ln(1-x)] \tag{16.16}$$

Differentiating equation (16.13b) with respect to λ_1 one also obtains

$$\frac{\partial \lambda_0}{\partial \lambda_1} = \frac{\partial}{\partial \lambda_1} \ln \Gamma(1-\lambda_1) - \frac{\partial}{\partial \lambda_1} \ln \Gamma(2-\lambda_1-\lambda_2) \tag{16.17}$$

Let

$$\alpha = 1-\lambda_1; \frac{\partial \alpha}{\alpha \lambda_1} = -1 \ ; \ \beta = 1-\lambda_2; \frac{\partial \beta}{\partial \lambda_2} = -1 \ ; \ \alpha + \beta = 2-\lambda_1-\lambda_2$$

Substituting the above quantities in equation (16.17), one gets

$$\frac{\partial \lambda_0}{\partial \lambda_1} = \frac{\partial}{\partial \alpha} \ln \Gamma(\alpha) \frac{\partial \alpha}{\partial \lambda_1} - \frac{\partial}{\partial \alpha} \ln \Gamma(\alpha+\beta) \frac{\partial \alpha}{\partial \lambda_1}$$

$$= \psi(\alpha)(-1) - \psi(\alpha+\beta)(-1) \ = \ -\psi(\alpha) + \psi(\alpha+\beta) \tag{16.18}$$

Differentiating equation (16.13b) with respect to λ_2, one obtains

$$\frac{\partial \lambda_0}{\partial \lambda_2} = \frac{\partial}{\partial \lambda_2} \ln \Gamma(1-\lambda_1) - \frac{\partial}{\partial \lambda_2} \ln \Gamma(2-\lambda_1-\lambda_2)$$

$$= \frac{\partial}{\partial \beta} \ln \Gamma(\beta) \frac{\partial \beta}{\partial \beta} - \frac{\partial}{\partial \beta} \ln \Gamma(\alpha+\beta) \frac{\partial \beta}{\partial \lambda_2} \ = \ -\psi(\beta) - \psi(\alpha+\beta) \tag{16.19}$$

Equating equations (16.15) and (16.18), as well as equations (16.16) and (16.19), we obtain

$$-E[\ln x] = -\ \psi(\alpha) + \psi(\alpha+\beta) \tag{16.20}$$

$$-E[\ln(1-x)] = -\psi(\beta) + \psi(\alpha+\beta) \tag{16.21}$$

or

$$E[\ln x] = \psi(\alpha) - \psi(\alpha+\beta) \tag{16.22}$$

$$E[\ln(1-x)] = \psi(\beta) - \psi(\alpha+\beta) \tag{16.23}$$

The left hand sides of equations (15.22) and (15.23) are known. Therefore, the values of α and β can be found, which are, in turn, related with λ_1 and λ_2.

16.1.4 RELATION BETWEEN LAGRANGE MULTIPLIERS AND PARAMETERS

Substituting equation (16.13b) in f(x) given by equation (16.10), one gets

$$f(x) = \exp[\ln\Gamma(1-\lambda_1)]^{-1}]\exp[\ln[\Gamma(1-\lambda_2)]^{-1}\exp[\ln\Gamma(2-\lambda_1-\lambda_2)]$$
$$\times \exp[\ln x^{-\lambda_1}]\exp[\ln(1-x)^{-\lambda_2}]$$
$$= \frac{\Gamma(2-\lambda_1-\lambda_2)}{\Gamma(1-\lambda_1)\Gamma(1-\lambda_2)}x^{-\lambda_1}(1-x)^{-\lambda_2} = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}$$

(16.24)

A comparison of equation (16.24) with equation (16.1) shows that a = α and b = β.

16.1.5 RELATION BETWEEN PARAMETERS AND CONSTRAINTS

Equations (16.22) and (16.23) relate the Lagrange multipliers to constraints and equation (16.24) shows the relation between the Lagrange multipliers and parameters. Eliminating the Lagrange multipliers between these equations, we get distribution parameters a and b in terms of the constraints E[ln x] and E[ln(1–x)] through equations (16.22) and (16.23). These equations are nonlinear but can be easily solved iteratively.

16.1.6 DISTRIBUTION ENTROPY

Equation (16.4) defines the distribution entropy.

$$I(x) = -\int_0^1 f(x)\ln f(x)dx \tag{16.25}$$

Substituting equation (16.1) in equation (16.25) and simplifying, one gets

$$I(x) = [-\ln\Gamma(a+b) + \ln\Gamma(a) + \ln\Gamma(b)]\int_0^1 f(x)dx$$

$$-(a-1)\int_0^1 \ln x f(x)dx - (b-1)\int_0^1 \ln(1-x)f(x)dx$$

$$I(x) = \ln[\frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha)+\beta}](a-1)E[\ln x] - (b-1)E[\ln(1-x)] \tag{16.26}$$

16.2 Parameter-Space Expansion Method

16.2.1 SPECIFICATION OF CONSTRAINTS

Following Singh and Rajagopal (1986), the constraints for this method are equation (16.5) and

$$\int_0^1 \ln x^{a-1} f(x)dx = E[\ln x^{a-1}] \tag{16.27}$$

$$\int_0^1 \ln(1-x)^{b-1} f(x)dx = E[\ln(1-x^b)^{b-1}] \tag{16.28}$$

16.2.2 DERIVATION OF ENTROPY FUNCTION

The pdf corresponding to the principle of maximum entropy (POME) and consistent with equations (16.5), (16.27), and (16.28) takes the form

$$f(x) = \exp[-\lambda_0 - \lambda_1 \ln x^{a-1} - \lambda_2 \ln(1-x)^{b-1}] \tag{16.29}$$

where λ_0, λ_1, and λ_2 are Lagrange multipliers. Insertion of equation (16.29) into equation (16.6) yields the entropy function:

$$\exp(\lambda_0) = \int_0^1 \exp[-\lambda_1 \ln x^{a-1} - \lambda_2 \ln(1-x)^{b-1}]$$

$$= \frac{\Gamma(1-\lambda_1(a-1))\Gamma(1-\lambda_2(b-1))}{\Gamma(2-\lambda_1(a-1)-\lambda_2(b-1))} \tag{16.30}$$

The zeroth Lagrange multiplier is given by equation (16.30) as

$$\lambda_0 = \ln\Gamma(1-\lambda_1(a-1)) + \ln\Gamma(1-\lambda_2(b-1)) - \ln\Gamma(2-\lambda_1(a-1)-\lambda_2(b-1)) \tag{16.31}$$

Substitution of equation (16.31) in equation (16.29) gives

$$f(x) = \frac{\Gamma(2-\lambda_1(a-1)-\lambda_2(b-1))}{\Gamma(1-\lambda_1(a-1))\Gamma(1-\lambda_2(b-1))} \exp[-\lambda_1 \ln x^{a-1} - \lambda_2 \ln(1-x)^{b-1}] \tag{16.32}$$

A comparison of equation (16.32) with equation (16.1) shows that $\lambda_1 = \lambda_2 = -1$. Taking logarithm of equation (16.32) one obtains

$$\ln f(x) = \ln\Gamma(2-\lambda_1(a-1)-\lambda_2(b-1)) - \ln\Gamma(1-\lambda_1(a-1)) - \ln\Gamma(1-\lambda_2(b-1))$$

$$- \lambda_1 \ln x^{a-1} - \lambda_2 \ln(1-x)^{b-1} \tag{16.33}$$

The entropy function of the beta distribution becomes

$$I(f) = -\ln\Gamma(2-\lambda_1(a-1)-\lambda_2(b-1)) + \ln\Gamma(1-\lambda_1(a-1)) + \ln\Gamma(1-\lambda_2(b-1))$$

$$+ \lambda_1 \ln x^{a-1} + \lambda_2 \ln(1-x)^{b-1} \tag{16.34}$$

16.2.3 RELATION BETWEEN PARAMETERS AND CONSTRAINTS

Taking partial derivatives of equation (16.34) with respect to λ_1, λ_2, a, and b separately and equating each derivative to zero, one obtains

$$\frac{\partial I}{\partial \lambda_1} = 0 = (a-1)\psi(K_1) - (a-1)\psi(K_2) + E[\ln x^{a-1}]$$

$$K_1 = \Gamma(2-\lambda_1(a-1)-\lambda_2(b-1)), K_2 = (1-\lambda_1(a-1)) \tag{16.35}$$

$$\frac{\partial I}{\partial \lambda_2} = 0 = (b-1)\psi(K_1) - (b-1)\psi(K_3) + E[\ln(1-x)^{b-1}],\ K_3 = \Gamma(1-\lambda_2(b-1)) \tag{16.36}$$

$$\frac{\partial I}{\partial a} = 0 = \lambda_1\,\psi(K_1) - \lambda_1\,\psi(K_2) + \lambda_1 E[\ln x] \tag{16.37}$$

$$\frac{\partial I}{\partial b} = 0 = (b-1)\psi(K_1) + (b-1)\psi(K_3) + \lambda_2 E[\ln(1-x)] \tag{16.38}$$

Simplification of equation (16.35) to (16.38), respectively, yields

$$\text{E}[\ln x] = \psi(K_2) - \psi(K_1) \tag{16.39}$$

$$\text{E}[\ln(1-x)] = \psi(K_3) - \psi(K_1) \tag{16.40}$$

$$\text{E}[\ln x] = \psi(K_2) - \psi(K_1) \tag{16.41}$$

$$\text{E}[\ln(1-x)] = \psi(K_3) - \psi(K_2) \tag{16.42}$$

Equations (16.39) and (16.41) are the same and so are equations (16.40) and (16.42). Therefore, equations (16.39) and (16.40) are the parameter estimation equations.

16.3 Other Methods of Parameter Estimation

Two other methods of parameter estimation are briefly outlined: Method of moments (MOM) and the method of maximum likelihood estimation (MLE).

16.3.1 METHOD OF MOMENTS

The beta distribution has two parameters. Therefore, two moments of X will suffice for the method of moments (MOM). The r-th moment of the distribution about the origin can be expressed as

$$M_r(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^1 x^{r+a+-1} (1-x)^{b-1} dx = \frac{\Gamma(a+b)\Gamma(a+r)}{\Gamma(a)\Gamma(a+b+r)} \tag{16.43}$$

Equation (16.43) yields the first two moments as

$$M_1 = \frac{a}{a+b} \tag{16.44}$$

$$M_2 = \frac{a(a+1)}{(a+b)(a+b+1)} \tag{16.45}$$

Equations (16.44) and (16.45) can be solved by recalling that the first moment is the mean and the second moment is equal to the sum of the variance and the square of the first moment.

16.3.2 METHOD OF MAXIMUM LIKELIHOOD ESTIMATION

For the method of maximum likelihood estimation (MLE), the likelihood function, L, for a sample of size n drawn from a beta population is given as

$$L = [\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}]^n \prod_{i=1}^{n} x_i^{a-1}(1-x_i)^{b-1} \tag{16.46}$$

The log likelihood function, ln L, is expressed from equation (16.46) as

$$\ln L = n \ln \Gamma(a+b) - n \ln \Gamma(a) - n\Gamma(b) + \sum_{i=1}^{n} \ln[x_i^{a-1}(1-x_i)^{b-1}] \quad (16.46)$$

Differentiating equation (16.46) with respect to parameters a and b, respectively, and equaing each derivative to zero yields the parameter estimation equations:

$$\frac{\partial[\ln \Gamma(a)]}{\partial a} - \frac{\partial[\ln \Gamma(a+b)]}{\partial a} = \frac{1}{n}\sum_{i=1}^{n} \ln x_i \tag{16.47}$$

$$\frac{\partial[\ln \Gamma(b)]}{\partial b} - \frac{\partial[\ln \Gamma(a+b)}{\partial b} = \frac{1}{n}\ln(1-x_i) \tag{16.48}$$

Equations (16.47) and (16.48) are solved iteratively. It should be noted that these equations are equivalent to equations (16.40) and (16.42) of the POME method. This means that POME and MLE would yield equivalent parameter estimates.

References

Fielitz, B.D. and Myers, B.L., 1975. Estimation of parameters in the beta distribution. *Decision Sciences*, Vol. **6**, No. 1, pp. 1-13.

Gnanadeshikan, R., Pinkham, R.S. and Hughes, L.P., 1967. Maximum likelihood estimation of the beta distribution from smallest order statistics. *Technometrics,* Vol. **9**, No.4, pp. 6-7- 620.

Harr, M.E., 1977. *Mechanics of Particulate Media: A Probabilistic Approach.* McGraw-Hill

Book Company, New York.

Oboni, F. and Bourdeau, P.L., 1985. Simplified use of the beta distribution and sensitivity to bounded locations. *Structural Safety*, Vol. **3**, pp. 63-66.

Romesburg, H.C., 1976. Estimation of parameters in the beta distribution: Comment. *Decision Sciences*, Vol. **7**, pp. 162.

CHAPTER 17

TWO-PARAMETER LOG-LOGISTIC DISTRIBUTION

The log-logistic distribution (LLD) is obtained by applying the logarithmic transformation to the logistic distribution (LD) in much the same way as the log-normal distribution is obtained from normal distribution or the log-Pearson distribution from the Pearson distribution. The log-logistic distribution is a special case of Burr's type-XII distribution (Burr, 1942), and also a special case of the "Kappa distributions" (Mielke and Johnson, 1973), that have been applied to precipitation and streamflow data.

When compared with the log-normal distribution, the LLD has a similar shape, but is mathematically more tractable. That is why Bennett (1983) chose it to analyze survival data. It can also be a good replacement for the Weibull distribution whose hazard function must be monotonic and, hence, may not be adequate in many practical cases. Furthermore, LLD is related to extremal distributions. As demonstrated by Lawless (1986), the Weibull distribution plays a central role in the field of reliability analysis because it is a unique distribution that belongs to two families of extreme distributions. Each of these two families possesses desirable attributes for analysis of proportional hazard and accelerated failure times. When these two families are generalized, the LLD has the attractive feature of being a member of both families. Because its distribution function has a closed form and its hazard function is quite flexible, this distribution has greater appeal and may be applicable to a wide variety of problems in many areas.

Shoukri, et al. (1988) applied the 2-parameter log-logistic distribution (LLD2) to analyze extensive Canadian precipitation data, and found it to be a suitable model for generating precipitation for the various Canadian regions. In their investigation of the methods of probability-weighted moments (PWM) and maximum likelihood estimation (MLE) for LLD2, Shoukri, et al. (1988) found PWM to produce smaller biases and variances in parameter estimates than MLE, even when sample sizes were as small as 15 or 25 observations. However, that was not true for efficiency of parameter estimates. Guo and Singh (1992) employed the principle of maximum entropy (POME) to derive a new method of parameter estimation for the LLD2. Monte Carlo simulated data were used to evaluate this method and compare it with the methods of moments (MOM), probability weighted moments (PWM), and maximum likelihood estimation (MLE). Simulation results showed that POME's performance was comparable to other methods.

The log-logistic distribution (LLD) is obtained by the logarithmic transformation of the logistic distribution. Thus, if Y is a random variable which has a standard logistic distribution with probability density function (pdf) $g(y) = \exp(y)(1+\exp(y))^{-2}$, then using the transformation $y = b \ln(x/a)$, the pdf of X can be expressed as

V. P. Singh, *Entropy-Based Parameter Estimation in Hydrology*

$$f(x)=\frac{(b/a)(x/a)^{b-1}}{[1+(x/a)^b]^2}, a, x \succ 0; b \geq 1 \tag{17.1}$$

A random variable X whose pdf is given by equation (17.1) is said to have a 2-parameter log-logistic distribution (LLD2). Its cumulative distribution function (cdf) and inverse cumulative distribution function are given, respectively, by

$$F(x)=\frac{(x/a)^b}{1+(x/a)^b} \tag{17.2}$$

$$x=a\,[F/(1-F)]^{1/b} \tag{17.3}$$

The shapes of the LLD2 for various values of a and b are illustrated in Figures 17.1 and 17.2.

17.1 Ordinary Entropy Method

17.1.1 SPECIFICATION OF CONSTRAINTS

Taking logarithms of equation (17.1) to the base 'e', one gets

$$\ln f(x) = \ln(b/a) + (b-1)\ln(x/a) - 2\ln[1+(\frac{x}{a})^b]$$

or

$$\ln f(x) = \ln b - \ln a + (b-1)\ln x - (b-1)\ln a - 2\ln[1+(x/a)^b] \tag{17.4}$$

Multiplying equation (17.4) by [-f(x)] and integrating between 0 and ∞, one obtains the entropy function:

$$\begin{aligned} I(f) &= -\int_0^\infty f(x)\ln f(x)\,dx = [-\ln(\frac{b}{a})]\int_0^\infty f(x)\,dx \\ &- (b-1)\int_0^\infty \ln x f(x)\,dx + 2\int_0^\infty f(x)\ln[1+(\frac{x}{a})^b]\,dx \end{aligned} \tag{17.5}$$

From equation (17.5), the constraints, appropriate for equation (17.1), can be written as

$$\int_0^\infty f(x)dx = 1 \tag{17.6}$$

$$\int_0^\infty \ln f(x)\,dx = E[\ln x] \tag{17.7}$$

$$\int_0^\infty \ln[1+(\frac{x}{a})^b]\,f(x)\,dx = E[\ln\{1+(\frac{x}{a})^b]] \tag{17.8}$$

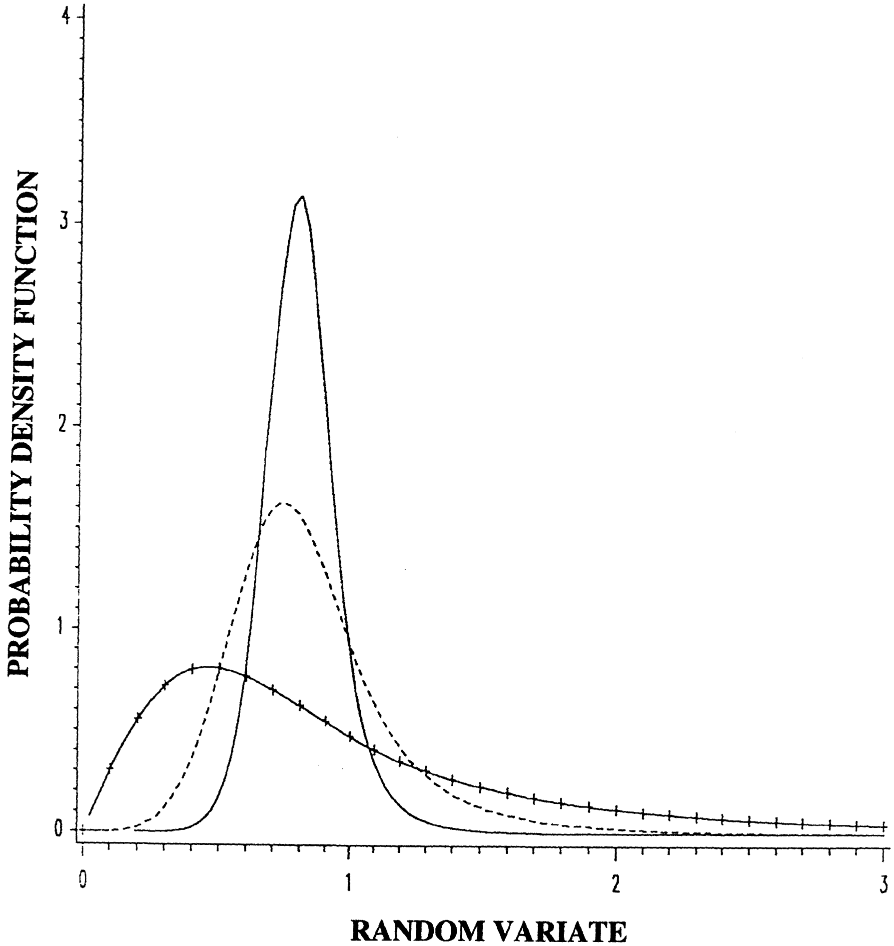

Figure 17.1 LLD2 density function with a = 0.8, b = 2, 5, 10; line: b = 10; dash: b = 5; and plus: b = 2.

in which E[•] denotes the expectation of the bracketed quantity. These constraints specify the information sufficient for LLD2. Because this information is determined from data in terms of expectations, the parameters and other statistics of the distribution can be physically interpreted.

17.1.2 CONSTRUCTION OF ZEROTH LAGRANGE MULTIPLIER

The least-biased pdf corresponding to POME and consistent with equations (17.6) to (18.8) takes the following form:

$$f(x) = \exp\{-a_0 - a_1 \ln(\frac{x}{a}) - a_2 \ln[1 + (\frac{x}{a})^b]\} \qquad (17.9)$$

where a_0, a_1, and a_2 are Lagrange multipliers. The mathematical rationale for equation (17.9) has been presented by Tribus (1969).

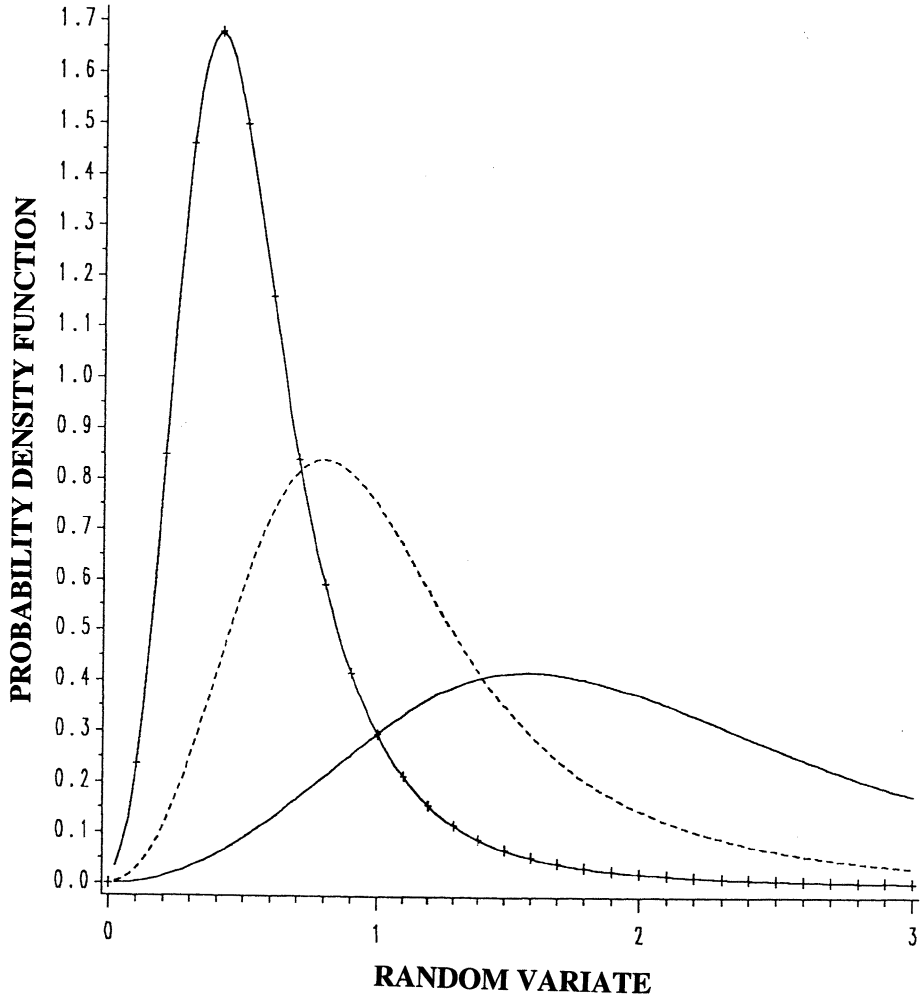

Figure 17.2 LLD2 density function with a = 0.5, 1, and 2; and b = 3; line: a = 2; dash: a = 1; and plus: a = 0.5

Insertion of equation (17.9) into equation (17.6) yields

$$\exp(a_0) = \int_0^\infty \exp(-a_1 \ln(\frac{x}{a}) - a_2 \ln[1+(\frac{x}{a})^b])\, dx$$
$$= \frac{b\,\Gamma(\frac{1-a_1}{b})\,\Gamma(a_2 - \frac{1-a_1}{b})}{\Gamma(a_2)} \tag{17.10}$$

which is called the partition function. $\Gamma(\bullet)$ is the gamma function. The zeroth Lagrange multiplier is given by equation (17.10) as

$$a_0 = \ln b + \ln\Gamma(\frac{1-a_1}{b}) + \ln\Gamma(a_2 - \frac{1-a_1}{b}) - \ln\Gamma(a_2) \tag{17.11}$$

One also gets the zeroth Lagrange multipler from equation (17.10) as

$$a_0 = \ln \int_0^\infty \exp(-a_1 \ln(\frac{x}{a}) - a_2 \ln[1+(\frac{x}{a})^b])\, dx \tag{17.12}$$

17.1.3 RELATION BETWEEN LAGRANGE MULTIPLIERS AND CONSTRAINTS

Differentiating equation (17.12) with respect to a_1 and a_2, respectively, one gets

$$\frac{\partial a_0}{\partial a_1} = \frac{\int_0^\infty \exp\{-a_1 \ln(\frac{x}{a}) - a_2 \ln[1+(\frac{x}{a})^b]\}\{-\ln(\frac{x}{a})\}\, dx}{\int_c^\infty \exp\{-a_1 \ln(\frac{x}{a}) - a_2 \ln[1+(\frac{x}{a})^b]\} dx}$$
$$= -E[\ln(\frac{x}{a})] \tag{17.13}$$

$$\frac{\partial a_0}{\partial a_2} = \frac{\int_0^\infty \exp\{-a_1 \ln(\frac{x}{a}) - a_2 \ln[1+(\frac{x}{a})^b]\}\{-\ln[1+(\frac{x}{a})^b]\}\, dx}{\int_c^\infty \exp\{-a_1 \ln(\frac{x}{a}) - a_2 \ln[1+(\frac{x}{a})^b]\} dx}$$
$$= -E\{\ln[1+(\frac{x}{a})^b]\} \tag{17.14}$$

Similarly, differentiation of equation (17.11) with respect to a_1 and a_2, respectively, yields

$$\frac{\partial a_0}{\partial a_1}=\frac{1}{b}[-\Psi(k_1)+\Psi(k_2)] \tag{17.15}$$

$$\frac{\partial a_0}{\partial a_2}=\Psi(k_2)-\Psi(a_2) \tag{17.16}$$

where $k_1 = (1-a_1)/b$ and $k_2 = [a_2 - (1- a_1)/b]$. Equating equation (17.13) to equation (17.15) and equation (17.14) to equation (17.16), we obtain

$$E[\ln(\frac{x}{a})]=\frac{1}{b}[\Psi(k_1)-\Psi(k_2)] \tag{17.17}$$

$$E[\ln\{1+(\frac{x}{a})^b\}]=\Psi(a_2)-\Psi(k_2) \tag{17.18}$$

17.1.4 RELATION BETWEEN PARAMETERS AND LAGRANGE MULTUPLIERS

Insertion of equation (17.11) into equation (17.10) yields

$$f(x)=\frac{b}{a}\frac{\Gamma(a_2)}{\Gamma[(1-a_1)/b]\Gamma[a_2-(1-a_1)/b]}(\frac{x}{a})^{-a_1}[1+(\frac{x}{a})^b]^{-a_2} \tag{17.19}$$

A comparison of equation (17.14) with equation (17.2) yields $a_2 = 2$ and $a_1 = 1-b$.

17.1.5 RELATION BETWEEN PARAMETERS AND CONSTRAINTS

The LLD2 distribution has three parameters which are related to the Lagrange multipliers by bequations (17.17) and (17.18) which, in turn, are related to the constraints through equation (17.19). Eliminating the Lagrange multipliers between these two sets of equations yields the relation between parameters and constraints. Hence, we get

$$E[\ln(\frac{x}{a})]=0 \tag{17.20}$$

$$E[\ln\{1+(\frac{x}{a})^b\}]=1 \tag{17.21}$$

Equations (17.20) and (17.21) are the estimation equations to get a unique determination of parameters a, and b.

17.1.6 DISTRIBUTION ENTROPY

The entropy of the LD2 distribution is given as

$$I(f) = \ln b - \ln a + (b-1) E[(\frac{x}{a})^b] - 2E\{\ln[1+(\frac{x}{a})^b]\} \quad (17.22)$$

17.2 Parameter-Space Expansion Method

17.2.1 SPECIFICATION OF CONSTRAINTS

Following Singh and Rajagopal (1986), the constraints for this method are given by equations (17.6)-(17.8).

17.2.2 DERIVATION OF ENTROPY FUNCTION

Taking logarithm of equation (17.19) gives

$$\ln f(x) = \ln(\frac{b}{a}) + \ln\Gamma(a_2) - \ln\Gamma(\frac{1-a_1}{b}) - \ln\Gamma(a_2 - \frac{1-a_1}{b}) - a_1 \ln(\frac{x}{a}) - a_2 \ln[1+(\frac{x}{a})^b] \quad (17.23)$$

Thus, the entropy function I(f) of LLD2 can be expressed using equation (17.15) as

$$I(f) = \ln a - \ln b + \ln\Gamma(\frac{1-a_1}{b}) + \ln\Gamma(a_2 - \frac{1-a_1}{b}) - \ln\Gamma(a_2) + a_1 E[\ln(\frac{x}{a})] + a_2 E[\ln(1+(\frac{x}{a})^b)] \quad (17.24)$$

17.2.3 RELATION BETWEEN PARAMETERS AND CONSTRAINTS

According to Singh and Rajagopal (1986), the relation between distribution parameters and constraints is obtained by taking partial derivatives of the entropy function with respect to Lagrange multipliers as well as distribution parameters and then equating these derivatives individually to zero. To that end, taking partial derivative of equation (17.24) with respect to a_1, a_2, a, and b separately, and equating each derivative to zero yields

$$\frac{\partial I}{a_1} = \frac{1}{b}[-\Psi(k_1) + \Psi(k_2)] + E[\ln(\frac{x}{a})] = 0 \quad (17.25)$$

$$\frac{\partial I}{\partial a_2} = \Psi(k_2) - \Psi(a_2) + E[\ln(1+(\frac{x}{a})^b] \quad (17.26)$$

$$\frac{\partial I}{\partial a} = \frac{1}{a} - \frac{a_1}{a} - \frac{b}{a} - a_2 E[\frac{(x/a)^b}{1+(x/a)^b}] = 0 \quad (17.27)$$

$$\frac{\partial I}{\partial b} = -\frac{1}{b} - \Psi(k_1)(\frac{1-a_1}{b^2}) + \Psi(k_2)(\frac{1-a_2}{b}) = 0 \tag{17.28}$$

where $k_1 = (1-a_1)/b$, $k_2=a_2-(1-a_1)/b$ and ψ is digamma function = $d[\ln \Gamma(x)]/dx$. Simplification of equations (17.25) to (17.28), respectively, yields

$$E[\ln(\frac{x}{a})] = \frac{1}{b}[\Psi(k_1) - \Psi(k_2)] \tag{17.29}$$

$$E[\ln\{1+(\frac{x}{a})^b\}] = \Psi(a_2) - \Psi(k_2) \tag{17.30}$$

$$a_2\, b\, E[\frac{(x/a)^b}{1+(\frac{x}{a})^b}] = 1 - a_1 \tag{17.31}$$

$$a_2\, E[\frac{(x/a)^b \ln(x/a)}{1+(x/a)^b}] = \frac{1}{b} + \frac{1-a_1}{b^2}\Psi(k_1) - \frac{1-a_2}{b^2}\Psi(k_2) \tag{17.32}$$

Note that $a_2=2$ and $a_1=1-b$. Therefore, equations (17.29) to (17.32) become

$$E[\ln(\frac{x}{a})] = 0 \tag{17.33}$$

$$E[\ln\{1+(\frac{x}{a})^b\}] = 1 \tag{17.34}$$

$$2\, E[\frac{(x/a)^b}{1+(x/a)^b}] = 1 \tag{17.35}$$

$$2b\, E[\frac{(x/a)^b \ln\{(x/a)\}}{1+(x/a)^b}] = 1 \tag{17.36}$$

Equations (17.35) and (17.36) are identities which can be proved as follows. For equation (17.35) we write

$$E[\frac{(x/a)^b}{1+(x/a)^b}] = \int_0^\infty \frac{(x/a)^b}{1+(x/a)^b} \frac{(b/a)(x/a)^{b-1}}{[1+(x/a)\}^b]^2} dx$$

$$= b\int_0^\infty \frac{y^{2b-1}}{(1+y^b)^3} dy, y = (x/a) \qquad (17.37)$$

$$= B(2,1) = 1/2, B(\bullet,\bullet) = beta\ function$$

Therefore,

$$2E[\frac{(x/a)^b}{1+(x/a)^b}] = 1 \qquad (17.38)$$

Similarly, we write for equation (17.36):

$$E[\frac{(x/a)^b \ln\{(x/a)\}}{1+(x/a)^b}] = E[\ln(\frac{x}{a})] - E[\frac{\ln\{(x)/a\}}{1+(x/a)^b}]$$

$$= -\int_0^\infty \frac{\ln\{(x/a)\}}{1+(x/a)^b} \frac{b(x/a)^{b-1}}{a[1+(x/a)^b]^2} dx$$

$$= -\int_0^\infty \frac{b\ln(y)y^{b-1}}{(1+y^b)^3} dy, y = (x/a)$$

$$= 0.5\int_0^\infty \ln(y)d(1+y^b)^{-2}$$

$$= \frac{1}{2b}\int_1^\infty \ln(z-1)dz^{-2},\ z = 1+y^b$$

$$= -\frac{1}{b}\lim_{z\to 0} \ln(z-1) - \frac{1}{2b}\int_1^\infty \frac{1}{z^2(z-1)} dz$$

$$= -\frac{1}{2b}\lim_{z\to 0} \ln(z-1) - \frac{1}{2b}\int_1^\infty (-\frac{1}{z} - \frac{1}{z^2} + \frac{1}{z-1}) dz$$

$$= \frac{1}{2b} \tag{17.39}$$

Thus,

$$2bE\left[\frac{(x/a)^b \ln\{(x/a)\}}{1+(x/a)^b}\right]=1 \tag{17.40}$$

Therefore, the POME-based parameter estimation equations are equations (17.33) and (17.34).

17.3 Other Methods of Parameter Estimation

17.3.1 METHOD OF MOMENTS

For LLD2, the moment estimation equations are (Guo and Singh, 1992):

$$E[x] = a\,B(1 + 1/b, 1 - 1/b),\ (b \geq 1) \tag{17.41}$$

$$B(m,n) = \frac{\Gamma(m)\Gamma(n)}{\Gamma(m+n)}$$

$$E[x^2] = a^2\,B(1 + 2/b, 1 - 2/b) \tag{17.42}$$

Therefore, the variance is obtained by

$$\mathrm{Var}[x] = a^2\,[B(1 + 2/b, 1 - 2/b) - B^2(1 + 1/b, 1 - 1/b)],\ (b \geq 2) \tag{17.43}$$

17.3.2 METHOD OF MAXIMUM LIKELIHOOD ESTIMATION

For the method of maximum likelihood estimation (MLE), the MLE estimation equations are

$$2\sum_{i=1}^{n}\left[\frac{(x_i/a)^b}{1+(x_i/a)^b}\right]=n \tag{17.44}$$

$$2b\sum_{i=1}^{n}\left[\frac{\ln(x_i/a)(x_i/a)^b}{1+(x_i/a)^b}\right]-b\sum_{i=1}^{n}\ln(x_i/a)-n=0 \tag{17.45}$$

17.3.3 METHOD OF PROBABILITY WEIGHTED MOMENTS

For the method of probability weighted moments (PWM), the estimation equations are given as

$$b = \frac{W_0}{W_0 - 2W_1} \tag{17.46}$$

$$a = \frac{W_0}{\Gamma(1 + 1/b)\,\Gamma(1 - 1/b)} \tag{17.47}$$

where

$$W_k = \frac{1}{n} \sum_{i=1}^{n} x_i \binom{n-1}{k} / \binom{n+1}{k} \tag{17.48}$$

17. 4 Comparative Evaluation of Parameter Estimation Methods

Guo and Singh (1992) and Singh et al. (1993) evaluated parameters and quantiles of the 2-parameter log-logistic distribution (LLG2) using the methods of moments (MOM), probability-weighted moments (PWM), maximum likelihood estimation MLE), and entropy (POME) for Monte Carlo generated samples. The performance of these estimators was statistically compared, with the objective of identifying the most robust estimator from amongst them. Their work is summarized here.

17.4.1 MONTE CARLO SIMULATION

To assess the performance of the parameter estimation methods outlined above, Monte Carlo sampling experiments were conducted. Four cases for LLD2, listed in Table 17.1, were considered. For each population case, 1000 random samples of size 10, 20, 50, 100, 200 and 500 were generated, and then parameters and quantiles were estimated. Figure 17.3 shows the unique relationship between coefficient of variation (CV) and parameter a.

Table 17.1 LLD2 population cases considered in sampling experiments (population mean, $\mu=1$).

LLD2 Population	Coefficient of Variation (CV)	a	b
Case 1	0.1	0.995	18.248
Case 2	0.5	0.907	4.137
Case 3	1.0	0.788	2.695
Case 4	3.0	0.663	2.088

17.4.2 PERFORMANCE INDICES

The performance of the estimation methods was evaluated using standardized bias (BIAS), root mean square error (RMSE), and robustness criteria. These indices are described in Chapter 2.

17.4.3 BIAS IN PARAMETER ESTIMATES

The results of parameter estimation for the LLD2 distribution showed that in general, when CV was small (CV<0.1), MOM, POME and MLE performed the best and were comparable in estimating both a and b. When CV>0.1, MLE performed in a superior manner in estimating both a and b, and POME was comparable. When CV>2.0, PWM produced the least BIAS in estimating b for all sample sizes.

17.4.4 RMSE IN PARAMETER ESTIMATES

In general, MLE performed the best in estimating a and b in terms of parameter RMSE for all sample sizes over all population cases, and POME was comparable. When CV>0.1 MOM was comparable, whereas when CV>1.0 PWM performed well and was comparable.

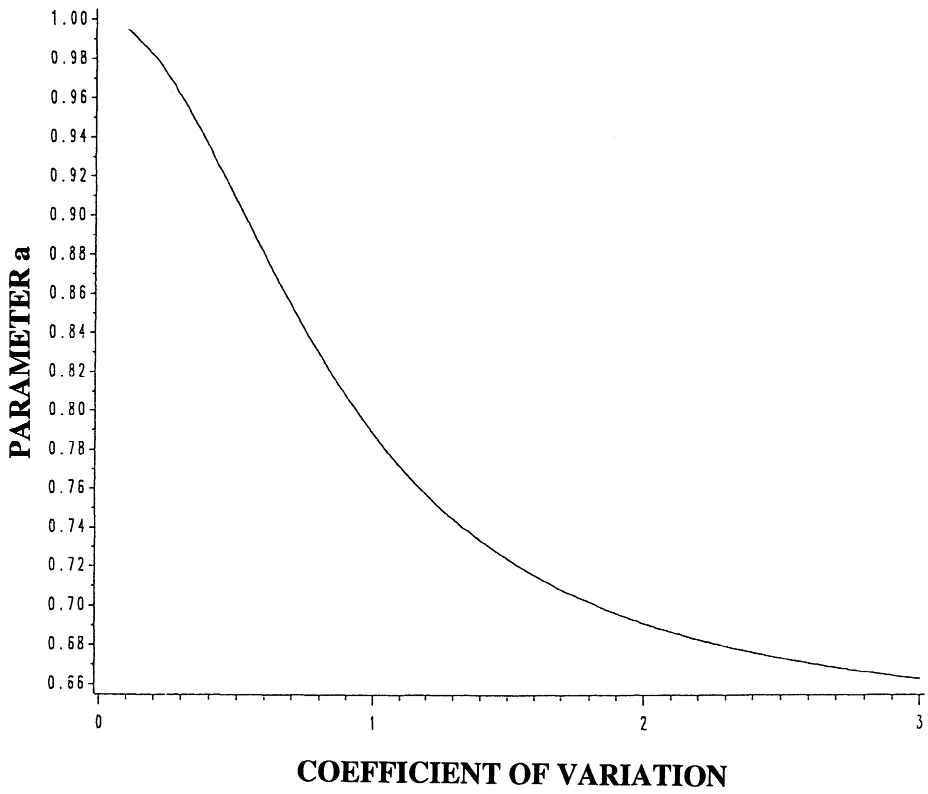

Figure 17.3 Parameter a versus CV for LLD2 distribution.

17.4.5 BIAS AND RMSE IN QUANTILE ESTIMATES

The results of quantile estimation of LLD2 showed that in general, POME and MLE performed in a superior manner in terms of quantile BIAS and RMSE for all sample sizes over all population cases. Again, for small CV, MOM was comparable and for large CV PWM was comparable.

17.4.6 ROBUSTNESS EVALUATION

The relative robustness of the parameter and quantile estimation methods clearly illustrated the generally superior performance of POME and MLE, and for small CV PWM performed worse, but for large CV, PWM performed better and MOM worse.

References

Bennett, S., 1983. Log-logistic regression models for survival data. *Applied Statistics*, Vol. **32**, pp. 165-171.

Burr, I. W., 1942. Cumulative frequency functions. *Annals of Mathematical Statist*istics, Vol. **13**, pp. 215-232.

Guo, H. and Singh, V. P., 1993. A comparative evaluation of estimators of log-logistic distributions by Monte Carlo simulation. *Technical Report* WRR**26**, Water Resources Program, Department of Civil Engineering, Louisiana State University, Baton Rouge, Louisiana, U.S.A.

Lawless, J. F., 1986. A note on lifetime regression models. *Biometrika,* Vol. **73,** No. 2, pp. 509-512.

Mielke, P. W. and Johnson, E. S., 1973. Three-parameter kappa distribution maximum likelihood estimates and likelihood ratio tests. *Monthly Weather Review*, Vol. **101**, pp. 701-709.

Shoukri, M. M., Mian, I. U. H. and Tracy, D. S., 1988. Sampling properties of estimators of the log-logistic distribution with application to Canadian precipitation data. *The Canadian Journal of Statistics*, Vol. **16**, No. 3, pp. 223-236.

Singh, V. P. and Rajagopal, A. K., 1986. A new method of parameter estimation for hydrologic frequency analysis. *Hydrological Science and Technology*, Vol. **2**, No. 3, pp. 33-40.

Tribus, M., 1969. *Rational Descriptors, Decisions and Designs*. Pergamon Press, New York.

CHAPTER 18

THREE-PARAMETER LOG-LOGISTIC DISTRIBUTION

Some general aspects of the log-logistic distribution (LLD) are discussed in Chapter 17. The three-parameter log-logistic distribution (LLD3) is a generalization of the two-parameter log-logistic distribution. The LLD3 has been applied to frequency analysis of precipitation and streamflow data. Ahmad, et al. (1988) employed it for flood frequency analysis of annual maximum series for part of Scotland, and compared its performance with the generalized extreme value, three-parameter log-normal and Pearson-type 3 distributions. They found LLD3 to consistently perform better than these three distributions. In a comparative study on statistical modeling of annual maximum flows of 112 Turkish rivers, Haktanir (1991) observed that LLD3 and log-Pearson type 3 distribution provided better fits than log-normal, Gumbel, SMEMAX, log-Boughton, and Pearson-type 3 distributions. He concurred with the findings of Ahmad, et al. (1988) as far as Turkish rivers were concerned.

Ahmad, et al. (1988) discussed several methods for estimating parameters of LLD3, including the methods of moments (MOM), maximum likelihood estimation (MLE), probability weighted moments (PWM), ordinary least squares (OLS), and generalized least squares (GLS). They found that for Scottish catchments OLS, GLS and MLE were in fairly close agreement. In their investigation of PWM and MLE for LLD2, Shoukri, et al. (1988) found PWM to produce smaller biases and variances in parameter estimates than MLE, even when sample sizes were as small as 15 or 25 observations. However, that was not true for efficiency of parameter estimates. In his comparative study on Turkish rivers, Haktanir (1991) used MOM, MLE and PWM for estimation of LLD3 parameters, and found that no one method was uniformly superior, although PWM was better for more rivers than MOM and MLE.

Guo and Singh (1992) and Singh et al. (1993) employed the principle of maximum entropy (POME) to derive a new method of parameter estimation for LLD3. Monte Carlo simulated data were used to evaluate this method and compare it with MOM, PWM, and MLE. Simulation results showed that POME's performance was superior in predicting quantiles of large recurrence intervals when population skew was greater than or equal to 2.0. In all other cases, POME's performance was comparable to other methods.

Let there be two random variables X and Y related through a logarithmic transformation as $Y = b \ln[(X-c)/a]$, where X is a positive random variable, and a, b and c are the three parameters. If Y has a logistic distribution (LD),

$$g(y) = \exp(y)\,[1 + \exp(y)]^{-2} \tag{18.1}$$

then X has a 3-parameter log-logistic distribution (LLD3) with probability density function (pdf, f(x)), cumulative distribution function (cdf, F(x)), and inverse cumulative distribution function (icdf, x(F)) expressed, respectively, as

V. P. Singh, *Entropy-Based Parameter Estimation in Hydrology*

$$f(x) = \frac{(b/a)\,[(x-c)/a]^{b-1}}{\{1+[(x-c)/a]^b\}^2}, \qquad a > 0,\ x > c;\ b \geq 1 \tag{18.2}$$

$$F(x) = \frac{[(x-c)/a]^b}{1 + [(x-\gamma)/a]^b} \tag{18.3}$$

$$x = \gamma + a\,[F/(1-F)]^{1/b} \tag{18.4}$$

Parameters a, b and c, respectively, are the shape, scale and location parameters; and x and y are values of X and Y, respectively. LLD3 can also be obtained by compounding the 3-parameter Weibull distribution (Shoukri, et al., 1988):

$$f(-x|z) = z(\frac{b}{a})(\frac{x-c}{a})^{b-1}\exp(-z(\frac{x-c}{a})^b \tag{18.5}$$

over the probability distribution of Z, that is taken as standard exponential. That is,

$$f(x) = \int_0^\infty f(x|z)\exp(-z)\,dz \tag{18.6}$$

The shape of LLD3 for various values of a and b and c = 0 are shown in Figures 18.1 and 18.2.

18.1 Ordinary Entropy Method

18.1.1 SPECIFICATION OF CONSTRAINTS

Following Jaynes (1968) and Tribus (1969), the constraints appropriate for equation (18.2) can be written as

$$\int_c^\infty f(x)\,dx = 1 \tag{18.7}$$

$$\int_c^\infty \ln(\frac{x-c}{a})\,f(x)\,dx = E[\ln(\frac{x-c}{a})] \tag{18.8}$$

$$\int_c^\infty \ln[1+(\frac{x-c}{a})^b]\,f(x)\,dx = E[1+(\frac{x-c}{a})^b] \tag{18.9}$$

in which E[•] denotes the expectation of the bracketed quantity. These constraints specify the information sufficient for LLD3. Because this information is determined from data in terms of expectations, the parameters and other statistics of the distribution can be physically interpreted.

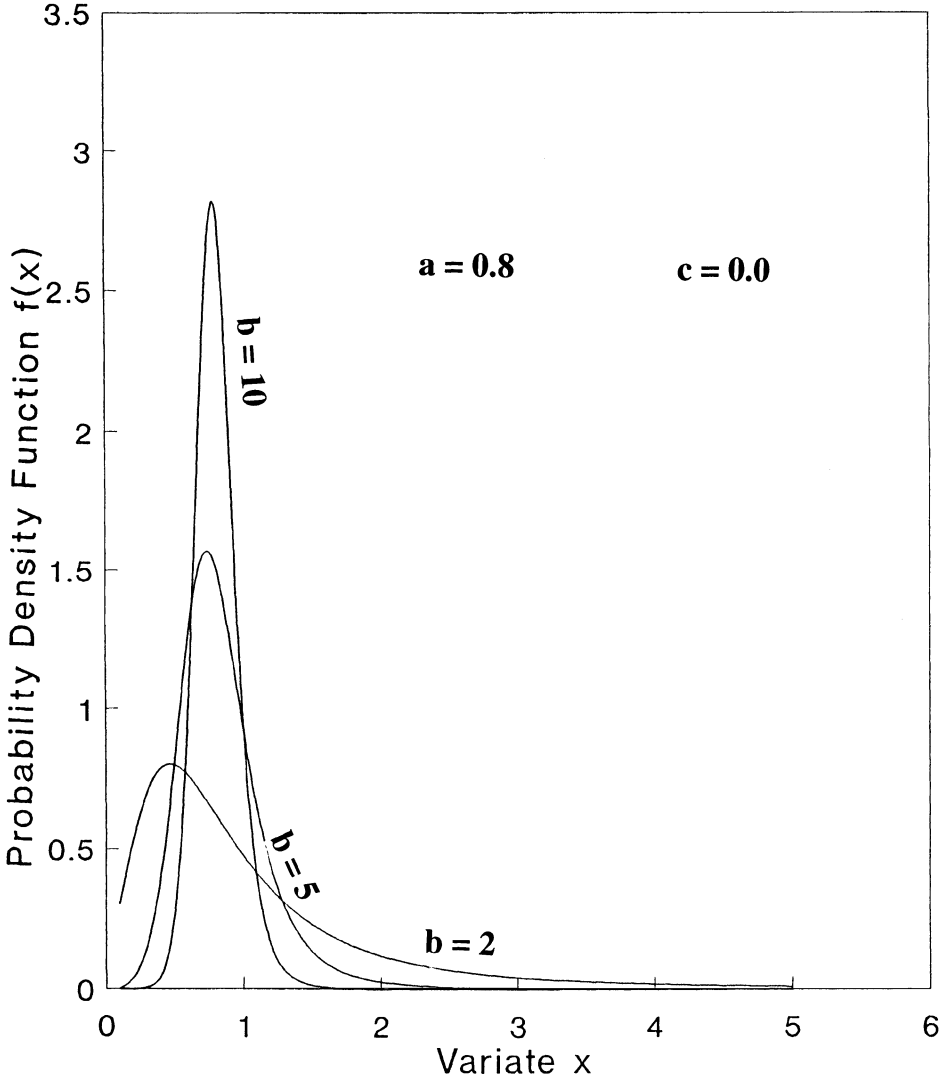

Figure 18.1 LLD3 density function with a = 0.8, and b = 2, 5, and 10.

18.1.2 CONSTRUCTION OF ZEROTH LAGRANGE MULTIPLIER

The least-biased pdf corresponding to POME and consistent with equations (18.7) to (18.9) takes the following form:

$$f(x) = \exp\{-a_0 - a_1 \ln(\frac{x-c}{a}) - a_2 \ln[1 + (\frac{x-c}{a})^b]\} \qquad (18.10)$$

where a_0, a_1, and a_2 are Lagrange multipliers. The mathematical rationale for equation (18.9) has been presented by Tribus (1969). Insertion of equation (18.10) into equation (18.7) yields the partition function:

$$\exp(a_0) = \int_c^\infty \exp\left(-a_1 \ln\left(\frac{x-c}{a}\right) - a_2 \ln\left[1 + \left(\frac{x-c}{a}\right)^b\right]\right) = \frac{b\,\Gamma\left(\frac{1-a_1}{b}\right)\Gamma\left(a_2 - \frac{1-a_1}{b}\right)}{\Gamma(a_2)} \tag{18.11}$$

where $\Gamma(\bullet)$ is the gamma function.

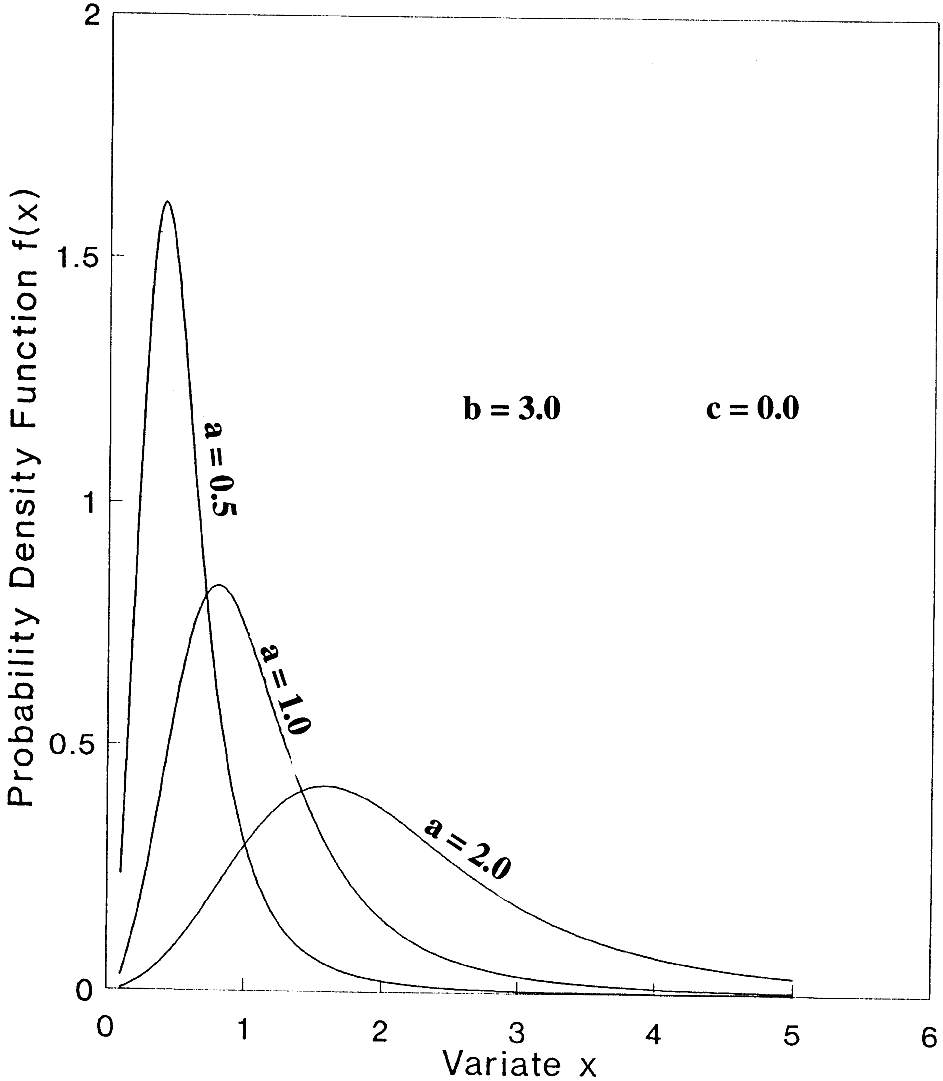

Figure 18.2 LLD3 density function with a = 0.5, 1.0, and 2.0; and b = 3.

The zeroth Lagrange multiplier is given by equation (18.11) as

$$a_0 = \ln b + \ln \Gamma\left(\frac{1-a_1}{b}\right) + \ln \Gamma\left(a_2 - \frac{1-a_1}{b}\right) - \ln \Gamma(a_2) \qquad (18.12)$$

One also gets the zeroth Lagrange multiplier from equation (18.11) as

$$a_0 = \ln \int_c^\infty \exp\left(-a_1 \ln\left(\frac{x-c}{a}\right) - a_2 \ln\left[1+\left(\frac{x-c}{a}\right)^b\right]\right) dx \qquad (18.13)$$

18.1.3 RELATION BETWEEN LAGRANGE MULTIPLIERS AND CONSTRAINTS

Differentiating equation (18.13) with respect to a_1 and a_2, respectively, one gets

$$\frac{\partial a_0}{\partial a_1} = \frac{\int_c^\infty \exp\{-a_1 \ln(\frac{x-c}{a}) - a_2 \ln[1+(\frac{x-c}{a})^b]\}\{-\ln(\frac{x-c}{a})\}dx}{\int_c^\infty \exp\{-a_1 \ln(\frac{x-c}{a}) - a_2 \ln[1+(\frac{x-c}{a})^b]\}dx} \qquad (18.14)$$

$$= -E\left[\ln\left(\frac{x-c}{a}\right)\right]$$

$$\frac{\partial a_0}{\partial a_2} = \frac{\int_c^\infty \exp\{-a_1 \ln(\frac{x-c}{a}) - a_2 \ln[1+(\frac{x-c}{a})^b]\}\{-\ln[1+(\frac{x-c}{a})^b]\}dx}{\int_c^\infty \exp\{-a_1 \ln(\frac{x-c}{a}) - a_2 \ln[1+(\frac{x-c}{a})^b]\}dx}$$

$$= -E\left\{\ln\left[1+\left(\frac{x-c}{a}\right)^b\right]\right\} \qquad (18.15)$$

Similarly, differentiation of equation (18.12) with respect to a_1 and a_2, respectively, yields

$$\frac{\partial a_0}{\partial a_1} = \frac{1}{b}[-\Psi(k_1) + \Psi(k_2)] \qquad (18.16)$$

$$\frac{\partial a_0}{\partial a_2} = \Psi(k_2) - \Psi(a_2) \qquad (18.17)$$

where $k_1 = [1-a_1]/b$ and $k_2 = [a_2 - (1-a_1)/b]$. Equating equation (18.14) to equation (18.16) and equation (18.15) to equation (18.17), we obtain

$$E\left[\ln\left(\frac{x-c}{a}\right)\right] = \frac{1}{b}[\Psi(k_1) - \Psi(k_2)] \qquad (18.18)$$

$$E\left[\ln\left\{1+\left(\frac{x-c}{a}\right)^b\right\}\right] = \Psi(a_2) - \Psi(k_2) \qquad (18.19)$$

18.1.4 RELATION BETWEEN PARAMETERS AND LAGRANGE MULTIPLIERS

Insertion of equation (18.12) into equation (18.10) yields

$$f(x)=\frac{b}{a}\frac{\Gamma(a_2)}{\Gamma[(1-a_1)/b]\Gamma[a_2-(1-a_1)/b]}(\frac{x-c}{a})^{-a_1}[1+(\frac{x-c}{a})^b]^{-a_2} \quad (18.20)$$

A comparison of equation (18.14) with equation (18.2) yields $a_2 = 2$ and $a_1 = 1\text{-}b$.

18.1.5 RELATION BETWEEN PARAMETERS AND CONSTRAINTS

The LLD3 has three parameers which are related to the Lagrange multipliers by bequations (18.18) and (18.19) which, in turn, are related to the constraints through equations (18.20). Eliminating the Lagrange multipliers between these two sets of equations yields the relation between parameters and constraints. Therefore, we get

$$E[\ln(\frac{x-c}{a})]=0 \quad (18.21)$$

$$E[\ln\{1+(\frac{x-c}{a})^b\}]=1 \quad (18.22)$$

However, equations (18.21) and (18.22) need to be supplemented by another equation to get a unique determination of parameters a, b, and c. This is obtained by recalling that

$$\frac{\partial^2 a_0}{\partial a_1^2}=Var[\ln(\frac{x-c}{a})] \quad (18.23)$$

Differentiation of equation (18.16) produces

$$\frac{\partial^2 a_0}{\partial a_1^2}=\frac{1}{b^2}[\frac{\partial}{\partial k_1}\{\Psi(k_1)\}+\frac{\partial}{\partial k_2}\{\Psi(k_2)\}] \quad (18.24)$$

Equating equation (18.23) to equation (18.24) gives

$$Var[\ln\{(x-c)/a\}]=\frac{1}{b^2}[\frac{\partial\Psi}{\partial k_1}+\frac{\partial\Psi}{\partial k_2}] \quad (18.25)$$

18.1.6 DISTRIBUTION ENTROPY

The entropy of the LLD3 is given as

$$I(f)=\ln b-\ln a+(b-1)E[(\frac{x-c}{a})^b]-2E\{\ln[1+(\frac{x-c}{a})^b]\} \quad (18.26)$$

18.2 Parameter-Space Expansion Method

18.2.1 SPECIFICATION OF CONSTRAINTS

The constraints for this method are give by equations (18.7)-(18.9).

18.2.2 DERIVATION OF ENTROPY FUNCTION

Taking logarithm of equation (18.20), we get

$$\ln f(x) = \ln(\frac{b}{a}) + \ln\Gamma(a_2) - \ln\Gamma(\frac{1-a_1}{b}) - \ln\Gamma(a_2 - \frac{1-a_1}{b}) - a_1 \ln(\frac{x-c}{a}) - a_2 \ln[1+(\frac{X-C}{a})^b] \quad (18.27)$$

Thus, the entropy function I(f) of LLD3 can be expressed as

$$I(f) = \ln a - \ln b + \ln\Gamma(\frac{1-a_1}{b}) + \ln\Gamma(a_2 - \frac{1-a_1}{b}) - \ln\Gamma(a_2) + a_1 E[\ln(\frac{x-c}{a})] + a_2 E[\ln(1+(\frac{x-c}{a})^b)] \quad (18.28)$$

18.2.3 RELATION BETWEEN PARAMETERS AND CONSTRAINTS

According to Singh and Rajagopal (1986), the relation between distribution parameters and constraints is obtained by taking partial derivatives of the entropy function with respect to Lagrange multipliers as well as distribution parameters and then equating these derivatives individually to zero. To that end, taking partial derivative of equation (18.28) with respect to a_1, a_2, a, b and c separately, and equating each derivative to zero yields

$$\frac{\partial I}{a_1} = \frac{1}{b}[-\Psi(k_1) + \Psi(k_2)] + E[\ln(\frac{x-c}{a})] = 0 \quad (18.29)$$

$$\frac{\partial I}{\partial a_2} = \Psi(k_2) - \Psi(a_2) + E[\ln(1+(\frac{x-c}{a})^b] = 0 \quad (18.30)$$

$$\frac{\partial I}{\partial a} = \frac{1}{a} - \frac{a_1}{a} - \frac{b}{a} a_2 E[\frac{\{(x-c)/a\}^b}{1+\{1+(x-c)/a\}^b}] = 0 \quad (18.31)$$

$$\frac{\partial I}{\partial b} = -\frac{1}{b} - \Psi(k_1)(\frac{1-a_1}{b^2}) + \Psi(k_2)(\frac{1-a_2}{b}) = 0 \quad (18.32)$$

$$\frac{\partial I}{\partial c} = -a_1\, E\left[\frac{1}{x-c}\right] - \frac{b}{a} a_2\, E\left[\frac{\{(x-c)/a\}^{b-1}}{1+\{(x-c)/a\}^b}\right] = 0 \qquad (18.33)$$

where $k_1 = (1-a_1)/b$, $k_2=a_2-(1-a_1)/b$ and ψ is digamma function = $d[\ln \Gamma(x)]/dx$. Simplification of equations (18.29) to (18.33), respectively, yields

$$E\left[\ln\left(\frac{x-c}{a}\right)\right] = \frac{1}{b}\left[\Psi(k_1) - \Psi(k_2)\right] \qquad (18.34)$$

$$E\left[\ln\left\{1+\left(\frac{x-c}{a}\right)^b\right\}\right] = \Psi(a_2) - \Psi(k_2) \qquad (18.35)$$

$$a_2\, b\, E\left[\frac{\{(x-c)/a\}^b}{1+\left(\frac{x-c}{a}\right)^b}\right] = 1 - a_1 \qquad (18.36)$$

$$a_2\, E\left[\frac{\{(x-c)/a\}^b \ln\{(x-c)/a\}}{1+\{(x-c)/a\}^b}\right] = \frac{1}{b} + \frac{1-a_1}{b^2}\Psi(k_1) - \frac{1-a_2}{b^2}\Psi(k_2) \qquad (18.37)$$

$$a_2\, b\, E\left[\frac{\{(x-c)/a\}^{b-1}}{1+\{(x-c)/a\}^b}\right] = -a\, a_1\, E\left[\frac{1}{x-c}\right] \qquad (18.38)$$

Note that $a_2=2$ and $a_1=1-b$. Therefore, equations (18.34) to (18.38) become

$$E\left[\ln\left(\frac{x-c}{a}\right)\right] = 0 \qquad (18.39)$$

$$E\left[\ln\left\{1+\left(\frac{x-c}{a}\right)^b\right\}\right] = 1 \qquad (18.40)$$

$$2\, E\left[\frac{\{(x-c)/a\}^b}{1+\{(x-c)/a\}^b}\right] = 1 \qquad (18.41)$$

$$2b\, E\left[\frac{\{(x-c)/a\}^b \ln\{(x-c)/a\}}{1+\{(x-c)/a\}^b}\right] = 1 \qquad (18.42)$$

$$2b\, E\left[\frac{\{(x-c)/a\}^{b-1}}{1+\{(x-c)/a\}^b}\right] = a(b-1)\, E\left[\frac{1}{x-c}\right] \qquad (18.43)$$

Equations (18.41) and (17.42) are identities which can be proved as follows. For equation (17.41) we write

$$E[\frac{\{(x-c)/a\}^b}{1+\{(x-c)/a\}^b}]=\int_c^\infty \frac{\{(x-c)/a\}^b}{1+(x-c)/a\}^b}\frac{(b/a)\{(x-c)/a\}^{b-1}}{[1+\{(x-c)/a\}^b]^2}dx$$

$$=b\int_0^\infty \frac{y^{2b-1}}{(1+y^b)^3}dy\,,y=(x-c)/a \tag{18.44}$$

$$=B(2,1)=1/2\,,B(\bullet,\bullet)=beta\ function$$

Therefore,

$$2\,E[\frac{\{(x-c)/a\}^b}{1+\{(x-c)/a\}^b}]=1 \tag{18.45}$$

Similarly, we write for equation (18.42):

$$E[\frac{\{(x-c)/a\}^b\ln\{(x-c)/a\}}{1+\{(x-c)/a\}^b}]=E[\ln(\frac{x-c}{a})]-E[\frac{\ln\{(x-c)/a\}}{1+\{(x-c)/a\}^b}]$$

$$=-\int_c^\infty \frac{\ln\{(x-c)/a\}}{1+\{(x-c)/a\}^b}\frac{b\{(x-c)/a\}^{b-1}}{a[1+\{(x-c)/a\}^b]^2}dx$$

$$=-\int_0^\infty \frac{b\ln(y)\,y^{b-1}}{(1+y^b)^3}dy\,,y=(x-c)/a$$

$$=0.5\int_0^\infty \ln(y)\,d\,(1+y^b)^{-2}$$

$$=\frac{1}{2b}\int_1^\infty \ln(z-1)\,d\,z^{-2},\ z=1+y^b$$

$$=-\frac{1}{b}\lim_{z\to 0}\ln(z-1)-\frac{1}{2b}\int_1^\infty \frac{1}{z^2(z-1)}dz$$

$$= -\frac{1}{2b}\lim_{z\to 0} \ln(z-1) - \frac{1}{2b}\int_1^{\infty} \left(-\frac{1}{z}-\frac{1}{z^2}+\frac{1}{z-1}\right)dz$$

$$= \frac{1}{2b} \tag{18.46}$$

Thus,

$$2bE\left[\frac{\{(x-c)/a\}^b \ln\{(x-c)/a\}}{1+\{(x-c)/a\}^b}\right]=1 \tag{18.47}$$

Thus, the POME-based estimation equations are equations (18.39), (18.40) and (18.43).

18.3 Other Methods of Parameter Estimation

Three other popular methods of parameter estimation are briefly outlined here: the methods of moments (MOM), probability-weighted moments (PWM) and maximum likelihood estimation (MLE).

18.3.1 METHOD OF MOMENTS

For the method of moments (MOM), the moment estimators can be expressed, following Ahmad, et al. (1988), as

$$E[x] = c + a\,B(1+1/b,\ 1-1/b) \tag{18.48}$$

$$Var[x] = a^2\,[B(1+2/b,\ 1-2/b) - B^2(1+1/b,\ 1-1/b)] \tag{18.49}$$

$$G(x) = B(1+3/b,1-3/b)-3B(1+2/b,1-2/b)B(1+1/b,1-1/b)+2B3(1+1/b,1-1/b) \tag{18.50}$$

where E[x], Var[x] and G[x] are the expectation, variance and skewness of X, respectively, and B(•,•) is the beta function. Figure 18.3 shows the relation between G and b. Parameters a, b, and c are estimated by replacing E(x), $\sqrt{Var(x)}$, and G(x) by the sample mean, sample standard deviation and sample coefficient of skewness, respectively.

18.3.2 PROBABILITY WEIGHTED MOMENTS (PWM)

For the probability weighted moments (PWM), the parameter estimators for LLD3 are given by (Guo and Singh, 1992):

$$b = \frac{2W_1-W_0}{6W_1-W_0-6W_2} \tag{18.51}$$

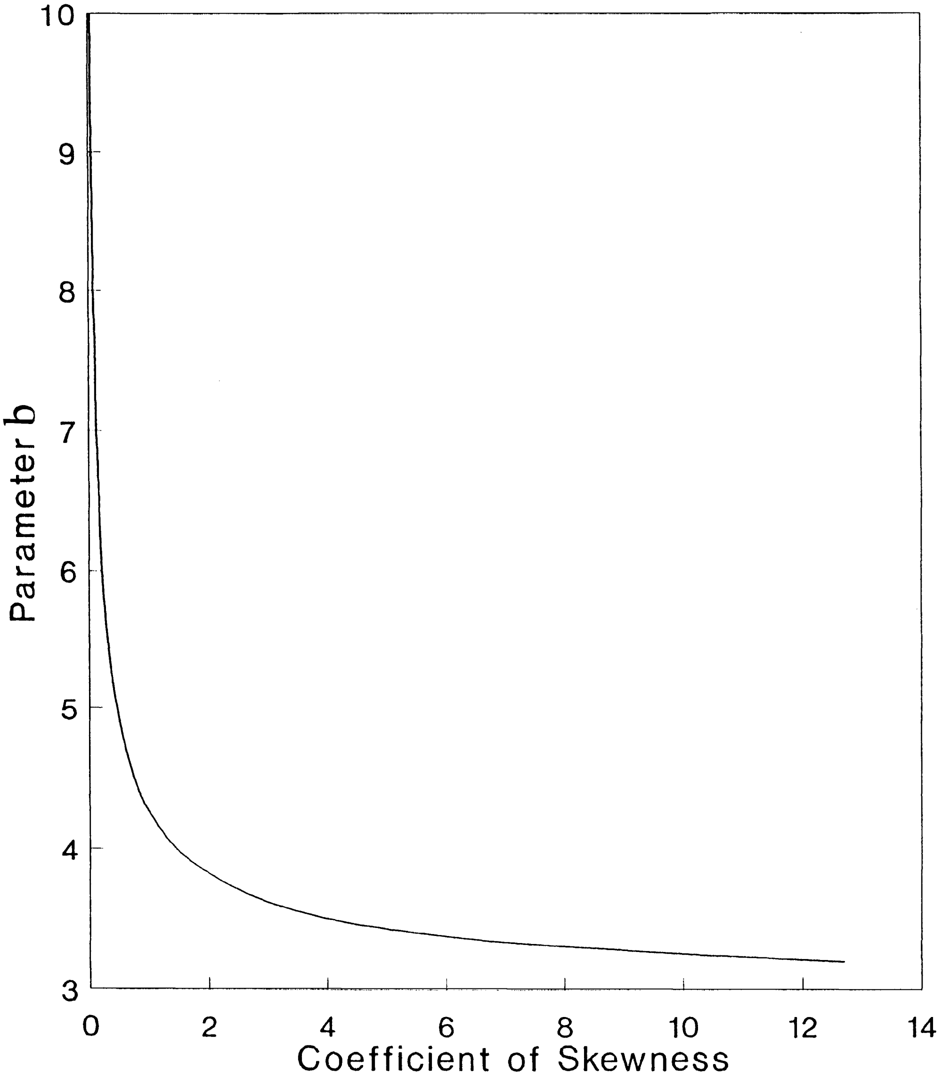

Figure 18.3 Parameter b versus skewness for LLD3 distribution.

$$a = \frac{(W_0-2W_1)b}{\Gamma(1+1/b)\Gamma(1-1/b)} \tag{18.52}$$

$$c = W_0 - a\,\Gamma(1+1/b)\Gamma(1-1/b) \tag{18.53}$$

where the k-th probability-weighted moment is

$$W_k = \frac{a\Gamma(k+1+1/b)\Gamma(1-1/b)}{\Gamma(k+2)} + \frac{c}{k+1}, \qquad k = 0, 1, 2 \ldots \quad (18.54)$$

Parameters a, b and c may be estimated by the sample probability-weighted moments W_k as

$$W_k = \frac{1}{n}\sum_{i=1}^{n} x_i \left(1 - \frac{i-0.35}{n}\right)^k, k = 0,1,2,\ldots\ldots \quad (18.55)$$

where x_i is an ordered random sample $x_1 \leq x_2 \ldots \leq x_n$, and n is the sample size.

18.3.3 METHOD OF MAXIMUM LIKELIHOOD ESTIMATION

For the method of maximum likelihood estimation (MLE), the estimation equations can be expressed (Singh and Guo, 1992) as

$$2\sum_{i=1}^{n}\left[\frac{\{(x_i - c)/a\}^b}{1+\{(x_i - c)/a\}^b}\right] = n \quad (18.56)$$

$$2b\sum_{i=1}^{n}\left[\frac{\ln\{((x_i - c)/a\}\{(x_i - c)/a\}^b}{1+\{(x_i - c)/a\}^b}\right] - b\sum_{i=1}^{n}\ln\{(x_i - c)/a\} - n = 0 \quad (18.57)$$

$$2b\sum_{i=1}^{n}\left[\frac{\{(x_i - c)/a\}^b}{1+\{(x_i - c)/a\}^b}\right] - a(b-1)\sum_{i=1}^{n}\left[\frac{1}{x_i - c}\right] = 0 \quad (18.58)$$

where n is the sample size. These three equations were solved using an iterative scheme. First, with an assumed value of b and c, equation (18.56) was solved for a. With this value of a and the initial guess of c, equation (18.57) was solved to give a new value of b. Then, a new value of c was calculated from equation (18.58). The iterative scheme was terminated when the parameters no longer changed significantly.

18.4 Comparative Evaluation of Parameter Estimation Methods

18.4.1 MONTE CARLO SIMULATED SAMPLE DATA

Guo and Singh (1992) and Singh et al. (1993) made a comparative assessment of MOM, MLE, PWM and POME using Monte Carlo simulated data. Their work is briefly summarized here. Two population cases, listed in Table 18.1, were considered. For each population case, 1000 random samples of size 20, 50 and 100 were generated, and then parameters and quantiles were estimated. The samples were generated using equation (18.3) and Figure 18.3, with the proviso that the population mean $\mu = 1$, the coefficient of variation = 0.5, and the coefficient of skewness = 0.5, and 2.0. This range of the skewness coefficient encompasses a broad class of hydrologic data.

18.4.2 PERFORMANCE INDICES

The performance of the parameter estimation methods was evaluated using the standardized bias (BIAS) and root mean square error (RMSE) of both parameters and quantiles. The number of samples of 1000 may arguably not be large enough to produce the true values of BIAS and RMSE, but will suffice to compare the performance of different estimation methods.

18.4.3 BIAS OF PARAMETER ESTIMATES

The values of BIAS in parameter estimates for LLD3 showed that for G = 0.5, of the four methods, POME yielded the least bias in parameter estimates for all sample sizes. However, its bias increased slightly with increasing sample size. The bias by PWM fluctuated with the sample size. MOM had less bias than MLE, but for both methods, the bias showed only a small reduction with increasing sample size. For G = 2.0, in absolute terms POME produced the least bias, PWM the second least bias except for n = 20, and MLE produced the third least bias. The highest bias was produced by MOM except for n = 20. With increasing sample size, the bias by MOM as well as by PWM decreased significantly. This was true for all three parameters. Thus, in terms of parameter bias, it is concluded that POME is the preferred estimator, regardless of the sample size and skewness.

Table 18.1 LLD3 population cases considered in sampling experiments.

LLD3 Population	CV	G	a	b	c
Case 1	0.5	0.5	4.843	13.773	-3.868
Case 2	0.5	2.0	1.387	5.653	-0.459

18.4.4 RMSE OF PARAMETER ESTIMATES

The values of RMSE of parameters estimated by the four methods showed that for G = 0.5, POME, MOM and MLE had comparable RMSE values in estimates of all three parameters for all sample sizes. PWM produced unrealistically high RMSE values. When G = 2.0, MLE and POME yielded comparable RMSE values for the three parameters. MOM had quite high RMSE values but they declined for large sample sizes. PWM did not perform well. Thus, overall POME or MLE would be preferable. However, for small values of skewness, MOM would be an equally good choice.

18.4.5 BIAS OF QUANTILE ESTIMATES

The results of bias in quantile estimation for LLD3 showed that for G = 0.5, MOM and PWM produced the least bias in quantile estimates for all sample sizes if the probability of nonexceedance (P) was less than or equal to 0.90. Of course, all four methods had fairly low bias. For $P \geq 0.99$, POME produced the highest bias, and PWM, MOM and MLE had low biases, with their comparative values depending upon the values of P and sample size n. For G = 2.0 and $P \leq 0.9$, POME had the least bias, and PWM had the second smallest bias, irrespective of the sample size. When P increased past 0.9, the bias of POME deteriorated for small sizes.

Overall, MOM, PWM, MLE and POME yielded quite low values of bias. It may, therefore, be inferred that for low values of G (≤ 0.5) and P (≤ 0.9), PWM is the preferred method but for high values of G, POME would be preferable. For large values of P (≥ 0.99), MLE or MOM would be preferred. For high values of G, POME would be preferable for $P \leq 0.90$, but PWM would be preferable for P past the value of 0.999.

18.4.6 RMSE OF QUANTILE ESTIMATES

The values of RMSE in quantile estimates of the four methods showed that for $G = 0.5$ and $P \leq 0.9$, all four methods produced comparable values of RMSE, with MOM having the lowest RMSE values and PWM the highest. When $P \geq 0.99$, RMSE of POME as well as MLE deteriorated significantly, especially for small sample sizes and low values of G, but MOM and MLE remained comparable. For $G = 2.0$, POME produced the least RMSE for all sample sizes. For $P \leq 0.9$, all estimators were comparable, and for $P \geq 0.99$, MLE and POME were comparable. Thus, it may be concluded that for high values of G, POME is the preferred method, especially for $P \geq 0.99$, but MLE is also a good choice. For low values of G, anyone of the four estimators would be adequate if $P \leq 0.9$, but MLE or MOM would be preferable for $P \geq 0.99$.

18.4.7 CONCLUDING REMARKS

Of the four methods, POME yielded the least parameter bias for all sample sizes. POME was comparable to MOM and MLE in terms of RMSE of parameters estimates. For high skewness ($G = 2.0$), the bias in quantile estimates by POME was comparable to that by MOM, MLE and PWM. For low values of skewness ($G = 0.5$), POME was comparable to the other 3 methods for lower values of probability of nonexceedance. However, POME performed poorly when P exceeded 0.99. In terms of RMSE in quantile estimates, POME was either better than or comparable to the other 3 methods, especially for large values of skewness and large return periods. Overall, POME performed the best for $G \geq 2.0$ and $P \geq 0.9$; MLE for $G \leq 0.5$ and $P > 0.9$, and PWM for $G \leq 0.5$ and $P \leq 0.9$, $G \geq 2$ and $P \leq 0.8$. In the cases tested, MOM performed reasonably well and usually in between the best and the worst estimators.

References

Ahmad, M. I., Sinclair, C. D. and Werritti, A., 1988. Log-logistic flood frequency analysis. *Journal of Hydrology*, Vol. **98**, pp. 205-224.

Bennett, S., 1983. Log-logistic regression models for survival data. *Applied Statistics*, Vol. **32**, pp. 165-171.

Burr, I. W., 1942. Cumulative frequency functions. *Annals of Mathematical Statistics*, Vol. **13**, pp. 215-232.

Guo, H. and Singh, V. P., 1993. A comparative evaluation of estimators of log-logistic distributions by Monte Carlo simulation. *Technical Report* WRR**26**, Water Resources Program, Department of Civil Engineering, Louisiana State University, Baton Rouge, Louisiana, U.S.A.

Haktanir, T., 1991. Statistical modeling of annual maximum flows in Turkish rivers. *Hydrological Sciences Journal*, Vol. **36**, No. 4, pp. 367-389.

Jaynes, E. T., 1968. Prior probabilities. *IEEE Transactions on Systems, Management and Cybern*etics, Vol. **3** (SSC-4), pp. 227-241.

Lawless, J. F., 1986. A note on lifetime regression models. *Biometrika*, Vol. **73,** No. 2, pp. 509-512.

Mielke, P. W. and Johnson, E. S., 1973. Three-parameter kappa distribution maximum likelihood estimates and likelihood ratio tests. *Monthly Weather Review*, Vol. **101**, pp. 701-709.

Shoukri, M. M., Mian, I. U. H. and Tracy, D. S., 1988. Sampling properties of estimators of the log-logistic distribution with application to Canadian precipitation data. *The Canadian Journal of Statistics,* Vol. **16**, No. 3, pp. 223-236.

Singh, V.P., Guo, H. and Yu, F.X., 1993. Parameter estimation for 3-parameter log-logistic distribution (LLD3) by POME. *Stochastic Hydrology and Hydraulics*, Vol. **7,** pp. 163-177.

Singh, V. P. and Rajagopal, A. K., 1986. A new method of parameter estimation for hydrologic frequency analysis. *Hydrological Science and Technology*, Vol. **2**, No. 3, pp. 33-40.

Tribus, M., 1969. *Rational Descriptors, Decisions and Designs*. Pergamon Press, New York.

CHAPTER 19

TWO-PARAMETER PARETO DISTRIBUTION

The Pareto distribution was introduced by Pickands (1975) and has since been applied to a number of areas including socio-economic phenomena, physical and biological processes (Saksena and Johnson, 1984), reliability studies and the analysis of environmental extremes. Davison and Smith (1990) pointed out that the Pareto distribution might form the basis of a broad modeling approach to high-level exceedances. DuMouchel (1983) applied it to estimate the stable index α to measure tail thickness, whereas Davison (1984a, 1984b) modeled contamination due to long-range atmospheric transport of radionuclides. van Montfort and Witter (1985, 1986, 1991) applied the Pareto distribution to model the peaks over threshold (POT) streamflows and rainfall series, and Smith (1984, 1987) applied it to analyze flood frequencies. Similarly, Joe (1987) employed it to estimate quantiles of the maximum of a set of observations. Wang (1991) applied it to develop a peak over threshold (POT) model for flood peaks with Poisson arrival time, whereas Rosbjerg et al. (1992) compared the use of the 2-parameter Pareto and exponential distributions as distribution models for exceedances with the parent distribution being a generalized Pareto distribution. In an extreme value analysis of the flow of Burbage Brook, Barrett (1992) used the Pareto distribution to model the POT flood series with Poisson interarrival times. Davison and Smith (1990) presented a comprehensive analysis of the extremes of data by use of the Pareto distribution for modeling the sizes and occurrences of exceedances over high thresholds.

Methods for estimating parameters of the 2-parameter Pareto distribution were reviewed by Hosking and Wallis (1987). Quandt (1966) used the method of moments (MOM), and Baxter (1980), and Cook and Mumme (1981) used the method of maximum likelihood estimation (MLE). MOM, MLE, and probability weighted moments (PWM) were included in the review. van Montfort and Witter (1986) used MLE to fit the Pareto distribution to represent the Dutch POT rainfall series, and used an empirical correction formula to reduce the bias of the scale and shape parameter estimates. Davison and Smith (1990) used MLE, PWM, a graphical method, and least squares to estimate the Pareto distribution parameters. Singh and Guo (1995) employed the principle of maximum entropy (POME) to derive a new method of parameter estimation (Singh and Rajagopal, 1986) for the 2-parameter Pareto distribution. Monte Carlo simulated data were used to evaluate this method and compare it with the methods of moments (MOM), probability weighted moments (PWM), and maximum likelihood estimation (MLE). The parameter estimates yielded by POME were either superior or comparable for small sample sizes when bias and root mean square error (RMSE) were used as the criteria, and were either comparable or adequate for large sample sizes. Their work is followed here.

For a random variable X, the two-parameter Pareto distribution (PD2) has the cumulative distribution function (cdf) given by

V. P. Singh, *Entropy-Based Parameter Estimation in Hydrology*

$$F(x) = 1 - \left(\frac{a}{x}\right)^b, \ x > a, b > 0 \tag{19.1a}$$

and the probability density function (pdf) given by

$$f(x) = b\, a^b\, x^{-b-1} \tag{19.1b}$$

where a is the location parameter and b is the shape parameter. The shapes of the Pareto distribution for various values of b are illustrated in Figure 19.1.

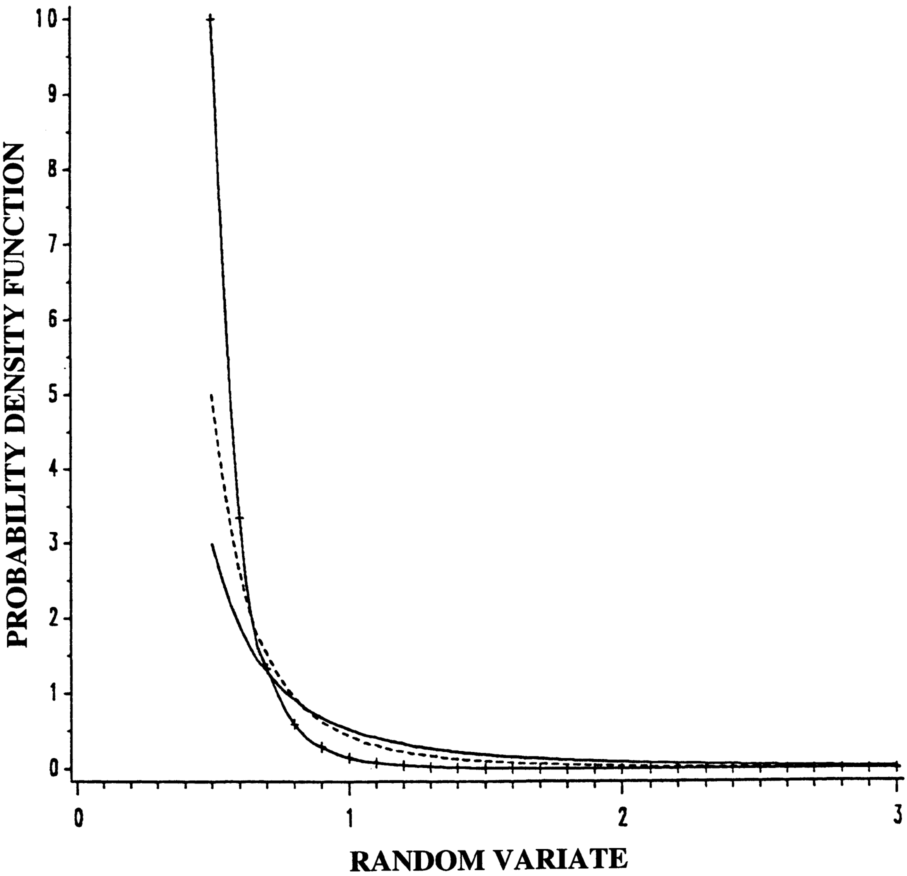

Figure 19.1 2-parameter Pareto density function with b = 1.5, 2.5, 5; line: b = 1.5, dash: b = 2.5, and plus: b = 5.

19.1 Ordinary Entropy Method

19.1.1 SPECIFICATION OF CONSTRAINTS

Taking logarithm to the base e of equation (19.1a), one gets

$$\ln f(x) = \ln b + b \ln a - (b+1) \ln x \qquad (19.2a)$$

Multiplying equation (19.2a) by [- f(x)] and then integrating, one obtains the entropy function:

$$\begin{aligned} I(f) &= -\int_a^\infty [\ln b + b \ln a - (b-1) \ln x] f(x) dx \\ &= -\ln b - b \ln a + (b-1) E[\ln x] \end{aligned} \qquad (19.2b)$$

Following Singh and Rajagopal (1986), the constraints from equation (19.2b), appropriate for equation (19.1b), are

$$\int_a^\infty f(x)\, dx = 1 \qquad (19.3)$$

$$\int_a^\infty \ln x\, f(x)\, dx = E[\ln x] \qquad (19.4)$$

These constraints specify the information sufficient for PD2. Because the information is determined from data, the parameters and other statistics of the distribution can be physically interpreted.

19.1.2 CONSTRUCTION OF ZEROTH LAGRANGE MULTIPLIER

The least-biased pdf of PD2 corresponding to POME and consistent with equations (19.3) and (19.4) takes the form

$$f(x) = \exp(-a_0 - a_1 \ln x) \qquad (19.5)$$

where a_0 and a_1 are Lagrange multipliers. The mathematical rationale for equation (19.5) has been presented by Tribus (1969), Levine and Tribus (1978), among others. By applying equation (19.5) to the total probability condition in equation (19.3), one obtains

$$\int_a^\infty f(x)\, dx = \int_a^\infty \exp(-a_0 - a_1 \ln x)\, dx = 1 \qquad (19.6)$$

Equation (19.6) produces the partition function as

$$\exp(a_0) = \int_a^\infty \exp(-a_i \ln x) dx \qquad (19.7a)$$

which yields

$$\exp(a_0) = \frac{1}{a_1 - 1} a^{-a_1+1}, \ (a_1 > 1) \qquad (19.7b)$$

The zeroth Lagrange multiplier is given equation (19.7b) by

$$a_0 = - \ln (a_1 - 1) - (a_1 - 1) \ln a \tag{19.8}$$

The zeroth Lagrange multiplier is also obtained from equation (19.7a) as

$$a_0 = \ln \left[\int_a^\infty \exp(- a_1 \ln x)\, dx \right] \tag{19.9}$$

19.1.3 RELATION BETWEEN LAGRANGE MULTIPLIERS AND CONSTRAINTS

The relation between Lagrange multipliers and constraints is obtained by taking partial derivatives of the zeroth Lagrange multiplier with respect to other Lagrange multipliers. Thus, differentiation of equation (19.8) with respect to a_1 yields

$$\frac{\partial a_0}{\partial a_1} = - \frac{1}{a_1 - 1} - \ln a \tag{19.10}$$

Differentiation of equation (19.10) with respect to a_1 gives

$$\frac{\partial^2 a_0}{\partial a_1^2} = \frac{1}{(a_1 - 1)^2} \tag{19.11}$$

Differentiating equation (19.9) with respect to a_1 gives

$$\frac{\partial a_0}{\partial a_1} = - \frac{\int_a^\infty \exp(- a_0 - a_1 \ln x) \ln x \, dx}{\int_a^\infty \exp(- a_0 - a_1 \ln x)\, dx}$$

$$= - \int_a^\infty \exp(- a_0 - a_1 \ln x) \ln x \, dx = - E[\ln x] \tag{19.12}$$

Differentiation of equation (19.12) with respect to a_1 yields

$$\begin{aligned} \frac{\partial^2 a_0}{\partial a_1^2} &= - \int_a^\infty \exp(- a_0 - a_1 \ln x) \ln x \left(- \frac{\partial a_0}{\partial a_1} - \ln x\right) dx \\ &= \frac{\partial a_0}{\partial a_1} \int_a^\infty \exp(- a_0 - a_1 \ln x) \ln x \, dx + \int_a^\infty \exp(- a_0 - a_1 \ln x) \ln^2 x \, dx \\ &= - E^2[\ln x] + E[\ln^2(x)] = Var(\ln x) \end{aligned} \tag{19.13}$$

where Var(•) denotes the variance of the quantity within brackets, (•). Equating equation (19.10) to equation (19.12), and equation (19.11) to equation (19.13) yields

$$- \frac{1}{a_1 - 1} - \ln a = - E[\ln x] \tag{19.14}$$

$$\frac{1}{(a_1 - 1)^2} = \mathrm{Var}(\ln x) \tag{19.15}$$

19.1.4 RELATION BETWEEN PARAMETERS AND LAGRANGE MULTIPLIERS

Insertion of equation (19.8) into equation (19.5) yields

$$f(x) = (a_1 - 1)\, a^{a_1 - 1}\, x^{-a_1} \tag{19.16}$$

A comparison of equation (19.16) with equation (19.1b) yields

$$a_1 - 1 = b \tag{19.17}$$

19.1.5 RELATION BETWEEN PARAMETERS AND CONSTRAINTS

The PD2 has two parameters a and b which are related to the Lagrange multipliers by equation (19.17) which, in turn, are related to the constraints by equations (19.14) and (19.15). Eliminating the Lagrange multipliers between these two sets of equations yields relations between parameters and constraints. Therefore,

$$\frac{1}{b} + \ln a = E[\ln x] \tag{19.18}$$

$$\frac{1}{b^2} = \mathrm{Var}(\ln x) \tag{19.19}$$

19.1.6 DISTRIBUTION ENTROPY

The entropy function I(x) of PD2 can be defined as

$$I(x) = -\ln b - b \ln a + (b + 1)\, E\{\ln x\} \tag{19.20}$$

where E[•] denotes the expectation of the quantity within brackets [•].

19.2 Parameter-Space Expansion Method

19.2.1 SPECIFICATION OF CONSTRAINTS

Following Singh and Rajagopal (1986), the constraints for this method from equation (19.2b) are given by equation and

$$\int_a^\infty (b+1)\ln x\, dx = E\,[\,(b+1)\ln x\,] \tag{19.21}$$

19.2.2 DERIVATION OF ENTROPY FUNCTION

The pdf corresponding to POME and consistent with equations (19.3) and (19.21) takes the form

$$f(x) = \exp[-a_0 - a_1(b+1)\ln x] \tag{19.22}$$

where a_0 and a_1 are Lagrange multipliers. Insertion of equation (19.22) in equation (19.3) yields the partition function:

$$\begin{aligned} \exp(a_0) &= \int_a^\infty \exp(-a_1(b+1)\ln x\, dx \\ &= \frac{a^{1-a_1(b+1)}}{-1+a_1(b+1)} \end{aligned} \tag{19.23}$$

The zeroth Lagrange multiplier is given by taking the logarithm of equation (19.23) as

$$a_0 = -\ln[a_1(b+1)-1] + [1 - a_1(b+1)\ln a] \tag{19.24}$$

The zeroth Lagrange multiplier is also obtained from equation (19.24) as

$$a_0 = \ln \int_a^\infty \exp[-a_1(b+1)\ln x]dx \tag{19.25}$$

Introduction of equation (19.24) in equation (19.22) gives

$$f(x) = [a_1(b+1)-1]a^{-[1-a_1(b++1)]}x^{-a_1(b+1)} \tag{19.26}$$

A comparison of equation (19.26) with equation (19.1b) shows that $a_1 = 1$.

The entropy function can be written as

$$\begin{aligned} I(f) = &-\ln[a_1(b+1)-1] + [1-a_1(b+1)]\ln a + \\ &+ a_1(b+1)E[x] \end{aligned} \tag{19.27}$$

19.2.3 RELATION BETWEEN PARAMETERS AND CONSTRAINTS

Taking partial derivatives equation (19.27) with respect to a_1, a, and b separately and equating each derivative to zero results in

$$\frac{\partial I}{\partial a_1} = -\frac{b+1}{a_1(b+1)-1} - \ln a(b+1) + (b+1)\ln E[x] = 0 \tag{19.28}$$

$$\frac{\partial I}{\partial a} = \frac{1-a_1(b+1)}{a} = 0 \tag{19.29}$$

$$\frac{\partial I}{\partial b} = -\frac{a_1}{a_1(b+1)-1} - a_1\ln a + a_1 E[x] = 0 \tag{19.30}$$

Simplification of equations (19.28)- (19.29) and noting that a_1 equals 1, respectively, leads to

$$\frac{1}{b} + \ln a = E[x] \tag{19.31}$$

$$b = 0 \tag{19.32}$$

$$\frac{1}{b} + \ln a = E[x] \tag{19.33}$$

Equations (19.31) and (19.33) are identical and equation (19.32) is trivial. Thus, only one useful equation is obtained and one more equation is needed. This is got as before. Therefore, the parameter estimation equations are same as for the ordinary method.

19.3 Other Methods of Parameter Estimation

19.3.1 METHOD OF MOMENTS

For the method of moments (MOM), the moment estimates are given as

$$b = 1 + (1 + \frac{1}{Cv^2})^{0.5} \tag{19.34}$$

$$a = \frac{\bar{x}(b - 1)}{b} \tag{19.35}$$

where $\bar{x}$ and Cv are, respectively, mean and coefficient of variation defined as

$$\bar{x} = \frac{ab}{b-1} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{19.36}$$

$$Cv = \frac{\sigma}{\bar{x}} = \frac{1}{[b(b-2)]^{0.5}}, \quad \sigma = [\sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n-1}]^{0.5} \tag{19.37}$$

where n is the number of observations (or sample size).

19.3.2 METHOD OF MAXIMUM LIKELIHOOD ESTIMATION

For the method of maximum likelihood estimation (MLE), the maximum likelihood estimate of b is

$$b = \frac{n}{\sum_{i=1}^{n} \ln x_i - n \ln a} \tag{19.38}$$

An MLE estimate cannot be obtained for a by differentiating the likelihood function (L) since L is unbounded with respect to a. Because a is the lower bound of the random variable X, L is maximized subject to the constraint $a \leq x_1$, and the lowest sample value x_1 indeed gives a, i.e., $a = x_1$.

19.3.3 METHOD OF PROBABILITY WEIGHTED MOMENTS

The probability weighted moment (PWM) estimates are given by

$$b = \frac{W_0 - W_1}{W_0 - 2W_1} \tag{19.39}$$

$$a = \frac{W_0 W_1}{W_0 - W_1} \tag{19.40}$$

where W_0 and W_1 are the probability-weighted moments defined as

$$W_0 = \int_0^1 a(1 - F)^{-1/b}\, dF = \frac{ab}{b - 1} \tag{19.41}$$

$$W_1 = \int_0^1 a(1 - F)^{-1/b}\, (1 - F)\, dF = \frac{ab}{2b - 1} \tag{19.42}$$

19.4 COMPARATIVE EVALUATION OF ESTIMATION METHODS USING MONTE CARLO EXPERIMENTS

19.4.1 MONTE CARLO SAMPLES

Guo and Singh (1992) and Guo and Singh (1995) assessed the performance of the POME, MOM, MLE, and PWM estimation methods using Monte Carlo sampling experiments. Three Pareto population cases, listed in Table 19.1, were considered. For each population case, 1000 random samples of size 20, 50, and 100 were generated, and then parameters and quantiles were estimated.

Table 19.1 Pareto population cases considered in sampling experiment (μ = 1).

Pareto Population	Cv	Parameter	
		a	b
Case 1	0.5	0.691	3.236
Case 2	1.0	5.86	2.414
Case 3	3.0	0.513	2.054

18.4.2 PERFORMANCE INDICES

The performance of the parameter estimation methods was evaluated using the following performance indices: standardized bias (BIAS) and root mean square error (RMSE) for both parameters and quantiles. It may be noted that the number of Monte Carlo samples (N) of 1,000 may arguably not be large enough to produce the true values of BIAS and RMSE, but will suffice to evaluate the performance of the POME estimation method and compare its performance with that of the other three methods.

18.4.3 BIAS IN PARAMETER ESTIMATES

Of the four methods, MOM produced the highest bias in estimates of both parameters a and b across all sample sizes and the range of the coefficient of variation. For large sample sizes ($N \geq 100$) POME, MLE and PWM yielded the very low but comparable values of bias. Except for MOM, the parameter bias did not increase with increasing Cv. Indeed the bias was little affected by Cv. For small sample sizes ($N \leq 20$), POME and MLE produced the least but comparable values of bias. With increasing sample size, PWM's performance improved significantly. Thus, it is inferred that if the sample size is less than or equal to 20, POME or MLE will be the preferred parameter estimation method. However, forN $\geq$ 50, PWM, POME or MLE with be comparable and either of these three methods could be used.

18.4.4 RMSE IN PARAMETER ESTIMATES

The method producing the highest RMSE across all sample sizes and the range of Cv was MOM. For the remaining 3 methods, the RMSE values were comparable for sample size $N \geq 50$ for all values of Cv. However, for $N \leq 20$, PWM produced high values of RMSE in estimate of parameter a, but not for b. The value of RMSE was not materially affected by the variation of Cv for any method. For $N \geq 50$, MLE resulted in the least RMSE in estimates of parameter b. In this case, the preference of a particular method should be decided by the sample size. If $N \leq 20$, MLE or POME would be the preferred method. For $N \geq 50$, either POME, PWM or MLE would be an acceptable choice.

18.4.5 BIAS IN QUANILE ESTIMATES

The performance of the four estimation methods varied with the probability of nonexceedance P, sample size n, and the coefficient of variation of variation Cv. For $P \leq 0.9$, MLE and POME produced the least bias, but MOM and PWM performed satisfactorily. Thus, regardless of n and Cv, any of the four methods could be used. For $P \geq 0.99$, MLE consistently produced the least bias and POME the second least bias across all sample sizes and the range of Cv. For small N, PWM's bias was the highest, and MOM's the second highest. For small sample sizes, either MLE or POME would be the preferred method.

18.4.6 RMSE IN QUANTILE ESTIMATES

For $P \leq 0.9$, the four methods were comparable for all sample sizes and the range of Cv. For $P \geq 0.99$, the lowest RMSE was produced by MOM and the second lowest by MLE. PWM yielded the highest RMSE and POME the second highest. Thus, in this case, MOM would be the preferred method, especially when N was small; for large N, MLE or POME would be satisfactory.

18.4.7 CONCLUDING REMARKS

To summarize, when sample size ($N \leq 20$) was small, POME produced less or comparable parameter bias. In terms of RMSE, POME was comparable to MLE and preferable to other methods for $N \leq 20$. For $P \leq 0.9$, in terms of bias in quantile estimates, POME was comparable to other three methods, and preferable to PWM and MOM for small sample sizes. For $P \leq 0.9$ in terms of RMSE, POME was comparable to the other 3 methods. For $P \geq 0.99$, POME was comparable to MLE for large sample sizes.

References

Barrett, J. H., 1992. An extreme value analysis of the flow of Burbage Brook. *Stochastic Hydrology and Hydraulics*, Vol. **6**, pp. 151-165.

Baxter, M. A., 1980. Minimum variance unbiased estimation of the parameter of the Pareto distribution. *Biometrika*, Vol. **27**, pp. 133-138.

Cook, W. L. and Mumme, D. C., 1981. Estimation of Pareto parameters by numerical methods. in: *Statistical Distributions in Scientific Work*, C. Taillie et al (eds), Vol. **5**, pp. 127-132.

Davison, A. C., 1984a. Modelling excesses over high thresholds, with an application. in: *Statistical Extremes and Applications*, J. Tiago de Oliveira (ed.), pp. 461-482.

Davison, A. C., 1984b. A statistical model for contamination due to long-range atmospheric transport of radionuclides. *unpublished Ph.D.* thesis, Department of Mathematics, Imperial College of Science and Technology, London, England.

Davison, A. C. and Smith, R. L., 1990. Models for exceedances over high thresholds. J. R. *Statist. Soc.* **B**, Vol. **52**, No. 3, pp. 393-442.

DuMouchel, W., 1983. Estimating the stable index α in order to measure tail thickness. *Ann. Statist.*, Vol. **11**, pp. 1019-1036.

Guo, H. and Singh, V.P., 1992. A comparative evaluation of estimators of Pareto distribution by Monte Carlo simulation. *Technical Report* **WRR25**, Water Resources Progam, Department of Civil and Environmental Engineering, Louisiana State University, Baton Rouge, Louisiana.

Hosking, J. R. M. and Wallis, J. R., 1987. Parameter and quantile estimation for the generalized Pareto distribution. *Technometrics*, Vol. **29**, No. 3, pp. 339-349.

Jaynes, E. T., 1968. Prior probabilities. *IEEE Transactions on Syst. Man. Cybern.*, Vol. **3 (SSC-4)**, pp. 227-241.

Joe, H., 1987. Estimation of quantiles of the maximum of N observations. *Biometrika*, Vol. **74**, pp. 347-354.

Levine, R. D. and Tribus, M., 1979. *The Maximum Entropy Formalism.* MIT Press, Cambridge, Mass.

Pickands, J., 1975. Statistical inference using extreme order statistics. *Ann. Statist.*, Vol. **3,** pp. 119-131.

Quandt, R. E., 1966. Old and new methods of estimation of the Pareto distribution. *Biometrika*, Vol. **10,** pp. 55-82.

Rosbjerg, D., Madsen, H. and Rasmussen, P. F., 1992. Prediction in partial duration series with generalized Pareto-distributed exceedances. *Water Resources Research*, Vol. **28**, No. 11, pp. 3001-3010.

Saksena, S. K. and Johnson, A. M., 1984. Best unbiased estimators for the parameters of a two-parameter Pareto distribution. *Biometrika*, Vol. **31**, pp. 77-83.

Singh, V.P. and Guo, H., 1995. Parameter estimation for 2-parameter Pareto distribution POME. *Water Resources Management*, Vol. **9**, pp. 81-93.

Singh, V. P. and Rajagopal, A. K., 1986. A new method of parameter estimation for hydrologic frequency analysis. *Hydrological Science and Technology,* Vol. **2**, No. 3, pp. 33-40.

Smith, J. A., 1991. Estimating the upper tail of flood frequency distributions. *Water Resources Research,* Vol. **23**, No. 8, pp. 1657-1666.

Smith, R. L., 1984. Threshold methods for sample extremes," in: *Statistical Extremes and Applications,* edited by J. Trago de Oliveita, pp. 621-638.

Smith, R. L., 1987. Estimating tails of probability distributions. *Ann. Statist.*, Vol. **15**, pp. 1174-1207.

Tribus, M., 1969. *Rational Descriptors, Decisions and Designs.* Pergamon Press, New York, N.Y.

van Montfort, M. A. J. and Witter, J. V., 1985. Testing exponentiality against generalized Pareto distribution. *Journal of Hydrology*, Vol. **78**, pp. 305-315.

van Montfort, M. A. J. and Witter, J. W., 1986. The generalized pareto distribution applied to rainfall depths. *Hydrological Sciences Journal*, Vol. **31**, No. 2, pp. 151-162.

van Montfort, M. A. J. and Witter, J. W., 1991. The first and the second e of the extreme value distribution, EV1. *Stochastic Hydrology and Hydraulics*, Vol. **5**, pp. 69-76.

Wang, Q. J., 1991. The POT model described by the generalized Pareto distribution with Poisson arrival rate. *Journal of Hydrology*, Vol. **129**, pp. 263-280.

CHAPTER 20

TWO-PARAMETER GENERALIZED PARETO DISTRIBUTION

The Pareto distribution has been introduced in Chapter 19. Also discussed in the chapter are a brief review of literature and the methods of estimating its parameters. Kotz and Johnson (1985) provided a detailed discusson of the Pareto distributin. Methods for estimating parameters of the 2-parameter generalized Pareto (GP2) distribution were reviewed by Hosking and Wallis (1987). The method of moments (MOM), maximum likelihood estimation (MLE), and probability weighted moments (PWM) were included in the review. Ashkar and Ouarda (1997) presented some methods of fitting the GP2 distribution using Monte Carlo generated data. They discussed six versions of the generalized method of moments. Wang (1991) derived PWMs for both known and unknown thresholds. van Montfort and Witter (1991) used the MLE method to fit the GP2 distribution to represent the Dutch POT rainfall series, and used an empirical correction formula to reduce bias of the scale and shape parameter estimates. Davison and Smith (1990) used MLE, PWM, and a graphical method to estimate the GP2 distribution parameters. Guo and Singh (1992) and Singh and Guo (1997) employed the principle of maximum entropy (POME) to derive a new method of parameter estimation (Singh and Rajagopal, 1986) for the GP2 distribution. They used Monte Carlo simulated data to evaluate this method and compare it with he MOM, PWM, and MLE methods. The parameter estimates yielded by POME were comparable or better within certain ranges of sample size and coefficient of variation.

Consider a random variable Y with the standard exponential distribution. Let a random variable X be defined as X = b(1 - exp (- aY))/a, where a and b are parameters. Then the distribution of X is the 2-parameter generalized Pareto (GP2) distribution which can be expressed as

$$F(x) = 1 - \left(1 - a\frac{x}{b}\right)^{1/a}, \qquad a \neq 0 \tag{20.1}$$

$$= 1 - \exp\left(-\frac{x}{b}\right), \qquad a = 0 \tag{20.2}$$

where b is a scale parameter, a is a shape parameter, and F(x) is the cumulative distribution function (cdf). The probability density function (pdf) of the GP2 distribution follows:

$$f(x) = \frac{1}{b}\left(1 - a\frac{x}{b}\right)^{1/a-1}, \qquad a \neq 0 \tag{20.3}$$

$$= \frac{1}{b}\exp\left(-\frac{x}{b}\right), \qquad a = 0 \tag{20.4}$$

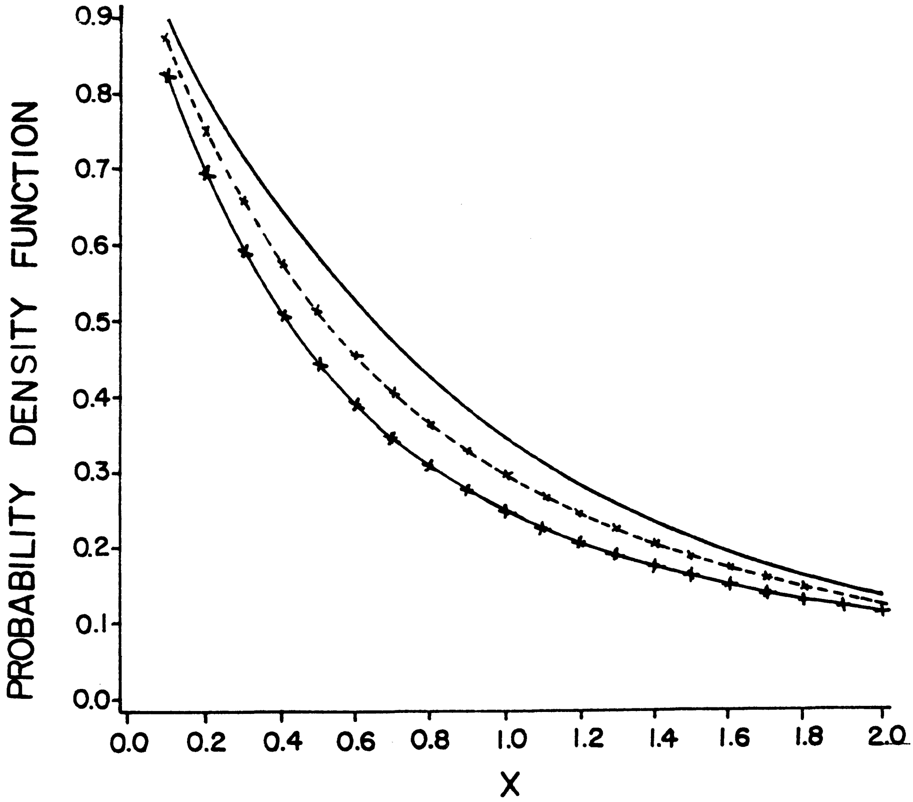

Figure 20.1 (a) Generalized Pareto density function with b = 1, a = -0.1, -0.5, and -1.0; line: a = -0.1, dash: a = -0.5, and plus: a = -1.0.

The Pareto distributions are obtained for $a < 0$. Figure 20.1 graphs the pdf for $b = 1.0$, and various values of a. Pickands (1975) has shown that the GP distribution given by equations (20.1) and (20.2) occurs as a limiting distribution for excesses over thresholds if and only if the parent distribution is in the domain of attraction of one of the extreme-value distributions. The GP2 distribution specializes into the exponential distribution for $a = 0$ and the uniform distribution on [0, b] for $a = 1$.

Some important properties of the GP2 distribution worth mentioning are: (1) By comparison with the exponential distribution, the GP2 distribution has a heavier tail for $a < 0$ (long-tailed distribution) and a lighter tail for $a > 0$ (short-tailed distribution). When $a < 0$, X has no upper limit, i.e., $0 \leq x < \infty$; there is an upper bound for $a > 0$, i.e., $0 \leq x \leq b/a$. This property makes GP2 distribution suitable for analysis of independent cluster peaks.

(2) In the context of the partial duration series, a truncated GP2 distribution remains a GP2 distribution with the original shape parameter a remaining unchanged. This property is popularly referred to as 'threshold stability' property (Smith, 1984, 1987). Consequently, if X has a GP distribution for a fixed threshold level Q_0, then the conditional distribution of $X - c$, given $x \geq c$, corresponding to a higher threshold $Q_0 + c$ also has a GP distribution. If the lower bound

c in the distribution is unknown then the 3-parameter GP distribution is obtained by replacing x in equations (20.3) and (20.4) by x-c. This is one of the properties that justifies the use of GP distribution to model excesses.

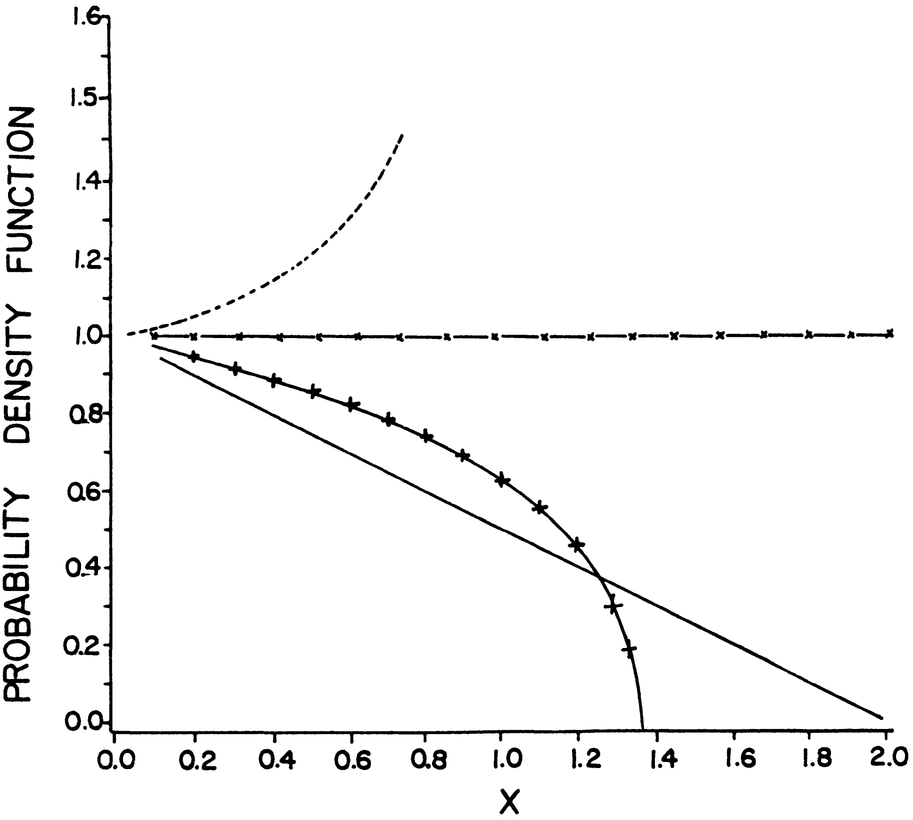

Figure 20.1 (b) Generalized Pareto distribution with b = 1, a = 0.5, 0.75, 1.0, and 1.25; line: a = 0.5, plus: a = 0.75, star: a = 1.0, and dash: a = 1.25.

(3) If the Wakeby distribution is parameterized as suggested by Hosking (1986), then this distribution can be considered as a generalization of the GP distribution.

(4) Let $Z = \max(0, X_1, X_2, ..., X_N)$, $N > 0$ is a number. If X_i, $i = 1, 2, ..., N$, are independent and identically distributed as GP distribution, and N has poisson distribution, then Z has a generalized extreme value (GEV) distribution (Smith, 1984; Jin and Stedinger, 1989; Wang, 1990), as defined by Jenkinson (1955). Thus, a Poisson process of exceedance times with generalized pareto excesses implies the classical extreme value distributions. As a special case, the maximum of a Poisson number of exponential variates has a Gumbel distribution. So exponential peaks lead to Gumbel maxima, and GP2 distribution peaks lead to GEV maxima. The GEV distribution, discussed in Chapter 11, can be expressed as

$$F(z) = \exp\left[-\left(1 - \delta \frac{z-\gamma}{\beta}\right)^{1/\delta}\right], \quad \delta \neq 0, z \geq 0 \tag{20.5a}$$

$$= \exp\left[-\exp\left(-\frac{z-\gamma}{\beta}\right)\right], \quad \delta = 0 \tag{20.5b}$$

where the parameters δ, β and γ are independent of z. Furthermore, $\delta = a$; that is, the shape parameters of the GEV and GP2 distributions are the same. Note that Z is not allowed to take on negative values, and $P(Z < 0) = 0$ and $P(Z = 0) = \exp(-\lambda)$, and only for $z \geq 0$ the cdf is modeled by the GEV distribution. This property makes the GP2 distribution suitable for modeling flood magnitudes exceeding a fixed threshold.

(5) The properties 2 and 3 characterize the GP2 distribution such that no other family has either property, and make it a practical family for statistical estimation, provided the threshold is assumed sufficiently high.

(6) The failure rate $r(x) = f(x)/\{1 - F(x)\}$ is expressed as $r(x) = 1/[b - ax]$ and is monotonic in X, decreasing if $a < 0$, constant if $a = 0$, and increasing if $a > 0$.

20.1 Ordinary Entropy Method

20.1.1 SPECIFICATION OF CONSTRAINTS B

The entropy of the GP distribution can be expressed using equation (20.3) as

$$I(x) = \ln b \int_0^\infty f(x)\,dx - \left(\frac{1}{a} - 1\right) \int_0^\infty \ln\left[1 - \frac{a}{b}x\right] f(x)\,dx \tag{20.6}$$

The constraints appropriate for equation (20.3) can be written (Singh and Rajagopal, 1986) as

$$\int_0^\infty f(x)dx = 1 \tag{20.7}$$

$$\int_0^\infty \ln\left[1 - a\frac{x}{b}\right] f(x)dx = E\left[\ln\left[1 - a\frac{x}{b}\right]\right] \tag{20.8}$$

in which E[•] denotes expectation of the bracketed quantity. These constraints are unique and specify the information sufficient for the GP2 distribution. The first constraint specifies the total probability. The second constraint specifies the mean of the logarithm of the inverse ratio of the scale parameter to the failure rate. Conceptually, this defines the expected value of the negative algorithm of the scale failure rate.

20.1.2 CONSTRUCTION OF ZEROTH LAGRANGE MULTIPLIER

The least-biased pdf of the GP2 distribution corresponding to POME and consistent with equations (20.7) and (20.8) takes the form:

$$f(x) = \exp\left[-a_0 - a_1 \ln\left(1 - a\frac{x}{b}\right)\right] \tag{20.9}$$

where a_0 and a_1 are Lagrange multipliers. The mathematical rationale for equation (20.9) has been presented by Tribus (1969). By applying equation (20.9) to the total probability condition in equation (20.7), one obtains

$$\exp(a_0) = \int_0^\infty \exp(-a_1 \ln[1 - a\frac{x}{b}])\,dx \tag{20.10}$$

which yields the partition function:

$$\exp(a_0) = \frac{b}{a}\frac{1}{1-a_1} \tag{20.11}$$

Taking logarithm of equation (20.11) yields the zeroth Lagrange multiplier given as

$$a_0 = \ln[\frac{b}{a}\frac{1}{1-a_1}] = \ln b - \ln a - \ln(1-a_1)] \tag{20.12}$$

The zeroth Lagrange multiplier is also obtained from equation (20.10) as

$$a_0 = \ln \int_0^\infty \exp(-a_1 \ln[1 - a\frac{x}{b}])\,dx \tag{20.13}$$

20.1.3 RELATION BETWEEN LAGRANGE MULTIPLIERS AND CONSTRAINTS

Differentiating equation (20.13) with respect to a_1:

$$\frac{\partial a_0}{\partial a_1} = -\frac{\int_0^\infty \exp\{-a_1 \ln[1-a(x/b)]\} \ln[1-a(x/b)]\,dx}{\int_0^\infty \exp[-a_0 \ln\{1-a(x/b)\}]dx}$$

$$= -\int_0^\infty \exp\{-a_0 - a_1 \ln[1-a(x/b)]\} \ln[1-a(x/b)]\,dx$$

$$= -E\{\ln[1-a(x/b)]\} \tag{20.14}$$

Similarly, differentiating equation (20.11) with respect to a_1 yields

$$\frac{\partial a_0}{\partial a_1} = \frac{1}{1-a_1} \tag{20.15}$$

Equating equations (20.14) and (20.15) gives

$$\frac{1}{1-a_1} = -E[\ln(1-\frac{ax}{b})] \tag{20.16}$$

The GP2 distribution has two parameters so two equations are needed. The second equation is obtained by noting that

$$\frac{\partial^2 a_0}{\partial a_1^2} = Var[\ln(1 - \frac{a\,x}{b})] \tag{20.17}$$

Differentiating equation (20.15), we get

$$\frac{\partial^2 a_0}{\partial a_1^2} = \frac{1}{(1 - a_1)^2} \tag{20.18}$$

Equating equations (20.17) and (20.18), we obtain

$$\frac{1}{(1 - a_1)^2} = Var[\ln(1 - \frac{ax}{b})] \tag{20.19}$$

20.1.4 RELATION BETWEEN PARAMETERS AND LAGRANGE MULTIPLIERS

Inserting equation (20.11) in equation (20.9) yields

$$f(x) = \frac{a(1 - a_1)}{b} (1 - a \frac{x}{b})^{a_1} \tag{20.20}$$

A comparison of equation (20.20) with equation (20.3) yields

$$1 - a_1 = \frac{1}{a} \tag{20.21}$$

20.1.5 RELATION BETWEEN PARAMETERS AND CONSTRAINTS

The GP2 distribution has two parameters, a and b, which are related to the Lagrange multipliers by equation (20.21) which, in turn, are related to the constraints by equations (20.16) and (20.18). Eliminating the Lagrange multipliers between these two sets of equations yields relations between parameters and constraints. Therefore, the parameter estimation equations are

$$a = - E[\ln(1 - \frac{a\,x}{b})] \tag{20.22}$$

$$a^2 = Var[\ln(1 - \frac{ax}{b})] \tag{20.23}$$

20.1.6 DISTRIBUTION ENTROPY

The entropy of the GP2 distribution is given by

$$I(f) = \ln b - \frac{(1-a)}{a} E[1 - \frac{a}{b} x] \tag{20.24}$$

20.2 Parameter-Space Expansion Method

20.2.1 SPECIFICATION OF CONSTRAINTS

Following Singh and Rajagopal (1986), the constraints for this method turn out to be the same as for the ordinary entropy method and are given by equations (20.7) and (20.8).

20.2.2 DERIVATION OF ENTROPY FUNCTION

Taking the natural logarithm of equation (20.20), we get

$$\ln f(x) = \ln a + \ln (1 - a_1) - \ln b - a_1 \ln [1 - a \frac{x}{b}] \tag{20.25a}$$

Therefore, the entropy I(f) of the GP2 distribution follows:

$$I(f) = - \ln a - \ln (1 - a_1) + \ln b + a_1 E \{\ln [1 - a \frac{x}{b}]\} \tag{20.25b}$$

20.2.3 RELATION BETWEEN PARAMETERS AND CONSTRAINTS

According to Singh and Rajagopal (1986), the relation between distribution parameters and constraints is obtained by taking partial derivatives of the entropy function I(f) with respect to Lagrange multipliers (other than zeroth) as well as distribution parameters, and then equating these derivatives to zero, and making use of the constraints. To that end, taking partial derivatives of equation (20.25) with respect to a_1, a, and b and equating each derivative to zero yields

$$\frac{\partial I}{\partial a_1} = \frac{1}{1 - a_1} + E\{\ln[1 - a \frac{x}{b}]\} = 0 \tag{20.26}$$

$$\frac{\partial I}{\partial a} = - \frac{1}{a} - a_1 E [\frac{x/b}{1 - a(x/b)}] = 0 \tag{20.27}$$

$$\frac{\partial I}{\partial b} = \frac{1}{b} + a_1 - E [\frac{x/b}{1 - a(x/b)}] = 0 \tag{20.28}$$

Simplification of equations (20.26) to (20.28) yields, respectively,

$$E \{\ln [1 - a \frac{x}{b}]\} = - \frac{1}{1 - a_1} \tag{20.29}$$

$$E\left[\frac{x/b}{1 - a\,(x/b)}\right] = -\frac{1}{a a_1} \tag{20.30}$$

$$E\left[\frac{x/b}{1 - a\,(x/b)}\right] = -\frac{1}{a a_1} \tag{20.31}$$

Clearly, equation (20.30) is the same as equation (20.31). With a_1 expressed by equation (20.21), equations (20.29) and (20.30) are the POME-based estimation equations

20.3 Other Methods of Parameter Estimation

Three other popular methods of parameter estimation are briefly outlined: the method of moments (MOM), the method of probability-weighted moments (PWM), and the method of maximum likelihood estimation (MLE).

20.3.1 METHOD OF MOMENTS

For the method of moments (MOM) the moment estimators of the GP2 distribution were derived by Hosking and Wallis (1987). Note that $E(1 - a\,(x/b)^r) = 1/(1 + ar)$ if $1 + ra > 0$. The rth moment of X exists if $a > -1/r$. Provided that they exist, then the moment estimators are

$$a = \frac{1}{2}\left(\frac{\bar{x}^2}{s^2} - 1\right) \tag{20.32}$$

$$b = \frac{1}{2}\,\bar{x}\left(\frac{\bar{x}^2}{s^2} + 1\right) \tag{20.33}$$

where $\bar{x}$ and s^2 are the mean and variance, respectively.

20.3.2 PROBABILITY WEIGHTED MOMENTS

For the method of probability weighted moments (PWM), the PWM estimators of the GP2 distribution were given by Hosking and Wallis (1987) as

$$a = \frac{W_0 - 4W_1}{2W_1 - W_0} \tag{20.34}$$

$$b = \frac{W_0\,W_1}{2W_1 - W_0} \tag{20.35}$$

where

$$W_r = E\{x(F)\,[1 - F(x)]^2\} = \frac{b}{a}\left[\frac{1}{r+1} - \frac{1}{a+r+1}\right],\ r = 0, 1, \ldots \tag{20.36}$$

20.3.3 METHOD OF MAXIMUM LIKELIHOOD ESTIMATION

The method of maximum likelihood estimation (MLE) estimators can be expressed as

$$\sum_{i=1}^{n} \frac{x_i}{1 - a\, x_i/b} = \frac{n}{1-a} \tag{20.37}$$

$$\frac{1}{a}\sum_{i=1}^{n} \ln\left[1 - a\, x_i/b\right] + (1-a)\sum_{i=1}^{n} \frac{x_i}{b - a x_i} = 0 \tag{20.38}$$

A successive iterative procedure based on the Golden search method was adopted to obtain the estimates of a and b. First, an initial value of a was assumed. Then, with the use of the Golden search method, the optimal value of b leading to the maximum of log-likelihood function (log L) was obtained. Then by fixing b, parameter a was re-established leading to the maximum log L. This procedure was continued until parameters a and b no longer significantly changed. As an example, for population parameters a = -0.444 and b = 0.556, the MLE estimates were found by this method for a sample size of 10.

20.4 Comparative Evaluation of Parameter Estimation Methods

20.4.1 MONTE CARLO SAMPLES

Guo and Singh (1992) and Singh and Guo (1997) assessed the performance of the POME, MOM, PWM and MLE methods using Monte Carlo sampling experiments. They considered three population cases as listed in Table 20.1. For each population case, 1000 random samples of size 20, 50, and 100 were generated, and then parameters and quantiles were estimated. The work of Guo and Singh is briefly summarized here.

Table 20.1. GP2 population cases considered in sampling experiment [mean, μ = 1; Cv = coefficient of variation]

GPD2 Population	Cv	Parameter	
		a	b
Case 1	1.5	-0.278	0.722
Case 2	2.0	-0.375	0.625
Case 3	3.0	-0.444	0.556

20.4.2 PERFORMANCE INDICES

The performance of estimation methods was evaluated using the standardized bias (BIAS) and root mean square error (RMSE) for both parameter and quantile estimates. It should be noted that the number (n) of Monte Carlo samples of 1,000 may arguably not be large enough to produce the true values of BIAS and RMSE, but will suffice to compare the performance of estimation methods.

20.4.3 BIAS IN PARAMETER ESTIMATES

The bias values of parameters estimated by the four methods showed that the parameter bias varied with sample size and Cv. In general, the parameter bias decreased with increasing sample size, but its variation with Cv was not consistent. For all sample sizes, MOM always produced the highest parameter bias in both parameters a and b, whereas POME, PWM, and MLE produced quite low parameter bias, especially for sample size $n \geq 50$. For $Cv \geq 1.5$ and $n \geq 50$, POME produced the least bias, followed by PWM and MLE. When $Cv \geq 2.0$ and $n \leq 20$, PWM produced the least bias, but the bias by MLE and POME was quite low. Thus, POME would be the preferred method for $Cv \geq 1.5$.

20.4.4 RMSE IN PARAMETER ESTIMATES

The values of RMSE of parameters estimated by the four methods showed that in general, RMSE decreased with increasing sample size for a specified Cv. However, for a given sample size, the variation of RMSE with Cv followed a decreasing trend for parameter a but did not follow a consistent pattern for parameter b. In general, PWM had the lowest RMSE and MOM the highest RMSE for both parameters a and b, but the differences between the RMSE value of PWM and those of POME and MLE, and even MOM (except for small sample sizes $n \leq 20$) were quite small. For $n \geq 50$, and $Cv \geq 1.5$, all four methods were comparable. Thus, for samples with $Cv \geq 1.5$, either of POME, MLE and PWM could be chosen.

20.4.5 BIAS IN QUANTILE ESTIMATES

The results of bias in quantile estimates for GP2 distribution showed that the bias of a given method, in general, varied with the sample size, the value of Cv, and the probability of non-exceedance (P). For $P \leq 0.90$ and $n \leq 20$, the bias by POME was either less than or comparable to that of PWM and MLE. For $Cv \geq 2.0$ and $n \geq 50$, MLE produced the least bias, and POME and PWM were comparable. For $0.99 \leq P \leq 0.999$ and $n \geq 50$, POME and PWM were comparable. When n was ≤ 20, POME as well as MLE did not perform well, especially for $Cv \geq 3.0$. Thus, for large sample sizes, anyone of the three methods--POME, PWM and MLE--could be used. However, the required sample size would be much larger for larger Cv. For small samples and high Cv, PWM would be the preferred method.

20.4.6 RMSE IN QUANTILE ESTIMATES

The values of RMSE in quantile estimates of the four methods are given in Table 20.5. RMSE of a given method significantly varied with sample size (n), the probability of nonexceedance (P), and the coefficient of variation (Cv). In general, POME and MLE produced the highest RMSE, especially for $Cv \geq 1.5$, and MOM the least RMSE across all sample sizes and the range of Cv. Thus, MOM would be the preferred method, and PWM the second preferred method. For $P \leq 0.8$, all methods were more or less comparable for all sample sizes and the range of Cv.

20.4.7 CONCLUDING REMARKS

To summarize, in terms of parameter bias, POME is comparable to PWM and MLE. POME was comparable to MLE or PWM in terms of RMSE of parameter estimates. For $P \leq 0.9$ and $n \leq 20$, POME produced the least bias in quantiles estimates. For $0.9 \leq P \leq 0.99$ and $n \geq 50$, POME and PWM were comparable. (5) For $P \leq 0.99$, the RMSE of POME was comparable to that of PWM and MLE but higher than that of MOM.

References

Ashkar, F. and Ouarda, T.B.M.J., 1996. On some methods of fitting the generalized Pareto distribution. *Journal of Hydrology*, Vol. **177**, pp. 117-141.

Davison, A. C. and Smith, R. L., 1990. Models for exceedances over high thresholds. *J. R. Statist. Soc.* B, Vol. **52**, No. 3, pp. 393-442.

Guo, H. and Singh, V.P., 1992. A comparative evaluation of estimators of Pareto distribution by Monte Carlo simulation. *Technical Report* **WRR25**, Water Resources Program, Department of Civil and Environmental Engineering, Louisiana State University, Baton Rouge, Louisiana.

Hosking, J. R. M., 1986. The theory of probability-weighted moments. *Research Report* **RC12210**, IBM, Yorktown Heights, NY.

Hosking, J. R. M. and Wallis, J. R., 1987. Parameter and quantile estimation for the generalized Pareto distribution. *Technometrics*, Vol. **29**, No. 3, pp. 339-349.

Jenkinson, A. F., 1955. The frequency of distrubution of the annual maximum (or minimum) of meteorological elements. *Quarterly Journal of Royal Meteorological Society*, Vol. **81**, pp. 158-171.

Jin, M. and Stendinger, J.R., 1989. Partial duration series analysis for a GEV annual flood distribution with systematic and historical flood information. *Unpublished Report*, Department of Civil Engineering, Pennsylvania State University, State College, Pennsylvania, 28 pp.

Pickands, J., 1975. Statistical inference using extreme order statistics. *Ann. Statist.*, Vol. **3**, pp. 119-131.

Singh, V.P. and Guo, H., 1997. Parameter estimation for 2-parameter generalized Pareto distribution by POME. *Stochastic Hydrology and Hydraulics*, Vol. **11**, pp. 211-227.

Singh, V. P. and Rajagopal, A. K., 1986. A new method of parameter estimation for hydrologic frequency analysis. *Hydrological Science and Technology*, Vol. **2,** No. 3, pp. 33-40.

Smith, R. L., 1984. Threshold methods for sample extremes. in: *Statistical Extremes and Applications*, edited by J. Trago de Oliveira, pp. 621-638, Reidel Publishing Company, Dordrecht, the Netherlands.

Smith, R. L., 1987. Estimating tails of probability distributions. *Ann. Statist.*, Vol. **15**, pp. 1174-1207.

van Montfort, M. A. J. and Otten, A., 1991. The first and the second e of the extreme value distribution, EV1. *Stochastic Hydrology and Hydraulics*, Vol. **5**, pp. 69-76.

Wang, Q. J., 1990. Studies on statistical methods of flood frequency analysis. *Unpublished Ph.D. dissertation*, 110 pp., National University of Ireland, Galway, Ireland.

Wang, Q. J., 1991. The POT model described by the generalized Pareto distribution with Poisson arrival rate. *Journal of Hydrology*, Vol. **129**, pp. 263-280.

CHAPTER 21

THREE-PARAMETER GENERALIZED PARETO DISTRIBUTION

The Pareto distribution has been introduced in Chapter 19. Also discussed there is a brief review of literature and methods of estimating its parameters. Further elaboration of the distribution is given in Chapter 20. Methods of parameter estimation were reviewed by Hosking and Wallis (1987). The methods of moments (MOM), maximum likelihood estimation (MLE) and probability weighted moments (PWM) were included in the review. Guo and Singh (1992) and Singh and Guo (1995) employed the principle of maximum entropy (POME) to develop a new competitive method of parameter estimation (Singh and Rajagopal, 1986) for the 3-parameter generalized Pareto (GP3) distribution and compared it with MOM, MLE and PWM using Monte Carlo simulated data. The parameter estimates yielded by POME were either superior or comparable for high skewness.

Consider a random variable Y with the standard exponential distribution. Let a random variable X be defined as $X = b(1 - \exp(-aY))/a$, where a and b are parameters. Then the distribution of X is the 2-parameter generalized Pareto distribution. If c is a threshold or lower bound of X, then the distribution of X is the 3-parameter generalized Pareto (GP) distribution which can be expressed as

$$F(x) = 1 - \left(1 - a\frac{x-c}{b}\right)^{1/a}, \qquad a \neq 0 \tag{21.1}$$

$$= 1 - \exp\left(-\frac{x-c}{b}\right), \qquad a = 0 \tag{21.2}$$

where c is a location parameter, b is a scale parameter, a is a shape parameter, and F(x) is the distribution function. The probability density function (pdf) of the GP distribution follows:

$$f(x) = \frac{1}{b}\left(1 - a\frac{x-c}{b}\right)^{1/a-1}, \qquad a \neq 0 \tag{21.3}$$

$$= \frac{1}{b}\exp\left(-\frac{x-c}{b}\right), \qquad a = 0 \tag{21.4}$$

Some of the properties of the Pareto distribution are discussed in the preceding chapter, apply to the GP3 distribution, and will therefore not be repeated here.

21.1 Ordinary Entropy Method

21.1.1 SPECIFICATION OF CONSTRAINTS

The entropy of the GP3 distribution can be expressed as

$$I(f) = \ln b \int_c^\infty f(x)\,dx + (\frac{1}{a} - 1) \int_c^\infty \ln [1 - \frac{a(x-c)}{b}]\, f(x)\,dx \tag{21.5}$$

From equation (21.5), the constraints appropriate for equation (21.3) can be written (Singh and Rajagopal, 1986) as

$$\int_c^\infty f(x)dx = 1 \tag{21.6}$$

$$\int_c^\infty \ln [1 - a\frac{x-c}{b}]\, f(x)dx = E[\ln [1 - a\frac{x-c}{b}]] \tag{21.7}$$

in which E[•] denotes expectation of the bracketed quantity. The first constraint specifies the total probability. The second constraint specifies the mean of the logarithm of the inverse ratio of the scale parameter to the failure rate. Conceptually, this defines the expected value of the negative logarithm of the scaled failure rate.

21.1.2 CONSTRUCTION OF ZEROTH LAGRANGE MULTIPLIER

The least-biased pdf of the GP3 distribution corresponding to POME and consistent with equations (21.6) and (21.7) takes the form:

$$f(x) = \exp [- a_0 - a_1 \ln (1 - a\frac{x-c}{b})] \tag{21.8}$$

where a_0 and a_1 are Lagrange multipliers. Applying equation (21.8) to the total probability condition in equation (21.6), one obtains

$$\exp (a_0) = \int_c^\infty \exp (- a_1 \ln [1 - a\frac{x-c}{b}])\, dx \tag{21.9}$$

which yields the partition function:

$$\exp (a_0) = \frac{b}{a}\frac{1}{1-a_1} \tag{21.10}$$

Taking logarithm of equation (21.10), we get the zeroth Lagrange multiplier as

$$a_0 = \ln [\frac{b}{a}\frac{1}{1-a_1}] \tag{21.11}$$

Equation (21.11) is recast as

$$a_0 = \ln (b/a) - \ln (1 - a_1) \tag{21.12}$$

The zeroth Lagrange multuplier is also obtained by taking logarithm of equation (21.9) as

$$a_0 = \ln \int_c^\infty \exp (- a_1 \ln [1 - a \frac{x-c}{b}])\, dx \tag{21.13}$$

21.1.3 RELATION BETWEEN LAGRANGE MULTUPLIERS AND CONSTRAINTS

Differentiating equation (21.13) with respect to a_1:

$$\frac{\partial a_0}{\partial a_1} = - \frac{\int_c^\infty \exp \{- a_1 \ln[1 - a(x-c)/b]\} \ln[1 - a(x-c)/b]\, dx}{\int_c^\infty \exp[- a_0 \ln\{1 - a(x-c)/b\}] dx}$$

$$= - \int_c^\infty \exp \{- a_0 - a_1 \ln[1 - a(x-c)/b]\} \ln[1 - a(x-c)/b]\, dx$$

$$= - E\{[1 - a(x-c)/b]\} \tag{21.14}$$

Following Tribus (1969), we can write

$$\frac{\partial^2 a_0}{\partial a_1^2} = \mathrm{Var}\, \{\ln [1 - a(x-c)/b]\} \tag{21.15}$$

where Var[•] is the variance of the bracketed quantity.

Differentiating equation (21.12) with respect to a_1 once and then again results in:

$$\frac{\partial a_0}{\partial a_1} = \frac{1}{1 - a_1} \tag{21.16}$$

$$\frac{\partial^2 a_0}{\partial a_1^2} = \frac{1}{(1 - a_1)^2} \tag{21.17}$$

Equating equation (21.16) to equation (21.14) leads to:

$$E [\ln (1 - a \frac{x-c}{b})] = - \frac{1}{1 - a_1} \tag{21.18}$$

When equation (21.17) is equated to equation (21.15), the following is obtained:

$$\mathrm{Var} [\ln (1 - a \frac{x-c}{b})] = \frac{1}{(1 - a_1)^2} \tag{21.19}$$

21.1.4 RELATION BETWEEN PARAMETERS AND LAGRANGE MULTIPLIERS

Inserting equation (21.10) into equation (21.8), we get

$$f(x) = \frac{a(1 - a_1)}{b} \left(1 - a \frac{x - c}{b}\right)^{-a_1} \tag{21.20}$$

A comparison of equation (21.20) with equation (21.3) yields

$$1 - a_1 = \frac{1}{a} \tag{21.21}$$

21.1.5 RELATION BETWEEN PARAMETERS AND CONSTRAINTS

The GP3 distribution has three parameters which are related to the Lagrange multipliers by equation (21.21) which, in turn, are related to the constraints by equations (21.18)-(21.19). Eliminating the Lagrange multipliers between these two sets of equations yields relations between parameters and constraints. Therefore,

$$E\left[\frac{1}{1 - a(x - c)/b}\right] = \frac{1}{1 - a} \tag{21.22}$$

$$\mathrm{Var}\left[\ln\left(1 - a \frac{x - c}{b}\right)\right] = a^2 \tag{21.23}$$

The GP3 distribution has three parameters; therefore, three equations are needed for estimation of its parameters. This means that equations (21.22) and (21.23) need to be supplemented. This is accomplished by setting parameter c as the lowest value of the observations in the sample.

21.1.6 DISTRIBUTION ENTROPY

The entropy of the GP3 distribution is given as

$$I(f) = \ln b + \left(\frac{1}{a} - 1\right) E\left\{\ln\left[1 - \frac{a(x-c)}{b}\right]\right\} \tag{21.24}$$

21.2 Parameter-Space Expansion Method

21.2.1 SPECIFICATION OF CONSTRAINTS

Following Singh and Rajagopal (1986), the constraints for this method are given by equations (21.6) and

$$\int_c^\infty \left(\frac{1}{a} - 1\right) \ln\left[1 - \frac{a(x-c)}{b}\right] f(x)\, dx = E\left[\left(\frac{1}{a} - 1\right) \ln\left\{1 - \frac{a(x-c)}{b}\right\}\right] dx \tag{21.25}$$

The pdf corresponding to POME and consistent with equations (21.6) and (21.25) takes the form:

$$f(x)=\exp\{-a_0-a_1(\frac{1}{a}-1)\ln[1-\frac{a(x-c)}{b}]\} \tag{21. 26}$$

Inserting equation (21.26) into equation (21.6) yields the partition function:

$$\exp(a_0)=\int_c^{\infty}\exp\{-a_1(\frac{1}{a}-1)\ln[1-\frac{a(x-c)}{b}]\}\,dx \tag{21.27}$$

Equation (21.27) simplifies to

$$\exp(a_0)=\frac{b}{a[1-a_1(\frac{1}{a}-1)]} \tag{21.28}$$

Taking logarithm of equation (21.28) gives the zeroth Lagrange multiplier:

$$a_0=\ln b-\ln a\ -\ln[1-a_1(\frac{1}{a}-1)] \tag{21.29}$$

Introduction of equation (21.29) in equation (21.26) gives the POME-based pdf:

$$f(x)=\frac{a}{b}\{1-a_1[\frac{1}{a}-1]\}\,[1-\frac{a(x-c)}{b}]^{-a_1(\frac{1}{a}-1)} \tag{21.30}$$

A comparison of equation (21.30) with equation (21.3) shows that a_1 =-1.

Taking logarithm to the base 'e' of equation (21.30) and multiplying it minus one by give

$$-\ln f(x)=-\ln a+\ln b-\ln[1-a_1(\frac{1}{a}-1)]+a_1(\frac{1}{a}-1)\ln[1-\frac{a(x-c)}{b}] \tag{21.31}$$

Therefore, the entropy function I(f) of the GP3 distribution follows:

$$I(f)=-\ln a+\ln b-\ln[1-a_1(\frac{1}{a}-1)]+a_1(\frac{1}{a}-1)E[\ln\{1-\frac{a(x-c)}{b}\}] \tag{21.32}$$

21.2.2 RELATION BETWEEN PARAMETERS AND CONSTRAINTS

According to Singh and Rajagopal (1986), the relation between distribution parameters and constraints is obtained by taking partial derivatives of the entropy I(f) with respect to Lagrange multipliers as well as distribution parameters, and then equating these derivatives to zero, and making use of the constraints. To that end, taking partial derivatives of equation (20.32) with respect to a_1, a, b and c separately and equating each derivative to zero yields

$$\frac{\partial I}{\partial a_1} = \frac{\frac{1}{a}-1}{1-a_1(\frac{1}{a}-1)} + (\frac{1}{a}-1)E\{\ln[1 - a\frac{x-c}{b}]\} = 0 \tag{21.33}$$

$$\frac{\partial I}{\partial a} = -\frac{1}{a} - \frac{a_1}{a^2}\frac{1}{[1-a_1(\frac{1}{a}-1)]} - a_1(\frac{1}{a}-1)\,E\,[\frac{(x-c)/b}{1-a(x-c)/b}] - \frac{a_1}{a^2}E[\ln(1-\frac{a(x-c)}{b})] = 0 \tag{21.34}$$

$$\frac{\partial I}{\partial b} = \frac{1}{b} + a_1(\frac{1}{a}-1)\,E\,[\frac{a(x-c)/b^2}{1-a(x-c)/b}] = 0 \tag{21.35}$$

$$\frac{\partial I}{\partial c} = a_1(\frac{1}{a}-1)\,E\,[\frac{a/b}{1-a(x-c)/b}] = 0 \tag{21.36}$$

Simplification of equations (21.33) to (21.36) yields, respectively,

$$E\{\ln[1 - a\frac{x-c}{b}]\} = -\frac{1}{1-a_1(\frac{1}{a}-1)} \tag{21.37}$$

$$E\,[\frac{(x-c)/b}{1-a(x-c)/b}] + \frac{1}{(1-a)a}E[\ln(1-\frac{a(x-c)}{b})] = \frac{1}{(a-1)a_1} + \frac{1}{a(1-a)}\frac{1}{1-a_1(\frac{1}{a}-1)} \tag{21.38}$$

$$E\,[\frac{(x-c)/b}{1-a(x-c)/b}] = -\frac{1}{aa_1(\frac{1}{a}-1)} \tag{21.39}$$

$$E\,[\frac{1}{1-a(x-c)/b}] = 0 \tag{21.40}$$

Clearly, equation (21.40) does not hold. Recall that $a_1 = -1$.Inserting $a_1 = 1 - 1/a$ from equation (21.21) into these three equations, one gets

$$E[\ln(1 - a\frac{x-c}{b})] = -a \tag{21.41}$$

From equation (21.38) we obtain

$$E[\frac{(x-c)/b}{1-a(x-c)/b}] + \frac{1}{a(1-a)}E[\ln\{1-\frac{a(x-c)}{b}\}] = 0 \tag{21.42}$$

Taking advantage of equation (21.39), equation (21.42) simplifies to

$$E\left[\frac{1}{1-a(x-c)/b}\right]=\frac{1}{1-a} \qquad (21.43)$$

Thus, the estimation equations are equations (21.41) and (21.43). In order to get a unique solution, an additional equation is needed and this is obtained in the same manners as in the ordinary entropy method. This means that equation (21.23) will hold in this case too. Therefore, the parameter estimation equations by POME consist of equations (21.41), (21.43) and (21.23).

21.3 Other Methods of Parameter Estimation

Three other popular methods of parameter estimation are briefly outlined: the method of moments (MOM), the method of probability-weighted moments (PWM), and the method of maximum likelihood estimation (MLE).

21.3.1 METHOD OF MOMENTS

For the method of moments (MOM), the moment estimators of the GP3 distribution were derived by Hosking and Wallis (1987). Note that $E(1 - a (x-c)/b)^r = 1/(1 + ar)$ if $1 + ra > 0$. The rth moment of X exists if $a > -1/r$. Provided that they exist, then the moment estimators are

$$\bar{x} = c + \frac{b}{1+a} \qquad (21.44)$$

$$S^2 = \frac{b^2}{(1+a)^2\,(1+2a)} \qquad (21.45)$$

$$G = \frac{2(1-a)\,(1+2a)^{0.5}}{1+3a} \qquad (21.46)$$

where $\bar{x}$, S^2 and G are the mean, variance and skewness, respectively. First, the moment estimate of a is obtained by solving equation (21.46). The relation between G and a is graphed in Figure 21.1. With a calculated, b and c follow from equation (21.44) and (21.45) as

$$b = S\,(1 + a)\,(1 + 2a)^{0.5} \qquad (21.47)$$

$$c = \bar{x} - \frac{b}{b+a} \qquad (21.48)$$

21.3.2 METHOD OF PROBABILITY-WEIGHTED MOMENTS

For the method of probability-weighted moments (PWM), the PWM estimators for the GP3 distribution (Hosking and Wallis, 1987) are given as

$$a = \frac{W_0 - 8W_1 + 9W_2}{-W_0 + 4W_1 - 3W_2} \qquad (21.49)$$

$$b = \frac{(W_0 - 2W_1)(W_0 - 3W_2)(-4W_1 + 6W_2)}{(-W_0 + 4W_1 - 3W_2)^2} \tag{21.50}$$

$$c = \frac{2W_0W_1 - 6W_0W_2 + 6W_1W_2}{-W_0 + 4W_1 - 3W_2} \tag{21.51}$$

where the r-th probability-weighted moment W_r is

$$W_r = E[x(F)(1 - F(x))^r] = \int_0^1 \{c + \frac{b}{a}[1 - (1 - F)^a]\}(1 - F)^r dF$$

$$= \frac{1}{r+1}(c + \frac{b}{a}) - \frac{b}{a}\frac{1}{a+r+1}, \quad r = 0, 1, 2, \ldots \tag{21.52}$$

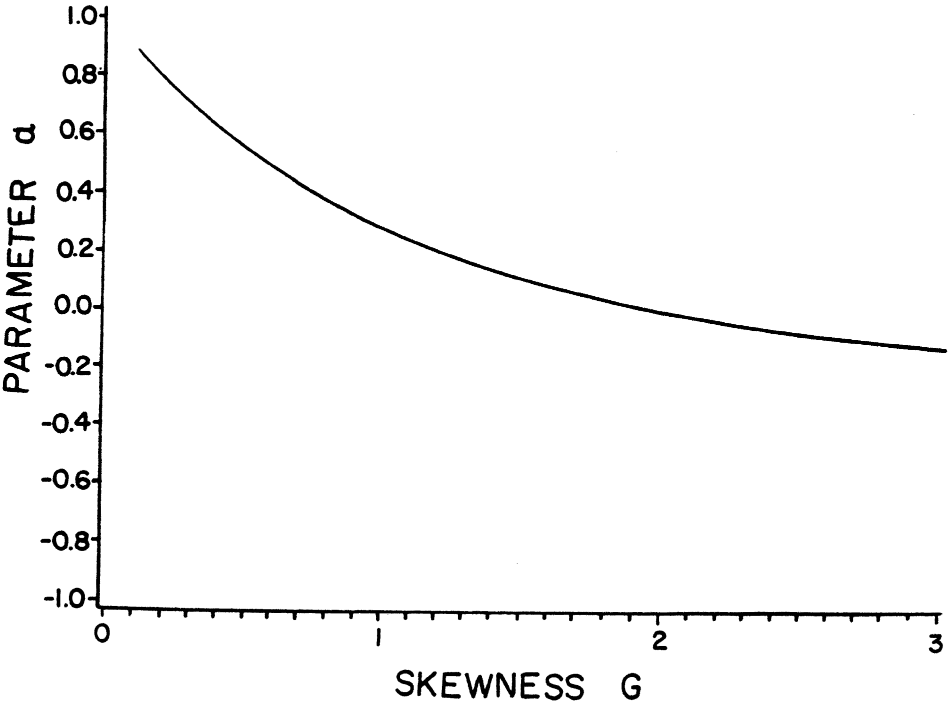

Figure 21.1. Parameter a versus skewness G for GPD3.

21.3.3 METHOD OF MAXIMUM LIKELIHOOD ESTIMATION

For the method of maximum likelihood estimation (MLE), the MLE estimators can be expressed as

$$\sum_{i=1}^{n} \frac{(x_i - c)/b}{1 - a(x_i - c)/b} = \frac{n}{1-a} \tag{21.53}$$

$$\sum_{i=1}^{n} \ln [1 - a (x_i - c)/b] = - na \tag{21.54}$$

A maximum likelihood estimator cannot be obtained for c, because the likelihood function is unbounded with respect to c, as shown in Figure 21.2. Since c is the lower bound of the random variable X, we may use the constraint $c \le x_1$, the lowest sample value. Clearly, the likelihood function is maximum with respect to c when $c = x_1$.

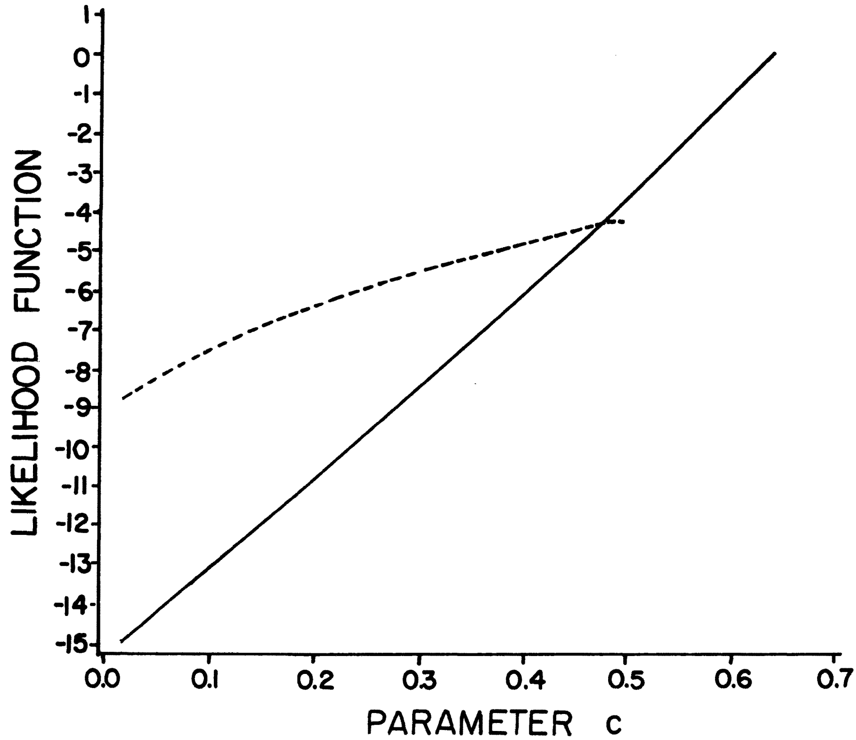

Figure 21.2. Likelihood function of GPD3 versus parameter c for sample size 10.
Line: a = - 0.116, b = 0.387, c = 0.562
Dash: a = 0.544, b = 1.116, c = 0.277

21.4 Comparative Evaluation of Parameter Estimation Methods

21.4.1 MONTE CARLO SAMPLES

Guo and Singh (1992) and Singh and Guo (1997) assessed the performance of the POME, MOM, PWM and MLE estimation methods using Monte Carlo sampling experiments. They considered two distribution population cases listed in Table 21.1. For each population case, 1000 random samples of size 20, 50, and 100 were generated, and then parameters and quantiles were estimated. Their work is summarized here.

21.4.2 PERFORMANCE INDICES

The performance of the parameter estimation methods was evaluated using the performance indices of standardized bias (BIAS) and root mean square error (RMSE). The number of samples (n) of 1,000 may arguably not be large enough to produce the true values of BIAS and RMSE, but will suffice to compare the performance of estimation methods.

Table 21.1. GP distribution population cases considered in the sampling experiment [Mean = 1, Cv = coefficient of variation, and G = coefficient of skewness].

GP Distribution Population	Cv	G	Parameters		
			a	b	c
Case 1	0.5	0.5	0.554	1.116	0.277
Case 2	0.5	2.5	-0.069	0.433	0.536

21.4.3 BIAS IN PARAMETER ESTIMATION

The bias values of parameters estimated by the four methods showed that for G = 0.5, in absolute terms MOM produced the least bias of the four methods for all sample sizes. MLE had the second least bias in parameter estimates. With increasing sample size, there was significant reduction in bias for all four methods. POME produced less bias in estimates of b and c for all sample sizes than PWM, but that was not uniformly true in case of the parameter a estimate. When G = 2.5, these methods performed quite differently. For all samples sizes, MLE and POME were comparable, producing the least bias. For the parameter a and c estimates, POME had the least bias, but MLE had the least bias for the parameter b estimate. PWM had the highest bias in all three parameter estimates for all sample sizes. Thus, if the value of G is high, POME or MLE may be the preferred method. For lower values of G, MOM or MLE may be preferable, especially when the sample size is small.

21.4.4 RMSE IN PARAMETER ESTIMATION

The values of RSME of parameters estimated by the four methods showed that for G = 0.5, of the four methods MOM produced the least RMSE in the parameter a estimate. However, as the sample size increased, MOM, PWM and MLE became comparable. In case of the parameters b and c estimates MLE had the least RMSE, but all four methods were comparable. For G = 2.5, the comparative behavior of the four methods was markedly different. In absolute terms, MOM and PWM produced the highest RMSE in parameter estimates for all sample sizes, with POME having the least bias in the parameter a estimate but MLE in the parameter b and c estimates. Thus, it may be concluded that for lower values of G, MOM or PWM may be the preferred method, but for higher values of G, MLE or POME is the preferred method.

21.4.5 BIAS IN QUANTILE ESTIMATION

The results of bias in quantile estimates by the GP3 distribution showed that the performance of the four estimation methods varied with the value of G, and probability of non-exceedance P. For G = 0.5, all four methods had comparable bias for $P \leq 0.9$ for all sample sizes. When $P \geq 0.99$, MOM and PWM produced the smallest bias and POME the highest, with MLE in the intermediate range. However, for G = 2.5, POME produced the least bias, especially when P was greater than 0.99. For all sample sizes, all four methods were somewhat comparable. In conclusion, for lower values of G, anyone of the four methods may be used for P < 0.99, but PWM, MOM or MLE may be preferable for P exceeding 0.99. For higher values of G, all four methods were comparable, but for P exceeding 0.99 POME is the preferred method.

21.4.6 RMSE IN QUANTILE ESTIMATION

The values of RMSE in quantile estimates of the four methods showed that for G = 0.5, for $P \leq 0.9$, all four methods produced comparable values of RMSE for all sample sizes; for $P \geq 0.99$, performance of POME deteriorated. When G = 2.5, all methods produced comparable values of RMSE for all sample sizes for $P \leq 0.9$; for $P \geq 0.99$ POME had the least RMSE. Thus, it is inferred that for smaller values of G, MOM, PWM or MLE may be used, but for higher values of G, POME may be the preferred method.

21.4.7 CONCLUDING REMARKS

To summarize, when the skewness was high (G = 2.5), POME yielded superior parameter estimates. For low skewness (G = 0.5), POME was better in parameter estimates than MLE and PWM but worse than MOM. However, for large sample sizes, its performance improved significantly. POME produced either better or comparable quantile estimates as compared with MOM, MLE and PWM for high skewness (G = 2.5). For low skewness (G = 0.5), POME was comparable to MOM, MLE and PWM for lower probabilities of nonexceedance for higher values, MOM or PWM was better than POME.

References

Guo, H. and Singh, V.P., 1992. A comparative evaluation of estimators of Pareto distribution by Monte Carlo simulation. *Technical Report* **WRR25**, Water Resources Program, Department of Civil and Environmental Engineering, Louisiana State University, Baton Rouge, Louisiana.

Hosking, J. R. M. and Wallis, J. R., 1987. Parameter and quantile estimation for the generalized Pareto distribution. *Technometrics*, Vol. **29**, No. 3, pp. 339-349.

Singh, V.P. and Guo, H., 1995. Parameter estimation for 3-parameter generalized Pareto distribution by the principle of maximum entropy (POME). *Hydrological Sciences Journal*, Vol. **40**, No. 2, pp. 165-181.

Singh, V. P. and Rajagopal, A. K., 1986. A new method of parameter estimation for hydrologic frequency analysis. *Hydrological Science and Technology*, Vol. **2**, No. 3, pp. 33-40.

Tribus, M., 1969. *Rational Descriptions, Decisions and Designs*. Pergamon Press, New York, N.Y.

CHAPTER 22

TWO-COMPONENT EXTREME VALUE DISTRIBUTION

It is well known that floods may be generated by different physical mechanisms. For instance, most of the annual flood maxima at a particular site might be the result of a primary mechanism, such as frontal storms. A smaller fraction of the events, however, might be associated with a secondary mechanism, such as rain on snow with frozen soils, that occasionally gives rise to floods larger than those associated with the primary mechanism. In this regard, Rossi et al. (1984) proposed a two-component extreme value distribution. This distribution belongs to the family of distributions of the annual maxima of a compound Poisson process, which forms a theoretical basis for annual flood series analysis. Single-component distribution methods of estimating return periods and probabilities of flood events do not work well when runoff originates from nonhomogeneous sources, i.e., when a mixture random variables is involved. The most important consideration in selecting a distribution for use in flood frequency analysis is the behavior of the right tail of the distribution. It is from the right tail that return periods and probabilities of rare events are determined. The two-component extreme value (TCEV) distribution permits a reasonable interpretation of the physical phenomenon which generates floods and is able to account for most of the characteristics of the real world flood data, important among them being the large variability of the sample skewness coefficient which mostly gives rise to the poor performance of many of the commonly used flood frequency distributions. The two component extreme value (TCEV) distribution has been shown to account for most of the characteristics of the real flood experience. The TCEV distribution also offers a practical approach to regional flood frequency estimation. Theoretical properties of the TCEV distribution have been widely investigated (Rossi, et al., 1984; Beran, et al., 1986; Rossi, et al., 1986; Fiorentino et al., 1987a, b). In his extensive review of a large number of commonly used distributions, Cunnane (1986) concluded that only the two-component extreme value distribution and the Wakeby distribution satisfied the important reproductive criterion--an ideal distribution must reproduce at least as much variability in flood characteristics as is observed in empirical data.

A random variable X is defined to have a two-component extreme value (TCEV) distribution if its probability density function (pdf) is given by

$$f(x) = [\frac{\lambda_1}{\theta_1}\exp(-x/\theta_1) + \frac{\lambda_2}{\theta_2}\exp(-x/\theta_2)]\exp[-\lambda_1\exp(-x/\theta_1) - \lambda_2\exp(-x/\theta_2)];\ x > 0 \tag{22.1}$$

V. P. Singh, *Entropy-Based Parameter Estimation in Hydrology*

$$= \exp(-\lambda_1 - \lambda_2); x = 0 \tag{22.2}$$

where $\lambda_1 > 0, \lambda_2 \geq 0, \theta_2 \geq \theta_1 > 0$ are parameters. Its cumulative density function (cdf) is

$$F(x) = \exp[-\lambda_1 \exp(-x/\theta_1) - \lambda_2 \exp(-x/\theta_2)]; x \geq 0 \tag{22.3}$$

For simplicity, it is assumed that equation (22.1) holds for non-positive values of X. This approximation is reasonable in all practical applications of the TCEV distribution (Rossi, et al., 1984).

The cdf of TCEV given in equation (22.3) has been shown (Versace, et al., 1982; Rossi, et al., 1984) to represent the distribution function of the annual maximum (represented by the random variable X) of a nonnegative random variate Z whose number of occurrences, K, in a year is a random variable when the following hypotheses hold: (1) Z is an independent, identically distributed (iid) random variable with probability density function defined by a mixture of two exponential distributions; (2) K is an iid Poisson distributed variate; and (3) Z and Z are independent of each other. The two components of the distribution of both Z and X are usually referred to as basic component (subscripts of parameters = 1) and outlying component (subscripts of parameters = 2). The basic component distribution generates ordinary floods, and the outlying component distribution exhibits a much greater variability than does the basic one and tends to generate rarer but more severe floods. The four parameters of TCEV distribution characterize the mean number of independent peaks in a year (λ_1 and λ_2) and the mean peak amplitude (θ_1 and θ_2) of the basic and outlying components. The outlier distribution is characterized by a mean number of events λ_2 much smaller than λ_1 and by a mean flood magnitude θ_2 larger than the corresponding parameter θ_1 of the basic distribution. The shapes of the distribution vary for different values of λ_1, λ_2, θ_1 and θ_2.

Three parameter estimation methods have been proposed for fitting the TCEV distribution to annual flood series. Canfield (1979) suggested a least squares technique, while Rossi, et al. (1984) presented a procedure based on the maximum likelihood estimation (MLE) method. Using the MLE method, Fiorentino, et al. (1985) developed a regional estimation algorithm. Small sample properties of the site-specific and regionalized TCEV-MLE procedure were assessed by Fiorentino and Gabriele (1985), and Arnell and Gabriele (1986). In particular, the latter compared the regionalized TCEV-MLE algorithm with other regional estimators. Although various features of the TCEV-MLE method exhibited a competitive performance, an improvement of the site-specific estimators was suggested by Fiorentino and Gabriele (1985). Furthermore, Fiorentino, et al. (1986) noted that regional estimates of some parameters could be still improved. Using the principle of maximum entropy (POME), Fiorentino et al. (1987a,b) developed another method for estimating parameters of the TCEV distribution. The POME method of parameter estimation is suitable for application in both the site-specific and regional cases and appears simpler than the maximum likelihood estimation method. Statistical properties of this regionalized estimation procedure were evaluated using a Monte Carlo approach and compared with those of the maximum likelihood regional estimators.

22.1 Ordinary Entropy Method

22.1.1 SPECIFICATION OF CONSTRAINTS

Taking logarithm to the base 'e' of equation (22.1), one gets

$$\ln f(x) = \ln \lambda_1 - \ln \theta_1 - \frac{x}{\theta_1} - \lambda_1 \exp(-\frac{x}{\theta_1})$$

$$+ \lambda_2 \exp(-\frac{x}{\theta_2}) + \ln[1 + \frac{(\lambda_2/\theta_2)\exp(-x/\theta_2)}{(\frac{\lambda_1}{\theta_1})\exp(-\frac{x}{\theta_1})}] \tag{22.4}$$

Multiplying equation (22.4) by [-f (x)] and integrating between 0 and ∞ yield the entropy function:

$$I(x) = -\int_{-\infty}^{\infty} \ln f(x) f(x)\, dx = (\ln \theta_1 - \ln \lambda_1) \int_{-\infty}^{\infty} f(x)\, dx +$$

$$+ \frac{1}{\theta_1} \int_{-\infty}^{\infty} x f(x)\, dx + \lambda_1 \int_{-\infty}^{\infty} \exp(-x/\theta_1) f(x)\, dx +$$

$$+ \lambda_2 \int_{-\infty}^{\infty} \exp(-x/\theta_2) f(x)\, dx -$$

$$- \int_{-\infty}^{\infty} \ln[1 + \frac{(\lambda_2/\theta_2)\exp(-x/\theta_2)}{(\lambda_1/\theta_1)\exp(-x/\theta_1)}] f(x)\, dx \tag{22.5}$$

From equation (22.5) the equations of constraints can be written as:

$$\int_{-\infty}^{\infty} f(x)\, dx = 1 \tag{22.6}$$

$$\int_{-\infty}^{\infty} x f(x)\, dx = E[x] \tag{22.7}$$

$$\int_{-\infty}^{\infty} \exp(-x/\theta_1) f(x)\, dx = E[\exp(-x/\theta_1)] \tag{22.8}$$

$$\int_{-\infty}^{\infty} \exp(-x/\theta_2) f(x)\, dx = E[\exp(-x/\theta_2)] \tag{22.9}$$

$$\int_{-\infty}^{\infty} \ln(1 + \frac{(\lambda_2/\theta_2)\exp(-x/\theta_2)}{(\lambda_1/\theta_1)\exp(-x/\theta_1)} f(x)\, dx$$

$$= E[\ln(1 + \frac{(\lambda_2/\theta_2)\exp(-x/\theta_2)}{(\lambda_1/\theta_1)\exp(-x/\theta_1)})] \tag{22.10}$$

The constraints are to be evaluated from data, directly or indirectly. It may be noted that the first three constraints are the same as those used for deriving EV1 distribution (Jowitt, 1979; Singh, et al., 1985), while the fourth constraint, which is analogous to the third one, provides information on the outlying component. The final constraint combines the information between the basic and outlying components.

22.1.2 CONSTRUCTION OF ZEROTH LAGRANGE MULTIPLIER

The least biased pdf, f (x), consistent with equations (22.6) - (22.10) and based on POME, takes the form:

$$f(x) = \exp(-a_0 - a_1 x - a_2 \exp(-x/\theta_1) - a_3 \exp(-x/\theta_2) -$$

$$- a_4 \ln[1 + \frac{\lambda_2 \theta_1}{\theta_2 \lambda_1} \exp[-x(\frac{1}{\theta_2} - \frac{1}{\theta_1})]) \tag{22.11}$$

where a_0, a_1, a_2, a_3, and a_4 are Lagrange multipliers. The zeroth Lagrange multiplier a_0 is determined as follows. Inserting equation (22.11) in equation (22.6), we get the partition function:

$$\exp(a_0) = \int_{-\infty}^{\infty} \exp[-a_1 x - a_2 \exp(-x/\theta_1) -$$

$$- a_3 \exp(-x/\theta_2)] [1 + \frac{\lambda_2 \theta_1}{\theta_2 \lambda_1} \cdot$$

$$\cdot \exp(-x(\frac{1}{\theta_2} - \frac{1}{\theta_1})^{-a_4})] dx \tag{22.12}$$

Let z = $z = \lambda_1 \exp(-x/\theta_1), \theta = \theta_2/\theta_1, \lambda = = \Lambda = \lambda_2/(\lambda_1^{1/\theta})$. After simple manipulation, the zeroth Lagrange multiplier is obtained:

$$a_0 = \ln \theta_1 - a_1 \theta_1 \ln \lambda_1 + \ln \int_0^{\infty} z^{a_1 \theta_1 - 1}$$

$$\cdot \exp(-a_2 x/\lambda_1) \exp[-a_3 z^{(1/\theta)} \lambda_1^{(-1/\theta)}] \cdot$$

$$\cdot [1 + \frac{\lambda}{\theta} z^{((1/\theta)-1)}]^{-a_4} dz \tag{22.13}$$

One also gets the zeroth Lagrange multiplier from equation (22.12) as

$$a_0 = \ln \int_{-\infty}^{\infty} \exp[-a_1 x - a_2 \exp(-x/\theta_1) - a_3 \exp(-x/\theta_2)]$$

$$[1 + \frac{\lambda_2 \theta_1}{\lambda_1 \theta_2} \exp\left\{-x(\frac{1}{\theta_2} - \frac{1}{\theta_1})^{-a_4}\right\}] dx \tag{22.14}$$

Let us refer to the integral in equation (22.13) as I_0. Inserting equation (22.13) into equation (22.11), we get

$$f(x) = \frac{1}{\theta_1} \lambda_1^{a_1 \theta_1} \exp[-a_1 x - a_2 \exp(-x/\theta_1) -$$

$$- a_3 \exp(-x/\theta_2)] \; [1 + \frac{(\lambda_2/\theta_2) \exp(-x/\theta_2)}{(\lambda_1/\theta_1) \exp(-x/\theta_1)}]^{-a_4} / I_0 \qquad (22.15)$$

When

$$a_1 = 1/\theta_1 \,, \quad a_2 = \lambda_1 \,, \quad a_3 = \lambda_2 \,, \quad \text{and} \quad a_4 = -1 \qquad (22.16)$$

integral I_0 becomes unity and equation (22.15) becomes equation (22.1)

22.1.3 RELATION BETWEEN CONSTRAINTS AND PARAMETERS

The relationship between the parameters of TCEV distribution and the constraints are specified by partially differentiating a_0 given by equation (22.13) with respect to a_1, a_2, a_3, and a_4 respectively.

$$\frac{\partial a_0}{\partial a_1} = -E[x] = -\theta_1 \ln \lambda_1 + \theta_1 \int_0^\infty \ln y \exp(-y - \lambda y^{(1/\theta)}) \cdot$$

$$\cdot (1 + \frac{\lambda}{\theta} y^{((1/\theta)-1)}) \, dy \qquad (22.17)$$

$$\frac{\partial a_0}{\partial a_2} = -E[\exp(-x/\theta_1)] = -\frac{1}{\lambda_1} \int_0^\infty y \exp(-y - \lambda y^{(1/\theta)}) \cdot$$

$$\cdot (1 + \frac{\lambda}{\theta} y^{((1/\theta)-1)}) \, dy \qquad (22.18)$$

$$\frac{\partial a_0}{\partial a_3} = -E[\exp(-x/\theta_2)] = -\frac{1}{\lambda^{1/\theta}} \int_0^\infty y^{(1/\theta)} \cdot$$

$$\cdot \exp(-y - \lambda y^{(1/\theta)}) \; (1 + \frac{\lambda}{\theta} y^{((1/\theta)-1)}) \, dy \qquad (22.19)$$

$$\frac{\partial a_0}{\partial a_4} = -E[\ln(1 + \frac{(\lambda_2/\theta_2) \exp(-x/\theta_2)}{(\lambda_1/\theta_1) \exp(-x/\theta_1)}]$$

$$= -\int_0^\infty \ln(1 + \frac{\lambda}{\theta} y^{((1/\theta)-1)}) \exp(-y - \lambda y^{1/\theta}) \cdot$$

$$\cdot (1 + \frac{\lambda}{\theta} y^{(1/\theta)-1)}) \, dy \qquad (22.20)$$

Solving integrals in equations (22.17) to (22.19) provides, one gets

$$E[x] = \theta_1 \ln \lambda_1 + \theta_1 \gamma - \theta_1 \sum_{j=1}^{\infty} \frac{(-1)^j \lambda^j}{j!} \Gamma(j/\theta) \tag{22.21}$$

$$E[\exp(-x/\theta_1)] = \frac{1}{\lambda_1}[1 + \frac{1}{\theta} \sum_{j=1}^{\infty} (-1)^j \frac{\lambda^j}{(j-1)!} \Gamma(j/\theta) \tag{22.22}$$

$$E[\exp(-x/\theta_2)] = -\frac{1}{\theta \lambda_2} \sum_{j=1}^{\infty} \frac{(-1)^j \lambda^j}{(j-1)!} \Gamma(j/\theta) \tag{22.23}$$

where $\gamma = 0.5772$ is the Euler's constant and $\Gamma(\cdot)$ is the gamma function. Equations (22.21) to (22.23) are also indicated by Beran et al. (1986). θ and λ are dimensionless parameters, already defined in terms of the four parameters of the TCEV distribution. The integral in equation (22.20) cannot be solved explicitly. But, for $\theta > 1.5$, it is closely approximated by the following function

$$\int_0^{\infty} \ln(1 + \frac{\lambda}{\theta} y^{((1/\theta)-1)}) \exp(-y - \lambda y^{1/\theta}) (1 + \frac{\lambda}{\theta} h^{((1/\theta)-1)}) dy$$

$$= 0.1 \exp(-1) (3 + \theta)^{2.059} \lambda^{(\ln 3 - 2(5.5)^{-\theta})} \tag{22.24}$$

Therefore, the fourth constraints can be related to the parameters by

$$E[\ln(1 + \frac{(\lambda_2/\theta_2)\exp(-x/\theta_2)}{(\lambda_1/\theta_1)\exp(-x/\theta_1)})] = E[Z] = 0.1 \exp(-1) (3 + \theta)^{0.059} \cdot$$

$$\cdot \lambda^{(\ln 3 - 2(5.5)^{\theta})} \tag{22.25}$$

The goodness of this approximation is shown in Figure 22.1. The curves approximating the expectation in equation (22.20) have not been plotted for $\theta < 1.5$ to avoid any confusion at the left-bottom where they tend to overlap with each other. Moreover, the goodness of the approximation deteriorates in this range.

Equations (22.21) - (22.23) and (22.25) show that constraints are related to the moments or moment-ratios of the distribution. In fact, besides the obvious case of the constraint $E[y_1(x)]$ representing the population mean of x, it is clear that $E[y_4(x)]$ depends on the dimensionless parameters θ and λ only, while both $E[y_3(x)]$ and $E[y_2(x)]$ depend on θ, λ and λ_1 (note that λ_2 is a function of λ_1 via θ and λ). A similar dependence is exhibited by the theoretical coefficients of skewness (and kurtosis) and variation respectively as can be easily shown using expressions of the moments given by Beran, et al. (1986). This implies that the estimation of the constraints will likely have a variability increasing with the rank.

Figure 22.2, where Λ=z, compares $E[y_4(x)] = E[Z]$ with the mean of the transformed variate y (or u)

$$y = u = \frac{x}{\theta_1} - \ln \lambda_1 \qquad (22.26)$$

which is also TCEV distributed and depends on θ and λ only. Skewness and kurtosis of both Y (or U) and X variates are the same, while the mean of Y is given by dividing the last two terms on the right-hand side of equation (22.21) by θ_1. One can note that $E[Y_4(x)]$ exhibits a shape similar to that of E [Y] and that it is more sensitive to changes in either θ or λ, particularly in the range of low values. This stipulates that entropy should provide dimensionless parameter estimates much less variable than those based on the method of moments.

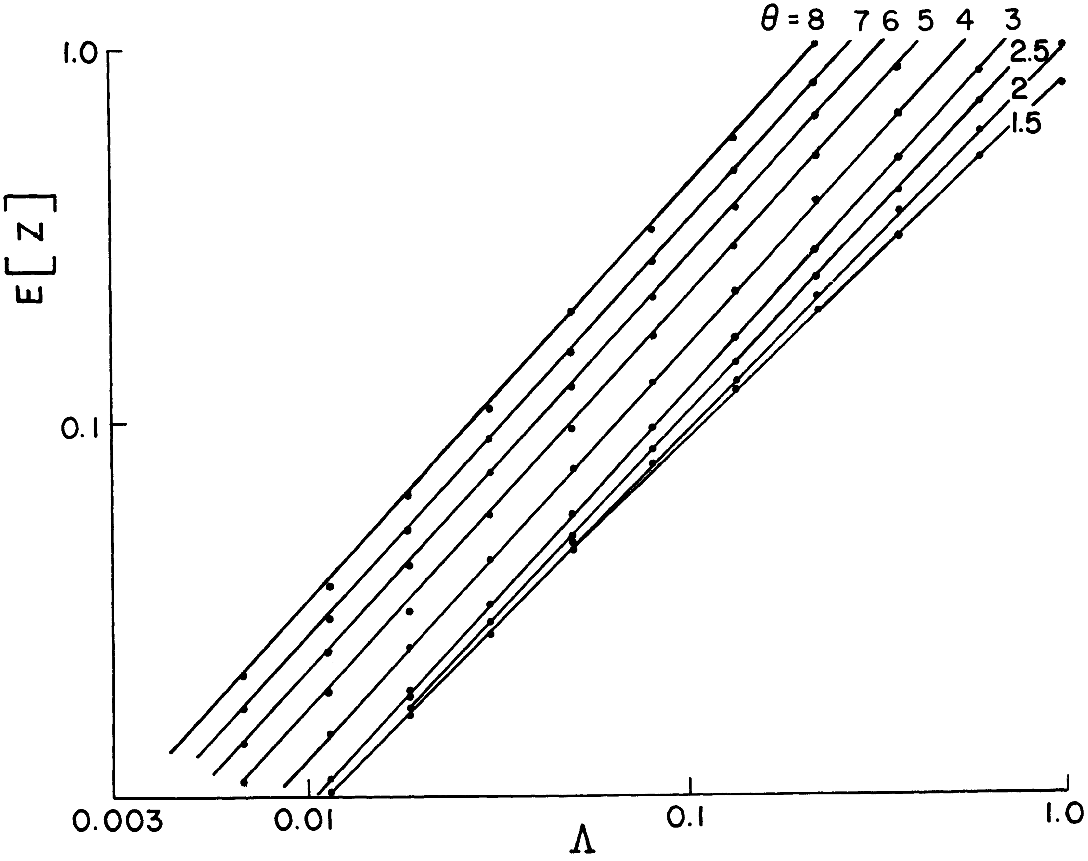

Figure 22.1 Mean of E[Z] versus Λ for various values of θ.

Furthermore, equation (22.23) shows that $E[y_3]$ is also related to the probability, P_2, that the annual maximum value of X comes from the outlying component, P_2 having been derived by Beran, et al. (1986) as

$$P_2 = -\frac{1}{\theta}\sum_{j=1}^{\infty}\frac{(-1)^j\,\lambda^j}{(j-1)!}\,\Gamma(j/\theta) \tag{22.27}$$

In fact, combining equations (22.23) and (22.27) gives $E[\exp(-x/\theta_2)] = P_2/\lambda_2$. Analogously, it can be shown that $E[\exp(-x/\theta_1)] = P_1/\lambda_1$, where P_1 represents the chance that the annual maximum value of X comes from the basis component.

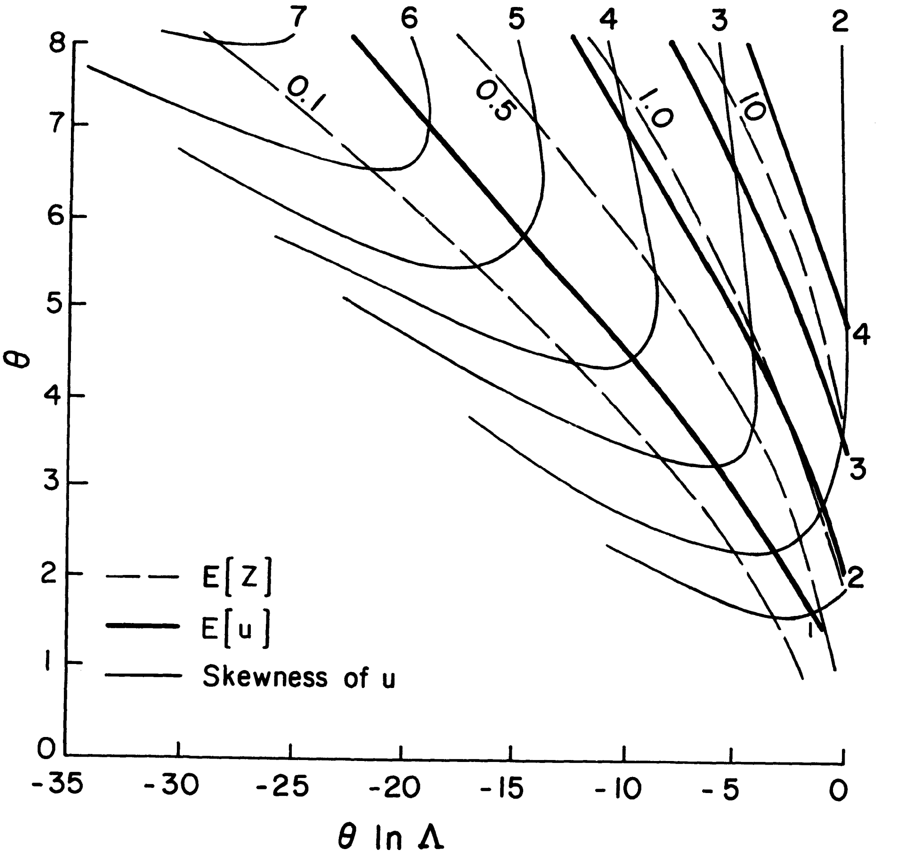

Figure 22.2 Variation of mean and skewness coefficient of the transformed TCEV variate, and mean of E[Z] with parameters Λ and θ.

A graph showing P_2 versus θ ln λ has been provided by Beran, et al. (1986). It shows that for a given value of P_2, θ is a quasi-linear function of θ ln λ in almost the entire difinition range of λ. This suggests that an approximate relationship solely between P_2 and λ (or Λ) can be confidently used for first order calculations. Figure 22.3 shows the goodness of this approximation, which has the following equation:

$$P_2 = 0.65\ \lambda^{0.85} \tag{22.28}$$

Thus, one can write

$$E[\exp(-x/\theta_2)] = \frac{0.65\ \lambda^{0.85}}{\lambda_2} \tag{22.29}$$

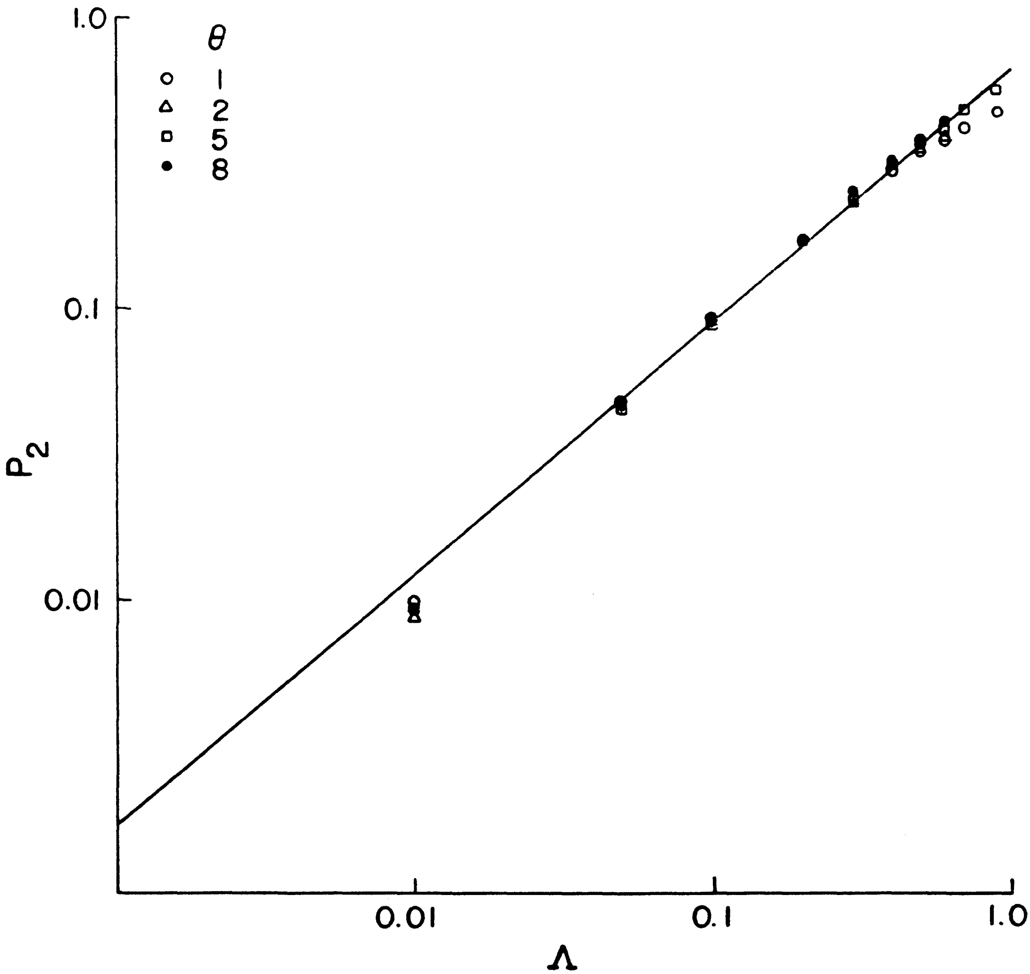

Figure 22.3 Outlier probability versus Λ for various values of θ.

22.1.4 ESTIMATION OF PARAMETERS

22.1.4.1 Point Estimation: Equation for estimation of parameters can be obtained by substituting sample values for the population means on the left-hand side of equations (22.21) to (22.23) and (22.25). The system of equations to be solved for giving estimates of the four parameters of the TCEV distribution thus takes the form:

$$\overline{x} = \theta_1 \ln \lambda_1 + \theta_1 \gamma - \theta_1 \sum_{j=1}^{\infty} \frac{(-1)^j \lambda^j}{j!} \Gamma(j/\theta) \tag{22.30}$$

$$\overline{\exp(-x/\theta_1)} = \frac{1}{\lambda_1} [1 + \frac{1}{\theta} \sum_{j=1}^{\infty} \frac{(-1)^j \lambda^j}{(j-1)!} \Gamma(j/\theta)] \tag{22.31}$$

$$\overline{\exp(-x/\theta_2)} = -\frac{1}{\theta \lambda_2} \sum_{j=1}^{\infty} \frac{(-1)^j \lambda^j}{(j-1)!} \Gamma(j/\theta) \tag{22.32}$$

$$\overline{\ln[1 + \frac{(\lambda_2/\theta_2)\exp(-x/\theta_2)}{(\lambda_1/\theta_1)\exp(-x/\theta_1)}]} = 0.1 \exp(-1)(3 + \theta)^{2.059} \lambda^{\ln 3 - 2(5.5^{-\theta})} \tag{22.33}$$

where the bar indicates that the sample mean of the underlying function is considered. For simplicity, the left-hand sides of equations (22.30) and (22.33) will be hereafter referred to as $\overline{Y}_1, \ldots, \overline{Y}_4$, respectively. Eliminating λ_2 by way of θ, λ and λ_1, and rearranging, we get

$$\lambda = \frac{1}{\overline{Y}_2} [1 + \frac{1}{\theta} \sum_{j=1}^{\infty} \frac{(-1)^j \lambda^j}{(j-1)!} \Gamma(j/\theta)] \tag{22.34}$$

$$\theta_1 = \overline{Y}_1/(\ln[1 + \frac{1}{\theta} \sum_{j=1}^{\infty} \frac{(-1)^j \lambda^j}{(j-1)!} \Gamma(j/\theta)]/\overline{Y}_2 + \gamma - \sum_{j=1}^{\infty} \frac{(-1)^j \lambda^j}{j!} \Gamma(j/\theta)) \tag{22.35}$$

$$\lambda = \frac{1 - \lambda_1 \overline{Y}_2}{\lambda_1^{1/\theta} \overline{Y}_3} \tag{22.36}$$

$$0.1 \exp(-1)(3 + \theta)^{2.059} (\frac{1 - \lambda_1 \overline{Y}_2}{\lambda_1^{1/\theta} \overline{Y}_3}) \ln 3 - 2(5.5^{-\theta}) = \overline{Y}_4 \tag{22.37}$$

Putting $\overline{Y}_3$ and $\overline{Y}_4$ respectively in the form:

$$\overline{Y}_3 = \overline{\exp[-x/(\theta_1 \theta)]} \tag{22.38}$$

$$\overline{Y}_4 = \overline{\ln(1 + \frac{1}{\theta \lambda_1} \frac{1 - \lambda_1 \overline{Y}_2}{\overline{Y}_3} \exp[-\frac{x}{\theta_1}(\frac{1}{\theta} - 1)])} \tag{22.39}$$

One obtains that the only unknown in equation (22.37) is θ. On the other hand: (1) λ does not appear on the right-hand side in equation (22.36); (2) θ_1 is the only unknown once θ and λ have been evaluated in equation (22.35); and (3) λ_1 does not appear on the right-hand side in equation (22.34). Therefore, a successive substitution iterative scheme can be developed for estimating the four parameters as follows. Assign tentative values to θ and λ , then successively estimate θ_1 by equation (22.35), λ_1 by equation (22.34), θ by equation (22.37), and λ by equation (22.36). Substitute the last values of θ and λ for those previously obtained and start again from estimation of θ_1. Stop when θ and λ no longer change. Note that the procedure is fast because equation (22.34) and (22.36) admit solution in closed form and equations (22.35) and (22.37), though not explicit, can be easily solved numerically, for each exhibits one unknown only.

22.1.4.2 Regional Estimation: A regional flood frequency estimation algorithm can be developed using equations (22.34) - (22.37) (obviously together with equations (22.38) and (22.39)), which can also be used to validate the regionalization model proposed by Fiorentino, et al. (1985) and also described in Fiorentino, et al. (1986). In short, this model assumes that dimensionless parameters θ and λ do not change over extensive regions, while parameter λ_1 is constant in smaller areas. In this chapter, a regionalization algorithm, based on POME, to estimate θ and λ is presented.

Suppose there are k gauged sites in a selected region which is assumed to be homogeneous with respect to θ and λ. Let each site have an annual flood series (AFS) with n years of record. At each site, one must estimate the basic component parameters θ_1 and λ_1, which vary from site to site, plus the two regional values of θ and λ. Hence, there are in practice 2k + 2 unknowns. An equal number of independent equations is then needed.

The first 2k equations of the algorithm proposed herein arise from writing equations (22.34) and (22.35) k times, once for each available AFS. The other two equations are derived by taking the average of left-hand sides of equations (22.32) and (22.33) over all k sites.

$$\frac{1}{kn}\sum_{r=1}^{k}\sum_{i=1}^{n}\lambda_{2r}\exp\left(-\frac{x_{ir}}{\theta_{2r}}\right) = -\frac{1}{\theta}\sum_{j=1}^{\infty}\frac{(-1)^j\lambda^j}{(j-1)!}\Gamma(j/\theta) \qquad (22.40)$$

$$\frac{1}{kn}\sum_{r=1}^{k}\sum_{i=1}^{n}\ln\left[1+\frac{(\lambda_2/\theta_2)_r\exp(-x_{ir}/\theta_{2r})}{(\lambda_1/\theta_1)_r\exp(-x_{ir}/\theta_{ir})}\right]$$

$$= 0.1\exp(-1)\,(3+\theta)^{2.059}\,\lambda^{\ln 3 - 2(5.5)^{-\theta}} \qquad (22.41)$$

Equations (22.40) and (22.41) can be written in a different manner taking into account the transformation:

$$y = \frac{x}{\theta_1} - \ln\lambda_1 \qquad (22.42)$$

which makes their left-hand sides dependent on θ and λ only. Since the values of θ and λ are assumed to be constant at every site, the following forms are thus obtained

$$\frac{1}{kn}\sum_{i=1}^{kn}\lambda\exp(-y_i/\theta) = \frac{1}{\theta}\sum_{j=1}^{\infty}\frac{(-1)^j\lambda^j}{(j-1)!}\Gamma(j/\theta) \qquad (22.43)$$

$$\frac{1}{kn}\sum_{i=1}^{kn}\ln\left(1+\frac{\lambda}{\theta}\exp\left[-\left(\frac{1}{\theta}-1\right)y_i\right]\right) \qquad (22.44)$$

The two procedures are mutually equivalent, at least when the available AFS's have the same length at every site. Whichever is used, estimates of θ_1 and λ_1 at any site need to be obtained together with regional estimates of θ and λ. In fact, both sets of equations depend on the basic component parameters, the former in an explicit manner and the latter through the transformation of equation (22.42). The iterative scheme proposed by Fiorentino and Gabriele (1985) for the regionalized TCEV-MLE procedure also successfully works using the POME-based estimation method. Details of this scheme can be readily found in Fiorentino, et al. (1986b).

For a comparison between the proposed procedure and the MLE method, two features of the former look favorable: (1) Estimation of the basic component parameters, once θ and λ have been evaluated, is relatively simple, for only equation (22.35) needs to be solved numerically. (2) The equations contain a smaller number of exponentials to be solved. However, only a large number of Monte Carlo experiments covering a wide range of situations, can confirm whether the TCEV-POME estimation procedure is competitive. The results based on a limited number of computer simulation experiments will be discussed later on.

22.2 Other Methods of Parameter Estimation

The two parameter estimation methods have been proposed for fitting the TCEV distribution to annual flood series. Rossi, et al, (1984) presented a procedure based on the maximum likelihood estimation (MLE), and Fiorentino, et al. (1987) derived the entropy-based parameter estimation method.

2.2.1 METHOD OF MAXIMUM LIKELIHOOD ESTIMATION

The TCEV distribution has a finite probability, $\exp(-\lambda_1-\lambda_2)$ when x=0. Since such a probability is negligible, therefore, the probability density function, Equation (22.1) can be reexpressed as

$$f(x) = F(x)\,\Psi(x) \tag{22.45}$$

where F(x) is the cumulative distribution function and

$$\Psi(x) = (\lambda_1/\theta_1)\exp(-x/\theta_1) + (\lambda_2/\theta_2)\exp(-x/\theta_2)$$

so that the logarithm of the likelihood function, L, may be written as

$$L = \sum_{i=1}^{n} \ln f(x_i) = \sum_{i=1}^{n} \ln F(x_i) + \sum_{i=1}^{n} \ln \Psi(x_i) \tag{22.46}$$

Partially differentiating equation (22.46) with respect to the four parameters to be estimated separately and equating each derivative to zero yield

$$\frac{\partial L}{\partial \lambda_1} = -\sum_{i=1}^{n} \exp(-x_i/\theta_i) + \frac{1}{\theta_1}\sum_{i=1}^{n} \frac{\exp(-x_i/\theta_1)}{\Psi(x_i)} = 0 \tag{22.47}$$

$$\frac{\partial L}{\partial \lambda_2} = -\sum_{i=1}^{n} \exp(-x_i/\theta_2) + \frac{1}{\theta_2}\sum_{i=1}^{n} \frac{\exp(-x_i/\theta_2)}{\Psi(x_i)} = 0 \qquad (22.48)$$

$$\frac{\partial L}{\partial \theta_1} = -\frac{\lambda_1}{\theta_1^2}[\sum_{i=1}^{n} x_i \exp(-x_i/\theta_1) + \sum_{i=1}^{n} \frac{\exp(-x_i/\theta_1)(1-x_i/\theta_1)}{\Psi(x_i)}] = 0 \qquad (22.49)$$

$$\frac{\partial L}{\partial \theta_2} = -\frac{\lambda_2}{\theta_2^2}[\sum_{i=1}^{n} x_i \exp(-x_i/\theta_2) + \sum_{i=1}^{n} \frac{\exp(-x_i/\theta_2)(1-x_i/\theta_2)}{\Psi(x_i)}] = 0 \qquad (22.50)$$

Therefore, the estimation equations given by Rossi et al. (1984) are

$$\lambda_1 = \lambda_1 [\sum_{i=1}^{n} \frac{\exp(x_i/\theta_1)}{\Psi(x_i)}]/[\theta_1 \sum_{i=1}^{n} \exp(-x_i/\theta_1) \qquad (22.51)$$

$$\lambda_2 = \lambda_2 [\sum_{i=1}^{n} \frac{\exp(x_i/\theta_2)}{\Psi(x_i)}]/[\theta_2 \sum_{i=1}^{n} \exp(-x_i/\theta_2) \qquad (22.52)$$

$$\theta_1 = [\sum_{i=1}^{n} \frac{x_i \exp(-x_i/\theta_1)}{\Psi(x_i)}]/[\sum_{i=1}^{n} x_i \exp(-x_i/\theta_1) + \sum_{i=1}^{n} \frac{\exp(-x_i/\theta_1)}{\Psi(x_1)} \qquad (22.53)$$

$$\theta_2 = [\sum_{i=1}^{n} \frac{x_i \exp(-x_i/\theta_2)}{\Psi(x_i)}]/[\sum_{i=1}^{n} x_i \exp(-x_i/\theta_2) + \sum_{i=1}^{n} \frac{\exp(-x_i/\theta_2)}{\Psi(x_1)} \qquad (22.54)$$

The four equations (22.51)-(22.54) can be solved by an iterative scheme involving successive substitution.

22.2.3 METHOD OF PROBABILITY WEIGHTED MOMENTS

For the TCEV distribution the probability weighted moments, PWM_r, are (Beran, et al., 1986):

$$PWM_r = E[x\{F(x)\}^r] = PWM_r^{(1)} + \frac{\theta_1}{r+1} \tag{22.55}$$

where

$$PWM_r^{(1)} = \frac{\theta_1}{r+1}[r + \log\lambda_1 + \log(r+1) \tag{22.56}$$

is the rth probability weighted moment of the basic series, and

$$T_r = \sum_{i=1}^{\infty} (-1)^{j-1}\lambda^j (r+1)^{j(1-1/\theta)}\Gamma(j/\theta)/j! \tag{22.57}$$

and γ is the Euler's constant. Thus, we have

$$W_0 = \theta_1[\gamma + \ln\lambda_1 + \sum_{i=1}^{n} (-1)^{j-1}\lambda^j \Gamma(j/\theta)/\Gamma(j+1)] \tag{22.58}$$

$$W_1 = \frac{\theta_1}{2}[\gamma + \ln\lambda_1 + \ln 2 + \sum_{i=1}^{n} (-1)^{j-1}\lambda^j 2^{j(1-1/\theta)}\Gamma(j/\theta)/\Gamma(j+1)] \tag{22.59}$$

$$W_2 = \frac{\theta_1}{3}[\gamma + \ln\lambda_1 + \ln 3 + \sum_{i=1}^{n} (-1)^{j-1}\lambda^j 3^{j(1-1/\theta)}\Gamma(j/\theta)/\Gamma(j+1)] \tag{22.60}$$

$$W_3 = \frac{\theta_1}{4}[\gamma + \ln\lambda_1 + \ln 4 + \sum_{i=1}^{n} (-1)^{j-1}\lambda^j 4^{j(1-1/\theta)}\Gamma(j/\theta)/\Gamma(j+1)] \tag{22.61}$$

Given a random sample of size n from the TCEV distribution, estimation of PWM_r is most conveniently based on the ordered sample $x_1 \le x_2 \le \ldots \le x_n$. The statistic

$$b_r = \frac{1}{n}\sum_{j=1}^{n} \frac{(j-1)(j-2)\ldots\ldots\ldots\ldots(j-r)}{(n-1)(n-2)\ldots\ldots\ldots\ldots\ldots(n-r)} \tag{22.62}$$

is an unbiased estimator of W_r (Landwehr, et al., 1986). Equations (22.58)-(22.61) can be solved by an iterative scheme involving successive substitution.

22.3 Comparative Evaluation of Parameter Estimators

It is important to evaluate the performance of all available estimators of a distribution, especially for small sample sizes, for which the variability of estimators is quite large and so is the marked difference in their performance. To minimize design losses, one would like to use the most efficient estimator. Approximate formulae can be derived for asymptotic standard error of several of the estimators. But one is chiefly interested in the sampling properties of the estimators for rather small sizes ($n \leq 50$) not covered by the asymptotic formulae. The sampling distribution of the estimator is generally intractable in the sample range of interest. Monte Carlo sampling experiments, therefore, offer an attractive procedure for evaluating and comparing the performance of estimators. Cunnane (1986) pointed out the simulation experiments that have been reported in most recent work on flood frequency analysis. Thus, the use of simulation has become a standard technique to evaluate the performance of competing estimators.

22.3.1 EXPERIMENTAL DESIGN

The estimation procedure, outlined above, was addressed, if only approximately, using the Monte Carlo technique, and generating synthetic series from a TCEV distribution with parameters $\theta_1 = 10$, $\lambda_1 = 10$, $\theta = 3.067$, and $\lambda = 0.173$, which is what was used to evaluate the TCEV-MLE procedure (Fiorentino and Gabriele, 1985; Arnell and Gabriele, 1986).Two measures of performance were used: the standardized bias (BIAS) and the standardized root mean square error (RMSE). Since the regionalization is the natural field for application of a distribution with a large number of parameters such as four, the attention was principally devoted to the assessment of the regionalized estimators. One hundred repetitions of forty synthetic series, each with forty years of record, were generated, i.e., 100 homogeneous regions, each with 40 gauged sites, were simulated. Then the regionalization algorithm described above was applied. For each repetition, a regional estimate of θ and λ together with forty on-site elements of θ_1 and λ_1 were obtained. BIAS and RMSE of parameter and quantile regional estimators were then evaluated. Of course, due to the very short number of experiments, these results are not expected to reproduce the true values of BIAS and RMSE, but they do not provide a first order approximation of the likely results.

22.3.2 BIAS IN PARAMETER ESTIMATION

The results of the parameter BIAS and RMSE analyses for each case showed that MLE always produced the highest BIAS in estimating λ_2 for all sample sizes over the two cases, and PWM produced the highest BIAS in estimating θ_2. For case 1, POME and PWM were comparable in estimating λ_2, and when $n \geq 100$ MLE performed better in estimating θ_1, θ_2 and λ_1, and when $n<100$ PWM performed better in estimating θ_1 and POME performed better in estimating θ_2 and POME and PWM were comparable in estimating λ_1; for case 2, PWM performed the best in estimating λ_2 and ENT performed better in estimating θ_1, θ_2 and λ_1.

22.3.3 RMSE IN PARAMETER ESTIMATION

The results of RMSE values in parameter estimation showed that MLE produced the highest RMSE in estimating λ_2 and PWM performed the worst in estimating θ_2 for all sample sizes over the two cases. For case 1, when $n \leq 100$ PWM performed better in estimating θ_1 and λ_1 while when $n>100$ MLE performed better, and ENT produced the least RMSE in estimating λ_2 and MLE performed the best in estimating θ_2 for all sample sizes; for case 2 PWM performed better in estimating θ_1, λ_1 and λ_2, and POME performed better in estimating θ_2.

22.3.4 BIAS AND RMSE IN QUANTILE ESTIMATION

The BIAS and RMSE values of the quantile estimates for the TCEV showed that, in general, MLE performed better in terms of quantile BIAS and RMSE for all sample sizes over the two cases.

22.3.5 CONCLUDING REMARKS

The results on relative performance of the three parameter estimation procedures showed that for case 1, PWM performed better when the sample size (n) ≤ 100, and MLE performed better when n>100; for case 2, POME performed the best for all sample sizes. For case 1, POME performed better for small sample sizes (n<100) and MLE performed better for large sample sizes ($n \geq 100$); for case 2, MLE performed better.

References

Beran, M., Hosking, J.R.M. and Arnell, N., 1986. Comment on 'Two-component extreme value distribution for flood frequency analysis' by Fabio Rossi, Mauro Fiorentino and Pasquale Versace. *Water Resources Research*, Vol. **20**, No. 7, pp. 847-856.

Cunnane, C., 1987. Review of statistical models or flood frequency estimation. In: V.P.Singh (ed.), *Hydrologic Frequency Modeling*, pp. 49-95, D. Reidel Publishing Company, Dordrecht, The Netherlands.

Fiorentino, M., Versace, P. and Rossi, F., 1985. Regional flood frequency estimation using the two-component extreme value distribution. *Hydrological Sciences Journal*, Vol. 30, pp. 51-64.

Landwehr, J. M., Matalas, N. C. and Wallis, J. R., 1979. Probability weighted moments compared with some traditional techniques in estimating Gumbel parameters and quantiles. *Water Resources Research*, Vol. **15**, No. 5, pp. 1055-1064.

Rossi, F., Fiorentino, M., Versace, P., 1984. Two-component extreme value distribution for flood frequency analysis. Water Resources Research, Vol. 20, No. 7, pp. 847-856.

Vogel, R.W. and McMartin, D.E., 1991. Probability plot goodness-of-fit and skewness estimation procedures for the Pearson type 3 distribution. *Water Resources Research*, Vol. 27, No. 12, pp. 3149-3158.

Wallis, J.R. and Wood, E.F., 1985. Relative accuracy of log Pearson procedures. Journal of Hydraulic Engineering, Vol.111, No. 7, pp. 1043-1056.

SUBJECT INDEX

Water Science and Technology Library

1. A.S. Eikum and R.W. Seabloom (eds.): *Alternative Wastewater Treatment.* Low-Cost Small Systems, Research and Development. Proceedings of the Conference held in Oslo, Norway (7–10 September 1981). 1982 ISBN 90-277-1430-4
2. W. Brutsaert and G.H. Jirka (eds.): *Gas Transfer at Water Surfaces.* 1984 ISBN 90-277-1697-8
3. D.A. Kraijenhoff and J.R. Moll (eds.): *River Flow Modelling and Forecasting.* 1986 ISBN 90-277-2082-7
4. World Meteorological Organization (ed.): *Microprocessors in Operational Hydrology.* Proceedings of a Conference held in Geneva (4–5 September 1984). 1986 ISBN 90-277-2156-4
5. J. Němec: *Hydrological Forecasting.* Design and Operation of Hydrological Forecasting Systems. 1986 ISBN 90-277-2259-5
6. V.K. Gupta, I. Rodríguez-Iturbe and E.F. Wood (eds.): *Scale Problems in Hydrology.* Runoff Generation and Basin Response. 1986 ISBN 90-277-2258-7
7. D.C. Major and H.E. Schwarz: *Large-Scale Regional Water Resources Planning.* The North Atlantic Regional Study. 1990 ISBN 0-7923-0711-9
8. W.H. Hager: *Energy Dissipators and Hydraulic Jump.* 1992 ISBN 0-7923-1508-1
9. V.P. Singh and M. Fiorentino (eds.): *Entropy and Energy Dissipation in Water Resources.* 1992 ISBN 0-7923-1696-7
10. K.W. Hipel (ed.): *Stochastic and Statistical Methods in Hydrology and Environmental Engineering.* A Four Volume Work Resulting from the International Conference in Honour of Professor T. E. Unny (21–23 June 1993). 1994
    10/1: Extreme values: floods and droughts ISBN 0-7923-2756-X
    10/2: Stochastic and statistical modelling with groundwater and surface water applications ISBN 0-7923-2757-8
    10/3: Time series analysis in hydrology and environmental engineering ISBN 0-7923-2758-6
    10/4: Effective environmental management for sustainable development ISBN 0-7923-2759-4
    Set 10/1–10/4: ISBN 0-7923-2760-8
11. S.N. Rodionov: *Global and Regional Climate Interaction: The Caspian Sea Experience.* 1994 ISBN 0-7923-2784-5
12. A. Peters, G. Wittum, B. Herrling, U. Meissner, C.A. Brebbia, W.G. Gray and G.F. Pinder (eds.): *Computational Methods in Water Resources X.* 1994 Set 12/1–12/2: ISBN 0-7923-2937-6
13. C.B. Vreugdenhil: *Numerical Methods for Shallow-Water Flow.* 1994 ISBN 0-7923-3164-8
14. E. Cabrera and A.F. Vela (eds.): *Improving Efficiency and Reliability in Water Distribution Systems.* 1995 ISBN 0-7923-3536-8
15. V.P. Singh (ed.): *Environmental Hydrology.* 1995 ISBN 0-7923-3549-X
16. V.P. Singh and B. Kumar (eds.): *Proceedings of the International Conference on Hydrology and Water Resources* (New Delhi, 1993). 1996
    16/1: Surface-water hydrology ISBN 0-7923-3650-X
    16/2: Subsurface-water hydrology ISBN 0-7923-3651-8

Water Science and Technology Library

16/3: Water-quality hydrology ISBN 0-7923-3652-6
16/4: Water resources planning and management ISBN 0-7923-3653-4
Set 16/1–16/4 ISBN 0-7923-3654-2

17. V.P. Singh: *Dam Breach Modeling Technology.* 1996 ISBN 0-7923-3925-8
18. Z. Kaczmarek, K.M. Strzepek, L. Somlyódy and V. Priazhinskaya (eds.): *Water Resources Management in the Face of Climatic/Hydrologic Uncertainties.* 1996 ISBN 0-7923-3927-4
19. V.P. Singh and W.H. Hager (eds.): *Environmental Hydraulics.* 1996 ISBN 0-7923-3983-5
20. G.B. Engelen and F.H. Kloosterman: *Hydrological Systems Analysis.* Methods and Applications. 1996 ISBN 0-7923-3986-X
21. A.S. Issar and S.D. Resnick (eds.): *Runoff, Infiltration and Subsurface Flow of Water in Arid and Semi-Arid Regions.* 1996 ISBN 0-7923-4034-5
22. M.B. Abbott and J.C. Refsgaard (eds.): *Distributed Hydrological Modelling.* 1996 ISBN 0-7923-4042-6
23. J. Gottlieb and P. DuChateau (eds.): *Parameter Identification and Inverse Problems in Hydrology, Geology and Ecology.* 1996 ISBN 0-7923-4089-2
24. V.P. Singh (ed.): *Hydrology of Disasters.* 1996 ISBN 0-7923-4092-2
25. A. Gianguzza, E. Pelizzetti and S. Sammartano (eds.): *Marine Chemistry.* An Environmental Analytical Chemistry Approach. 1997 ISBN 0-7923-4622-X
26. V.P. Singh and M. Fiorentino (eds.): *Geographical Information Systems in Hydrology.* 1996 ISBN 0-7923-4226-7
27. N.B. Harmancioglu, V.P. Singh and M.N. Alpaslan (eds.): *Environmental Data Management.* 1998 ISBN 0-7923-4857-5
28. G. Gambolati (ed.): *CENAS. Coastline Evolution of the Upper Adriatic Sea Due to Sea Level Rise and Natural and Anthropogenic Land Subsidence.* 1998 ISBN 0-7923-5119-3
29. D. Stephenson: *Water Supply Management.* 1998 ISBN 0-7923-5136-3
30. V.P. Singh: *Entropy-Based Parameter Estimation in Hydrology.* 1998 ISBN 0-7923-5224-6

Kluwer Academic Publishers – Dordrecht / Boston / London