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

$$f(x) = \frac{1}{(x-a)\, c\sqrt{2\pi}}\, exp\, [\frac{-[ln(x-a)-b]^2}{2c^2}] \tag{7.1a}$$

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

$$F(x) = \int_a^x \frac{1}{(x-a)c\sqrt{2\pi}}\, exp\, [-\frac{(ln(x-a)-b)^2}{2c^2}]\, dx \tag{7.2}$$

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:

$$\int_a^\infty f(x)dx = \frac{1}{c\sqrt{2\pi}} \int_a^\infty \frac{1}{(x-a)} exp\, [\frac{-[ln(x-a)\ -\ b]^2}{2c^2}]dx \tag{7.3}$$

Let

$$z = \frac{ln(x-a)\ -\ \bar{y}}{c};\ \frac{dz}{dx} = \frac{1}{(x-a)c} \tag{7.4}$$